(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 641 175**

(51) Int. CI.:

G10L 19/008 (2013.01)
G10L 19/16 (2013.01)
G06F 17/16 (2006.01)
H04S 7/00 (2006.01)
G10L 19/20 (2013.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **29.05.2014 PCT/US2014/040048**

(87) Fecha y número de publicación internacional: **04.12.2014 WO14194110**

(96) Fecha de presentación y número de la solicitud europea: **29.05.2014 E 14736510 (0)**

(97) Fecha y número de publicación de la concesión europea: **21.06.2017 EP 3005361**

(54) Título: **Compresión de las representaciones descompuestas de un campo sonoro**

(30) Prioridad:

29.05.2013 US 201361828445 P
29.05.2013 US 201361828615 P
30.05.2013 US 201361829174 P
30.05.2013 US 201361829155 P
30.05.2013 US 201361829182 P
31.05.2013 US 201361829846 P
31.05.2013 US 201361829791 P
03.10.2013 US 201361886617 P
03.10.2013 US 201361886605 P
01.11.2013 US 201361899041 P
01.11.2013 US 201361899034 P
08.01.2014 US 201461925074 P
08.01.2014 US 201461925112 P
08.01.2014 US 201461925126 P
08.01.2014 US 201461925158 P
30.01.2014 US 201461933706 P
30.01.2014 US 201461933721 P
27.05.2014 US 201462003515 P
28.05.2014 US 201414289522

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
08.11.2017

(73) Titular/es:

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

(72) Inventor/es:

SEN, DIPANJAN y
RYU, SANG-UK

(74) Agente/Representante:

FORTEA LAGUNA, Juan José

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

Compresión de las representaciones descompuestas de un campo sonoro

CAMPO TÉCNICO

[0001] Esta divulgación se refiere a datos de audio y, más específicamente, a la compresión de datos de audio.

ANTECEDENTES

[0002] 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 este campo sonoro de una manera que es 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 hacia atrás, ya que esta señal SHC puede reproducirse en formatos multicanal bien conocidos y altamente 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 acomoda la compatibilidad hacia atrás.

[0003] En el documento EP 2665208 A1, se estiman las direcciones de sonido dominantes y la representación de señales HOA se descompone en señales direccionales dominantes en dominio de tiempo y en información de dirección relacionada, y un componente ambiental en dominio HOA seguido de la compresión del componente ambiental reduciendo su orden. El componente ambiental de orden reducido se transforma en el dominio espacial, y se codifica perceptualmente junto con las señales direccionales. En el lado del receptor, las señales direccionales codificadas y el componente ambiental codificado de orden reducido se descomprimen perceptualmente, las señales ambientales perceptualmente descomprimidas se transforman en una representación de dominio HOA de orden reducido, seguido de extensión de orden. La representación HOA total se vuelve a componer a partir de las señales direccionales, la información de dirección correspondiente, y el componente HOA ambiental de orden original.

RESUMEN

[0004] En general, se describen técnicas para la compresión y descompresión de datos de audio ambisónicos de orden superior.

[0005] De acuerdo con un primer aspecto, se proporciona un procedimiento que comprende: obtener un flujo de bits que comprende una versión comprimida de un componente espacial de un campo sonoro, estando el componente espacial definido en un dominio armónico esférico y generado realizando una descomposición lineal con respecto a una pluralidad de coeficientes armónicos esféricos, en el que la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un campo que especifica un modo de predicción utilizado al comprimir el componente espacial.

[0006] De acuerdo con un segundo aspecto, se proporciona un dispositivo que comprende uno o más procesadores configurados para: obtener un flujo de bits que comprende una versión comprimida de un componente espacial de un campo sonoro, el componente espacial definido en un dominio armónico esférico y generado realizando una descomposición lineal con respecto a una pluralidad de coeficientes armónicos esféricos, en el que la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un campo que especifica un modo de predicción utilizado al comprimir el componente espacial.

[0007] En un modo de realización, la versión comprimida del componente espacial se representa además en el flujo de bits utilizando, al menos en parte, información de tabla Huffman que especifica una tabla Huffman usada al comprimir el componente espacial.

[0008] En un modo de realización, la versión comprimida del componente espacial se representa además en el flujo de bits utilizando, al menos en parte, un campo que indica un valor que expresa un tamaño del paso de cuantización o una variable del mismo utilizado al comprimir el componente espacial. Preferiblemente, el flujo de bits comprende una versión comprimida de una pluralidad de componentes espaciales del campo sonoro de la que está incluida la versión comprimida del componente espacial, y en el que el valor expresa el tamaño del paso de cuantización o una variable del mismo utilizada al comprimir la pluralidad de componentes espaciales.

[0009] En un modo de realización, la versión comprimida del componente espacial se representa además en el flujo de bits utilizando, al menos en parte, un código Huffman para representar un identificador de categoría que identifica una categoría de compresión a la que corresponde el componente espacial.

[0010] En un modo de realización, la versión comprimida del componente espacial se representa además en el flujo de bits utilizando, al menos en parte, un bit de signo que identifica si el componente espacial es un valor positivo o un valor negativo.

[0011] En un modo de realización, la versión comprimida del componente espacial se representa además en el flujo de bits utilizando, al menos en parte, un código Huffman para representar un valor residual del componente espacial.

[0012] En un modo de realización, la descomposición comprende una descomposición de valor singular.

[0013] De acuerdo con un aspecto adicional, se proporciona un procedimiento que comprende: generar un flujo de bits que comprende una versión comprimida de un componente espacial de un campo sonoro, estando el componente espacial definido en un dominio armónico esférico y generado realizando una descomposición lineal con respecto a una pluralidad de coeficientes armónicos esféricos, en el que la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un campo que especifica un modo de predicción utilizado al comprimir el componente espacial.

[0014] De acuerdo con un aspecto adicional, se proporciona un dispositivo que comprende uno o más procesadores configurados para: generar un flujo de bits que comprende una versión comprimida de un componente espacial de un campo sonoro, estando el componente espacial definido en un dominio armónico esférico y generado realizando una descomposición lineal con respecto a una pluralidad de coeficientes armónicos esféricos, en el que la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un campo que especifica un modo de predicción utilizado al comprimir el componente espacial.

[0015] En un modo de realización, el uno o más procesadores están configurados para generar el flujo de bits para incluir un campo que indica un valor que expresa un tamaño de paso de cuantización o una variable del mismo utilizado al comprimir el componente espacial, en los que el uno o más procesadores están configurados para generar el flujo de bits para incluir una versión comprimida de una pluralidad de componentes espaciales del campo sonoro del que está incluida la versión comprimida del componente espacial, y en los que el valor expresa el tamaño del paso de cuantización o una variable del mismo, utilizado al comprimir la pluralidad de componentes espaciales.

[0016] En un modo de realización, el uno o más procesadores están configurados para generar el flujo de bits para incluir información de tabla Huffman que especifica una tabla Huffman usada al comprimir el componente espacial.

[0017] En un modo de realización, el uno o más procesadores están configurados para generar el flujo de bits para incluir un código Huffman para representar un identificador de categoría que identifica una categoría de compresión con la que se corresponde el componente espacial.

[0018] En un modo de realización, el uno o más procesadores están configurados para generar el flujo de bits para incluir un código Huffman para representar una valor residual del componente espacial.

[0019] En un modo de realización, el uno o más procesadores están configurados para generar el flujo de bits para incluir un bit de signo que identifica si el componente espacial es un valor positivo o un valor negativo.

[0020] De acuerdo con un aspecto adicional, se proporciona un medio de almacenamiento no transitorio legible por ordenador que comprende instrucciones que, cuando se ejecutan, hacen que uno o más procesadores realicen uno de los procedimientos descritos anteriormente.

[0021] Los detalles de uno o más aspectos de las técnicas se exponen en los dibujos adjuntos y en la siguiente descripción. Otras características, objetos y ventajas de estas técnicas resultarán evidentes a partir de la descripción y de los dibujos, y a partir de las reivindicaciones.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0022]

Las Figs. 1 y 2 son diagramas que ilustran funciones de base armónicas esféricas de varios órdenes y subórdenes.

La Fig. 3 es un diagrama de bloques que ilustra un dispositivo que puede implementar una o más técnicas ejemplares descritas en esta divulgación.

La Fig. 4 es un diagrama de bloques que ilustra, más en detalle, un ejemplo del codificador de audio que se muestra en el ejemplo de la Fig. 3 que puede realizar diversos aspectos de las técnicas descritas en esta divulgación.

La Fig. 5 es un diagrama de bloques que ilustra el dispositivo de descodificación de audio que se muestra en el ejemplo de la Fig. 3 con mayor detalle.

La Fig. 6 es un diagrama de flujo que ilustra el funcionamiento ejemplar de una unidad de análisis de contenido de un dispositivo de codificación de audio en la realización de varios aspectos de las técnicas descritas en esta divulgación.

La Fig. 7 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.

La Fig. 8 es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de descodificación de audio en la realización de varios aspectos de las técnicas descritas en esta divulgación.

Las Figs. 9A-9L son diagramas de bloques que ilustran diversos aspectos del dispositivo de codificación de audio del ejemplo de la Fig. 4 con mayor detalle.

Las Figs. 10A-10O(ii) son diagramas que ilustran una parte de la información del canal de bits o del canal lateral que puede especificar los componentes espaciales comprimidos con más detalle.

Las Figs. 11A-11G son diagramas de bloques que ilustran, con más detalle, varias unidades del dispositivo de descodificación de audio mostrado en el ejemplo de la Fig. 5.

La Fig. 12 es un diagrama de bloques que ilustra un ecosistema de audio ejemplar que puede realizar diferentes aspectos de las técnicas descritas en esta divulgación.

La Fig. 13 es un diagrama de bloques que ilustra un ejemplo del ecosistema de audio de la Fig. 12 con mayor detalle.

La Fig. 14 es un diagrama de bloques que ilustra un ejemplo del ecosistema de audio de la Fig. 12 con mayor detalle.

Las Figs. 15A y 15B son diagramas que ilustran otros ejemplos del ecosistema de audio de la Fig. 12 con mayor detalle.

La Fig. 16 es un diagrama de bloques que ilustra un dispositivo de codificación de vídeo ejemplar que puede realizar diferentes aspectos de las técnicas descritas en esta divulgación.

La Fig. 17 es un diagrama de bloques que ilustra un ejemplo del dispositivo de codificación de audio de la Fig. 16 con mayor detalle.

La Fig. 18 es un diagrama de bloques que ilustra un dispositivo de descodificación de audio ejemplar que puede implementar diferentes aspectos de las técnicas descritas en esta divulgación.

La Fig. 19 es un diagrama de bloques que ilustra un ejemplo del dispositivo de descodificación de audio de la Fig. 18 con mayor detalle.

Las Figs. 20A-20G son diagramas que ilustran ejemplos de dispositivos de adquisición de audio que pueden realizar diversos aspectos de las técnicas descritas en esta divulgación.

Las Figs. 21A-21E son diagramas que ilustran ejemplos de dispositivos de reproducción de audio que pueden realizar diversos aspectos de las técnicas descritas en esta divulgación.

Las Figs. 22A-22H son diagramas que ilustran ejemplos de entornos de reproducción de audio de acuerdo con una o más técnicas descritas en esta divulgación.

La Fig. 23 es un diagrama que ilustra un ejemplo de caso de uso en el que un usuario puede experimentar un campo sonoro 3D de un juego deportivo mientras lleva auriculares de acuerdo con una o más técnicas descritas en esta divulgación.

La Fig. 24 es un diagrama que ilustra un estadio deportivo al que se puede grabar un campo sonoro 3D de acuerdo con una o más técnicas descritas en esta divulgación.

La Fig. 25 es un diagrama de flujo que ilustra una técnica para reproducir un campo sonoro 3D basado en un paisaje de audio local de acuerdo con una o más técnicas descritas en esta divulgación.

La Fig. 26 es un diagrama de flujo que ilustra un juego de estudio ejemplar de acuerdo con una o más técnicas descritas en esta divulgación.

La Fig. 27 es un diagrama que ilustra una pluralidad de sistemas de juego que incluyen motores de reproducción de acuerdo con una o más técnicas descritas en esta divulgación.

La Fig. 28 es un diagrama que ilustra una configuración de altavoz que puede ser simulada con auriculares de acuerdo con una o más técnicas descritas en esta divulgación.

La Fig. 29 es un diagrama que ilustra una pluralidad de dispositivos móviles que pueden usarse para adquirir y/o editar un campo sonoro 3D de acuerdo con una o más técnicas descritas en esta divulgación.

La Fig. 30 es un diagrama que ilustra una trama de vídeo asociada con un campo sonoro 3D que puede procesarse de acuerdo con una o más técnicas descritas en esta divulgación.

Las Figs. 31A-31M son diagramas que ilustran gráficas que muestran diversos resultados de la simulación de realizar una categorización sintética o registrada del campo sonoro de acuerdo con diversos aspectos de las técnicas descritas en esta divulgación.

La Fig. 32 es un diagrama que ilustra un gráfico de valores singulares de una matriz S descompuesta a partir de coeficientes ambisónicos de orden superior de acuerdo con las técnicas descritas en esta divulgación.

Las Figs. 33A y 33B son diagramas que ilustran gráficas respectivas que muestran un impacto potencial que tiene la reordenación cuando se codifican los vectores que describen los componentes de primer plano del campo sonoro de acuerdo con las técnicas descritas en esta divulgación.

Las Figs. 34 y 35 son diagramas conceptuales que ilustran diferencias entre únicamente la identificación basada en energía y la direccionalidad de distintos objetos de audio, de acuerdo con esta divulgación.

Las Figs. 36A-36G son diagramas que ilustran proyecciones de, al menos, una parte de la versión descompuesta de los coeficientes armónicos esféricos en el dominio espacial para realizar la interpolación de acuerdo con diversos aspectos de las técnicas descritas en esta divulgación.

La Fig. 37 ilustra una representación de técnicas para obtener una interpolación espacio-temporal como se describe en el presente documento.

La Fig. 38 es un diagrama de bloques que ilustra las matrices US artificiales, US_1 y US_2, para los bloques SVD secuenciales de una señal multidimensional de acuerdo con las técnicas descritas en el presente documento.

La Fig. 39 es un diagrama de bloques que ilustra la descomposición de las tramas subsiguientes de una señal ambisónica de orden superior (HOA) usando la descomposición en valores singulares y el suavizado de los componentes espacio-temporales de acuerdo con las técnicas descritas en esta divulgación.

Las Figs. 40A-40J son cada una un diagrama de bloques que ilustra ejemplos de dispositivos de codificación de audio que pueden realizar diversos aspectos de las técnicas descritas en esta divulgación para comprimir coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales.

La Fig. 41A-41D son diagramas de bloques que ilustran cada uno un ejemplo de dispositivo de descodificación de audio que puede realizar diversos aspectos de las técnicas descritas en esta divulgación para descodificar coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales.

Las Figs. 42A-42C son diagramas de bloques que ilustran la unidad de reducción de orden mostrada en los ejemplos de las Figs. 40B-40J con más detalle.

La Fig. 43 es un diagrama que ilustra la unidad de compresión V mostrada en la Fig. 1 con mayor detalle.

La Fig. 44 es una ilustración del diagrama de operaciones ejemplares realizadas por el dispositivo de codificación de audio para compensar el error de cuantización de acuerdo con diversos aspectos de las técnicas descritas en esta divulgación.

Las Figs. 45A y 45B son diagramas que ilustran la interpolación de subtramas de partes de dos tramas de acuerdo con diversos aspectos de las técnicas descritas en esta divulgación.

Las Figs. 46A-46E son diagramas que ilustran una sección transversal de una proyección de uno o más vectores de una versión descompuesta de una pluralidad de coeficientes armónicos esféricos que han sido interpolados de acuerdo con las técnicas descritas en esta divulgación.

La Fig. 47 es un diagrama de bloques que ilustra, con más detalle, la unidad de extracción de los dispositivos de descodificación de audio mostrados en los ejemplos de las Figs. 41A-41D.

La Fig. 48 es un diagrama de bloques que ilustra la unidad de reproducción de audio del dispositivo de descodificación de audio mostrado en los ejemplos de la Figs. 41A-41D con más detalle.

Las Figs. 49A-49E(ii) son diagramas que ilustran sistemas de codificación de audio respectivos que pueden implementar diversos aspectos de las técnicas descritas en esta divulgación.

Las Figs. 50A y 50B son diagramas de bloques que ilustran cada uno de dos enfoques diferentes para reducir potencialmente el orden del contenido de fondo de acuerdo con las técnicas descritas en esta divulgación.

La Fig. 51 es un diagrama de bloques que ilustra ejemplos de una trayectoria de compresión de componente distinta de un dispositivo de codificación de audio que puede implementar diversos aspectos de las técnicas descritas en esta divulgación para comprimir coeficientes armónicos esféricos.

Las Figs. 52 es un diagrama de bloques que ilustra otro ejemplo de un dispositivo de descodificación de audio que puede implementar diversos aspectos de las técnicas descritas en esta divulgación para reconstruir o casi reconstruir coeficientes armónicos esféricos (SHC).

La Fig. 53 es un diagrama de bloques que ilustra otro ejemplo de un dispositivo de codificación de audio que puede realizar diversos aspectos de las técnicas descritas en esta divulgación.

La Fig. 54 es un diagrama de bloques que ilustra, con más detalle, el dispositivo de codificación de audio mostrado en el ejemplo de la Fig. 53.

Las Figs. 55A y 55B son diagramas que ilustran un ejemplo de realización de diversos aspectos de las técnicas descritas en esta divulgación para girar un campo sonoro.

La Fig. 56 es un diagrama que ilustra un campo sonoro ejemplar capturado de acuerdo con una primera trama de referencia que se hace girar de acuerdo con las técnicas descritas en esta divulgación para expresar el campo sonoro en términos de una segunda trama de referencia.

Las Figs. 57A-57E son cada una un diagrama que ilustra los flujos de bits formados de acuerdo con las técnicas descritas en esta divulgación.

La Fig. 58 es un diagrama de flujo que ilustra el ejemplo de funcionamiento del dispositivo de codificación de audio mostrado en el ejemplo de la Fig. 53 en la implementación de los aspectos de rotación de las técnicas descritas en esta divulgación.

La Fig. 59 es un diagrama de flujo que ilustra el ejemplo de funcionamiento del dispositivo de codificación de audio mostrado en el ejemplo de la Fig. 53 en la realización de los aspectos de transformación de las técnicas descritas en esta divulgación.

DESCRIPCIÓN DETALLADA

[0023] La evolución del sonido envolvente ha puesto a disposición 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 "canales" basados en que, implícitamente, especifican las señales a los altavoces en ciertas coordenadas geométricas. Estos 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 de sonido envolvente izquierdo, posterior derecho o de sonido envolvente derecho y 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 Alta Definición). Los formatos no consumidores pueden abarcar cualquier número de altavoces (en geometrías simétricas y no simétricas) a menudo denominados "configuraciones de sonido envolvente". Un ejemplo de una configuración de este tipo incluye 32 altavoces colocados en coordenadas en las esquinas de un icosaedro truncado.

[0024] La entrada a un futuro codificador MPEG es, opcionalmente, uno de tres formatos posibles:

(i) audio basado en un canal tradicional (como se ha comentado anteriormente), que está destinado a ser reproducido a través de altavoces en posiciones preestablecidas; (ii) audio basado en objetos, que implica datos de modulación por código de pulso (PCM) discretos 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"). Este futuro codificador MPEG se puede describir con mayor detalle en un documento titulado "Convocatoria de propuestas para audio 3D", de la Organización Internacional para la Estandarización/Comisión

Electrotécnica Internacional (ISO)/(IEC) JTC1/SC29/WG11/N13411, publicado en enero 2013 en Ginebra, Suiza, y disponible en http://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/w13411.zip.

[0025] Hay varios formatos de "sonido envolvente" en el mercado. Por ejemplo, desde el sistema de cine en casa 5.1 (que es el que ha tenido más éxito en cuanto a entrar en las salas de estar más allá del estéreo) hasta el sistema 22.2 desarrollado por NHK (Nippon Hoso Kyokai o Japan Broadcasting Corporation). Los creadores de contenido (por ejemplo, los estudios de Hollywood) desean producir la banda sonora de una película una vez y no gastar los esfuerzos en volver a mezclarla para cada configuración de altavoz. Recientemente, las Organizaciones de Desarrollo de Normas han estado considerando maneras de proporcionar una codificación en un flujo de bits estandarizado y una descodificación subsiguiente que sea adaptable e indiferente con respecto a la geometría (y número) del altavoz y a las condiciones acústicas en la ubicación de la reproducción (que implica un reproductor).

[0026] 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 hace más detallada, aumentando la resolución.

[0027] 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:

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

[0028] Esta expresión muestra que la presión p_i en cualquier punto $\{r_r, \theta_r, \varphi_r\}$ del campo sonoro, en el tiempo t, puede representarse únicamente por el 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 de base armónicas esféricas de orden n y suborden 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 tiempo-frecuencia, tales como la transformada discreta de Fourier (DFT), la transformada de coseno discreta (DCT) o una transformada de ondícula. Otros ejemplos de conjuntos jerárquicos incluyen conjuntos de coeficientes de transformada de ondícula y otros conjuntos de coeficientes de funciones de base a partir de multirresolución.

[0029] 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 subórdenes m que se muestran pero no se anotan explícitamente en el ejemplo de la Fig. 1 para facilitar la ilustración.

[0030] La Fig. 2 es otro diagrama que ilustra las funciones de base armónicas esféricas desde el orden cero ($n = 0$) hasta el cuarto orden ($n = 4$). En la Fig. 2, las funciones de la base armónica esférica se muestran en el espacio de coordenadas tridimensional con el orden y el suborden mostrados.

[0031] El SHC $A_n^m(k)$ puede ser físicamente adquirido (por ejemplo, registrado) con varias configuraciones de micrófono o, alternativamente, se puede obtener de descripciones basadas en canales o basadas en objetos del campo sonoro. El SHC representa el audio basado en escenas, donde el SHC puede ser introducido en un codificador de audio para obtener el SHC codificado que puede fomentar una transmisión o un almacenamiento más eficiente. Por ejemplo, puede utilizarse una representación de cuarto orden que implique coeficientes $(1+4)^2$ (25 y, por lo tanto, de cuarto orden).

[0032] Como se ha indicado anteriormente, el SHC puede obtenerse de una grabación con micrófono utilizando un micrófono. En Poletti, M., "Three-Dimensional Surround Sound Systems Based on Spherical Harmonics", se describen varios ejemplos de cómo se puede obtener el SHC a partir de configuraciones de micrófonos, J. Audio Eng. Soc., Vol. 53, nº 11, 2005 Noviembre, págs. 1004 - 1025.

[0033] Para ilustrar cómo estos SHCs se pueden obtener a partir de una descripción basada en objetos, considere la siguiente ecuación. Los coeficientes $A_n^m(k)$ del 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 $\sqrt{-1}$, $h_n^{(2)}(\cdot)$ es la función esférica de Hankel (del segundo tipo) de orden n, y $\{r_s, \theta_s, \varphi_s\}$ es la ubicación del objeto. El hecho de conocer la fuente de energía del objeto $g(\omega)$ como una función de la frecuencia (por ejemplo, utilizando técnicas de análisis de tiempo-frecuencia, tales como la realización de una transformada rápida de Fourier en el flujo PCM) nos permite convertir cada objeto PCM y su ubicación en el SHC $A_n^m(k)$. Además, se puede demostrar (ya 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 PCM puede ser representada con los coeficientes $A_n^m(k)$ (por ejemplo, como una suma de los vectores de coeficientes para los objetos individuales). Esencialmente, estos coeficientes contienen información sobre el campo sonoro (la presión como una función de las coordenadas 3D), 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.

[0034] La Fig. 3 es un diagrama de bloques 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. 3, el sistema 10 incluye un creador de contenido 12 y un consumidor de contenido 14. Aunque se describen en el contexto del creador de contenido 12 y del 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 creador de contenido 12 puede representar cualquier forma de dispositivo informático capaz de implementar las técnicas descritas en esta divulgación, incluido un teléfono móvil (o teléfono celular), un ordenador de tableta, un teléfono inteligente o un ordenador de sobremesa para proporcionar algunos ejemplos. Del mismo modo, el consumidor de contenido 14 puede representar cualquier forma de dispositivo informático capaz de implementar las técnicas descritas en esta divulgación, incluido un teléfono móvil, una tableta, un teléfono inteligente, un descodificador o un ordenador de sobremesa para proporcionar algunos ejemplos.

[0035] El creador de contenido 12 puede representar un estudio de cine u otra entidad que puede generar contenido de audio multicanal para el consumo de los consumidores de contenido, tal como el consumidor de contenido 14. En algunos ejemplos, el creador de contenido 12 puede representar a 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 consumidor de contenido 14 representa un individuo que posee o tiene acceso a un sistema de reproducción de audio, que puede referirse a cualquier forma de sistema de reproducción de audio capaz de reproducir SHC para reproducirlo como contenido de audio multicanal. En el ejemplo de la Fig. 3, el consumidor de contenido 14 incluye un sistema de reproducción de audio 16.

[0036] El creador de contenido 12 incluye un sistema de edición de audio 18. El creador de contenido 12 obtiene grabaciones en directo 7 en diversos formatos (incluidos directamente como coeficientes HOA) y objetos de audio 9, que el creador de contenido 12 puede editar utilizando el sistema de edición de audio 18. El creador de contenido puede, durante el proceso de edición, reproducir los coeficientes HOA 11 de los objetos de audio 9, escuchando las señales de los altavoces reproducidas en un intento de identificar varios aspectos del campo sonoro que requieran una edición adicional. El creador de contenido 12 puede entonces editar los coeficientes HOA 11 (potencialmente indirectamente a través de la manipulación de diferentes objetos de audio 9 a partir de los cuales pueden obtenerse los coeficientes HOA originales de la manera descrita anteriormente). El 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 estos datos de audio como uno o más coeficientes armónicos esféricos originales.

[0037] Cuando se completa el proceso de edición, el creador de contenido 12 puede generar un flujo de bits 21 basado en los coeficientes HOA 11. Es decir, el creador de contenido 12 incluye un dispositivo de codificación de audio 20 que representa un dispositivo configurado para codificar o, de algún modo, comprimir los coeficientes HOA 11 de acuerdo con 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 la 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.

[0038] 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 debe realizar la metodología de síntesis basada en vectores o una metodología de síntesis basada en la dirección, el dispositivo de codificación de audio 20 puede determinar, en base al menos en parte en los coeficientes HOA 11, si los coeficientes HOA 11 se generaron a través de una grabación natural de un campo sonoro (por ejemplo, grabación en directo 7) o se produjeron 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 síntesis basada en la dirección. Cuando los coeficientes HOA 11 se

capturaron en directo usando, por ejemplo, un eigenmike, el dispositivo de codificación de audio 20 puede codificar los coeficientes HOA 11 basados en la metodología de síntesis basada en vectores. La distinción anterior representa un ejemplo de dónde puede implantarse la metodología de síntesis basada en vectores o direccional. Puede haber otros casos en que uno o ambos pueden ser útiles para las grabaciones naturales, el contenido generado artificialmente o una mezcla de ambos (contenido híbrido). Además, también es posible utilizar ambas metodologías simultáneamente para codificar un solo intervalo de tiempo de los coeficientes HOA.

[0039] Asumiendo con fines de ilustración que el dispositivo de codificación de audio 20 determina que los coeficientes HOA 11 se capturaron en directo o, de algún modo, representan grabaciones en directo, tales como la grabación en directo 7, el dispositivo de codificación de audio 20 puede configurarse para codificar los coeficientes HOA 11 utilizando un vector basado en la aplicación de una transformada lineal invertible (LIT). Un ejemplo de la transformada lineal invertible 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 la reordenación 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 dicha reordenación, 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 a través de las tramas de los coeficientes HOA (donde una trama comúnmente incluye M muestras de los coeficientes HOA 11 y M es, en algunos ejemplos, establecido en 1024). Después de reordenar la versión descompuesta de los coeficientes HOA 11, el dispositivo de codificación de audio 20 puede seleccionar aquellos coeficientes de 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 direccional asociada.

[0040] 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 aquellos coeficientes de los coeficientes HOA 11 representativos de uno o más componentes de fondo del campo sonoro. El dispositivo de codificación de audio 20 puede realizar compensación de energía con respecto a los componentes de fondo dado que, en algunos ejemplos, los componentes de fondo solo pueden incluir un subconjunto de cualquier muestra dada de los coeficientes HOA 11 (por ejemplo, los correspondientes a las funciones de base esférica de cero y primer orden y no las correspondientes a las funciones de base esférica de segundo 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 fondo restantes de los coeficientes HOA 11 para compensar el cambio en la energía total que resulta de realizar la reducción de orden.

[0041] El dispositivo de codificación de audio 20 puede realizar, a continuación, una forma de codificación psicoacústica (tal como la envolvente 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 fondo y de 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 por 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 por orden, que proporciona información direccional de primer plano codificada. En algunos casos, esta cuantización puede comprender una cuantización escalar/entrópica. El dispositivo de codificación de audio 20 puede formar entonces el flujo de bits 21 para incluir los componentes de fondo 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, de algún modo, emitir el flujo de bits 21 al consumidor de contenido 14.

[0042] Como se muestra en la Fig. 3 como transmitido directamente al consumidor de contenido 14, el creador de contenido 12 puede emitir el flujo de bits 21 a un dispositivo intermedio situado entre el creador de contenido 12 y el consumidor de contenido 14. Este dispositivo intermedio puede almacenar el flujo de bits 21 para su posterior suministro al consumidor de contenido 14, que puede solicitar este flujo de bits. El dispositivo intermedio puede comprender un servidor de archivos, un servidor web, un ordenador de sobremesa, un ordenador portátil, un ordenador de tabletas, un teléfono móvil, un teléfono inteligente o cualquier otro dispositivo capaz de almacenar el flujo de bits 21 para su posterior recuperación mediante un descodificador 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 flujo de bits de datos de vídeo correspondiente) a los suscriptores, como el consumidor de contenido 14 que solicita el flujo de bits 21.

[0043] Alternativamente, el 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 transitorios legibles por ordenador. En este contexto, el canal de transmisión puede referirse a aquellos canales a través de los cuales se transmiten contenidos almacenados en estos medios (y pueden incluir tiendas minoristas y otros mecanismos de suministro basados en almacén). En cualquier caso, las técnicas de esta divulgación no deben, por lo tanto, estar limitadas a este respecto al ejemplo de la Fig. 3.

[0044] Como se muestra adicionalmente en el ejemplo de la Fig. 3, el 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 procesadores 22. Los reproductores 22 pueden proporcionar cada uno una forma diferente de reproducción, en la que 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".

[0045] El sistema de reproducción de audio 16 puede incluir además un dispositivo de descodificación de audio 24. El dispositivo de descodificación de audio 24 puede representar un dispositivo configurado para descodificar 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 pérdida (por ejemplo, cuantización) y/o la transmisión a través del canal de transmisión. Es decir, el dispositivo de descodificación de audio 24 puede desquantizar la información direccional de primer plano especificada en el flujo de bits 21, al mismo tiempo que realiza la descodificación 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 fondo. El dispositivo de descodificación de audio 24 puede realizar además la interpolación con respecto a la información direccional de primer plano descodificada y luego determinar los coeficientes HOA representativos de los componentes de primer plano basados en los objetos de audio de primer plano descodificados y la información direccional de primer plano interpolada. El dispositivo de descodificación 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 descodificados representativos de los componentes de fondo.

[0046] El sistema de reproducción de audio 16 puede, después de descodificar el flujo de bits 21, obtener los coeficientes HOA 11' y reproducir los coeficientes HOA 11' en las señales de salida de los altavoces 25. Las señales de los altavoces 25 pueden accionar uno o más altavoces (que no se muestran en el ejemplo de la Fig. 3 para facilitar la ilustración).

[0047] Para seleccionar el reproductor apropiado o, en algunos casos, generar un reproductor apropiado, 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 que determina dinámicamente la información de los altavoces 13. En otros casos o, junto con la determinación dinámica de la información de los altavoces 13, el sistema de reproducción de audio 16 puede pedirle a un usuario que interconecte con el sistema de reproducción de audio 16 e introduzca la información de los altavoces 16.

[0048] El sistema de reproducción de audio 16 puede entonces seleccionar una información de los reproductores de audio 22 en base a la información de los altavoces 13. En algunos casos, el sistema de reproducción de audio 16 puede, cuando ninguno de los procesadores de audio 22 están dentro de una medida de similitud de umbral (altavoz geométrica) con respecto a la especificada en la información de los altavoces 13, el sistema de reproducción de audio 16 puede generar la información de reproducción de audio 22 en base a la información de los altavoces 13. El sistema de reproducción de audio 16 puede, en algunos casos, generar la información de los procesadores de audio 22 en base a la información de los altavoces 13 sin intentar primero seleccionar una información existente de los procesadores de audio 22.

[0049] La Fig. 4 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. 3 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 metodología de síntesis basada en vectores 27 y una unidad de metodología de síntesis basada en la dirección 28.

[0050] La unidad de análisis de contenido 26 representa una unidad configurada para analizar el contenido de los coeficientes HOA 11 y que identifica si los coeficientes HOA 11 representan contenido generado a partir de una grabación en directo 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 audio artificial. La unidad de análisis de contenido 26 puede realizar esta determinación de varias maneras. Por ejemplo, la unidad de análisis de contenido 26 puede codificar $(N+1)^2$-1 canales y predecir el último canal restante (que puede representarse como un vector). La unidad de análisis de contenido 26 puede aplicar escalares a, al menos algunos

de los canales $(N+1)^2$-1, y añadir los valores resultantes para determinar el último canal restante. Además, en este ejemplo, la unidad de análisis de contenido 26 puede determinar una precisión del canal predicho. En este ejemplo, si la precisión del canal predicho es relativamente alta (por ejemplo, la precisión excede un umbral particular), es probable que los coeficientes HOA 11 se generen a partir de un objeto de audio sintético. Por el contrario, si la precisión del canal predicho es relativamente baja (por ejemplo, la precisión está por debajo del umbral particular), es más probable que los coeficientes HOA 11 representen un campo sonoro registrado. Por ejemplo, en este ejemplo, si una relación señal-ruido (SNR) del canal predicho es superior a 100 decibeles (dbs), los coeficientes HOA 11 son más propensos a representar un campo sonoro generado a partir de un objeto de audio sintético. En contraste, la SNR de un campo sonoro grabado utilizando un micrófono propio puede ser de 5 a 20 dbs. Por lo tanto, puede haber una demarcación aparente en las relaciones SNR entre el campo sonoro representado con los coeficientes HOA 11 generados a partir de una grabación directa real y desde un objeto de audio sintético.

[0051] Más específicamente, la unidad de análisis de contenido 26 puede, al determinar si los coeficientes HOA 11 representativos de un campo sonoro se generan a partir de un objeto de audio sintético, obtener unos coeficientes HOA por trama, que puede ser del tamaño 25 por 1024 para una representación de cuarto orden (es decir, N = 4). Después de obtener los coeficientes HOA por trama (que también se pueden indicar en el presente documento como una matriz de SHC por trama 11 y las matrices SHC por trama subsiguiente se pueden indicar como matrices SHC por trama 27B, 27C, etc.). La unidad de análisis de contenido 26 puede excluir entonces el primer vector de los coeficientes en HOA por trama 11 para generar coeficientes HOA reducidos por trama. En algunos ejemplos, este primer vector excluido de los coeficientes HOA por trama 11 puede corresponder a los de los coeficientes HOA 11 asociados con la función de base armónica esférica de orden cero y suborden cero.

[0052] La unidad de análisis de contenido 26 puede entonces predecir el primer vector distinto de cero de los coeficientes HOA reducidos por trama a partir de los vectores restantes de los coeficientes HOA reducidos por trama. El primer vector distinto de cero puede referirse a un primer vector que va desde el primer orden (y considerando cada uno de los subórdenes dependientes del orden) al cuarto orden (y considerando cada uno de los subórdenes dependientes del orden) que tiene valores distintos de cero. En algunos ejemplos, el primer vector distinto de cero de los coeficientes HOA reducidos por trama se refiere a los de los coeficientes HOA 11 asociados con la función de base armónica esférica de primer orden y suborden cero. Aunque se describen con respecto al primer vector distinto de cero, las técnicas pueden predecir otros vectores de los coeficientes HOA reducidos por trama de los vectores restantes de los coeficientes HOA reducidos por trama. Por ejemplo, la unidad de análisis de contenido 26 puede predecir los de los coeficientes HOA reducidos por trama asociados con una función de base armónica esférica de primer orden y primer suborden o una función de base armónica esférica de primer orden y primer orden negativo. Aún otros ejemplos, la unidad de análisis de contenido 26 puede predecir los de los coeficientes HOA reducidos por trama asociados con una función de base armónica esférica de segundo orden y orden cero.

[0053] Para predecir el primer vector distinto de cero, la unidad de análisis de contenido 26 puede funcionar de acuerdo con la siguiente ecuación:

$$\sum_i (\alpha_i v_i),$$

donde *i* es de 1 a $(N + 1)^2$-2, que es 23 para una representación de cuarto orden, α_i denota alguna constante para el i-ésimo vector y v_i se refiere al *i*- ésimo vector. Después de predecir el primer vector distinto de cero, la unidad de análisis de contenido 26 puede obtener un error basado en el primer vector distinto de cero predicho y el vector distinto de cero real. En algunos ejemplos, la unidad de análisis de contenido 26 substrae el primer vector distinto de cero predicho del primer vector distinto de cero real para derivar el error. La unidad de análisis de contenido 26 puede calcular el error como una suma del valor absoluto de las diferencias entre cada entrada en el primer vector distinto de cero predicho y el primer vector distinto de cero real.

[0054] Una vez obtenido el error, la unidad de análisis de contenido 26 puede calcular una relación basada en una energía del primer vector distinto de cero real y del error. La unidad de análisis de contenido 26 puede determinar esta energía cuadrando cada entrada del primer vector distinto de cero y añadiendo las entradas cuadradas una a otra. La unidad de análisis de contenido 26 puede entonces comparar esta relación con un umbral. Cuando la relación no excede el umbral, la unidad de análisis de contenido 26 puede determinar que los coeficientes HOA por trama 11 se generan a partir de una grabación e indica en el flujo de bits que la representación codificada correspondiente de los coeficientes HOA 11 se generó a partir de una grabación. Cuando la relación excede el umbral, la unidad de análisis de contenido 26 puede determinar que los coeficientes HOA por trama 11 se generan a partir de un objeto de audio sintético e indica en el flujo de bits que la representación codificada correspondiente de los coeficientes HOA por trama 11 se generó a partir de un objeto de audio sintético.

[0055] La indicación de si los coeficientes HOA por trama 11 se generaron a partir de una grabación o de un objeto de audio sintético puede comprender un solo bit para cada trama. El bit único puede indicar que se usaron diferentes codificaciones para cada trama de manera eficaz, intercambiando entre diferentes maneras de codificar la trama correspondiente. En algunos casos, cuando los coeficientes HOA por trama 11 se generaron a partir de una

grabación, la unidad de análisis de contenido 26 pasa los coeficientes HOA 11 a la unidad de síntesis basada en vectores 27. En algunos casos, cuando los coeficientes HOA por trama 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.

[0056] En otras palabras, las técnicas se basan en la codificación de los coeficientes HOA utilizando un clasificador de entrada. El clasificador puede funcionar de la siguiente manera:

Comienza con una matriz SH por trama (supongamos de 4º orden, tamaño de trama de 1024, que también se pueden denominar coeficientes HOA por trama o como coeficientes HOA) - en la que se obtiene una matriz de tamaño 25 x 1024.

[0057] Excluye el primer vector (SH de orden 0) - de manera que hay una matriz de tamaño 24 x 1024.

[0058] Predice el primer vector distinto de cero en la matriz (vector de tamaño a 1 x 1024) - del resto de los vectores de la matriz (23 vectores de tamaño 1x1024).

[0059] La predicción es la siguiente: vector predicho = suma-sobre-i [alfa-i x vector-I] (donde la suma sobre I se hace sobre 23 índices, i=1... 23).

[0060] A continuación, compruebe el error: vector real - vector predicho = error.

[0061] Si la relación de la energía del vector/error es grande (es decir, el error es pequeño), entonces el campo sonoro subyacente (en esa trama) es escaso/sintético. De lo contrario, el campo sonoro subyacente es un campo sonoro grabado (usando una configuración de micrófonos).

[0062] En función de la decisión grabada vs. sintética, lleva a cabo la codificación / descodificación (que puede referirse a la compresión del ancho de banda) de diferentes maneras. La decisión es una decisión de 1 bit que se envía a través del flujo de bits para cada trama.

[0063] Como se muestra en el ejemplo de la Fig. 4, la unidad de síntesis basada en vectores 27 puede incluir una unidad de transformada lineal invertible (LIT) 30, una unidad de cálculo de parámetros 32, una unidad de reordenación 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 del flujo de bits 42, una unidad de análisis del campo sonoro 44, una unidad de reducción de coeficientes 46, una unidad de selección de fondo 48 (BG), una unidad de interpolación espacio-temporal 50 y una unidad de cuantización 52.

[0064] La unidad de transformada lineal invertible (LIT) 30 recibe los coeficientes HOA 11 en forma de canales HOA, cada canal representativo de un bloque o trama de un coeficiente asociado con un orden dado, suborden 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 D dimensiones: $M \times (N+1)^2$.

[0065] Es decir, la unidad LIT 30 puede representar una unidad configurada para realizar una forma de análisis denominada descomposición en valores singulares. Aunque se describe 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 compacta y no correlacionada linealmente. También, la referencia a "conjuntos" en esta divulgación generalmente se refiere a conjuntos diferentes de cero a menos que se indique específicamente lo contrario y no se pretende referir a la definición matemática clásica de conjuntos que incluye el denominado "conjunto vacío".

[0066] 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 linealmente no correlacionadas denominadas componentes principales. Las variables linealmente no correlacionadas representan variables que no tienen una relación estadística lineal (o dependencia) entre sí. Estos componentes principales pueden ser descritos como teniendo 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 tiene la varianza más grande posible (o, en otras palabras, representa la mayor parte de la variabilidad en los datos como sea posible) y cada componente sucesivo a su vez tiene la mayor varianza posible bajo la restricción de que este componente sucesivo sea ortogonal (lo que puede ser reexpresado 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 lugar a la compresión de los coeficientes HOA 11. En función del contexto, el PCA puede ser referido con varios nombres diferentes, tales como la transformada discreta de Karhunen-Loeve, la

transformada de Hotelling, la descomposición ortogonal apropiada (POD) y la descomposición del valor propio (EVD) para nombrar algunos ejemplos. Las propiedades de dichas operaciones que conducen a la meta subyacente de compresión de datos de audio son la "compactación energética" y la "descorrelación" de los datos de audio multicanal.

[0067] En cualquier caso, la unidad LIT 30 realiza una descomposición en valores singulares (que, de nuevo, puede denominarse "SVD") para 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. 4, la unidad 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 real y-por-z o compleja X (donde X puede representar datos de audio multicanal como los coeficientes HOA 11) de la siguiente forma:

$$X = USV^*$$

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

[0068] Aunque se describe en esta divulgación como aplicable 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 de la matriz U, la matriz S y la matriz V.

[0069] En algunos ejemplos, la matriz V* en la expresión matemática SVD mencionada anteriormente se denomina transposición 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, el conjugado complejo de la matriz V (o, en otras palabras, la matriz V*) puede considerarse como la transposición de la matriz V. A continuación se supone, para facilitar la ilustración, que los coeficientes HOA 11 comprenden números reales, con el resultado de que la matriz V se emite a través de la SVD en lugar de la matriz V*. Además, aunque se denotan como la matriz V en esta divulgación, la referencia a la matriz V debe entenderse que hace referencia a la transposición de la matriz V cuando sea apropiado. 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 deben 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 tienen componentes complejos para generar una matriz V*.

[0070] En cualquier caso, la unidad LIT 30 puede realizar una forma en bloque de SVD con respecto a cada bloque (que puede referirse a una trama) de datos de audio ambisónicos (HOA) de orden superior (donde estos 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 a este valor típico para M, las técnicas de esta divulgación no deben limitarse a este valor típico para M. La unidad LIT 30 puede realizar, por lo tanto, una SVD en bloque con respecto a un bloque, con los coeficientes HOA 11 que tienen M-por-$(N + 1)^2$ coeficientes HOA, donde N, de nuevo, indica el orden de los datos de audio HOA. La unidad LIT 30 puede generar, mediante la realización de esta 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 transformada lineal invertible 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 S vectores y los U vectores) que tienen D dimensiones: Vectores Mx $(N+1)^2$ y V[k] 35 con D dimensiones: $(N+1)^2$ x $(N+1)^2$. 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).*

[0071] Un análisis de las matrices U, S y V puede revelar que estas 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 longitud M muestras) puede representar señales de audio normalizadas separadas en 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 anchura de la forma espacial y la posición (r, theta, phi) pueden

representarse en su lugar por vectores i^{th} individuales, $v^{(i)}(k)$ en la matriz V (cada uno de longitud $(N+1)^2$). Ambos vectores de la matriz U y la matriz V se normalizan de tal manera que sus energías de valor cuadrático medio son iguales a la unidad. La energía de las señales de audio en U está representada por los elementos diagonales en S. Multiplicando U y S para formar US[*k*] (con los elementos vectoriales individuales $X_{PS}(k)$), representan 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) pueden soportar varios aspectos de las técnicas descritas en esta divulgación. Además, este 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 "metodología de síntesis basada en vectores" que se utiliza en todo este documento.

[0072] Aunque se describe que se realiza directamente con respecto a los coeficientes HOA 11, la unidad LIT 30 puede aplicar la transformada lineal invertible a los derivados de los coeficientes HOA 11. Por ejemplo, la unidad LIT 30 puede aplicar una SVD con respecto a una matriz de densidad espectral de potencia obtenida 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 transposición del hoaFrame al hoaFrame, como se indica en el pseudocódigo que sigue a continuación. La notación HoaFrame se refiere a una trama de los coeficientes HOA 11.

[0073] La unidad LIT 30 puede, después de aplicar la SVD (svd) a la PSD, puede obtener una matriz $S[k]^2$ (S_squared) y V[*k*]. La matriz $S[k]^2$ puede indicar una matriz S[*k*] al cuadrado, con lo que la unidad LIT 30 puede aplicar una operación de raíz cuadrada a la matriz $S[k]^2$ para obtener la matriz S[*k*]. La unidad 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 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 LIT 30 puede obtener a continuación la pseudoinversa (pinv) de la matriz SV[*k*]' y luego multiplicar los coeficientes HOA 11 por la pseudoinversa de la matriz SV[*k*]' para obtener la matriz U[*k*]. Lo anterior puede estar representado por el siguiente pseudocódigo:

```
PSD = hoaFrame'* hoaFrame;

[V, S_squared] = svd(PSD, 'econ');

S = sqrt(S_squared);

U = hoaFrame * pinv(S*V');
```

[0074] 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 LIT 30 puede reducir potencialmente la complejidad computacional de realizar la SVD en términos de uno o más ciclos de procesador y espacio de almacenamiento, al mismo tiempo que 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 de cálculo porque la SVD se realiza sobre una matriz F*F (con F el número de coeficientes HOA). En comparación con una matriz M*F con M es la longitud de tramas, es decir, 1024 o más muestras. La complejidad de una SVD puede ahora, a través de la aplicación en 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 grande-O de la complejidad computacional común a las artes de la informática).

[0075] 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 estos parámetros para la trama actual puede indicarse como *R*[*k*], θ[*k*], φ[*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 estos parámetros. La unidad de cálculo de parámetros 32 también puede determinar estos parámetros para la trama anterior, donde los parámetros de trama anteriores pueden indicarse *R*[*k*-1], θ[*k*-1], φ [*k*-1], *r*[*k*-1] y *e*[*k*-1] en base a la trama anterior de vectores US[*k*-1] y 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.

[0076] Es decir, la unidad de cálculo de parámetros 32 puede realizar un análisis de energía con respecto a cada uno de los L primeros US[*k*] vectores 33 correspondientes a una primera vez y cada uno de los segundos vectores US[*k*-1] 33 correspondientes a una segunda vez, calculando una energía media cuadrática para al menos una parte de la primera trama de audio (pero a menudo la totalidad) y una parte de la segunda trama de audio (pero a menudo la totalidad) y de este modo generar 2L energías, una para cada uno de los L primeros vectores US[*k*] 33 de la primera trama de audio y uno para cada uno de los segundos vectores US[*k*-1] 33 de la segunda trama de audio.

[0077] En otros ejemplos, la unidad de cálculo de parámetros 32 puede realizar una correlación cruzada entre una parte (si no la totalidad) del conjunto de muestras para cada uno de los primeros vectores US[*k*] 33 y cada uno de los segundos vectores US[*k*-1] 33. La correlación cruzada puede referirse a la correlación cruzada tal como se

entiende en las técnicas de procesamiento de señales. En otras palabras, la correlación cruzada puede referirse a una medida de similitud entre dos formas de onda (que en este caso se define como un conjunto discreto de M muestras) como una función de un retardo de tiempo aplicada a una de ellas. En algunos ejemplos, para realizar una correlación cruzada, la unidad de cálculo de parámetros 32 compara las últimas L muestras de cada uno de los primeros vectores US[*k*] 27, por orden, con las L primeras muestras de cada uno de los restantes de los segundos vectores US[*k*-1] 33 para determinar un parámetro de correlación. Tal como se utiliza en el presente documento, una operación "por orden" se refiere a una operación elemento por elemento realizada con respecto a un primer conjunto de elementos y un segundo conjunto de elementos, en el que la operación extrae un elemento de cada uno de los conjuntos primero y segundo de elementos "uno tras otro" de acuerdo con una ordenación de los conjuntos.

[0078] La unidad de cálculo de parámetros 32 también puede analizar los vectores V[*k*] y/o V[*k*-1] 35 para determinar los parámetros de propiedad direccionales. Estos parámetros de propiedad direccionales pueden proporcionar una indicación de movimiento y localización del objeto de audio representado por los vectores correspondientes US[*k*] y/o US[*k*-1] 33. La unidad de cálculo de parámetros 32 puede proporcionar cualquier combinación de los parámetros actuales anteriores 37 (determinada con respecto a los vectores US[*k*] 33 y/o los vectores V[*k*] 35) y cualquier combinación de los parámetros anteriores 39 (determinada con respecto a los vectores US[*k*-1] 33 y/o los vectores V[*k*-1] 35) a la unidad de reordenación 34.

[0079] La descomposición SVD no garantiza que la señal/objeto de audio representada por el p-ésimo vector en los vectores US[*k*-1] 33, que se pueda indicar como vector US [*k*-1][p] (o, alternativamente, como $X_{PS}^{(p)}(k\text{-}1)$), será la misma señal/objeto de audio (progresado en el tiempo) representado por el vector p-ésimo en los vectores US[*k*] 33, que también se puede 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 reordenación 34 para reordenar los objetos de audio para representar su evaluación natural o continuidad en el tiempo.

[0080] Es decir, la unidad de reordenación 34 puede entonces comparar cada uno de los parámetros 37 de los primeros vectores US[*k*] 33 por orden frente a cada uno de los parámetros 39 de los segundos vectores US[*k*-1] 33. La unidad de reordenación 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 en base a los parámetros actuales 37 y los parámetros anteriores 39 para emitir una matriz US[*k*] reordenada 33' (que se puede designar matemáticamente como $\overline{US}[k]$) y una matriz V[*k*] reordenada 35' (que se puede designar matemáticamente como $\overline{V}[k]$) 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.

[0081] En otras palabras, la unidad de reordenación 34 puede representar una unidad configurada para reordenar los vectores dentro de la matriz US[*k*] 33 para generar la matriz US[*k*] reordenada 33'. La unidad de reordenación 34 puede reordenar la matriz US[*k*] 33 porque el orden de los vectores US[*k*] 33 (donde, de nuevo, cada vector de los vectores US[*k*] 33, que de nuevo se puede indicar como $X_{PS}^{(p)}(k)$, puede representar uno o más objetos monoaudio distintos (o, en otras palabras, predominantes) presentes en el campo sonoro) puede variar a partir de partes de los datos de audio. Es decir, dado que el dispositivo de codificación de audio 12, en algunos ejemplos, opera en estas partes de los datos de audio denominadas generalmente tramas de audio, la posición de los vectores correspondientes a estos distintos objetos monoaudio como se representa en la matriz US[*k*] 33 tal como se obtiene, puede variar desde la trama de audio a la trama de audio debido a la aplicación de la SVD a las tramas y la notabilidad variable de cada forma de objeto de audio trama a trama.

[0082] Pasar los vectores dentro de la matriz US[*k*] 33 directamente a la unidad codificadora de audio psicoacústico 40 sin reordenar los vectores dentro de la matriz US[*k*] 33 desde la trama de audio a la trama de audio puede reducir el alcance de la compresión alcanzable para algunos esquemas de compresión, tales como los esquemas de compresión heredados que funcionan mejor cuando los objetos monoaudio son continuos (en el sentido del canal, que se define en este ejemplo por el orden de posición de los vectores dentro de la matriz US[*k*] 33 relativos unos a otros) a través de tramas de audio. Además, cuando no se vuelve a ordenar, la codificación de los vectores dentro de la matriz US[*k*] 33 puede reducir la calidad de los datos de audio cuando se descodifican. Por ejemplo, los codificadores AAC que pueden representarse en el ejemplo de la Fig. 3 con la unidad codificadora de audio psicoacústico 40 pueden comprimir más eficazmente uno o más de los vectores reordenados dentro de la matriz US[*k*] 33' trama a trama en comparación con la compresión lograda al codificar directamente los vectores dentro de la matriz US[*k*] 33 trama a trama. Aunque se ha descrito anteriormente con respecto a los codificadores AAC, las técnicas pueden realizarse con respecto a cualquier codificador que proporcione mejor compresión cuando se especifican objetos monoaudio a través de tramas en un orden o posición específicos (sentido del canal).

[0083] Varios aspectos de las técnicas pueden, de este modo, permitir al dispositivo de codificación de audio 12 reordenar uno o más vectores (por ejemplo, los vectores dentro de la matriz US[*k*] 33 para generar uno o más vectores reordenados dentro de la matriz US[*k*] reordenada 33' y, de este modo, facilitar la compresión de los vectores dentro de la matriz US[*k*] 33 mediante un codificador de audio heredado tal como la unidad codificadora de audio psicoacústico 40).

[0084] Por ejemplo, la unidad de reordenación 34 puede reordenar uno o más vectores dentro de la matriz US[*k*] 33 desde una primera trama de audio subsiguiente en el tiempo a la segunda trama a la que uno o más vectores segundos dentro de la matriz US[*k*-1] 33 corresponden en base a los parámetros actuales 37 y los parámetros anteriores 39. Aunque se describe en el contexto de una primera trama de audio que es posterior en el tiempo a la segunda trama de audio, la primera trama de audio puede preceder en el tiempo a la segunda trama de audio. Por consiguiente, las técnicas no deben limitarse al ejemplo descrito en esta divulgación.

[0085] Para ilustrar considere la siguiente Tabla 1 donde cada uno de los p vectores dentro de la matriz US [*k*] 33 se denomina US[*k*][*p*], donde k indica si el vector correspondiente es de la k-ésima trama o la trama anterior (*k*-1) y p indica la fila del vector relativa a los vectores de la misma trama de audio (donde la matriz US[*k*] tiene $(N+1)^2$ de dichos vectores). Como se ha indicado anteriormente, suponiendo que se determina que N es uno, p puede indicar los vectores uno (1) a (4).

Tabla 1

Energía bajo consideración	Comparado con
US[*k*-1][1]	US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4]
US[*k*-1][2]	US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4]
US[*k*-1][3]	US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4]
US[*k*-1][4]	US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4]

[0086] En la Tabla 1 anterior, la unidad de reordenación 34 compara la energía calculada para US[*k*-1][1] con la energía calculada para cada uno de US[*k*][1], US[*k*] [2], US[*k*][3], US[*k*][4], la energía calculada para US[*k*-1][2] a la energía calculada para cada uno de US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4], etc. La unidad de reordenación 34 puede entonces descartar uno o más de los segundos vectores US[*k*-1] 33 de la segunda trama de audio anterior (en el tiempo). Para ilustrar, considere la siguiente Tabla 2 que muestra los segundos vectores US[*k*-1] restantes 33:

Tabla 2

Vector bajo consideración	Restante bajo consideración
US[*k*-1][1]	US[*k*][1], US[*k*][2]
US[*k*-1][2]	US[*k*][1], US[*k*][2]
US[*k*-1][3]	US[*k*][3], US[*k*][4]
US[*k*-1][4]	US[*k*][3], US[*k*][4]

[0087] En la Tabla 2 anterior, la unidad de reordenación 34 puede determinar, en base a la comparación de energía, que la energía calculada para US[*k*-1][1] es similar a la energía calculada para cada uno de US[*k*][1] y US[*k*][2], la energía calculada para US[*k*-1][2] es similar a la energía calculada para cada uno de US[*k*][1] y US[*k*][2], la energía calculada para US[*k*-1][3] es similar a la energía calculada para cada uno de US[*k*][3] y US[*k*][4], y la energía calculada para US[*k*-1][4] es similar a la energía calculada para cada uno de US[*k*][3] y US[*k*][4]. En algunos ejemplos, la unidad de reordenación 34 puede realizar un análisis de energía adicional para identificar una similitud entre cada uno de los primeros vectores de la matriz US[*k*] 33 y cada uno de los segundos vectores de la matriz US[*k*-1] 33.

[0088] En otros ejemplos, la unidad de reordenación 32 puede reordenar los vectores basándose en los parámetros actuales 37 y los parámetros anteriores 39 relativos a la correlación cruzada. En estos ejemplos, haciendo referencia de nuevo a la Tabla 2 anterior, la unidad de reordenación 34 puede determinar la siguiente correlación ejemplar expresada en la Tabla 3 basándose en estos parámetros de correlación cruzada:

Tabla 3

Vector bajo consideración	Correlaciona con
US[*k*-1][1]	US[*k*][2]
US[*k*-1][2]	US[*k*][1]
US[*k*-1][3]	US[*k*][3]
US[*k*-1][4]	US[*k*][4]

[0089] A partir de la Tabla 3 anterior, la unidad de reordenación 34 determina, por ejemplo, que el vector US[*k*-1][1] correlaciona con el vector US[*k*][2] posicionado de forma diferente, el vector US[*k*-1][2] correlaciona con el vector US[*k*][1] posicionado de forma diferente, el vector US[*k*-1][3] correlaciona con el vector US[*k*][3] posicionado de forma similar y el US[*k*-1][4] correlaciona con el vector US[*k*][4] posicionado de forma similar. En otras palabras, la unidad de reordenación 34 determina lo que puede denominarse información de reordenación que describe cómo reordenar los primeros vectores de la matriz US[*k*] 33 de manera que el vector US[*k*][2] se reposiciona en la primera fila de los primeros vectores de la matriz US[*k*] 33 y el vector US[*k*][1] se reposicionan en la segunda fila de los primeros vectores US[*k*] 33. La unidad de reordenación 34 puede entonces reordenar los primeros vectores de la matriz US[*k*] 33 basándose en esta información de reordenación para generar la matriz US[*k*] reordenada 33'.

[0090] Además, la unidad de reordenación 34 puede, aunque no se muestra en el ejemplo de la Fig. 4, proporcionar esta información de reordenación al dispositivo de generación de flujo de bits 42, que puede generar el flujo de bits 21 para incluir esta información de reordenación de modo que el dispositivo de descodificación de audio, tal como el dispositivo de descodificación de audio 24 mostrado en el ejemplo de las Figs. 3 y 5, puede determinar cómo reordenar los vectores reordenados de la matriz US[*k*] 33' para recuperar los vectores de la matriz US[*k*] 33.

[0091] Aunque se ha descrito anteriormente que realizan un proceso de dos pasos que implica un análisis basado primero en parámetros específicos de energía y luego en parámetros de correlación cruzada, la unidad de reordenación 32 solo puede realizar este análisis solamente con respecto a parámetros de energía para determinar la información de reordenación, realizar este análisis solo con respecto a los parámetros de correlación cruzada para determinar la información de reordenación, o realizar el análisis con respecto tanto a los parámetros de energía como a los parámetros de correlación cruzada de la manera descrita anteriormente. Además, las técnicas pueden emplear otros tipos de procesos para determinar la correlación que no implican realizar una o ambas de una comparación de energía y/o una correlación cruzada. Por consiguiente, las técnicas no deben estar limitadas a este respecto a los ejemplos expuestos anteriormente. Además, pueden usarse también otros parámetros obtenidos a partir de la unidad de cálculo de parámetros 32 (tales como los parámetros de posición espaciales obtenidos a partir de los vectores V o la correlación de los vectores en V[*k*] y V[*k*-1]) (simultáneamente/conjuntamente o secuencialmente) con la energía y los parámetros de correlación cruzada obtenidos de US[*k*] y US[*k*-1] para determinar el orden correcto de los vectores en US.

[0092] Como ejemplo de utilización de la correlación de los vectores de la matriz V, la unidad de cálculo de parámetros 34 puede determinar que los vectores de la matriz V[*k*] 35 están correlacionados tal como se especifica en la siguiente Tabla 4:

Tabla 4

Vector bajo consideración	Correlaciona con
V[*k*-1][1]	V[*k*][2]
V[*k*-1][2]	V[*k*][1]
V[*k*-1][3]	V[*k*][3]
V[*k*-1][4]	V[*k*][4]

A partir de la Tabla 4 anterior, la unidad de reordenación 34 determina, por ejemplo, que el vector V[*k*-1][1] correlaciona con el vector V[*k*][2] posicionado de forma diferente, el vector V[*k*-1][2] correlaciona con el vector V[*k*] [1] posicionado de forma diferente, el vector V[*k*-1][3] correlaciona con el vector V[*k*][3] posicionado de forma similar y el vector V[*k*-1][4] correlaciona con el vector V[*k*][4] posicionado de forma similar. La unidad de reordenación 34 puede emitir la versión reordenada de los vectores de la matriz V[*k*] 35 como una matriz V[*k*] reordenada 35'.

[0093] En algunos ejemplos, la misma reordenación que se aplica a los vectores de la matriz US se aplica también a los vectores de la matriz V. En otras palabras, cualquier análisis usado para reordenar los vectores V puede usarse junto con cualquier análisis usado para reordenar los vectores US. Para ilustrar un ejemplo en el que la información de reordenación no se determina únicamente con respecto a los parámetros de energía y/o los parámetros de correlación cruzada con respecto a los vectores US[*k*] 35, la unidad de reordenación 34 también puede realizar este análisis con respecto a los vectores V[*k*] 35 en base a los parámetros de correlación cruzada y los parámetros energéticos de una manera similar a la descrita anteriormente con respecto a los vectores V[*k*] 35. Además, aunque los vectores US[*k*] 33 no tienen propiedades direccionales, los vectores V[*k*] 35 pueden proporcionar información relativa a la direccionalidad de los correspondientes vectores US[*k*] 33. En este sentido, la unidad de reordenación 34 puede identificar correlaciones entre los vectores V[*k*] 35 y los vectores V[*k*-1] 35 en base a en un análisis de los correspondientes parámetros de propiedades direccionales. Es decir, en algunos ejemplos, el objeto de audio se mueve dentro de un campo sonoro de una manera continua cuando se mueve o permanece en una posición relativamente estable. Como tal, la unidad de reordenación 34 puede identificar aquellos vectores de la matriz V[*k*] 35 y la matriz V[*k*-1] 35 que exhiben algún movimiento realista físicamente conocido o que permanecen estacionarios dentro del campo sonoro como correlacionados, reordenando los vectores US[*k*] 33 y los vectores V[*k*]

35 basados en esta correlación de propiedades direccionales. En cualquier caso, la unidad de reordenación 34 puede emitir los vectores US[*k*] reordenados 33' y los vectores V[*k*] reordenados 35' a la unidad de selección del primer plano 36.

[0094] Además, las técnicas pueden emplear otros tipos de procesos para determinar el orden correcto que no implican realizar una comparación de energía y/o una correlación cruzada, o ambas. Por consiguiente, las técnicas no deben estar limitadas a este respecto a los ejemplos expuestos anteriormente.

[0095] Aunque se describen anteriormente como reordenando los vectores de la matriz V para reflejar la reordenación de los vectores de la matriz US, en ciertos casos, los V vectores pueden reordenarse de manera diferente que los vectores US, donde se pueden generar elementos sintácticos separados para indicar la reordenación de los vectores US y la reordenación de los vectores V. En algunos casos, los vectores V pueden no ser reordenados y solo los vectores US pueden ser reordenados dado que los vectores V pueden no estar codificados psicoacústicamente.

[0096] Un modo de realización en el que la reordenación de los vectores de la matriz V y los vectores de la matriz US son diferentes, es cuando la intención es intercambiar objetos de audio en el espacio, es decir, alejarlos de la posición original registrada (cuando el campo sonoro subyacente era una grabación natural) o la posición artísticamente deseada (cuando el campo sonoro subyacente es una mezcla artificial de objetos). Por ejemplo, supongamos que hay dos fuentes de audio A y B, A puede ser el sonido de un gato "miau" que emana de la parte "izquierda" del campo sonoro y B puede ser el sonido de un perro "guau" que emana de la parte "derecha" del campo sonoro. Cuando la reordenación de V y US es diferente, la posición de las dos fuentes de sonido es intercambiada. Después del intercambio, A (el "miau") emana de la parte derecha del campo sonoro, y B ("el guau") emana de la parte izquierda del campo sonoro.

[0097] 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 objetivo 41. La unidad de análisis del campo sonoro 44 puede, en base a este análisis y/o en una velocidad de bits objetivo recibida 41, determinar el número total de instanciaciones del codificador psicoacústico (que puede ser una función del número total de canales ambientales o de fondo (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 numHOATransportChannels. La unidad de análisis del campo sonoro 44 también puede determinar, de nuevo para alcanzar potencialmente la velocidad de bits objetivo 41, el número total de canales de primer plano (nFG) 45, el orden mínimo del campo sonoro de fondo (o, en otras palabras, ambiente) (N_{BG} o, alternativamente, MinAmbHoaOrder), el número correspondiente de canales reales representativos del orden mínimo del campo sonoro de fondo ($nBGa = (MinAmbHoaOrder+1)^2$), y los índices (i) de los canales adicionales de BG HOA a enviar (que pueden ser colectivamente indicados como información de canal de fondo 43 en el ejemplo de la Fig. 4). La información de canal de fondo 42 también puede denominarse información de canal ambiental 43. Cada uno de los canales que permanece de numHOATransportChannels - nBGa, puede ser un "canal de fondo/ambiental adicional", un "canal predominante basado en un vector activo", una "señal predominante basada en la dirección activa" o "completamente inactivo". En un modo de realización, estos tipos de canales se pueden indicar (como un "ChannelType") un elemento de sintaxis con dos bits (por ejemplo, 00: canal de fondo adicional, 01: señal predominante basada en vector, 10: señal inactiva, 11: señal basada la dirección). El número total de señales de fondo o de ambiente, nBGa, puede ser dado por $(MinAmbHoaOrder +1)^2$ + el número de veces que el índice 00 (en el ejemplo anterior) aparece como un tipo de canal en el flujo de bits para esa trama.

[0098] En cualquier caso, la unidad de análisis del campo sonoro 44 puede seleccionar el número de canales de fondo (o, en otras palabras, ambientales) y el número de canales de primer plano (o, en otras palabras, predominantes) en base a la velocidad de bits objetivo 41, seleccionado más canales de fondo y/o de primer plano cuando la velocidad de bits objetivo 41 es relativamente mayor (por ejemplo, cuando la velocidad de bits objetivo 41 es igual o mayor que 512 Kbps). En un modo de realización, el numHOATransportChannels se puede establecer en 8 mientras que el MinAmbHoaOrder puede ser puesto a 1 en la sección de encabezamiento del flujo de bits (que se describe con más detalle con respecto a las Figs. 10-10O (ii)). En este escenario, en cada trama se pueden dedicar cuatro canales para representar la parte de fondo o ambiente del campo sonoro, mientras que los otros 4 canales pueden, en una base trama a trama, variar según el tipo de canal, por ejemplo, un canal de fondo/ambiente 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.

[0099] En algunos casos, el número total de señales predominantes basadas en vectores para una trama, puede darse con el número de veces que el índice ChannelType es 01, en el flujo de bits de dicha trama, en el ejemplo anterior. En la realización anterior, para cada canal de fondo/ambiente adicional (por ejemplo, correspondiente a un ChannelType de 00), se puede representar en ese canal una información correspondiente de los posibles coeficientes HOA (más allá de los cuatro primeros). Esta información, para el contenido HOA de cuarto orden, puede ser un índice para indicar entre 5-25 (los primeros cuatro 1-4 se pueden enviar todo el tiempo cuando minAmbHoaOrder se establece en 1, por lo tanto, solo es necesario indicar uno entre 5-25). Esta información podría

ser enviada utilizando un elemento de sintaxis de 5 bits (para el contenido de 4º orden), que puede denominarse "CodedAmbCoeffIdx".

[0100] En una segunda forma de realización, todas las señales de primer plano/predominantes son señales basadas en vectores. En esta segunda forma de realización, el número total de señales en primer plano / predominantes puede darse con nFG = numHOATransportChannels - [(MinAmbHoaOrder +1)^2 + el número de veces del índice 00].

[0101] La unidad de análisis del campo sonoro 44 envía la información de canal de fondo 43 y los coeficientes HOA 11 a la unidad de selección de fondo 46, la información de canal de fondo 43 a la unidad de reducción de coeficientes 46 y la unidad de generación del flujo de bits 42 y el nFG 45 a una unidad de selección de primer plano 36.

[0102] En algunos ejemplos, la unidad de análisis del campo sonoro 44 puede seleccionar, basándose en un análisis de los vectores de la matriz US[*k*] 33 y de la velocidad de bits objetivo 41, un número variable de nFG de estos componentes que tiene el valor máximo. En otras palabras, la unidad de análisis del campo sonoro 44 puede determinar un valor para una variable A (que puede ser similar o sustancialmente similar a N_{BG}), que separa dos subespacios, analizando la pendiente de la curva creada por los valores diagonales descendentes de los vectores de la matriz S[*k*] 33, donde los grandes valores singulares representan sonidos de primer plano o distintos y los valores singulares bajos singular representan los componentes de fondo del campo sonoro. Es decir, la variable A puede segmentar el campo sonoro global en un subespacio de primer plano y un subespacio de fondo.

[0103] En algunos ejemplos, la unidad de análisis del campo sonoro 44 puede usar una primera y una segunda derivada de la curva de valor singular. La unidad de análisis del campo sonoro 44 puede también limitar el valor para la variable A a estar entre uno y cinco. Como otro ejemplo, la unidad de análisis del campo sonoro 44 puede limitar el valor de la variable A a estar entre uno y $(N+1)^2$. Alternativamente, la unidad de análisis del campo sonoro 44 puede predefinir el valor para la variable A, por ejemplo a un valor de cuatro. En cualquier caso, basándose en el valor de A, la unidad de análisis del campo sonoro 44 determina el número total de canales de primer plano (nFG) 45, el orden del campo sonoro de fondo (N_{BG}) y el número (nBGa) y los índices (i) de los canales BG HOA adicionales a enviar.

[0104] Además, la unidad de análisis del campo sonoro 44 puede determinar la energía de los vectores de la matriz V[*k*] 35 sobre una base por vector. La unidad de análisis del campo sonoro 44 puede determinar la energía para cada uno de los vectores de la matriz V[*k*] 35 e identificar aquellos que tienen una alta energía como componentes de primer plano.

[0105] Además, la unidad de análisis del campo sonoro 44 puede realizar otros análisis diversos con respecto a los coeficientes HOA 11, que incluyen un análisis de energía espacial, un análisis de enmascaramiento espacial, un análisis de difusión u otras formas de análisis auditivo. La unidad de análisis del campo sonoro 44 puede realizar el análisis de energía espacial a través de la transformación de los coeficientes HOA 11 en el dominio espacial e identificar áreas de alta energía representativas de componentes direccionales del campo sonoro que deben conservarse. La unidad de análisis del campo sonoro 44 puede realizar el análisis de enmascaramiento espacial perceptivo de una manera similar a la del análisis de energía espacial, excepto que la unidad de análisis del campo sonoro 44 puede identificar áreas espaciales que están enmascaradas por sonidos de energía superior espacialmente próximos. La unidad de análisis del campo sonoro 44 puede entonces, en base a áreas perceptualmente enmascaradas, identificar en menor número de componentes de primer plano en algunos casos. La unidad de análisis del campo sonoro 44 puede realizar además un análisis de difusión con respecto a los coeficientes HOA 11 para identificar áreas de energía difusa que pueden representar componentes de fondo del campo sonoro.

[0106] La unidad de análisis del campo sonoro 44 también puede representar una unidad configurada para determinar la notabilidad, la distinción o el predominio de los datos de audio que representan un campo sonoro, utilizando información basada en la direccionalidad asociada con los datos de audio. Mientras que las determinaciones basadas en energía pueden mejorar la reproducción de un campo sonoro descompuesto por SVD para identificar componentes de audio distintos del campo sonoro, las determinaciones basadas en energía también pueden hacer que un dispositivo identifique erróneamente componentes de audio de fondo como componentes de audio distintos, en aquellos casos en los que los componentes de audio de fondo exhiben un alto nivel de energía. Es decir, una separación de componentes de audio distintos y de fondo, exclusivamente basada en energía, puede no ser robusta, ya que los componentes de audio de fondo energéticos (por ejemplo, más altos) pueden identificarse incorrectamente como componentes de audio distintos. Para distinguir más fuertemente entre componentes de audio distintos y de fondo del campo sonoro, diversos aspectos de las técnicas descritas en esta divulgación pueden permitir que la unidad de análisis del campo sonoro 44 realice un análisis basado en direccionalidad de los coeficientes HOA 11 para separar los componentes de audio de primer plano y ambientales de las versiones descompactadas de los coeficientes HOA 11.

[0107] A este respecto, la unidad de análisis de campos sonoros 44 puede representar una unidad configurada o, de algún modo, operable para identificar elementos distintos (o de primer plano) a partir de elementos de fondo incluidos en uno o más de los vectores de la matriz US[*k*] 33 y los vectores de la matriz V[*k*] 35. De acuerdo con algunas técnicas basadas en SVD, los componentes más energéticos (por ejemplo, los primeros vectores de una o más de la matriz US[*k*] 33 y la matriz V[*k*] 35 o vectores derivados de la misma) pueden tratarse como componentes distintos. Sin embargo, los componentes más enérgicos (que están representados por vectores) de uno o más de los vectores de la matriz US[*k*] 33 y los vectores de la matriz V[*k*] 35 pueden no representar en todos los escenarios, los componentes/señales que son los más direccionales.

[0108] La unidad de análisis del campo sonoro 44 puede implementar uno o más aspectos de las técnicas descritas aquí para identificar elementos de primer plano/directo/predominantes basados en la direccionalidad de los vectores de uno o más de los vectores de la matriz US[*k*] 33 y los vectores de la matriz V[*k*] 35 o los vectores derivados de la misma. En algunos ejemplos, la unidad de análisis del campo sonoro 44 puede identificar o seleccionar como componentes de audio distintos (donde los componentes también pueden denominarse "objetos"), uno o más vectores basados tanto en energía como en la direccionalidad de los vectores. Por ejemplo, la unidad de análisis del campo sonoro 44 puede identificar aquellos vectores de uno o más de los vectores de la matriz US[*k*] 33 y los vectores de la matriz V[*k*] 35 (o vectores derivados de la misma) que muestran tanto una alta energía como una alta direccionalidad (por ejemplo, representada como un cociente de direccionalidad) como componentes de audio distintos. Como resultado, si la unidad de análisis del campo sonoro 44 determina que un vector particular es relativamente menos direccional cuando se compara con otros vectores de uno o más de los vectores de la matriz US[*k*] 33 y los vectores de la matriz V[*k*] 35 (o vectores derivados de la misma), entonces independientemente del nivel de energía asociado con el vector particular, la unidad de análisis del campo sonoro 44 puede determinar que el vector particular representa componentes de audio de fondo (o ambiente) del campo sonoro representados por los coeficientes HOA 11.

[0109] En algunos ejemplos, la unidad de análisis del campo sonoro 44 puede identificar objetos de audio distintos (que, como se ha indicado anteriormente, también pueden denominarse "componentes") basándose en la direccionalidad, realizando las siguientes operaciones. La unidad de análisis del campo sonoro 44 puede multiplicar (por ejemplo, usando uno o más procesos de multiplicación de matrices) los vectores de la matriz S[*k*] (que pueden derivarse a partir de los vectores US[*k*] 33 o, aunque no se muestra en el ejemplo de la Fig. 4, emitidos por separado mediante la unidad LIT 30) con los vectores de la matriz V[*k*] 35. Mediante la multiplicación de la matriz V[*k*] 35 y de los vectores S[*k*], la unidad de análisis del campo sonoro 44 puede obtener la matriz VS[*k*]. Además, la unidad de análisis del campo sonoro 44 puede ser cuadrada (es decir, elevar exponencialmente por una potencia de dos) al menos algunas de las entradas de cada uno de los vectores de la matriz VS[*k*]. En algunos casos, la unidad de análisis del campo sonoro 44 puede sumar aquellas entradas al cuadrado de cada vector que están asociadas con un orden mayor que 1.

[0110] Por ejemplo, si cada vector de la matriz VS[*k*], que incluye 25 entradas, la unidad de análisis del campo sonoro 44 puede, con respecto a cada vector, elevar al cuadrado las entradas de cada vector comenzando por la entrada quinta y terminando por la entrada vigésimo quinta, sumando las entradas al cuadrado para determinar un cociente de direccionalidad (o un indicador de direccionalidad). Cada operación de suma puede dar como resultado un cociente de direccionalidad para un vector correspondiente. En este ejemplo, la unidad de análisis del campo sonoro 44 puede determinar que aquellas entradas de cada fila que están asociadas con un orden menor o igual a 1, concretamente, de la primera a la cuarta entradas, se dirigen más generalmente a la cantidad de energía y menos a la direccionalidad de esas entradas. Es decir, los ambisónicos de orden inferior asociados a un orden de cero o uno corresponden a funciones de base esférica que, como se ilustra en la Fig. 1 y la Fig. 2, no proporcionan mucho en términos de la dirección de la onda de presión, sino que proporcionan algo de volumen (que es representativo de la energía).

[0111] Las operaciones descritas en el ejemplo anterior también se pueden expresar de acuerdo con el siguiente pseudocódigo. El pseudocódigo siguiente incluye anotaciones, en forma de declaraciones de comentarios que se incluyen en las instancias consecutivas de las cadenas de caracteres "/*" y "*/" (sin comillas).

```
[U, S, V] = svd(audioframe, 'ecom');

VS = V*S;

/* La línea siguiente se dirige a analizar cada fila independientemente y a sumar los valores de la primera fila (como un ejemplo) de la quinta entrada a la entrada vigésimo quinta para determinar un cociente de direccionalidad o métrica de direccionalidad para un vector correspondiente. Elevar al cuadrado las entradas antes de sumar. Las entradas en cada fila que están asociadas con un orden mayor que 1 están asociadas con ambisónicos de orden superior y, por lo tanto, tienen más probabilidades de ser direccionales. */

sumVS = sum(VS(5:end,:).^2,1);
```

/* La línea siguiente se dirige a ordenar la suma de los cuadrados de la matriz VS generada, y a seleccionar un conjunto de los valores más grandes (por ejemplo, tres o cuatro de los valores más grandes) */

[~,idxVS] = sort(sumVS,'descend');

U = U(:,idxVS);

V = V(:,idxVS);

S = S(idxVS,idxVS);

[0112] En otras palabras, de acuerdo con el pseudocódigo anterior, la unidad de análisis del campo sonoro 44 puede seleccionar entradas de cada vector de la matriz VS[*k*] descompuesta de las de los coeficientes HOA 11 correspondientes a una función de base esférica que tiene un orden mayor que uno. La unidad de análisis del campo sonoro 44 puede entonces elevar al cuadrado estas entradas para cada vector de la matriz VS[*k*], sumando las entradas elevadas al cuadrado para identificar, calcular o, de algún modo, determinar una métrica o cociente de direccionalidad para cada vector de la matriz VS[*k*]. A continuación, la unidad de análisis del campo sonoro 44 puede ordenar los vectores de la matriz VS[*k*] en base a las respectivas métricas de direccionalidad de cada uno de los vectores. La unidad de análisis del campo sonoro 44 puede clasificar estos vectores en un orden descendente de métricas de direccionalidad, de tal manera que los vectores con la direccionalidad correspondiente más alta son los primeros y aquellos vectores con la direccionalidad correspondiente más baja son últimos. La unidad de análisis del campo sonoro 44 puede entonces seleccionar el subconjunto diferente de cero de los vectores que tienen la métrica de direccionalidad relativa más alta.

[0113] La unidad de análisis del campo sonoro 44 puede realizar cualquier combinación de los análisis anteriores para determinar el número total de instanciaciones del codificador psicoacústico (que puede ser una función del número total de canales ambientales o de fondo (BG_{TOT}) y el número de canales de primer plano. La unidad de análisis del campo sonoro 44 puede, en base a cualquier combinación de los análisis anteriores, determinar el número total de canales de primer plano (nFG) 45, el orden del campo sonoro de fondo (N_{BG}) y el número (nBGa) y los índices (i) de los canales BG HOA adicionales para enviar (que pueden indicarse colectivamente como información de canal de fondo 43 en el ejemplo de la Fig. 4).

[0114] En algunos ejemplos, la unidad de análisis del campo sonoro 44 puede realizar este análisis cada M-muestras, las cuales pueden ser reexpresadas en una base de trama por trama. A este respecto, el valor de A puede variar de trama en trama. Un ejemplo de un flujo de bits donde se toma la decisión cada M-muestras se muestra en las Figs. 10-10O(ii). En otros ejemplos, la unidad de análisis del campo sonoro 44 puede realizar este análisis más de una vez por trama, analizando dos o más partes de la trama. Por consiguiente, las técnicas no deben estar limitadas a este respecto a los ejemplos descritos en esta divulgación.

[0115] La unidad de selección de fondo 48 puede representar una unidad configurada para determinar los coeficientes HOA de fondo o ambientales 47 en base a la información de canal de fondo (por ejemplo, el campo sonoro de fondo (N_{BG}) y el número (nBGa) y los índices (i) de los canales BG HOA adicionales para enviar). Por ejemplo, cuando N_{BG} es igual a uno, la unidad de selección de fondo 48 puede seleccionar los coeficientes HOA 11 para cada muestra de la trama de audio que tiene un orden igual o inferior a uno. La unidad de selección de fondo 48 puede, en este ejemplo, seleccionar entonces los coeficientes HOA 11 que tienen un índice identificado por uno de los índices (i) como coeficientes BG HOA adicionales, en donde el nBGa se proporciona a la unidad de generación del flujo de bits 42 que se especificará en el flujo de bits 21 para permitir que el dispositivo de descodificación de audio, tal como el dispositivo de descodificación de audio 24 que se muestra en el ejemplo de la Fig. 3, para analizar los coeficientes BG HOA 47 del flujo de bits 21. La unidad de selección de fondo 48 puede emitir entonces los coeficientes HOA ambientales 47 a la unidad de compensación de energía 38. Los coeficientes HOA ambientales 47 pueden tener D dimensiones: $M \times [(N_{BG}+1)^2+nBGa]$.

[0116] La unidad de selección del primer plano 36 puede representar una unidad configurada para seleccionar aquellos vectores de la matriz US[*k*] reordenada 33' y la matriz V[*k*] reordenada 35' que representan componentes de primer plano o distintos del campo sonoro en base a nFG 45 (que pueden representar uno o más índices que identifican estos vectores de primer plano). La unidad de selección del primer plano 36 puede emitir señales nFG 49 (que se pueden designar como $US[k]_{1, ..., nFG}$ 49, $FG_{1, ..., nFG}[k]$ 49, o $X_{PS}^{(1..nFG)}$ (k) 49) a la unidad codificadora de audio psicoacústico 40, en la que las señales nFG 49 pueden tener dimensiones D:*Mx* nFG y cada una representa objetos monoaudio. 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 aquellos componentes de la matriz V[*k*] reordenada 35 correspondiente a los componentes de primer plano se puede indicar como matriz V[*k*] de primer plano 51 $_k$ (que se puede designar matemáticamente como $V_{1, ..., nFC}[k]$) con dimensiones D: $(N+1)^2$ x nFG.

[0117] 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 y compensar la pérdida de energía

debido a la eliminación de varios de los canales HOA por la unidad de selección de fondo 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 señales nFG 49, los vectores V[*k*] 51_k y los coeficientes HOA ambientales 47 y luego realizar una compensación de energía basada en este análisis de energía para generar los coeficientes HOA ambientales compensados de energía 47 '. La unidad de compensación de energía 38 puede emitir los coeficientes HOA ambientales compensados de energía 47' a la unidad codificadora de audio psicoacústico 40.

[0118] Efectivamente, la unidad de compensación de energía 38 puede utilizarse para compensar posibles reducciones en la energía global de los componentes de sonido de fondo del campo sonoro causadas por la reducción del orden de los componentes ambientales del campo sonoro descritos por los coeficientes HOA 11 y generar los coeficientes HOA ambientales de orden reducido 47 (que, en algunos ejemplos, tienen un orden menor que N en términos de coeficientes solo incluidos que corresponden a funciones de base esférica que tienen los siguientes órdenes/subórdenes: $[(N_{BG}+1)^2+nBGa]$). En algunos ejemplos, la unidad de compensación de energía 38 compensa esta pérdida de energía determinando una ganancia de compensación en forma de valores de amplificación que se aplican a cada una de las columnas $[(N_{BG}+1)^2 + nBGa]$ de los coeficientes HOA ambientales 47 con el fin de aumentar la energía cuadrática media (RMS) de los coeficientes HOA ambientales 47 para igualar o, al menos aproximar, aproximadamente el RMS de los coeficientes HOA 11 (determinado a través del análisis de energía agregado de uno o más de las matrices US[*k*] reordenadas 33', la matriz V[*k*] reordenada 35', las señales nFG 49, los vectores V[*k*] de primer plano 51_k y los coeficientes HOA ambientales de orden reducido 47) antes de emitir los coeficientes HOA ambientales 47 a la unidad codificadora de audio psicoacústico 40.

[0119] En algunos casos, la unidad de compensación de energía 38 puede identificar el RMS para cada fila y/o columna de una o más entre la matriz US[*k*] reordenada 33' y la matriz V[*k*] reordenada 35'. La unidad de compensación de energía 38 también puede identificar el RMS para cada fila y/o columna de uno o más de los canales de primer plano seleccionados, los cuales pueden incluir las señales nFG 49 y los vectores V[*k*] de primer plano 51_k y los coeficientes HOA ambientales de orden reducido 47. El RMS para cada fila y/o columna de una o más entre la matriz US[*k*] reordenada 33' y la matriz V[*k*] reordenada 35' puede almacenarse en un vector denotado $RMS_{COMPLETO}$, mientras que el RMS para cada fila y/o columna de una o más de las señales nFG 49, los vectores V[*k*] de primer plano 51_k y los coeficientes HOA ambientales de orden reducido 47 pueden almacenarse en un vector denotado $RMS_{REDUCIDO}$. La unidad de compensación de energía 38 puede entonces calcular un vector de valores de amplificación Z de acuerdo con la siguiente ecuación: $Z = RMS_{COMPLETO}/RMS_{REDUCIDO}$. La unidad de compensación de energía 38 puede entonces aplicar este vector de valores de amplificación Z, o varias partes del mismo, a uno o más entre las señales nFG 49, los vectores V[*k*] de primer plano 51_k y los coeficientes HOA ambientales de orden reducido 47. En algunos casos, el vector de valor de amplificación Z se aplica solo a los coeficientes HOA ambientales de orden reducido 47 mediante la siguiente ecuación $HOA_{BG-RED}' = HOA_{BG-RED}Z^T$, donde HOA_{BG-RED} indica los coeficientes HOA ambiental de orden reducido 47, HOA_{BG-RED}' indica los coeficientes HOA ambientales reducidos y compensados en energía 47' y Z^T representa la transposición del vector Z.

[0120] En algunos ejemplos, para determinar cada RMS de las respectivas filas y/o columnas de una o más entre la matriz US[*k*] reordenada 33', la matriz V[*k*] reordenada 35', las señales nFG 49, los vectores V[*k*] de primer plano 51_k, y los coeficientes HOA ambientales de orden reducido 47, la unidad de compensación de energía 38 puede aplicar primero un coeficiente de armónicos esféricos de referencia (SHC) a las columnas. La aplicación del reproductor SHC de referencia con la unidad de compensación de energía 38 permite la determinación de RMS en el dominio SHC para determinar la energía del campo sonoro global descrito por cada fila y/o columna de la trama representado por filas y/o columnas de uno o más entre la matriz US[*k*] reordenada 33', la matriz V[*k*] reordenada 35', las señales nFG 49, los vectores V[*3*] de primer plano 51_k y los coeficientes HOA ambientales de orden reducido 47, como se describe en más detalle a continuación.

[0121] 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 trama k y los vectores de primer plano V[*k*-1] 51_{k-1} para la trama anterior (de ahí el 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 señales nFG 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 y generar unas señales nFG interpoladas 49'. La unidad de interpolación espacio-temporal 50 también puede emitir los de 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 descodificación de audio, tal como el dispositivo de descodificación de audio 24 puede 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 descodificador (para crear los vectores interpolados V[*k*]) se pueden usar versiones cuantizadas/descuantizadas de éstos en el codificador y el descodificador.

[0122] A este respecto, la unidad de interpolación espacio-temporal 50 puede representar una unidad que interpola una primera parte de una primera trama de audio a partir de algunas otras partes de la primera trama de

audio y una segunda trama de audio temporalmente subsiguiente o anterior. En algunos ejemplos, las partes pueden indicarse como subtramas, donde la interpolación realizada con respecto a las subtramas se describe con más detalle a continuación con respecto a las Figs. 45-46E. En otros ejemplos, la unidad de interpolación espacio-temporal 50 puede funcionar con respecto a algún último número de muestras de la trama anterior y algún primer número de muestras de la siguiente trama, como se describe con más detalle con respecto a las Figs. 37-39. La unidad de interpolación espacio-temporal 50 puede, al realizar esta interpolación, reducir el número de muestras de los vectores V[*k*] de primer plano 51_k que se requieren para ser especificados en el flujo de bits 21, ya que solo aquellos de los vectores V[*k*] de primer plano 51k que se utilizan para generar los vectores V[*k*] interpolados representan un subconjunto de los vectores V[*k*] de primer plano 51_k. Es decir, con el fin de potencialmente hacer más eficiente la compresión de los coeficientes HOA 11 (reduciendo el número de los vectores V[*k*] de primer plano 51_k que se especifican en el flujo de bits 21), diversos aspectos de las técnicas descritas en esta divulgación pueden proporcionar la interpolación de una o más partes de la primera trama de audio, donde cada una de las partes puede representar versiones descompuestas de los coeficientes HOA 11.

[0123] La interpolación espacio-temporal puede resultar en varias ventajas. En primer lugar, las señales nFG 49 pueden no ser continuas de una trama a otra debido a la naturaleza en bloque en la que se realiza la SVD u otro LIT. Dicho de otro modo, dado que la unidad LIT 30 aplica la SVD en base a cada fotograma, pueden existir ciertas discontinuidades en los coeficientes HOA transformados resultantes como evidencia, por ejemplo, la naturaleza no ordenada de la matriz US[*k*] 33 y la matriz V[*K*] 35. Mediante la realización de esta interpolación, la discontinuidad puede reducirse dado que la interpolación puede tener un efecto suavizante que potencialmente reduce los artefactos introducidos debido a los límites de la trama (o, en otras palabras, la segmentación de los coeficientes HOA 11 en tramas). Usando los vectores V[*k*] de primer plano 51_k para realizar esta interpolación y después generar las señales nFG interpoladas 49' basadas en los vectores V[*3*] de primer plano 51_k interpolados a partir de los coeficientes HOA reordenados recuperados pueden suavizar al menos algunos efectos debidos a la operación trama por trama, así como debido a la reordenación de las señales nFG 49.

[0124] En funcionamiento, la unidad de interpolación espacio-temporal 50 puede interpolar una o más subtramas de una primera trama de audio desde 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 una o más subtramas.

[0125] 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.

[0126] En otras palabras, el audio 3D basado en armónicos esféricos puede ser una representación paramétrica del campo de presión 3D en términos de funciones de base ortogonales sobre una esfera. Cuanto mayor sea el orden N de la representación, la resolución espacial será potencialmente más alta 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. Estas 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 SVD puede descomponer cada trama de coeficientes en tres matrices U, S y V. En algunos ejemplos, las técnicas pueden manejar algunos de los vectores de la matriz US[*k*] como componentes de primer plano del campo sonoro subyacente. Sin embargo, cuando se manejan de esta manera, estos vectores (en la matriz US[*k*]) son discontinuos de trama en trama, aunque representan el mismo componente de audio distinto. Estas discontinuidades pueden conducir a artefactos significativos cuando los componentes son alimentados a través de codificadores de audio de transformadas.

[0127] Las técnicas descritas en esta divulgación pueden abordar esta discontinuidad. Es decir, las técnicas pueden basarse 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 al eje espacial ortogonal (V[*k*]) que cambia cada trama y, por lo tanto, son discontinuas en sí mismas. Esto es diferente a la descomposición similar, como la Transformada 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 por coincidencia. Las técnicas descritas en esta divulgación pueden permitir que la unidad de interpolación espacio-temporal 50 mantenga la continuidad entre las funciones de base (V[*k*]) de una trama a otra mediante interpolación entre ellas.

[0128] Como se ha indicado anteriormente, la interpolación puede realizarse con respecto a las muestras. Este caso se generaliza en la descripción anterior cuando las subtramas comprenden un solo conjunto de muestras.

Tanto en el caso de la interpolación sobre muestras como sobre subtramas, 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 V*v*(*k*-1), que en un modo de realización podría representar V vectores 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, ..., *T* (donde *T* es la longitud de las muestras sobre las que se realiza la interpolación y sobre la cual los vectores interpolados de salida, $\overline{v(l)}$ son necesarios y también indica que la salida de este proceso produce *l* de estos vectores). Alternativamente, *l* podría indicar subtramas consistentes en múltiples muestras. Cuando, por ejemplo, una trama está dividida en cuatro subtramas, *l* puede comprender valores de 1, 2, 3 y 4 para cada una de las subtramas. El valor de *l* se puede señalar como un campo denominado "CodedSpatialInterpolationTime" a través de un flujo de bits, de modo que la operación de interpolación puede ser replicada en el descodificador. El *w*(*l*) puede comprender valores de los pesos de interpolación. Cuando la interpolación es lineal, *w*(*l*) puede variar linealmente y monotónicamente entre 0 y 1, en función de *l*. En otros casos, *w*(*l*) puede variar entre 0 y 1 de una manera no lineal pero monótona (tal como un cuarto de ciclo de un coseno elevado) en función de /. La función, *w*(*l*), puede ser indexada entre unas pocas posibilidades diferentes de funciones y señalada en el flujo de bits como un campo denominado "SpatialInterpolationMethod" de tal manera que la operación de interpolación idéntica pueda ser replicada por el descodificador. Cuando *w*(*l*) es un valor cercano a 0, la salida, $\overline{v(l)}$ puede estar altamente ponderado o influenciado por *v*(*k*-1). Mientras que w(/) es un valor cercano a 1, asegura que la salida, $\overline{v(l)}$, está altamente ponderada o influenciada por *v*(*k*-1).

[0129] La unidad de reducción de coeficientes 46 puede representar una unidad configurada para realizar una reducción de coeficiente con respecto a los vectores V[*k*] de primer plano restantes 53 basados en la información de canal de fondo 43 para emitir los vectores reducidos de primer plano V[*k*] 55 a la unidad de cuantización 52. Los vectores V[*k*] de primer plano reducidos pueden tener dimensiones D: $[(N+1)^2-(N_{BG}+1)^2-nBGa] \times nFG$.

[0130] La unidad de reducción de coeficientes 46 puede, a este respecto, representar una unidad configurada para reducir el número de coeficientes de los vectores V[*k*] restantes de primer plano 53. En otras palabras, la unidad de reducción de coeficientes 46 puede representar una unidad configurada para eliminar aquellos coeficientes de los vectores V[*k*] de primer plano (que forman los vectores V[*k*] de primer plano restantes) que tienen poca o ninguna información direccional. Como se ha descrito anteriormente, en algunos ejemplos, 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 (a través de un proceso que se puede denominar "reducción de coeficientes"). En este ejemplo, se puede proporcionar una mayor flexibilidad para identificar no solo estos coeficientes que corresponden a N_{BG} sino para identificar canales HOA adicionales (que pueden indicarse con la variable TotalOfAddAmbHOAChan) a partir del conjunto de $[(N_{BG}+1)^2+1, (N+1, (N+1)^2]$. 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 TotalOfAddAmbHOAChan, que se puede denominar colectivamente como la información de canal de fondo 43. La unidad de reducción de coeficientes 46 puede entonces eliminar aquellos coeficientes correspondientes a $(N_{BG}+1)^2$ y el TotalOfAddAmbHOAChan de los vectores V[*k*] restantes de primer plano 53 para generar una matriz V[*k*] dimensional menor 55 de tamaño $((N+1)^2-(BG_{TOT}) \times nFG$, que también se puede denominar los vectores V[*k*] de primer plano reducidos 55.

[0131] 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, que emiten estos 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 de este ejemplo. A efectos de ejemplo, se supone que los vectores V[*k*] de primer plano reducidos 55 incluyen dos vectores de fila que tienen, como resultado de la reducción del coeficiente, 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 entrópica, la unidad de cuantización 52 puede realizar cualquier forma de cuantización que resulte de la compresión de los vectores V[*k*] de primer plano reducidos 55.

[0132] 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. Este esquema de compresión puede implicar cualquier esquema de compresión concebible para comprimir elementos de un vector o datos en general, y no debe 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 a representaciones enteras de cada elemento de los vectores reducidos V[*k*] de primer plano 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.

[0133] En algunos ejemplos, varios entre uno o más procesos de este esquema de compresión pueden ser controlados dinámicamente por parámetros para conseguir, o casi conseguir, por ejemplo, una velocidad de bits objetivo 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 el mismo modo de codificación (definido por varios submodos).

[0134] En cualquier caso, tal como se ha indicado anteriormente, este esquema de codificación puede implicar en primer lugar la transformación de las representaciones en coma flotante de cada elemento (que es, en algunos ejemplos, un número de coma flotante de 32 bits) de cada uno de los vectores de fondo reducidos V[*k*] 55 en una representación entera de 16 bits. La unidad de cuantización 52 puede realizar esta transformación de coma flotante a entero multiplicando cada elemento de uno dado de los vectores reducidos V[*k*] de primer plano 55 por 2^{15}, que es, en algunos ejemplos, realizado por un desplazamiento a la derecha de 15.

[0135] La unidad de cuantización 52 puede entonces realizar una cuantización uniforme con respecto a todos los elementos del uno dado de los reducidos vectores V[*k*] de primer plano 55. La unidad de cuantización 52 puede identificar un tamaño de paso de cuantización basado en un valor, que se puede indicar como un parámetro de *nbits*. La unidad de cuantización 52 puede determinar dinámicamente este parámetro de *nbits* basado en la velocidad de bits objetivo 41. La unidad de cuantización 52 puede determinar el tamaño del paso de cuantización como una función de este parámetro de *nbits*. Por ejemplo, la unidad de cuantización 52 puede determinar el tamaño del paso de cuantización (denominado "delta" o "Δ" en esta divulgación) como igual a $2^{16\text{-nbits}}$. En este ejemplo, si *nbits* es igual a seis, delta es igual a 2^{10} y hay 2^6 niveles de cuantización. A este respecto, para un elemento vectorial v, el elemento vectorial cuantizado v_q es igual a [*v*/Δ] y $-2^{nbits-1} < v_q < 2^{nbits-1}$.

[0136] La unidad de cuantización 52 puede realizar entonces la categorización y la codificación residual de los elementos vectoriales cuantizados. Por ejemplo, la unidad de cuantización 52 puede, para un elemento vectorial cuantizado dado v_q, identificar una categoría (mediante la determinación de un identificador de categoría *cid*) a la que este elemento se corresponde usando la siguiente ecuación:

$$cid = \begin{cases} 0, & if\ v_q = 0 \\ \lfloor log_2 |v_q| \rfloor + 1, & if\ v_q \neq 0 \end{cases}$$

La unidad de cuantización 52 puede entonces codificar con Huffman este índice de categoría *cid*, al mismo tiempo que también identifica un bit de signo que indica si v_q es un valor positivo o un valor negativo. La unidad de cuantización 52 puede identificar a continuación un residuo en esta categoría. Por ejemplo, la unidad de cuantización 52 puede determinar este residuo de acuerdo con la siguiente ecuación:

$$residual = |v_q| - 2^{cid-1}$$

La unidad de cuantización 52 puede entonces bloquear el código de este residuo con *cid* -1 bits.

[0137] El siguiente ejemplo ilustra un ejemplo simplificado de este proceso de categorización y codificación residual. Primero, suponga que *nbits* es igual a seis, de modo que v_q ε [-31,31]. A continuación, asuma lo siguiente:

cid	vq	Código Huffman para cid
0	0	"1"
1	-1, 1	"01"
2	-3, -2, 2,3	"000"
3	-7, -6, -5, -4, 4, 5, 6, 7	"0010"
4	-15, -14, ..., -8, 8, ..., 14, 15	"00110"
5	-31, -30, ..., -16, 16, ..., 30, 31	"00111"

Además, asuma lo siguiente:

cid	Código de bloque para residuales
0	N/A
1	0, 1
2	01, 00, 10, 11
3	011, 010, 001, 000, 100, 101, 110, 111
4	0111, 0110..., 0000, 1000, ..., 1110, 1111
5	01111, ..., 00000, 10000, ..., 11111

Por lo tanto, para un v_q = [6, -17, 0, 0, 3], se puede determinar lo siguiente:

» cid = 3, 5, 0, 0, 2
» Signo = 1, 0, x, x, 1
» Residual = 2, 1, x, x, 1
» Bits para 6 = "0010" + "1" + "10"
» Bits para -17 = "00111" + "0" + "0001"
» Bits para 0 = "0"
» Bits para 0 = "0"
» Bits para 3 = "000" + "1" + "1"
» Total de bits = 7 + 10 + 1 + 1 + 5 = 24
» Promedio de bits = 24/5 = 4,8

[0138] Aunque no se muestra en el ejemplo simplificado anterior, la unidad de cuantización 52 puede seleccionar diferentes libros de código Huffman para diferentes valores de *nbits* cuando codifica el *cid*. En algunos ejemplos, la unidad de cuantización 52 puede proporcionar una tabla de codificación Huffman *diferente* para los valores de nbits 6, ..., 15. Además, la unidad de cuantización 52 puede incluir cinco libros de código Huffman diferentes para cada uno de los *diferentes* valores de nbits que van desde 6, ..., 15 para un total de 50 libros de códigos Huffman. A este respecto, la unidad de cuantización 52 puede incluir una pluralidad de diferentes libros de códigos Huffman para acomodar la codificación del *cid* en varios contextos estadísticos diferentes.

[0139] Para ilustrar, la unidad de cuantización 52 puede incluir, para cada uno de los valores de *nbits*, un primer libro de códigos Huffman para codificar los elementos vectoriales de uno a cuatro, un segundo libro de códigos Huffman para codificar los elementos vectoriales de cinco a nueve, un tercer libro de códigos Huffman para codificar los elementos vectoriales de nueve y por encima. Estos primeros tres libros de código Huffman pueden usarse cuando el vector entre los vectores reducidos V[*k*] de primer plano 55 a comprimir no se predice a partir de un correspondiente correlativo temporal posterior de los vectores reducidos V[*k*] reducidos de primer plano 55 y no es representativo de la Información espacial de un objeto de audio sintético (uno definido, por ejemplo, originalmente por un objeto de audio modulado por impulsos codificados (PCM)). La unidad de cuantización 52 puede incluir adicionalmente, para cada uno de los valores de *nbits*, un cuarto libro de códigos Huffman para codificar uno de los vectores reducidos V[*k*] de primer plano 55 cuando se predice este vector entre los vectores reducidos V[*k*] de primer plano 55 se predice a partir de un vector subsiguiente temporal que se corresponde con uno de los vectores reducidos V[*k*] de primer plano 55. La unidad de cuantización 52 puede incluir también, para cada uno de los valores de *nbits*, un quinto libro de códigos Huffman para codificar uno de los vectores reducidos V[*k*] de primer plano 55 cuando este vector entre los vectores reducidos V[*k*] de primer plano 55 es representativo de un objeto de audio sintético. Los diversos libros de código Huffman pueden desarrollarse para cada uno de estos diferentes contextos estadísticos, es decir, el contexto no predicho y no sintético, el contexto predicho y el contexto sintético en este ejemplo.

[0140] La tabla siguiente ilustra la selección de la tabla Huffman y los bits que se especificarán en el flujo de bits para permitir que la unidad de descompresión seleccione la tabla Huffman adecuada:

Modo Pred	Info HT	Tabla HT
0	0	HT5
0	1	HT{1,2,3}
1	0	HT4
1	1	HT5

En la tabla anterior, el modo de predicción ("Modo Pred") indica si se realizó la predicción para el vector actual, mientras que la Tabla Huffman ("Info HT") indica la información adicional del libro de códigos Huffman (o tabla) usada para seleccionar una tabla Huffman de uno a cinco.

[0141] La tabla siguiente ilustra este proceso de selección de tablas Huffman dados diversos contextos o escenarios estadísticos.

	Grabación	**Sintética**
W/O Pred	HT{1,2,3}	HT5
Con Pred	HT4	HT5

En la tabla anterior, la columna "Grabación" indica el contexto de codificación cuando el vector es representativo de un objeto de audio que fue grabado mientras que la columna "sintética" indica un contexto de codificación cuando el vector es representativo de un objeto de audio sintético. La fila "W/O Pred" indica el contexto de codificación cuando la predicción no se realiza con respecto a los elementos vectoriales, mientras que la fila "Con Pred" indica el contexto de codificación cuando se realiza la predicción con respecto a los elementos vectoriales. Como se muestra en esta tabla, la unidad de cuantización 52 selecciona HT {1, 2, 3} cuando el vector es representativo de un objeto de audio grabado y no se realiza la predicción con respecto a los elementos de vector. La unidad de cuantización 52 selecciona HT5 cuando el objeto de audio es representativo de un objeto de audio sintético y no se realiza la predicción con respecto a los elementos de vector. La unidad de cuantización 52 selecciona HT4 cuando el vector es representativo de un objeto de audio grabado y se realiza la predicción con respecto a los elementos vectoriales. La unidad de cuantización 52 selecciona HT5 cuando el objeto de audio es representativo de un objeto de audio sintético y se realiza la predicción con respecto a los elementos de vector.

[0142] A este respecto, la unidad de cuantización 52 puede realizar la cuantización escalar y/o la codificación Huffman antes mencionadas para comprimir los vectores reducidos V[*k*] de primer plano 55, que emiten los vectores codificados V[*k*] de primer plano 57, que se pueden referir como información de canal lateral 57. Esta información de canal lateral 57 puede incluir elementos de sintaxis usados para codificar los vectores [V][*k*] de primer plano restantes. La unidad de cuantización 52 puede emitir la información de canal lateral 57 de una manera similar a la que se muestra en el ejemplo de una de las Figs. 10B y 10C.

[0143] Como se ha indicado anteriormente, la unidad de cuantización 52 puede generar elementos de sintaxis para la información de canal lateral 57. Por ejemplo, la unidad de cuantización 52 puede especificar un elemento de sintaxis en un encabezamiento de una unidad de acceso (que puede incluir uno o más tramas) que indica cuál de la pluralidad de modos de configuración ha sido seleccionado. Aunque se describe que se especifica en una base de unidad de acceso, la unidad de cuantización 52 puede especificar este elemento sintáctico sobre una base por trama o cualquier otra base periódica o no periódica (como una vez para el flujo de bits completo). En cualquier caso, este elemento de sintaxis puede comprender dos bits que indican cuál de los cuatro modos de configuración fueron seleccionados para especificar el conjunto de coeficientes diferente de cero de los vectores reducidos V[*k*] de primer plano 55 para representar los aspectos direccionales de este componente distinto. El elemento de sintaxis se puede indicar como "codedVVecLength". De esta manera, la unidad de cuantización 52 puede señalar o especificar de algún modo en el flujo de bits cuál de los cuatro modos de configuración se usaron para especificar los vectores codificados V[*k*] de primer plano 57 en el flujo de bits. Aunque se describen con respecto a cuatro modos de configuración, las técnicas no deben limitarse a cuatro modos de configuración, sino a cualquier número de modos de configuración, incluido un único modo de configuración o una pluralidad de modos de configuración. La unidad de cuantización escalar/entrópica 53 también puede especificar el indicador 63 como otro elemento sintáctico en la información del canal lateral 57.

[0144] La unidad codificadora de audio psicoacústico 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 de energía 47' y las señales nFG interpoladas 49' para generar los coeficientes HOA ambientales codificados 59 y señales nFG codificadas 61. La unidad codificadora de audio psicoacústico 40 puede emitir los coeficientes HOA ambientales codificados 59 y las señales nFG codificadas 61 a la unidad de generación de flujo de bits 42.

[0145] En algunos casos, esta unidad codificadora de audio psicoacústico 40 puede representar una o más instancias de una unidad codificadora de codificación de audio avanzada (AAC). La unidad codificadora de audio psicoacústico 40 puede codificar cada columna o fila de los coeficientes HOA ambientales compensados de energía 47' y las señales nFG interpoladas 49. A menudo, la unidad codificadora de audio psicoacústico 40 puede invocar una instancia de una unidad de codificación AAC para cada una de las combinaciones de orden/suborden que permanecen en los coeficientes HOA ambientales compensados de energía 47' y las señales nFG interpoladas 49'. Puede encontrarse más información sobre cómo se pueden codificar los coeficientes armónicos esféricos de fondo 31 mediante una unidad de codificación AAC en un documento de convención de Eric Hellerud, y col., titulado

"Encoding Higher Order Ambisonics with AAC", presentado en la 124ª Convención, 2008 17-20 de mayo y disponible en: http://ro.uow.edu.au/cgi/viewcontent.cgi?article=8025&context=engpapers. En algunos casos, la unidad de codificación de audio 14 puede codificar en audio los coeficientes HOA ambientales compensados de energía 47" usando una velocidad binaria objetivo inferior que la utilizada para codificar las señales nFG interpoladas 49', comprimiendo potencialmente los coeficientes HOA ambientales compensados de energía 47' más en comparación con las señales nFG interpoladas 49'.

[0146] 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 descodificación), generando de este modo el flujo de bits basado en vectores 21. La unidad de generación de flujo de bits 42 puede representar un multiplexor en algunos ejemplos, que pueden recibir los vectores codificados V[*k*] de primer plano 57, los coeficientes de HOA codificados 59, las señales nFG codificadas 61 y la información de canal de fondo 43. La unidad de generación de 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 codificados 59, las señales nFG codificadas 61 y la información de canal de fondo 43. El flujo de bits 21 puede incluir un flujo de bits principal o principal y uno o más flujos de bits de canal lateral.

[0147] Aunque no se muestra en el ejemplo de la Fig. 4, el dispositivo de codificación de audio 20 también puede incluir una unidad de salida de flujo de bits que conmuta la salida de flujo de bits del 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) en base a si la trama actual va a codificarse usando la síntesis basada en la dirección o la síntesis basada en vectores. Esta unidad de salida de flujo de bits puede realizar esta conmutación basada en el elemento de sintaxis 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 la 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 encabezado correcta para indicar esta conmutación o codificación actual utilizada para la trama actual junto con la respectiva de los flujos de bits 21.

[0148] En algunos casos, varios aspectos de las técnicas también pueden permitir que el dispositivo de codificación de audio 20 determine si los coeficientes HOA 11 se generaron a partir de un objeto de audio sintético. Estos aspectos de las técnicas pueden permitir que el dispositivo de codificación de audio 20 sea configurado para obtener una indicación de si se generan coeficientes armónicos esféricos representativos de un campo sonoro a partir de un objeto de audio sintético.

[0149] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para determinar si los coeficientes armónicos esféricos se generan a partir del objeto de audio sintético.

[0150] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para excluir un primer vector de una matriz de coeficientes armónicos esféricos por trama que almacena al menos una parte de los coeficientes armónicos esféricos representativos del campo sonoro para obtener una matriz de coeficientes armónicos esféricos reducida por trama.

[0151] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para excluir un primer vector de una matriz de coeficientes armónicos esféricos por trama que almacena al menos una parte de los coeficientes armónicos esféricos representativos del campo sonoro para obtener una matriz de coeficientes armónicos esféricos reducida por trama, y predecir un vector de la matriz de coeficientes armónicos esféricos reducida por trama en base a los vectores restantes de la matriz de coeficientes armónicos esféricos reducida.

[0152] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para excluir un primer vector de una matriz de coeficientes armónicos esféricos por trama que almacena al menos una parte de los coeficientes armónicos esféricos representativos del campo sonoro para obtener una matriz de coeficientes armónicos esféricos enmarcada reducida, Y predecir un vector de la matriz de coeficientes armónicos esféricos enmarcada reducida basada, al menos en parte, en una suma de vectores restantes de la matriz de coeficientes armónicos esféricos enmarcada reducida.

[0153] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para predecir un vector de una matriz de coeficientes armónicos esféricos enmarcada que almacena al menos una parte de los coeficientes armónicos esféricos basados, al menos en parte, en una suma de vectores restantes del vector enmarcado Matriz de coeficientes armónicos esféricos.

[0154] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para configurarse adicionalmente para predecir un vector de una matriz de coeficientes armónicos esféricos enmarcada que almacena al menos una parte de los coeficientes armónicos esféricos basados, al menos en parte, en una suma de vectores restantes De la matriz de coeficientes armónicos esféricos por trama, y calcular un error basado en el vector predicho.

[0155] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para configurarse para predecir un vector de una matriz de coeficientes armónicos esféricos enmarcada que almacena al menos una parte de los coeficientes armónicos esféricos basado, al menos en parte, en una suma de vectores restantes de La matriz de coeficientes armónicos esféricos por trama, y calcular un error basado en el vector predicho y el correspondiente vector de la matriz de coeficientes armónicos esféricos por trama.

[0156] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para configurarse para predecir un vector de una matriz de coeficientes armónicos esféricos enmarcada que almacena al menos una parte de los coeficientes armónicos esféricos basado, al menos en parte, en una suma de vectores restantes de La matriz de coeficientes armónicos esféricos por trama y calcular un error como una suma del valor absoluto de la diferencia del vector predicho y el correspondiente vector de la matriz de coeficientes armónicos esféricos por trama.

[0157] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para configurarse para predecir un vector de una matriz de coeficientes armónicos esféricos enmarcada que almacena al menos una parte de los coeficientes armónicos esféricos basado, al menos en parte, en una suma de vectores restantes de La matriz de coeficientes armónicos esféricos por trama, calcular un error basado en el vector predicho y el correspondiente vector de la matriz de coeficientes armónicos esféricos por trama, calcular una razón basada en una energía del vector correspondiente de la matriz de coeficientes armónicos esféricos por trama y el error, y Comparar la relación con un umbral para determinar si los coeficientes armónicos esféricos representativos del campo sonoro se generan a partir del objeto de audio sintético.

[0158] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para configurarse para especificar la indicación en un flujo de bits 21 que almacena una versión comprimida de los coeficientes armónicos esféricos.

[0159] En algunos casos, las diversas técnicas pueden permitir que el dispositivo de codificación de audio 20 realice una transformación con respecto a los coeficientes HOA 11. En estos y otros casos, el dispositivo de codificación de audio 20 puede configurarse para obtener uno o más primeros vectores que describen componentes distintos del campo sonoro y uno o más segundos vectores que describen componentes de fondo del campo sonoro, tanto uno como más primeros vectores y el uno O más segundos vectores generados al menos realizando una transformación con respecto a la pluralidad de coeficientes armónicos esféricos 11.

[0160] En estos y otros casos, el dispositivo de codificación de audio 20, en el que la transformación comprende una descomposición en valores singulares que genera una matriz U representativa de los vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos 11.

[0161] En estos y otros casos, el dispositivo de codificación de audio 20, en el que uno o más primeros vectores comprenden uno o más vectores U_{DIST} * S_{DIST} codificados de audio que, antes de la codificación de audio, se generaron multiplicando uno o más vectores U_{DIST} codificados de audio de una matriz U por uno o más vectores S_{DIST} de una matriz S, y en el que la matriz U y la matriz S se generan al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos.

[0162] En estos y otros casos, el dispositivo de codificación de audio 20, en el que uno o más de los primeros vectores comprenden uno o más vectores U_{DIST} * S_{DIST} codificados de audio que, antes de la codificación de audio, se generaron multiplicando uno o más vectores U_{DIST} codificados de audio de una matriz U por uno o más vectores S_{DIST} de una matriz S y uno o más vectores V^T_{DIST} de una transposición de una matriz V, y en el que la matriz U y la matriz S y la matriz V se generan al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos 11.

[0163] En estos y otros casos, el dispositivo de codificación de audio 20, en el que uno o más primeros vectores comprenden uno o más vectores U_{DIST} * S_{DIST} que, antes de la codificación de audio, se generaron multiplicando uno o más vectores U_{DIST} codificados por audio de una matriz U por uno o más vectores S_{DIST} de una matriz S y uno o más vectores V^T_{DIST} de una transposición de una matriz V, en donde la matriz U, la matriz S y la matriz V se generaron al menos realizando la descomposición de valor singular con respecto a la pluralidad de coeficientes armónicos esféricos y en el que el dispositivo de codificación de audio 20 está configurado adicionalmente para obtener un valor D que indica el número de vectores a extraer de un flujo de bits para formar uno o más vectores U_{DIST} *S_{DIST} y uno o más vectores V^T_{DIST}.

[0164] En estos y otros casos, el dispositivo de codificación de audio 20, en el que uno o más primeros vectores comprenden uno o más vectores U_{DIST} * S_{DIST} que, antes de la codificación por audio, se generaron multiplicando uno o más vectores U_{DIST} codificados por audio de una matriz U por uno o más vectores S_{DIST} de una matriz S y uno o más vectores V^T_{DIST} de una transposición de una matriz V, en donde la matriz U, la matriz S y la matriz V se generaron al menos realizando la descomposición singular con respecto a la pluralidad de coeficientes armónicos esféricos, en el que el dispositivo de codificación de audio 20 está configurado adicionalmente para obtener un valor

D sobre una base de trama de audio por trama de audio que indica el número de vectores que se van a extraer de un flujo de bits para formar uno o más vectores $U_{DIST} * S_{DIST}$ y uno o más vectores V^T_{DIST}.

[0165] En estos y otros casos, el dispositivo de codificación de audio 20, en el que la transformación comprende un análisis de componente principal para identificar los distintos componentes del campo sonoro y los componentes de fondo del campo sonoro.

[0166] Varios aspectos de las técnicas descritas en esta divulgación pueden proporcionar el dispositivo de codificación de audio 20 configurado para compensar el error de cuantización.

[0167] En algunos casos, el dispositivo de codificación de audio 20 puede estar configurado para cuantificar uno o más primeros vectores representativos de uno o más componentes de un campo sonoro y compensar el error introducido debido a la cuantización de uno o más de los primeros vectores en uno o más de los segundos vectores que son también representativos del mismo o más componentes del campo sonoro.

[0168] En estos y otros casos, el dispositivo de codificación de audio está configurado para cuantificar uno o más vectores de una transposición de una matriz V generada, al menos en parte, realizando una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos que describen el campo sonoro.

[0169] En estos y otros casos, el dispositivo de codificación de audio está configurado además para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de los vectores singulares izquierdos de la pluralidad de armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y configurada para cuantificar uno o más vectores a partir de una transposición de la matriz V.

[0170] En estos y otros casos, el dispositivo de codificación de audio está configurado además para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de los vectores singulares izquierdos de la pluralidad de armónicos esféricos, una matriz S representativa de los valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, configurada para cuantificar uno o más vectores de una transposición de la matriz V, configurada para compensar el error introducido debido a la cuantización en uno o más vectores U * S calculados multiplicando uno o más vectores U de la matriz U por uno o más vectores S de la matriz S.

[0171] En estos y otros casos, el dispositivo de codificación de audio está configurado además para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de los vectores singulares izquierdos de la pluralidad de armónicos esféricos, una matriz S representativa de los valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, determinan uno o más vectores U_{DIST} de la matriz U, cada uno de los cuales corresponde a un componente distinto del campo sonoro, determinan uno o más vectores S_{DIST} de la matriz S, cada uno de los cuales corresponde a la misma componente distinta del campo sonoro, y determinan uno o más vectores V^T_{DIST} de una transposición de la matriz V, cada una de las cuales corresponde a la misma componente distinta del campo sonoro, configurada para cuantificar uno o más vectores V^T_{DIST} para generar uno o más vectores $V^T_{Q_DIST}$ configurados para compensar el error introducido debido a la cuantización en uno o más vectores $U_{DIST} * S_{DIST}$ calculados multiplicando uno o más vectores U_{DIST} de la matriz U por uno o más vectores S_{DIST} de la matriz S para generar uno o más vectores $U_{DIST} * S_{DIST}$ compensados de errores.

[0172] En estos y otros casos, el dispositivo de codificación de audio está configurado para determinar coeficientes armónicos esféricos distintos basados en uno o más vectores U_{DIST}, uno o más vectores S_{DIST} y uno o más vectores V^T_{DIST} y realizar una pseudo inversa con respecto a los vectores $V^T_{Q_DIST}$ para dividir los distintos coeficientes armónicos esféricos por uno o más de los vectores $V^T_{Q_DIST}$ y de este modo generar uno o más vectores $U_{C_DIST} * S_{C_DIST}$ compensados de errores que compensan, al menos en parte, el error introducido a través de la cuantización de los vectores V^T_{DIST}.

[0173] En estos y otros casos, el dispositivo de codificación de audio está configurado además para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de los vectores singulares izquierdos de la pluralidad de armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, determinan uno o más vectores U_{BG} de la matriz U que describen uno o más componentes de fondo del campo sonoro y uno o más vectores U_{DIST} de la matriz U que describen uno o más componentes distintos del campo sonoro, determinan uno o más vectores S_{BG} de la matriz S que describen uno o más componentes de fondo del campo sonoro y uno o más vectores S_{DIST} de la matriz S que describen uno o más

componentes distintos del campo sonoro, y determinan uno o más vectores V^T_{DIST} y uno o más vectores V^T_{BG} de una transposición de la matriz V, en el que los vectores V^T_{DIST} describen uno o más componentes distintos del campo sonoro y el V^T_{BG} describe uno o más componentes de fondo del campo sonoro, configurados para cuantificar uno o más vectores V^T_{DIST} para generar uno o más vectores $V^T_{Q_DIST}$, configurados para compensar el error introducido debido a la cuantización en los coeficientes armónicos esféricos de fondo formados multiplicando uno o más vectores U_{BG} por uno o más vectores S_{BG} y luego por uno o más vectores V^T_{BG} para generar coeficientes armónicos esféricos de fondo compensados de error.

[0174] En estos y otros casos, el dispositivo de codificación de audio está configurado para determinar el error basado en los vectores V^T_{DIST} y uno o más vectores $U_{DIST} * S_{DIST}$ formados multiplicando los vectores U_{DIST} por los vectores S_{DIST} y añadir el error determinado a los coeficientes armónicos esféricos de fondo para generar los coeficientes armónicos esféricos compensados de error.

[0175] En estos y otros casos, el dispositivo de codificación de audio está configurado para compensar el error introducido debido a la cuantización de uno o más de los primeros vectores en uno o más de los segundos vectores que son también representativos del mismo o más componentes del campo sonoro para generar uno o más segundos vectores compensados de error, configurados adicionalmente para generar un flujo de bits que incluya uno o más segundos vectores compensados de error y uno o más de los primeros vectores cuantizados.

[0176] En estos y otros casos, el dispositivo de codificación de audio está configurado para compensar el error introducido debido a la cuantización de uno o más de los primeros vectores en uno o más de los segundos vectores que son también representativos del mismo o más componentes del campo sonoro para generar uno o más segundos vectores compensados de error, configurados adicionalmente para codificar en audio uno o más vectores secundarios compensados de error, y generar un flujo de bits que incluya el audio codificado de uno o más segundos vectores compensados de error y uno o más de los primeros vectores cuantizados.

[0177] Los diversos aspectos de las técnicas pueden permitir además que el dispositivo de codificación de audio 20 genere coeficientes armónicos esféricos reducidos o descomposiciones de los mismos. En algunos casos, el dispositivo de codificación de audio 20 puede configurarse para realizar, en base a una velocidad de bits objetivo, una reducción de orden con respecto a una pluralidad de coeficientes armónicos esféricos o descomposiciones de los mismos para generar coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos, en los que la pluralidad de los coeficientes armónicos esféricos representan un campo sonoro.

[0178] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para, antes de realizar la reducción de orden, realizar una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos para identificar uno o más de los primeros vectores que describen componentes distintos del campo sonoro, y uno o más de los segundos vectores que identifican componentes de fondo del campo sonoro, configurados para realizar la reducción de orden con respecto a uno o más de los primeros vectores, uno o más de los segundos vectores o, ambos, uno o más de los primeros vectores y uno o más de los segundos vectores.

[0179] En estos y otros casos, el dispositivo de codificación de audio 20 está además configurado para realizar un análisis de contenido con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos, y configurado para realizar, basándose en la velocidad de bits objetivo y el análisis de contenido, una reducción de orden con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos para generar los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos.

[0180] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para realizar un análisis espacial con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos.

[0181] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para realizar un análisis de difusión con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos.

[0182] En estos y otros casos, el dispositivo de codificación de audio 20 es uno o más procesadores que están configurados para realizar un análisis espacial y un análisis de difusión con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos.

[0183] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para especificar uno o más órdenes y/o una o más subórdenes de funciones de base esférica con las que los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de éstos se corresponden en un flujo de bits que incluye los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos.

[0184] En estos y otros casos, los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos tienen valores menores que la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos.

[0185] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para eliminar aquellos de la pluralidad de coeficientes armónicos esféricos o vectores de las descomposiciones de los mismos que tienen un orden y/o suborden especificados para generar los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos.

[0186] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para poner a cero los de la pluralidad de coeficientes armónicos esféricos o aquellos vectores de su descomposición que tienen un orden y/o suborden especificados para generar los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos.

[0187] Varios aspectos de las técnicas también pueden permitir que el dispositivo de codificación de audio 20 sea configurado para representar componentes distintos del campo sonoro. En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para obtener un primer conjunto diferente de cero de coeficientes de un vector que se utilizará para representar un componente distinto de un campo sonoro, en el que el vector se descompone a partir de una pluralidad de coeficientes armónicos esféricos que describen el campo sonoro.

[0188] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para determinar el primer conjunto distinto de cero de los coeficientes del vector para incluir todos los coeficientes.

[0189] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para determinar el primer conjunto de coeficientes distinto de cero como los de los coeficientes correspondientes a un orden mayor que el orden de una función de base con la cual uno o más de la pluralidad de coeficientes armónicos esféricos se corresponden.

[0190] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para determinar el primer conjunto de coeficientes distinto de cero para incluir los de los coeficientes correspondientes a un orden mayor que el orden de una función base con la cual uno o más de la pluralidad de coeficientes armónicos esféricos se corresponden y excluyen al menos uno de los coeficientes que corresponde a un orden mayor que el orden de la función de base con la cual se corresponden uno o más de la pluralidad de coeficientes armónicos esféricos.

[0191] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para determinar el primer conjunto de coeficientes distinto de cero para incluir todos los coeficientes excepto por lo menos uno de los coeficientes correspondientes a un orden mayor que el orden de una función de base con la cual uno o más de la pluralidad de coeficientes armónicos esféricos se corresponden.

[0192] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para especificar el primer conjunto distinto de cero de los coeficientes del vector en la información de canal lateral.

[0193] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para especificar el primer conjunto distinto de cero de los coeficientes del vector en la información de canal lateral sin codificación de audio del primer conjunto distinto de cero de los coeficientes del vector.

[0194] En estos y otros casos, el vector comprende un vector descompuesto a partir de la pluralidad de coeficientes armónicos esféricos utilizando la síntesis basada en vectores.

[0195] En estos y otros casos, la síntesis basada en vectores comprende una descomposición en valores singulares.

[0196] En estos y otros casos, el vector comprende un vector V descompuesto a partir de la pluralidad de coeficientes armónicos esféricos utilizando descomposición en valores singulares.

[0197] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para seleccionar uno de una pluralidad de modos de configuración mediante los cuales se especifica el conjunto de coeficientes del vector distinto de cero y se especifica el conjunto distinto de cero de los coeficientes del vector basado en el modo seleccionado de la pluralidad de modos de configuración.

[0198] En estos y otros casos, el modo de la pluralidad de modos de configuración indica que el conjunto distinto de cero de los coeficientes incluye todos los coeficientes.

[0199] En estos y otros casos, el modo de la pluralidad de modos de configuración indica que el conjunto de coeficientes distinto de cero incluye los de los coeficientes correspondientes a un orden mayor que el orden de una función base con la cual uno o más de la pluralidad de coeficientes armónicos esféricos se corresponden.

[0200] En estos y otros casos, el modo de la pluralidad de modos de configuración indica que el conjunto distinto de cero de los coeficientes incluye los de los coeficientes correspondientes a un orden mayor que el orden de una

función base con la cual uno o más de la pluralidad de coeficientes armónicos esféricos se corresponden y excluyen al menos uno de los coeficientes correspondientes a un orden mayor que el orden de la función de base con la cual se corresponden uno o más de la pluralidad de coeficientes armónicos esféricos,

[0201] En estos y otros casos, el modo de la pluralidad de modos de configuración indica que el conjunto de coeficientes distinto de cero incluye todos los coeficientes excepto por lo menos uno de los coeficientes.

[0202] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para especificar el modo seleccionado de la pluralidad de modos de configuración en un flujo de bits.

[0203] Varios aspectos de las técnicas descritas en esta descripción pueden permitir también que el dispositivo de codificación de audio 20 sea configurado para representar de forma diferente dicho componente distinto del campo sonoro. En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para obtener un primer conjunto distinto de cero de coeficientes de un vector que representa un componente distinto de un campo sonoro, habiéndose descompuesto el vector a partir de una pluralidad de coeficientes armónicos esféricos que describen el campo sonoro.

[0204] En estos y otros casos, el primer conjunto distinto de cero de los coeficientes incluye todos los coeficientes del vector.

[0205] En estos y otros casos, el primer conjunto de coeficientes distinto de cero incluye los de los coeficientes correspondientes a un orden mayor que el orden de una función de base con la que se corresponden uno o más de la pluralidad de coeficientes armónicos esféricos.

[0206] En estos y otros casos, el primer conjunto distinto de cero de los coeficientes incluye los de los coeficientes correspondientes a un orden mayor que el orden de una función base con la cual uno o más de la pluralidad de coeficientes armónicos esféricos se corresponden y excluyen al menos uno de los coeficientes correspondientes a un orden mayor que el orden de la función de base con la cual se corresponden uno o más de la pluralidad de coeficientes armónicos esféricos.

[0207] En estos y otros casos, el primer conjunto de coeficientes distinto de cero incluye todos los coeficientes excepto por lo menos uno de los coeficientes identificados porque no tienen suficiente información direccional.

[0208] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para extraer el primer conjunto distinto de cero de los coeficientes como una primera parte del vector.

[0209] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para extraer el primer conjunto distinto de cero del vector de la información de canal lateral y obtener una versión recompuesta de la pluralidad de coeficientes armónicos esféricos basándose en el primer conjunto distinto de cero de los coeficientes del vector.

[0210] En estos y otros casos, el vector comprende un vector descompuesto a partir de la pluralidad de coeficientes armónicos esféricos utilizando la síntesis basada en vectores.

[0211] En estos y otros casos, la síntesis basada en vectores comprende la descomposición en valores singulares.

[0212] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para determinar uno modo de una pluralidad de modos de configuración mediante los cuales extraer el conjunto de coeficientes del vector distinto de cero de acuerdo con el modo de la pluralidad de modos de configuración y extraer el conjunto distinto de cero de los coeficientes del vector basado en el modo obtenido de la pluralidad de modos de configuración.

[0213] En estos y otros casos, el modo de la pluralidad de modos de configuración indica que el conjunto distinto de cero de los coeficientes incluye todos los coeficientes.

[0214] En estos y otros casos, el modo de la pluralidad de modos de configuración indica que el conjunto de coeficientes distinto de cero incluye los de los coeficientes correspondientes a un orden mayor que el orden de una función base con la cual uno o más de la pluralidad de coeficientes armónicos esféricos se corresponden.

[0215] En estos y otros casos, el modo de la pluralidad de modos de configuración indica que el conjunto distinto de cero de los coeficientes incluye los de los coeficientes correspondientes a un orden mayor que el orden de una función base con la cual uno o más de la pluralidad de coeficientes armónicos esféricos se corresponden y excluyen al menos uno de los coeficientes correspondientes a un orden mayor que el orden de la función de base con la cual se corresponden uno o más de la pluralidad de coeficientes armónicos esféricos,

[0216] En estos y otros casos, el modo de la pluralidad de modos de configuración indica que el conjunto de coeficientes distinto de cero incluye todos los coeficientes excepto por lo menos uno de los coeficientes.

[0217] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para determinar el modo de la pluralidad de modos de configuración basándose en un valor señalado en un flujo de bits.

[0218] Varios aspectos de las técnicas también pueden, en algunos casos, permitir que el dispositivo de codificación de audio 20 identifique uno o más objetos de audio distintos (o, en otras palabras, objetos de audio predominantes). En algunos casos, el dispositivo de codificación de audio 20 puede configurarse para identificar uno o más objetos de audio distintos a partir de uno o más coeficientes armónicos esféricos (SHC) asociados con los objetos de audio basados en una direccionalidad determinada para uno o más de los objetos de audio.

[0219] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para determinar la direccionalidad de uno o más objetos de audio basándose en los coeficientes armónicos esféricos asociados con los objetos de audio.

[0220] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para realizar una descomposición en valores singulares con respecto a los coeficientes armónicos esféricos para generar una matriz U representativa de los vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y representar la pluralidad de coeficientes armónicos esféricos como una función de al menos una parte de una o más de la matriz U, la matriz S y la matriz V, en la que el dispositivo de codificación de audio 20 está configurado para determinar la direccionalidad respectiva de uno o más objetos de audio y se basa al menos en parte en la matriz V.

[0221] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para reordenar uno o más vectores de la matriz V de modo que los vectores que tienen un cociente de direccionalidad mayor estén situados por encima de los vectores que tienen un cociente de direccionalidad menor en la matriz V reordenada.

[0222] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para determinar que los vectores que tienen el mayor cociente de direccionalidad incluyen una mayor información direccional que los vectores que tienen el cociente de direccionalidad menor.

[0223] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para multiplicar la matriz V por la matriz S para generar una matriz VS, la matriz VS que incluye uno o más vectores.

[0224] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para seleccionar entradas de cada fila de la matriz VS que están asociadas con un orden mayor de 14, cuadrar cada una de las entradas seleccionadas para formar entradas cuadráticas correspondientes y, para cada fila de la matriz VS, sumar todas las entradas al cuadrado para determinar un cociente de direccionalidad para un vector correspondiente.

[0225] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para seleccionar las entradas de cada fila de la matriz VS asociada con el orden mayor que 14 comprende seleccionar todas las entradas comenzando en una entrada decimoctava de cada fila de la matriz VS y terminando en una entrada 38 de cada fila de la matriz VS.

[0226] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para seleccionar un subconjunto de los vectores de la matriz VS para representar los distintos objetos de audio. En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para seleccionar cuatro vectores de la matriz VS, y en el que los cuatro vectores seleccionados tienen los cuatro cocientes de direccionalidad más grandes de todos los vectores de la matriz VS.

[0227] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para determinar que el subconjunto seleccionado de los vectores representa los distintos objetos de audio y se basa tanto en la direccionalidad como en una energía de cada vector.

[0228] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para realizar una comparación de energía entre uno o más primeros vectores y uno o más segundos vectores representativos de los distintos objetos de audio para determinar uno o más primeros vectores reordenados, en el que uno o más primeros vectores describen los distintos objetos de audio una primera parte de datos de audio y uno o más segundos vectores describen los distintos objetos de audio en una segunda parte de los datos de audio.

[0229] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para realizar una correlación cruzada entre uno o más primeros vectores y uno o más segundos vectores representativos de los distintos objetos de audio para determinar uno o más primeros vectores reordenados, en el que uno o más

primeros vectores describen los distintos objetos de audio una primera parte de datos de audio y uno o más segundos vectores describen los distintos objetos de audio en una segunda parte de los datos de audio.

[0230] Varios aspectos de las técnicas también pueden, en algunos casos, permitir que el dispositivo de codificación de audio 20 esté configurado para realizar compensación de energía con respecto a las descomposiciones de los coeficientes HOA 11. En estos y otros casos, el dispositivo de codificación de audio 20 puede estar configurado para realizar una síntesis basada en vectores con respecto a una pluralidad de coeficientes armónicos esféricos para generar representaciones descompuestas de la pluralidad de coeficientes armónicos esféricos representativos de uno o más objetos de audio y la correspondiente información direccional, en el que los coeficientes armónicos esféricos están asociados con un orden y describen un campo sonoro, determinan información direccional distinta y de fondo a partir de la información direccional, reducen un orden de la información direccional asociada con los objetos de audio de fondo para generar información de dirección de fondo transformada, aplican compensación para incrementar los valores de la información direccional transformada para preservar una energía global del campo sonoro.

[0231] En estos y otros casos, el dispositivo de codificación de audio 20 puede estar configurado para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos para generar una matriz U y una matriz S representativa de los objetos de audio y una matriz V representativa de la información direccional, determinar vectores de columna distintos de la matriz V y vectores de columna de fondo de la matriz V, reducir un orden de los vectores de columna de fondo de la matriz V para generar vectores de columna de fondo transformados de la matriz V y aplicar la compensación para aumentar los valores de los vectores de columna de fondo transformados de la matriz V para conservar una energía global del campo sonoro.

[0232] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para determinar un número de valores singulares relevantes de la matriz S, en el que un número de los vectores de columna distintos de la matriz V es el número de valores singulares relevantes de la matriz S.

[0233] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para determinar un orden reducido para los coeficientes armónicos esféricos y valores cero para filas de los vectores de columna de fondo de la matriz V asociados a un orden que es mayor que el orden reducido.

[0234] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para combinar columnas de fondo de la matriz U, columnas de fondo de la matriz S y transposición de las columnas de fondo transformadas de la matriz V para generar coeficientes armónicos esféricos modificados.

[0235] En estos y otros casos, los coeficientes armónicos esféricos modificados describen uno o más componentes de fondo del campo sonoro.

[0236] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para determinar una primera energía de un vector de los vectores de columna de fondo de la matriz V y una segunda energía de un vector de los vectores de columna de fondo transformados de la matriz V y aplicar un valor de amplificación a cada elemento del vector de los vectores de columna de fondo transformados de la matriz V, en el que el valor de amplificación comprende una relación entre la primera energía y la segunda energía.

[0237] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para determinar una primera energía media cuadrática de un vector de los vectores de columna de fondo de la matriz V y una segunda energía media cuadrática de un vector de los vectores de la columna de fondo transformada de la matriz V y aplicar un valor de amplificación a cada elemento del vector de los vectores de columna de fondo transformados de la matriz V, en el que el valor de amplificación comprende una relación entre la primera energía y la segunda energía.

[0238] Varios aspectos de las técnicas descritas en esta descripción pueden también permitir que el dispositivo de codificación de audio 20 realice interpolación con respecto a versiones descompuestas de los coeficientes HOA 11. En algunos casos, el dispositivo de codificación de audio 20 puede configurarse para obtener coeficientes armónicos esféricos interpolados descompuestos para un segmento de tiempo, realizando al menos en parte una interpolación con respecto a una primera descomposición de una primera pluralidad de coeficientes armónicos esféricos y una segunda descomposición de una segunda pluralidad de coeficientes armónicos esféricos.

[0239] En estos y otros casos, la primera descomposición comprende una primera matriz V representativa de los vectores singulares derechos de la primera pluralidad de coeficientes armónicos esféricos.

[0240] En estos y otros ejemplos, la segunda descomposición comprende una segunda matriz V representativa de los vectores singulares derechos de la segunda pluralidad de coeficientes armónicos esféricos.

[0241] En estos y otros casos, la primera descomposición comprende una primera matriz V representativa de los vectores singulares derechos de la primera pluralidad de coeficientes armónicos esféricos y la segunda

descomposición comprende una segunda matriz V representativa de los vectores singulares derechos de la segunda pluralidad de coeficientes armónicos esféricos.

[0242] En estos y otros casos, el segmento de tiempo comprende una subtrama de una trama de audio.

[0243] En estos y otros casos, el segmento de tiempo comprende una muestra de tiempo de una trama de audio.

[0244] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para obtener una descomposición interpolada de la primera descomposición y la segunda descomposición para un coeficiente armónico esférico de la primera pluralidad de coeficientes armónicos esféricos.

[0245] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para obtener descomposiciones interpoladas de la primera descomposición para una primera parte de la primera pluralidad de coeficientes armónicos esféricos incluidos en la primera trama y la segunda descomposición para una segunda parte de la segunda pluralidad de los coeficientes armónicos esféricos incluidos en la segunda trama, y el dispositivo de codificación de audio 20 está configurado además para aplicar las descomposiciones interpoladas a un primer componente de tiempo de la primera parte de la primera pluralidad de coeficientes armónicos esféricos incluidos en la primera trama para generar un primer componente artificial de tiempo de la primera pluralidad de coeficientes armónicos esféricos y aplicar las respectivas descomposiciones interpoladas a un segundo componente de tiempo de la segunda parte de la segunda pluralidad de coeficientes armónicos esféricos incluidos en la segunda trama para generar un segundo componente artificial de tiempo de la segunda pluralidad de coeficientes armónicos esféricos incluidos.

[0246] En estos y otros casos, el primer componente de tiempo se genera realizando una síntesis basada en vectores con respecto a la primera pluralidad de coeficientes armónicos esféricos.

[0247] En estos y otros casos, el segundo componente de tiempo se genera realizando una síntesis basada en vectores con respecto a la segunda pluralidad de coeficientes armónicos esféricos.

[0248] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para recibir el primer componente artificial de tiempo y el segundo componente artificial de tiempo, calcular las descomposiciones interpoladas de la primera descomposición para la primera parte de la primera pluralidad de coeficientes armónicos esféricos y la segunda descomposición para la segunda parte de la segunda pluralidad de coeficientes armónicos esféricos y aplicar inversos de las descomposiciones interpoladas al primer componente artificial de tiempo para recuperar el componente de primer tiempo y al segundo componente artificial de tiempo para recuperar el segundo componente de tiempo.

[0249] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para interpolar un primer componente espacial de la primera pluralidad de coeficientes armónicos esféricos y el segundo componente espacial de la segunda pluralidad de coeficientes armónicos esféricos.

[0250] En estos y otros casos, el primer componente espacial comprende una primera matriz U representativa de los vectores singulares izquierdos de la primera pluralidad de coeficientes armónicos esféricos.

[0251] En estos y otros casos, el segundo componente espacial comprende una segunda matriz U representativa de los vectores singulares izquierdos de la segunda pluralidad de coeficientes armónicos esféricos.

[0252] En estos y otros casos, el primer componente espacial es representativo de M segmentos temporales de coeficientes armónicos esféricos para la primera pluralidad de coeficientes armónicos esféricos y el segundo componente espacial es representativo de M segmentos temporales de coeficientes armónicos esféricos para la segunda pluralidad de coeficientes armónicos esféricos.

[0253] En estos y otros casos, el primer componente espacial es representativo de M segmentos temporales de coeficientes armónicos esféricos para la primera pluralidad de coeficientes armónicos esféricos y el segundo componente espacial es representativo de M segmentos temporales de coeficientes armónicos esféricos para la segunda pluralidad de coeficientes armónicos esféricos y el dispositivo de codificación de audio 20 está configurado para interpolar los últimos N elementos del primer componente espacial y los primeros N elementos del segundo componente espacial.

[0254] En estos y otros casos, la segunda pluralidad de coeficientes armónicos esféricos es posterior a la primera pluralidad de coeficientes armónicos esféricos en el dominio del tiempo.

[0255] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para descomponer la primera pluralidad de coeficientes armónicos esféricos para generar la primera descomposición de la primera pluralidad de coeficientes armónicos esféricos.

[0256] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para descomponer la segunda pluralidad de coeficientes armónicos esféricos para generar la segunda descomposición de la segunda pluralidad de coeficientes armónicos esféricos.

[0257] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para realizar una descomposición en valores singulares con respecto a la primera pluralidad de coeficientes armónicos esféricos para generar una matriz U representativa de los vectores singulares izquierdos de la primera pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la primera pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la primera pluralidad de coeficientes armónicos esféricos.

[0258] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para realizar una descomposición en valores singulares con respecto a la segunda pluralidad de coeficientes armónicos esféricos para generar una matriz U representativa de los vectores singulares izquierdos de la segunda pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la segunda pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la segunda pluralidad de coeficientes armónicos esféricos.

[0259] En estos y otros casos, la primera y la segunda pluralidad de coeficientes armónicos esféricos representan cada una una representación de onda plana del campo sonoro.

[0260] En estos y otros casos, la primera y la segunda pluralidad de coeficientes armónicos esféricos representan cada uno uno o más objetos monoaudio mezclados entre sí.

[0261] En estos y otros casos, la primera y la segunda pluralidad de coeficientes armónicos esféricos comprenden cada uno un primer y segundo coeficientes armónicos esféricos respectivos que representan un campo sonoro tridimensional.

[0262] En estos y otros casos, la primera y la segunda pluralidad de coeficientes armónicos esféricos están asociadas cada una con al menos una función de base esférica que tiene un orden mayor que uno.

[0263] En estos y otros casos, la primera y la segunda pluralidad de coeficientes armónicos esféricos están asociadas cada una con al menos una función de base esférica que tiene un orden igual a cuatro.

[0264] En estos y otros casos, la interpolación es una interpolación ponderada de la primera descomposición y segunda descomposición, en la que los pesos de la interpolación ponderada aplicada a la primera descomposición son inversamente proporcionales a un tiempo representado por los vectores de la primera y segunda descomposición y en donde los pesos de la interpolación ponderada aplicada a la segunda descomposición es proporcional a un tiempo representado por vectores de la primera y segunda descomposición.

[0265] En estos y otros casos, los coeficientes armónicos esféricos interpolados descompuestos suavizan al menos uno de componentes espaciales y los componentes temporales de la primera pluralidad de coeficientes armónicos esféricos y la segunda pluralidad de coeficientes armónicos esféricos.

[0266] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para calcular US[n] = HOA (n) * (V_vec [n]) - 1 para obtener un escalar.

[0267] En estos y otros casos, la interpolación comprende una interpolación lineal. En estos y otros casos, la interpolación comprende una interpolación no lineal. En estos y otros casos, la interpolación comprende una interpolación de coseno. En estos y otros casos, la interpolación comprende una interpolación de coseno ponderada. En estos y otros casos, la interpolación comprende una interpolación cúbica. En estos y otros casos, la interpolación comprende una interpolación de spline adaptativa. En estos y otros casos, la interpolación comprende una interpolación de curvatura mínima.

[0268] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para generar un flujo de bits que incluye una representación de los coeficientes armónicos esféricos interpolados descompuestos para el segmento de tiempo y una indicación de un tipo de interpolación.

[0269] En estos y otros casos, la indicación comprende uno o más bits que se corresponden con el tipo de interpolación.

[0270] De esta manera, varios aspectos de las técnicas descritas en esta descripción pueden permitir que el dispositivo de codificación de audio 20 sea configurado para obtener un flujo de bits que incluya una representación de los coeficientes armónicos esféricos interpolados descompuestos para el segmento de tiempo y una indicación de un tipo de interpolación.

[0271] En estos y otros casos, la indicación comprende uno o más bits que se corresponden con el tipo de interpolación.

[0272] A este respecto, el dispositivo de codificación de audio 20 puede representar una realización de las técnicas en que el dispositivo de codificación de audio 20 puede, en algunos casos, configurarse para generar un flujo de bits que comprende una versión comprimida de un componente espacial de un campo sonoro, generado mediante la realización de una síntesis basada en vectores con respecto a una pluralidad de coeficientes armónicos esféricos.

[0273] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para generar el flujo de bits para incluir un campo que especifica un modo de predicción utilizado al comprimir el componente espacial.

[0274] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para generar el flujo de bits para incluir información de tabla Huffman que especifica una tabla Huffman usada al comprimir el componente espacial.

[0275] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para generar el flujo de bits para incluir un campo que indica un valor que expresa un tamaño del paso de cuantización o una variable del mismo utilizado al comprimir el componente espacial.

[0276] En estos y otros casos, el valor comprende un valor *nbits*.

[0277] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para generar el flujo de bits para incluir una versión comprimida de una pluralidad de componentes espaciales del campo sonoro de la que está incluida la versión comprimida del componente espacial, donde el valor expresa el tamaño del paso de cuantización o una variable de la misma utilizada al comprimir la pluralidad de componentes espaciales.

[0278] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para generar el flujo de bits para incluir un código Huffman para representar un identificador de categoría que identifica una categoría de compresión con la que se corresponde el componente espacial.

[0279] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para generar el flujo de bits para incluir un bit de signo que identifica si el componente espacial es un valor positivo o un valor negativo.

[0280] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para generar el flujo de bits para incluir un código Huffman para representar un valor residual del componente espacial.

[0281] En estos y otros casos, la síntesis basada en vectores comprende una descomposición en valores singulares.

[0282] A este respecto, el dispositivo de codificación de audio 20 puede implementar adicionalmente diversos aspectos de las técnicas en que el dispositivo de codificación de audio 20 puede, en algunos casos, configurarse para identificar un libro de códigos Huffman para usar al comprimir un componente espacial de una pluralidad de componentes espaciales en base a un orden del componente espacial relativo a los restantes de la pluralidad de componentes espaciales, el componente espacial generado realizando una síntesis basada en vectores con respecto a una pluralidad de coeficientes armónicos esféricos.

[0283] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para identificar el libro de códigos Huffman basado en un modo de predicción utilizado al comprimir el componente espacial.

[0284] En estos y otros casos, una versión comprimida del componente espacial se representa en un flujo de bits utilizando, al menos en parte, información de la tabla Huffman que identifica el libro de códigos Huffman.

[0285] En estos y otros casos, una versión comprimida del componente espacial se representa en un flujo de bits utilizando, al menos en parte, un campo que indica un valor que expresa un tamaño del paso de cuantización o una variable del mismo utilizado al comprimir el componente espacial.

[0286] En estos y otros casos, el valor comprende un valor *nbits*.

[0287] En estos y otros casos, el flujo de bits comprende una versión comprimida de una pluralidad de componentes espaciales del campo sonoro de la que está incluida la versión comprimida del componente espacial y el valor expresa el tamaño del paso de cuantización o una variable del mismo utilizada al comprimir la pluralidad de componentes espaciales.

[0288] En estos y otros casos, una versión comprimida del componente espacial se representa en un flujo de bits utilizando, al menos en parte, un código Huffman seleccionado del libro de códigos Huffman identificado para representar un identificador de categoría que identifica una categoría de compresión con la que se corresponde el componente espacial.

[0289] En estos y otros casos, una versión comprimida del componente espacial se representa en un flujo de bits utilizando, al menos en parte, un bit de signo que identifica si el componente espacial es un valor positivo o un valor negativo.

[0290] En estos y otros casos, una versión comprimida del componente espacial se representa en un flujo de bits utilizando, al menos en parte, un código Huffman seleccionado del libro de códigos Huffman identificado para representar un valor residual del componente espacial.

[0291] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para comprimir el componente espacial basado en el libro de códigos Huffman identificado para generar una versión comprimida del componente espacial y generar el flujo de bits para incluir la versión comprimida del componente espacial.

[0292] Además, el dispositivo de codificación de audio 20 puede, en algunos casos, implementar diversos aspectos de las técnicas en que el dispositivo de codificación de audio 20 puede configurarse para determinar un tamaño del paso de cuantización que se utilizará al comprimir un componente espacial de un campo sonoro, el componente espacial generado mediante la realización de una síntesis basada en vectores con respecto a una pluralidad de coeficientes armónicos esféricos.

[0293] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para determinar el tamaño del paso de cuantización basado en una velocidad de bits objetivo.

[0294] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para determinar una estimación de un número de bits utilizados para representar el componente espacial y determinar el tamaño del paso de cuantización basándose en una diferencia entre la estimación y una velocidad de bits objetivo.

[0295] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para determinar una estimación de un número de bits utilizados para representar el componente espacial, determinar una diferencia entre la estimación y una velocidad de bits objetivo y determinar el tamaño del paso de cuantización añadiendo la diferencia a la velocidad de bits objetivo.

[0296] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para calcular la estimación del número de bits que se van a generar para el componente espacial dado un libro de códigos correspondiente a la velocidad de bits objetivo.

[0297] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para calcular la estimación del número de bits que se van a generar para el componente espacial dado un modo de codificación utilizado al comprimir el componente espacial.

[0298] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para calcular una primera estimación del número de bits que se van a generar para el componente espacial dado un primer modo de codificación que se utilizará al comprimir el componente espacial, calcular una segunda estimación del número de bits que se van a generar para el componente espacial dado un segundo modo de codificación que se utilizará al comprimir el componente espacial, seleccionar el de la primera estimación y la segunda estimación que tenga el menor número de bits que se va a utilizar como la estimación determinada del número de bits.

[0299] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado para identificar un identificador de categoría que identifica una categoría a la cual corresponde el componente espacial, identificar una longitud de bits de un valor residual para el componente espacial que resultaría al comprimir el componente espacial correspondiente a la categoría y determinar la estimación del número de bits mediante, al menos en parte, la adición de un número de bits utilizados para representar el identificador de categoría a la longitud de bits del valor residual.

[0300] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para seleccionar uno de una pluralidad de libros de códigos que se utilizarán al comprimir el componente espacial.

[0301] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para determinar una estimación de un número de bits utilizados para representar el componente espacial utilizando cada uno de la pluralidad de libros de códigos y seleccionar el libro de la pluralidad de libros de código que resultó en la estimación determinada que tiene el menor número de bits.

[0302] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado además para determinar una estimación de un número de bits utilizados para representar el componente espacial utilizando uno o más de la pluralidad de libros de códigos, uno o más de la pluralidad de libros de códigos seleccionados basándose en un orden de elementos del componente espacial que se comprimirá con relación a otros elementos del componente espacial.

[0303] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para determinar una estimación de un número de bits utilizados para representar el componente espacial utilizando uno de la pluralidad de libros de códigos diseñados para usarse cuando el componente espacial no se predice desde un componente espacial subsiguiente.

[0304] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para determinar una estimación de un número de bits utilizados para representar el componente espacial utilizando uno de la pluralidad de libros de códigos diseñados para usarse cuando el componente espacial se predice a partir de una componente espacial subsiguiente.

[0305] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para determinar una estimación de un número de bits utilizados para representar el componente espacial utilizando uno de la pluralidad de libros de códigos diseñados para usarse cuando el componente espacial es representativo de un objeto de audio sintético en el campo sonoro.

[0306] En estos y otros casos, el objeto de audio sintético comprende un objeto de audio modulado por impulsos codificados (PCM).

[0307] En estos y otros casos, el dispositivo de codificación de audio 20 está configurado adicionalmente para determinar una estimación de un número de bits utilizados para representar el componente espacial utilizando uno de la pluralidad de libros de códigos diseñados para usarse cuando el componente espacial es representativo de un objeto de audio grabado en el campo sonoro.

[0308] En cada una de las diversas instancias descritas anteriormente, debe entenderse que el dispositivo de codificación de audio 20 puede realizar un procedimiento o comprender de otro modo medios para realizar cada paso del procedimiento para el cual el dispositivo de codificación de audio 20 está configurado para llevar a cabo en algunos casos, estos medios pueden comprender uno o más procesadores. En algunos casos, 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 instrucciones que, cuando se ejecutan, hacen que uno o más procesadores realicen el procedimiento para el cual el dispositivo de codificación de audio 20 se ha configurado para llevar a cabo.

[0309] La Fig. 5 es un diagrama de bloques que ilustra el dispositivo de descodificación de audio 24 que se muestra en el ejemplo de la Fig. 3 con mayor detalle. Como se muestra en el ejemplo de la Fig. 5, 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.

[0310] 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 de sintaxis anotado anteriormente (por ejemplo, el elemento de sintaxis ChannelType que se muestra en los ejemplos de las Figs. 10E y 10H (i)-10O(ii) si los coeficientes HOA 11 se han codificado a través de las diversas versiones. Cuando se ha 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 de sintaxis asociados con esta versión codificada (que se designa como información basada en la dirección 91 en el ejemplo de la Fig. 5), pasando esta información basada en la dirección 91 a la unidad de reconstrucción basada en la dirección 90. Esta unidad de reconstrucción basada en la dirección 90 puede representar una unidad configurada para reconstruir los coeficientes HOA en forma de coeficientes HOA 11' basados en la información basada en la dirección 91. El flujo de bits y la disposición de elementos de sintaxis dentro del flujo de bits se describen a continuación con más detalle con respecto al ejemplo de las Figs. 10-10O(ii) y 11.

[0311] Cuando el elemento de sintaxis indica que los coeficientes HOA 11 se han codificado utilizando una síntesis basada en vectores, la unidad de extracción 72 puede extraer los vectores [*k*] codificados de primer plano 57, los coeficientes HOA codificados 59 y las señales nFG codificadas 59. La unidad de extracción 72 puede pasar los vectores codificados V[*k*] de primer plano 57 a la unidad de cuantización 74 y los coeficientes HOA ambientales codificados 59 junto con las señales nFG codificadas 61 a la unidad de decodificación psicoacústica 80.

[0312] Para extraer los vectores de primer plano codificados V[*k*] 57, los coeficientes HOA ambientales codificados 59 y las señales nFG codificadas 59, la unidad de extracción 72 puede obtener la información de canal lateral 57, que incluye el elemento de sintaxis denominado codedVVecLength. La unidad de extracción 72 puede analizar la codedVVecLength a partir de la información de canal lateral 57. La unidad de extracción 72 puede estar configurada para funcionar en cualquiera de los modos de configuración descritos anteriormente basándose en el elemento de sintaxis codedVVecLength.

[0313] La unidad de extracción 72 funciona entonces de acuerdo con cualquiera de los modos de configuración para analizar una forma comprimida de los vectores V[*k*] de primer plano reducidos 55_k a partir de la información del canal lateral 57. La unidad de extracción 72 puede operar de acuerdo con la instrucción de conmutación que se presenta en el pseudocódigo siguiente con la sintaxis que se presenta en la siguiente tabla de sintaxis para VVectorData:

```
switch CodedVVecLength {
        caso 0:
                VVec Longitud = NumOfHoaCoeffs;
                para (m = 0; m <VVecLength; ++m){
                        VVecCoeffId [m] = m;
                }
                break;
        caso 1:
                VVecLength = NumOfHoaCoeffs - MinNumOfCoeffsForAmbHOA -
        NumOfContAddHoaChans;

                n=0
                para (m=MinNumOfCoeffsForAmbHOA; m<NumOfHoaCoeffs; ++m){
                        CoeffIdx = m+
                si (CoeffIdx isNotMemberOf ContAddHoaCoeff) {
                VVecCoeffId[n] = CoeffIdx-1;
                        n++;
                }
        }
                break;
    caso 2:
                VVecLength = NumOfHoaCoeffs - MinNumOfCoeffsForAmbHOA;
                para (m=0; m< VVecLength; ++m){
                        VVecCoeffId[m] = m + MinNumOfCoeffsForAmbHOA;
                }
        break;
        caso 3:
        VVecLength = NumOfHoaCoeffs - NumOfContAddHoaChans;
        n=0
        para (m=0; m<NumOfHoaCoeffs; ++m){
                        c = m+1;
                si (c isNotMemberOf ContAddHoaCoeff){
                        VVecCoeffId[n] = c-1;
                        n++;
                }
        }
    }
```

Sintaxis	No. de bits	Mnemotécnico

VVectorData(i)
{
si (NbitsQ(k)[i] == 5){
para (m=0; m< VVecLength; ++m){
VVec[i][VVecCoeffId[m]](k) = (**VecVal**/128.0)- 1.0;
}
sino si (NbitsQ(k)[i] >= 6){
para (m=0; m< VVecLength; ++m){
huffIdx = *huffSelect*(VVecCoeffId[m], PFlag[i],
CbFlag[i]);
	8	**uimsbf**

```
        cid = huffDecode{NbitsQ[i], huffIdx, huffVal);                    dinámica    huffDecode
        aVal[i][m] = 0.0;
        si (cid > 0) {
            aVal[i][m] = sgn = (sgnVal * 2)-1;                            1           bslbf
            si (cid > 1){
               aVal[i][m] = sgn * (2.0^(cid-1) + intAddVal);              cid-1       uimsbf
              }
        }
        VVec[i][WecCoeffId[m]](k) = aVal[i][m] *(2^(16
-
              NbitsQ(k)[i])*aVal[i][m])/2^15;
         si (PFlag(k)[i] ==1){
            VVec[i] [VVecCoeffId[m]](k)+=
            VVec[i][WecCoeffId[m]](k-1)
          }
    }
}
```

[0314] En la tabla de sintaxis anterior, la primera instrucción de conmutación con los cuatro casos (casos 0-3) proporciona una forma mediante la cual se determina la longitud del vector V^T_{DIST} en términos del número (VVecLength) y los índices de coeficientes (VVecCoeffId). El primer caso, caso 0, indica que se especifican todos los coeficientes para los vectores V^T_{DIST} (NumOfHoaCoeffs). El segundo caso, caso 1, indica que solo se especifican aquellos coeficientes del vector $V^T_{DIST\ que}$ corresponden al número mayor que un MinNumOfCoeffsForAmbHOA, que puede indicar lo que se denomina $(N_{DIST} + 1)^2 - (N_{BG} + 1)^2$ anteriormente. Además, estos coeficientes NumOfContAddAmbHoaChan identificados en ContAddAmbHoaChan se suprimen. La lista ContAddAmbHoaChan especifica canales adicionales (donde "canales" se refiere a un coeficiente particular correspondiente a un cierta combinación de orden y suborden) correspondiente a un orden que excede el orden MinAmbHoaOrder. El tercer caso, caso 2, indica que se especifican aquellos coeficientes del vector V^T_{DIST} que corresponden al número mayor que un MinNumOfCoeffsForAmbHOA, que puede indicar lo que se denomina $(N_{DIST} + 1)^2 - (N_{BG} + 1)^2$ anteriormente. El cuarto caso, caso 3, indica que se especifican aquellos coeficientes del vector V^T_{DIST} que quedan después de eliminar los coeficientes identificados por NumOfContAddAmbHoaChan. Tanto la VVecLength como la lista VVecCoeffId son válidas para todos los VVectors dentro de HOAFrame.

[0315] Después de esta instrucción de conmutación, la decisión de realizar una decuantización uniforme puede ser controlada por NbitsQ (o, como se ha designado anteriormente *nbits*) que, si es igual a 5, se realiza una decuantización escalar uniforme de 8 bits. Por el contrario, un valor de NbitsQ mayor o igual a 6 puede resultar en la aplicación de la decodificación de Huffman. El valor *cid* mencionado anteriormente puede ser igual a los dos bits menos significativos del valor NbitsQ. El modo de predicción tratado anteriormente se designa como PFlag en la tabla de sintaxis anterior, mientras que el bit de información HT se designa como CbFlag en la tabla de sintaxis anterior. La sintaxis restante especifica cómo ocurre la decodificación de una manera sustancialmente similar a la descrita anteriormente. Varios ejemplos del flujo de bits 21 que se adapta a cada uno de los diversos casos mencionados anteriormente se describen con más detalle a continuación con respecto a las Figs. 10H (i)-10O(ii).

[0316] 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 síntesis basada en vectores 27 para reconstruir los coeficientes HOA 11'. La unidad de reconstrucción basada en vectores 92 puede incluir una unidad de cuantización 74, una unidad de interpolación espacio-temporal 76, una unidad de formulación del primer plano 78, una unidad de decodificación psicoacústica 80, una unidad de formulación de coeficientes HOA 82 y una unidad de reordenación 84.

[0317] La unidad de cuantización 74 puede representar una unidad configurada para operar de manera recíproca con la unidad de cuantización 52 que se muestra en el ejemplo de la Fig. 4 a fin de decuantizar los vectores V[*k*] de primer plano codificados y, de este modo, generar V[*k*] vectores de primer plano reducidos 55_k. La unidad de decuantización 74 puede, en algunos ejemplos, realizar una forma de decodificación de entropía y de decuantización escalar de una manera recíproca a la descrita anteriormente con respecto a la unidad de cuantización 52. La unidad de decuantización 74 puede enviar los vectores V[*k*] de primer plano reducidos 55_k a la unidad de reordenación 84.

[0318] La unidad de decodificación psicoacústica 80 puede operar de manera recíproca con la unidad de codificación de audio psicoacústica 40 que se muestra en el ejemplo de la Fig. 4 para descifrar los coeficientes HOA codificados 59 y las señales nFG 61 codificadas y, de este modo, generar coeficientes HOA ambientales compensados de energía 47' y las señales nFG interpoladas 49' (que también se pueden denominar objetos de audio nFG interpolados 49'). La unidad de decodificación psicoacústica 80 puede pasar los coeficientes HOA

ambientales compensados de energía 47' a la unidad de formulación de coeficientes HOA 82 y las señales nFG 49' a la reordenación 84.

[0319] La unidad de reordenación 84 puede representar una unidad configurada para operar de una manera similar a la descrita anteriormente con respecto a la unidad de reordenación 34. La unidad de reordenación 84 puede recibir elementos de sintaxis indicativos del orden original de los componentes en primer plano de los coeficientes HOA 11. La unidad de reordenación 84 puede, con base en estos elementos de sintaxis de reordenación, reordenar las señales nFG interpoladas 49' y los vectores V[*k*] de primer plano reducidos 55_k para generar señales nFG reordenadas 49' y vectores [*k*] de primer plano reordenados 55_k'. La unidad de reordenación 84 puede emitir las señales nFG reordenadas 49" a la unidad de formulación del primer plano 78 y los vectores V[*k*] de primer plano reordenados 55_k' a la unidad de interpolación espacio-temporal 76.

[0320] 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*] reordenados 55_k' y realizar la interpolación espacio-temporal con respecto a los vectores V[*k*] reordenados de primer plano 55_k' y los vectores V[*k*-1] reordenados 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 enviar los vectores V[*k*] de primer plano interpolados 55_k" a la unidad de formulación del primer plano 78.

[0321] 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*] interpolados de primer plano 55_k" y las señales nFG reordenadas 49" para generar los coeficientes HOA de primer plano 65. La unidad de formulación del primer plano 78 puede realizar una multiplicación matricial de las señales nFG reordenadas 49" por los vectores V[*k*] de primer plano interpolados 55_k".

[0322] La unidad de formulación de coeficientes HOA 82 puede representar una unidad configurada para añadir los coeficientes HOA de primer plano 65 a los canales HOA ambientales 47' para obtener los coeficientes HOA 11', donde la notación primaria refleja que estos 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 transmisión sobre un medio de transmisión con pérdidas, cuantización u otras operaciones con pérdidas.

[0323] De esta manera, las técnicas pueden permitir que un dispositivo de decodificación de audio, tal como el dispositivo de decodificación de audio 24, determine, a partir de un flujo de bits, información direccional cuantizada, un objeto de audio de primer plano codificado y coeficientes ambisónicos de orden superior (HOA) codificados, en el que la información direccional cuantizada y el objeto de audio de primer plano codificado representan coeficientes HOA de primer plano que describen un componente en primer plano de un campo sonoro, y en el que los coeficientes HOA ambientales codificados describen un componente ambiental del campo sonoro, decuantizan la información direccional cuantizada para generar información direccional, realizan la interpolación espacio-temporal para generar una información direccional interpolada, decodifican por audio el objeto de audio de primer plano codificado para generar un objeto de audio de primer plano y los coeficientes HOA ambientales codificados para generar coeficientes HOA ambientales, determinan los coeficientes HOA de primer plano en función de la información direccional interpolada y el objeto de audio de primer plano, y determinan los coeficientes HOA como una función de los coeficientes HOA de primer plano y los coeficientes HOA ambientales.

[0324] De esta manera, varios aspectos de las técnicas pueden permitir que un dispositivo de decodificación de audio unificado 24 conmute entre dos esquemas de descompresión diferentes. En algunos casos, el dispositivo de decodificación de audio 24 puede estar configurado para seleccionar uno de una pluralidad de esquemas de descompresión basándose en la indicación de si se genera una versión comprimida de coeficientes armónicos esféricos representativos de un campo sonoro a partir de un objeto de audio sintético y descomprimir la versión comprimida de los coeficientes armónicos esféricos utilizando el esquema seleccionado de la pluralidad de esquemas de descompresión. En estos y otros casos, el dispositivo de decodificación de audio 24 comprende un decodificador integrado.

[0325] En algunos casos, el dispositivo de decodificación de audio 24 puede configurarse para obtener una indicación de si se generan coeficientes armónicos esféricos representativos de un campo sonoro a partir de un objeto de audio sintético.

[0326] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado para obtener la indicación desde un flujo de bits que almacena una versión comprimida de los coeficientes armónicos esféricos.

[0327] De esta manera, varios aspectos de las técnicas pueden permitir que el dispositivo de decodificación de audio 24 obtenga vectores que describen componentes distintos y de fondo del campo sonoro. En algunos casos, el dispositivo de decodificación de audio 24 puede estar configurado para determinar uno o más primeros vectores que describen componentes distintos del campo sonoro y uno o más segundos vectores que describen componentes de fondo del campo sonoro, tanto uno como los más primeros vectores y uno o los más segundos vectores generados al menos realizando una transformación con respecto a la pluralidad de coeficientes armónicos esféricos.

[0328] En estos y otros casos, el dispositivo de decodificación de audio 24, en el que la transformación comprende una descomposición en valores singulares que genera una matriz U representativa de los vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos.

[0329] En estos y otros casos, el dispositivo de decodificación de audio 24, en el que uno o más primeros vectores comprenden uno o más vectores $U_{DIST} * S_{DIST}$ codificados en audio que, antes de la codificación de audio, se generaron multiplicando uno o más vectores U_{DIST} codificados en audio de una matriz U por uno o más vectores S_{DIST} de una matriz S, y en el que la matriz U y la matriz S se generan al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos.

[0330] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado adicionalmente para decodificar en audio uno o más vectores $U_{DIST} * S_{DIST}$ codificados en audio para generar una versión decodificada en audio de uno o más vectores $U_{DIST} * S_{DIST}$ codificados en audio.

[0331] En estos y otros casos, el dispositivo de decodificación de audio 24, en el que uno o más primeros vectores comprenden uno o más vectores $U_{DIST} * S_{DIST}$ codificados en audio que, antes de la codificación por audio, se generaron multiplicando uno o más vectores U_{DIST} codificados en audio de una matriz U por uno o más vectores S_{DIST} de una matriz S y uno o más vectores V^T_{DIST} de una transposición de una matriz V, y en el que la matriz U y la matriz S y la matriz V se generan al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos.

[0332] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado adicionalmente para decodificar en audio uno o más vectores $U_{DIST} * S_{DIST}$ codificados en audio para generar una versión decodificada en audio de uno o más vectores $U_{DIST} * S_{DIST}$ codificados en audio.

[0333] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado además para multiplicar los vectores $U_{DIST} * S_{DIST}$ por los vectores V^T_{DIST} para recuperar los de la pluralidad de armónicos esféricos representativos de los distintos componentes del campo sonoro.

[0334] En estos y otros casos, el dispositivo de decodificación de audio 24, en el que el uno o más segundos vectores comprenden uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ codificados en audio que, antes de la codificación de audio, se generaron multiplicando los vectores U_{BG} incluidos dentro de una matriz U por los vectores S_{BG} incluidos dentro de una matriz S y luego por vectores V^T_{BG} incluidos dentro de una transposición de una matriz V, y en la que la matriz S, la matriz U y la matriz V se generaron, al menos, realizando una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos.

[0335] En estos y otros casos, el dispositivo de decodificación de audio 24, en el que el uno o más segundos vectores comprenden uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ codificados en audio que, antes de la codificación de audio, se generaron multiplicando los vectores U_{BG} incluidos dentro de una matriz U por los vectores S_{BG} incluidos dentro de una matriz S y luego por los vectores V^T_{BG} incluidos dentro de una transposición de una matriz V, en la que la matriz S, la matriz U y la matriz V se generaron al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos y en el que el dispositivo de decodificación de audio 24 está configurado adicionalmente para decodificar en audio uno o más vectores $_{BG} * S_{BG} * V^T_{BG}$ codificados en audio para generar uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$.

[0336] En estos y otros casos, el dispositivo de decodificación de audio 24, en el que uno o más primeros vectores comprenden uno o más vectores $U_{DIST} * S_{DIST}$ codificados en audio que, antes de la codificación de audio, se generaron multiplicando uno o más vectores U_{DIST} codificados en audio de una matriz U por uno o más vectores S_{DIST} de una matriz S, y uno o más vectores V^T_{DIST} de una transposición de una matriz V, en el que la matriz U, la matriz S y la matriz V se generaron al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos y en el que el dispositivo de decodificación de audio 24 está configurado además para decodificar en audio uno o más vectores $U_{DIST} * S_{DIST}$ codificados en audio para generar uno o más vectores $U_{DIST} * S_{DIST}$ y multiplicar los vectores $U_{DIST} * S_{DIS}T$ por los vectores V^T_{DIST} para recuperar los de la pluralidad de coeficientes armónicos esféricos que describen los distintos componentes del campo sonoro, en el que uno o más segundos vectores comprenden uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ codificados en audio que, antes de la codificación de audio, se generaron multiplicando los vectores U_{BG} incluidos en la matriz U por los vectores S_{BG} incluidos en la matriz S y luego por los vectores V^T_{BG} incluidos dentro de la transposición de la matriz V, y en el que el dispositivo de decodificación de audio 24 está configurado adicionalmente para decodificar en audio uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ codificados en audio para recuperar al menos una parte de la pluralidad de coeficientes armónicos esféricos que describen componentes de fondo del campo sonoro y añadir la pluralidad de coeficientes armónicos esféricos que describen los distintos componentes del campo sonoro a la, al menos parte, de la pluralidad de coeficientes armónicos esféricos que describen componentes de fondo del campo sonoro para generar una versión reconstruida de la pluralidad de coeficientes armónicos esféricos.

[0337] En estos y otros casos, el dispositivo de decodificación de audio 24, en el que uno o más primeros vectores comprenden uno o más vectores $U_{DIST} * S_{DIST}$ que, antes de la codificación de audio, se generaron multiplicando uno o más vectores U_{DIST} codificados en audio de una matriz U por uno o más vectores S_{DIST} de una matriz S y uno o más vectores V^T_{DIST} de una transposición de una matriz V, en donde la matriz U, la matriz S y la matriz V se generaron al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, y en el que el dispositivo de decodificación de audio 20 está configurado además para obtener un valor D que indica el número de vectores a extraer de un flujo de bits para formar uno o más vectores $U_{DIST} * S_{DIST}$ y el uno o más vectores V^T_{DIST}.

[0338] En estos y otros casos, el dispositivo de decodificación de audio 24, en el que uno o más primeros vectores comprenden uno o más vectores $U_{DIST} * S_{DIST}$ que, antes de la codificación de audio, se generaron multiplicando uno o más vectores U_{DIST} codificados en audio de una matriz U por uno o más vectores S_{DIST} de una matriz S y uno o más vectores V^T_{DIST} de una transposición de una matriz V, en donde la matriz U, la matriz S y la matriz V se generaron al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, y en el que el dispositivo de decodificación de audio 24 está configurado además para obtener un valor D sobre una base de trama de audio por trama de audio que indica el número de vectores que se van a extraer de un flujo de bits para formar uno o más vectores $U_{DIST} * S_{DIST}$ y el uno o más vectores V^T_{DIST}.

[0339] En estos y otros casos, el dispositivo de decodificación de audio 24, en el que la transformación comprende un análisis de componentes principales para identificar los distintos componentes del campo sonoro y los componentes de fondo del campo sonoro.

[0340] Varios aspectos de las técnicas descritas en esta descripción pueden también permitir que el dispositivo de codificación de audio 24 realice interpolación con respecto a las versiones descompuestas de los coeficientes HOA. En algunos casos, el dispositivo de decodificación de audio 24 puede configurarse para obtener coeficientes armónicos esféricos interpolados descompuestos para un segmento de tiempo, realizando al menos en parte una interpolación con respecto a una primera descomposición de una primera pluralidad de coeficientes armónicos esféricos y una segunda descomposición de una segunda pluralidad de coeficientes armónicos esféricos.

[0341] En estos y otros casos, la primera descomposición comprende una primera matriz V representativa de los vectores singulares derechos de la primera pluralidad de coeficientes armónicos esféricos.

[0342] En estos y otros ejemplos, la segunda descomposición comprende una segunda matriz V representativa de los vectores singulares derechos de la segunda pluralidad de coeficientes armónicos esféricos.

[0343] En estos y otros casos, la primera descomposición comprende una primera matriz V representativa de los vectores singulares derechos de la primera pluralidad de coeficientes armónicos esféricos y la segunda descomposición comprende una segunda matriz V representativa de los vectores singulares derechos de la segunda pluralidad de coeficientes armónicos esféricos.

[0344] En estos y otros casos, el segmento de tiempo comprende una subtrama de una trama de audio.

[0345] En estos y otros casos, el segmento de tiempo comprende una muestra de tiempo de una trama de audio.

[0346] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado para obtener una descomposición interpolada de la primera descomposición y la segunda descomposición de un coeficiente armónico esférico de la primera pluralidad de coeficientes armónicos esféricos.

[0347] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado para obtener descomposiciones interpoladas de la primera descomposición de una primera parte de la primera pluralidad de coeficientes armónicos esféricos incluidos en la primera trama y la segunda descomposición de una segunda parte de la segunda pluralidad de los coeficientes armónicos esféricos incluidos en la segunda trama, y el dispositivo de decodificación de audio 24 está configurado además para aplicar las descomposiciones interpoladas a un primer componente de tiempo de la primera parte de la primera pluralidad de coeficientes armónicos esféricos incluidos en la primera trama para generar una primera componente artificial de tiempo de la primera pluralidad de coeficientes armónicos esféricos y aplicar las respectivas descomposiciones interpoladas a una segunda componente de tiempo de la segunda parte de la segunda pluralidad de coeficientes armónicos esféricos incluidos en la segunda trama para generar un segundo componente artificial de tiempo de la segunda pluralidad de coeficientes armónicos esféricos incluidos.

[0348] En estos y otros casos, el primer componente de tiempo se genera realizando una síntesis basada en vectores con respecto a la primera pluralidad de coeficientes armónicos esféricos.

[0349] En estos y otros casos, el segundo componente de tiempo se genera realizando una síntesis basada en vectores con respecto a la segunda pluralidad de coeficientes armónicos esféricos.

[0350] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado adicionalmente para recibir el primer componente artificial de tiempo y el segundo componente artificial de tiempo, calcular las descomposiciones interpoladas de la primera descomposición para la primera parte de la primera pluralidad de coeficientes armónicos esféricos y la segunda Descomponer para la segunda parte de la segunda pluralidad de coeficientes armónicos esféricos y aplicar inversos de las descomposiciones interpoladas en el primer componente artificial de tiempo para recuperar el componente de primer tiempo y en el segundo componente artificial de tiempo para recuperar el segundo componente de tiempo.

[0351] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado para interpolar un primer componente espacial de la primera pluralidad de coeficientes armónicos esféricos y el segundo componente espacial de la segunda pluralidad de coeficientes armónicos esféricos.

[0352] En estos y otros casos, el primer componente espacial comprende una primera matriz U representativa de los vectores singulares izquierdos de la primera pluralidad de coeficientes armónicos esféricos.

[0353] En estos y otros casos, el segundo componente espacial comprende una segunda matriz U representativa de los vectores singulares izquierdos de la segunda pluralidad de coeficientes armónicos esféricos.

[0354] En estos y otros casos, el primer componente espacial es representativo de M segmentos temporales de coeficientes armónicos esféricos para la primera pluralidad de coeficientes armónicos esféricos y el segundo componente espacial es representativo de M segmentos temporales de coeficientes armónicos esféricos para la segunda pluralidad de coeficientes armónicos esféricos.

[0355] En estos y otros casos, el primer componente espacial es representativo de M segmentos temporales de coeficientes armónicos esféricos para la primera pluralidad de coeficientes armónicos esféricos y el segundo componente espacial es representativo de M segmentos temporales de coeficientes armónicos esféricos para la segunda pluralidad de coeficientes armónicos esféricos y el dispositivo de decodificación de audio 24 está configurado para interpolar los últimos N elementos del primer componente espacial y los primeros N elementos del segundo componente espacial.

[0356] En estos y otros casos, la segunda pluralidad de coeficientes armónicos esféricos es posterior a la primera pluralidad de coeficientes armónicos esféricos en el dominio del tiempo.

[0357] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado además para descomponer la primera pluralidad de coeficientes armónicos esféricos para generar la primera descomposición de la primera pluralidad de coeficientes armónicos esféricos.

[0358] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado además para descomponer la segunda pluralidad de coeficientes armónicos esféricos para generar la segunda descomposición de la segunda pluralidad de coeficientes armónicos esféricos.

[0359] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado además para realizar una descomposición en valores singulares con respecto a la primera pluralidad de coeficientes armónicos esféricos para generar una matriz U representativa de los vectores singulares izquierdos de la primera pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la primera pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la primera pluralidad de coeficientes armónicos esféricos.

[0360] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado además para realizar una descomposición en valores singulares con respecto a la segunda pluralidad de coeficientes armónicos esféricos para generar una matriz U representativa de los vectores singulares izquierdos de la segunda pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la segunda pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la segunda pluralidad de coeficientes armónicos esféricos.

[0361] En estos y otros casos, la primera y la segunda pluralidad de coeficientes armónicos esféricos representan cada una una representación de onda plana del campo sonoro.

[0362] En estos y otros casos, la primera y la segunda pluralidad de coeficientes armónicos esféricos representan cada uno uno o más objetos monoaudio mezclados entre sí.

[0363] En estos y otros casos, la primera y la segunda pluralidad de coeficientes armónicos esféricos comprenden cada uno un primer y segundo coeficientes armónicos esféricos respectivos que representan un campo sonoro tridimensional.

[0364] En estos y otros casos, la primera y la segunda pluralidad de coeficientes armónicos esféricos están asociadas cada una con al menos una función de base esférica que tiene un orden mayor que uno.

[0365] En estos y otros casos, la primera y la segunda pluralidad de coeficientes armónicos esféricos están asociadas cada una con al menos una función de base esférica que tiene un orden igual a cuatro.

[0366] En estos y otros casos, la interpolación es una interpolación ponderada de la primera descomposición y segunda descomposición, en la que los pesos de la interpolación ponderada aplicada a la primera descomposición son inversamente proporcionales a un tiempo representado por los vectores de la primera y segunda descomposición y en donde los pesos de la interpolación ponderada aplicada a la segunda descomposición es proporcional a un tiempo representado por vectores de la primera y segunda descomposición.

[0367] En estos y otros casos, los coeficientes armónicos esféricos interpolados descompuestos suavizan al menos uno de componentes espaciales y los componentes temporales de la primera pluralidad de coeficientes armónicos esféricos y la segunda pluralidad de coeficientes armónicos esféricos.

[0368] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado Us[n] = HOA(n) * (V_vec[n])-1 y obtener un escalar.

[0369] En estos y otros casos, la interpolación comprende una interpolación lineal. En estos y otros casos, la interpolación comprende una interpolación no lineal. En estos y otros casos, la interpolación comprende una interpolación de coseno. En estos y otros casos, la interpolación comprende una interpolación de coseno ponderada. En estos y otros casos, la interpolación comprende una interpolación cúbica. En estos y otros casos, la interpolación comprende una interpolación de spline adaptativa. En estos y otros casos, la interpolación comprende una interpolación de curvatura mínima.

[0370] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado además para generar un flujo de bits que incluye una representación de los coeficientes armónicos esféricos interpolados descompuestos para el segmento de tiempo y una indicación de un tipo de interpolación.

[0371] En estos y otros casos, la indicación comprende uno o más bits que se corresponden con el tipo de interpolación.

[0372] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado adicionalmente para obtener un flujo de bits que incluye una representación de los coeficientes armónicos esféricos interpolados descompuestos para el segmento de tiempo y una indicación de un tipo de interpolación.

[0373] En estos y otros casos, la indicación comprende uno o más bits que se corresponden con el tipo de interpolación.

[0374] Varios aspectos de las técnicas pueden, en algunos casos, permitir además que el dispositivo de decodificación de audio 24 sea configurado para obtener 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.

[0375] En estos y otros casos, la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un campo que especifica un modo de predicción utilizado al comprimir el componente espacial.

[0376] En estos y otros casos, la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, información de tabla Huffman que especifica una tabla Huffman usada al comprimir el componente espacial.

[0377] En estos y otros casos, la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un campo que indica un valor que expresa un tamaño del paso de cuantización o una variable del mismo utilizado al comprimir el componente espacial.

[0378] En estos y otros casos, el valor comprende un valor nbits.

[0379] En estos y otros casos, el flujo de bits comprende una versión comprimida de una pluralidad de componentes espaciales del campo sonoro de la que está incluida la versión comprimida del componente espacial y el valor expresa el tamaño del paso de cuantización o una variable del mismo utilizada al comprimir la pluralidad de componentes espaciales.

[0380] En estos y otros casos, la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un código Huffman para representar un identificador de categoría que identifica una categoría de compresión a la que corresponde el componente espacial.

[0381] En estos y otros casos, la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un bit de signo que identifica si el componente espacial es un valor positivo o un valor negativo.

[0382] En estos y otros casos, la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un código Huffman para representar un valor residual del componente espacial.

[0383] En estos y otros casos, el dispositivo comprende un dispositivo de decodificación de audio.

[0384] Varios aspectos de las técnicas también pueden permitir que el dispositivo de decodificación de audio 24 identifique un libro de códigos Huffman para usar al descomprimir una versión comprimida de un componente espacial de una pluralidad de componentes espaciales comprimidos basado en un orden de la versión comprimida del componente espacial con relación a componentes restantes de la pluralidad de componentes espaciales comprimidos, generando el componente espacial al realizar una síntesis basada en vectores con respecto a una pluralidad de coeficientes armónicos esféricos.

[0385] En estos y otros casos, el dispositivo de decodificación de audio 24 está configurado para obtener un flujo de bits que comprende la versión comprimida de un componente espacial de un campo sonoro y descomprimir la versión comprimida del componente espacial utilizando, al menos en parte, el libro de códigos Huffman identificado para obtener el componente espacial.

[0386] En estos y otros casos, la versión comprimida del componente espacial está representada en el flujo de bits utilizando, al menos en parte, un campo que especifica un modo de predicción utilizado al comprimir el componente espacial y el dispositivo de decodificación de audio 24 está configurado para descomprimir la versión comprimida del componente espacial basada, al menos en parte, en el modo de predicción para obtener el componente espacial.

[0387] En estos y otros casos, la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, información de tabla Huffman que especifica una tabla Huffman usada al comprimir el componente espacial y el dispositivo de decodificación de audio 24 está configurado para descomprimir la versión comprimida del componente espacial basado, al menos en parte, en la información de la tabla Huffman.

[0388] En estos y otros casos, la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un campo que indica un valor que expresa un tamaño del paso de cuantización o una variable del mismo utilizado al comprimir el componente espacial y el dispositivo de decodificación de audio 24 está configurado para descomprimir la versión comprimida del componente espacial basado, al menos en parte, en el valor.

[0389] En estos y otros casos, el valor comprende un valor *nbits*.

[0390] En estos y otros casos, el flujo de bits comprende una versión comprimida de una pluralidad de componentes espaciales del campo sonoro de la que está incluida la versión comprimida del componente espacial, el valor expresa el tamaño del paso de cuantización o una variable del mismo utilizado al comprimir la pluralidad de componentes espaciales y el dispositivo de decodificación de audio 24 está configurado para descomprimir la pluralidad de la versión comprimida del componente espacial basado, al menos en parte, en el valor.

[0391] En estos y otros casos, la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un código Huffman para representar un identificador de categoría que identifica una categoría de compresión a la que corresponde el componente espacial y el dispositivo de decodificación de audio 24 está configurado para descomprimir la versión comprimida del componente espacial basado, al menos en parte, en el código Huffman.

[0392] En estos y otros casos, la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un bit de signo que identifica si el componente espacial es un valor positivo o negativo y el dispositivo de decodificación de audio 24 está configurado para descomprimir la versión comprimida del componente espacial basado, al menos en parte, en el bit de signo.

[0393] En estos y otros casos, la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un código Huffman para representar un valor residual del componente espacial y el dispositivo de decodificación de audio 24 está configurado para descomprimir la versión comprimida del componente espacial basado, al menos en parte, en el código Huffman incluido en el libro de códigos Huffman identificado.

[0394] En cada una de las diversas instancias descritas anteriormente, debe entenderse que el dispositivo de decodificación de audio 24 puede realizar un procedimiento o comprender de otro modo medios para realizar cada paso del procedimiento para el que está configurado el dispositivo de decodificación de audio 24. En algunos casos, estos medios pueden comprender uno o más procesadores. En algunos casos, 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 el cual el dispositivo de decodificación de audio 24 se ha configurado para realizar.

[0395] La Fig. 6 es un diagrama de flujo que ilustra el funcionamiento ejemplar de una unidad de análisis de contenido de un dispositivo de codificación de audio, tal como la unidad de análisis de contenido 26 que se muestra en el ejemplo de la Fig. 4, en la realización de diversos aspectos de las técnicas descritas en esta descripción.

[0396] La unidad de análisis de contenido 26 puede, al determinar si los coeficientes HOA 11 representativos de un campo sonoro se generan a partir de un objeto de audio sintético, obtener un entramado de coeficientes HOA (93), que puede ser de tamaño 25 por 1024 para una representación de cuarto orden (es decir, N = 4). Después de obtener los coeficientes HOA entramados (que también se pueden denominar en este documento como una matriz de SHC entramada 11 y las matrices de SHC entramadas subsiguientes pueden designarse como matrices de SHC entramadas 27B, 27C, etc.), la unidad de análisis de contenido 26 puede entonces excluir el primer vector de los coeficientes HOA entramados 11 para generar coeficientes HOA entramados reducidos (94).

[0397] La unidad de análisis de contenido 26 puede entonces predecir el primer vector distinto de cero de los coeficientes HOA entramados reducidos a partir de los vectores restantes de los coeficientes HOA entramados reducidos (95). Después de predecir el primer vector distinto de cero, la unidad de análisis de contenido 26 puede obtener un error basado en el primer vector distinto de cero predicho y el vector distinto de cero real (96). Una vez obtenido el error, la unidad de análisis de contenido 26 puede calcular una relación basada en una energía del primer vector distinto de cero real y del error (97). La unidad de análisis de contenido 26 puede entonces comparar esta relación con un umbral (98). Cuando la relación no excede el umbral ("NO", 98), la unidad de análisis de contenido 26 puede determinar que la matriz de SHC entramada 11 se genera a partir de un registro e indicar en el flujo de bits que la representación codificada correspondiente de la matriz SHC 11 se generó a partir de una grabación (100, 101). Cuando la relación excede el umbral ("SÍ", 98), la unidad de análisis de contenido 26 puede determinar que la matriz de SHC entramada 11 se genera a partir de un objeto de audio sintético e indica en el flujo de bits que la representación codificada correspondiente de la matriz SHC 11 se generó a partir de un objeto de audio sintético (102, 103). En algunos casos, cuando se genera la matriz de SHC entramada 11 a partir de una grabación, la unidad de análisis de contenido 26 pasa la matriz de SHC entramada 11 a la unidad de síntesis basada en vectores 27 (101). En algunos casos, cuando la matriz de SHC entramada 11 se generó a partir de un objeto de audio sintético, la unidad de análisis de contenido 26 pasa la matriz de SHC entramada 11 a la unidad de síntesis basada en la dirección 28 (104).

[0398] La Fig. 7 es un diagrama de flujo que ilustra una operación a modo de ejemplo de un dispositivo de codificación de audio, tal como el dispositivo de codificación de audio 20 que se muestra en la Fig. 4, en la realización de diversos aspectos de las técnicas de síntesis basadas en vectores descritas en esta descripción. 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 LIT 30, que puede aplicar una LIT con respecto a los coeficientes HOA para dar salida a los coeficientes HOA transformados (por ejemplo, en el caso de SVD, los coeficientes HOA transformados pueden comprender los vectores US[*k*] 33 y los vectores V[*k*] 35) (107).

[0399] 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 determinar al menos un parámetro basado en un análisis de los coeficientes HOA transformados 33/35 (108).

[0400] El dispositivo de codificación de audio 20 puede entonces invocar la unidad de reordenación 34, que puede reordenar los coeficientes HOA transformados (que, de nuevo en el contexto de SVD, pueden referirse a los vectores US[*k*] 33 y los vectores V[*k*] 35 basados 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 u 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 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 fondo (N_{BG}) y el número (nBGa) y los índices (i) de canales BG HOA adicionales a enviar (que pueden designarse colectivamente como información de canal de fondo 43 en el ejemplo de la Fig. 4) (109).

[0401] El dispositivo de codificación de audio 20 también puede invocar la unidad de selección de fondo 48. La unidad de selección de fondo 48 puede determinar coeficientes HOA de fondo o ambientales 47 basados en la información de canal de fondo 43 (110). El dispositivo de codificación de audio 20 puede invocar adicionalmente la unidad de selección del primer plano 36, la cual puede seleccionar las de los vectores US[*k*] reordenados 33' y los vectores V[*k*] reordenados 35 que representan componentes en primer plano o distintos del campo sonoro basados en nFG 45 (que puede representar uno o más índices que identifican estos vectores de primer plano) (112).

[0402] 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 debido a la eliminación de varios de los canales HOA por la unidad de selección de fondo 48 (114) y generar así coeficientes HOA ambientales compensados de energía 47'.

[0403] El dispositivo de codificación de audio 20 también invoca entonces 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 "señales nFG interpoladas 49") y la información direccional restante en primer plano 53 (que también se puede denominar 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 vectores V[*k*] de primer plano restantes basados en la información de canal de fondo 43 para obtener información direccional reducida en primer plano 55 (que también se puede denominar vectores de primer plano reducidos V[*k*] 55) (118).

[0404] 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*] 57 de primer plano codificados.

[0405] 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 de energía 47' y las señales nFG interpoladas 49' para generar coeficientes HOA codificados 59 y señales nFG codificadas 61. El dispositivo de codificación de audio puede entonces invocar la unidad de generación del flujo de bits 42. La unidad de generación del flujo de bits 42 puede generar el flujo de bits 21 basado en la información direccional codificada en primer plano 57, los coeficientes HOA ambientales codificados 59, las señales nFG codificadas 61 y la información de canal de fondo 43.

[0406] La Fig. 8 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. 5, al realizar diversos aspectos de las técnicas descritas en esta descripción. 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 este 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 este flujo de bits para recuperar la información anotada anteriormente, pasando esta información a la unidad de reconstrucción basada en vectores 92.

[0407] En otras palabras, la unidad de extracción 72 puede extraer la información direccional codificada en primer plano 57 (que, de nuevo, también se puede denominar 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 señales nFG 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).

[0408] El dispositivo de decodificación de audio 24 puede invocar adicionalmente la unidad de cuantización 74. La unidad de cuantización 74 puede decodificar por entropía y decuantizar 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 psicoacústica 80 puede decodificar los coeficientes HOA ambientales codificados 59 y las señales de primer plano codificadas 61 para obtener coeficientes HOA ambientales compensados de energía 47' y las señales de primer plano interpoladas 49' (138). La unidad de decodificación psicoacústica 80 puede pasar los coeficientes HOA ambientales compensados de energía 47' a la unidad de formulación de coeficientes HOA 82 y las señales nFG 49' a la unidad de reordenación 84.

[0409] La unidad de reordenación 84 puede recibir elementos de sintaxis indicativos del orden original de los componentes en primer plano de los coeficientes HOA 11. La unidad de reordenación 84 puede, con base en estos elementos de sintaxis de reordenación, reordenar las señales nFG 49' interpoladas y los vectores de primer plano V[*k*] reducidos 55_k para generar señales nFG reordenadas 49' y vectores V[*k*] de primer plano reordenados 55_k' (140). La unidad de reordenación 84 puede emitir las señales nFG reordenadas 49" a la unidad de formulación del primer plano 78 y los vectores V[*k*] de primer plano reordenados 55_k' a la unidad de interpolación espacio-temporal 76.

[0410] 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" (142). La unidad de interpolación espacio-temporal 76 puede enviar los vectores V[*k*] de primer plano interpolados 55_k" a la unidad de formulación del primer plano 78.

[0411] 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 señales de primer plano interpoladas 49" por la información direccional de primer plano interpolada 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 canales ambientales HOA 47' para obtener los coeficientes HOA 11' (146).

[0412] Las Figs. 9A-9L son diagramas de bloques que ilustran varios aspectos del dispositivo de codificación de audio 20 del ejemplo de la Fig. 4 con mayor detalle. La Fig. 9A es un diagrama de bloques que ilustra la unidad LIT 30 del dispositivo de codificación de audio 20 con más detalle. Como se muestra en el ejemplo de la Fig. 9A, la unidad LIT 30 puede incluir múltiples transformadas reversibles lineales diferentes 200-200N. La unidad LIT 30 puede incluir, para proporcionar unos pocos ejemplos, una transformación 200A ("SVD 200A") de una descomposición en valores singulares (SVD 200A), una transformación de análisis de componentes principales (PCA) 200B ("PCA 200B"), una transformada de Karhunen-Loève (KLT) 200C ("KLT 200C"), una transformada de Fourier rápida (FFT) 200D ("FFT 200D") y una transformada de coseno discreta (DCT) 200N ("DCT 200N"). La unidad LIT 30 puede invocar cualquiera de estas transformadas lineales reversibles 200 para aplicar la transformación respectiva con respecto a los coeficientes HOA 11 y generar los coeficientes HOA transformados respectivos 33/35.

[0413] Aunque se describe que se realiza directamente con respecto a los coeficientes HOA 11, la unidad LIT 30 puede aplicar las transformadas lineales reversibles 200 a derivados de los coeficientes HOA 11. Por ejemplo, la unidad LIT 30 puede aplicar la SVD 200 con respecto a una matriz de densidad espectral de potencia derivada de los coeficientes HOA 11. La matriz de densidad espectral de potencia se puede designarse como PSD y se obtiene mediante la multiplicación matricial de la transposición del hoaFrame al hoaFrame, como se indica en el pseudocódigo que sigue a continuación. La notación HoaFrame se refiere a una trama de los coeficientes HOA 11.

[0414] La unidad LIT 30 puede, después de aplicar la SVD 200 (svd) al PSD, puede obtener una matriz $S[k]^2$ (S_squared) y una matriz V[*k*]. La matriz $S[k]^2$ puede designar una matriz S[*k*] al cuadrado, con lo que la unidad LIT 30 (o, alternativamente, la unidad SVD 200 como un ejemplo) puede aplicar una operación de raíz cuadrada a la matriz $S[k]^2$ para obtener la matriz S[*k*]. La unidad SVD 200 puede, en algunos casos, realizar la cuantización con respecto a la matriz V[*k*] para obtener una matriz V[*k*] cuantizada (que puede estar designada como matriz V[*k*]'). La unidad 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 LIT 30 puede obtener a continuación la pseudoinversa (pinv) de la matriz SV[*k*]' y luego multiplicar los coeficientes HOA 11 por la pseudoinversa de la matriz SV[*k*]' para obtener la matriz U[*k*]. Lo anterior puede estar representado por el siguiente pseudocódigo:

```
PSD = hoaFrame' *hoaFrame;

[V, S_squared] = svd(PSD,'econ');

S = sqrt(S_squared);

U = hoaFrame * pinv(S*V');
```

[0415] Mediante la realización de SVD con respecto a la densidad espectral de potencia (PSD) de los coeficientes HOA en lugar de los coeficientes en sí, la unidad LIT 30 puede reducir potencialmente la complejidad computacional de realizar la SVD en términos de uno o más ciclos de procesador y espacio de almacenamiento, mientras que 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 de cálculo porque la SVD se realiza sobre una matriz F*F (con F el número de coeficientes HOA). En comparación con una matriz M*F con M es la longitud de tramas, es decir, 1024 o más muestras. La complejidad de una SVD puede ahora, a través de 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 grande-O de la complejidad computacional común a las ciencias informáticas).

[0416] La Fig. 9B es un diagrama de bloques que ilustra la unidad de cálculo de parámetros 32 del dispositivo de codificación de audio 20 con más detalle. La unidad de cálculo de parámetros 32 puede incluir una unidad de análisis de energía 202 y una unidad de correlación cruzada 204. La unidad de análisis de energía 202 puede

realizar el análisis de energía descrito anteriormente con respecto a uno o más de los vectores US[*k*] 33 y los vectores V[*k*] 35 para generar uno o más de los parámetros de correlación (*R*), los parámetros de propiedades direccionales (*θ, φ, r*) y la propiedad de la energía (*e*) para uno o más de la trama actual (*k*) o de la trama anterior (*k*-1). Del mismo modo, la unidad de correlación cruzada 204 puede realizar la correlación cruzada descrita anteriormente con respecto a uno o más de los vectores US[*k*] 33 y los vectores V[*k*] 35 para generar uno o más de los parámetro de correlación (*R*), los parámetros de propiedades direccionales (*θ, φ, r*) y la propiedad de la energía (*e*) para una o más de la trama actual (*k*) o de la trama anterior (*k*-1). La unidad de cálculo de parámetros 32 puede emitir los parámetros de las tramas actuales 37 y los parámetros de las tramas anteriores 39.

[0417] La Fig. 9C es un diagrama de bloques que ilustra la unidad de reordenación 34 del dispositivo de codificación de audio 20 con más detalle. La unidad de reordenación 34 incluye una unidad de evaluación de parámetros 206 y una unidad de reordenación de vectores 208. La unidad de evaluación de parámetros 206 representa una unidad configurada para evaluar los parámetros de las tramas anteriores 39 y los parámetros de la trama actual 37 de la manera descrita anteriormente para generar índices de reordenación 205. Los índices de reordenación 205 incluyen índices que identifican cómo se reordenarán los vectores de los vectores US[*k*] 33 y los vectores de los vectores V[*k*] 35 (por ejemplo, por pares de índice con el primer índice del par que identifica el índice de la ubicación actual del vector y el segundo índice del par que identifica la ubicación reordenada del vector). La unidad de reordenación de vectores 208 representa una unidad configurada para reordenar los vectores US[*k*] 33 y los vectores V[*k*] 35 de acuerdo con los índices de reordenación 205. La unidad de reordenación 34 puede emitir los vectores US[*k*] reordenados 33' y los vectores V[*k*] reordenados 35', al pasar también los índices de reordenación 205 como uno o más elementos de sintaxis a la unidad de generación del flujo de bits 42.

[0418] La Fig. 9D es un diagrama de bloques que ilustra la unidad de análisis del campo sonoro 44 del dispositivo de codificación de audio 20 con más detalle. Como se muestra en el ejemplo de la Fig. 9D, la unidad de análisis del campo sonoro 44 puede incluir una unidad de análisis de los valores singulares 210A, una unidad de análisis de energía 210B, una unidad de análisis espacial 210C, una unidad de análisis de enmascaramiento espacial 210D, una unidad de análisis de difusión 210E y una unidad de análisis direccional 210F. La unidad de análisis de los valores singulares 210A puede representar una unidad configurada para analizar la pendiente de la curva creada por los valores diagonales descendentes de S vectores (que forman parte de los vectores US[*k*] 33), donde los grandes valores singulares representan sonidos de primer plano o distintos sonidos y los valores singulares bajos representan componentes de fondo del campo sonoro, como se ha descrito anteriormente. La unidad de análisis de energía 210B puede representar una unidad configurada para determinar la energía de los vectores V[*k*] 35 sobre una base por vector.

[0419] La unidad de análisis espacial 210C puede representar una unidad configurada para realizar el análisis de energía espacial descrito anteriormente mediante la transformación de los coeficientes HOA 11 en el dominio espacial e identificar áreas de alta energía representativas de los componentes direccionales del campo sonoro que deben ser preservadas. La unidad de análisis de enmascaramiento espacial 210D puede representar una unidad configurada para realizar el análisis de enmascaramiento espacial de una manera similar a la del análisis de energía espacial, excepto que la unidad de análisis de enmascaramiento espacial 210D puede identificar áreas espaciales que están enmascaradas por sonidos de energía superior espacialmente próximos. La unidad 210E de análisis de difusión puede representar una unidad configurada para realizar el análisis de difusión descrito anteriormente con respecto a los coeficientes HOA 11 para identificar áreas de energía difusa que pueden representar componentes de fondo del campo sonoro. La unidad de análisis direccional 210F puede representar una unidad configurada para realizar el análisis direccional anotado anteriormente que implica el cálculo de los vectores VS[*k*], y el cuadrado y la suma de cada entrada de cada uno de estos vectores VS[*k*] para identificar un cociente de direccionalidad. La unidad de análisis direccional 210F puede proporcionar este cociente de direccionalidad para cada uno de los vectores VS[*k*] a la unidad de identificación de fondo/primer plano (BG/FG) (ID) 212.

[0420] La unidad de análisis del campo sonoro 44 puede incluir también la unidad de ID BG/FG 212, que puede representar una unidad configurada para determinar el número total de canales de primer plano (nFG) 45, el orden del campo sonoro de fondo (N_{BG}) y el número (nBGa) y los índices (i) de canales BG HOA adicionales para enviar (que se pueden designar colectivamente como información de canal de fondo 43 en el ejemplo de la Fig. 4) basado en cualquier combinación de la salida de análisis por cualquier combinación de unidades de análisis 210-210F. La unidad de ID de BG/FG 212 puede determinar el nFG 45 y la información de canal de fondo 43 para alcanzar la velocidad de bits objetivo 41.

[0421] La Fig. 9E es un diagrama de bloques que ilustra la unidad de selección del primer plano 36 del dispositivo de codificación de audio 20 con más detalle. La unidad de selección del primer plano 36 incluye una unidad de análisis de vectores 214 que puede analizar o extraer de otro modo los vectores US[*k*] de primer plano 49 y los vectores V[*k*] de primer plano 51_k identificados por el elemento de sintaxis nFG 45 a partir de los vectores US[*k*] reordenados 33' y los vectores V[*k*] reordenados 35'. La unidad de análisis de vectores 214 puede analizar los diversos vectores representativos de los componentes en primer plano del campo sonoro identificados por la unidad de análisis del campo sonoro 44 y especificados por el elemento de sintaxis de nFG 45 (que también se puede denominar información de canal de primer plano 45). Como se muestra en el ejemplo de la Fig. 9E, la unidad de análisis de vectores 214 puede seleccionar, en algunos casos, vectores no consecutivos dentro de los vectores US[*k*]

de primer plano 49 y los vectores V[*k*] de primer plano 51_k para representar los componentes en primer plano del campo sonoro. Además, la unidad de análisis de vectores 214 puede seleccionar, en algunos casos, los mismos vectores (posición) de los vectores US[*k*] de primer plano 49 y los vectores V[*k*] de primer plano 51_k para representar los componentes en primer plano del campo sonoro.

[0422] La Fig. 9F es un diagrama de bloques que ilustra la unidad de selección de fondo 48 del dispositivo de codificación de audio 20 con más detalle. La unidad de selección de fondo 48 puede determinar coeficientes HOA de fondo o ambientales 47 basados en la información de canal de fondo (por ejemplo, el campo sonoro de fondo (N_{BG}) y el número (nBGa) y los índices (i) de canales BG HOA adicionales para enviar). Por ejemplo, cuando N_{BG} es igual a uno, la unidad de selección de fondo 48 puede seleccionar los coeficientes HOA 11 para cada muestra de la trama de audio que tiene un orden igual o inferior a uno. La unidad de selección de fondo 48 puede, en este ejemplo, seleccionar entonces los coeficientes HOA 11 que tienen un índice identificado por uno de los índices (i) como coeficientes adicionales de BG HOA, en donde el nBGa se proporciona a la unidad de generación del flujo de bits 42 que se especificará 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 la Fig. 5, pueda analizar los coeficientes BG HOA 47 del flujo de bits 21. La unidad de selección de fondo 48 puede emitir entonces los coeficientes HOA ambientales 47 a la unidad de compensación de energía 38. Los coeficientes HOA ambientales 47 pueden tener D dimensiones: $M \times [(N_{BG}+1)^2+nBGa]$.

[0423] La Fig. 9G es un diagrama de bloques que ilustra la unidad de compensación de energía 38 del dispositivo de codificación de audio 20 con más detalle. 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 debido a la eliminación de varios de los canales HOA por la unidad de selección de fondo 48. La unidad de compensación de energía 38 puede incluir una unidad de determinación de energía 218, una unidad de análisis de energía 220 y una unidad de amplificación de energía 222.

[0424] La unidad 218 de determinación de energía puede representar una unidad configurada para identificar el RMS para cada fila y/o columna de una o más de la matriz reordenada US[*k*] 33' y la matriz reordenada V[*k*] 35'. La unidad de determinación de energía 38 también puede identificar el RMS para cada fila y/o columna de uno o más de los canales de primer plano seleccionados, los cuales pueden incluir las señales nFG 49 y los vectores V[*k*] de primer plano 51_k y los coeficientes HOA ambientales de orden reducido 47. El RMS para cada fila y/o columna de una o más de la matriz US[*k*] reordenada 33' y la matriz V[*k*] reordenada 35' puede almacenarse en un vector indicado $RMS_{COMPLETO}$, mientras que el RMS para cada fila y/o la columna de una o más de las señales nFG 49, los vectores V[*k*] primer plano 51_k y los coeficientes HOA ambientales de orden reducido 47 pueden almacenarse en un vector indicado $RMS_{REDUCIDO}$.

[0425] En algunos ejemplos, para determinar cada RMS de las respectivas filas y/o columnas de una o más de la matriz US[*k*] reordenada 33 ', la matriz V[*k*] reordenada 35', las señales nFG 49, el primer plano V[*k*] Vectores 51_k, y los coeficientes HOA ambientales de orden reducido 47, la unidad de determinación de energía 218 puede aplicar primero unos coeficiente de armónicos esféricos (SHC) de referencia a las columnas. La aplicación del renderizador de los SHC de referencia por la unidad 218 de determinación de energía permite la determinación de RMS en el dominio SHC para determinar la energía del campo sonoro total descrito por cada fila y/o columna de la trama representado por filas y/o columnas de uno o más más de la matriz US[*k*] reordenada 33', la matriz V[*k*] reordenada 35', las señales nFG 49, los vectores V[*k*] 51_k de primer plano y los coeficientes HOA ambientales de orden reducido 47. La unidad de determinación de energía 38 puede pasar estos vectores $RMS_{COMPLETO}$ y $RMS_{REDUCIDO}$ a la unidad 220 de análisis de energía.

[0426] La unidad de análisis de energía 220 puede representar una unidad configurada para calcular un vector de valores de amplificación *Z*, de acuerdo con la siguiente ecuación: $Z = RMS_{COMPLETO}/RMS_{REDUCIDO}$. La unidad de análisis de energía 220 puede entonces pasar este vector de valores de amplificación Z a la unidad de amplificación de energía 222. La unidad de amplificación de energía 222 puede representar una unidad configurada para aplicar este vector de valores de amplificación Z o varias partes del mismo a una o más de las señales nFG 49, los vectores V[*k*] de primer plano 51_k y los coeficientes HOA ambientales de orden reducido 47. En algunos casos, el vector de valores de amplificación Z se aplica solo a los coeficientes HOA ambientales de orden reducido 47 por la siguiente ecuación $HOA_{BG\text{-}RED'} = HOA_{BG\text{-}RED}Z^T$, donde $HOA_{BG\text{-}RED}$ indica los coeficientes HOA ambientales de orden reducido 47, $HOA_{BG\text{-}RED}'$ indica los coeficientes HOA ambientales reducidos y reducidos en energía 47' y Z^T representa la transposición del vector Z.

[0427] La Fig. 9H es un diagrama de bloques que ilustra, con más detalle, la unidad de interpolación espacio-temporal 50 del dispositivo de codificación de audio 20 que se muestra en el ejemplo de la Fig. 4. 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 quinta trama y los vectores V[*k*-1] de primer plano 51_{k-1} para la trama anterior (de ahí el k-1) y realizar interpolación espacio-temporal para generar los vectores V[*k*] de primer plano interpolados. La unidad de interpolación espacio-temporal 50 puede incluir una unidad de interpolación V 224 y una unidad de adaptación de primer plano 226.

[0428] La unidad de interpolación V 224 puede seleccionar una parte de los vectores V[*k*] actuales de primer plano 51_k para interpolar basándose en las partes restantes de los vectores V[*k*] actuales de primer plano 51_k y los vectores anteriores V[*k*-1] de primer plano 55_{k-1}. La unidad de interpolación V 224 puede seleccionar la parte para ser una o más de las subtramas anotadas anteriormente o solo una única parte indefinida que puede variar según una base de trama a trama. La unidad de interpolación V 224 puede, en algunos casos, seleccionar una única 128 parte de muestras de las muestras 1024 de los vectores V[*k*] de primer plano actuales 51_k para interpolar. La unidad de interpolación V 224 puede entonces convertir cada uno de los vectores en los vectores V[*k*] de primer plano actuales 51_k y los vectores anteriores V[*k*-1] de primer plano 51_{k-1} para separar mapas espaciales proyectando los vectores en una esfera (utilizando una matriz de proyección tal como una matriz de diseño T). La unidad de interpolación V 224 puede entonces interpretar los vectores en V como formas en una esfera. Para interpolar las matrices V para la parte de 256 muestras, la unidad de interpolación V 224 puede interpolar entonces estas formas espaciales y luego transformarlas de nuevo a los vectores de dominio armónico esférico a través de la inversa de la matriz de proyección. Las técnicas de esta descripción pueden, de esta manera, proporcionar una transición suave entre las matrices V. La unidad de interpolación V 224 puede entonces generar los vectores restantes V[*k*] 53, que representan los vectores V[*k*] de primer plano 51_k después de haber sido modificados para eliminar la parte interpolada de los vectores V[*k*] de primer plano 51_k. La unidad de interpolación V 224 puede entonces pasar los vectores V[*k*] interpolados de primer plano 51_k' a la unidad de adaptación de nFG 226.

[0429] Cuando se selecciona una sola parte a interpolación, la unidad de interpolación V 224 puede generar un elemento sintáctico denominado CodedSpatialInterpolationTime 254, que identifica la duración o, en otras palabras, el tiempo de la interpolación (por ejemplo, en términos de un número de muestras). Cuando se selecciona una sola parte de realizar la interpolación de subtrama, la unidad de interpolación V 224 también puede generar otro elemento de sintaxis denominado SpatialInterpolationMethod 255, que puede identificar un tipo de interpolación realizada (o, en algunos casos, si la interpolación fue realizada o no). La unidad de interpolación espacio-temporal 50 puede emitir estos elementos de sintaxis 254 y 255 a la unidad de generación del flujo de bits 42.

[0430] La unidad de adaptación de nFG 226 puede representar una unidad configurada para generar las señales de nFG adaptadas 49'. La unidad de adaptación de nFG 226 puede generar las señales de nFG adaptadas 49' obteniendo primero los coeficientes HOA de primer plano mediante la multiplicación de las señales nFG 49 por los vectores V[*k*] 51_k de primer plano. Después de obtener los coeficientes HOA de primer plano, la unidad de adaptación de nFG 226 puede dividir los coeficientes HOA de primer plano por los vectores V[*k*] interpolados en primer plano 53 para obtener las señales nFG adaptadas 49' (que se pueden denominar las señales nFG interpoladas 49' dado que estas señales se derivan de los vectores V[*k*] interpolados de primer plano 51_k').

[0431] La Fig. 9I es un diagrama de bloques que ilustra, con más detalle, la unidad de reducción de coeficientes 46 del dispositivo de codificación de audio 20 que se muestra en el ejemplo de la Fig. 4. La unidad de reducción de coeficientes 46 puede representar una unidad configurada para realizar una reducción de coeficiente con respecto a los vectores V[*k*] de primer plano restantes 53 basados en la información de canal de fondo 43 para emitir los vectores de primer plano reducidos V[*k*] 55 a la unidad de cuantización 52. Los vectores de primer plano reducidos V[*k*] 55 pueden tener D dimensiones: $[(N+1)^2 - (N_{BG}+1)^2 - nBGa] \times nFG$.

[0432] La unidad de reducción de coeficientes 46 puede incluir una unidad de minimización de coeficientes 228, que puede representar una unidad configurada para reducir o minimizar de otro modo el tamaño de cada uno de los vectores V[*k*] de primer plano restantes 53 eliminando cualquier coeficiente que se contabilice en los coeficientes HOA de fondo 47 (identificados por la información de canal de fondo 43). La unidad de minimización de coeficientes 228 puede eliminar aquellos coeficientes identificados por la información de canal de fondo 43 para obtener los vectores V[*k*] de primer plano reducidos 55.

[0433] La Fig. 9J es un diagrama de bloques que ilustra, con más detalle, la unidad codificadora de audio psicoacústico 40 del dispositivo de codificación de audio 20 que se muestra en el ejemplo de la Fig. 4. La unidad codificadora de audio psicoacústico 40 puede representar una unidad configurada para realizar la codificación psicoacústica con respecto a los coeficientes HOA de fondo compensados de energía 47' y las señales nFG interpoladas 49'. Como se muestra en el ejemplo de la Fig. 9H, la unidad codificadora de audio psicoacústico 40 puede invocar múltiples instancias de un codificador de audio psicoacústico 40A-40N para codificar en audio cada uno de los canales de los coeficientes HOA de fondo compensado de energía 47' (donde un canal en este contexto hace referencia a coeficientes para todas las muestras en la trama correspondiente a una función de base esférica y orden/suborden particular) y cada señal de las señales nFG interpoladas 49'. En algunos ejemplos, la unidad codificadora de audio psicoacústico 40 instancia, o de otra manera incluye, (cuando se implementa en hardware) codificadores de audio 40A-40N de número suficiente para codificar por separado cada canal de los coeficientes HOA de fondo compensados de energía 47'(o nBGa más el número total de índices (i)) y cada señal de las señales nFG interpoladas 49' (o nFG) para un total de nBGa más el número total de índices (i) de canales HOA ambientales adicionales más nFG. Los codificadores de audio 40A-40N pueden emitir los coeficientes HOA de fondo codificados 59 y las señales nFG codificadas 61.

[0434] La Fig. 9K es un diagrama de bloques que ilustra, con más detalle, la unidad de cuantización 52 del dispositivo de codificación de audio 20 que se muestra en el ejemplo de la Fig. 4. En el ejemplo de la Fig. 9K, la

unidad de cuantización 52 incluye una unidad de cuantización uniforme 230, una unidad de nbits 232, una unidad de predicción 234, una unidad de modo de predicción 236 ("Pred Mode Unit 236"), una unidad de codificación de categoría y residuo 238 y una unidad de selección de tabla Huffman 240. La unidad de cuantización uniforme 230 representa una unidad configurada para llevar a cabo la cuantización uniforme descrita anteriormente con respecto a uno de los componentes espaciales (que puede representar cualquiera de los vectores V[k] de primer plano reducidos 55). La unidad de nbits 232 representa una unidad configurada para determinar el parámetro o valor de nbits.

[0435] La unidad de predicción 234 representa una unidad configurada para realizar la predicción con respecto al componente espacial cuantizado. La unidad de predicción 234 puede realizar la predicción realizando una sustracción de elementos del actual de los vectores V[*k*] de primer plano reducidos 55 por un vector correspondiente temporal posterior de los V[*k*] vectores de primer plano reducidos 55 (que se pueden designar como vectores reducidos en primer plano V[*k*-1] 55). El resultado de esta predicción puede ser referido como un componente espacial predicho.

[0436] La unidad de modo de predicción 236 puede representar una unidad configurada para seleccionar el modo de predicción. La unidad de selección de tabla Huffman 240 puede representar una unidad configurada para seleccionar una tabla Huffman apropiada para la codificación del cid. La unidad de modo de predicción 236 y la unidad de selección de tabla Huffman 240 pueden funcionar, como un ejemplo, de acuerdo con el siguiente pseudocódigo:

```
Para un determinado nbits, recuperar todas las tablas Huffman con nbits

    B00 = 0; B01 = 0; B _ {10} = 0; B11 = 0; // inicializar para calcular los bits esperados por modo de
    codificación
    para m = 1: (# elementos en el vector)
        // calcular el número esperado de bits para un elemento vectorial v (m)
        // sin predicción y utilizando tabla Huffman 5
        B00 = B00 + calculate_bits(v (m), HT5);
        // sin predicción y utilizando tabla Huffman {1, 2, 3}
        B01 = B01 + calculate_bits(v (m), HTq); q in {1, 2, 3}
        // calcular el número esperado de bits para el residuo de predicción e (m)
        e (m) = v (m) - vp (m); // vp (m): elemento vectorial de trama anterior
        // con predicción y uso de tabla Huffman 4
        B10 = B10 + calculate_bits(e(m), HT4);
        // con predicción y uso de Huffman Tabla 5
        B11 = B11 + calculate_bits (e (m), HT5);
     fin
     // encuentra un mejor modo de predicción y una tabla Huffman que producen bits mínimos
     // el mejor modo de predicción y la tabla Huffman son marcados por pflag y Htflag,
     respectivamente.
      [Be, id] = min ([B00 B01 B10 B11]);
     Conmutar id
              caso 1: pflag = 0; HTflag = 0;
              caso 2: pflag = 0; HTflag = 1;
              caso 3: pflag = 1; HTflag = 0;
              caso 4: pflag = 1; HTflag = 1;
     fin
```

[0437] La unidad de codificación de categoría y residuo 238 puede representar una unidad configurada para realizar la categorización y la codificación residual de un componente espacial previsto o el componente espacial cuantizado (cuando la predicción está desactivada) de la manera descrita con más detalle anteriormente.

[0438] Como se muestra en el ejemplo de la Fig. 9K, la unidad de cuantización 52 puede emitir varios parámetros o valores para su inclusión en el flujo de bits 21 o información lateral (que puede ser un flujo de bits separado del flujo de bits 21). Asumiendo que la información se especifica en la información de canal lateral, la unidad de cuantización escalar/entropía 50 puede emitir el valor de nbits como valor de nbits 233, el modo de predicción como modo de predicción 237 y la información de tabla Huffman como información de tabla Huffman 241 a unidad de generación del flujo de bits 42 junto con la versión comprimida del componente espacial (que se muestra como vectores de primer plano codificados V[*k*] 57 en el ejemplo de la Fig. 4), que en este ejemplo puede referirse al código Huffman seleccionado para codificar el cid, el bit de signo y el residuo codificado por bloques. El valor de nbits se puede especificar una vez en la información de canal lateral para todos los vectores V[*k*] codificados de primer plano 57, mientras que el modo de predicción y la información de tabla Huffman pueden especificarse para cada uno de los vectores V[*k*] codificados de primer plano 57. La parte del flujo de bits que especifica la versión comprimida del componente espacial se muestra más en el ejemplo de las Figs. 10B y/o 10C.

[0439] La Fig. 9L es un diagrama de bloques que ilustra, con más detalle, la unidad de generación del flujo de bits 42 del dispositivo de codificación de audio 20 que se muestra en el ejemplo de la Fig. 4. La unidad de generación del flujo de bits 42 puede incluir una unidad de generación de información de canal principal 242 y una unidad de generación de información de canal lateral 244. La unidad de generación de información de canal principal 242 puede generar un flujo de bits principal 21 que incluye uno o más, si no todos, los índices de reordenación 205, el elemento de sintaxis CodedSpatialInterpolationTime 254, el elemento de sintaxis SpatialInterpolationMethod 255, los coeficientes HOA de fondo codificados 59 y las señales nFG codificadas 61. La unidad de generación de información de canal lateral 244 puede representar una unidad configurada para generar un flujo de bits de canal lateral 21B que puede incluir uno o más, si no todos, del valor de nbits 233, el modo de predicción 237, la información de tabla Huffman 241 y los vectores V[*k*] de primer plano codificados 57. Los flujos de bits 21 y 21B pueden denominarse en conjunto como el flujo de bits 21. En algunos contextos, el flujo de bits 21 solo puede referirse al flujo de bits de canal principal 21, mientras que el flujo de bits 21B puede denominarse información de canal lateral 21B.

[0440] Las Figs. 10A-10O(ii) son diagramas que ilustran partes de la información del canal de bits o del canal lateral que puede especificar los componentes espaciales comprimidos con más detalle. En el ejemplo de la Fig. 10A, una parte 250 incluye un campo identificador del renderizador ("renderer ID") 251 y un campo HOADecoderConfig 252. El campo ID del renderizador 251 puede representar un campo que almacena un ID del renderizador que se ha utilizado para la mezcla del contenido HOA. El campo HOADecoderConfig 252 puede representar un campo configurado para almacenar la información para inicializar el decodificador espacial HOA.

[0441] El campo HOADecoderConfig 252 incluye además un campo de información direccional ("direction info") 253, un campo CodedSpatialInterpolationTime 254, un campo SpatialInterpolationMethod 255, un campo CodedVVecLength 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 direccional. El campo CodedSpatialInterpolationTime 254 puede representar un campo que almacena un tiempo de la interpolación espacio-temporal de las señales basadas en vectores. El campo SpatialInterpolationMethod 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 CodedVVecLength 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.

[0442] En el ejemplo de la Fig. 10B, 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 número de campos de bytes 260 y un campo de bits 261. El número de bytes del campo 260 puede representar un campo para expresar el número de bytes incluidos en la trama para especificar los componentes espaciales v1 a vn que incluyen los ceros para el campo de alineación de bytes 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-vn.

[0443] Como se muestra adicionalmente en el ejemplo de la Fig. 10B, la parte 258A puede incluir flujos de bits secundarios para v1-vn, cada una de las cuales incluye un campo de modo de predicción 262, un campo de información de tabla Huffman 263 y uno correspondiente de los componentes espaciales comprimidos v1-vn. El campo de modo de predicción 262 puede representar un campo para almacenar una indicación de si la predicción se ha realizado con respecto al correspondiente de los componentes espaciales comprimidos v1-vn. El campo de información de tabla Huffman 263 representa un campo para indicar, al menos en parte, qué tabla Huffman se va a utilizar para decodificar diversos aspectos de los correspondientes componentes espaciales comprimidos v1-vn.

[0444] 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.

[0445] La Fig. 10C es un diagrama que ilustra un ejemplo alternativo de una parte 258B de la información de canal lateral que puede especificar los componentes espaciales comprimidos con más detalle. En el ejemplo de la Fig. 10C, la parte 258B incluye una cabecera de trama 259 que incluye un campo Nbits 261. El campo Nbits 261 representa un campo que puede especificar un valor nbits identificado para su uso en la descompresión de los componentes espaciales v1-vn.

[0446] Como se muestra adicionalmente en el ejemplo de la Fig. 10C, la parte 258B puede incluir flujos de bits secundarios para v1-vn, cada uno de las cuales incluye un campo de modo de predicción 262, un campo de información de tabla Huffman 263 y uno correspondiente de los componentes espaciales comprimidos v1-vn. El campo de modo de predicción 262 puede representar un campo para almacenar una indicación de si la predicción se ha realizado con respecto al correspondiente de los componentes espaciales comprimidos v1-vn. El campo de información de tabla Huffman 263 representa un campo para indicar, al menos en parte, qué tabla Huffman se va a utilizar para decodificar diversos aspectos de los correspondientes componentes espaciales comprimidos v1-vn.

[0447] El campo Nbits 261 en el ejemplo ilustrado incluye los subcampos A 265, B 266 y C 267. En este ejemplo, A 265 y B 266 son cada uno subcampos de 1 bit, mientras que C 267 es un subcampo de 2 bits. Otros ejemplos pueden incluir subcampos de tamaño diferente 265, 266 y 267. El campo A 265 y el campo B 266 pueden representar campos que almacenan primer y segundo bits más significativos del campo Nbits 261, mientras que el campo C 267 puede representar un campo que almacena los bits menos significativos del campo Nbits 261.

[0448] La parte 258B también puede incluir un campo AddAmbHoaInfoChannel 268. El campo AddAmbHoaInfoChannel 268 puede representar un campo que almacena información para los coeficientes HOA ambientales adicionales. Como se muestra en el ejemplo de la Fig. 10C, el AddAmbHoaInfoChannel 268 incluye un campo CodedAmbCoeffIdx 246, un campo AmbCoeffIdxTransition 247. El campo CodedAmbCoeffIdx 246 puede representar un campo que almacena un índice de un coeficiente HOA ambiental adicional. El campo AmbCoeffIdxTransition 247 puede representar un campo configurado para almacenar datos indicativos de si, en esta trama, un coeficiente HOA ambiente adicional está o bien se ha desvanecido o ha desaparecido.

[0449] La Fig. 10C (i) es un diagrama que ilustra un ejemplo alternativo de una parte 258B ' de la información de canal lateral que puede especificar los componentes espaciales comprimidos con más detalle. En el ejemplo de la Fig. 10C (i), la parte 258B' incluye una cabecera de trama 259 que incluye un campo Nbits 261. El campo Nbits 261 representa un campo que puede especificar un valor nbits identificado para su uso en la descompresión de los componentes espaciales v1-vn.

[0450] Como se muestra adicionalmente en el ejemplo de la Fig. 10C (i), la parte 258B' puede incluir flujos de bits secundarios para v1-vn, cada uno de los cuales incluye un campo 263 de información de la tabla Huffman y uno correspondiente de los componentes direccionales comprimidos v1-vn sin incluir el campo de modo de predicción 262. En todos los demás aspectos, la parte 258B' puede ser similar a la parte 258B.

[0451] La Fig. 10D es un diagrama que ilustra una parte 258C del flujo de bits 21 con más detalle. La parte 258C es similar a la parte 258, excepto que la cabecera de trama 259 y la alineación de cero bytes 264 han sido eliminadas, mientras que el campo de Nbits 261 se ha añadido antes de cada uno de los flujos de bits para v1-vn, como se muestra en el ejemplo De la Fig. 10D.

[0452] La Fig. 10D (i) es un diagrama que ilustra una parte 258C 'del flujo de bits 21 con más detalle. La parte 258C 'es similar a la parte 258C excepto que la parte 258C' no incluye el campo de modo de predicción 262 para cada uno de los V vectores v1-vn.

[0453] La Fig. 10E es un diagrama que ilustra una parte 258D del flujo de bits 21 con más detalle. La parte 258D es similar a la parte 258B, excepto que se han eliminado el encabezado de trama 259 y la alineación de cero bytes 264, mientras que el campo de Nbits 261 se ha añadido antes de cada uno de los flujos de bits para v1-vn, como se muestra en el ejemplo de la Fig. 10E.

[0454] La Fig. 10E(i) es un diagrama que ilustra una parte 258D' del flujo de bits 21 con más detalle. La parte 258D' es similar a la parte 258D excepto que la parte 258D' no incluye el campo de modo de predicción 262 para cada uno de los V vectores v1-vn. A este respecto, el dispositivo de codificación de audio 20 puede generar un flujo de bits 21 que no incluye el campo de modo de predicción 262 para cada vector V comprimido, como se demuestra con respecto a los ejemplos de las Figs. 10C (i), 10D (i) y 10E (i).

[0455] La Fig. 10F es un diagrama que ilustra, de manera diferente, la parte 250 del flujo de bits 21 ilustrada en el ejemplo de la Fig. 10A. La parte 250 que se muestra en el ejemplo de la Fig. 10D, incluye un campo HOAOrder (que no se ha mostrado en el ejemplo de la Fig. 10F para facilitar la ilustración), un campo MinAmbHoaOrder (que de nuevo no se ha mostrado en el ejemplo de la Fig. 10 para facilitar la ilustración), el campo de información direccional 253, el campo CodedSpatialInterpolationTime 254, el campo SpatialInterpolationMethod 255, el campo CodedVVecLength 256 y el campo de información de ganancia 257. Como se muestra en el ejemplo de la Fig. 10F, el campo CodedSpatialInterpolationTime 253 puede comprender un campo de tres bits, el campo SpatialInterpolationMethod 255 puede comprender un campo de un bit y el campo CodedVVecLength 256 puede comprender un campo de dos bits.

[0456] La Fig. 10G es un diagrama que ilustra una parte 248 del flujo de bits 21 con más detalle. La parte 248 representa una carga útil tridimensional (3D) de codificador de voz/audio unificado (USAC) que incluye un campo HOAframe 249 (que también se puede designar como información de banda lateral, información de canal lateral o flujo de bits de canal lateral). Como se muestra en el ejemplo de la Fig. 10E, la vista ampliada del campo HOAFrame 249 puede ser similar a la parte 258B del flujo de bits 21 que se muestra en el ejemplo de la Fig. 10C. El "ChannelSideInfoData" incluye un campo ChannelType 269, que no se ha mostrado en el ejemplo de la Fig. 10C para facilitar la ilustración, el campo A 265 designado como "ba" en el ejemplo de la Fig. 10E, el campo B 266 designado como "bb" en el ejemplo de la Fig. 10E y el campo C 267 designado como "unitC" en el ejemplo de la Fig. 10E. El campo ChannelType indica si el canal es una señal basada en la dirección, una señal basada en vectores o un coeficiente HOA ambiental adicional. Entre diferentes ChannelSideInfoData hay campos AddAmbHoaInfoChannel 268 con los diferentes flujos de bits de vectores V indicados en gris (por ejemplo, "flujo de bits para v1" y "flujo de

bits para v2").

[0457] Las Figs. 10H-10O(ii) son diagramas que ilustran otras varias partes de ejemplo 248H-248O del flujo de bits 21 junto con las partes de HOAconfig correspondientes 250H-2500 con más detalle. Las Figs. 10H(i) y 10H(ii) ilustran un primer flujo de bits de ejemplo 248H y la parte de HOAconfig correspondiente 250H que se ha generado para corresponder con el caso 0 en el pseudocódigo anterior. En el ejemplo de la Fig. 10H(i), la parte de HOAconfig 250H incluye un elemento de sintaxis CodedVVecLength 256 establecido para indicar que todos los elementos de un vector V están codificados, por ejemplo, todos los 16 elementos vectoriales V. La parte de HOAconfig 250H también incluye un elemento de sintaxis SpatialInterpolationMethod 255 establecido para indicar que la función de interpolación de la interpolación espacio-temporal es un coseno elevado. La parte de HOAconfig 250H además incluye un conjunto CodedSpatialInterpolationTime 254 para indicar una duración de muestra interpolada de 256. La parte de HOAconfig 250H incluye además un elemento de sintaxis MinAmbHoaOrder 150 establecido para indicar que el orden de MinimumHOA del contenido de HOA ambiental es uno, donde el dispositivo de decodificación de audio 24 puede obtener un elemento de sintaxis MinNumofCoeffsForAmbHOA igual a $(1+1)^2$ o cuatro. La parte de HOAconfig 250H incluye un elemento de sintaxis HoaOrder 152 configurado para indicar que el orden de HOA del contenido es igual a tres (o, en otras palabras, N = 3), donde el dispositivo de decodificación de audio 24 puede obtener un NumOf-HoaCoeffs a ser Igual a $(N+1)^2$ o 16.

[0458] Como se muestra adicionalmente en el ejemplo de la Fig. 10H(i), la parte 248H incluye una trama de audio tridimensional unificada (USAC-3D) de codificación de voz y audio (USAC-3D) en la que dos tramas HOA 249A y 249B se almacenan en una carga útil de extensión USAC dado que se almacenan dos tramas de audio Ddntro 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 de sintaxis numHOATransportChannels y un elemento de sintaxis MinNumOfCoeffsForAmbHOA. En los siguientes ejemplos, se supone que el elemento de sintaxis numHOATransportChannels es igual a 7 y el elemento de sintaxis MinNumOfCoeffsForAmbHOA es igual a cuatro, donde el número de canales de transporte flexibles es igual al elemento de sintaxis numHOATransportChannels menos el elemento de sintaxis MinNumOfCoeffsForAmbHOA (o tres).

[0459] La Fig. 10H(ii) ilustra las tramas 249A y 249B con más detalle. Como se muestra en el ejemplo de la Fig. 10H(ii), la trama 249A incluye los campos 154-154C de ChannelSideInfoData (CSID), los campos HOAGainCorrectionData (HOAGCD), los campos VVectorData 156 y 156B y los campos HOAPredictionInfo. El campo CSID 154 incluye la unitC 267, bb 266 y ba265 junto con el ChannelType 269, cada uno de los cuales se ajusta a los valores correspondientes 01, 1, 0 y 01 que se muestran en el ejemplo de la Fig. 10H(i). El campo CSID 154B incluye la unitC 267, bb 266 y ba265 junto con el ChannelType 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. 10H(ii). El campo CSID 154C incluye el campo ChannelType 269 que tiene un valor de 3. Cada uno de los campos CSID 154-154C corresponde al canal respectivo de los canales de transporte 1, 2 y 3. En efecto, cada campo CSID 154-154C indica si la carga útil correspondiente 156 y 156B son señales basadas en la dirección (cuando el ChannelType correspondiente es igual a cero), señales basadas en vectores (cuando el ChannelType correspondiente es igual a uno), un coeficiente HOA ambiental adicional (cuando el ChannelType correspondiente es igual a dos) o vacío (cuando el ChannelType es igual a tres).

[0460] En el ejemplo de la Fig. 10H(ii), la trama 249A incluye dos señales basadas en vectores (dado el ChannelType 269 igual a 1 en los campos CSID 154 y 154B) y un vacío (dado el ChannelType 269 igual a 3 en los campos CSID 154C). Dada la parte de HOAconfig anterior 250H, el dispositivo de decodificación de audio 24 puede determinar que todos los 16 elementos vectoriales V están codificados. Por tanto, los VVectorData 156 y 156B incluyen cada uno de los 16 elementos vectoriales, cada uno de ellos uniformemente cuantizado con 8 bits. Como se indica en la nota 1, el número y los índices de los elementos VVectorData codificados se especifican mediante el parámetro CodedVVecLength = 0. Además, como se señala con el asterisco único (*), el esquema de codificación se señala con NbitsQ = 5 en el campo CSID para el canal de transporte correspondiente.

[0461] En la trama 249B, el campo CSID 154 y 154B son los mismos que en la trama 249, mientras que el campo CSID 154C de la trama 249B ha conmutado a un ChannelType de uno. El campo CSID 154C de la trama 249B incluye por lo tanto el Cbflag 267, el Pflag 267 (que indica la codificación de Huffman) y el Nbits 261 (igual a doce). Como resultado, la trama 249B incluye un tercer campo VVectorData 156C que incluye 16 elementos vectoriales de V, cada uno de ellos cuantizado uniformemente con 12 bits y codificado por Huffman. Como se ha indicado anteriormente, el número y los índices de los elementos VVectorData codificados se especifican mediante el parámetro CodedVVecLength = 0, mientras que el esquema de codificación de Huffman está señalizado por el NbitsQ = 12, CbFlag = 0 y Pflag = 0 en el campo CSID 154C para este particular canal de transporte (por ejemplo, canal de transporte nº 3).

[0462] El ejemplo de las Figs. 10I(i) y 10I(ii) ilustran un segundo flujo de bits de ejemplo 248I y la parte de HOAconfig correspondiente 250I que ha sido generada para corresponder con el caso 0 en el pseudocódigo anterior. En el ejemplo de la Fig. 10I(i), la parte de configuración HOA 250I incluye un elemento de sintaxis CodedVVecLength 256 configurado para indicar que todos los elementos de un vector V están codificados, por

ejemplo, todos los 16 elementos vectoriales de V. La parte de HOAconfig 250I también incluye un elemento de sintaxis SpatialInterpolationMethod 255 establecido para indicar que la función de interpolación de la interpolación espacio-temporal es un coseno elevado. La parte de HOAconfig 250I además incluye un CodedSpatialInterpolationTime 254 configurado para indicar una duración de la muestra interpolada de 256.

[0463] La parte de HOAconfig 250I incluye además un elemento sintaxis MinAmbHoaOrder 150 configurado para indicar que el orden MinimumHOA del contenido de HOA ambiental es uno, donde el dispositivo de decodificación de audio 24 puede obtener un elemento sintáctico MinNumof-CoeffsForAmbHOA para que sea igual a $(1+1)^2$ o cuatro. El dispositivo de decodificación de audio 24 puede obtener también un elemento de sintaxis MaxNoofAddActiveAmbCoeffs configurado como una diferencia entre el elemento de sintaxis NumOfHoaCoeff y el MinNumOfCoeffsForAmbHOA, que se supone en este ejemplo igual a 16-4 o 12. El dispositivo de decodificación de audio 24 también puede obtener un elemento de sintaxis AmbAsignmbits configurado como ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4. La parte de HOAconfig 250H incluye un elemento de sintaxis HoaOrder 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 NumOfHoaCoeffs igual a $(N + 1)^2$ ó 16.

[0464] Como se muestra adicionalmente en el ejemplo de la Fig. 10I(i), la parte 248H incluye una trama de audio USAC-3D en la que dos tramas HOA 249C y 249D están almacenadas 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 en función de un elemento de sintaxis numHOATransportChannels y de un elemento de sintaxis de MinNumOfCoeffs-ForAmbHOA. En los siguientes ejemplos, se supone que el elemento de sintaxis numHOATransportChannels es igual a 7 y el elemento de sintaxis MinNumOfCoeffsForAmbHOA es igual a cuatro, donde el número de canales de transporte flexibles es igual al elemento de sintaxis numHOATransportChannels menos el elemento de sintaxis MinNumOfCoeffsForAmbHOA (o tres).

[0465] La Fig. 10I(ii) ilustra las tramas 249C y 249D con más detalle. Como se muestra en el ejemplo de la Fig. 10I (ii), la trama 249C incluye los campos de CSID 154-154C y campos de VVectorData 156. El campo CSID 154 incluye el CodedAmb-CoeffIdx 246, el AmbCoeffIdxTransition 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 AmbCoeffIdxTransitionState = 2, lo que da lugar a que el campo de bits CodedAmbCoeffIdx se indica o se especifica de otro modo en el flujo de bits), y el ChannelType 269 (que es igual a dos, indicando que la carga útil correspondiente es un coeficiente HOA ambiental adicional). El dispositivo de decodificación de audio 24 puede obtener el AmbCoeffIdx como igual al CodedAmbCoeffIdx + 1 + MinNumOfCoeffsForAmbHOA o 5 en este ejemplo. El campo CSID 154B incluye unitC 267, bb 266 y ba 265, junto con ChannelType 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. 10I(ii). El campo CSID 154C incluye el campo ChannelType 269 que tiene un valor de 3.

[0466] En el ejemplo de la Fig. 10I(ii), la trama 249C incluye una única señal basada en vectores (dado que el ChannelType 269 es igual a 1 en los campos CSID 154B) y un vacío (dado el ChannelType 269 igual a 3 en los campos CSID 154C). Dada la parte de HOAconfig anterior 250I, el dispositivo de decodificación de audio 24 puede determinar que todos los 16 elementos vectoriales V están codificados. Por lo tanto, el VVectorData 156 incluye todos los 16 elementos vectoriales, cada uno de ellos uniformemente cuantizado con 8 bits. Como se indica en la nota 1, el número y los índices de los elementos VVectorData codificados se especifican mediante el parámetro CodedVVecLength = 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.

[0467] En la trama 249D, el campo CSID 154 incluye un AmbCoeffIdxTransition 247 que indica que no ha ocurrido ninguna transición y, por lo tanto, el CodedAmbCoeffIdx 246 puede estar implícito desde la trama anterior y no necesita ser señalado o de otra manera especificado de nuevo. El campo CSID 154B y 154C de la trama 249D es el mismo que el de la trama 249C y por lo tanto, al igual que la trama 249C, la trama 249D incluye un único campo VVectorData 156, que incluye todos los 16 elementos vectoriales, cada uno de ellos cuantizado uniformemente con 8 bits.

[0468] Las Figs. 10J(i) y 10J(ii) ilustran un primer flujo de bits de ejemplo 248J y la parte de HOA config correspondiente 250J que han sido generados para corresponder con el caso 1 en el pseudocódigo anterior. En el ejemplo de la Fig. La parte de HOAconfig 250J incluye un elemento de sintaxis CodedVVecLength 256 establecido para indicar que todos los elementos de un vector V están codificados, excepto los elementos de sintaxis del elemento 1 hasta MinNumOfCoeffsForAmbHOA y aquellos elementos especificados en un elemento de sintaxis ContAddAmbHoaChan (se supone que es cero en este ejemplo). La parte de HOAconfig 250J también incluye un elemento de sintaxis SpatialInterpolationMethod 255 establecido para indicar que la función de interpolación de la interpolación espacio-temporal es un coseno elevado. La parte de HOAconfig 250J incluye además un CodedSpatialInterpolationTime 254 establecido para indicar una duración de muestra interpolada de 256. La parte de HOAconfig 250J incluye además un elemento de sintaxis MinAmbHoaOrder 150 configurado para indicar que el orden de MinimumHOA del contenido de HOA ambiental es uno, donde el dispositivo de decodificación de audio 24 puede obtener un elemento de sintaxis MinNumofCoeffsForAmbHOA igual a $(1+1)^2$ o cuatro. La parte de HOAconfig

250J incluye un elemento de sintaxis HoaOrder 152 configurado para indicar que el orden de HOA del contenido es igual a tres (o, en otras palabras, N =3), donde el dispositivo de decodificación de audio 24 puede obtener un NumOfHoaCoeffs igual a $(N+1)^2$ ó 16.

[0469] Como se muestra adicionalmente en el ejemplo de la Fig. 10J(i), la parte 248J incluye una trama de audio USAC-3D en la que dos tramas HOA 249E y 249F 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 en función de un elemento de sintaxis numHOATransportChannels y de un elemento de sintaxis de MinNumOfCoeffs-ForAmbHOA. En los siguientes ejemplos, se supone que el elemento de sintaxis numHOATransportChannels es igual a 7 y el elemento de sintaxis MinNumOfCoeffsForAmbHOA es igual a cuatro, donde el número de canales de transporte flexibles es igual al elemento de sintaxis numHOATransportChannels menos el elemento de sintaxis MinNumOfCoeffsForAmbHOA (o tres).

[0470] La Fig. 10J(ii) ilustra las tramas 249E y 249F con más detalle. Como se muestra en el ejemplo de la Fig. 10J(ii), la trama 249E incluye los campos CSID 154-154C y los campos VVectorData 156 y 156B. El campo CSID 154 incluye la unitC 267, bb 266 y ba265 junto con el ChannelType 269, cada uno de los cuales se ajusta a los valores correspondientes 01, 1, 0 y 01 que se muestran en el ejemplo de la Fig. 10J(i). El campo CSID 154B incluye la unitC 267, bb 266 y ba265 junto con el ChannelType 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. 10J(ii). El campo CSID 154C incluye el campo ChannelType 269 que tiene un valor de 3. Cada uno de los campos CSID 154-154C corresponde al canal respectivo de los canales de transporte 1, 2 y 3.

[0471] En el ejemplo de la Fig. 10J(ii), la trama 249E incluye dos señales basadas en vectores (dado que el ChannelType 269 es igual a 1 en los campos CSID 154 y 154B) y un vacío (dado el ChannelType 269 igual a 3 en los campos CSID 154C). Dada la parte de HOAconfig anterior 250H, el dispositivo de decodificación de audio 24 puede determinar que todos los 12 elementos vectoriales V están codificados (donde 12 se obtiene como HOAOrder + 1)^2 - (MinNumOfCoeffsForAmbHOA) - (ContAddAmb - HoaChan) = 16 - 4 - 0 = 12). Por tanto, los VVectorData 156 y 156B incluyen cada uno de los 12 elementos vectoriales, cada uno de ellos uniformemente cuantizado con 8 bits. Como se indica en la nota 1, el número y los índices de los elementos VVectorData codificados se especifican mediante el parámetro CodedVVecLength = 0. Además, como se señala con el asterisco único (*), el esquema de codificación se señala con NbitsQ = 5 en el campo CSID para el canal de transporte correspondiente.

[0472] En la trama 249F, el campo CSID 154 y 154B son los mismos que en la trama 249E, mientras que el campo CSID 154C de la trama 249F ha conmutado a un ChannelType de uno. El campo CSID 154C de la trama 249B incluye por lo tanto el Cbflag 267, el Pflag 267 (que indica la codificación de Huffman) y el Nbits 261 (igual a doce). Como resultado, la trama 249F incluye un tercer campo de VVectorData 156C que incluye 12 elementos vectoriales de V, cada uno de ellos cuantizado uniformemente con 12 bits y codificado por Huffman. Como se ha indicado anteriormente, el número y los índices de los elementos VVectorData codificados se especifican mediante el parámetro CodedVVecLength = 0, mientras que el esquema de codificación de Huffman está señalizado por el NbitsQ = 12, CbFlag = 0 y Pflag = 0 en el campo CSID 154C para este particular canal de transporte (por ejemplo, canal de transporte nº 3).

[0473] El ejemplo de las Figs. 10K(i) y 10K(ii) ilustran un segundo flujo de bits de ejemplo 248K y la parte de HOAconfig correspondiente 250K que se han generado para corresponder con el caso 1 en el pseudocódigo anterior. En el ejemplo de la Fig. 10K(i), la parte de HOAconfig 250K incluye un elemento de sintaxis CodedVVecLength 256 configurado para indicar que todos los elementos de un vector V están codificados, excepto los elementos de sintaxis del elemento 1 al MinNumOfCoeffsForAmbHOA y aquellos elementos especificados en un elemento de sintaxis ContAddAmbHoaChan (se supone que es uno en este ejemplo). La parte de HOAconfig 250K también incluye un elemento de sintaxis SpatialInterpolationMethod 255 establecido para indicar que la función de interpolación de la interpolación espacio-temporal es un coseno elevado. La parte de HOAconfig 250K además incluye un CodedSpatial-InterpolationTime 254 configurado para indicar una duración de muestra interpolada de 256.

[0474] La parte de HOAconfig 250K incluye además un elemento de sintaxis MinAmbHoaOrder 150 establecido para indicar que el orden de MinimumHOA del contenido de HOA ambiental es uno, donde el dispositivo de decodificación de audio 24 puede obtener un elemento de sintaxis MinNumof-CoeffsForAmbHOA igual a $(1+1)^2$ o cuatro. El dispositivo de decodificación de audio 24 puede obtener también un elemento de sintaxis MaxNoofAddActiveAmbCoeffs configurado como una diferencia entre el elemento de sintaxis NumOfHoaCoeff y el MinNumOfCoeffsForAmbHOA, que se supone en este ejemplo igual a 16-4 ó 12. El dispositivo de decodificación de audio 24 también puede obtener un elemento de sintaxis AmbAsignmbits configurado como ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4. La parte de HOAconfig 250K incluye un elemento de sintaxis HoaOrder 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 NumOfHoaCoeffs que es igual a $(N+1)^2$ ó 16.

[0475] Como se muestra adicionalmente en el ejemplo de la Fig. 10K(i), 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 en función de un elemento de sintaxis numHOATransportChannels y de un elemento de sintaxis MinNumOf-CoeffsForAmbHOA. En los siguientes ejemplos, se supone que el elemento de sintaxis numHOATransportChannels es igual a 7 y el elemento de sintaxis MinNumOfCoeffsForAmbHOA es igual a cuatro, donde el número de canales de transporte flexibles es igual al elemento de sintaxis numHOATransportChannels menos el elemento de sintaxis MinNumOfCoeffsForAmbHOA (o tres).

[0476] La Fig. 10K(ii) ilustra las tramas 249G y 249H con más detalle. Como se muestra en el ejemplo de la Fig. 10K(ii), la trama 249G incluye los campos CSID 154-154C y los campos VVectorData 156. El campo CSID 154 incluye el CodedAmb-CoeffIdx 246, la AmbCoeffIdxTransition 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 AmbCoeffIdxTransitionState = 2, lo que da lugar a que el campo de bits CodedAmbCoeffIdx se indica o se especifica de otro modo en el flujo de bits) y el ChannelType 269 (que es igual a dos, indicando que la carga útil correspondiente es un coeficiente HOA ambiental adicional). El dispositivo de decodificación de audio 24 puede obtener el AmbCoeffIdx igual que el CodedAmbCoeffIdx+1+MinNumOfCoeffsForAmbHOA o 5 en este ejemplo. El campo CSID 154B incluye unitC 267, bb 266 y ba 265, junto con ChannelType 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. 10K(ii). El campo CSID 154C incluye el campo ChannelType 269 que tiene un valor de 3.

[0477] En el ejemplo de la Fig. 10K(ii), la trama 249G incluye una única señal basada en vectores (dado el ChannelType 269 igual a 1 en los campos CSID 154B) y un vacío (dado el ChannelType 269 igual a 3 en los campos CSID 154C). Dada la parte de HOAconfig anterior 250K, el dispositivo de decodificación de audio 24 puede determinar que se codifiquen 11 elementos vectoriales V (donde 12 se deriva como $(HOAOder + 1)^2$ - (MinNumOfCoeffsForAmbHOA) - (ContAddAmbHoaChan) = 16-4-1 = 11). Por lo tanto, el VVectorData 156 incluye todos los 11 elementos vectoriales, cada uno de ellos uniformemente cuantizado con 8 bits. Como se indica en la nota 1, el número y los índices de los elementos VVectorData codificados se especifican mediante el parámetro CodedVVecLength = 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.

[0478] En la trama 249H, el campo CSID 154 incluye una AmbCoeffIdxTransition 247 que indica que no ha ocurrido ninguna transición y por lo tanto el CodedAmbCoeffIdx 246 puede estar implícito en la trama anterior y no necesita ser señalado o especificado otra vez de nuevo. 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 VVectorData 156, el cual incluye 11 elementos vectoriales, cada uno de ellos cuantizado uniformemente con 8 bits

[0479] Las Figs. 10L(i) y 10L(ii) ilustran un primer flujo de bits de ejemplo 248L y la parte de HOAconfig correspondiente 250L que han sido generados para corresponder con el caso 2 en el pseudocódigo anterior. En el ejemplo de la Fig. 10L(i), la parte de HOAconfig 250A incluye un elemento de sintaxis CodedVVecLength 256 establecido para indicar que todos los elementos de un vector V están codificados, excepto los elementos del orden cero hasta el orden especificado por el elemento de sintaxis MinAmbHoaOrder 150 (que es igual A $(HoaOrder + 1)^2$ - $(MinAmbHoaOrder + 1)^2 = 16 - 4 = 12$ en este ejemplo). La parte de HOAconfig 250L también incluye un elemento de sintaxis SpatialInterpolationMethod 255 establecido para indicar que la función de interpolación de la interpolación espacio-temporal es un coseno elevado. La parte de HOAconfig 250L además incluye un CodedSpatialInterpolation-Time 254 para indicar una duración de muestra interpolada de 256. La parte de HOAconfig 250L incluye además un elemento de sintaxis MinAmbHoaOrder 150 configurado para indicar que el orden de MinimumHOA del contenido de HOA ambiental es uno, donde el dispositivo de decodificación de audio 24 puede obtener un elemento de sintaxis MinNumofCoeffsForAmbHOA igual a $(1+1)^2$ o cuatro. La parte de HOAconfig 250A incluye un elemento 152 de sintaxis HoaOrder establecido 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 NumOfHoaCoeffs igual a $(N+1)^2$ o 16.

[0480] Como se muestra adicionalmente en el ejemplo de la Fig. 10L(i), la parte 248L incluye una trama de audio USAC -3D en la cual dos tramas HOA 249I y 249J 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 en función de un elemento de sintaxis numHOATransportChannels y de un elemento de sintaxis MinNumOf-CoeffsForAmbHOA. En los siguientes ejemplos, se supone que el elemento de sintaxis numHOATransportChannels es igual a 7 y el elemento de sintaxis MinNumOfCoeffsForAmbHOA es igual a cuatro, donde el número de canales de transporte flexibles es igual al elemento de sintaxis numHOATransportChannels menos el elemento de sintaxis MinNumOfCoeffsForAmbHOA (o tres).

[0481] La Fig. 10L(ii) ilustra las tramas 249I y 249J con más detalle. Como se muestra en el ejemplo de la Fig. 10L(ii), la trama 249I incluye los campos CSID 154-154C y los campos VVectorData 156 y 156B. El campo CSID 154 incluye la unitC 267, bb 266 y ba265 junto con el ChannelType 269, cada uno de los cuales se ajusta a los valores

correspondientes 01, 1, 0 y 01 que se muestran en el ejemplo de la Fig. 10J(i). El campo CSID 154B incluye la unitC 267, bb 266 y ba265 junto con el ChannelType 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. 10L(ii). El campo CSID 154C incluye el campo ChannelType 269 que tiene un valor de 3. Cada uno de los campos CSID 154-154C corresponde al canal respectivo de los canales de transporte 1, 2 y 3.

[0482] En el ejemplo de la Fig. 10L(ii), la trama 249I incluye dos señales basadas en vectores (dado el ChannelType 269 igual a 1 en los campos CSID 154 y 154B) y un vacío (dado el ChannelType 269 igual a 3 en los campos CSID 154C). Dada la parte de HOAconfig 250H, el dispositivo de decodificación de audio 24 puede determinar que se codifiquen 12 elementos vectoriales de V. Por tanto, los VVectorData 156 y 156B incluyen cada uno 12 elementos vectoriales, cada uno de ellos uniformemente cuantizado con 8 bits. Como se indica en la nota 1, el número y los índices de los elementos VVectorData codificados se especifican mediante el parámetro CodedVVecLength = 0. Además, como se señala con el asterisco único (*), el esquema de codificación se señala con NbitsQ = 5 en el campo CSID para el canal de transporte correspondiente.

[0483] En la trama 249J, el campo CSID 154 y 154B son los mismos que en la trama 249I, mientras que el campo CSID 154C de la trama 249F ha conmutado a un ChannelType de uno. El campo CSID 154C de la trama 249B incluye por lo tanto el Cbflag 267, el Pflag 267 (que indica la codificación de Huffman) y el Nbits 261 (igual a doce). Como resultado, la trama 249F incluye un tercer campo de VVectorData 156C que incluye 12 elementos vectoriales de V, cada uno de ellos cuantizado uniformemente con 12 bits y codificado por Huffman. Como se ha indicado anteriormente, el número y los índices de los elementos VVectorData codificados se especifican mediante el parámetro CodedVVecLength = 0, mientras que el esquema de codificación de Huffman está señalizado por el NbitsQ = 12, CbFlag = 0 y Pflag = 0 en el campo CSID 154C para este particular canal de transporte (por ejemplo, canal de transporte nº 3).

[0484] El ejemplo de las Figs. 10M(i) y 10M(ii) ilustran un segundo ejemplo de flujo de bits 248M y la parte de HOAconfig correspondiente 250M que ha sido generada para corresponder con el caso 2 en el pseudocódigo anterior. En el ejemplo de la Fig. 10M(i), la parte de HOAconfig 250M incluye un elemento de sintaxis CodedVVecLength 256 establecido para indicar que todos los elementos de un vector V están codificados, excepto los elementos del orden cero hasta el orden especificado por el elemento de sintaxis MinAmb-HoaOrder 150 (que es igual a $(\text{HoaOrder}+1)^2 - (\text{MinAmbHoaOrder} + 1)^2 = 16 - 4 = 12$ en este ejemplo). La parte de HOAconfig 250M también incluye un elemento de sintaxis SpatialInterpolationMethod 255 establecido para indicar que la función de interpolación de la interpolación espacio-temporal es un coseno elevado. La parte de HOAconfig 250M además incluye un CodedSpatialInterpolationTime 254 para indicar una duración de muestra interpolada de 256.

[0485] La parte de HOAconfig 250M incluye además un elemento de sintaxis MinAmbHoaOrder 150 configurado para indicar que el orden de MinimumHOA del contenido de HOA ambiental es uno, donde el dispositivo de decodificación de audio 24 puede obtener un elemento de sintaxis MinNumof-CoeffsForAmbHOA igual a $(1+1)^2$ o cuatro. El dispositivo de decodificación de audio 24 puede obtener también un elemento de sintaxis MaxNoofAddActiveAmbCoeffs configurado como una diferencia entre el elemento de sintaxis NumOfHoaCoeff y el MinNumOfCoeffsForAmbHOA, que se supone en este ejemplo igual a 16-4 o 12. El dispositivo de decodificación de audio 24 también puede obtener un elemento de sintaxis AmbAsignmbits configurado como ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4. La parte de HOAconfig 250M incluye un elemento de sintaxis HoaOrder 152 configurado para indicar que el orden HOA del contenido sea igual a tres (o, en otras palabras, $N = 3$), donde el dispositivo de decodificación de audio 24 puede obtener un NumOfHoaCoeffs igual a $(N+1)^2$ o 16.

[0486] Como se muestra adicionalmente en el ejemplo de la Fig. 10M(i), la parte 248M incluye una trama de audio USAC-3D en la que dos tramas HOA 249K y 249L 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 en función de un elemento de sintaxis numHOATransportChannels y de un elemento de sintaxis MinNumOf-CoeffsForAmbHOA. En los siguientes ejemplos, se supone que el elemento de sintaxis numHOATransportChannels es igual a 7 y el elemento de sintaxis MinNumOfCoeffsForAmbHOA es igual a cuatro, donde el número de canales de transporte flexibles es igual al elemento de sintaxis numHOATransportChannels menos el elemento de sintaxis MinNumOfCoeffsForAmbHOA (o tres).

[0487] La Fig. 10M(ii) ilustran las tramas 249K y 249L con más detalle. Como se muestra en el ejemplo de la Fig. 10M(ii), la trama 249K incluye los campos CSID 154-154C y un campo VVectorData 156. El campo CSID 154 incluye el CodedAmb-CoeffIdx 246, el AmbCoeffIdxTransition 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 AmbCoeffIdxTransitionState = 2, lo que da lugar a que el campo de bits CodedAmbCoeffIdx se indica o se especifica de otro modo en el flujo de bits), y el ChannelType 269 (que es igual a dos, indicando que la carga útil correspondiente es un coeficiente HOA ambiental adicional). El dispositivo de decodificación de audio 24 puede obtener el AmbCoeffIdx igual que el CodedAmbCoeffIdx+1+MinNumOfCoeffsForAmbHOA o 5 en este ejemplo. El campo CSID 154B incluye unitC 267, bb 266 y ba 265, junto con ChannelType 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. 10M(ii). El campo CSID 154C incluye el campo ChannelType 269 que tiene un valor de 3.

[0488] En el ejemplo de la Fig. 10M(ii), la trama 249K incluye una única señal basada en vectores (dado que el ChannelType 269 es igual a 1 en los campos CSID 154B) y un vacío (dado el ChannelType 269 igual a 3 en los campos CSID 154C). Dada la parte de HOAconfig anterior 250M, el dispositivo de decodificación de audio 24 puede determinar que se codifiquen 12 elementos vectoriales de V. Por lo tanto, el VVectorData 156 incluye 12 elementos vectoriales, cada uno de ellos uniformemente cuantizado con 8 bits. Como se indica en la nota 1, el número y los índices de los elementos VVectorData codificados se especifican mediante el parámetro CodedVVecLength = 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.

[0489] En la trama 249L, el campo CSID 154 incluye una AmbCoeffIdxTransition 247 que indica que no ha ocurrido ninguna transición y por lo tanto el CodedAmbCoeffIdx 246 puede estar implícito desde la trama anterior y no necesita ser señalado o especificado otra vez de nuevo. El campo CSID 154B y 154C de la trama 249L es el mismo que el de la trama 249K y, por lo tanto, al igual que la trama 249K, la trama 249L incluye un único campo VVectorData 156, que incluye 12 elementos vectoriales, cada uno de ellos cuantizado uniformemente con 8 bits

[0490] Las Figs. 10N(i) y 10N(ii) ilustran un primer ejemplo de flujo de bits 248N y la parte de HOAconfig correspondiente 250N que se han generado para corresponder con el caso 3 en el pseudocódigo anterior. En el ejemplo de la Fig. 10N(i), la parte de HOAconfig 250N incluye un elemento de sintaxis CodedVVecLength 256 establecido para indicar que todos los elementos de un vector V están codificados, excepto aquellos elementos especificados en un elemento de sintaxis ContAddAmbHoaChan (que se supone que es cero en este ejemplo). La parte de HOAconfig 250N también incluye un elemento de sintaxis SpatialInterpolationMethod 255 establecido para indicar que la función de interpolación de la interpolación espacio-temporal es un coseno elevado. La parte de HOAconfig 250N además incluye un CodedSpatialInterpolationTime 254 para indicar una duración de muestra interpolada de 256. La parte de HOAconfig 250N incluye además un elemento de sintaxis MinAmbHoaOrder 150 establecido para indicar que el orden de MinimumHOA del contenido de HOA ambiental es uno, donde el dispositivo de decodificación de audio 24 puede obtener un elemento de sintaxis MinNumofCoeffsForAmbHOA igual a $(1+1)^2$ o cuatro. La parte de HOAconfig 250N incluye un elemento de sintaxis HOAOrder 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 NumOfHoaCoeffs igual a $(N + 1)^2$ o 16.

[0491] Como se muestra adicionalmente en el ejemplo de la Fig. 10N(i), la parte 248N incluye una trama de audio USAC-3D en la que dos tramas HOA 249M y 249N están almacenadas 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 en función de un elemento de sintaxis numHOATransportChannels y de un elemento de sintaxis MinNumOf-CoeffsForAmbHOA. En los siguientes ejemplos, se supone que el elemento de sintaxis numHOATransportChannels es igual a 7 y el elemento de sintaxis MinNumOfCoeffsForAmbHOA es igual a cuatro, donde el número de canales de transporte flexibles es igual al elemento de sintaxis numHOATransportChannels menos el elemento de sintaxis MinNumOfCoeffsForAmbHOA (o tres).

[0492] La Fig. 10N(ii) ilustra las tramas 249M y 249N con más detalle. Como se muestra en el ejemplo de la Fig. 10N(ii), la trama 249M incluye los campos CSID 154-154C y los campos VVectorData 156 y 156B. El campo CSID 154 incluye la unitC 267, bb 266 y ba265 junto con el ChannelType 269, cada uno de los cuales se ajusta a los valores correspondientes 01, 1, 0 y 01 que se muestran en el ejemplo de la Fig. 10J(i). El campo CSID 154B incluye la unitC 267, bb 266 y ba265 junto con el ChannelType 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. 10N(ii). El campo CSID 154C incluye el campo ChannelType 269 que tiene un valor de 3. Cada uno de los campos CSID 154-154C corresponde al canal respectivo de los canales de transporte 1, 2 y 3.

[0493] En el ejemplo de la Fig. 10N(ii), la trama 249M incluye dos señales basadas en vectores (dado el ChannelType 269 igual a 1 en los campos CSID 154 y 154B) y un vacío (dado el ChannelType 269 igual a 3 en los campos CSID 154C). Dada la parte de HOAconfig anterior 250M, el dispositivo de decodificación de audio 24 puede determinar que se codifiquen 16 elementos vectoriales de V. Por tanto, los VVectorData 156 y 156B incluyen cada uno 16 elementos vectoriales, cada uno de ellos uniformemente cuantizado con 8 bits. Como se indica en la nota 1, el número y los índices de los elementos VVectorData codificados se especifican mediante el parámetro CodedVVecLength = 0. Además, como se señala con el asterisco único (*), el esquema de codificación se señala con NbitsQ = 5 en el campo CSID para el canal de transporte correspondiente.

[0494] En la trama 249N, el campo CSID 154 y 154B son los mismos que en la trama 249M, mientras que el campo CSID 154C de la trama 249F ha conmutado a un ChannelType de uno. El campo CSID 154C de la trama 249B incluye por lo tanto el Cbflag 267, el Pflag 267 (que indica la codificación de Huffman) y el Nbits 261 (igual a doce). Como resultado, la trama 249F incluye un tercer campo de VVectorData 156C que incluye 16 elementos vectoriales de V, cada uno de ellos cuantizado uniformemente con 12 bits y codificado por Huffman. Como se ha

indicado anteriormente, el número y los índices de los elementos VVectorData codificados se especifican mediante el parámetro CodedVVecLength = 0, mientras que el esquema de codificación de Huffman está señalizado por el NbitsQ = 12, CbFlag = 0 y Pflag = 0 en el campo CSID 154C para este particular canal de transporte (por ejemplo, canal de transporte nº 3).

[0495] El ejemplo de las Figs. 10O(i) y 10O(ii) ilustran un segundo ejemplo de flujo de bits 2480 y la parte de HOAconfig correspondiente 2500 que han sido generados para corresponder con el caso 3 en el pseudocódigo anterior. En el ejemplo de la Fig. 10O(i), la parte de HOAconfig 2500 incluye un elemento de sintaxis CodedVVecLength 256 establecido para indicar que todos los elementos de un vector V están codificados, excepto aquellos elementos especificados en un elemento de sintaxis ContAddAmbHoaChan (que se supone que es uno en este ejemplo). La parte de HOAconfig 2500 también incluye un elemento de sintaxis SpatialInterpolationMethod 255 establecido para indicar que la función de interpolación de la interpolación espacio-temporal es un coseno elevado. La parte de HOAconfig 2500 además incluye un CodedSpatialInterpolationTime 254 para indicar una duración de la muestra interpolada de 256.

[0496] La parte de HOAconfig 2500 incluye además un elemento de sintaxis MinAmbHoaOrder 150 establecido para indicar que el orden de MinimumHOA del contenido de HOA ambiental es uno, donde el dispositivo de decodificación de audio 24 puede obtener un elemento de sintaxis MinNumof-CoeffsForAmbHOA igual a $(1+1)^2$ o cuatro. El dispositivo de decodificación de audio 24 puede obtener también un elemento de sintaxis MaxNoofAddActiveAmbCoeffs configurado como una diferencia entre el elemento de sintaxis NumOfHoaCoeff y el MinNumOfCoeffsForAmbHOA, que se supone en este ejemplo igual a 16-4 o 12. El dispositivo de decodificación de audio 24 también puede obtener un elemento de sintaxis AmbAsignmbits configurado como ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4. La parte de HOAconfig 2500 incluye un elemento de sintaxis HOAOrder 152 configurado para indicar que el orden de HOA del contenido es igual a tres (o, en otras palabras, $N = 3$), donde el dispositivo de decodificación de audio 24 puede obtener un NumOfHoaCoeffs igual a $(N+1)^2$ ó 16.

[0497] Como se muestra adicionalmente en el ejemplo de la Fig. 10O(i), la parte 2480 incluye una trama de audio USAC-3D en la que dos tramas HOA 2490 y 249P 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 en función de un elemento de sintaxis numHOATransportChannels y de un elemento de sintaxis MinNumOf-CoeffsForAmbHOA. En los siguientes ejemplos, se supone que el elemento de sintaxis numHOATransportChannels es igual a 7 y el elemento de sintaxis MinNumOfCoeffsForAmbHOA es igual a cuatro, donde el número de canales de transporte flexibles es igual al elemento de sintaxis numHOATransportChannels menos el elemento de sintaxis MinNumOfCoeffsForAmbHOA (o tres).

[0498] La Fig. 10O(ii) ilustra las tramas 2490 y 249P con más detalle. Como se muestra en el ejemplo de la Fig. 10O(ii), la trama 2490 incluye los campos CSID 154-154C y un campo VVectorData 156. El campo CSID 154 incluye el CodedAmb-CoeffIdx 246, el AmbCoeffIdxTransition 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 AmbCoeffIdxTransitionState = 2, lo que da lugar a que el campo de bits CodedAmbCoeffIdx se indica o se especifica de otro modo en el flujo de bits), y el ChannelType 269 (que es igual a dos, indicando que la carga útil correspondiente es un coeficiente HOA ambiental adicional). El dispositivo de decodificación de audio 24 puede obtener el AmbCoeffIdx como igual al CodedAmbCoeffIdx + 1 + MinNumOfCoeffsForAmbHOA o 5 en este ejemplo. El campo CSID 154B incluye unitC 267, bb 266 y ba 265, junto con ChannelType 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. 10O(ii). El campo CSID 154C incluye el campo ChannelType 269 que tiene un valor de 3.

[0499] En el ejemplo de la Fig. 10O(ii), la trama 2490 incluye una única señal basada en vectores (dada el ChannelType 269 igual a 1 en los campos CSID 154B) y un vacío (dado el ChannelType 269 igual a 3 en los campos CSID 154C). Dada la parte de HOAconfig anterior 2500, el dispositivo de decodificación de audio 24 puede determinar que 16 menos el especificado por el elemento de sintaxis ContAddAmbHoaChan (por ejemplo, donde el elemento vectorial asociado con un índice de 6 se especifica como el elemento de sintaxis ContAddAmbHoaChan) o 15 elementos vectoriales V están codificados. Por tanto, el VVectorData 156 incluye 15 elementos vectoriales, cada uno de ellos uniformemente cuantizado con 8 bits. Como se indica en la nota 1, el número y los índices de los elementos VVectorData codificados se especifican mediante el parámetro CodedVVecLength = 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.

[0500] En la trama 249P, el campo CSID 154 incluye una AmbCoeffIdxTransition 247 que indica que no ha ocurrido ninguna transición y por lo tanto el CodedAmbCoeffIdx 246 puede estar implícito desde la trama anterior y no necesita ser señalado o especificado otra vez de nuevo. El campo CSID 154B y 154C de la trama 249P son iguales a los de la trama 2490 y, por lo tanto, al igual que la trama 2490, la trama 249P incluye un único campo VVectorData 156, que incluye 15 elementos vectoriales, cada uno de ellos cuantizado uniformemente con 8 bits.

[0501] Las Figs. 11A-11G son diagramas de bloques que ilustran, con más detalle, varias unidades del dispositivo de decodificación de audio 24 que se muestra en el ejemplo de la Fig. 5. La Fig. 11A es un diagrama de bloques que ilustra, con más detalle, la unidad de extracción 72 del dispositivo de decodificación de audio 24. Como se muestra en el ejemplo de la Fig. 11A, la unidad de extracción 72 puede incluir una unidad de análisis de modo 270, una unidad de configuración de modo 272 ("mode config unit 272") y una unidad de extracción configurable 274.

[0502] La unidad de análisis de modo 270 puede representar una unidad configurada para analizar el elemento de sintaxis anotado anteriormente indicativo de un modo de codificación (por ejemplo, el elemento de sintaxis de ChannelType que se muestra en el ejemplo de la Fig. 10E) usado para codificar los coeficientes HOA 11 para formar el flujo de bits 21. La unidad de análisis de modo 270 puede pasar el elemento de sintaxis de determinación a la unidad de configuración de modo 272. La unidad de configuración de modo 272 puede representar una unidad configurada para configurar la unidad de extracción configurable 274 en base al elemento de sintaxis analizado. La unidad de configuración de modo 272 puede configurar la unidad de extracción configurable 274 para extraer una representación codificada basada en la dirección de los coeficientes HOA 11 desde el flujo de bits 21, o extraer una representación codificada basada en los vectores de los coeficientes HOA 11 del flujo de bits 21 basado en el análisis del elemento de sintaxis.

[0503] Cuando se ha realizado una codificación basada en la dirección, la unidad de extracción configurable 274 puede extraer la versión basada en la dirección de los coeficientes HOA 11 y los elementos sintácticos asociados con esta versión codificada (que se designa como información basada en la dirección 91 en el ejemplo de la Fig. 11A). Esta información basada en la dirección 91 puede incluir la información direccional 253 que se muestra en el ejemplo de la Fig. 10D y SideChannelInfoData basados en la dirección que se muestran en el ejemplo de la Fig. 10E como se define por un ChannelType igual a cero.

[0504] Cuando el elemento de sintaxis indica que los coeficientes HOA 11 se han codificado utilizando una síntesis basada en vectores (por ejemplo, cuando el elemento de sintaxis de ChannelType es igual a uno), la unidad de extracción configurable 274 puede extraer los vectores V[*k*] de primer plano codificados 57, los coeficientes HOA ambientales codificados 59 y las señales nFG codificadas 59. La unidad de extracción configurable 274 también puede, al determinar que el elemento de sintaxis indica que los coeficientes HOA 11 se han codificado utilizando una síntesis basada en vectores, extraer el elemento de sintaxis CodedSpatialInterpolationTime 254 y el elemento de sintaxis SpatialInterpolationMethod 255 del flujo de bits 21, pasando estos elementos de sintaxis 254 y 255 a la unidad de interpolación espacio-temporal 76.

[0505] La Fig. 11B es un diagrama de bloques que ilustra, con más detalle, la unidad de cuantización 74 del dispositivo de decodificación de audio 24 que se muestra en el ejemplo de la Fig. 5. La unidad de cuantización 74 puede representar una unidad configurada para operar de manera recíproca con la unidad de cuantización 52 que se muestra en el ejemplo de la Fig. 4 con el fin de decodificar por entropía y decuantizar los vectores V[k] de primer plano codificados 57 y, de este modo, generar vectores los V[k] de primer plano reducidos 55_k. La unidad de decuantización escalar/entrópica 984 puede incluir una unidad de decodificación de categoría/residuo 276, una unidad de predicción 278 y una unidad de decuantización uniforme 280.

[0506] La unidad de decodificación de categoría/residuo 276 puede representar una unidad configurada para llevar a cabo la decodificación de Huffman con respecto a los vectores de primer plano V[k] codificados 57 utilizando la tabla Huffman identificada mediante la información de la tabla Huffman 241 (que es, como se ha indicado anteriormente, expresada como un elemento de sintaxis en el flujo de bits 21). La unidad de decodificación de categoría/residuo 276 puede emitir vectores de primer plano cuantizados V[k] a la unidad de predicción 278. La unidad de predicción 278 puede representar una unidad configurada para realizar la predicción con respecto a los vectores V[k] de primer plano cuantizados basados en el modo de predicción 237, que emiten vectores V[*k*] de primer plano cuantizados aumentados a la unidad de decuantización uniforme 280. La unidad de decuantización uniforme 280 puede representar una unidad configurada para realizar la decuantización con respecto a los vectores V[*k*] de primer plano cuantizados aumentados basándose en el valor de *nbits* 233, produciendo los vectores V[k] de primer plano reducidos 55_k

[0507] La Fig. 11C es un diagrama de bloques que ilustra, con más detalle, la unidad de decodificación psicoacústica 80 del dispositivo de decodificación de audio 24 que se muestra en el ejemplo de la Fig. 5. Como se ha indicado anteriormente, la unidad de decodificación psicoacústica 80 puede operar de una manera recíproca con la unidad de codificación de audio psicoacústica 40 que se muestra en el ejemplo de la Fig. 4 a fin de decodificar los coeficientes HOA 59 codificados y las señales nFG 61 codificadas y, de este modo, generar coeficientes HOA ambientales compensados de energía 47' y las señales nFG interpoladas 49' (que también se pueden denominar objetos de audio nFG interpolados 49'). La unidad de decodificación psicoacústica 80 puede pasar los coeficientes HOA ambientales compensados de energía 47' a la unidad de formulación de coeficientes HOA 82 y las señales nFG 49' a la reordenación 84. La unidad de decodificación psicoacústica 80 puede incluir una pluralidad de decodificadores de audio 80-80N similares a la unidad de codificación de audio psicoacústica 40. Los decodificadores de audio 80-80N pueden ser instanciados por o incluidos de otro modo dentro de la unidad de codificación de audio psicoacústica 40 en cantidad suficiente para permitir, como se ha indicado anteriormente, la decodificación concurrente de cada canal de los coeficientes HOA de fondo 47' y cada señal de las señales nFG 49'.

[0508] La Fig. 11D es un diagrama de bloques que ilustra, con más detalle, la unidad de reordenación 84 del dispositivo de decodificación de audio 24 que se muestra en el ejemplo de la Fig. 5. La unidad de reordenación 84 puede representar una unidad configurada para operar de una manera similar a la descrita anteriormente con respecto a la unidad de reordenación 34. La unidad de reordenación 84 puede incluir una unidad de reordenación de vectores 282 que puede representar una unidad configurada para recibir elementos de sintaxis 205 indicativos del orden original de los componentes en primer plano de los coeficientes HOA 11. La unidad de extracción 72 puede analizar estos elementos de sintaxis 205 del flujo de bits 21 y pasar el elemento de sintaxis 205 a la unidad de reordenación 84. La unidad de reordenación de vectores 282 puede, basándose en estos elementos 205 de sintaxis de reordenación, reordenar las señales nFG interpoladas 49' y los vectores V[*k*] de primer plano reducidos 55_k para generar señales nFG reordenadas 49' y los vectores V[*k*] de primer plano reordenados 55_k'. La unidad de reordenación 84 puede emitir las señales nFG reordenadas 49" a la unidad de formulación del primer plano 78 y los vectores V[*k*] de primer plano reordenados 55_k' a la unidad de interpolación espacio-temporal 76.

[0509] La Fig. 11E es un diagrama de bloques que ilustra, con más detalle, la unidad de interpolación espacio-temporal 76 del dispositivo de decodificación de audio 24 que se muestra en el ejemplo de la Fig. 5. 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 incluir una unidad de interpolación V 284, que puede representar una unidad configurada para recibir los vectores V[*k*] reordenados 55_k' y realizar la interpolación espacio-temporal con respecto a los vectores V[*k*] de primer plano reordenados 55_k' y los vectores V[*k*-1] de primer plano reordenados 55_{k-1}' para generar los vectores V[*k*] de primer plano interpolados 55_k". La unidad de interpolación V 284 puede realizar interpolación basada en el elemento de sintaxis CodedSpatialInterpolationTime 254 y el elemento de sintaxis SpatialInterpolationMethod 255. En algunos ejemplos, la unidad de interpolación V 285 puede interpolar los vectores V durante la duración especificada por el elemento de sintaxis CodedSpatialInterpolationTime 254 utilizando el tipo de interpolación identificado por el elemento de sintaxis SpatialInterpolationMethod 255. La unidad de interpolación espacio-temporal 76 puede enviar los vectores V[*k*] de primer plano interpolados 55_k" a la unidad de formulación del primer plano 78.

[0510] La Fig. 11F es un diagrama de bloques que ilustra, con más detalle, la unidad de formulación del primer plano 78 del dispositivo de decodificación de audio 24 que se muestra en el ejemplo de la Fig. 5. La unidad de formulación del primer plano 78 puede incluir una unidad de multiplicación 286 que puede representar una unidad configurada para realizar la multiplicación matricial con respecto a los vectores V[*k*] interpolados de primer plano 55_k" y las señales nFG reordenadas 49" para generar los coeficientes HOA de primer plano 65.

[0511] La Fig. 11G es un diagrama de bloques que ilustra, con más detalle, la unidad de formulación de coeficientes HOA 82 del dispositivo de decodificación de audio 24 que se muestra en el ejemplo de la Fig. 5. La unidad de formulación de coeficientes HOA 82 puede incluir una unidad de adición 288, que puede representar una unidad configurada para añadir los coeficientes HOA de primer plano 65 a los canales ambientales HOA 47' para obtener los coeficientes HOA 11'.

[0512] La Fig. 12 es un diagrama que ilustra un ecosistema de audio ejemplar que puede realizar diferentes aspectos de las técnicas descritas en esta divulgación. Como se ilustra en la Fig. 12, el ecosistema de audio 300 puede incluir la adquisición 301, edición 302, codificación, 303, transmisión 304 y reproducción 305.

[0513] La adquisición 301 puede representar las técnicas del ecosistema de audio 300 donde se adquiere el contenido de audio. Los ejemplos de adquisición 301 incluyen, pero no se limitan a, grabación de sonido (por ejemplo, sonido en directo), generación de audio (por ejemplo, objetos de audio, producción de foley, síntesis de sonido, simulaciones) y similares. En algunos ejemplos, el sonido puede ser grabado en conciertos, eventos deportivos, y cuando se lleva a cabo vigilancia. En algunos ejemplos, se puede generar audio al realizar simulaciones, y crear/mezclar (por ejemplo, movimientos, juegos). Los objetos de audio pueden ser utilizados en Hollywood (por ejemplo, estudios IMAX). En algunos ejemplos, la adquisición 301 puede ser realizada por un creador de contenido, tal como el creador de contenido 12 de la Fig. 3.

[0514] La edición 302 puede representar las técnicas del ecosistema de audio 300 donde el contenido de audio se edita y/o se modifica. Como ejemplo, el contenido de audio se puede editar combinando múltiples unidades de contenido de audio en una sola unidad de contenido de audio. Como otro ejemplo, el contenido de audio puede ser editado ajustando el contenido de audio real (por ejemplo, ajustando los niveles de uno o más componentes de frecuencia del contenido de audio). En algunos ejemplos, la edición 302 puede realizarse mediante un sistema de edición de audio, tal como el sistema de edición de audio 18 de la Fig. 3. En algunos ejemplos, la edición 302 puede realizarse en un dispositivo móvil, tal como uno o más de los dispositivos móviles ilustrados en la Fig. 29.

[0515] La codificación 303 puede representar las técnicas del ecosistema de audio 300 donde el contenido de audio está codificado en una representación del contenido de audio. En algunos ejemplos, la representación del contenido de audio puede ser un flujo de bits, tal como el flujo de bits 21 de la Fig. 3. En algunos ejemplos, la codificación 302 puede realizarse mediante un dispositivo de codificación de audio, tal como el dispositivo de codificación de audio 20 de la Fig. 3.

[0516] La transmisión 304 puede representar los elementos del ecosistema de audio 300 donde el contenido de audio se transporta desde un creador de contenido a un consumidor de contenido. En algunos ejemplos, el contenido de audio puede ser transportado en tiempo real o casi en tiempo real. Por ejemplo, el contenido de audio puede transmitirse al consumidor de contenido. En algunos ejemplos, el contenido de audio puede ser transportado codificando el contenido de audio en un medio, tal como un medio de almacenamiento legible por ordenador. Por ejemplo, el contenido de audio puede almacenarse en un disco, unidad y similares (por ejemplo, un disco Blu-ray, una tarjeta de memoria, un disco duro, etc.)

[0517] La reproducción 305 puede representar las técnicas del ecosistema de audio 300 donde el contenido de audio se reproduce y reproduce para el consumidor de contenido. En algunos ejemplos, la reproducción 305 puede incluir renderizar un campo sonoro 3D basado en uno o más aspectos de un entorno de reproducción. En otras palabras, la reproducción 305 puede estar basada en un paisaje acústico local.

[0518] La Fig. 13 es un diagrama que ilustra un ejemplo de un ecosistema de audio de la Fig. 12 con mayor detalle. Como se ilustra en la Fig. 13, el ecosistema de audio 300 puede incluir contenido de audio 308, estudios de cine 310, estudios de música 311, estudios de audio de juegos 312, contenido de audio basado en canales 313, motores de codificación 314, derivaciones de audio de juegos 315, motores 316 de codificación/317. Un ejemplo de estudio de audio de juegos 312 se ilustra en la Fig. 26. Algunos ejemplos de motores 316 de codificación/reproducción de audio de juegos se ilustran en la Fig. 27.

[0519] Como se ilustra en la Fig. 13, los estudios de cine 310, los estudios de música 311 y los estudios de audio de juegos 312 pueden recibir contenido de audio 308. En algún ejemplo, el contenido de audio 308 puede representar la salida de la adquisición 301 de la Fig. 12. Los estudios de películas 310 pueden emitir contenido de audio basado en canales 313 (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 310 pueden emitir contenido de audio basado en canales 313 (por ejemplo, en 2.0 y 5.1), tal como utilizando un DAW. En cualquier caso, los motores de codificación 314 pueden recibir y codificar el contenido de audio basado en el canal 313 basado 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 317. De esta manera, los motores de codificación 314 pueden ser un ejemplo de codificación 303 de la Fig. 12. Los estudios de audio de juegos 312 pueden emitir uno o más fragmentos de audios de juego 315, por ejemplo, mediante un DAW. Los motores 316 de codificación/reproducción de audios de juego pueden codificar y/o hacer que las derivaciones de audio 315 se conviertan en contenido de audio basado en canales para la salida por los sistemas de suministro 317. En algunos ejemplos, la salida de estudios de películas 310, estudios de música 311 y estudios de audio de juegos 312 puede representar la salida de edición 302 de la Fig. 12. En algunos ejemplos, la salida de los motores de codificación 314 y/o motores 316 de codificación/reproducción de audios de juego puede ser transportada a sistemas de suministro 317 a través de las técnicas de transmisión 304 de la Fig. 12.

[0520] La Fig. 14 es un diagrama que ilustra otro ejemplo del ecosistema de audio de la Fig. 12 con mayor detalle. Como se ilustra en la Fig. 14, el ecosistema de audio 300B puede incluir objetos de audio de grabación de difusión 319, sistemas de audio profesionales 320, captura de dispositivos de consumo 322, formato de audio HOA 323, reproducción de dispositivos 324, audio de consumo, TV y accesorios 325 y sistemas de audio para automóviles 326.

[0521] Como se ilustra en la Fig. 14, los objetos de audio de grabación de difusión 319, los sistemas de audio profesionales 320 y la captura de dispositivos de consumo 322 pueden codificar todos sus resultados utilizando el formato de audio HOA 323. De esta manera, el contenido de audio puede codificarse utilizando el formato de audio HOA 323 en una representación única que puede reproducirse utilizando la representación en dispositivo 324, audio de consumo, TV y accesorios 325 y sistemas de audio para automóviles 326. 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.).

[0522] Las Figs. 15A y 15B son diagramas que ilustran otros ejemplos del ecosistema de audio de la Fig. 12 con mayor detalle. Como se ilustra en la Fig. 15A, el ecosistema de audio 300C puede incluir elementos de adquisición 331 y elementos de reproducción 336. Los elementos 331 de adquisición pueden incluir dispositivos 332 de adquisición cableados y/o inalámbricos (por ejemplo, micrófonos Eigen), captura de sonido envolvente en dispositivo 334 y dispositivos móviles 335 (por ejemplo, teléfonos inteligentes y tabletas). En algunos ejemplos, los dispositivos de adquisición cableados y/o inalámbricos 332 pueden estar acoplados al dispositivo móvil 335 a través de canales de comunicación cableados y/o inalámbricos 333.

[0523] De acuerdo con una o más técnicas de esta descripción, el dispositivo móvil 335 puede usarse para adquirir un campo sonoro. Por ejemplo, el dispositivo móvil 335 puede adquirir un campo sonoro a través de dispositivos de adquisición cableados y/o inalámbricos 332 y/o captura de sonido envolvente en el dispositivo 334 (por ejemplo, una pluralidad de micrófonos integrados en el dispositivo móvil 335). El dispositivo móvil 335 puede entonces codificar el campo sonoro adquirido en las HOA 337 para su reproducción mediante uno o más de los elementos de reproducción 336. Por ejemplo, un usuario del dispositivo móvil 335 puede grabar (adquirir un campo sonoro de) un evento en directo (por ejemplo, una reunión, una conferencia, una obra de teatro, un concierto, etc.) y codificar la

grabación en HOAs.

[0524] El dispositivo móvil 335 también puede utilizar uno o más de los elementos de reproducción 336 para reproducir el campo sonoro codificado por HOA. Por ejemplo, el dispositivo móvil 335 puede decodificar el campo sonoro codificado por HOA y emitir una señal a uno o más de los elementos de reproducción 336 que hace que uno o más de los elementos de reproducción 336 recreen el campo sonoro. Como un ejemplo, el dispositivo móvil 335 puede utilizar 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 335 puede utilizar soluciones de acoplamiento 339 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/o hogares). Como otro ejemplo, el dispositivo móvil 335 puede utilizar la renderización de auriculares 340 para enviar la señal a un conjunto de auriculares, por ejemplo, para crear un sonido binaural realista.

[0525] En algunos ejemplos, un dispositivo móvil particular 335 puede adquirir un campo sonoro 3D y reproducir el mismo campo sonoro 3D en un momento posterior. En algunos ejemplos, el dispositivo móvil 335 puede adquirir un campo sonoro 3D, codificar el campo sonoro 3D en HOA y transmitir el campo sonoro 3D codificado a uno o más dispositivos (por ejemplo, otros dispositivos móviles y/u otros dispositivos no móviles) para su reproducción.

[0526] Como se ilustra en la Fig. 15B, el ecosistema de audio 300D puede incluir contenido de audio 343, estudios de juego 344, contenido de audio codificado 345, motores de renderización 346 y sistemas de entrega 347. En algunos ejemplos, los estudios de juego 344 pueden incluir uno o más DAWs que pueden permitir la edición de señales HOA. Por ejemplo, uno o más DAWs pueden incluir plugins 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 344 pueden producir nuevos formatos de derivación que soporten HOA. En cualquier caso, los estudios de juego 344 pueden emitir contenido de audio codificado 345 a los motores de representación 346 que pueden proporcionar un campo sonoro para su reproducción mediante los sistemas de entrega 347.

[0527] La Fig. 16 es un diagrama que ilustra un dispositivo de codificación de vídeo ejemplar que puede realizar diferentes aspectos de las técnicas descritas en esta divulgación. Como se ilustra en la Fig. 16, el ecosistema de audio 300E puede incluir contenido de audio 3D original 351, codificador 352, flujo de bits 353, decodificador 354, renderizador 355 y elementos de reproducción 356. Como se ilustra adicionalmente en la Fig. 16, el codificador 352 puede incluir análisis y descomposición de campo sonoro 357, extracción de fondo 358, determinación de relevancia de fondo 359, codificación de audio 360, extracción de audio de primer plano/distinto 361 y codificación de audio 362. En algunos ejemplos, el codificador 352 puede estar configurado para realizar operaciones similares al dispositivo de codificación de audio 20 de las Figs. 3 y 4. En algunos ejemplos, el análisis y la descomposición del campo sonoro 357 pueden estar configurados para realizar operaciones similares a la unidad de análisis del campo sonoro 44 de la Fig. 4. En algunos ejemplos, la extracción de fondo 358 y la determinación de relevancia de fondo 359 pueden estar configuradas para realizar operaciones similares a la unidad de selección de $_{BG}$ 48 de la Fig. 4. En algunos ejemplos, la codificación de audio 360 y la codificación de audio 362 pueden configurarse para llevar a cabo operaciones similares a la unidad codificadora de audio psicoacústico 40 de la Fig. 4. En algunos ejemplos, la extracción de audio de primer plano/distinto 363 puede configurarse para llevar a cabo operaciones similares a la unidad de selección del primer plano 36 de la Fig. 4.

[0528] En algunos ejemplos, la extracción de audio de primer plano/distinto 361 puede analizar el contenido de audio correspondiente a la trama de vídeo 390 de la Fig. 33. Por ejemplo, la extracción de audio en primer plano/distinto 361 puede determinar que el contenido de audio correspondiente a las regiones 391A-391C es audio de primer plano.

[0529] Como se ilustra en la Fig. 16, el codificador 352 puede configurarse para codificar el contenido original 351, que puede tener una velocidad de bits de 25-75 Mbps, en el flujo de bits 353, que puede tener un velocidad de bits de 256 kbps - 1,2 Mbps. La Fig. 17 es un diagrama que ilustra un ejemplo del dispositivo de codificación de audio de la Fig. 16 con mayor detalle.

[0530] La Fig. 18 es un diagrama que ilustra un dispositivo de descodificación de audio ejemplar que puede implementar diferentes aspectos de las técnicas descritas en esta divulgación. Como se ilustra en la Fig. 18, el ecosistema de audio 300E puede incluir contenido de audio 3D original 351, codificador 352, flujo de bits 353, decodificador 354, renderizador 355 y elementos de reproducción 356. Como se ilustra adicionalmente en la Fig. 16, el decodificador 354 puede incluir decodificador de audio 363, decodificador de audio 364, reconstrucción de primer plano 365 y mezcla 366. En algunos ejemplos, el decodificador 354 puede estar configurado para realizar operaciones similares al dispositivo de decodificación de audio 24 de las Figs. 3 y 5. En algunos ejemplos, el decodificador de audio 363, el decodificador de audio 364 pueden estar configurados para realizar operaciones similares a la unidad de decodificación psicoacústica 80 de la Fig. 5. En algunos ejemplos, la reconstrucción de primer plano 365 puede estar configurada para realizar operaciones similares a la unidad de formulación del primer plano 78 de la Fig. 5.

[0531] Como se ilustra en la Fig. 16, el decodificador 354 puede estar configurado para recibir y decodificar el flujo de bits 353 y dar salida al campo sonoro 3D reconstruido resultante al renderizador 355 que puede entonces hacer que, uno o más de los elementos de reproducción 356, produzcan una representación del contenido 3D original 351. La Fig. 19 es un diagrama que ilustra un ejemplo del dispositivo de descodificación de audio de la Fig. 18 con mayor detalle.

[0532] Las Figs. 20A a 20G son diagramas que ilustran ejemplos de dispositivos de adquisición de audio que pueden realizar diversos aspectos de las técnicas descritas en esta descripción. La Fig. 20A ilustra el micrófono Eigen 370 que puede incluir una pluralidad de micrófonos que están configurados colectivamente para grabar un campo sonoro 3D. En algunos ejemplos, la pluralidad de micrófonos del micrófono Eigen 370 puede estar situada en la superficie de una bola sustancialmente esférica con un radio de aproximadamente 4 cm. En algunos ejemplos, el dispositivo de codificación de audio 20 puede estar integrado en el micrófono de Eigen de manera que emita un flujo de bits 17 directamente desde el micrófono 370.

[0533] La Fig. 20B ilustra el camión de producción 372 que puede estar configurado para recibir una señal de uno o más micrófonos, tal como uno o más micrófonos Eigen 370. El camión de producción 372 también puede incluir un codificador de audio, tal como el codificador de audio 20 de la Fig. 3.

[0534] Las Figs. 20C-20E ilustran el dispositivo móvil 374 que puede incluir una pluralidad de micrófonos que están configurados colectivamente para grabar un campo sonoro 3D. En otras palabras, la pluralidad de micrófonos puede tener diversidad X, Y, Z. En algunos ejemplos, el dispositivo móvil 374 puede incluir un micrófono 376 que puede ser girado para proporcionar diversidad X, Y, Z con respecto a uno o más micrófonos del dispositivo móvil 374. El dispositivo móvil 374 también puede incluir un codificador de audio, tal como el codificador de audio 20 de la Fig. 3.

[0535] La Fig. 20F ilustra un dispositivo de captura de vídeo robusto 378 que puede configurarse para grabar un campo sonoro 3D. En algunos ejemplos, el dispositivo de captura de vídeo robusto 378 puede estar unido a un casco de un usuario que participa en una actividad. Por ejemplo, el dispositivo de captura de vídeo reforzado 378 puede estar unido a un casco de un usuario de rafting en aguas bravas. De esta manera, el dispositivo de captura de vídeo robusto 378 puede capturar un campo sonoro 3D que representa la acción alrededor del usuario (por ejemplo, agua que se estrella detrás del usuario, otro rafter que habla delante del usuario, etc.).

[0536] La Fig. 20G ilustra un dispositivo móvil mejorado secundario 380 que puede configurarse para grabar un campo sonoro 3D. En algunos ejemplos, el dispositivo móvil 380 puede ser similar al dispositivo móvil 335 de la Fig. 15, con la adición de uno o más accesorios. Por ejemplo, un micrófono Eigen puede estar unido al dispositivo móvil 335 de la Fig. 15 para formar un dispositivo móvil mejorado secundario 380. De esta manera, el dispositivo móvil mejorado secundario 380 puede capturar una versión de calidad superior del campo sonoro 3D que utilizar únicamente componentes de captura de sonido integrados en el dispositivo móvil mejorado secundario 380.

[0537] Las Figs. 21A-21E son diagramas que ilustran ejemplos de dispositivos de reproducción de audio que pueden realizar diversos aspectos de las técnicas descritas en esta descripción. Las Figs. 21A y 21B ilustran una pluralidad de altavoces 382 y barras de sonido 384. De acuerdo con una o más técnicas de esta descripción, los altavoces 382 y/o las barras de sonido 384 pueden disponerse en cualquier configuración arbitraria mientras se sigue reproduciendo un campo sonoro 3D. Las Figs. 21C-21E ilustran una pluralidad de dispositivos de reproducción de auriculares 386-386C. Los dispositivos de reproducción de auriculares 386-386C pueden estar acoplados a un decodificador a través de una conexión cableada o inalámbrica. De acuerdo con una o más técnicas de esta descripción, se puede utilizar una única representación genérica de un campo sonoro para renderizar el campo sonoro sobre cualquier combinación de altavoces 382, barras de sonido 384 y dispositivos de reproducción de auriculares 386-386C.

[0538] Las Figs. 22A a 22H son diagramas que ilustran ejemplos de entornos de reproducción de audio de acuerdo con una o más técnicas descritas en esta descripción. Por ejemplo, la Fig. 22A ilustra un entorno de reproducción de altavoces 5.1, la Fig. 22B ilustra un entorno de reproducción de altavoces 2.0 (por ejemplo, estéreo), la Fig. 22C ilustra un entorno de reproducción de altavoces 9.1 con altavoces delanteros de altura completa, las Figs. 22D y 22E ilustran cada uno un entorno de reproducción de altavoces 22.2. 22F ilustra un entorno de reproducción de altavoces 16.0, la Fig. 22G ilustra un entorno de reproducción de altavoces para automóviles, y la Fig. 22H ilustra un dispositivo móvil con un entorno de reproducción de oído.

[0539] De acuerdo con una o más técnicas de esta descripción, se puede utilizar una única representación genérica de un campo sonoro para representar el campo sonoro en cualquiera de los entornos de reproducción ilustrados en las Figs. 22A - 22H. Adicionalmente, las técnicas de esta descripción permiten a un renderizador renderizar un campo sonoro desde una representación genérica para reproducción en entornos de reproducción distintos de los ilustrados en las Figs. 22A - 22H. Por ejemplo, si las consideraciones de diseño prohíben la colocación correcta de los altavoces de acuerdo con un entorno de reproducción de altavoces 7.1 (por ejemplo, si no es posible colocar un altavoz envolvente derecho), las técnicas de esta descripción permiten que un renderizador compense con los otros 6 altavoces tales que la reproducción se puede lograr en un entorno de reproducción de

altavoces 6.1.

[0540] Como se ilustra en la Fig. 23, un usuario puede ver un juego de deportes mientras utiliza los auriculares 386. De acuerdo con una o más técnicas de esta descripción, se puede adquirir el campo sonoro 3D del juego deportivo (por ejemplo, uno o más micrófonos Eigen se pueden colocar en y/o alrededor del estadio de béisbol ilustrado en la Fig. 24), los coeficientes HOA correspondientes al campo sonoro 3D pueden obtenerse y transmitirse a un decodificador, el decodificador puede determinar reconstruir el campo sonoro 3D basado en los coeficientes HOA y emitir el campo sonoro 3D a un renderizador, el renderizador puede obtener una indicación como el tipo de entorno de reproducción (por ejemplo, los auriculares) y convertir el campo sonoro 3D en señales que hacen que los auriculares emitan una representación del campo sonoro 3D del juego deportivo. En algunos ejemplos, el renderizador puede obtener una indicación del tipo de entorno de reproducción de acuerdo con las técnicas de la Fig. 25. De esta manera, el renderizador puede "adaptarse" a diferentes ubicaciones de altavoces, tipo de números, tamaño e idealmente ecualizar para el entorno local.

[0541] La Fig. 28 es un diagrama que ilustra una configuración de altavoz que puede ser simulada por auriculares de acuerdo con una o más técnicas descritas en esta descripción. Como se ilustra en la Fig. 28, las técnicas de esta descripción pueden permitir que un usuario que lleve audífonos 389 experimente un campo sonoro como si el campo sonoro fuera reproducido por los altavoces 388. De esta manera, un usuario puede escuchar un campo sonoro en 3D sin que se emita el sonido en un área grande.

[0542] La Fig. 30 es un diagrama que ilustra una trama de vídeo asociada con un campo sonoro 3D que puede procesarse de acuerdo con una o más técnicas descritas en esta descripción.

[0543] Las Figs. 31A-31M son diagramas que ilustran los gráficos 400A-400M que muestran diversos resultados de simulación de realizar categorización sintética o grabada del campo sonoro de acuerdo con diversos aspectos de las técnicas descritas en esta descripción. En el ejemplo de la Fig. 31A a 31M, cada uno de los gráficos 400A-400M incluye un umbral 402 que está indicado por una línea de puntos y un objeto de audio respectivo 404A-404M(colectivamente, "los objetos de audio 404") indicados por una línea discontinua.

[0544] Cuando se determina que los objetos de audio 404 a través del análisis descrito anteriormente con respecto a la unidad de análisis de contenido 26 están por debajo del umbral 402, la unidad de análisis de contenido 26 determina que el objeto correspondiente de los objetos de audio 404 representa un objeto de audio que ha sido grabado. Como se ilustra en el ejemplo de la Figs. 31B, 31D-31H y 31J-31L, la unidad de análisis de contenido 26 determina que los objetos de audio 404B, 404D-404H, 404J-404L están por debajo del umbral 402 (al menos más del 90% del tiempo y a menudo el 100% del tiempo) y por lo tanto representan objetos de audio grabados. Como se ilustra en el ejemplo de la Figs. 31A, 31C y 31I, la unidad de análisis de contenido 26 determina que los objetos de audio 404A, 404C y 404I exceden el umbral 402 y por lo tanto representan objetos de audio sintéticos.

[0545] En el ejemplo de la Fig. 31M, el objeto de audio 404M representa un objeto de audio sintético/grabado mezclado, que tiene algunas partes sintéticas (por ejemplo, por encima del umbral 402) y algunas partes sintéticas (por ejemplo, por debajo del umbral 402). La unidad de análisis de contenido 26 identifica en este caso las partes sintéticas y grabadas del objeto de audio 404M con el resultado de que el dispositivo de codificación de audio 20 genera el flujo de bits 21 para incluir tanto datos de audio codificados basados en la direccionalidad como datos de audio codificados basados en vector.

[0546] La Fig. 32 es un diagrama que ilustra un gráfico 406 de valores singulares de una matriz S descompuesta a partir de coeficientes ambisónicos de orden superior de acuerdo con las técnicas descritas en esta descripción. Como se muestra en la Fig. 32, los valores singulares distintos de cero que tienen valores grandes son pocos. La unidad de análisis del campo sonoro 44 de la Fig. 4 puede analizar estos valores singulares para determinar los componentes de primer plano nFG (o, en otras palabras, predominantes) de los vectores US[*1*] reordenados 33' y los vectores V[*k*] reordenados 35'.

[0547] Las Figs. 33A y 33B son diagramas que ilustran los respectivos gráficos 410A y 410B que muestran un impacto potencial que tiene la reordenación cuando se codifican los vectores que describen componentes en primer plano del campo sonoro de acuerdo con las técnicas descritas en esta descripción. El gráfico 410A muestra el resultado de la codificación de al menos algunos de los vectores US[*k*] no ordenados 33 (o, en otras palabras, los originales), mientras que el gráfico 410B muestra el resultado de la codificación de los correspondientes vectores US[*k*] ordenados 33'. La gráfica superior en cada una de las gráficas 410A y 410B muestra el error en la codificación, donde es probable que solo se observe un error notable en el gráfico 410B en los límites de la trama. Por consiguiente, las técnicas de reordenación descritas en esta descripción pueden facilitar o de otra manera promover la codificación de objetos monoaudio utilizando un codificador de audio heredado.

[0548] Las Figs. 34 y 35 son diagramas conceptuales que ilustran diferencias entre la identificación basada en energía y la direccionalidad de distintos objetos de audio, de acuerdo con esta descripción. En el ejemplo de la Fig. 34, los vectores que presentan mayor energía se identifican como siendo distintos objetos de audio, independientemente de la direccionalidad. Como se muestra en la Fig. 34, se determina que los objetos de audio

que están posicionados de acuerdo con valores de energía más altos (dibujados en un eje y) están "en primer plano", independientemente de la direccionalidad (por ejemplo, representados por cocientes de direccionalidad dibujados en un eje x).

[0549] La Fig. 35 ilustra la identificación de objetos de audio distintos basándose tanto en la direccionalidad como en la energía, tal como de acuerdo con las técnicas implementadas por la unidad de análisis del campo sonoro 44 de la Fig. 4. Como se muestra en la Fig. 35, los cocientes de mayor direccionalidad se trazan hacia la izquierda del eje x, y los mayores niveles de energía se representan en la parte superior del eje y. En este ejemplo, la unidad de análisis del campo sonoro 44 puede determinar que, objetos de audio distintos (por ejemplo, que están "en primer plano"), están asociados con datos vectoriales dibujados relativamente hacia la parte superior izquierda del gráfico. Como ejemplo, la unidad de análisis del campo sonoro 44 puede determinar que los vectores que están representados en el cuadrante superior izquierdo del gráfico están asociados con distintos objetos de audio.

[0550] Las Figs. 36A a 36F son diagramas que ilustran proyecciones de al menos una parte de la versión descompuesta de coeficientes armónicos esféricos en el dominio espacial para realizar interpolación de acuerdo con diversos aspectos de las técnicas descritas en esta descripción. La Fig. 36A es un diagrama que ilustra la proyección de uno o más de los V[*k*] vectores 35 en una esfera 412. En el ejemplo de la Fig. 36A, cada número identifica un coeficiente armónico esférico diferente proyectado sobre la esfera (posiblemente asociado con una fila y/o columna de la matriz V 19 '). Los diferentes colores sugieren una dirección de un componente de audio distinto, donde el color más claro (y progresivamente más oscuro) indica la dirección primaria del componente distinto. La unidad de interpolación espacio-temporal 50 del dispositivo de codificación de audio 20 que se muestra en el ejemplo de la Fig. 4 puede realizar interpolación espacio-temporal entre cada uno de los puntos rojos para generar la esfera que se muestra en el ejemplo de la Fig. 36A.

[0551] La Fig. 36B es un diagrama que ilustra la proyección de uno o más de los V[*k*] vectores 35 sobre una viga. La unidad de interpolación espacio-temporal 50 puede proyectar una fila y/o una columna de los vectores V[*k*] 35 o múltiples filas y/o columnas de los vectores V[*k*] 35 para generar el haz 414 que se muestra en el ejemplo de la Fig. 36B.

[0552] La Fig. 36C es un diagrama que ilustra una sección transversal de una proyección de uno o más vectores de uno o más de los vectores V[*k*] 35 en una esfera, tal como la esfera 412 que se muestra en el ejemplo de la Fig. 36.

[0553] En las Figs. 36D-36G son ejemplos de instantáneas de tiempo (más de 1 trama de unos 20 milisegundos) cuando se pueden ilustrar diferentes fuentes de sonido (abejas, helicópteros, música electrónica y personas en un estadio) en un espacio tridimensional.

[0554] Las técnicas descritas en esta descripción permiten que la representación de estas diferentes fuentes de sonido sea identificada y representada utilizando un único vector US[*k*] y un único vector V[*k*]. La variabilidad temporal de las fuentes sonoras se representa en el vector US[*k*], mientras que la distribución espacial de cada fuente sonora está representada por el único vector V[*k*]. Un vector V[*k*] puede representar el ancho, la ubicación y el tamaño de la fuente de sonido. Además, el único vector V[*k*] puede representarse como una combinación lineal de funciones de base armónica esférica. En las gráficas de las Figs. 36D-36G, la representación de las fuentes de sonido se basa en la transformación del único vector V en un sistema de coordenadas espaciales. Procedimientos similares de ilustración de fuentes sonoras se utilizan en las Figs. 36-36C.

[0555] La Fig. 37 ilustra una representación de técnicas para obtener una interpolación espacio-temporal como se describe en este documento. La unidad de interpolación espacio-temporal 50 del dispositivo de codificación de audio 20 que se muestra en el ejemplo de la Fig. 4 puede realizar la interpolación espacio-temporal descrita más adelante con más detalle. La interpolación espacio-temporal puede incluir la obtención de componentes espaciales de mayor resolución en las dimensiones espaciales y temporales. Los componentes espaciales pueden estar basados en una descomposición ortogonal de una señal multidimensional compuesta por coeficientes ambisónicos de orden superior (HOA) (o, como coeficientes HOA también pueden ser referidos, "coeficientes armónicos esféricos").

[0556] En el gráfico ilustrado, los vectores V_1 y V_1 representan los vectores correspondientes de dos componentes espaciales diferentes de una señal multidimensional. Los componentes espaciales pueden obtenerse mediante una descomposición en bloques de la señal multidimensional. En algunos ejemplos, los componentes espaciales resultan de realizar una forma en bloques de SVD con respecto a cada bloque (que puede referirse a una trama) de datos de audio ambisónicos (HOA) de orden superior (donde estos datos de audio ambisónicos incluyen bloques, muestras o cualquier otra forma de datos de audio multicanal). Se puede utilizar una variable M para indicar la longitud de una trama de audio en muestras.

[0557] Por consiguiente, V_1 y V_2 pueden representar los vectores correspondientes de los vectores V[*k*] de primer plano 51_k y los vectores V[*k*-1] de primer plano 51_{k-1} para bloques secuenciales de los coeficientes HOA 11. Por ejemplo, V_1 puede representar un primer vector de los vectores V[*k*-1] de primer plano 51_{k-1} para una primera trama (*k*-1), mientras que V_2 puede representar un primer vector de los vectores V[*k*] de primer plano 51_k para una segunda

y siguiente trama (*k*). V_1 y V_2 pueden representar un componente espacial para un único objeto de audio incluido en la señal multidimensional.

[0558] Los vectores interpolados V_x para cada *x* se obtienen ponderando V_1 y V_2 de acuerdo con un número de segmentos temporales o "muestras de tiempo", *x*, para una componente temporal de la señal multidimensional a la que se pueden aplicar los vectores interpolados V_x para suavizar el componente temporal (y, en consecuencia, en algunos casos el espacial). Suponiendo una composición SVD, como se ha descrito anteriormente, se puede obtener el suavizado de las señales nFG 49 haciendo una división vectorial de cada vector de muestra de tiempo (por ejemplo, una muestra de los coeficientes HOA 11) con la correspondiente V_x interpolada. Esto es, US[n] = HOA[n] * $V_x[n]^{-1}$, donde esto representa un vector de fila multiplicado por un vector de columna, produciendo así un elemento escalar para US. $V_x [n]^{-1}$ se puede obtener como un pseudoinversa de $V_x[n]$.

[0559] Con respecto a la ponderación de V_1 y V_2, V_1 se pondera proporcionalmente más bajo a lo largo de la dimensión de tiempo debido a que el V_2 se produce subsiguiente en el tiempo a V_1. Es decir, aunque los vectores V[*k*-1] de primer plano 51_{k-1} son componentes espaciales de la descomposición, los vectores V[*k*] de primer plano secuencial temporalmente 51_k representan valores diferentes del componente espacial a lo largo del tiempo. Por consiguiente, el peso de V_1 disminuye mientras que el peso de V_2 crece a medida que *x* aumenta a lo largo de *t*. Aquí, d_1 y d_2 representan pesos.

[0560] La Fig. 38 es un diagrama de bloques que ilustra las matrices US artificiales, US_1 y US_2, para bloques SVD secuenciales para una señal multidimensional de acuerdo con las técnicas descritas en la presente memoria. Pueden aplicarse vectores V interpolados a los vectores de fila de las matrices US artificiales para recuperar la señal multidimensional original. Más específicamente, la unidad de interpolación espacio-temporal 50 puede multiplicar la pseudoinversa de los vectores V[*k*] de primer plano interpolados al resultado de multiplicar las señales nFG 49 por los vectores V[*k*] de primer plano 51_k (que pueden designarse como coeficientes HOA de primer plano) para obtener muestras interpoladas K/2, que pueden utilizarse en lugar de las muestras K/2 de las señales nFG como las primeras muestras K/2 como se muestra en el ejemplo de la Fig. 38 de la matriz U_2.

[0561] La Fig. 39 es un diagrama de bloques que ilustra la descomposición de tramas subsiguientes de una señal ambisónica de orden superior (HOA) utilizando la descomposición en valores singulares y el suavizado de los componentes espacio-temporales de acuerdo con las técnicas descritas en esta descripción. La trama n-1 y la trama n (que también se pueden designar como trama n y trama n + 1) representan tramas posteriores en el tiempo, con cada trama que comprende 1024 segmentos de tiempo y que tiene un orden HOA de 4, dando $(4+1)^2 = 25$ coeficientes. Las matrices US que son matrices U suavizadas artificialmente en la trama n-1 y la trama n pueden obtenerse mediante la aplicación de vectores V interpolados como se ilustra. Cada fila gris o columna de vectores representa un objeto de audio.

Cálculo de la representación HOA de señales activas basadas en vectores

[0562] El CVEC*k* instantáneo se crea tomando cada una de las señales basadas en vectores representadas en XVEC*k* y multiplicándolas por su vector espacial (decuantizado) correspondiente, *V*VEC*k*. Cada *V*VEC*k* está representado en MVECk. Por lo tanto, para una señal HOA de orden *L* y *M* señales basadas en vectores, habrá *M* señales basadas en vectores, cada una de las cuales tendrá una dimensión dada por la longitud de trama P. Estas señales pueden representarse así: ***X***VEC*kmn*, *n=0,..P-1*; *m=0,..M-1*. Correspondientemente, habrá *M* vectores espaciales, *V*VEC*k* de dimensión $(L+1)^2$. Estos pueden representarse como *M*VEC*kml*, *l=0,..*, $(L+1)^{2-1}$; *m=0,..,M-1*. La representación HOA para cada señal basada en vectores, *C*VEC*km*, es una multiplicación de vectores matricial dada por:

$$\boldsymbol{C}\text{VEC}km=(\boldsymbol{X}\text{VEC}km(\boldsymbol{M}\text{VEC}km)\ T)\ T$$

que produce una matriz de $(L+1)^2$ por P. La representación HOA completa se da sumando la contribución de cada señal basada en vectores como sigue:

$$\boldsymbol{C}\text{VEC}k=m=0M-1\ \boldsymbol{C}\text{VEC}k[m]$$

Interpolaciones espacio-temporales de V vectores

[0563] Sin embargo, con el fin de mantener la continuidad espacio-temporal suave, el cálculo anterior solo se lleva a cabo para la parte de la longitud de trama *PB*. Las primeras *B* muestras de una matriz HOA, se llevan a cabo en su lugar utilizando un conjunto interpolado de *M*VEC*kml*, *m = 0,..*, *M-1* ; *L = 0,..*, $(L+1)^2$, obtenido del ***M***VEC*km* actual y los valores ***M***VECk *-lm* anteriores. Esto da como resultado un vector espacial de mayor densidad de tiempo a medida que obtenemos un vector para cada muestra de tiempo *p*, como sigue:

$$\boldsymbol{M}\text{VEC}kmp=pB-1\boldsymbol{M}\text{VEC}km+B-1-pB-1\boldsymbol{M}\text{VEC}k-1m,\ p=0,..,B-1.$$

Para cada muestra de tiempo p, se calcula un nuevo vector HOA de $(L+1)^2$ dimensiones tal como:

$$\boldsymbol{C}\mathrm{VEC}kp=(\boldsymbol{X}\mathrm{VEC}kmp)\mathbf{M}\mathrm{VEC}kmp,\ p=0,..,B-1$$

Estas primeras *B* muestras se aumentan con las *P-B* muestras de la sección anterior para dar como resultado la representación HOA completa, C*V*EC*km* de la *mésima* señal basada en vectores.

[0564] En el decodificador (por ejemplo, el dispositivo de decodificación de audio 24 que se muestra en el ejemplo de la Fig. 5), para ciertos sonidos distintos, de primer plano o de base vectorial predominante, el vector V de la trama anterior y el vector V de la trama actual pueden interpolarse utilizando una interpolación lineal (o no lineal) para producir un vector V de mayor resolución (en el tiempo) interpolado sobre un segmento de tiempo particular. La unidad de interpolación espacial temporal 76 puede realizar esta interpolación, en la que la unidad de interpolación espacio-temporal 76 puede entonces multiplicar el vector US en la trama actual con el vector V de mayor resolución interpolado para producir la matriz HOA sobre ese segmento temporal particular.

[0565] Alternativamente, la unidad de interpolación espacio-temporal 76 puede multiplicar el vector US con el vector V de la trama actual para crear una primera matriz HOA. El decodificador puede multiplicar adicionalmente el vector US con el vector V desde la trama anterior para crear una segunda matriz HOA. La unidad de interpolación espacio-temporal 76 puede entonces aplicar una interpolación lineal (o no lineal) a la primera y segunda matrices HOA sobre un segmento de tiempo particular. La salida de esta interpolación puede coincidir con la de la multiplicación del vector US con un vector V interpolado, proporcionando matrices/vectores de entrada comunes.

[0566] A este respecto, las técnicas pueden permitir que el dispositivo de codificación de audio 20 y/o el dispositivo de decodificación de audio 24 estén configurados para funcionar de acuerdo con las siguientes cláusulas.

Cláusula 135054-1C. Un dispositivo, tal como el dispositivo de codificación de audio 20 o el dispositivo de decodificación de audio 24, que comprende: uno o más procesadores configurados para obtener una pluralidad de componentes espaciales de mayor resolución en espacio y tiempo, en donde los componentes espaciales se basan en una descomposición ortogonal de una señal multidimensional compuesta de coeficientes armónicos esféricos.

Cláusula 135054-1D. Un dispositivo, tal como el dispositivo de codificación de audio 20 o el dispositivo de decodificación de audio 24, que comprende: uno o más procesadores configurados para suavizar al menos uno de los componentes espaciales y componentes temporales de la primera pluralidad de coeficientes armónicos esféricos y la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-1E. Un dispositivo, tal como el dispositivo de codificación de audio 20 o el dispositivo de decodificación de audio 24, que comprende: uno o más procesadores configurados para obtener una pluralidad de componentes espaciales de mayor resolución en espacio y tiempo, en donde los componentes espaciales se basan en una descomposición ortogonal de una señal multidimensional compuesta de coeficientes armónicos esféricos.

Cláusula 135054-1G. Un dispositivo, tal como el dispositivo de codificación de audio 20 o el dispositivo de decodificación de audio 24, que comprende: uno o más procesadores configurados para obtener coeficientes armónicos esféricos de resolución incrementada descompuestos para un segmento de tiempo, al menos en parte, incrementando una resolución con respecto a una primera descomposición de una primera pluralidad de coeficientes armónicos esféricos y una segunda descomposición de una segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-2G. El dispositivo de la cláusula 135054-1G, en el que la primera descomposición comprende una primera matriz V representativa de los vectores singulares derechos de la primera pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-3G. El dispositivo de la cláusula 135054-1G, en el que la segunda descomposición comprende una segunda matriz V representativa de los vectores singulares derechos de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-4G. El dispositivo de la cláusula 135054-1G, en el que la primera descomposición comprende una primera matriz V representativa de los vectores singulares derechos de la primera pluralidad de coeficientes armónicos esféricos, y en la que la segunda descomposición comprende una segunda matriz V representativa de los vectores singulares derechos de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-5G. El dispositivo de la cláusula 135054-1G, en el que el segmento de tiempo comprende una subtrama de una trama de audio.

Cláusula 135054-6G. El dispositivo de la cláusula 135054-1G, en el que el segmento de tiempo comprende una muestra de tiempo de una trama de audio.

Cláusula 135054-7G. El dispositivo de la cláusula 135054-1G, en el que el uno o más procesadores están configurados para obtener una descomposición interpolada de la primera descomposición y la segunda descomposición para un coeficiente armónico esférico de la primera pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-8G. El dispositivo de la cláusula 135054-1G, en el que el uno o más procesadores están configurados para obtener descomposiciones interpoladas de la primera descomposición para una primera parte de la primera pluralidad de coeficientes armónicos esféricos incluidos en la primera trama y la segunda descomposición para una segunda parte de la segunda pluralidad de coeficientes armónicos esféricos incluidos en la segunda trama, en el que el uno o más procesadores están configurados además para aplicar las descomposiciones interpoladas a un primer componente de tiempo de la primera parte de la primera pluralidad de coeficientes armónicos esféricos incluidos en la primera trama para generar un primer componente artificial de tiempo de la primera pluralidad de coeficientes armónicos esféricos y aplicar las respectivas descomposiciones interpoladas a un segundo componente de tiempo de la segunda parte de la segunda pluralidad de coeficientes armónicos esféricos incluidos en la segunda trama para generar un segundo componente artificial de tiempo de la segunda pluralidad de coeficientes armónicos esféricos incluidos.

Cláusula 135054-9G. El dispositivo de la cláusula 135054-8G, en el que el primer componente de tiempo se genera realizando una síntesis basada en vectores con respecto a la primera pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-10G. El dispositivo de la cláusula 135054-8G, en el que el segundo componente de tiempo se genera realizando una síntesis basada en vectores con respecto a la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-11G. El dispositivo de la cláusula 135054-8G, en el que el uno o más procesadores están configurados además para recibir el primer componente artificial de tiempo y el segundo componente artificial de tiempo, calcular las descomposiciones interpoladas de la primera descomposición para la primera parte de la primera pluralidad de coeficientes armónicos esféricos y la segunda descomposición para la segunda parte de la segunda pluralidad de coeficientes armónicos esféricos, y aplicar inversos de las descomposiciones interpoladas al primer componente artificial de tiempo para recuperar el componente de primer tiempo y al segundo componente artificial de tiempo para recuperar el segundo componente de tiempo.

Cláusula 135054-12G. El dispositivo de la cláusula 135054-1G, en el que el uno o más procesadores están configurados para interpolar un primer componente espacial de la primera pluralidad de coeficientes armónicos esféricos y el segundo componente espacial de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-13G. El dispositivo de la cláusula 135054-12G, en el que el primer componente espacial comprende una primera matriz U representativa de los vectores singulares izquierdos de la primera pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-14G. El dispositivo de la cláusula 135054-12G, en el que el segundo componente espacial comprende una segunda matriz U representativa de los vectores singulares izquierdos de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-15G. El dispositivo de la cláusula 135054-12G, en el que el primer componente espacial es representativo de M segmentos temporales de coeficientes armónicos esféricos para la primera pluralidad de coeficientes armónicos esféricos y el segundo componente espacial es representativo de M segmentos temporales de coeficientes armónicos esféricos para la segunda pluralidad de los coeficientes armónicos esféricos.

Cláusula 135054-16G. El dispositivo de la cláusula 135054-12G, en el que el primer componente espacial es representativo de M segmentos temporales de coeficientes armónicos esféricos para la primera pluralidad de coeficientes armónicos esféricos y el segundo componente espacial es representativo de M segmentos temporales de coeficientes armónicos esféricos para la segunda pluralidad de coeficientes armónicos esféricos y en el que uno o más procesadores están configurados para obtener los coeficientes armónicos esféricos interpolados descompuestos para el segmento de tiempo comprende interpolar los últimos N elementos del primer componente espacial y los primeros N elementos del segundo componente espacial.

Cláusula 135054-17G. El dispositivo de la cláusula 135054-1G, en el que la segunda pluralidad de coeficientes armónicos esféricos es posterior a la primera pluralidad de coeficientes armónicos esféricos en el dominio de tiempo.

Cláusula 135054-18G. El dispositivo de la cláusula 135054-1G, en el que el uno o más procesadores están configurados además para descomponer la primera pluralidad de coeficientes armónicos esféricos para generar

la primera descomposición de la primera pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-19G. El dispositivo de la cláusula 135054-1G, en el que el uno o más procesadores están configurados además para descomponer la segunda pluralidad de coeficientes armónicos esféricos para generar la segunda descomposición de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-20G. El dispositivo de la cláusula 135054-1G, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a la primera pluralidad de coeficientes armónicos esféricos para generar una matriz U representativa de los vectores singulares izquierdos de la primera pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la primera pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la primera pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-21G. El dispositivo de la cláusula 135054-1G, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a la segunda pluralidad de coeficientes armónicos esféricos para generar una matriz U representativa de los vectores singulares izquierdos de la segunda pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la segunda pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-22G. El dispositivo de la cláusula 135054-1G, en el que la primera y la segunda pluralidad de coeficientes armónicos esféricos representan, cada una, una representación de onda plana del campo sonoro.

Cláusula 135054-23G. El dispositivo de la cláusula 135054-1G, en el que la primera y la segunda pluralidad de coeficientes armónicos esféricos representan cada uno uno o más objetos monoaudio mezclados entre sí.

Cláusula 135054-24G. El dispositivo de la cláusula 135054-1G, en el que la primera y la segunda pluralidad de coeficientes armónicos esféricos comprenden cada uno un primer y segundo coeficientes armónicos esféricos respectivos que representan un campo sonoro tridimensional.

Cláusula 135054-25G. El dispositivo de la cláusula 135054-1G, en el que la primera y la segunda pluralidad de coeficientes armónicos esféricos están asociadas cada una con al menos una función de base esférica que tiene un orden mayor que uno.

Cláusula 135054-26G. El dispositivo de la cláusula 135054-1G, en el que la primera y la segunda pluralidad de coeficientes armónicos esféricos están asociadas cada una con al menos una función de base esférica que tiene un orden igual a cuatro.

Cláusula 135054-27G. El dispositivo de la cláusula 135054-1G, en el que la interpolación es una interpolación ponderada de la primera descomposición y segunda descomposición, en la que los pesos de la interpolación ponderada aplicada a la primera descomposición son inversamente proporcionales a un tiempo representado por vectores de la primera y segunda descomposición y en la que los pesos de la interpolación ponderada aplicada a la segunda descomposición son proporcionales a un tiempo representado por vectores de la primera y segunda descomposición.

Cláusula 135054-28G. El dispositivo de la cláusula 135054-1G, en el que los coeficientes armónicos esféricos interpolados descompuestos suavizan al menos uno de los componentes espaciales y componentes temporales de la primera pluralidad de coeficientes armónicos esféricos y la segunda pluralidad de coeficientes armónicos esféricos.

[0567] Las Figs. 40A-40J son cada una un diagrama de bloques que ilustra ejemplos de dispositivos de codificación de audio 510A-510J que pueden realizar diversos aspectos de las técnicas descritas en esta descripción para comprimir coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales. En los ejemplos de las Fig. 40A a 40J, los dispositivos de codificación de audio 510A y 510B representan, en algunos ejemplos, cualquier dispositivo capaz de codificar datos de audio, tales como un ordenador de sobremesa, un ordenador portátil, una estación de trabajo, una tableta o un ordenador de pizarra, un dispositivo de grabación de audio dedicado, un teléfono móvil (incluidos los llamados "teléfonos inteligentes"), un dispositivo de reproductor multimedia personal, un dispositivo de juego personal o cualquier otro tipo de dispositivo capaz de codificar datos de audio.

[0568] Aunque se muestra como un dispositivo único, es decir, los dispositivos 510A-510J en los ejemplos de las Figs. 40A a 40J, los diversos componentes o unidades a los que se hace referencia a continuación como incluidos en los dispositivos 510A-510J pueden en realidad formar dispositivos separados que son externos a los dispositivos 510A-510J. En otras palabras, aunque se describe en esta descripción como realizada por un solo dispositivo, es decir, los dispositivos 510A-510J en los ejemplos de las Figs. 40A a 40J, las técnicas pueden ser implementadas o realizadas de otro modo por un sistema que comprende múltiples dispositivos, donde cada uno de estos dispositivos

puede incluir cada uno uno o más de los diversos componentes o unidades descritos con más detalle a continuación. En consecuencia, las técnicas no deben limitarse a los ejemplos de la Fig. 40A - 40J.

[0569] En algunos ejemplos, los dispositivos de codificación de audio 510A-510J representan dispositivos de codificación de audio alternativos a los descritos anteriormente con respecto a los ejemplos de las Figs. 3 y 4. A lo largo de la siguiente descripción de dispositivos de codificación de audio 510A-510J se observan diversas similitudes en términos de funcionamiento con respecto a las diversas unidades 30-52 del dispositivo de codificación de audio 20 descrito anteriormente con respecto a la Fig. 4. En muchos aspectos, los dispositivos de codificación de audio 510A-510J pueden, como se describe más adelante, operar de una manera sustancialmente similar al dispositivo de codificación de audio 20 aunque con ligeras derivaciones o modificaciones.

[0570] Como se muestra en el ejemplo de la Fig. 40A, el dispositivo de codificación de audio 510A comprende una unidad de compresión de audio 512, una unidad de codificación de audio 514 y una unidad de generación del flujo de bits 516. La unidad de compresión de audio 512 puede representar una unidad que comprime los coeficientes armónicos esféricos (SHC) 511 ("SHC 511"), que también se pueden designar como coeficientes ambisónicos de orden superior (HOA) 511. La unidad de compresión de audio 512 puede en algunos casos, la unidad de compresión de audio 512 representar una unidad que puede comprimir sin pérdidas o realizar compresión con pérdida con respecto a los SHC 511. Los SHC 511 pueden representar una pluralidad de SHC, donde al menos uno de la pluralidad de SHC corresponde a una función de base esférica que tiene un orden mayor que uno (donde los SHC de esta variedad se denominan ambisónicos de orden superior (HOA) para distinguirse de los ambisónicos de orden inferior de los cuales un ejemplo es el llamado "formato B"), como se describe con más detalle anteriormente. Si bien la unidad de compresión de audio 512 puede comprimir los SHC 511 sin pérdidas, en algunos ejemplos, la unidad de compresión de audio 512 elimina las de los SHC 511 que no son relevantes o relevantes para describir el campo sonoro cuando se reproducen (en que algunos no pueden ser escuchados por el sistema auditivo humano). En este sentido, la naturaleza con pérdidas de esta compresión no puede afectar excesivamente la calidad percibida del campo sonoro cuando se reproduce a partir de la versión comprimida de los SHC 511.

[0571] En el ejemplo de la Fig. 40A, la unidad de compresión de audio incluye una unidad de descomposición 518 y una unidad de extracción de componentes del campo sonoro 520. La unidad de descomposición 518 puede ser similar a la unidad de transformada reversible lineal 30 del dispositivo de codificación de audio 20. Es decir, la unidad de descomposición 518 puede representar una unidad configurada para realizar una forma de análisis denominada descomposición en valores singulares. Aunque se describe con respecto a la SVD, las técnicas pueden realizarse con respecto a cualquier transformación o descomposición similar que proporcione conjuntos de datos linealmente no correlacionados. También, la referencia a "conjuntos" en esta descripción pretende referirse a conjuntos "distintos de cero" a menos que se indique específicamente lo contrario y no se pretende referirse a la definición matemática clásica de conjuntos que incluye el denominado "conjunto vacío."

[0572] En cualquier caso, la unidad de descomposición 518 realiza una descomposición en valores singulares (que, de nuevo, se puede designar por su inicialismo "SVD") para transformar los coeficientes armónicos esféricos 511 en dos o más conjuntos de coeficientes armónicos esféricos transformados. En el ejemplo de la Fig. 40, la unidad de descomposición 518 puede realizar la SVD con respecto a los SHC 511 para generar una matriz V llamada 519, una matriz S 519B y una matriz U 519C. En el ejemplo de la Fig. 40, la unidad de descomposición 518 emite cada una de las matrices por separado en lugar de emitir los vectores US[*k*] en forma combinada como se ha tratado anteriormente con respecto a la unidad de transformada lineal reversible 30.

[0573] Como se ha indicado anteriormente, la matriz V* en la expresión matemática SVD mencionada anteriormente se designa como transposición 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, el conjugado complejo de la matriz V (o, en otras palabras, la matriz V*) puede considerarse igual a la matriz V. A continuación se supone, para facilitar la ilustración, que los SHC 511 comprenden números reales con el resultado de que la matriz V se emite a través de SVD en lugar de la matriz V*. Aunque se supone que es la matriz V, las técnicas pueden aplicarse de una manera similar a los SHC 511 que tienen coeficientes complejos, donde la salida de la SVD es la matriz V*. Por consiguiente, las técnicas no deberían limitarse a este respecto a proporcionar solamente la aplicación de SVD para generar una matriz V, sino que pueden incluir la aplicación de SVD a los SHC 511 que tienen componentes complejos para generar una matriz V*.

[0574] En cualquier caso, la unidad de descomposición 518 puede realizar una forma en bloque de SVD con respecto a cada bloque (que puede referirse a una trama) de datos de audio ambisónicos (HOA) de orden superior (donde estos datos de audio ambisónicos incluyen bloques o muestras de los SHC 511 o cualquier otra forma de datos de audio multicanal). Se puede utilizar 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. La unidad de descomposición 518 puede realizar por lo tanto una SVD en bloque con respecto a un bloque, los SHC 511 que tiene M-by-$(N+1)^2$ SHC, donde N, nuevamente, indica el orden de los datos de audio HOA. La unidad de descomposición 518 puede generar, mediante la realización de esta SVD, la matriz V 519, la matriz S 519B y la matriz U 519C, donde cada una de las matrices 519-519C ("matrices 519") puede representar las respectivas matrices V, S y U descritas en más detalle arriba. La unidad de descomposición 518 puede pasar o emitir estas

matrices 519A a la unidad de extracción de componentes del campo sonoro 520. La matriz V 519A puede ser de tamaño $(N+1)^2$ por $(N+1)^2$, la matriz S 519B puede ser de tamaño $(N+1)^2$ por $(N+1)^2$ y la matriz U puede ser de tamaño M por $(N+1)^2$, donde M se refiere al número de muestras en una trama de audio. Un valor típico para M es 1024, aunque las técnicas de esta descripción no deben limitarse a este valor típico para M.

[0575] La unidad de extracción de componentes del campo sonoro 520 puede representar una unidad configurada para determinar y luego extraer componentes distintos del campo sonoro y componentes de fondo del campo sonoro, separando eficazmente los distintos componentes del campo sonoro de los componentes de fondo del campo sonoro. A este respecto, la unidad de extracción de componentes del campo sonoro 520 puede realizar muchas de las operaciones descritas anteriormente con respecto a la unidad de análisis del campo sonoro 44, la unidad de selección de fondo 48 y la unidad de selección del primer plano 36 del dispositivo de codificación de audio 20 que se muestra en el ejemplo de la Fig. 4. Dado que los distintos componentes del campo sonoro, en algunos ejemplos, requieren funciones de orden superior (relativas a los componentes de fondo del campo sonoro) (y por lo tanto más SHC) para representar con precisión la naturaleza distinta de estos componentes, separar los distintos componentes de los componentes de fondo puede permitir que se asignen más bits a los componentes distintos y se asignen menos bits (relativamente, hablando) a los componentes de fondo. Por consiguiente, a través de la aplicación de esta transformación (en forma de SVD o cualquier otra forma de transformación, incluyendo PCA), las técnicas descritas en esta descripción pueden facilitar la asignación de bits a varios SHC y, por lo tanto, la compresión de los SHC 511.

[0576] Además, las técnicas también pueden permitir, como se describe con más detalle a continuación con respecto a la Fig. 40B, la reducción del orden de los componentes de fondo del campo sonoro dado que las funciones de base de orden superior no son, en algunos ejemplos, necesarias para representar estas partes de fondo del campo sonoro dada la naturaleza difusa o de fondo de estos componentes. Por lo tanto, las técnicas pueden permitir la compresión de aspectos difusos o de fondo del campo sonoro mientras que se preservan los componentes o aspectos distintos relevantes del campo sonoro mediante la aplicación de SVD a los SHC 511.

[0577] Como se muestra adicionalmente en el ejemplo de la Fig. 40, la unidad de extracción de componentes del campo sonoro 520 incluye una unidad de transposición 522, una unidad de análisis de componentes relevantes 524 y una unidad de matemáticas 526. La unidad de transposición 522 representa una unidad configurada para transponer la matriz V 519A para generar una transposición de la matriz V 519, que se designa como "matriz V^T 523". La unidad de transposición 522 puede emitir esta matriz V^T 523 a la unidad matemática 526. La matriz V^T 523 puede ser de tamaño $(N+1)^2$ por $(N+1)^2$.

[0578] La unidad de análisis de componentes relevantes 524 representa una unidad configurada para realizar un análisis de relevancia con respecto a la matriz S 519B. La unidad de análisis de componentes relevantes 524 puede, a este respecto, realizar operaciones similares a las descritas anteriormente con respecto a la unidad de análisis del campo sonoro 44 del dispositivo de codificación de audio 20 que se muestra en el ejemplo de la Fig. 4. La unidad de análisis de componentes relevantes 524 puede analizar los valores diagonales de la matriz S 519B, seleccionando un número variable D de estos componentes que tienen el valor más alto. En otras palabras, la unidad de análisis de componentes relevantes 524 puede determinar el valor D, que separa los dos subespacios (por ejemplo, el subespacio de primer plano o predominante y el subespacio de fondo o ambiente), analizando la pendiente de la curva creada por los valores diagonales descendentes de S, donde los grandes valores singulares representan sonidos de primer plano o distintos y los valores de bajo singular representan componentes de fondo del campo sonoro. En algunos ejemplos, la unidad de análisis de componentes relevantes 524 puede utilizar una primera y una segunda derivada de la curva de valor singular. La unidad de análisis de componentes relevantes 524 puede limitar también el número D a estar entre uno y cinco. Como otro ejemplo, la unidad de análisis de componentes relevantes 524 puede limitar el número D a estar entre uno y $(N+1)^2$. Alternativamente, la unidad de análisis de componentes relevantes 524 puede predefinir el número D, tal como a un valor de cuatro. En cualquier caso, una vez que se ha estimado el número D, la unidad 24 de análisis de componentes relevantes extrae el subespacio de primer plano y fondo de las matrices U, V y S.

[0579] En algunos ejemplos, la unidad de análisis de componentes relevantes 524 puede realizar este análisis cada M muestras, las cuales pueden ser reestructuradas como una base de trama por trama. A este respecto, D puede variar de trama a trama. En otros ejemplos, la unidad de análisis de componentes relevantes 24 puede realizar este análisis más de una vez por trama, analizando dos o más partes de la trama. Por consiguiente, las técnicas no deben estar limitadas a este respecto a los ejemplos descritos en esta descripción.

[0580] En efecto, la unidad de análisis de componentes relevantes 524 puede analizar los valores singulares de la matriz diagonal, que se designa como matriz S 519B en el ejemplo de la Fig. 40, identificando aquellos valores que tienen un valor relativo mayor que los otros valores de la matriz diagonal S 519B. La unidad de análisis de componentes relevantes 524 puede identificar valores D, extrayendo estos valores para generar la matriz S_{DIST} 525A y la matriz S_{BG} 525B. La matriz S_{DIST} 525A puede representar una matriz diagonal que comprende columnas D que tienen $(N+1)^2$ de la matriz S original 519B. En algunos casos, la matriz S_{BG} 525B puede representar una matriz que tiene $(N+1)^2$-D columnas, cada una de las cuales incluye $(N+1)^2$ coeficientes armónicos esféricos transformados de la matriz S original 519B. Aunque se describe como una matriz S_{DIST} que representa una matriz que comprende

columnas D que tienen valores $(N+1)^2$ de la matriz S original 519B, la unidad de análisis de componentes relevantes 524 puede truncar esta matriz para generar una matriz S_{DIST} que tiene D columnas que tienen D valores de la matriz S original 519B, dado que la matriz S 519B es una matriz diagonal y los valores $(N+1)^2$ de las D columnas después del D^{avo} en cada columna es a menudo un valor de cero. Aunque se describe con respecto a una matriz S_{DIST} completa 525A y una matriz S_{BG} completa 525B, las técnicas pueden implementarse con respecto a versiones truncadas de esta matriz S_{DIST} 525A y una versión truncada de esta matriz S_{BG}525B. Por consiguiente, las técnicas de esta descripción no deben estar limitadas a este respecto.

[0581] En otras palabras, la matriz S_{DIST} 525A puede tener un tamaño de D por $(N+1)^2$, mientras que la matriz S_{BG} 525B puede tener un tamaño de $(N+1)^2$-D por $(N+1)^2$. La matriz S_{DIST} 525A puede incluir aquellos componentes principales o, en otras palabras, valores singulares que se determinan como relevantes en términos de componentes de audio distintos (DIST) del campo sonoro, mientras que la matriz S_{BG} 525B puede incluir aquellos valores singulares que se determinan como fondo ($_{BG}$) o, en otras palabras, componentes de audio ambiental o no distinto del campo sonoro. Aunque 525A y 525B se muestran como matrices separadas en el ejemplo de la Fig. 40, las matrices 525A y 525B pueden especificarse como una matriz única utilizando la variable D para indicar el número de columnas (de izquierda a derecha) de esta única matriz que representa la matriz S_{DIST} 525. En algunos ejemplos, la variable D puede establecerse en cuatro.

[0582] La unidad de análisis de componentes relevantes 524 también puede analizar la matriz U 519C para generar la matriz U_{DIST} 525C y la matriz U_{BG} 525D. A menudo, la unidad de análisis de componentes relevantes 524 puede analizar la matriz S 519B para identificar la variable D, generando la matriz U_{DIST} 525C y la matriz U_{BG} 525B basada en la variable D. Es decir, después de identificar las columnas D de la matriz S 519B que son relevantes, la unidad de análisis de componentes relevantes 524 puede dividir la matriz U 519C en base a esta variable D determinada. En este caso, la unidad de análisis de componentes relevantes 524 puede generar la matriz U_{DIST} 525C para incluir las D columnas (de izquierda a derecha) de los $(N+1)^2$ coeficientes armónicos esféricos transformados de la matriz U original 519C y la matriz U_{BG} 525D para incluir las columnas restantes $(N+1)^2$-D de los $(N+1)^2$ coeficientes armónicos esféricos transformados de la matriz U original 519C. La matriz U_{DIST} 525C puede tener un tamaño de M por D, mientras que la matriz U_{BG} 525D puede tener un tamaño de M por $(N+1)^2$-D. Aunque 525C y 525D se muestran como matrices separadas en el ejemplo de la Fig. 40, las matrices 525C y 525D pueden especificarse como una matriz única utilizando la variable D para indicar el número de columnas (de izquierda a derecha) de esta única matriz que representa la matriz U_{DIST} 525B.

[0583] La unidad de análisis de componentes relevantes 524 también puede analizar la matriz V^T 523 para generar la matriz V^T_{DIST} 525E y la matriz V^T_{BG} 525F. A menudo, la unidad de análisis de componentes relevantes 524 puede analizar la matriz S 519B para identificar la variable D, generando la matriz V^T_{DIST} 525E y la matriz V_{BG} 525F basada en la variable D. Es decir, después de identificar las columnas D de la matriz S 519B que son relevantes, la unidad de análisis de componentes relevantes 254 puede dividir la matriz V 519A en base a esta variable determinada D. En este caso, la unidad de análisis de componentes relevantes 524 puede generar la matriz V^T_{DIST} 525E para incluir las $(N+1)^2$ filas (de arriba a abajo) de los D valores de la matriz original V^T 523 y de la matriz V^T_{BG} 525F para incluir las restantes $(N+1)^2$ filas de los $(N+1)^2$-D valores de la matriz V^T original 523. La matriz V^T_{DIST} 525E puede tener un tamaño de $(N+1)^2$ por D, mientras que la matriz V^T_{BG} 525D puede tener un tamaño de $(N+1)^2$ por $(N+1)^2$-D. Aunque se muestran como matrices separadas 525E y 525F en el ejemplo de la Fig. 40, las matrices 525E y 525F pueden especificarse como una matriz única utilizando la variable D para indicar el número de columnas (de izquierda a derecha) de esta única matriz que representa la matriz V_{DIST} 525E. La unidad de análisis de componentes relevantes 524 puede emitir la matriz S_{DIST} 525, la matriz S_{BG} 525B, la matriz U_{DIST} 525C, la matriz U_{BG} 525D y la matriz V^T_{BG} 525F en la unidad matemática 526, al mismo tiempo que emite la matriz V^T_{DIST} 525E a la unidad de generación del flujo de bits 516.

[0584] La unidad matemática 526 puede representar una unidad configurada para realizar multiplicaciones matriciales o cualquier otra operación matemática capaz de realizarse con respecto a una o más matrices (o vectores). Más específicamente, como se muestra en el ejemplo de la Fig. 40, la unidad matemática 526 puede representar una unidad configurada para realizar una multiplicación matricial para multiplicar la matriz U_{DIST} 525C por la matriz S_{DIST} 525A para generar los vectores $U_{DIST} * S_{DIST}$ 527 de tamaño M por D. La unidad matemática matricial 526 también puede representar una unidad configurada para realizar una multiplicación matricial para multiplicar la matriz U_{BG} 525D por la matriz S_{BG} 525B y luego por la matriz V^T_{BG} 525F para generar la matriz $U_{BG} * S_{BG} * V^T_{BG}$ 525F para generar coeficientes armónicos esféricos de fondo 531 de tamaño M por $(N+1)^2$ (que pueden representar los de los coeficientes armónicos esféricos 511 representativos de los componentes de fondo del campo sonoro). La unidad matemática 526 puede emitir los vectores $U_{DIST} * S_{DIST}$ 527 y los coeficientes armónicos esféricos de fondo 531 a la unidad de codificación de audio 514.

[0585] El dispositivo de codificación de audio 510 difiere por lo tanto del dispositivo de codificación de audio 20 en que el dispositivo de codificación de audio 510 incluye esta unidad de matemáticas 526 configurada para generar los vectores $U_{DIST} * S_{DIST}$ 527 y los coeficientes armónicos esféricos de fondo 531 a través de la multiplicación matricial al final del proceso de codificación. La unidad de transformación lineal reversible 30 del dispositivo de codificación de audio 20 realiza la multiplicación de las matrices U y S para emitir los vectores US[*k*] 33 en el comienzo relativo del proceso de codificación, lo que puede facilitar las operaciones posteriores, tales como la reordenación, que no que

se muestra en el ejemplo de la Fig. 40. Además, el dispositivo de codificación de audio 20, en lugar de recuperar los SHC de fondo 531 al final del proceso de codificación, selecciona los coeficientes HOA de fondo 47 directamente a partir de los coeficientes HOA 11, potencialmente evitando de este modo multiplicaciones matriciales para recuperar los SHC de fondo 531.

[0586] La unidad de codificación de audio 514 puede representar una unidad que realiza una forma de codificación para comprimir adicionalmente los vectores U_{DIST} * S_{DIST} 527 y los coeficientes armónicos esféricos de fondo 531. La unidad de codificación de audio 514 puede operar de una manera sustancialmente similar a la unidad codificadora de audio psicoacústico 40 del dispositivo de codificación de audio 20 que se muestra en el ejemplo de la Fig. 4. En algunos casos, esta unidad de codificación de audio 514 puede representar una o más instancias de una unidad codificadora de codificación de audio avanzada (AAC). La unidad de codificación de audio 514 puede codificar cada columna o fila de los vectores U_{DIST} * S_{DIST} 527. A menudo, la unidad de codificación de audio 514 puede invocar una instancia de una unidad de codificación AAC para cada una de las combinaciones de orden/suborden que permanecen en los coeficientes armónicos esféricos de fondo 531. Se puede encontrar más información sobre cómo se pueden codificar los coeficientes armónicos esféricos de fondo 531 utilizando una unidad de codificación AAC en un documento de convención de Eric Hellerud y col., titulado "Encoding Higher Order Ambisonics with AAC", presentado en la 124ª Convención de 2008 el 17-20 de mayo y disponible en: http://ro.uow.edu.au/cgi/viewcontent.cgi?article=8025&context=engpapers. La unidad de codificación de audio 14 puede emitir una versión codificada de los vectores U_{DIST} *S_{DIST} 527 (denominados "vectores U_{DIST}* S_{DIST} codificados 515") y una versión codificada de los coeficientes armónicos esféricos de fondo 531 (denominados "coeficientes armónicos esféricos de fondo codificados 515B ") a la unidad de generación del flujo de bits 516. En algunos casos, la unidad de codificación de audio 514 puede codificar en audio los coeficientes armónicos esféricos de fondo 531 utilizando una velocidad de bits objetivo inferior a la utilizada para codificar los vectores U_{DIST} * S_{DIST} 527, de este modo comprimiendo potencialmente más los coeficientes armónicos esféricos de fondo 531 en comparación con los vectores U_{DIST} *S_{DIST} 527.

[0587] La unidad de generación del flujo de bits 516 representa una unidad que formatea datos para ajustarse a un formato conocido (que puede referirse a un formato conocido por un dispositivo de decodificación), generando de este modo el flujo de bits 517. La unidad de generación del flujo de bits 42 puede operar de una manera sustancialmente similar a la descrita anteriormente con respecto a la unidad de generación del flujo de bits 42 del dispositivo de codificación de audio 24 que se muestra en el ejemplo de la Fig. 4. La unidad de generación del flujo de bits 516 puede incluir un multiplexor que multiplexa los vectores codificados U_{DIST} * S_{DIST} 515, los coeficientes armónicos esféricos de fondo codificados 515B y la matriz V^T_{DIST} 525E.

[0588] La Fig. 40B es un diagrama de bloques que ilustra un ejemplo de dispositivo de codificación de audio 510B que puede realizar diversos aspectos de las técnicas descritas en esta descripción para comprimir los coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales. El dispositivo de codificación de audio 510B puede ser similar al dispositivo de codificación de audio 510 en que el dispositivo de codificación de audio 510B incluye una unidad de compresión de audio 512, una unidad de codificación de audio 514 y una unidad de generación del flujo de bits 516. Además, la unidad de compresión de audio 512 del dispositivo de codificación de audio 510B puede ser similar a la del dispositivo de codificación de audio 510 en que la unidad de compresión de audio 512 incluye una unidad de descomposición 518. La unidad de compresión de audio 512 del dispositivo de codificación de audio 510B puede diferir de la unidad de compresión de audio 512 del dispositivo de codificación de audio 510 porque la unidad de extracción de los componentes del campo sonoro 520 incluye una unidad adicional designada como unidad de reducción del orden 528A "). Por esta razón, la unidad de extracción de componentes del campo sonoro 520 del dispositivo de codificación de audio 510B se designa como "unidad de extracción de componentes del campo sonoro 520B".

[0589] La unidad de reducción del orden 528A representa una unidad configurada para realizar una reducción del orden adicional de los coeficientes armónicos esféricos de fondo 531. En algunos casos, la unidad de reducción del orden 528A puede girar el campo sonoro representado los coeficientes armónicos esféricos de fondo 531 para reducir el número de los coeficientes armónicos esféricos de fondo 531 necesarios para representar el campo sonoro. En algunos casos, dado que los coeficientes armónicos esféricos de fondo 531 representan componentes de fondo del campo sonoro, la unidad de reducción del orden 528A puede borrar, eliminar o suprimir de otro modo (a menudo por reducción a cero) los de los coeficientes armónicos esféricos de fondo correspondientes a las funciones de base esférica y orden superior. A este respecto, la unidad de reducción del orden 528A puede realizar operaciones similares a la unidad de selección de fondo 48 del dispositivo de codificación de audio 20 que se muestra en el ejemplo de la Fig. 4. La unidad de reducción del orden 528A puede emitir una versión reducida de los coeficientes armónicos esféricos de fondo 531 (designados como "coeficientes armónicos esféricos de fondo reducidos 529") a la unidad de codificación de audio 514, que pueden realizar la codificación en audio de la manera descrita anteriormente para codificar los coeficientes armónicos esféricos de fondo reducidos 529 y, de este modo, generar los coeficientes armónicos esféricos de fondo reducidos codificados 515B.

[0590] Las diversas cláusulas enumeradas a continuación pueden presentar diversos aspectos de las técnicas descritas en esta descripción.

Cláusula 132567-1. Un dispositivo tal como el dispositivo de codificación de audio 510 o el dispositivo de codificación de audio 510B, que comprende: uno o más procesadores configurados para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de los vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y representar la pluralidad de coeficientes armónicos esféricos como una función de al menos una parte de, una o más, de la matriz U, la matriz S y la matriz V.

Cláusula 132567-2. El dispositivo de la cláusula 132567-1, en el que el uno o más procesadores están configurados además para generar un flujo de bits para incluir la representación de la pluralidad de coeficientes armónicos esféricos como uno o más vectores de la matriz U, la matriz S y la matriz V incluyendo combinaciones de las mismas o derivadas de las mismas.

Cláusula 132567-3. El dispositivo de la cláusula 132567-1, en el que el uno o más procesadores están configurados adicionalmente para, al representar la pluralidad de coeficientes armónicos esféricos, determinar uno o más vectores U_{DIST} incluidos dentro de la matriz U que describen componentes distintos del campo sonoro.

Cláusula 132567-4. El dispositivo de la cláusula 132567-1, en el que el uno o más procesadores están configurados adicionalmente para, al representar la pluralidad de coeficientes armónicos esféricos, determinar uno o más vectores U_{DIST} incluidos dentro de la matriz U que describen componentes distintos del campo sonoro, determinar uno o más vectores S_{DIST} incluidos dentro de la matriz S que también describen los distintos componentes del campo sonoro y multiplican uno o más vectores U_{DIST} y uno o más uno o más vectores S_{DIST} para generar vectores $U_{DIST} * S_{DIST}$.

Cláusula 132567-5. El dispositivo de la cláusula 132567-1, en el que el uno o más procesadores están configurados adicionalmente para, al representar la pluralidad de coeficientes armónicos esféricos, determinar uno o más vectores U_{DIST}, incluidos dentro de la matriz U que describen componentes distintos del campo sonoro, determinar uno o más vectores S_{DIST} incluidos dentro de la matriz S que también describen los distintos componentes del campo sonoro y multiplicar uno o más vectores U_{DIST} y uno o más uno o más vectores S_{DIST} para generar uno o más $U_{DIST} * S_{DIST}$, y en el que el uno o más procesadores están configurados adicionalmente para codificar en audio uno o más vectores $U_{DIST} * S_{DIST}$ para generar una versión codificada en audio de uno o más vectores $U_{DIST}* S_{DIST}$.

Cláusula 132567-6. El dispositivo de la cláusula 132567-1, en el que el uno o más procesadores están configurados adicionalmente para, al representar la pluralidad de coeficientes armónicos esféricos, determinar uno o más vectores U_{BG} incluidos dentro de la matriz U.

Cláusula 132567-7. El dispositivo de la cláusula 132567-1, en el que el uno o más procesadores están configurados adicionalmente para, al representar la pluralidad de coeficientes armónicos esféricos, analizar la matriz S para identificar componentes distintos y de fondo del campo sonoro.

Cláusula 132567-8. El dispositivo de la cláusula 132567-1, en el que el uno o más procesadores están configurados adicionalmente para, al representar la pluralidad de coeficientes armónicos esféricos, analizar la matriz S para identificar componentes distintos y de fondo del campo sonoro y determinar, basándose en el análisis de la matriz S, uno o más vectores U_{DIST} de la matriz U que describen componentes distintos del campo sonoro y uno o más vectores U_{BG} de la matriz U que describen componentes de fondo del campo sonoro.

Cláusula 132567-9. El dispositivo de la cláusula 132567-1, en el que el uno o más procesadores están configurados adicionalmente para, al representar la pluralidad de coeficientes armónicos esféricos, analizar la matriz S para identificar componentes distintos y de fondo del campo sonoro en una trama de audio por trama de audio, y determinar basándose en el análisis de trama de audio por trama de audio de la matriz S, uno o más vectores U_{DIST} de la matriz U que describen componentes distintos del campo sonoro y uno o más vectores U_{BG} de la matriz U que describen los componentes de fondo del campo sonoro.

Cláusula 132567-10. El dispositivo de la cláusula 132567-1, en el que el uno o más procesadores están configurados adicionalmente para, al representar la pluralidad de coeficientes armónicos esféricos, analizar la matriz S para identificar componentes distintos y de fondo del campo sonoro, determinar, basándose en el análisis de la matriz S, uno o más vectores U_{DIST} de la matriz U que describen componentes distintos del campo sonoro y uno o más vectores U_{BG} de la matriz U que describen componentes de fondo del campo sonoro, que determinan, a partir del análisis de la matriz S, uno o más vectores S_{DIST} y uno o más vectores S_{BG} de la matriz S correspondientes a uno o más vectores U_{DIST} y uno o más vectores U_{BG} y determinar, a partir del análisis de la matriz S, uno o más vectores V^T_{DIST} y uno o más vectores V^T_{BG} de una transposición de la matriz V correspondiente a uno o más vectores U_{DIST} y uno o más vectores U_{BG}.

Cláusula 132567-11. El dispositivo de la cláusula 132567-10, en el que el uno o más procesadores están configurados además para, al representar la pluralidad de coeficientes armónicos esféricos, multiplicar uno o más vectores U_{BG} por uno o más vectores S_{BG} y luego por uno o más vectores V^T_{BG} para generar uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$, y en el que uno o más procesadores están configurados adicionalmente para codificar en audio los vectores $U_{BG} * S_{BG} * V^T_{BG}$ para generar una versión codificada en audio de los vectores $U_{BG} * S_{BG} * V^T_{BG}$.

Cláusula 132567-12. El dispositivo de la cláusula 132567-10, en el que el uno o más procesadores están configurados además para, al representar la pluralidad de coeficientes armónicos esféricos, multiplicar uno o más vectores U_{BG} por uno o más vectores S_{BG} y luego por uno o más vectores V^T_{BG} para generar uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$, y realizar un proceso de reducción del orden para eliminar aquellos de los coeficientes de uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ asociados con uno o más órdenes de funciones de base armónica esférica y, de este modo, generar una versión de orden reducido de uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$.

Cláusula 132567-13. El dispositivo de la cláusula 132567-10, en el que el uno o más procesadores están configurados además para, al representar la pluralidad de coeficientes armónicos esféricos, multiplicar uno o más vectores U_{BG} por uno o más vectores S_{BG} y luego por uno o más vectores V^T_{BG} para generar uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$, y realizar un proceso de reducción del orden para eliminar aquellos de los coeficientes de uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ asociados con uno o más órdenes de funciones de base armónica y esférica y, de este modo, generar una versión ordenada reducida de uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$, y en el que uno o más procesadores están configurados adicionalmente para codificar en audio la versión reducida en orden de uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ para generar una versión codificada en audio de orden reducido de uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$.

Cláusula 132567-14. El dispositivo de la cláusula 132567-10, en el que el uno o más procesadores están configurados además para, al representar la pluralidad de coeficientes armónicos esféricos, multiplicar uno o más vectores U_{BG} por uno o más vectores S_{BG} y luego por uno o más vectores V^T_{BG} para generar uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$, realizar un proceso de reducción del orden para eliminar aquellos de los coeficientes de uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ asociados con uno o más órdenes mayores que una de las funciones de base armónicas esféricas y, de este modo, generar una versión de orden reducido de uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$, y codificar en audio la versión de orden reducido de uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ para generar una versión codificada en audio de uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ de orden reducido.

Cláusula 132567-15. El dispositivo de la cláusula 132567-10, en el que el uno o más procesadores están configurados además para generar un flujo de bits para incluir uno o más vectores V^T_{DIST}.

Cláusula 132567-16. El dispositivo de la cláusula 132567-10, en el que el uno o más procesadores están configurados además para generar un flujo de bits para incluir uno o más vectores V^T_{DIST} sin codificación de audio de uno o más vectores V^T_{DIST}.

Cláusula 132567-1F. Un dispositivo, tal como el dispositivo de codificación de audio 510 o 510B, que comprende uno o más procesadores para 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 representar los datos de audio multicanal en función de, al menos una parte, de una o más de la matriz U, la matriz S y la matriz V.

Cláusula 132567-2F. El dispositivo de la cláusula 132567-1F, en el que los datos de audio multicanal comprenden una pluralidad de coeficientes armónicos esféricos.

Cláusula 132567-3F. El dispositivo de la cláusula 132567-2F, en el que el uno o más procesadores están configurados adicionalmente tal como se indica mediante cualquier combinación de las cláusulas 132567-2 a 132567-16.

[0591] De cada una de las diversas cláusulas descritas anteriormente, debe entenderse que cualquiera de los dispositivos de codificación de audio 510A-510J puede realizar un procedimiento o comprender de otro modo medios para realizar cada paso del procedimiento para el cual está configurado el dispositivo de codificación de audio 510A-510J. En algunos casos, estos medios pueden comprender uno o más procesadores. En algunos casos, 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 el cual el dispositivo de codificación de audio 510A-510J se ha configurado para realizar.

[0592] Por ejemplo, una cláusula 132567-17 puede derivarse de la cláusula anterior 132567-1 como un procedimiento que comprende realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y que representan la pluralidad de coeficientes armónicos esféricos en función de al menos una porción de una o más de la matriz U, la matriz S y la matriz V.

[0593] Como otro ejemplo, una cláusula 132567-18 puede derivarse de la cláusula anterior 132567-1 que es un dispositivo, tal como el dispositivo de codificación de audio 510B, que comprende medios para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y medios para representar la pluralidad de coeficientes armónicos esféricos en función de al menos una porción de una o más de la matriz U, la matriz S y la matriz V.

[0594] Como otro ejemplo, una cláusula 132567-18 puede derivarse de la cláusula 132567-1 anterior como un medio de almacenamiento no transitorio legible por ordenador que tiene almacenadas en el mismo instrucciones que, cuando se ejecutan, hacen que uno o más procesadores realicen una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y representan la pluralidad de coeficientes armónicos esféricos en función de al menos una porción de una o más de la matriz U, la matriz S y la matriz V.

[0595] De igual modo, pueden derivarse diversas cláusulas de las cláusulas 132567-2 a 132567-16 para los diversos dispositivos, procedimientos y medios de almacenamiento no transitorios legibles por ordenador derivados como se ha ilustrado anteriormente. Lo mismo puede hacerse para las otras cláusulas enumeradas a lo largo de esta divulgación.

[0596] La FIG. 40C es un diagrama de bloques que dispositivos de codificación de audio ejemplares 510C que pueden realizar diversos aspectos de las técnicas descritas en esta divulgación para comprimir coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales. El dispositivo de codificación de audio 510C puede ser similar al dispositivo de codificación de audio 510B en que el dispositivo de codificación de audio 510C incluye una unidad de compresión de audio 512, una unidad de codificación de audio 514 y una unidad de generación de flujo de bits 516. Además, la unidad de compresión de audio 512 del dispositivo de codificación de audio 510C puede ser similar a la del dispositivo de codificación de audio 510B porque la unidad de compresión de audio 512 incluye una unidad de descomposición 518.

[0597] Sin embargo, la unidad de compresión de audio 512 del dispositivo de codificación de audio 510C puede diferir de la unidad de compresión de audio 512 del dispositivo de codificación de audio 510B porque la unidad de extracción de componente de campo sonoro 520 incluye una unidad adicional, indicada como unidad de reordenación de vectores 532. Por esta razón, la unidad de extracción de componentes del campo sonoro 520 del dispositivo de codificación de audio 510C se denomina "unidad de extracción de componentes del campo sonoro 520C".

[0598] La unidad de reordenación de vectores 532 puede representar una unidad configurada para reordenar los vectores U_{DIST} * S_{DIST} 527 para generar uno o más vectores U_{DIST} * S_{DIST} reordenados 533. A este respecto, la unidad de reordenación de vectores 532 puede operar de una manera similar a la descrita anteriormente con respecto a la unidad de reordenación 34 del dispositivo de codificación de audio 20 mostrado en el ejemplo de la FIG. 4. La unidad de extracción de componentes del campo sonoro 520C puede llamar a la unidad de reordenación de vectores 532 para reordenar los vectores U_{DIST} S_{DIST} 527 porque el orden de los vectores U_{DIST} * S_{DIST} 527 (donde cada vector de los vectores U_{DIST} * S_{DIST} 527 puede representar uno o más objetos monoaudio distintos presentes en el campo sonoro) puede variar desde partes de los datos de audio por la razón indicada anteriormente. Es decir, dado que la unidad de compresión de audio 512, en algunos ejemplos, opera en estas porciones de los datos de audio denominados generalmente tramas de audio (que pueden tener M muestras de los coeficientes armónicos esféricos 511, donde M es, en algunos ejemplos, ajustado a 1024), la posición de los vectores correspondientes a estos objetos monoaudio distintos, como se representan en la matriz U 519C a partir de la cual se derivan los vectores U_{DIST} * S_{DIST} 527, puede variar de trama de audio a trama de audio.

[0599] El paso de estos vectores U_{DIST} * S_{DIST} 527 directamente a la unidad de codificación de audio 514 sin reordenar estos vectores U_{DIST} * S_{DIST} 527 de trama de audio a trama de audio puede reducir el alcance de la

compresión alcanzable para algunos esquemas de compresión, tales como la compresión heredada que se realizan mejor cuando los objetos monoaudio están correlacionados (sentido del canal, que se define en este ejemplo por el orden de los vectores U_{DIST} * S_{DIST} 527 uno con relación al otro) a través de tramas de audio. Además, cuando no se reordena, la codificación de los vectores U_{DIST} * S_{DIST} 527 puede reducir la calidad de los datos de audio cuando se recuperan. Por ejemplo, los codificadores AAC que pueden representarse en el ejemplo de la FIG. 40C por la unidad de codificación de audio 514, pueden comprimir más eficientemente uno o más vectores U_{DIST} * S_{DIST} reordenados 533 de trama a trama en comparación con la compresión lograda al codificar directamente los vectores U_{DIST} * S_{DIST} 527 de trama a trama. Aunque se ha descrito anteriormente con respecto a los codificadores AAC, las técnicas pueden realizarse con respecto a cualquier codificador que proporcione mejor compresión cuando se especifican objetos monoaudio a través de tramas en un orden o posición específicos (sentido del canal).

[0600] Como se describe con más detalle a continuación, las técnicas pueden permitir que el dispositivo de codificación de audio 510C reordene uno o más vectores (es decir, los vectores U_{DIST} * S_{DIST} 527 para generar uno o más vectores reordenados U_{DIST} * S_{DIST} 533 y así facilitar la compresión de los vectores U_{DIST} * S_{DIST} 527 por un codificador de audio heredado, tal como la unidad de codificación de audio 514. El dispositivo de codificación de audio 510C puede realizar adicionalmente las técnicas descritas en esta divulgación para codificar audio del uno o más vectores U_{DIST} * S_{DIST} reordenados 533 usando la unidad de codificación de audio 514 para generar una versión codificada 515A del uno o más vectores U_{DIST} * S_{DIST} reordenados 533.

[0601] Por ejemplo, la unidad de extracción de componentes del campo sonoro 520C puede recurrir a la unidad de reordenación de vectores 532 para reordenar uno o más primeros vectores U_{DIST} * S_{DIST} 527 de una primera trama de audio posterior en el tiempo con respecto a la segunda trama a la que corresponden uno o más segundos vectores U_{DIST} * S_{DIST} 527. Aunque se describe en el contexto de una primera trama de audio que es posterior en el tiempo a la segunda trama de audio, la primera trama de audio puede preceder en el tiempo a la segunda trama de audio. Por consiguiente, las técnicas no deben limitarse al ejemplo descrito en esta divulgación.

[0602] La unidad de reordenación de vectores 532 puede realizar primero un análisis de energía con respecto a cada uno de los primeros vectores U_{DIST} * S_{DIST} 527 y los segundos vectores U_{DIST} * S_{DIST} 527, calculando una energía cuadrática media para al menos una porción de (pero a menudo la totalidad de) la primera trama de audio y una porción de (pero a menudo la totalidad de) una segunda trama de audio y generar así (suponiendo que D sea cuatro) ocho energías, una para cada uno de los primeros vectores U_{DIST} * S_{DIST} 527 de la primera trama de audio y uno para cada uno de los segundos vectores U_{DIST} * S_{DIST} 527 de la segunda trama de audio. La unidad de reordenación de vectores 532 puede entonces comparar cada energía de los primeros vectores U_{DIST} * S_{DIST} 527 de forma giratoria contra cada uno de los segundos vectores U_{DIST} * S_{DIST} 527 como se ha descrito anteriormente con respecto a las Tablas 1-4.

[0603] En otras palabras, cuando se usa descomposición SVD basada en tramas (o procedimientos relacionados tales como KLT y PCA) en señales HoA, puede no garantizarse que el ordenamiento de los vectores de trama a trama sea consistente. Por ejemplo, si hay dos objetos en el campo acústico subyacente, la descomposición (que, cuando se lleva a cabo correctamente, puede denominarse "descomposición ideal") puede dar como resultado la separación de los dos objetos de tal manera que un vector represente un objeto en la matriz U. Sin embargo, incluso cuando la descomposición puede representarse como una "descomposición ideal", los vectores pueden alternarse de posición en la matriz U (y correspondientemente en la matriz S y V) de trama a trama. Además, pueden existir también diferencias de fase, en las que la unidad de reordenación de vectores 532 puede invertir la fase usando inversión de fase (multiplicando por puntos cada elemento del vector invertido por menos o uno negativo). Para alimentar estos vectores, la trama por trama en el mismo "motor de codificación AAC/Audio" puede requerir que se identifique el orden (o, en otras palabras, que las señales se emparejen), que la fase se rectifique, y que se aplique una interpolación cuidadosa en los límites de trama. Sin esto, el códec de audio subyacente puede producir artefactos extremadamente discordantes incluyendo los conocidos como "difuminado temporal" o "pre-eco".

[0604] De acuerdo con diversos aspectos de las técnicas descritas en esta divulgación, el dispositivo de codificación de audio 510C puede aplicar múltiples metodologías para identificar/combinar vectores, utilizando energía y correlación cruzada en los límites de trama de los vectores. El dispositivo de codificación de audio 510C también puede garantizar que se corrija un cambio de fase de 180 grados -que a menudo aparece en los límites de trama-. La unidad de reordenación de vectores 532 puede aplicar una forma de ventana de interpolación de entrada en fundido/fundirse entre los vectores para asegurar una transición suave entre las tramas.

[0605] De esta manera, el dispositivo de codificación de audio 530C puede reordenar uno o más vectores para generar uno o más primeros vectores reordenados y facilitar así la codificación por un codificador de audio heredado, en el que uno o más vectores describen componentes distintos de un campo sonoro, y codificación de audio del uno o más vectores reordenados usando el codificador de audio heredado para generar una versión codificada de uno o más vectores reordenados.

[0606] Varios aspectos de las técnicas descritas en esta divulgación pueden permitir que el dispositivo de codificación de audio 510C funcione de acuerdo con las siguientes cláusulas.

Cláusula 133143-1A. Un dispositivo, tal como el dispositivo de codificación de audio 510C, que comprende: uno o más procesadores configurados para realizar una comparación de energía entre uno o más primeros vectores y uno o más segundos vectores para determinar uno o más primeros vectores reordenados y facilitar la extracción de uno o ambos del uno o más primeros vectores y el uno o más segundos vectores, en el que uno o más primeros vectores describen componentes distintos de un campo sonoro en una primera porción de datos de audio y el uno o más segundos vectores describen componentes distintos de un campo sonoro en una segunda porción de los datos de audio.

Cláusula 133143-2A. El dispositivo de la cláusula 133143-1A, en el que el uno o más primeros vectores no representan componentes de fondo del campo sonoro en la primera porción de los datos de audio, y en el que uno o más segundos vectores no representan componentes de fondo del campo sonoro en la segunda porción de los datos de audio.

Cláusula 133143-3A. El dispositivo de la cláusula 133143-1A, en el que el uno o más procesadores están configurados adicionalmente para, después de realizar la comparación de energía, realizar una correlación cruzada entre uno o más primeros vectores y uno o más segundos vectores para identificar el uno o más primeros vectores que se correlacionan con el uno o más segundos vectores.

Cláusula 133143-4A. El dispositivo de la cláusula 133143-1A, en el que el uno o más procesadores están configurados adicionalmente para descartar uno o más de los segundos vectores basados en la comparación de energía para generar uno o más segundos vectores reducidos que tienen menos vectores que el uno o más segundos vectores, realizar una correlación cruzada entre al menos uno del uno o más primeros vectores y el uno o más segundos vectores reducidos para identificar uno del uno o más segundos vectores reducidos que se correlaciona con al menos uno del uno o más primeros vectores, y reordenar al menos uno del uno o más primeros vectores basándose en la correlación cruzada para generar el uno o más de los primeros vectores reordenados.

Cláusula 133143-5A. El dispositivo de la cláusula 133143-1A, en el que el uno o más procesadores están configurados adicionalmente para descartar uno o más de los segundos vectores basados en la comparación de energía para generar uno o más segundos vectores reducidos que tienen menos vectores que el uno o más segundos vectores, realizar una correlación cruzada entre al menos uno del uno o más primeros vectores y el uno o más segundos vectores reducidos para identificar uno del uno o más segundos vectores reducidos que se correlaciona con al menos uno del uno o más primeros vectores, reordenar al menos uno del uno o más primeros vectores basándose en la correlación cruzada para generar el uno o más primeros vectores reordenados, y codificar el uno o más primeros vectores reordenados para generar la versión codificada de audio del uno o más de los primeros vectores reordenados.

Cláusula 133143-6A. El dispositivo de la cláusula 133143-1A, en el que el uno o más procesadores están configurados adicionalmente para descartar uno o más de los segundos vectores basados en la comparación de energía para generar uno o más segundos vectores reducidos que tienen menos vectores que el uno o más segundos vectores, realizar una correlación cruzada entre al menos uno del uno o más primeros vectores y el uno o más segundos vectores reducidos para identificar uno del uno o más segundos vectores reducidos que se correlaciona con al menos uno del uno o más primeros vectores, reordenar al menos uno del uno o más primeros vectores basándose en la correlación cruzada para generar el uno o más primeros vectores reordenados, codificar el uno o más primeros vectores reordenados para generar la versión codificada de audio del uno o más primeros vectores reordenados, y generar y flujo de bits para incluir la versión codificada del uno o más primeros vectores reordenados.

Cláusula 133143-7A. El dispositivo de las reivindicaciones 3A-6A, en el que la primera porción de los datos de audio comprende una primera trama de audio que tiene M muestras, en el que la segunda porción de los datos de audio comprende una segunda trama de audio que tiene el mismo número, M, de muestras, en el que el uno o más procesadores están configurados adicionalmente para, al realizar la correlación cruzada, realizar la correlación cruzada con respecto a los últimos valores M-Z del al menos uno de uno o más primeros vectores y los primeros valores M-Z de cada uno de los uno o más segundos vectores reducidos para identificar uno de los uno o más segundos vectores reducidos que se correlacionan con al menos uno del uno o más primeros vectores, y en el que Z es menor que M.

Cláusula 133143-8A. El dispositivo de las reivindicaciones 3A-6A, en el que la primera porción de los datos de audio comprende una primera trama de audio que tiene M muestras, en el que la segunda porción de los datos de audio comprende una segunda trama de audio que tiene el mismo número, M, de muestras, en el que el uno o más procesadores están configurados adicionalmente para, al realizar la correlación cruzada, realizar la correlación cruzada con respecto a los últimos valores M-Y del al menos uno de uno o más primeros vectores y los primeros valores M-Z de cada uno de los uno o más segundos vectores reducidos para identificar uno de los uno o más segundos vectores reducidos que se correlacionan con al menos uno del uno o más primeros vectores, y en el que tanto Z como Y son menores que M.

Cláusula 133143-9A. El dispositivo de las reivindicaciones 3A-6A, en el que el uno o más procesadores están configurados adicionalmente para, al realizar la correlación cruzada, invertir al menos uno del uno o más primeros vectores y el uno o más segundos vectores. Cláusula 133143-10A. El dispositivo de la cláusula 133143-1A, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos del campo sonoro para generar el uno o más primeros vectores y el uno o más segundos vectores.

Cláusula 133143-11A. El dispositivo de la cláusula 133143-1A, en el que el uno o más procesadores están configurados adicionalmente para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos del campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y generar el uno o más primeros vectores y el uno o más de segundos vectores en función de una o más de la matriz U, la matriz S y la matriz V.

Cláusula 133143-12A. El dispositivo de la cláusula 133143-1A, en el que el uno o más procesadores están configurados adicionalmente para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos del campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, realizar un análisis de saliencia con respecto a la matriz S para identificar uno o más vectores UDIST de la matriz U y uno o más vectores SDIST de la matriz S, y determinar el uno o más primeros vectores y uno o más segundos vectores multiplicando, al menos en parte, el uno o más vectores UDIST por el uno o más vectores SDIST.

Cláusula 133143-13A. El dispositivo de la cláusula 133143-1A, en el que la primera parte de los datos de audio se produce en el tiempo antes de la segunda porción de los datos de audio.

Cláusula 133143-14A. El dispositivo de la cláusula 133143-1A, en el que la primera parte de los datos de audio se produce en el tiempo después de la segunda porción de los datos de audio.

Cláusula 133143-15A. El dispositivo de la cláusula 133143-1A, en el que el uno o más procesadores están configurados adicionalmente para, al realizar la comparación de energía, calcular una energía cuadrática media para cada uno del uno o más primeros vectores y el uno o más segundos vectores, y comparar la energía cuadrática media calculada para al menos uno del uno o más primeros vectores con respecto a la energía media cuadrática media calculada para cada uno del uno o más segundos vectores.

Cláusula 133143-16A. El dispositivo de la cláusula 133143-1A, en el que el uno o más procesadores están configurados además para reordenar al menos uno del uno o más primeros vectores basándose en la comparación de energía para generar el uno o más primeros vectores reordenados, y en el que el uno o más procesadores están configurados adicionalmente para, al reordenar los primeros vectores, aplicar una ventana de interpolación de entrar en fundido/fundido entre el uno o más primeros vectores para asegurar una transición suave cuando se generan los uno o más primeros vectores reordenados.

Cláusula 133143-17A. El dispositivo de la cláusula 133143-1A, en el que el uno o más procesadores están configurados además para reordenar uno o más primeros vectores basados en al menos la comparación de energía para generar el uno o más primeros vectores reordenados, generar un flujo de bits para incluir el uno o más primeros vectores reordenados o una versión codificada del uno o más primeros vectores reordenados, y especificar información de reordenación en el flujo de bits que describe cómo se reordenaron el uno o más primeros vectores.

Cláusula 133143-18A. El dispositivo de la cláusula 133143-1A, en el que la comparación de energía facilita la extracción del uno o ambos del uno o más primeros vectores y uno o más segundos vectores con el fin de promover la codificación de audio del uno o ambos del uno o más primeros vectores y el uno o más segundos vectores.

Cláusula 133143-1B. El dispositivo, tal como el dispositivo de codificación de audio 510C, que comprende: uno o más procesadores configurados para realizar una correlación cruzada con respecto a uno o más primeros vectores y uno o más segundos vectores para determinar uno o más primeros vectores reordenados y facilitar la extracción de uno o ambos del uno o más primeros vectores y el uno o más segundos vectores, en el que uno o más primeros vectores describen componentes distintos de un campo sonoro en una primera porción de datos de audio y el uno o más segundos vectores describen componentes distintos de un campo sonoro en una segunda porción de los datos de audio.

Cláusula 133143-2B. El dispositivo de la cláusula 133143-1B, en el que el uno o más primeros vectores no

representan componentes de fondo del campo sonoro en la primera porción de los datos de audio, y en el que uno o más segundos vectores no representan componentes de fondo del campo sonoro en la segunda porción de los datos de audio.

Cláusula 133143-3B. El dispositivo de la cláusula 133143-1B, en el que el uno o más procesadores están configurados adicionalmente para, antes de realizar la correlación cruzada, realizar una comparación de energía entre el uno o más primeros vectores y el uno o más segundos vectores para generar uno o más segundos vectores reducidos con menos vectores que el uno o más vectores, y en el que el uno o más procesadores están configurados adicionalmente para, al realizara la correlación cruzada, realizar la correlación cruzada entre el uno o más primeros vectores y uno o más segundos vectores reducidos para facilitar la codificación de audio de uno o ambos del uno o más primeros vectores y el uno o más segundos vectores.

Cláusula 133143-4B. El dispositivo de la cláusula 133143-3B, en el que el uno o más procesadores están configurados adicionalmente para, al realizar la comparación de energía, calcular una energía cuadrática media para cada uno del uno o más primeros vectores y el uno o más segundos vectores, y comparar la energía cuadrática media calculada para al menos uno del uno o más primeros vectores con respecto a la energía media cuadrática media calculada para cada uno del uno o más segundos vectores.

Cláusula 133143-5B. El dispositivo de la cláusula 133143-3B, en el que el uno o más procesadores están configurados adicionalmente para descartar uno o más de los segundos vectores basándose en la comparación de energía para generar uno o más segundos vectores reducidos con menos vectores que el uno o más segundos vectores, en el que el uno o más procesadores están configurados adicionalmente para, al realizara la correlación cruzada, realizar la correlación cruzada entre al menos uno del uno o más primeros vectores y el uno o más segundos vectores reducidos para identificar uno del uno o más segundos vectores reducidos que se correlacionan con el al menos uno del uno o más primeros vectores, y en el que el uno o más procesadores están configurados adicionalmente para reordenar al menos uno del uno o más primeros vectores basándose en la correlación cruzada para generar el uno o más primeros vectores reordenados.

Cláusula 133143-6B. El dispositivo de la cláusula 133143-3B, en el que el uno o más procesadores están configurados adicionalmente para descartar uno o más de los segundos vectores basándose en la comparación de energía para generar uno o más segundos vectores reducidos con menos vectores que el uno o más segundos vectores, en el que el uno o más procesadores están configurados adicionalmente para, al realizara la correlación cruzada, realizar la correlación cruzada entre al menos uno del uno o más primeros vectores y el uno o más segundos vectores reducidos para identificar uno del uno o más segundos vectores reducidos que se correlacionan con el al menos uno del uno o más primeros vectores, y en el que el uno o más procesadores están configurados adicionalmente para reordenar al menos uno del uno o más primeros vectores basándose en la correlación cruzada para generar el uno o más primeros vectores reordenados, y codificar el uno o más primeros vectores reordenados para generar la versión codificada de audio del uno o más primeros vectores reordenados.

Cláusula 133143-7B. El dispositivo de la cláusula 133143-3B, en el que el uno o más procesadores están configurados adicionalmente para descartar uno o más de los segundos vectores basándose en la comparación de energía para generar uno o más segundos vectores reducidos con menos vectores que el uno o más segundos vectores, en el que el uno o más procesadores están configurados adicionalmente para, al realizara la correlación cruzada, realizar la correlación cruzada entre al menos uno del uno o más primeros vectores y el uno o más segundos vectores reducidos para identificar uno del uno o más segundos vectores reducidos que se correlacionan con el al menos uno del uno o más primeros vectores, y en el que el uno o más procesadores están configurados adicionalmente para reordenar al menos uno del uno o más primeros vectores basándose en la correlación cruzada para generar el uno o más primeros vectores reordenados, codificar el uno o más primeros vectores reordenados para generar la versión codificada de audio del uno o más primeros vectores reordenados, y generar un flujo de bits para incluir la versión codificada del uno o más primeros vectores reordenados.

Cláusula 133143-8B. El dispositivo de las reivindicaciones 3B-7B, en el que la primera porción de los datos de audio comprende una primera trama de audio que tiene M muestras, en el que la segunda porción de los datos de audio comprende una segunda trama de audio que tiene el mismo número, M, de muestras, en el que el uno o más procesadores están configurados adicionalmente para, al realizar la correlación cruzada, realizar la correlación cruzada con respecto a los últimos valores M-Z del al menos uno de uno o más primeros vectores y los primeros valores M-Z de cada uno de los uno o más segundos vectores reducidos para identificar uno de los uno o más segundos vectores reducidos que se correlacionan con al menos uno del uno o más primeros vectores, y en el que Z es menor que M.

Cláusula 133143-9B. El dispositivo de las reivindicaciones 3B-7B, en el que la primera porción de los datos de audio comprende una primera trama de audio que tiene M muestras, en el que la segunda porción de los datos de audio comprende una segunda trama de audio que tiene el mismo número, M, de muestras, en el que el uno o más procesadores están configurados adicionalmente para, al realizar la correlación cruzada, realizar la correlación cruzada con respecto a los últimos valores M-Y del al menos uno de uno o más primeros vectores y

los primeros valores M-Z de cada uno de los uno o más segundos vectores reducidos para identificar uno de los uno o más segundos vectores reducidos que se correlacionan con al menos uno del uno o más primeros vectores, y en el que tanto Z como Y son menores que M.

Cláusula 133143-10B. El dispositivo de las reivindicaciones 1B, en el que el uno o más procesadores están configurados adicionalmente para, al realizar la correlación cruzada, invertir al menos uno del uno o más primeros vectores y el uno o más segundos vectores.

Cláusula 133143-11B. El dispositivo de la cláusula 133143-1B, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos del campo sonoro para generar el uno o más primeros vectores y el uno o más segundos vectores.

Cláusula 133143-12B. El dispositivo de la cláusula 133143-1B, en el que el uno o más procesadores están configurados adicionalmente para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos del campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y generar el uno o más primeros vectores y el uno o más de segundos vectores en función de una o más de la matriz U, la matriz S y la matriz V.

Cláusula 133143-13B. El dispositivo de la cláusula 133143-1B, en el que el uno o más procesadores están configurados adicionalmente para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos del campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, realizar un análisis de saliencia con respecto a la matriz S para identificar uno o más vectores U_{DIST} de la matriz U y uno o más vectores S_{DIST} de la matriz S, y determinar el uno o más primeros vectores y uno o más segundos vectores multiplicando, al menos en parte, el uno o más vectores U_{DIST} por el uno o más vectores S_{DIST}.

Cláusula 133143-14B. El dispositivo de la cláusula 133143-1B, en el que el uno o más procesadores están configurados adicionalmente para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos del campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y al determinar el uno o más primeros vectores y el uno o más segundos vectores, realizar un análisis de saliencia con respecto a la matriz S para identificar uno o más vectores V_{DIST} de la matriz V como al menos uno del uno o más primeros vectores y el uno o más segundos vectores.

Cláusula 133143-15B. El dispositivo de la cláusula 133143-1B, en el que la primera parte de los datos de audio se produce en el tiempo antes de la segunda porción de los datos de audio.

Cláusula 133143-16B. El dispositivo de la cláusula 133143-1B, en el que la primera parte de los datos de audio se produce en el tiempo después de la segunda porción de los datos de audio.

Cláusula 133143-17B. El dispositivo de la cláusula 133143-1B, en el que el uno o más procesadores están configurados además para reordenar al menos uno del uno o más primeros vectores basándose en la correlación cruzada para generar el uno o más primeros vectores reordenados, y al reordenar los primeros vectores, aplicar una ventana de interpolación de entrar en fundido/fundido entre el uno o más primeros vectores para asegurar una transición suave cuando se generan los uno o más primeros vectores reordenados.

Cláusula 133143-18B. El dispositivo de la cláusula 133143-1B, en el que el uno o más procesadores están configurados además para reordenar uno o más primeros vectores basados en al menos la correlación cruzada para generar el uno o más primeros vectores reordenados, generar un flujo de bits para incluir el uno o más primeros vectores reordenados o una versión codificada del uno o más primeros vectores reordenados, y especificar en el flujo de bits cómo se reordenaron el uno o más primeros vectores.

Cláusula 133143-19B. El dispositivo de la cláusula 133143-1B, en el que la correlación cruzada facilita la extracción del uno o ambos del uno o más primeros vectores y uno o más segundos vectores con el fin de promover la codificación de audio del uno o ambos del uno o más primeros vectores y el uno o más segundos vectores.

[0607] La FIG. 40D es un diagrama de bloques que ilustra un ejemplo de dispositivo de codificación de audio 510D

que puede realizar diversos aspectos de las técnicas descritas en esta descripción para comprimir los coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales. El dispositivo de codificación de audio 510D puede ser similar al dispositivo de codificación de audio 510C en que el dispositivo de codificación de audio 510D incluye una unidad de compresión de audio 512, una unidad de codificación de audio 514 y una unidad de generación de flujo de bits 516. Además, la unidad de compresión de audio 512 del dispositivo de codificación de audio 510D puede ser similar a la del dispositivo de codificación de audio 510C porque la unidad de compresión de audio 512 incluye una unidad de descomposición 518.

[0608] Sin embargo, la unidad de compresión de audio 512 del dispositivo de codificación de audio 510D puede diferir de la unidad de compresión de audio 512 del dispositivo de codificación de audio 510C en que la unidad de extracción de componente de campo sonoro 520 incluye una unidad adicional, indicada como unidad de cuantización 534 ("unidad cuant. 534"). Por esta razón, la unidad de extracción de componentes del campo sonoro 520 del dispositivo de codificación de audio 510D se denomina "unidad de extracción de componentes del campo sonoro 520D".

[0609] La unidad de cuantización 534 representa una unidad configurada para cuantizar uno o más vectores V^T_{DIST} 525E y/o uno o más vectores V^T_{BG} 525F para generar uno o más vectores $V^T_{Q_DIST}$ 525G correspondientes y/o uno o más vectores $V^T_{Q_BG}$ 525H. La unidad de cuantización 534 puede cuantizar (que es un término de procesamiento de señales para el redondeo matemático a través de la eliminación de bits utilizados para representar un valor) uno o más vectores V^T_{DIST} 525E para reducir el número de bits que se utilizan para representar el uno o más vectores V^T_{DIST} 525E en el flujo de bits 517. En algunos ejemplos, la unidad de cuantización 534 puede cuantizar los valores de 32 bits de uno o más vectores V^T_{DIST} 525E, reemplazando estos valores de 32 bits por valores de 16 bits redondeados para generar uno o más vectores $V^T_{Q_DIST}$ 525G. A este respecto, la unidad de cuantización 534 puede operar de una manera similar a la descrita anteriormente con respecto a la unidad de cuantización 52 del dispositivo de codificación de audio 20 mostrado en el ejemplo de la FIG. 4.

[0610] La cuantización de esta naturaleza puede introducir un error en la representación del campo sonoro que varía de acuerdo con la tosquedad de la cuantización. En otras palabras, los bits más utilizados para representar uno o más vectores V^T_{DIST} 525E pueden dar lugar a menos error de cuantización. El error de cuantización debido a la cuantización de los vectores V^T_{DIST} 525E (que se puede denominar "E_{DIST}") puede determinarse restando uno o más vectores V^T_{DIST} 525E a uno o más vectores $V^T_{Q_DIST}$ 525G.

[0611] De acuerdo con las técnicas descritas en esta divulgación, el dispositivo de codificación de audio 510D puede compensar uno o más de los errores de cuantización E_{DIST} proyectando el error E_{DIST} o modificando de otro modo uno o más de los vectores $U_{DIST} * S_{DIST}$ 527 o los coeficientes armónicos esféricos de fondo 531 generados multiplicando uno o más vectores U_{BG} 525D por uno o más vectores S_{BG} 525B y luego por uno o más vectores V^T_{BG} 525F. En algunos ejemplos, el dispositivo de codificación de audio 510D sólo puede compensar el error E_{DIST} en los vectores $U_{DIST} * S_{DIST}$ 527. En otros ejemplos, el dispositivo de codificación de audio 510D sólo puede compensar el error E_{BG} en los coeficientes armónicos esféricos de fondo. En otros ejemplos, el dispositivo de codificación de audio 510D puede compensar el error E_{DIST} tanto en los vectores $U_{DIST} * S_{DIST}$ 527 como en los coeficientes armónicos esféricos de fondo.

[0612] En funcionamiento, la unidad de análisis de componentes relevantes 524 puede configurarse para emitir el uno o más vectores S_{DIST} 525, uno o más vectores S_{BG} 525B, uno o más vectores U_{DIST} 525C, uno o más vectores U_{BG} 525D, el uno o más vectores V^T_{DIST} 525E y uno o más vectores V^T_{BG} 525F a la unidad matemática 526. La unidad de análisis de componentes relevantes 524 también puede emitir el uno o más vectores V^T_{DIST} 525E a la unidad de cuantización 534. La unidad de cuantización 534 puede cuantizar uno o más vectores V^T_{DIST} 525E para generar uno o más vectores $V^T_{Q_DIST}$ 525G. La unidad de cuantización 534 puede proporcionar uno o más vectores $V^T_{Q_DIST}$ 525G a la unidad matemática 526, proporcionando al mismo tiempo uno o más vectores $V^T_{Q_DIST}$ 525G a la unidad de reordenación de vectores 532 (como se ha descrito anteriormente). La unidad de reordenación de vectores 532 puede operar con respecto a uno o más vectores $V^T_{Q_DIST}$ 525G de una manera similar a la descrita anteriormente con respecto a los vectores $V^T_{Q_DIST}$ 525E.

[0613] Al recibir estos vectores 525-525G ("vectores 525"), la unidad matemática 526 puede determinar primero coeficientes armónicos esféricos distintos que describen componentes distintos de los coeficientes armónicos esféricos de campo sonoro y de fondo que describen componentes de fondo del campo sonoro. La unidad matemática matricial 526 puede configurarse para determinar los distintos coeficientes armónicos esféricos multiplicando el uno o más vectores U_{DIST} 525C por el uno o más vectores S_{DIST} 525A y luego por uno o más vectores V^T_{DIST} 525E. La unidad matemática 526 puede configurarse para determinar los coeficientes armónicos esféricos de fondo multiplicando el uno o más vectores U_{BG} 525D por uno o más vectores S_{BG} 525A y luego por uno o más vectores V^T_{BG} 525E.

[0614] La unidad matemática 526 puede entonces determinar uno o más vectores $U_{DIST} * S_{DIST}$ compensados 527' (que pueden ser similares a los vectores $U_{DIST} * S_{DIST}$ 527 excepto que estos vectores incluyen valores para compensar el error E_{DIST}) realizando una operación de pseudoinvertida con respecto a uno o más vectores $V^T_{Q_DIST}$ 525G y multiplicando luego los distintos armónicos esféricos por la pseudoinvertida del uno o más vectores $V^T_{Q_DIST}$

525G. La unidad de reordenación de vectores puede operar de la manera descrita anteriormente para generar vectores reordenados 527', los cuales son entonces codificados de audio por la unidad de codificación de audio S 15A para generar vectores 515' reordenados codificados de audio, de nuevo como se ha descrito anteriormente.

[0615] La unidad matemática 526 puede proyectar a continuación el error E_{DIST} a los coeficientes esféricos armónicos de fondo. La unidad matemática 526 puede, para realizar esta proyección, determinar o recuperar de otro modo los coeficientes esféricos armónicos originales 511 añadiendo los coeficientes armónicos esféricos distintos a los coeficientes esféricos de armónicos de fondo. La unidad matemática 526 puede entonces restar los coeficientes armónicos esféricos distintos cuantificados (que pueden generarse multiplicando los vectores U_{DIST} 525C por los vectores S_{DIST} 525A y luego por los vectores $V^T_{Q_DIST}$ 525G) y los coeficientes armónicos esféricos de fondo de los coeficientes armónicos esféricos 511 para determinar el error restante debido a la cuantización de los vectores V^T_{DIST} 519. La unidad matemática 526 puede entonces añadir este error a los coeficientes armónicos esféricos de fondo cuantificados para generar coeficientes armónicos esféricos de fondo cuantificados compensados 531'.

[0616] En cualquier caso, la unidad de reducción de órdenes 528A puede actuar como se ha descrito anteriormente para reducir los coeficientes armónicos esféricos de fondo cuantificados compensados 531'a coeficientes armónicos esféricos de fondo reducidos 529', que pueden codificarse por audio por la unidad de codificación de audio 514 de la manera descrita anteriormente para generar coeficientes armónicos esféricos de fondo reducidos codificados de audio 515B'.

[0617] De este modo, las técnicas pueden permitir que el dispositivo de codificación de audio 510D cuantifique uno o más primeros vectores, tales como vectores V^T_{DIST} 525E, representativos de uno o más componentes de un campo sonoro y compensen el error introducido debido a la cuantización del uno o más primeros vectores en uno o más segundos vectores, tales como los vectores $U_{DIST} * S_{DIST}$ 527 y/o los vectores de coeficientes armónicos esféricos de fondo 531, que son también representativos del mismo uno o más componentes del campo sonoro.

[0618] Además, las técnicas pueden proporcionar esta compensación de error de cuantización de acuerdo con las siguientes cláusulas.

Cláusula 133146-1B. Un dispositivo, tal como el dispositivo de codificación de audio 510D, que comprende: uno o más procesadores configurados para cuantizar uno o más primeros vectores representativos de uno o más componentes distintos de un campo sonoro, y compensar el error introducido debido a la cuantización del uno o más primeros vectores en uno o más segundos vectores que son también representativos del mismo uno o más componentes distintos del campo sonoro.

Cláusula 133146-2B. El dispositivo de la cláusula 133146-1B, en el que el uno o más procesadores están configurados para cuantizar uno o más vectores de una transposición de una matriz V generada, al menos en parte, realizando una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos que describen el campo sonoro.

Cláusula 133146-3B. El dispositivo de la cláusula 133146-1B, en el que el uno o más procesadores están configurados adicionalmente para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de los vectores singulares izquierdos de la pluralidad de armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y en el que el uno o más procesadores están configurados para cuantizar uno o más vectores a partir de una transposición de la matriz V.

Cláusula 133146-4B. El dispositivo de la cláusula 133146-1B, en el que el uno o más procesadores están configurados para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de los vectores singulares izquierdos de la pluralidad de armónicos esféricos, una matriz S representativa de los valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, en el que uno o más procesadores están configurados para cuantizar uno o más vectores de una transposición de la matriz V, en el que el uno o más procesadores están configurados para compensar el error introducido debido a la cuantización en uno o más vectores U * S calculados multiplicando uno o más vectores U de la matriz U por uno o más vectores S de la matriz S.

Cláusula 133146-5B. El dispositivo de la cláusula 133146-1B, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, determinar

uno o más vectores U_{DIST} de la matriz U, cada uno de los cuales corresponde a uno de los distintos componentes del campo sonoro, determinar uno o más vectores S_{DIST} de la matriz S, cada uno de los cuales corresponde al mismo de los distintos componentes del campo sonoro, y determinar uno o más vectores v^T_{DIST} de una transposición de la matriz V, cada uno de los cuales corresponde al mismo de los distintos componentes del campo sonoro,
en el que uno o más procesadores están configurados para cuantizar uno o más vectores $_{VTDIST}$ para generar uno o más vectores $V^T_{Q_DIST}$ y en el que uno o más procesadores están configurados para compensar el error introducido debido a la cuantización en uno o más $U_{DIST} * S_{DIST}$ vectores calculados multiplicando uno o más vectores U_{DIST} de la matriz U por uno o más vectores S_{DIST} de la matriz S para generar uno o más vectores $U_{DIST} * S_{DIST}$ compensados con error.

Cláusula 133146-6B. El dispositivo de la cláusula 133146-5B, en el que el uno o más procesadores están configurados para determinar coeficientes armónicos esféricos distintos basados en uno o más vectores U_{DIST}, uno o más vectores S_{DIST} y uno o más vectores $_{VTDIST}$, y realizar una pseudoinvertida con respecto a los vectores $V^T_{Q_DIST}$ para dividir los distintos coeficientes armónicos esféricos por uno o más vectores $V^T_{Q_DIST}$ y de este modo generar uno o más vectores $U_{C_DIST} * S_{C_DIST}$ compensados por error que se compensan al menos en parte por el error introducido a través de la cuantización de los vectores v^T_{DIST}.

Cláusula 133146-7B. El dispositivo de la cláusula 133146-5B, en el que el uno o más procesadores están configurados adicionalmente para codificar por audio los uno o más vectores $U_{DIST} * S_{DIST}$ compensados por error.

Cláusula 133146-8B. El dispositivo de la cláusula 133146-1B, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, determinar uno o más vectores U_{BG} de la matriz U que describen uno o más componentes de fondo del campo sonoro y uno o más vectores U_{DIST} de la matriz U que describen uno o más componentes distintos del campo sonoro, determinar uno o más vectores S_{BG} de la matriz S que describen el uno o más componentes de fondo del campo sonoro y uno o más vectores S_{DIST} de la matriz S que describen uno o más componentes distintos del campo sonoro, y determinar uno o más vectores V^T_{DIST} y uno o más vectores V^T_{BG} de una transposición de la matriz V, en el que los vectores V^T_{DIST} describen uno o más componentes distintos del campo sonoro y el V^T_{BG} describen uno o más componentes de fondo del campo sonoro, en el que el uno o más procesadores están configurados para cuantizar el uno o más vectores V^T_{DIST} para generar uno o más vectores $V^T_{Q_DIST}$, y en el que uno o más procesadores están configurados además para compensar al menos una porción del error introducido debido a la cuantización en los coeficientes armónicos esféricos de fondo formados multiplicando el uno o más vectores U_{BG} por uno o más vectores S_{BG} y luego por uno o más vectores V^T_{BG} para generar coeficientes armónicos esféricos de fondo compensados por error.

Cláusula 133146-9B. El dispositivo de la cláusula 133146-8B, en el que el uno o más procesadores están configurados para determinar el error basado en los vectores $_{VTDIST}$ y uno o más vectores $U_{DIST} * S_{DIST}$ formados multiplicando los vectores U_{DIST} por los vectores S_{DIST}, y añadir el error determinado a los coeficientes armónicos esféricos de fondo para generar los coeficientes armónicos esféricos de fondo compensados por error.

Cláusula 133146-10B. El dispositivo de la cláusula 133146-8B, en el que el uno o más procesadores están configurados adicionalmente para codificar por audio los coeficientes armónicos esféricos compensados por error.

Cláusula 133146-11B. El dispositivo de la cláusula 133146-1B,
en el que el uno o más procesadores están configurados para compensar el error introducido debido a la cuantización del uno o más primeros vectores en uno o más segundos vectores que son también representativos del mismo o más componentes del campo sonoro para generar uno o más segundos vectores compensados por error, y en el que el uno o más procesadores están configurados además para generar un flujo de bits para incluir el uno o más segundos vectores compensados de error y uno o más primeros vectores cuantificados.

Cláusula 133146-12B. El dispositivo de la cláusula 133146-1B, en el que el uno o más procesadores están configurados para compensar el error introducido debido a la cuantización del uno o más primeros vectores en uno o más segundos vectores que son también representativos del mismo o más componentes del campo sonoro para generar uno o más segundos vectores compensados de error, y en el que el uno o más procesadores están configurados adicionalmente para codificar por audio los uno o más segundos vectores compensados de error, y generar un flujo de bits para incluir el audio codificado de uno o más segundos vectores compensados de error y el uno o más primeros vectores cuantificados.

Cláusula 133146-1C. Un dispositivo, tal como el dispositivo de codificación de audio 510D, que comprende: uno o más procesadores configurados para cuantizar uno o más primeros vectores representativos de uno o más componentes distintos de un campo sonoro, y compensar el error introducido debido a la cuantización del uno o

más primeros vectores en uno o más segundos vectores que son representativos de uno o más componentes de fondo del campo sonoro.

Cláusula 133146-2C. El dispositivo de la cláusula 133146-1C, en el que el uno o más procesadores están configurados para cuantizar uno o más vectores de una transposición de una matriz V generada, al menos en parte, realizando una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos que describen el campo sonoro.

Cláusula 133146-3C. El dispositivo de la cláusula 133146-1C, en el que el uno o más procesadores están configurados adicionalmente para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de los vectores singulares izquierdos de la pluralidad de armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y en el que el uno o más procesadores están configurados para cuantizar uno o más vectores a partir de una transposición de la matriz V.

Cláusula 133146-4C. El dispositivo de la cláusula 133146-1C, en el que el uno o más procesadores están configurados adicionalmente para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, determinar uno o más vectores U_{DIST} de la matriz U, cada uno de los cuales corresponde a uno de los distintos componentes del campo sonoro, determinar uno o más vectores S_{DIST} de la matriz S, cada uno de los cuales corresponde al mismo de los distintos componentes del campo sonoro, y determinar uno o más vectores V^T_{DIST} de una transposición de la matriz V, cada uno de los cuales corresponde al mismo de los distintos componentes del campo sonoro, en el que el uno o más procesadores están configurados para cuantizar el uno o más vectores V^T_{DIST} para generar uno o más vectores $V^T_{Q_DIST}$, y compensar al menos una porción del error introducido debido a la cuantización en uno o más vectores $U_{DIST} * S_{DIST}$ calculados multiplicando el uno o más vectores U_{DIST} de la matriz U por uno o más vectores S_{DIST} de la matriz S para generar uno o más vectores $U_{DIST} * S_{DIST}$ compensados de error.

Cláusula 133146-5C. El dispositivo de la cláusula 133146-4C, en el que el uno o más procesadores están configurados para determinar coeficientes armónicos esféricos distintos basadosen uno o más vectores U_{DIST}, uno o más vectores S_{DIST} y uno o más vectores V^T_{DIST}, y realizar una pseudoinvertida con respecto a los vectores $V^T_{Q_DIST}$ para dividir los distintos coeficientes armónicos esféricos por uno o más vectores $V^T_{Q_DIST}$ y de este modo generar uno o más vectores $U_{C_DIST} * S_{C_DIST}$ que compensan al menos en parte el error introducido mediante la cuantización de los vectores V^T_{DIST}.

Cláusula 133146-6C. El dispositivo de la cláusula 133146-4C, en el que el uno o más procesadores están configurados adicionalmente para codificar por audio los uno o más vectores $U_{DIST} * S_{DIST}$ compensados por error.

Cláusula 133146-7C. El dispositivo de la cláusula 133146-1C, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos representativos de un campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, determinar uno o más vectores U_{BG} de la matriz U que describen uno o más componentes de fondo del campo sonoro y uno o más vectores U_{DIST} de la matriz U que describen uno o más componentes distintos del campo sonoro, determinar uno o más vectores S_{BG} de la matriz S que describen el uno o más componentes de fondo del campo sonoro y uno o más vectores S_{DIST} de la matriz S que describen uno o más componentes distintos del campo sonoro, y determinar uno o más vectores V^T_{DIST} y uno o más vectores V^T_{BG} de una transposición de la matriz V, en el que los vectores V^T_{DIST} describen uno o más componentes distintos del campo sonoro y el V^T_{BG} describen uno o más componentes de fondo del campo sonoro, en el que el uno o más procesadores están configurados para cuantizar el uno o más vectores V^T_{DIST} para generar uno o más vectores $V^T_{Q_DIST}$, y en el que uno o más procesadores están configurados para compensar el error introducido debido a la cuantización en los coeficientes armónicos esféricos de fondo formados multiplicando el uno o más vectores U_{BG} por uno o más vectores S_{BG} y luego por uno o más vectores V^T_{BG} para generar coeficientes armónicos esféricos de fondo compensados por error.

Cláusula 133146-8C. El dispositivo de la cláusula 133146-7C, en el que el uno o más procesadores están configurados para determinar el error basado en los vectores V^T_{DIST} y uno o más vectores $U_{DIST} * S_{DIST}$ formados multiplicando los vectores U_{DIST} por los vectores S_{DIST}, y añadir el error determinado a los coeficientes armónicos esféricos de fondo para generar los coeficientes armónicos esféricos de fondo compensados por error.

Cláusula 133146-9C. El dispositivo de la cláusula 133146-7C, en el que el uno o más procesadores están configurados adicionalmente para codificar por audio los coeficientes armónicos esféricos compensados de error.

Cláusula 133146-10C. El dispositivo de la cláusula 133146-1C, en el que el uno o más procesadores están configurados adicionalmente para compensar el error introducido debido a la cuantización del uno o más primeros vectores en uno o más segundos vectores que son también representativos del mismo o más componentes del campo sonoro para generar uno o más segundos vectores compensados por error, y generar un flujo de bits para incluir el uno o más segundos vectores compensados de error y uno o más primeros vectores cuantificados.

Cláusula 133146-11C. El dispositivo de la cláusula 133146-1C, en el que el uno o más procesadores están configurados adicionalmente para compensar el error introducido debido a la cuantización del uno o más primeros vectores en uno o más segundos vectores que son también representativos del mismo o más componentes del campo sonoro para generar uno o más segundos vectores compensados de error, codificar por audio los uno o más segundos vectores compensados de error, y generar un flujo de bits para incluir el audio codificado de uno o más segundos vectores compensados de error y el uno o más primeros vectores cuantificados.

[0619] En otras palabras, cuando se usa descomposición SVD basada en tramas (o procedimientos relacionados tales como KLT y PCA) en señales HoA con el propósito de reducción de ancho de banda, las técnicas descritas en esta divulgación pueden permitir que el dispositivo de codificación de audio 10D cuantifique los primeros vectores de la matriz U (multiplicada por los valores singulares correspondientes de la matriz S), así como los correspondientes vectores del vector V. Esto comprenderá los componentes de "primer plano" o "distintos" del campo sonoro. Las técnicas pueden entonces permitir que el dispositivo de codificación de audio 510D codifique los vectores U*S usando un motor de codificación de audio de "caja negra", tal como un codificador AAC. El V vector puede ser escalar o un vector cuantificado.

[0620] Además, algunos de los vectores restantes en la matriz U pueden multiplicarse con los valores singulares correspondientes de la matriz S y de la matriz V y también codificarse utilizando un motor de codificación de audio de "caja negra". Estos comprenderán los componentes "de fondo" del campo sonoro. Una simple cuantización escalar de 16 bits de los vectores V puede dar como resultado una sobrecarga de aproximadamente 80 kbps para el 4º orden (25 coeficientes) y 160 kbps para el 6º orden (49 coeficientes). Una cuantización más gruesa puede dar lugar a errores de cuantización mayores. Las técnicas descritas en esta divulgación pueden compensar el error de cuantización de los vectores V, mediante la "proyección" del error de cuantización del vector V sobre los componentes de primer plano y de fondo.

[0621] Las técnicas en esta divulgación pueden incluir el cálculo de una versión cuantificada del vector V real. Este V vector cuantificado puede denominarse "V" (donde V' = V + e). La señal HoA subyacente -para los componentes de primer plano- que las técnicas están intentando recrear está dada por H_f = USV, donde U, S y V contienen solamente los elementos de primer plano. Para el propósito de este análisis, US será reemplazado por un solo conjunto de vectores U. Por lo tanto, H_f = UV. Dado que hay un V' erróneo, las técnicas están intentando recrear H_f lo más cerca posible. Por lo tanto, las técnicas pueden permitir que el dispositivo de codificación de audio 10D encuentre U' de tal forma que H_f = U'V'. El dispositivo de codificación de audio 10D puede utilizar una metodología pseudoinvertida que permite que U' = H_f[V']^(-1). Utilizando el denominado motor de codificación de audio de "caja negra" para codificar U', las técnicas pueden minimizar el error en H, causado por lo que se puede denominar como el vector V' erróneo.

[0622] De manera similar, las técnicas también pueden permitir que el dispositivo de codificación de audio proyecte el error debido a la cuantización de V en los elementos de fondo. El dispositivo de codificación de audio 510D puede configurarse para recrear la señal HoA total que es una combinación de señales HoA de primer plano y de fondo, es decir, H = H_f + H_b. Esto puede nuevamente ser modelado como H = H_f + e + H_b, debido al error de cuantización en V'. De esta manera, en lugar de poner el H_b a través del "codificador de audio de caja negra", se pone (e + H_b) a través del codificador de audio, compensando de hecho el error en V'. En la práctica, esto compensa el error sólo hasta el orden determinado por el dispositivo de codificación de audio 510D para enviar los elementos de fondo.

[0623] La FIG. 40E es un diagrama de bloques que ilustra un ejemplo de dispositivo de codificación de audio 510E que puede realizar diversos aspectos de las técnicas descritas en esta descripción para comprimir los coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales. El dispositivo de codificación de audio 510E puede ser similar al dispositivo de codificación de audio 510D en que el dispositivo de codificación de audio 510E incluye una unidad de compresión de audio 512, una unidad de codificación de audio 514 y una unidad de generación de flujo de bits 516. Además, la unidad de compresión de audio 512 del dispositivo de codificación de audio 510E puede ser similar a la del dispositivo de codificación de audio 510D porque la unidad de compresión de audio 512 incluye una unidad de descomposición 518.

[0624] Sin embargo, la unidad de compresión de audio 512 del dispositivo de codificación de audio 510E puede diferir de la unidad de compresión de audio 512 del dispositivo de codificación de audio 510D en que la unidad matemática 526 de la unidad de extracción de componentes del campo sonoro 520 realiza aspectos adicionales de las técnicas descritas en esta divulgación para reducir aún más la matriz V 519A antes de incluir la versión reducida de la transposición de la matriz V 519A en el flujo de bits 517. Por esta razón, la unidad de extracción de componentes del campo sonoro 520 del dispositivo de codificación de audio 510E se denomina "unidad de extracción de componentes del campo sonoro 520E".

[0625] En el ejemplo de la FIG. 40E, la unidad de reducción de órdenes 528, en lugar de reenviar los coeficientes armónicos esféricos de fondo reducidos 529' a la unidad de codificación de audio 514, devuelve los coeficientes armónicos esféricos de fondo reducidos 529' a la unidad matemática 526. Como se ha indicado anteriormente, estos coeficientes armónicos esféricos de fondo reducidos 529' pueden haberse reducido eliminando los de los coeficientes correspondientes a funciones de base esférica que tienen uno o más órdenes y/o subórdenes identificados. El orden reducido de los coeficientes armónicos esféricos de fondo reducidos 529' puede estar representado por la variable N_{BG}.

[0626] Dado que la unidad de extracción de componentes del campo sonoro 520E no puede realizar reducción de orden con respecto a uno o más vectores $U_{DIST} * S_{DIST}$ reordenados 533', el orden de esta descomposición de los coeficientes armónicos esféricos que describen componentes distintos del campo sonoro (que puede representarse por la variable N_{DIST}) puede ser mayor que el orden de fondo, N_{BG}. En otras palabras, N_{BG} puede ser comúnmente menor que N_{DIST}. Una posible razón de que N_{BG} pueda ser menor que N_{DIST} es que se asume que los componentes de fondo no tienen mucha direccionalidad de tal forma que no se requieren funciones de base esférica de orden superior, permitiendo de este modo la reducción de orden y dando lugar a que N_{BG} sea menor que N_{DIST}.

[0627] Dado que el uno o más de los vectores $V^T_{Q_DIST}$ reordenados 539 previamente se enviaron abiertamente, sin codificación de audio de estos vectores 539 en el flujo de bits 517, como se muestra en los ejemplos de las FIGS. 40A-40D, el uno o más vectores $V^T_{Q_DIST}$ reordenados 539 pueden consumir un ancho de banda considerable. Como un ejemplo, cada uno de los vectores $V^T_{Q_DIST}$ 539, cuando se cuantifican a valores escalares de 16 bits, pueden consumir aproximadamente 20 Kbps para los datos de audio ambisónicos de cuarto orden (donde cada vector tiene 25 coeficientes) y 40 Kbps para los datos de audio ambisónicos de sexto orden (donde cada vector tiene 49 coeficientes).

[0628] De acuerdo con diversos aspectos de las técnicas descritas en esta divulgación, la unidad de extracción de componentes del campo sonoro 520E puede reducir la cantidad de bits que necesitan ser especificados para coeficientes armónicos esféricos o descomposiciones de los mismos, tales como los uno o más vectores $V^T_{Q_DIST}$ reordenados 539. En algunos ejemplos, la unidad matemática 526 puede determinar, basándose en los coeficientes armónicos esféricos reducidos ordenados 529', los de los vectores $V^T_{Q_DIST}$ reordenados 539 que se van a eliminar y recombinar con los coeficientes armónicos esféricos reducidos ordenados 529' y los de los vectores $V^T_{Q_DIST}$ reordenados 539 que son para formar los vectores $V^T_{PEQUEÑO}$ 521. Es decir, la unidad matemática 526 puede determinar un orden del orden reducido de los coeficientes armónicos esféricos 529', donde este orden se puede representar N_{BG}. Los vectores $V^T_{Q_DIST}$ reordenados 539 pueden ser de un orden representado por la variable N_{DIST}, donde N_{DIST} es mayor que el orden N_{BG}.

[0629] La unidad matemática 526 puede entonces analizar los primeros N_{BG} órdenes de los vectores $V^T_{Q_DIST}$ reordenados 539, eliminando los vectores que especifican coeficientes armónicos esféricos descompuestos que corresponden a funciones de base esféricas que tienen un orden menor o igual que N_{BG}. Estos vectores $V^T_{Q_DIST}$ reordenados eliminados 539 pueden usarse entonces para formas coeficientes armónicos esféricos intermedios multiplicando los de los vectores $U_{DIST} * S_{DIST}$ reordenados 533' representativo de las versiones descompuestas de los coeficientes armónicos esféricos 511 correspondientes a las funciones de base esférica que tienen un orden menor de o igual a N_{BG} por los vectores $V^T_{Q_DIST}$ reordenados eliminados 539 para formar los coeficientes armónicos esféricos distintos intermedios. La unidad matemática 526 puede entonces generar coeficientes armónicos esféricos de fondo modificados 537 añadiendo los coeficientes armónicos esféricos distintos intermedios a los coeficientes armónicos esféricos reducido ordenados 529'. La unidad matemática 526 puede entonces pasar estos coeficientes armónicos esféricos de fondo modificados 537 a la unidad de codificación de audio 514, cuyo audio codifica estos coeficientes 537 para formar coeficientes armónicos esféricos de fondo modificado codificados de audio 515B'.

[0630] La unidad matemática 526 puede entonces pasar el uno o más vectores $V^T_{PEQUEÑO}$ 521, que pueden representar los vectores 539 representativos de una forma descompuesta de los coeficientes armónicos esféricos 511 correspondientes a funciones de base esféricas que tienen un orden mayor que N_{BG} y menor que o igual a N_{DIST}. A este respecto, la unidad matemática 526 puede realizar operaciones similares a la unidad de reducción de coeficientes 46 del dispositivo de codificación de audio 20 mostrado en el ejemplo de la FIG. 4. La unidad matemática 526 puede pasar el uno o más vectores $V^T_{PEQUEÑO}$ 521 a la unidad de generación de flujo de bits 516, que puede generar el flujo de bits 517 para incluir los vectores $V^T_{PEQUEÑO}$ 521 a menudo en su forma codificada sin audio original. Dado que los vectores $V^T_{PEQUEÑO}$ 521 incluyen menos vectores que los vectores $V^T_{Q_DIST}$ reordenados 539, las técnicas pueden facilitar la asignación de menos bits a los vectores $V^T_{Q_DIST}$ reordenados 539 especificando únicamente los vectores $V^T_{PEQUEÑO}$ 521 en el flujo de bits 517.

[0631] Aunque se muestra sin cuantificar, en algunos casos, el dispositivo de codificación de audio 510E puede cuantizar los vectores V^T_{BG} 525F. En algunos casos, tal como cuando la unidad de codificación de audio 514 no se utiliza para comprimir coeficientes armónicos esféricos de fondo, el dispositivo de codificación de audio 510E puede cuantizar los vectores V^T_{BG} 525F.

[0632] De este modo, las técnicas pueden permitir que el dispositivo de codificación de audio 510E determine al menos uno de uno o más vectores descompuestos a partir de coeficientes armónicos esféricos a recombinar con coeficientes armónicos esféricos de fondo para reducir la cantidad de bits que se requiere asignar a uno o más vectores en un flujo de bits, en el que los coeficientes armónicos esféricos describen un campo sonoro, y en el que los coeficientes armónicos esféricos de fondo describen uno o más componentes de fondo del mismo campo sonoro.

[0633] Es decir, las técnicas pueden permitir que el dispositivo de codificación de audio 510E sea configurado de una manera indicada por las siguientes cláusulas.

Cláusula 133149-1A. Un dispositivo, tal como el dispositivo de codificación de audio 510E, que comprende: uno o más procesadores configurados para determinar al menos uno de uno o más vectores descompuestos a partir de coeficientes armónicos esféricos para ser recombinados con coeficientes armónicos esféricos de fondo para reducir una cantidad de bits que se requiere para asignarse a uno o más vectores en un flujo de bits, en el que los coeficientes armónicos esféricos describen un campo sonoro, y en el que los coeficientes armónicos esféricos de fondo describen uno o más componentes de fondo del mismo campo sonoro.

Cláusula 133149-2A. El dispositivo de la cláusula 133149-1A, en el que el uno o más procesadores están configurados además para generar un conjunto reducido de uno o más vectores eliminando el determinado al menos uno de uno o más vectores del uno o más vectores.

Cláusula 133149-3A. El dispositivo de la cláusula 133149-1A, en el que el uno o más procesadores están configurados además para generar un conjunto reducido de uno o más vectores eliminando el determinado al menos uno de uno o más vectores del uno o más vectores, recombinar el al menos uno de uno o más vectores eliminados con los coeficientes armónicos esféricos de fondo para generar coeficientes armónicos esféricos de fondo modificados, y generar el flujo de bits para incluir el conjunto reducido de uno o más vectores y los coeficientes armónicos esféricos de fondo modificados.

Cláusula 133149-4A. El dispositivo de la cláusula 133149-3A, en el que el conjunto reducido de uno o más vectores está incluido en el flujo de bits sin que primero se codifique con audio.

Cláusula 133149-5A. El dispositivo de la cláusula 133149-1A, en el que el uno o más procesadores están configurados adicionalmente para generar un conjunto reducido de uno o más vectores eliminando el determinado al menos uno de uno o más vectores del uno o más vectores, recombinar el al menos uno eliminado de uno o más vectores con los coeficientes armónicos esféricos de fondo para generar coeficientes armónicos esféricos de fondo modificados, codificación de audio de los coeficientes armónicos esféricos de fondo modificados, y generar el flujo de bits para incluir el conjunto reducido de uno o más vectores y y los coeficientes armónicos esféricos de fondo modificados codificados con audio.

Cláusula 133149-6A. El dispositivo de la cláusula 133149-1A, en el que el uno o más vectores comprenden vectores representativos de al menos algún aspecto de uno o más componentes distintos del campo sonoro.

Cláusula 133149-7A. El dispositivo de la cláusula 133149-1A, en el que el uno o más vectores comprenden uno o más vectores de una transposición de una matriz V generada, al menos en parte, realizando una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos que describen el campo sonoro.

Cláusula 133149-8A. El dispositivo de la cláusula 133149-1A, en el que el uno o más procesadores están configurados adicionalmente para realizar una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos para generar una matriz U representativa de los vectores singulares izquierdos de la pluralidad de armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y en el que el uno o más vectores comprenden uno o más vectores a partir de una transposición de la matriz V.

Cláusula 133149-9A. El dispositivo de la cláusula 133149-1A, en el que el uno o más procesadores están configurados además para realizar una reducción de orden con respecto a los coeficientes armónicos esféricos de fondo para eliminar aquellos de los coeficientes armónicos esféricos de fondo correspondientes a funciones de base esféricas que tienen un orden y/o suborden identificado, en el que los coeficientes armónicos esféricos de fondo corresponden a un orden N_{BG}.

Cláusula 133149-10A. El dispositivo de la cláusula 133149-1A, en el que el uno o más procesadores están configurados además para realizar una reducción de orden con respecto a los coeficientes armónicos esféricos de fondo para eliminar aquellos de los coeficientes armónicos esféricos de fondo correspondientes a funciones de base esféricas que tienen un orden y/o suborden identificado, en el que los coeficientes armónicos esféricos de fondo corresponden a un orden N_{BG} que es menor del orden de coeficientes armónicos esféricos distintos, N_{DIST}, y en el que los coeficientes armónicos esféricos distintos representan distintos componentes del campo sonoro.

Cláusula 133149-11A. El dispositivo de la cláusula 133149-1A, en el que el uno o más procesadores están configurados además para realizar una reducción de orden con respecto a los coeficientes armónicos esféricos de fondo para eliminar aquellos de los coeficientes armónicos esféricos de fondo correspondientes a funciones de base esféricas que tienen un orden y/o suborden identificado, en el que los coeficientes armónicos esféricos de fondo corresponden a un orden N_{BG} que es menor del orden de coeficientes armónicos esféricos distintos, N_{DIST}, y en el que los coeficientes armónicos esféricos distintos representan distintos componentes del campo sonoro y no están sujetos a la reducción de orden.

Cláusula 133149-12A. El dispositivo de la cláusula 133149-1A, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y determinar uno o más vectores V^T_{DIST} y uno o más V^T_{BG} de una transposición de la matriz V, uno o más vectores V^T_{DIST} describen uno o más componentes distintos del campo sonoro y uno o más vectores V^T_{BG} describen uno o más componentes de fondo del campo sonoro, y en el que el uno o más vectores incluyen uno o más vectores V^T_{DIST}.

Cláusula 133149-13A. El dispositivo de la cláusula 133149-1A, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, determinar uno o más vectores V^T_{DIST} y uno o más V^T_{BG} de una transposición de la matriz V, uno o más vectores V_{DIST} describen uno o más componentes distintos del campo sonoro y uno o más vectores V_{BG} describen uno o más componentes de fondo del campo sonoro, y cuantifican uno o más vectores V^T_{DIST} para generar uno o más vectores $V^T_{Q_DIST}$, y en el que uno o más vectores incluye uno o más vectores $V^T_{Q_DIST}$.

Cláusula 133149-14A. El dispositivo de la cláusula 133149-12A o la cláusula 133149-13A, en el que el uno o más procesadores están configurados adicionalmente para determinar uno o más vectores U_{DIST} y uno o más vectores U_{BG} de la matriz U, uno o más vectores U_{DIST} describen uno o más componentes distintos del campo sonoro y uno o más vectores U_{BG} describen uno o más componentes de fondo del campo sonoro, y determinan uno o más vectores S_{DIST} y uno o más vectores S_{BG} de la matriz S, uno o más vectores S_{DIST} describen uno o más componentes distintos del campo sonoro y uno o más vectores S_{BG} describen uno o más componentes de fondo del campo sonoro.

Cláusula 133149-15A. El dispositivo de la cláusula 133149-14A, en el que el uno o más procesadores están configurados además para determinar los coeficientes armónicos esféricos de fondo en función de uno o más vectores U_{BG}, uno o más vectores S_{BG}, y uno o más V^T_{BG}, realizan reducción de orden con respecto a los coeficientes armónicos esféricos de fondo para generar coeficientes armónicos esféricos de fondo reducidos que tienen un orden igual a N_{BG}, multiplicar uno o más U_{DIST} por uno o más vectores S_{DIST} para generar uno o más vectores $U_{DIST} * S_{DIST}$, eliminar el menos uno determinado de uno o más vectores del uno o más vectores para generar un conjunto reducido de uno o más vectores, multiplicar uno o más vectores $U_{DIST} * S_{DIST}$ por el al menos uno eliminado de uno o más vectores V^T_{DIST} o uno o más vectores $V^T_{Q_DIST}$ para generar coeficientes armónicos esféricos distintos intermedios, y añadir los coeficientes armónicos esféricos distintos intermedios al coeficiente armónico esférico de fondo para recombinar el al menos uno eliminado de uno o más vectores V^T_{DIST} o uno o más vectores $V^T_{Q_DIST}$ con los coeficientes esféricos de armónicos de fondo.

Cláusula 133149-16A. El dispositivo de la cláusula 133149-14A, en el que el uno o más procesadores están configurados además para determinar los coeficientes armónicos esféricos de fondo en función de uno o más vectores U_{BG}, uno o más vectores S_{BG} y uno o más V^T_{BG}, realizar una reducción de orden con respecto a los coeficientes armónicos esféricos de fondo para generar coeficientes armónicos esféricos de fondo reducido que tienen un orden igual a N_{BG}, multiplicar uno o más U_{DIST} por uno o más vectores S_{DIST} para generar uno o más $U_{DIST} * S_{DIST}$, reordenar uno o más vectores $U_{DIST} * S_{DIST}$ para generar uno o más vectores $U_{DIST} * S_{DIST}$ reordenados, eliminar al menos uno determinado de uno o más vectores del uno o más vectores para generar un conjunto reducido de uno o más vectores, multiplicar el uno o más vectores $U_{DIST} * S_{DIST}$ reordenados por el al menos uno eliminado de uno o más vectores V^T_{DIST} o uno o más vectores $V^T_{Q_DIST}$ para generar coeficientes

armónicos esféricos distintos intermedios, y añadir los coeficientes armónicos esféricos distintos intermedios al coeficiente armónico esférico de fondo para recombinar el al menos uno eliminado de uno o más vectores V^T_{DIST} o uno o más vectores $V^T_{Q_DIST}$ con los coeficientes armónicos esféricos de fondo.

Cláusula 133149-17A. El dispositivo de la cláusula 133149-15A o la cláusula 133149-16A, en el que el uno o más procesadores están configurados adicionalmente para codificar en audio los coeficientes armónicos esféricos de fondo después de añadir los coeficientes armónicos esféricos distintos intermedios a los coeficientes armónicos esféricos de fondo, y generar el flujo de datos para incluir los coeficientes armónicos esféricos de fondo codificados en audio.

Cláusula 133149-18A. El dispositivo de la cláusula 133149-1A, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos, y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, determinar uno o más vectores V^T_{DIST} y uno o más V^T_{BG} de una transposición de la matriz V, uno o más vectores V_{DIST} describen uno o más componentes distintos del campo sonoro, y uno o más vectores V_{BG} describen uno o más componentes de fondo del campo sonoro, cuantificar uno o más vectores V^T_{DIST} para generar uno o más vectores $V^T_{Q_DIST}$, y reordenar uno o más vectores $V^T_{Q_DIST}$ para generar uno o más vectores $V^T_{Q_DIST}$ reordenados, y en el que el uno o más vectores incluye el uno o más vectores $V^T_{Q_DIST}$ reordenados.

[0634] La FIG. 40F es un diagrama de bloques que ilustra un dispositivo de codificación de audio ejemplar 510F que puede realizar diversos aspectos de las técnicas descritas en esta divulgación para comprimir los coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales. El dispositivo de codificación de audio 510F puede ser similar al dispositivo de codificación de audio 510C en que el dispositivo de codificación de audio 510F incluye una unidad de compresión de audio 512, una unidad de codificación de audio 514 y una unidad de generación de flujo de bits 516. Además, la unidad de compresión de audio 512 del dispositivo de codificación de audio 510F puede ser similar a la del dispositivo de codificación de audio 510C porque la unidad de compresión de audio 512 incluye una unidad de descomposición 518 y una unidad de reordenación de vectores 532, que pueden operar de forma similar a las unidades del dispositivo de codificación de audio 510C. En algunos ejemplos, el dispositivo de codificación de audio 510F puede incluir una unidad de cuantización 534, como se ha descrito con respecto a las FIGS. 40D y 40E, para cuantizar uno o más vectores de cualquiera de los vectores U_{DIST} 525C, los vectores U_{BG} 525D, los vectores V^T_{DIST} 525E y los vectores V^T_{BG} 525J.

[0635] Sin embargo, la unidad de compresión de audio 512 del dispositivo de codificación de audio 510F puede diferir de la unidad de compresión de audio 512 del dispositivo de codificación de audio 510C en que la unidad de análisis de componentes relevantes 524 de la unidad de extracción de componentes de campo sonoro 520 puede realizar un análisis de contenido para seleccionar el número de componentes de primer plano, denominado D en el contexto de las FIGS. 40A-40J. En otras palabras, la unidad de análisis de componentes relevantes 524 puede operar con respecto a las matrices U, S y V 519 de la manera descrita anteriormente para identificar si las versiones descompuestas de los coeficientes armónicos esféricos se generaron a partir de objetos de audio sintéticos o de una grabación natural con un micrófono. La unidad de análisis de componentes relevantes 524 puede entonces determinar D sobre la base de esta determinación sintética.

[0636] Además, la unidad de compresión de audio 512 del dispositivo de codificación de audio 510F puede diferir de la unidad de compresión de audio 512 del dispositivo de codificación de audio 510C en que la unidad de extracción de componentes del campo sonoro 520 puede incluir una unidad adicional, una unidad de reducción de orden y conservación de energía 528F (ilustrada como "unidad de red. orden y conserv. de energía 528F"). Por estas razones, la unidad de extracción de componentes del campo sonoro 520 del dispositivo de codificación de audio 510F se denomina "unidad de extracción de componentes del campo sonoro 520F".

[0637] La unidad de reducción de orden y conservación de energía 528F representa una unidad configurada para realizar la reducción de orden de los componentes de fondo de la matriz V_{BG} 525H representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos 511 manteniendo la energía global (y la presión de sonido concomitante) del campo sonoro descrito en parte por la matriz V_{BG} completa 525H. A este respecto, la unidad de reducción de orden y conservación de energía 528F puede realizar operaciones similares a las descritas anteriormente con respecto a la unidad de selección de fondo 48 y la unidad de compensación de energía 38 del dispositivo de codificación de audio 20 mostrado en el ejemplo de la FIG. 4.

[0638] La matriz V_{BG} completa 525H tiene una dimensionalidad $(N+1)^2$ x $(N+1)^2$ - D, donde D representa varios componentes principales o, en otras palabras, valores singulares que se determinan como relevantes en términos de ser componentes de audio distintos del campo sonoro. Es decir, la matriz V_{BG} completa 525H incluye aquellos valores singulares que se determinan como fondo ($_{BG}$) o, en otras palabras, componentes de sonido ambiental o audio no distinto del campo sonoro.

[0639] Como se ha descrito anteriormente con respecto, por ejemplo, a la unidad de reducción de orden 524 de las FIGS. 40B-40E, la unidad de reducción de orden y conservación de energía 528F puede extraer, eliminar o borrar de otro modo (a menudo por reducción a cero) los valores singulares de de fondo de la matriz V_{BG} 525H correspondientes a funciones de base esférica de orden superior. La unidad de reducción de orden y conservación de energía 528F puede emitir una versión reducida de la matriz V_{BG} 525H (designada como "matriz V_{BG}' 525I" y denominada en lo sucesivo "matriz V_{BG}' reducida 525I") para transponer la unidad 522. La matriz V_{BG}' reducida 525I puede tener una dimensionalidad $(ñ+1)^2$ x $(N+1)^2$ - D, con ñ <N. La unidad de transposición 522 aplica una operación de transposición a la matriz V_{BG}' reducida 525I para generar y producir una matriz V^T_{BG}' reducida transpuesta 525J a la unidad matemática 526, que puede operar para reconstruir los componentes de sonido de fondo del campo sonoro calculando U_{BG} * S_{BG} * V^T_{BG} usando la matriz U_{BG} 525D, la matriz S_{BG} 525B, y la matriz V^T_{BG}' reducida transpuesta 525J.

[0640] De acuerdo con las técnicas descritas en el presente documento, la unidad de reducción de orden y conservación de energía 528F está configurada además para compensar posibles reducciones en la energía total de los componentes de sonido de fondo del campo sonoro causadas por la reducción del orden de la matriz V_{BG} completa 525H para generar la matriz V_{BG}' reducida 525I. En algunos ejemplos, la unidad de reducción de orden y conservación de energía 528F compensa determinando una ganancia de compensación en forma de valores de amplificación que se aplican a cada una de las columnas $(N+1)^2$ - D de la matriz V_{BG}' reducida 525I con el fin de aumentar la energía cuadrática media (RMS) de la matriz V_{BG}' reducida 525I para igualar o al menos aproximar la RMS de la matriz V_{BG} completa 525H, antes de emitir la matriz V_{BG}' reducida 525I para transponer la unidad 522.

[0641] En algunos casos, la unidad de reducción de orden y conservación de energía 528F puede determinar la energía RMS de cada columna de la matriz V_{BG} completa 525H y la energía RMS de cada columna de la matriz V_{BG}' reducida 525I, a continuación, determinar el valor de amplificación para la columna como la relación de la primera con respecto a la última, como se indica en la siguiente ecuación:

$$\propto = v_{BG} / v_{BG}',$$

donde α es el valor de amplificación para una columna, V_{BG} representa una única columna de la matriz V_{BG} 525H, y v_{BG}' representa la columna única correspondiente de la matriz V_{BG}' 525I. Esto puede representarse en la notación matricial como:

$$A = V_{BG}^{RMS} / V_{BG}^{',RMS},$$

$$A = [\ \propto_1 ... \propto_{(N+1)^2-D}\ ],$$

donde V_{BG}^{RMS} es un vector RMS que tiene elementos que representan la RMS de cada columna de matriz V_{BG} 525H, $V_{BG}^{',RMS}$ es un vector RMS que tiene elementos que representan la RMS de cada columna de la matriz V_{BG}' reducida 525I, y A es un vector de valor de amplificación que tiene elementos para cada columna de la matriz V_{BG} 525H. La unidad de reducción de orden y conservación de energía 528F aplica una multiplicación escalar a cada columna de matriz V_{BG} reducida 525I usando el correspondiente valor de amplificación α, o en forma vectorial:

$$V''_{BG} = V_{BG}' A^T,$$

donde V''_{BG} representa una matriz V_{BG}' reducida 525I que incluye compensación de energía. La unidad de reducción de orden y conservación de energía 528F puede producir una matriz V_{BG}' reducida 525I que incluye una compensación de energía para transponer la unidad 522 para igualar (o casi igualar) la RMS de la matriz V_{BG}' reducida 525I con la de la matriz V_{BG} completa 525H. La dimensionalidad de salida de la matriz V_{BG}' reducida 525I que incluye la compensación de energía puede ser $(ñ+1)^2 x (N+1)^2 - D$.

[0642] En algunos ejemplos, para determinar cada RMS de columnas respectivas de matriz V_{BG}' reducida 525I y matriz V_{BG} completa 525H, la unidad de reducción de orden y conservación de energía 528F puede aplicar primero un reproductor de coeficientes armónicos esféricos de referencia (SHC) a las columnas. La aplicación del reproductor de SHC de referencia por la unidad de reducción de orden y conservación de energía 528F permite la determinación de RMS en el dominio SHC para determinar la energía del campo acústico global descrito por cada columna de la trama representada por la matriz V_{BG}' reducida 525I y la matriz V_{BG} completa 525H. Por lo tanto, en tales ejemplos, la unidad de reducción de orden y conservación de energía 528F puede aplicar el reproductor de SHC de referencia a cada columna de la matriz V_{BG} completa 525H y a cada columna reducida de la matriz V_{BG}' reducida 525I, determinar los valores de RMS respectivos para la columna y la columna reducida, y determinar el valor de amplificación para la columna como la relación del valor de RMS para la columna con respecto al valor de RMS para la columna reducida. En algunos ejemplos, la reducción de orden a la matriz V_{BG}' reducida 525I procede en dirección de la columna coincidente con la conservación de energía. Esto puede expresarse en pseudocódigo como se indica a continuación:

```
R = ReferenceRenderer;
para m = numDist+1 : numChannels
```

```
    fullV = V(:,m); //toma una columna de V => fullV
  reducedV =[fullV(1 :numBG); zeros(numChannels-numBG,1)];
    alpha=sqrt(sum((fullV'*R).^2)/sum((reducedV'*R).^2));
  if isnan(alpha) ||isinf(alpha), alpha = 1; end;
  V_out(:,m) = reducedV * alpha;
fin
```

[0643] En el pseudocódigo anterior, numChannels puede representar $(N+1)^2$ - D, numBG puede representar $(ñ+1)^2$, V puede representar la matriz V_{BG} 525H, y V_out puede representar la matriz V_{BG}' reducida 525I, y R puede representar el reproductor de SHC de referencia de la unidad de reducción de orden y conservación de energía 528F. La dimensionalidad de V puede ser $(N+1)^2$ x $(N+1)^2$ - D y la dimensionalidad de V_out puede ser $(ñ+1)^2$ x $(N+1)^2$ - D.

[0644] Como resultado, el dispositivo de codificación de audio 510F puede, al representar la pluralidad de coeficientes armónicos esféricos 511, reconstruir los componentes de sonido de fondo usando una matriz V_{BG}' reducida ordenada 525I que incluye compensación de energía que puede perderse como resultado del proceso de reducción de orden.

[0645] La FIG. 40G es un diagrama de bloques que ilustra un dispositivo de codificación de audio ejemplar 510G que puede realizar diversos aspectos de las técnicas descritas en esta divulgación para comprimir los coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales. En el ejemplo de la FIG. 40G, el dispositivo de codificación de audio 510G incluye una unidad de extracción de componentes del campo sonoro 520F. A su vez, la unidad de extracción de componentes del campo sonoro 520F incluye una unidad de análisis de componentes relevantes 524G.

[0646] La unidad de compresión de audio 512 del dispositivo de codificación de audio 510G puede, sin embargo, diferir de la unidad de compresión de audio 512 del dispositivo de codificación de audio 10F en que la unidad de compresión de audio 512 del dispositivo de codificación de audio 510G incluye una unidad de análisis de componentes relevantes 524G. La unidad de análisis de componentes relevantes 524G puede representar una unidad configurada para determinar la saliencia o distinción de los datos de audio que representan un campo sonoro, usando información basada en direccionalidad asociada con los datos de audio.

[0647] Mientras que las determinaciones basadas en energía pueden mejorar la reproducción de un campo sonoro descompuesto por SVD para identificar componentes de audio distintos del campo sonoro, las determinaciones basadas en energía también pueden hacer que un dispositivo identifique erróneamente componentes de audio de fondo como componentes de audio distintos, en aquellos casos en los que los componentes de audio de fondo exhiben un alto nivel de energía. Es decir, una separación de componentes de audio distintos y de fondo, exclusivamente basada en energía, puede no ser robusta, ya que los componentes de audio de fondo energéticos (por ejemplo, más altos) pueden identificarse incorrectamente como componentes de audio distintos. Para distinguir más fuertemente entre componentes de audio distintos y de fondo del campo sonoro, diversos aspectos de las técnicas descritas en esta divulgación pueden permitir que la unidad de análisis de componentes relevantes 524G realice un análisis basado en direccionalidad de SHC 511 para separar componentes de audio distintos y de fondo de versiones descompuestas de SHC 511.

[0648] La unidad de análisis de componentes relevantes 524G puede, en el ejemplo de la FIG. 40H, representar una unidad configurada o operable de otro modo para separar elementos distintos (o de primer plano) de elementos de fondo incluidos en una o más de la matriz V 519, la matriz S 519B y la matriz U 519C, similar a las unidades de análisis de componentes relevantes 524 de los dispositivos de codificación de audio descritos anteriormente 510-510F. De acuerdo con algunas técnicas basadas en SVD, los componentes más energéticos (por ejemplo, los primeros vectores de una o más de las matrices V, S y U 519-519C, o una matriz derivada de las mismas) pueden tratarse como componentes distintos. Sin embargo, los componentes más enérgicos (que están representados por vectores) de una o más de las matrices 519-519C no pueden, en ninguno de los escenarios, representar los componentes/señales que son los más direccionales.

[0649] A diferencia de las unidades de análisis de componentes relevantes descritas anteriormente 524, la unidad de análisis de componentes relevantes 524G puede implementar uno o más aspectos de las técnicas descritas en el presente documento para identificar elementos de primer plano basados en la direccionalidad de los vectores de una o más de las matrices 519-519C o una matriz derivada de las mismas. En algunos ejemplos, la unidad de análisis de componentes relevantes 524G puede identificar o seleccionar como componentes de audio distintos (donde los componentes también pueden denominarse "objetos"), uno o más vectores basados tanto en energía como en la direccionalidad de los vectores. Por ejemplo, la unidad de análisis de componentes relevantes 524G puede identificar aquellos vectores de una o más de las matrices 519-519C (o una matriz derivada de las mismas) que muestran tanto alta energía como alta direccionalidad (por ejemplo, representada como cociente de direccionalidad) como componentes de audio distintos. Como resultado, si la unidad de análisis de componentes relevantes 524G determina que un vector particular es relativamente menos direccional cuando se compara con otros vectores de una o más de las matrices 519-519C (o una matriz derivada de las mismas), independientemente entonces del nivel

de energía asociado con el vector particular, la unidad de análisis de componentes relevantes 524G puede determinar que el vector particular representa componentes de audio de fondo (o ambientales) del campo sonoro representado por los SHC 511. A este respecto, la unidad de análisis de componentes relevantes 524G puede realizar operaciones similares a las descritas anteriormente con respecto a la unidad de análisis de campo sonoro 44 del dispositivo de codificación de audio 20 mostrado en el ejemplo de la FIG. 4.

[0650] En algunas implementaciones, la unidad de análisis de componentes relevantes 524G puede identificar objetos de audio distintos (que, como se ha indicado anteriormente, también pueden denominarse "componentes") basándose en la direccionalidad, realizando las siguientes operaciones. La unidad de análisis de componentes relevantes 524G puede multiplicar (por ejemplo, utilizando uno o más procesos de multiplicación de matrices) la matriz V 519A por la matriz S 519B. Mediante la multiplicación de la matriz V 519A y la matriz S 519B, la unidad de análisis de componentes relevantes 524G puede obtener una matriz VS. Además, la unidad de análisis de componentes relevantes 524G puede cuadrar (es decir, exponenciar por una potencia de dos) al menos algunas de las entradas de cada uno de los vectores (que pueden ser una fila) de la matriz VS. En algunos casos, la unidad de análisis de componentes relevantes 524G puede sumar aquellas entradas al cuadrado de cada vector que están asociadas con un orden mayor que 1. Como un ejemplo, si cada vector de la matriz incluye 25 entradas, la unidad de análisis de componentes relevantes 524G puede, con respecto a cada vector, elevar al cuadrado las entradas de cada vector comenzando por la entrada quinta y terminando por la entrada vigésimo quinta, sumando las entradas al cuadrado para determinar un cociente de direccionalidad (o un indicador de direccionalidad). Cada operación de suma puede dar como resultado un cociente de direccionalidad para un vector correspondiente. En este ejemplo, la unidad de análisis de componentes relevantes 524G puede determinar que aquellas entradas de cada fila que están asociadas con un orden menor o igual a 1, concretamente, de la primera a la cuarta entradas, se dirigen más generalmente a la cantidad de energía y menos a la direccionalidad de esas entradas. Es decir, los ambisónicos de orden inferior asociados a un orden de cero o uno corresponden a funciones de base esférica que, como se ilustra en la FIG. 1 y la FIG. 2, no proporcionan mucho en términos de la dirección de la onda de presión, sino que proporcionan algo de volumen (que es representativo de la energía).

[0651] Las operaciones descritas en el ejemplo anterior también se pueden expresar de acuerdo con el siguiente pseudocódigo. El pseudocódigo siguiente incluye anotaciones, en forma de declaraciones de comentarios que se incluyen en las instancias consecutivas de las cadenas de caracteres "/*" y "*/" (sin comillas).

```
[U,S,V] = svd(audioframe,'ecom');
VS = V*S;
/* La línea siguiente se dirige a analizar cada fila independientemente y a sumar los valores de la primera fila (por
ejemplo) de la quinta entrada a la entrada vigésimo quinta para determinar un cociente de direccionalidad o métrica
de direccionalidad para un vector correspondiente. Elevar al cuadrado las entradas antes de sumar. Las entradas en
cada fila que están asociadas con un orden mayor que 1 están asociadas con ambisónicos de orden superior y, por
lo tanto, tienen más probabilidades de ser direccionales. */
SumVS = sum(VS(5:end,:).^2,1);
/* La línea siguiente se dirige a ordenar la suma de los cuadrados de la matriz VS generada, y a seleccionar un
conjunto de los valores más grandes (por ejemplo, tres o cuatro de los valores más grandes) */
[~,idxVS] = sort(sumVS,'descend');
U = U(:,idxVS);
V = V(:,idxVS);
S = S(idxVS,idxVS);
```

[0652] En otras palabras, de acuerdo con el pseudocódigo anterior, la unidad de análisis de componentes relevantes 524G puede seleccionar entradas de cada vector de la matriz VS descompuesta de las de los SHC 511 correspondientes a una función de base esférica que tiene un orden mayor que uno. La unidad de análisis de componentes relevantes 524G puede entonces elevar al cuadrado estas entradas para cada vector de la matriz VS, sumando las entradas elevadas al cuadrado para identificar, calcular o, de algún modo, determinar una métrica o cociente de direccionalidad para cada vector de la matriz VS. A continuación, la unidad de análisis de componentes relevantes 524G puede ordenar los vectores de la matriz VS en base a las respectivas métricas de direccionalidad de cada uno de los vectores. La unidad de análisis de componentes relevantes 524G puede clasificar estos vectores en un orden descendente de métricas de direccionalidad, de tal manera que los vectores con la direccionalidad correspondiente más alta son los primeros y aquellos vectores con la direccionalidad correspondiente más baja son últimos. La unidad de análisis de componentes relevantes 524G puede entonces seleccionar el subconjunto diferente de cero de los vectores que tienen la métrica de direccionalidad relativa más alta.

[0653] De acuerdo con algunos aspectos de las técnicas descritas en el presente documento, el dispositivo de codificación de audio 510G, o uno o más de sus componentes, puede identificar o utilizar de otra manera un número predeterminado de los vectores de la matriz VS como componentes de audio distintos. Por ejemplo, después de seleccionar las entradas 5 a 25 de cada fila de la matriz VS y de elevar al cuadrado y sumar las entradas seleccionadas para determinar la métrica de direccionalidad relativa para cada vector respectivo, la unidad de análisis de componentes relevantes 524G puede implementar una selección adicional entre los vectores para identificar vectores que representan distintos componentes de audio. En algunos ejemplos, la unidad de análisis de

componentes relevantes 524G puede seleccionar un número predeterminado de los vectores de la matriz VS, comparando los cocientes de direccionalidad de los vectores. Como ejemplo, la unidad de análisis de componentes relevantes 524G puede seleccionar los cuatro vectores representados en la matriz VS que tienen los cuatro cocientes de direccionalidad más elevados (y que son los primeros cuatro vectores de la matriz VS ordenada). A su vez, la unidad de análisis de componentes relevantes 524G puede determinar que los cuatro vectores seleccionados representan los cuatro objetos de audio más distintos asociados con la representación SHC correspondiente del campo sonoro.

[0654] En algunos ejemplos, la unidad de análisis de componentes relevantes 524G puede reordenar los vectores derivados de la matriz VS, para reflejar la distinción de los cuatro vectores seleccionados, como se ha descrito anteriormente. En un ejemplo, la unidad de análisis de componentes relevantes 524G puede reordenar los vectores de tal manera que las cuatro entradas seleccionadas se reubican en la parte superior de la matriz VS. Por ejemplo, la unidad de análisis de componentes relevantes 524G puede modificar la matriz VS de modo que las cuatro entradas seleccionadas estén situadas en una primera fila (o superior) de la matriz VS reordenada resultante. Aunque se describe en el presente documento con respecto a la unidad de análisis de componentes relevantes 524G, en diversas implementaciones, otros componentes del dispositivo de codificación de audio 510G, tal como la unidad de reordenación de vectores 532, pueden realizar la reordenación.

[0655] La unidad de análisis de componentes relevantes 524G puede comunicar la matriz resultante (es decir, la matriz VS, reordenada o no, según sea el caso) a la unidad de generación de flujo de bits 516. A su vez, la unidad de generación de flujo de bits 516 puede usar la matriz VS 525K para generar el flujo de bits 517. Por ejemplo, si la unidad de análisis de componentes relevantes 524G ha reordenado la matriz VS 525K, la unidad de generación de flujo de bits 516 puede usar la fila superior de la versión reordenada de matriz VS 525K como objetos de audio distintos, por ejemplo cuantificando o descartando los vectores restantes de la versión reordenada de la matriz VS 525K. Mediante la cuantización de los vectores restantes de la versión reordenada de la matriz VS 525K, la unidad de generación de flujo de bits 16 puede tratar los vectores restantes como datos de audio ambiental o de fondo.

[0656] En los ejemplos en los que la unidad de análisis de componentes relevantes 524G no ha reordenado la matriz VS 525K, la unidad de generación de flujo de bits 516 puede distinguir datos de audio distintos de datos de audio de fondo, basados en entradas particulares (por ejemplo, las entradas 5ª a 25ª) de cada fila de la matriz VS 525K, según se seleccionan por la unidad de análisis de componentes relevantes 524G. Por ejemplo, la unidad de generación de flujo de bits 516 puede generar el flujo de bits 517 cuantificando o descartando las primeras cuatro entradas de cada fila de la matriz VS 525K.

[0657] De esta manera, el dispositivo de codificación de audio 510G y/o sus componentes, tal como la unidad de análisis de componentes relevantes 524G, pueden implementar técnicas de esta divulgación para determinar o utilizar de otra manera las relaciones de las energías de coeficientes mayores e inferiores de datos de audio, con el fin de distinguir entre distintos objetos de audio y datos de audio de fondo representativos del campo sonoro. Por ejemplo, como se describe, la unidad de análisis de componentes relevantes 524G puede utilizar las relaciones de energía basadas en valores de las diversas entradas de la matriz VS 525K generadas por la unidad de análisis de componentes relevantes 524H. Mediante la combinación de datos proporcionados por la matriz V 519A y la matriz S 519B, la unidad de análisis de componentes relevantes 524G puede generar la matriz VS 525K para proporcionar información sobre la direccionalidad y la energía global de los diversos componentes de los datos de audio, en la forma de vectores y datos relacionados (por ejemplo, cocientes de direccionalidad). Más específicamente, la matriz V 519A puede proporcionar información relacionada con determinaciones de direccionalidad, mientras que la matriz S 519B puede proporcionar información relacionada con determinaciones energéticas globales para los componentes de los datos de audio.

[0658] En otros ejemplos, la unidad de análisis de componentes relevantes 524G puede generar la matriz VS 525K usando los vectores V^{T}_{DIST} reordenados 539. En estos ejemplos, la unidad de análisis de componentes relevantes 524G puede determinar la distinción basada en la matriz V 519, antes de cualquier modificación basada en la matriz S 519B. En otras palabras, de acuerdo con estos ejemplos, la unidad de análisis de componentes relevantes 524G puede determinar la direccionalidad usando sólo la matriz V 519, sin realizar la etapa de generación de la matriz VS 525K. Más específicamente, la matriz V 519A puede proporcionar información sobre la forma en que se mezclan componentes (por ejemplo, vectores de la matriz V 519) de los datos de audio y, potencialmente, información sobre diversos efectos sinérgicos de los datos transportados por los vectores. Por ejemplo, la matriz V 519A puede proporcionar información sobre la "dirección de llegada" de diversos componentes de audio representados por los vectores, tales como la dirección de llegada de cada componente de audio, según se transmiten al dispositivo de codificación de audio 510G por un EigenMike®. Como se utiliza en el presente documento, la expresión "componente de datos de audio" se puede usar indistintamente con "entrada" de cualquiera de las matrices 519 o cualquier matriz derivada de las mismas.

[0659] De acuerdo con algunas implementaciones de las técnicas de esta descripción, la unidad de análisis de componentes relevantes 524G puede complementar o aumentar las representaciones de SHC con información extraña para realizar diversas determinaciones descritas en el presente documento. Como un ejemplo, la unidad de análisis de componentes relevantes 524G puede aumentar el SHC con información extraña con el fin de determinar

la saliencia de varios componentes de audio representados en las matrices 519-519C. Como otro ejemplo, la unidad de análisis de componentes relevantes 524G y/o la unidad de reordenación de vectores 532 pueden aumentar el HOA con datos extraños para distinguir entre objetos de audio distintos y datos de audio de fondo.

[0660] En algunos ejemplos, la unidad de análisis de componentes relevantes 524G puede detectar que partes (por ejemplo, objetos de audio distintos) de los datos de audio muestran energía keynesiana. Un ejemplo de tales objetos distintos puede estar asociado con una voz humana que modula. En el caso de los datos de audio basados en voz que modula, la unidad de análisis de componentes relevantes 524G puede determinar que la energía de los datos de modulación, como relación con respecto a las energías de los componentes restantes, permanece aproximadamente constante (por ejemplo, constante dentro de un intervalo umbral), o aproximadamente estacionaria en el tiempo. Tradicionalmente, si las características de energía de distintos componentes de audio con energía keynesiana (por ejemplo, los asociados con la voz de modulación) cambian de una trama de audio a otra, un dispositivo puede no ser capaz de identificar la serie de componentes de audio como una sola señal. Sin embargo, la unidad de análisis de componentes relevantes 524G puede implementar técnicas de esta divulgación para determinar una direccionalidad o una apertura del objeto de distancia representado como un vector en las diversas matrices.

[0661] Más específicamente, la unidad de análisis de componentes relevantes 524G puede determinar que es improbable que las características tales como la direccionalidad y/o la apertura cambien sustancialmente a través de las tramas de audio. Como se utiliza en el presente documento, la apertura representa una relación entre los coeficientes de orden superior y los coeficientes de orden inferior, dentro de los datos de audio. Cada fila de la matriz V 519A puede incluir vectores que corresponden a SHC particulares. La unidad de análisis de componentes relevantes 524G puede determinar que los SHC de orden inferior (por ejemplo, asociado con un orden menor o igual a 1) tienden a representar datos ambientales, mientras que las entradas de orden superior tienden a representar datos distintos. Además, la unidad de análisis de componentes relevantes 524G puede determinar que, en muchos casos, los SHC de orden superior (por ejemplo, asociados con un orden mayor de 1) muestran una mayor energía, y que la relación de energía de los SHC de orden superior con respecto a orden inferior permanece sustancialmente similar (o aproximadamente constante) de trama de audio a trama de audio.

[0662] Uno o más componentes de la unidad de análisis de componentes relevantes 524G pueden determinar características de los datos de audio tales como direccionalidad y apertura, utilizando la matriz V 519. De esta manera, los componentes del dispositivo de codificación de audio 510G, tal como la unidad de análisis de componentes relevantes 524G, pueden implementar las técnicas descritas en el presente documento para determinar saliencia y/o distinguir objetos de audio distintos del audio de fondo, utilizando información basada en direccionalidad. Mediante la utilización de la direccionalidad para determinar la saliencia y/o la distinción, la unidad de análisis de componentes relevantes 524G puede llegar a determinaciones más robustas que en los casos de un dispositivo configurado para determinar la saliencia y/o distinción usando sólo datos basados en energía. Aunque se ha descrito anteriormente con respecto a las determinaciones basadas en la direccionalidad de la saliencia y/o distinción, la unidad de análisis de componentes relevantes 524G puede implementar las técnicas de esta divulgación para usar la direccionalidad además de otras características, tales como la energía, para determinar la saliencia y/o la distinción de componentes particulares de los datos de audio, según se representa por vectores de una o más de las matrices 519-519C (o cualquier matriz derivada de las mismas).

[0663] En algunos casos, un procedimiento incluye identificar uno o más objetos de audio distintos de uno o más coeficientes armónicos esféricos (SHC) asociados con los objetos de audio basados en una direccionalidad determinada para uno o más de los objetos de audio. En un ejemplo, el procedimiento incluye además determinar la direccionalidad de uno o más objetos de audio basándose en los coeficientes armónicos esféricos asociados con los objetos de audio. En algunos ejemplos, el procedimiento incluye además realizar una descomposición en valores singulares con respecto a los coeficientes armónicos esféricos para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos; y representar la pluralidad de coeficientes armónicos esféricos en función de al menos una parte de una o más de la matriz U, la matriz S y la matriz V, en el que la determinación de la direccionalidad respectiva de uno o más objetos de audio se basa, al menos en parte, en la matriz V.

[0664] En un ejemplo, el procedimiento incluye además reordenar uno o más vectores de la matriz V de modo que los vectores que tienen un mayor cociente de direccionalidad se posicionan por encima de los vectores que tienen un cociente de direccionalidad menor en la matriz V reordenada. En un ejemplo, el procedimiento incluye además determinar que los vectores que tienen el mayor cociente de direccionalidad incluyen mayor información direccional que los vectores que tienen el cociente de direccionalidad menor. En un ejemplo, el procedimiento incluye además multiplicar la matriz V por la matriz S para generar una matriz VS, incluyendo la matriz VS uno o más vectores. En un ejemplo, el procedimiento incluye además seleccionar entradas de cada fila de la matriz VS que están asociadas con un orden mayor de 1, elevar al cuadrado cada una de las entradas seleccionadas para formar entradas cuadráticas correspondientes, y para cada fila de la matriz VS, sumar todas las entradas al cuadrado para determinar un cociente de direccionalidad para un vector correspondiente.

[0665] En algunos ejemplos, cada fila de la matriz VS incluye 25 entradas. En un ejemplo, seleccionar las entradas de cada fila de la matriz VS asociada con el orden mayor de 1 incluye seleccionar todas las entradas que empiezan en una 5ª entrada de cada fila de la matriz VS y que terminan en una 25ª entrada de cada fila de la VS. En un ejemplo, el procedimiento incluye además seleccionar un subconjunto de los vectores de la matriz VS para representar los distintos objetos de audio. En algunos ejemplos, seleccionar el subconjunto incluye seleccionar cuatro vectores de la matriz VS, y los cuatro vectores seleccionados tienen los cuatro cocientes de direccionalidad más grandes de todos los vectores de la matriz VS. En un ejemplo, la determinación de que el subconjunto seleccionado de los vectores representan los objetos de audio distintos se basa tanto en la direccionalidad como en una energía de cada vector.

[0666] En algunos casos, un procedimiento incluye identificar uno o más objetos de audio distintos de uno o más coeficientes armónicos esféricos asociados con los objetos de audio basados en una direccionalidad y una energía determinadas para uno o más de los objetos de audio. En un ejemplo, el procedimiento incluye además determinar una o ambas de la direccionalidad y la energía de uno o más objetos de audio basándose en los coeficientes armónicos esféricos asociados con los objetos de audio. En algunos ejemplos, el procedimiento incluye además realizar una descomposición en valores singulares con respecto a los coeficientes armónicos esféricos representativos del campo sonoro para generar una matriz U representativa de vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y representar la pluralidad de coeficientes armónicos esféricos en función de al menos una parte de una o más de la matriz U, la matriz S y la matriz V, en el que la determinación de la direccionalidad respectiva de uno o más objetos de audio se basa, al menos en parte, en la matriz V, y en el que la determinación de la energía respectiva del uno o más objetos de audio se basa, al menos en parte, en la matriz S.

[0667] En un ejemplo, el procedimiento incluye además multiplicar la matriz V por la matriz S para generar una matriz VS, incluyendo la matriz VS uno o más vectores. En algunos ejemplos, el procedimiento incluye además seleccionar entradas de cada fila de la matriz VS que están asociadas con un orden mayor de 1, elevar al cuadrado cada una de las entradas seleccionadas para formar entradas cuadráticas correspondientes, y para cada fila de la matriz VS, sumar todas las entradas al cuadrado para generar un cociente de direccionalidad para un vector correspondiente de la matriz VS. En algunos ejemplos, cada fila de la matriz VS incluye 25 entradas. En un ejemplo, seleccionar las entradas de cada fila de la matriz VS asociada con el orden mayor de 1 comprende seleccionar todas las entradas que empiezan en una 5ª entrada de cada fila de la matriz VS y que terminan en una 25ª entrada de cada fila de la VS. En algunos ejemplos, el procedimiento incluye además seleccionar un subconjunto de los vectores para representar objetos de audio distintos. En un ejemplo, seleccionar el subconjunto comprende seleccionar cuatro vectores de la matriz VS, y los cuatro vectores seleccionados tienen los cuatro cocientes de direccionalidad más grandes de todos los vectores de la matriz VS. En algunos ejemplos, la determinación de que el subconjunto seleccionado de los vectores representan los objetos de audio distintos se basa tanto en la direccionalidad como en una energía de cada vector.

[0668] En algunos casos, un procedimiento incluye determinar, usando información basada en la direccionalidad, uno o más primeros vectores que describen componentes distintos del campo sonoro y uno o más segundos vectores que describen componentes de fondo del campo sonoro, tanto uno como los más primeros vectores y uno o los más segundos vectores generados al menos realizando una transformación con respecto a la pluralidad de coeficientes armónicos esféricos. En un ejemplo, la transformación comprende una descomposición en valores singulares que genera una matriz U representativa de los vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos. En un ejemplo, la transformación comprende un análisis de componentes principales para identificar los distintos componentes del campo sonoro y los componentes de fondo del campo sonoro.

[0669] En algunos ejemplos, un dispositivo está configurado o es operativo de otro modo, para llevar a cabo cualquiera de las técnicas descritas en el presente documento o cualquier combinación de las técnicas. En algunos ejemplos, se codifica un medio de almacenamiento legible por ordenador con instrucciones que, cuando se ejecutan, hacen que uno o más procesadores realicen cualquiera de las técnicas descritas en el presente documento o cualquier combinación de las técnicas. En algunos ejemplos, un dispositivo incluye medios para llevar a cabo cualquiera de las técnicas descritas en el presente documento o cualquier combinación de las técnicas.

[0670] Es decir, los aspectos anteriores de las técnicas pueden permitir que el dispositivo de codificación de audio 510G sea configurado para funcionar de acuerdo con las siguientes cláusulas.

Cláusula 134954-1B. Un dispositivo, tal como el dispositivo de codificación de audio 510G, que comprende: uno o más procesadores configurados para identificar uno o más objetos de audio distintos a partir de uno o más coeficientes armónicos esféricos asociados con los objetos de audio, basados en una direccionalidad y una energía determinada para uno o más de los objetos de audio.

Cláusula 134954-2B. El dispositivo de la cláusula 134954-1B, en el que el uno o más procesadores están configurados adicionalmente para determinar una o ambas de la direccionalidad y la energía de uno o más objetos de audio basándose en los coeficientes armónicos esféricos asociados con los objetos de audio.

Cláusula 134954-3B. El dispositivo de cualquiera de las reivindicaciones 1B o 2B o combinación de las mismas, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a los coeficientes armónicos esféricos representativos del campo sonoro para generar una matriz U representativa de los vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la pluralidad de coeficientes armónicos esféricos, y una matriz V representativa de vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos, y representar la pluralidad de coeficientes armónicos esféricos en función de al menos una parte de una o más de la matriz U, la matriz S y la matriz V, en el que el uno o más procesadores están configurados para determinar la direccionalidad respectiva de uno o más objetos de audio basados,al menos en parte, en la matriz V, y en el que uno o más procesadores están configurados para determinar la energía respectiva de uno o más objetos de audio que se basa, al menos en parte, en la matriz S.

Cláusula 134954-4B. El dispositivo de la cláusula 134954-3B, en el que el uno o más procesadores están configurados adicionalmente para multiplicar la matriz V por la matriz S para generar una matriz VS, incluyendo la matriz VS uno o más vectores.

Cláusula 134954-5B. El dispositivo de la cláusula 134954-4B, en el que el uno o más procesadores están configurados además para seleccionar entradas de cada fila de la matriz VS que están asociadas con un orden mayor de 1, elevar al cuadrado cada una de las entradas seleccionadas para formar entradas cuadradas correspondientes, y para cada fila de la matriz VS, sumar todas las entradas al cuadrado para generar un cociente de direccionalidad para un vector correspondiente de la matriz VS.

Cláusula 134954-6B. El dispositivo de cualquiera de las reivindicaciones 5B y 6B o una combinación de las mismas, en el que cada fila de la matriz VS incluye 25 entradas.

Cláusula 134954-7B. El dispositivo de la cláusula 134954-6B, en el que el uno o más procesadores están configurados para seleccionar todas las entradas que comienzan en una quinta entrada de cada fila de la matriz VS y que terminan en una 25ª entrada de cada fila de la matriz VS.

Cláusula 134954-8B. El dispositivo de cualquiera de las cláusulas 134954-6B y la cláusula 134954-7B, o una combinación de las mismas, en el que el uno o más procesadores están configurados además para seleccionar un subconjunto de los vectores para representar objetos de audio distintos.

Cláusula 134954-9B. El dispositivo de la cláusula 134954-8B, en el que el uno o más procesadores están configurados para seleccionar cuatro vectores de la matriz VS, y en el que los cuatro vectores seleccionados tienen los cuatro cocientes de direccionalidad más grandes de todos los vectores de la matriz VS.

Cláusula 134954-10B. El dispositivo de cualquiera de la cláusula 134954-8B y la cláusula 134954-9B o combinación de las mismas, en el que el uno o más procesadores están configurados adicionalmente para determinar que el subconjunto seleccionado de los vectores representan los distintos objetos de audio que se basa tanto en la direccionalidad como en una energía de cada vector.

Cláusula 134954-1C. Un dispositivo, tal como el dispositivo de codificación de audio 510G, que comprende: uno o más procesadores configurados para determinar, usando información basada en direccionalidad, uno o más primeros vectores que describen componentes distintos del campo sonoro y uno o más segundos vectores que describen componentes de fondo del campo sonido, tanto el uno o más primeros vectores y uno o más segundos vectores generados al menos realizando una transformación con respecto a la pluralidad de coeficientes armónicos esféricos.

Cláusula 134954-2C. El procedimiento de la cláusula 134954-1C, en el que la transformación comprende una descomposición en valores singulares que genera una matriz U representativa de los vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos.

Cláusula 134954-3C. El procedimiento de la cláusula 134954-2C, que comprende además las operaciones enumeradas por cualquier combinación de la cláusula 134954-1A a la cláusula 134954-12A y la cláusula 134954-1B a la cláusula 134954-9B.

Cláusula 134954-4C. El procedimiento de la cláusula 134954-1C, en el que la transformación comprende un análisis de componentes principales para identificar los distintos componentes del campo sonoro y los

componentes de fondo del campo sonoro.

[0671] La FIG. 40H es un diagrama de bloques que ilustra un dispositivo de codificación de audio ejemplar 510H que puede realizar diversos aspectos de las técnicas descritas en esta divulgación para comprimir los coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales. El dispositivo de codificación de audio 510H puede ser similar al dispositivo de codificación de audio 510G en que el dispositivo de codificación de audio 510H incluye una unidad de compresión de audio 512, una unidad de codificación de audio 514 y una unidad de generación de flujo de bits 516. Además, la unidad de compresión de audio 512 del dispositivo de codificación de audio 510H puede ser similar a la del dispositivo de codificación de audio 510C porque la unidad de compresión de audio 512 incluye una unidad de descomposición 518 y una unidad de extracción de componentes del campo sonoro 520G, que pueden operar de forma similar a las unidades del dispositivo de codificación de audio 510G. En algunos ejemplos, el dispositivo de codificación de audio 510H puede incluir una unidad de cuantización 534, como se ha descrito con respecto a las FIGS. 40D-40E, para cuantizar uno o más vectores de cualquiera de los vectores U_{DIST} 525C, los vectores U_{BG} 525D, los vectores V^T_{DIST} 525E y los vectores V^T_{BG} 525J.

[0672] La unidad de compresión de audio 512 del dispositivo de codificación de audio 510H puede, sin embargo, diferir de la unidad de compresión de audio 512 del dispositivo de codificación de audio 510G en que la unidad de compresión de audio 512 del dispositivo de codificación de audio 510H incluye una unidad adicional denominada como unidad de interpolación 550. La unidad de interpolación 550 puede representar una unidad que interpola subtramas de una primera trama de audio de las subtramas de la primera trama de audio y una segunda trama de audio temporalmente posterior o anterior, como se describe con más detalle a continuación con respecto a las FIGS. 45 y 45B. La unidad de interpolación 550 puede reducir, al realizar esta interpolación, la complejidad computacional (en términos de ciclos de procesamiento y/o consumo de memoria) reduciendo potencialmente la medida en que la unidad de descomposición 518 se requiere para descomponer SHC 511. A este respecto, la unidad de interpolación 550 puede realizar operaciones similares a las descritas anteriormente con respecto a la unidad de interpolación espacio-temporal 50 del dispositivo de codificación de audio 24 mostrado en el ejemplo de la FIG. 4.

[0673] Es decir, la descomposición en valores singulares realizada por la unidad de descomposición 518 es potencialmente muy intensiva en el procesador y/o la memoria, mientras que en algunos ejemplos, toma cantidades extensas de tiempo para descomponer los SHC 511, especialmente conforme aumenta el orden de los SHC 511. Con el fin de reducir la cantidad de tiempo y hacer que la compresión de los SHC 511 sea más eficiente (en términos de ciclos de procesamiento y/o consumo de memoria), las técnicas descritas en esta divulgación pueden proporcionar la interpolación de una o más subtramas de la primera trama de audio, donde cada una de las subtramas puede representar versiones descompuestas de los SHC 511. En lugar de realizar la SVD con respecto a toda la trama, las técnicas pueden permitir que la unidad de descomposición 518 descomponga una primera subtrama de una primera trama de audio, generando una matriz V 519'.

[0674] La unidad de descomposición 518 también puede descomponer una segunda subtrama de una segunda trama de audio, en la que esta segunda trama de audio puede ser temporalmente posterior o temporalmente anterior a la primera trama de audio. La unidad de descomposición 518 puede emitir una matriz V 519' para esta subtrama de la segunda trama de audio. La unidad de interpolación 550 puede entonces interpolar las subtramas restantes de la primera trama de audio basándose en las matrices V 519' descompuestas de las primera y segunda subtramas, dando la matriz V 519, la matriz S 519B y la matriz U 519C, donde las descomposiciones para las subtramas restantes se pueden calcular basándose en los SHC 511, la matriz V 519A para la primera trama de audio y las matrices V 519 interpoladas para las subtramas restantes de la primera trama de audio. Por lo tanto, la interpolación puede evitar el cálculo de las descomposiciones para las subtramas restantes de la primera trama de audio.

[0675] Además, como se ha indicado anteriormente, la matriz U 519C puede no ser continua de una trama a otra, donde pueden especificarse componentes distintos de la matriz U 519C descompuestos de una primera trama de audio de los SHC 511 en filas y/o columnas diferentes que en la matriz U 519C descompuesta a partir de una segunda trama de audio de los SHC 511. Mediante la realización de esta interpolación, la discontinuidad puede reducirse dado que la interpolación lineal puede tener un efecto suavizante que puede reducir los artefactos introducidos debido a los límites de la trama (o, en otras palabras, la segmentación de los SHC 511 en tramas). El uso de la matriz V 519' para realizar esta interpolación y después recuperar las matrices U 519C basadas en la matriz V 519' interpolada de los SHC 511 puede suavizar cualquier efecto de reorganizar la matriz U 519C.

[0676] En funcionamiento, la unidad de interpolación 550 puede interpolar una o más subtramas de una primera trama de audio a partir de una primera descomposición, por ejemplo, la matriz V 519', de una parte de una primera pluralidad de coeficientes armónicos esféricos 511 incluidos en la primera trama y una segunda descomposición, por ejemplo, la matriz V 519', de una porción de una segunda pluralidad de coeficientes armónicos esféricos 511 incluidos en una segunda trama para generar coeficientes armónicos esféricos interpolados descompuestos para la una o más subtramas.

[0677] En algunos ejemplos, la primera descomposición comprende la primera matriz V 519' representativa de vectores singulares derechos de la porción de la primera pluralidad de coeficientes armónicos esféricos 511. De manera similar, en algunos ejemplos, la segunda descomposición comprende la segunda matriz V 519'

representativa de vectores singulares derechos de la porción de la segunda pluralidad de coeficientes armónicos esféricos.

[0678] La unidad de interpolación 550 puede realizar una interpolación temporal con respecto a una o más subtramas basadas en la primera matriz V 519' y la segunda matriz V 519'. Es decir, la unidad de interpolación 550 puede interpolar temporalmente, por ejemplo, la segunda, tercera y cuarta subtramas de cuatro subtramas totales para la primera trama de audio basadas en una matriz V 519' descompuesta a partir de la primera subtrama de la primera trama de audio y la matriz V 519' descompuesta a partir de la primera subtrama de la segunda trama de audio. En algunos ejemplos, esta interpolación temporal es una interpolación temporal lineal, en la que la matriz V 519' descompuesta a partir de la primera subtrama de la primera trama de audio se pondera más pesadamente cuando se interpola la segunda subtrama de la primera trama de audio que cuando se interpola la cuarta subtrama de la primera trama de audio. Cuando se interpola la tercera subtrama, las matrices V 519' pueden ponderarse uniformemente. Cuando se interpola la cuarta subtrama, la matriz V 519' descompuesta a partir de la primera subtrama de la segunda trama de audio puede ponderarse con mayor peso que la matriz V 519' descompuesta de la primera subtrama de la primera trama de audio.

[0679] En otras palabras, la interpolación temporal lineal puede ponderar las matrices V 519' dada la proximidad de una de las subtramas de la primera trama de audio a interpolar. Para que la segunda subtrama sea interpolada, la matriz V 519' descompuesta a partir de la primera subtrama de la primera trama de audio se pondera en mayor medida dada su proximidad a la segunda subtrama a interpolar que la matriz V 519' descompuesta a partir de la primera subtrama de la segunda trama de audio. Los pesos pueden ser equivalentes por este motivo al interpolar la tercera subtrama basada en las matrices V 519'. El peso aplicado a la matriz V 519' descompuesta a partir de la primera subtrama de la segunda trama de audio puede ser mayor que el aplicado a la matriz V 519' descompuesta a partir de la primera subtrama de la primera trama de audio dado que la cuarta subtrama a interpolar está más próxima a la primera subtrama de la segunda trama de audio que la primera subtrama de la primera trama de audio.

[0680] Aunque, en algunos ejemplos, sólo se utiliza una primera subtrama de cada trama de audio para realizar la interpolación, la parte de la primera pluralidad de coeficientes armónicos esféricos puede comprender dos de cuatro subtramas de la primera pluralidad de coeficientes armónicos esféricos 511. En estos y otros ejemplos, la porción de la segunda pluralidad de coeficientes armónicos esféricos 511 comprende dos de cuatro subtramas de la segunda pluralidad de coeficientes armónicos esféricos 511.

[0681] Como se ha indicado anteriormente, un dispositivo individual, por ejemplo, el dispositivo de codificación de audio 510H, puede realizar la interpolación, al tiempo que también descompone la parte de la primera pluralidad de coeficientes armónicos esféricos para generar las primeras descomposiciones de la porción de la primera pluralidad de coeficientes armónicos esféricos. En estos y otros ejemplos, la unidad de descomposición 518 puede descomponer la porción de la segunda pluralidad de coeficientes armónicos esféricos para generar las segundas descomposiciones de la porción de la segunda pluralidad de coeficientes armónicos esféricos. Aunque se describen con respecto a un único dispositivo, dos o más dispositivos pueden realizar las técnicas descritas en esta divulgación, donde uno de los dos dispositivos realiza la descomposición y otro de los dispositivos realiza la interpolación de acuerdo con las técnicas descritas en esta divulgación.

[0682] En otras palabras, el audio 3D basado en armónicos esféricos puede ser una representación paramétrica del campo de presión 3D en términos de funciones de base ortogonales sobre una esfera. Cuanto mayor sea el orden N de la representación, la resolución espacial será potencialmente más alta 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. Estas 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 SVD puede descomponer cada trama de coeficientes en tres matrices U, S y V. En algunos ejemplos, las técnicas pueden manejar algunos de los vectores en U como componentes direccionales del campo sonoro subyacente. Sin embargo, cuando se manejan de esta manera, estos vectores (en U) son discontinuos de trama en trama, aunque representan el mismo componente de audio distinto. Estas discontinuidades pueden conducir a artefactos significativos cuando los componentes son alimentados a través de codificadores de audio de transformadas.

[0683] Las técnicas descritas en esta divulgación pueden abordar esta discontinuidad. Es decir, las técnicas pueden basarse en la observación de que la matriz V puede interpretarse como ejes espaciales ortogonales en el dominio de las armónicas esféricas. La matriz U puede representar una proyección de los datos de armónicos esféricos (HOA) en términos de las funciones de base, donde la discontinuidad puede atribuirse a las funciones de base (V) que cambian cada trama y, por lo tanto, son discontinuas en sí mismas. Esto es diferente a la descomposición similar, como la Transformada 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 por coincidencia. Las técnicas descritas en esta divulgación pueden permitir que la unidad de interpolación 550 mantenga la continuidad entre las funciones de base (V) de una trama a otra mediante interpolación entre ellas.

[0684] En algunos ejemplos, las técnicas permiten a la unidad de interpolación 550 dividir la trama de datos SH en cuatro subtramas, como se describió anteriormente y se describe adicionalmente a continuación con respecto a las FIGS. 45 y 45B. La unidad de interpolación 550 puede entonces calcular la SVD para la primera subtrama. De forma similar, se calcula la SVD para la primera subtrama de la segunda trama. Para cada una de la primera trama y la segunda trama, la unidad de interpolación 550 puede convertir los vectores en V en un mapa espacial proyectando los vectores sobre una esfera (utilizando una matriz de proyección tal como una matriz de diseño en T). La unidad de interpolación 550 puede entonces interpretar los vectores en V como formas en una esfera. Para interpolar las matrices V para las tres subtramas entre la primera subtrama de la primera trama, la primera subtrama de la trama siguiente, la unidad de interpolación 550 puede interpolar entonces estas formas espaciales y luego transformarlas de nuevo en los vectores SH a través de la inversa de la matriz de proyección. Las técnicas de esta descripción pueden, de esta manera, proporcionar una transición suave entre las matrices V.

[0685] De esta manera, el dispositivo de codificación de audio 540H puede configurarse para realizar diversos aspectos de las técnicas expuestas a continuación con respecto a las siguientes cláusulas.

Cláusula 135054-1A. Un dispositivo, tal como el dispositivo de codificación de audio 510H, que comprende: uno o más procesadores configurados para interpolar una o más subtramas de una primera trama de una primera descomposición de una porción de una primera pluralidad de coeficientes armónicos esféricos incluidos en la primera trama y una segunda descomposición de una porción de una segunda pluralidad de coeficientes armónicos esféricos incluidos en una segunda trama para generar coeficientes armónicos esféricos interpolados descompuestos para la una o más subtramas.

Cláusula 135054-2A. El dispositivo de la cláusula 135054-1A, en el que la primera descomposición comprende una primera matriz V representativa de los vectores singulares derechos de porción de la primera pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-3A. El dispositivo de la cláusula 135054-1A, en el que la segunda descomposición comprende una segunda matriz V representativa de vectores singulares derechos de la porción de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-4A. El dispositivo de la cláusula 135054-1A, en el que la primera descomposición comprende una primera matriz V representativa de vectores singulares derechos de la porción de la primera pluralidad de coeficientes armónicos esféricos y en la que la segunda descomposición comprende una segunda matriz V representativa de vectores singulares derechos de la porción de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-5A. El dispositivo de la cláusula 135054-1A, en el que el uno o más procesadores están configurados adicionalmente para, al interpolar la una o más subtramas, interpolar temporalmente la una o más subtramas basándose en la primera descomposición y la segunda descomposición.

Cláusula 135054-6A. El dispositivo de la cláusula 135054-1A, en el que el uno o más procesadores están configurados adicionalmente para, al interpolar la una o más subtramas, proyectar la primera descomposición en un dominio espacial para generar las primeras descomposiciones proyectadas, proyectar la segunda descomposición en el dominio espacial para generar segundas descomposiciones proyectadas, interpolar espacialmente las primeras descomposiciones proyectadas y las segundas descomposiciones proyectadas para generar una primera descomposición proyectada espacialmente interpolada y una segunda descomposición proyectada interpolada espacialmente, e interpolar temporalmente la una o más subtramas basándose en la primera descomposición proyectada espacialmente interpolada y la segunda descomposición proyectada espacialmente interpolada.

Cláusula 135054-7A. El dispositivo de la cláusula 135054-6A, en el que el uno o más procesadores están configurados además para proyectar los coeficientes armónicos esféricos interpolados temporalmente resultantes de la interpolación de uno o más subtramas de vuelta a un dominio armónico esférico.

Cláusula 135054-8A. El dispositivo de la cláusula 135054-1A, en el que la porción de la primera pluralidad de coeficientes armónicos esféricos comprende una única subtrama de la primera pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-9A. El dispositivo de la cláusula 135054-1A, en el que la porción de la segunda pluralidad de coeficientes armónicos esféricos comprende una única subtrama de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-10A. El dispositivo de la cláusula 135054-1A, en el que la primera trama está dividida en cuatro subtramas, y
en el que la porción de la primera pluralidad de coeficientes armónicos esféricos comprende solamente la primera subtrama de la primera pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-11A. El dispositivo de la cláusula 135054-1A, en el que la segunda trama está dividida en cuatro subtramas, y
en el que la porción de la segunda pluralidad de coeficientes armónicos esféricos comprende solamente la primera subtrama de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-12A. El dispositivo de la cláusula 135054-1A, en el que la porción de la primera pluralidad de coeficientes armónicos esféricos comprende dos de cuatro subtramas de la primera pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-13A. El dispositivo de la cláusula 135054-1A, en el que la porción de la segunda pluralidad de coeficientes armónicos esféricos comprende dos de cuatro subtramas de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-14A. El dispositivo de la cláusula 135054-1A, en el que el uno o más procesadores están configurados además para descomponer la porción de la primera pluralidad de coeficientes armónicos esféricos para generar las primeras descomposiciones de la porción de la primera pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-15A. El dispositivo de la cláusula 135054-1A, en el que el uno o más procesadores están configurados además para descomponer la porción de la segunda pluralidad de coeficientes armónicos esféricos para generar las segundas descomposiciones de la porción de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-16A. El dispositivo de la cláusula 135054-1A, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a la porción de la primera pluralidad de coeficientes armónicos esféricos para generar una matriz U representativa de vectores singulares izquierdos de la primera pluralidad de coeficientes armónicos esféricos, una matriz S representativa de valores singulares de la primera pluralidad de coeficientes armónicos esféricos y una matriz V representativa de vectores singulares derechos de la primera pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-17A. El dispositivo de la cláusula 135054-1A, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a la porción de la segunda pluralidad de coeficientes armónicos esféricos para generar una matriz U representativa de los vectores singulares izquierdos de la segunda pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la segunda pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la segunda pluralidad de coeficientes armónicos esféricos.

Cláusula 135054-18A. El dispositivo de la cláusula 135054-1A, en el que la primera y la segunda pluralidad de coeficientes armónicos esféricos representan cada una una representación de onda plana del campo sonoro.

Cláusula 135054-19A. El dispositivo de la cláusula 135054-1A, en el que la primera y la segunda pluralidad de coeficientes armónicos esféricos representan cada uno uno o más objetos monoaudio mezclados entre sí.

Cláusula 135054-20A. El dispositivo de la cláusula 135054-1A, en el que la primera y la segunda pluralidad de coeficientes armónicos esféricos comprenden cada uno un primer y segundo coeficientes armónicos esféricos respectivos que representan un campo sonoro tridimensional.

Cláusula 135054-21A. El dispositivo de la cláusula 135054-1A, en el que la primera y la segunda pluralidad de coeficientes armónicos esféricos están asociadas cada una con al menos una función de base esférica que tiene un orden mayor que uno.

Cláusula 135054-22A. El dispositivo de la cláusula 135054-1A, en el que la primera y la segunda pluralidad de coeficientes armónicos esféricos están asociadas cada una con al menos una función de base esférica que tiene un orden igual a cuatro.

[0686] Aunque se ha descrito anteriormente como realizado por el dispositivo de codificación de audio 510H, los diversos dispositivos de decodificación de audio 24 y 540 también pueden realizar cualquiera de los diversos aspectos de las técnicas expuestas anteriormente con respecto a las cláusulas 135054-1A a 135054-22A.

[0687] La FIG. 40I es un diagrama de bloques que ilustra un dispositivo de codificación de audio ejemplar 510I que puede realizar diversos aspectos de las técnicas descritas en esta divulgación para comprimir los coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales. El dispositivo de codificación de audio 510I puede ser similar al dispositivo de codificación de audio 510H en que el dispositivo de codificación de audio 510I incluye una unidad de compresión de audio 512, una unidad de codificación de audio 514 y una unidad de generación de flujo de bits 516. Además, la unidad de compresión de audio 512 del dispositivo de

codificación de audio 510I puede ser similar a la del dispositivo de codificación de audio 510H porque la unidad de compresión de audio 512 incluye una unidad de descomposición 518 y una unidad de extracción de componentes del campo sonoro 520, que pueden operar de forma similar a las unidades del dispositivo de codificación de audio 510H. En algunos ejemplos, el dispositivo de codificación de audio 10I puede incluir una unidad de cuantización 34, como se ha descrito con respecto a las FIGS. 3D-3E, para cuantizar uno o más vectores de cualquiera de U_{DIST} 25C, U_{BG} 25D, V^T_{DIST} 25E, y V^T_{BG} 25J.

[0688] Sin embargo, aunque tanto la unidad de compresión de audio 512 del dispositivo de codificación de audio 510I como la unidad de compresión de audio 512 del dispositivo de codificación de audio 10H incluyen una unidad de extracción de componente de campo de sonido, la unidad de extracción de componente de campo sonoro 520I del dispositivo de codificación de audio 510I puede incluir una módulo adicional denominado unidad de compresión V 552. La unidad de compresión V 552 puede representar una unidad configurada para comprimir un componente espacial del campo sonoro, es decir, uno o más de los vectores V^T_{DIST} 539 en este ejemplo. Es decir, la descomposición de valores singulares realizada con respecto al SHC puede descomponer el SHC (que es representativo del campo sonoro) en componentes energéticos representados por vectores de la matriz S, componentes temporales representados por la matriz U y componentes espaciales representados por la matriz V. La unidad de compresión V 552 puede realizar operaciones similares a las descritas anteriormente con respecto a la unidad de cuantización 52.

[0689] A efectos de ejemplo, se supone que los vectores V^T_{DIST} 539 comprenden dos vectores de fila que tienen 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^T_{DIST} 539 hasta $(n + 1)^2$, donde n indica el orden de la representación HOA del campo sonoro.

[0690] La unidad de compresión V 552 puede recibir los vectores V^T_{DIST} 539 y realizar un esquema de compresión para generar representaciones de vector V^T_{DIST} comprimidas 539'. Este esquema de compresión puede implicar cualquier esquema de compresión concebible para comprimir elementos de un vector o datos en general, y no debe limitarse al ejemplo descrito más adelante con más detalle.

[0691] La unidad de compresión V 552 puede realizar, como ejemplo, un esquema de compresión que incluye uno o más de representaciones de coma flotante de transformación de cada elemento de los vectores V^T_{DIST} 539 a representaciones enteras de cada elemento de los vectores V^T_{DIST} 539, cuantización uniforme de las representaciones enteras de los vectores V^T_{DIST} 539 y la categorización y codificación de las representaciones enteras cuantificadas de los vectores V^T_{DIST} 539. Varios entre uno o más procesos de este esquema de compresión pueden ser controlados dinámicamente por parámetros para conseguir, o casi conseguir, por ejemplo, una velocidad de bits objetivo para el flujo de bits resultante 517.

[0692] Dado que cada uno de los vectores V^T_{DIST} 539 son ortonormales entre sí, cada uno de los vectores V^T_{DIST} 539 puede codificarse independientemente. En algunos ejemplos, como se describe con más detalle a continuación, cada elemento de cada vector V^T_{DIST} 539 puede codificarse utilizando el mismo modo de codificación (definido por diversos sub-modos).

[0693] En cualquier caso, como se ha indicado anteriormente, este esquema de codificación puede implicar primero transformar las representaciones de coma flotante de cada elemento (que es, en algunos ejemplos, un número de coma flotante de 32 bits) de cada uno de los vectores V^T_{DIST} 539 a una representación entera de 16 bits. La unidad de compresión V 552 puede realizar esta transformación de coma flotante a entero multiplicando cada elemento de uno dado de los vectores V^T_{DIST} 539 por 2^{15}, que es, en algunos ejemplos, realizado por un desplazamiento a la derecha por 15.

[0694] La unidad de compresión V 552 puede realizar entonces una cuantización uniforme con respecto a todos los elementos del uno dado de los vectores V^T_{DIST} 539. La unidad de compresión V 552 puede identificar un tamaño de paso de cuantización basado en un valor, que se puede indicar como un parámetro de *nbits*. La unidad de compresión V 552 puede determinar dinámicamente este parámetro de *nbits* basado en una velocidad de bits diana. La unidad de compresión V 552 puede determinar el tamaño del paso de cuantización como una función de este parámetro de *nbits*. Como ejemplo, la unidad de compresión V 552 puede determinar el tamaño del paso de cuantización (designado como "delta" o "Δ" en esta divulgación) como igual a $2^{16 - nbits}$. En este ejemplo, si *nbits* es igual a seis, delta es igual a 2^{10} y hay 2^6 niveles de cuantización. A este respecto, para un elemento vectorial v, el elemento vectorial cuantizado v_q equivale a [v/Δ] y $-2^{nbits-1} < v_q < 2^{nbits-1}$.

[0695] La unidad de compresión V 552 puede realizar entonces la categorización y la codificación residual de los elementos vectoriales cuantizados. Por ejemplo, la unidad de compresión V 552 puede, para un elemento vectorial cuantizado dado v_q, identificar una categoría (mediante la determinación de un identificador de categoría *cid*) a la que este elemento se corresponde usando la siguiente ecuación:

$$cid = \begin{cases} 0, & if\ v_q = 0 \\ \lfloor log_2|v_q| \rfloor + 1, & if\ v_q \neq 0 \end{cases}$$

La unidad de compresión V 552 puede entonces codificar con Huffman este índice de categoría *cid*, al mismo tiempo que también identifica un bit de signo que indica si v_q es un valor positivo o un valor negativo. La unidad de compresión V 552 puede identificar a continuación un residuo en esta categoría. Como un ejemplo, la unidad de compresión V 552 puede determinar este residuo de acuerdo con la siguiente ecuación:

$$residual = |v_q| - 2^{cid-1}$$

La unidad de compresión V 552 puede entonces bloquear el código de este residuo con *cid* -1 bits.

[0696] El siguiente ejemplo ilustra un ejemplo simplificado de este proceso de categorización y codificación residual. En primer lugar, se asume que *nbits* equivale a seis, de manera que $v_q \in$ [-31,31]. A continuación, se asume lo siguiente:

cid	vq	Código Huffman para *cid*
0	0	'1'
1	-1, 1	'01'
2	-3, -2, 2, 3	'000'
3	-7, -6, -5, -4, 4, 5, 6, 7	'0010'
4	-15, -14, ..., -8, 8, ...,14, 15	'00110'
5	-31, -30, ..., -16, 16, ..., 30, 31	'00111'

Además, se asume lo siguiente:

cid	Código de bloque para residuales
0	No disponible
1	0, 1
2	01,00, 10, 11
3	011, 010, 001, 000, 100, 101, 110, 111
4	0111, 0110...,0000, 1000, ...,1110, 1111
5	01111, ..., 00000, 10000, ..., 11111

Por lo tanto, para v_q = [6, -17, 0, 0, 3], puede determinarse lo siguiente:

- » *cid* = 3,5,0,0,2
- » signo =1,0,x,x,1
- » residual = 2,1,x,x,1
- » Bits para 6 = '0010' + '1' + '10'
- » Bits para -17 = '00111' + '0' + '0001'
- » Bits para 0 = '0'
- » Bits para 0 = '0'
- » Bits para 3 = '000' + '1' + '1'
- » Bits totales = 7+10+1+1+5 = 24
- » Promedio de bits = 24/5 = 4.8

[0697] Aunque no se muestra en el ejemplo simplificado anterior, la unidad de compresión V 552 puede seleccionar diferentes libros de código Huffman para diferentes valores de *nbits* cuando codifica el *cid*. En algunos ejemplos, la unidad de compresión V 552 puede proporcionar una tabla de codificación Huffman *diferente* para los valores de nbits 6, ..., 15. Además, la unidad de compresión V 552 puede incluir cinco libros de códigos Huffman diferentes para cada uno de los *diferentes* valores de nbits que van desde 6, ..., 15 para un total de 50 libros de códigos Huffman. A este respecto, la unidad de compresión V 552 puede incluir una pluralidad de diferentes libros de códigos Huffman para acomodar la codificación del *cid* en varios contextos estadísticos diferentes.

[0698] Para ilustrar, la unidad de compresión V 552 puede incluir, para cada uno de los valores de *nbits*, un primer libro de códigos Huffman para codificar los elementos vectoriales de uno a cuatro, un segundo libro de códigos Huffman para codificar los elementos vectoriales de cinco a nueve, un tercer libro de códigos Huffman para codificar los elementos vectoriales de nueve y sucesivos. Estos primeros tres libros de códigos Huffman pueden usarse cuando uno de los vectores V^T_{DIST} 539 a comprimir no se predice a partir de uno correspondiente temporalmente posterior de los vectores V^T_{DIST} 539 y no es representativo de la información espacial de un objeto de audio sintético (uno definido, por ejemplo, originalmente por un objeto de audio modulado por impulsos codificados (PCM)). La unidad de compresión V 552 puede incluir adicionalmente, para cada uno de los valores de *nbits*, un cuarto libro de códigos Huffman para codificar uno de los vectores V^T_{DIST} 539 cuando este de los vectores V^T_{DIST} 539 se predice a partir de uno correspondiente temporalmente posterior de los vectores V^T_{DIST} 539. La unidad de compresión V 552 también puede incluir, para cada uno de los valores de *nbits*, un quinto libro de códigos Huffman para codificar uno de los vectores V^T_{DIST} 539 cuando este uno de los vectores V^T_{DIST} 539 es representativo de un objeto de audio sintético. Los diversos libros de códigos Huffman pueden desarrollarse para cada uno de estos diferentes contextos estadísticos, es decir, el contexto no predicho y no sintético, el contexto predicho y el contexto sintético en este ejemplo.

[0699] La tabla siguiente ilustra la selección de la tabla Huffman y los bits que se especificarán en el flujo de bits para permitir que la unidad de descompresión seleccione la tabla Huffman adecuada:

Modo Pred	Info HT	Tabla HT
0	0	HT5
0	1	HT{1,2,3}
1	0	HT4
1	1	HT5

En la tabla anterior, el modo de predicción ("Modo Pred") indica si se realizó la predicción para el vector actual, mientras que la Tabla Huffman ("Info HT") indica la información adicional del libro de códigos Huffman (o tabla) usada para seleccionar una tabla Huffman de uno a cinco.

[0700] La tabla siguiente ilustra este proceso de selección de tablas Huffman dados diversos contextos o escenarios estadísticos.

	Grabación	Sintético
W/O Pred	HT{1,2,3}	HT5
Con Pred	HT4	HT5

En la tabla anterior, la columna "Grabación" indica el contexto de codificación cuando el vector es representativo de un objeto de audio que fue grabado mientras que la columna "sintética" indica un contexto de codificación cuando el vector es representativo de un objeto de audio sintético. La fila "W/O Pred" indica el contexto de codificación cuando la predicción no se realiza con respecto a los elementos vectoriales, mientras que la fila "Con Pred" indica el contexto de codificación cuando se realiza la predicción con respecto a los elementos vectoriales. Como se muestra en esta tabla, la unidad de compresión V 552 selecciona HT{1, 2, 3} cuando el vector es representativo de un objeto de audio grabado y no se realiza la predicción con respecto a los elementos de vector. La unidad de compresión V 552 selecciona HT5 cuando el objeto de audio es representativo de un objeto de audio sintético y no se realiza la predicción con respecto a los elementos de vector. La unidad de compresión 552 selecciona HT4 cuando el vector es representativo de un objeto de audio grabado y se realiza la predicción con respecto a los elementos vectoriales. La unidad de compresión V 552 selecciona HT5 cuando el objeto de audio es representativo de un objeto de audio sintético y se realiza la predicción con respecto a los elementos de vector.

[0701] De esta manera, las técnicas pueden permitir que un dispositivo de compresión de audio comprima un componente espacial de un campo sonoro, donde el componente espacial se generó realizando una síntesis basada en vectores con respecto a una pluralidad de coeficientes armónicos esféricos.

[0702] La FIG. 43 es un diagrama que ilustra la unidad de compresión V 552 mostrada en la FIG. 40I en más detalle. En el ejemplo de la FIG. 9K, la unidad de compresión V 552 incluye una unidad de cuantización uniforme 600, una unidad de nbits 602, una unidad de predicción 604, una unidad de modo de predicción 606 ("Unidad modo prep. 606"), una unidad de codificación de categoría y residual 608 y una unidad de selección de tabla Huffman 610. La unidad de cuantización uniforme 600 representa una unidad configurada para llevar a cabo la cuantización uniforme descrita anteriormente con respecto a uno de los componentes espaciales representado como *v* en el ejemplo de la FIG. 43 (que puede representar uno cualquiera de los vectores V^T_{DIST} 539). La unidad de nbits 602 representa una unidad configurada para determinar el parámetro o valor de *nbits*.

[0703] La unidad de predicción 604 representa una unidad configurada para realizar la predicción con respecto al componente espacial cuantizado indicado como v_q en el ejemplo de la FIG. 43. La unidad de predicción 604 puede realizar la predicción realizando una resta de elementos del actual de los vectores V^T_{DIST} 539 por uno temporalmente posterior correspondiente de los vectores V^T_{DIST} 539. El resultado de esta predicción puede ser referido como un componente espacial predicho.

[0704] La unidad de modo de predicción 606 puede representar una unidad configurada para seleccionar el modo de predicción. La unidad de selección de tabla Huffman 610 puede representar una unidad configurada para seleccionar una tabla Huffman apropiada para la codificación del *cid*. La unidad de modo de predicción 606 y la unidad de selección de tabla Huffman 610 pueden funcionar, como un ejemplo, de acuerdo con el siguiente pseudocódigo:

```
Para un determinado nbits, recuperar todas las tablas Huffman con nbits
    B00 = 0; B01 = 0; B10 = 0; B11 = 0; // inicializar para calcular los bits esperados
    modo de codificación
    para m = 1: (# Elementos en el vector)
        // calcular el número esperado de bits para un elemento vectorial v(m)
        // sin predicción y utilizando tabla Huffman 5
        B00 = B00 + calculate_bits(v(m), HT5);
        // sin predicción y usando tabla Huffman {1,2,3}
        B01 = B01 + calculate_bits(v(m), HTq); q en {1,2,3}
        // calcular el número esperado de bits para residuo de predicción e(m)
        e(m) = v(m) - vp(m); // vp(m): elemento vectorial de trama anterior
        // con predicción y uso de tabla Huffman 4
        B10 = B10 + calculate_bits(e(m), HT4);
        // con predicción y uso de tabla Huffman 5
        B11 = B11 + calculate_bits(e(m), HT5);
    end
    // encontrar un mejor modo de predicción y una tabla Huffman que produzcan bits mínimos
    // el mejor modo de predicción y la tabla Huffman son marcados por pflag y Htflag,
    respectivamente
    [Be, id] = min([B00 B01 B10 B11]);
    Conmutar id
        caso 1: pflag = 0; HTflag = 0;
        caso 2: pflag = 0; HTflag = 1;
        caso 3: pflag = 1; HTflag = 0;
        caso 4: pflag = 1; HTflag = 1;
end
```

[0705] La unidad de codificación de categoría y residual 608 puede representar una unidad configurada para realizar la codificación de categorización y residual de un componente espacial previsto o el componente espacial cuantizado (cuando la predicción está desactivada) de la manera descrita con más detalle anteriormente.

[0706] Como se ilustra en el ejemplo de la FIG. 9K, la unidad de compresión V 552 puede emitir varios parámetros o valores para su inclusión en el flujo de bits 517 o información lateral (que puede ser un flujo de bits separado del flujo de bits 517). Asumiendo que la información se especifica en el flujo de bits 517, la unidad de compresión V 552 puede emitir el valor de *nbits*, el modo de predicción y la información de tabla Huffman a la unidad de generación de flujo de bits 516 junto con la versión comprimida del componente espacial (mostrado como el componente espacial comprimido 539' en el ejemplo de la FIG. 40I), que en este ejemplo puede referirse al código Huffman seleccionado para codificar el *cid*, el bit de signo y el residuo codificado por bloques. El valor de *nbits* se puede especificar una vez en el flujo de bits 517 para todos los vectores V^T_{DIST} 539, mientras que el modo de predicción y la información de tabla Huffman pueden especificarse para cada uno de los vectores V^T_{DIST} 539. La parte del flujo de bits que especifica la versión comprimida del componente espacial se muestra en el ejemplo de las FIGS. 10B y 10C.

[0707] De esta manera, el dispositivo de codificación de audio 510H puede realizar diversos aspectos de las técnicas expuestas a continuación con respecto a las siguientes cláusulas.

Cláusula 141541-1A. Un dispositivo, tal como el dispositivo de codificación de audio 510H, que comprende: uno o más procesadores configurados para obtener 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.

Cláusula 141541-2A. El dispositivo de las cláusulas 141541-1A, en el que la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un campo que especifica un modo de predicción utilizado al comprimir el componente espacial.

Cláusula 141541-3A. El dispositivo de cualquier combinación de la cláusula 141541-1A y la cláusula 141541-2A, en el que la versión comprimida del componente espacial está representada en el flujo de bits utilizando, al menos en parte, información de tabla Huffman que especifica una tabla Huffman usada al comprimir el componente espacial.

Cláusula 141541-4A. El dispositivo de cualquier combinación de la cláusula 141541-1A a la cláusula 141541-3A, en el que la versión comprimida del componente espacial está representada en el flujo de bits utilizando, al menos en parte, un campo que indica un valor que expresa un tamaño de paso de cuantización o una variable del mismo que se utiliza al comprimir el componente espacial.

Cláusula 141541-5A. El dispositivo de la cláusula 141541-4A, en el que el valor comprende un valor de nbits.

Cláusula 141541-6A. El dispositivo de cualquier combinación de la cláusula 141541-4A y la cláusula 141541-5A, en el que el flujo de bits comprende una versión comprimida de una pluralidad de componentes espaciales del campo sonoro de la que está incluida la versión comprimida del componente espacial, y en el que el valor expresa el tamaño del paso de cuantización o una variable del mismo utilizada al comprimir la pluralidad de componentes espaciales.

Cláusula 141541-7A. El dispositivo de cualquier combinación de la cláusula 141541-1A a la cláusula 141541-6A, en el que la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un código Huffman para representar un identificador de categoría que identifica una categoría de compresión a la que corresponde el componente espacial.

Cláusula 141541-8A. El dispositivo de cualquier combinación de la cláusula 141541-1A a la cláusula 141541-7A, en el que la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un bit de signo que identifica si el componente espacial es un valor positivo o un valor negativo.

Cláusula 141541-9A. El dispositivo de cualquier combinación de la cláusula 141541-1A a la cláusula 141541-8A, en el que la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un código Huffman para representar un valor residual del componente espacial.

Cláusula 141541-10A. El dispositivo de cualquier combinación de la cláusula 141541-1A a la cláusula 141541-9A, en el que el dispositivo comprende un dispositivo de codificación de audio, un dispositivo de generación de flujo de bits.

Cláusula 141541-12A. El dispositivo de cualquier combinación de la cláusula 141541-1A a la cláusula 141541-11A, en el que la síntesis basada en vectores comprende una descomposición en valores singulares.

[0708] Aunque se describen como realizadas por el dispositivo de codificación de audio 510H, las técnicas también pueden realizarse por cualquiera de los dispositivos de decodificación de audio 24 y/o 540.

[0709] De esta manera, el dispositivo de codificación de audio 510H puede realizar adicionalmente diversos aspectos de las técnicas expuestas a continuación con respecto a las siguientes cláusulas.

Cláusula 141541-1D. Un dispositivo, tal como el dispositivo de codificación de audio 510H, que comprende: uno o más procesadores configurados para generar 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.

Cláusula 141541-2D. El dispositivo de la cláusula 141541-1D, en el que el uno o más procesadores están configurados adicionalmente para, al generar el flujo de bits, generar el flujo de bits para incluir un campo que especifica un modo de predicción usado al comprimir el componente espacial.

Cláusula 141541-3D. El dispositivo de cualquier combinación de la cláusula 141541-1D y la cláusula 141541-2D, en el que el uno o más procesadores están configurados adicionalmente para, al generar el flujo de bits, generar el flujo de bits para incluir información de tabla Huffman que especifica una tabla Huffman usada al comprimir el componente espacial.

Cláusula 141541-4D. El dispositivo de cualquier combinación de la cláusula 141541-1D a la cláusula 141541-3D, en el que el uno o más procesadores están configurados adicionalmente para, al generar el flujo de bits, generar el flujo de bits para incluir un campo que indica un valor que expresa un tamaño de paso de cuantización o una variable del mismo que se utiliza al comprimir el componente espacial.

Cláusula 141541-5D. El dispositivo de la cláusula 141541-4D, en el que el valor comprende un valor de *nbits*.

Cláusula 141541-6D. El dispositivo de cualquier combinación de la cláusula 141541-4D y la cláusula 141541-5D,

en el que el uno o más procesadores están configurados adicionalmente para, al generar la corriente de bits, generar la corriente de bits para incluir una versión comprimida de una pluralidad de componentes espaciales del campo sonoro de los cuales se incluye la versión comprimida del componente espacial, y en el que el valor expresa el tamaño del paso de cuantización o una variable del mismo que se utiliza al comprimir la pluralidad de componentes espaciales.

Cláusula 141541-7D. El dispositivo de cualquier combinación de la cláusula 141541-1D a la cláusula 141541-6D, en el que el uno o más procesadores están configurados adicionalmente para, al generar el flujo de bits, generar el flujo de bits para incluir un código Huffman para representar un identificador de categoría que identifique una categoría de compresión a la que corresponde el componente espacial.

Cláusula 141541-8D. El dispositivo de cualquier combinación de la cláusula 141541-1D a la cláusula 141541-7D, en el que el uno o más procesadores están configurados adicionalmente para, al generar el flujo de bits, generar el flujo de bits para incluir un bit de signo que identifica si el componente espacial es un valor positivo o un valor negativo.

Cláusula 141541-9D. El dispositivo de cualquier combinación de la cláusula 141541-1D a la cláusula 141541-8D, en el que el uno o más procesadores están configurados adicionalmente para, al generar el flujo de bits, generar el flujo de bits para incluir un código Huffman para representar un valor residual del componente espacial.

Cláusula 141541-10D. El dispositivo de cualquier combinación de la cláusula 141541-1D a la cláusula 141541-10D, en el que la síntesis basada en vectores comprende una descomposición en valores singulares.

[0710] El dispositivo de codificación de audio 510H puede configurarse además para implementar diversos aspectos de las técnicas como se establece en las siguientes cláusulas.

Cláusula 141541-1E. Un dispositivo, tal como el dispositivo de codificación de audio 510H, que comprende: uno o más procesadores configurados para comprimir 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.

Cláusula 141541-2E. El dispositivo de la cláusula 141541-1E, en el que el uno o más procesadores están configurados adicionalmente para, al comprimir el componente espacial, convertir el componente espacial de una representación de coma flotante a una representación entera.

Cláusula 141541-3E. El dispositivo de cualquier combinación de la cláusula 141541-1E y la cláusula 141541-2E, en el que el uno o más procesadores están configurados adicionalmente para, al comprimir el componente espacial, determinar dinámicamente un valor indicativo de un tamaño de paso de cuantización, y cuantizar el componente espacial basado en el valor para generar un componente espacial cuantificado.

Cláusula 141541-4E. El dispositivo de cualquier combinación de las reivindicaciones 1E-3E, en el que el uno o más procesadores están configurados adicionalmente para, al comprimir el componente espacial, identificar una categoría a la cual corresponde el componente espacial.

Cláusula 141541-5E. El dispositivo de cualquier combinación de la cláusula 141541-1E a la cláusula 141541-4E, en el que el uno o más procesadores están configurados adicionalmente para, al comprimir el componente espacial, identificar un valor residual para el componente espacial.

Cláusula 141541-6E. El dispositivo de cualquier combinación de la cláusula 141541-1E a la cláusula 141541-5E, en el que el uno o más procesadores están configurados adicionalmente para, al comprimir el componente espacial, realizar una predicción con respecto al componente espacial y un componente espacial posterior para generar un componente espacial predicho.

Cláusula 141541-7E. El dispositivo de cualquier combinación de la cláusula 141541-1E, en el que el uno o más procesadores están configurados adicionalmente para, al comprimir el componente espacial, convertir el componente espacial de una representación de coma flotante en una representación entera, determinar dinámicamente un valor indicativo de un tamaño del paso de cuantización, cuantizar la representación entera del componente espacial basándose en el valor para generar un componente espacial cuantificado, identificar una categoría a la que corresponde el componente espacial basándose en el componente espacial cuantificado para generar un identificador de categoría, determinar un signo del componente espacial, identificar un valor residual para el componente espacial basándose en el componente espacial cuantificado y el identificador de categoría, y generar una versión comprimida del componente espacial basándose en el identificador de categoría, el signo y el valor residual.

Cláusula 141541-8E. El dispositivo de cualquier combinación de la cláusula 141541-1E, en el que el uno o más procesadores están configurados adicionalmente para, al comprimir el componente espacial, convertir el

componente espacial de una representación de coma flotante en una representación entera, determinar dinámicamente un valor indicativo de un tamaño del paso de cuantización, cuantizar la representación entera del componente espacial basándose en el valor para generar un componente espacial cuantificado, realizar una predicción con respecto al componente espacial y un componente espacial posterior para generar un componente espacial predicho, identificar una categoría a la que corresponde el componente espacial predicho basándose en el componente espacial cuantificado para generar un identificador de categoría, determinar un signo del componente espacial, identificar un valor residual para el componente espacial basándose en el componente espacial cuantificado y el identificador de categoría, y generar una versión comprimida del componente espacial basándose en el identificador de categoría, el signo y el valor residual.

Cláusula 141541-9E. El dispositivo de cualquier combinación de la cláusula 141541-1E a la cláusula 141541-8E, en el que la síntesis basada en vectores comprende una descomposición en valores singulares.

[0711] Diversos aspectos de las técnicas pueden además permitir que el dispositivo de codificación de audio 510H esté configurado para funcionar como se establece en las siguientes cláusulas.

Cláusula 141541-1F. Un dispositivo, tal como el dispositivo de codificación de audio 510H, que comprende: uno o más procesadores configurados para identificar un libro de códigos Huffman para usar al comprimir un componente espacial actual de una pluralidad de componentes espaciales basándose en un orden del componente espacial actual con respecto a los restantes de la pluralidad de componentes espaciales, el componente espacial generado realizando una síntesis basada en vectores con respecto a una pluralidad de coeficientes armónicos esféricos.

Cláusula 141541-2F. El dispositivo de la cláusula 141541-3F, en el que el uno o más procesadores están configurados además para realizar cualquier combinación de las etapas enumeradas en la cláusula 141541-1A a la cláusula 141541-12A, la cláusula 141541-1B a la cláusula 141541-10B y la cláusula 141541-1C a la cláusula 141541-9C.

[0712] Diversos aspectos de las técnicas pueden además permitir que el dispositivo de codificación de audio 510H esté configurado para funcionar como se establece en las siguientes cláusulas.

Cláusula 141541-1H. Un dispositivo, tal como el dispositivo de codificación de audio 510H, que comprende: uno o más procesadores configurados para determinar un tamaño de paso de cuantización que se usará al comprimir 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.

Cláusula 141541-2H. El dispositivo de la cláusula 141541-1H, en el que el uno o más procesadores están configurados adicionalmente para determinar, al determinar el tamaño del paso de cuantización, el tamaño del paso de cuantización basándose en una velocidad de bits objetivo.

Cláusula 141541-3H. El dispositivo de la cláusula 141541-1H, en el que el uno o más procesadores están configurados para, al seleccionar una de la pluralidad de tamaños del paso de cuantización, determinar una estimación de un número de bits utilizados para representar el componente espacial y determinar el tamaño del paso de cuantización basándose en una diferencia entre la estimación y una velocidad de bits objetivo.

Cláusula 141541-4H. El dispositivo de la cláusula 141541-1H, en el que el uno o más procesadores están configurados para, al seleccionar una de la pluralidad de tamaños del paso de cuantización, determinar una estimación de un número de bits utilizados para representar el componente espacial, determinar una diferencia entre la estimación y una velocidad de bits objetivo, y determinar el tamaño del paso de cuantización añadiendo la diferencia a la velocidad de bits objetivo.

Cláusula 141541-5H. El dispositivo de la cláusula 141541-3H o la cláusula 141541-4H, en el que el uno o más procesadores están configurados adicionalmente para, al determinar la estimación del número de bits, calcular la estimación del número de bits que se van a generar para el componente espacial que recibe un libro de códigos correspondiente a la velocidad de bits objetivo.

Cláusula 141541-6H. El dispositivo de la cláusula 141541-3H o la cláusula 141541-4H, en el que el uno o más procesadores están configurados adicionalmente para, al determinar la estimación del número de bits, calcular la estimación del número de bits que se van a generar para el componente espacial que recibe un modo de codificación usando al comprimir el componente espacial.

Cláusula 141541-7H. El dispositivo de la cláusula 141541-3H o la cláusula 141541-4H, en el que el uno o más procesadores están configurados para, al determinar la estimación del número de bits, calcular una primera estimación del número de bits que se van a generar para el componente espacial dado un primer modo de codificación que se utilizará al comprimir el componente espacial, calcular una segunda estimación del número de bits que se van a generar para el componente espacial dado un segundo modo de codificación que se

utilizará al comprimir el componente espacial, seleccionar el de la primera estimación y la segunda estimación que tenga el menor número de bits que se va a utilizar como la estimación determinada del número de bits.

Cláusula 141541-8H. El dispositivo de la cláusula 141541-3H o la cláusula 141541-4H, en el que el uno o más procesadores están configurados para, al determinar la estimación del del número de bits, identificar un identificador de categoría que identifica una categoría a la cual corresponde el componente espacial, identificar una longitud de bits de un valor residual para el componente espacial que resultaría al comprimir el componente espacial correspondiente a la categoría y determinar la estimación del número de bits mediante, al menos en parte, la adición de un número de bits utilizados para representar el identificador de categoría a la longitud de bits del valor residual.

Cláusula 141541-9H. El dispositivo de cualquier combinación de la cláusula 141541-1H a la cláusula 141541-8H, en el que la síntesis basada en vectores comprende una descomposición en valores singulares.

[0713] Aunque se describen como realizadas por el dispositivo de codificación de audio 510H, las técnicas expuestas en las cláusulas anteriores, la cláusula 141541-1H a la cláusula 141541-9H también pueden realizarse por el dispositivo de decodificación de audio 540D.

[0714] Adicionalmente, diversos aspectos de las técnicas pueden permitir que el dispositivo de codificación de audio 510H esté configurado para funcionar como se establece en las siguientes cláusulas.

Cláusula 141541-1J. Un dispositivo, tal como el dispositivo de codificación de audio 510J, que comprende: uno o más procesadores configurados para seleccionar uno de una pluralidad de libros de códigos a usar al comprimir 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.

Cláusula 141541-2J. El dispositivo de la cláusula 141541-1J, en el que el uno o más procesadores están configurados además para, al seleccionar uno de la pluralidad de libros de códigos, determinar una estimación de un número de bits utilizados para representar el componente espacial utilizando cada uno de la pluralidad de libros de códigos, y seleccionar el libro de la pluralidad de libros de código que dio como resultado la estimación determinada que tiene el menor número de bits.

Cláusula 141541-3J. El dispositivo de la cláusula 141541-1J, en el que el uno o más procesadores están configurados adicionalmente para, al seleccionar uno de la pluralidad de libros de códigos, determinar una estimación de un número de bits utilizados para representar el componente espacial utilizando uno o más de la pluralidad de libros de códigos, uno o más de la pluralidad de libros de códigos seleccionados basándose en un orden de elementos del componente espacial que se comprimirá con relación a otros elementos del componente espacial.

Cláusula 141541-4J. El dispositivo de la cláusula 141541-1J, en el que el uno o más procesadores están configurados además para, al seleccionar uno de la pluralidad de libros de códigos, determinar una estimación de un número de bits utilizados para representar el componente espacial utilizando uno de la pluralidad de libros de códigos designados para usarse cuando el componente espacial no se predice a partir de un componente espacial posterior.

Cláusula 141541-5J. El dispositivo de la cláusula 141541-1J, en el que el uno o más procesadores están configurados además para, al seleccionar uno de la pluralidad de libros de códigos, determinar una estimación de un número de bits utilizados para representar el componente espacial utilizando uno de la pluralidad de libros de códigos designados para usarse cuando el componente espacial se predice a partir de un componente espacial posterior.

Cláusula 141541-6J. El dispositivo de la cláusula 141541-1J, en el que el uno o más procesadores están configurados además para, al seleccionar uno de la pluralidad de libros de códigos, determinar una estimación de un número de bits utilizados para representar el componente espacial utilizando uno de la pluralidad de libros de códigos designados para usarse cuando el componente espacial es representativo de un objeto de audio sintético en el campo sonoro.

Cláusula 141541-7J. El dispositivo de la cláusula 141541-1J, en el que el objeto de audio sintético comprende un objeto de audio modulado por impulsos codificados (PCM).

Cláusula 141541-8J. El dispositivo de la cláusula 141541-1J, en el que el uno o más procesadores están configurados además para, al seleccionar uno de la pluralidad de libros de códigos, determinar una estimación de un número de bits utilizados para representar el componente espacial utilizando uno de la pluralidad de libros de códigos designados para usarse cuando el componente espacial es representativo de un objeto de audio grabado en el campo sonoro.

Cláusula 141541-9J. El dispositivo de cualquier combinación de las reivindicaciones 1J-8J. en el que la síntesis basada en vectores comprende una descomposición en valores singulares.

[0715] En cada una de las diversas instancias descritas anteriormente, debe entenderse que el dispositivo de codificación de audio 510 puede realizar un procedimiento o comprender de otro modo medios para realizar cada paso del procedimiento para el que está configurado el dispositivo de codificación de audio 510. En algunos casos, estos medios pueden comprender uno o más procesadores. En algunos casos, 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 instrucciones que, cuando se ejecutan, hacen que uno o más procesadores realicen el procedimiento para el cual el dispositivo de codificación de audio 510 se ha configurado para llevar a cabo.

[0716] La FIG. 40J es un diagrama de bloques que ilustra un ejemplo de dispositivo de codificación de audio 510J que puede realizar diversos aspectos de las técnicas descritas en esta descripción para comprimir los coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales. El dispositivo de codificación de audio 510J puede ser similar al dispositivo de codificación de audio 510G en que el dispositivo de codificación de audio 510J incluye una unidad de compresión de audio 512, una unidad de codificación de audio 514 y una unidad de generación de flujo de bits 516. Además, la unidad de compresión de audio 512 del dispositivo de codificación de audio 510J puede ser similar a la del dispositivo de codificación de audio 510G porque la unidad de compresión de audio 512 incluye una unidad de descomposición 518 y una unidad de extracción de componentes del campo sonoro 520, que pueden operar de forma similar a las unidades del dispositivo de codificación de audio 510I. En algunos ejemplos, el dispositivo de codificación de audio 510J puede incluir una unidad de cuantización 534, como se ha descrito con respecto a las FIGS. 40D-40E, para cuantizar uno o más vectores de cualquiera de los vectores U_{DIST} 525C, los vectores U_{BG} 525D, los vectores V^T_{DIST} 525E y los vectores V^T_{BG} 525J.

[0717] La unidad de compresión de audio 512 del dispositivo de codificación de audio 510J puede, sin embargo, diferir de la unidad de compresión de audio 512 del dispositivo de codificación de audio 510G en que la unidad de compresión de audio 512 del dispositivo de codificación de audio 510J incluye una unidad adicional denominada como unidad de interpolación 550. La unidad de interpolación 550 puede representar una unidad que interpola subtramas de una primera trama de audio de las subtramas de la primera trama de audio y una segunda trama de audio temporalmente posterior o anterior, como se describe con más detalle a continuación con respecto a las FIGS. 45 y 45B. La unidad de interpolación 550 puede reducir, al realizar esta interpolación, la complejidad computacional (en términos de ciclos de procesamiento y/o consumo de memoria) reduciendo potencialmente la medida en que la unidad de descomposición 518 se requiere para descomponer SHC 511. La unidad de interpolación 550 puede operar de una manera similar a la descrita anteriormente con respecto a la unidad de interpolación 550 de los dispositivos de codificación de audio 510H y 510I mostrados en los ejemplos de las FIGS. 40H y 40I.

[0718] En funcionamiento, la unidad de interpolación 200 puede interpolar una o más subtramas de una primera trama de audio a partir de una primera descomposición, por ejemplo, la matriz V 19', de una parte de una primera pluralidad de coeficientes armónicos esféricos 11 incluidos en la primera trama y una segunda descomposición, por ejemplo, la matriz V 19', de una porción de una segunda pluralidad de coeficientes armónicos esféricos 11 incluidos en una segunda trama para generar coeficientes armónicos esféricos interpolados descompuestos para la una o más subtramas.

[0719] La unidad de interpolación 550 puede obtener coeficientes armónicos esféricos interpolados descompuestos para un segmento de tiempo, realizando al menos en parte una interpolación con respecto a una primera descomposición de una primera pluralidad de coeficientes armónicos esféricos y una segunda descomposición de una segunda pluralidad de coeficientes armónicos esféricos. La unidad de suavizado 554 puede aplicar los coeficientes armónicos esféricos interpolados descompuestos para suavizar al menos uno de componentes espaciales y los componentes temporales de la primera pluralidad de coeficientes armónicos esféricos y la segunda pluralidad de coeficientes armónicos esféricos. La unidad de suavizado 554 puede generar matrices U_{DIST} suavizadas 525C' como se ha descrito anteriormente con respecto a las FIGS. 37-39. La primera y segunda descomposiciones pueden referirse a V_1^T 556, V_2^T 556B en la FIG. 40J.

[0720] En algunos casos, V^T u otros vectores V o matrices V pueden ser emitidos en una versión cuantificada para interpolación. De esta manera, los vectores V para la interpolación pueden ser idénticos a los vectores V en el decodificador, que también realiza la interpolación de vectores V, por ejemplo, para recuperar la señal multidimensional.

[0721] En algunos ejemplos, la primera descomposición comprende la primera matriz V 519' representativa de vectores singulares derechos de la porción de la primera pluralidad de coeficientes armónicos esféricos 511. De manera similar, en algunos ejemplos, la segunda descomposición comprende la segunda matriz V 519' representativa de vectores singulares derechos de la porción de la segunda pluralidad de coeficientes armónicos esféricos.

[0722] La unidad de interpolación 550 puede realizar una interpolación temporal con respecto a una o más subtramas basadas en la primera matriz V 519' y la segunda matriz V 19'. Es decir, la unidad de interpolación 550 puede interpolar temporalmente, por ejemplo, la segunda, tercera y cuarta subtramas de cuatro subtramas totales para la primera trama de audio basadas en una matriz V 519' descompuesta a partir de la primera subtrama de la primera trama de audio y la matriz V 519' descompuesta a partir de la primera subtrama de la segunda trama de audio. En algunos ejemplos, esta interpolación temporal es una interpolación temporal lineal, en la que la matriz V 519' descompuesta a partir de la primera subtrama de la primera trama de audio se pondera más pesadamente cuando se interpola la segunda subtrama de la primera trama de audio que cuando se interpola la cuarta subtrama de la primera trama de audio. Cuando se interpola la tercera subtrama, las matrices V 519' pueden ponderarse uniformemente. Cuando se interpola la cuarta subtrama, la matriz V 519' descompuesta a partir de la primera subtrama de la segunda trama de audio puede ponderarse con mayor peso que la matriz V 519' descompuesta de la primera subtrama de la primera trama de audio.

[0723] En otras palabras, la interpolación temporal lineal puede ponderar las matrices V 519' dada la proximidad de una de las subtramas de la primera trama de audio a interpolar. Para que la segunda subtrama sea interpolada, la matriz V 519' descompuesta a partir de la primera subtrama de la primera trama de audio se pondera en mayor medida dada su proximidad a la segunda subtrama a interpolar que la matriz V 519' descompuesta a partir de la primera subtrama de la segunda trama de audio. Los pesos pueden ser equivalentes por este motivo al interpolar la tercera subtrama basada en las matrices V 519'. El peso aplicado a la matriz V 519' descompuesta a partir de la primera subtrama de la segunda trama de audio puede ser mayor que el aplicado a la matriz V 519' descompuesta a partir de la primera subtrama de la primera trama de audio dado que la cuarta subtrama a interpolar está más próxima a la primera subtrama de la segunda trama de audio que la primera subtrama de la primera trama de audio.

[0724] En algunos ejemplos, la unidad de interpolación 550 puede proyectar la primera matriz V 519' descompuesta de la primera subtrama de la primera trama de audio a un dominio espacial para generar las primeras descomposiciones proyectadas. En algunos ejemplos, esta proyección incluye una proyección en una esfera (por ejemplo, utilizando una matriz de proyección, tal como una matriz de diseño T). La unidad de interpolación 550 puede entonces proyectar la segunda matriz V 519' descompuesta de la primera subtrama de la segunda trama de audio al dominio espacial para generar segundas descomposiciones proyectadas. La unidad de interpolación 550 puede entones interpolar espacialmente (que de nuevo puede ser una interpolación lineal) las primeras descomposiciones proyectadas y las segundas descomposiciones proyectadas para generar una primera descomposición proyectada espacialmente interpolada y una segunda descomposición proyectada espacialmente interpolada. La unidad de interpolación 550 puede entonces interpolar temporalmente la una o más subtramas basándose en la primera descomposición proyectada espacialmente interpolada y la segunda descomposición proyectada espacialmente interpolada.

[0725] En los ejemplos, en los que la unidad de interpolación 550 proyecta espacialmente y después temporalmente las matrices V 519', la unidad de interpolación 550 puede proyectar los coeficientes armónicos esféricos interpolados temporalmente resultantes de la interpolación de uno o más subtramas de vuelta a un dominio armónico esférico, generando así la matriz V 519, la matriz S 519B y la matriz U 519C.

[0726] En algunos ejemplos, la porción de la primera pluralidad de coeficientes armónicos esféricos comprende una única subtrama de la primera pluralidad de coeficientes armónicos esféricos 511. En algunos ejemplos, la parte de la segunda pluralidad de coeficientes armónicos esféricos comprende una única subtrama de la segunda pluralidad de coeficientes armónicos esféricos 511. En algunos ejemplos, esta subtrama única de la que se descomponen las matrices V 19' es la primera subtrama.

[0727] En algunos ejemplos, la primera trama se divide en cuatro subtramas. En estos y otros ejemplos, la parte de la primera pluralidad de coeficientes armónicos esféricos comprende solamente la primera subtrama de la pluralidad de coeficientes armónicos esféricos 511. En estos y otros ejemplos, la segunda trama está dividida en cuatro subtramas, y la parte de la segunda pluralidad de coeficientes armónicos esféricos 511 comprende solamente la primera subtrama de la segunda pluralidad de coeficientes armónicos esféricos 511.

[0728] Aunque, en algunos ejemplos, sólo se utiliza una primera subtrama de cada trama de audio para realizar la interpolación, la parte de la primera pluralidad de coeficientes armónicos esféricos puede comprender dos de cuatro subtramas de la primera pluralidad de coeficientes armónicos esféricos 511. En estos y otros ejemplos, la porción de la segunda pluralidad de coeficientes armónicos esféricos 511 comprende dos de cuatro subtramas de la segunda pluralidad de coeficientes armónicos esféricos 511.

[0729] Como se ha indicado anteriormente, un dispositivo individual, por ejemplo, el dispositivo de codificación de audio 510J, puede realizar la interpolación, al tiempo que también descompone la parte de la primera pluralidad de coeficientes armónicos esféricos para generar las primeras descomposiciones de la porción de la primera pluralidad de coeficientes armónicos esféricos. En estos y otros ejemplos, la unidad de descomposición 518 puede descomponer la porción de la segunda pluralidad de coeficientes armónicos esféricos para generar las segundas descomposiciones de la porción de la segunda pluralidad de coeficientes armónicos esféricos. Aunque se describen

con respecto a un único dispositivo, dos o más dispositivos pueden realizar las técnicas descritas en esta divulgación, donde uno de los dos dispositivos realiza la descomposición y otro de los dispositivos realiza la interpolación de acuerdo con las técnicas descritas en esta divulgación.

[0730] En algunos ejemplos, la unidad de descomposición 518 puede realizar una descomposición en valores singulares con respecto a la porción de la primera pluralidad de coeficientes armónicos esféricos 511 para generar una matriz V 519' (así como una matriz S 519B' y una matriz U 519C', que no se muestran con fines de ilustración) representativa de vectores singulares derechos de la primera pluralidad de coeficientes armónicos esféricos 511. En estos y otros ejemplos, la unidad de descomposición 518 puede realizar la descomposición en valores singulares con respecto a la porción de la segunda pluralidad de coeficientes armónicos esféricos 511 para generar una matriz V 519' (así como una matriz S 519B' y una matriz U 519C', que no se muestran con fines de ilustración) representativa de vectores singulares derechos de la segunda pluralidad de coeficientes armónicos esféricos.

[0731] En algunos ejemplos, como se ha indicado anteriormente, la primera y la segunda pluralidad de coeficientes armónicos esféricos representan, cada una, una representación de ondas planas del campo sonoro. En estos y otros ejemplos, la primera y la segunda pluralidad de coeficientes armónicos esféricos 511 representan cada uno uno o más objetos monoaudio mezclados entre sí.

[0732] En otras palabras, el audio 3D basado en armónicos esféricos puede ser una representación paramétrica del campo de presión 3D en términos de funciones de base ortogonales sobre una esfera. Cuanto mayor sea el orden N de la representación, la resolución espacial será potencialmente más alta 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. Estas 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 SVD puede descomponer cada trama de coeficientes en tres matrices U, S y V. En algunos ejemplos, las técnicas pueden manejar algunos de los vectores en U como componentes direccionales del campo sonoro subyacente. Sin embargo, cuando se manejan de esta manera, estos vectores (en U) son discontinuos de trama en trama, aunque representan el mismo componente de audio distinto. Estas discontinuidades pueden conducir a artefactos significativos cuando los componentes son alimentados a través de codificadores de audio de transformadas.

[0733] Las técnicas descritas en esta divulgación pueden abordar esta discontinuidad. Es decir, las técnicas pueden basarse en la observación de que la matriz V puede interpretarse como ejes espaciales ortogonales en el dominio de las armónicas esféricas. La matriz U puede representar una proyección de los datos de armónicos esféricos (HOA) en términos de las funciones de base, donde la discontinuidad puede atribuirse a las funciones de base (V) que cambian cada trama y, por lo tanto, son discontinuas en sí mismas. Esto es diferente a la descomposición similar, como la Transformada 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 por coincidencia. Las técnicas descritas en esta divulgación pueden permitir que la unidad de interpolación 550 mantenga la continuidad entre las funciones de base (V) de una trama a otra mediante interpolación entre ellas.

[0734] En algunos ejemplos, las técnicas permiten a la unidad de interpolación 550 dividir la trama de datos SH en cuatro subtramas, como se describió anteriormente y se describe adicionalmente a continuación con respecto a las FIGS. 45 y 45B. La unidad de interpolación 550 puede entonces calcular la SVD para la primera subtrama. De forma similar, se calcula la SVD para la primera subtrama de la segunda trama. Para cada una de la primera trama y la segunda trama, la unidad de interpolación 550 puede convertir los vectores en V en un mapa espacial proyectando los vectores sobre una esfera (utilizando una matriz de proyección tal como una matriz de diseño en T). La unidad de interpolación 550 puede entonces interpretar los vectores en V como formas en una esfera. Para interpolar las matrices V para las tres subtramas entre la primera subtrama de la primera trama, la primera subtrama de la trama siguiente, la unidad de interpolación 550 puede interpolar entonces estas formas espaciales y luego transformarlas de nuevo en los vectores SH a través de la inversa de la matriz de proyección. Las técnicas de esta descripción pueden, de esta manera, proporcionar una transición suave entre las matrices V.

[0735] La FIG. 41-41D son diagramas de bloques que ilustran cada uno un dispositivo de decodificación de audio ejemplar 540A-540D que puede realizar diversos aspectos de las técnicas descritas en esta divulgación para decodificar coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales. El dispositivo de decodificación de audio 540A puede representar cualquier dispositivo capaz de decodificar datos de audio, tales como un ordenador de sobremesa, un ordenador portátil, una estación de trabajo, una tableta o un ordenador de pizarra, un dispositivo de grabación de audio dedicado, un teléfono móvil (incluidos los llamados "teléfonos inteligentes"), un dispositivo de reproductor multimedia personal, un dispositivo de juego personal o cualquier otro tipo de dispositivo capaz de decodificar datos de audio.

[0736] En algunos ejemplos, el dispositivo de decodificación de audio 540A realiza un proceso de decodificación de audio que es recíproco al proceso de codificación de audio realizado por cualquiera de los dispositivos de codificación de audio 510 o 510B con la excepción de realizar la reducción de orden (como se ha descrito

anteriormente con respecto a los ejemplos de las FIGS. 40B-40J), que es, en algunos ejemplos, utilizado por los dispositivos de codificación de audio 510B-510J para facilitar la eliminación de datos irrelevantes extraños.

[0737] Aunque se muestra como un dispositivo individual, es decir, el dispositivo 540A en el ejemplo de la FIG. 41, los diversos componentes o unidades a los que se hace referencia a continuación como incluidos en el dispositivo 540A pueden formar dispositivos separados que son externos al dispositivo 540. En otras palabras, aunque se describe en esta divulgación como realizadas por un solo dispositivo, es decir, el dispositivo 540A en el ejemplo de la FIG. las técnicas pueden ser implementadas o realizadas de otro modo por un sistema que comprende múltiples dispositivos, donde cada uno de estos dispositivos puede incluir cada uno uno o más de los diversos componentes o unidades descritos con más detalle a continuación. Por consiguiente, las técnicas no deben estar limitadas a este respecto al ejemplo de la FIG. 41.

[0738] Como se ilustra en el ejemplo de la FIG. 41, el dispositivo de decodificación de audio 540A comprende una unidad de extracción 542, una unidad de decodificación de audio 544, una unidad matemática 546 y una unidad de representación de audio 548. La unidad de extracción 542 representa una unidad configurada para extraer los coeficientes armónicos esféricos de fondo reducidos codificados 515B, los vectores U_{DIST} * SDIST codificados 515A y los vectores V^{T}_{DIST} 525E del flujo de bits 517. La unidad de extracción 542 emite los coeficientes armónicos esféricos de fondo reducidos codificados 515B y los vectores U_{DIST} * S_{DIST} codificados 515A a la unidad de decodificación de audio 544, a la vez que emite la matriz V^{T}_{DIST} 525E a la unidad matemática 546. A este respecto, la unidad de extracción 542 puede funcionar de una manera similar a la unidad de extracción 72 del dispositivo de decodificación de audio 24 mostrado en el ejemplo de la FIG. 5.

[0739] La unidad de decodificación de audio 544 representa una unidad para decodificar los datos de audio codificados (a menudo de acuerdo con un esquema de decodificación de audio recíproco, tal como un esquema de decodificación AAC) para recuperar los vectores U_{DIST} * S_{DIST} 527 y los coeficientes armónicos esféricos de fondo reducidos 529. La unidad de decodificación de audio 544 emite los vectores U_{DIST} * S_{DIST} 527 y los coeficientes armónicos esféricos de fondo reducidos 529 a la unidad matemática 546. A este respecto, la unidad de decodificación de audio 544 puede funcionar de una manera similar a la unidad de decodificación psicoacústica 80 del dispositivo de decodificación de audio 24 mostrado en el ejemplo de la FIG. 5.

[0740] La unidad matemática 546 puede representar una unidad configurada para realizar multiplicación y adición matricial (así como, en algunos ejemplos, cualquier otra operación matricial matemática). La unidad matemática 546 puede realizar primero una multiplicación matricial de los vectores U_{DIST} * S_{DIST} 527 por la matriz V^{T}_{DIST} 525E. La unidad matemática 546 puede entonces añadir el resultado de la multiplicación de los vectores U_{DIST} * S_{DIST} 527 por la matriz V^{T}_{DIST} 525E por los coeficientes armónicos esféricos de fondo reducidos 529 (que, de nuevo, pueden referirse al resultado de la multiplicación de la matriz U_{BG} 525D por la matriz SBG 525B y luego por la matriz V^{T}_{BG} 525F) al resultado de la multiplicación matricial de los vectores U_{DIST} * S_{DIST} 527 por la matriz V^{T}_{DIST} 525E para generar la versión reducida de los coeficientes armónicos esféricos originales 11, que se representa como coeficientes armónicos esféricos recuperados 547. La unidad matemática 546 puede emitir los coeficientes armónicos esféricos recuperados 547 a la unidad de representación de audio 548. A este respecto, la unidad matemática 546 puede funcionar de una manera similar a la unidad de formulación de primer plano 78 y la unidad de formulación de coeficiente HOA 82 del dispositivo de decodificación de audio 24 mostrado en el ejemplo de la FIG. 5.

[0741] La unidad de representación de audio 548 representa una unidad configurada para representar los canales 549A-549N (los "canales 549", que también pueden denominarse generalmente los "datos de audio multicanal 549" o las "señales de altavoz 549"). La unidad de representación de audio 548 puede aplicar una transformación (a menudo expresada en forma de matriz) a los coeficientes armónicos esféricos recuperados 547. Debido a que los coeficientes armónicos esféricos recuperados 547 describen el campo sonoro en tres dimensiones, los coeficientes armónicos esféricos recuperados 547 representan un formato de audio que facilita la reproducción de los datos de audio multicanal 549A de una manera que es capaz de alojar la mayoría de las geometrías de decodificador-altavoz local (que puede referirse a la geometría de los altavoces que reproducirán los datos de audio multicanal 549). Se ha descrito anteriormente más información sobre la reproducción de los datos de audio multicanal 549A con respecto a la FIG. 48.

[0742] Aunque se describe en el contexto de los datos de audio multicanal 549A que son datos de audio multicanal de sonido envolvente 549, la unidad de reproducción de audio 48 también puede realizar una forma de binauralización para binauralizar los coeficientes armónicos esféricos recuperados 549A y generar así dos canales binauralmente reproducidos 549. Por consiguiente, las técnicas no deben limitarse a formas de sonido envolvente de datos de audio multicanal, pero pueden incluir datos de audio multicanal binauralizados.

[0743] Las diversas cláusulas enumeradas a continuación pueden presentar diversos aspectos de las técnicas descritas en esta descripción.

Cláusula 132567-1B. Un dispositivo, tal como el dispositivo de decodificación de audio 540, que comprende: uno o más procesadores configurados para determinar uno o más primeros vectores que describen componentes distintos del campo sonoro y uno o más segundos vectores que describen componentes de fondo del campo

acústico, tanto el uno o más primeros vectores como el uno o más segundos vectores generados al menos realizando una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos.

Cláusula 132567-2B. El dispositivo de la cláusula 132567-1B, en el que uno o más primeros vectores comprenden uno o más vectores $U_{DIST} * S_{DIST}$ codificados en audio que, antes de la codificación de audio, se generaron multiplicando uno o más vectores U_{DIST} codificados de audio de una matriz U por uno o más vectores S_{DIST} de una matriz S, en el que la matriz U y la matriz S se generan al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, y en el que uno o más procesadores están configurados adicionalmente para decodificar audio de uno o más vectores $U_{DIST} * S_{DIST}$ codificados de audio para generar una versión decodificada de audio de uno o más vectores $U_{DIST} * S_{DIST}$ codificados de audio.

Cláusula 132567-3B. El dispositivo de la cláusula 132567-1B, en el que uno o más primeros vectores comprenden uno o más vectores $U_{DIST} * S_{DIST}$ codificados de audio que, antes de la codificación de audio, se generaron multiplicando uno o más vectores U_{DIST} codificados de audio de una matriz U por uno o más vectores S_{DIST} de una matriz S, y uno o más vectores V^T_{DIST} de una transposición de una matriz V, en el que la matriz U y la matriz S y la matriz V se generan al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, y en el que uno o más procesadores están configurados adicionalmente para decodificar audio uno o más vectores $U_{DIST} * S_{DIST}$ codificados de audio para generar una versión decodificada de audio de uno o más vectores $U_{DIST} * S_{DIST}$ codificados de audio.

Cláusula 132567-4B. El dispositivo de la cláusula 132567-3B, en el que el uno o más procesadores están configurados además para multiplicar los vectores $U_{DIST} * S_{DIST}$ por los vectores V^T_{DIST} para recuperar los de la pluralidad de armónicos esféricos representativos de los distintos componentes del campo sonoro.

Cláusula 132567-5B. El dispositivo de la cláusula 132567-1B, en el que el uno o más segundos vectores comprenden uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ codificados en audio que, antes de la codificación de audio, se generaron multiplicando los vectores U_{BG} incluidos dentro de una matriz U por los vectores S_{BG} incluidos dentro de una matriz S y luego por vectores V^T_{BG} incluidos dentro de una transposición de una matriz V, y en la que la matriz S, la matriz U y la matriz V se generaron, al menos, realizando una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos.

Cláusula 132567-6B. El dispositivo de la cláusula 132567-1B, en el que el uno o más segundos vectores comprenden uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ codificados en audio que, antes de la codificación de audio, se generaron multiplicando los vectores U_{BG} incluidos dentro de una matriz U por los vectores S_{BG} incluidos dentro de una matriz S y luego por los vectores V^T_{BG} incluidos dentro de una transposición de una matriz V, y en el que la matriz S, la matriz U y la matriz V se generaron al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, y en el que el uno o más procesadores están configurados adicionalmente para decodificar en audio uno o más vectores $_{BG} * S_{BG} * V^T_{BG}$ codificados en audio para generar uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$.

Cláusula 132567-7B. El dispositivo de la cláusula 132567-1B, en el que uno o más primeros vectores comprenden uno o más vectores $U_{DIST} * S_{DIST}$ codificados en audio que, antes de la codificación de audio, se generaron multiplicando uno o más vectores U_{DIST} codificados en audio de una matriz U por uno o más vectores S_{DIST} de una matriz S, y uno o más vectores V^T_{DIST} de una transposición de una matriz V, en el que la matriz U, la matriz S y la matriz V se generaron al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, y en el que el uno o más procesadores están configurados adicionalmente para decodificar en audio uno o más vectores $U_{DIST} * S_{DIST}$ codificados en audio para generar uno o más vectores $U_{DIST} * S_{DIST}$ y multiplicar los vectores $U_{DIST} * S_{DIS}T$ por los vectores V^T_{DIST} para recuperar los de la pluralidad de coeficientes armónicos esféricos que describen los distintos componentes del campo sonoro, en el que uno o más segundos vectores comprenden uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ codificados en audio que, antes de la codificación de audio, se generaron multiplicando los vectores U_{BG} incluidos en la matriz U por los vectores S_{BG} incluidos en la matriz S y luego por los vectores V^T_{BG} incluidos dentro de la transposición de la matriz V, y en el que el uno o más procesadores están configurados adicionalmente para decodificar en audio uno o más vectores $U_{BG} * S_{BG} * V^T_{BG}$ codificados en audio para recuperar al menos una parte de la pluralidad de coeficientes armónicos esféricos que describen componentes de fondo del campo sonoro y añadir la pluralidad de coeficientes armónicos esféricos que describen los distintos componentes del campo sonoro a la, al menos parte, de la pluralidad de coeficientes armónicos esféricos que describen componentes de fondo del campo sonoro para generar una versión reconstruida de la pluralidad de coeficientes armónicos esféricos.

Cláusula 132567-8B. El dispositivo de la cláusula 132567-1B, en el que uno o más primeros vectores comprenden uno o más vectores $U_{DIST} * S_{DIST}$ que, antes de la codificación de audio, se generaron multiplicando uno o más vectores U_{DIST} codificados en audio de una matriz U por uno o más vectores S_{DIST} de una matriz S y uno o más vectores V^T_{DIST} de una transposición de una matriz V, en donde la matriz U, la matriz S y la matriz V se generaron al menos realizando la descomposición en valores singulares con respecto a la pluralidad de

coeficientes armónicos esféricos, y en el que el uno o más procesadores están configurados además para determinar un valor D que indica el número de vectores a extraer de un flujo de bits para formar uno o más vectores U_{DIST} * S_{DIST} y el uno o más vectores V^T_{DIST}.

Cláusula 132567-9B. El dispositivo de la cláusula 132567-10B, en el que uno o más primeros vectores comprenden uno o más vectores U_{DIST} * S_{DIST} que, antes de la codificación de audio, se generaron multiplicando uno o más vectores U_{DIST} codificados en audio de una matriz U por uno o más vectores S_{DIST} de una matriz S y uno o más vectores V^T_{DIST} de una transposición de una matriz V, en donde la matriz U, la matriz S y la matriz V se generaron al menos realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, y en el que el uno o más procesadores están configurados además para determinar un valor D en una base de trama de audio a trama de audio que indica el número de vectores a extraer de un flujo de bits para formar uno o más vectores U_{DIST} * S_{DIST} y el uno o más vectores V^T_{DIST}.

Cláusula 132567-1G. Un dispositivo, tal como el dispositivo de decodificación de audio 540, que comprende: uno o más procesadores configurados para determinar uno o más primeros vectores que describen componentes distintos de un campo sonoro y uno o más segundos vectores que describen componentes de fondo del campo acústico, tanto el uno o más primeros vectores como el uno o más segundos vectores generados al menos realizando una descomposición en valores singulares con respecto a la datos de audio multicanal representativos de al menos una porción del campo sonoro.

Cláusula 132567-2G. El dispositivo de la cláusula 132567-1G, en el que los datos de audio multicanal comprenden una pluralidad de coeficientes armónicos esféricos.

Cláusula 132567-3G. El dispositivo de la cláusula 132567-2G, en el que el uno o más procesadores están configurados adicionalmente para realizar cualquier combinación de la cláusula 132567-2B a la cláusula 132567-9B.

[0744] De cada una de las diversas cláusulas descritas anteriormente, debe entenderse que cualquiera de los dispositivos de decodificación de audio 540A-540D puede realizar un procedimiento o comprender de otro modo medios para realizar cada paso del procedimiento para el cual está configurado el dispositivo de decodificación de audio 540A-540D. En algunos casos, estos medios pueden comprender uno o más procesadores. En algunos casos, 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 el cual los dispositivos de decodificación de audio 540A-540D se han configurado.

[0745] Por ejemplo, una cláusula 132567-10B puede obtenerse a partir de la anterior cláusula 132567-1B que será un procedimiento que comprende: determinar uno o más primeros vectores que describen componentes distintos de un campo sonoro y uno o más segundos vectores que describen componentes de fondo del campo acústico, tanto el uno o más primeros vectores como el uno o más segundos vectores generados al menos realizando una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos que representan el campo sonoro.

[0746] Como otro ejemplo, una cláusula 132567-11B puede derivarse de la cláusula 132567-1B anterior para ser un dispositivo, tal como el dispositivo de decodificación de audio 540, que comprende medios para determinar uno o más primeros vectores que describen componentes distintos del campo sonoro y uno o más segundos vectores que describen componentes de fondo del campo sonoro, tanto uno o más primeros vectores como uno o más segundos vectores generados al menos realizando una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos; y medios para almacenar el uno o más primeros vectores y uno o más segundos vectores.

[0747] Como otro ejemplo más, una cláusula 132567-12B puede derivarse de la cláusula 132567-1B anterior como un medio de almacenamiento no transitorio legible por ordenador que tiene almacenadas en el mismo instrucciones que, cuando se ejecutan, hacen que uno o más procesadores determinen uno o más primeros vectores que describen componentes distintos de un campo sonoro y uno o más segundos vectores que describen componentes de fondo del campo sonoro, tanto uno o más primeros vectores como uno o más segundos vectores generados al menos realizando una descomposición en valores singulares con respecto a una pluralidad de coeficientes armónicos esféricos incluidos dentro de datos de audio ambisónicos de orden superior que describen el campo sonoro.

[0748] Asimismo, pueden derivarse diversas cláusulas de las cláusulas 132567-2B a 132567-9B para los diversos dispositivos, procedimientos y medios de almacenamiento no transitorios legibles por ordenador derivados como se ha ilustrado anteriormente. Lo mismo puede hacerse para las otras cláusulas enumeradas a lo largo de esta divulgación.

[0749] La FIG. 41B es un diagrama de bloques que ilustra un ejemplo de dispositivo de decodificación de audio 540B que puede realizar diversos aspectos de las técnicas descritas en esta descripción para decodificar los coeficientes armónicos esféricos que describen campos sonoros bidimensionales o tridimensionales. El dispositivo de decodificación de audio 540B puede ser similar al dispositivo de decodificación de audio 540, excepto que, en algunos ejemplos, la unidad de extracción 542 puede extraer vectores V^T_{DIST} reordenados 539 en lugar de vectores V^T_{DIST} 525E. En otros ejemplos, la unidad de extracción 542 puede extraer los vectores V^T_{DIST} 525E y luego reordenar estos vectores V^T_{DIST} 525E basándose en la información de reordenación especificada en el flujo de bits o inferida (mediante el análisis de otros vectores) para determinar los vectores V^T_{DIST} reordenados 539. A este respecto, la unidad de extracción 542 puede funcionar de una manera similar a la unidad de extracción 72 del dispositivo de decodificación de audio 24 mostrado en el ejemplo de la FIG. 5. En cualquier caso, la unidad de extracción 542 puede emitir los vectores V^T_{DIST} reordenados 539 a la unidad matemática 546, donde el proceso descrito anteriormente con respecto a la recuperación de los coeficientes armónicos esféricos puede realizarse con respecto a estos vectores V^T_{DIST} reordenados 539.

[0750] De esta manera, las técnicas pueden permitir que el dispositivo de decodificación de audio 540B decodifique audio de uno o más vectores reordenados representativos de componentes distintos de un campo sonoro, habiéndose reordenado uno o más de los vectores reordenados para facilitar la compresión del uno o más vectores. En estos y otros ejemplos, el dispositivo de decodificación de audio 540B puede recombinar el uno o más vectores reordenados con uno o más vectores adicionales reordenados adicionales para recuperar coeficientes armónicos esféricos representativos de componentes distintos del campo sonoro. En estos y otros ejemplos, el dispositivo de decodificación de audio 540B puede entonces recuperar una pluralidad de coeficientes armónicos esféricos basados en los coeficientes armónicos esféricos representativos de componentes distintos del campo sonoro y coeficientes armónicos esféricos representativos de componentes de fondo del campo sonoro.

[0751] Es decir, diversos aspectos de las técnicas pueden proporcionar el dispositivo de decodificación de audio 540B configurado para decodificar uno o más vectores reordenados de acuerdo con las siguientes cláusulas.

Cláusula 133146-1F. Un dispositivo, tal como el dispositivo de codificación de audio 540B, que comprende: uno o más procesadores configurados para determinar varios vectores correspondientes a componentes en el campo sonoro.

Cláusula 133146-2F. El dispositivo de la cláusula 133146-1F, en el que el uno o más procesadores están configurados para determinar el número de vectores después de realizar la reducción de orden de acuerdo con cualquier combinación de los casos descritos anteriormente.

Cláusula 133146-3F. El dispositivo de la cláusula 133146-1F, en el que el uno o más procesadores están configurados adicionalmente para realizar la reducción de orden de acuerdo con cualquier combinación de los casos descritos anteriormente.

Cláusula 133146-4F. El dispositivo de la cláusula 133146-1F, en el que el uno o más procesadores están configurados para determinar el número de vectores a partir de un valor especificado en un flujo de bits, y en el que uno o más procesadores están configurados adicionalmente para analizar el flujo de bits basándose en el número determinado de vectores para identificar uno o más vectores en el flujo de bits que representan componentes distintos del campo sonoro.

Cláusula 133146-5F. El dispositivo de la cláusula 133146-1F, en el que el uno o más procesadores están configurados para determinar el número de vectores a partir de un valor especificado en un flujo de bits, y en el que uno o más procesadores están configurados adicionalmente para analizar el flujo de bits basándose en el número determinado de vectores para identificar uno o más vectores en el flujo de bits que representan componentes de fondo del campo sonoro.

Cláusula 133143-1C. Un dispositivo, tal como el dispositivo de decodificación de audio 540B, que comprende: uno o más procesadores configurados para reordenar uno o más vectores reordenados representativos de componentes distintos de un campo sonoro.

Cláusula 133143-2C. El dispositivo de la cláusula 133143-1C, en el que el uno o más procesadores están configurados además para determinar los uno o más vectores reordenados, y determinar información de reordenación que describe cómo se reordenaron uno o más de los vectores reordenados, en el que el uno o más procesadores están configurados adicionalmente para, al reordenar el uno o más vectores reordenados, reordenar el uno o más vectores reordenados basándose en la información de reordenación determinada.

Cláusula 133143-3C. El dispositivo de 1C, en el que el uno o más vectores reordenados comprenden uno o más primeros vectores reordenados indicados por cualquier combinación de las reivindicaciones 1A-18A o cualquier combinación de las reivindicaciones 1B-19B, y en el que uno o más primeros vectores se determinan de acuerdo con el procedimiento indicado por cualquier combinación de las reivindicaciones 1A-18A o cualquier combinación

de las reivindicaciones 1B-19B.

Cláusula 133143-4D. Un dispositivo, tal como el dispositivo de decodificación de audio 540B, que comprende: uno o más procesadores configurados para decodificar en audio uno o más vectores reordenados representativos de componentes distintos de un campo sonoro, habiéndose reordenado el uno o más vectores reordenados para facilitar la compresión de uno o más vectores.

Cláusula 133143-5D. El dispositivo de la cláusula 133143-4D, en el que el uno o más procesadores están configurados además para recombinar el uno o más vectores reordenados con uno o más vectores adicionales reordenados para recuperar coeficientes armónicos esféricos representativos de componentes distintos del campo sonoro.

Cláusula 133143-6D. El dispositivo de la cláusula 133143-5D, en el que el uno o más procesadores están configurados además para recuperar una pluralidad de coeficientes armónicos esféricos basados en los coeficientes armónicos esféricos representativos de componentes distintos del campo sonoro y coeficientes armónicos esféricos representativos de los componentes de fondo del campo sonoro.

Cláusula 133143-1E. Un dispositivo, tal como el dispositivo de decodificación de audio 540B, que comprende: uno o más procesadores configurados para reordenar uno o más vectores para generar uno o más primeros vectores reordenados y facilitar así la codificación por un codificador de audio heredado, en el que uno o más vectores describen componentes distintos de un campo sonoro, y codificar en audio el uno o más vectores reordenados utilizando el codificador de audio heredado para generar una versión codificada del uno o más vectores reordenados.

Cláusula 133143-2E. El dispositivo de 1E, en el que el uno o más vectores reordenados comprenden uno o más primeros vectores reordenados indicados por cualquier combinación de las reivindicaciones 1A-18A o cualquier combinación de las reivindicaciones 1B-19B, y en el que uno o más primeros vectores se determinan de acuerdo con el procedimiento indicado por cualquier combinación de las reivindicaciones 1A-18A o cualquier combinación de las reivindicaciones 1B-19B.

[0752] La FIG. 41C es un diagrama de bloques que ilustra otro dispositivo de codificación de audio ejemplar 540C. El dispositivo de decodificación de audio 540C puede representar cualquier dispositivo capaz de decodificar datos de audio, tales como un ordenador de sobremesa, un ordenador portátil, una estación de trabajo, una tableta o un ordenador de pizarra, un dispositivo de grabación de audio dedicado, un teléfono móvil (incluidos los llamados "teléfonos inteligentes"), un dispositivo de reproductor multimedia personal, un dispositivo de juego personal o cualquier otro tipo de dispositivo capaz de decodificar datos de audio.

[0753] En el ejemplo de la FIG. 41C, el dispositivo de decodificación de audio 540C realiza un proceso de decodificación de audio que es recíproco al proceso de codificación de audio realizado por cualquiera de los dispositivos de codificación de audio 510B-510E con la excepción de realizar la reducción de orden (como se ha descrito anteriormente con respecto a los ejemplos de las FIGS. 40B-40J), que es, en algunos ejemplos, utilizado por el dispositivo de codificación de audio 510B-510J para facilitar la eliminación de datos irrelevantes extraños.

[0754] Aunque se muestra como un dispositivo individual, es decir, el dispositivo 540C en el ejemplo de la FIG. 41C, los diversos componentes o unidades a los que se hace referencia a continuación como incluidos en el dispositivo 540C pueden formar dispositivos separados que son externos al dispositivo 540C. En otras palabras, aunque se describe en esta descripción como realizada por un solo dispositivo, es decir, el dispositivo 540C en el ejemplo de la FIG. 41C, las técnicas pueden ser implementadas o realizadas de otro modo por un sistema que comprende múltiples dispositivos, donde cada uno de estos dispositivos puede incluir cada uno uno o más de los diversos componentes o unidades descritos con más detalle a continuación. Por consiguiente, las técnicas no deben estar limitadas a este respecto al ejemplo de la FIG. 41C.

[0755] Además, el dispositivo de codificación de audio 540C puede ser similar al dispositivo de codificación de audio 540B. Sin embargo, la unidad de extracción 542 puede determinar uno o más vectores $V^{T}_{PEQUEÑO}$ 521 del flujo de bits 517 en lugar de vectores $V^{T}_{Q_DIST}$ reordenados 539 o vectores V^{T}_{DIST} 525E (como es el caso descrito con respecto al dispositivo de codificación de audio 510 de la FIG. 40). Como resultado, la unidad de extracción 542 puede pasar los vectores $V^{T}_{PEQUEÑO}$ 521 a la unidad matemática 546.

[0756] Además, la unidad de extracción 542 puede determinar coeficientes armónicos esféricos de fondo modificados codificados de audio 515B' desde la corriente de bits 517, pasando estos coeficientes 515B' a la unidad de decodificación de audio 544, que puede decodificar en audio los coeficientes armónicos esféricos de fondo modificados codificados 515B para recuperar los coeficientes armónicos esféricos de fondo modificados 537. La unidad de decodificación de audio 544 puede pasar estos coeficientes armónicos esféricos de fondo modificados 537 a la unidad matemática 546.

[0757] La unidad matemática 546 puede entonces multiplicar los vectores $U_{DIST} * S_{DIST}$ decodificados de audio (y

posiblemente no ordenados) 527' por uno o más vectores $V^T_{PEQUEÑO}$ 521 para recuperar los coeficientes armónicos esféricos distintos de orden superior. La unidad matemática 546 puede entonces añadir los coeficientes armónicos esféricos distintos de orden superior a los coeficientes armónicos esféricos de fondo modificados 537 para recuperar la pluralidad de los coeficientes armónicos esféricos 511 o algún derivado de los mismos (que puede ser un derivado debido a la reducción de orden realizada en la unidad codificadora 510E).

[0758] De este modo, las técnicas pueden permitir que el dispositivo de decodificación de audio 540C determine, a partir de un flujo de bits, al menos uno de uno o más vectores descompuestos a partir de coeficientes armónicos esféricos que se recombinaron con coeficientes armónicos esféricos de fondo para reducir la cantidad de bits que se requiere asignar a uno o más vectores en el flujo de bits, en el que los coeficientes armónicos esféricos describen un campo sonoro, y en el que los coeficientes armónicos esféricos de fondo describen uno o más componentes de fondo del mismo campo sonoro.

[0759] Varios aspectos de las técnicas pueden permitir a este respecto que el dispositivo de decodificación de audio 540C, en algunos casos, se configure para determinar, a partir de un flujo de bits, al menos uno de uno o más vectores descompuestos a partir de coeficientes armónicos esféricos que se recombinaron con coeficientes armónicos esféricos de fondo en el que los coeficientes armónicos esféricos describen un campo sonoro, y en el que los coeficientes armónicos esféricos de fondo describían uno o más componentes de fondo del mismo campo sonoro.

[0760] En estos y otros casos, el dispositivo de decodificación de audio 540C está configurado para obtener, a partir del flujo de bits, una primera porción de los coeficientes armónicos esféricos que tienen un orden igual a N_{BG}.

[0761] En estos y otros casos, el dispositivo de decodificación de audio 540C está configurado además para obtener, a partir del flujo de bits, una primera parte codificada de audio, teniendo los coeficientes armónicos esféricos un orden igual a N_{BG}, y decodificar en audio la primera porción codificada de audio de los coeficientes armónicos esféricos para generar una primera porción de los coeficientes armónicos esféricos.

[0762] En estos y otros casos, el al menos uno de uno o más vectores comprende uno o más vectores $V^T_{PEQUEÑO}$, habiéndose determinado uno o más vectores $V^T_{PEQUEÑO}$ a partir de una transposición de una matriz V generada realizando una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos.

[0763] En estos y otros casos, el al menos uno de uno o más vectores comprende uno o más vectores $V^T_{PEQUEÑO}$, habiéndose determinado uno o más vectores $V^T_{PEQUEÑO}$ a partir de una transposición de una matriz V generada realizando una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, y el dispositivo de decodificación de audio 540C está configurado adicionalmente para obtener, a partir del flujo de bits, uno o más vectores $U_{DIST} * S_{DIST}$ que se han derivado de una matriz U y una matriz S, ambos de los cuales se generaron realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, y multiplicando los vectores $U_{DIST} * S_{DIST}$ por los vectores $V^T_{PEQUEÑO}$.

[0764] En estos y otros casos, el al menos uno de uno o más vectores comprende uno o más vectores $V^T_{PEQUEÑO}$, habiéndose determinado uno o más vectores $V^T_{PEQUEÑO}$ a partir de una transposición de una matriz V generada realizando una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, y el dispositivo de decodificación de audio 540C está configurado adicionalmente para obtener, a partir del flujo de bits, uno o más vectores $U_{DIST} * S_{DIST}$ que se han derivados de una matriz U y una matriz S, ambos de los cuales se generaron realizando la descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, multiplicar los vectores $U_{DIST} * S_{DIST}$ por los vectores $V^T_{PEQUEÑO}$ para recuperar los coeficientes armónicos esféricos de fondo distintos de orden superior, y sumar los coeficientes armónicos esféricos de fondo que incluyen los coeficientes armónicos esféricos de fondo distinto de orden inferior a los coeficientes armónicos esféricos de fondo distinto de orden superior para recuperar, al menos en parte, la pluralidad de coeficientes armónicos esféricos.

[0765] En estos y otros casos, el al menos uno de uno o más vectores comprende uno o más vectores $V^T_{PEQUEÑO}$, habiéndose determinado uno o más vectores $V^T_{PEQUEÑO}$ a partir de una transposición de una matriz V generada realizando una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, y el dispositivo de decodificación de audio 540C está configurado adicionalmente para obtener, a partir del flujo de bits, uno o más vectores $U_{DIST} * S_{DIST}$ que han sido derivados de una matriz U y una matriz S, ambos de los cuales se generaron realizando la descomposición de valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos, multiplicar los vectores $U_{DIST} * S_{DIST}$ por los vectores $V^T_{PEQUEÑO}$ para recuperar coeficientes armónicos esféricos de fondo diferentes de orden superior, añadir los coeficientes esféricos de armónicos de fondo que incluyen los coeficientes armónicos esféricos de fondo distinto de orden inferior a los coeficientes armónicos esféricos de fondo distintos de orden superior para recuperar, al menos en parte, la pluralidad de coeficientes armónicos esféricos, y reproducir la pluralidad recuperada de coeficientes armónicos esféricos.

[0766] La FIG. 41D es un diagrama de bloques que ilustra otro dispositivo de codificación de audio ejemplar 540D. El dispositivo de decodificación de audio 540D puede representar cualquier dispositivo capaz de decodificar datos de

audio, tales como un ordenador de sobremesa, un ordenador portátil, una estación de trabajo, una tableta o un ordenador de pizarra, un dispositivo de grabación de audio dedicado, un teléfono móvil (incluidos los llamados "teléfonos inteligentes"), un dispositivo de reproductor multimedia personal, un dispositivo de juego personal o cualquier otro tipo de dispositivo capaz de decodificar datos de audio.

[0767] En el ejemplo de la FIG. 41D, el dispositivo de decodificación de audio 540D realiza un proceso de decodificación de audio que es recíproco al proceso de codificación de audio realizado por cualquiera de los dispositivos de codificación de audio 510B-510J con la excepción de realizar la reducción de orden (como se ha descrito anteriormente con respecto a los ejemplos de las FIGS. 40B-40J), que es, en algunos ejemplos, utilizado por los dispositivos de codificación de audio 510B-510J para facilitar la eliminación de datos irrelevantes extraños.

[0768] Aunque se muestra como un dispositivo individual, es decir, el dispositivo 540D en el ejemplo de la FIG. 41D, los diversos componentes o unidades a los que se hace referencia a continuación como incluidos en el dispositivo 540D pueden formar dispositivos separados que son externos al dispositivo 540D. En otras palabras, aunque se describe en esta descripción como realizada por un solo dispositivo, es decir, el dispositivo 540D en el ejemplo de la FIG. 41D, las técnicas pueden ser implementadas o realizadas de otro modo por un sistema que comprende múltiples dispositivos, donde cada uno de estos dispositivos puede incluir cada uno uno o más de los diversos componentes o unidades descritos con más detalle a continuación. Por consiguiente, las técnicas no deben estar limitadas a este respecto al ejemplo de la FIG. 41D.

[0769] Además, el dispositivo de decodificación de audio 540D puede ser similar al dispositivo de decodificación de audio 540B, excepto que el dispositivo de decodificación de audio 540D realiza una descompresión V adicional que es generalmente recíproca a la compresión realizada por la unidad de compresión V 552 descrita anteriormente con respecto a la FIG. 40I. En el ejemplo de la FIG. 41D, la unidad de extracción 542 incluye una unidad de descompresión V 555 que realiza esta descompresión V de los componentes espaciales comprimidos 539' incluidos en el flujo de bits 517 (y generalmente se especifica de acuerdo con el ejemplo mostrado en una de las FIGS. 10B y 10C). La unidad de descompresión V 555 puede descomprimir vectores V^{T}_{DIST} 539 basándose en la siguiente ecuación:

$$\hat{v}_q = \begin{cases} 0, & if\ cid = 0 \\ sgn * (2^{cid-1} + residual), & if\ cid \neq 0 \end{cases}$$

En otras palabras, la unidad de descompresión V 555 puede analizar primero el valor de *nbits* del flujo de bits 517 e identificar el conjunto apropiado de cinco tablas de códigos Huffman para usar al decodificar el código Huffman representativo de *cid*. Basándose en el modo de predicción y en la información de codificación de Huffman especificada en el flujo de bits 517 y posiblemente en el orden del elemento del componente espacial con respecto a los otros elementos del componente espacial, la unidad de descompresión V 555 puede identificar la correcta de las cinco tablas Huffman definidas para el valor de *nbits* analizados. Utilizando esta tabla Huffman, la unidad de descompresión V 555 puede decodificar el valor *cid* del código Huffman. La unidad de descompresión V 555 puede entonces analizar el bit de signo y el código de bloque residual, decodificando el código de bloque residual para identificar el *residuo*. De acuerdo con la ecuación anterior, la unidad de descompresión V 555 puede decodificar uno de los vectores V^{T}_{DIST} 539.

[0770] Lo anterior puede resumirse en la siguiente tabla de sintaxis:

Tabla - Vectores decodificados

Sintaxis	Nº de bits	Mnemónica

decodeVVec(i)
{
switch codedVVecLength {
caso 0: //Vector completo
VVecLength = NumOfHoaCoeffs;
para (m=0; m< VVecLength; ++m){ VecCoeff[m] = m+1; }
break;
caso 1: //se eliminan los órdenes inferiores
VVecLength = NumOfHoaCoeffs –
MinNumOfCoeffsForAmbHOA;
 	
para (m=0; m< VVecLength; ++m){
VecCoeff[m] = m + MinNumOfCoeffsForAmbHOA +
1;
}		

```
break;
    caso 2:
     VVecLength = NumOfHoaCoeffs–
MinNumOfCoeffsForAmbHOA - NumOfAddAmbHoaChan;
     n = 0;
     para (m=0;m<NumOfHoaCoeffs–
MinNumOfCoeffsForAmbHOA; ++m){
       c = m + MinNumOfCoeffsForAmbHOA + 1;
       if (ismember(c, AmbCoeffIdx) == 0){
         VecCoeff[n] = c;
         n++;
       }
     }
     break;
    caso 3:
     VVecLength = NumOfHoaCoeffs -
NumOfAddAmbHoaChan;
     n = 0;
     para (m=0; m<NumOfHoaCoeffs; ++m){
       c = m + 1;
       if (ismember(c, AmbCoeffIdx) == 0){
         VecCoeff[n] = c;
         n++;
       }
     }
}
if (NbitsQ[i] == 5) { /* cuantizador uniforme */
     para (m=0; m< VVecLength; ++m){
       VVec(k)[i][m] = (VecValue /128.0) – 1.0;                    8          uimsbf
     }
}
else { /* decodificación Huffman */
     para (m=0; m< VVecLength; ++m){
       Idx = 5;
       If(CbFlag[i] == 1){
         idx = (min(3, max(I, ceil(sqrt(VecCoeff[m]) – 1)));
       }
       else if (PFlag[i] == 1) {idx = 4;}
         cid =                                                     dinámico 1 huffDecode
huffDecode(huffmannTable[NbitsQ].codebook[idx]; huffVal);
       if(cid>0) {
         aVal = sgn = (sgnVal * 2) – 1;                            1          bslbf
        if (cid > 1) {
         aVal = sgn * (2.0A(cid -1) + intAddVal);                  cid-1      uimsbf
         }
       } else {aVal = 0.0;}
     }
    }
}
```

NOTA:
- La función del codificador para el cuantificador uniforme es min(255, round((x+ 1.0) *128.0))
- El n.º de bits para huffDecode nemónico es dinámico

[0771] En la tabla de sintaxis anterior, la primera instrucción de conmutación con los cuatro casos (casos 0-3) proporciona una forma mediante la cual se determina la longitud del vector V^T_{DIST} en términos del número de coeficientes. El primer caso, caso 0, indica que se especifican todos los coeficientes para los vectores V^T_{DIST}. El segundo caso, caso 1, indica que sólo se especifican los coeficientes del vector V^T_{DIST} que corresponden a un orden mayor que un MinNumOfCoeffsForAmbHOA, lo que puede indicar lo que se denomina (N_{DIST} + 1) - (N_{BG} + 1)

anteriormente. El tercer caso, caso 2, es similar al segundo caso, pero resta aún más los coeficientes identificados por NumOfAddAmbHoaChan, que indica una variable para especificar canales adicionales (donde "canales" se refiere a un coeficiente particular que corresponde a un cierto orden, combinación de suborden) correspondiente a un orden que excede el orden N_{BG}. El cuarto caso, caso 3, indica que únicamente se especifican aquellos coeficientes del vector V^{T}_{DIST} que quedan después de eliminar los coeficientes identificados por NumOfAddAmbHoaChan.

[0772] Después de esta instrucción de conmutación, la decisión de realizar una descuantización unificada es controlada por NbitsQ (o, como se ha indicado anteriormente, *nbits*), que si no es igual a 5, da como resultado la aplicación de la decodificación Huffman. El valor cid mencionado anteriormente es igual a los dos bits menos significativos del valor NbitsQ. El modo de predicción analizado anteriormente se designa como PFlag en la tabla de sintaxis anterior, mientras que el bit de información HT se designa como CbFlag en la tabla de sintaxis anterior. La sintaxis restante especifica cómo ocurre la decodificación de una manera sustancialmente similar a la descrita anteriormente.

[0773] De esta manera, las técnicas de esta divulgación pueden permitir que el dispositivo de decodificación de audio 540D obtenga un flujo de bits que comprenda una versión comprimida de un componente espacial de un campo sonoro, estando el componente espacial generando realizando una síntesis basada en vectores con respecto a una pluralidad de coeficientes armónicos esféricos, y descomprimir la versión comprimida del componente espacial para obtener el componente espacial.

[0774] Además, las técnicas pueden permitir que el dispositivo de decodificación de audio 540D descomprima 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 de armónicos esféricos.

[0775] De esta manera, el dispositivo de codificación de audio 540D puede realizar diversos aspectos de las técnicas expuestas a continuación con respecto a las siguientes cláusulas.

Cláusula 141541-1B. Un dispositivo que comprende:

uno o más procesadores configurados para obtener un flujo de bits que comprende una versión comprimida de un componente espacial de un campo de sonido, el componente espacial generado realizando una síntesis basada en vectores con respecto a una pluralidad de coeficientes de armónicos esféricos, y descomprimir la versión comprimida del componente espacial para obtener el componente espacial.

Cláusula 141541-2B. El dispositivo de la cláusula 141541-1B, en el que la versión comprimida del componente espacial está representada en el flujo de bits utilizando, al menos en parte, un campo que especifica un modo de predicción utilizado al comprimir el componente espacial, y en el que uno o más procesadores están configurados además para, al descomprimir la versión comprimida del componente espacial, descomprimir la versión comprimida del componente espacial basada, al menos en parte, en el modo de predicción para obtener el componente espacial.

Cláusula 141541-3B. El dispositivo de cualquier combinación de la cláusula 141541-1B y la cláusula 141541-2B, en el que la versión comprimida del componente espacial está representada en el flujo de bits utilizando, al menos en parte, información de tabla Huffman que especifica una tabla Huffman usada al comprimir el componente espacial, y en el que uno o más procesadores están configurados adicionalmente para, al descomprimir la versión comprimida del componente espacial, descomprimir la versión comprimida del componente espacial basada, al menos en parte, en la información de la tabla Huffman.

Cláusula 141541-4B. El dispositivo de cualquier combinación de la cláusula 141541-1B a la cláusula 141541-3B, en el que la versión comprimida del componente espacial está representada en el flujo de bits utilizando, al menos en parte, un campo que indica un valor que expresa un tamaño de paso de cuantización o una variable del mismo utilizado al comprimir el componente espacial, y en el que uno o más procesadores están configurados además para, al descomprimir la versión comprimida del componente espacial, descomprimir la versión comprimida del componente espacial basada, al menos en parte, en el valor.

Cláusula 141541-5B. El dispositivo de la cláusula 141541-4B, en el que el valor comprende un valor de nbits.

Cláusula 141541-6B. El dispositivo de cualquier combinación de las cláusulas 141541-4B y 141541-5B, en el que el flujo de bits comprende una versión comprimida de una pluralidad de componentes espaciales del campo sonoro cuya versión comprimida del componente espacial se incluye, en el que el valor expresa el tamaño de paso de cuantización o una variable del mismo utilizada cuando se comprime la pluralidad de componentes espaciales y en el que uno o más procesadores están configurados adicionalmente para, al descomprimir la versión comprimida del componente espacial, descomprimir la pluralidad de versión comprimida del componente espacial basada, al menos en parte, en el valor.

Cláusula 141541-7B. El dispositivo de cualquier combinación de la cláusula 141541-1B a la cláusula 141541-6B, en el que la versión comprimida del componente espacial está representada en el flujo de bits utilizando, al menos en parte, un código Huffman para representar un identificador de categoría que identifica una categoría de compresión a la que corresponde el componente espacial, y en el que uno o más procesadores están configurados además para, al descomprimir la versión comprimida del componente espacial, descomprimir la versión comprimida del componente espacial basada, al menos en parte, en el código Huffman.

Cláusula 141541-8B. El dispositivo de cualquier combinación de la cláusula 141541-1B a la cláusula 141541-7B, en el que la versión comprimida del componente espacial está representada en el flujo de bits utilizando, al menos en parte, un bit de signo que identifica si el componente espacial es un valor positivo o un valor negativo, y el que el uno o más procesadores están configurados además para, al descomprimir la versión comprimida del componente espacial, descomprimir la versión comprimida del componente espacial basada, al menos en parte, en el bit de signo.

Cláusula 141541-9B. El dispositivo de cualquier combinación de la cláusula 141541-1B a la cláusula 141541-8B, en el que la versión comprimida del componente espacial está representada en el flujo de bits utilizando, al menos en parte, un código Huffman para representar un valor residual del componente espacial, y en el que el uno o más procesadores están configurados además para, al descomprimir la versión comprimida del componente espacial, descomprimir la versión comprimida del componente espacial basada, al menos en parte, en el código Huffman.

Cláusula 141541-10B. El dispositivo de cualquier combinación de la cláusula 141541-1B a la cláusula 141541-10B, en el que la síntesis basada en vectores comprende una descomposición en valores singulares.

[0776] Además, el dispositivo de decodificación de audio 540D puede configurarse para realizar diversos aspectos de las técnicas expuestas a continuación con respecto a las siguientes cláusulas.

Cláusula 141541-1C. Un dispositivo, tal como el dispositivo de decodificación de audio 540D, que comprende: uno o más procesadores configurados para descomprimir 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.

Cláusula 141541-2C. El dispositivo de cualquier combinación de la cláusula 141541-1C y la cláusula 141541-2C, en el que el uno o más procesadores están configurados adicionalmente para, al descomprimir la versión comprimida del componente espacial, obtener un identificador de categoría que identifica una categoría en la que se clasificó el componente espacial al comprimirse, obtener un signo que identifique si el componente espacial es un valor positivo o negativo, obtener un valor residual asociado con la versión comprimida del componente espacial, y descomprimir la versión comprimida del componente espacial basándose en el identificador de categoría, el signo y el valor residual.

Cláusula 141541-3C. El dispositivo de la cláusula 141541-2C, en el que el uno o más procesadores están configurados adicionalmente para, al obtener el identificador de categoría, obtener un código Huffman representativo del identificador de categoría, y decodificar el código Huffman para obtener el identificador de categoría.

Cláusula 141541-4C. El dispositivo de la cláusula 141541-3C, en el que el uno o más procesadores están configurados además para, al decodificar el código Huffman, identificar una tabla Huffman usada para decodificar el código Huffman basado en, al menos en parte, una posición relativa del componente espacial en un vector que especifica una pluralidad de componentes espaciales.

Cláusula 141541-5C. El dispositivo de cualquier combinación de la cláusula 141541-3C y la cláusula 141541-4C, en el que el uno o más procesadores están configurados además para, al decodificar el código Huffman, identificar una tabla Huffman usada para decodificar el código Huffman basado en, al menos en parte, un modo de predicción utilizado al comprimir el componente espacial.

Cláusula 141541-6C. El dispositivo de cualquier combinación de la cláusula 141541-3C a la cláusula 141541-5C, en el que el uno o más procesadores están configurados adicionalmente para, al decodificar el código Huffman, identificar una tabla Huffman usada para decodificar el código Huffman basándose, al menos en parte, en información de la tabla Huffman asociada con la versión comprimida del componente espacial.

Cláusula 141541-7C. El dispositivo de la cláusula 141541-3C, en el que el uno o más procesadores están configurados además para, al decodificar el código Huffman, identificar una tabla Huffman usada para decodificar el código Huffman basado en, al menos en parte, una posición relativa del componente espacial en un vector que especifica una pluralidad de componentes espaciales, un modo de predicción usado al comprimir el componente espacial, e información de tabla Huffman asociada a la versión comprimida del componente espacial.

Cláusula 141541-8C. El dispositivo de la cláusula 141541-2C, en el que el uno o más procesadores están configurados adicionalmente para, al obtener el valor residual, decodificar un código de bloque representativo del valor residual para obtener el valor residual.

Cláusula 141541-9C. El dispositivo de cualquier combinación de la cláusula 141541-1C a la cláusula 141541-8C, en el que la síntesis basada en vectores comprende una descomposición en valores singulares.

[0777] Además, el dispositivo de decodificación de audio 540D puede configurarse para realizar diversos aspectos de las técnicas expuestas a continuación con respecto a las siguientes cláusulas.

Cláusula 141541-1G. Un dispositivo, tal como el dispositivo de decodificación de audio 540D que comprende: uno o más procesadores configurados para identificar un libro de códigos Huffman para usar cuando se descomprime una versión comprimida de un componente espacial actual de una pluralidad de componentes espaciales comprimidos basándose en un orden de la versión comprimida del componente espacial actual con respecto a los restantes de la pluralidad de componentes espaciales comprimidos, realizando el componente espacial generado una síntesis basada en vectores con respecto a una pluralidad de coeficientes armónicos esféricos.

Cláusula 141541-2G. El dispositivo de la cláusula 141541-1G, en el que el uno o más procesadores están configurados además para realizar cualquier combinación de las etapas enumeradas en la cláusula 141541-1D a la cláusula 141541-10D y la cláusula 141541-1E a la cláusula 141541-9E.

[0778] Las FIGS. 42-42C son diagramas de bloques que ilustran la unidad de reducción de orden 528A mostrada en los ejemplos de las FIGS. 40B-40J en más detalle. La FIG. 42 es un diagrama de bloques que ilustra una unidad de reducción de orden 528, que puede representar un ejemplo de la unidad de reducción de orden 528A de las FIGS. 40B-40J. La unidad de reducción de órdenes 528A puede recibir o determinar de otro modo una velocidad de bits objetivo 535 y realizar una reducción de orden con respecto a los coeficientes armónicos esféricos de fondo 531 basándose únicamente en esta velocidad de bits objetivo 535. En algunos ejemplos, la unidad de reducción de órdenes 528A puede acceder a una tabla u otra estructura de datos utilizando la velocidad de bits objetivo 535 para identificar las órdenes y/o subórdenes que han de eliminarse de los coeficientes de armónicos esféricos de fondo 531 para generar coeficientes armónicos esféricos de fondo reducido 529.

[0779] De este modo, las técnicas pueden permitir que un dispositivo de codificación de audio, tal como los dispositivos de codificación de audio 510B-410J, realicen, sobre la base de una velocidad de bits objetivo 535, una reducción de orden con respecto a una pluralidad de coeficientes armónicos esféricos o descomposiciones de los mismos, tales como los coeficientes armónicos esféricos de fondo 531, para generar coeficientes armónicos esféricos reducidos 529 o las descomposiciones reducidas de los mismos, en el que la pluralidad de coeficientes armónicos esféricos representa un campo sonoro.

[0780] En cada una de las diversas instancias descritas anteriormente, debe entenderse que el dispositivo de decodificación de audio 540 puede realizar un procedimiento o comprender de otro modo medios para realizar cada paso del procedimiento para el que está configurado el dispositivo de decodificación de audio 540. En algunos casos, estos medios pueden comprender uno o más procesadores. En algunos casos, 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 el cual el dispositivo de decodificación de audio 540 se ha configurado para realizar.

[0781] La FIG. 42B es un diagrama de bloques que ilustra una unidad de reducción de orden 528B, que puede representar un ejemplo de la unidad de reducción de orden 528A de las FIGS. 40B-40J. En el ejemplo de la FIG. 42B, en lugar de realizar una reducción de orden basada solamente en una velocidad de bits objetivo 535, la unidad de reducción de orden 528B puede realizar una reducción de orden basada en un análisis de contenido de los coeficientes armónicos esféricos de fondo 531. La unidad de reducción de órdenes 528B puede incluir una unidad de análisis de contenido 536A que realiza este análisis de contenido.

[0782] En algunos ejemplos, la unidad de análisis de contenido 536A puede incluir una unidad de análisis espacial 536A que realiza una forma de análisis de contenido referida al análisis espacial. El análisis espacial puede implicar analizar los coeficientes de armónicos esféricos de fondo 531 para identificar información espacial que describe la forma u otras propiedades espaciales de los componentes de fondo del campo sonoro. En base a esta información espacial, la unidad de reducción de órdenes 528B puede identificar las órdenes y/o subórdenes que han de eliminarse de los coeficientes de armónicos esféricos de fondo 531 para generar coeficientes armónicos esféricos de fondo reducido 529.

[0783] En algunos ejemplos, la unidad de análisis de contenido 536A puede incluir una unidad de análisis de

difusión 536B que realiza una forma de análisis de contenido referida al análisis de difusión. El análisis de difusión puede implicar analizar los coeficientes de armónicos esféricos de fondo 531 para identificar información de difusión que describe la difluoroetilo de los componentes de fondo del campo sonoro. En base a esta información de difusión, la unidad de reducción de órdenes 528B puede identificar las órdenes y/o subórdenes que han de eliminarse de los coeficientes de armónicos esféricos de fondo 531 para generar coeficientes armónicos esféricos de fondo reducido 529.

[0784] Aunque se muestra incluyendo tanto la unidad de análisis espacial 536A como la unidad de análisis de difusión 36B, la unidad de análisis de contenido 536A puede incluir solamente la unidad de análisis espacial 536, solamente la unidad de análisis de difusión 536B o tanto la unidad de análisis espacial 536A como la unidad de análisis de difusión 536B. En algunos ejemplos, la unidad de análisis de contenido 536A puede realizar otras formas de análisis de contenido además o como una alternativa a uno o ambos del análisis espacial y el análisis de difusión. Por consiguiente, las técnicas descritas en esta divulgación no deben limitarse en este aspecto.

[0785] De esta manera, las técnicas pueden permitir que un dispositivo de codificación de audio, tal como los dispositivos de codificación de audio 510B-510J, realice, basándose en un análisis de contenido de una pluralidad de coeficientes armónicos esféricos o descomposiciones de los mismos que describen un campo sonoro, la reducción de orden con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos, para generar coeficientes armónicos esféricos reducidos o descomposiciones reducidas de los mismos.

[0786] En otras palabras, las técnicas pueden permitir que un dispositivo, tal como los dispositivos de codificación de audio 510B-510J, se configure de acuerdo con las siguientes cláusulas.

Cláusula 133146-1E. Un dispositivo, tal como cualquiera de los dispositivos de codificación de audio 510B-510J, que comprende uno o más procesadores configurados para realizar, basándose en un análisis de contenido de una pluralidad de coeficientes armónicos esféricos o descomposiciones de los mismos que describen un campo sonoro, una reducción de orden con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos para generar coeficientes armónicos esféricos reducidos o descomposiciones reducidas de los mismos.

Cláusula 133146-2E. El dispositivo de la cláusula 133146-1E, en el que el uno o más procesadores están configurados adicionalmente para, antes de realizar la reducción de orden, realizar una descomposición en valores singulares con respecto a la pluralidad de coeficientes armónicos esféricos para identificar uno o más primeros vectores que describen distintos componentes del campo sonoro y uno o más segundos vectores que identifican componentes de fondo del campo sonoro, y en el que el uno o más procesadores están configurados para realizar la reducción de orden con respecto a uno o más primeros vectores, uno o más segundos vectores, o tanto uno o más primeros vectores como uno o más segundos vectores.

Cláusula 133146-3E. El dispositivo de la cláusula 133146-1E, en el que el uno o más procesadores están configurados adicionalmente para realizar el análisis de contenido con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos.

Cláusula 133146-4E. El dispositivo de la cláusula 133146-3E, en el que el uno o más procesadores están configurados para realizar un análisis espacial con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos.

Cláusula 133146-5E. El dispositivo de la cláusula 133146-3E, en el que realizar el análisis de contenido comprende realizar un análisis de difusión con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos.

Cláusula 133146-6E. El dispositivo de la cláusula 133146-3E, en el que el uno o más procesadores están configurados para realizar un análisis espacial y un análisis de difusión con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos.

Cláusula 133146-7E. El dispositivo de la reivindicación 1, en el que el uno o más procesadores están configurados para realizar, basándose en el análisis de contenido de la pluralidad de coeficientes armónicos esféricos o sus descomposiciones y una velocidad de bits objetivo, la reducción de orden con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos para generar los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos.

Cláusula 133146-8E. El dispositivo de la cláusula 133146-1E, en el que el uno o más procesadores están configurados adicionalmente para codificar de en audio los coeficientes armónicos esféricos reducidos o las descomposiciones de los mismos.

Cláusula 133146-9E. El dispositivo de la cláusula 133146-1E, en el que el uno o más procesadores están configurados adicionalmente para codificar en audio los coeficientes armónicos esféricos reducidos o las

descomposiciones reducidas de los mismos, y generar un flujo de bits para incluir los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos.

Cláusula 133146-10E. El dispositivo de la cláusula 133146-1E, en el que el uno o más procesadores están configurados adicionalmente para especificar uno o más órdenes y/o una o más subórdenes de funciones de base esférica con las que los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de éstos se corresponden en un flujo de bits que incluye los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos.

Cláusula 133146-11E. El dispositivo de la cláusula 133146-1E, en el que los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos tienen valores menores que la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos.

Cláusula 133146-12E. El dispositivo de la cláusula 133146-1E, en el que el uno o más procesadores están configurados adicionalmente para eliminar aquellos de la pluralidad de coeficientes armónicos esféricos o vectores de las descomposiciones de los mismos que tienen un orden y/o suborden especificados para generar los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos.

Cláusula 133146-13E. El dispositivo de la cláusula 133146-1E, en el que el uno o más procesadores están configurados para poner a cero la pluralidad de coeficientes armónicos esféricos o aquellos vectores de la descomposición de los mismos, que tienen un orden y/o suborden especificado para generar los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos.

[0787] La FIG. 42C es un diagrama de bloques que ilustra una unidad de reducción de orden 528C, que puede representar un ejemplo de la unidad de reducción de orden 528A de las FIGS. 40B-40J. La unidad de reducción de órdenes 528C de la FIG. 42B es sustancialmente la misma que la unidad de reducción de órdenes 528B, pero puede recibir o determinar de otro modo una velocidad de bits objetivo 535 de la manera descrita anteriormente con respecto a la unidad de reducción de órdenes 528A de la FIG. 42, mientras que también realiza el análisis de contenido de la manera descrita anteriormente con respecto a la unidad de reducción de órdenes 528B de la FIG. 42B. La unidad de reducción de órdenes 528C puede realizar entonces una reducción de orden con respecto a los coeficientes de armónicos esféricos de fondo 531 basados en esta velocidad de bits objetivo 535 y el análisis de contenido.

[0788] De este modo, las técnicas pueden permitir que un dispositivo de codificación de audio, tal como los dispositivos de codificación de audio 510B-510J, realice un análisis de contenido con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos. Al realizar la reducción de orden, los dispositivos de codificación de audio 510B-510J pueden realizar, basándose en la velocidad de bits objetivo 535 y en el análisis de contenido, la reducción de orden con respecto a la pluralidad de coeficientes armónicos esféricos o las descomposiciones de los mismos para generar los coeficientes armónicos esféricos reducidos o las descomposiciones reducidas de los mismos.

[0789] Dado que se eliminan uno o más vectores, los dispositivos de codificación de audio 510B-510J pueden especificar el número de vectores en el flujo de bits como datos de control. Los dispositivos de codificación de audio 510B-510J pueden especificar este número de vectores en el flujo de bits para facilitar la extracción de los vectores del flujo de bits por el dispositivo de decodificación de audio.

[0790] La FIG. 44 es una ilustración del diagrama de operaciones ejemplares realizadas por el dispositivo de codificación de audio 410D para compensar el error de cuantización de acuerdo con diversos aspectos de las técnicas descritas en esta divulgación. En el ejemplo de la FIG. 44, la unidad matemática 526 del dispositivo de codificación de audio 510D se muestra como un bloque de líneas discontinuas para indicar que las operaciones matemáticas pueden ser realizadas por la unidad matemática 526 del dispositivo de decodificación de audio 510D.

[0791] Como se ilustra en el ejemplo de la FIG. 44, la unidad matemática 526 puede multiplicar primero los vectores U_{DIST} * S_{DIST} 527 por los vectores V^T_{DIST} 525E para generar coeficientes armónicos esféricos distintos (designados como "vectores H_{DIST} 630"). La unidad matemática 526 puede entonces dividir los vectores H_{DIST} 630 por la versión cuantificada de los vectores V^T_{DIST} 525E (que se indican, de nuevo, como "vectores $V^T_{Q_DIST}$ 525G"). La unidad matemática 526 puede realizar esta división determinando una pseudoinvertida de los vectores $V^T_{Q_DIST}$ 525G y multiplicando luego los vectores H_{DIST} por la pseudoinvertida de los vectores $V^T_{Q_DIST}$ 525G, emitiendo una versión compensada de error de U_{DIST} * S_{DIST} (que pueden abreviarse como "vectores US_{DIST}" o "US_{DIST}"). La versión compensada de error de US_{DIST} puede representarse como vectores US^*_{DIST} 527' en el ejemplo de la FIG. 44. De esta manera, las técnicas pueden proyectar eficazmente el error de cuantización, al menos en parte, a los vectores US_{DIST} 527, generando los vectores US^*_{DIST} 527'.

[0792] La unidad matemática 526 puede entonces restar los vectores US^*_{DIST} 527' de los vectores U_{DIST} * S_{DIST} 527 para determinar los vectores USERR 634 (que pueden representar al menos una porción del error debido a la cuantización proyectada en los vectores U_{DIST} * S_{DIST} 527). La unidad matemática 526 puede entonces multiplicar los

vectores USERR 634 por los vectores $V^T_{Q_DIST}$ 525G para determinar los vectores H_{ERR} 636. Matemáticamente, los vectores H_{ERR} 636 pueden ser equivalentes a los vectores US_{DIST} 527 - vectores US^*_{DIST} 527', cuyo resultado se multiplica entonces por los vectores V^T_{DIST} 525E. La unidad matemática 526 puede entonces añadir los vectores H_{ERR} 636 a los coeficientes armónicos esféricos de fondo 531 (designados como vectores H_{BG} 531 en el ejemplo de la FIG. 44) calculados multiplicando los vectores U_{BG} 525D por los vectores S_{BG} 525B y luego por los vectores V^T_{BG} 525F. La unidad matemática 526 puede sumar los vectores H_{ERR} 636 a los vectores H_{BG} 531, proyectando eficazmente al menos una porción del error de cuantización en los vectores H_{BG} 531 para generar los vectores H_{BG} 531' compensados. De esta manera, las técnicas pueden proyectar al menos una porción del error de cuantización en los vectores H_{BG} 531.

[0793] Las FIGS. 45 y 45B son diagramas que ilustran la interpolación de subtramas de partes de dos tramas de acuerdo con diversos aspectos de las técnicas descritas en esta divulgación. En el ejemplo de la FIG. 45, se muestra una primera trama 650 y una segunda trama 652. La primera trama 650 puede incluir coeficientes armónicos esféricos ("SH[1]") que se pueden descomponer en las matrices U[1], S[1] y V'[1]. La segunda trama 652 puede incluir coeficientes armónicos esféricos ("SH[2]"). Estos SH[1] y SH[2] pueden identificar diferentes tramas de los SHC 511 descritos anteriormente.

[0794] En el ejemplo de la FIG. 45B, la unidad de descomposición 518 del dispositivo de codificación de audio 510H mostrado en el ejemplo de la FIG. 40H puede separar cada una de las tramas 650 y 652 en cuatro subtramas 651A-651D y 653A-653D respectivas. La unidad de descomposición 518 puede entonces descomponer la primera subtrama 651A (designada como "SH[1,1]") de la trama 650 en las matrices U[1, 1], S[1, 1] y V[1, 1], emitiendo la matriz V[1, 1] 519' a la unidad de interpolación 550. La unidad de descomposición 518 puede entonces descomponer la segunda subtrama 653A (designada como "SH[2,1]") de la trama 652 en las matrices U[1, 1], S[1, 1] y V[1, 1], emitiendo la matriz V[2, 1] 519' a la unidad de interpolación 550. La unidad de descomposición 518 también puede emitir SH[1, 1], SH[1, 2], SH[1, 3] y SH[1, 4] de los SHC 11 y SH[2, 1] 2], SH[2, 3] y SH[2, 4] de los SHC 511 a la unidad de interpolación 550.

[0795] La unidad de interpolación 550 puede entonces realizar las interpolaciones identificadas en la parte inferior de la ilustración mostrada en el ejemplo de la FIG. 45B. Es decir, la unidad de interpolación 550 puede interpolar V'[1, 2] en base a V'[1, 1] y V'[2, 1]. La unidad de interpolación 550 también puede interpolar V'[1, 3] en base a V'[1, 1] y V'[2, 1]. Además, la unidad de interpolación 550 también puede interpolar V'[1, 4] en base a V'[1, 1] y V'[2, 1]. Estas interpolaciones pueden implicar una proyección de V'[1, 1] y V'[2, 1] en el dominio espacial, como se muestra en el ejemplo de las FIGS. 46-46E, seguido por una interpolación temporal y luego una proyección de nuevo en el dominio armónico esférico.

[0796] La unidad de interpolación 550 puede obtener a continuación U[1, 2]S[1, 2] multiplicando SH[1, 2] por (V'[1, 2])-1, U[1, 3]S[1, 3] multiplicando SH[1, 3] por $(V'[1, 3])^{-1}$, y U[1, 4]S[1, 4] multiplicando SH[1, 4] por $(V'[1, 4])^{-1}$. La unidad de interpolación 550 puede entonces reformar la trama en forma descompuesta que emite la matriz V 519, la matriz S 519B y la matriz U 519C.

[0797] Las FIGS. 46A-46E son diagramas que ilustran una sección transversal de una proyección de uno o más vectores de una versión descompuesta de una pluralidad de coeficientes armónicos esféricos que han sido interpolados de acuerdo con las técnicas descritas en esta divulgación. La FIG. 46A ilustra una sección transversal de una proyección de uno o más primeros vectores de una primera matriz V 19' que se ha descompuesto de SHC 511 de una primera subtrama a partir de una primera trama a través de un proceso de SVD. La FIG. 46B ilustra una sección transversal de una proyección de uno o más segundos vectores de una segunda matriz V 519' que se ha descompuesto de SHC 511 de una primera subtrama a partir de una segunda trama a través de un proceso de SVD.

[0798] La FIG. 46C ilustra una sección transversal de una proyección de uno o más vectores interpolados para una matriz V 519A representativa de una segunda subtrama de la primera trama, habiéndose interpolado estos vectores de acuerdo con las técnicas descritas en esta divulgación a partir de la matriz V 519' descompuesta a partir de la primera subtrama de la primera trama de los SHC 511 (es decir, uno o más vectores de la matriz V 519' mostrada en el ejemplo de la FIG. 46 en este ejemplo) y la primera subtrama de la segunda trama de los SHC 511 (es decir, uno o más vectores de la matriz V 519' mostrada en el ejemplo de la FIG. 46B en este ejemplo).

[0799] La FIG. 46D ilustra una sección transversal de una proyección de uno o más vectores interpolados para una matriz V 519A representativa de una tercera subtrama de la primera trama, habiéndose interpolado estos vectores de acuerdo con las técnicas descritas en esta divulgación a partir de la matriz V 519' descompuesta a partir de la primera subtrama de la primera trama de los SHC 511 (es decir, uno o más vectores de la matriz V 519' mostrada en el ejemplo de la FIG. 46 en este ejemplo) y la primera subtrama de la segunda trama de los SHC 511 (es decir, uno o más vectores de la matriz V 519' mostrada en el ejemplo de la FIG. 46B en este ejemplo).

[0800] La FIG. 46E ilustra una sección transversal de una proyección de uno o más vectores interpolados para una matriz V 519A representativa de una cuarta subtrama de la primera trama, habiéndose interpolado estos vectores de acuerdo con las técnicas descritas en esta divulgación a partir de la matriz V 519' descompuesta a partir de la primera subtrama de la primera trama de los SHC 511 (es decir, uno o más vectores de la matriz V 519' mostrada en

el ejemplo de la FIG. 46 en este ejemplo) y la primera subtrama de la segunda trama de los SHC 511 (es decir, uno o más vectores de la matriz V 519' mostrada en el ejemplo de la FIG. 46B en este ejemplo).

[0801] La FIG. 47 es un diagrama de bloques que ilustra, con más detalle, la unidad de extracción 542 de los dispositivos de decodificación de audio 540A-540D mostrados en los ejemplos de las FIGS. 41-41D. En algunos ejemplos, la unidad de extracción 542 puede representar un extremo frontal a lo que se puede denominar "decodificador integrado", que puede realizar dos o más esquemas de decodificación (donde mediante la realización de estos dos o más esquemas, el decodificador puede considerarse para "integrar" los dos o más esquemas). Como se ilustra en el ejemplo de la FIG. 44, la unidad de extracción 542 incluye un multiplexor 620 y las subunidades de extracción 622A y 622B ("subunidades de extracción 622"). El multiplexor 620 identifica aquellas matrices de SHC entramadas codificadas 547-547N a enviar a la subunidad de extracción 622A y a la subunidad de extracción 622B basándose en la indicación correspondiente de si las matrices de SHC entramadas codificadas asociadas 547-547N se generan a partir de un objeto de audio sintético o una grabación. Cada una de las subunidades de extracción 622A puede realizar un esquema de decodificación diferente (que se puede denominar "descompresión") que está, en algunos ejemplos, adaptado a los SHC generados a partir de un objeto de audio sintético o SHC generados a partir de una grabación. Cada una de las subunidades de extracción 622A puede realizar uno de estos esquemas de descompresión respectivos con el fin de generar tramas de SHC 547, que se emiten a SHC 547.

[0802] Por ejemplo, la unidad de extracción 622A puede realizar un esquema de descompresión para reconstruir SA a partir de una señal predominante (PS) usando la siguiente fórmula:

$$HOA = DirV \times PS,$$

donde DirV es un vector direccional (representativo de varias direcciones y anchuras), que puede ser transmitido a través de un canal lateral. La unidad de extracción 622B puede, en este ejemplo, realizar un esquema de descompresión que reconstruye la matriz HOA de la PS usando la siguiente fórmula:

$$HOA = \text{sqrt}(4\pi) * Ynm (theta, phi) * PS,$$

donde Ynm es la función armónica esférica y la información de theta y phi puede ser enviada a través del canal lateral.

[0803] A este respecto, las técnicas permiten a la unidad de extracción 358 seleccionar uno de una pluralidad de esquemas de descompresión basándose en la indicación de si se genera una versión comprimida de coeficientes armónicos esféricos representativos de un campo sonoro a partir de un objeto de audio sintético y descomprimir la versión comprimida de los coeficientes armónicos esféricos utilizando el esquema seleccionado de la pluralidad de esquemas de descompresión. En algunos ejemplos, el dispositivo comprende un decodificador integrado.

[0804] La FIG. 48 es un diagrama de bloques que ilustra la unidad de reproducción de audio 48 del dispositivo de decodificación de audio 540A-540D mostrado en los ejemplos de la FIGS. 41A-41D con más detalle. La FIG. 48 ilustra una conversión de los coeficientes armónicos esféricos recuperados 547 en los datos de audio multicanal 549A que son compatibles con una geometría de decodificador-altavoz local. Para algunas geometrías de altavoces locales (que, de nuevo, pueden referirse a una geometría de altavoz en el decodificador), algunas transformaciones que aseguran la invertibilidad pueden dar como resultado una calidad de audio-imagen menos que deseable. Es decir, la reproducción del sonido no siempre puede dar como resultado una correcta localización de los sonidos en comparación con el audio que se está capturando. Para corregir esta calidad de imagen menos que deseable, las técnicas pueden aumentarse adicionalmente para introducir un concepto que puede denominarse como "altavoces virtuales".

[0805] En lugar de requerir que uno o más altavoces sean reposicionados o posicionados en regiones particulares o definidas de espacio que tengan ciertas tolerancias angulares especificadas por una norma, tal como la citada anteriormente ITU-R BS.775-1, la trama anterior puede modificarse para incluir alguna forma de panoramización, tal como la panoramización de amplitud basada en vectores (VBAP), la panoramización de amplitud basada en la distancia, u otras formas de panoramización. Centrándose en la VBAP con fines de ilustración, VBAP puede introducir eficazmente lo que se puede caracterizar como "altavoces virtuales". VBAP puede modificar una alimentación a uno o más altavoces para que estos uno o más altavoces emitan de manera eficaz un sonido que parece originarse desde un altavoz virtual en una o más de una ubicación y ángulo diferente de al menos uno de la ubicación y/o el ángulo del uno o más altavoces que soportan el altavoz virtual.

[0806] Para ilustrar, la siguiente ecuación para determinar las alimentaciones de altavoz en términos de los SHC puede ser la siguiente:

$$\begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ \dots \\ A_{(Order+1)(Order+1)}^{-(Order+1)(Order+1)}(\omega) \end{bmatrix} = -ik \begin{bmatrix} VBAP \\ MATRIX \\ MxN \end{bmatrix} \begin{bmatrix} D \\ Nx(Order+1)^2 \end{bmatrix} \begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ g_3(\omega) \\ \dots \\ g_{\mathrm{M}}(\omega) \end{bmatrix}.$$

[0807] En la ecuación anterior, la matriz de VBAP es de tamaño de M filas por N columnas, donde M representa el número de altavoces (y será igual a cinco en la ecuación anterior) y N representa el número de altavoces virtuales. La matriz VBAP se puede calcular en función de los vectores desde la posición definida del oyente a cada una de las posiciones de los altavoces y los vectores desde la posición definida del oyente a cada una de las posiciones de los altavoces virtuales. La matriz D en la ecuación anterior puede ser de un tamaño de N filas por (orden+1)^2 columnas, donde el orden puede referirse al orden de las funciones de SH. La matriz *D* puede representar lo siguiente

$$\text{matrix:} \begin{bmatrix} h_0^{(2)}(kr_1)Y_0^{0^*}(\theta_1,\varphi_1) & h_0^{(2)}(kr_2)Y_0^{0^*}(\theta_2,\varphi_2) & . & . & . \\ h_1^{(2)}(kr_1)Y_1^{1^*}(\theta_1,\varphi_1). & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \end{bmatrix}.$$

[0808] La matriz ***g*** (o vector, dado que sólo hay una sola columna) puede representar la ganancia para las alimentaciones de altavoz para los altavoces dispuestos en la geometría decodificador-local. En la ecuación, la matriz ***g*** es de tamaño M. La matriz ***A*** (o vector, dado que sólo hay una sola columna) puede representar los SHC 520, y es de tamaño (Orden+1)(orden+1), que también puede representarse como (Orden+1)^2.

[0809] De hecho, la matriz VBAP es una matriz MxN que proporciona lo que se puede denominar un "ajuste de ganancia" que influye en la ubicación de los altavoces y la posición de los altavoces virtuales. La introducción de la panoramización de esta manera puede dar como resultado una mejor reproducción del audio multicanal que da como resultado una imagen de mejor calidad cuando se reproduce por la geometría del altavoz local. Además, al incorporar VBAP en esta ecuación, las técnicas pueden superar geometrías de altavoz deficientes que no se alinean con las especificadas en varios estándares.

[0810] En la práctica, la ecuación puede invertirse y emplearse para transformar los SHC de nuevo en las alimentaciones multicanal para una geometría o configuración particular de altavoces, que de nuevo se puede denominar geometría de decodificador-local en esta divulgación. Es decir, la ecuación puede invertirse para resolver la matriz **g**. La ecuación invertida puede ser como se indica a continuación:

$$\begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ g_3(\omega) \\ \dots \\ g_{\mathrm{M}}(\omega) \end{bmatrix} = -ik \begin{bmatrix} VBAP \\ MATRIX^{-1} \\ MxN \end{bmatrix} \begin{bmatrix} D^{-1} \\ Nx(Order+1)^2 \end{bmatrix} \begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ \dots \\ A_{(Order+1)(Order+1)}^{-(Order+1)(Order+1)}(\omega) \end{bmatrix}.$$

[0811] La matriz ***g*** puede representar la ganancia del altavoz para, en este ejemplo, cada uno de los cinco altavoces en una configuración de 5.1 altavoces. Las ubicaciones de los altavoces virtuales utilizadas en esta configuración pueden corresponder a las ubicaciones definidas en una especificación o estándar de formato multicanal 5.1. La ubicación de los altavoces que puede soportar cada uno de estos altavoces virtuales se puede determinar usando cualquier número de técnicas de localización de audio conocidas, muchas de las cuales implican reproducir un tono que tiene una frecuencia particular para determinar una ubicación de cada altavoz con respecto a una unidad de cabecera (tal como un receptor de audio/vídeo (receptor A/V), televisión, sistema de juegos, sistema de disco de vídeo digital u otros tipos de sistemas de cabecera). Alternativamente, un usuario de la unidad de cabecera puede especificar manualmente la ubicación de cada uno de los altavoces. En cualquier caso, dadas estas ubicaciones conocidas y ángulos posibles, la unidad de cabecera puede resolver las ganancias, asumiendo una configuración ideal de altavoces virtuales por medio de VBAP.

[0812] A este respecto, un dispositivo o aparato puede realizar una panoramización de amplitud basada en vectores u otra forma de panoramización en la pluralidad de canales virtuales para producir una pluralidad de canales que impulsan a los altavoces en una geometría decodificador-local para emitir sonidos que parecen originarse a partir de altavoces virtuales configurados en una geometría local diferente. Por lo tanto, las técnicas pueden permitir al dispositivo de decodificación de audio 40 realizar una transformación en la pluralidad de coeficientes armónicos esféricos, tales como los coeficientes armónicos esféricos recuperados 47, para producir una pluralidad de canales. Cada una de la pluralidad de canales puede estar asociada con una región de espacio correspondiente diferente. Además, cada una de la pluralidad de canales puede comprender una pluralidad de canales virtuales, donde la pluralidad de canales virtuales puede estar asociada con la correspondiente región de espacio diferente. Por lo tanto, un dispositivo puede realizar una panoramización de amplitud basada en vectores en los canales virtuales para producir la pluralidad de canales de los datos de audio multicanal 49.

[0813] Las FIGS. 49A-49E(ii) son diagramas que ilustran sistemas de codificación de audio respectivos 560A-560C, 567D, 569D, 571E y 573E que pueden implementar diversos aspectos de las técnicas descritas en esta divulgación. Como se ilustra en el ejemplo de la FIG. 49A, el sistema de codificación de audio 560A puede incluir un dispositivo de codificación de audio 562 y un dispositivo de decodificación de audio 564. El dispositivo de codificación de audio 562 puede ser similar a cualquiera de los dispositivos de codificación de audio 20 y 510A-510D mostrados en el ejemplo de las FIGS. 4 y 40A-40D, respectivamente. El dispositivo de decodificación de audio 564 puede ser similar al dispositivo de decodificación de audio 24 y 40 mostrado en el ejemplo de las FIGS. 5 y 41.

[0814] Como se ha descrito anteriormente, el ambisónico de orden superior (HOA) es una manera mediante la cual describir toda la información direccional de un campo sonoro basado en una transformada espacial de Fourier. En algunos ejemplos, cuanto mayor es el orden de ambisónicos, N, mayor es la resolución espacial y mayor es el número de coeficientes armónicos esféricos (SH) $(N+1)^2$. Por lo tanto, cuanto más alto sea el orden de los ambisónicos N, en algunos ejemplos, se obtienen requisitos de ancho de banda mayores para transmitir y almacenar los coeficientes. Debido a que los requisitos de ancho de banda de HOA son bastante altos en comparación, por ejemplo, con datos de audio de sonido envolvente 5.1 o 7.1, se puede desear una reducción de ancho de banda para muchas aplicaciones.

[0815] De acuerdo con las técnicas descritas en esta divulgación, el sistema de codificación de audio 560A puede realizar un procedimiento basado en separar los elementos distintos (de primer plano) de los no distintos (fondo o ambiente) en una escena de sonido espacial. Esta separación puede permitir que el sistema de codificación de audio 560A procese elementos de primer plano y de fondo independientemente entre sí. En este ejemplo, el sistema de codificación de audio 560A explora la propiedad de que los elementos de primer plano pueden llamar más la atención (por el oyente) y pueden ser más fáciles de localizar (de nuevo, por el oyente) en comparación con los elementos de fondo. Como resultado, el sistema de codificación de audio 560A puede almacenar o transmitir contenido HOA de manera más eficiente.

[0816] En algunos ejemplos, el sistema de codificación de audio 560A puede conseguir esta separación empleando el proceso de descomposición en valores singulares (SVD). El proceso SVD puede separar una trama de coeficientes HOA en 3 matrices (U, S, V). La matriz U contiene los vectores singulares izquierdos y la matriz V contiene los vectores singulares derechos. La matriz Diagonal S contiene los valores singulares no negativos clasificados en su diagonal. Una reconstrucción de los coeficientes HOA generalmente buena (o, en algunos casos, perfecta suponiendo una precisión ilimitada en la representación de los coeficientes HOA) será dada por U*S*V'. Sólo reconstruyendo el subespacio con los D valores singulares más grandes: U(:,1:D)*S(1:D,:)*V', el sistema de codificación de audio 560A puede extraer la información espacial más sobresaliente de esta trama HOA, es decir, elementos de sonido de primer plano (y tal vez algunas reflexiones de la sala tempranas fuertes). El resto U(:,D+1:end)*S(D+1:end,:)*V' puede reconstruir elementos de fondo y reverberación a partir del contenido.

[0817] El sistema de codificación de audio 560A puede determinar el valor D, que separa los dos subespacios, analizando la pendiente de la curva creada por los valores diagonales descendentes de S: los valores singulares grandes representan sonidos de primer plano, los valores singulares bajos representan valores de fondo. El sistema de codificación de audio 560A puede utilizar una primera y una segunda derivada de la curva de valor singular. El sistema de codificación de audio 560A puede también limitar el número D entre uno y cinco. Alternativamente, el sistema de codificación de audio 560A puede predefinir el número D, tal como a un valor de cuatro. En cualquier caso, una vez estimado el número D, el sistema de codificación de audio 560A extrae el subespacio de primer plano y fondo de las matrices U y S.

[0818] El sistema de codificación de audio 560A puede entonces reconstruir los coeficientes HOA de la escena de fondo mediante U(:,D+1:end)*S(D+1:end,:)*V', dando como resultado $(N+1)^2$ canales de coeficientes HOA. Como se sabe que los elementos de fondo, en algunos ejemplos, no tan relevantes y no son tan localizables con respecto a los elementos de primer plano, el sistema de codificación de audio 560A puede truncar el orden de los canales HOA. Además, el sistema de codificación de audio 560A puede comprimir estos canales con códecs de audio con pérdidas o sin pérdida, tales como AAC, u opcionalmente con un códec de audio más agresivo en comparación con el utilizado para comprimir los elementos relevantes de primer plano. En algunos casos, para ahorrar ancho de banda, el sistema de codificación de audio 560A puede transmitir los elementos de primer plano de manera diferente. Es

decir, el sistema de codificación de audio puede transmitir los vectores singulares izquierdos U(:,1:D) después de comprimirse con códecs de audio con pérdidas o sin pérdida (tales como AAC) y transmitir estos valores singulares izquierdos comprimidos junto con la matriz de reconstrucción R = S(1:D,:)*V'. R puede representar una matriz D x $(N+1)^2$, que puede diferir entre tramas.

[0819] En el lado receptor del sistema de codificación de audio 560, el sistema de codificación de audio puede multiplicar estas dos matrices para reconstruir una trama de $(N+1)^2$ canales HOA. Una vez que se suman los canales HOA de fondo y de primer plano, el sistema de codificación de audio 560A puede reproducir cualquier configuración de altavoz utilizando cualquier reproductor ambisónico apropiado. Dado que las técnicas proporcionan la separación de elementos de primer plano (sonido directo o distinto) de los elementos de fondo, una persona con discapacidad auditiva podrá controlar la mezcla de elementos de primer plano con elementos de fondo para aumentar la inteligibilidad. Además, pueden aplicarse otros efectos de audio, por ejemplo, un compresor dinámico en los elementos de primer plano.

[0820] La FIG. 49B es un diagrama de bloques que ilustra el sistema de codificación de audio 560B con más detalle. Como se ilustra en el ejemplo de la FIG. 49B, el sistema de codificación de audio 560B puede incluir un dispositivo de codificación de audio 566 y un dispositivo de decodificación de audio 568. El dispositivo de codificación de audio 566 puede ser similar a los dispositivos de codificación de audio 24 y 510E mostrados en el ejemplo de las FIGS. 4 y 40E. El dispositivo de decodificación de audio 568 puede ser similar al dispositivo de decodificación de audio 24 y 540B mostrado en el ejemplo de las FIGS. 5 y 41B.

[0821] De acuerdo con las técnicas descritas en esta divulgación, cuando se utiliza la descomposición SVD basada en tramas (o procedimientos relacionados tales como KLT y PCA) en señales HoA, con el fin de reducir el ancho de banda, el dispositivo de codificación de audio 66 puede cuantizar los primeros pocos vectores de la U (multiplicado por los valores singulares correspondientes de la matriz S), así como los vectores correspondientes del vector V^T. Esto comprenderá los componentes "de primer plano" del campo sonoro. Las técnicas pueden permitir que el dispositivo de codificación de audio 566 codifique el vector $U_{DIST} * S_{DIST}$ utilizando un motor de codificación de audio de "caja negra". El V vector puede ser escalar o un vector cuantificado. Además, algunos o todos los vectores restantes en la matriz U pueden multiplicarse con los valores singulares correspondientes de la matriz S y de la matriz V y también codificarse utilizando un motor de codificación de audio de "caja negra". Estos comprenderán los componentes "de fondo" del campo sonoro.

[0822] Puesto que los componentes auditivos más ruidosos se descomponen en los "componentes de primer plano", el dispositivo de codificación de audio 566 puede reducir el orden ambisónico de los componentes "de fondo" antes de utilizar un motor de codificación de audio de "caja negra" porque (se asume) que el fondo no contiene contenido localizable importante. Dependiendo del orden ambisónico de los componentes de primer plano, la unidad de codificación de audio 566 puede transmitir el o los vectores correspondientes, que pueden ser bastante grandes. Por ejemplo, una cuantización escalar simple de 16 bits de los vectores V dará como resultado una sobrecarga de aproximadamente 20 kbps para el 4º orden (25 coeficientes) y 40 kbps para el 6º orden (49 coeficientes) por componente de primer plano. Las técnicas descritas en esta divulgación pueden proporcionar un procedimiento para reducir esta sobrecarga del vector V.

[0823] Para ilustración, se asume que el orden ambisónico de los elementos de primer plano es N_{DIST} y el orden ambisónico de los elementos de fondo N_{BG}, como se ha descrito anteriormente. Dado que el dispositivo de codificación de audio 566 puede reducir el orden ambisónico de los elementos de fondo como se ha descrito anteriormente, N_{BG} puede ser menor que N_{DIST}. La longitud del vector V de primer plano que necesita transmitirse para reconstruir los elementos de primer plano en el lado del receptor tiene la longitud de $(N_{DIST}+1)^2$ por elemento de primer plano, mientras que los primeros coeficientes $((N_{DIST}+1)^2)$ - $((N_{BG}+1)^2)$ se pueden utilizar para reconstruir los componentes de primer plano o distintos hasta el orden N_{BG}. Usando las técnicas descritas en esta divulgación, el dispositivo de codificación de audio 566 puede reconstruir el primer plano hasta el orden N_{BG} y fusionar los canales $(N_{BG}+1)^2$ resultantes con los canales de fondo, dando como resultado un campo sonoro completo hasta el orden NBG. El dispositivo de codificación de audio 566 puede entonces reducir el vector V a aquellos coeficientes con el índice mayor de $(N_{BG}+1)^2$ para la transmisión, (donde estos vectores pueden denominarse "$V^T_{PEQUEÑO}$"). En el lado del receptor, la unidad de decodificación de audio 568 puede reconstruir los canales de audio de primer plano para el orden ambisónico mayor que N_{BG} multiplicando los elementos de primer plano por los vectores $V^T_{PEQUEÑO}$.

[0824] La FIG. 49c es un diagrama de bloques que ilustra el sistema de codificación de audio 560C con más detalle. Como se ilustra en el ejemplo de la FIG. 49C, el sistema de codificación de audio 560B puede incluir un dispositivo de codificación de audio 567 y un dispositivo de decodificación de audio 569. El dispositivo de codificación de audio 567 puede ser similar a los dispositivos de codificación de audio 20 y 510F mostrados en el ejemplo de las FIGS. 4 y 40F. El dispositivo de decodificación de audio 569 puede ser similar a los dispositivos de decodificación de audio 24 y 540B mostrados en el ejemplo de las FIGS. 5 y 41B.

[0825] De acuerdo con las técnicas descritas en esta divulgación, cuando se utiliza la descomposición SVD basada en tramas (o procedimientos relacionados tales como KLT y PCA) en señales HoA, con el fin de reducir el ancho de banda, el dispositivo de codificación de audio 567 puede cuantizar los primeros pocos vectores de la U

(multiplicado por los valores singulares correspondientes de la matriz S), así como los vectores correspondientes del vector V^T. Esto comprenderá los componentes "de primer plano" del campo sonoro. Las técnicas pueden permitir que el dispositivo de codificación de audio 567 codifique el vector $U_{DIST} * S_{DIST}$ utilizando un motor de codificación de audio de "caja negra". El V vector puede ser escalar o un vector cuantificado. Además, algunos o todos los vectores restantes en la matriz U pueden multiplicarse con los valores singulares correspondientes de la matriz S y de la matriz V y también codificarse utilizando un motor de codificación de audio de "caja negra". Estos comprenderán los componentes "de fondo" del campo sonoro.

[0826] Puesto que los componentes auditivos más ruidosos se descomponen en los "componentes de primer plano", el dispositivo de codificación de audio 567 puede reducir el orden ambisónico de los componentes "de fondo" antes de utilizar un motor de codificación de audio de "caja negra" porque (se asume) que el fondo no contiene contenido localizable importante. El dispositivo de codificación de audio 567 puede reducir el orden de tal manera que conserve la energía total del campo sonoro de acuerdo con las técnicas descritas en el presente documento. Dependiendo del orden ambisónico de los componentes de primer plano, la unidad de codificación de audio 567 puede transmitir el o los vectores correspondientes, que pueden ser bastante grandes. Por ejemplo, una cuantización escalar simple de 16 bits de los vectores V dará como resultado una sobrecarga de aproximadamente 20 kbps para el 4º orden (25 coeficientes) y 40 kbps para el 6º orden (49 coeficientes) por componente de primer plano. Las técnicas descritas en esta divulgación pueden proporcionar un procedimiento para reducir esta sobrecarga del uno o más vectores V.

[0827] Para ilustración, se asume que el orden ambisónico de los elementos de primer plano y de los elementos de fondo es N. El dispositivo de codificación de audio 567 puede reducir el orden ambisónico de los elementos de fondo del uno o más vectores V de N a ñ de tal forma que ñ <N. El dispositivo de codificación de audio 67 aplica además una compensación para aumentar los valores de los elementos de fondo del uno o más vectores V para conservar la energía total del campo sonoro descrito por los SHC. Se han descrito anteriormente técnicas de aplicación de la compensación con respecto a la FIG. 40F. En el lado del receptor, la unidad de decodificación de audio 569 puede reconstruir los canales de audio de fondo para el orden ambisónico.

[0828] Las FIGS. 49D (i) y 49D (ii) ilustran un dispositivo de codificación de audio 567D y un dispositivo de decodificación de audio 569D respectivamente. El dispositivo de codificación de audio 567D y el dispositivo de decodificación de audio 569D pueden estar configurados para llevar a cabo una o más determinaciones de distinción basadas en la direccionalidad, de acuerdo con aspectos de esta divulgación. El ambisónico de orden superior (HOA) es un procedimiento para describir toda la información direccional de un campo sonoro basado en la transformada espacial de Fourier. Cuanto mayor es el orden ambisónico N, mayor es la resolución espacial, cuanto mayor es el número de coeficientes armónicos esféricos (SH) (N+1)^2, mayor será el ancho de banda necesario para transmitir y almacenar los datos. Debido a que los requisitos de ancho de banda de HOA son bastante altos, para muchas aplicaciones se desea una reducción de ancho de banda.

[0829] Las descripciones anteriores han descrito cómo la SVD (descomposición en valores singulares) o procesos relacionados pueden ser utilizados para la compresión de audio espacial. Las técnicas descritas en el presente documento presentan un algoritmo mejorado para seleccionar los elementos relevantes también conocidos como los elementos de primer plano. Después de una descomposición basada en SVD de una trama de audio de HOA en sus matrices U, S y V, las técnicas basan la selección de los K elementos relevantes exclusivamente en los primeros K canales de la matriz U [U(:,1:K)*S(1:K,1:K)]. Esto da como resultado la selección de los elementos de audio con mayor energía. Sin embargo, no está garantizado que estos elementos sean también direccionales. Por lo tanto, las técnicas se dirigen a encontrar los elementos sonoros que tienen alta energía y también son direccionales. Esto se logra potencialmente ponderando la matriz V con la matriz S. A continuación, para cada fila de esta matriz resultante, los elementos indexados superiores (que están asociados con los coeficientes HOA de orden más alto) se elevan al cuadrado y se suman, dando como resultado un valor por fila [sumVS en el pseudocódigo descrito con respecto a la FIG. 40H]. De acuerdo con el flujo de trabajo representado en el pseudocódigo, se consideran los coeficientes ambisónicos de orden superior que comienzan en el 5º índice. Estos valores se clasifican según su tamaño y el índice de clasificación se utiliza para reordenar la matriz U, S y V original en consecuencia. El algoritmo de compresión basado en SVD descrito anteriormente en esta divulgación puede entonces aplicarse sin modificación adicional.

[0830] Las FIGS. 49E(i) y 49E(ii) son diagramas de bloques que ilustran un dispositivo de codificación de audio 571E y un dispositivo de decodificación de audio 573E respectivamente. El dispositivo de codificación de audio 571E y el dispositivo de decodificación de audio 573E pueden realizar diversos aspectos de las técnicas descritas anteriormente con respecto a los ejemplos de las FIGS. 49-49D(ii), excepto que el dispositivo de codificación de audio 571E puede realizar la descomposición en valores singulares con respecto a una matriz de densidad espectral de potencia (PDS) de los coeficientes HOA para generar una matriz S^2 y una matriz V. La matriz S^2 puede representar una matriz S al cuadrado, con lo que la matriz S^2 puede experimentar una operación de raíz cuadrada para obtener la matriz S. El dispositivo de codificación de audio 571E, en algunos casos, puede realizar la cuantización con respecto a la matriz V para obtener una matriz V cuantizada (que puede estar designada como matriz V').

[0831] El dispositivo de codificación de audio 571E puede obtener la matriz U multiplicando primero la matriz S por la matriz cuantizada V' para generar una matriz SV'. El dispositivo de codificación de audio 571E puede obtener a continuación la pseudoinvertida (pinv) de la matriz SV' y luego multiplicar los coeficientes HOA 11 por la pseudoinvertida de la matriz SV' para obtener la matriz U. Mediante la realización de SVD con respecto a la densidad espectral de potencia de los coeficientes HOA en lugar de los coeficientes en sí, el dispositivo de codificación de audio 571E puede reducir potencialmente la complejidad computacional de realizar la SVD en términos de uno o más ciclos de procesador y espacio de almacenamiento, mientras que se consigue la misma eficacia en la codificación de la fuente de audio como si la SVD se aplicara directamente a los coeficientes HOA.

[0832] El dispositivo de decodificación de audio 573E puede ser similar a los dispositivos de decodificación de audio descritos anteriormente, excepto que el dispositivo de decodificación de audio 573 puede reconstruir los coeficientes HOA de las descomposiciones de los coeficientes HOA obtenidos mediante la aplicación de la SVD a la densidad espectral de potencia de los coeficientes HOA en lugar de los coeficientes HOA directamente.

[0833] Las FIGS. 50A y 50B son diagramas de bloques que ilustran cada uno de dos enfoques diferentes para reducir potencialmente el orden del contenido de fondo de acuerdo con las técnicas descritas en esta divulgación. Como se ilustra en el ejemplo de la FIG. 50, el primer enfoque puede emplear la reducción de orden con respecto a los vectores $U_{BG} * S_{BG} * V^T$ para reducir el orden de N a ñ, donde ñ es menor que (<) N. Es decir, la unidad de reducción de órdenes 528A mostrada en los ejemplos de la FIG. 40B-40J puede realizar una reducción de orden para truncar o reducir de otro modo el orden N de los vectores $U_{BG} * S_{BG} * V^T$ a ñ, donde ñ es menor que (<) N.

[0834] Como un enfoque alternativo, la unidad de reducción de órdenes 528A, como se muestra en el ejemplo de la FIG. 50B, puede realizar este truncamiento con respecto a V^T eliminando las filas que serán $(ñ+1)^2$, lo cual no se ilustra en el ejemplo de la FIG. 40B para facilitar la ilustración. En otras palabras, la unidad de reducción de órdenes 528A puede eliminar uno o más órdenes de la matriz V^T para generar eficazmente una matriz V_{BG}. El tamaño de esta matriz V_{BG} es $(ñ+1)^2 \times (N+1)^2 - D$, donde esta matriz V_{BG} se usa entonces en lugar de la matriz V^T al generar los vectores $U_{BG} * S_{BG} * V^T$, realizando eficazmente el truncamiento para generar los vectores $U_{BG} * S_{BG} * V^T$ del tamaño $M \times (ñ+1)^2$.

[0835] La FIG. 51 es un diagrama de bloques que ilustra ejemplos de una trayectoria de compresión de componente distinta de un dispositivo de codificación de audio 700A que puede implementar diversos aspectos de las técnicas descritas en esta divulgación para comprimir coeficientes armónicos esféricos 701. En el ejemplo de la FIG. 51, la trayectoria de compresión de componente distinta puede referirse a una trayectoria de procesamiento del dispositivo de codificación de audio 700A que comprime los distintos componentes del campo sonoro representados por SHC 701. Otra trayectoria, que puede denominarse trayectoria de compresión de componente de fondo, puede representar una trayectoria de procesamiento del dispositivo de codificación de audio 700A que comprime los componentes de fondo de los SHC 701.

[0836] Aunque no se muestra para facilitar la ilustración, la trayectoria de compresión del componente de fondo puede operar con respecto a los SHC 701 directamente en lugar de las descomposiciones de los SHC 701. Es similar a lo descrito anteriormente con respecto a las FIGS. 49-49C, excepto que en lugar de recomponer los componentes de fondo de las matrices U_{BG}, S_{BG} y V_{BG} y luego realizar alguna forma de codificación psicoacústica (por ejemplo, usando un codificador AAC) de estos componentes de fondo recompuestos, la trayectoria de procesamiento de componente de fondo puede funcionar con respecto al SHC 701 directamente (como se ha descrito anteriormente con respecto al dispositivo de codificación de audio 20 mostrado en el ejemplo de la FIG. 4), comprimiendo estos componentes de fondo utilizando el codificador psicoacústico. Mediante la realización de la codificación psicoacústica con respecto a los SHC 701 directamente, las discontinuidades pueden reducirse al mismo tiempo que se reduce la complejidad computacional (en términos de operaciones requeridas para comprimir los componentes de fondo) en comparación con la realización de la codificación psicoacústica con respecto a los componentes de fondo recompuestos. Aunque se hace referencia en términos de un fondo distinto y de fondo, el término "prominente" puede usarse en lugar de "distinto" y el término "ambiente" puede usarse en lugar de "fondo" en esta divulgación.

[0837] En cualquier caso, los coeficientes armónicos esféricos 701 ("SHC 701") pueden comprender una matriz de coeficientes que tienen un tamaño de $M \times (N+1)^2$, donde M representa el número de muestras (y es, en algunos ejemplos, 1024) en una trama de audio y N representa el orden más alto de la función de base a la que corresponden los coeficientes. Como se ha indicado anteriormente, N se fija comúnmente a cuatro (4) para un total de 1024 x 25 coeficientes. Cada uno de los SHC 701 correspondientes a una combinación de suborden, de orden particular, puede denominarse canal. Por ejemplo, todos los M coeficientes de muestra correspondientes a una función básica de primer orden, de sub-orden cero orden pueden representar un canal, mientras que los coeficientes correspondientes a la función de orden cero, sub-orden cero pueden representar otro canal, etc. Los SHC 701 también se pueden denominar en esta divulgación como contenido ambisónico de orden superior (HOA) 701 o como una señal de SH 701.

[0838] Como se ilustra en el ejemplo de la FIG. 51, el dispositivo de codificación de audio 700A incluye una unidad de análisis 702, una unidad de síntesis basada en vectores 704, una unidad de reducción de vectores 706, una

unidad de codificación psicoacústica 708, una unidad de reducción de coeficientes 710 y una unidad de compresión 712 ("unidad compr. 712"). La unidad de análisis 702 puede representar una unidad configurada para realizar un análisis con respecto a los SHC 701 para identificar componentes distintos del campo sonoro (D) 703 y un número total de componentes de fondo (BGTOT) 705. En comparación con los dispositivos de codificación de audio descritos anteriormente, el dispositivo de codificación de audio 700A no realiza esta determinación con respecto a las descomposiciones de los SHC 701, sino directamente con respecto a los SHC 701.

[0839] La unidad de síntesis basada en vectores 704 representa una unidad configurada para realizar alguna forma de síntesis basada en vectores con respecto a los SHC 701, tal como SVD, KLT, PCA o cualquier otra síntesis basada en vectores, para generar, en los casos de SVD, una matriz [US] 707 que tiene un tamaño de M x $(N+1)^2$ y una matriz [V] 709 que tiene un tamaño de $(N+1)^2$ x $(N+1)^2$. La matriz [US] 707 puede representar una matriz resultante de una multiplicación matricial de la matriz [U] y de la matriz [S] generadas mediante la aplicación de SVD a los SHC 701.

[0840] La unidad de reducción de vectores 706 puede representar una unidad configurada para reducir el número de vectores de la matriz [US] 707 y la matriz [V] 709 de tal forma que cada uno de los vectores restantes de la matriz [US] 707 y de la matriz [V] 709 identifiquen un componente distinto o prominente del campo sonoro. La unidad de reducción de vectores 706 puede realizar esta reducción en base al número de componentes D distintos 703. El número de componentes distintos D 703 puede, en efecto, representar una matriz de números, donde cada número identifica diferentes vectores distintos de las matrices 707 y 709. La unidad de reducción de vectores 706 puede producir una matriz [US] reducida 711 de tamaño M x D y una matriz [V] reducida 713 de tamaño $(N+1)^2$ x D.

[0841] Aunque no se muestra para facilitar la ilustración, la interpolación de la matriz [V] 709 puede ocurrir antes de la reducción de la matriz [V] 709 de manera similar a la descrita con más detalle anteriormente. Además, aunque no se muestra con fines fáciles de ilustración, reordenación de la matriz [US] reducida 711 y/o la matriz [V] reducida 712 de la manera descrita con más detalle anteriormente. Por consiguiente, las técnicas no deben estar limitadas en estos y otros aspectos (tales como proyección de errores o cualquier otro aspecto de las técnicas anteriores descritas anteriormente, pero no mostradas en el ejemplo de la FIG. 51).

[0842] La unidad de codificación psicoacústica 708 representa una unidad configurada para realizar la codificación psicoacústica con respecto a la matriz [US] 711 para generar un flujo de bits 715. La unidad de reducción de coeficientes 710 puede representar una unidad configurada para reducir el número de canales de la matriz [V] reducida 713. En otras palabras, la unidad de reducción de coeficientes 710 puede representar una unidad configurada para eliminar aquellos coeficientes de los vectores V distintos (que forman la matriz [V] reducida 713) que tienen poca o ninguna información direccional. Como se ha descrito anteriormente, en algunos ejemplos, los coeficientes de los distintos vectores V que corresponden a funciones de base de primer orden y de orden cero (representados como N_{BG} anteriores) proporcionan poca información direccional y por lo tanto pueden ser eliminados de los distintos vectores V (a través de lo que se conoce como "reducción de orden" anteriormente). En este ejemplo, se puede proporcionar una mayor flexibilidad para identificar no solo estos coeficientes que corresponden a N_{BG} sino para identificar canales HOA adicionales (que pueden indicarse con la variable TotalOfAddAmbHOAChan) a partir del conjunto de $[(N_{BG}+1)^2+1, (N+1)^2]$. La unidad de análisis 702 puede analizar los SHC 701 para determinar BG_{TOT}, que puede identificar no solo el $(N_{BG}+1)^2$, sino el TotalOfAddAmbHOAChan. La unidad de reducción de coeficientes 710 puede entonces eliminar aquellos coeficientes correspondientes a $(N_{BG}+1)^2$ y el TotalOfAddAmbHOAChan de la matriz [V] reducida 713 para generar una matriz [V] pequeña 717 de tamaño $((N+1)^2 - (BG_{TOT})$ x D.

[0843] La unidad de compresión 712 puede entonces realizar la codificación escalar escalar y/o la codificación Huffman indicada anteriormente para comprimir la matriz [V] pequeña 717, que emite la matriz [V] pequeña comprimida 717 como información de canal lateral 719 ("información de canal lateral 719"). La unidad de compresión 712 puede emitir la información de canal lateral 719 de una manera similar a la que se muestra en el ejemplo de las FIGS. 10-10O(ii). En algunos ejemplos, una unidad de generación de flujo de bits similar a las descritos anteriormente puede incorporar la información de canal lateral 719 en el flujo de bits 715. Además, aunque se denomina flujo de bits 715, el dispositivo de codificación de audio 700A puede, como se ha indicado anteriormente, incluir una trayectoria de procesamiento de componentes de fondo que da como resultado otro flujo de bits, donde una unidad de generación de flujo de bits similar a las descritas anteriormente puede generar un flujo de bits similar al flujo de bits 17 descrito anteriormente que incluye el flujo de bits 715 y el flujo de bits transmitido por la trayectoria de procesamiento del componentes de fondo.

[0844] De acuerdo con las técnicas descritas en esta divulgación, la unidad de análisis 702 puede configurarse para determinar un primer conjunto distinto de cero de coeficientes de un vector, es decir, los vectores de la matriz [V] reducida 713 en este ejemplo, que se utilizarán para representar el componente distinto del campo sonoro. En algunos ejemplos, la unidad de análisis 702 puede determinar que todos los coeficientes de cada vector que forma la matriz [V] reducida 713 deben incluirse en la información de canal lateral 719. La unidad de análisis 702 puede por lo tanto ajustar BG_{TOT} igual a cero.

[0845] Por lo tanto, el dispositivo de codificación de audio 700A puede actuar eficazmente de una manera recíproca a la descrita anteriormente con respecto a la Tabla indicada como "Vectores decodificados". Además, el

dispositivo de codificación de audio 700A puede especificar un elemento de sintaxis en un encabezamiento de una unidad de acceso (que puede incluir uno o más tramas) de la que la pluralidad de modos de configuración se seleccionó. Aunque se describe que se especifica en una base de unidad de acceso, la unidad de análisis 702 puede especificar este elemento sintáctico sobre una base por trama o cualquier otra base periódica o no periódica (como una vez para el flujo de bits completo). En cualquier caso, este elemento de sintaxis puede comprender dos bits que indican cuál de los cuatro modos de configuración fueron seleccionados para especificar el conjunto de coeficientes diferente de cero de la matriz [V] reducida 713 para representar los aspectos direccionales de este componente distinto. El elemento de sintaxis se puede indicar como "codedVVecLength". De esta manera, el dispositivo de codificación de audio 700A puede señalar o especificar de otro modo en el flujo de bits cuál de los cuatro modos de configuración se usó para especificar la matriz [V] pequeña 717 en el flujo de bits. Aunque se describen con respecto a cuatro modos de configuración, las técnicas no deben limitarse a cuatro modos de configuración, sino a cualquier número de modos de configuración, incluido un único modo de configuración o una pluralidad de modos de configuración.

[0846] Diversos aspectos de las técnicas pueden además permitir que el dispositivo de codificación de audio 700A esté configurado para funcionar de acuerdo con las siguientes cláusulas.

Cláusula 133149-1F. Un dispositivo que comprende: uno o más procesadores configurados para seleccionar uno de una pluralidad de modos de configuración mediante los cuales se especifica un conjunto distinto de cero de coeficientes de un vector, habiéndose descompuesto el vector de una pluralidad de coeficientes armónicos esféricos que describen un campo sonoro y representando un componente distinto del campo sonoro, y especificar el conjunto distinto de cero de los coeficientes del vector basado en el seleccionado de la pluralidad de modos de configuración.

Cláusula 133149-2F. El dispositivo de la cláusula 133149-1F, en el que el modo de la pluralidad de modos de configuración indica que el conjunto distinto de cero de los coeficientes incluye todos los coeficientes.

Cláusula 133149-3F. El dispositivo de la cláusula 133149-1F, en el que el modo de la pluralidad de modos de configuración indica que el conjunto de coeficientes distinto de cero incluye los de los coeficientes correspondientes a un orden mayor que el orden de una función base con la cual uno o más de la pluralidad de coeficientes armónicos esféricos se corresponden.

Cláusula 133149-4F. El dispositivo de la cláusula 133149-1F, en el que el modo de la pluralidad de modos de configuración indica que el conjunto distinto de cero de los coeficientes incluye los de los coeficientes correspondientes a un orden mayor que el orden de una función base con la cual uno o más de la pluralidad de coeficientes armónicos esféricos se corresponden y excluyen al menos uno de los coeficientes correspondientes a un orden mayor que el orden de la función de base con la cual se corresponden uno o más de la pluralidad de coeficientes armónicos esféricos.

Cláusula 133149-5F. El dispositivo de la cláusula 133149-1F, en el que el modo de la pluralidad de modos de configuración indica que el conjunto de coeficientes distinto de cero incluye todos los coeficientes excepto por lo menos uno de los coeficientes.

Cláusula 133149-6F. El dispositivo de la cláusula 133149-1F, en el que el uno o más procesadores están configurados además para especificar el modo seleccionado de la pluralidad de modos de configuración en un flujo de bits.

Cláusula 133149-1G. Un dispositivo que comprende: uno o más procesadores configurados para determinar uno de una pluralidad de modos de configuración mediante los cuales se extrae un conjunto distinto de cero de coeficientes de un vector de acuerdo con una de una pluralidad de modos de configuración, habiéndose descompuesto el vector de una pluralidad de coeficientes armónicos esféricos que describen un campo sonoro y representando un componente distinto del campo sonoro, y extraer el conjunto distinto de cero de los coeficientes del vector basado en el obtenido de la pluralidad de modos de configuración.

Cláusula 133149-2G. El dispositivo de la cláusula 133149-1G, en el que el modo de la pluralidad de modos de configuración indica que el conjunto distinto de cero de los coeficientes incluye todos los coeficientes.

Cláusula 133149-3G. El dispositivo de la cláusula 133149-1G, en el que el modo de la pluralidad de modos de configuración indica que el conjunto de coeficientes distinto de cero incluye los de los coeficientes correspondientes a un orden mayor que el orden de una función base con la cual uno o más de la pluralidad de coeficientes armónicos esféricos se corresponden.

Cláusula 133149-4G. El dispositivo de la cláusula 133149-1G, en el que el modo de la pluralidad de modos de configuración indica que el conjunto distinto de cero de los coeficientes incluye los de los coeficientes correspondientes a un orden mayor que el orden de una función base con la cual uno o más de la pluralidad de coeficientes armónicos esféricos se corresponden y excluyen al menos uno de los coeficientes correspondientes

a un orden mayor que el orden de la función de base con la cual se corresponden uno o más de la pluralidad de coeficientes armónicos esféricos.

Cláusula 133149-5G. El dispositivo de la cláusula 133149-1G, en el que el modo de la pluralidad de modos de configuración indica que el conjunto de coeficientes distinto de cero incluye todos los coeficientes excepto por lo menos uno de los coeficientes.

Cláusula 133149-6G. El dispositivo de la cláusula 133149-1G, en el que el uno o más procesadores están configurados adicionalmente para determinar, cuando se determina el modo de la pluralidad de modos de configuración, el modo de la pluralidad de modos de configuración basado en un valor señalado en un flujo de bits.

[0847] Las FIGS. 52 es un diagrama de bloques que ilustra otro ejemplo de un dispositivo de decodificación 750A de audio que puede implementar diversos aspectos de las técnicas descritas en esta divulgación para reconstruir o casi reconstruir los SHC 701. En el ejemplo de la FIG. 52, el dispositivo de decodificación de audio 750A es similar al dispositivo de decodificación de audio 540D mostrado en el ejemplo de la FIG. 41D, excepto que la unidad de extracción 542 recibe el flujo de bits 715' (que es similar al flujo de bits 715 descrito anteriormente con respecto al ejemplo de la FIG. 51, excepto que el flujo de bits 715' incluye también la versión codificada de audio de SHC_{BG} 752) y la información de canal lateral 719. Por esta razón, la unidad de extracción se denomina "unidad de extracción 542'".

[0848] Además, la unidad de extracción 542' difiere de la unidad de extracción 542 en que la unidad de extracción 542' incluye una forma modificada de la unidad de descompresión V 555 (que se muestra como "unidad de descompresión V 555" en el ejemplo de la FIG. 52). la unidad de descompresión V 555' recibe la información de canal lateral 719 y el elemento de sintaxis representado codedVVecLength 754. La unidad de extracción 542' analiza la codedVVecLength 754 del flujo de bits 715' (y, en un ejemplo, del encabezado de la unidad de acceso incluido dentro del flujo de bits 715'). La unidad de descompresión V 555' incluye una unidad de configuración de modo 756 ("unidad de config de modo 756") y una unidad de análisis sintáctico 758 configurable para funcionar de acuerdo con uno cualquiera de los modos de configuración descritos anteriormente 760.

[0849] La unidad de configuración de modo 756 recibe el elemento de sintaxis 754 y selecciona uno de los modos de configuración 760. La unidad de configuración de modo 756 configura entonces la unidad de análisis sintáctico 758 con el modo seleccionado de los modos de configuración 760. La unidad de análisis sintáctico 758 representa una unidad configurada para operar de acuerdo con uno cualquiera de los modos de configuración 760 para analizar una forma comprimida de los vectores [V] pequeños 717 a partir de la información de canal lateral 719. La unidad de análisis sintáctico 758 puede funcionar de acuerdo con la declaración de conmutación presentada en la siguiente Tabla.

Tabla - Vectores decodificados

Sintaxis	Nº de bits	Mnemónica

decodeVVec(i)
{
switch codedVVecLength {
caso 0: //Vector completo
VVecLength = NumOfHoaCoeffs;
para (m=0; m< VVecLength; ++m){ VecCoeff[m] = m+1; }
break;
caso 1: //se eliminan los órdenes inferiores
VVecLength = NumOfHoaCoeffs - MinNumOfCoeffsForAmbHOA;
para (m=0; m< VVecLength; ++m){
VecCoeff[m] = m + MinNumOfCoeffsForAmbHOA + 1;
}
break;
caso 2:
VVecLength = NumOfHoaCoeffs - MinNumOfCoeffsForAmbHOA - NumOfAddAmbHoaChan;
n = 0;
para (m=0;m<NumOfHoaCoeffs-MinNumOfCoeffsForAmbHOA; ++m){
c = m + MinNumOfCoeffsForAmbHOA + 1;
if (*ismember(c,* AmbCoeffIdx) == 0){
VecCoeff[n] = c;
n++;		

}
}
break;
caso 3:
VVecLength = NumOfHoaCoeffs - NumOfAddAmbHoaChan;
n = 0;
para (m=0; m<NumOfHoaCoeffs; ++m){
c = m + 1;
if (*ismember(c,* AmbCoeffIdx) == 0){
VecCoeff[n] = c;
n++;
}
}
}
if (NbitsQ[i] == 5) { /* cuantizador uniforme */
para (m=0; m< VVecLength; ++m){
VVec(k)[i][m] = **(VecValue** /128.0) – 1.0;
	8	**uimsbf**

}
}
else { /* decodificación Huffman */
para (m=0; m< VVecLength; ++m){
Idx = 5;
If (CbFlag[i] == 1){
idx = (min(3, max(I, ceil(sqrt(VecCoeff[m]) - 1)));
}
else if (PFlag[i] == 1) {idx = 4;}
cid = *huffDecode*(huffmannTable[NbitsQ].codebook[idx]; **huffVal);**
	dinámica	**huffDecode**

if (cid>0) { aVal = sgn = **(sgnVal** * 2) - 1;	**1**	**bslbf**

if (cid > 1) {
aVal = sgn * (2.0^(cid -1) **+ intAddVal);**
}
	cid-1	**uimsbf**

} else {aVal = 0.0;} } } }		

NOTA:

- La función del codificador para el cuantificador uniforme es min(255, round((x + 1.0) * 128.0))
- El n.º de bits para huffDecode nemónico es dinámico

[0850] En la tabla de sintaxis anterior, la primera instrucción de conmutación con los cuatro casos (caso 0-3) proporciona una forma mediante la cual se determinan las longitudes de cada vector de la matriz [V] pequeña 717 en términos del número de coeficientes. El primer caso, caso 0, indica que se especifican todos los coeficientes para los vectores V^T_{DIST}. El segundo caso, caso 1, indica que sólo se especifican los coeficientes del vector V^T_{DIST} que corresponden a un orden mayor que un MinNumOfCoeffsForAmbHOA, lo que puede indicar lo que se denomina $(N_{DIST}+1) - (N_{BG}+1)$ anteriormente. El tercer caso, caso 2, es similar al segundo caso, pero resta aún más los coeficientes identificados por NumOfAddAmbHoaChan, que indica una variable para especificar canales adicionales (donde "canales" se refiere a un coeficiente particular que corresponde a un cierto orden, combinación de suborden) correspondiente a un orden que excede el orden N_{BG}. El cuarto caso, caso 3, indica que únicamente se especifican aquellos coeficientes del vector V^T_{DIST} que quedan después de eliminar los coeficientes identificados por NumOfAddAmbHoaChan.

[0851] A este respecto, el dispositivo de decodificación de audio 750A puede funcionar de acuerdo con las técnicas descritas en esta divulgación para determinar un primer conjunto distinto de cero de coeficientes de un vector que representan un componente distinto del campo sonoro, habiéndose descompuesto el vector de una pluralidad de coeficientes armónicos esféricos que describen un campo sonoro.

[0852] Además, el dispositivo de decodificación de audio 750A puede estar configurado para operar de acuerdo con las técnicas descritas en esta divulgación para determinar uno de una pluralidad de modos de configuración mediante los cuales se extrae un conjunto distinto de cero de coeficientes de un vector de acuerdo con una de una pluralidad de modos de configuración, habiéndose descompuesto el vector de una pluralidad de coeficientes armónicos esféricos que describen un campo sonoro y representando un componente distinto del campo sonoro, y

extraer el conjunto distinto de cero de los coeficientes del vector basado en el obtenido de la pluralidad de modos de configuración.

[0853] La FIG. 53 es un diagrama de bloques que ilustra otro ejemplo de un dispositivo de codificación de audio 570 que puede realizar diversos aspectos de las técnicas descritas en esta divulgación. En el ejemplo de la FIG. 53, el dispositivo de codificación de audio 570 puede ser similar a uno o más de los dispositivos de codificación de audio 510A-510J (donde se supone que la unidad de reducción de órdenes 528A está incluida dentro de la unidad de extracción de componentes de campo sonoro 20 pero no mostrada para facilitar la ilustración). Sin embargo, el dispositivo de codificación de audio 570 puede incluir una unidad de transformación más general 572 que puede comprender una unidad de descomposición 518 en algunos ejemplos.

[0854] La FIG. 54 es un diagrama de bloques que ilustra, con más detalle, una implementación ejemplar del dispositivo de codificación de audio 570 mostrado en el ejemplo de la FIG. 53. Como se ilustra en el ejemplo de la FIG. 54, la unidad de transformación 572 del dispositivo de codificación de audio 570 incluye una unidad de rotación 654. La unidad de extracción de componentes de campo sonoro 520 del dispositivo de codificación de audio 570 incluye una unidad de análisis espacial 650, una unidad de análisis de características de contenido 652, una unidad de componentes coherentes de extracción 656 y una unidad de componentes difusos de extracto 658. La unidad de codificación de audio 514 del dispositivo de codificación de audio 570 incluye un motor de codificación AAC 660 y un motor de codificación AAC 162. La unidad de generación de flujo de bits 516 del dispositivo de codificación de audio 570 incluye un multiplexor (MUX) 164.

[0855] El ancho de banda -en términos de bits/segundo- necesario para representar datos de audio en 3D en forma de SHC puede hacer que sea prohibitivo en términos de consumo. Por ejemplo, cuando se utiliza una frecuencia de muestreo de 48 kHz y con 32 bits/misma resolución, una representación SHC de cuarto orden representa un ancho de banda de 36 Mbits/segundo (25 x 48000 x 32 bps). Cuando se compara con la codificación de audio del estado de la técnica para las señales estéreo, que es típicamente aproximadamente de 100 kbits/segundo, esta es una cifra grande. Las técnicas implementadas en el ejemplo de la FIG. 54 pueden reducir el ancho de banda de las representaciones de audio 3D.

[0856] La unidad de análisis espacial 650, la unidad de análisis de características de contenido 652 y la unidad de rotación 654 pueden recibir SHC 511. Como se describe en otra parte en esta divulgación, el SHC 511 puede ser representativo de un campo sonoro. En el ejemplo de la FIG. 54, la unidad de análisis espacial 650, la unidad de análisis de características de contenido 652 y la unidad de rotación 654 pueden recibir veinticinco SHC para una representación de cuarto orden ($n = 4$) del campo sonoro.

[0857] La unidad de análisis espacial 650 puede analizar el campo sonoro representado por el SHC 511 para identificar componentes distintos del campo sonoro y componentes difusos del campo sonoro. Los distintos componentes del campo sonoro son sonidos que se perciben como provenientes de una dirección identificable o que son distintos de los componentes de fondo o difusos del campo sonoro. Por ejemplo, el sonido generado por un instrumento musical individual puede ser percibido como procedente de una dirección identificable. Por el contrario, los componentes difusos o de fondo del campo sonoro no se perciben como procedentes de una dirección identificable. Por ejemplo, el sonido del viento a través de un bosque puede ser un componente difuso de un campo sonoro.

[0858] La unidad de análisis espacial 650 puede identificar uno o más componentes distintos que intentan identificar un ángulo óptimo mediante el cual girar el campo sonoro para alinear aquellos de los distintos componentes que tienen la mayor energía con el eje vertical y/u horizontal (con respecto a un supuesto micrófono que grabó este campo sonoro). La unidad de análisis espacial 650 puede identificar este ángulo óptimo de manera que el campo sonoro pueda girarse de tal manera que estos distintos componentes se alineen mejor con las funciones de base esféricas subyacentes mostradas en los ejemplos las FIGS. 1 y 2.

[0859] En algunos ejemplos, la unidad de análisis espacial 650 puede representar una unidad configurada para realizar una forma de análisis de difusión para identificar un porcentaje del campo sonoro representado por el SHC 511 que incluye sonidos difusos (que pueden referirse a sonidos que tienen bajos niveles de dirección o SHC de menor orden, es decir, aquellos de SHC 511 que tienen una orden menor o igual a uno). Como ejemplo, la unidad de análisis espacial 650 puede realizar un análisis de difusión de una manera similar a la descrita en un documento de Ville Pulkki, titulado "Spatial Sound Reproduction with Directional Audio Coding", publicado en J. Audio Eng. Soc., Vol. 55, nº 6, con fecha de junio de 2007. En algunos casos, la unidad de análisis espacial 650 sólo puede analizar un subconjunto distinto de cero de los coeficientes HOA, tales como los de orden cero y de primer orden e los SHC 511, al realizar el análisis de difusión para determinar el porcentaje de difusión.

[0860] La unidad de análisis de características de contenido 652 puede determinar basándose, al menos en parte, en los SHC 511, si los SHC 511 se generaron a través de una grabación natural de un campo sonoro o se produjeron artificialmente (es decir, sintéticamente) a partir de, como un ejemplo, un objeto de audio, tal como un objeto PCM. Además, la unidad de análisis de características de contenido 652 puede determinar entonces, basándose al menos en parte en si los SHC 511 se generaron a través de una grabación real de un campo sonoro o

de un objeto de audio artificial, el número total de canales a incluir en el flujo de bits 517. Por ejemplo, la unidad de análisis de características de contenido 652 puede determinar, basándose al menos en parte en si los SHC 511 se generaron a partir de una grabación de un campo sonoro real o de un objeto de audio artificial, que el flujo de bits 517 debe incluir dieciséis canales. Cada uno de los canales puede ser un canal mono. La unidad de análisis de características de contenido 652 puede realizar adicionalmente la determinación del número total de canales a incluir en el flujo de bits 517 basado en una velocidad de bits de salida del flujo de bits 517, por ejemplo, de 1,2 Mbps.

[0861] Además, la unidad de análisis de características de contenido 652 puede determinar, basándose al menos en parte en si los SHC 511 se generaron a partir de una grabación de un campo sonoro real o de un objeto de audio artificial, cuántos de los canales a asignar a campos coherentes o, en otras palabras, componentes distintos del campo sonoro y cuántos de los canales a asignar a componentes difusos o, en otras palabras, de fondo del campo sonoro. Por ejemplo, cuando los SHC 511 se generaron a partir de una grabación de un campo sonoro real utilizando, como un ejemplo, un micrófono dedicado, la unidad de análisis de características de contenido 652 puede asignar tres de los canales a componentes coherentes del campo sonoro y puede asignar los canales restantes a componentes difusos del campo sonoro. En este ejemplo, cuando los SHC 511 se generaron a partir de un objeto de audio artificial, la unidad de análisis de características de contenido 652 puede asignar cinco de los canales a componentes coherentes del campo sonoro y puede asignar los canales restantes a componentes difusos del campo sonoro. De esta manera, el bloque de análisis de contenido (es decir, la unidad de análisis de características de contenido 652) puede determinar el tipo de campo sonoro (por ejemplo, difuso/direccional, etc.) y a su vez determinar el número de componentes coherentes/difusos a extraer.

[0862] La velocidad de bits objetivo puede afectar al número de componentes y la velocidad de bits de los motores de codificación AAC individuales (por ejemplo, motores de codificación AAC 660, 662). En otras palabras, la unidad de análisis de características de contenido 652 puede realizar además la determinación de cuántos canales asignar a componentes coherentes y cuántos canales asignar a componentes difusos basándose en una velocidad de bits de salida del flujo de bits 517, por ejemplo, 1,2 Mbps.

[0863] En algunos ejemplos, los canales asignados a componentes coherentes del campo sonoro pueden tener mayores velocidades de bits que los canales asignados a componentes difusos del campo sonoro. Por ejemplo, una velocidad de bits máxima del flujo de bits 517 puede ser de 1,2 Mb/s. En este ejemplo, puede haber cuatro canales asignados a componentes coherentes y 16 canales asignados a componentes difusos. Además, en este ejemplo, cada uno de los canales asignados a los componentes coherentes puede tener una velocidad de bits máxima de 64 kb/s. En este ejemplo, cada uno de los canales asignados a los componentes difusos puede tener una velocidad de bits máxima de 48 kb/s.

[0864] Como se ha indicado anteriormente, la unidad de análisis de características de contenido 652 puede determinar si los SHC 511 se generaron a partir de una grabación de un campo sonoro real o a partir de un objeto de audio artificial. La unidad de análisis de características de contenido 652 puede realizar esta determinación de varias maneras. Por ejemplo, el dispositivo de codificación de audio 570 puede utilizar SHC de 4º orden. En este ejemplo, la unidad de análisis de características de contenido 652 puede codificar 24 canales y predecir un 25º canal (que puede representarse como un vector). La unidad de análisis de características de contenido 652 puede aplicar escalares a, al menos, algunos de los 24 canales y añadir los valores resultantes para determinar el 25º vector. Además, en este ejemplo, la unidad de análisis de características de contenido 652 puede determinar una precisión del 25º canal predicho. En este ejemplo, si la precisión del 25º canal predicho es relativamente alta (por ejemplo, la precisión excede un umbral particular), es probable que los SHC 511 se generen a partir de un objeto de audio sintético. Por el contrario, si la precisión del 25º canal predicho es relativamente baja (por ejemplo, la precisión está por debajo del umbral particular), es más probable que los SHC 511 representen un campo sonoro registrado. Por ejemplo, en este ejemplo, si una relación señal-ruido (SNR) del 25º canal predicho es superior a 100 decibelios (dbs), los SHC 511 son más propensos a representar un campo sonoro generado a partir de un objeto de audio sintético. En contraste, la SNR de un campo sonoro grabado utilizando un micrófono propio puede ser de 5 a 20 dbs. Por lo tanto, puede haber una demarcación aparente en las relaciones SNR entre el campo sonoro representado con los SHC 511 generados a partir de una grabación directa real y desde un objeto de audio sintético.

[0865] Además, la unidad de análisis de características de contenido 652 puede seleccionar, basándose al menos en parte en si los SHC 511 se generaron a partir de una grabación de un campo sonoro real o de un objeto de audio artificial, libros de código para cuantizar el vector V. En otras palabras, la unidad de análisis de características de contenido 652 puede seleccionar diferentes libros de códigos para su uso en la cuantización del vector V, dependiendo de si el campo sonoro representado por los coeficientes HOA está grabado o es sintético.

[0866] En algunos ejemplos, la unidad de análisis de características de contenido 652 puede determinar, en una base recurrente, si los SHC 511 se generaron a partir de una grabación de un campo sonoro real o a partir de un objeto de audio artificial. En algunos de estos ejemplos, la base recurrente puede ser cada trama. En otros ejemplos, la unidad de análisis de características de contenido 652 puede realizar esta determinación una vez. Además, la unidad de análisis de características de contenido 652 puede determinar, en una base recurrente, el número total de canales y la asignación de canales de componentes coherentes y canales de componentes difusos. En algunos de estos ejemplos, la base recurrente puede ser cada trama. En otros ejemplos, la unidad de análisis de características

de contenido 652 puede realizar esta determinación una vez. En algunos ejemplos, la unidad de análisis de características de contenido 652 puede seleccionar, en una base recurrente, libros de códigos para su uso en la cuantización del vector V. En algunos de estos ejemplos, la base recurrente puede ser cada trama. En otros ejemplos, la unidad de análisis de características de contenido 652 puede realizar esta determinación una vez.

[0867] La unidad de rotación 654 puede realizar una operación de rotación de los coeficientes HOA. Como se analiza en otra parte de esta divulgación (por ejemplo, con respecto a la FIG. 55 y 55B), la realización de la operación de rotación puede reducir el número de bits requeridos para representar el SHC 511. En algunos ejemplos, el análisis de rotación realizado por la unidad de rotación 652 es un caso de un análisis de descomposición en valores singulares ("SVD"). El análisis de componentes principales ("PCA"), el análisis de componentes independientes ("ICA") y la transformación de Karhunen-Loeve ("KLT") son técnicas relacionadas que pueden ser aplicables.

[0868] En el ejemplo de la FIG. 54, la unidad de componentes coherentes de extracción 656 recibe SHC 511 girados de la unidad de rotación 654. Además, la unidad de componentes coherentes de extracción 656 extrae, a partir de los SHC 511 girados, los de los SHC 511 girados asociados con los componentes coherentes del campo sonoro.

[0869] Además, la unidad de componentes coherentes de extracción 656 genera uno o más canales de componentes coherentes. Cada uno de los canales de componentes coherentes puede incluir un subconjunto diferente de los SHC 511 girados asociados con los coeficientes coherentes del campo sonoro. En el ejemplo de la FIG. 54, la unidad de componentes coherentes de extracción 656 puede generar de uno a 16 canales de componentes coherentes. El número de canales de componentes coherentes generados por la unidad de componentes coherentes de extracción 656 puede determinarse por el número de canales asignados por la unidad de análisis de características de contenido 652 a los componentes coherentes del campo sonoro. Las velocidades de bits de los canales de componentes coherentes generados por la unidad de componentes coherentes de extracción 656 pueden ser las determinadas por la unidad de análisis de características de contenido 652.

[0870] De forma similar, en el ejemplo de la FIG. 54, la unidad de componentes difusos de extracción 658 recibe SHC 511 girados de la unidad de rotación 654. Además, la unidad de componentes difusos de extracción 658 extrae, a partir de los SHC 511 girados, los de los SHC 511 girados asociados con los componentes difusos del campo sonoro.

[0871] Además, la unidad de componentes difusos de extracción 658 genera uno o más canales de componentes difusos. Cada uno de los canales de componentes difusos puede incluir un subconjunto diferente de los SHC 511 girados asociados con los coeficientes difusos del campo sonoro. En el ejemplo de la FIG. 54, la unidad de componentes difusos de extracción 658 puede generar de uno a 9 canales de componentes difusos. El número de canales de componentes difusos generados por la unidad de componentes difusos de extracción 658 puede determinarse por el número de canales asignados por la unidad de análisis de características de contenido 652 a los componentes difusos del campo sonoro. Las velocidades de bits de los canales de componentes difusos generados por la unidad de componentes difusos de extracción 658 pueden ser las determinadas por la unidad de análisis de características de contenido 652.

[0872] En el ejemplo de la FIG. 54, la unidad de codificación AAC 660 puede usar un códec AAC para codificar los canales de componentes coherentes generados por la unidad de componentes coherentes de extracción 656. De forma similar, la unidad de codificación AAC 662 puede usar un códec AAC para codificar los canales de componentes difusos generados por la unidad de componentes difusos de extracción 658. El multiplexor 664 ("MUX 664") puede multiplexar los canales de componentes coherentes codificados y los canales de componentes difusos codificados, junto con datos secundarios (por ejemplo, un ángulo óptimo determinado por la unidad de análisis espacial 650), para generar el flujo de bits 517.

[0873] De este modo, las técnicas pueden permitir que el dispositivo de codificación de audio 570 determine si se generan coeficientes armónicos esféricos representativos de un campo sonoro a partir de un objeto de audio sintético.

[0874] En algunos ejemplos, el dispositivo de codificación de audio 570 puede determinar, basándose en si los coeficientes armónicos esféricos se generan a partir de un objeto de audio sintético, un subconjunto de los coeficientes armónicos esféricos representativos de componentes distintos del campo sonoro. En estos y otros ejemplos, el dispositivo de codificación de audio 570 puede generar un flujo de bits para incluir el subconjunto de los coeficientes armónicos esféricos. El dispositivo de codificación de audio 570 puede, en algunos casos, codificar en audio el subconjunto de los coeficientes armónicos esféricos, y generar un flujo de bits para incluir el subconjunto codificado en audio de los coeficientes armónicos esféricos.

[0875] En algunos ejemplos, el dispositivo de codificación de audio 570 puede determinar, basándose en si los coeficientes armónicos esféricos se generan a partir de un objeto de audio sintético, un subconjunto de los coeficientes armónicos esféricos representativos de componentes de fondo del campo sonoro. En estos y otros

ejemplos, el dispositivo de codificación de audio 570 puede generar un flujo de bits para incluir el subconjunto de los coeficientes armónicos esféricos. En estos y otros ejemplos, el dispositivo de codificación de audio 570 puede codificar en audio el subconjunto de los coeficientes armónicos esféricos, y generar un flujo de bits para incluir el subconjunto codificado en audio de los coeficientes armónicos esféricos.

[0876] En algunos ejemplos, el dispositivo de codificación de audio 570 puede realizar un análisis espacial con respecto a los coeficientes armónicos esféricos para identificar un ángulo por el que girar el campo sonoro representado por los coeficientes armónicos esféricos y realizar una operación de rotación para girar el campo sonoro por el ángulo identificado para generar coeficientes armónicos esféricos rotados.

[0877] En algunos ejemplos, el dispositivo de codificación de audio 570 puede determinar, basándose en si los coeficientes armónicos esféricos se generan a partir de un objeto de audio sintético, un primer subconjunto de los coeficientes armónicos esféricos representativos de componentes distintos del campo sonoro, y determinar, en base a si los coeficientes armónicos esféricos se generan a partir de un objeto de audio sintético, un segundo subconjunto de los coeficientes armónicos esféricos representativos de los componentes de fondo del campo sonoro. En estos y otros ejemplos, el dispositivo de codificación de audio 570 puede codificar en audio el primer subconjunto de los coeficientes armónicos esféricos que tienen una velocidad de bits objetivo más alta que la utilizada para codificar en audio el segundo subconjunto de los coeficientes armónicos esféricos.

[0878] De esta manera, varios aspectos de las técnicas pueden permitir que el dispositivo de codificación de audio 570 determine si SCH 511 se generan a partir de un objeto de audio sintético de acuerdo con las siguientes cláusulas.

Cláusula 132512-1. Un dispositivo, tal como el dispositivo de codificación de audio 570, que comprende: en el que uno o más procesadores están configurados adicionalmente para determinar si se generan coeficientes armónicos esféricos representativos de un campo sonoro a partir de un objeto de audio sintético.

Cláusula 132512-2. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están configurados además para, al determinar si los coeficientes armónicos esféricos representativos del campo sonoro se generan a partir del objeto de audio sintético, excluir un primer vector de una matriz de coeficientes esféricos armónicos entramada almacenando al menos una parte de los coeficientes armónicos esféricos representativos del campo sonoro para obtener una matriz de coeficientes armónicos esféricos entramada reducida.

Cláusula 132512-3. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están configurados además para, al determinar si los coeficientes armónicos esféricos representativos del campo sonoro se generan a partir del objeto de audio sintético, excluir un primer vector de una matriz de coeficientes esféricos armónicos entramada almacenando al menos una parte de los coeficientes armónicos esféricos representativos del campo sonoro para obtener una matriz de coeficientes armónicos esféricos entramada reducida, y predecir un vector de la matriz de coeficientes armónicos esféricos entramada reducida en base a los vectores restantes de la matriz de coeficientes armónicos esféricos entramada reducida.

Cláusula 132512-4. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están configurados además para, al determinar si los coeficientes armónicos esféricos representativos del campo sonoro se generan a partir del objeto de audio sintético, excluir un primer vector de una matriz de coeficientes esféricos armónicos entramada almacenando al menos una parte de los coeficientes armónicos esféricos representativos del campo sonoro para obtener una matriz de coeficientes armónicos esféricos entramada reducida, y predecir un vector de la matriz de coeficientes armónicos esféricos entramada reducida en base, al menos en parte, a una suma de los vectores restantes de la matriz de coeficientes armónicos esféricos entramada reducida.

Cláusula 132512-5. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están configurados además para, al determinar si los coeficientes armónicos esféricos representativos del campo sonoro se generan a partir del objeto de audio sintético, predecir un vector de una matriz de coeficientes armónicos esféricos entramada que almacena al menos una parte de los coeficientes armónicos esféricos en base, al menos en parte, a una suma de vectores restantes de la matriz de coeficientes armónicos esféricos entramada.

Cláusula 132512-6. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están configurados además para, al determinar si los coeficientes armónicos esféricos representativos del campo sonoro se generan a partir del objeto de audio sintético, predecir un vector de una matriz de coeficientes armónicos esféricos entramada que almacena al menos una parte de los coeficientes armónicos esféricos en base, al menos en parte, a una suma de vectores restantes de la matriz de coeficientes armónicos esféricos entramada, y calcular un error basado en el vector predicho.

Cláusula 132512-7. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están

configurados además para, al determinar si los coeficientes armónicos esféricos representativos del campo sonoro se generan a partir del objeto de audio sintético, predecir un vector de una matriz de coeficientes armónicos esféricos entramada que almacena al menos una parte de los coeficientes armónicos esféricos en base, al menos en parte, a una suma de vectores restantes de la matriz de coeficientes armónicos esféricos entramada, y calcular un error basado en el vector predicho y el vector correspondiente de la matriz de coeficientes armónicos esféricos entramada.

Cláusula 132512-8. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están configurados además para, al determinar si los coeficientes armónicos esféricos representativos del campo sonoro se generan a partir del objeto de audio sintético, predecir un vector de una matriz de coeficientes armónicos esféricos entramada que almacena al menos una parte de los coeficientes armónicos esféricos en base, al menos en parte, a una suma de vectores restantes de la matriz de coeficientes armónicos esféricos entramada, y calcular un error como una suma del valor absoluto de la diferencia del vector predicho y el vector correspondiente de la matriz de coeficientes armónicos esféricos entramada.

Cláusula 132512-9. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están configurados, además para, al determinar si los coeficientes armónicos esféricos representativos del campo sonoro se generan a partir del objeto de audio sintético, predecir un vector de una matriz de coeficientes armónicos esféricos enmarcada que almacena al menos una parte de los coeficientes armónicos esféricos basado, al menos en parte, en una suma de vectores restantes de La matriz de coeficientes armónicos esféricos por trama, calcular un error basado en el vector predicho y el correspondiente vector de la matriz de coeficientes armónicos esféricos por trama, calcular una razón basada en una energía del vector correspondiente de la matriz de coeficientes armónicos esféricos por trama y el error, y Comparar la relación con un umbral para determinar si los coeficientes armónicos esféricos representativos del campo sonoro se generan a partir del objeto de audio sintético

Cláusula 132512-10. El dispositivo de cualquiera de las reivindicaciones 4-9, en el que el uno o más procesadores están configurados adicionalmente para, al predecir el vector, predecir un primer vector distinto de cero de la matriz de coeficientes armónicos esféricos entramada que almacena al menos la porción de los coeficientes armónicos esféricos.

Cláusula 132512-11. El dispositivo de cualquiera de las reivindicaciones 1-10, en el que el uno o más procesadores están configurados además para especificar una indicación de si los coeficientes armónicos esféricos se generan a partir del objeto de audio sintético en un flujo de bits que almacena una versión comprimida de los coeficientes armónicos esféricos.

Cláusula 132512-12. El dispositivo de la cláusula 132512-11, en el que la indicación es un bit único.

Cláusula 132512-13. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están configurados además para determinar, basándose en si los coeficientes armónicos esféricos se generan a partir de un objeto de audio sintético, un subconjunto de los coeficientes armónicos esféricos representativos de componentes distintos del campo sonoro.

Cláusula 132512-14. El dispositivo de la cláusula 132512-13, en el que el uno o más procesadores están configurados además para generar un flujo de bits para incluir el subconjunto de los coeficientes armónicos esféricos.

Cláusula 132512-15. El dispositivo de la cláusula 132512-13, en el que el uno o más procesadores están configurados adicionalmente para codificar en audio el subconjunto de los coeficientes armónicos esféricos, y generar un flujo de bits para incluir el subconjunto codificado en audio de los coeficientes armónicos esféricos.

Cláusula 132512-16. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están configurados además para determinar, basándose en si los coeficientes armónicos esféricos se generan a partir de un objeto de audio sintético, un subconjunto de los coeficientes armónicos esféricos representativos de componentes de fondo del campo sonoro.

Cláusula 132512-17. El dispositivo de la cláusula 132512-16, en el que el uno o más procesadores están configurados además para generar un flujo de bits para incluir el subconjunto de los coeficientes armónicos esféricos.

Cláusula 132512-18. El dispositivo de la cláusula 132512-15, en el que el uno o más procesadores están configurados adicionalmente para codificar en audio el subconjunto de los coeficientes armónicos esféricos, y generar un flujo de bits para incluir el subconjunto codificado en audio de los coeficientes armónicos esféricos.

Cláusula 132512-19. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están configurados adicionalmente para realizar un análisis espacial con respecto a los coeficientes armónicos

esféricos para identificar un ángulo mediante el cual girar el campo sonoro representado por los coeficientes armónicos esféricos, y realizar una operación de rotación para girar el campo sonoro por el ángulo identificado para generar coeficientes armónicos esféricos girados.

Cláusula 132512-20. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están configurados adicionalmente para determinar, basándose en si los coeficientes armónicos esféricos se generan a partir de un objeto de audio sintético, un primer subconjunto de los coeficientes armónicos esféricos representativos de componentes distintos del campo sonoro, y determinar, basándose en si los coeficientes armónicos esféricos se generan a partir de un objeto de audio sintético, un segundo subconjunto de los coeficientes armónicos esféricos representativos de los componentes de fondo del campo sonoro.

Cláusula 132512-21. El dispositivo de la cláusula 132512-20, en el que el uno o más procesadores están configurados adicionalmente para codificar en audio el primer subconjunto de los coeficientes armónicos esféricos que tienen una velocidad de bits objetivo mayor que la usada para codificar en audio al segundo subconjunto de los coeficientes armónicos esféricos.

Cláusula 132512-22. El dispositivo de la cláusula 132512-1, en el que el uno o más procesadores están configurados además para realizar una descomposición en valores singulares con respecto a los coeficientes armónicos esféricos para generar una matriz U representativa de los vectores singulares izquierdos de la pluralidad de coeficientes armónicos esféricos, una matriz S representativa de los valores singulares de la pluralidad de coeficientes armónicos esféricos y una matriz V representativa de los vectores singulares derechos de la pluralidad de coeficientes armónicos esféricos.

Cláusula 132512-23. El dispositivo de la cláusula 132512-22, en el que el uno o más procesadores están configurados adicionalmente para determinar, basándose en si los coeficientes armónicos esféricos se generan a partir de un objeto de audio sintético, aquellas porciones de una o más de la matriz U, la matriz S y la matriz V representativas de los componentes distintos del campo sonoro.

Cláusula 132512-24. El dispositivo de la cláusula 132512-22, en el que el uno o más procesadores están configurados adicionalmente para determinar, basándose en si los coeficientes armónicos esféricos se generan a partir de un objeto de audio sintético, aquellas porciones de una o más de la matriz U, la matriz S y la matriz V representativas de los componentes de fondo del campo sonoro.

Cláusula 132512-1C. Un dispositivo, tal como el dispositivo de codificación de audio 570, que comprende: uno o más procesadores configurados para determinar si se generan coeficientes armónicos esféricos representativos de un campo sonoro a partir de un objeto de audio sintético en base a una relación calculada en función de, al menos, una energía de un vector de los coeficientes armónicos esféricos y un error derivado en base a una versión predicha del vector de los coeficientes armónicos esféricos y el vector de los coeficientes armónicos esféricos.

[0879] En cada una de las diversas instancias descritas anteriormente, debe entenderse que el dispositivo de codificación de audio 570 puede realizar un procedimiento o comprender de otro modo medios para realizar cada paso del procedimiento para el que está configurado el dispositivo de codificación de audio 570. En algunos casos, estos medios pueden comprender uno o más procesadores. En algunos casos, 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 instrucciones que, cuando se ejecutan, hacen que uno o más procesadores realicen el procedimiento para el cual el dispositivo de codificación de audio 570 se ha configurado para llevar a cabo.

[0880] Las FIGS. 55 y 55B son diagramas que ilustran un ejemplo de realización de diversos aspectos de las técnicas descritas en esta divulgación para girar un campo sonoro 640. La FIG. 55 es un diagrama que ilustra el campo sonoro 640 antes de la rotación de acuerdo con los diversos aspectos de las técnicas descritas en esta divulgación. En el ejemplo de la FIG. 55, el campo sonoro 640 incluye dos ubicaciones de alta presión, designadas como ubicación 642A y 642B. Esta ubicación 642A y 642B ("ubicaciones 642") residen a lo largo de una línea 644 que tiene una pendiente distinta de cero (que es otra forma de referirse a una línea que no es horizontal, ya que las líneas horizontales tienen una pendiente de cero). Dado que las ubicaciones 642 tienen una coordenada z, además de las coordenadas x e y, pueden requerirse funciones de base esférica de orden superior para representar correctamente este campo sonoro 640 (ya que estas funciones de base esférica de orden superior describen las partes superior e inferior o no horizontales del campo sonoro. En lugar de reducir el campo sonoro 640 directamente a los SHC 511, el dispositivo de codificación de audio 570 puede girar el campo sonoro 640 hasta que la línea 644 que conecta las ubicaciones 642 es horizontal.

[0881] La FIG. 55B es un diagrama que ilustra el campo sonoro 640 después de ser girado hasta que la línea 644 que conecta las ubicaciones 642 es horizontal. Como resultado de la rotación del campo sonoro 640 de esta manera,

los SHC 511 pueden obtenerse de tal manera que los órdenes superiores de los SHC 511 se especifican como ceros dado que el campo sonoro girado 640 ya no tiene posiciones de presión (o energía) con coordenadas z. De esta manera, el dispositivo de codificación de audio 570 puede girar, trasladar, más generalmente, ajustar el campo sonoro 640 para reducir el número de SHC 511 con valores distintos de cero. En relación con otros aspectos de las técnicas, el dispositivo de codificación de audio 570 puede entonces, en lugar de señalar un número firmado de 32 bits que identifica que estos de orden superior de SHC 511 tienen valores cero, señalan en un campo del flujo de bits 517 que estos de orden superior de SHC 511 no están señalizados. El dispositivo de codificación de audio 570 también puede especificar información de rotación en el flujo de bits 517 que indica cómo se giró el campo sonoro 640, a menudo mediante la expresión de un acimut y elevación de la manera descrita anteriormente. Un dispositivo de extracción, tal como el dispositivo de codificación de audio, puede entonces implicar que los no señalados de los SHC 511 tienen un valor de cero y, al reproducir el campo sonoro 640 basado en SHC 511, realizar la rotación para girar el campo sonoro 640 de manera que el campo sonoro 640 se asemeje al campo sonoro 640 mostrado en el ejemplo de la FIG. 55. De esta manera, el dispositivo de codificación de audio 570 puede reducir el número de SHC 511 requeridos que se especifiquen en el flujo de bits 517 de acuerdo con las técnicas descritas en esta divulgación.

[0882] Se puede usar un algoritmo de "compactación espacial" para determinar la rotación óptima del campo sonoro. En una realización, el dispositivo de codificación de audio 570 puede realizar el algoritmo para iterar a través de todas las posibles combinaciones en acimut y elevación (es decir, combinaciones 1024 x 512 en el ejemplo anterior), girando el campo sonoro para cada combinación, y calculando el número de SHC 511 que están por encima del valor umbral. La combinación acimut/elevación candidata que produce el menor número de SHC 511 por encima del valor umbral puede considerarse lo que se puede denominar la "rotación óptima". En esta forma rotada, el campo sonoro puede requerir el menor número de SHC 511 para representar el campo sonoro y puede entonces considerarse compactado. En algunos casos, el ajuste puede comprender esta rotación óptima y la información de ajuste descrita anteriormente puede incluir esta información de rotación (que se puede denominar "rotación óptima") (en términos de ángulos de acimut y elevación).

[0883] En algunos casos, en lugar de solamente especificar el ángulo de acimut y el ángulo de elevación, el dispositivo de codificación de audio 570 puede especificar ángulos adicionales en la forma, como un ejemplo, de ángulos de Euler. Los ángulos de Euler especifican el ángulo de rotación en torno al eje z, el eje x anterior y el eje z anterior. Aunque se describen en esta divulgación con respecto a combinaciones de ángulos de acimut y de elevación, las técnicas de esta divulgación no deben limitarse a especificar únicamente los ángulos de acimut y de elevación, pero pueden incluir especificar cualquier número de ángulos, incluyendo los tres ángulos de Euler mencionados anteriormente. En este sentido, el dispositivo de codificación de audio 570 puede girar el campo sonoro para reducir un número de la pluralidad de elementos jerárquicos que proporcionan información relevante para describir el campo sonoro y especificar ángulos de Euler como información de rotación en el flujo de bits. Los ángulos de Euler, como se señaló anteriormente, pueden describir cómo se hizo girar el campo sonoro. Cuando se utilizan ángulos de Euler, el dispositivo de extracción de flujo de bits puede analizar el flujo de bits para determinar la información de rotación que incluye los ángulos de Euler y, al reproducir el campo sonoro en base a aquellos de la pluralidad de elementos jerárquicos que proporcionan información relevante para describir el campo sonoro, girar el campo sonoro en base a los ángulos de Euler.

[0884] Además, en algunos casos, en lugar de especificar explícitamente estos ángulos en el flujo de bits 517, el dispositivo de codificación de audio 570 puede especificar un índice (que puede denominarse "índice de rotación") asociado con combinaciones predefinidas del uno o más ángulos que especifican la rotación. En otras palabras, la información de rotación puede, en algunos casos, incluir el índice de rotación. En estos casos, un valor dado del índice de rotación, tal como un valor de cero, puede indicar que no se realizó ninguna rotación. Este índice de rotación se puede utilizar en relación con una tabla de rotación. Es decir, el dispositivo de codificación de audio 570 puede incluir una tabla de rotación que comprende una entrada para cada una de las combinaciones del ángulo de acimut y el ángulo de elevación.

[0885] Alternativamente, la tabla de rotación puede incluir una entrada para cada transformación de matriz representativa de cada combinación del ángulo de acimut y el ángulo de elevación. Es decir, el dispositivo de codificación de audio 570 puede almacenar una tabla de rotación que tiene una entrada para cada transformación de matriz para girar el campo sonoro por cada una de las combinaciones de ángulos de acimut y elevación. Típicamente, el dispositivo de codificación de audio 570 recibe los SHC 511 y deriva los SHC 511', cuando se realiza la rotación, de acuerdo con la siguiente ecuación:

$$\begin{bmatrix} SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (25x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x25) \end{bmatrix} \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

En la ecuación anterior, los SHC 511' se calculan en función de una matriz de codificación para codificar un campo sonoro en términos de una segunda trama referencia (*EncMat*$_2$), una matriz de inversión para revertir los SHC 511

de vuelta a un campo sonoro en términos de una primera trama de referencia ($InvMat_1$), y SHC 511. $EncMat_2$ es de tamaño 25 x 32, mientras que $InvMat_2$ es de tamaño 32 x 25. Tanto SHC 511' como SHC 511 son de tamaño 25, donde SHC 511' puede reducirse aún más debido a la eliminación de aquellos que no especifican información de audio relevante. $EncMat_2$ puede variar para cada combinación de ángulo de acimut y de elevación, mientras que $InvMat_1$ puede permanecer estático con respecto a cada combinación de ángulo de azimut y de elevación. La tabla de rotación puede incluir una entrada que almacena el resultado de multiplicar cada $EncMat_2$ diferente a $InvMat_1$.

[0886] La FIG. 56 es un diagrama que ilustra un campo sonoro ejemplar capturado de acuerdo con una primera trama de referencia que se hace girar de acuerdo con las técnicas descritas en esta divulgación para expresar el campo sonoro en términos de una segunda trama de referencia. En el ejemplo de la FIG. 56, el campo sonoro que rodea un micrófono dedicado 646 se captura asumiendo una primera trama de referencia, que se representa por los ejes X1, Y1 y Z1 en el ejemplo de la FIG. 56. SHC 511 describen el campo sonoro en términos de esta primera trama de referencia. $InvMat_1$ transforma los SHC 511 de nuevo en el campo de sonido, permitiendo que el campo sonoro se gire a la segunda trama de referencia representada por los ejes X2, Y2 y Z2 en el ejemplo de la FIG. 56. $EncMat_2$ descrito anteriormente puede girar el campo sonoro y generar SHC 511' que describen este campo sonoro girado en términos de la segunda trama de referencia.

[0887] En cualquier caso, la ecuación anterior puede derivarse como se indica a continuación. Dado que el campo sonoro se registra con un cierto sistema de coordenadas, de modo que el frente se considera la dirección del eje x, las 32 posiciones de micrófono de un micrófono dedicado (u otras configuraciones de micrófono) se definen a partir de este sistema de coordenadas de referencia. La rotación del campo sonoro puede entonces considerarse como una rotación de esa trama de referencia. Para la supuesta trama de referencia, SHC 511 pueden calcularse como se indica a continuación:

$$\begin{bmatrix} SHC \\ 27 \end{bmatrix} = \begin{bmatrix} Y_0^0(Pos_1) & Y_0^0(Pos_2) & \cdots & Y_0^0(Pos_{32}) \\ Y_1^{-1}(Pos_1) & \cdots & & Y_1^{-1}(Pos_{32}) \\ \vdots & \ddots & & \\ Y_4^4(Pos_1) & & & Y_4^4(Pos_{32}) \end{bmatrix} \begin{bmatrix} mic_1(t) \\ mic_2(t) \\ \vdots \\ mic_{32}(t) \end{bmatrix}$$

En la ecuación anterior, Y_n^m representan las funciones de base esférica en la posición (Pos_i) del i -ésimo micrófono (donde i puede ser 1-32 en este ejemplo). El vector mic_i indica la señal de micrófono para el i-ésimo micrófono durante un tiempo t. Las posiciones (Pos_i) se refieren a la posición del micrófono en la primera trama de referencia (es decir, la trama de referencia antes de la rotación en este ejemplo).

[0888] La ecuación anterior puede expresarse alternativamente en términos de las expresiones matemáticas indicadas anteriormente como:

$$[SHC_27] = [E_S(\theta,\varphi)][m_i(t)].$$

[0889] Para girar el campo sonoro (o en la segunda trama de referencia), la posición (Pos_i) se calculará en la segunda trama de referencia. Mientras estén presentes las señales originales del micrófono, el campo sonoro puede girarse arbitrariamente. Sin embargo, las señales de micrófono originales ($mic_i(t)$) a menudo no están disponibles. El problema entonces puede ser cómo recuperar las señales de micrófono ($mic_i(t)$) de los SHC 511. Si se utiliza un diseño en T (como en un micrófono dedicado de 32 micrófonos), la solución a este problema se puede lograr resolviendo la siguiente ecuación:

$$\begin{bmatrix} mic_1(t) \\ mic_2(t) \\ \vdots \\ mic_{32}(t) \end{bmatrix} = \left[InvMat_1 \right] \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

Este *InvMat1* puede especificar las funciones de base armónicas esféricas calculadas de acuerdo con la posición de los micrófonos como se especifica con respecto a la primera trama de referencia. Esta ecuación también puede expresarse como $[m_i(t)] = [E_S(\theta,\varphi)]^{-1}[SHC]$, como se ha indicado anteriormente.

[0890] Una vez que se recuperan las señales de micrófono ($mic_i(t)$) de acuerdo con la ecuación anterior, las señales de micrófono ($mic_i(t)$) que describen el campo sonoro pueden girarse para calcular los SHC 511' correspondientes a la segunda trama de referencia, dando como resultado la siguiente ecuación:

$$\begin{bmatrix} SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (25x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x25) \end{bmatrix} \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

EncMat₂ especifica las funciones de base armónica esférica desde una posición girada (*Pos$_i$*). De esta manera, *EncMat₂* puede especificar eficazmente una combinación del ángulo de acimut y de elevación. Por lo tanto, cuando la tabla de rotación almacena el resultado de $\begin{bmatrix} EncMat_2 \\ (25x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x25) \end{bmatrix}$ para cada combinación de ángulos de acimut y de elevación, la tabla de rotación especifica eficazmente cada combinación de los ángulos de acimut y de elevación. La señal recibida puede expresarse entonces como:

$$[SHC_27'] = [E_S(\theta_2,\varphi_2)]\,[E_S(\theta_1,\varphi_1)]^{-1}\,[SHC\ 27],$$

donde θ_2, φ_2 representan un segundo ángulo de acimut y un segundo ángulo de elevación diferente del primer ángulo de acimut y el ángulo de elevación representado por θ_1, φ_1. θ_1, φ_1 corresponden a la primera trama de referencia, mientras que θ_2, φ_2 corresponden a la segunda trama de referencia. Por lo tanto, *InvMat₁* puede corresponder a $[E_S(\theta_1,\varphi_1)]^{-1}$, mientras que *EncMat₂* puede corresponder a $[E_S(\theta_2,\varphi_2)]$

[0891] Lo anterior puede representar una versión más simplificada del cálculo que no considera la operación de filtrado, representada anteriormente en varias ecuaciones que indican la derivación de los SHC 511 en el dominio de frecuencia por la función $j_n(.)$, que se refiere a la función esférica de Bessel de orden *n*. En el dominio del tiempo, esta función $j_n(.)$ representa una operación de filtrado que es específica de un orden particular, *n*. Con el filtrado, la rotación se puede realizar por orden. Para más ilustración, considérense las siguientes ecuaciones:

$$a_n^k(t) = b_n(t) * \left\langle \left[Y_n^m \right], \left[m_i(t) \right] \right\rangle$$

$$a_n^k(t) = \left\langle \left[Y_n^m \right], b_n(t) * \left[m_i(t) \right] \right\rangle$$

[0892] A partir de estas ecuaciones, los SHC 511' girados para órdenes se realizan por separado ya que $b_n(t)$ son diferentes para cada orden. Como resultado, la ecuación anterior puede ser alterada de la siguiente manera para calcular los de primer orden de los SHC 511' girados:

$$\begin{bmatrix} 1^{st} \\ Order \\ SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (3x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x3) \end{bmatrix} \begin{bmatrix} 1^{st} \\ Order \\ SHC \\ 27 \end{bmatrix}$$

Dado que hay tres de primer orden de SHC 511, cada uno de los vectores SHC 511' y 511 son 3 de tamaño en la ecuación anterior. Igualmente, para el segundo orden se puede aplicar la siguiente ecuación:

$$\begin{bmatrix} 2^{nd} \\ Order \\ SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (5x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x5) \end{bmatrix} \begin{bmatrix} 2^{nd} \\ Order \\ SHC \\ 27 \end{bmatrix}$$

De nuevo, dado que hay cinco SHC 511 de segundo orden, cada uno de los vectores SHC 511' y 511 son 5 de tamaño en la ecuación anterior. Las ecuaciones restantes para los demás órdenes, es decir, los órdenes tercero y cuarto, pueden ser similares a las descritas anteriormente, siguiendo el mismo patrón con respecto a los tamaños de las matrices (dado que el número de filas de $EncMat_2$, el número de columnas de $InvMat_1$ y los tamaños de los vectores SHC 511 y SHC 511' de tercer y cuarto orden son iguales al número de subórdenes (m veces dos más 1) de cada una de las funciones de base armónicas esféricas de tercer y cuarto orden.

[0893] Por tanto, el dispositivo de codificación de audio 570 puede realizar esta operación de rotación con respecto a cada combinación de ángulo de acimut y de elevación en un intento de identificar la llamada rotación óptima. El dispositivo de codificación de audio 570 puede, después de realizar esta operación de rotación, calcular el número de SHC 511' por encima del valor de umbral. En algunos casos, el dispositivo de codificación de audio 570 puede realizar esta rotación para derivar una serie de SHC 511' que representan el campo sonoro durante una duración de tiempo, tal como una trama de audio. Al realizar esta rotación para derivar la serie del SHC 511' que representan el campo sonoro durante esta duración de tiempo, el dispositivo de codificación de audio 570 puede reducir el número de operaciones de rotación que tienen que realizarse en comparación con hacer esto para cada conjunto de los SHC 511 que describen el campo sonoro para duraciones de tiempo menores que una trama u otra longitud. En cualquier caso, el dispositivo de codificación de audio 570 se puede ahorrar, a lo largo de este proceso, los SHC 511' que tienen el número mínimo de SHC 511' mayores que el valor umbral.

[0894] Sin embargo, la realización de esta operación de rotación con respecto a cada combinación de ángulo de acimut y elevación puede ser intensiva en el procesador o requerir mucho tiempo. Como resultado, el dispositivo de codificación de audio 570 no puede realizar lo que podría caracterizarse como esta implementación de "fuerza bruta" del algoritmo de rotación. En cambio, el dispositivo de codificación de audio 570 puede realizar rotaciones con respecto a un subconjunto de combinaciones posiblemente conocidas (estadísticamente) de ángulo de acimut y de elevación que ofrecen una compactación generalmente buena, realizando rotaciones adicionales con respecto a combinaciones alrededor de las de este subconjunto proporcionando mejor compactación en comparación con otras combinaciones en el subconjunto.

[0895] Como otra alternativa, el dispositivo de codificación de audio 570 puede realizar esta rotación con respecto solamente al subconjunto conocido de combinaciones. Como otra alternativa, el dispositivo de codificación de audio 570 puede seguir una trayectoria (espacialmente) de combinaciones, realizando las rotaciones con respecto a esta trayectoria de combinaciones. Como otra alternativa, el dispositivo de codificación de audio 570 puede especificar un umbral de compactación que define un número máximo de SHC 511' que tienen valores distintos de cero por encima del valor de umbral. Este umbral de compactación puede establecer eficazmente un punto de parada para la búsqueda, de tal forma que, cuando el dispositivo de codificación de audio 570 realiza una rotación y determina que el número de SHC 511' con un valor por encima del umbral establecido es menor o igual (o menor que en algunos casos) que el umbral de compactación, el dispositivo de codificación de audio 570 deja de realizar cualquier operación de rotación adicional con respecto a las combinaciones restantes. Como otra alternativa, el dispositivo de codificación de audio 570 puede atravesar un árbol (u otra estructura de datos) dispuesto jerárquicamente de combinaciones, realizando las operaciones de rotación con respecto a la combinación actual y recorriendo el árbol a la derecha o a la izquierda (por ejemplo, para árboles binarios) dependiendo del número de SHC 511' con un valor distinto de cero mayor que el valor umbral.

[0896] En este sentido, cada una de estas alternativas implica realizar una primera y segunda operación de rotación y comparar el resultado de realizar la primera y segunda operación de rotación para identificar una de las primera y segunda operaciones de rotación que da como resultado el menor número de los SHC 511' con un valor distinto de cero mayor que el valor umbral. Por consiguiente, el dispositivo de codificación de audio 570 puede realizar una primera operación de rotación en el campo sonoro para girar el campo sonoro de acuerdo con un primer ángulo de acimut y un primer ángulo de elevación y determinar un primer número de la pluralidad de elementos jerárquicos representativos del campo sonoro girado de acuerdo con el primer ángulo de acimut y el primer ángulo de elevación que proporcionan información relevante para describir el campo sonoro. El dispositivo de codificación de audio 570 también puede realizar una segunda operación de rotación en el campo sonoro para girar el campo sonoro de acuerdo con un segundo ángulo de acimut y un segundo ángulo de elevación y determinar un segundo número de la pluralidad de elementos jerárquicos representativos del campo sonoro girado de acuerdo con el segundo ángulo de acimut y el segundo ángulo de elevación que proporcionan información relevante para describir el campo sonoro. Además, el dispositivo de codificación de audio 570 puede seleccionar la primera operación de rotación o la segunda operación de rotación basándose en una comparación del primer número de la pluralidad de elementos jerárquicos y el segundo número de la pluralidad de elementos jerárquicos.

[0897] En algunos casos, el algoritmo de rotación se puede realizar con respecto a una duración de tiempo, donde las invocaciones posteriores del algoritmo de rotación pueden realizar operaciones de rotación basadas en invocaciones pasadas del algoritmo de rotación. En otras palabras, el algoritmo de rotación puede ser adaptativo basándose en la información de rotación pasada determinada al girar el campo sonoro durante una duración de tiempo anterior. Por ejemplo, el dispositivo de codificación de audio 570 puede hacer girar el campo sonoro durante una primera duración de tiempo, por ejemplo, una trama de audio, para identificar el SHC 511' durante esta primera duración de tiempo. El dispositivo de codificación de audio 570 puede especificar la información de rotación y los SHC 511' en el flujo de bits 517 en cualquiera de las formas descritas anteriormente. Esta información de rotación puede denominarse información de primera rotación ya que describe la rotación del campo sonoro durante la primera duración de tiempo. El dispositivo de codificación de audio 570 puede entonces, basándose en esta primera información de rotación, hacer girar el campo sonoro durante una segunda duración de tiempo, por ejemplo, una segunda trama de audio, para identificar los SHC 511' durante esta segunda duración de tiempo. El dispositivo de codificación de audio 570 puede utilizar esta primera información de rotación al realizar la segunda operación de

rotación durante la segunda duración de tiempo para inicializar una búsqueda de la combinación "óptima" de ángulos de acimut y elevación, como un ejemplo. El dispositivo de codificación de audio 570 puede entonces especificar los SHC 511' y la información de rotación correspondiente para la segunda duración de tiempo (que se puede denominar "segunda información de rotación") en el flujo de bits 517.

[0898] Aunque se ha descrito anteriormente con respecto a varias maneras diferentes de implementar el algoritmo de rotación para reducir el tiempo de procesamiento y/o el consumo, las técnicas pueden realizarse con respecto a cualquier algoritmo que pueda reducir o de otro modo acelerar la identificación de lo que puede denominarse como la "rotación óptima". Además, las técnicas pueden realizarse con respecto a cualquier algoritmo que identifique rotaciones no óptimas pero que pueda mejorar el rendimiento en otros aspectos, a menudo medidos en términos de velocidad o procesador u otra utilización de recursos.

[0899] Las FIGS. 57-57E son cada una un diagrama que ilustra los flujos de bits 517A-517E formados de acuerdo con las técnicas descritas en esta divulgación. En el ejemplo de la FIG. 57A, el flujo de bits 517A puede representar un ejemplo del flujo de bits 517 mostrado en la FIG. 53 anterior. El flujo de bits 517A incluye un campo actual SHC 670 y un campo que almacena SHC 511' (donde el campo se denomina "SHC 511'"). El campo actual SHC 670 puede incluir un bit correspondiente a cada uno de SHC511. Los SHC 511' pueden representar los de SHC 511 que se especifican en el flujo de bits, que pueden ser menores en número que el número de los SHC 511. Típicamente, cada uno de los SHC 511' son aquellos de SHC 511 que tienen valores distintos de cero. Como se ha indicado anteriormente, para una representación de cuarto orden de cualquier campo sonoro dado, se requieren $(1+4)^2$ o 25 SHC. Eliminar uno o más de estos SHC y reemplazar estos SHC de valor cero con un único bit puede ahorrar 31 bits, los cuales pueden ser asignados para expresar otras porciones del campo sonoro con más detalle o eliminarse de otro modo para facilitar la utilización eficiente del ancho de banda.

[0900] En el ejemplo de la FIG. 57B, el flujo de bits 517B puede representar un ejemplo del flujo de bits 517 mostrado en la FIG. 53 anterior. El flujo de bits 517B incluye un campo de información de transformación 672 ("información de transformación 672") y un campo que almacena SHC 511' (donde el campo se denomina "SHC 511'"). La información de transformación 672, como se ha indicado anteriormente, puede comprender información de traducción, información de rotación y/o cualquier otra forma de información que indique un ajuste a un campo sonoro. En algunos casos, la información de transformación 672 también puede especificar un orden mayor de SHC 511 que se especifica en el flujo de bits 517B como SHC 511'. Es decir, la información de transformación 672 puede indicar un orden de tres, que el dispositivo de extracción puede entender que indica que SHC 511' incluye los de SHC 511 hasta e incluyendo los de SHC 511 que tienen un orden de tres. El dispositivo de extracción puede entonces configurarse para establecer que SHC 511 tenga un orden de cuatro o más a cero, eliminando potencialmente la señalización explícita de SHC 511 de orden cuatro o superior en el flujo de bits.

[0901] En el ejemplo de la FIG. 57C, el flujo de bits 517C puede representar un ejemplo del flujo de bits 517 mostrado en la FIG. 53 anterior. El flujo de bits 517C incluye el campo de información de transformación 672 ("información de transformación 672"), el campo actual de SHC 670 y un campo que almacena SHC 511' (donde el campo se denomina "SHC 511'"). En lugar de estar configurado para comprender qué orden de SHC 511 no se señaliza como se ha descrito anteriormente con respecto a la FIG. 57B, el campo actual SHC 670 puede indicar explícitamente cuáles de los SHC 511 se especifican en el flujo de bits 517C como SHC 511'

[0902] En el ejemplo de la FIG. 57D, el flujo de bits 517D puede representar un ejemplo del flujo de bits 517 mostrado en la FIG. 53 anterior. El flujo de bits 517D incluye un campo de orden 674 ("orden 60"), el campo actual de SHC 670, un indicador de acimut 676 ("AZF 676"), un indicador de elevación 678 ("ELF 678"), un campo de ángulo de acimut 680 ("acimut 680"), un campo de ángulo de elevación 682 ("elevación 682 ") y un campo que almacena SHC 511' (donde, de nuevo, el campo se denomina "SHC 511'"). El campo de orden 674 especifica el orden de SHC 511', es decir, el orden representado por n anteriormente para el orden más alto de la función de base esférica utilizada para representar el campo sonoro. El campo de orden 674 se muestra como un campo de 8 bits, pero puede ser de otros tamaños de bits diferentes, tal como tres (que es el número de bits requerido para especificar el cuarto orden). El campo actual de SHC 670 se muestra como un campo de 25 bits. De nuevo, sin embargo, el campo actual de SHC 670 puede ser de otros tamaños de bits diferentes. El campo actual de SHC 670 se muestra como de 25 bits para indicar que el campo presente de SHC 670 puede incluir un bit para cada uno de los coeficientes armónicos esféricos correspondientes a una representación de cuarto orden del campo sonoro.

[0903] El indicador de acimut 676 representa un indicador de un bit que especifica si el campo de acimut 680 está presente en el flujo de bits 517D. Cuando el indicador de acimut 676 se ajusta a uno, el campo de acimut 680 para SHC 511' está presente en el flujo de bits 517D. Cuando el indicador de acimut 676 se ajusta a cero, el campo de acimut 680 para SHC 511' no está presente o especificado de otro modo en el flujo de bits 517D. Del mismo modo, el indicador de elevación 678 representa un indicador de un bit que especifica si el campo de elevación 682 está presente en el flujo de bits 517D. Cuando el indicador de elevación 678 está ajustado a uno, el campo de elevación 682 para SHC 511' está presente en el flujo de bits 517D. Cuando el indicador de elevación 678 se ajusta a cero, el campo de elevación 682 para SHC 511' no está presente o especificado de otro modo en el flujo de bits 517D. Aunque se describe como una señalización de que el campo correspondiente está presente y cero indica que el campo correspondiente no está presente, la convención puede invertirse de tal manera que un cero especifique que

el campo correspondiente se especifica en el flujo de bits 517D y uno especifica que el campo correspondiente no se especifica en el flujo de bits 517D. Por lo tanto, las técnicas descritas en esta divulgación no deben limitarse en este aspecto.

[0904] El campo de acimut 680 representa un campo de 10 bits que especifica, cuando está presente en el flujo de bits 517D, el ángulo de acimut. Aunque se muestra como un campo de 10 bits, el campo de acimut 680 puede tener otros tamaños de bits. El campo de elevación 682 representa un campo de 9 bits que especifica, cuando está presente en el flujo de bits 517D, el ángulo de elevación. El ángulo de acimut y el ángulo de elevación especificados en los campos 680 y 682, respectivamente, pueden, junto con los indicadores 676 y 678, representar la información de rotación descrita anteriormente. Esta información de rotación se puede usar para girar el campo sonoro para recuperar los SHC 511 en la trama de referencia original.

[0905] El campo de SHC 511' se muestra como un campo variable que es de tamaño X. El campo de SHC 511' puede variar debido al número de SHC 511' especificado en el flujo de bits como se indica por el campo actual de SHC 670. El tamaño X puede derivarse en función del número de unidades en el campo actual de SHC 670 en veces de 32 bits (que es el tamaño de cada SHC 511').

[0906] En el ejemplo de la FIG. 57E, el flujo de bits 517E puede representar otro ejemplo del flujo de bits 517 mostrado en la FIG. 53 anterior. El flujo de bits 517E incluye un campo de orden 674 ("orden 60"), un campo actual de SHC 670 y un campo de índice de rotación 684, y un campo que almacena SHC 511' (donde, de nuevo, el campo se denomina "SHC 511"). El campo de orden 674, el campo actual de SHC 670 y el campo de SHC 511' pueden ser sustancialmente similares a los descritos anteriormente. El campo de índice de rotación 684 puede representar un campo de 20 bits utilizado para especificar una de las combinaciones de 1024 x 512 (o, en otras palabras, 524288) de los ángulos de elevación y acimut. En algunos casos, sólo se pueden utilizar 19 bits para especificar este campo de índice de rotación 684, y el dispositivo de codificación de audio 570 puede especificar un indicador adicional en el flujo de bits para indicar si se ha realizado una operación de rotación (y, por tanto, si el campo de índice de rotación 684 está presente en el flujo de bits). Este campo de índice de rotación 684 especifica el índice de rotación indicado anteriormente, que puede referirse a una entrada en una tabla de rotación común tanto para el dispositivo de codificación de audio 570 como para el dispositivo de extracción de flujo de bits. Esta tabla de rotación puede, en algunos casos, almacenar las diferentes combinaciones de ángulos de acimut y elevación. Alternativamente, la tabla de rotación puede almacenar la matriz descrita anteriormente, que almacena eficazmente las diferentes combinaciones de los ángulos de acimut y elevación en forma de matriz.

[0907] La FIG. 58 es un diagrama de flujo que ilustra el ejemplo de funcionamiento del dispositivo de codificación de audio 570 mostrado en el ejemplo de la FIG. 53 en la implementación de los aspectos de rotación de las técnicas descritas en esta divulgación. Inicialmente, el dispositivo de codificación de audio 570 puede seleccionar una combinación de ángulo de acimut y ángulo de elevación de acuerdo con uno o más de los diversos algoritmos de rotación descritos anteriormente (800). El dispositivo de codificación de audio 570 puede entonces girar el campo sonoro de acuerdo con el ángulo de acimut y elevación seleccionado (802). Como se ha descrito anteriormente, el dispositivo de codificación de audio 570 puede derivar primero el campo sonoro de SHC 511 usando $InvMat_1$ indicado anteriormente. El dispositivo de codificación de audio 570 también puede determinar SHC511' que representan el campo sonoro girado (804). Aunque se describen como etapas u operaciones separadas, el dispositivo de codificación de audio 570 puede aplicar una transformación (que puede representar el resultado de $[EncMat_2][InvMat_1]$) que representa la selección de la combinación del ángulo de acimut y de ángulo de elevación, derivando el campo sonoro de los SHC 511, girando el campo sonoro y determinando los SHC 511' que representan el campo sonoro girado.

[0908] En cualquier caso, el dispositivo de codificación de audio 570 puede entonces calcular un número de los SHC 511' determinados que sean mayores que un valor umbral, comparando este número con un número calculado para una iteración previa con respecto a una combinación anterior de ángulo de acimut y ángulo de elevación (806, 808). En la primera iteración con respecto a la primera combinación de ángulo de acimut y ángulo de elevación, esta comparación puede ser con respecto a un número predefinido anterior (que puede ajustarse a cero). En cualquier caso, si el número determinado de los SHC 511' es menor que el número anterior ("SÍ" 808), el dispositivo de codificación de audio 570 almacena los SHC 511', el ángulo de acimut y el ángulo de elevación, reemplazando a menudo los anteriores SHC511', ángulo de acimut y ángulo de elevación almacenados a partir de una iteración anterior del algoritmo de rotación (810).

[0909] Si el número determinado de los SHC 511' no es menor que el número anterior ("NO" 808) o después de almacenar los SHC 511', el ángulo de acimut y el ángulo de elevación en lugar de los SHC 511', el ángulo de acimut y el ángulo de elevación previamente almacenados, el dispositivo de codificación de audio 570 puede determinar si el algoritmo de rotación ha terminado (812). Es decir, el dispositivo de codificación de audio 570 puede, como un ejemplo, determinar si se han evaluado todas las combinaciones disponibles de ángulo de acimut y ángulo de elevación. En otros ejemplos, el dispositivo de codificación de audio 570 puede determinar si se cumplen otros criterios (tales como que se ha realizado todo un subconjunto definido de combinación, si se ha recorrido una trayectoria dada, si se ha recorrido un árbol jerárquico a un nodo de hoja, etc.) de tal forma que el dispositivo de codificación de audio 570 haya terminado de realizar el algoritmo de rotación. Si no se ha terminado ("NO" 812), el

dispositivo de codificación de audio 570 puede realizar el proceso anterior con respecto a otra combinación seleccionada (800-812). Si se ha terminado ("SÍ" 812), el dispositivo de codificación de audio 570 puede especificar los SHC 511', el ángulo de ángulo y el ángulo elevación almacenados en el flujo de bits 517 en una de las diversas formas descritas anteriormente (814).

[0910] La FIG. 59 es un diagrama de flujo que ilustra el ejemplo de funcionamiento del dispositivo de codificación de audio 570 mostrado en el ejemplo de la FIG. 53 en la realización de los aspectos de transformación de las técnicas descritas en esta divulgación. Inicialmente, el dispositivo de codificación de audio 570 puede seleccionar una matriz que represente una transformación lineal reversible (820). Un ejemplo de una matriz que representa una transformación lineal reversible puede ser la matriz mostrada anteriormente que es el resultado de $[EncMat_1][IncMat_1]$. El dispositivo de codificación de audio 570 puede entonces aplicar la matriz al campo sonoro para transformar el campo sonoro (822). El dispositivo de codificación de audio 570 también puede determinar SHC 511' que representan el campo sonoro girado (824). Aunque se describen como etapas u operaciones separadas, el dispositivo de codificación de audio 570 puede aplicar una transformación (que puede representar el resultado de $[EncMat_2][InvMat_1]$), derivando el campo sonoro de los SHC 511, transformar el campo sonoro y determinar los SHC 511' que representan el campo sonoro de transformación.

[0911] En cualquier caso, el dispositivo de codificación de audio 570 puede entonces calcular un número de los SHC 511' determinados que sean mayores que un valor umbral, comparando este número con un número calculado para una iteración previa con respecto a una aplicación de una matriz de transformación (826, 828). Si el número determinado de los SHC 511' es menor que el número anterior ("SÍ" 828), el dispositivo de codificación de audio 570 almacena los SHC 511' y la matriz (o algún derivado de la misma, tal como un índice asociado con la matriz), reemplazando a menudo los SHC 511' anteriores y la matriz (o un derivado de la misma) almacenados de una iteración anterior del algoritmo de rotación (830).

[0912] Si el número determinado de los SHC 511' no es menor que el número anterior ("NO" 828) o después de almacenar los SHC 511' y la matriz en lugar de los SHC 511' y la matriz previamente almacenados, el dispositivo de codificación de audio 570 puede determinar si el algoritmo de transformación ha terminado (832). Es decir, el dispositivo de codificación de audio 570 puede, como un ejemplo, determinar si se han evaluado todas las matrices de transformación. En otros ejemplos, el dispositivo de codificación de audio 570 puede determinar si se cumplen otros criterios (tales como que se ha realizado todo un subconjunto definido de las matrices de transformación disponibles, si se ha recorrido una trayectoria dada, si se ha recorrido un árbol jerárquico a un nodo de hoja, etc.) de tal forma que el dispositivo de codificación de audio 570 haya terminado de realizar el algoritmo de transformación. Si no se ha terminado ("NO" 832), el dispositivo de codificación de audio 570 puede realizar el proceso anterior con respecto a otra matriz de transformación seleccionada (820-832). Si se ha terminado ("SÍ" 832), el dispositivo de codificación de audio 570 puede especificar los SHC 511' y la matriz almacenados en el flujo de bits 517 en una de las diversas formas descritas anteriormente (834).

[0913] En algunos ejemplos, el algoritmo de transformación puede realizar una sola iteración, evaluando una sola matriz de transformación. Es decir, la matriz transformada puede comprender cualquier matriz que represente una transformación lineal reversible. En algunos casos, la transformación lineal reversible puede transformar el campo sonoro del dominio espacial en el dominio de frecuencia. Los ejemplos de tal transformación lineal reversible pueden incluir una transformada discreta de Fourier (DFT). La aplicación de la DFT sólo puede implicar una única iteración y, por lo tanto, no necesariamente incluirá pasos para determinar si el algoritmo de transformación ha terminado. Por consiguiente, las técnicas no deben limitarse al ejemplo de la FIG. 59.

[0914] En otras palabras, un ejemplo de una transformación lineal reversible es una transformada discreta de Fourier (DFT). Los veinticinco SHC 511' pueden operarse por la DFT para formar un conjunto de veinticinco coeficientes complejos. El dispositivo de codificación de audio 570 también puede recibir una relleno nulo. Los veinticinco SHC 511' son un múltiplo entero de 2, con el fin de aumentar potencialmente la resolución del tamaño de intervalo de la DFT, y potencialmente tienen una implementación más eficiente de la DFT, por ejemplo, aplicando una transformada de Fourier rápida (FFT). En algunos casos, el aumento de la resolución de la DFT más allá de 25 puntos no se requiere necesariamente. En el dominio de transformación, el dispositivo de codificación de audio 570 puede aplicar un umbral para determinar si hay alguna energía espectral en un intervalo particular. El dispositivo de codificación de audio 570 puede, en este contexto, descartar o anular energía de coeficiente espectral que está por debajo de este umbral, y el dispositivo de codificación de audio 570 puede aplicar una transformada inversa para recuperar SHC 511' que tiene uno o más de los SHC 511' descartados o anulados. Es decir, después de aplicar la transformada inversa, los coeficientes por debajo del umbral no están presentes, y como resultado, se pueden usar menos bits para codificar el campo sonoro.

[0915] En uno o más ejemplos, las funciones descritas pueden implementarse en hardware, software, firmware o cualquier combinación de estos. Si se implementan en software, las funciones pueden almacenarse en un medio legible por ordenador, o transmitirse a este, como una o más instrucciones o código, 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 unos medios de almacenamiento de datos, o medios de comunicación que incluyen cualquier medio que facilite la transferencia de

un programa informático de un lugar a otro, por ejemplo, de acuerdo con un protocolo de comunicación. De esta manera, los medios legibles por ordenador pueden corresponder, en general, a (1) unos medios de almacenamiento tangibles y legibles por ordenador que sean no transitorios o (2) un medio de comunicación tal como una señal u onda portadora. Los medios de almacenamiento de datos pueden ser cualquier medio disponible al que se puede acceder desde uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para implementar las técnicas descritas en la presente divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

[0916] A modo de ejemplo, y no de manera limitativa, dichos 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 utilizarse para almacenar código de programa deseado en forma de instrucciones o estructuras de datos y al que pueda accederse mediante un ordenador. Además, cualquier conexión puede denominarse de forma correcta con el término medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde un sitio web, un servidor u otra fuente remota mediante un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio. 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, como se usan en el presente documento, incluyen el disco compacto (CD), el disco de láser, el disco óptico, el disco versátil digital (DVD), el disco flexible y el disco Blu-ray, donde algunos discos normalmente reproducen datos de manera magnética, mientras que otros discos reproducen los datos de manera óptica con láser. Las combinaciones de lo anterior también deberían incluirse dentro del alcance de los medios legibles por ordenador.

[0917] 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), matrices lógicas programables in situ (FPGA) u otros circuitos lógicos integrados o discretos equivalentes. Por consiguiente, el término «procesador», como se utiliza en el presente documento, puede referirse a cualquier estructura anterior 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 específicos 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.

[0918] Las técnicas de la presente divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, que incluyen un teléfono inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). En la presente divulgación se describen varios componentes, módulos o unidades para enfatizar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no requieren necesariamente su realización mediante diferentes unidades de hardware. En su lugar, 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 adecuado.

[0919] Se han descrito diversas realizaciones de las técnicas. La invención está definida por las reivindicaciones siguientes.

REIVINDICACIONES

1. Un procedimiento, que comprende:

obtener un flujo de bits (21) que comprende una versión comprimida de un componente espacial de un campo sonoro (57), el componente espacial definido en un dominio armónico esférico y generado realizando una descomposición lineal con respecto a una pluralidad de coeficientes armónicos esféricos,
en el que la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un campo (262) que especifica un modo de predicción (237) utilizado al comprimir el componente espacial.

2. El procedimiento de la reivindicación 1, en el que la versión comprimida del componente espacial está representada adicionalmente en el flujo de bits utilizando, al menos en parte, uno de:

información de tabla Huffman que especifica una tabla Huffman utilizada al comprimir el componente espacial;
un código Huffman para representar un identificador de categoría que identifica una categoría de compresión a la que corresponde el componente espacial;
un bit de signo que identifica si el componente espacial es un valor positivo o un valor negativo; y
un código Huffman para representar un valor residual del componente espacial.

3. El procedimiento de la reivindicación 1, en el que la descomposición comprende una descomposición en valores singulares.

4. Un dispositivo (16, 24) que comprende uno o más procesadores configurados para:

obtener un flujo de bits (21) que comprende una versión comprimida de un componente espacial de un campo sonoro (57), el componente espacial definido en un dominio armónico esférico y generado realizando una descomposición lineal con respecto a una pluralidad de coeficientes armónicos esféricos,
en el que la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un campo (262) que especifica un modo de predicción (237) utilizado al comprimir el componente espacial.

5. El dispositivo de la reivindicación 4, en el que la versión comprimida del componente espacial se representa además en el flujo de bits utilizando, al menos en parte, información de tabla Huffman que especifica una tabla Huffman usada al comprimir el componente espacial.

6. El dispositivo de la reivindicación 4, en el que la versión comprimida del componente espacial se representa además en el flujo de bits utilizando, al menos en parte, un código Huffman para representar un identificador de categoría que identifica una categoría de compresión a la que corresponde el componente espacial.

7. El dispositivo de la reivindicación 4, en el que la versión comprimida del componente espacial se representa además en el flujo de bits utilizando, al menos en parte, un bit de signo que identifica si el componente espacial es un valor positivo o un valor negativo.

8. El dispositivo de la reivindicación 4, en el que la versión comprimida del componente espacial se representa además en el flujo de bits utilizando, al menos en parte, un código Huffman para representar un valor residual del componente espacial.

9. El dispositivo de la reivindicación 4, en el que la descomposición comprende una descomposición en valores singulares.

10. Un procedimiento, que comprende:

generar un flujo de bits (21) que comprende una versión comprimida de un componente espacial de un campo sonoro (57), el componente espacial definido en un dominio armónico esférico y generado realizando una descomposición lineal con respecto a una pluralidad de coeficientes armónicos esféricos,
en el que la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un campo (262) que especifica un modo de predicción (237) utilizado al comprimir el componente espacial.

11. El procedimiento de la reivindicación 10, en el que generar el flujo de bits comprende generar el flujo de bits para incluir uno de:

información de tabla Huffman que especifica una tabla Huffman utilizada al comprimir el componente espacial;

un código Huffman para representar un identificador de categoría que identifica una categoría de compresión a la que corresponde el componente espacial;
un bit de signo que identifica si el componente espacial es un valor positivo o un valor negativo; y
un código Huffman para representar un valor residual del componente espacial.

12. Un dispositivo (16, 24) que comprende uno o más procesadores configurados para:

generar un flujo de bits (21) que comprende una versión comprimida de un componente espacial de un campo sonoro (57), el componente espacial definido en un dominio armónico esférico y generado realizando una descomposición lineal con respecto a una pluralidad de coeficientes armónicos esféricos,
en el que la versión comprimida del componente espacial se representa en el flujo de bits utilizando, al menos en parte, un campo (262) que especifica un modo de predicción (237) utilizado al comprimir el componente espacial.

13. El dispositivo de la reivindicación 12, en el que el uno o más procesadores están configurados para generar el flujo de bits para incluir uno de:

información de tabla Huffman que especifica una tabla Huffman utilizada al comprimir el componente espacial;
un código Huffman para representar un identificador de categoría que identifica una categoría de compresión a la que corresponde el componente espacial;
un bit de signo que identifica si el componente espacial es un valor positivo o un valor negativo; y
un código Huffman para representar un valor residual del componente espacial.

14. El dispositivo de la reivindicación 12, en el que la descomposición comprende una descomposición en valores singulares.

15. Un medio de almacenamiento no transitorio legible por ordenador que comprende instrucciones que, cuando se ejecutan, hacen que uno o más procesadores realicen cualquiera de las reivindicaciones 1 a 3, 10 y 11.

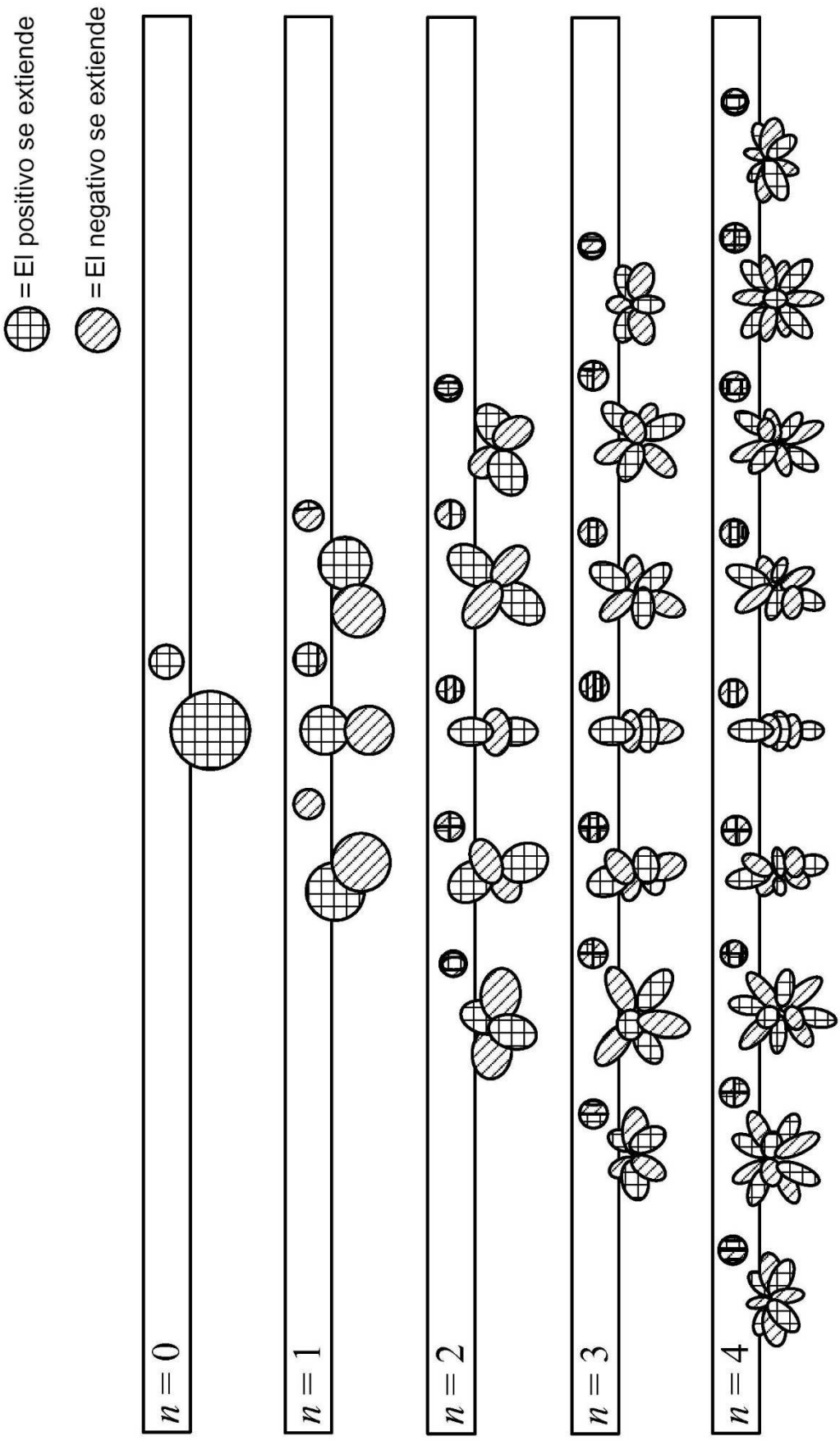

= El positivo se extiende
= El negativo se extiende
n = 0
n = 1
n = 2
n = 3
n = 4

FIG. 1

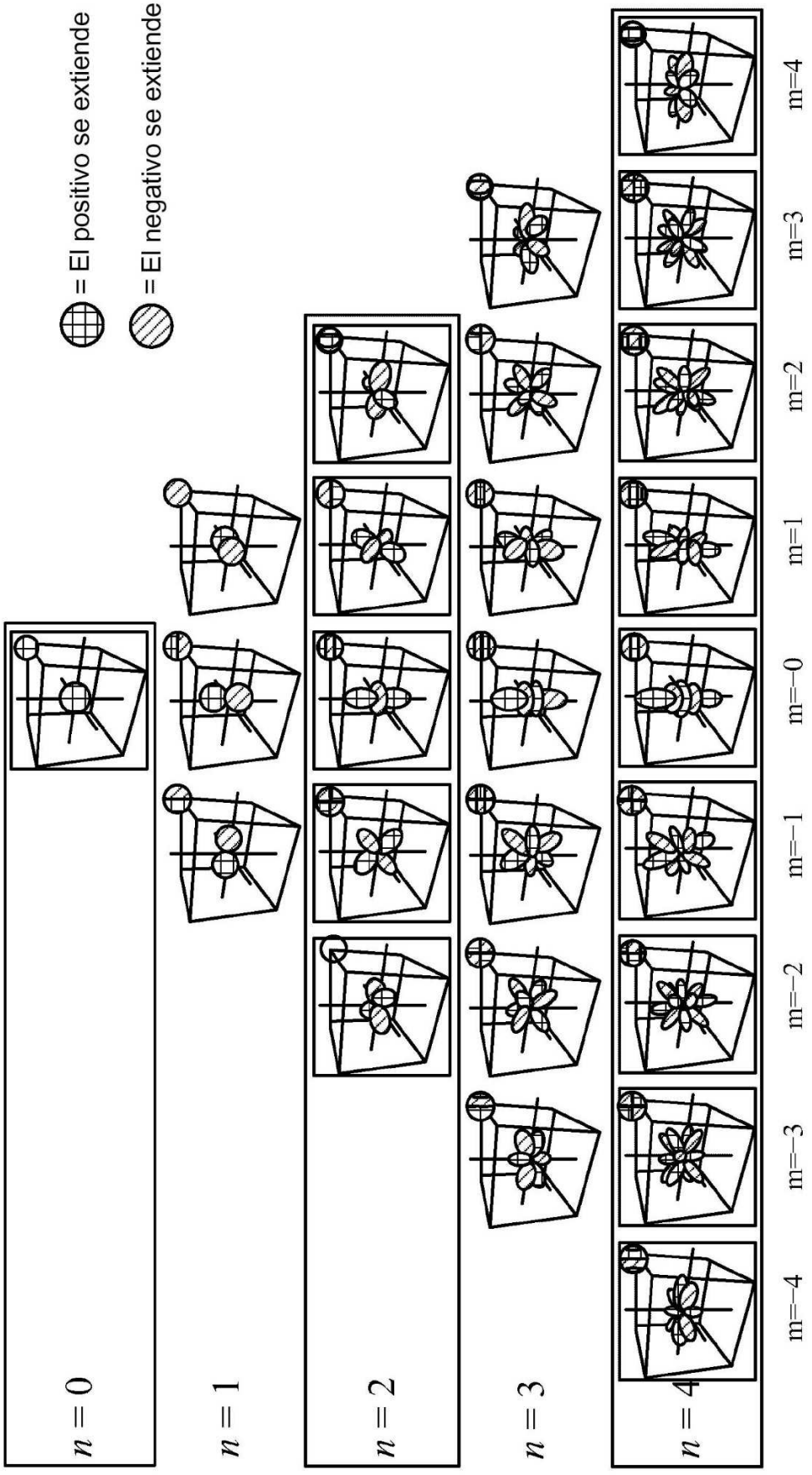

= El positivo se extiende
= El negativo se extiende
n = 0
n = 1
n = 2
n = 3
n = 4
m=−4
m=−3
m=−2
m=−1
m=−0
m=1
m=2
m=3
m=4

FIG. 2

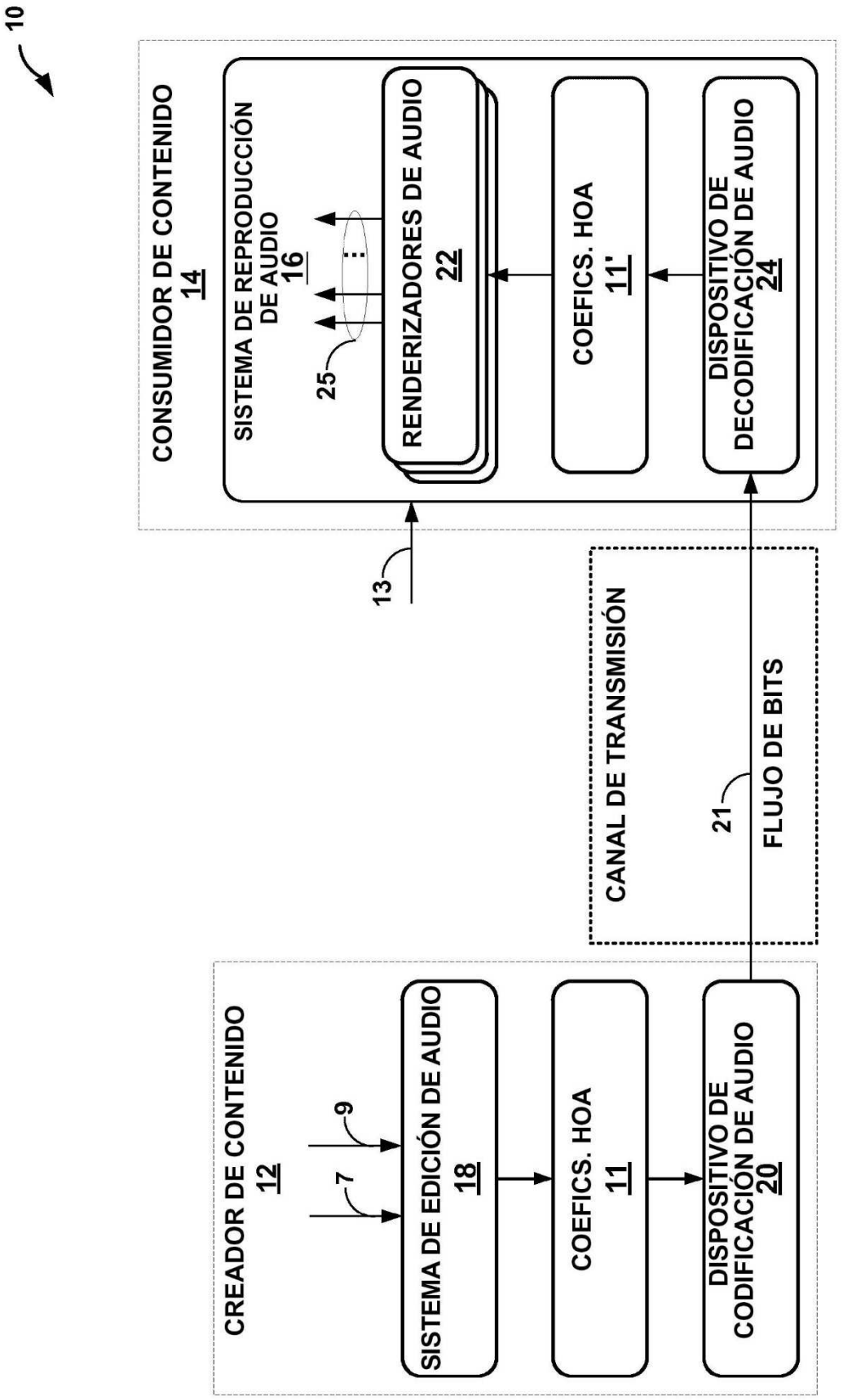

10
CONSUMIDOR DE CONTENIDO
14
SISTEMA DE REPRODUCCIÓN DE AUDIO
16
RENDERIZADORES DE AUDIO
22
25
COEFICS. HOA
11'
DISPOSITIVO DE DECODIFICACIÓN DE AUDIO
24
13
CANAL DE TRANSMISIÓN
21
FLUJO DE BITS
CREADOR DE CONTENIDO
12
9
7
SISTEMA DE EDICIÓN DE AUDIO
18
COEFICS. HOA
11
DISPOSITIVO DE CODIFICACIÓN DE AUDIO
20

FIG. 3

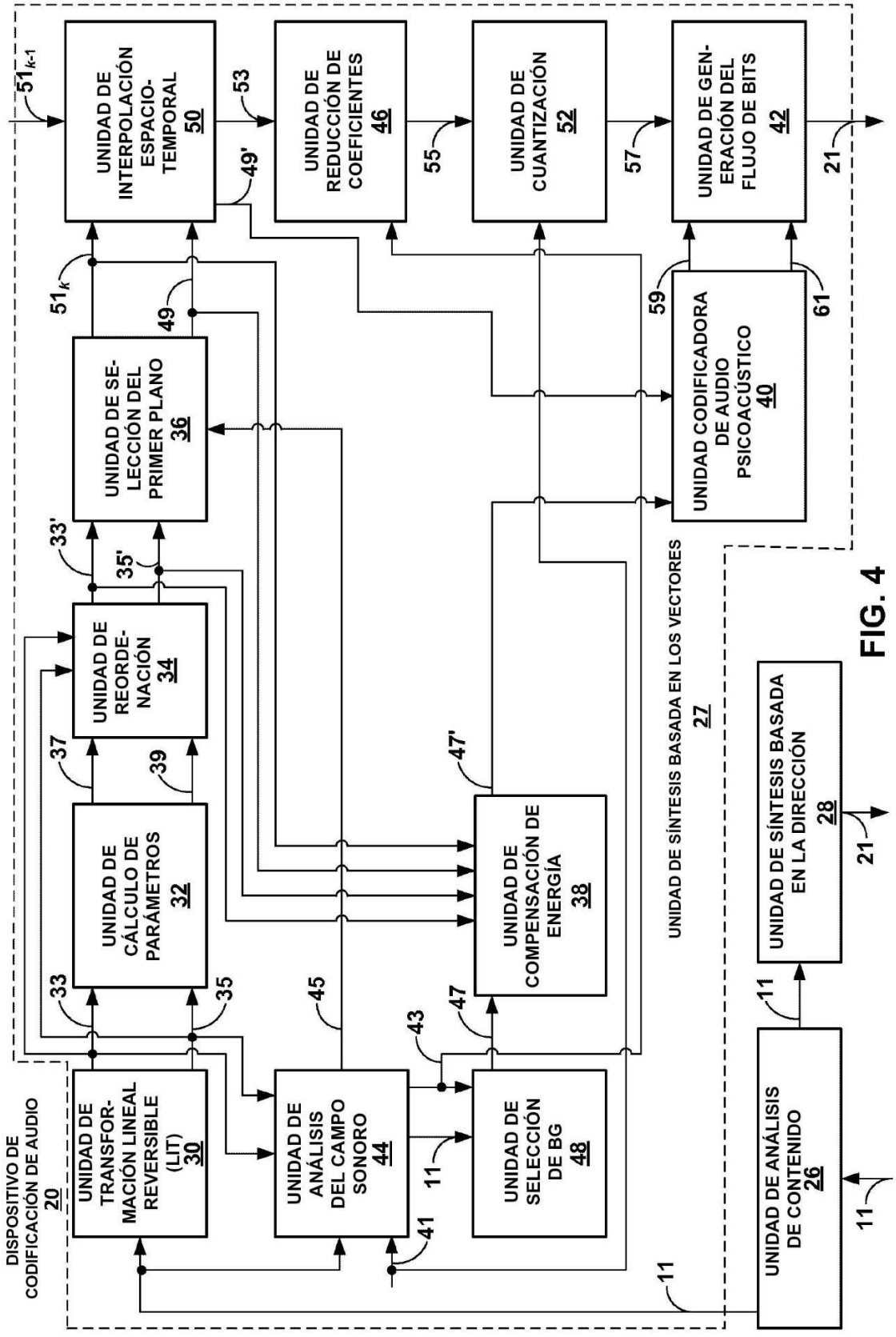

DISPOSITIVO DE CODIFICACIÓN DE AUDIO 20
UNIDAD DE TRANSFORMACIÓN LINEAL REVERSIBLE (LIT) 30
UNIDAD DE CÁLCULO DE PARÁMETROS 32
UNIDAD DE REORDENACIÓN 34
UNIDAD DE SELECCIÓN DEL PRIMER PLANO 36
UNIDAD DE INTERPOLACIÓN ESPACIO-TEMPORAL 50
UNIDAD DE REDUCCIÓN DE COEFICIENTES 46
UNIDAD DE CUANTIZACIÓN 52
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS 42
UNIDAD CODIFICADORA DE AUDIO PSICOACÚSTICO 40
UNIDAD DE ANÁLISIS DEL CAMPO SONORO 44
UNIDAD DE SELECCIÓN DE BG 48
UNIDAD DE COMPENSACIÓN DE ENERGÍA 38
UNIDAD DE SÍNTESIS BASADA EN LOS VECTORES 27
UNIDAD DE SÍNTESIS BASADA EN LA DIRECCIÓN 28
UNIDAD DE ANÁLISIS DE CONTENIDO 26
51_{k-1}
51_k
53
49'
49
55
57
59
61
21
33'
35'
33
35
37
39
45
43
47
47'
41
11

FIG. 4

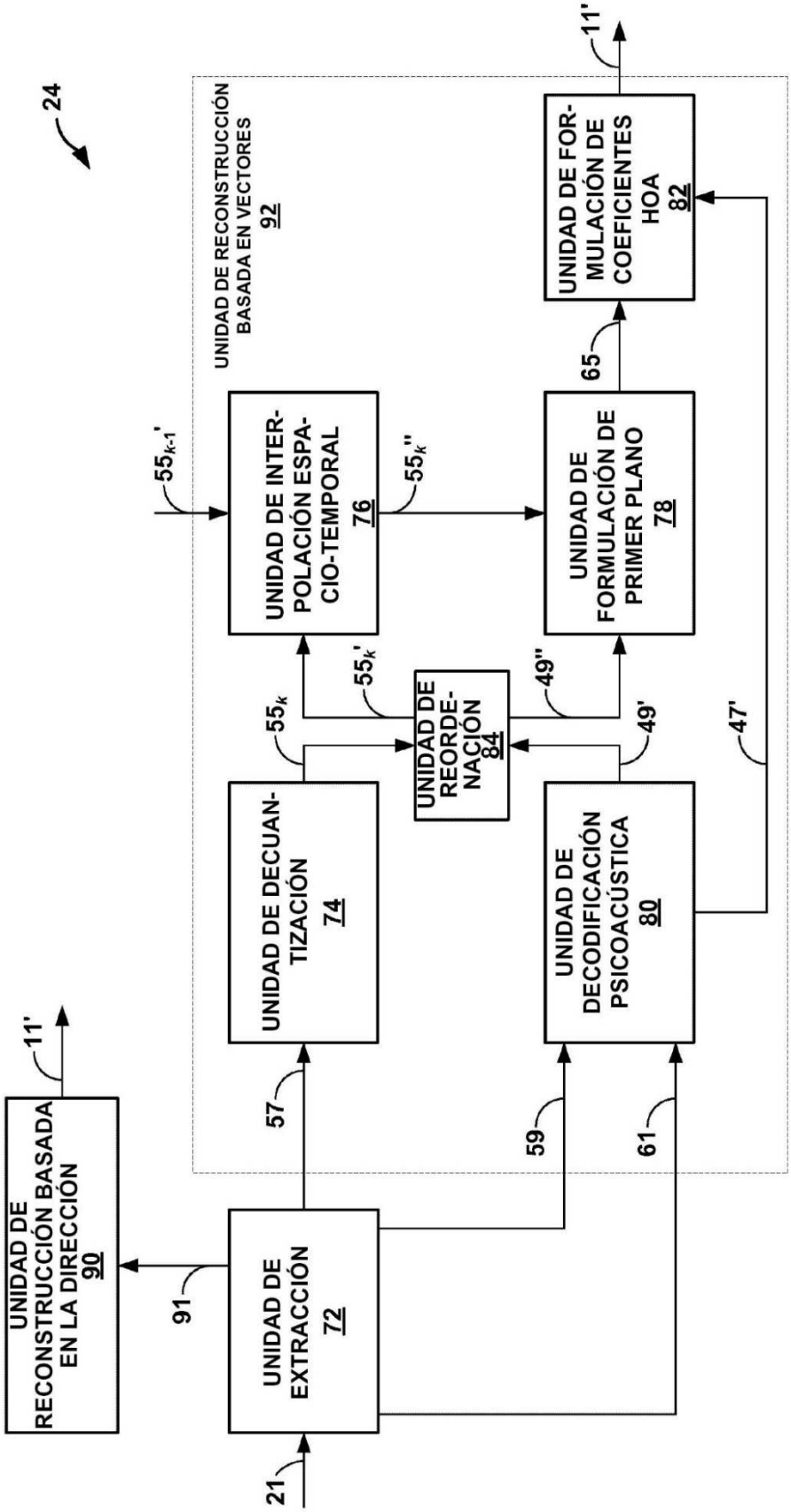

24
UNIDAD DE RECONSTRUCCIÓN BASADA EN VECTORES 92
UNIDAD DE RECONSTRUCCIÓN BASADA EN LA DIRECCIÓN 90
UNIDAD DE EXTRACCIÓN 72
UNIDAD DE DECUANTIZACIÓN 74
UNIDAD DE INTERPOLACIÓN ESPACIO-TEMPORAL 76
UNIDAD DE REORDENACIÓN 84
UNIDAD DE DECODIFICACIÓN PSICOACÚSTICA 80
UNIDAD DE FORMULACIÓN DE PRIMER PLANO 78
UNIDAD DE FORMULACIÓN DE COEFICIENTES HOA 82
11'
21
57
59
61
91
65
47'
49'
49''
55k
55k'
55k''
55k-1'

FIG. 5

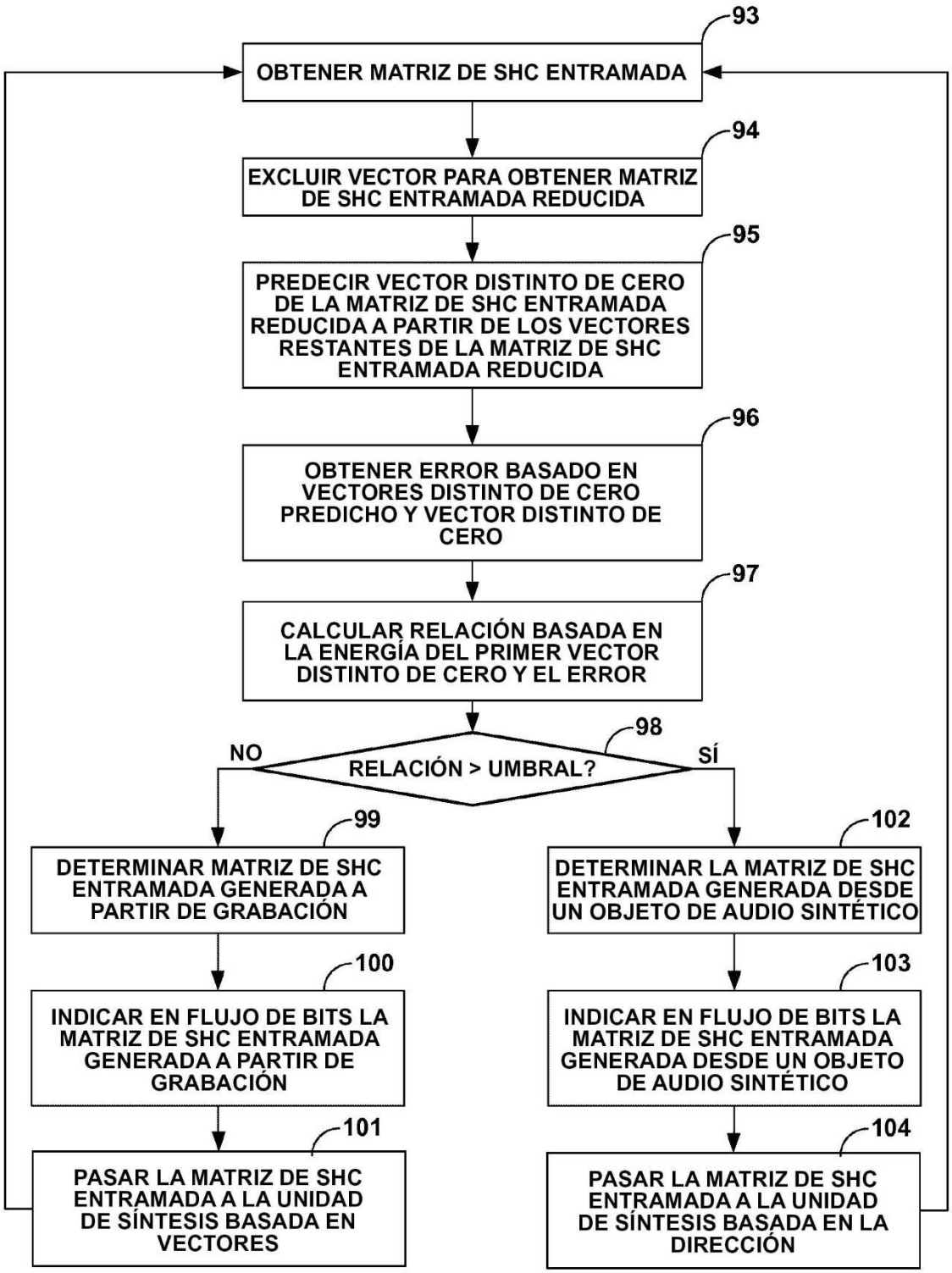

93
OBTENER MATRIZ DE SHC ENTRAMADA
94
EXCLUIR VECTOR PARA OBTENER MATRIZ DE SHC ENTRAMADA REDUCIDA
95
PREDECIR VECTOR DISTINTO DE CERO DE LA MATRIZ DE SHC ENTRAMADA REDUCIDA A PARTIR DE LOS VECTORES RESTANTES DE LA MATRIZ DE SHC ENTRAMADA REDUCIDA
96
OBTENER ERROR BASADO EN VECTORES DISTINTO DE CERO PREDICHO Y VECTOR DISTINTO DE CERO
97
CALCULAR RELACIÓN BASADA EN LA ENERGÍA DEL PRIMER VECTOR DISTINTO DE CERO Y EL ERROR
98
RELACIÓN > UMBRAL?
NO
SÍ
99
DETERMINAR MATRIZ DE SHC ENTRAMADA GENERADA A PARTIR DE GRABACIÓN
100
INDICAR EN FLUJO DE BITS LA MATRIZ DE SHC ENTRAMADA GENERADA A PARTIR DE GRABACIÓN
101
PASAR LA MATRIZ DE SHC ENTRAMADA A LA UNIDAD DE SÍNTESIS BASADA EN VECTORES
102
DETERMINAR LA MATRIZ DE SHC ENTRAMADA GENERADA DESDE UN OBJETO DE AUDIO SINTÉTICO
103
INDICAR EN FLUJO DE BITS LA MATRIZ DE SHC ENTRAMADA GENERADA DESDE UN OBJETO DE AUDIO SINTÉTICO
104
PASAR LA MATRIZ DE SHC ENTRAMADA A LA UNIDAD DE SÍNTESIS BASADA EN LA DIRECCIÓN

FIG. 6


RECIBIR COEFICIENTES (COEFICS) HOA
106
APLICAR LIT CON RESPECTO A LOS COEFICS. HOA PARA EMITIR LOS COEFICS. HOA TRANSFORMADOS
107
DETERMINAR PARÁMETRO BASADO EN EL ANÁLISIS DE COEFICS. HOA TRANSFORMADOS (TRANSF)
108
REORDENAR COEFICS. HOA TRANSF BASADOS EN PARÁMETROS PARA GENERAR COEFICS. HOA TRANSF REORDENADOS
109
REALIZAR EL ANÁLISIS DEL CAMPO SONORO CON RESPECTO A LOS COEFICS. HOA TRANSF Y ORIGINALES PARA DETERMINAR nFG Y LA INFORMACIÓN DE CANAL DE FONDO (BCI)
110
DETERMINAR LOS COEFICS. HOA AMBIENTALES BASADOS EN EL BCI
112
SELECCIONAR LOS COEFICS. HOA DE PRIMER PLANO TRANSFORMADOS BASADOS EN nFG
113
REALIZAR UNA COMPENSACIÓN DE ENERGÍA CON RESPECTO A LOS COEFICS. HOA AMBIENTALES PARA GENERAR LOS COEFICS. HOA AMBIENTALES COMPENSADOS DE ENERGÍA
114
REALIZAR LA INTERPOLACIÓN CON RESPECTO A LOS COEFICS. HOA TRANSF PARA OBTENER LAS SEÑALES nFG INTERPOLADAS Y LA INFORMACIÓN DIRECCIONAL DE PRIMER PLANO RESTANTE
116
REALIZAR LA REDUCCIÓN DE COEFICIENTES CON RESPECTO A LA INFO DIRECCIONAL DE PRIMER PLANO RESTANTE BASADA EN BCI PARA OBTENER INFORMACIÓN DIRECCIONAL DE PRIMER PLANO REDUCIDA
118
COMPRIMIR INFO DIRECCIONAL DE PRIMER PLANO REDUCIDA Y GENERAR INFORMACIÓN DIRECCIONAL DE PRIMER PLANO CODIFICADA
120
CODIFICACIÓN PSICOACÚSTICA DE LOS COEFICS. HOA AMBIENTALES COMPENSADOS DE ENERGÍA Y LAS SEÑALES DE PRIMER PLANO INTERPOLADAS PARA GENERAR COEFICS. HOA AMBIENTALES CODIFICADOS Y SEÑALES DE PRIMER PLANO CODIFICADAS
122
GENERAR FLUJO DE BITS BASADO EN INFORMACIÓN DIRECCIONAL DE PRIMER PLANO CODIFICADA, COEFICS. HOA AMBIENTALES CODIFICADOS, SEÑALES DE PRIMER PLANO CODIFICADAS Y BCI
124

FIG. 7

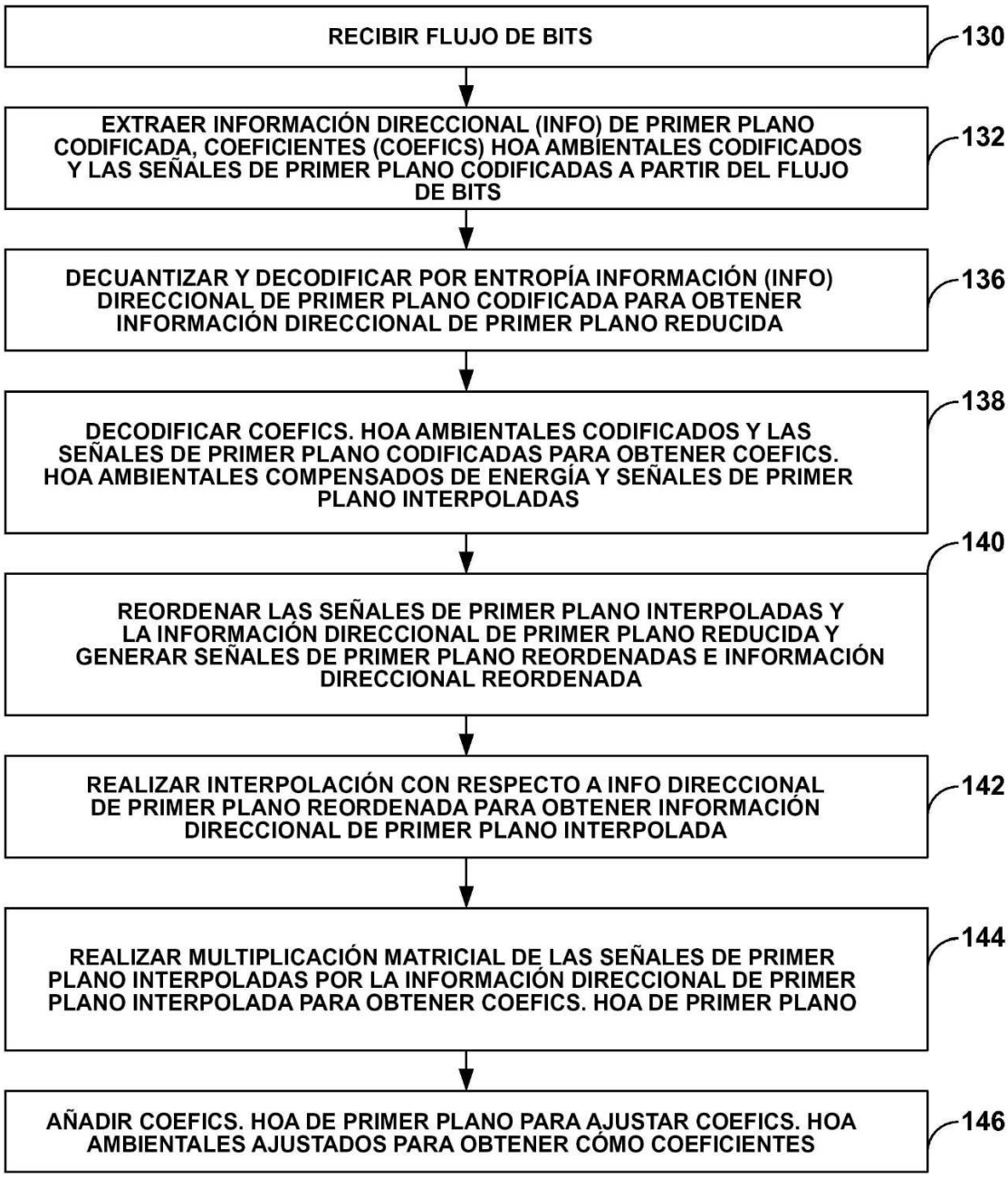

RECIBIR FLUJO DE BITS
130
EXTRAER INFORMACIÓN DIRECCIONAL (INFO) DE PRIMER PLANO CODIFICADA, COEFICIENTES (COEFICS) HOA AMBIENTALES CODIFICADOS Y LAS SEÑALES DE PRIMER PLANO CODIFICADAS A PARTIR DEL FLUJO DE BITS
132
DECUANTIZAR Y DECODIFICAR POR ENTROPÍA INFORMACIÓN (INFO) DIRECCIONAL DE PRIMER PLANO CODIFICADA PARA OBTENER INFORMACIÓN DIRECCIONAL DE PRIMER PLANO REDUCIDA
136
DECODIFICAR COEFICS. HOA AMBIENTALES CODIFICADOS Y LAS SEÑALES DE PRIMER PLANO CODIFICADAS PARA OBTENER COEFICS. HOA AMBIENTALES COMPENSADOS DE ENERGÍA Y SEÑALES DE PRIMER PLANO INTERPOLADAS
138
REORDENAR LAS SEÑALES DE PRIMER PLANO INTERPOLADAS Y LA INFORMACIÓN DIRECCIONAL DE PRIMER PLANO REDUCIDA Y GENERAR SEÑALES DE PRIMER PLANO REORDENADAS E INFORMACIÓN DIRECCIONAL REORDENADA
140
REALIZAR INTERPOLACIÓN CON RESPECTO A INFO DIRECCIONAL DE PRIMER PLANO REORDENADA PARA OBTENER INFORMACIÓN DIRECCIONAL DE PRIMER PLANO INTERPOLADA
142
REALIZAR MULTIPLICACIÓN MATRICIAL DE LAS SEÑALES DE PRIMER PLANO INTERPOLADAS POR LA INFORMACIÓN DIRECCIONAL DE PRIMER PLANO INTERPOLADA PARA OBTENER COEFICS. HOA DE PRIMER PLANO
144
AÑADIR COEFICS. HOA DE PRIMER PLANO PARA AJUSTAR COEFICS. HOA AMBIENTALES AJUSTADOS PARA OBTENER CÓMO COEFICIENTES
146

FIG. 8

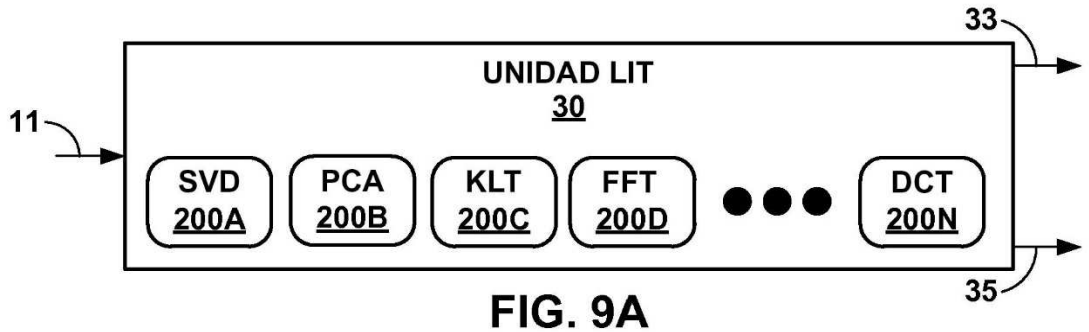


FIG. 9A

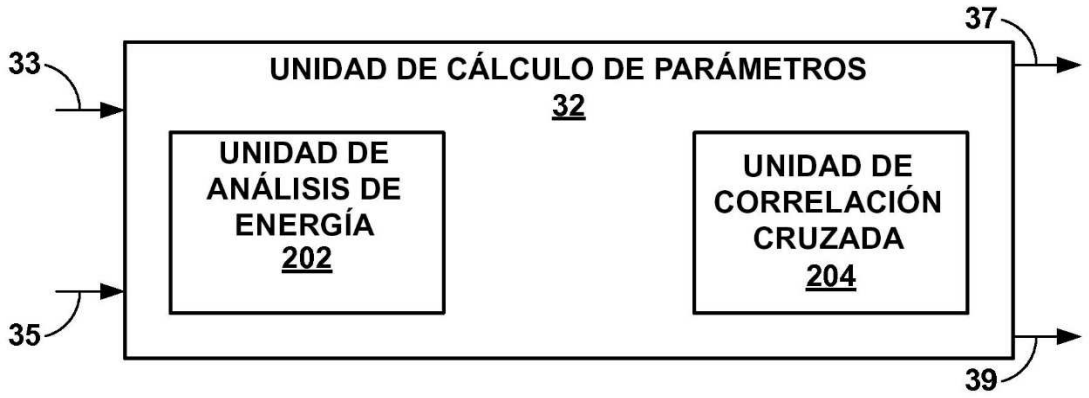


FIG. 9B

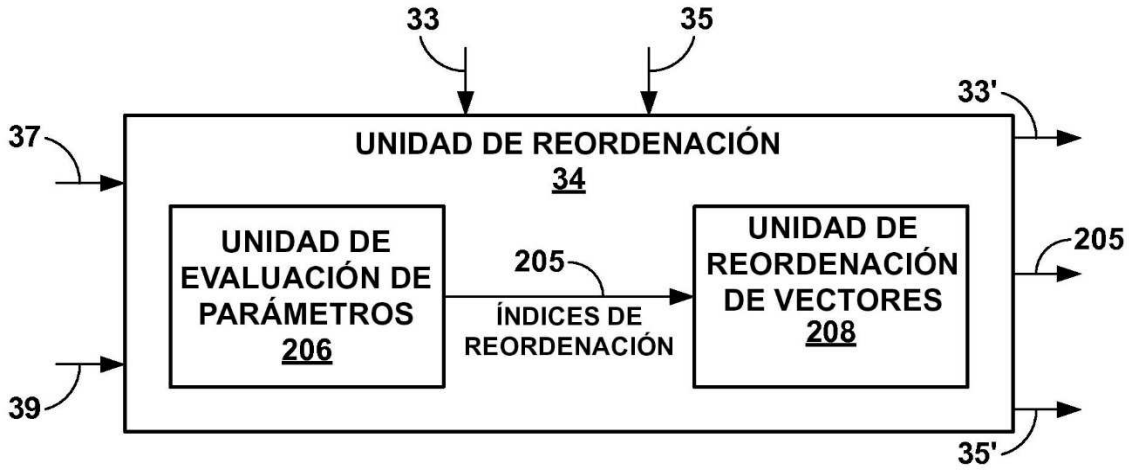


FIG. 9C

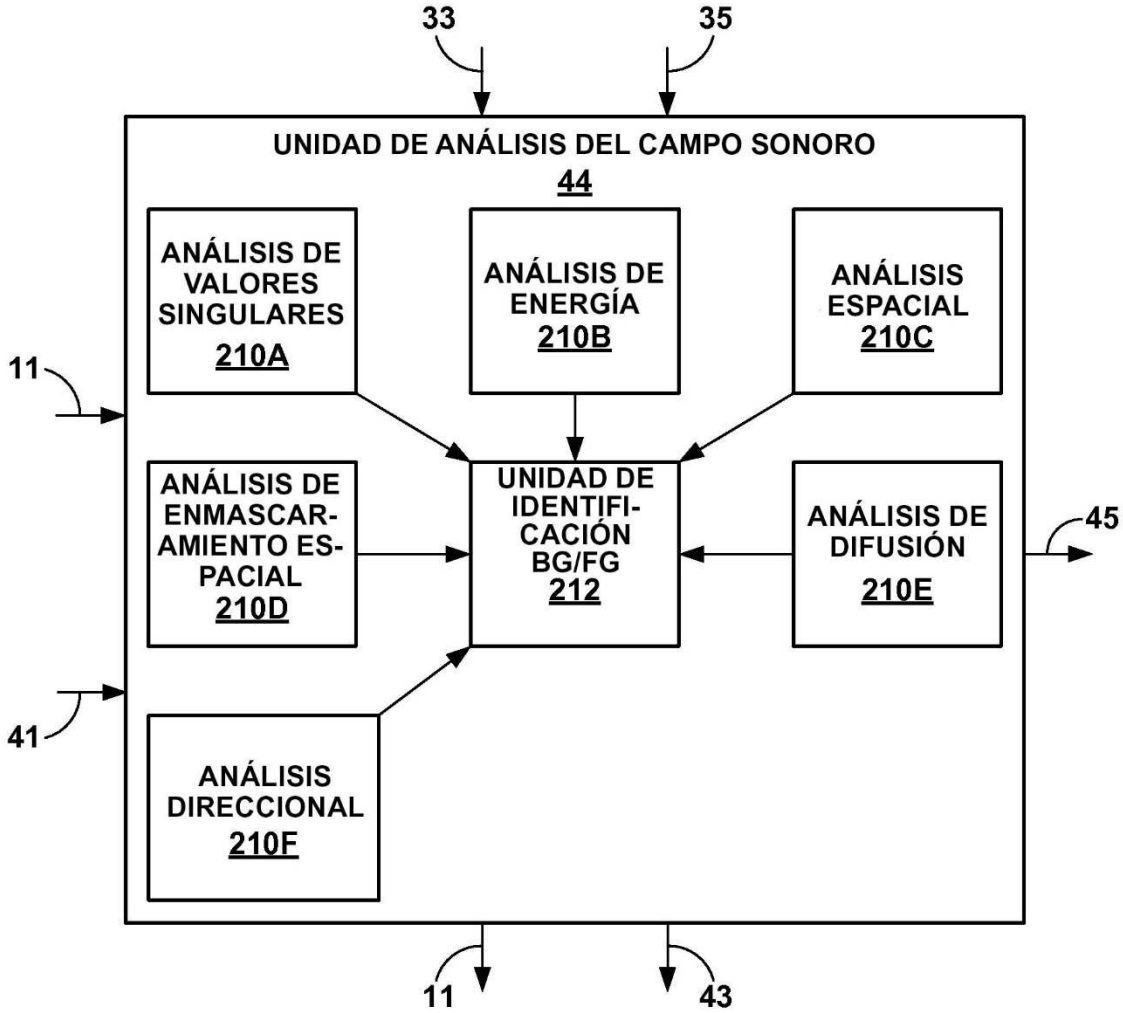

33
35
UNIDAD DE ANÁLISIS DEL CAMPO SONORO
44
ANÁLISIS DE VALORES SINGULARES
210A
ANÁLISIS DE ENERGÍA
210B
ANÁLISIS ESPACIAL
210C
11
ANÁLISIS DE ENMASCAR-AMIENTO ES-PACIAL
210D
UNIDAD DE IDENTIFI-CACIÓN BG/FG
212
ANÁLISIS DE DIFUSIÓN
210E
45
41
ANÁLISIS DIRECCIONAL
210F
11
43

FIG. 9D

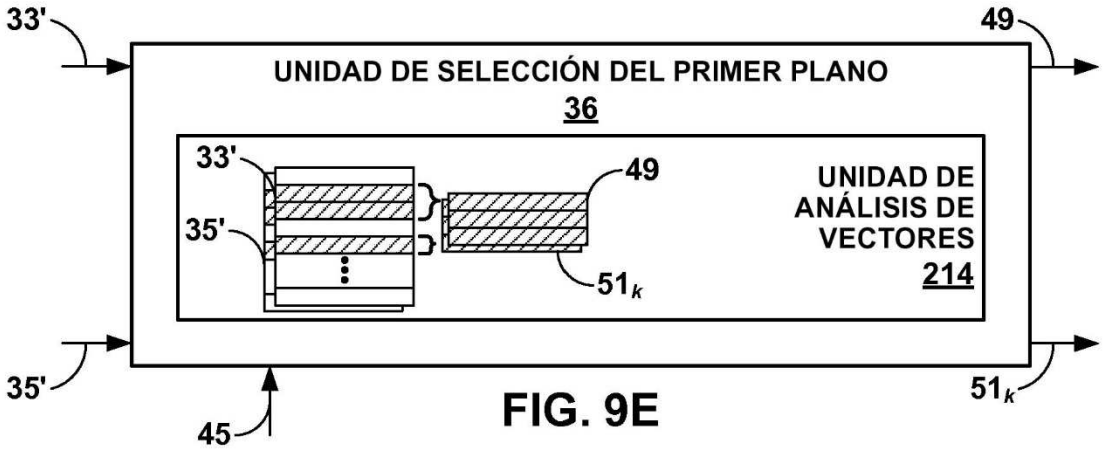

33'
49
UNIDAD DE SELECCIÓN DEL PRIMER PLANO
36
33'
49
35'
51_k
UNIDAD DE ANÁLISIS DE VECTORES
214
35'
45
51_k

FIG. 9E

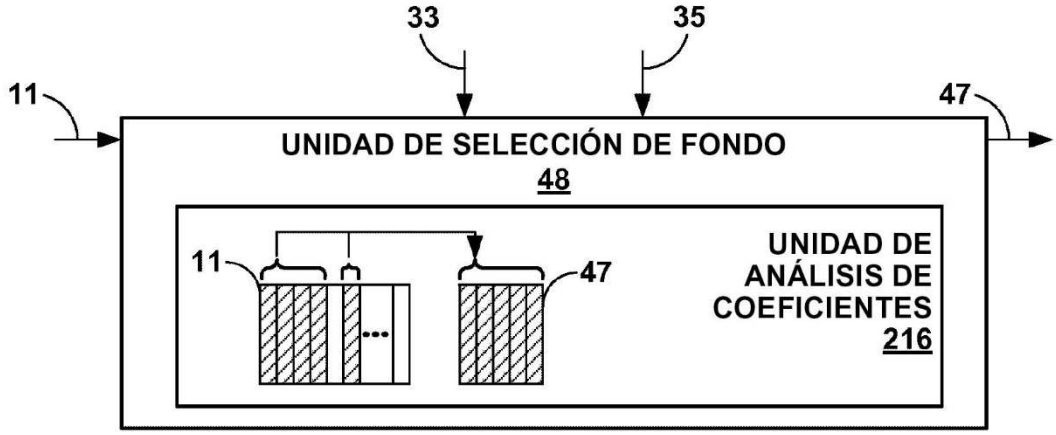


FIG. 9F

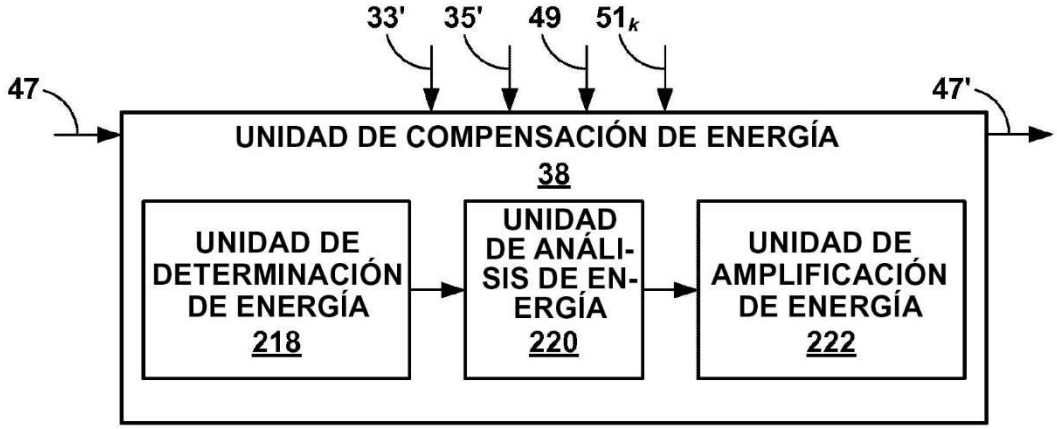


FIG. 9G

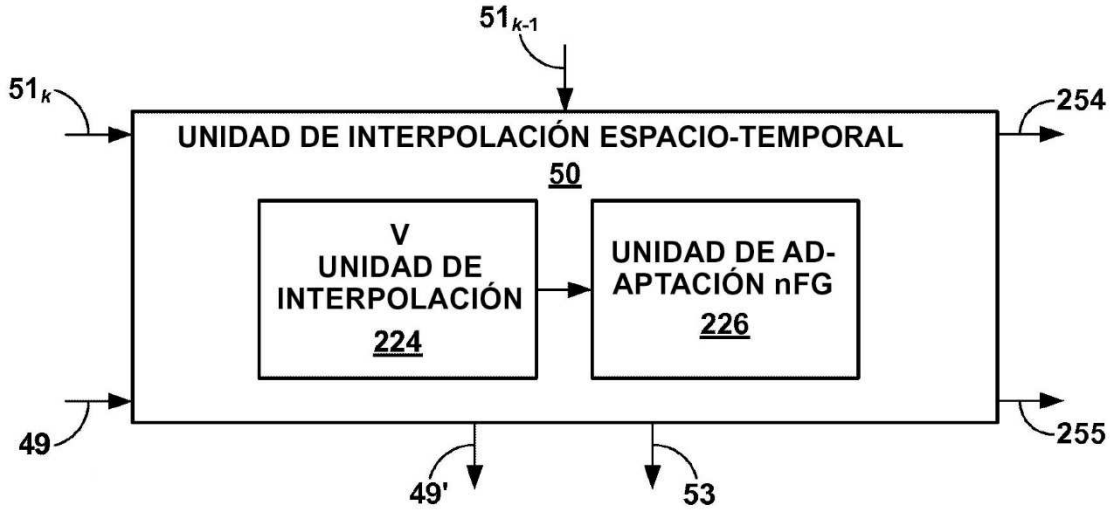


FIG. 9H

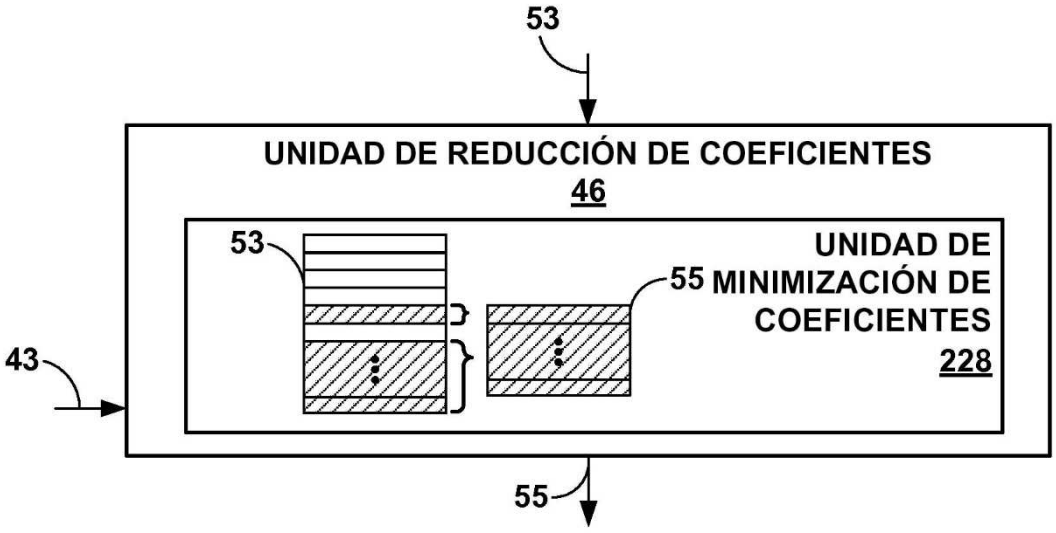

53
UNIDAD DE REDUCCIÓN DE COEFICIENTES
46
53
55
UNIDAD DE MINIMIZACIÓN DE COEFICIENTES
228
43
55

FIG. 9I

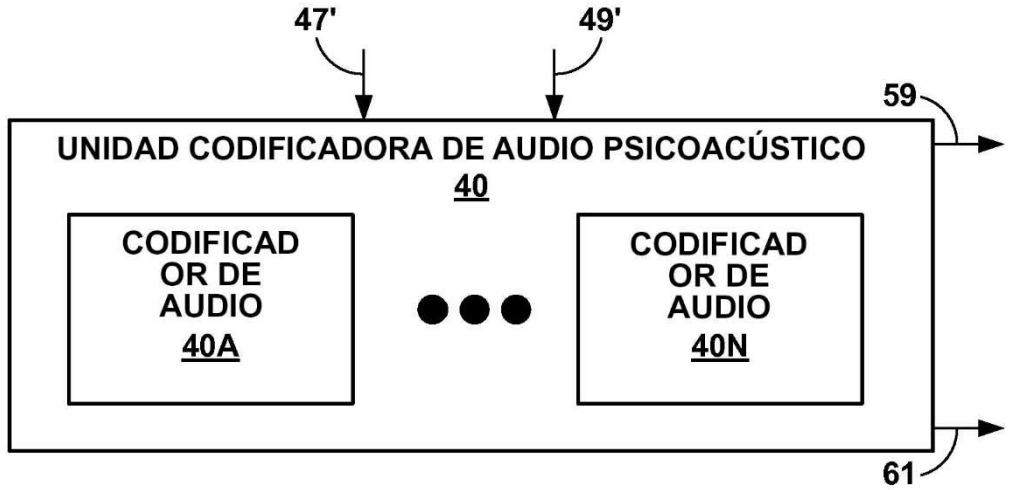

47'
49'
59
UNIDAD CODIFICADORA DE AUDIO PSICOACÚSTICO
40
CODIFICADOR DE AUDIO
40A
CODIFICADOR DE AUDIO
40N
61

FIG. 9J

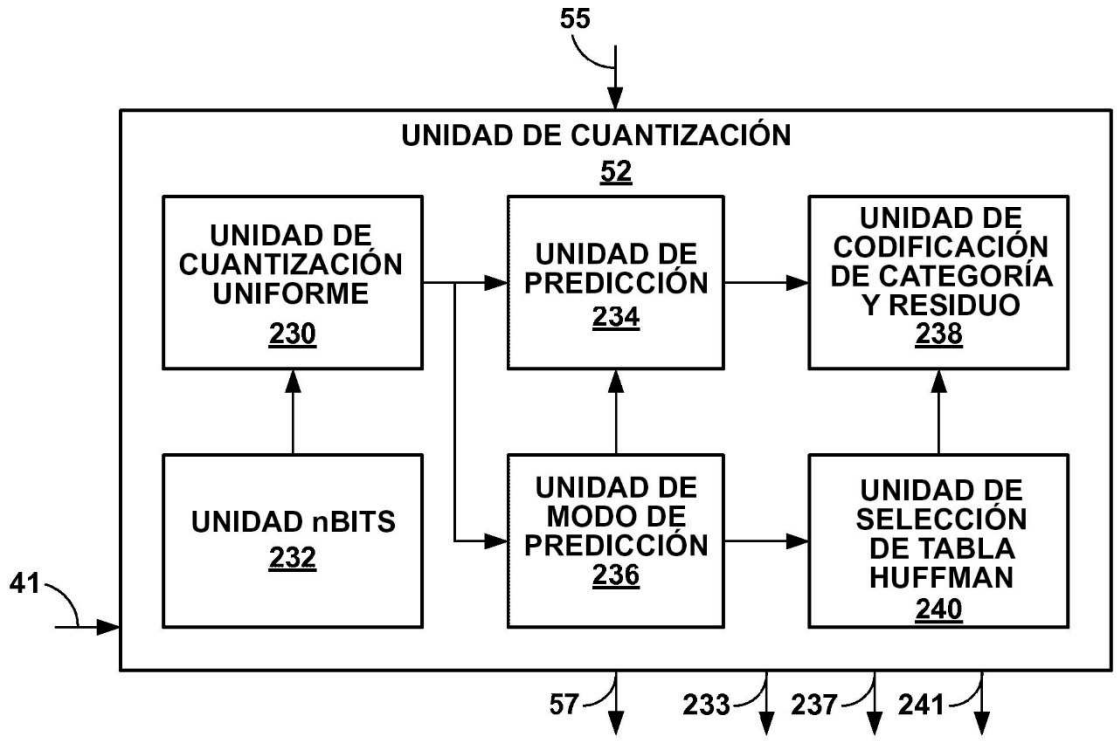

55
UNIDAD DE CUANTIZACIÓN
52
UNIDAD DE CUANTIZACIÓN UNIFORME
230
UNIDAD DE PREDICCIÓN
234
UNIDAD DE CODIFICACIÓN DE CATEGORÍA Y RESIDUO
238
UNIDAD nBITS
232
UNIDAD DE MODO DE PREDICCIÓN
236
UNIDAD DE SELECCIÓN DE TABLA HUFFMAN
240
41
57
233
237
241

FIG. 9K

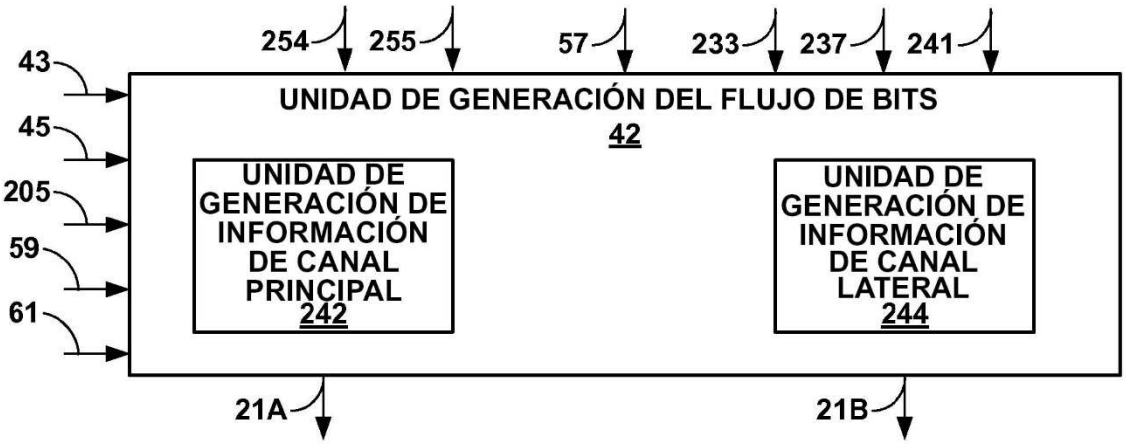

254
255
57
233
237
241
43
45
205
59
61
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS
42
UNIDAD DE GENERACIÓN DE INFORMACIÓN DE CANAL PRINCIPAL
242
UNIDAD DE GENERACIÓN DE INFORMACIÓN DE CANAL LATERAL
244
21A
21B

FIG. 9L

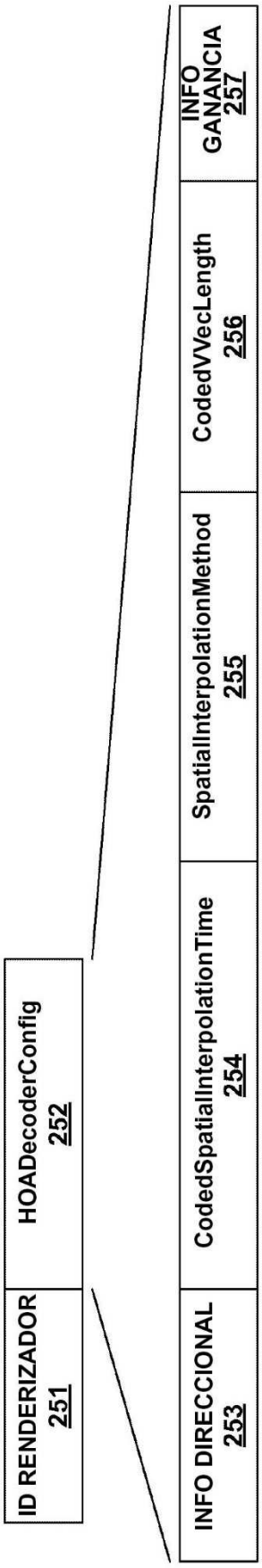


FIG. 10A

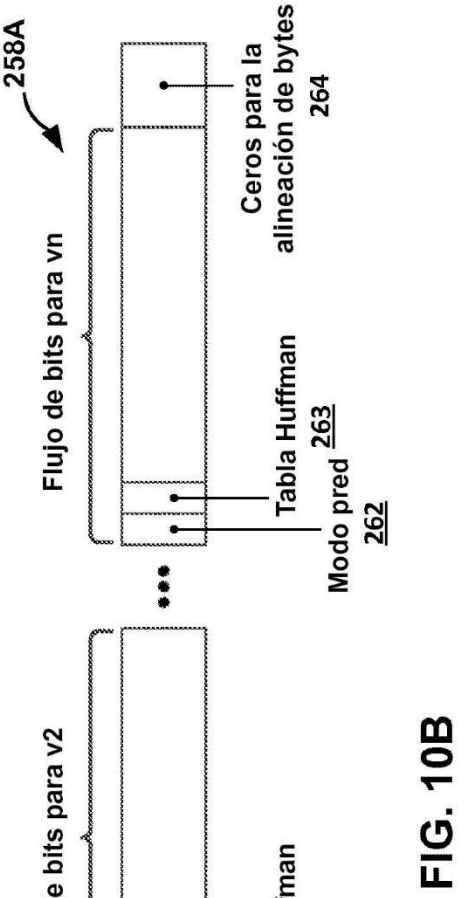


FIG. 10B

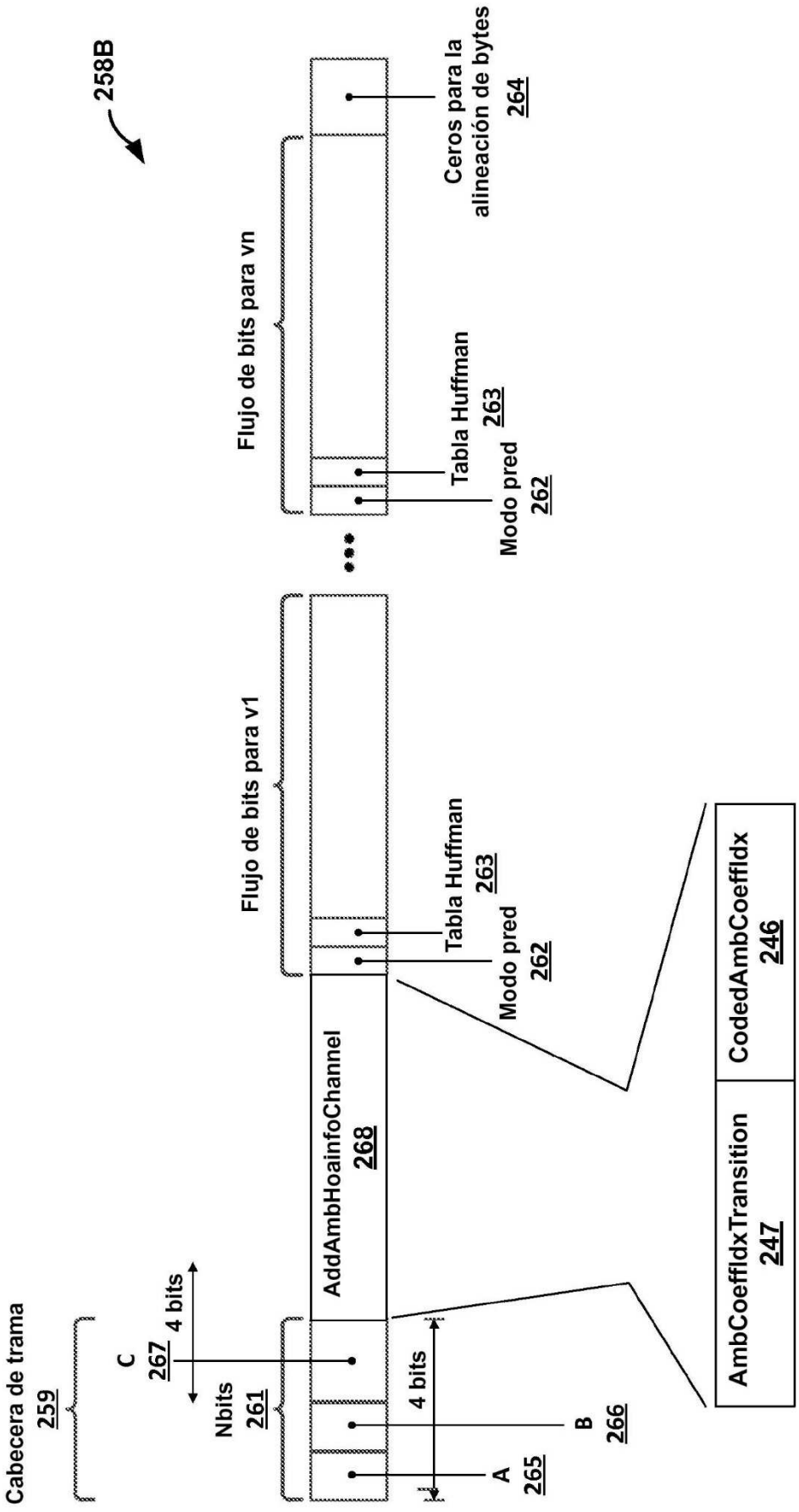

258B
Cabecera de trama
259
C
267
4 bits
Nbits
261
AddAmbHoainfoChannel
268
4 bits
A
265
B
266
Flujo de bits para v1
Tabla Huffman
263
Modo pred
262
Flujo de bits para vn
Tabla Huffman
263
Modo pred
262
Ceros para la alineación de bytes
264
AmbCoeffIdxTransition
247
CodedAmbCoeffIdx
246

FIG. 10C

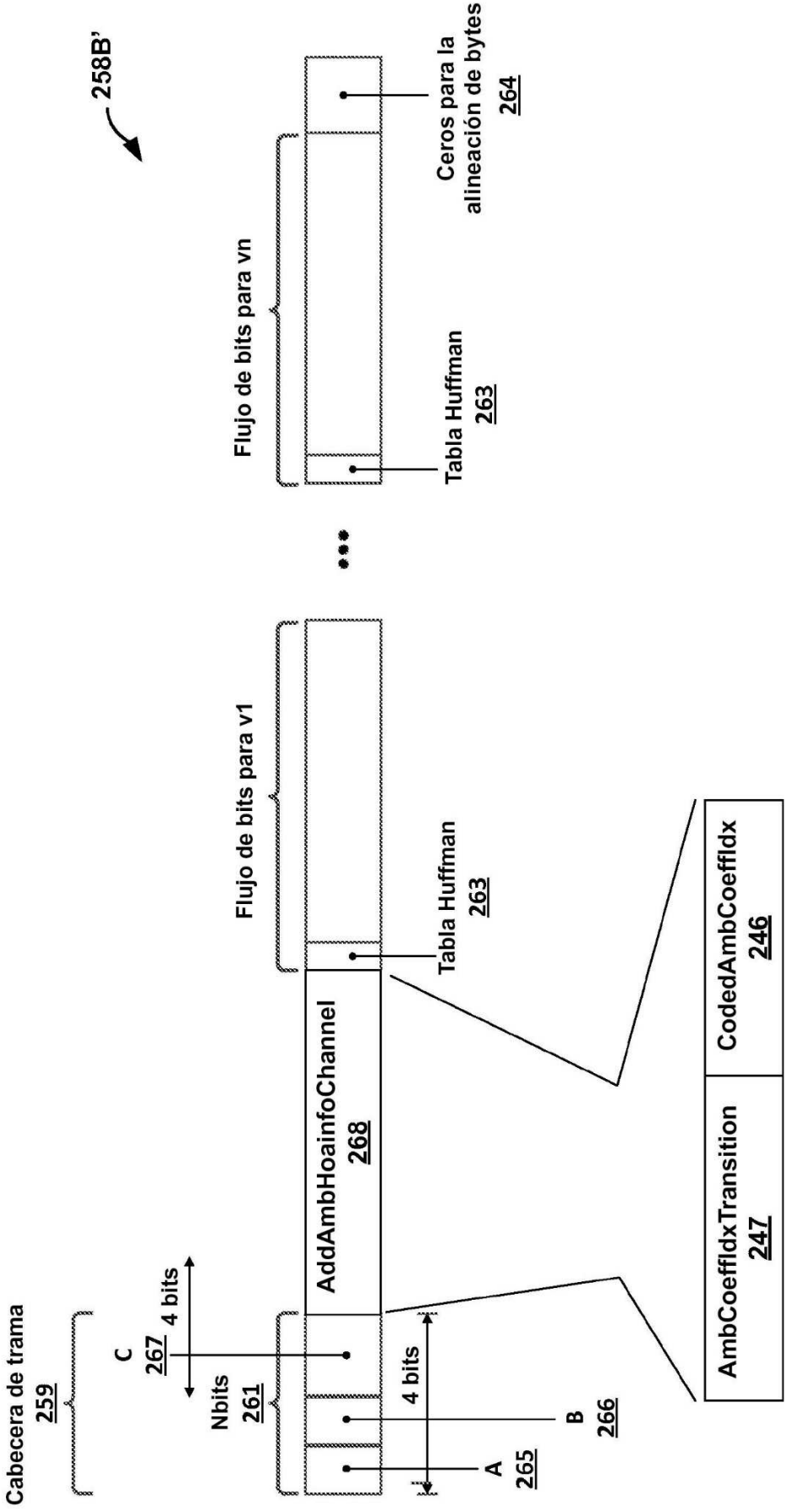

258B'
Cabecera de trama
259
C
267
4 bits
Nbits
261
A
265
4 bits
B
266
AddAmbHoainfoChannel
268
Flujo de bits para v1
Tabla Huffman
263
Flujo de bits para vn
Tabla Huffman
263
Ceros para la
alineación de bytes
264
AmbCoeffIdxTransition
247
CodedAmbCoeffIdx
246

FIG. 10C(i)

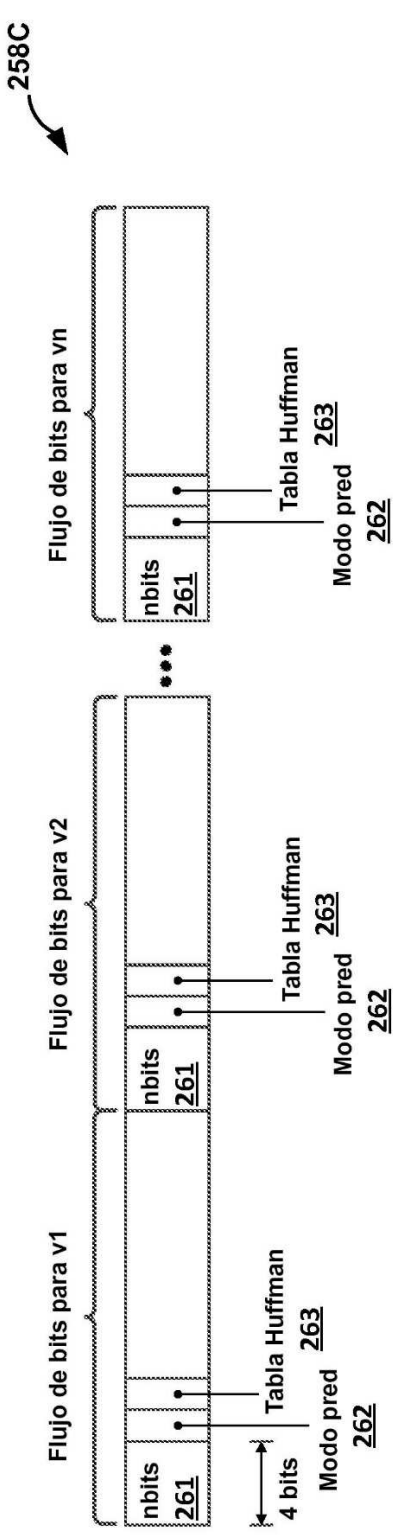


FIG. 10D

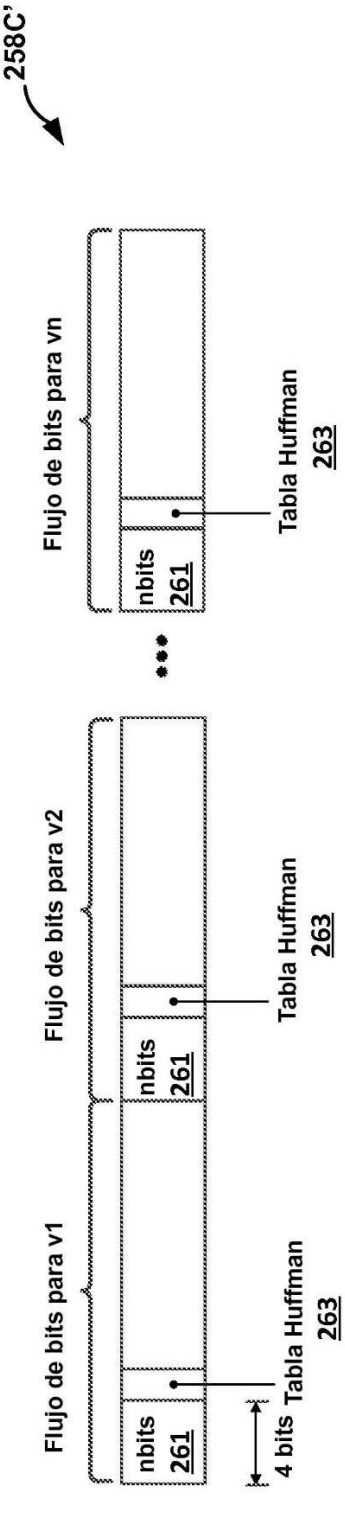


FIG. 10D(i)

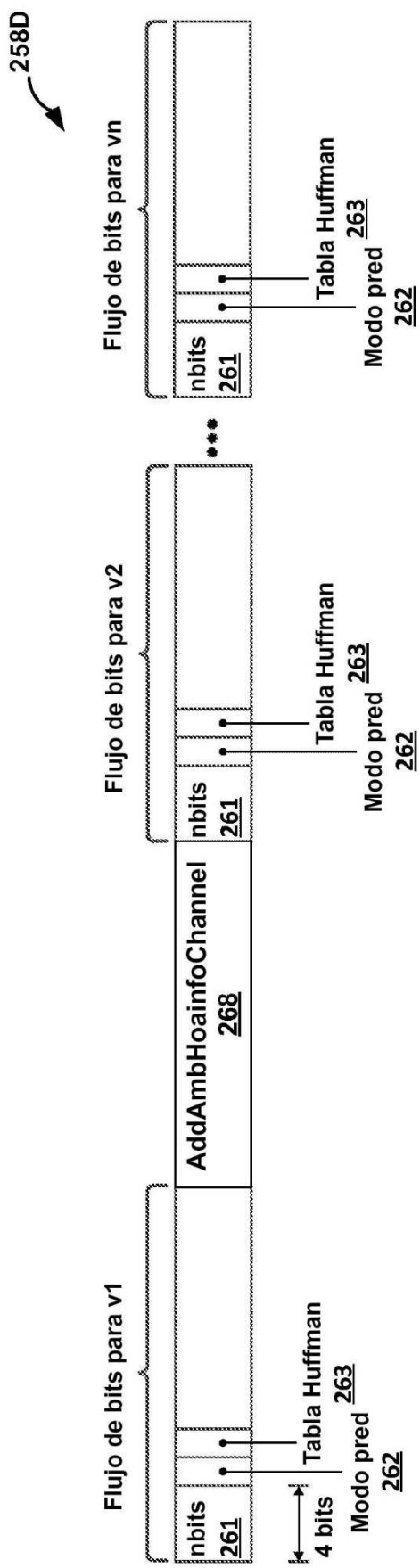


FIG. 10E

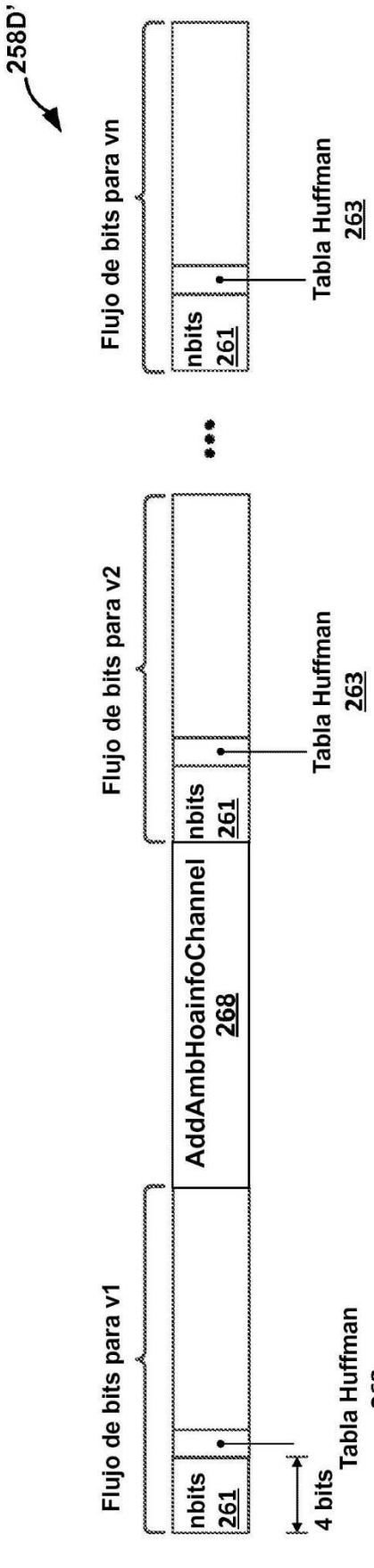


FIG. 10E(i)

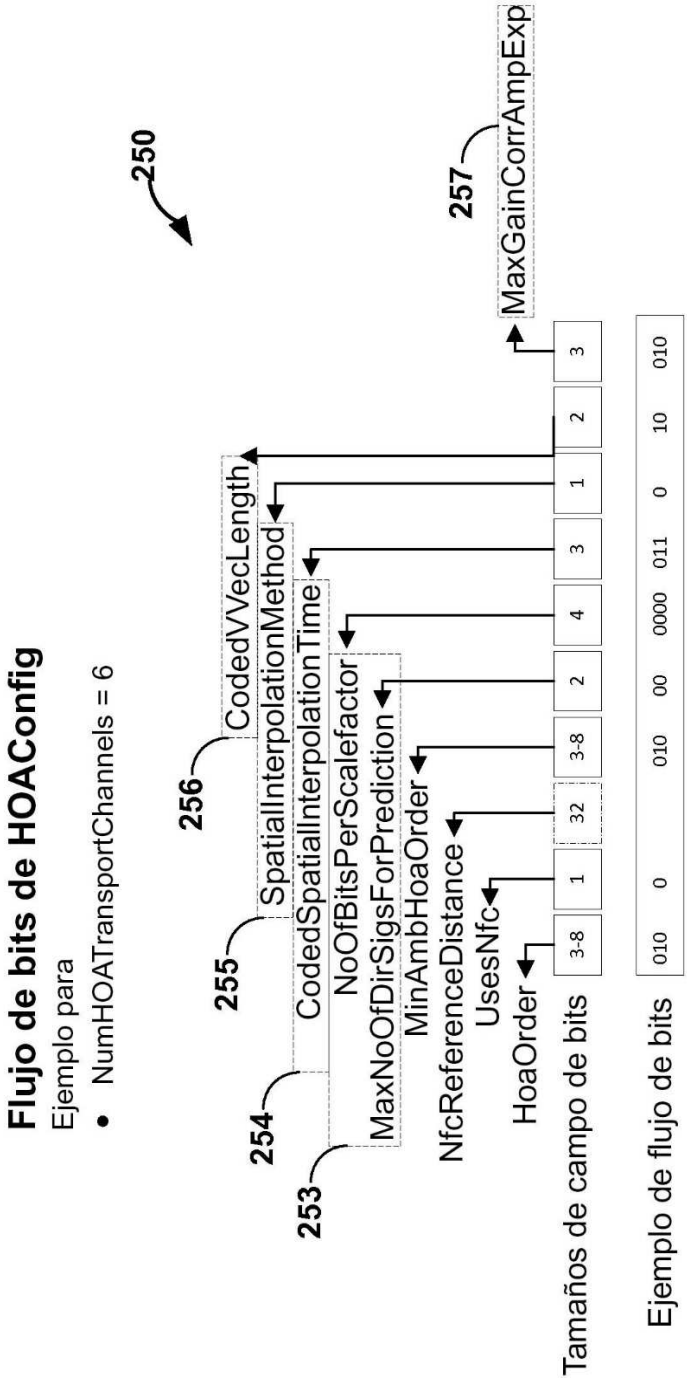

Flujo de bits de HOAConfig
Ejemplo para
• NumHOATransportChannels = 6
250
256
255
254
253
257
CodedVVecLength
SpatialInterpolationMethod
CodedSpatialInterpolationTime
NoOfBitsPerScalefactor
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
HoaOrder
MaxGainCorrAmpExp
Tamaños de campo de bits
3-8
1
32
3-8
2
4
3
1
2
3
Ejemplo de flujo de bits
010
0
010
00
0000
011
0
10
010

FIG. 10F

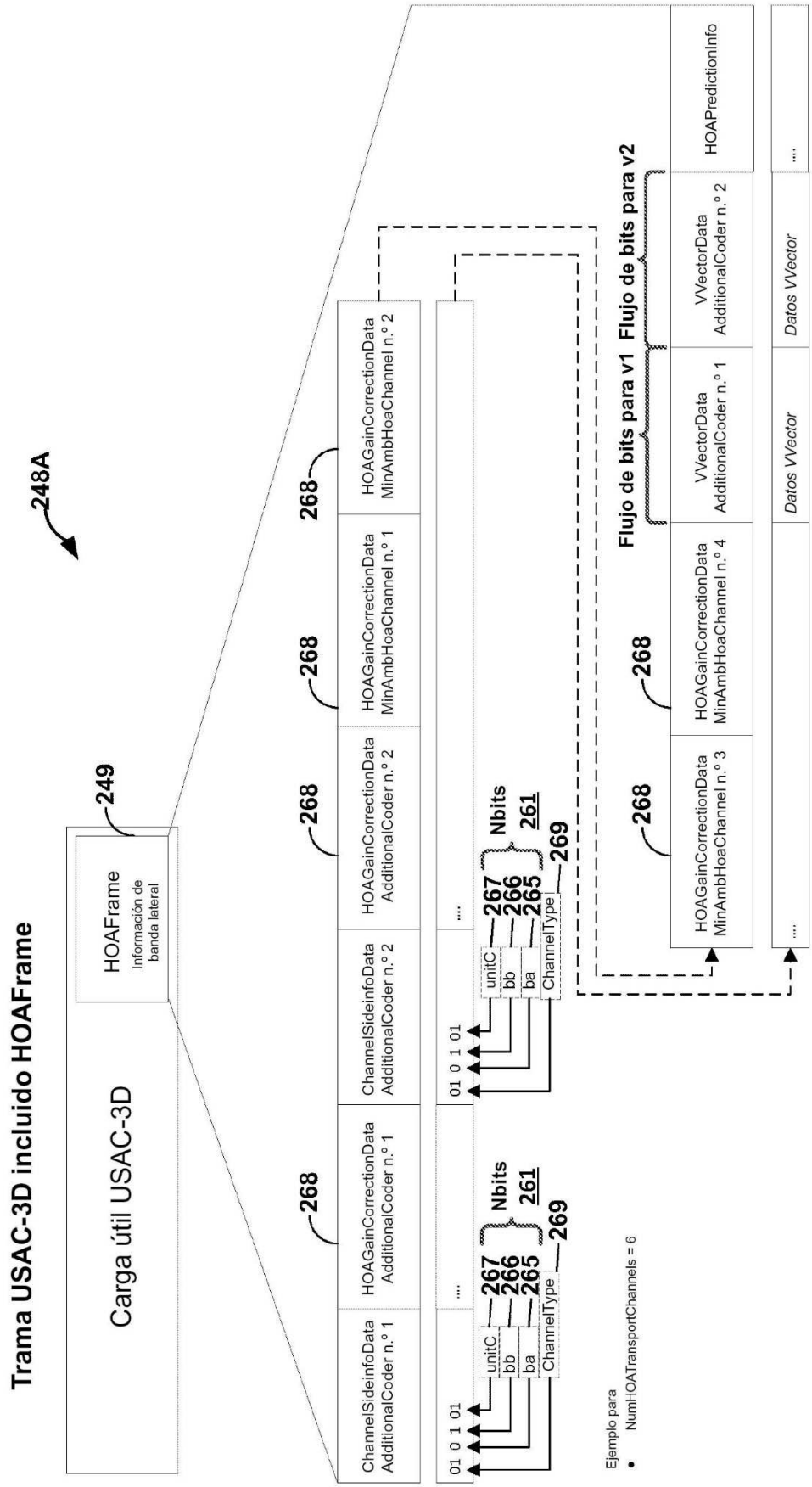

Trama USAC-3D incluido HOAFrame
248A
Carga útil USAC-3D
HOAFrame
Información de banda lateral
249
ChannelSideinfoData AdditionalCoder n.º 1
HOAGainCorrectionData AdditionalCoder n.º 1
ChannelSideinfoData AdditionalCoder n.º 2
HOAGainCorrectionData AdditionalCoder n.º 2
HOAGainCorrectionData MinAmbHoaChannel n.º 1
HOAGainCorrectionData MinAmbHoaChannel n.º 2
HOAGainCorrectionData MinAmbHoaChannel n.º 3
HOAGainCorrectionData MinAmbHoaChannel n.º 4
VVectorData AdditionalCoder n.º 1
VVectorData AdditionalCoder n.º 2
HOAPredictionInfo
Datos VVector
268
Nbits
261
269
267
266
265
unitC
bb
ba
ChannelType
01 0 1 01
....
Flujo de bits para v1
Flujo de bits para v2
Ejemplo para
NumHOATransportChannels = 6

FIG. 10G

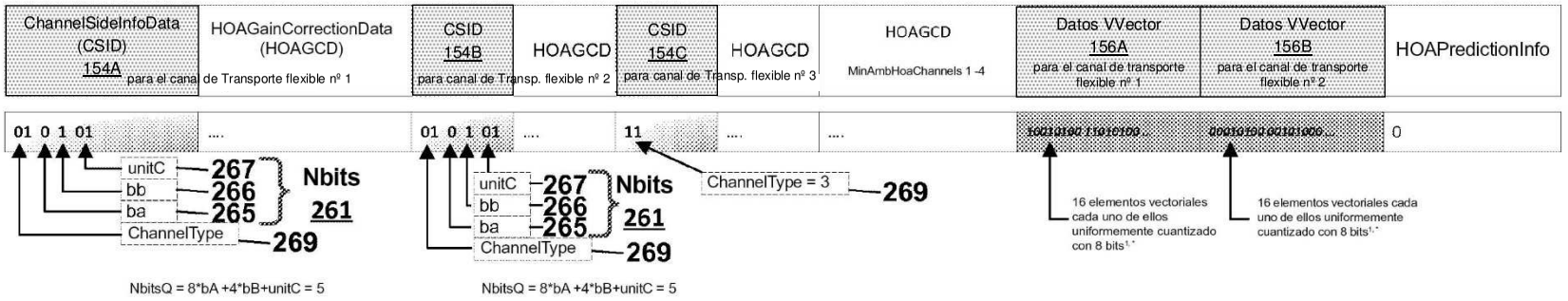
250H
Flujo de bits de HOAConfig
Ejemplo para
• NumHOATransportChannels = 7
256
CodedVVecLength
255
SpatialInterpolationMethod
254
CodedSpatialInterpolationTime
253
NoOfBitsPerScalefactor
150
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
152
UsesNfc
HoaOrder
257
MaxGainCorrAmpExp
Tamaños de campo de bits
Ejemplo de flujo de bits
3-8 1 32 3-8 2 4 1 2 3
010 0 010 00 0000 011 0 00 010
Significado de campos de bits
• Todos los elementos de un Vvector están codificados, por ejemplo, 16 elementos Vvector
• Función de interpolación de la interpolación espacio-temporal: coseno elevado
• La interpolación espacio-temporal de las señales de primer plano basadas en vectores: 256 muestras
• Orden HOA mínimo del contenido HOA ambiental: 1,
MinNumOfCoeffsForAmbHOA = (1+1)² = 4
• Orden HOA del contenido: N = 3, NumOfHoaCoeffs = (N+1)²=16
248H
Trama USAC-3D con SBR incluido HOAFrame
2 HOAFrames consecutivos se almacenan en la carga útil de extensión USAC, ya que dos tramas de audio se almacenan dentro de una trama USAC-3D cuando SBR está habilitada. Flexible Transport channels = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 canales de transporte flexibles
Carga útil USAC-3D
HOAFrame
Trama con información de banda lateral k
249A
HOAFrame
Trama con información de banda lateral k+1
249B

FIG. 10H(i)

249A

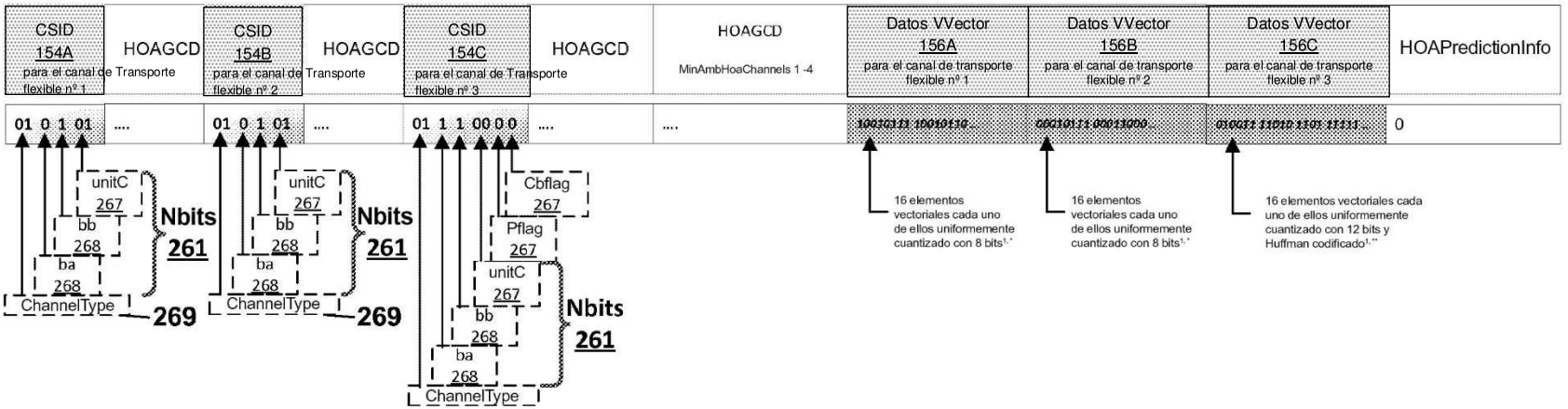
Trama k

En este ejemplo, los canales de transporte flexibles contienen dos canales de primer plano basados en vectores en la trama k y tres vectores de primer plano basados en vectores en la segunda trama k+1



249B

Trama k+1



Número e índices [1] de elementos de VVectorData codificados especificados por el parámetro CodedVVecLength = 0

* Esquema de codificación señalizado por NbitsQ = 5 en el ChannelSideInfo para este canal de transporte en particular

** Esquema de codificación de Huffman señalizado por NbitsQ = 12, CbFlag = 0 y Pflag = 0 en el ChannelSideInfo para este canal de transporte particular

FIG. 10H(ii)

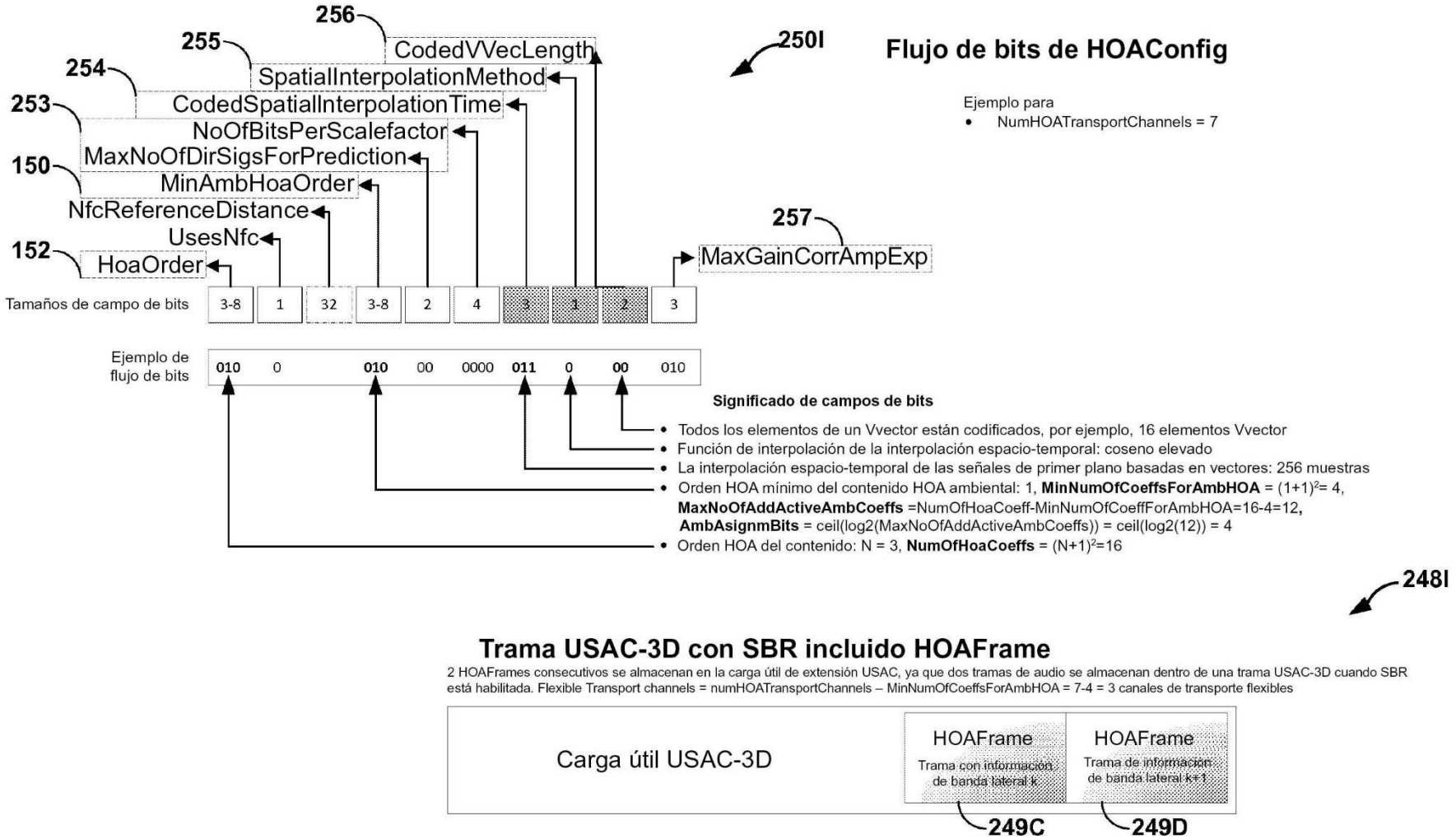
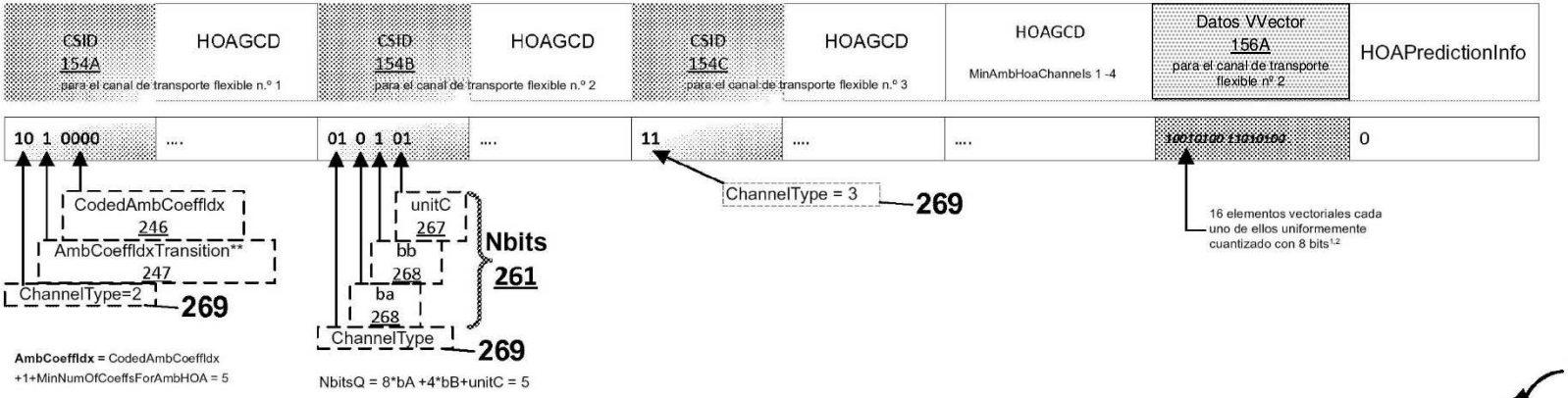

256
255
254
253
150
152
CodedVVecLength
SpatialInterpolationMethod
CodedSpatialInterpolationTime
NoOfBitsPerScalefactor
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
HoaOrder
257
MaxGainCorrAmpExp
250I
Flujo de bits de HOAConfig
Ejemplo para
• NumHOATransportChannels = 7
Tamaños de campo de bits
3-8
1
32
3-8
2
4
3
1
2
3
Ejemplo de flujo de bits
010 0 010 00 0000 011 0 00 010
Significado de campos de bits
• Todos los elementos de un Vvector están codificados, por ejemplo, 16 elementos Vvector
• Función de interpolación de la interpolación espacio-temporal: coseno elevado
• La interpolación espacio-temporal de las señales de primer plano basadas en vectores: 256 muestras
• Orden HOA mínimo del contenido HOA ambiental: 1, MinNumOfCoeffsForAmbHOA = (1+1)²= 4, MaxNoOfAddActiveAmbCoeffs =NumOfHoaCoeff-MinNumOfCoeffForAmbHOA=16-4=12, AmbAsignmBits = ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4
• Orden HOA del contenido: N = 3, NumOfHoaCoeffs = (N+1)²=16
248I
Trama USAC-3D con SBR incluido HOAFrame
2 HOAFrames consecutivos se almacenan en la carga útil de extensión USAC, ya que dos tramas de audio se almacenan dentro de una trama USAC-3D cuando SBR está habilitada. Flexible Transport channels = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 canales de transporte flexibles
Carga útil USAC-3D
HOAFrame
Trama con información de banda lateral k
HOAFrame
Trama de información de banda lateral k+1
249C
249D

FIG. 10I(i)

249C

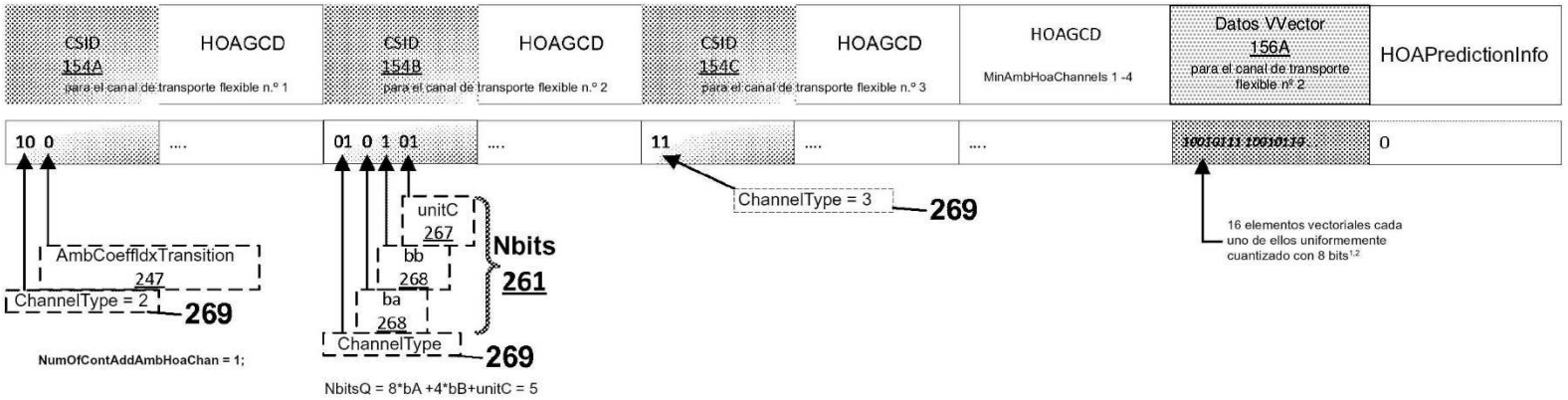
Trama k

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



249D

Trama k+1



Número e índices [1] de elementos VVectorData codificados especificados por el parámetro CodedVVecLength = 0

Esquema de codificación [2] señalizado por NbitsQ = 5 en el ChannelSideinfo para el canal de transporte flexible n.º 2

** Para el canal de transporte flexible n.º1 se supone que el estado interno del decodificador es AmbCoeffIdxTransitionState = 2. Por lo tanto, se señaliza el campo de bits CodedAmbCoeffIdx.

FIG. 10I(ii)

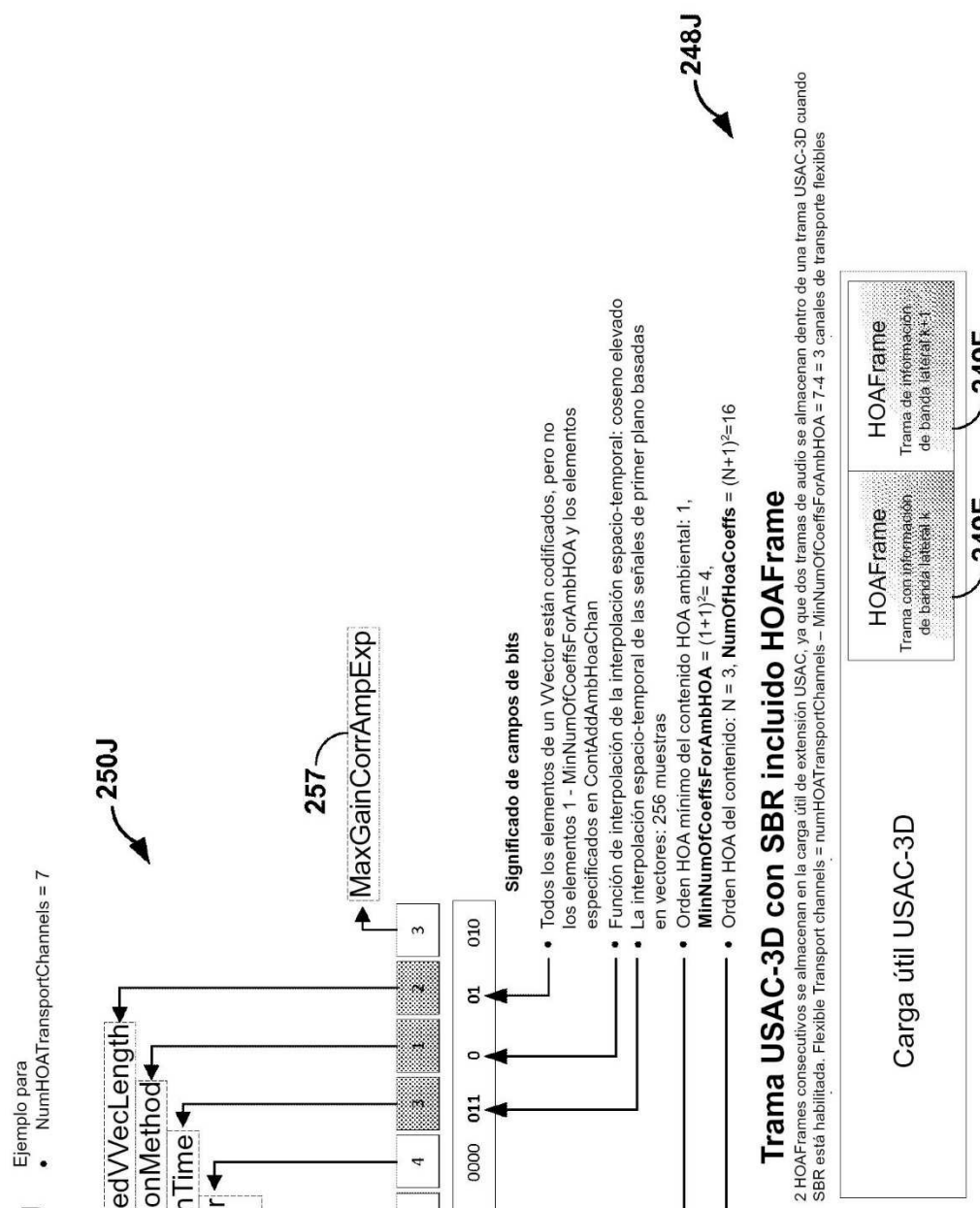

Flujo de bits de HOAConfig
Ejemplo para
• NumHOATransportChannels = 7
250J
256
CodedVVecLength
255
SpatialInterpolationMethod
254
CodedSpatialInterpolationTime
253
NoOfBitsPerScalefactor
150
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
152
UsesNfc
HoaOrder
257
MaxGainCorrAmpExp
Tamaños de campo de bits
3-8
1
32
3-8
2
4
3
1
2
3
Ejemplo de flujo de bits
010
0
010
00
0000
011
0
01
010
Significado de campos de bits
• Todos los elementos de un VVector están codificados, pero no los elementos 1 - MinNumOfCoeffsForAmbHOA y los elementos especificados en ContAddAmbHoaChan
• Función de interpolación de la interpolación espacio-temporal: coseno elevado
• La interpolación espacio-temporal de las señales de primer plano basadas en vectores: 256 muestras
• Orden HOA mínimo del contenido HOA ambiental: 1, MinNumOfCoeffsForAmbHOA = (1+1)2= 4,
• Orden HOA del contenido: N = 3, NumOfHoaCoeffs = (N+1)2=16
248J
Trama USAC-3D con SBR incluido HOAFrame
2 HOAFrames consecutivos se almacenan en la carga útil de extensión USAC, ya que dos tramas de audio se almacenan dentro de una trama USAC-3D cuando SBR está habilitada. Flexible Transport channels = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 canales de transporte flexibles
Carga útil USAC-3D
HOAFrame
Trama con información de banda lateral k
249E
HOAFrame
Trama de información de banda lateral k+1
249F

FIG. 10J(i)

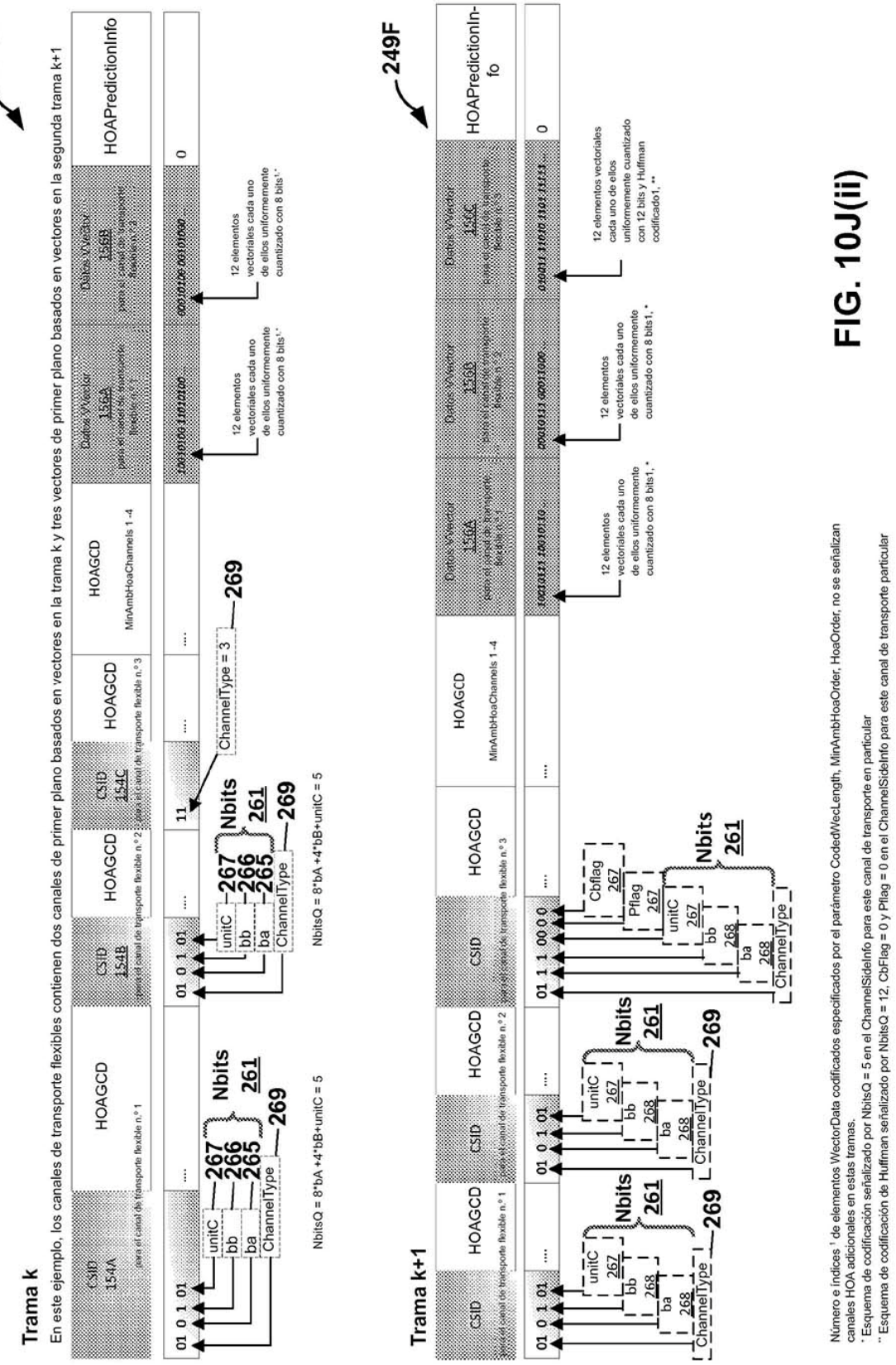


FIG. 10J(ii)

Número e índices [1] de elementos WectorData codificados especificados por el parámetro CodedWecLength, MinAmbHoaOrder, HoaOrder, no se señalizan canales HOA adicionales en estas tramas.
* Esquema de codificación señalizado por NbitsQ = 5 en el ChannelSideInfo para este canal de transporte en particular
** Esquema de codificación de Huffman señalizado por NbitsQ = 12, CbFlag = 0 y Pflag = 0 en el ChannelSideInfo para este canal de transporte particular

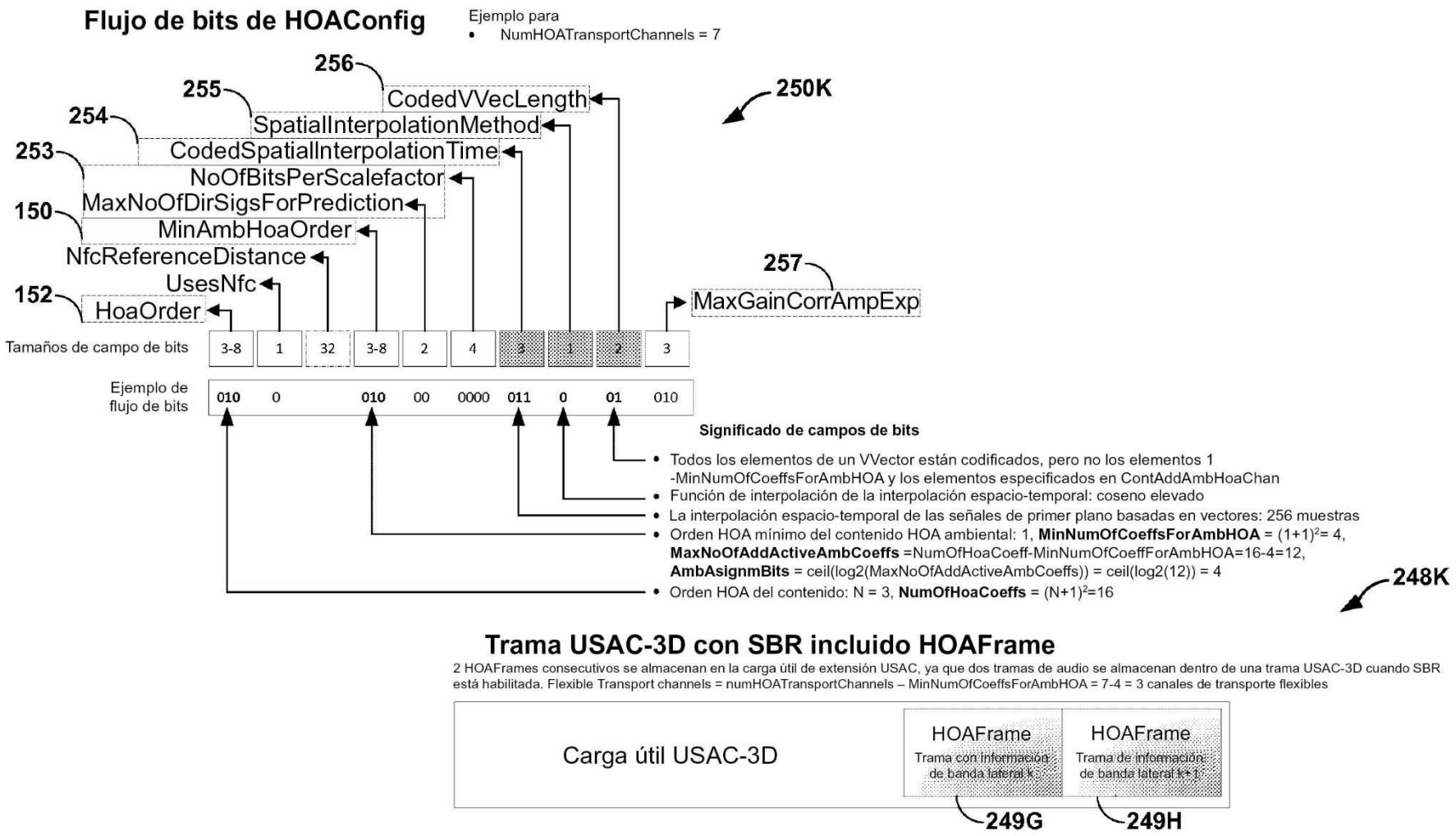
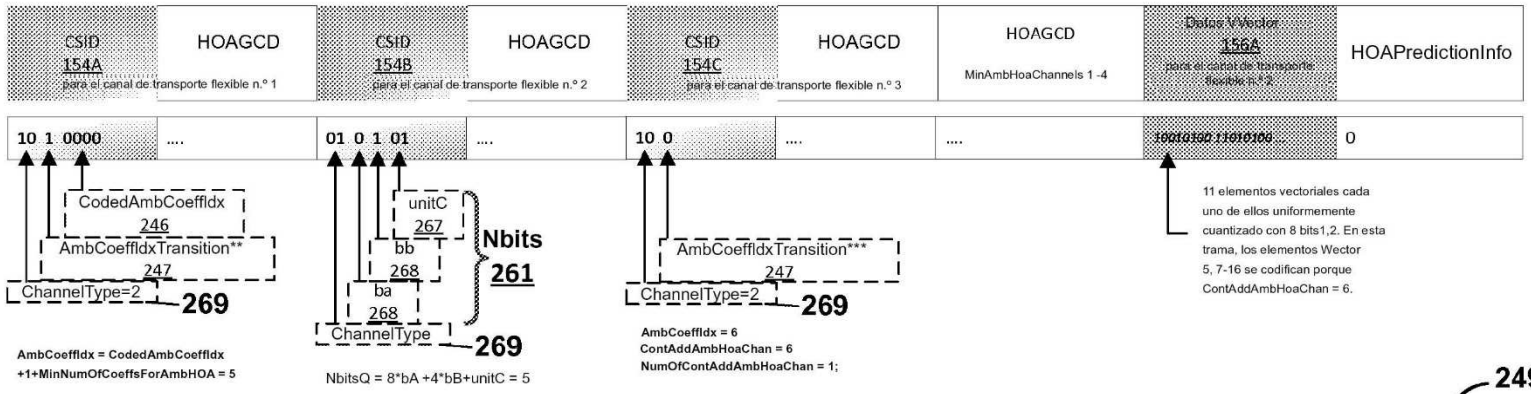

Flujo de bits de HOAConfig
Ejemplo para
• NumHOATransportChannels = 7
256
255
254
253
150
152
257
250K
CodedVVecLength
SpatialInterpolationMethod
CodedSpatialInterpolationTime
NoOfBitsPerScalefactor
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
HoaOrder
MaxGainCorrAmpExp
Tamaños de campo de bits
3-8
1
32
3-8
2
4
3
1
2
3
Ejemplo de flujo de bits
010
0
010
00
0000
011
0
01
010
Significado de campos de bits
• Todos los elementos de un VVector están codificados, pero no los elementos 1 -MinNumOfCoeffsForAmbHOA y los elementos especificados en ContAddAmbHoaChan
• Función de interpolación de la interpolación espacio-temporal: coseno elevado
• La interpolación espacio-temporal de las señales de primer plano basadas en vectores: 256 muestras
• Orden HOA mínimo del contenido HOA ambiental: 1, MinNumOfCoeffsForAmbHOA = (1+1)²= 4, MaxNoOfAddActiveAmbCoeffs =NumOfHoaCoeff-MinNumOfCoeffForAmbHOA=16-4=12, AmbAsignmBits = ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4
• Orden HOA del contenido: N = 3, NumOfHoaCoeffs = (N+1)²=16
248K
Trama USAC-3D con SBR incluido HOAFrame
2 HOAFrames consecutivos se almacenan en la carga útil de extensión USAC, ya que dos tramas de audio se almacenan dentro de una trama USAC-3D cuando SBR está habilitada. Flexible Transport channels = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 canales de transporte flexibles
Carga útil USAC-3D
HOAFrame
Trama con información de banda lateral k
HOAFrame
Trama de información de banda lateral k+1
249G
249H

FIG. 10K(i)

Trama k

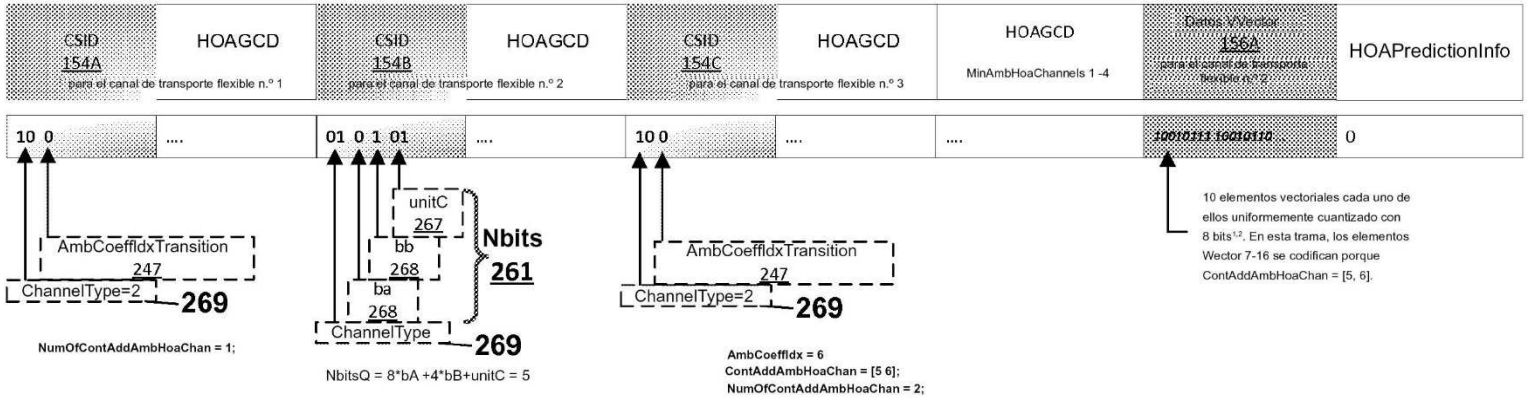
249G

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



249H

Trama k+1



Número e índices [1] de los elementos VVectorData codificados especificados por el parámetro CodedVVecLength, MinNumOfCoeffsForAmbHOA, NumOfHoaCoeff

Esquema de codificación [2] señalizado por NbitsQ = 5 en el ChannelSideinfo para el canal de transporte flexible n.º 2

** Para el canal de transporte flexible n.º1 se supone que el estado interno del decodificador es AmbCoeffIdxTransitionState = 2. Por lo tanto, se señaliza el campo de bits CodedAmbCoeffIdx.

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

FIG. 10K(ii)

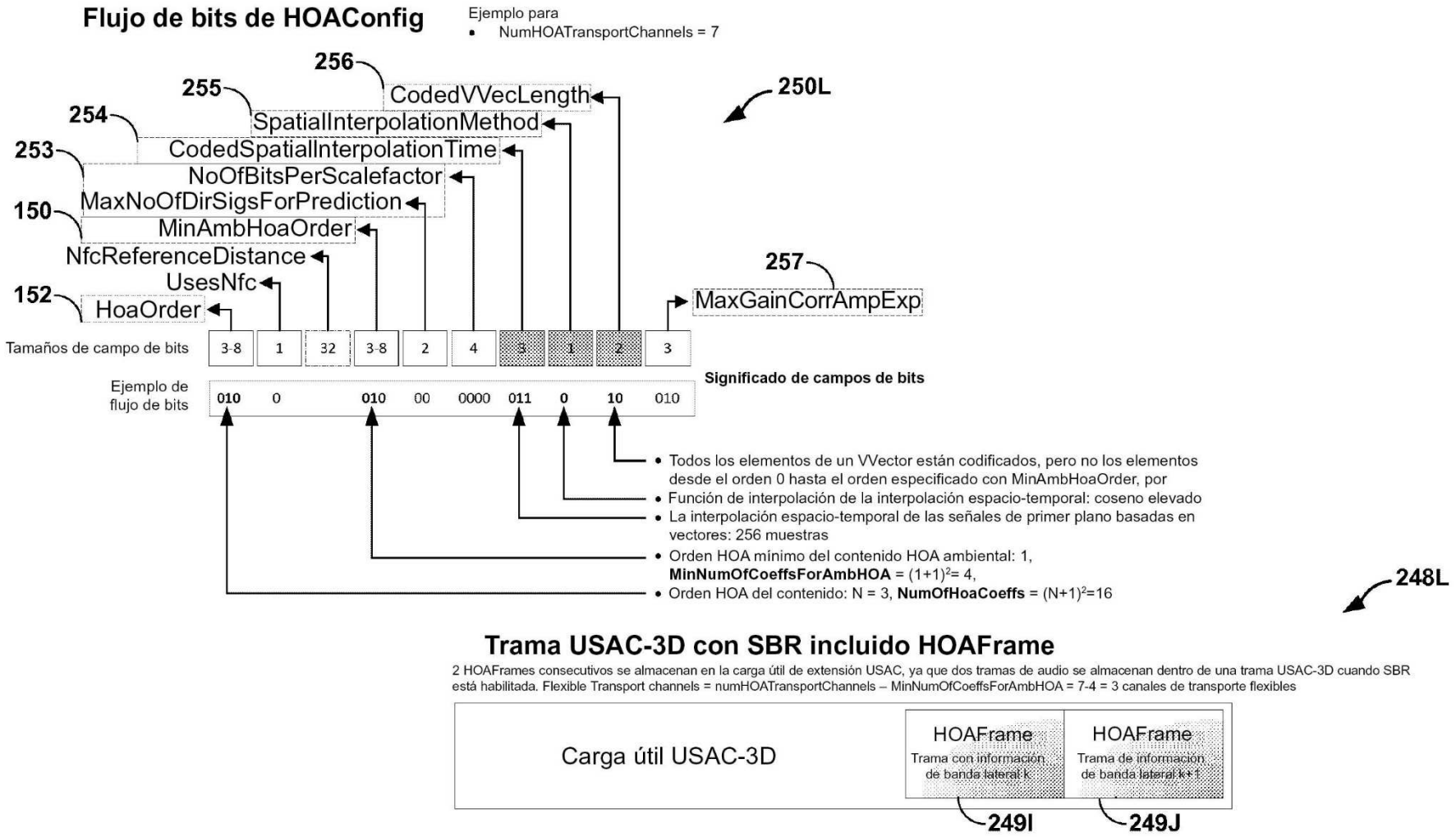
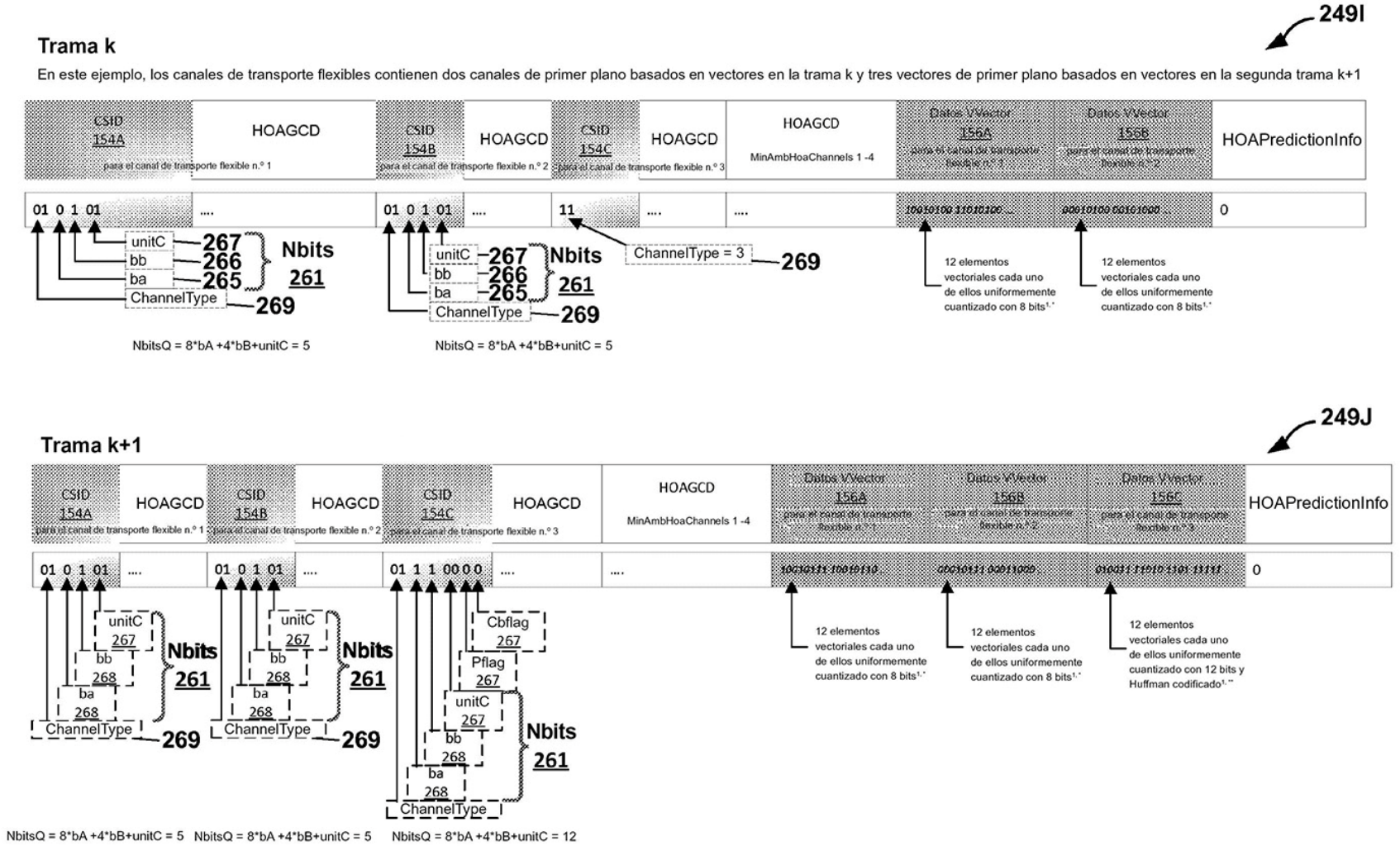

Flujo de bits de HOAConfig
Ejemplo para
• NumHOATransportChannels = 7
256
CodedVVecLength
255
SpatialInterpolationMethod
254
CodedSpatialInterpolationTime
253
NoOfBitsPerScalefactor
MaxNoOfDirSigsForPrediction
150
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
152
HoaOrder
250L
257
MaxGainCorrAmpExp
Tamaños de campo de bits
3-8
1
32
3-8
2
4
5
1
2
3
Ejemplo de flujo de bits
010
0
010
00
0000
011
0
10
010
Significado de campos de bits
• Todos los elementos de un VVector están codificados, pero no los elementos desde el orden 0 hasta el orden especificado con MinAmbHoaOrder, por
• Función de interpolación de la interpolación espacio-temporal: coseno elevado
• La interpolación espacio-temporal de las señales de primer plano basadas en vectores: 256 muestras
• Orden HOA mínimo del contenido HOA ambiental: 1, MinNumOfCoeffsForAmbHOA = (1+1)²= 4,
• Orden HOA del contenido: N = 3, NumOfHoaCoeffs = (N+1)²=16
248L
Trama USAC-3D con SBR incluido HOAFrame
2 HOAFrames consecutivos se almacenan en la carga útil de extensión USAC, ya que dos tramas de audio se almacenan dentro de una trama USAC-3D cuando SBR está habilitada. Flexible Transport channels = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 canales de transporte flexibles
Carga útil USAC-3D
HOAFrame
Trama con información de banda lateral k
HOAFrame
Trama de información de banda lateral k+1
249I
249J

FIG. 10L(i)



Número e índices [1] de elementos VVectorData codificados especificados por el parámetro CodedVVecLength, MinAmbHoaOrder, HoaOrder

[*] Esquema de codificación señalizado por NbitsQ = 5 en el ChannelSideInfo para este canal de transporte en particular

[**] Esquema de codificación de Huffman señalizado por NbitsQ = 12, CbFlag = 0 y Pflag = 0 en el ChannelSideInfo para este canal de transporte particular

FIG. 10L(ii)

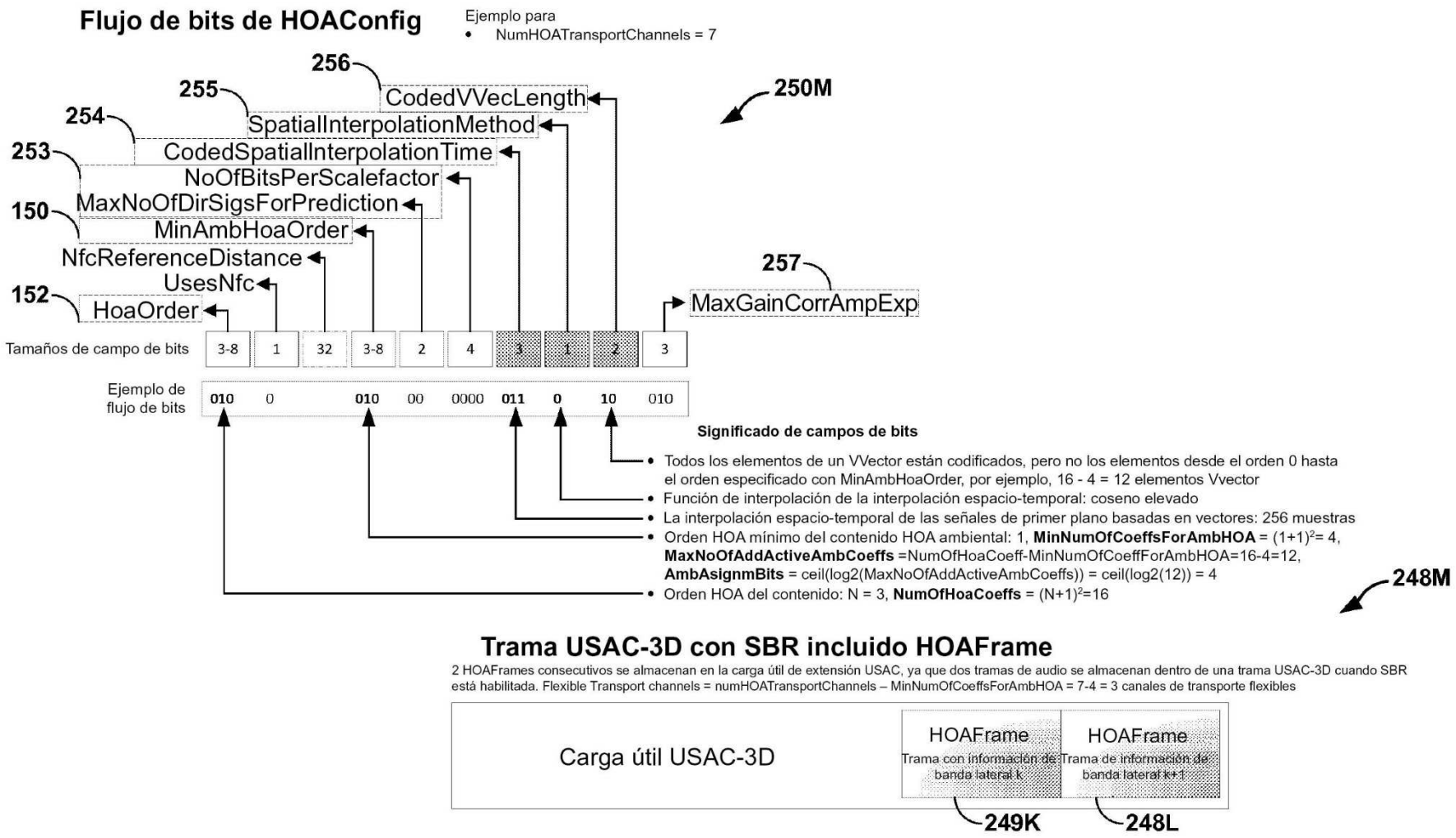

Flujo de bits de HOAConfig
Ejemplo para
• NumHOATransportChannels = 7
256
255
254
253
150
152
257
250M
CodedVVecLength
SpatialInterpolationMethod
CodedSpatialInterpolationTime
NoOfBitsPerScalefactor
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
HoaOrder
MaxGainCorrAmpExp
Tamaños de campo de bits
3-8
1
32
3-8
2
4
3
1
2
3
Ejemplo de flujo de bits
010
0
010
00
0000
011
0
10
010
Significado de campos de bits
• Todos los elementos de un VVector están codificados, pero no los elementos desde el orden 0 hasta el orden especificado con MinAmbHoaOrder, por ejemplo, 16 - 4 = 12 elementos Vvector
• Función de interpolación de la interpolación espacio-temporal: coseno elevado
• La interpolación espacio-temporal de las señales de primer plano basadas en vectores: 256 muestras
• Orden HOA mínimo del contenido HOA ambiental: 1, MinNumOfCoeffsForAmbHOA = (1+1)²= 4, MaxNoOfAddActiveAmbCoeffs =NumOfHoaCoeff-MinNumOfCoeffForAmbHOA=16-4=12, AmbAsignmBits = ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4
• Orden HOA del contenido: N = 3, NumOfHoaCoeffs = (N+1)²=16
248M
Trama USAC-3D con SBR incluido HOAFrame
2 HOAFrames consecutivos se almacenan en la carga útil de extensión USAC, ya que dos tramas de audio se almacenan dentro de una trama USAC-3D cuando SBR está habilitada. Flexible Transport channels = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 canales de transporte flexibles
Carga útil USAC-3D
HOAFrame
Trama con información de banda lateral k
HOAFrame
Trama de información de banda lateral k+1
249K
248L

FIG. 10M(i)

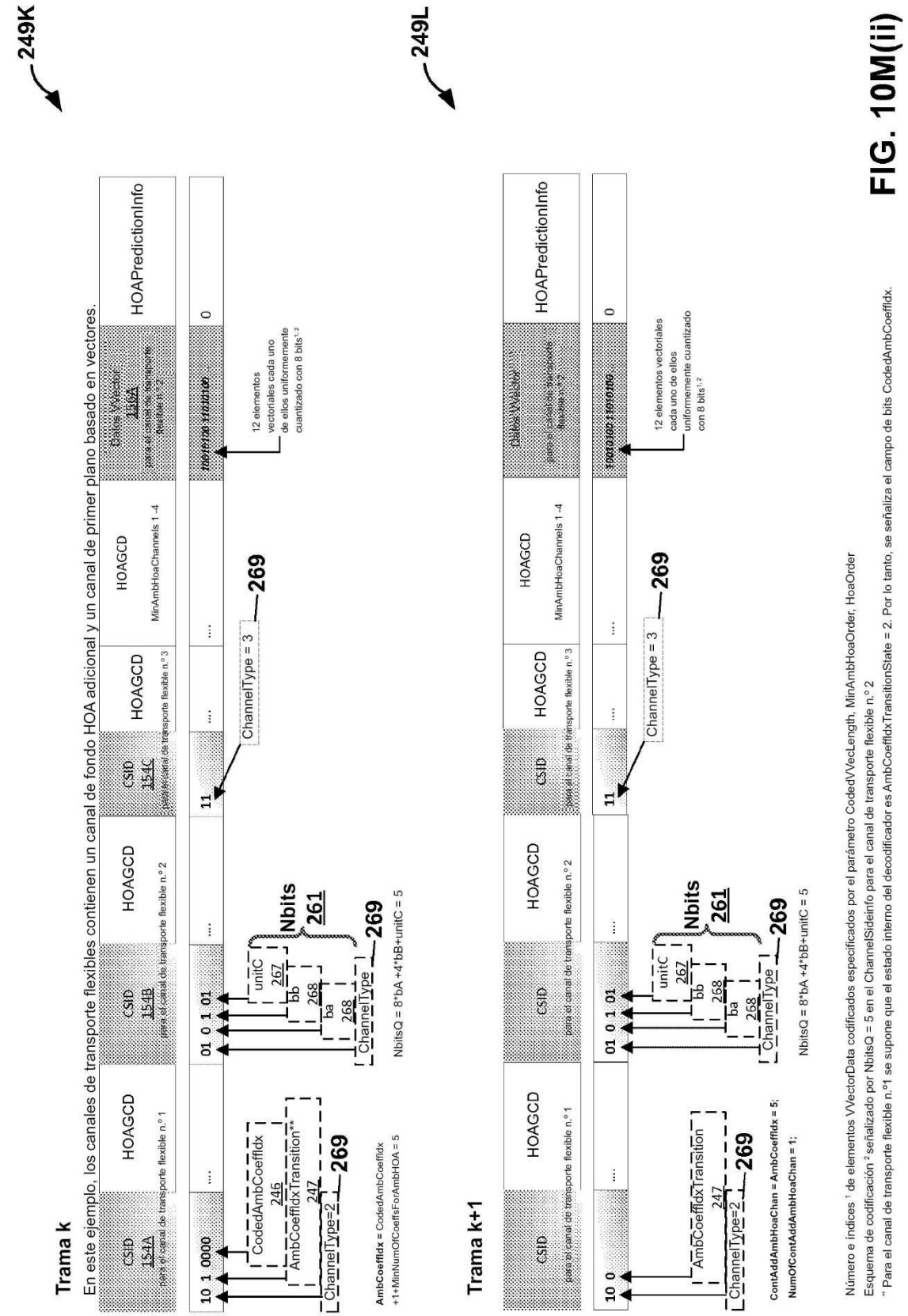

249K
Trama k
En este ejemplo, los canales de transporte flexibles contienen un canal de fondo HOA adicional y un canal de primer plano basado en vectores.
CSID 154A para el canal de transporte flexible n.º 1
HOAGCD
CSID 154B para el canal de transporte flexible n.º 2
HOAGCD
CSID 154C para el canal de transporte flexible n.º 3
HOAGCD MinAmbHoaChannels 1 -4
Datos VVector 156A para el canal de transporte flexible nº 2
HOAPredictionInfo
10 1 0000
....
01 0 1 01
....
11
....
....
0
CodedAmbCoeffIdx
246
AmbCoeffIdxTransition**
247
ChannelType=2
269
unitC
267
bb
268
ba
268
ChannelType
269
Nbits
261
ChannelType = 3
269
12 elementos vectoriales cada uno de ellos uniformemente cuantizado con 8 bits*,2
AmbCoeffIdx = CodedAmbCoeffIdx +1+MinNumOfCoeffsForAmbHOA = 5
NbitsQ = 8*bA +4*bB+unitC = 5
249L
Trama k+1
CSID para el canal de transporte flexible n.º 1
HOAGCD
CSID para el canal de transporte flexible n.º 2
HOAGCD
CSID para el canal de transporte flexible n.º 3
HOAGCD MinAmbHoaChannels 1 -4
Datos VVector para el canal de transporte flexible nº 2
HOAPredictionInfo
10 0
....
01 0 1 01
....
11
....
....
0
AmbCoeffIdxTransition
247
ChannelType=2
269
unitC
267
bb
268
ba
268
ChannelType
269
Nbits
261
ChannelType = 3
269
12 elementos vectoriales cada uno de ellos uniformemente cuantizado con 8 bits*,2
ContAddAmbHoaChan = AmbCoeffIdx = 5;
NumOfContAddAmbHoaChan = 1;
NbitsQ = 8*bA +4*bB+unitC = 5
Número e índices * de elementos VVectorData codificados especificados por el parámetro CodedVVecLength, MinAmbHoaOrder, HoaOrder
Esquema de codificación ² señalizado por NbitsQ = 5 en el ChannelSideinfo para el canal de transporte flexible n.º 2
** Para el canal de transporte flexible n.º1 se supone que el estado interno del decodificador es AmbCoeffIdxTransitionState = 2. Por lo tanto, se señaliza el campo de bits CodedAmbCoeffIdx.

FIG. 10M(ii)

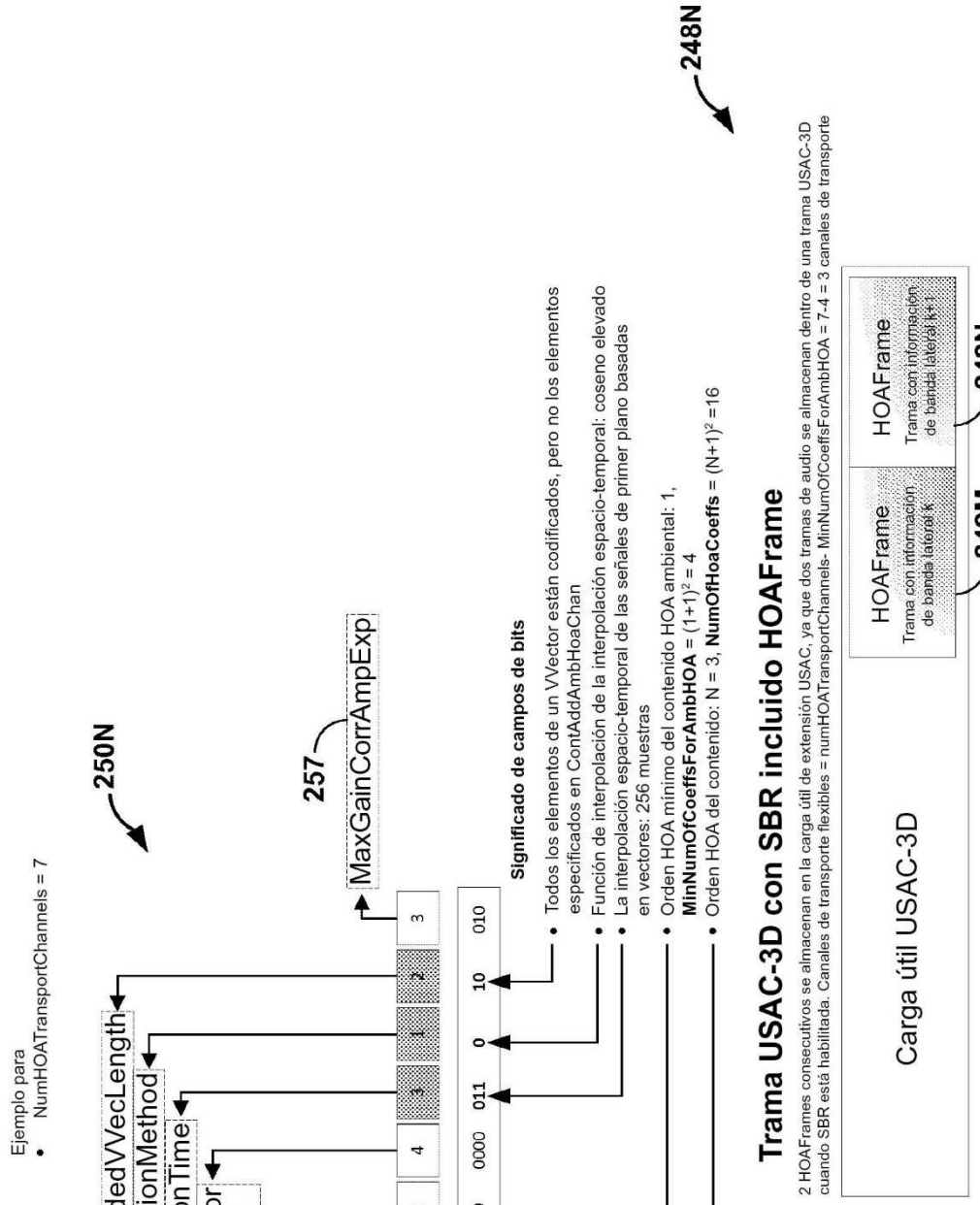
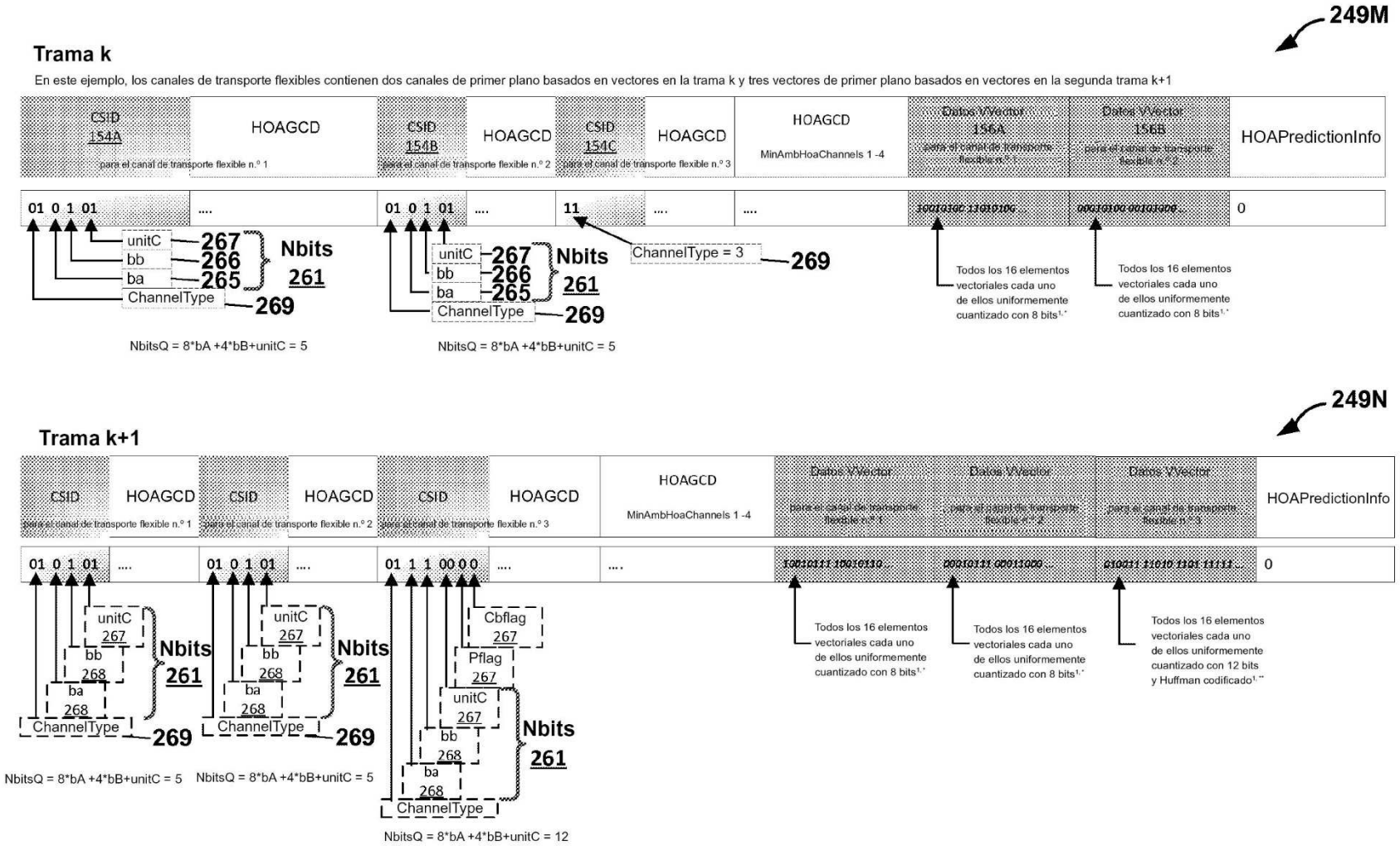

Flujo de bits de HOAConfig
Ejemplo para
• NumHOATransportChannels = 7
250N
256
255
254
253
150
152
CodedVVecLength
SpatialInterpolationMethod
CodedSpatialInterpolationTime
NoOfBitsPerScalefactor
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
HoaOrder
257
MaxGainCorrAmpExp
Tamaños de campo de bits
3-8 1 32 3-8 2 4 3 1 2 3
Ejemplo de flujo de bits
010 0 010 00 0000 011 0 10 010
Significado de campos de bits
• Todos los elementos de un VVector están codificados, pero no los elementos especificados en ContAddAmbHoaChan
• Función de interpolación de la interpolación espacio-temporal: coseno elevado
• La interpolación espacio-temporal de las señales de primer plano basadas en vectores: 256 muestras
• Orden HOA mínimo del contenido HOA ambiental: 1, MinNumOfCoeffsForAmbHOA = (1+1)² = 4
• Orden HOA del contenido: N = 3, NumOfHoaCoeffs = (N+1)² =16
248N
Trama USAC-3D con SBR incluido HOAFrame
2 HOAFrames consecutivos se almacenan en la carga útil de extensión USAC, ya que dos tramas de audio se almacenan dentro de una trama USAC-3D cuando SBR está habilitada. Canales de transporte flexibles = numHOATransportChannels- MinNumOfCoeffsForAmbHOA = 7-4 = 3 canales de transporte
Carga útil USAC-3D
HOAFrame
Trama con información de banda lateral k
HOAFrame
Trama con información de banda lateral k+1
249M
249N

FIG. 10N(i)



Número e índices de [1]los elementos VVectorData codificados especificados por el parámetro CodedVVecLength, MinAmbHoaOrder, HoaOrder, no se señalizan canales HOA adicionales en estas tramas.

* Esquema de codificación señalizado por NbitsQ = 5 en el ChannelSideInfo para este canal de transporte en particular

** Esquema de codificación de Huffman señalizado por NbitsQ = 12, CbFlag = 0 y Pflag = 0 en el ChannelSideInfo para este canal de transporte particular

FIG. 10N(ii)

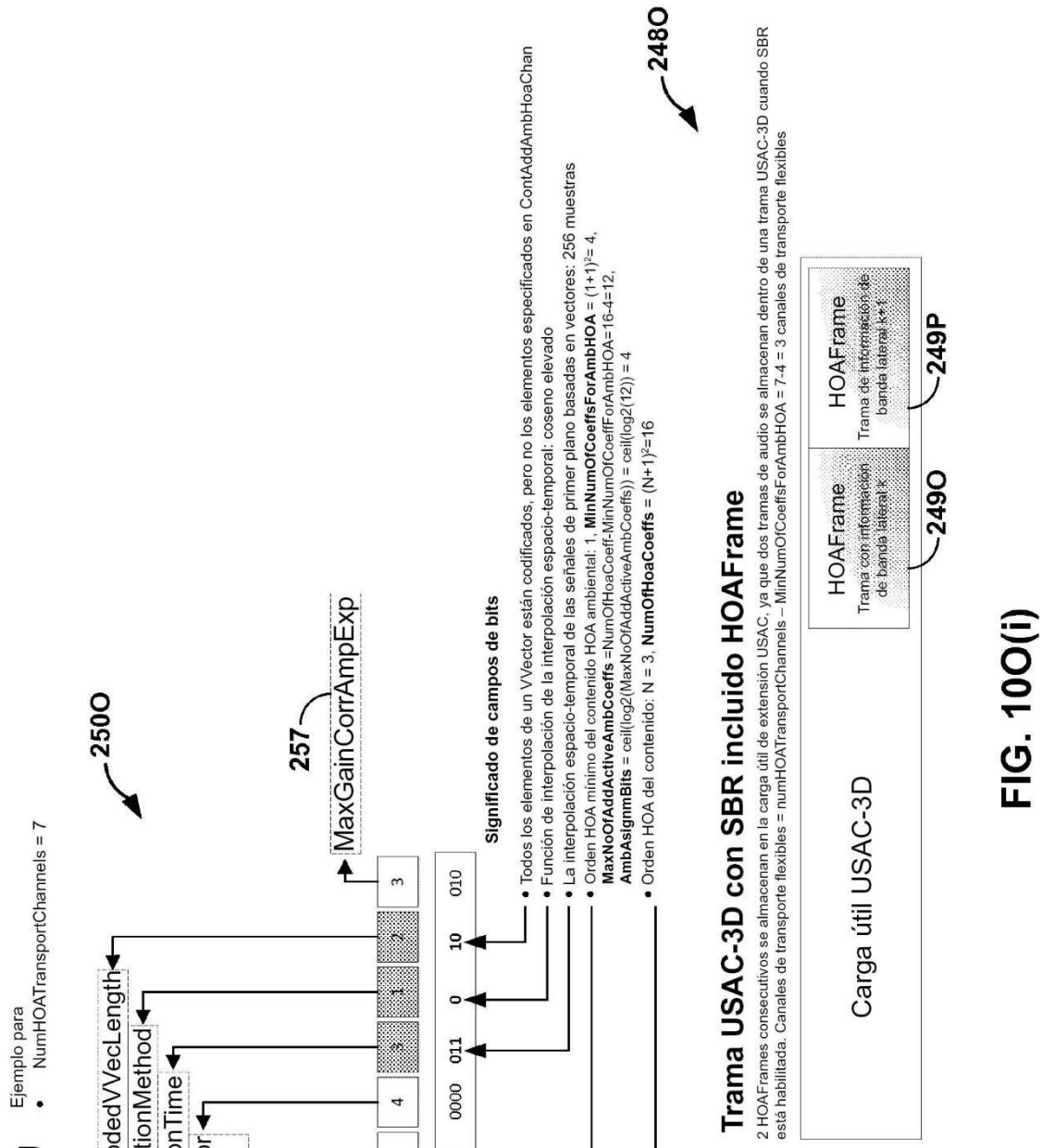
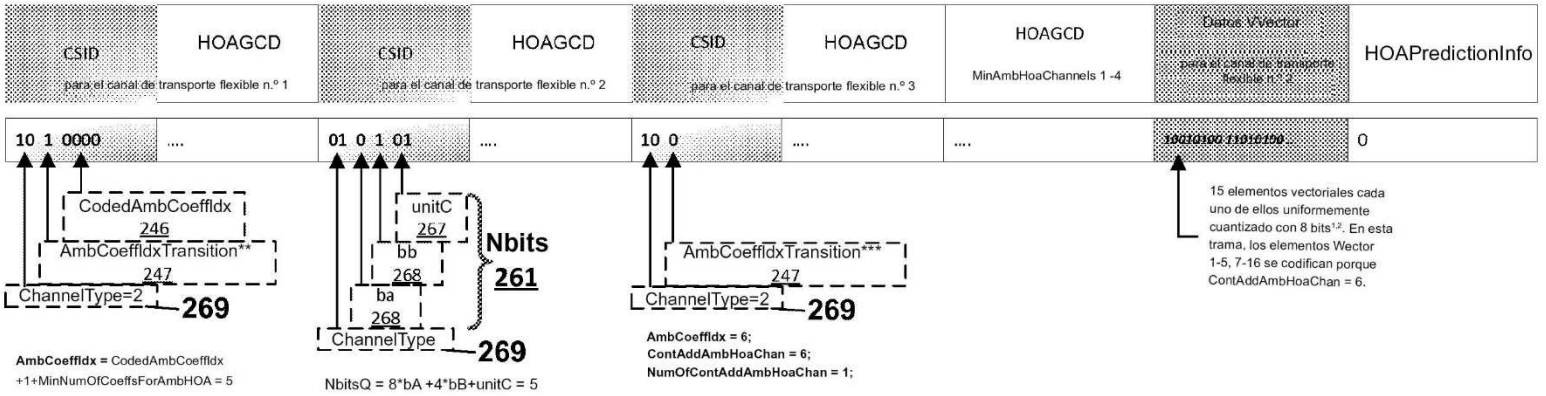

Flujo de bits de HOAConfig
Ejemplo para
• NumHOATransportChannels = 7
250O
256
255
254
253
150
152
CodedVVecLength
SpatialInterpolationMethod
CodedSpatialInterpolationTime
NoOfBitsPerScalefactor
MaxNoOfDirSigsForPrediction
MinAmbHoaOrder
NfcReferenceDistance
UsesNfc
HoaOrder
257
MaxGainCorrAmpExp
Tamaños de campo de bits
3-8 1 32 3-8 2 4 3 1 2 3
Ejemplo de flujo de
010 0 010 00 0000 011 0 10 010
Significado de campos de bits
• Todos los elementos de un VVector están codificados, pero no los elementos especificados en ContAddAmbHoaChan
• Función de interpolación de la interpolación espacio-temporal: coseno elevado
• La interpolación espacio-temporal de las señales de primer plano basadas en vectores: 256 muestras
• Orden HOA mínimo del contenido HOA ambiental: 1, MinNumOfCoeffsForAmbHOA = (1+1)²= 4,
MaxNoOfAddActiveAmbCoeffs =NumOfHoaCoeff-MinNumOfCoeffForAmbHOA=16-4=12,
AmbAsIgnmBits = ceil(log2(MaxNoOfAddActiveAmbCoeffs)) = ceil(log2(12)) = 4
• Orden HOA del contenido: N = 3, NumOfHoaCoeffs = (N+1)²=16
248O
Trama USAC-3D con SBR incluido HOAFrame
2 HOAFrames consecutivos se almacenan en la carga útil de extensión USAC, ya que dos tramas de audio se almacenan dentro de una trama USAC-3D cuando SBR está habilitada. Canales de transporte flexibles = numHOATransportChannels – MinNumOfCoeffsForAmbHOA = 7-4 = 3 canales de transporte flexibles
Carga útil USAC-3D
HOAFrame
Trama con información de banda lateral k
249O
HOAFrame
Trama de información de banda lateral k+1
249P

FIG. 10O(i)

249O

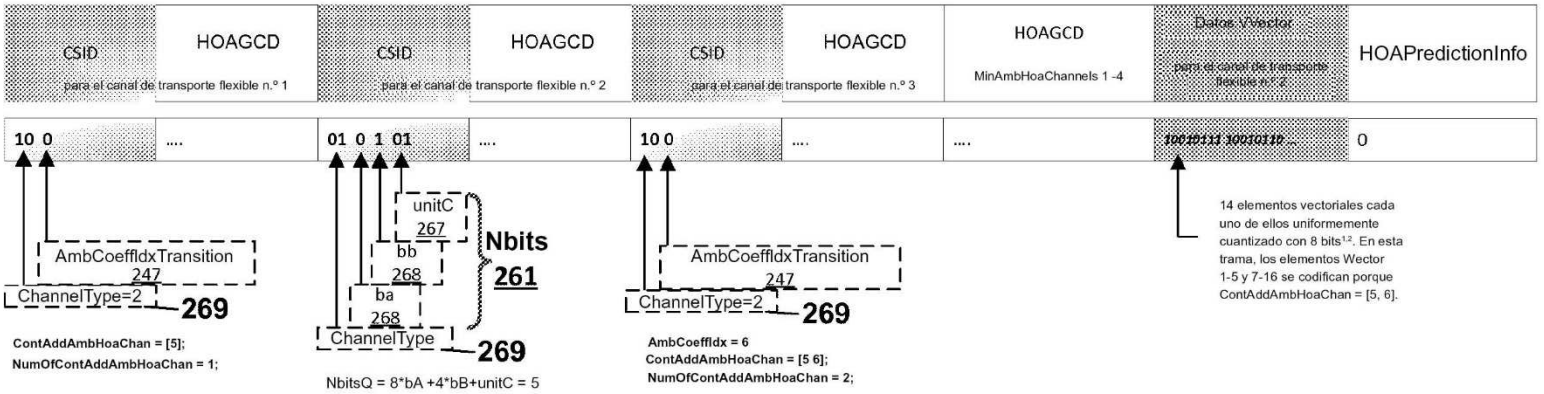
Trama k

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



249P

Trama k+1



Número e índices de [1] los elementos VVectorData codificados especificados por el parámetro CodedVVecLength, MinNumOfCoeffsForAmbHOA, NumOfHoaCoeff

Esquema de codificación [2] señalizado por NbitsQ = 5 en el ChannelSideinfo para el canal de transporte flexible n.º 2

** Para el canal de transporte flexible n.º1 se supone que el estado interno del decodificador AmbCoeff IdxTransitionState = 2. Por lo tanto, se señaliza el campo de bits CodedAmbCoeffIdx.

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

FIG. 10O(ii)

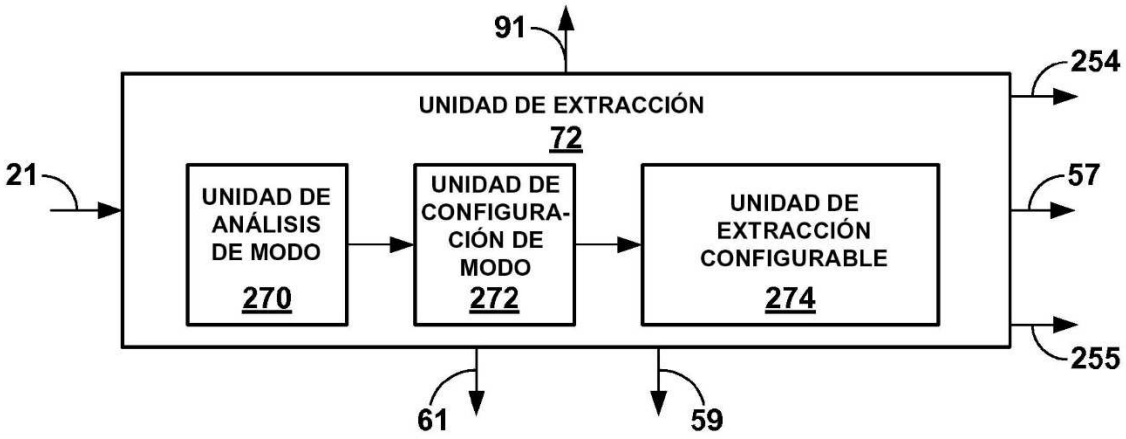


FIG. 11A

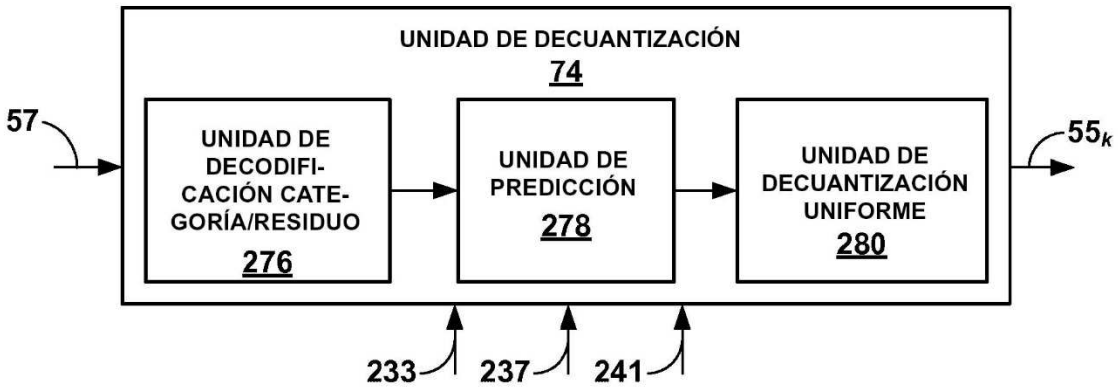


FIG. 11B



FIG. 11C



FIG. 11D

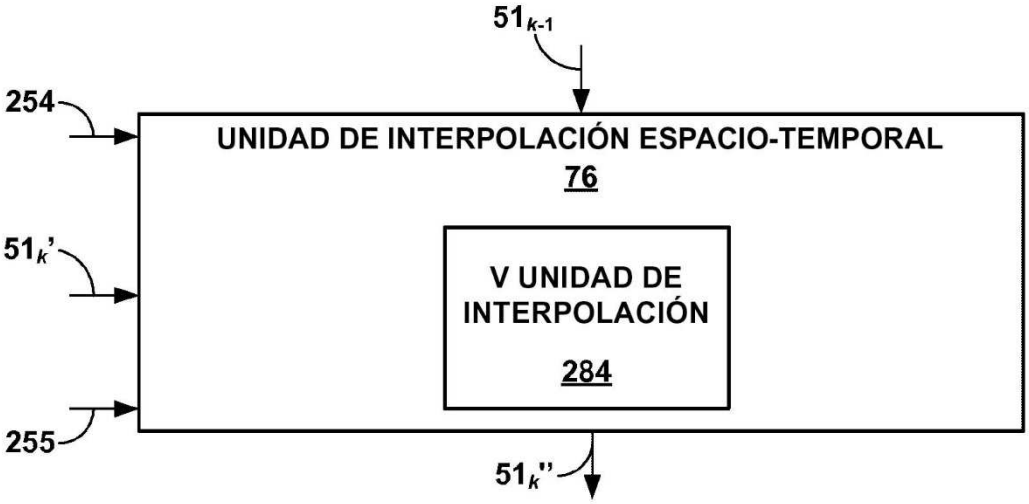


FIG. 11E

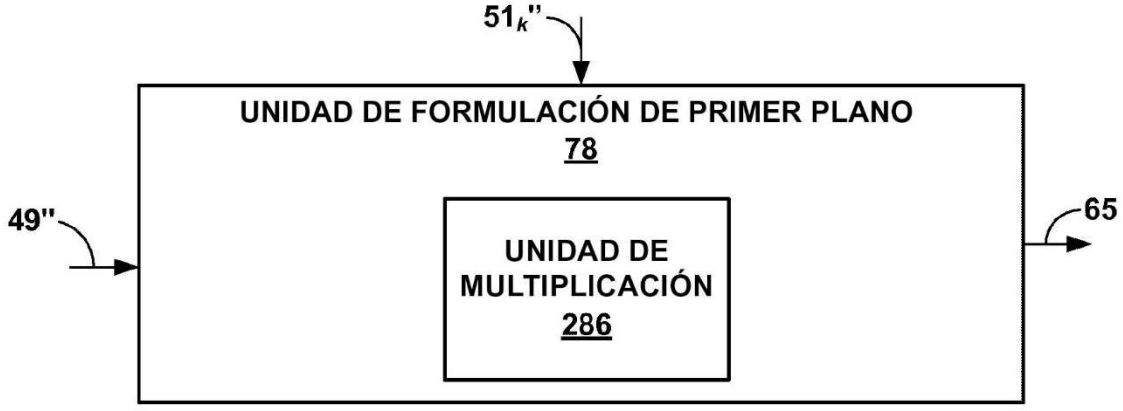


FIG. 11F

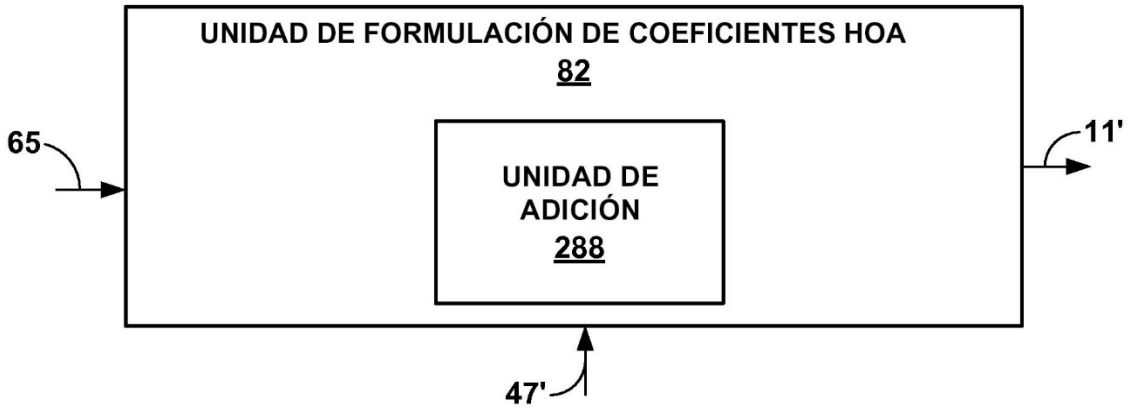

UNIDAD DE FORMULACIÓN DE COEFICIENTES HOA
82
UNIDAD DE ADICIÓN
288
65
11'
47'

FIG. 11G

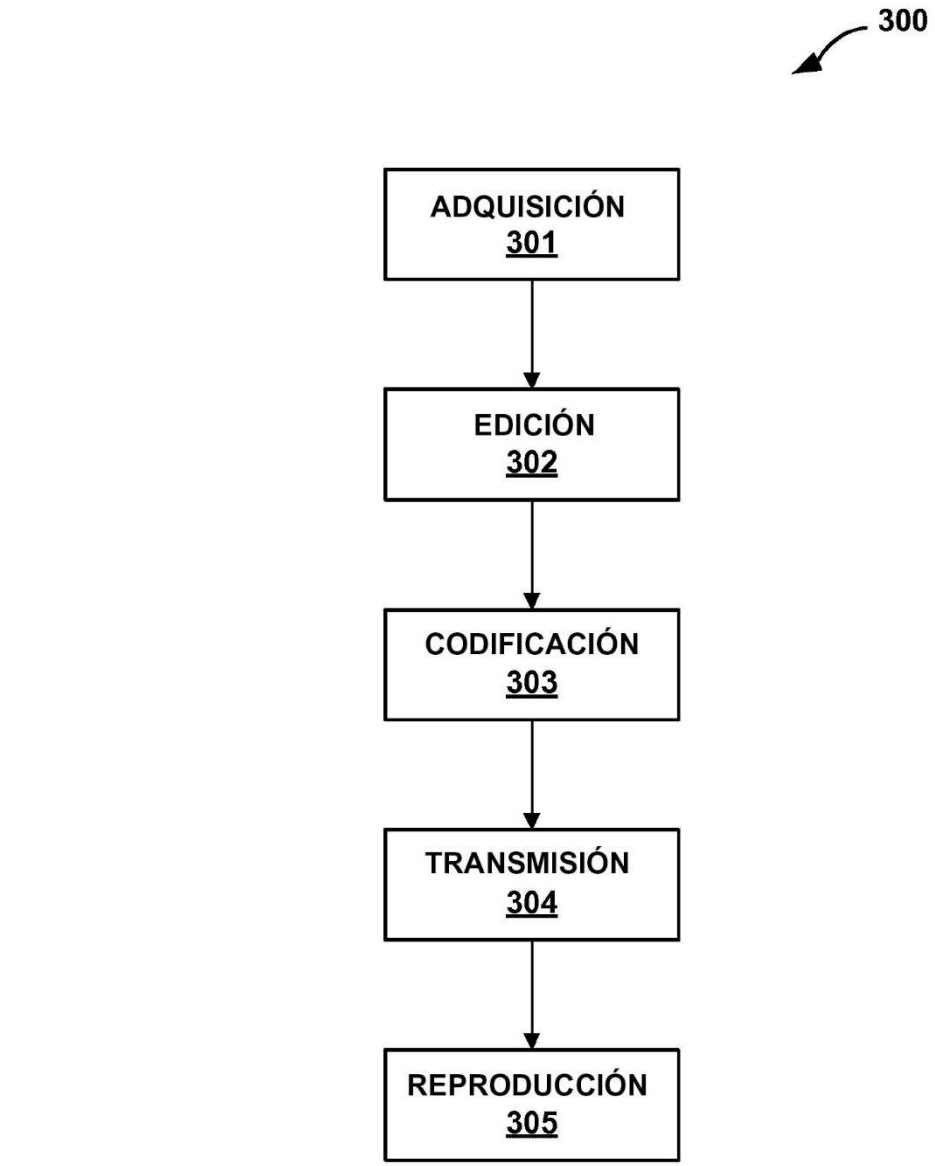

300
ADQUISICIÓN
301
EDICIÓN
302
CODIFICACIÓN
303
TRANSMISIÓN
304
REPRODUCCIÓN
305

FIG. 12

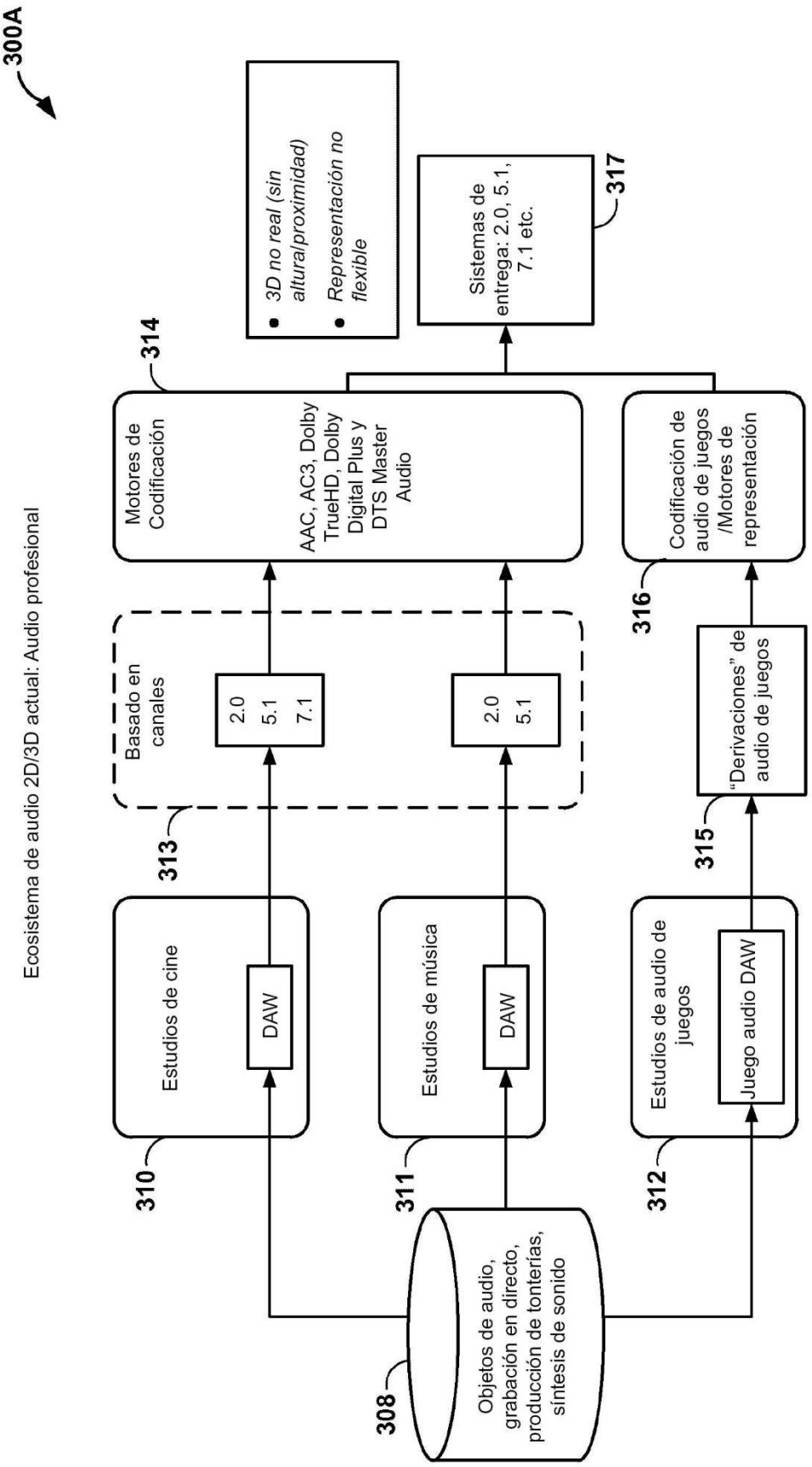

300A
Ecosistema de audio 2D/3D actual: Audio profesional
Objetos de audio, grabación en directo, producción de tonterías, síntesis de sonido
308
Estudios de cine
DAW
310
Estudios de música
DAW
311
Estudios de audio de juegos
Juego audio DAW
312
Basado en canales
2.0
5.1
7.1
2.0
5.1
313
"Derivaciones" de audio de juegos
315
Codificación de audio de juegos /Motores de representación
316
Motores de Codificación
AAC, AC3, Dolby TrueHD, Dolby Digital Plus y DTS Master Audio
314
Sistemas de entrega: 2.0, 5.1, 7.1 etc.
317
3D no real (sin altura/proximidad)
Representación no flexible

FIG. 13

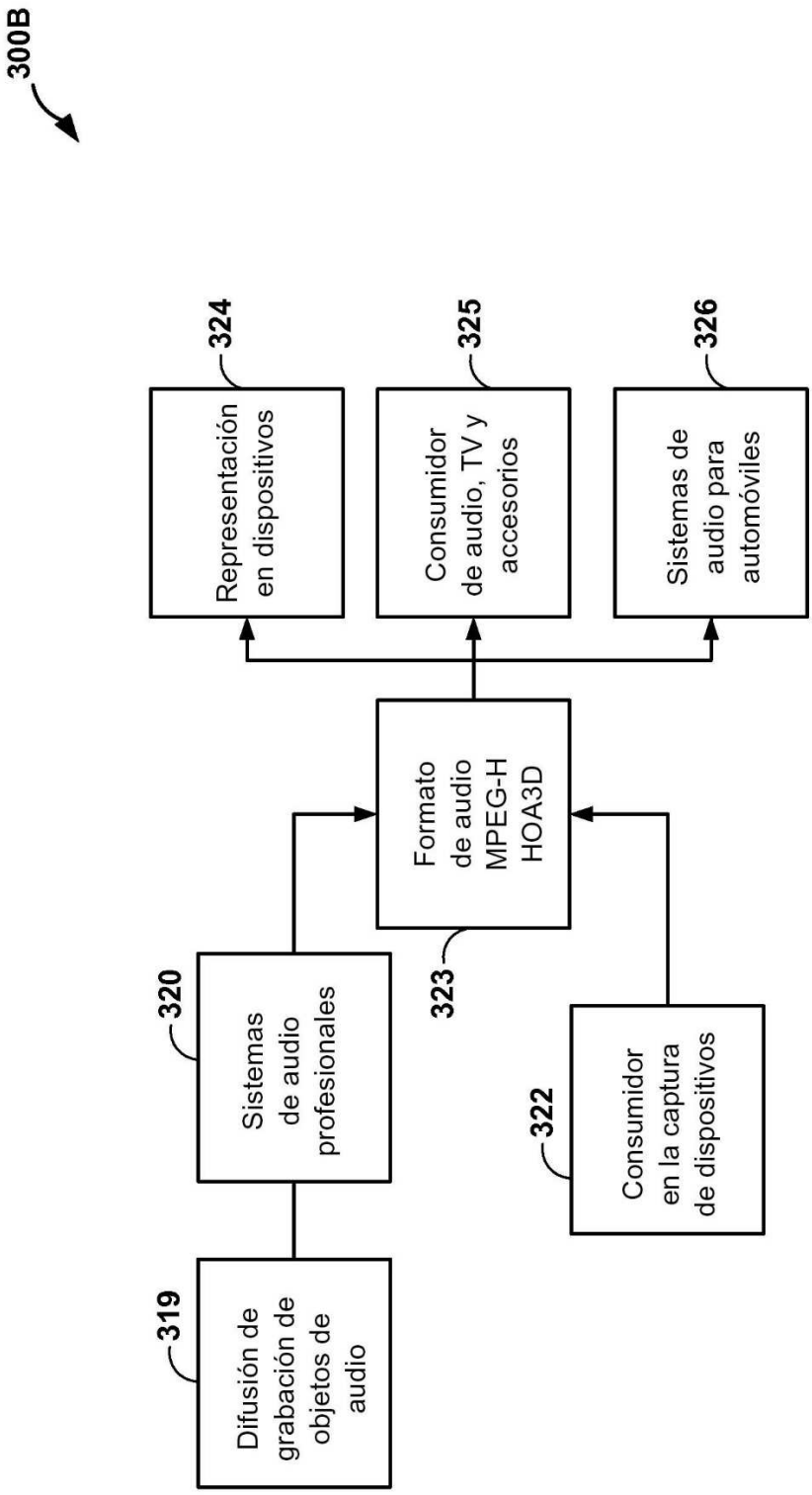

300B
319
Difusión de grabación de objetos de audio
320
Sistemas de audio profesionales
322
Consumidor en la captura de dispositivos
323
Formato de audio MPEG-H HOA3D
324
Representación en dispositivos
325
Consumidor de audio, TV y accesorios
326
Sistemas de audio para automóviles

FIG. 14

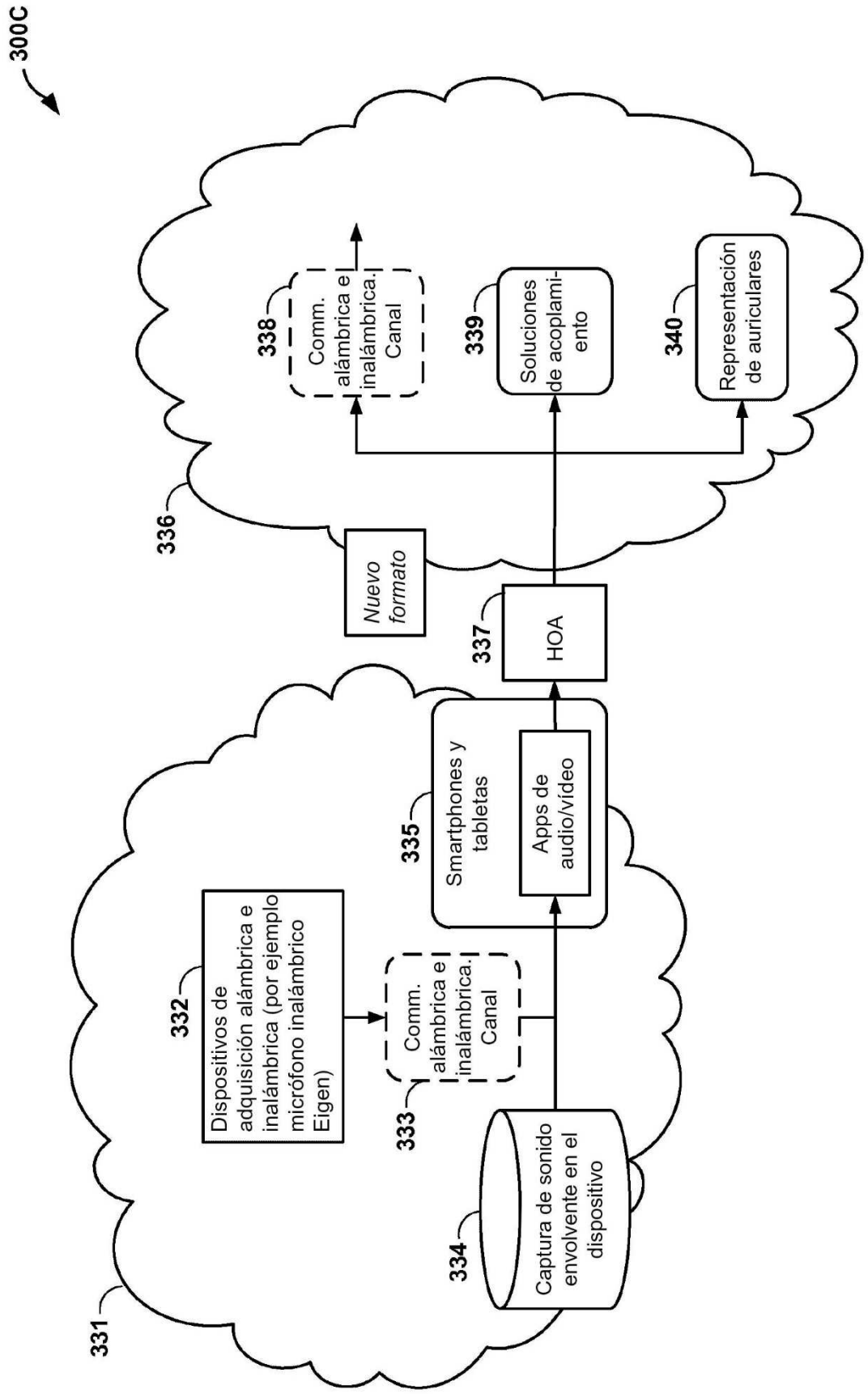

300C
331
Captura de sonido envolvente en el dispositivo
334
Comm. alámbrica e inalámbrica. Canal
333
Dispositivos de adquisición alámbrica e inalámbrica (por ejemplo micrófono inalámbrico Eigen)
332
Apps de audio/vídeo
Smartphones y tabletas
335
HOA
337
Nuevo formato
336
Comm. alámbrica e inalámbrica. Canal
338
Soluciones de acoplami-ento
339
Representación de auriculares
340

FIG. 15A

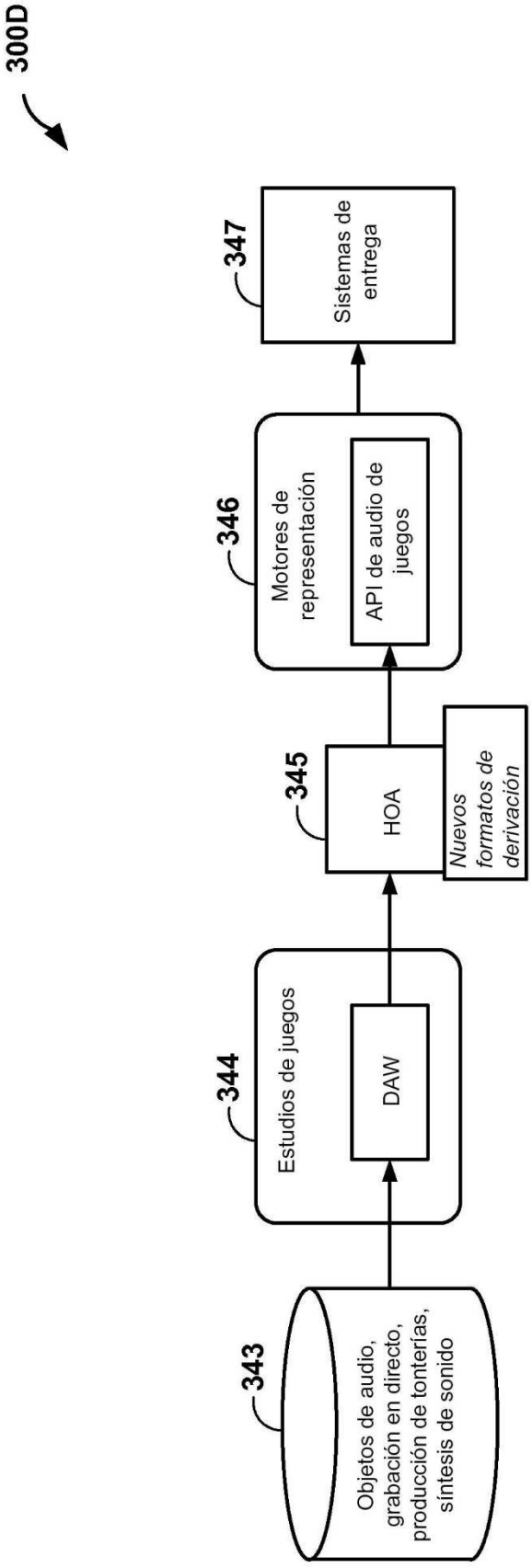

300D
343
Objetos de audio, grabación en directo, producción de tonterías, síntesis de sonido
344
Estudios de juegos
DAW
345
HOA
Nuevos formatos de derivación
346
Motores de representación
API de audio de juegos
347
Sistemas de entrega

FIG. 15B

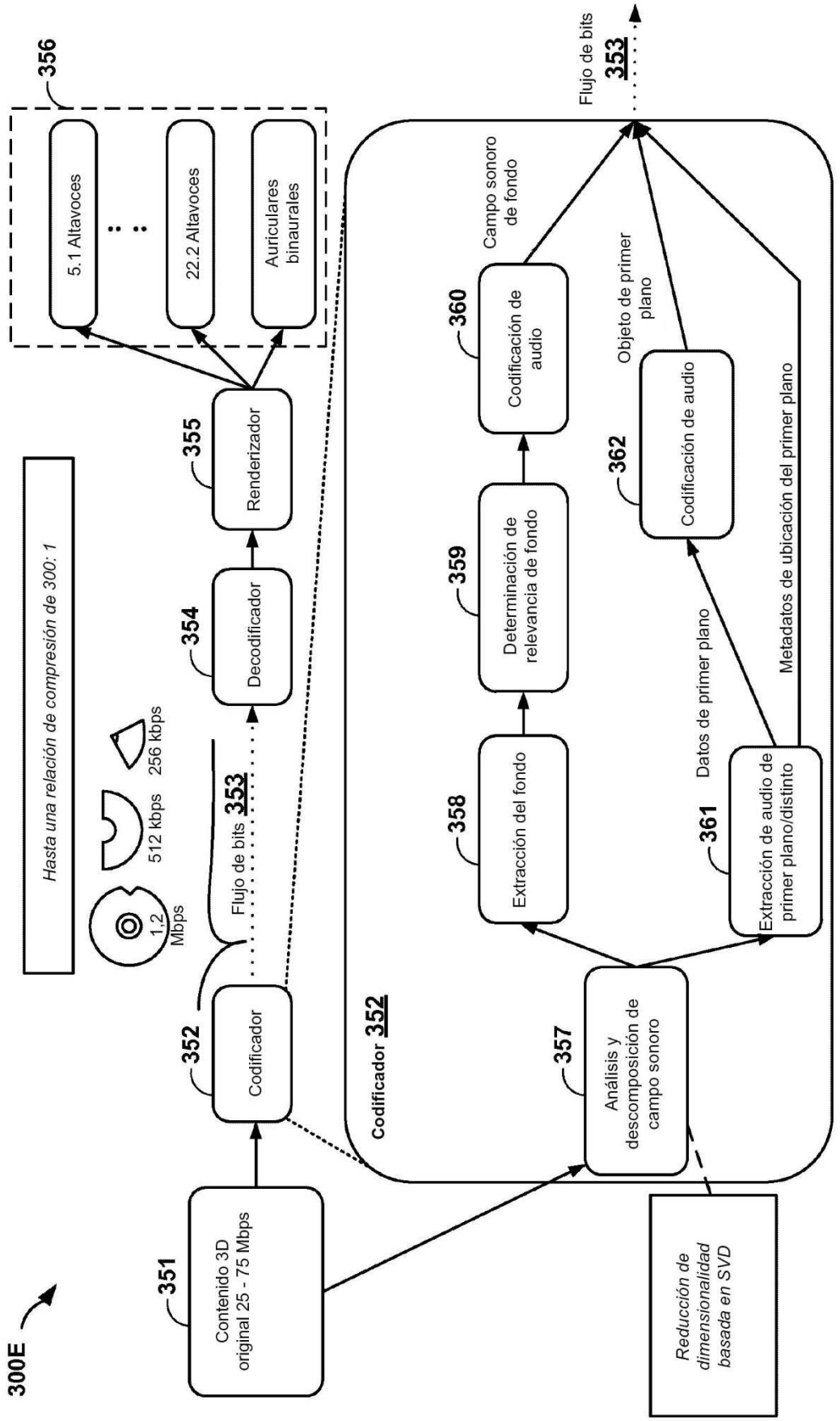

300E
351
Contenido 3D original 25 - 75 Mbps
352
Codificador
Flujo de bits 353
1,2 Mbps
512 kbps
256 kbps
Hasta una relación de compresión de 300: 1
354
Decodificador
355
Renderizador
356
5.1 Altavoces
22.2 Altavoces
Auriculares binaurales
Codificador 352
357
Análisis y descomposición de campo sonoro
358
Extracción del fondo
359
Determinación de relevancia de fondo
360
Codificación de audio
Campo sonoro de fondo
361
Extracción de audio de primer plano/distinto
Datos de primer plano
362
Codificación de audio
Objeto de primer plano
Metadatos de ubicación del primer plano
Flujo de bits 353
Reducción de dimensionalidad basada en SVD

FIG. 16

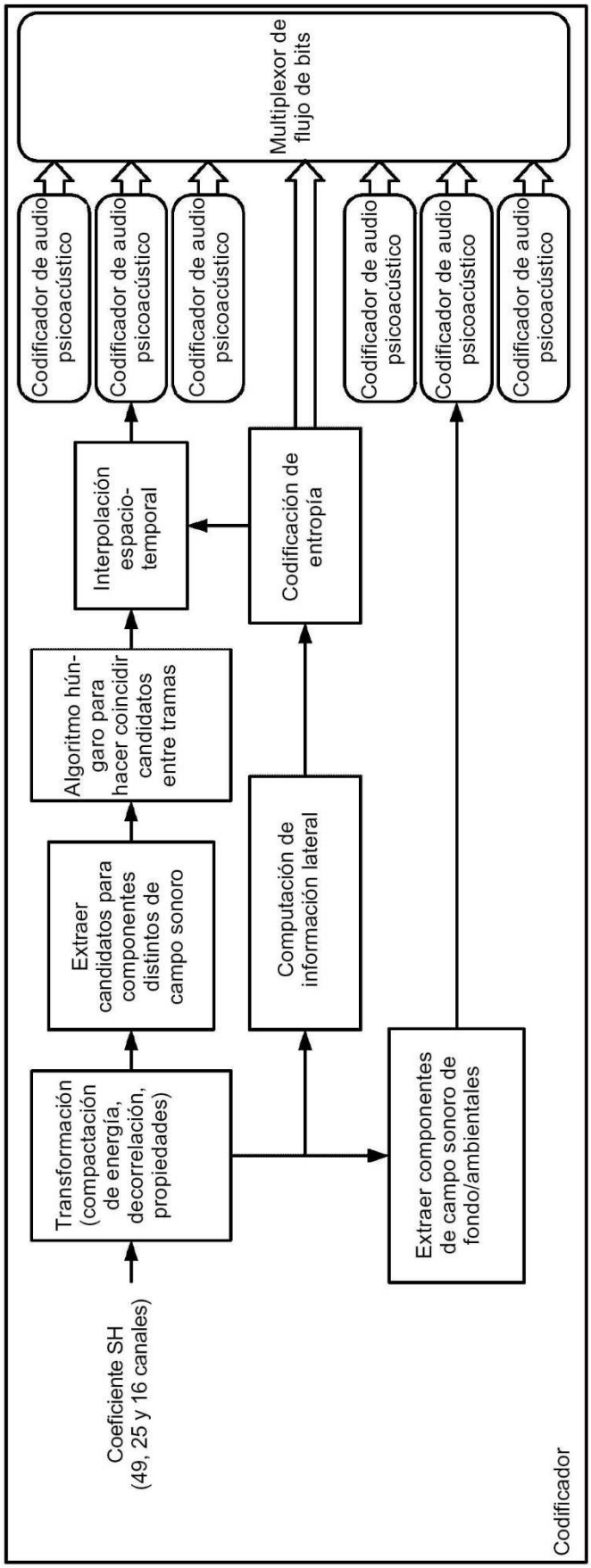

Codificador
Coeficiente SH
(49, 25 y 16 canales)
Transformación
(compactación
de energía,
decorrelación,
propiedades)
Extraer
candidatos para
componentes
distintos de
campo sonoro
Algoritmo hún-
garo para
hacer coincidir
candidatos
entre tramas
Interpolación
espacio-
temporal
Computación de
información lateral
Codificación de
entropía
Extraer componentes
de campo sonoro de
fondo/ambientales
Multiplexor de
flujo de bits
Codificador de audio
psicoacústico
Codificador de audio
psicoacústico
Codificador de audio
psicoacústico
Codificador de audio
psicoacústico
Codificador de audio
psicoacústico
Codificador de audio
psicoacústico

FIG. 17

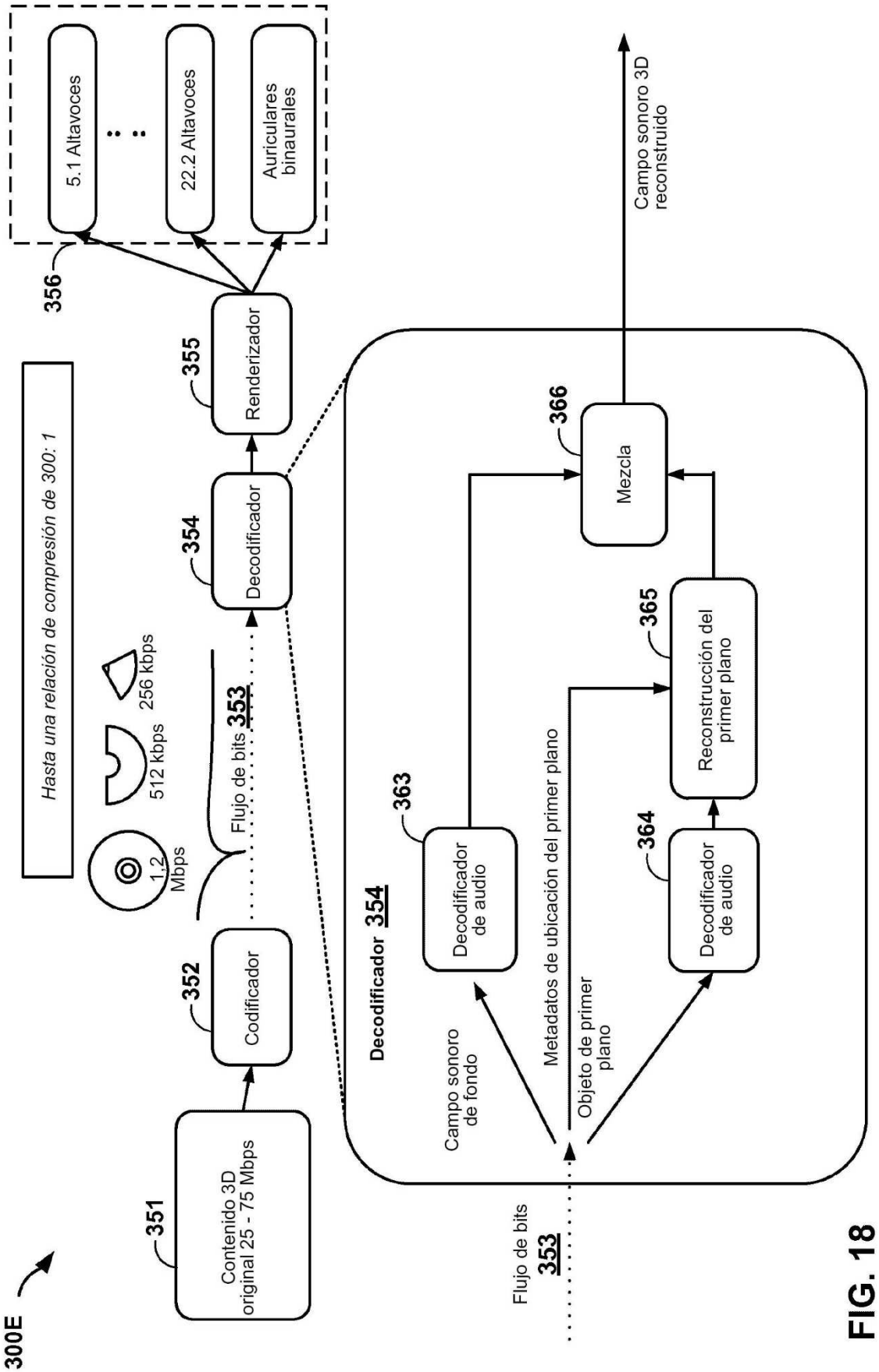

300E
351
Contenido 3D original 25 - 75 Mbps
352
Codificador
Hasta una relación de compresión de 300: 1
1.2 Mbps
512 kbps
256 kbps
Flujo de bits 353
354
Decodificador
355
Renderizador
356
5.1 Altavoces
22.2 Altavoces
Auriculares binaurales
Decodificador 354
363
Decodificador de audio
364
Decodificador de audio
365
Reconstrucción del primer plano
366
Mezcla
Campo sonoro de fondo
Metadatos de ubicación del primer plano
Objeto de primer plano
Flujo de bits 353
Campo sonoro 3D reconstruido

FIG. 18

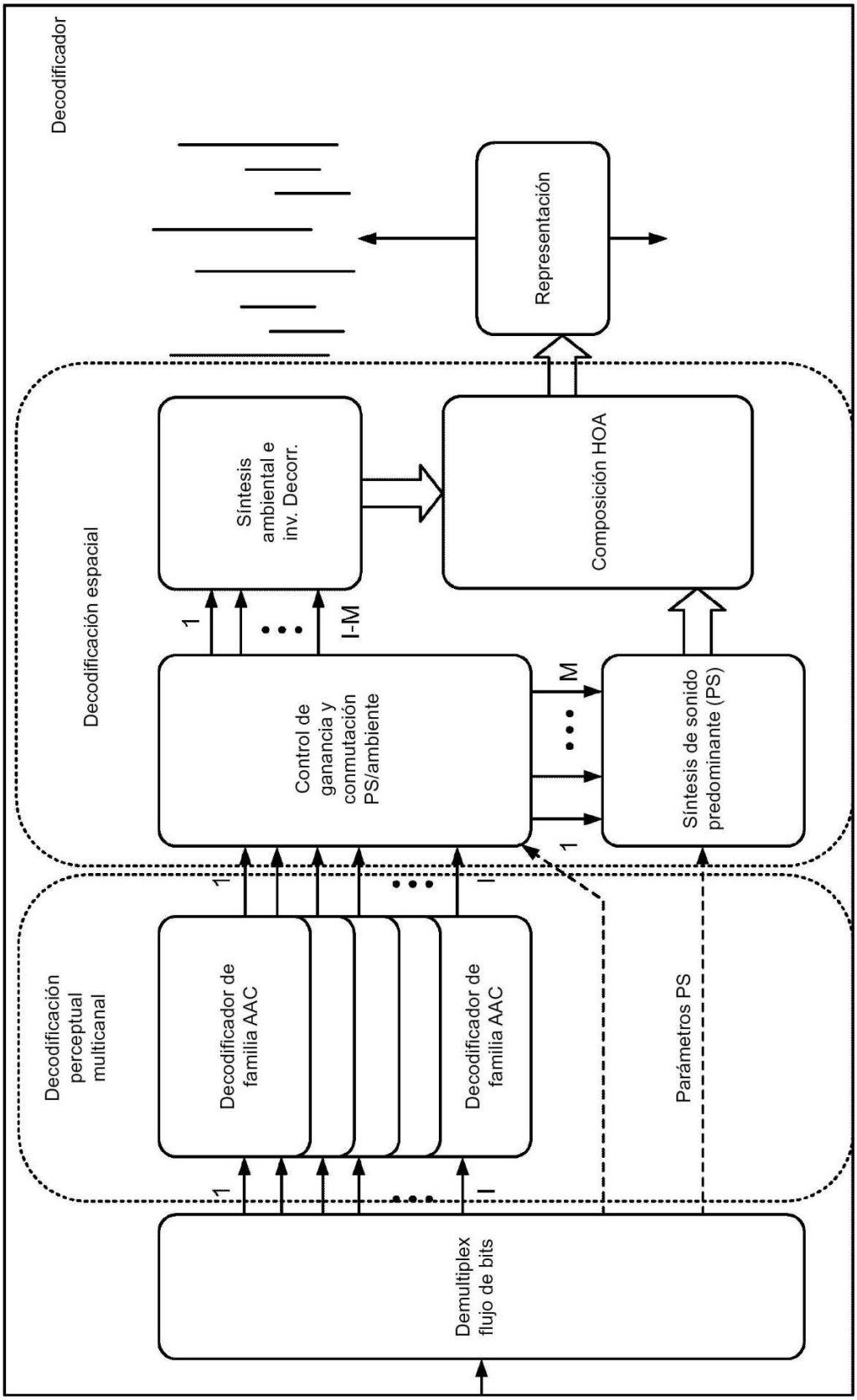

Decodificador
Representación
Síntesis ambiental e inv. Decorr.
Composición HOA
Decodificación espacial
1
I-M
Control de ganancia y conmutación PS/ambiente
M
Síntesis de sonido predominante (PS)
1
1
I
Decodificación perceptual multicanal
Decodificador de familia AAC
Decodificador de familia AAC
Parámetros PS
1
I
Demultiplex flujo de bits

FIG. 19

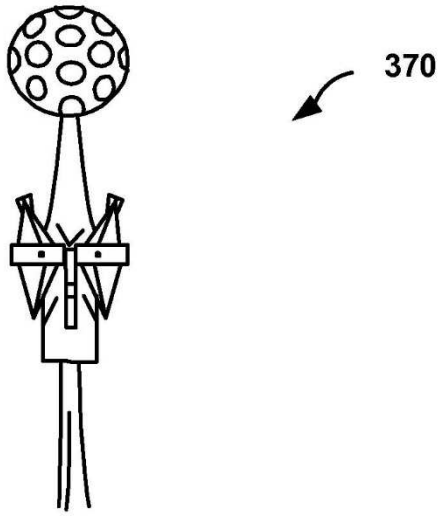


FIG. 20A

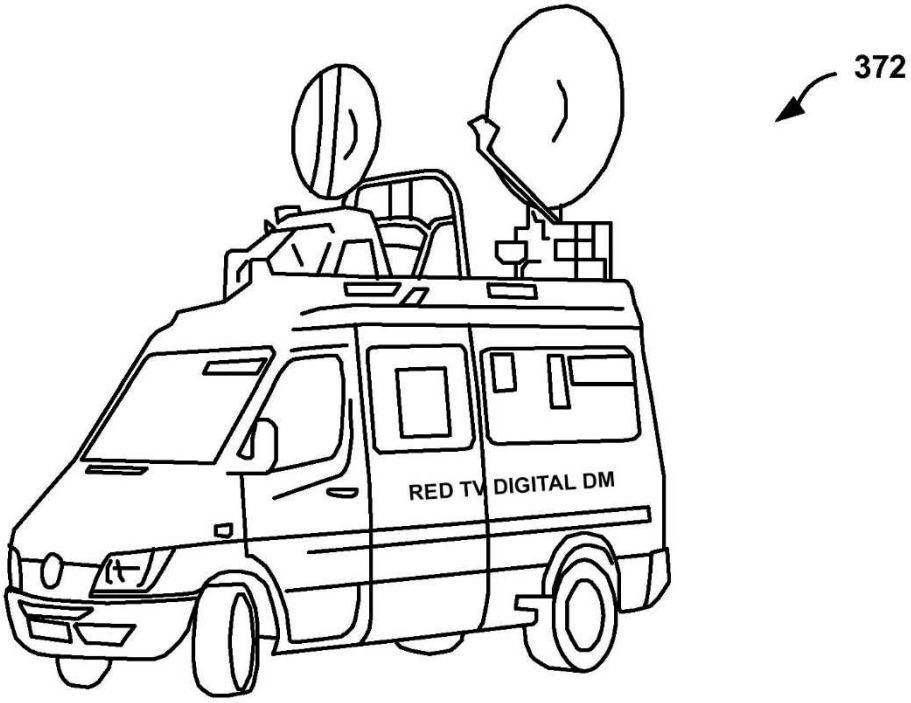


FIG. 20B

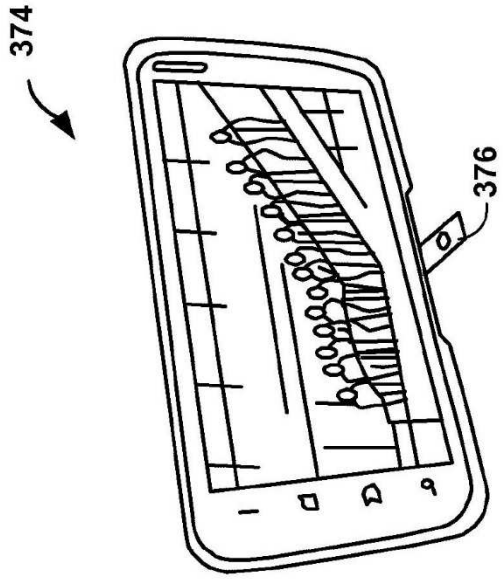


FIG. 20E

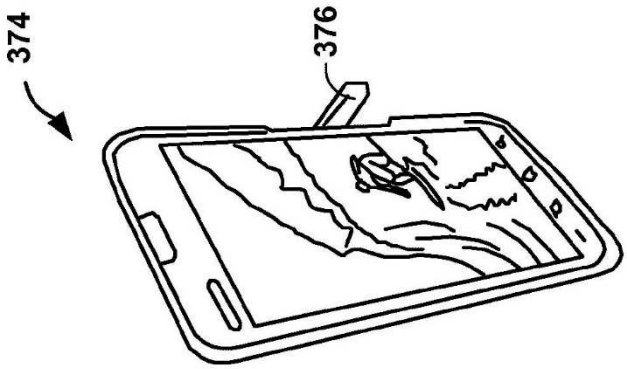


FIG. 20D

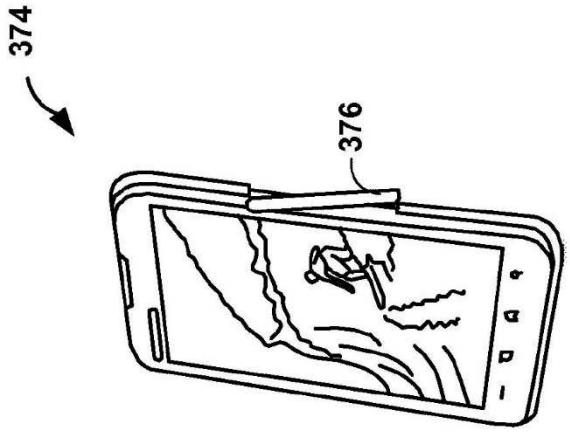


FIG. 20C

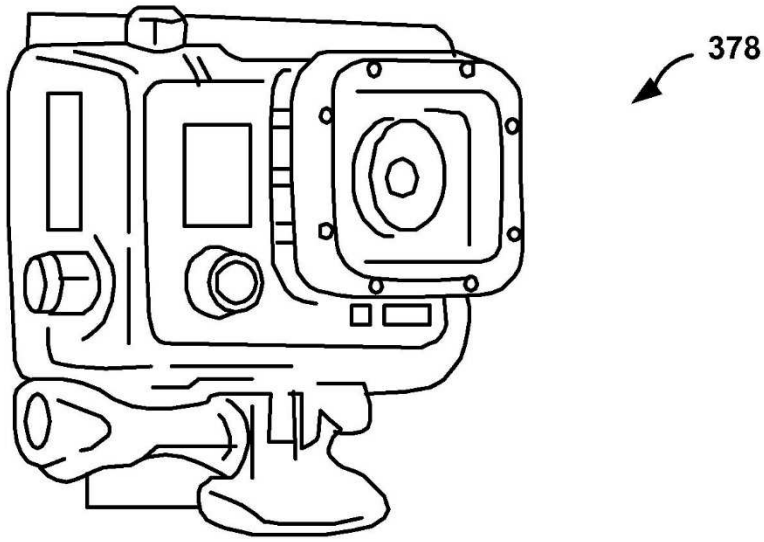


FIG. 20F

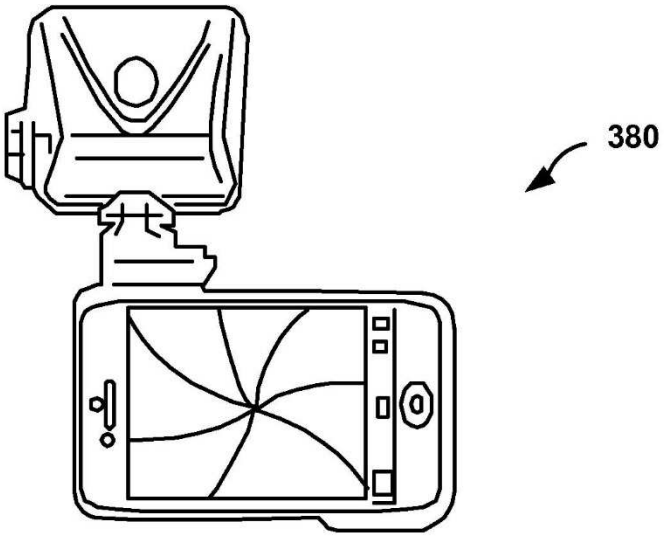


FIG. 20G

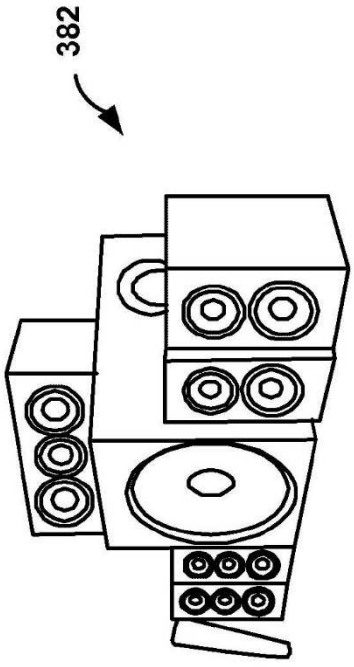


FIG. 21A

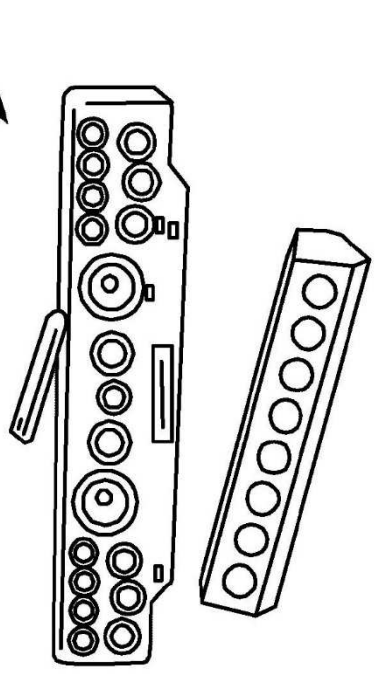


FIG. 21B

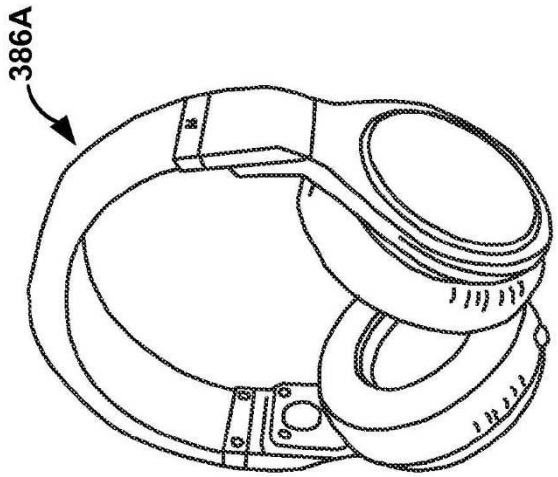


FIG. 21C

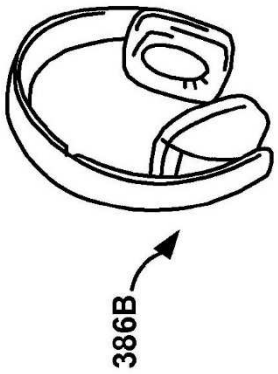


FIG. 21D

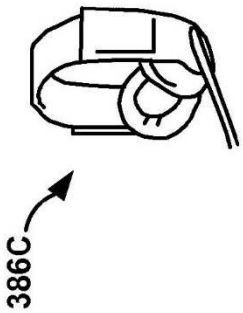


FIG. 21E

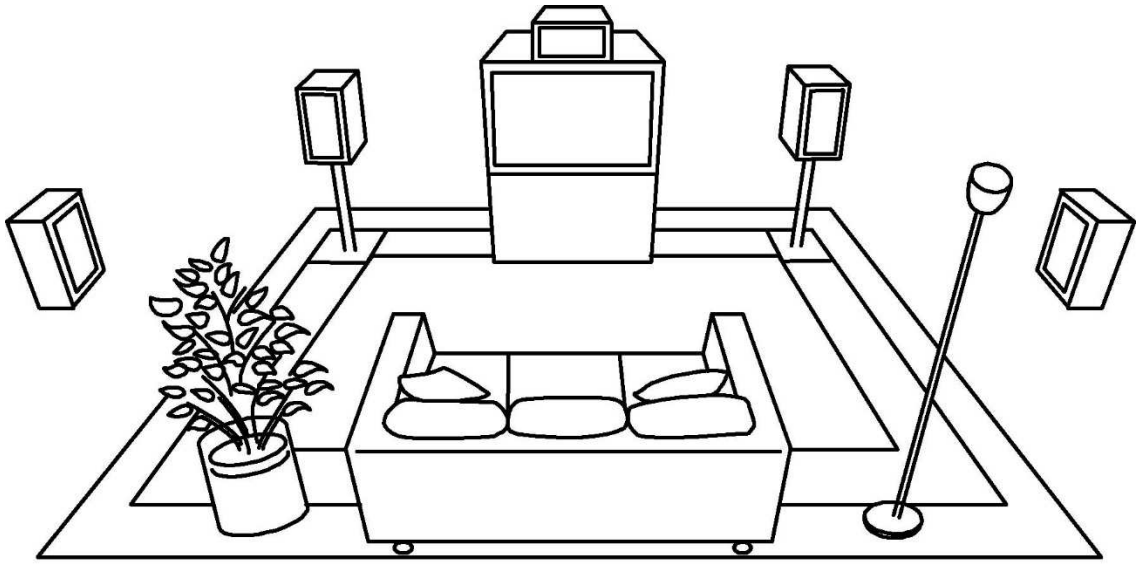


FIG. 22A

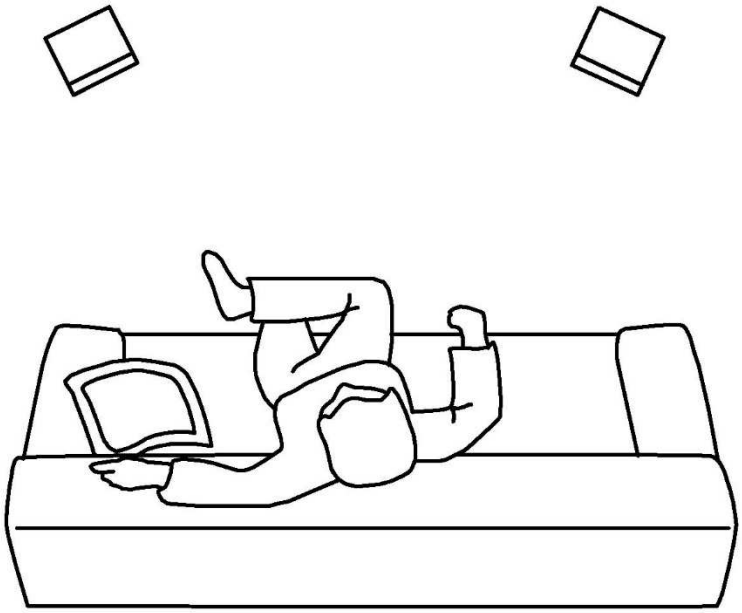
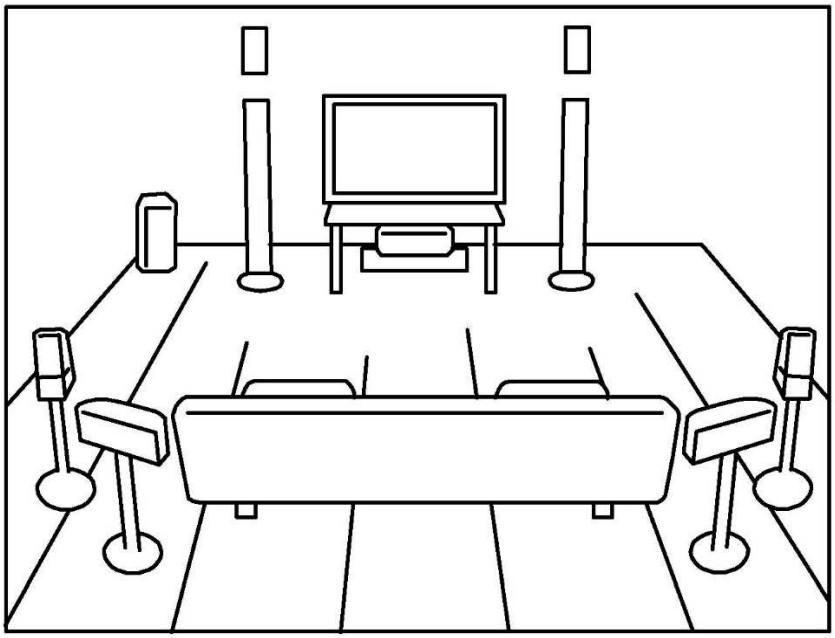


FIG. 22B



9.1 con altura completa

FIG. 22C

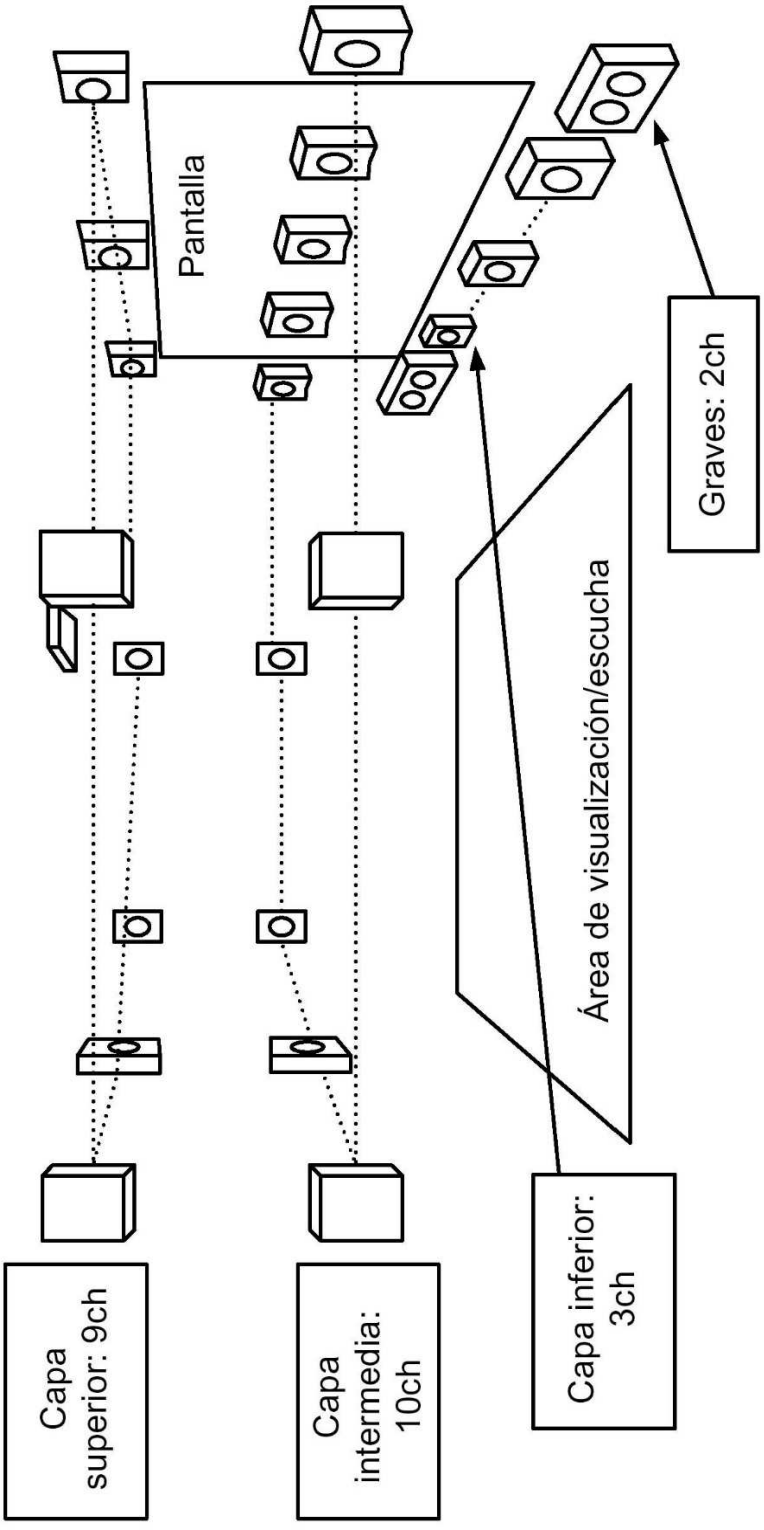

Pantalla
Graves: 2ch
Área de visualización/escucha
Capa inferior: 3ch
Capa intermedia: 10ch
Capa superior: 9ch
22.2

FIG. 22D

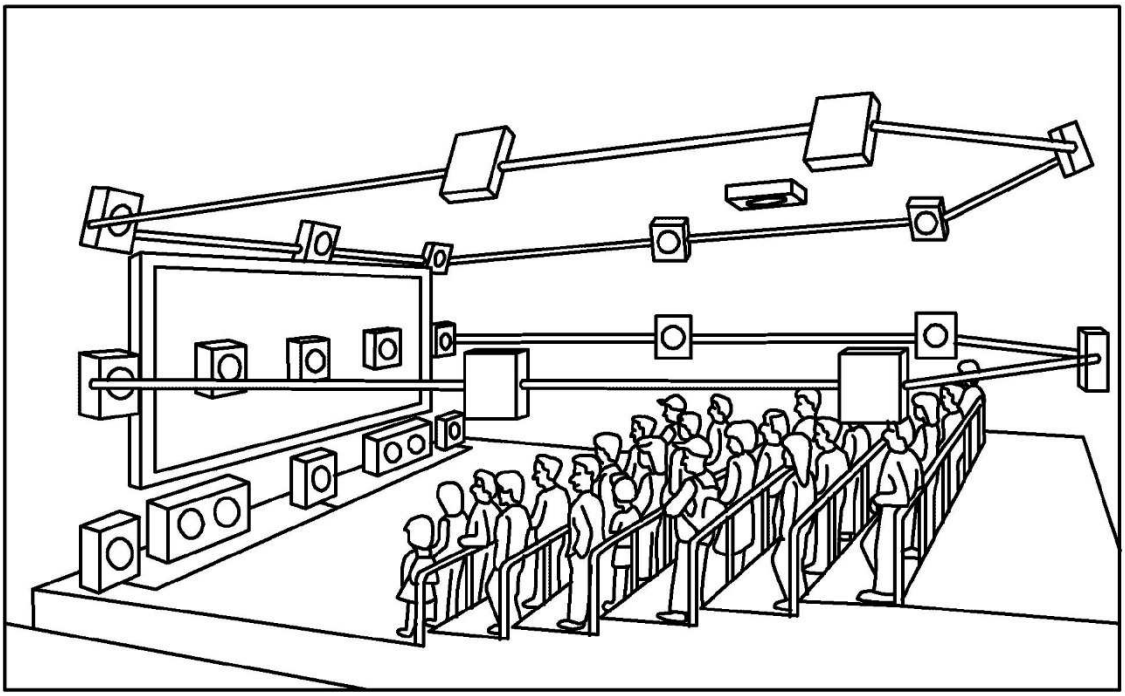


FIG. 22E

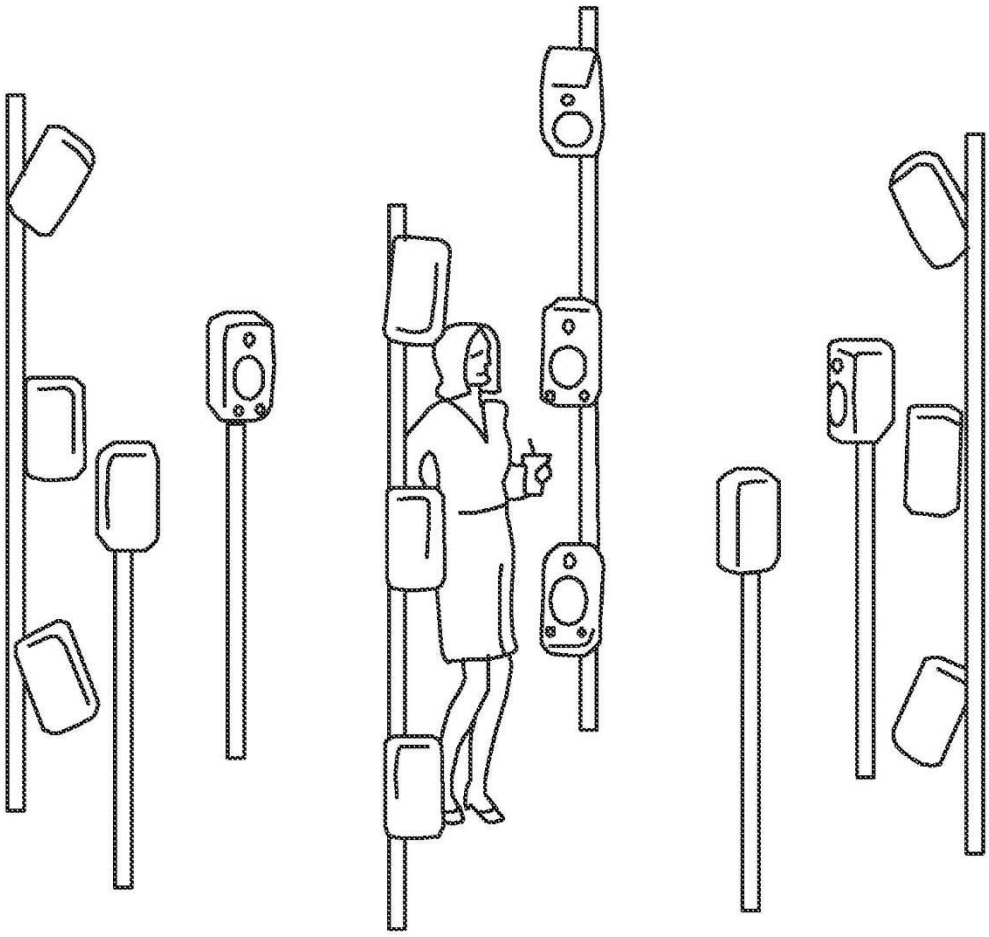


FIG. 22F

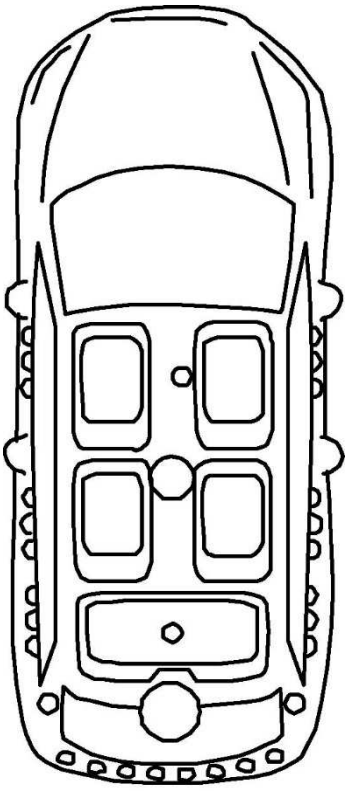


FIG. 22G

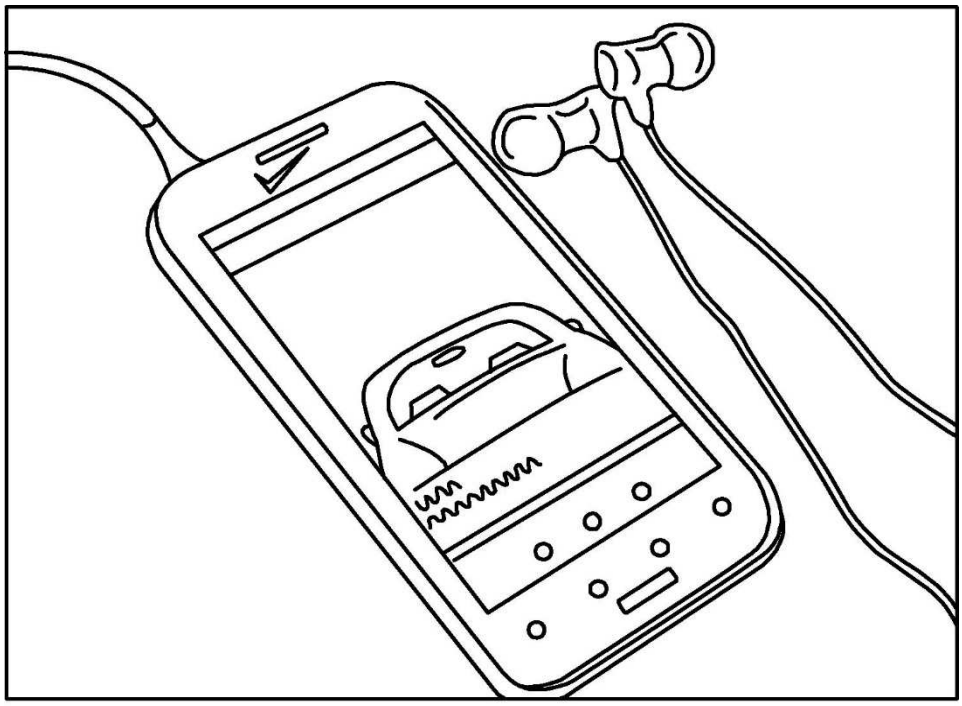


FIG. 22H

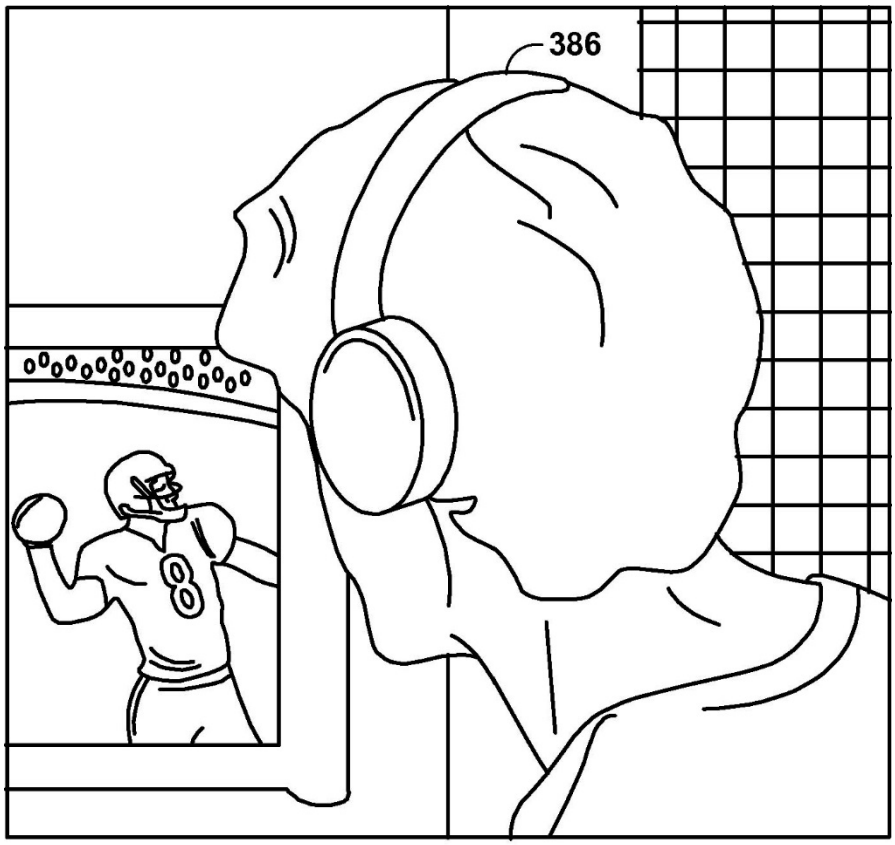


FIG. 23

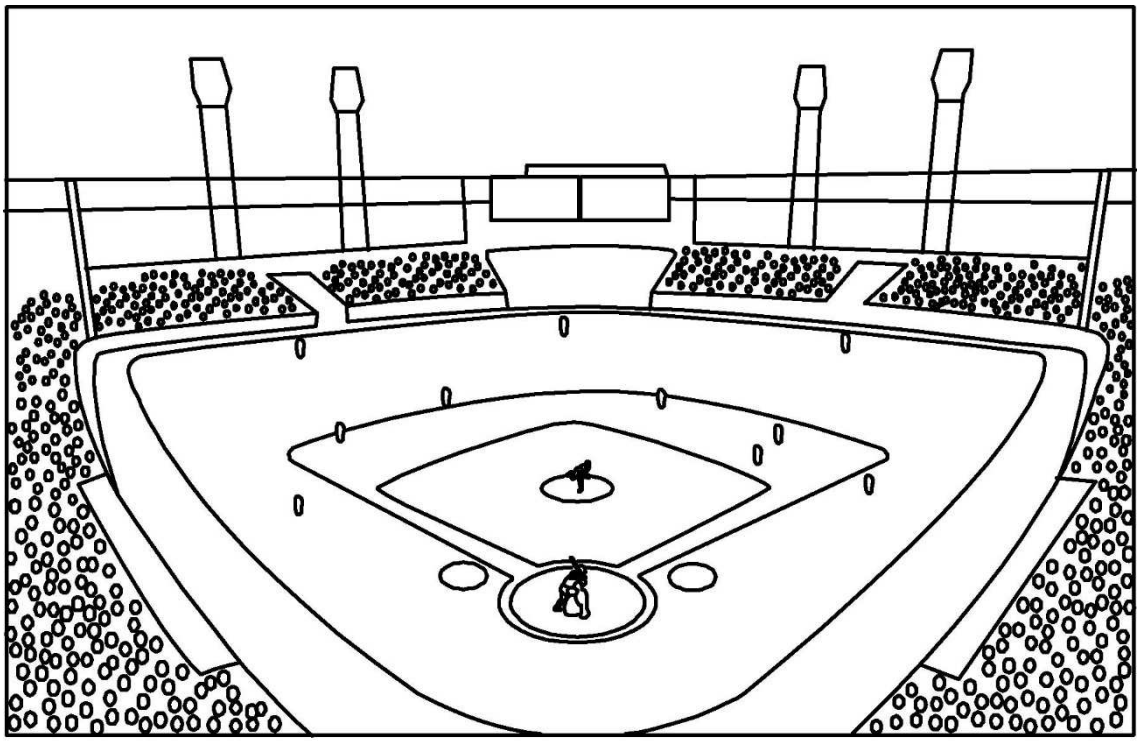


FIG. 24

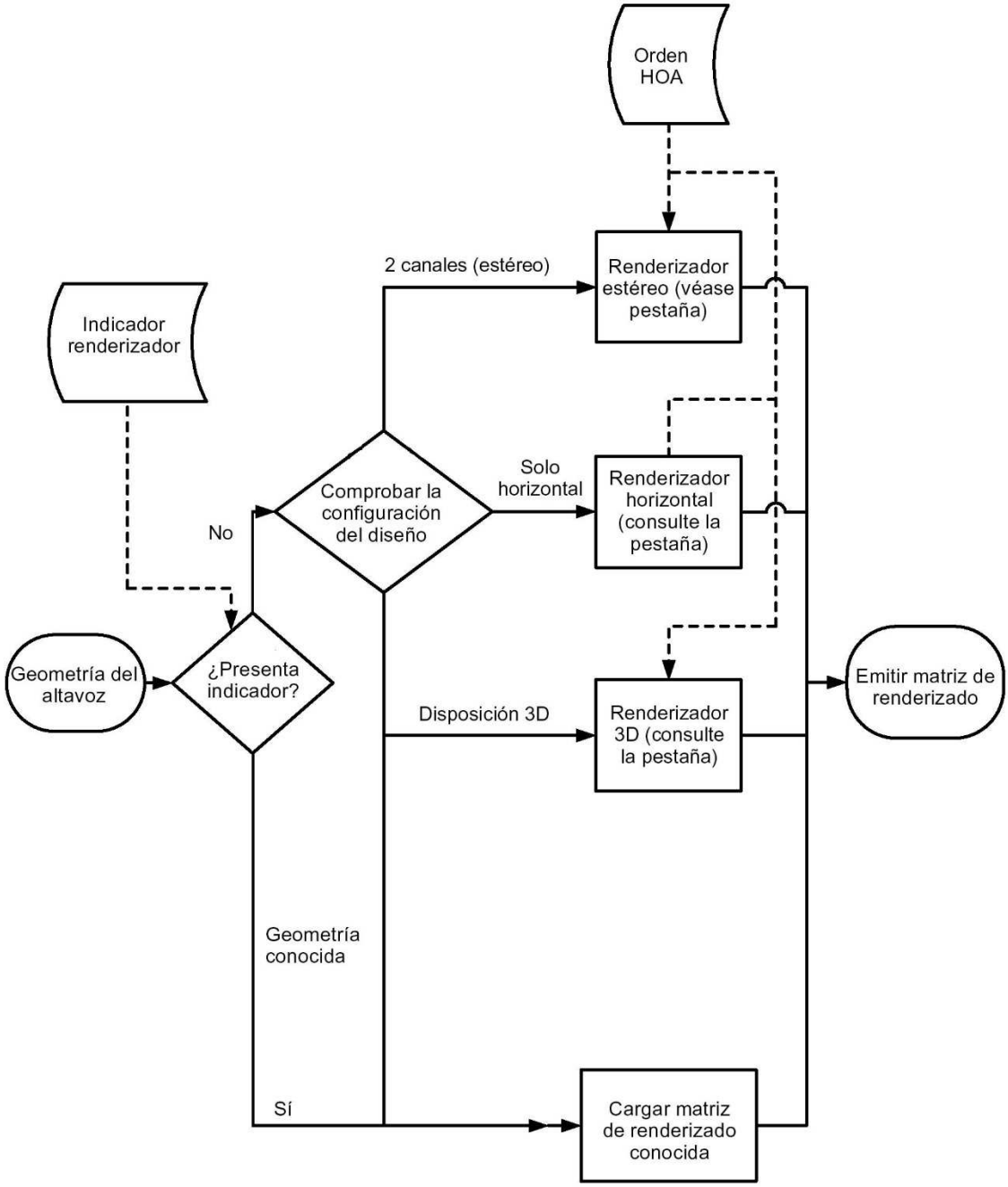

Orden HOA
2 canales (estéreo)
Renderizador estéreo (véase pestaña)
Indicador renderizador
Comprobar la configuración del diseño
Solo horizontal
Renderizador horizontal (consulte la pestaña)
No
Geometría del altavoz
¿Presenta indicador?
Disposición 3D
Renderizador 3D (consulte la pestaña)
Emitir matriz de renderizado
Geometría conocida
Sí
Cargar matriz de renderizado conocida

FIG. 25

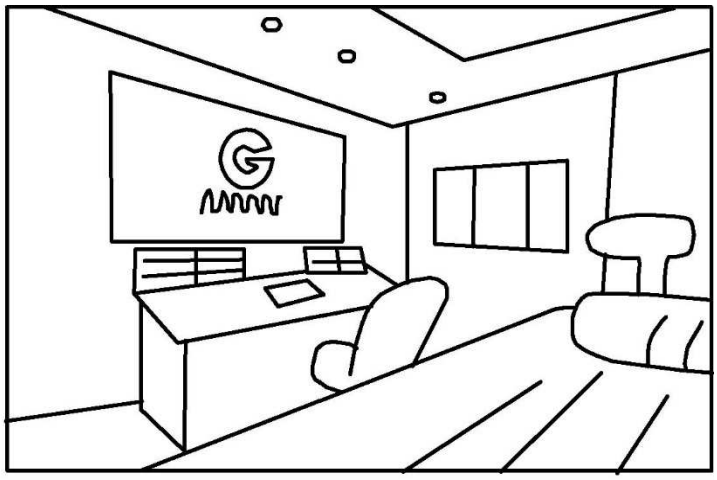


FIG. 26

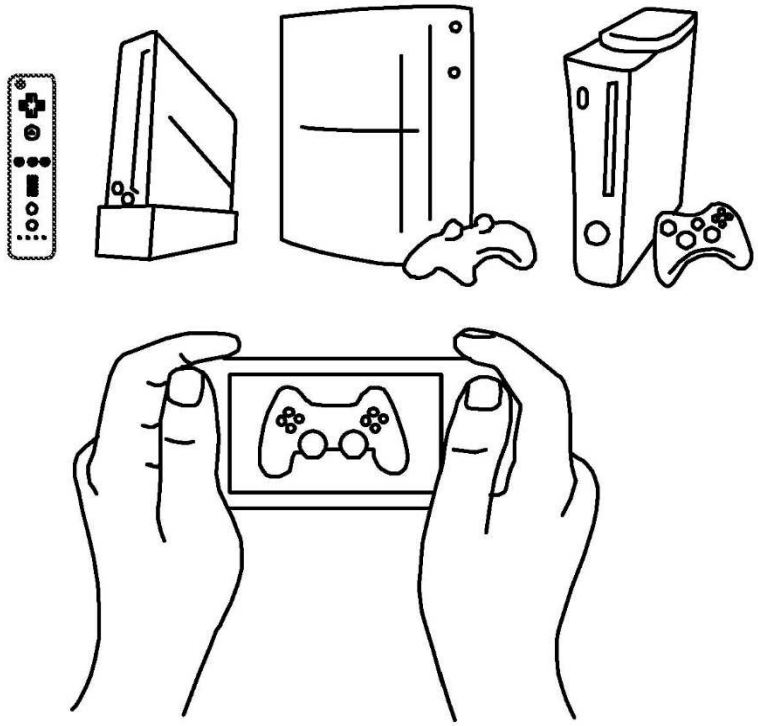


FIG. 27

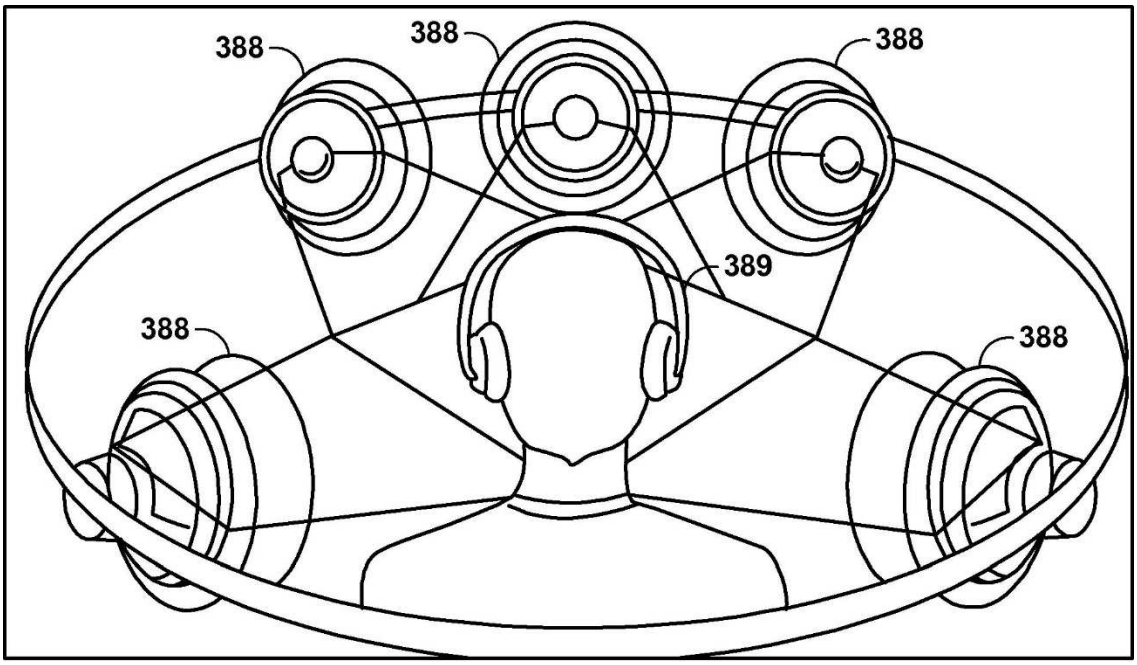


FIG. 28

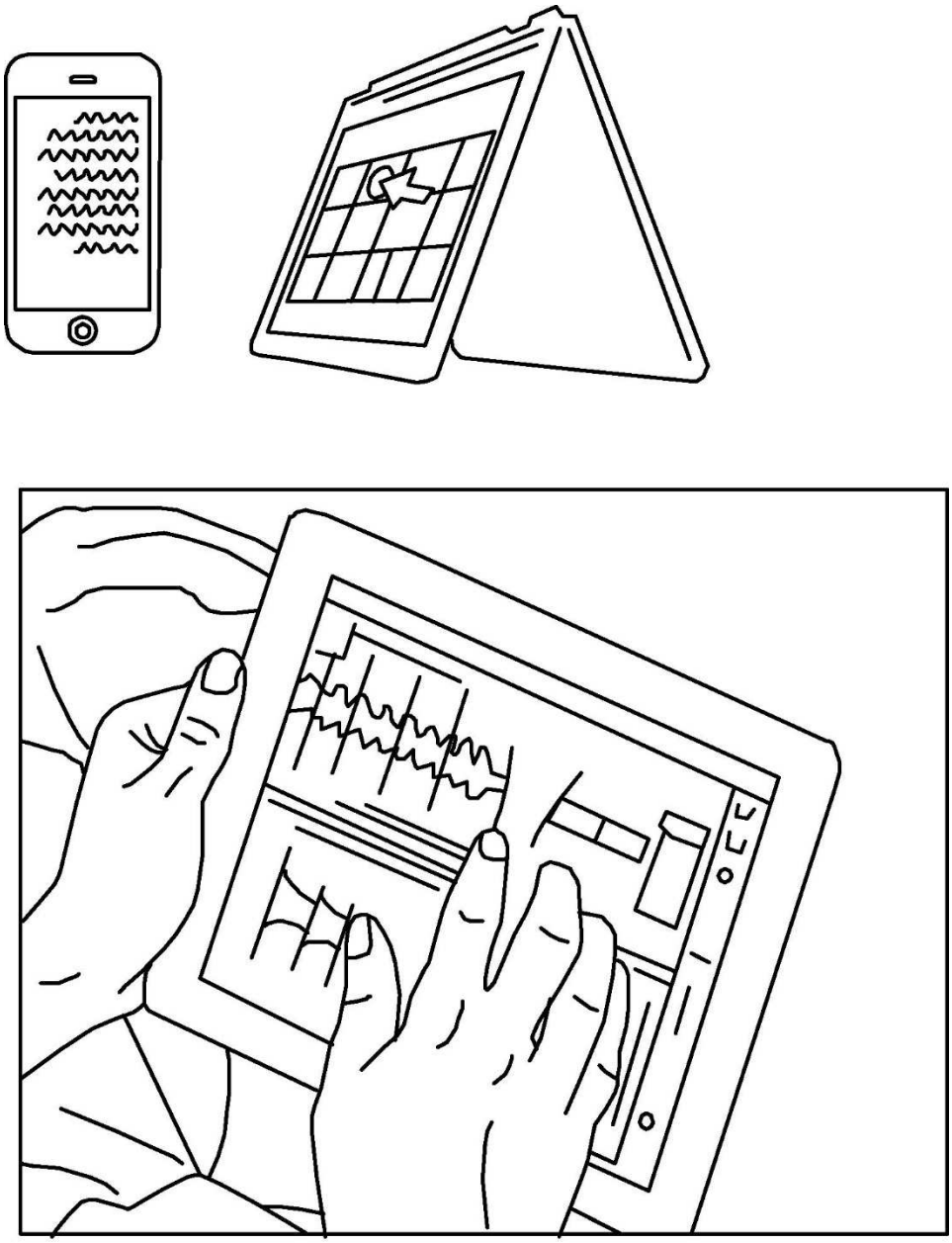


FIG. 29

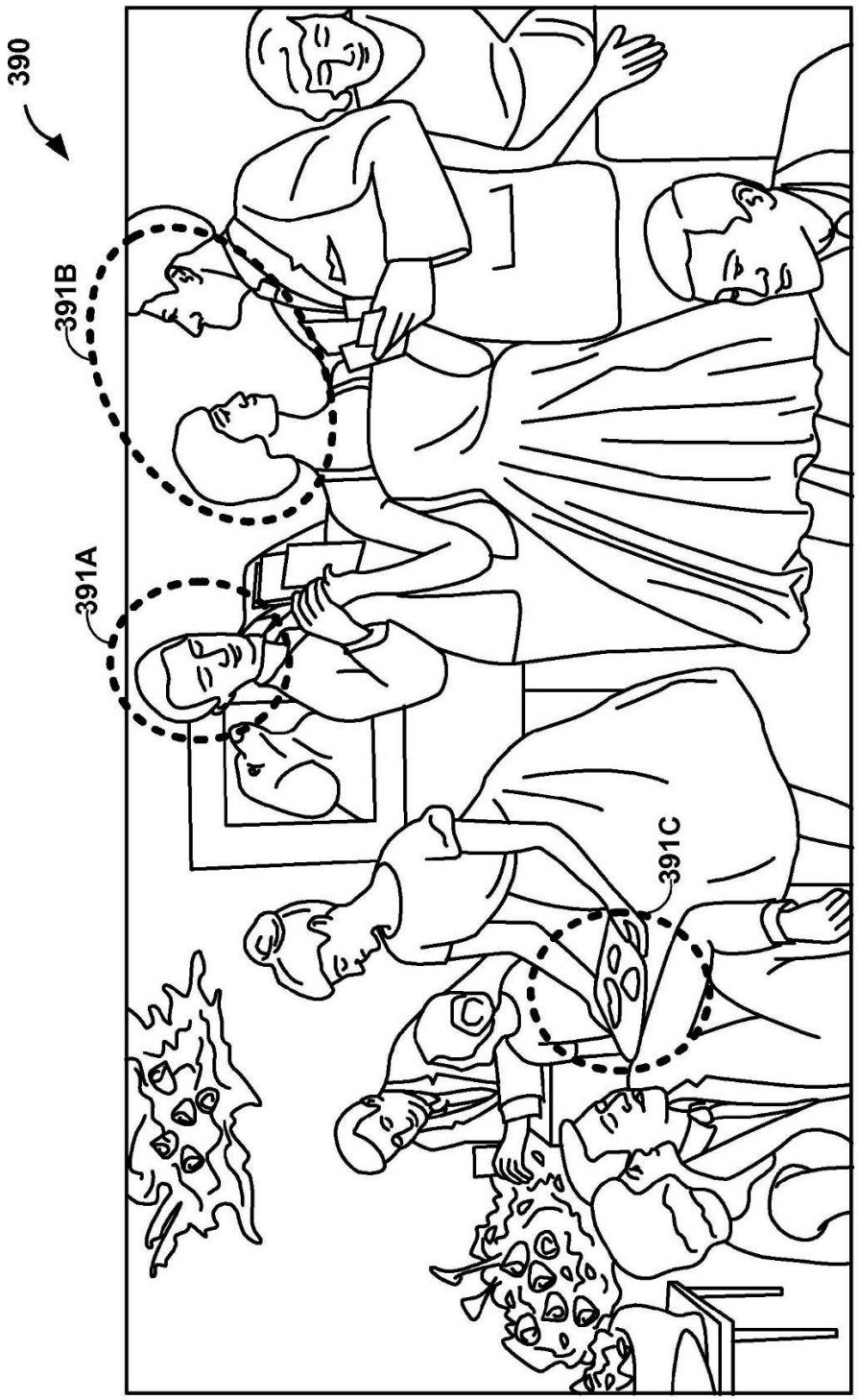


FIG. 30

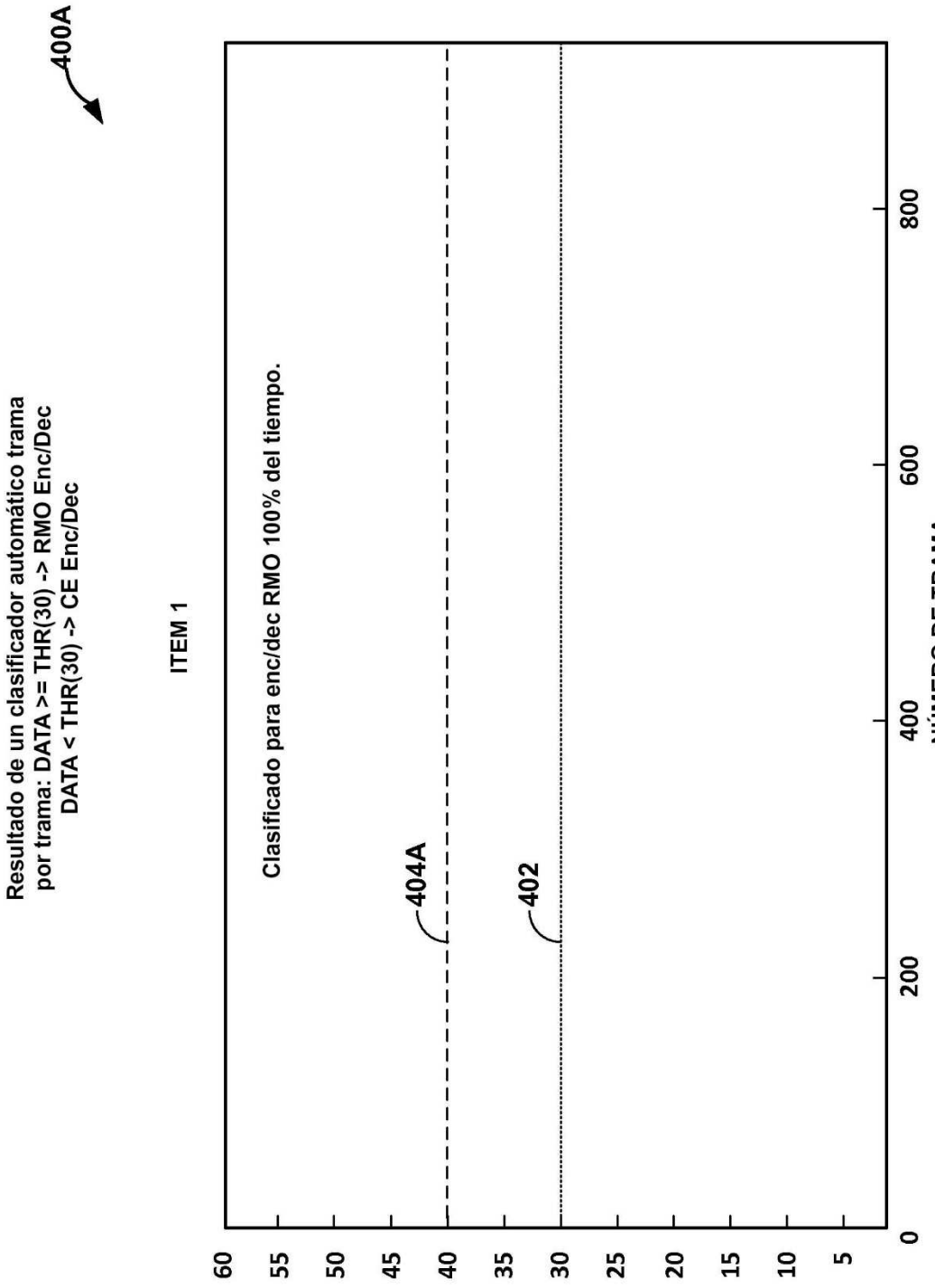

400A
Resultado de un clasificador automático trama por trama: DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 1
Clasificado para enc/dec RMO 100% del tiempo.
404A
402
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
NÚMERO DE TRAMA

FIG. 31A

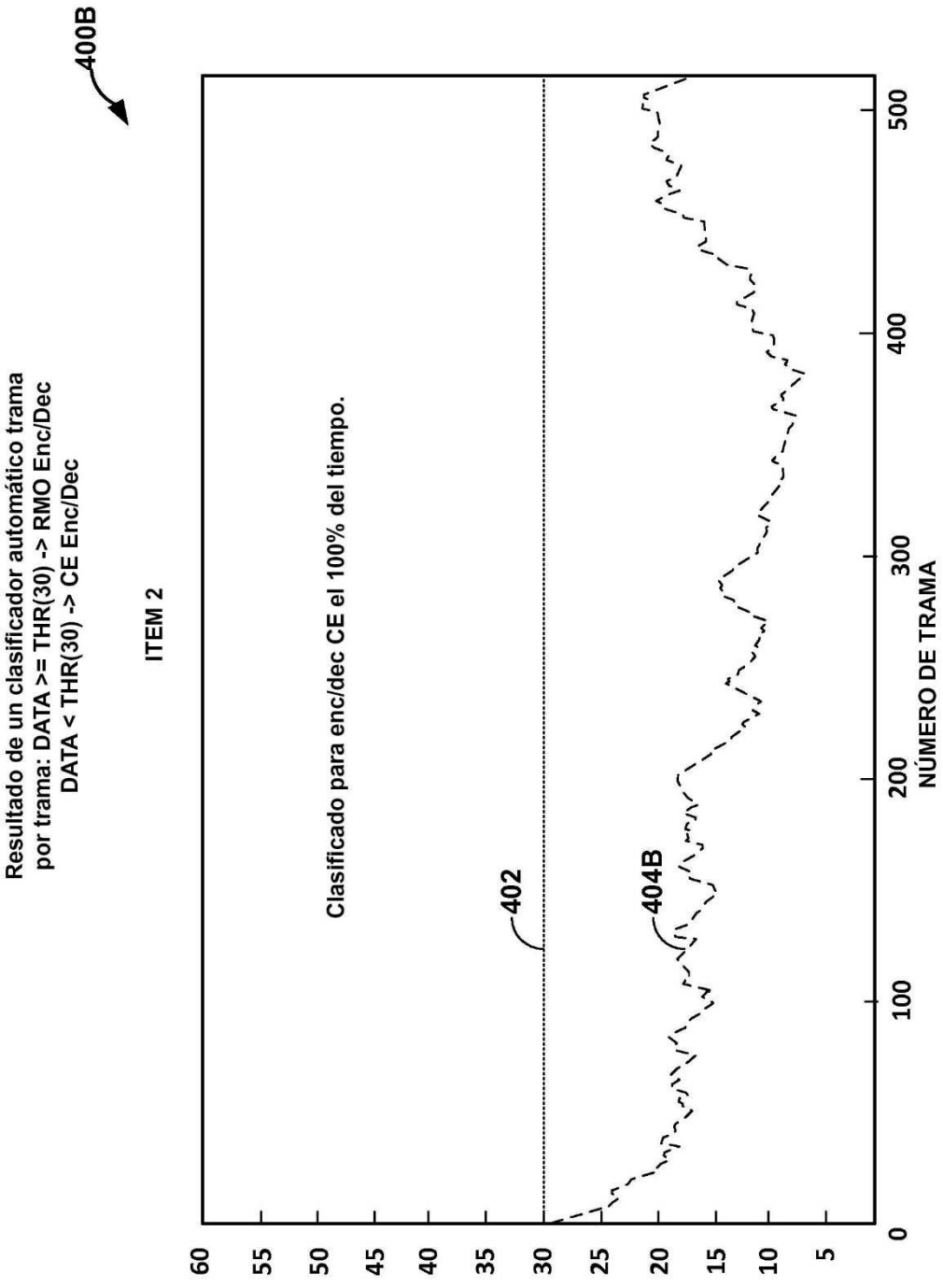

400B
Resultado de un clasificador automático trama por trama: DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 2
Clasificado para enc/dec CE el 100% del tiempo.
402
404B
NÚMERO DE TRAMA
0
100
200
300
400
500
5
10
15
20
25
30
35
40
45
50
55
60

FIG. 31B

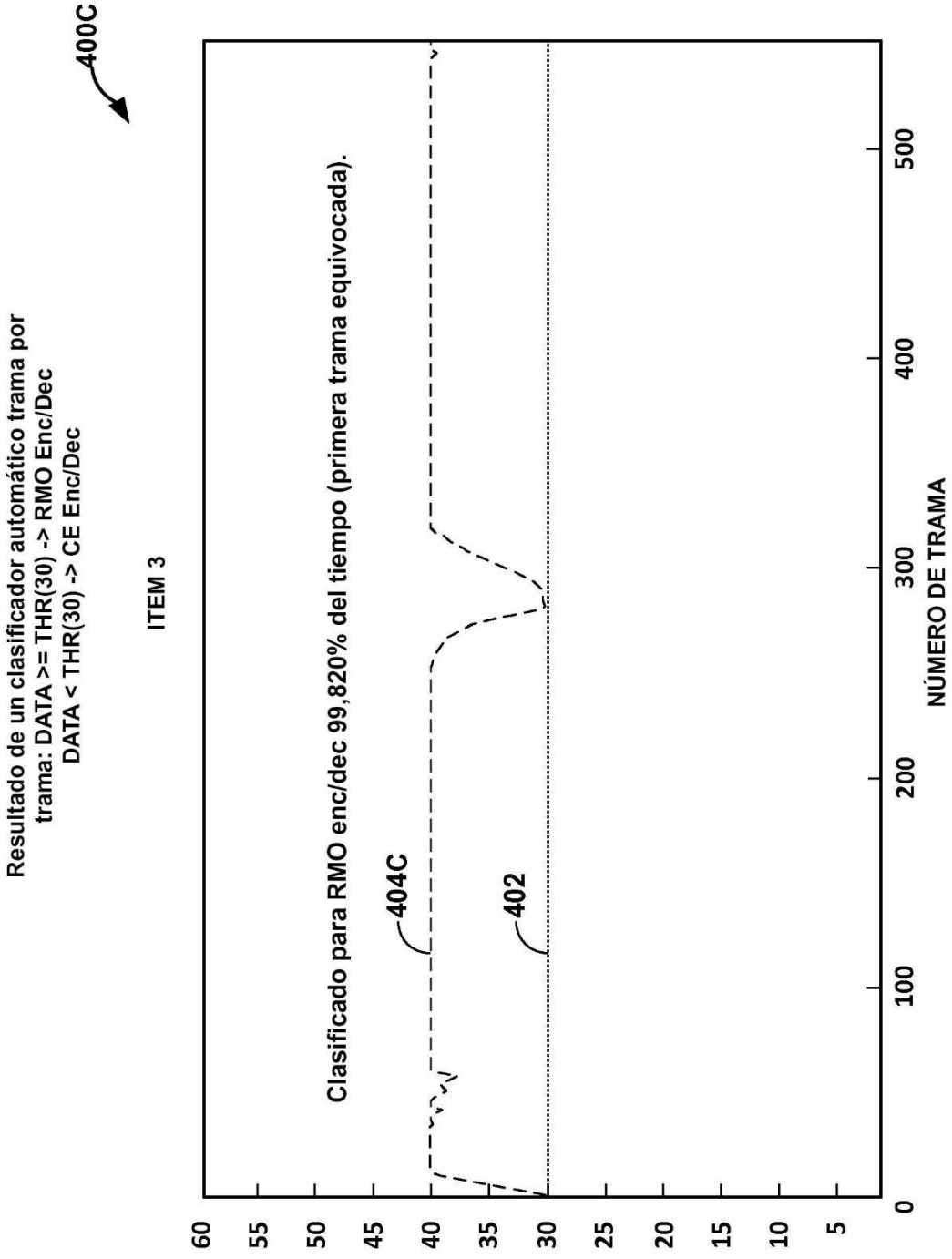

400C
Resultado de un clasificador automático trama por trama: DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 3
Clasificado para RMO enc/dec 99,820% del tiempo (primera trama equivocada).
404C
402
60
55
50
45
40
35
30
25
20
15
10
5
0
100
200
300
400
500
NÚMERO DE TRAMA

FIG. 31C

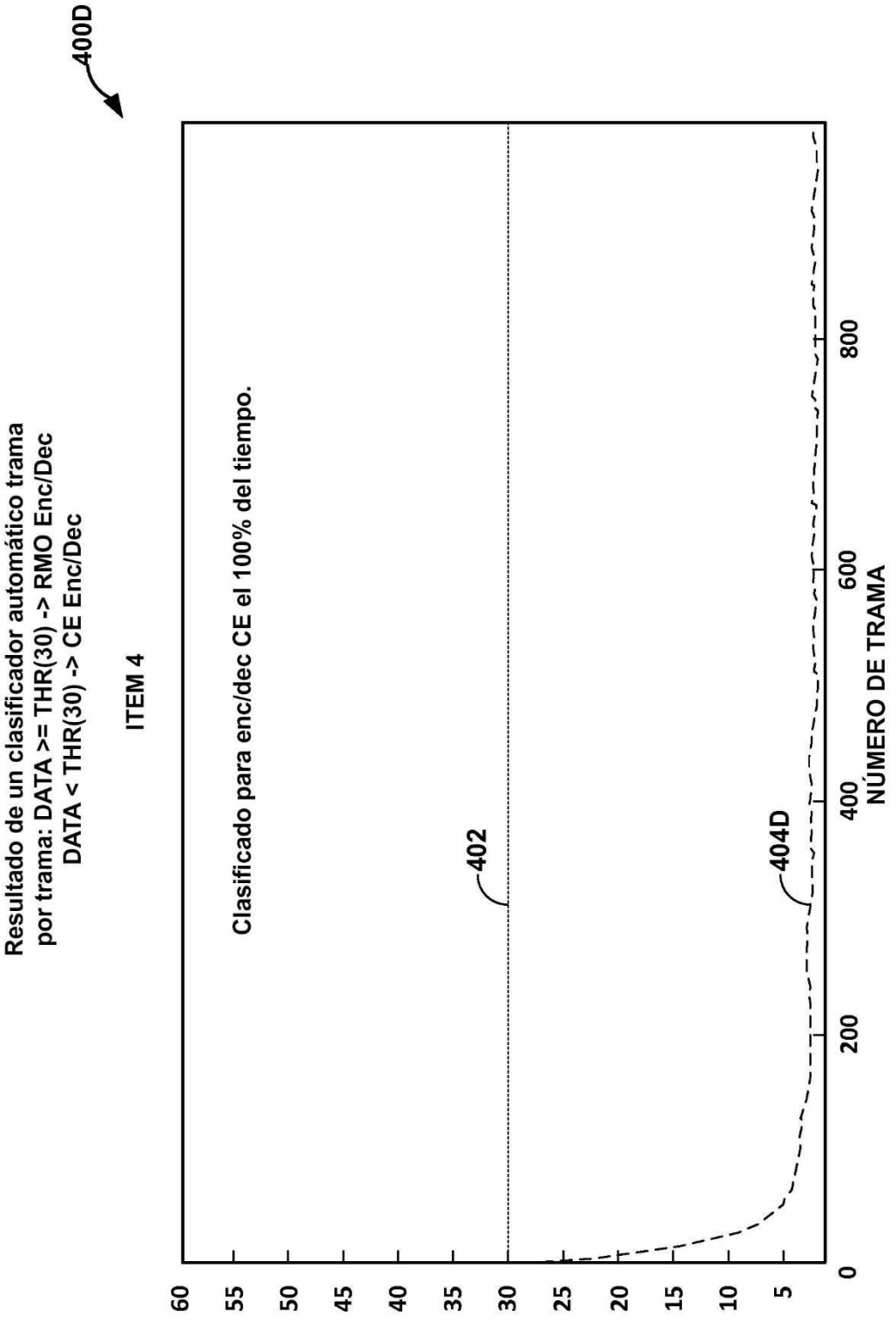

400D
Resultado de un clasificador automático trama por trama: DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 4
Clasificado para enc/dec CE el 100% del tiempo.
402
404D
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
NÚMERO DE TRAMA

FIG. 31D

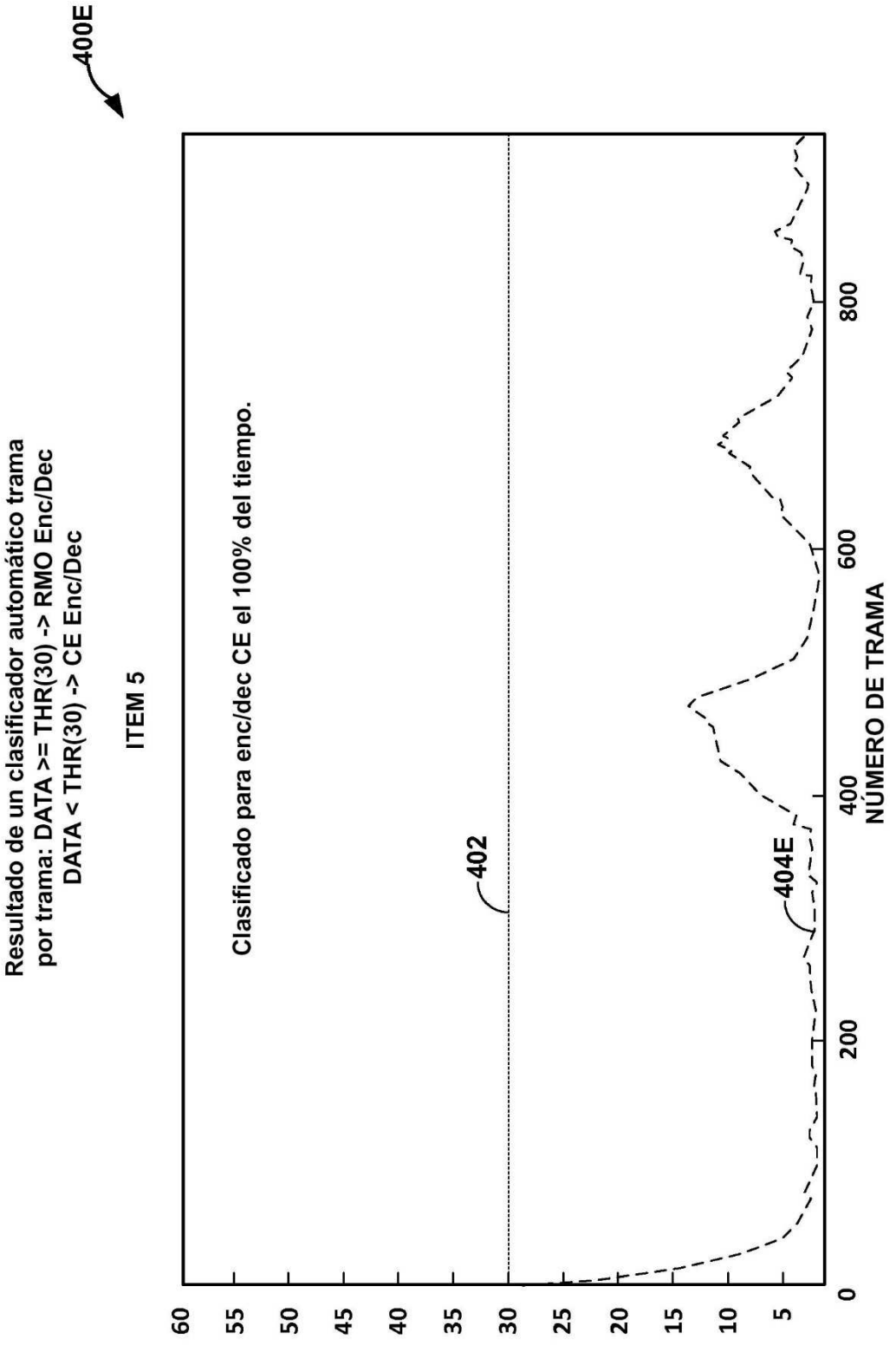

400E
Resultado de un clasificador automático trama por trama: DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 5
Clasificado para enc/dec CE el 100% del tiempo.
402
404E
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
NÚMERO DE TRAMA

FIG. 31E

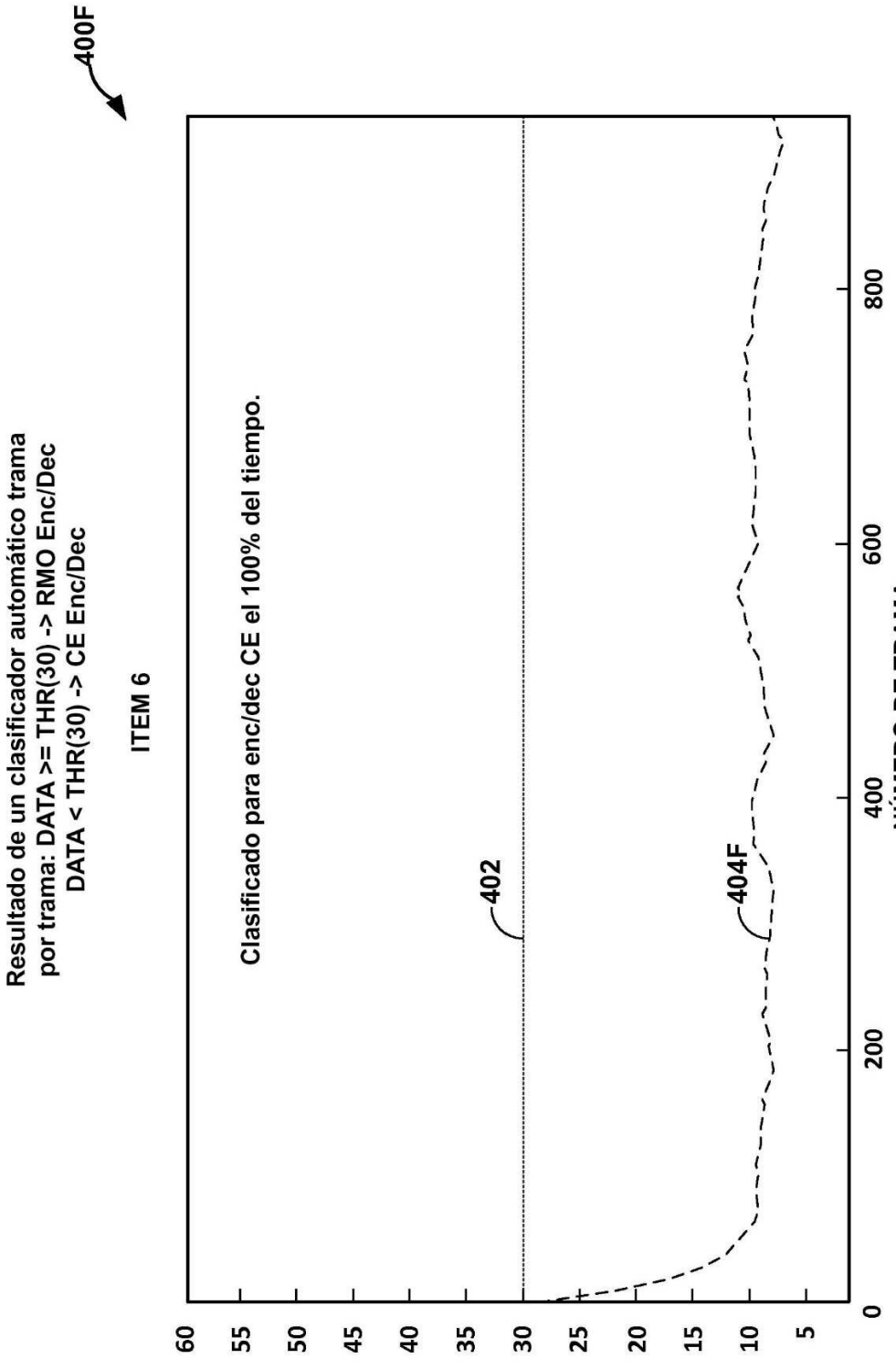

400F
Resultado de un clasificador automático trama por trama: DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 6
Clasificado para enc/dec CE el 100% del tiempo.
402
404F
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
NÚMERO DE TRAMA

FIG. 31F

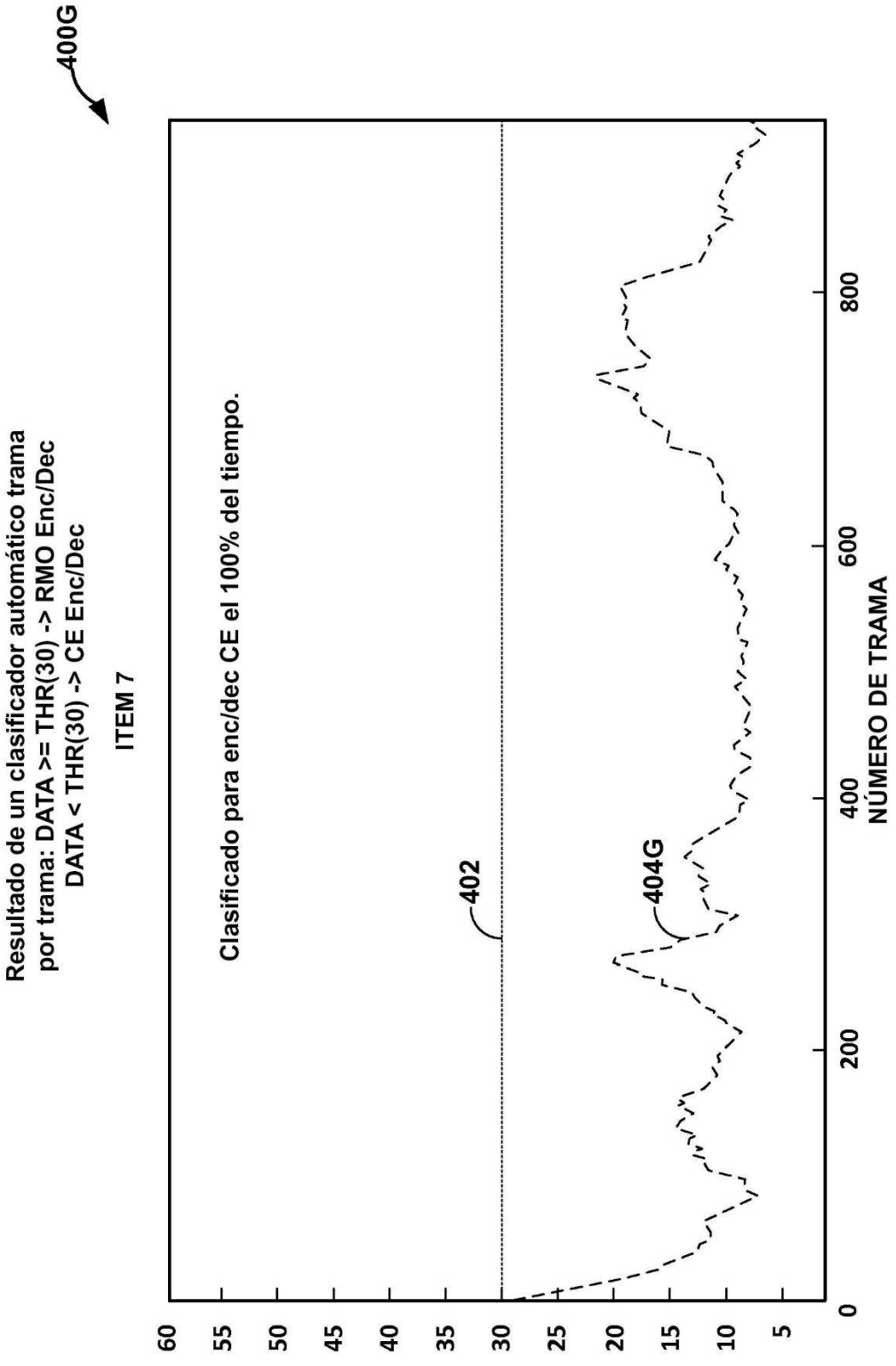

400G
Resultado de un clasificador automático trama por trama: DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 7
Clasificado para enc/dec CE el 100% del tiempo.
402
404G
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
NÚMERO DE TRAMA

FIG. 31G

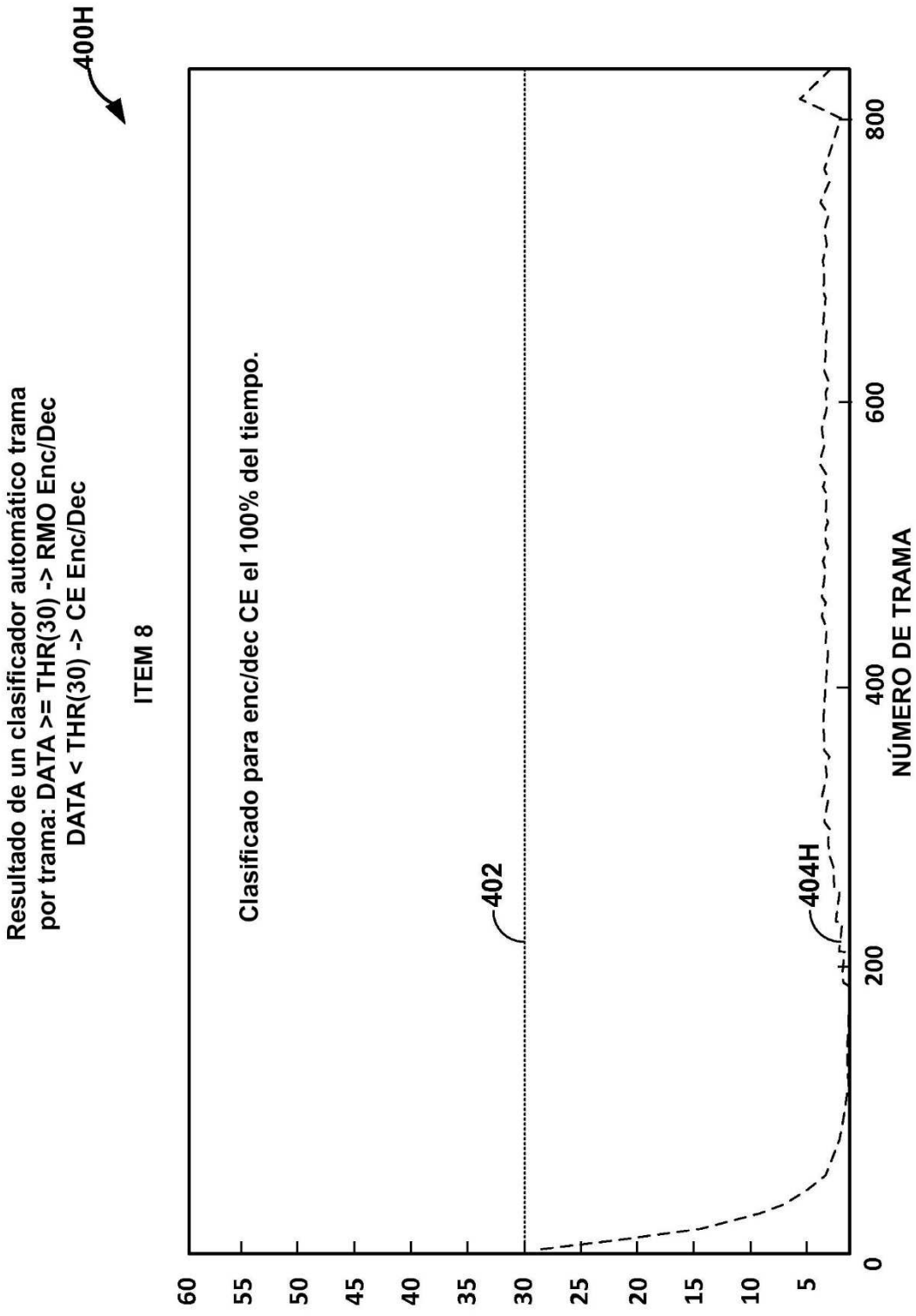

400H
Resultado de un clasificador automático trama por trama: DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 8
Clasificado para enc/dec CE el 100% del tiempo.
402
404H
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
NÚMERO DE TRAMA

FIG. 31H

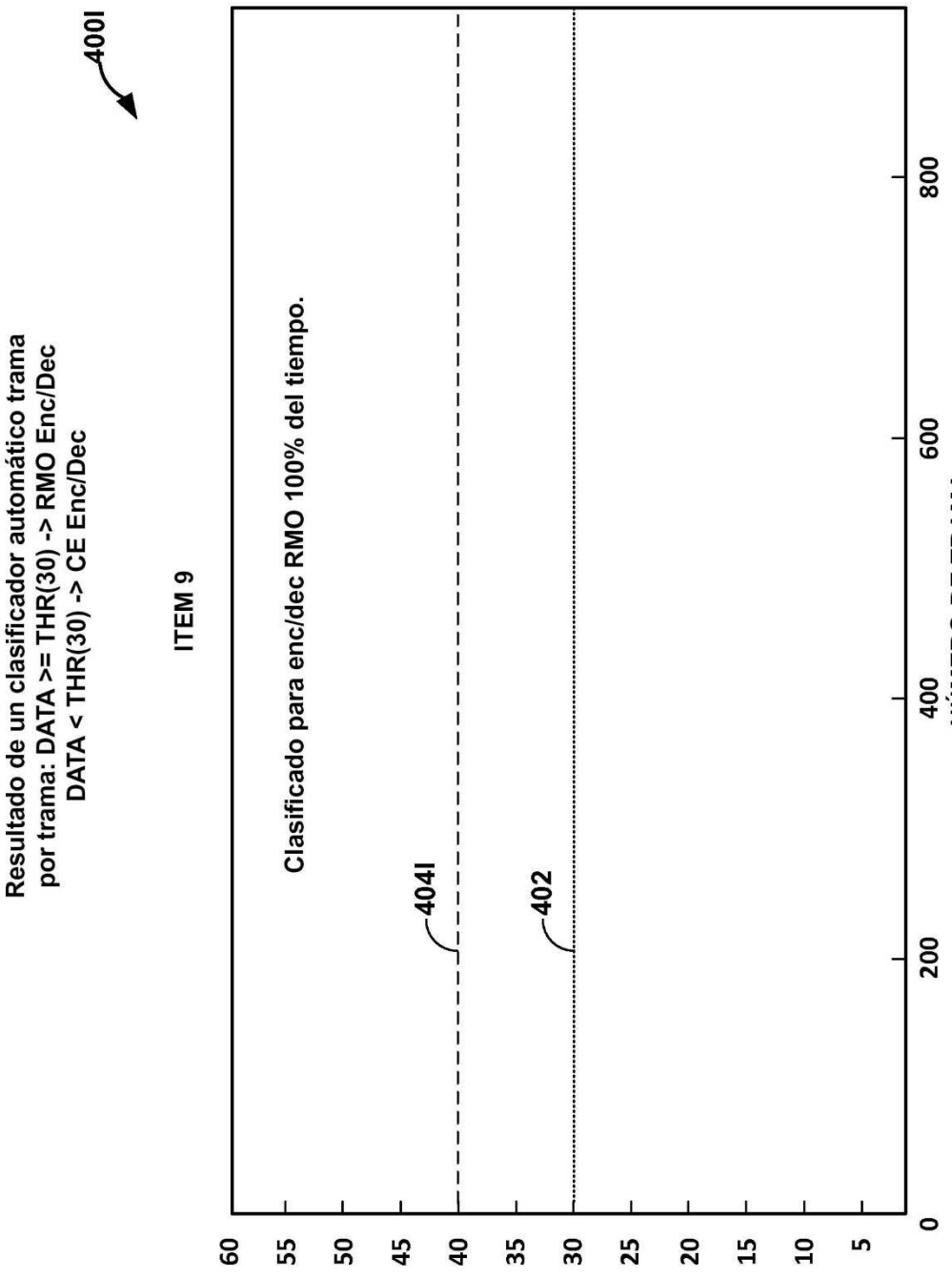

400I
Resultado de un clasificador automático trama por trama: DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 9
Clasificado para enc/dec RMO 100% del tiempo.
404I
402
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
NÚMERO DE TRAMA

FIG. 31I

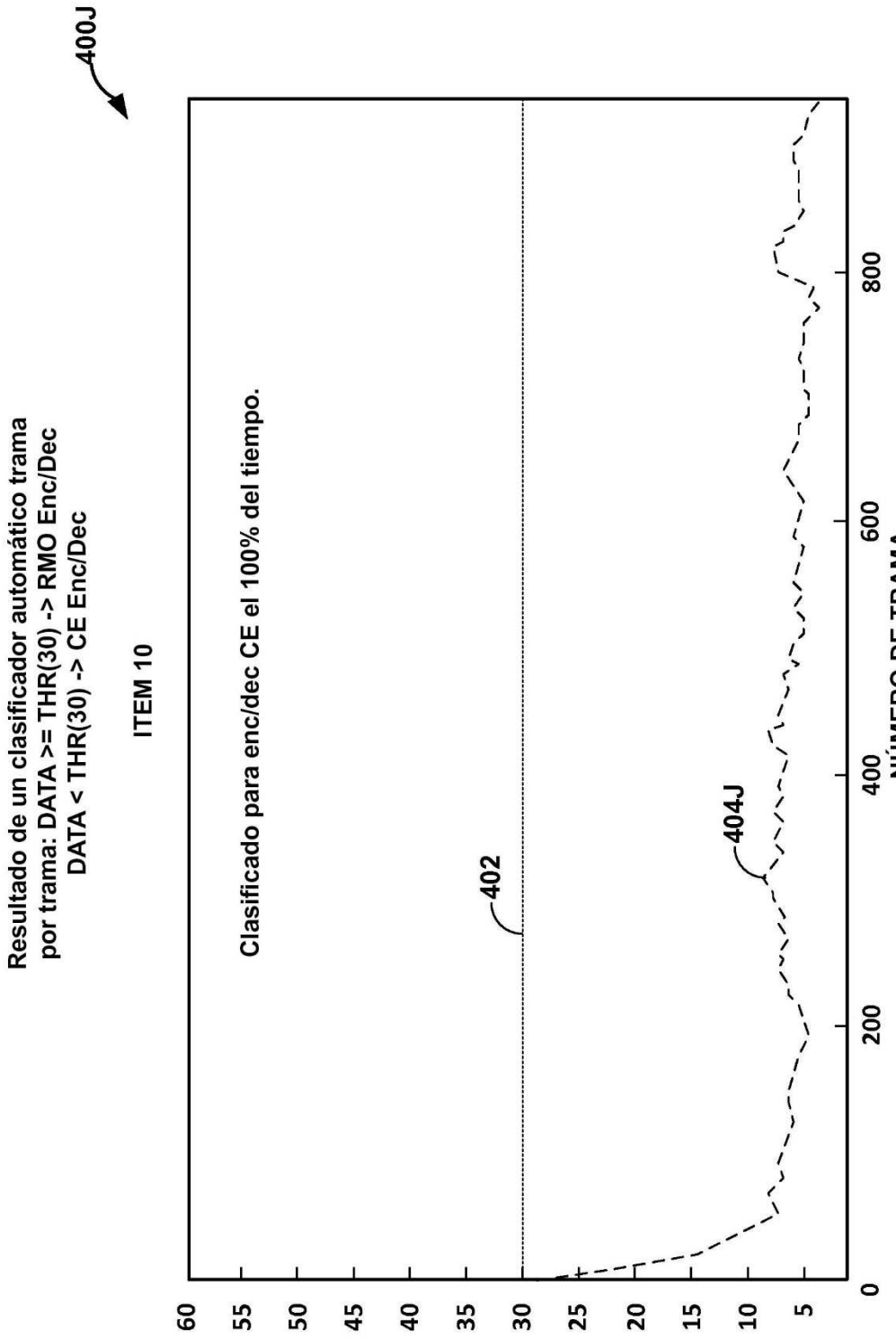

400J
Resultado de un clasificador automático trama por trama: DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 10
Clasificado para enc/dec CE el 100% del tiempo.
402
404J
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
NÚMERO DE TRAMA

FIG. 31J

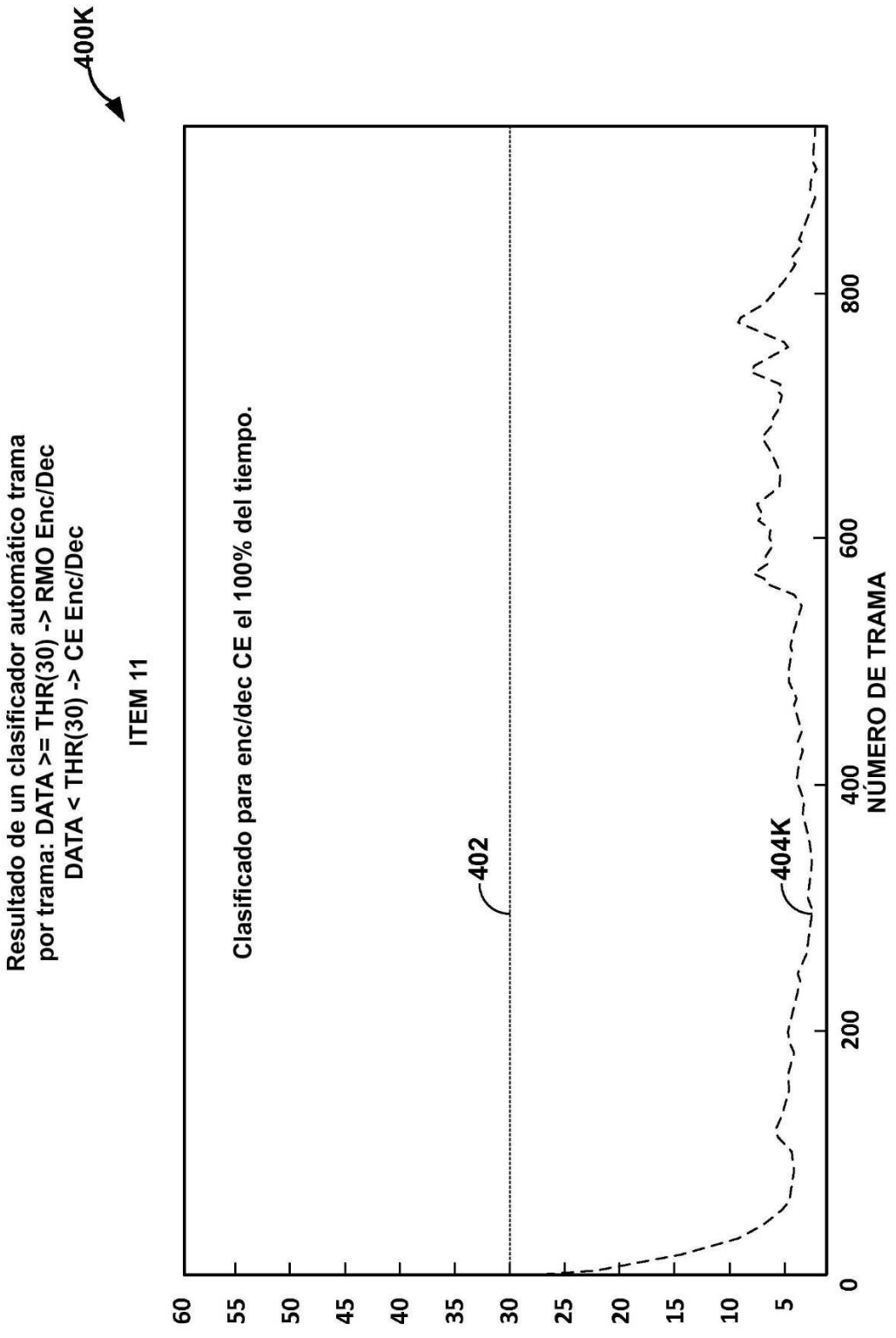

400K
Resultado de un clasificador automático trama por trama: DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 11
Clasificado para enc/dec CE el 100% del tiempo.
402
404K
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
NÚMERO DE TRAMA

FIG. 31K

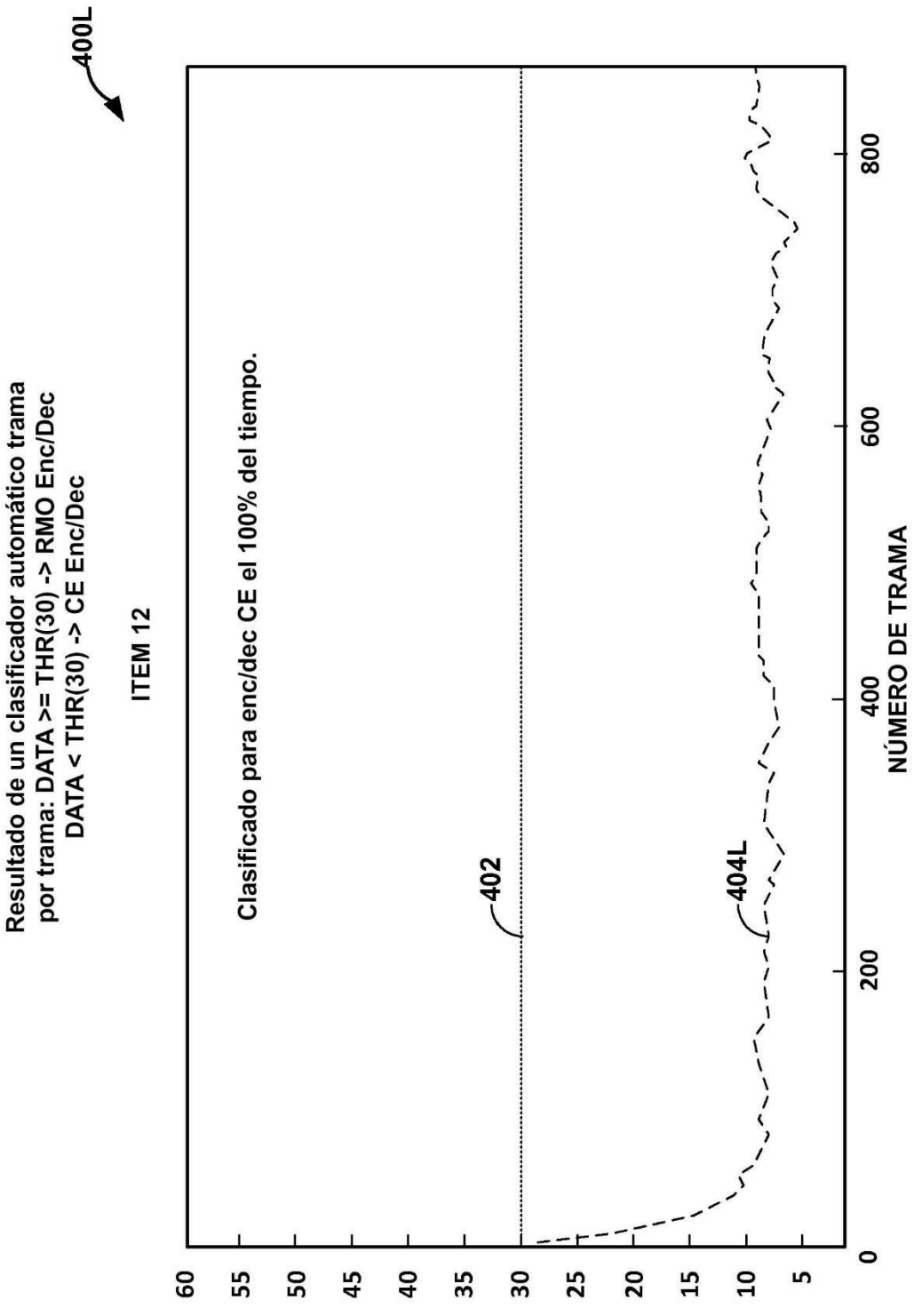

400L
Resultado de un clasificador automático trama por trama: DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
ITEM 12
Clasificado para enc/dec CE el 100% del tiempo.
402
404L
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
NÚMERO DE TRAMA

FIG. 31L

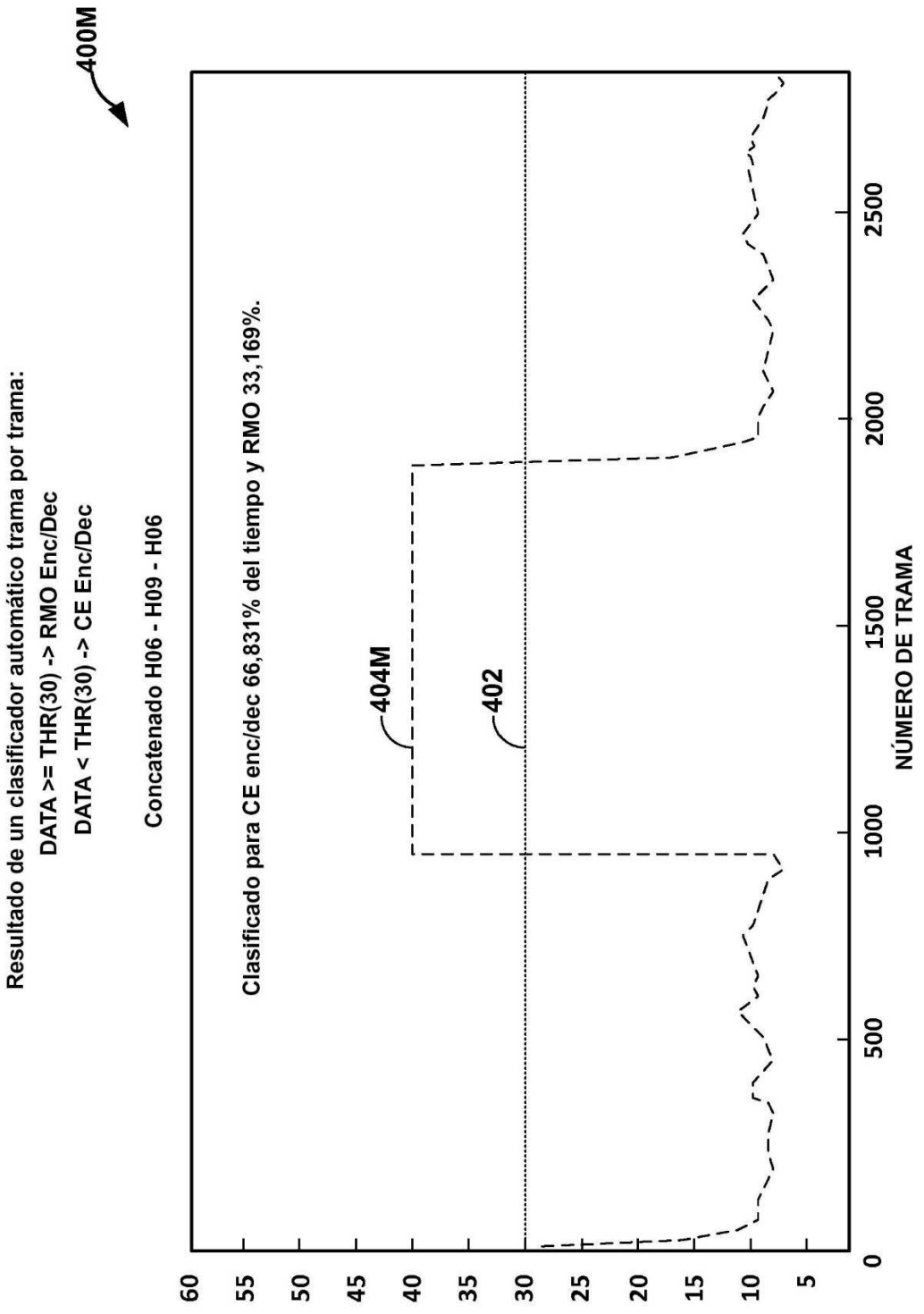

400M
Resultado de un clasificador automático trama por trama:
DATA >= THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec
Concatenado H06 - H09 - H06
Clasificado para CE enc/dec 66,831% del tiempo y RMO 33,169%.
404M
402
60
55
50
45
40
35
30
25
20
15
10
5
0
500
1000
1500
2000
2500
NÚMERO DE TRAMA

FIG. 31M

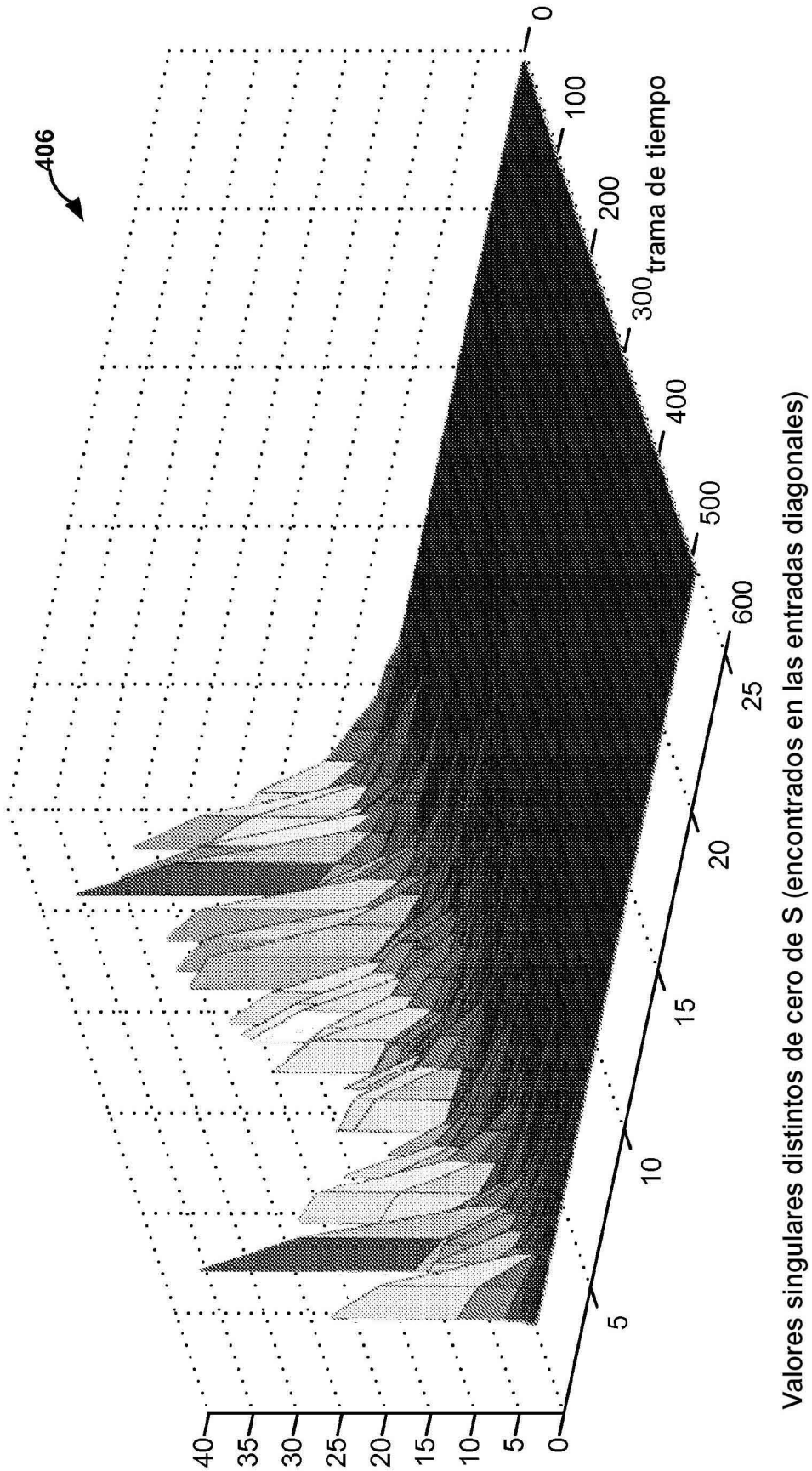

406
0
100
200
300
trama de tiempo
400
500
600
5
10
15
20
25
Valores singulares distintos de cero de S (encontrados en las entradas diagonales)
0
5
10
15
20
25
30
35
40

FIG. 32

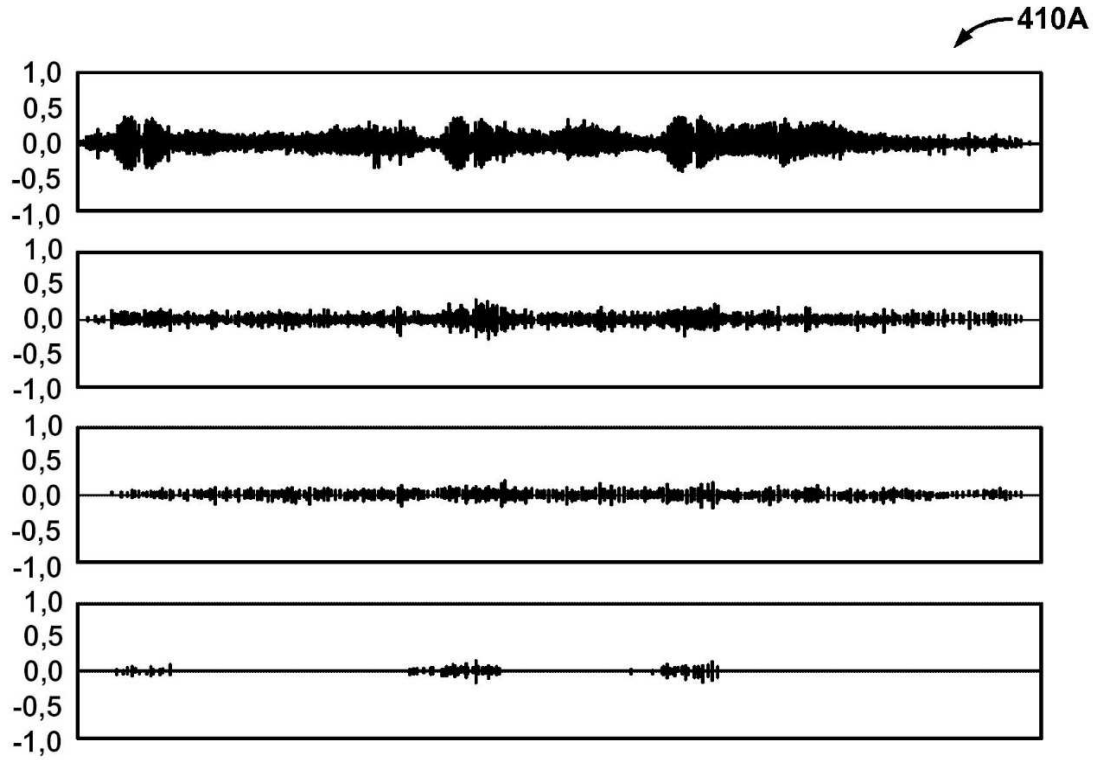


FIG. 33A

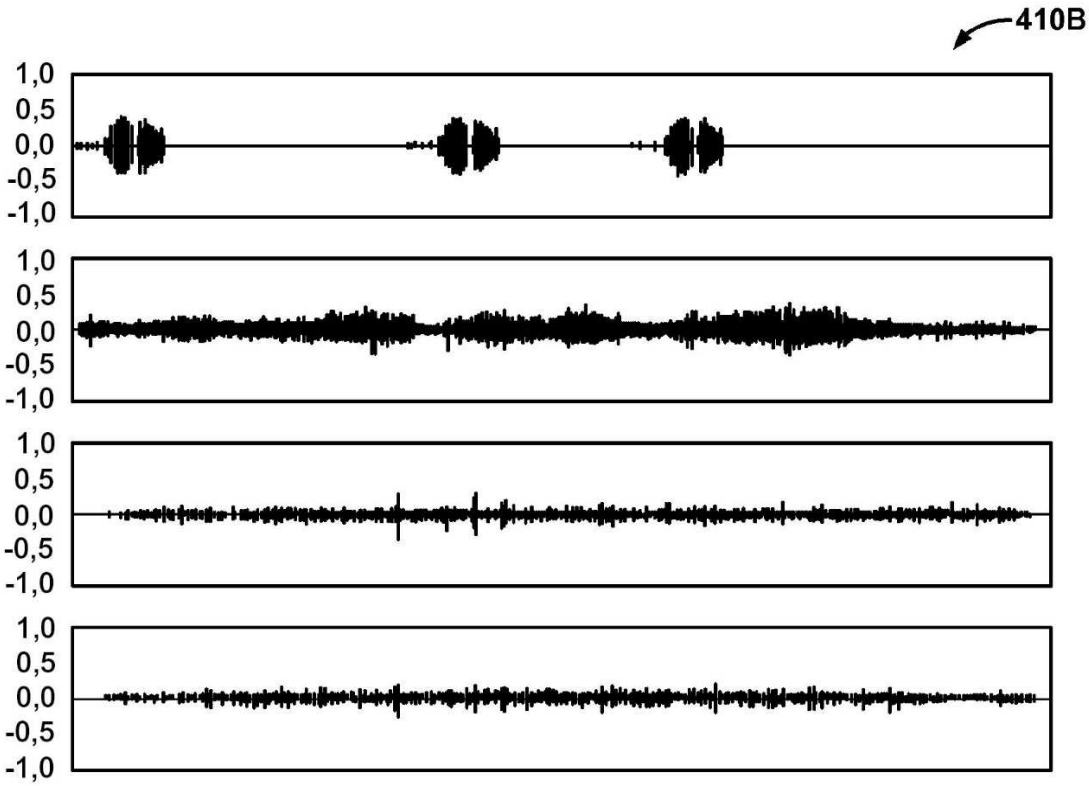


FIG. 33B

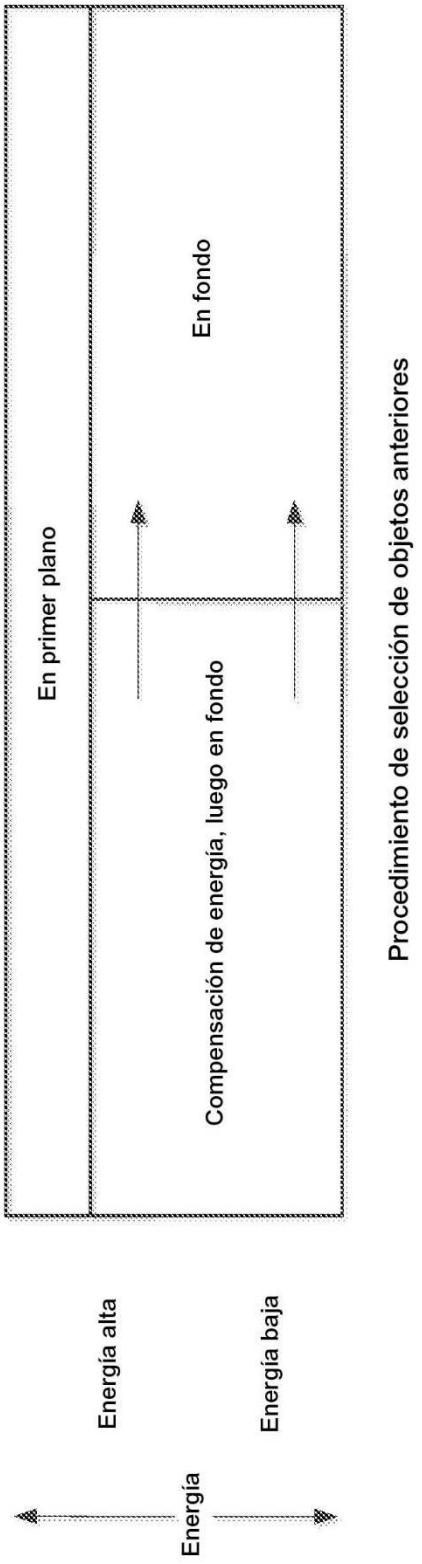

Direccionalidad
Alta direccionalidad
Baja direccionalidad
En primer plano
En fondo
Compensación de energía, luego en fondo
Procedimiento de selección de objetos anteriores
Energía
Energía alta
Energía baja

FIG. 34

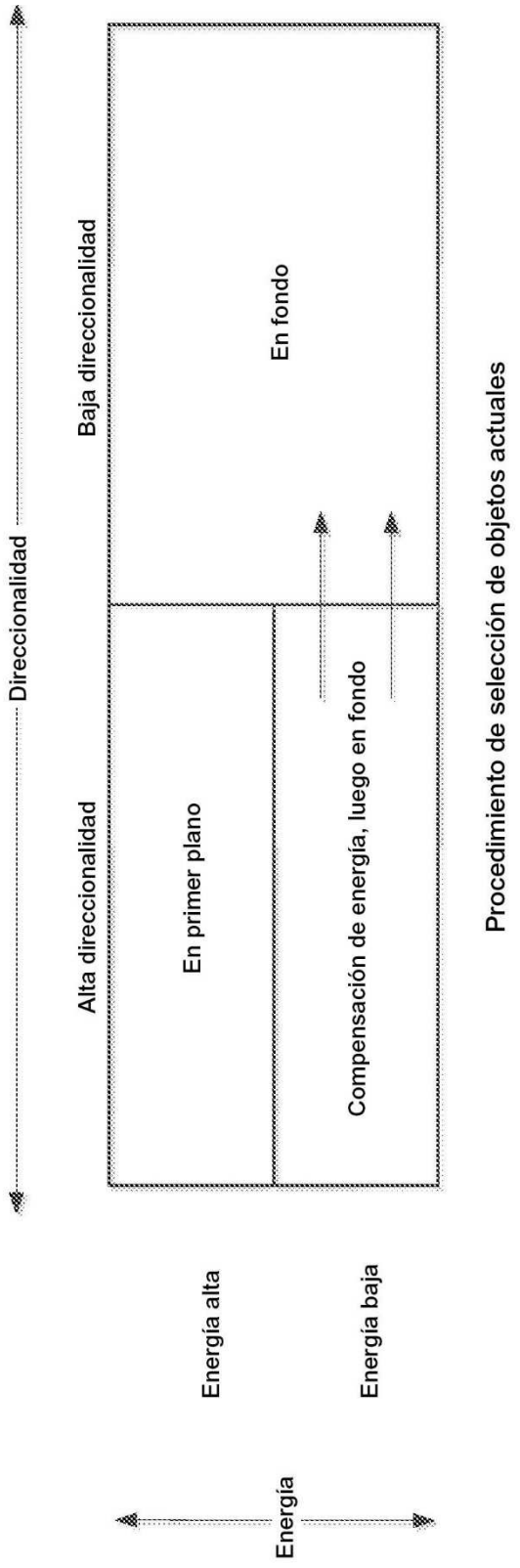

Direccionalidad
Baja direccionalidad
En fondo
Alta direccionalidad
En primer plano
Compensación de energía, luego en fondo
Procedimiento de selección de objetos actuales
Energía alta
Energía baja
Energía

FIG. 35

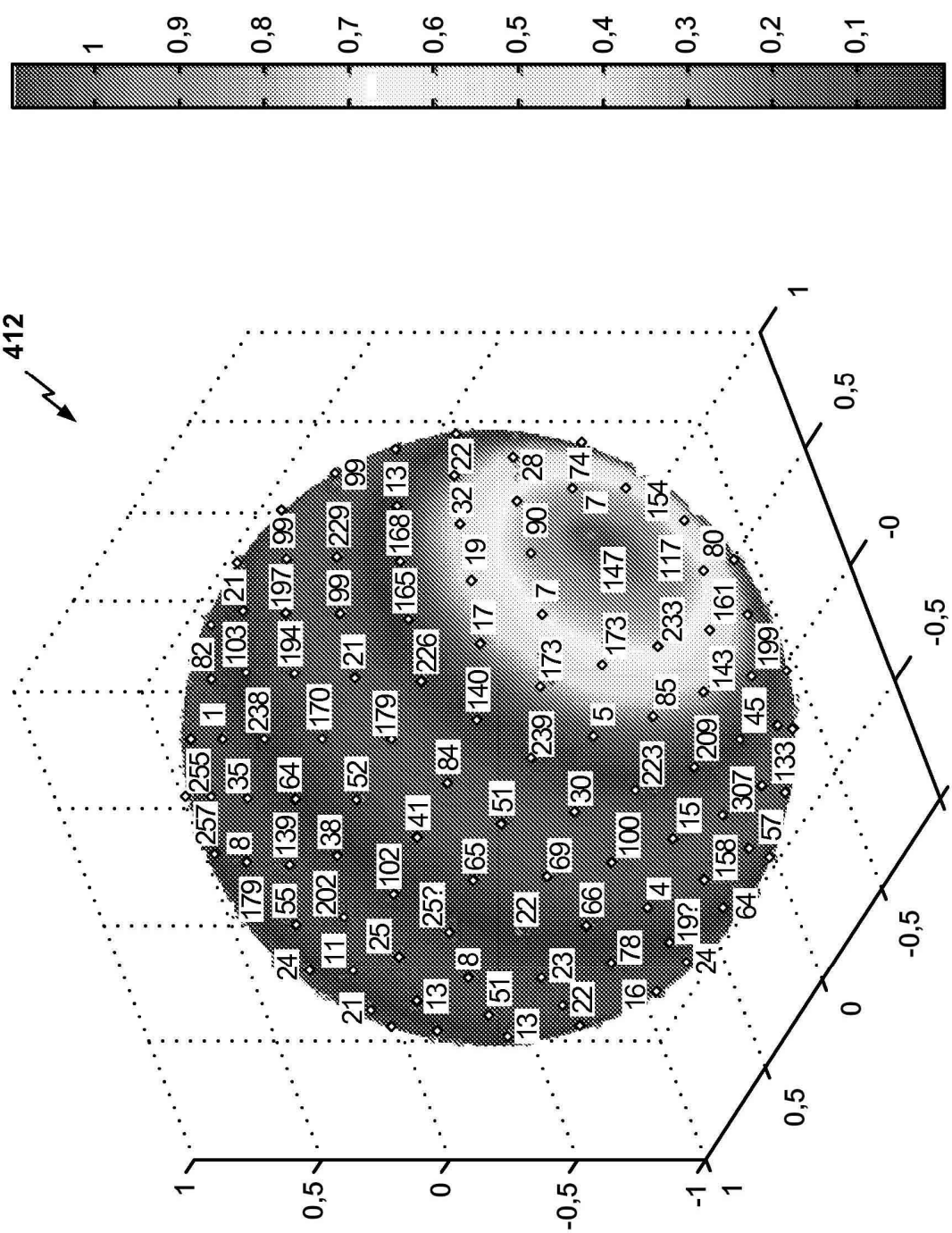


FIG. 36A

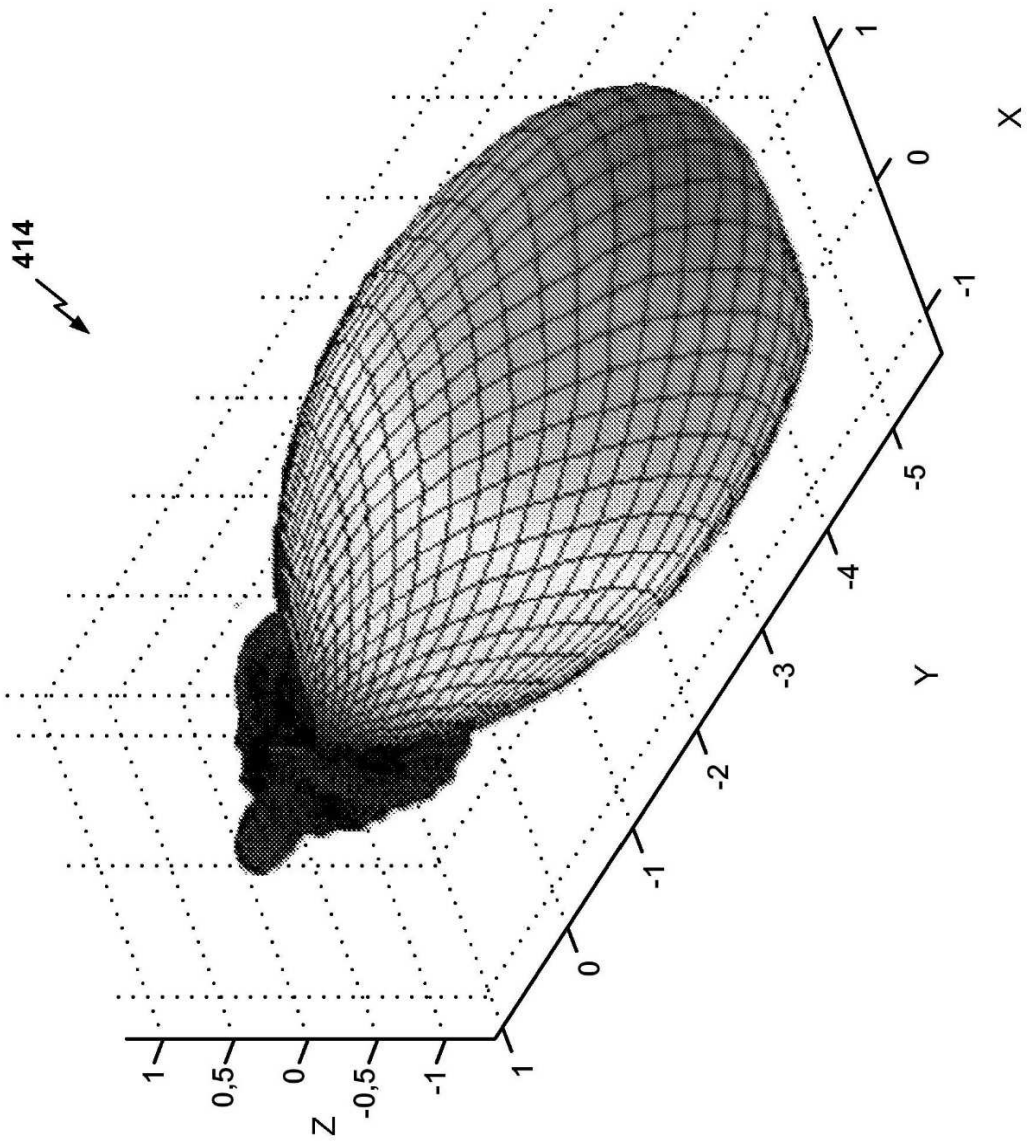


FIG. 36B

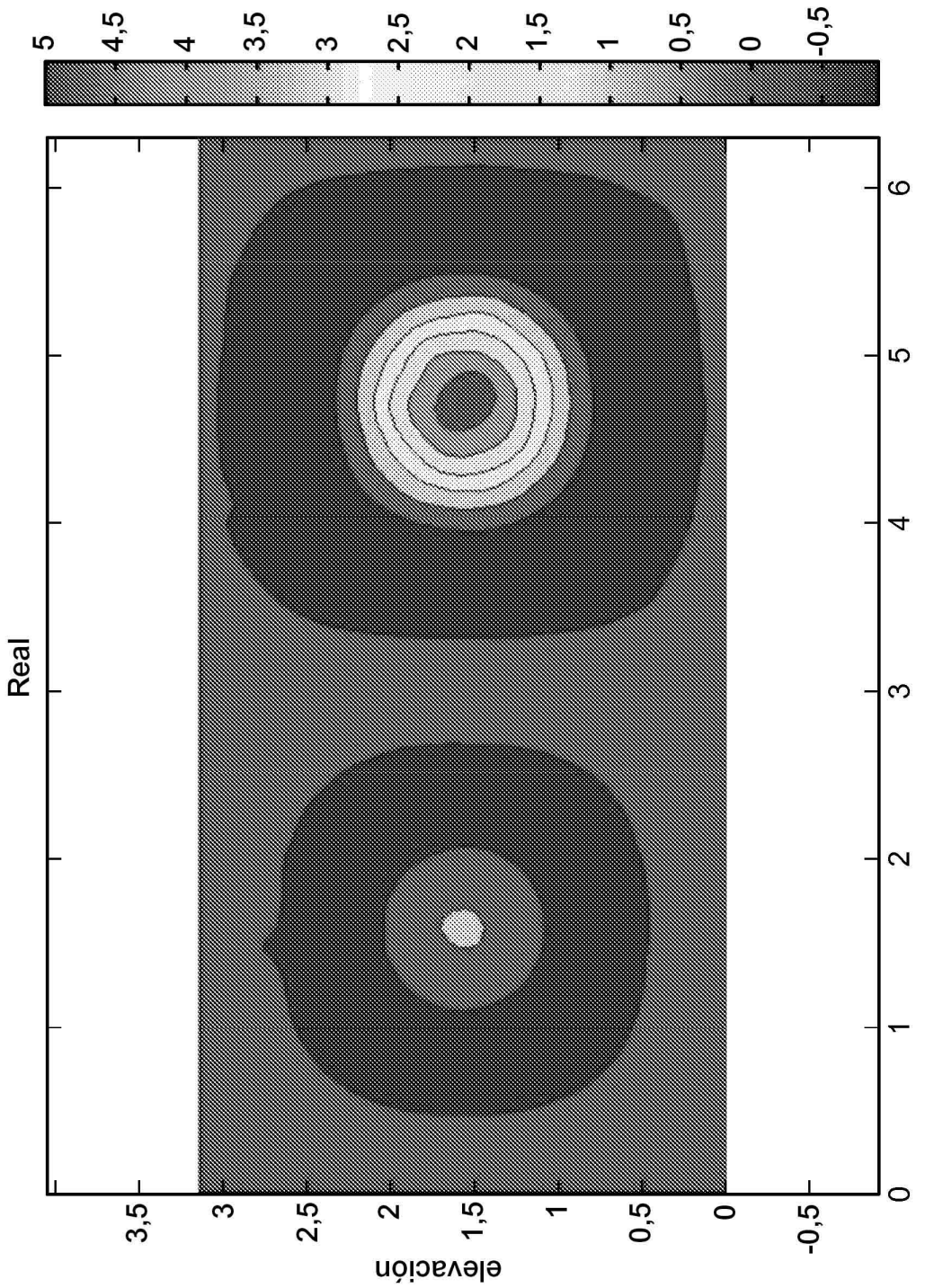

Real
elevación
azimut
5
4,5
4
3,5
3
2,5
2
1,5
1
0,5
0
-0,5
3,5
3
2,5
2
1,5
1
0,5
0
-0,5
0
1
2
3
4
5
6

FIG. 36C

El sonido de la abeja procede del azimut = 0° y elevación = 45°.

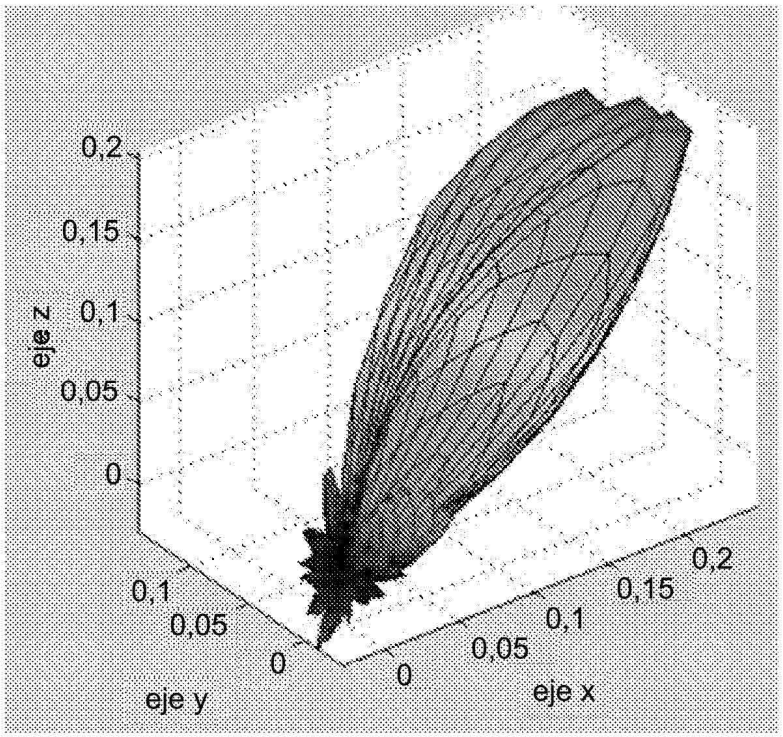


FIG. 36D

Sonido de helicóptero procede del cielo.

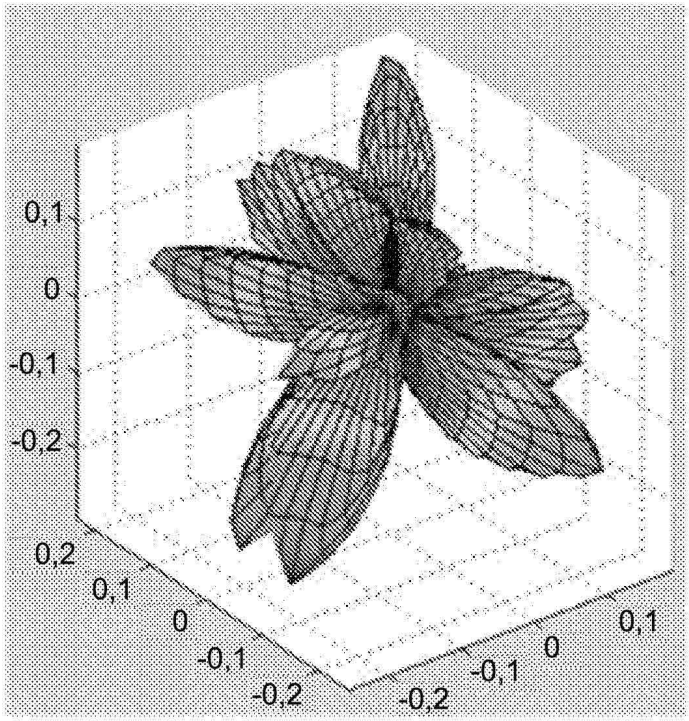


FIG. 36E

La música electrónica moderna procede de dos direcciones diferentes.

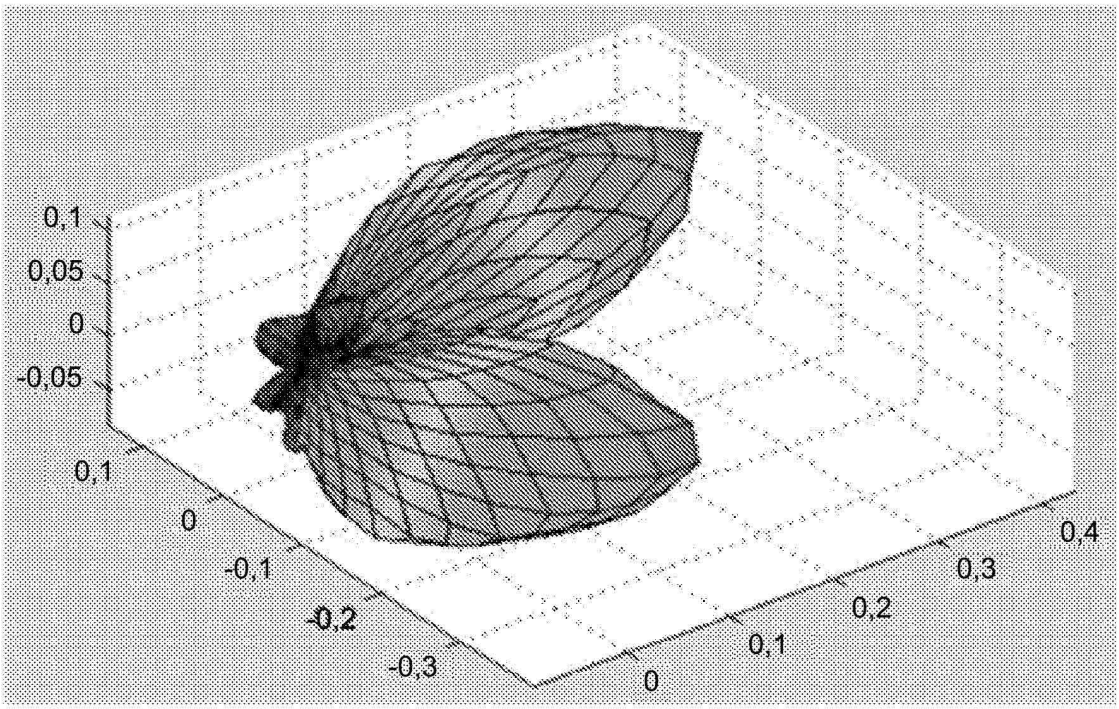


FIG. 36F

La gente está gritando en un estadio.

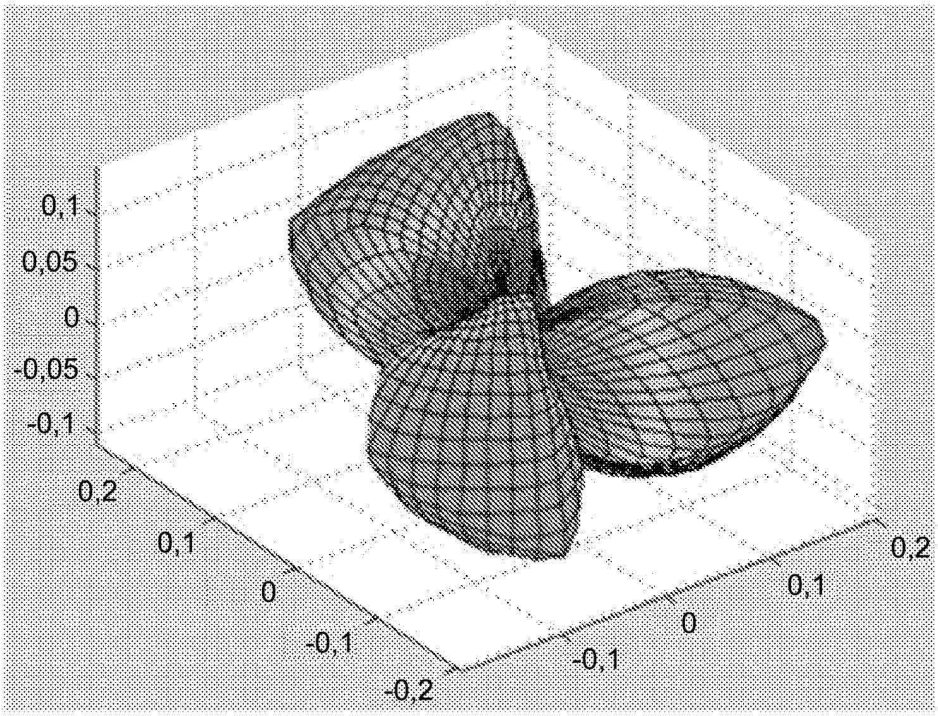


FIG. 36G

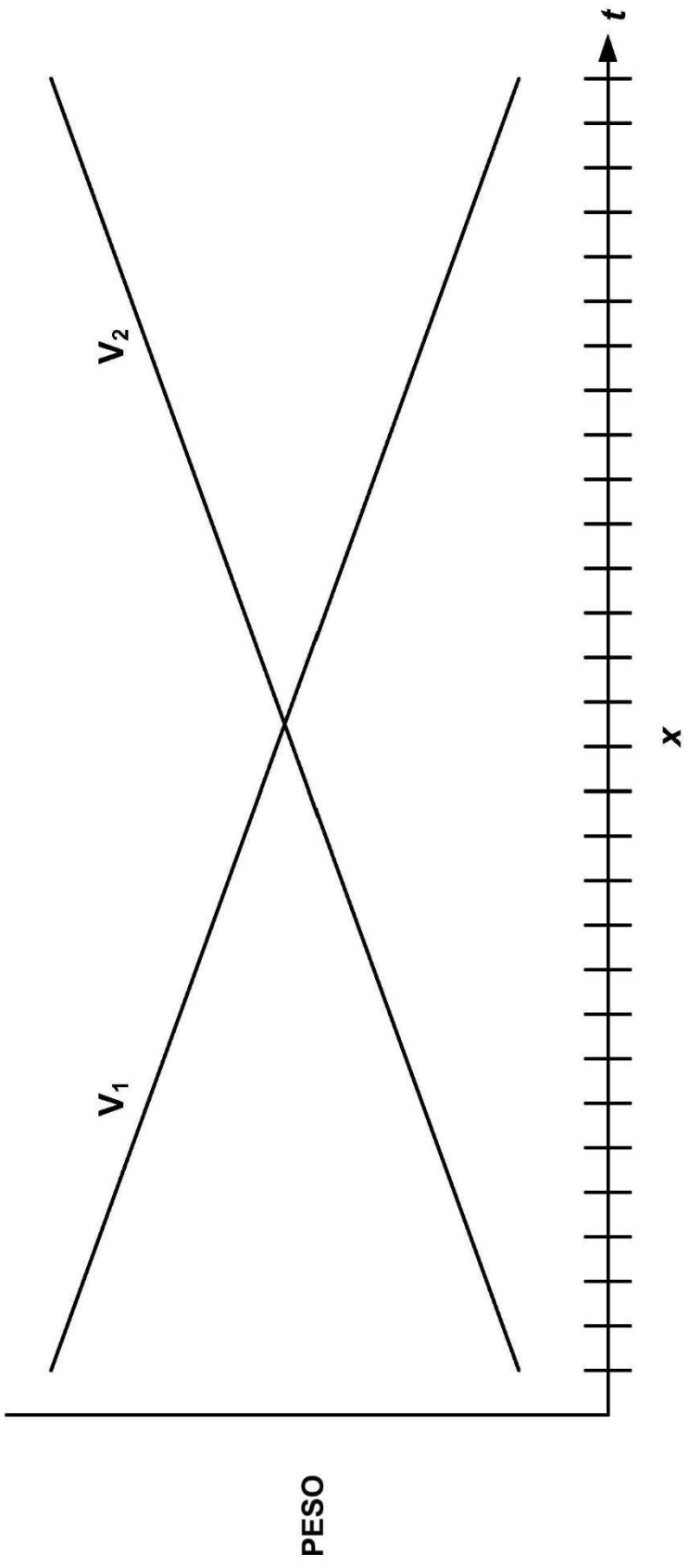


FIG. 37

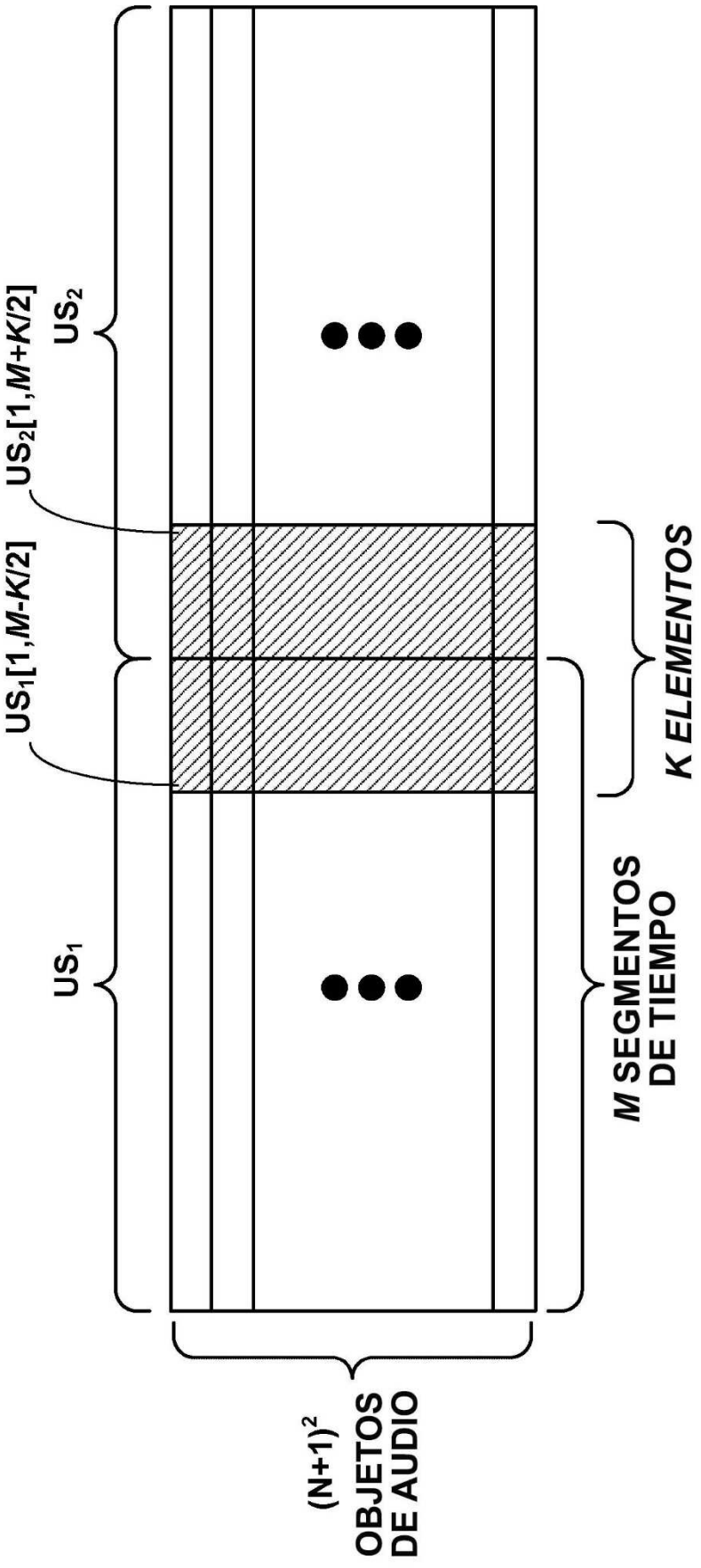

US_2
$US_2[1,M+K/2]$
$US_1[1,M-K/2]$
US_1
$(N+1)^2$ OBJETOS DE AUDIO
M SEGMENTOS DE TIEMPO
K ELEMENTOS

FIG. 38

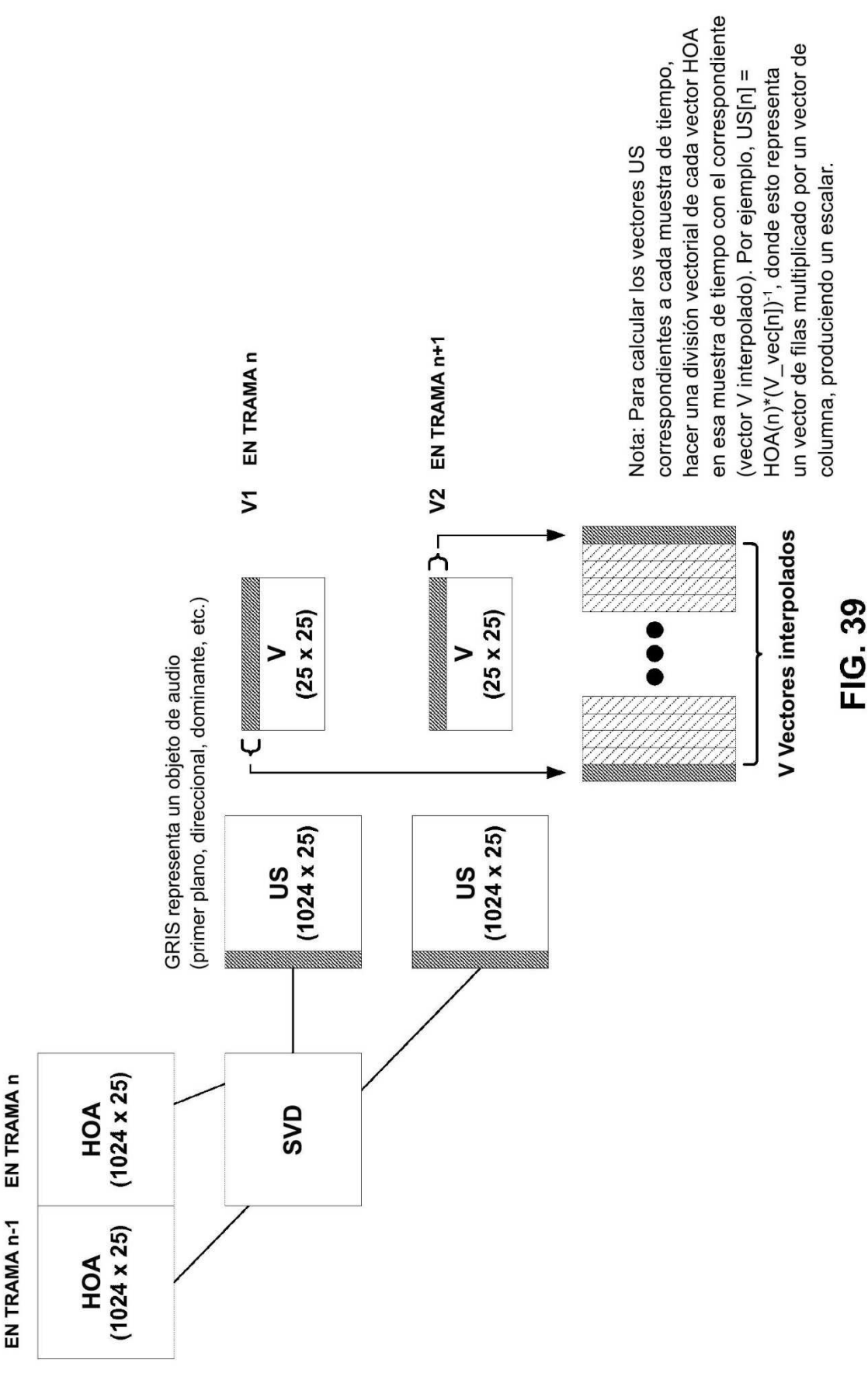

EN TRAMA n-1
EN TRAMA n
HOA (1024 x 25)
HOA (1024 x 25)
SVD
US (1024 x 25)
US (1024 x 25)
GRIS representa un objeto de audio (primer plano, direccional, dominante, etc.)
V1 EN TRAMA n
V (25 x 25)
V2 EN TRAMA n+1
V (25 x 25)
V Vectores interpolados
Nota: Para calcular los vectores US correspondientes a cada muestra de tiempo, hacer una división vectorial de cada vector HOA en esa muestra de tiempo con el correspondiente (vector V interpolado). Por ejemplo, US[n] = HOA(n)*(V_vec[n])-1, donde esto representa un vector de filas multiplicado por un vector de columna, produciendo un escalar.

FIG. 39

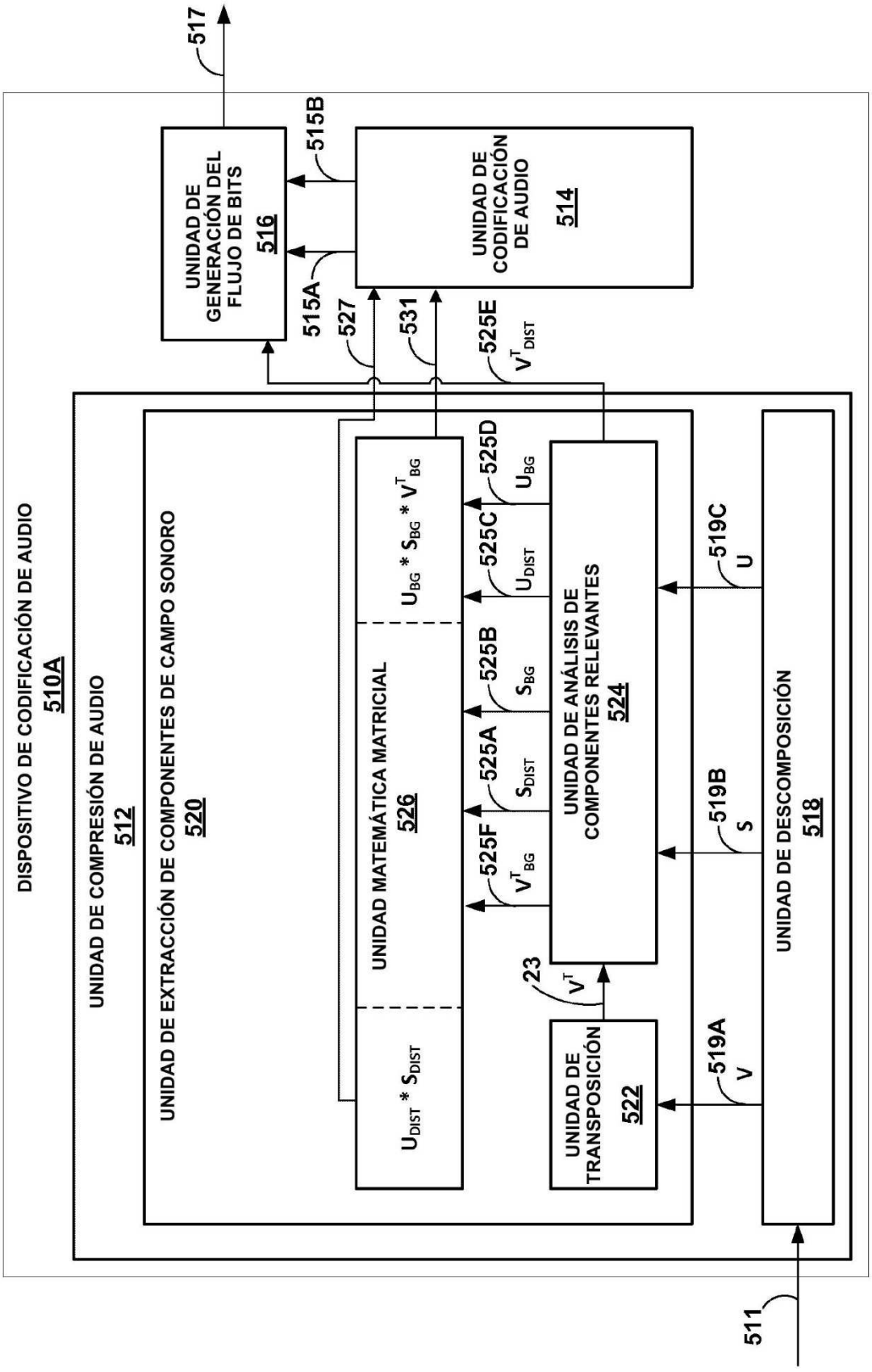

DISPOSITIVO DE CODIFICACIÓN DE AUDIO
510A
UNIDAD DE COMPRESIÓN DE AUDIO
512
UNIDAD DE EXTRACCIÓN DE COMPONENTES DE CAMPO SONORO
520
UNIDAD MATEMÁTICA MATRICIAL
526
$U_{DIST} * S_{DIST}$
$U_{BG} * S_{BG} * V^T_{BG}$
UNIDAD DE TRANSPOSICIÓN
522
UNIDAD DE ANÁLISIS DE COMPONENTES RELEVANTES
524
UNIDAD DE DESCOMPOSICIÓN
518
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS
516
UNIDAD DE CODIFICACIÓN DE AUDIO
514
511
517
515A
515B
519A
V
519B
S
519C
U
23
V^T
525A
S_{DIST}
525B
S_{BG}
525C
U_{DIST}
525D
U_{BG}
525E
V^T_{DIST}
525F
V^T_{BG}
527
531

FIG. 40A

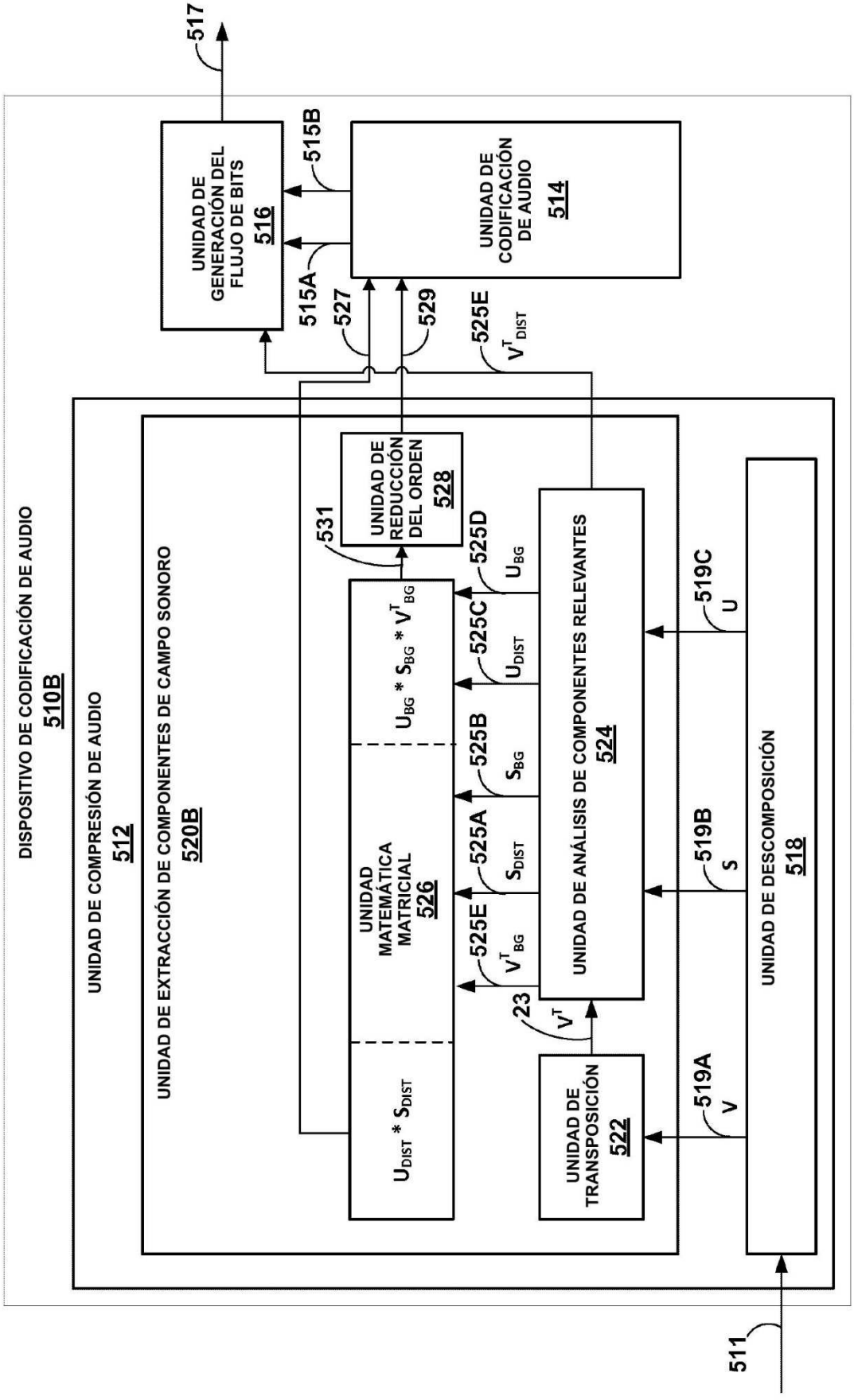

DISPOSITIVO DE CODIFICACIÓN DE AUDIO
510B
UNIDAD DE COMPRESIÓN DE AUDIO
512
UNIDAD DE EXTRACCIÓN DE COMPONENTES DE CAMPO SONORO
520B
UNIDAD MATEMÁTICA MATRICIAL
526
UNIDAD DE ANÁLISIS DE COMPONENTES RELEVANTES
524
UNIDAD DE TRANSPOSICIÓN
522
UNIDAD DE REDUCCIÓN DEL ORDEN
528
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS
516
UNIDAD DE CODIFICACIÓN DE AUDIO
514
UNIDAD DE DESCOMPOSICIÓN
518
511
517
515A
515B
519A
519B
519C
23
525A
525B
525C
525D
525E
527
529
531

FIG. 40B

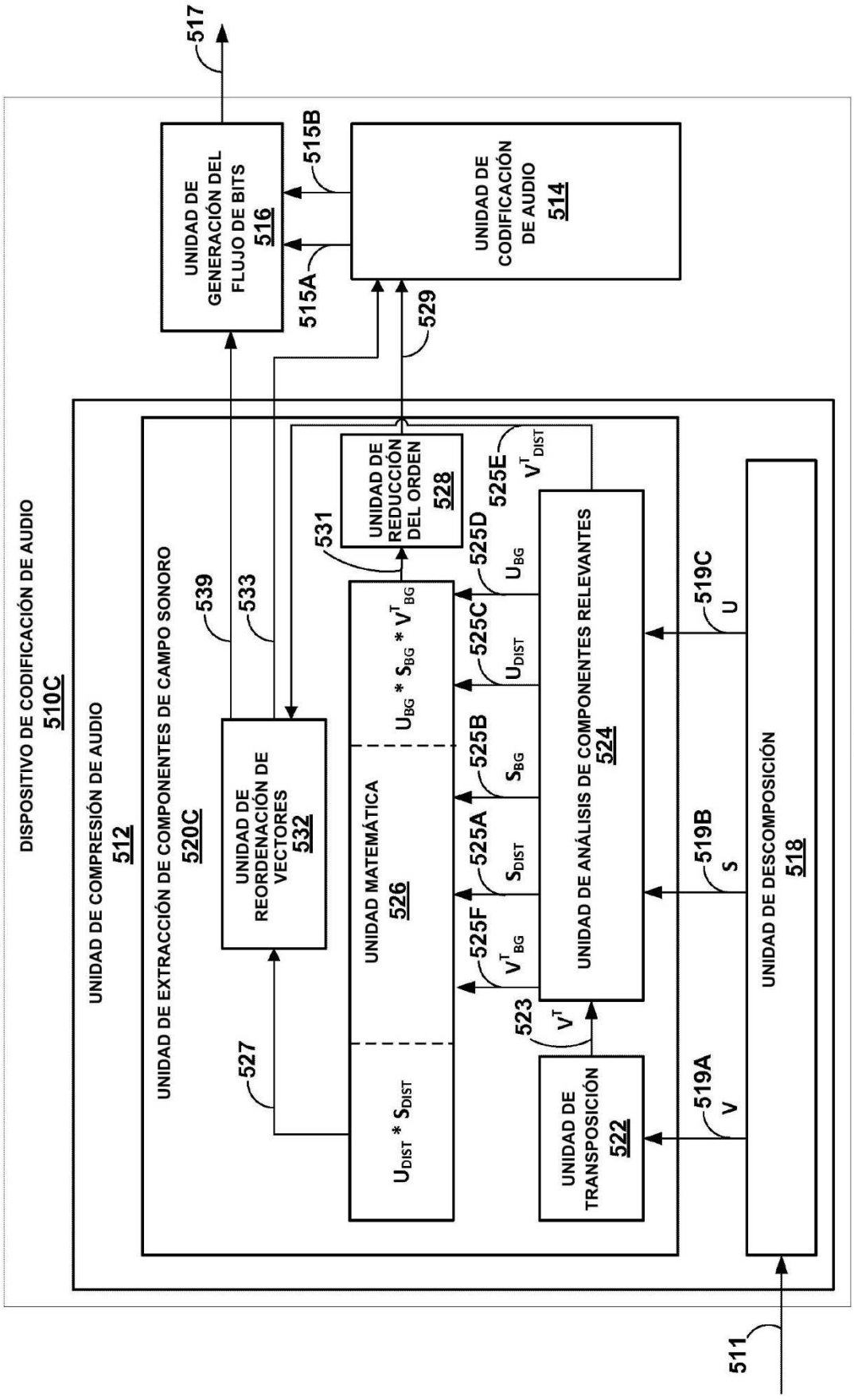

DISPOSITIVO DE CODIFICACIÓN DE AUDIO
510C
UNIDAD DE COMPRESIÓN DE AUDIO
512
UNIDAD DE EXTRACCIÓN DE COMPONENTES DE CAMPO SONORO
520C
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS
516
UNIDAD DE CODIFICACIÓN DE AUDIO
514
UNIDAD DE REORDENACIÓN DE VECTORES
532
UNIDAD MATEMÁTICA
526
UNIDAD DE REDUCCIÓN DEL ORDEN
528
UNIDAD DE ANÁLISIS DE COMPONENTES RELEVANTES
524
UNIDAD DE TRANSPOSICIÓN
522
UNIDAD DE DESCOMPOSICIÓN
518
517
515B
515A
529
531
539
533
527
523
519A
519B
519C
511
525A
525B
525C
525D
525E
525F
$U_{DIST} * S_{DIST}$
$U_{BG} * S_{BG} * V^T_{BG}$
S_{DIST}
S_{BG}
U_{DIST}
U_{BG}
V^T_{DIST}
V^T_{BG}
V^T
V
S
U

FIG. 40C

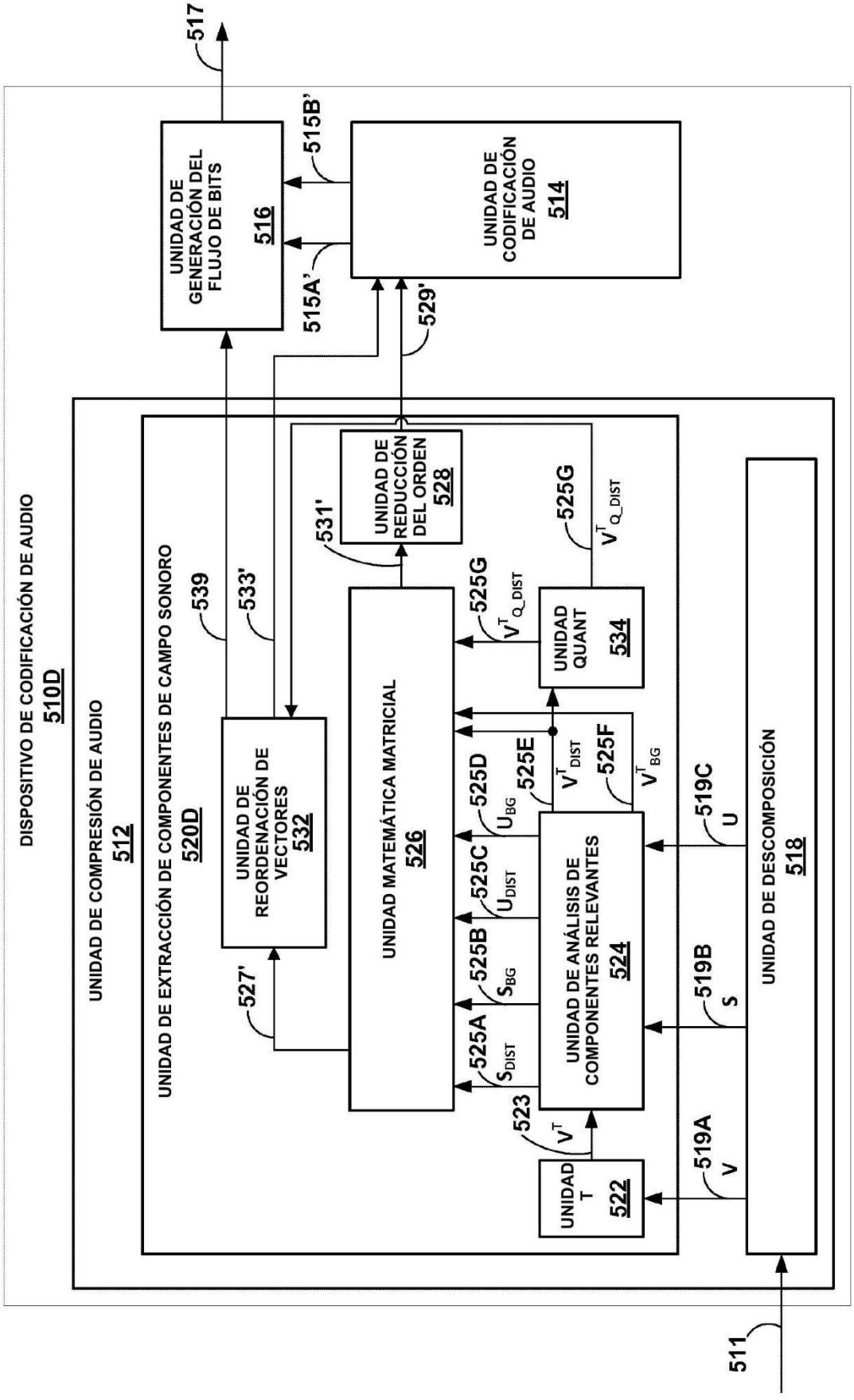

DISPOSITIVO DE CODIFICACIÓN DE AUDIO
510D
UNIDAD DE COMPRESIÓN DE AUDIO
512
UNIDAD DE EXTRACCIÓN DE COMPONENTES DE CAMPO SONORO
520D
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS
516
UNIDAD DE CODIFICACIÓN DE AUDIO
514
UNIDAD DE REORDENACIÓN DE VECTORES
532
UNIDAD MATEMÁTICA MATRICIAL
526
UNIDAD DE REDUCCIÓN DEL ORDEN
528
UNIDAD QUANT
534
UNIDAD DE ANÁLISIS DE COMPONENTES RELEVANTES
524
UNIDAD T
522
UNIDAD DE DESCOMPOSICIÓN
518
517
515B'
515A'
529'
531'
539
533'
527'
525G
525E
525F
525D
525C
525B
525A
523
519A
519B
519C
511

FIG. 40D

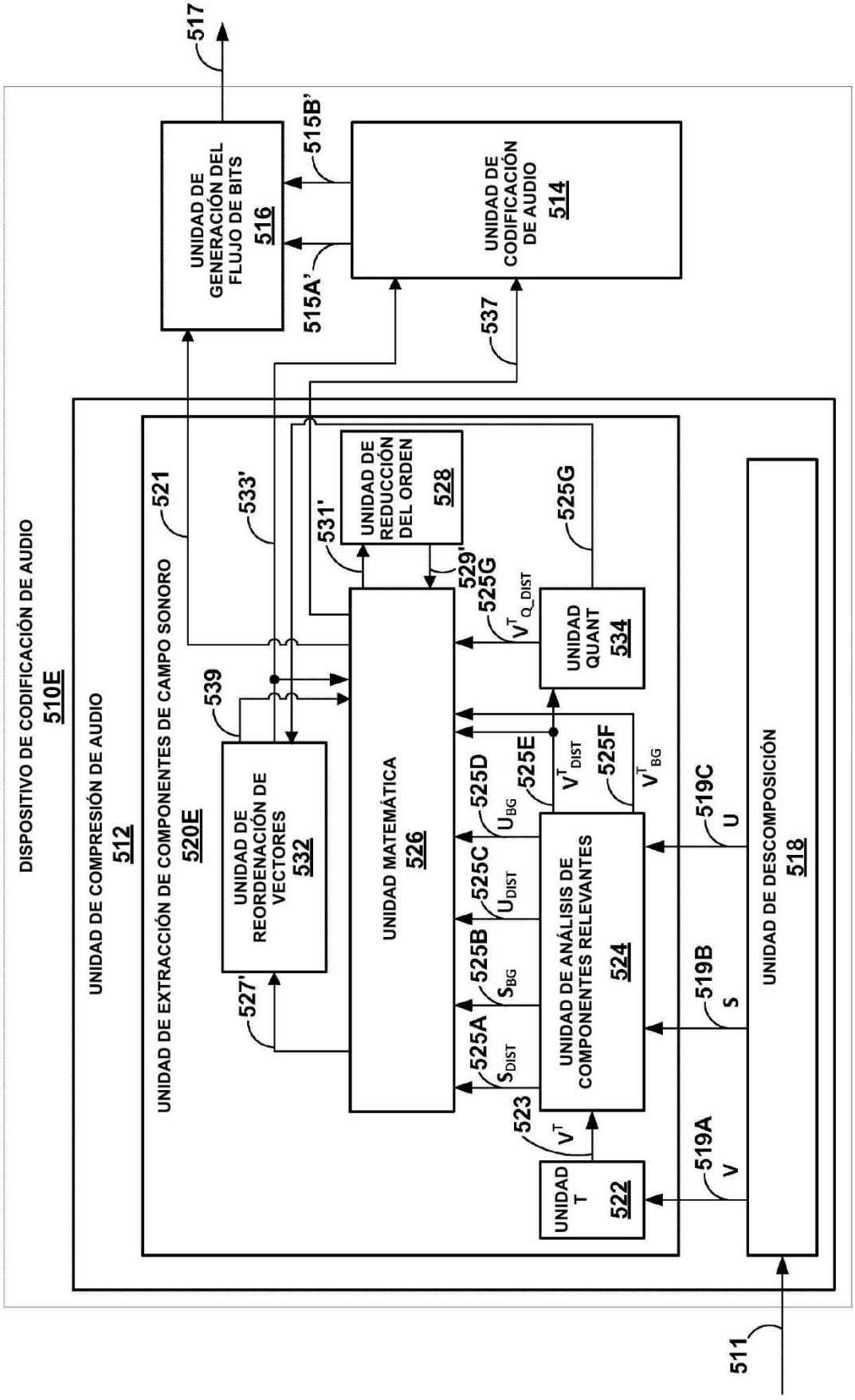

DISPOSITIVO DE CODIFICACIÓN DE AUDIO
510E
UNIDAD DE COMPRESIÓN DE AUDIO
512
UNIDAD DE EXTRACCIÓN DE COMPONENTES DE CAMPO SONORO
520E
521
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS
516
517
UNIDAD DE CODIFICACIÓN DE AUDIO
514
515B'
515A'
537
533'
531'
UNIDAD DE REDUCCIÓN DEL ORDEN
528
529'
525G
UNIDAD QUANT
534
539
UNIDAD DE REORDENACIÓN DE VECTORES
532
527'
UNIDAD MATEMÁTICA
526
525D
525E
525F
525C
525B
525A
UNIDAD DE ANÁLISIS DE COMPONENTES RELEVANTES
524
519C
519B
519A
523
UNIDAD T
522
UNIDAD DE DESCOMPOSICIÓN
518
511

FIG. 40E

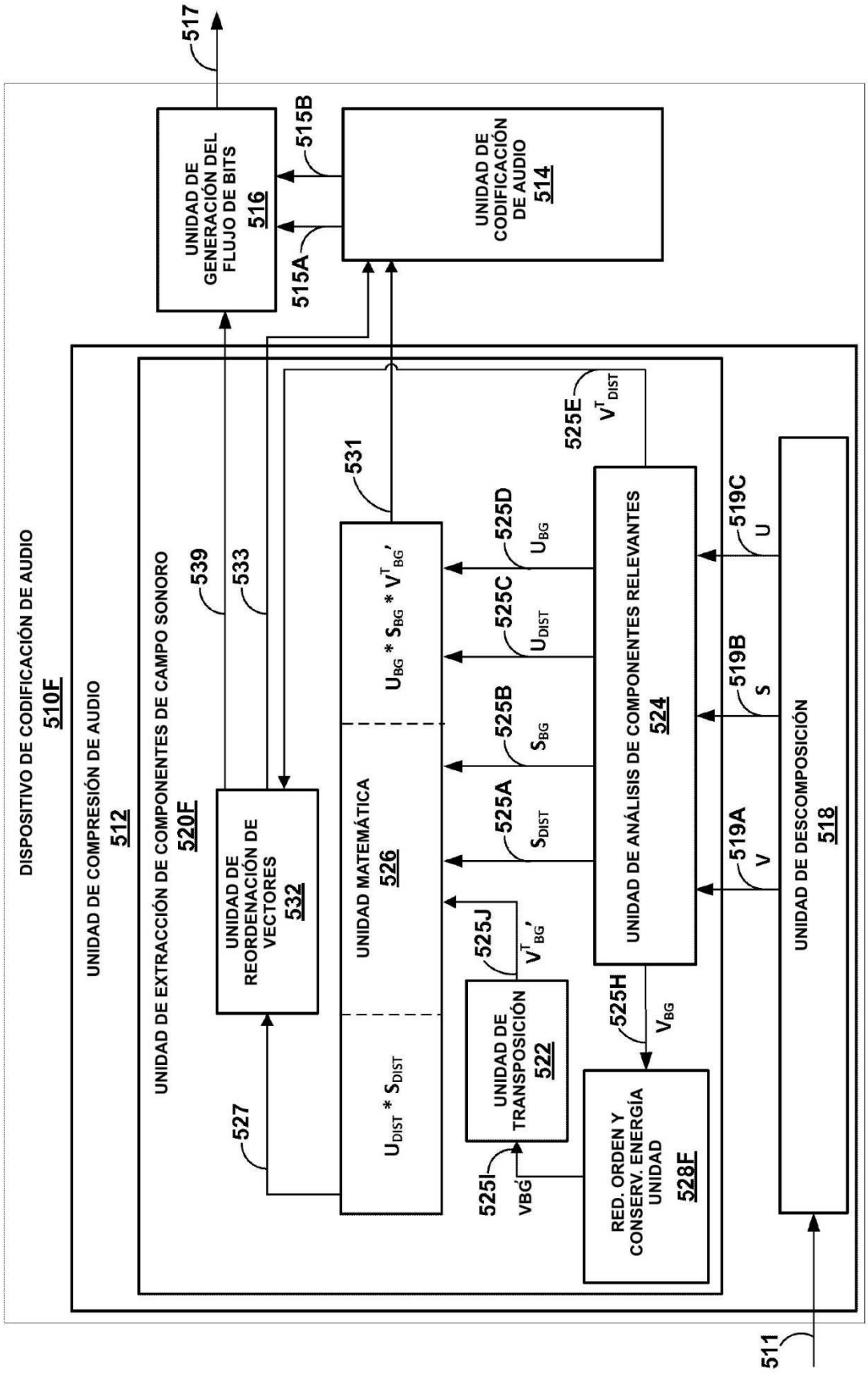

DISPOSITIVO DE CODIFICACIÓN DE AUDIO
510F
UNIDAD DE COMPRESIÓN DE AUDIO
512
UNIDAD DE EXTRACCIÓN DE COMPONENTES DE CAMPO SONORO
520F
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS
516
UNIDAD DE CODIFICACIÓN DE AUDIO
514
UNIDAD DE REORDENACIÓN DE VECTORES
532
UNIDAD MATEMÁTICA
526
UNIDAD DE TRANSPOSICIÓN
522
RED. ORDEN Y CONSERV. ENERGÍA UNIDAD
528F
UNIDAD DE ANÁLISIS DE COMPONENTES RELEVANTES
524
UNIDAD DE DESCOMPOSICIÓN
518
517
515A
515B
539
533
531
527
525A
525B
525C
525D
525E
525H
525I
525J
519A
519B
519C
511

FIG. 40F

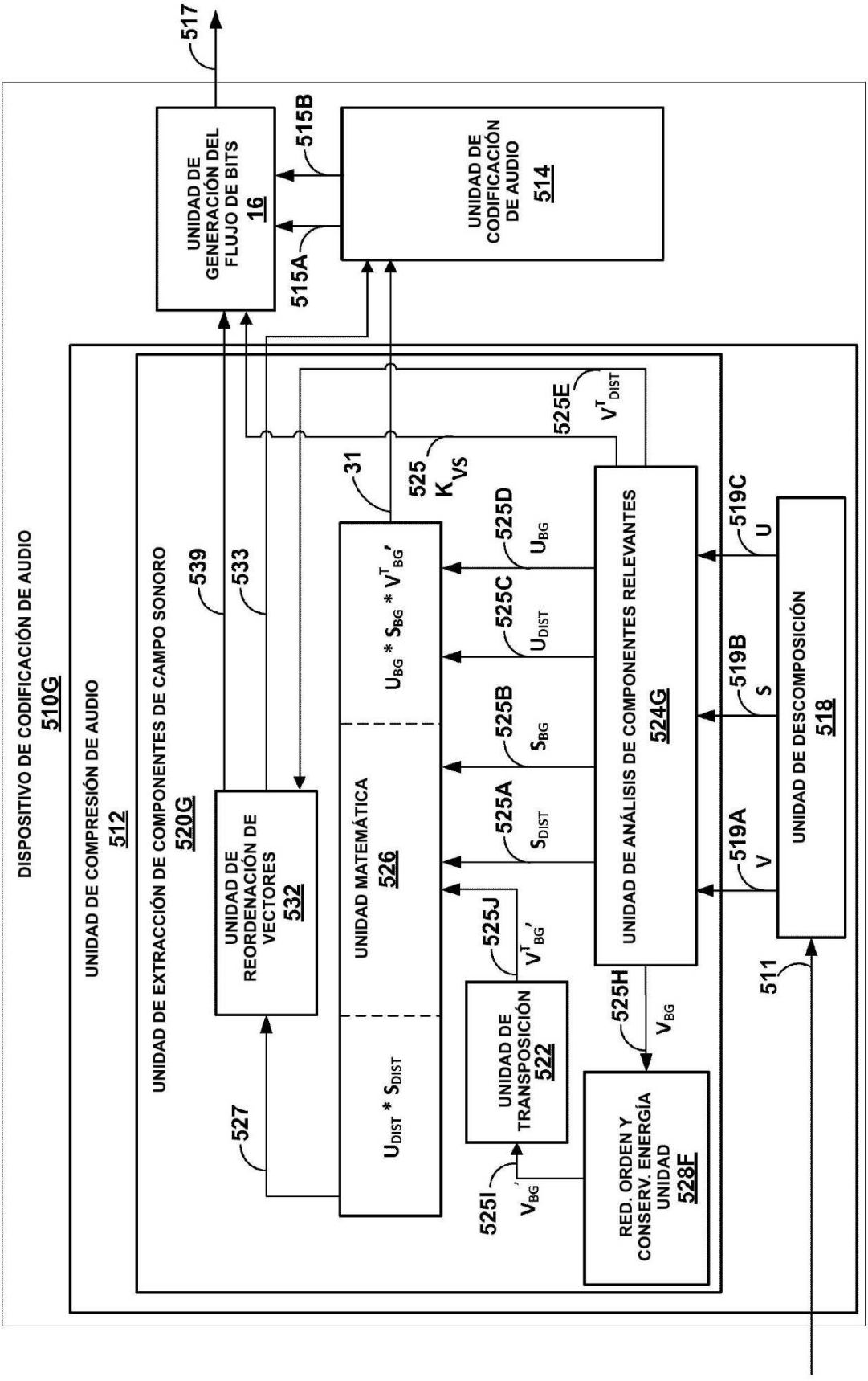

DISPOSITIVO DE CODIFICACIÓN DE AUDIO
510G
UNIDAD DE COMPRESIÓN DE AUDIO
512
UNIDAD DE EXTRACCIÓN DE COMPONENTES DE CAMPO SONORO
520G
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS
16
UNIDAD DE CODIFICACIÓN DE AUDIO
514
517
515A
515B
539
533
UNIDAD DE REORDENACIÓN DE VECTORES
532
UNIDAD MATEMÁTICA
526
31
525
K_{VS}
525E
V^T_{DIST}
525D
U_{BG}
525C
U_{DIST}
525B
S_{BG}
525A
S_{DIST}
$U_{BG} * S_{BG} * V^T_{BG}{}'$
$U_{DIST} * S_{DIST}$
UNIDAD DE ANÁLISIS DE COMPONENTES RELEVANTES
524G
519C
U
519B
S
519A
V
UNIDAD DE DESCOMPOSICIÓN
518
511
525J
$V^T_{BG}{}'$
UNIDAD DE TRANSPOSICIÓN
522
525H
V_{BG}
527
525I
$V_{BG}{}'$
RED. ORDEN Y CONSERV. ENERGÍA UNIDAD
528F

FIG. 40G

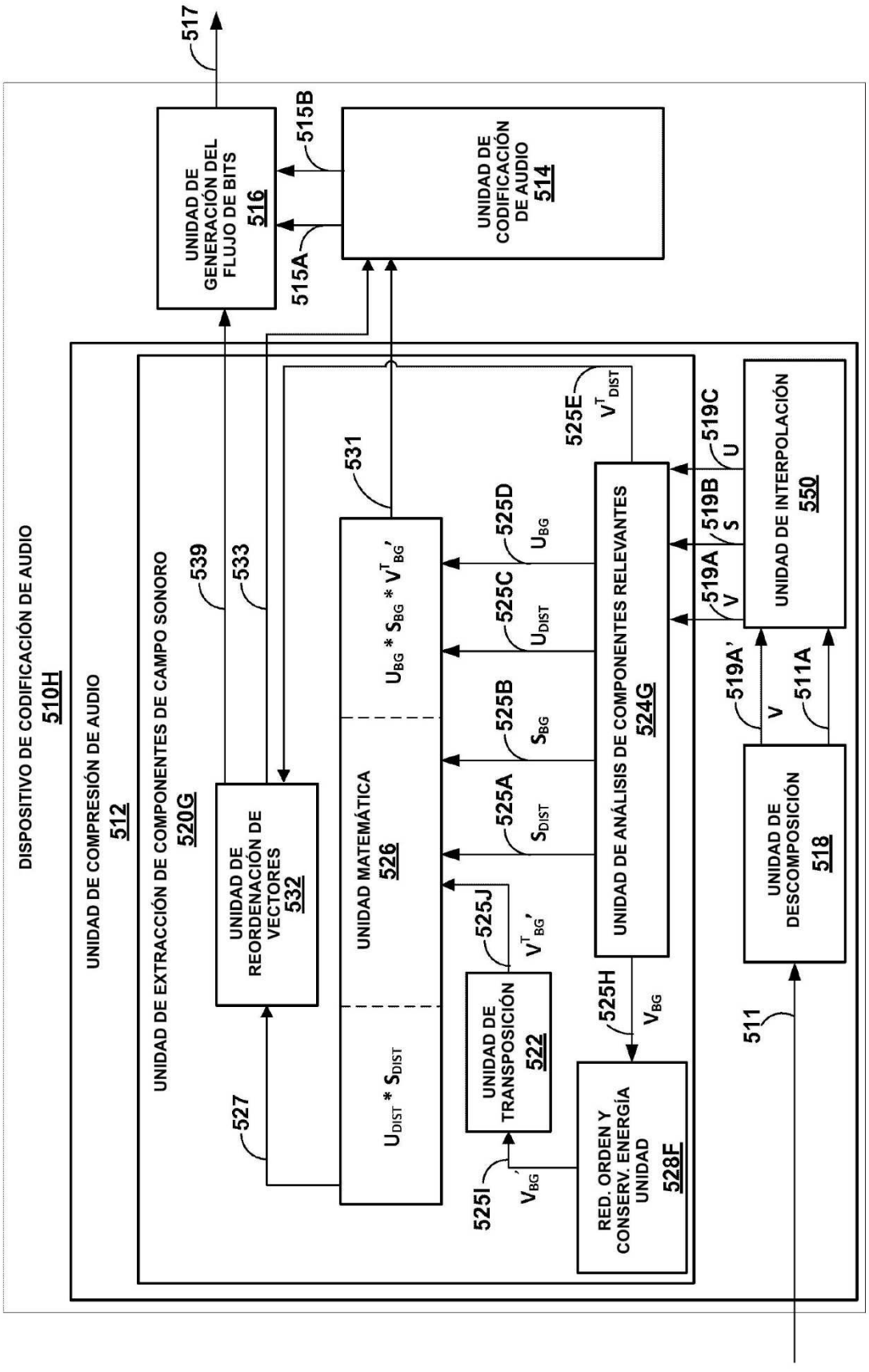

DISPOSITIVO DE CODIFICACIÓN DE AUDIO
510H
UNIDAD DE COMPRESIÓN DE AUDIO
512
UNIDAD DE EXTRACCIÓN DE COMPONENTES DE CAMPO SONORO
520G
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS
516
UNIDAD DE CODIFICACIÓN DE AUDIO
514
UNIDAD DE REORDENACIÓN DE VECTORES
532
UNIDAD MATEMÁTICA
526
UNIDAD DE TRANSPOSICIÓN
522
RED. ORDEN Y CONSERV. ENERGÍA UNIDAD
528F
UNIDAD DE ANÁLISIS DE COMPONENTES RELEVANTES
524G
UNIDAD DE DESCOMPOSICIÓN
518
UNIDAD DE INTERPOLACIÓN
550
517
515A
515B
531
533
539
527
525A
525B
525C
525D
525E
525H
525I
525J
519A
519A'
519B
519C
511
511A
$U_{DIST} * S_{DIST}$
$U_{BG} * S_{BG} * V^T_{BG}$'
S_{DIST}
S_{BG}
U_{DIST}
U_{BG}
V^T_{DIST}
V^T_{BG}'
V_{BG}'
V_{BG}
V
S
U

FIG. 40H

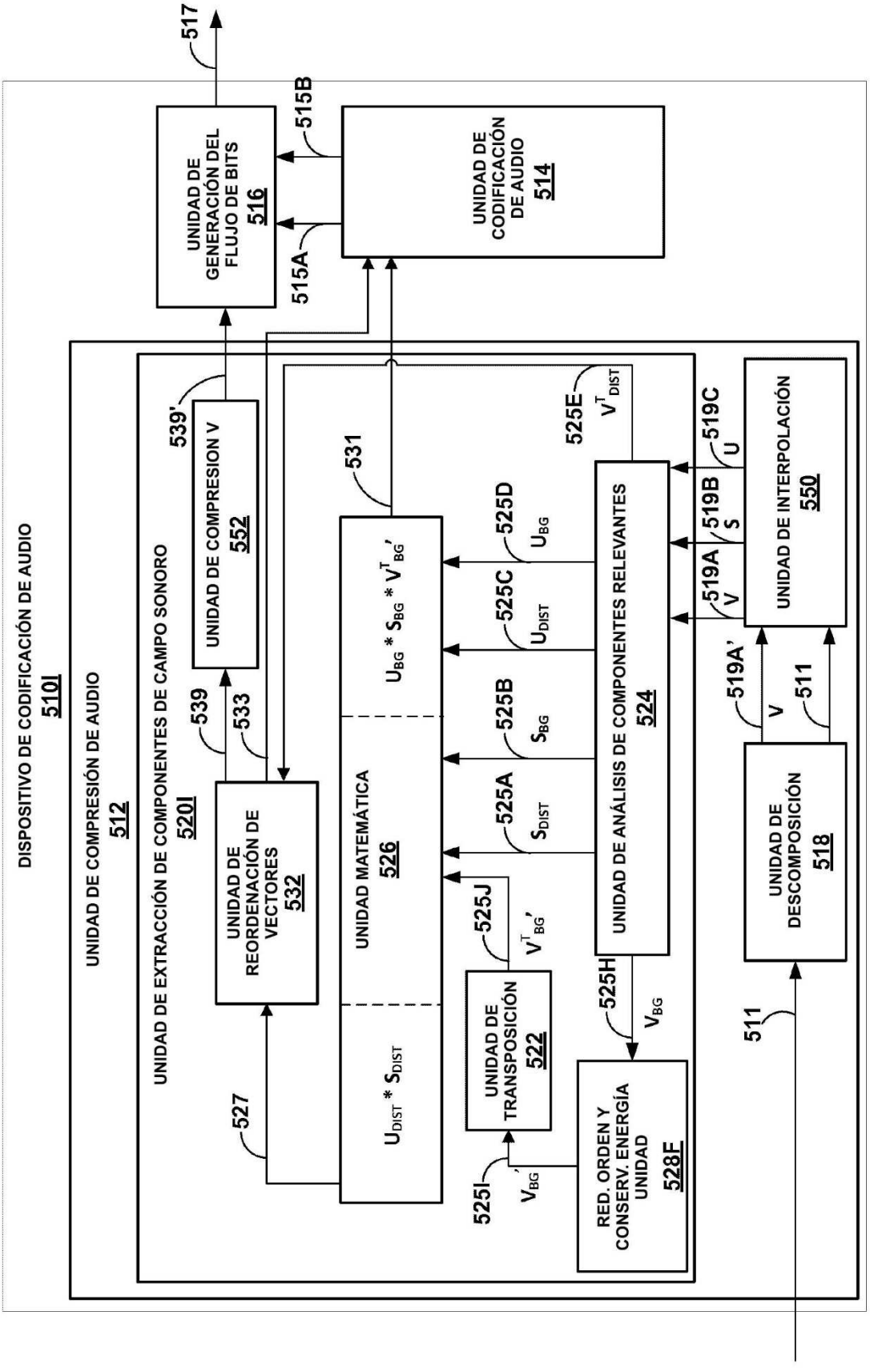

DISPOSITIVO DE CODIFICACIÓN DE AUDIO
510I
UNIDAD DE COMPRESIÓN DE AUDIO
512
UNIDAD DE EXTRACCIÓN DE COMPONENTES DE CAMPO SONORO
520I
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS
516
UNIDAD DE CODIFICACIÓN DE AUDIO
514
UNIDAD DE COMPRESION V
552
UNIDAD DE REORDENACIÓN DE VECTORES
532
UNIDAD MATEMÁTICA
526
UNIDAD DE TRANSPOSICIÓN
522
RED. ORDEN Y CONSERV. ENERGÍA UNIDAD
528F
UNIDAD DE ANÁLISIS DE COMPONENTES RELEVANTES
524
UNIDAD DE DESCOMPOSICIÓN
518
UNIDAD DE INTERPOLACIÓN
550
517
515A
515B
539'
539
533
527
531
525A
525B
525C
525D
525E
525H
525I
525J
519A
519A'
519B
519C
511

FIG. 40I

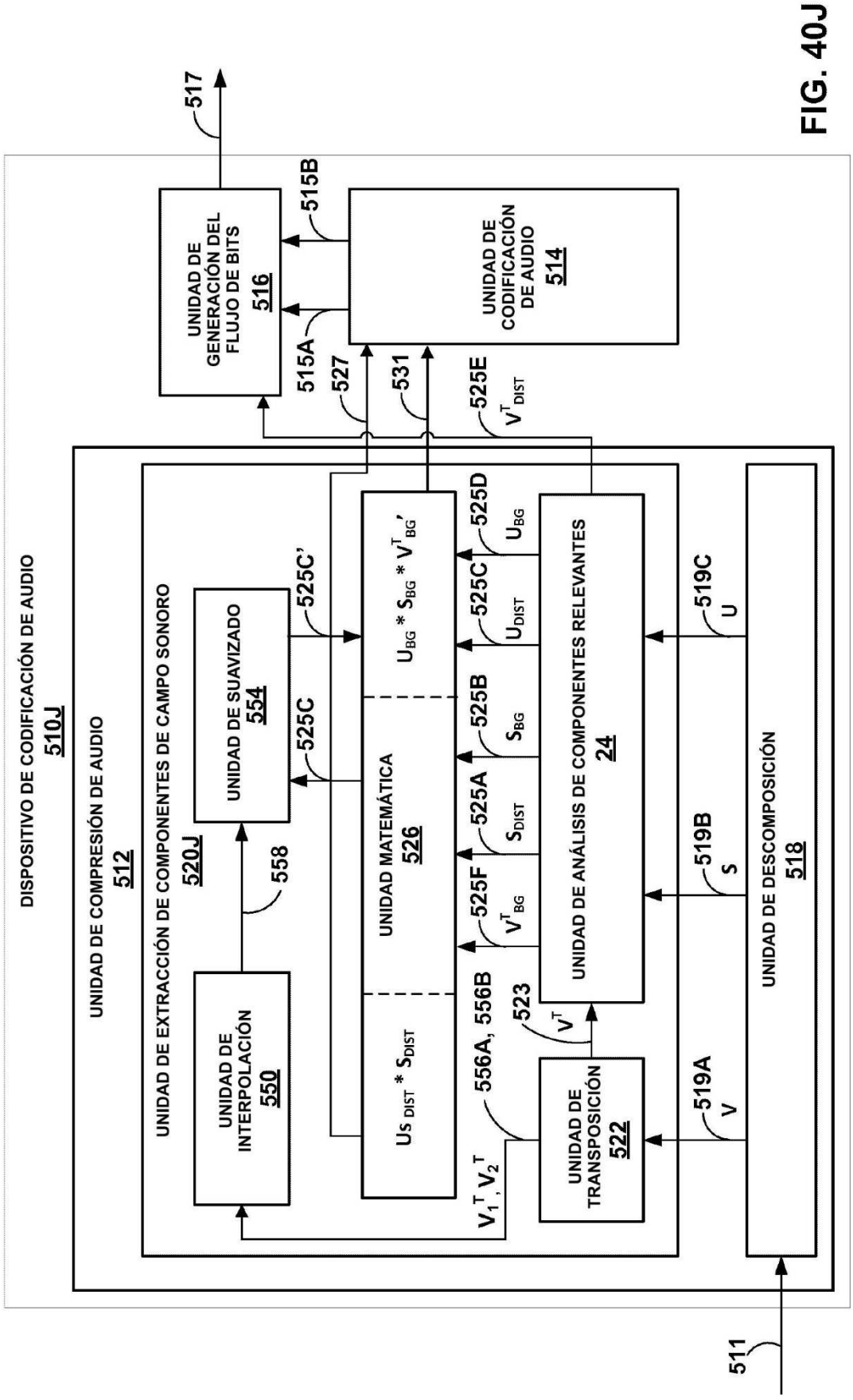

DISPOSITIVO DE CODIFICACIÓN DE AUDIO
510J
UNIDAD DE COMPRESIÓN DE AUDIO
512
UNIDAD DE EXTRACCIÓN DE COMPONENTES DE CAMPO SONORO
520J
UNIDAD DE INTERPOLACIÓN
550
UNIDAD DE SUAVIZADO
554
UNIDAD MATEMÁTICA
526
UNIDAD DE ANÁLISIS DE COMPONENTES RELEVANTES
24
UNIDAD DE TRANSPOSICIÓN
522
UNIDAD DE DESCOMPOSICIÓN
518
UNIDAD DE CODIFICACIÓN DE AUDIO
514
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS
516
511
517
515A
515B
527
531
525E
V^T_{DIST}
525C'
525C
525D
525C
525B
525A
525F
558
$U_{BG} * S_{BG} * V^T_{BG}{}'$
$Us_{DIST} * S_{DIST}$
U_{DIST}
S_{DIST}
S_{BG}
U_{BG}
V^T_{BG}
519C
U
519B
S
519A
V
523
V^T
556A, 556B
V^T_1, V^T_2

FIG. 40J

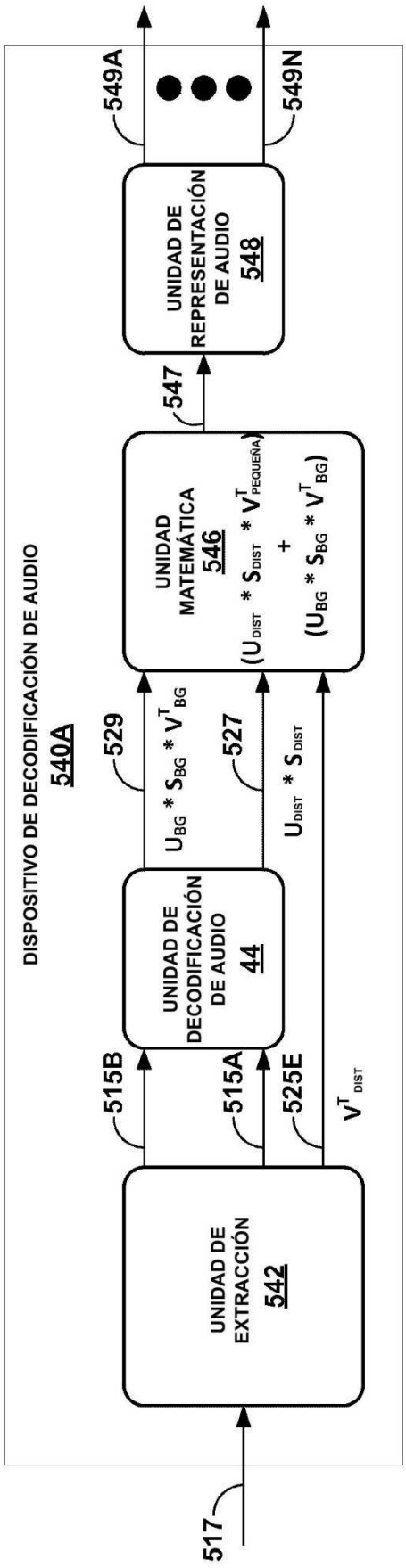

DISPOSITIVO DE DECODIFICACIÓN DE AUDIO
540A
517
UNIDAD DE EXTRACCIÓN
542
515B
515A
525E
V^T_{DIST}
UNIDAD DE DECODIFICACIÓN DE AUDIO
44
529
$U_{BG} * S_{BG} * V^T_{BG}$
527
$U_{DIST} * S_{DIST}$
UNIDAD MATEMÁTICA
546
$(U_{DIST} * S_{DIST} * V^T_{PEQUEÑA})$
+
$(U_{BG} * S_{BG} * V^T_{BG})$
547
UNIDAD DE REPRESENTACIÓN DE AUDIO
548
549A
549N

FIG. 41A

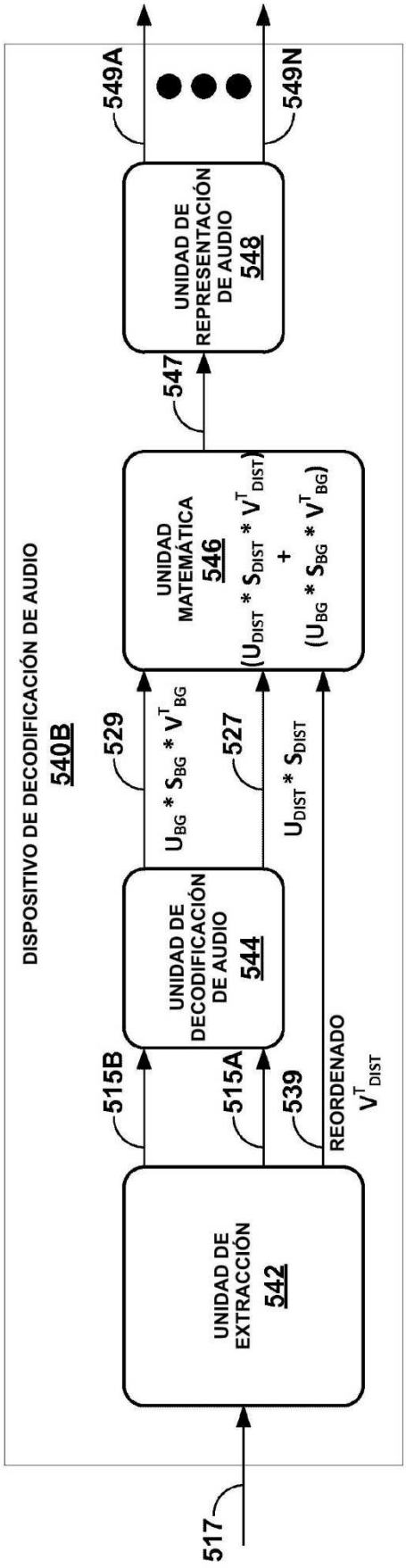

DISPOSITIVO DE DECODIFICACIÓN DE AUDIO
540B
UNIDAD DE EXTRACCIÓN
542
UNIDAD DE DECODIFICACIÓN DE AUDIO
544
UNIDAD MATEMÁTICA
546
UNIDAD DE REPRESENTACIÓN DE AUDIO
548
517
515A
515B
539
529
527
547
549A
549N
REORDENADO
V^T_{DIST}
$U_{BG} * S_{BG} * V^T_{BG}$
$U_{DIST} * S_{DIST}$
$(U_{DIST} * S_{DIST} * V^T_{DIST}) + (U_{BG} * S_{BG} * V^T_{BG})$

FIG. 41B

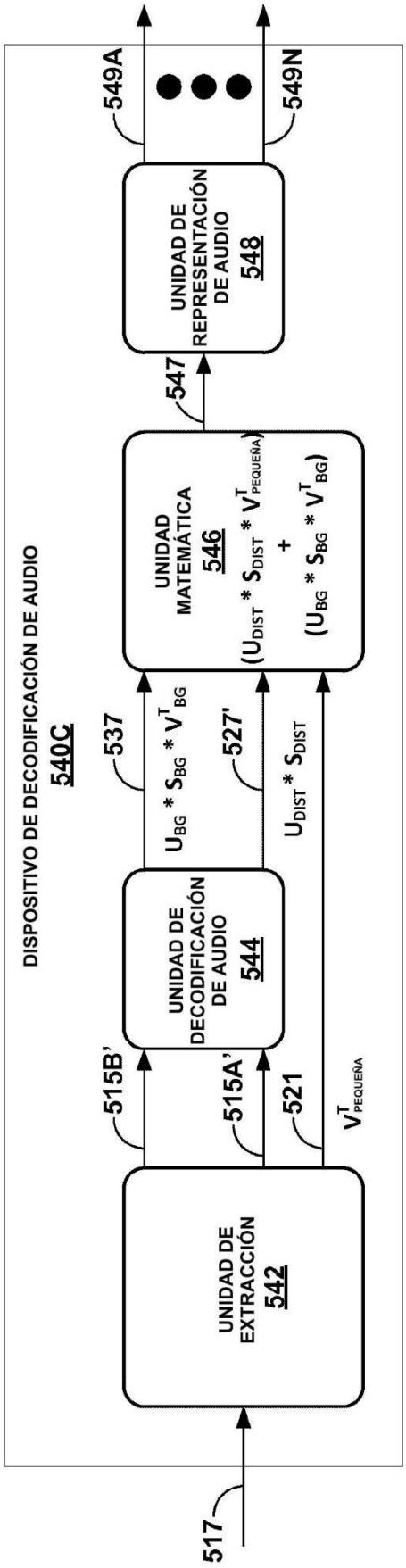

DISPOSITIVO DE DECODIFICACIÓN DE AUDIO
540C
UNIDAD DE EXTRACCIÓN
542
UNIDAD DE DECODIFICACIÓN DE AUDIO
544
UNIDAD MATEMÁTICA
546
$(U_{DIST} * S_{DIST} * V^T_{PEQUEÑA}) + (U_{BG} * S_{BG} * V^T_{BG})$
UNIDAD DE REPRESENTACIÓN DE AUDIO
548
517
515B'
515A'
521
537
527'
547
549A
549N
$V^T_{PEQUEÑA}$
$U_{DIST} * S_{DIST}$
$U_{BG} * S_{BG} * V^T_{BG}$

FIG. 41C

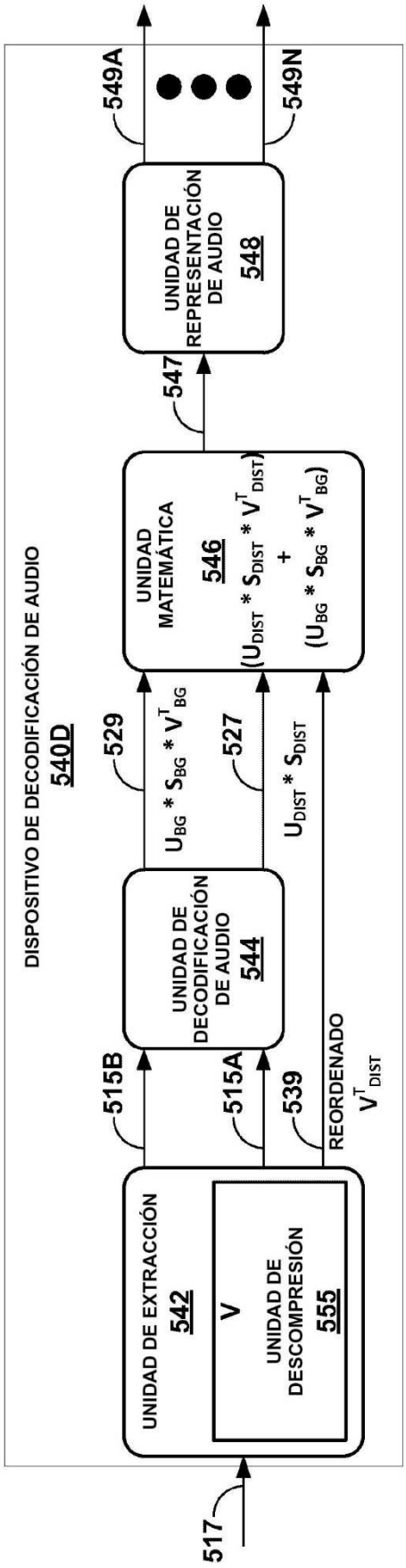

DISPOSITIVO DE DECODIFICACIÓN DE AUDIO
540D
517
UNIDAD DE EXTRACCIÓN
542
V
UNIDAD DE DESCOMPRESIÓN
555
515B
515A
539
REORDENADO
V^T_{DIST}
UNIDAD DE DECODIFICACIÓN DE AUDIO
544
529
$U_{BG} * S_{BG} * V^T_{BG}$
527
$U_{DIST} * S_{DIST}$
UNIDAD MATEMÁTICA
546
$(U_{DIST} * S_{DIST} * V^T_{DIST})$
+
$(U_{BG} * S_{BG} * V^T_{BG})$
547
UNIDAD DE REPRESENTACIÓN DE AUDIO
548
549A
549N

FIG. 41D



FIG. 42A

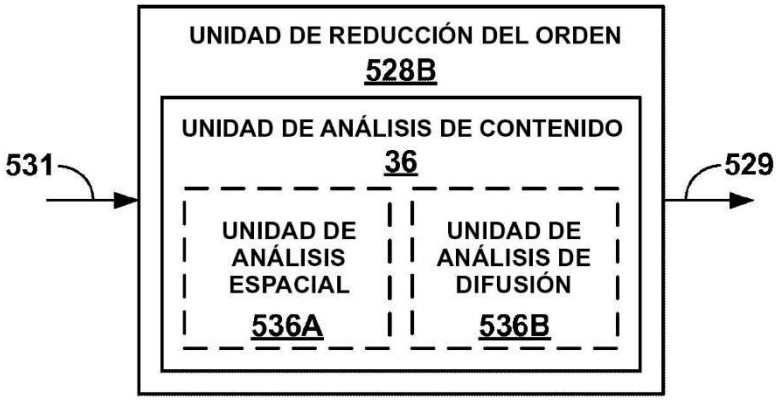


FIG. 42B

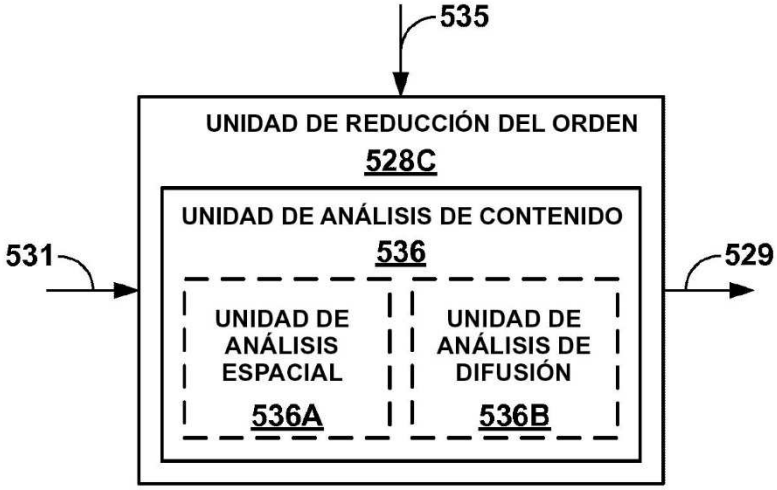


FIG. 42C

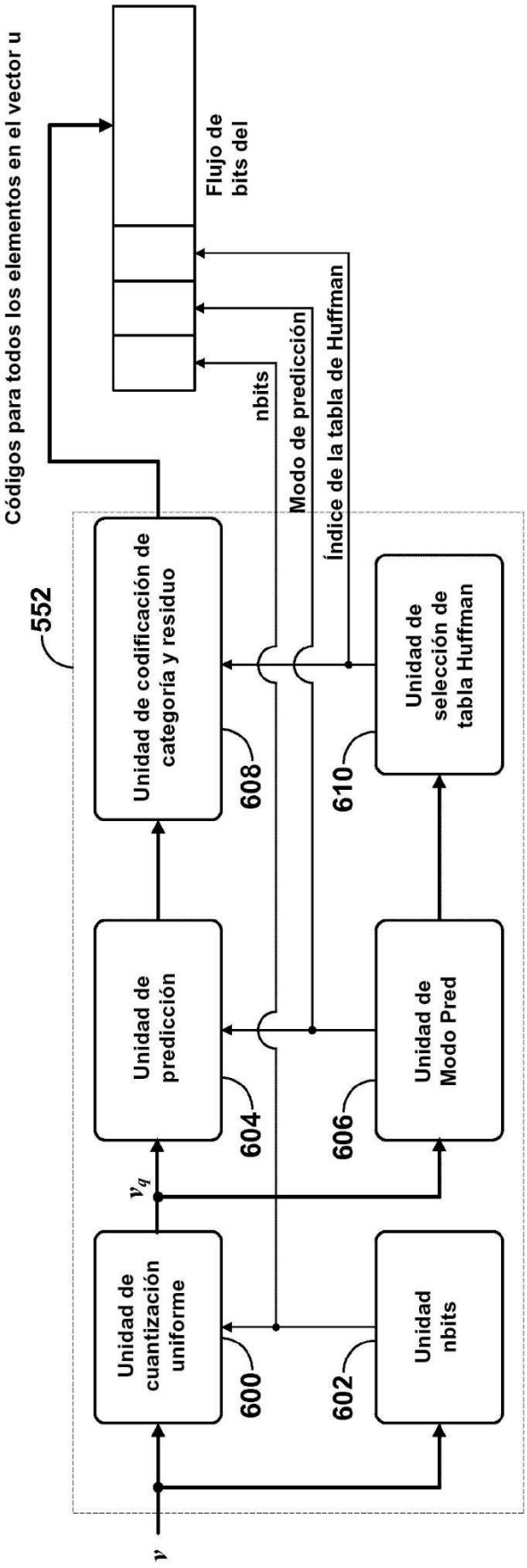

Códigos para todos los elementos en el vector u
Flujo de bits del
nbits
Modo de predicción
Índice de la tabla de Huffman
552
Unidad de codificación de categoría y residuo
608
Unidad de selección de tabla Huffman
610
Unidad de predicción
604
Unidad de Modo Pred
606
v_q
Unidad de cuantización uniforme
600
Unidad nbits
602
v

FIG. 43

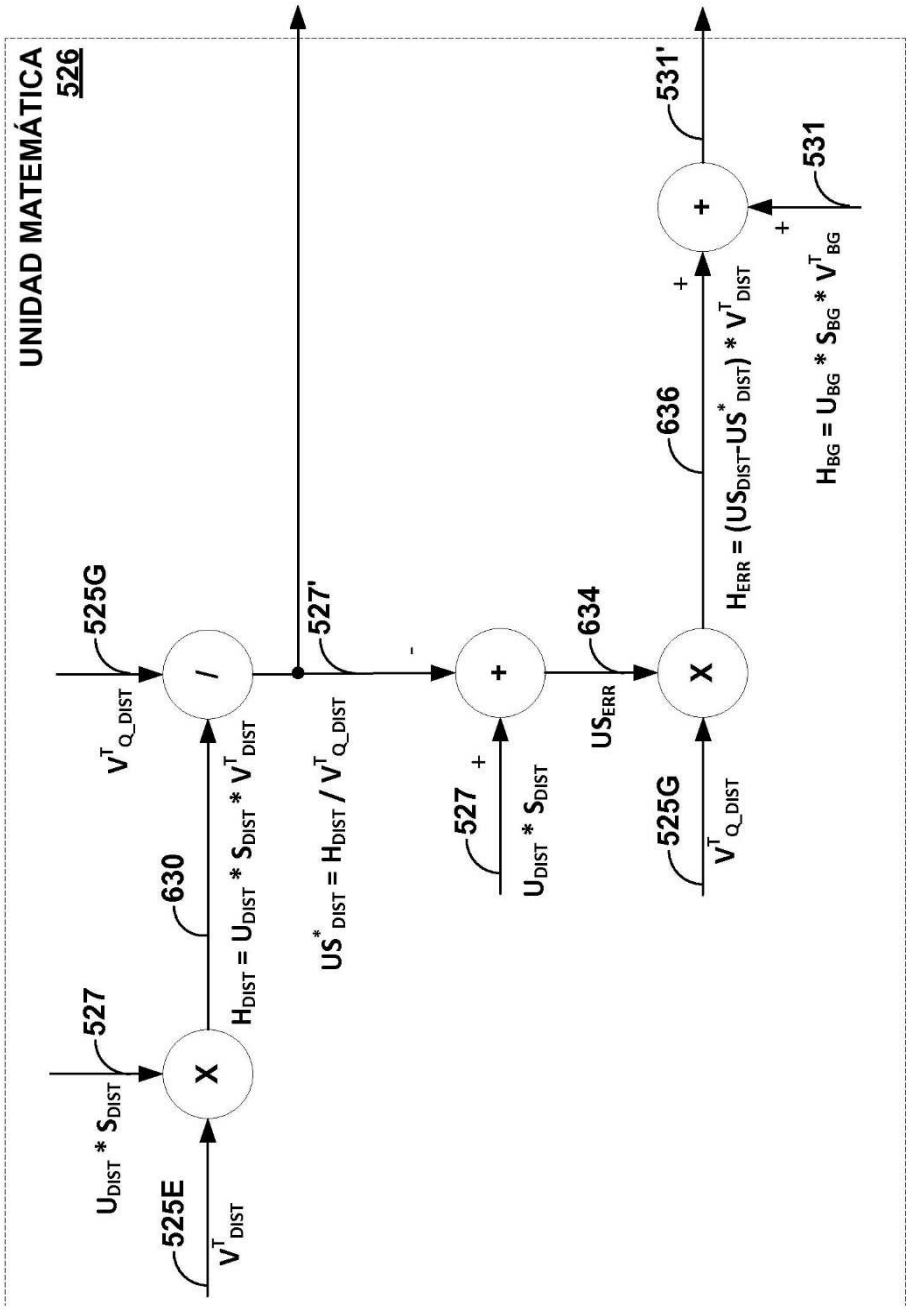

UNIDAD MATEMÁTICA
526
525E
V^T_{DIST}
$U_{DIST} * S_{DIST}$
527
X
630
$H_{DIST} = U_{DIST} * S_{DIST} * V^T_{DIST}$
525G
$V^T_{Q_DIST}$
/
527'
$US^*_{DIST} = H_{DIST} / V^T_{Q_DIST}$
-
527
+
$U_{DIST} * S_{DIST}$
+
634
US_{ERR}
525G
$V^T_{Q_DIST}$
X
636
$H_{ERR} = (US_{DIST} - US^*_{DIST}) * V^T_{DIST}$
+
531'
531
+
$H_{BG} = U_{BG} * S_{BG} * V^T_{BG}$

FIG. 44

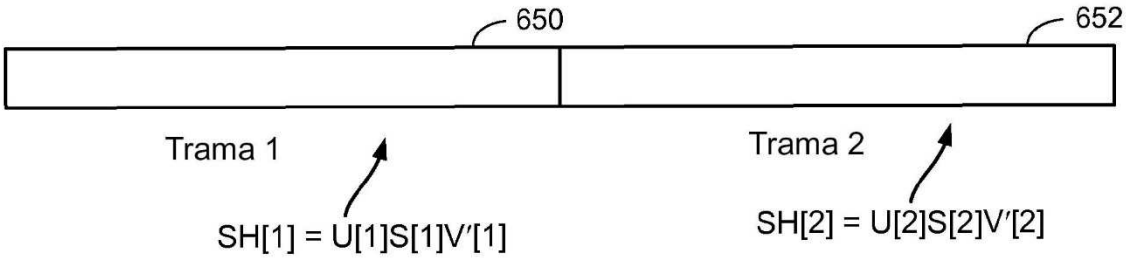
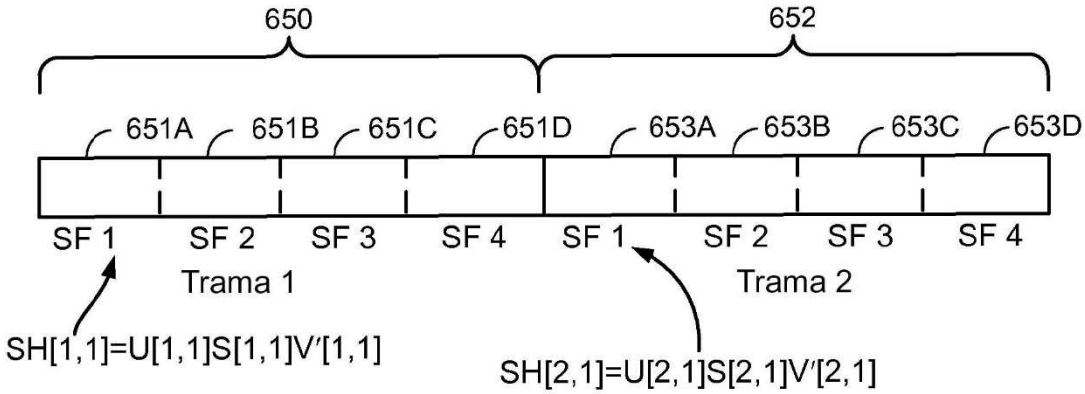


FIG. 45A



V'[1,2] = interpolación(V'[1,1], V'[2,1]) → U[1,2]S[1,2] = SH[1,2](V'[1,2])-1

V'[1,3] = interpolación(V'[1,1], V'[2,1]) → U[1,3]S[1,3] = SH[1,3](V'[1,3])-1

V'[1,4] = interpolación(V'[1,1], V'[2,1]) → U[1,3]S[1,3] = SH[1,3](V'[1,3])-1

FIG. 45B

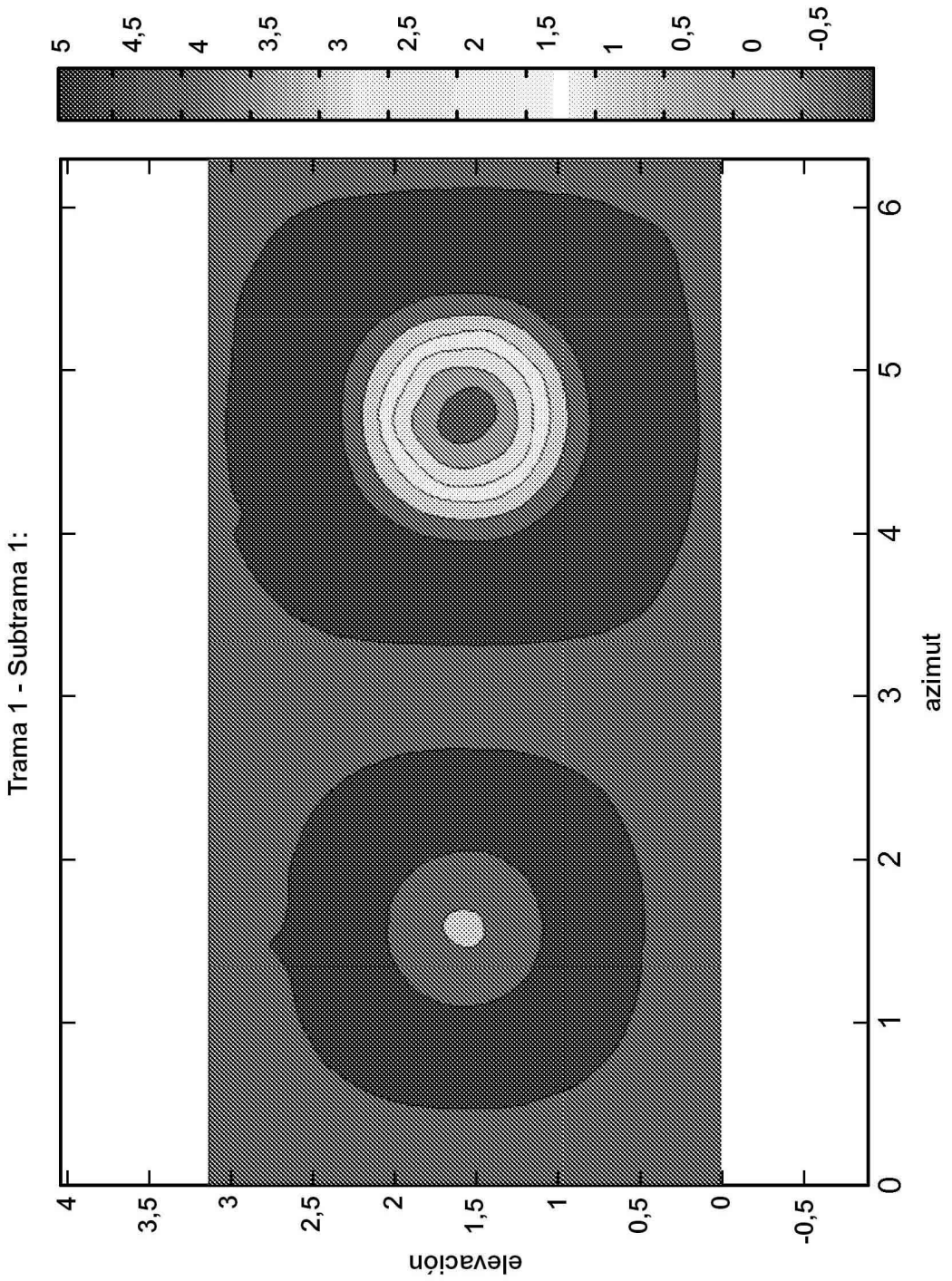

Trama 1 - Subtrama 1:
elevación
azimut
5
4,5
4
3,5
3
2,5
2
1,5
1
0,5
0
-0,5
4
3,5
3
2,5
2
1,5
1
0,5
0
-0,5
0
1
2
3
4
5
6

FIG. 46A

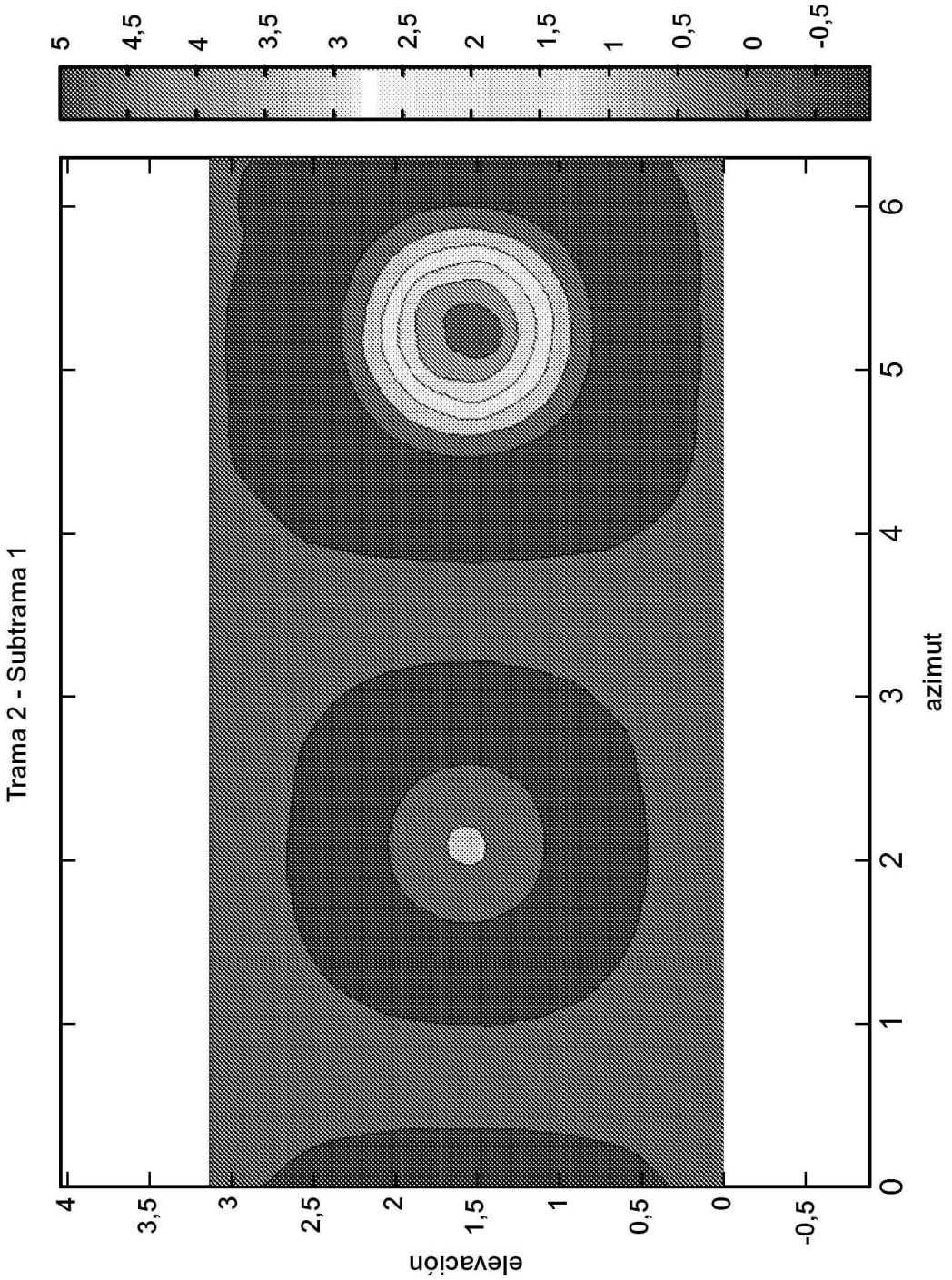

Trama 2 - Subtrama 1
elevación
azimut
4
3,5
3
2,5
2
1,5
1
0,5
0
-0,5
0
1
2
3
4
5
6
5
4,5
4
3,5
3
2,5
2
1,5
1
0,5
0
-0,5

FIG. 46B

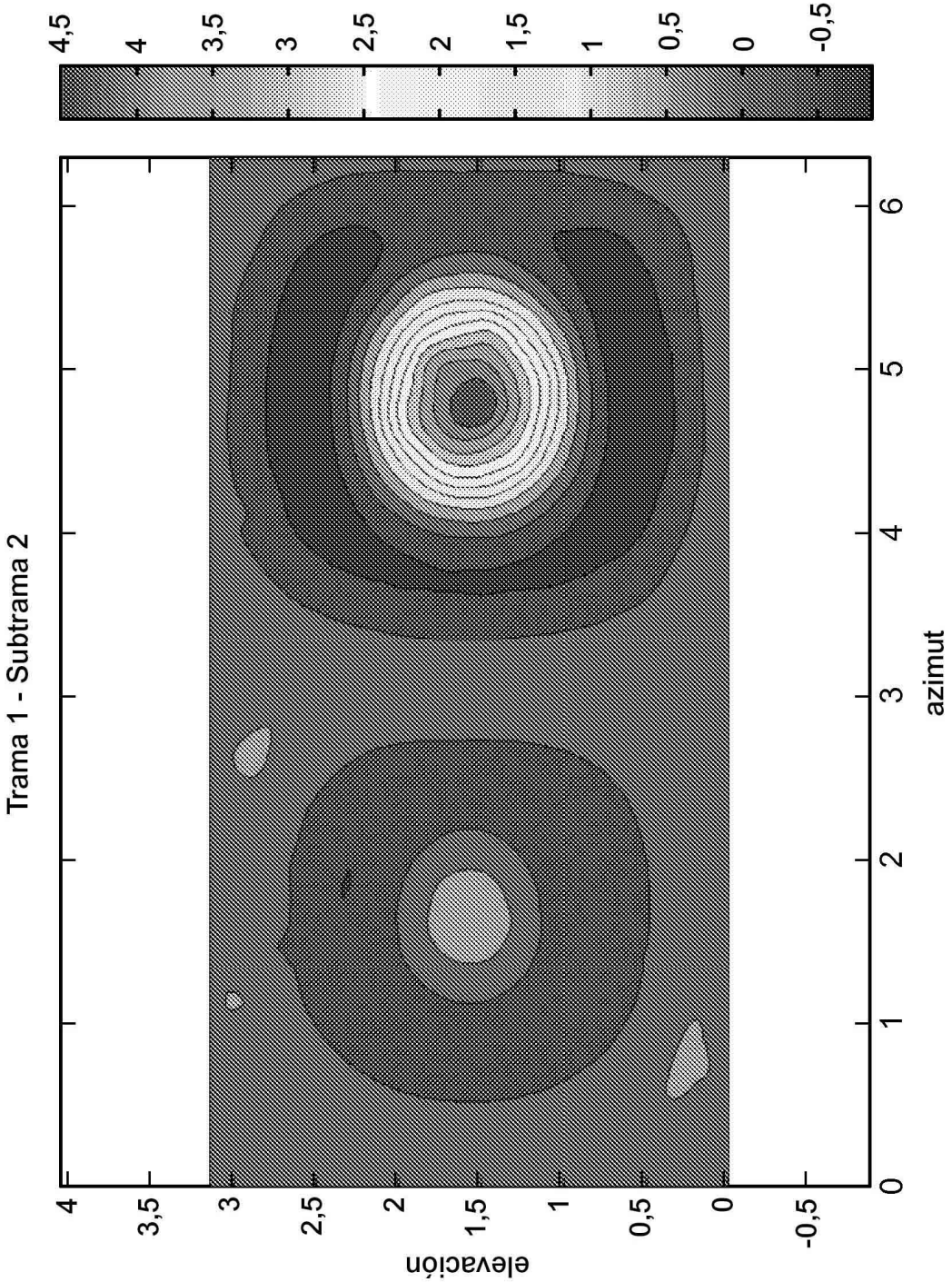

Trama 1 - Subtrama 2
elevación
azimut
4,5
4
3,5
3
2,5
2
1,5
1
0,5
0
-0,5
0
1
2
3
4
5
6

FIG. 46C

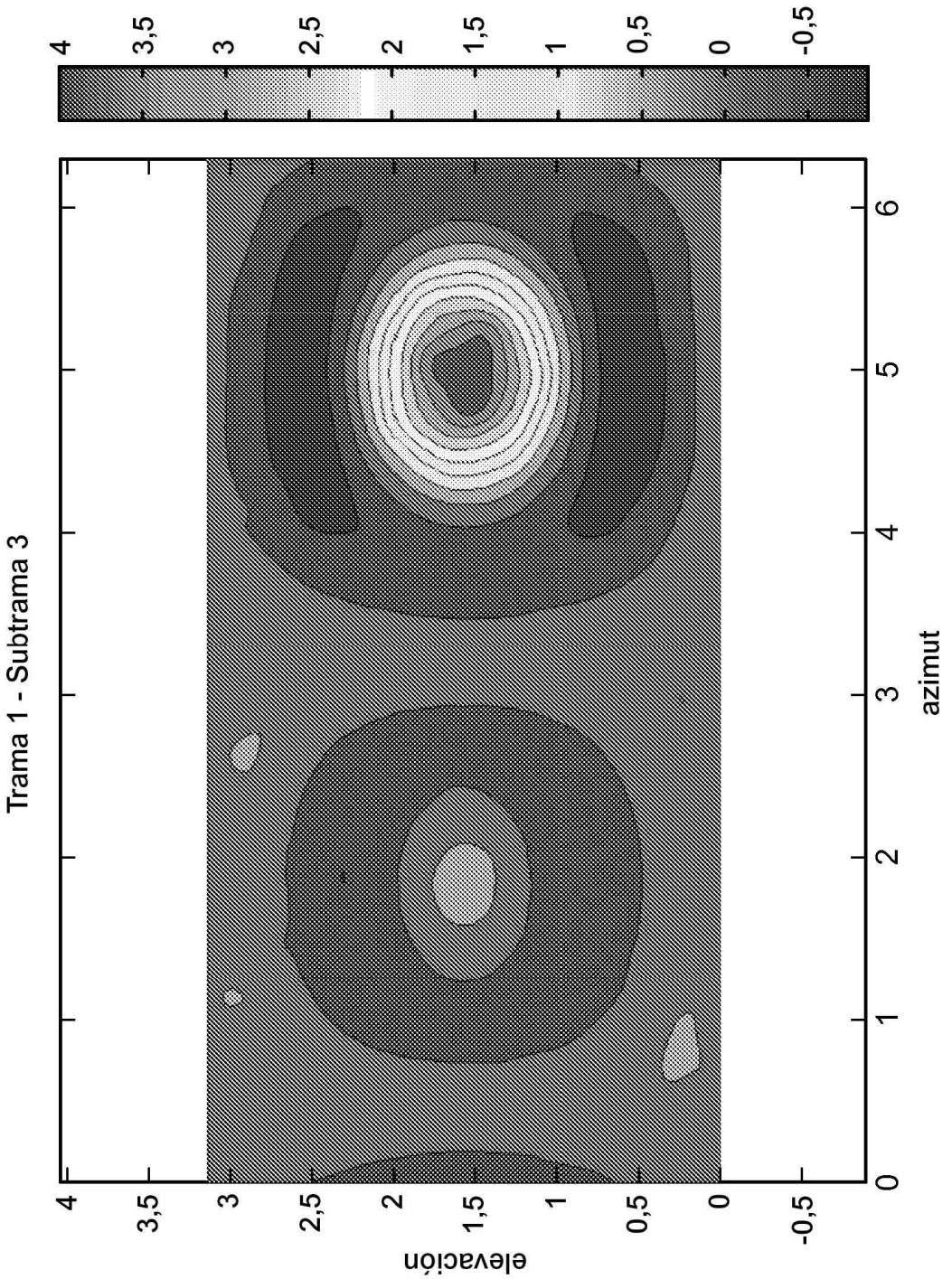

Trama 1 - Subtrama 3
elevación
azimut
4
3,5
3
2,5
2
1,5
1
0,5
0
-0,5
0
1
2
3
4
5
6

FIG. 46D

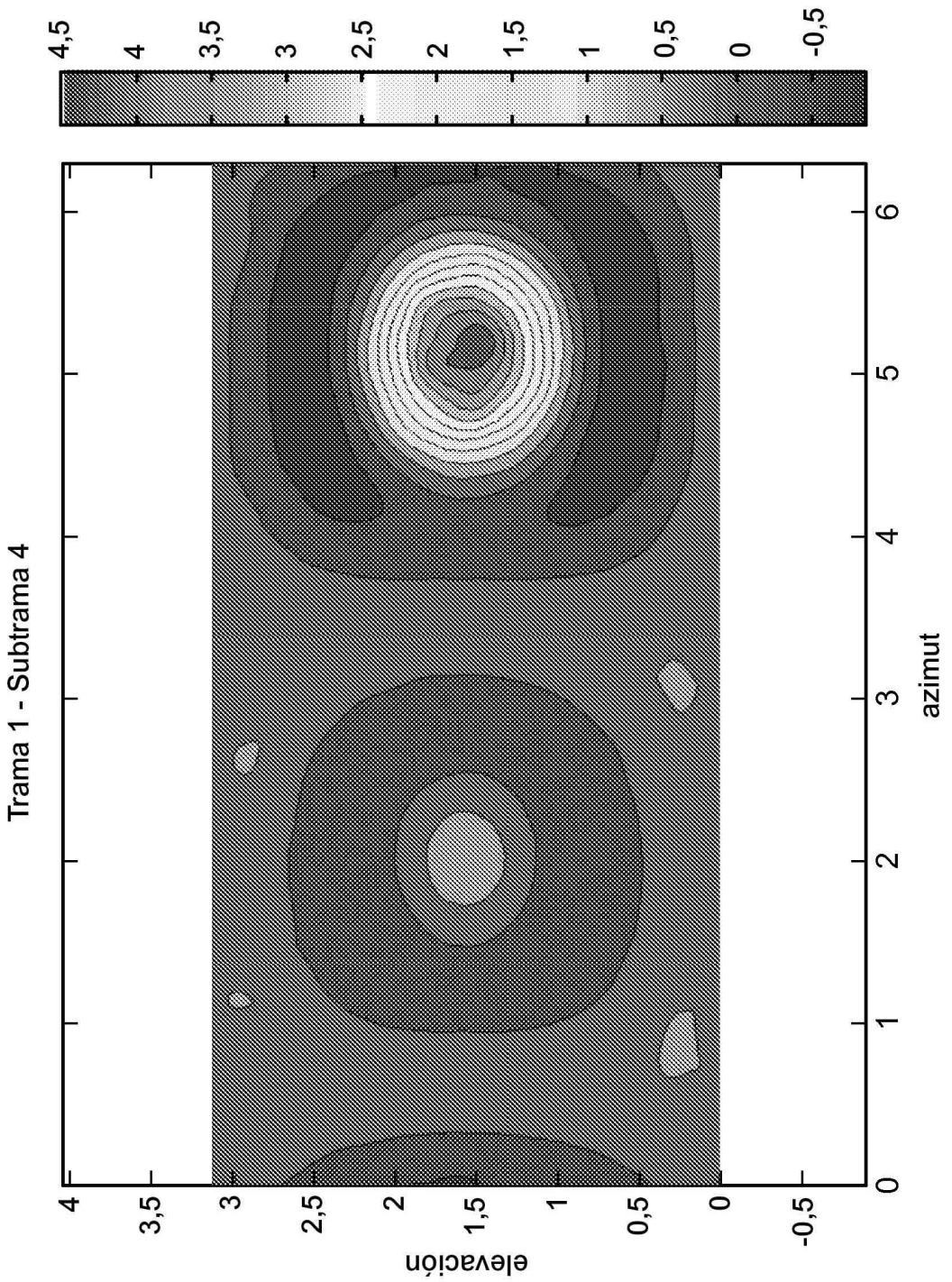

Trama 1 - Subtrama 4
elevación
azimut
4,5
4
3,5
3
2,5
2
1,5
1
0,5
0
-0,5
0
1
2
3
4
5
6

FIG. 46E

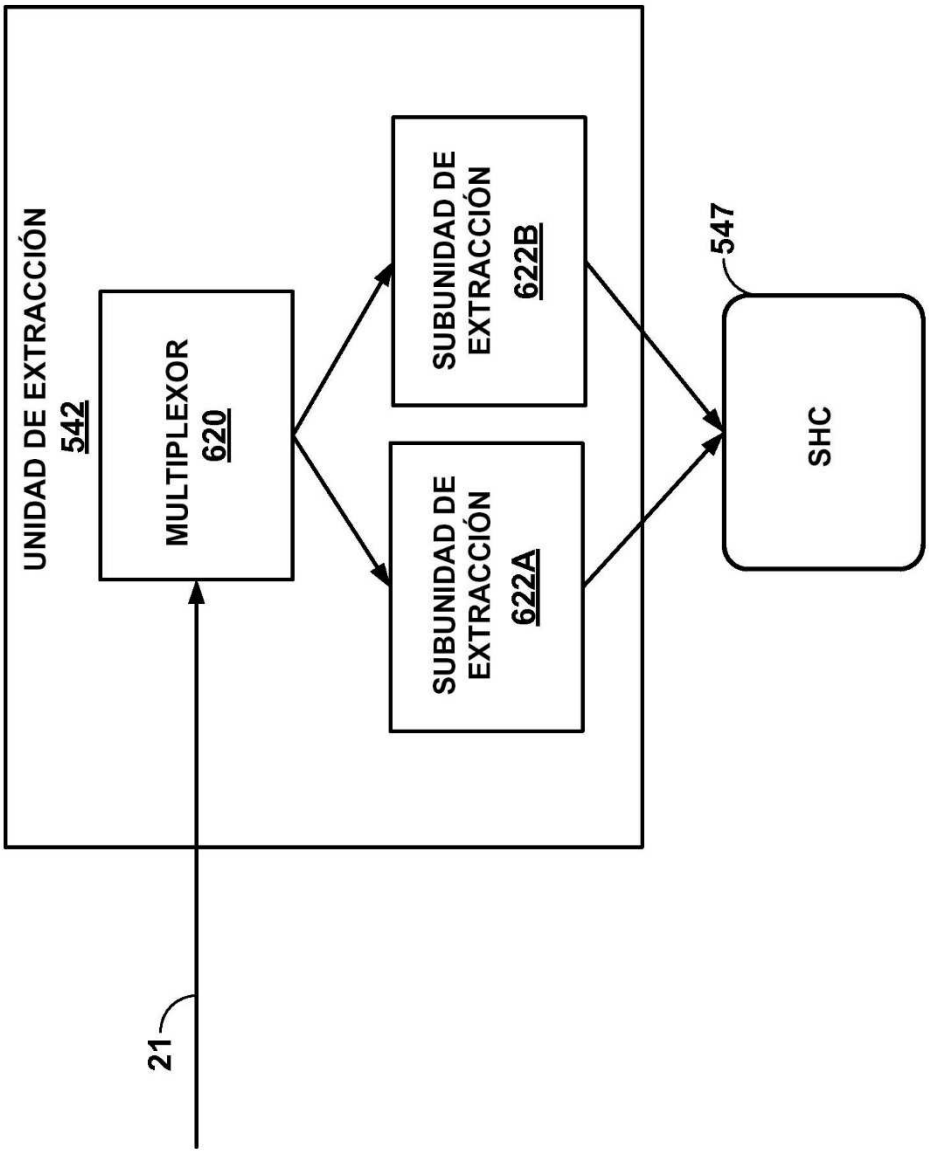

UNIDAD DE EXTRACCIÓN
542
MULTIPLEXOR
620
SUBUNIDAD DE EXTRACCIÓN
622B
SUBUNIDAD DE EXTRACCIÓN
622A
547
SHC
21

FIG. 47

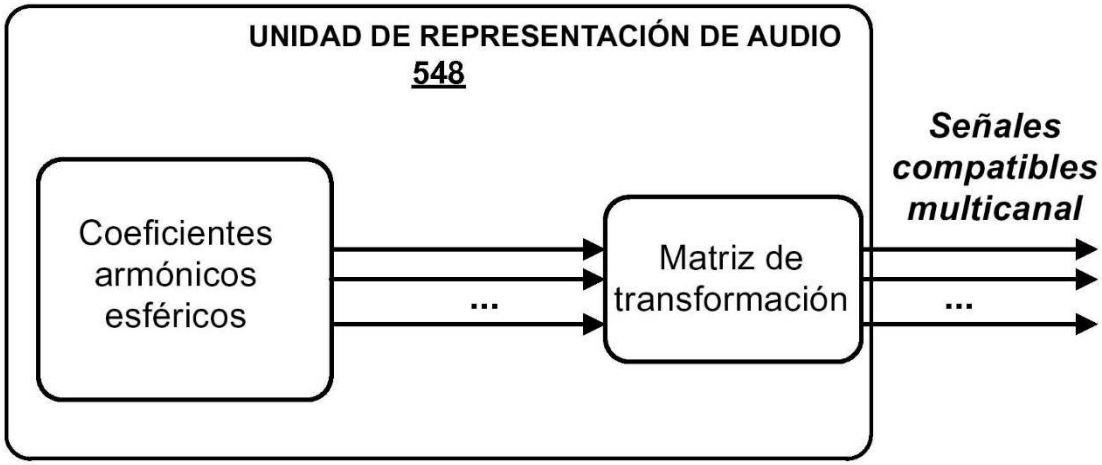

UNIDAD DE REPRESENTACIÓN DE AUDIO
548
Coeficientes armónicos esféricos
...
Matriz de transformación
...
Señales compatibles multicanal

FIG. 48

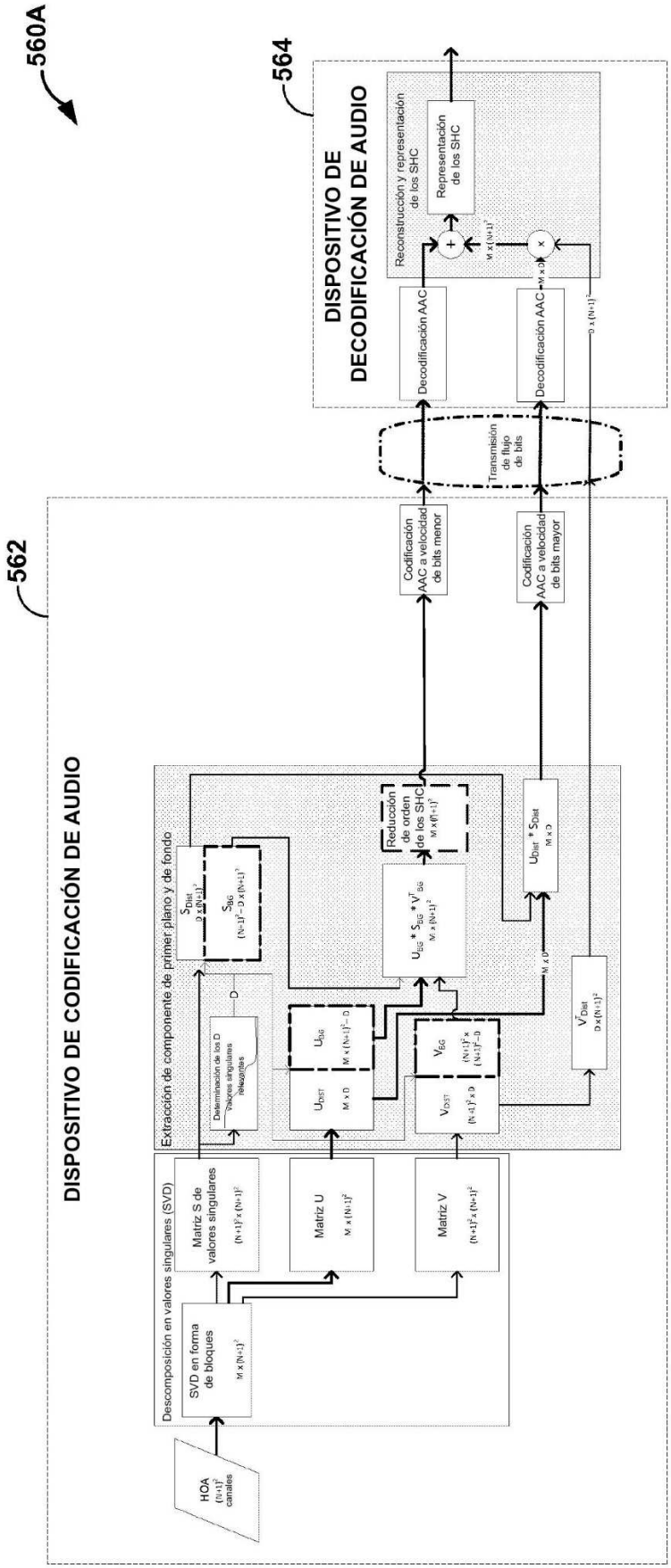

560A
562
564
DISPOSITIVO DE CODIFICACIÓN DE AUDIO
DISPOSITIVO DE DECODIFICACIÓN DE AUDIO
HOA
Descomposición en valores singulares (SVD)
SVD en forma de bloques
Matriz S de valores singulares
Matriz U
Matriz V
Extracción de componente de primer plano y de fondo
Determinación de los D valores singulares relevantes
Reducción de orden de los SHC
Codificación AAC a velocidad de bits menor
Codificación AAC a velocidad de bits mayor
Transmisión de flujo de bits
Decodificación AAC
Decodificación AAC
Reconstrucción y representación de los SHC
Representación de los SHC

FIG. 49A

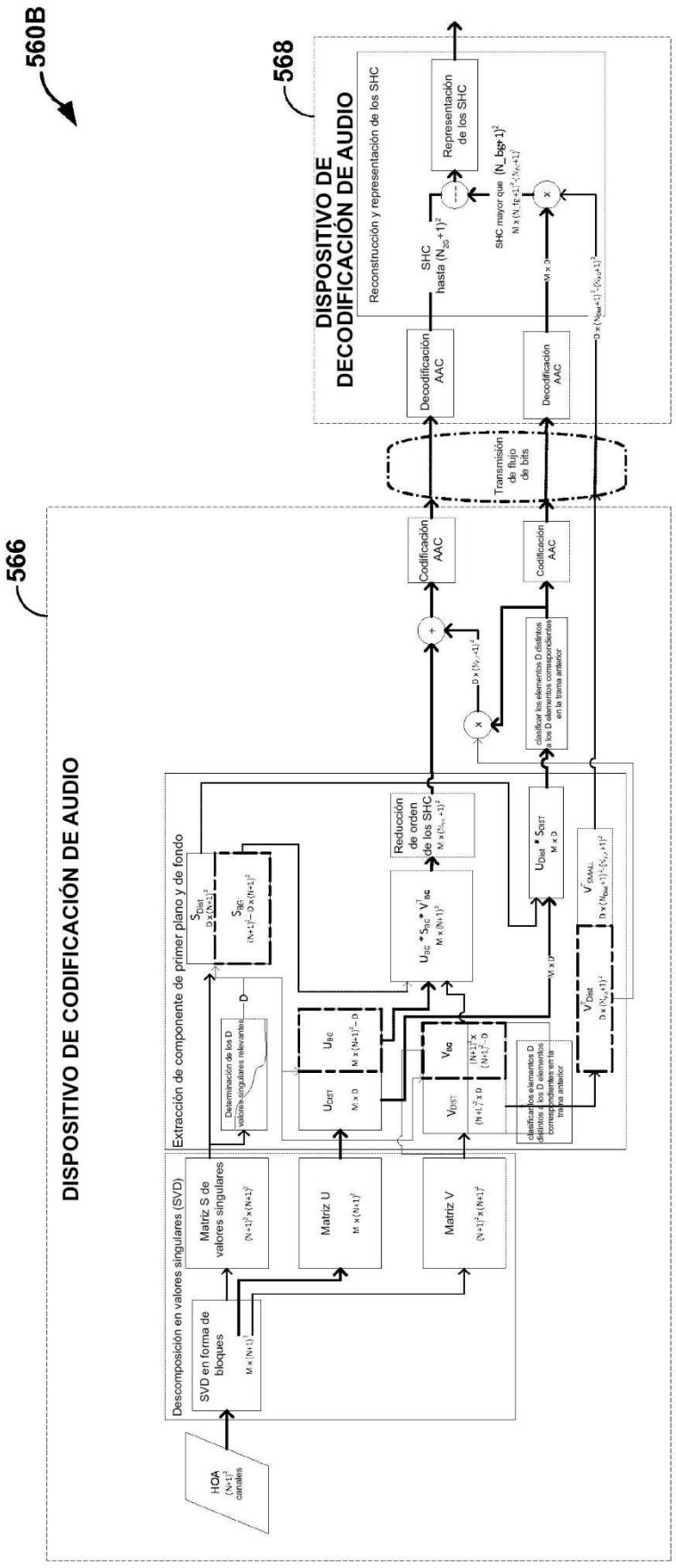

560B
568
DISPOSITIVO DE DECODIFICACIÓN DE AUDIO
Reconstrucción y representación de los SHC
Representación de los SHC
SHC hasta $(N_{BG}+1)^2$
SHC mayor que $(N_bgr-1)^2$
Decodificación AAC
Decodificación AAC
Transmisión de flujo de bits
566
DISPOSITIVO DE CODIFICACIÓN DE AUDIO
Codificación AAC
Codificación AAC
Extracción de componente de primer plano y de fondo
Descomposición en valores singulares (SVD)
HOA $(N+1)^2$ canales
SVD en forma de bloques
Matriz S de valores singulares
Matriz U
Matriz V
Determinación de los D valores singulares relevantes
Reducción de orden de los SHC
clasificar los elementos D distintos de los D elementos correspondientes en la toma anterior
clasificar los elementos D distintos de los D elementos correspondientes en la toma anterior

FIG. 49B

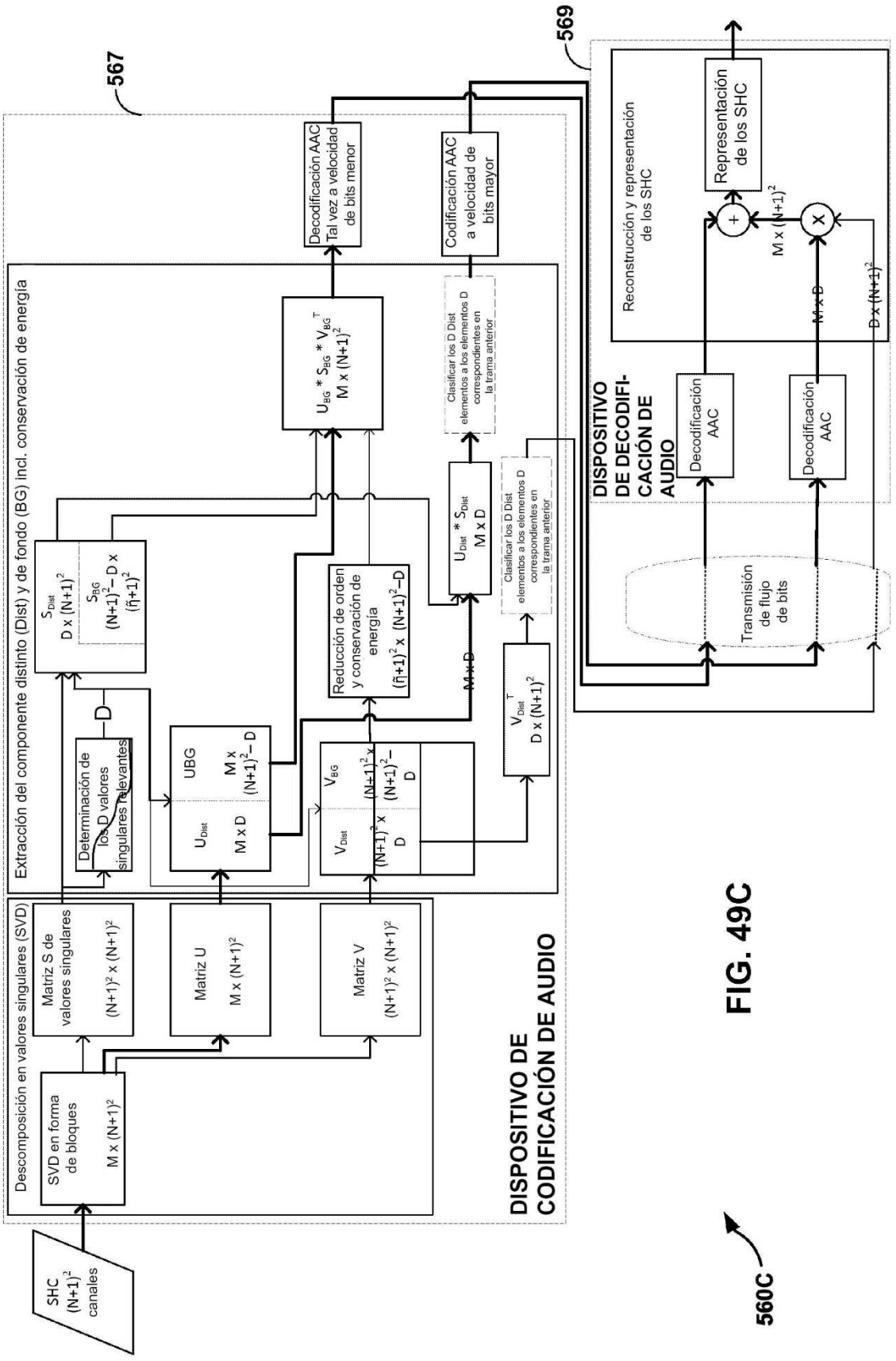
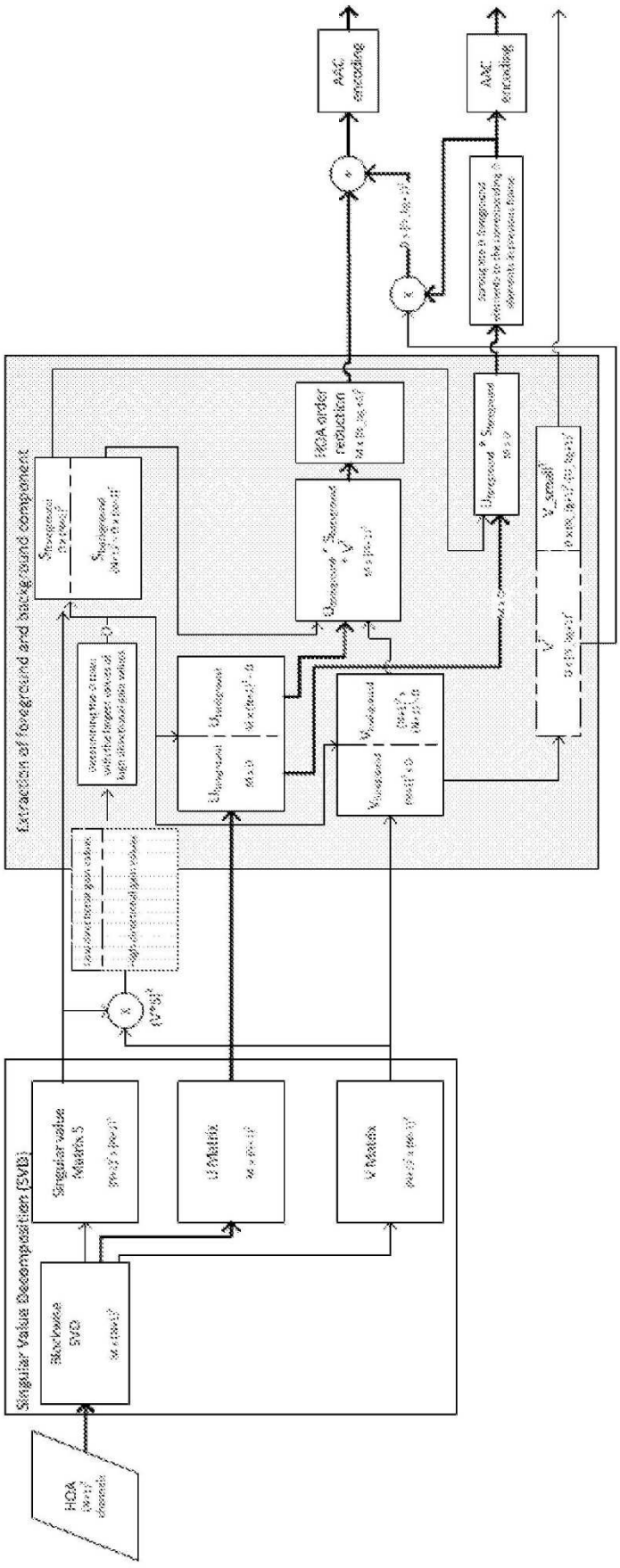

SHC
(N+1)²
canales
560C
DISPOSITIVO DE CODIFICACIÓN DE AUDIO
Descomposición en valores singulares (SVD)
SVD en forma de bloques
M x (N+1)²
Matriz S de valores singulares
(N+1)² x (N+1)²
Matriz U
M x (N+1)²
Matriz V
(N+1)² x (N+1)²
567
Extracción del componente distinto (Dist) y de fondo (BG) incl. conservación de energía
Determinación de los D valores singulares relevantes
D
UBG
U_Dist
M x D
M x (N+1)² – D
V_BG
V_Dist
(N+1)² x D
(N+1)² x (N+1)² – D
S_Dist
D x (N+1)²
S_BG
(N+1)² – D x (ñ+1)²
Reducción de orden y conservación de energía
(ñ+1)² x (N+1)²–D
M x D
U_Dist * S_Dist
M x D
V_Dist^T
D x (N+1)²
Clasificar los D Dist elementos a los elementos D correspondientes en la trama anterior
U_BG * S_BG * V_BG^T
M x (N+1)²
Decodificación AAC Tal vez a velocidad de bits menor
Codificación AAC a velocidad de bits mayor
569
DISPOSITIVO DE DECODIFICACIÓN DE AUDIO
Decodificación AAC
Decodificación AAC
Transmisión de flujo de bits
Reconstrucción y representación de los SHC
Representación de los SHC
+
X
M x (N+1)²
M x D
D x (N+1)²

FIG. 49C



567D

FIG. 49D(i)

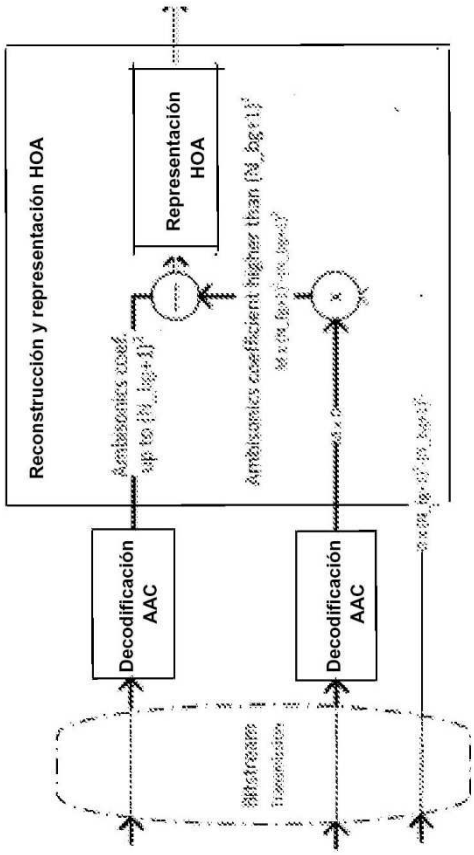

Reconstrucción y representación HOA
Representación HOA
Decodificación AAC
Decodificación AAC
569D

FIG. 49D(ii)

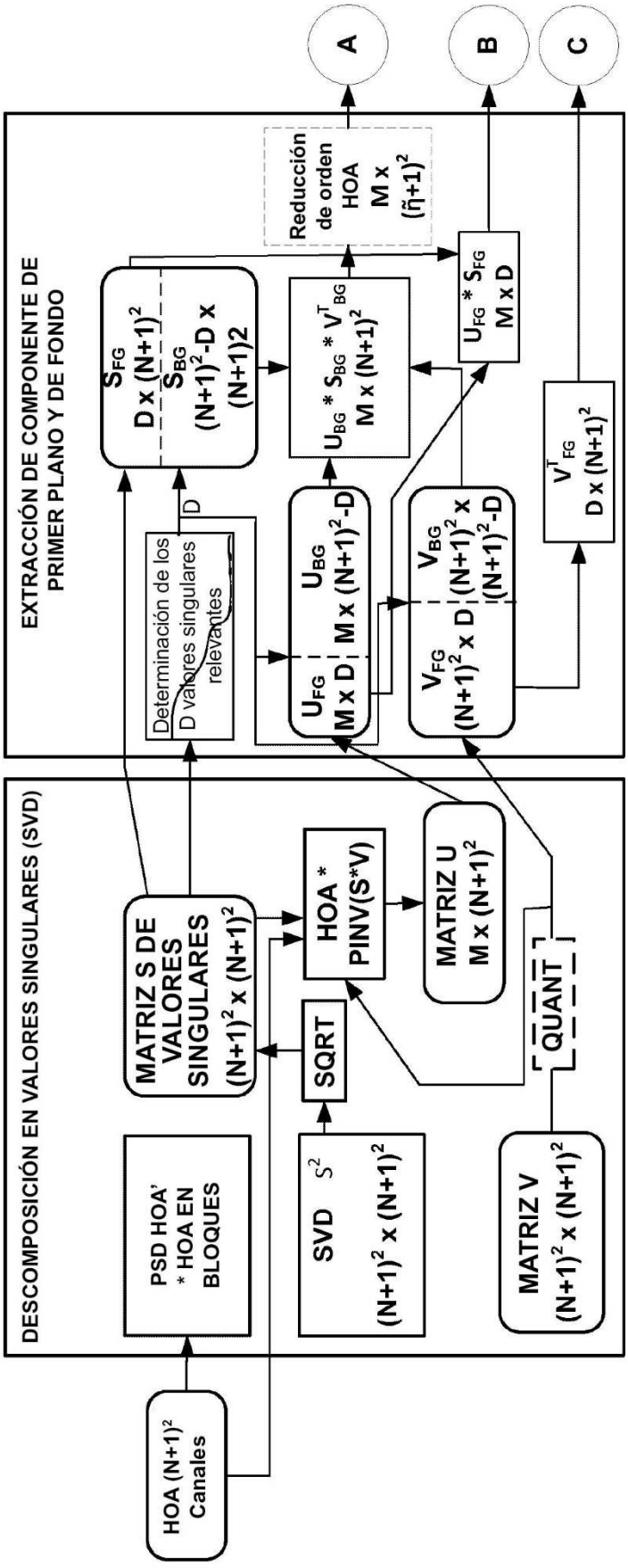

HOA (N+1)² Canales
DESCOMPOSICIÓN EN VALORES SINGULARES (SVD)
PSD HOA' * HOA EN BLOQUES
SVD S² (N+1)² x (N+1)²
SQRT
MATRIZ S DE VALORES SINGULARES (N+1)² x (N+1)²
HOA * PINV(S*V)
MATRIZ U M x (N+1)²
QUANT
MATRIZ V (N+1)² x (N+1)²
EXTRACCIÓN DE COMPONENTE DE PRIMER PLANO Y DE FONDO
Determinación de los D valores singulares relevantes
D
U_FG M x D
U_BG M x (N+1)²-D
V_FG (N+1)² x D
V_BG (N+1)² x (N+1)²-D
S_FG D x (N+1)²
S_BG (N+1)²-D x (N+1)2
U_BG * S_BG * V^T_BG M x (N+1)²
Reducción de orden HOA M x (ñ+1)²
U_FG * S_FG M x D
V^T_FG D x (N+1)²
A
B
C
571E

FIG. 49E(i)

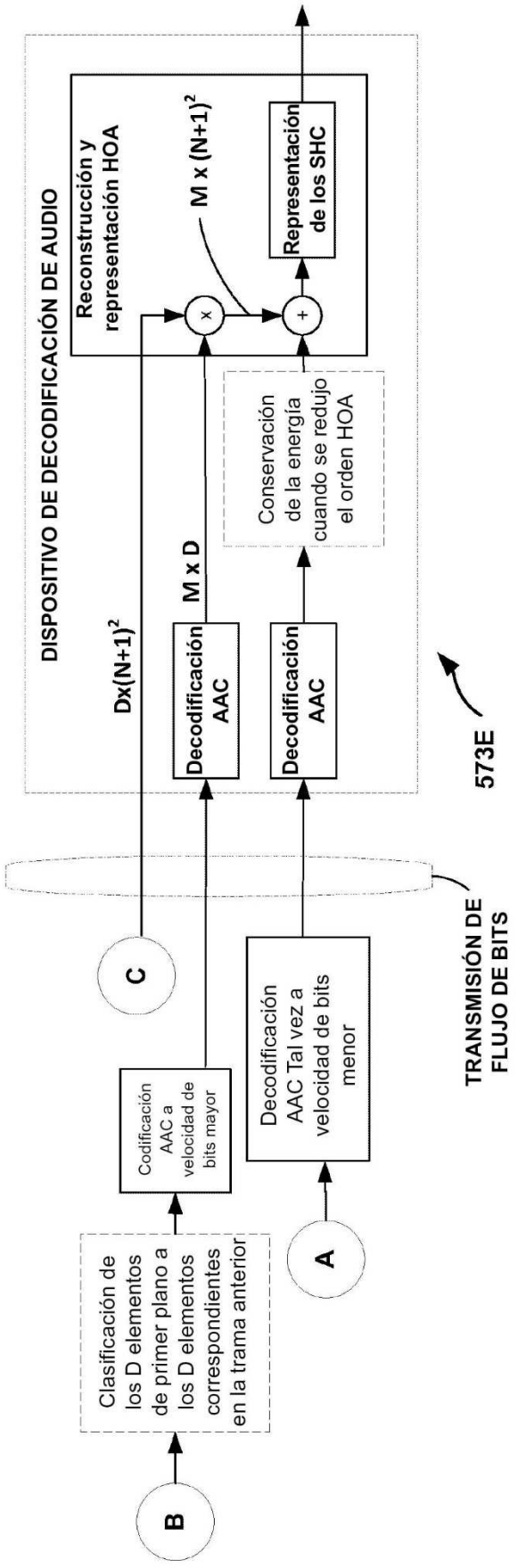

DISPOSITIVO DE DECODIFICACIÓN DE AUDIO
Reconstrucción y representación HOA
M x (N+1)²
Representación de los SHC
Dx(N+1)²
M x D
Decodificación AAC
Decodificación AAC
Conservación de la energía cuando se redujo el orden HOA
573E
TRANSMISIÓN DE FLUJO DE BITS
C
Codificación AAC a velocidad de bits mayor
Decodificación AAC Tal vez a velocidad de bits menor
A
B
Clasificación de los D elementos de primer plano a los D elementos correspondientes en la trama anterior

FIG. 49E(ii)

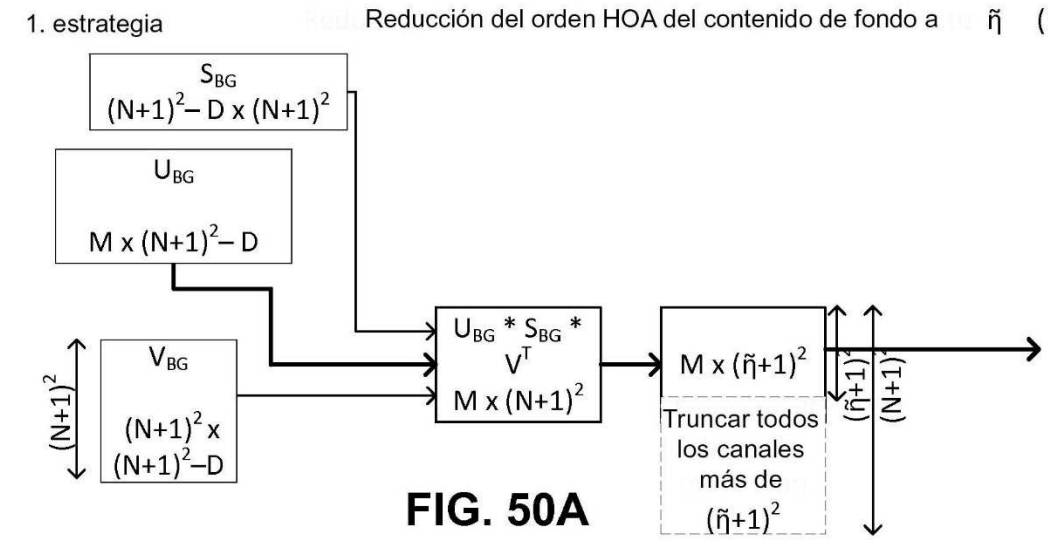


FIG. 50A

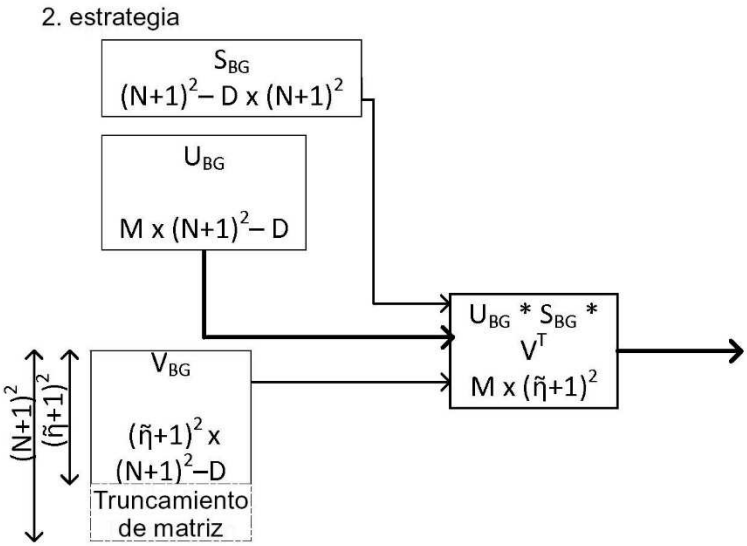


FIG. 50B

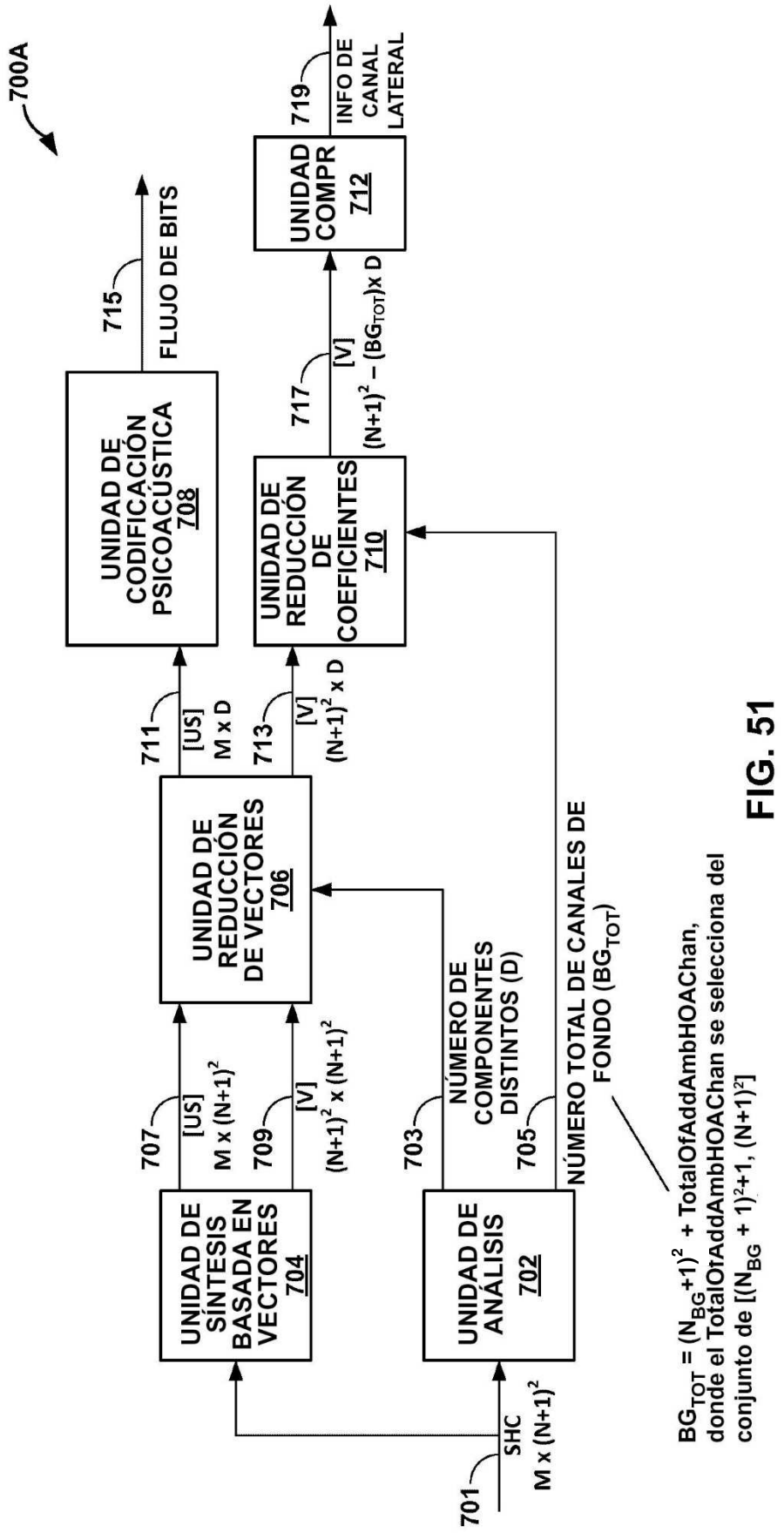

700A
UNIDAD DE CODIFICACIÓN PSICOACÚSTICA 708
715
FLUJO DE BITS
UNIDAD COMPR 712
719
INFO DE CANAL LATERAL
UNIDAD DE REDUCCIÓN DE COEFICIENTES 710
717
[V] (N+1)² – (BG_TOT)x D
UNIDAD DE REDUCCIÓN DE VECTORES 706
711
[US] M x D
713
[V] (N+1)² x D
UNIDAD DE SÍNTESIS BASADA EN VECTORES 704
707
[US] M x (N+1)²
709
[V] (N+1)² x (N+1)²
UNIDAD DE ANÁLISIS 702
703
NÚMERO DE COMPONENTES DISTINTOS (D)
705
NÚMERO TOTAL DE CANALES DE FONDO (BG_TOT)
701
SHC M x (N+1)²
BG_TOT = (N_BG+1)² + TotalOfAddAmbHOAChan, donde el TotalOfAddAmbHOAChan se selecciona del conjunto de [(N_BG + 1)²+1, (N+1)²]

FIG. 51

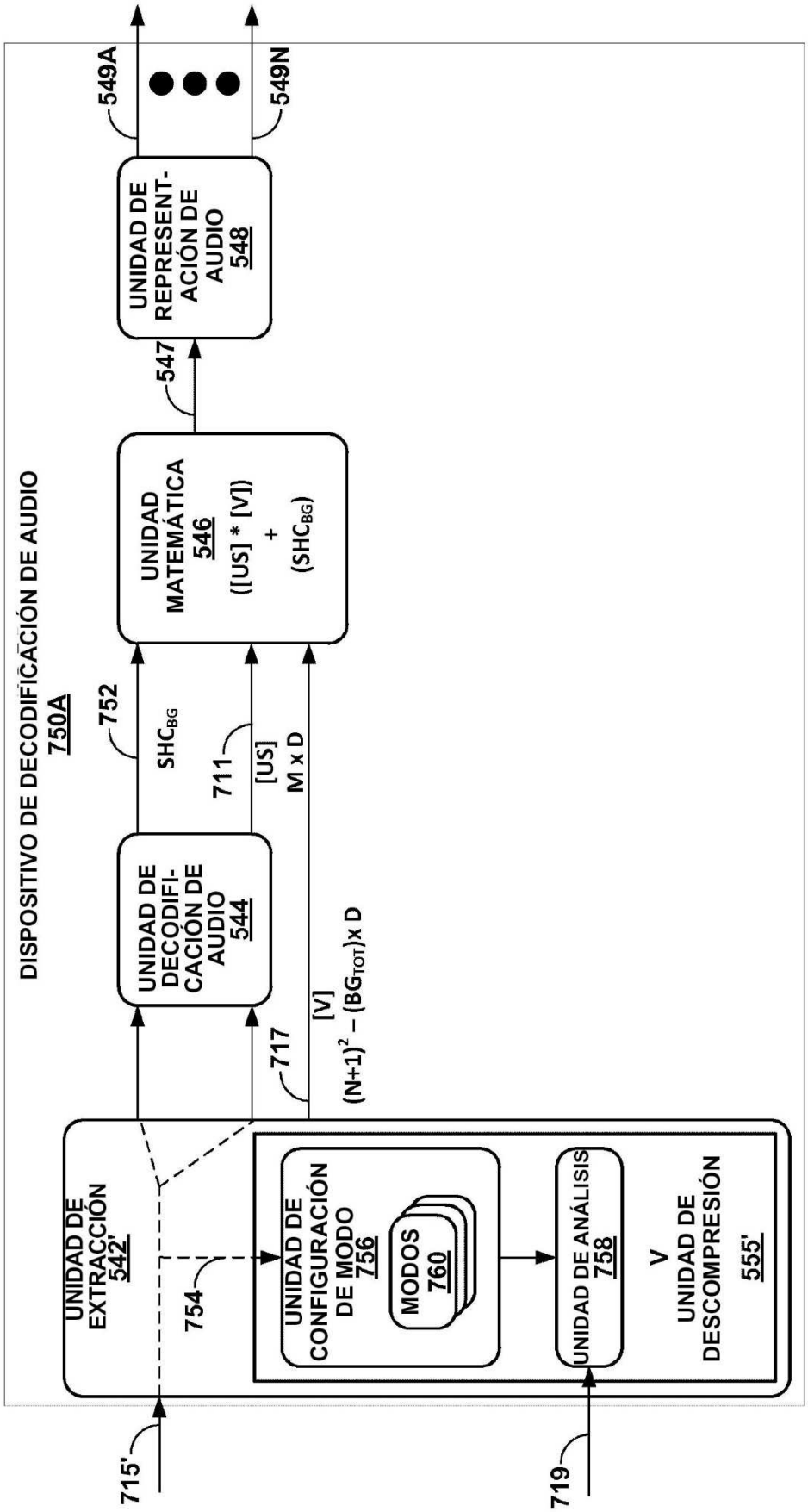

DISPOSITIVO DE DECODIFICACIÓN DE AUDIO
750A
UNIDAD DE EXTRACCIÓN 542'
UNIDAD DE DECODIFI-CACIÓN DE AUDIO 544
UNIDAD MATEMÁTICA 546
([US] * [V]) + (SHC_{BG})
UNIDAD DE REPRESENT-ACIÓN DE AUDIO 548
UNIDAD DE CONFIGURACIÓN DE MODO 756
MODOS 760
UNIDAD DE ANÁLISIS 758
V
UNIDAD DE DESCOMPRESIÓN 555'
715'
719
754
752
SHC_{BG}
711
[US]
M x D
717
[V]
$(N+1)^2 - (BG_{TOT})$ x D
547
549A
549N

FIG. 52

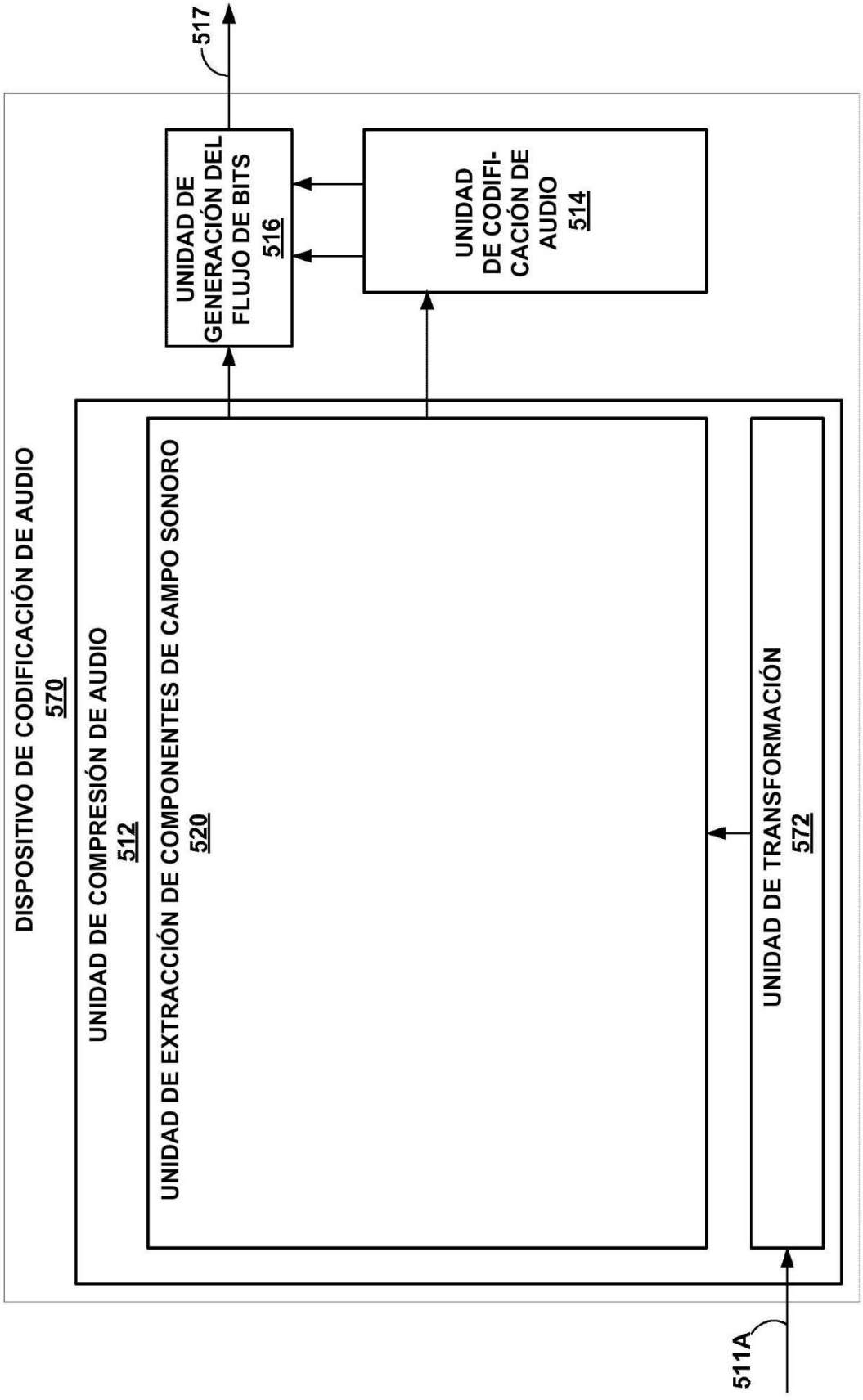

517
UNIDAD DE GENERACIÓN DEL FLUJO DE BITS 516
UNIDAD DE CODIFI-CACIÓN DE AUDIO 514
DISPOSITIVO DE CODIFICACIÓN DE AUDIO 570
UNIDAD DE COMPRESIÓN DE AUDIO 512
UNIDAD DE EXTRACCIÓN DE COMPONENTES DE CAMPO SONORO 520
UNIDAD DE TRANSFORMACIÓN 572
511A

FIG. 53

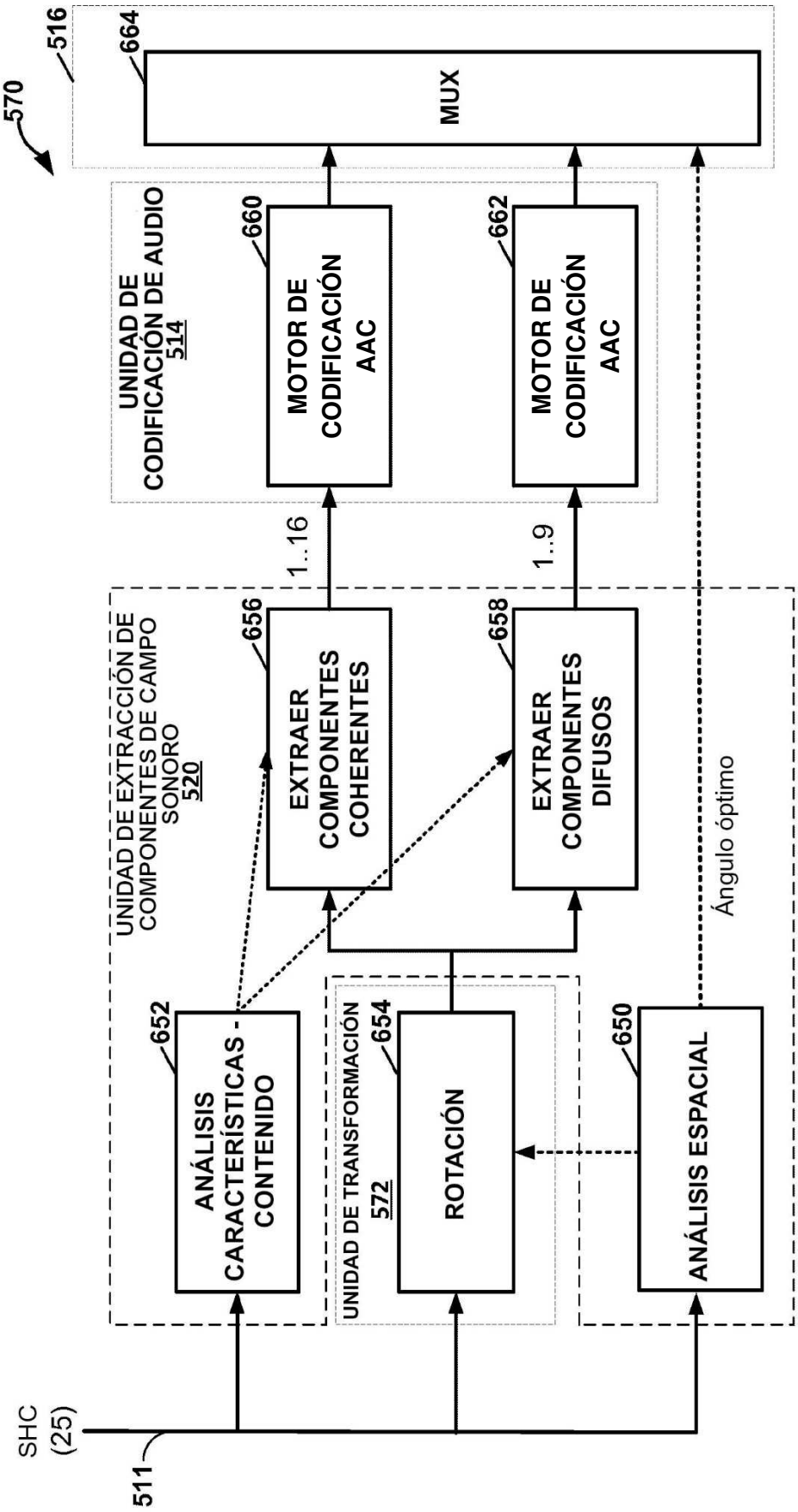

570
516
664
MUX
UNIDAD DE CODIFICACIÓN DE AUDIO 514
660
MOTOR DE CODIFICACIÓN AAC
662
MOTOR DE CODIFICACIÓN AAC
1..16
1..9
UNIDAD DE EXTRACCIÓN DE COMPONENTES DE CAMPO SONORO 520
656
EXTRAER COMPONENTES COHERENTES
658
EXTRAER COMPONENTES DIFUSOS
Ángulo óptimo
652
ANÁLISIS CARACTERÍSTICAS - CONTENIDO
UNIDAD DE TRANSFORMACIÓN 572
654
ROTACIÓN
650
ANÁLISIS ESPACIAL
SHC (25)
511

FIG. 54

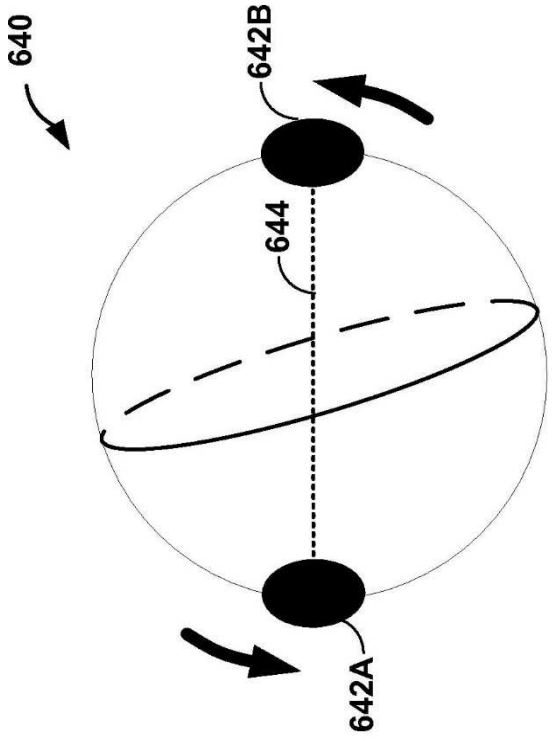


FIG. 55B

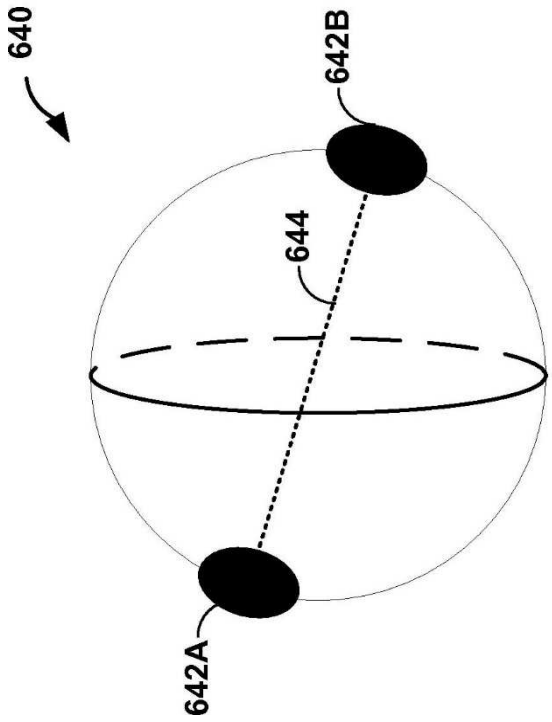


FIG. 55A

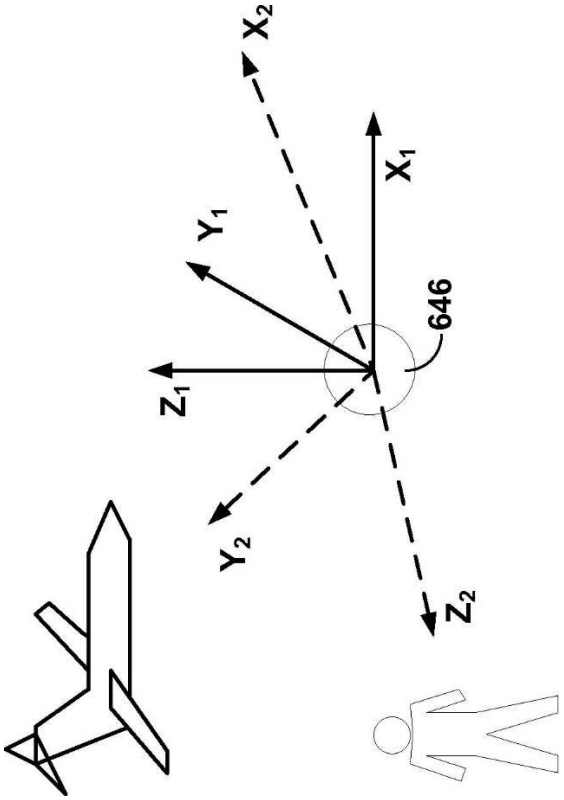


FIG. 56

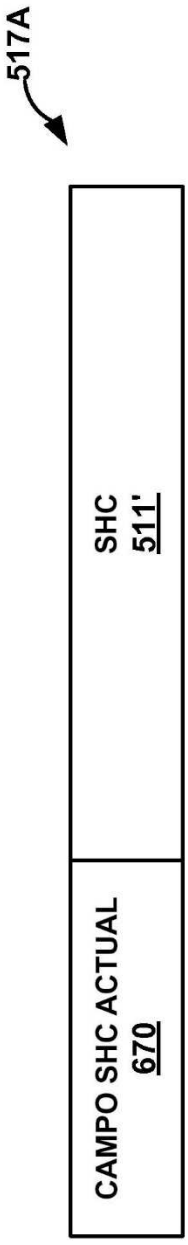


FIG. 57A

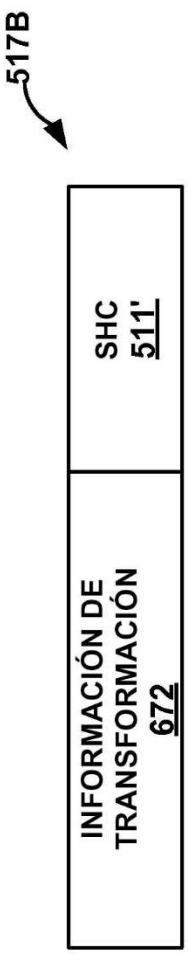


FIG. 57B

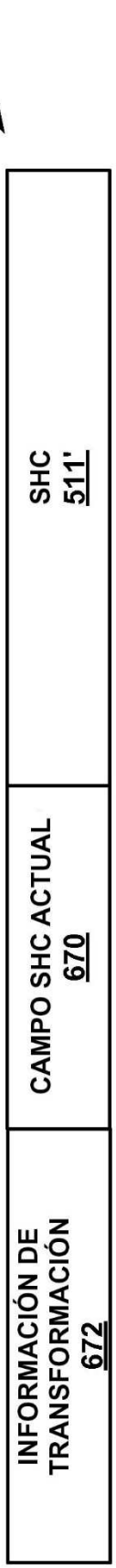


FIG. 57C

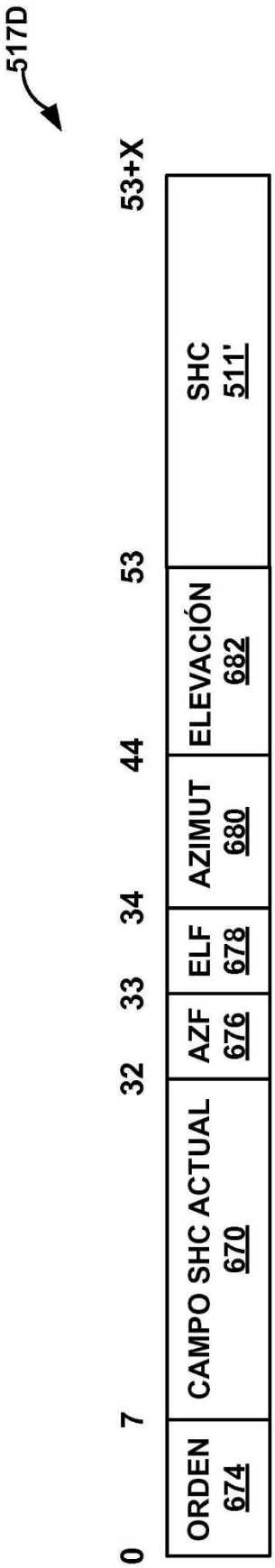


FIG. 57D

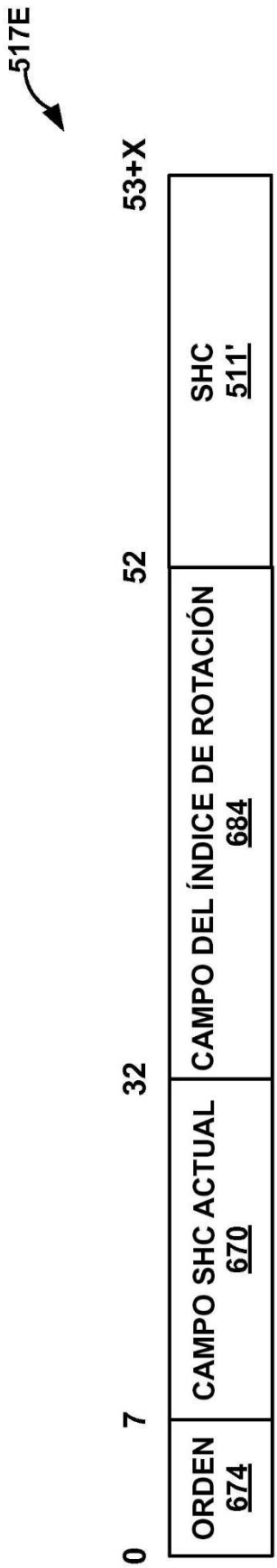


FIG. 57E

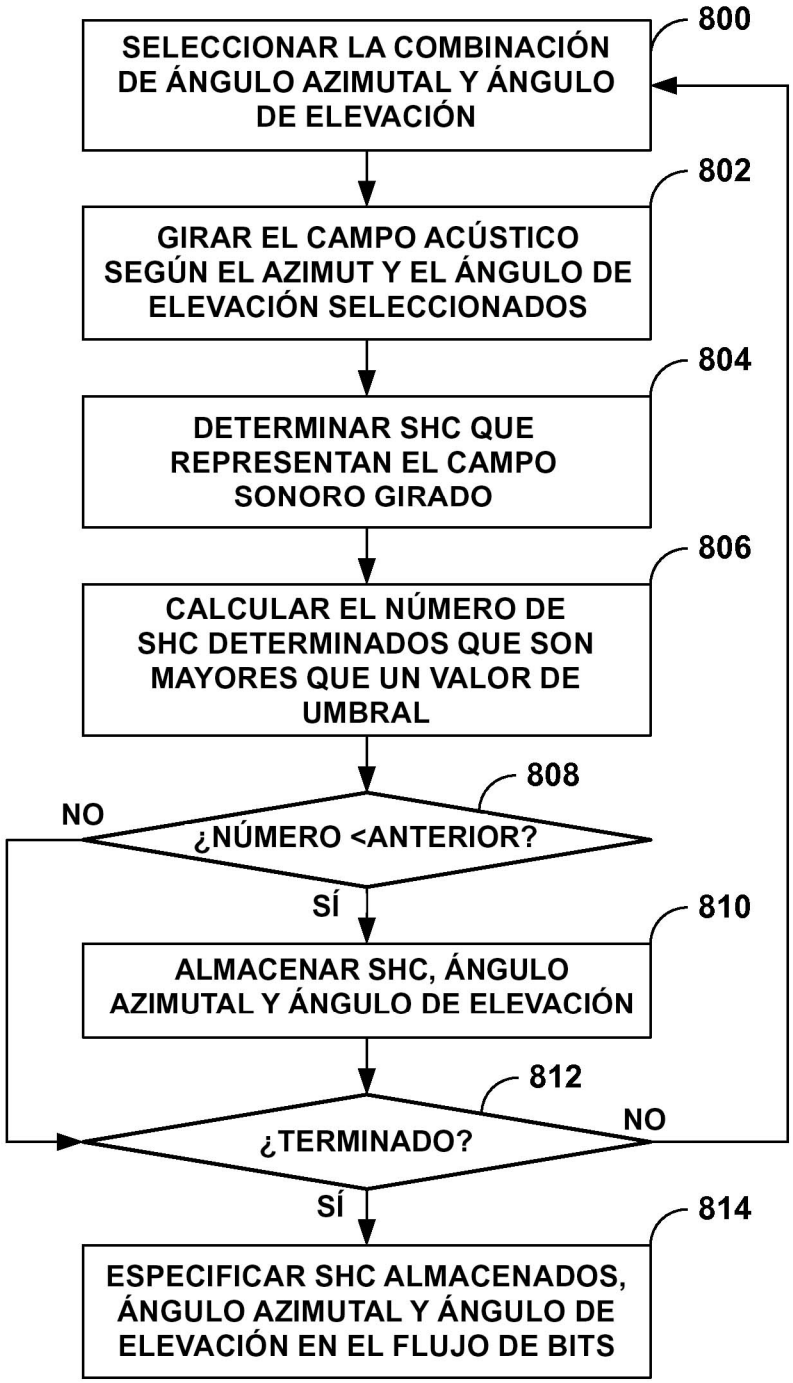

800
SELECCIONAR LA COMBINACIÓN DE ÁNGULO AZIMUTAL Y ÁNGULO DE ELEVACIÓN
802
GIRAR EL CAMPO ACÚSTICO SEGÚN EL AZIMUT Y EL ÁNGULO DE ELEVACIÓN SELECCIONADOS
804
DETERMINAR SHC QUE REPRESENTAN EL CAMPO SONORO GIRADO
806
CALCULAR EL NÚMERO DE SHC DETERMINADOS QUE SON MAYORES QUE UN VALOR DE UMBRAL
808
¿NÚMERO <ANTERIOR?
NO
SÍ
810
ALMACENAR SHC, ÁNGULO AZIMUTAL Y ÁNGULO DE ELEVACIÓN
812
¿TERMINADO?
NO
SÍ
814
ESPECIFICAR SHC ALMACENADOS, ÁNGULO AZIMUTAL Y ÁNGULO DE ELEVACIÓN EN EL FLUJO DE BITS

FIG. 58

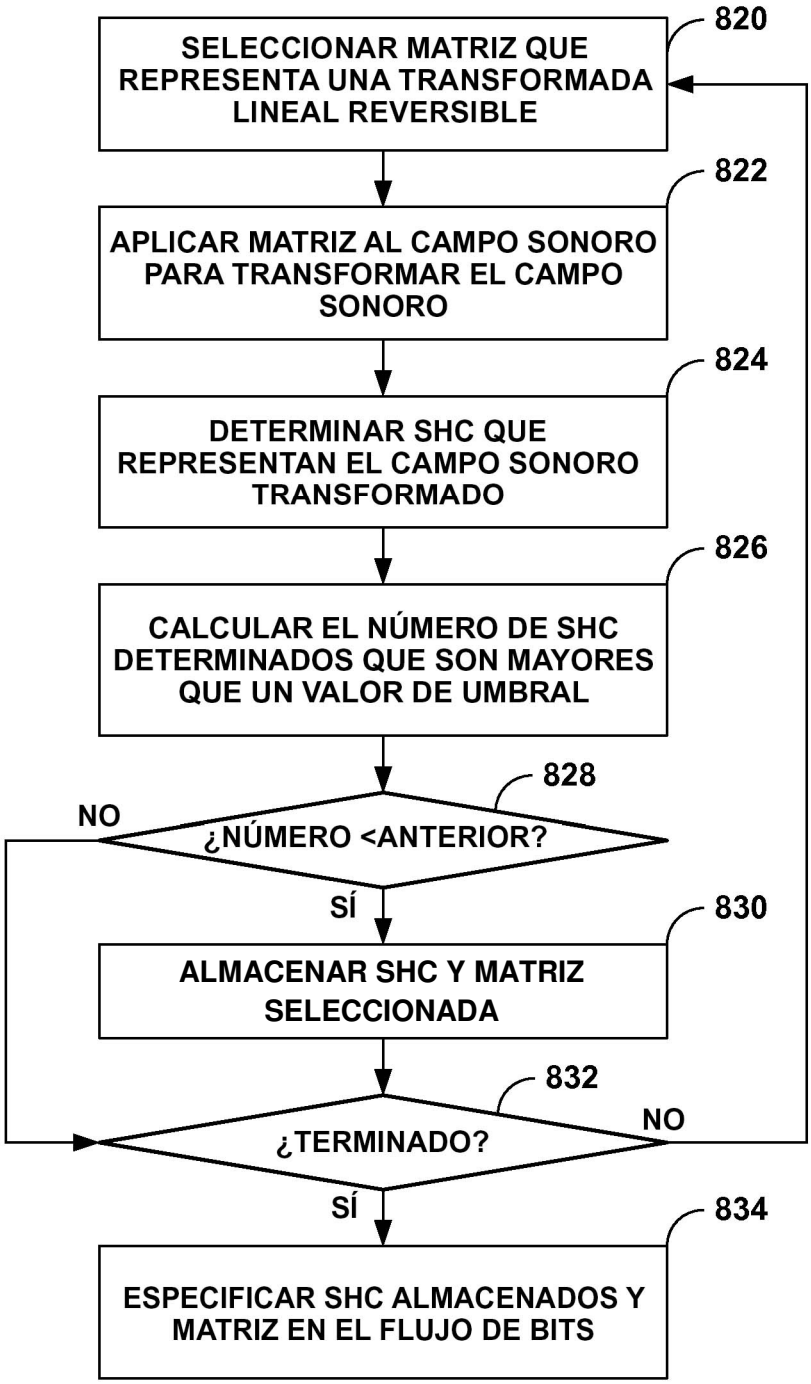

820
SELECCIONAR MATRIZ QUE REPRESENTA UNA TRANSFORMADA LINEAL REVERSIBLE
822
APLICAR MATRIZ AL CAMPO SONORO PARA TRANSFORMAR EL CAMPO SONORO
824
DETERMINAR SHC QUE REPRESENTAN EL CAMPO SONORO TRANSFORMADO
826
CALCULAR EL NÚMERO DE SHC DETERMINADOS QUE SON MAYORES QUE UN VALOR DE UMBRAL
828
NO
¿NÚMERO <ANTERIOR?
SÍ
830
ALMACENAR SHC Y MATRIZ SELECCIONADA
832
NO
¿TERMINADO?
SÍ
834
ESPECIFICAR SHC ALMACENADOS Y MATRIZ EN EL FLUJO DE BITS

FIG. 59