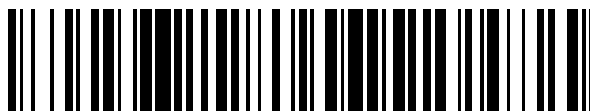


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 738 490**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/16 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.02.2014** **PCT/US2014/019446**

87 Fecha y número de publicación internacional: **04.09.2014** **WO14134462**

96 Fecha de presentación y número de la solicitud europea: **28.02.2014** **E 14713289 (8)**

97 Fecha y número de publicación de la concesión europea: **24.04.2019** **EP 2962298**

54 Título: **Especificación de coeficientes ambisónicos de orden superior y/o armónicos esféricos en flujos de bits**

30 Prioridad:

01.03.2013 US 201361771677 P

30.07.2013 US 201361860201 P

27.02.2014 US 201414192819

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

23.01.2020

73 Titular/es:

QUALCOMM INCORPORATED (100.0%)

5775 Morehouse Drive

San Diego, CA 92121-1714, US

72 Inventor/es:

SEN, DIPANJAN y

MORRELL, MARTIN JAMES

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 738 490 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Especificación de coeficientes ambisónicos de orden superior y/o armónicos esféricos en flujos de bits

5 CAMPO TÉCNICO

[0001] La presente divulgación se refiere a la codificación de audio y, más específicamente, a los flujos de bits que especifican datos de audio codificados.

10 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 usado para reproducir una señal de audio multicanal reproducida a partir de esta señal SHC. Esta señal SHC también puede facilitar la retrocompatibilidad, ya que esta señal SHC se puede reproducir 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 de SHC puede, por lo tanto, hacer posible una mejor representación de un campo sonoro que también asimila la retrocompatibilidad.

[0003] En el documento "Codificación de audio multicanal basada en mínimos ángulos audibles" de Daniel *et al*, 40.^a Conferencia Internacional: Audio espacial: percibir el sonido del espacio (octubre de 2010), se describe un procedimiento que proporciona un esquema para codificar señales de audio multicanal que representan una escena auditiva espacial basada en la percepción del sonido humana en el espacio. Se basa en un modelo psicoacústico basado en medidas de mínimos ángulos audibles (MAA) en presencia de fuentes de sonido que distraen. Una ganancia de compresión se obtiene al truncar el orden de representación de la escena auditiva en el dominio ambisónico de orden superior (HOA) de acuerdo con el modelo psicoacústico. Se llevaron a cabo simulaciones numéricas para vincular el error de representación del campo con una distorsión angular de la dirección aparente de las fuentes de sonido.

30 SUMARIO

[0004] En general, se describen diversas técnicas para señalar la información de audio en un flujo de bits representativo de datos de audio y para realizar una transformación con respecto a los datos de audio. En algunos aspectos, se describen técnicas para señalar cuáles de una pluralidad de elementos jerárquicos, tales como los coeficientes ambisónicos de orden superior (HOA) (que también se pueden denominar coeficientes armónicos esféricos), se incluyen en el flujo de bits. Dado que algunos de los coeficientes HOA pueden no proporcionar información pertinente al describir un campo sonoro, el codificador de audio puede reducir la pluralidad de los coeficientes HOA a un subconjunto distinto de cero de los coeficientes HOA que proporcionan información pertinente al describir el campo sonoro, aumentando de este modo la eficacia de la codificación. Como resultado, diversos aspectos de las técnicas pueden hacer posible especificar en el flujo de bits que incluye los coeficientes HOA y/o las versiones codificadas de los mismos, los de los coeficientes HOA que se incluyen realmente en el flujo de bits (por ejemplo, el subconjunto distinto de cero de los coeficientes HOA que incluye al menos uno de los coeficientes HOA pero no la totalidad de los coeficientes). La información que identifica el subconjunto de los coeficientes HOA se puede especificar en el flujo de bits como se indica anteriormente, o en algunos casos, en la información de canal lateral.

[0005] En otros aspectos, se describen las técnicas para transformar SHC a fin de reducir un número de SHC que se van a especificar en el flujo de bits y de este modo aumentar la eficacia de la codificación. Es decir, las técnicas pueden realizar alguna forma de una transformada invertible lineal con respecto a los SHC con el resultado de reducir el número de SHC que se van a especificar en el flujo de bits. Los ejemplos de una transformada invertible lineal incluyen rotación, traslación, una transformada de coseno discreta (DCT), una transformada de Fourier discreta (DFT), descomposición de valores singulares y análisis de componentes principales. A continuación, las técnicas pueden especificar "información de transformación" que identifique la transformación realizada con respecto a los SHC. Por ejemplo, cuando se realiza una rotación con respecto a los SHC, las técnicas pueden proporcionar la especificación de la información de rotación que identifica la rotación (a menudo en términos de diversos ángulos de rotación). Cuando se realiza SVD como otro ejemplo, las técnicas pueden proporcionar un indicador que indica que se realizó SVD.

[0006] En un ejemplo, se proporciona un procedimiento de generación de un flujo de bits representativo de contenido de audio, comprendiendo el procedimiento: determinar que uno o más de una pluralidad de elementos jerárquicos tienen información pertinente al describir un campo sonoro en tres dimensiones, en el que la pluralidad de elementos jerárquicos comprende coeficientes armónicos esféricos representativos del campo sonoro, comprendiendo la determinación determinar si un coeficiente armónico esférico está por encima de un valor umbral, siendo el valor umbral mayor que cero; especificar, en el flujo de bits y basado en la determinación, cuáles de los coeficientes armónicos esféricos se incluyen en el flujo de bits; y especificar, en el flujo de bits, el uno o más de la pluralidad de coeficientes armónicos esféricos determinados.

[0007] En otro ejemplo, se proporciona un dispositivo configurado para generar un flujo de bits representativo de contenido de audio, comprendiendo el dispositivo: medios para determinar que uno o más de una pluralidad de elementos jerárquicos tengan información pertinente al describir un campo sonoro en tres dimensiones, en el que la pluralidad de elementos jerárquicos son coeficientes armónicos esféricos representativos del campo sonoro, comprendiendo la determinación determinar si un coeficiente armónico esférico está por encima de un valor umbral, siendo el valor umbral mayor que cero; medios para especificar, en el flujo de bits y basado en la determinación, cuáles de los coeficientes armónicos esféricos se incluyen en el flujo de bits; y medios para especificar, en el flujo de bits, el uno o más de la pluralidad de coeficientes armónicos esféricos determinados.

[0008] En otro ejemplo, se proporciona un procedimiento de procesamiento de un flujo de bits representativo de contenido de audio, comprendiendo el procedimiento: identificar, a partir del flujo de bits, que uno o más de una pluralidad de coeficientes armónicos esféricos se incluyen en el flujo de bits, en el que se ha determinado que el uno o más de la pluralidad de coeficientes armónicos esféricos proporcionan información pertinente al describir un campo sonoro en tres dimensiones al estar por encima de un valor umbral, siendo el valor umbral mayor que cero, y en el que la pluralidad de coeficientes armónicos esféricos son representativos del campo sonoro; y analizar sintácticamente el flujo de bits para determinar el uno o más de la pluralidad de coeficientes armónicos esféricos identificados.

[0009] En otro ejemplo, se proporciona un dispositivo configurado para procesar un flujo de bits representativo de contenido de audio, comprendiendo el dispositivo: medios para identificar, a partir del flujo de bits, que uno o más de una pluralidad de coeficientes armónicos esféricos se incluyen en el flujo de bits, en el que se ha determinado que el uno o más de la pluralidad de coeficientes armónicos esféricos proporcionan información pertinente al describir un campo sonoro en tres dimensiones al estar por encima de un valor umbral, siendo el valor umbral mayor que cero, y en el que la pluralidad de coeficientes armónicos esféricos son representativos del campo sonoro; y medios para analizar sintácticamente el flujo de bits para determinar la pluralidad de coeficientes armónicos esféricos identificados.

[0010] En otro ejemplo, un medio de almacenamiento legible por ordenador no transitorio ha almacenado en el mismo instrucciones que, cuando se ejecutan, hacen que uno o más procesadores lleven a cabo uno de los procedimientos descritos anteriormente.

[0011] Los detalles de uno o más aspectos de las técnicas se exponen en los dibujos adjuntos y en la descripción a continuación. Otras características, objetivos y ventajas de estas técnicas serán evidentes a partir de la descripción y dibujos, y a partir de las reivindicaciones.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0012]

Las FIGS. 1 y 2 son diagramas que ilustran funciones de base armónica esférica de diversos órdenes y subórdenes.

La FIG. 3 es un diagrama que ilustra un sistema que puede implementar diversos aspectos de las técnicas descritas en la presente divulgación.

La FIG. 4A y 4B son diagramas de bloques que ilustran implementaciones de ejemplo del dispositivo de generación de flujo de bits mostrado en el ejemplo de la FIG. 3.

Las FIGS. 5A y 5B son diagramas que ilustran un ejemplo de realización de diversos aspectos de las técnicas descritas en la presente divulgación para girar un campo sonoro.

La FIG. 6 es un diagrama que ilustra un campo sonoro de ejemplo capturado de acuerdo con una primera trama de referencia que a continuación se gira de acuerdo con las técnicas descritas en la presente divulgación para expresar el campo sonoro en términos de una segunda trama de referencia.

Las FIGS. 7A-7E ilustran ejemplos de un flujo de bits formado de acuerdo con las técnicas descritas en la presente divulgación.

La FIG. 8 es un diagrama de flujo que ilustra el funcionamiento de ejemplo del dispositivo de generación de flujo de bits de la FIG. 3 al realizar los aspectos de rotación de las técnicas descritas en la presente divulgación.

La FIG. 8 es un diagrama de flujo que ilustra el funcionamiento de ejemplo del dispositivo de generación de flujo de bits de la FIG. 3 al realizar los aspectos de rotación de las técnicas descritas en la presente divulgación.

La FIG. 9 es un diagrama de flujo que ilustra el funcionamiento de ejemplo del dispositivo de generación de flujo de bits mostrado en el ejemplo de la FIG. 3 al realizar los aspectos de transformación de las técnicas descritas en la presente divulgación.

La FIG. 10 es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de extracción al realizar diversos aspectos de las técnicas descritas en la presente divulgación.

La FIG. 11 es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de generación de flujo de bits y un dispositivo de extracción al realizar diversos aspectos de las técnicas descritas en la presente divulgación.

DESCRIPCIÓN DETALLADA

[0013] La evolución del sonido envolvente ha puesto a la disposición muchos formatos de salida para el entretenimiento hoy en día. Los ejemplos de dichos formatos de sonido envolvente incluyen el popular formato 5.1 (que incluye los siguientes seis canales: frontal izquierdo (FL), frontal derecho (FR), central o frontal central, posterior izquierdo o envolvente izquierdo, posterior derecho o envolvente derecho y los efectos de baja frecuencia (LFE)), el formato 7.1 creciente y el próximo formato 22.2 (por ejemplo, para su uso con la norma de televisión de ultra alta definición). Otros ejemplos incluyen formatos para una matriz armónica esférica.

[0014] La entrada a un futuro codificador de MPEG es opcionalmente uno de tres formatos posibles: (i) audio tradicional basado en canales, que está destinado a reproducirse a través de altavoces en posiciones preespecificadas; (ii) audio basado en objetos, que implica datos discretos de modulación de código de pulsos (PCM) para objetos de audio individuales con metadatos asociados que contienen sus coordenadas de ubicación (entre otra información); y (iii) audio basado en la escena, que implica representar el campo sonoro usando coeficientes de funciones de base armónica esférica (también llamados "coeficientes armónicos esféricos" o SHC).

[0015] Existen diversos formatos de 'sonido envolvente' en el mercado. Varían, por ejemplo, desde el sistema de cine en casa 5.1 (que es el que ha tenido el mayor éxito en cuanto a irrumpir en las salas de estar, más allá del estéreo) hasta el sistema 22.2 desarrollado por NHK (Nippon Hoso Kyokai o Corporación Difusora de Japón). Los creadores de contenido (por ejemplo, los estudios de Hollywood) desean producir la banda sonora de una película una vez y no dedicar esfuerzos en remezclarla para cada configuración de altavoz. Recientemente, los comités de normas han estado considerando maneras de proporcionar una codificación en un flujo de bits estandarizado y una decodificación posterior que sea adaptable e indiferente a la geometría del altavoz y las condiciones acústicas en la ubicación del reproductor.

[0016] Para proporcionar dicha flexibilidad para los creadores de contenido, se puede usar un conjunto jerárquico de elementos para representar un campo sonoro. El conjunto jerárquico de elementos se puede referir a un conjunto de elementos en el que los elementos se ordenan de modo 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 vuelve más detallada.

[0017] 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:

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

Esta expresión muestra que la presión p_i en cualquier punto $\{r_r, \theta_r, \phi_r\}$ del campo sonoro se puede representar únicamente por los SHC $A_n^m(k)$. En este caso, $k = \frac{\omega}{c}$, es la velocidad del sonido (~ 343 m/s), $\{r_r, \theta_r, \phi_r\}$ es un punto de referencia (o punto de observación), $j_n(\cdot)$ es la función esférica de Bessel de orden n , y $Y_n^m(\theta_r, \phi_r)$ son las funciones de base armónica esférica 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, $S(\omega, r_r, \theta_r, \phi_r)$) que se puede aproximar mediante diversas transformaciones de tiempo-frecuencia, tales como la transformada de Fourier discreta (DFT), la transformada de coseno discreta (DCT) o una transformada de ondículas. Otros ejemplos de conjuntos jerárquicos incluyen conjuntos de coeficientes de transformada de ondículas y otros conjuntos de coeficientes de funciones de base de multirresolución.

[0018] La FIG. 1 es un diagrama que ilustra las funciones de base armónica esférica desde el orden cero ($n = 0$) hasta el cuarto orden ($n = 4$). Como se puede observar, para cada orden, existe una expansión de los subórdenes m que se muestran pero no se indican explícitamente en el ejemplo de la FIG. 1, para propósitos de facilidad de ilustración.

[0019] La FIG. 2 es otro diagrama que ilustra las funciones de base armónica esférica 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 tanto con el orden como con el suborden mostrados.

[0020] En cualquier caso, los SHC $A_n^m(k)$ se pueden adquirir físicamente (por ejemplo, grabarse) mediante diversas configuraciones de formaciones de micrófonos o, de forma alternativa, se pueden derivar de descripciones basadas en canales o basadas en objetos del campo sonoro. Lo anterior representa una entrada de audio basada en la escena a un codificador. Por ejemplo, se puede usar una representación de cuarto orden que implique coeficientes $1+2^4$ (25 y, por consiguiente, de cuarto orden).

[0021] Para ilustrar cómo estos SHC se pueden derivar de una descripción basada en objetos, considérese la siguiente ecuación. Los coeficientes $A_n^m(k)$ para el campo sonoro correspondiente a un objeto de audio individual se pueden expresar como

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

donde i es $\sqrt{-1}$, $h_n^{(2)}(\cdot)$ es la función esférica de Hankel (del segundo tipo) de orden n y $\{r_s, \theta_s, \phi_s\}$ es la ubicación del objeto. Conociendo la energía fuente $g(\omega)$ como una función de la frecuencia (por ejemplo, usando técnicas de análisis de tiempo-frecuencia, tales como realizar una transformada rápida de Fourier en el flujo PCM) 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 se puede representar mediante 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 en una representación del campo sonoro global en las proximidades del punto de observación $\{r, \theta, \phi\}$. Las cifras restantes se describen a continuación en el contexto de la codificación de audio basada en objetos y basada en SHC.

[0022] Si bien los SHC se pueden derivar de los objetos PCT, los SHC también se pueden derivar de una grabación de una formación de micrófonos como sigue:

$$a_n^m(t) = b_n(r_i, t) * \langle Y_n^m(\theta_i, \phi_i), m_i(t) \rangle$$

dónde, $a_n^m(t)$ son el equivalente en el dominio del tiempo de $A_n^m(k)$ (el SHC), el $*$ representa una operación de convolución, el $\langle, \rangle$ representa un producto interno, $b_n(r_i, t)$ representa una función de filtro del dominio del tiempo dependiente de r_i , $m_i(t)$ son la i ésima señal del micrófono, donde el i ésimo transductor del micrófono está ubicado en el radio r_i , ángulo de elevación θ_i y ángulo de acimut ϕ_i . Por tanto, si existen 32 transductores en la formación de micrófonos y cada micrófono se sitúa en una esfera de modo que, $r_i = a$, es una constante (tal como las de un dispositivo Eigenmike EM32 de mhAcoustics), los 25 SHC se pueden derivar usando una operación de matrices como sigue:

$$\begin{bmatrix} a_0^0(t) \\ a_1^{-1}(t) \\ \vdots \\ a_4^4(t) \end{bmatrix} = \begin{bmatrix} b_0(a, t) \\ b_1(a, t) \\ \vdots \\ b_4(a, t) \end{bmatrix} * \begin{bmatrix} Y_0^0(\theta_1, \phi_1) & Y_0^0(\theta_2, \phi_2) & \dots & Y_0^0(\theta_{32}, \phi_{32}) \\ Y_1^{-1}(\theta_1, \phi_1) & Y_1^{-1}(\theta_2, \phi_2) & \dots & Y_1^{-1}(\theta_{32}, \phi_{32}) \\ \vdots & \vdots & \ddots & \vdots \\ Y_4^4(\theta_1, \phi_1) & Y_4^4(\theta_2, \phi_2) & \dots & Y_4^4(\theta_{32}, \phi_{32}) \end{bmatrix} \begin{bmatrix} m_1(a, t) \\ m_2(a, t) \\ \vdots \\ m_{32}(a, t) \end{bmatrix}.$$

La matriz en la ecuación anterior se puede denominar más en general $E_s(\theta, \phi)$, donde el subíndice s puede indicar que la matriz es para un determinado conjunto de geometría de transductores, s . La convolución en la ecuación anterior (indicada por el $*$), se realiza fila por fila, de modo que, por ejemplo, la salida $a_0^0(t)$ es el resultado de la convolución entre $b_0(a, t)$ y la serie de tiempo que resulta de la multiplicación vectorial de la primera fila de la matriz $E_s(\theta, \phi)$, y la columna de señales del micrófono (que varía como una función del tiempo; teniendo en cuenta el hecho de que el resultado de la multiplicación vectorial es una serie de tiempo). El cómputo puede ser más exacto cuando las posiciones del transductor de la formación de micrófonos están en las denominadas geometrías de diseño en T (que están muy cerca de la geometría de transductores Eigenmike). Una característica de la geometría de diseño en T puede ser que la matriz $E_s(\theta, \phi)$ que resulta de la geometría, tiene una inversa (o pseudoinversa) con muy buen comportamiento y, además, que la inversa a menudo se puede aproximar muy bien por la transposición de la matriz, $E_s(\theta, \phi)$. Si la operación de filtrado con $b_n(a, t)$ se ignorara, esta propiedad puede permitir la recuperación de las señales del micrófono a partir de los SHC (es decir, $[m_i(t)] = [E_s(\theta, \phi)]^{-1} [SHC]$ en este ejemplo). Las figuras restantes se describen a continuación en el contexto de la codificación de audio basada en SHC.

[0023] En general, las técnicas descritas en la presente divulgación pueden proporcionar un enfoque sólido para la transformación direccional de un campo sonoro a través del uso de un dominio de armónicos esféricos a la transformada de dominio espacial y una transformada inversa coincidente. La transformada direccional del campo sonoro se puede controlar por medio de rotación, inclinación y caída. En algunos casos, solo los coeficientes de un orden dado se combinan para crear los nuevos coeficientes, lo que significa que no existen dependencias entre órdenes tales como las que se pueden producir cuando se usan filtros. La transformada resultante entre el dominio

armónico esférico y el espacial se puede representar entonces como una operación de matrices. La transformación direccional puede, como resultado, ser completamente reversible ya que esta transformación direccional se puede cancelar mediante el uso de un reproductor transformado direccionalmente de igual manera. Una aplicación de esta transformación direccional puede ser reducir el número de coeficientes armónicos esféricos requeridos para representar un campo sonoro subyacente. La reducción se puede conseguir alineando la región de energía más alta con la dirección del campo sonoro que requiere el menor número de coeficientes armónicos esféricos para representar el campo sonoro girado. Incluso se puede lograr una reducción adicional del número de coeficientes empleando un umbral de energía. Este umbral de energía puede reducir el número de coeficientes requeridos sin la correspondiente pérdida de información perceptible. Esto puede ser beneficioso para aplicaciones que requieren la transmisión (o almacenamiento) de material de audio basado en armónicos esféricos al eliminar información espacial redundante en lugar de información espectral redundante.

[0024] La FIG. 3 es un diagrama que ilustra un sistema 20 que puede realizar las técnicas descritas en la presente divulgación para representar datos de audio potencialmente de forma más eficaz usando coeficientes armónicos esféricos. Como se muestra en el ejemplo de la FIG. 3, el sistema 20 incluye un creador de contenido 22 y un consumidor de contenido 24. Si bien se describen en el contexto del creador de contenido 22 y el consumidor de contenido 24, las técnicas se pueden implementar en cualquier contexto en el que los SHC o cualquier otra representación jerárquica de un campo sonoro se codifiquen para formar un flujo de bits representativo de los datos de audio.

[0025] El creador de contenido 22 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 24. A menudo, este creador de contenido genera contenido de audio junto con el contenido de vídeo. El consumidor de contenido 24 representa un individuo que posee o tiene acceso a un sistema de reproducción de audio, que se puede referir a cualquier forma de sistema de reproducción de audio que pueda reproducir SHC para reproducirlo como contenido de audio multicanal. En el ejemplo de la FIG. 3, el consumidor de contenido 24 incluye un sistema de reproducción de audio 32.

[0026] El creador de contenido 22 incluye un sistema de edición de audio 30. El reproductor de audio 26 puede representar una unidad de procesamiento de audio que reproduce o de otro modo genera fuentes de altavoz (que también se pueden denominar "fuentes de baffle", "señales de altavoz" o "señales de baffle"). Cada fuente de altavoz puede corresponder a una fuente de altavoz que reproduce sonido para un canal particular de un sistema de audio multicanal. En el ejemplo de la FIG. 3, el reproductor 28 puede reproducir fuentes de altavoz para formatos de sonido envolvente 5.1, 7.1 o 22.2 convencionales, generando una fuente de altavoz para cada uno de los 5, 7 o 22 altavoces en los sistemas de altavoces de sonido envolvente 5.1, 7.1 o 22.2. De forma alternativa, el reproductor 28 se puede configurar para reproducir las fuentes de altavoz a partir de los coeficientes armónicos esféricos fuente para cualquier configuración de altavoces que tenga cualquier número de altavoces, dadas las propiedades de los coeficientes armónicos esféricos fuente analizados anteriormente. El reproductor de audio 28 puede, de esta manera, generar un número de fuentes de altavoz, que se indican en la FIG. 3 como fuentes de altavoz 29.

[0027] El creador de contenido puede, durante el proceso de edición, reproducir coeficientes armónicos esféricos 27 ("SHC 27"), escuchando a las fuentes de altavoz reproducidas en un intento de identificar los aspectos del campo sonoro que no tienen alta fidelidad o que no proporcionan una experiencia de sonido envolvente convincente. El creador de contenido 22 puede a continuación editar los coeficientes armónicos esféricos fuente (a menudo indirectamente a través de la manipulación de diferentes objetos de los cuales se pueden derivar los coeficientes armónicos esféricos fuente de la manera descrita anteriormente). El creador de contenido 22 puede emplear el sistema de edición de audio 30 para editar los coeficientes armónicos esféricos 27. El sistema de edición de audio 30 representa cualquier sistema que pueda editar datos de audio y emitir estos datos de audio como uno o más coeficientes armónicos esféricos fuente.

[0028] Cuando se completa el proceso de edición, el creador de contenido 22 puede generar un flujo de bits 31 en base a los coeficientes armónicos esféricos 27. Es decir, el creador de contenido 22 incluye un dispositivo de generación de flujo de bits 36, que puede representar cualquier dispositivo que pueda generar el flujo de bits 31, por ejemplo, para la transmisión a través de un canal de transmisión, que puede ser un canal cableado o inalámbrico, un dispositivo de almacenamiento de datos, o similares, como se describe con más detalle a continuación. En algunos casos, el dispositivo de generación de flujo de bits 36 puede representar un codificador que comprime el ancho de banda (a través de la codificación de entropía, como un ejemplo) de los coeficientes armónicos esféricos 27 y que dispone la versión codificada por entropía de los coeficientes armónicos esféricos 27 en un formato aceptado para formar el flujo de bits 31. En otros casos, el dispositivo de generación de flujo de bits 36 puede representar un codificador de audio (posiblemente, uno que cumpla con una norma de codificación de audio conocida, tal como envolvente MPEG, o un derivado del mismo) que codifica el contenido de audio multicanal 29 usando, como un ejemplo, procedimientos similares a los de los procedimientos de codificación de sonido envolvente de audio convencionales para comprimir el contenido de audio multicanal o derivados de los mismos. A continuación, el contenido de audio multicanal comprimido 29 se puede codificar por entropía o codificarse de alguna otra manera para comprimir el ancho de banda del contenido 29 y disponerse de acuerdo con un formato acordado (o, en otras palabras, especificado) para formar el flujo de bits 31. Ya sea comprimido directamente para formar el flujo de bits 31 o

reproducido y a continuación comprimido para formar el flujo de bits 31, el creador de contenido 22 puede transmitir el flujo de bits 31 al consumidor de contenido 24.

[0029] Si bien se muestra en la FIG. 3 como transmitido directamente al consumidor de contenido 24, el creador de contenido 22 puede emitir el flujo de bits 31 a un dispositivo intermedio situado entre el creador de contenido 22 y el consumidor de contenido 24. Este dispositivo intermedio puede almacenar el flujo de bits 31 para su posterior entrega al consumidor de contenido 24, que puede solicitar este flujo de bits. El dispositivo intermedio puede comprender un servidor de archivos, un servidor de la Red, un ordenador de escritorio, un ordenador portátil, un ordenador de tipo tableta, un teléfono móvil, un teléfono inteligente o cualquier otro dispositivo que pueda almacenar el flujo de bits 31 para su posterior recuperación mediante un descodificador de audio. Este dispositivo intermedio puede residir en una red de entrega de contenido que pueda transmitir en flujo continuo el flujo de bits 31 (y posiblemente junto con la transmisión de un flujo de bits de datos de vídeo correspondiente) a los abonados, tal como el consumidor de contenido 24 que solicita el flujo de bits 31.

[0030] De forma alternativa, el creador de contenido 22 puede almacenar el flujo de bits 31 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 se pueden leer por un ordenador y, por lo tanto, se pueden denominar medios de almacenamiento legibles por ordenador o medios de almacenamiento legibles por ordenador no transitorios. En este contexto, el canal de transmisión se puede referir a los canales mediante los que se transmiten contenidos almacenados en estos medios (y pueden incluir tiendas minoristas y otros mecanismos de entrega basados en almacén). En cualquier caso, las técnicas de la presente divulgación, por lo tanto, no se deben limitar, en este sentido, al ejemplo de la FIG. 3.

[0031] Como se muestra además en el ejemplo de la FIG. 3, el consumidor de contenido 24 incluye el sistema de reproducción de audio 32. El sistema de reproducción de audio 32 puede representar cualquier sistema de reproducción de audio que pueda reproducir datos de audio multicanal. El sistema de reproducción de audio 32 puede incluir un número de diferentes reproductores 34. Los reproductores 34 pueden proporcionar cada uno una forma diferente de reproducción, donde las diferentes formas de reproducción pueden incluir una o más de las diversas maneras de realizar la exploración de amplitud de base vectorial (VBAP) y/o una o más de las diversas maneras de realizar la síntesis del campo sonoro.

[0032] El sistema de reproducción de audio 32 puede incluir además un dispositivo de extracción 38. El dispositivo de extracción 38 puede representar cualquier dispositivo que pueda extraer los coeficientes armónicos esféricos 27' ("SHC 27'", que pueden representar una forma modificada o un duplicado de los coeficientes armónicos esféricos 27) a través de un procedimiento que en general puede ser recíproco al del dispositivo de generación de flujo de bits 36. En cualquier caso, el sistema de reproducción de audio 32 puede recibir los coeficientes armónicos esféricos 27' y puede seleccionar uno de los reproductores 34. El seleccionado de los reproductores 34 puede a continuación reproducir los coeficientes armónicos esféricos 27' para generar un número de fuentes de altavoz 35 (correspondientes al número de altavoces acoplados de forma eléctrica o posiblemente inalámbrica al sistema de reproducción de audio 32, que no se muestran en el ejemplo de la FIG. 3 para propósitos de facilidad de ilustración).

[0033] Típicamente, cuando el dispositivo de generación de flujo de bits 36 codifica directamente SHC 27, el dispositivo de generación de flujo de bits 36 codifica la totalidad de SHC 27. El número de SHC 27 enviados para cada representación del campo sonoro depende del orden y se puede expresar matemáticamente como $(1+n)^2$ /muestra, donde n de nuevo indica el orden. Para lograr una representación de cuarto orden del campo sonoro, como un ejemplo, se pueden derivar 25 SHC. Típicamente, cada uno de los SHC se expresa como un número de punto flotante con signo de 32 bits. Por tanto, para expresar una representación de cuarto orden del campo sonoro, se requiere un total de 25×32 u 800 bits/muestra en este ejemplo. Cuando se usa una frecuencia de muestreo de 48 kHz, esto representa 800×48.000 o 38.400.000 bits/segundo. En algunos casos, uno o más de los SHC 27 no pueden especificar información relevante (que se puede referir a información que contiene información de audio audible o importante al describir el campo sonoro cuando se reproduce en el consumidor de contenido 24). La codificación de estos no relevantes de los SHC 27 puede dar como resultado un uso ineficiente del ancho de banda a través del canal de transmisión (suponiendo un tipo de red de entrega de contenido de mecanismo de transmisión). En una aplicación que implica el almacenamiento de estos coeficientes, lo anterior puede representar un uso ineficiente del espacio de almacenamiento.

[0034] En algunos casos, cuando se identifica un subconjunto de los SHC 27 que se incluyen en el flujo de bits 31, el dispositivo de generación de flujo de bits 36 puede especificar un campo que tiene una pluralidad de bits con uno diferente de la pluralidad de bits que identifica si uno correspondiente de los SHC 27 se incluye en el flujo de bits 31. En algunos casos, cuando se identifica un subconjunto de los SHC 27 que se incluyen en el flujo de bits 31, el dispositivo de generación de flujo de bits 36 puede especificar un campo que tenga una pluralidad de bits igual a $(n+1)^2$ bits, donde n indica un orden del conjunto jerárquico de elementos que describen el campo sonoro, y donde cada uno de la pluralidad de bits identifica si uno correspondiente de los SHC 27 se incluye en el flujo de bits 31.

[0035] En algunos casos, el dispositivo de generación de flujo de bits 36 puede, cuando se identifica un subconjunto de los SHC 27 que se incluyen en el flujo de bits 31, especificar un campo en el flujo de bits 31 que tiene una pluralidad

de bits con uno diferente de la pluralidad de bits que identifica si uno correspondiente de los SHC 27 se incluye en el flujo de bits 31. Cuando se especifica el subconjunto identificado de los SHC 27, el dispositivo de generación de flujo de bits 36 puede especificar, en el flujo de bits 31, el subconjunto identificado de los SHC 27 directamente después del campo que tiene la pluralidad de bits.

[0036] En algunos casos, el dispositivo de generación de flujo de bits 36 puede determinar adicionalmente que uno o más de los SHC 27 tiene información pertinente al describir el campo sonoro. Cuando se identifica el subconjunto de los SHC 27 que se incluyen en el flujo de bits 31, el dispositivo de generación de flujo de bits 36 puede identificar que el uno o más de los SHC 27 determinados que tienen información pertinente al describir el campo sonoro se incluyen en el flujo de bits 31.

[0037] En algunos casos, el dispositivo de generación de flujo de bits 36 puede determinar adicionalmente que uno o más de los SHC 27 tienen información pertinente al describir el campo sonoro. Cuando se identifica el subconjunto de los SHC 27 que se incluyen en el flujo de bits 31, el dispositivo de generación de flujo de bits 36 puede identificar, en el flujo de bits 31, que el uno o más de los SHC 27 determinados que tienen información pertinente al describir el campo sonoro se incluyen en el flujo de bits 31, e identifica, en el flujo de bits 31, que los restantes de los SHC 27 que tienen información no pertinente al describir el campo sonoro no se incluyen en el flujo de bits 31.

[0038] En algunos casos, el dispositivo de generación de flujo de bits 36 puede determinar que uno o más de los valores de SHC 27 están por debajo de un valor umbral. Cuando se identifica el subconjunto de los SHC 27 que se incluyen en el flujo de bits 31, el dispositivo de generación de flujo de bits 36 puede identificar, en el flujo de bits 31, que el uno o más de los SHC 27 determinados que están por encima de este valor umbral se especifican en el flujo de bits 31. Si bien el umbral a menudo puede tener un valor de cero, para implementaciones prácticas, el umbral se puede establecer en un valor que represente un piso de ruido (o energía ambiental) o algún valor proporcional a la energía de la señal actual (que puede hacer dependiente la señal umbral).

[0039] En algunos casos, el dispositivo de generación de flujo de bits 36 puede ajustar o transformar el campo sonoro para reducir un número de los SHC 27 que proporcionan información pertinente al describir el campo sonoro. El término "ajustar" se puede referir a la aplicación de cualquier matriz o matrices que representen una transformada invertible lineal. En estos casos, el dispositivo de generación de flujo de bits 36 puede especificar información de ajuste (que también se puede denominar "información de transformación") en el flujo de bits 31 que describe cómo se ajustó el campo sonoro o, en otras palabras, se transformó. Si bien se describe que especifica esta información además de la información que identifica el subconjunto de los SHC 27 que se especifican posteriormente en el flujo de bits, este aspecto de las técnicas se puede realizar como una alternativa a especificar la información que identifica el subconjunto de los SHC 27 que se incluyen en el flujo de bits. Por lo tanto, las técnicas no se deben limitar en este sentido.

[0040] En algunos casos, el dispositivo de generación de flujo de bits 36 puede girar el campo sonoro para reducir un número de los SHC 27 que proporcionan información pertinente al describir el campo sonoro. En estos casos, el dispositivo de generación de flujo de bits 36 puede especificar información de rotación en el flujo de bits 31 que describe cómo se giró el campo sonoro. La información de rotación puede comprender un valor de acimut (que puede señalar 360 grados) y un valor de elevación (que puede señalar 180 grados). En algunos casos, el valor de acimut comprende uno o más bits, y típicamente incluye 10 bits. En algunos casos, el valor de elevación comprende uno o más bits y típicamente incluye al menos 9 bits. Esta elección de bits permite, en el modo de realización más simple, una resolución de 180/512 grados (tanto en elevación como en acimut). En algunos casos, la transformación puede comprender la rotación y la información de transformación descrita anteriormente incluye la información de rotación. En algunos casos, el dispositivo de generación de flujo de bits 36 puede transformar el campo sonoro para reducir un número de los SHC 27 que proporcionan información pertinente al describir el campo sonoro. En estos casos, el dispositivo de generación de flujo de bits 36 puede especificar información de transformación en el flujo de bits 31 que describe cómo se transformó el campo sonoro. En algunos casos, el ajuste puede comprender la transformación y la información de ajuste descrita anteriormente incluye la información de transformación.

[0041] En algunos casos, el dispositivo de generación de flujo de bits 36 puede ajustar el campo sonoro para reducir un número de los SHC 27 que tienen valores distintos de cero por encima de un valor umbral y especificar la información de ajuste en el flujo de bits 31 que describe cómo se ajustó el campo sonoro. En algunos casos, el dispositivo de generación de flujo de bits 36 puede girar el campo sonoro para reducir un número de los SHC 27 que tienen valores distintos de cero por encima de un valor umbral, y especificar información de rotación en el flujo de bits 31 que describe cómo se giró el campo sonoro. En algunos casos, el dispositivo de generación de flujo de bits 36 puede transformar el campo sonoro para reducir un número de los SHC 27 que tienen valores distintos de cero por encima de un valor umbral, y especificar información de transformación en el flujo de bits 31 que describe cómo se transformó el campo sonoro.

[0042] Al identificar el flujo de bits 31 el subconjunto de los SHC 27 que se incluyen en el flujo de bits 31, el dispositivo de generación de flujo de bits 36 puede promover un uso más eficaz del ancho de banda ya que el subconjunto de los SHC 27 que no incluyen información pertinente para la descripción del campo sonoro (tales como los de valor cero de los SCH 27) no se especifican en el flujo de bits, es decir, no se incluyen en el flujo de bits. Asimismo, al ajustar

adicionalmente o de forma alternativa el campo sonoro cuando se generan los SHC 27 para reducir el número de SHC 27 que especifican información pertinente para la descripción del campo sonoro, el dispositivo de generación de flujo de bits 36 puede proporcionar de nuevo o adicionalmente un uso de ancho de banda potencialmente más eficaz. De esta manera, el dispositivo de generación de flujo de bits 31 puede reducir el número de SHC 27 que se requiere que se especifiquen en el flujo de bits 31, potencialmente mejorando de este modo la utilización del ancho de banda en sistemas de tasa no fija (que se pueden referir a técnicas de codificación de audio que no tengan una tasa de bits objetivo o proporcionen un presupuesto de bits por trama o muestra para proporcionar algunos ejemplos) o, en el sistema de tasa fija, dando como resultado potencialmente la adjudicación de bits a la información que es más pertinente al describir el campo sonoro.

[0043] Adicionalmente o de forma alternativa, el dispositivo de generación de flujo de bits 36 puede funcionar de acuerdo con las técnicas descritas en la presente divulgación para asignar diferentes tasas de bits a diferentes subconjuntos de los coeficientes armónicos esféricos transformados. En virtud de transformar, por ejemplo, girar, el campo sonoro, el dispositivo de generación de flujo de bits 36 puede alinear las partes más relevantes (a menudo identificadas a través del análisis de energía en diversas ubicaciones espaciales del campo sonoro) con un eje, tal como el eje Z, estableciendo eficazmente las partes de mayor energía por encima del oyente en el campo sonoro. En otras palabras, el dispositivo de generación de flujo de bits 36 puede analizar la energía del campo sonoro para identificar la parte del campo sonoro que tiene la energía más alta. Si dos o más partes del campo sonoro tienen alta energía, el dispositivo de generación de flujo de bits 36 puede comparar estas energías para identificar la que tiene la energía más alta. A continuación, el dispositivo de generación de flujo de bits 36 puede identificar uno o más ángulos mediante los que girar el campo sonoro a fin de alinear la parte de energía más alta del campo sonoro con el eje Z.

[0044] Esta rotación u otra transformación se puede considerar como una transformación de una trama de referencia en la que se establecen las funciones de base esférica. En lugar de mantener el eje Z, tal como los que se muestran en el ejemplo de la FIG. 2, que es recto hacia arriba y hacia abajo, este eje Z se puede transformar por uno o más ángulos para apuntar en la dirección de la parte de energía más alta del campo sonoro. Las funciones de base que tienen algún componente direccional, tal como la función de base esférica de orden uno y suborden cero que se alinea con el eje Z, entonces se pueden girar. El campo sonoro se puede entonces expresar usando estas funciones de base esférica transformadas, por ejemplo, giradas. El dispositivo de generación de flujo de bits 36 puede girar esta trama de referencia de manera que el eje Z se alinee con la parte de energía más alta del campo sonoro. Esta rotación puede dar como resultado que la energía más alta del campo sonoro se exprese principalmente por esas funciones de base de suborden cero, mientras que las funciones de base de suborden distinto de cero pueden no contener tanta información relevante.

[0045] Una vez girado de esta manera, el dispositivo de generación de flujo de bits 36 puede determinar coeficientes armónicos esféricos transformados, que se refiere a los coeficientes armónicos esféricos asociados con las funciones de base esférica transformadas. Dado que las funciones de base esférica de suborden cero pueden representar principalmente el campo sonoro, el dispositivo de generación de flujo de bits 36 puede asignar una primera tasa de bits para expresar estos coeficientes armónicos esféricos transformados de suborden cero (que se pueden referir a los coeficientes armónicos esféricos transformados correspondientes a funciones de base de suborden cero) en el flujo de bits 31, mientras asigna una segunda tasa de bits para expresar los coeficientes armónicos esféricos transformados de suborden distinto de cero (que se pueden referir a los coeficientes armónicos esféricos transformados correspondientes a funciones de base de suborden distinto de cero) en el flujo de bits 36, donde esta primera tasa de bits es mayor que la segunda tasa de bits. En otras palabras, debido a que los coeficientes armónicos esféricos transformados de suborden cero describen las partes más relevantes del campo sonoro, el dispositivo de generación de flujo de bits 36 puede asignar una tasa de bits más alta para expresar estos coeficientes transformados en el flujo de bits, mientras que asigna una tasa de bits más baja (en relación con la tasa de bits más alta) para expresar estos coeficientes en el flujo de bits.

[0046] Cuando se asignan estas tasas de bits a lo que se puede denominar el primer subconjunto de los coeficientes armónicos esféricos transformados (por ejemplo, los coeficientes armónicos esféricos transformados de suborden cero) y el segundo subconjunto de los coeficientes armónicos esféricos transformados (por ejemplo, los coeficientes armónicos esféricos transformados de suborden distinto de cero), el dispositivo de generación de flujo de bits 36 puede utilizar una función ventana, tal como una función ventana de Hanning, una función ventana de Hamming, una función ventana rectangular o una función ventana triangular. Si bien se describe con respecto al primer y segundo subconjuntos de los coeficientes armónicos esféricos transformados, el dispositivo de generación de flujo de bits 36 puede identificar dos, tres, cuatro y a menudo hasta $2 \cdot n + 1$ (donde n se refiere al orden) subconjuntos de los coeficientes armónicos esféricos. Típicamente, cada suborden para el orden puede representar otro subconjunto de los coeficientes armónicos esféricos transformados a los que el dispositivo de generación de flujo de bits 36 asigna una tasa de bits diferente.

[0047] En este sentido, el dispositivo de generación de flujo de bits 36 puede asignar dinámicamente diferentes tasas de bits a aquellos diferentes de los SHC 27 por orden y/o suborden. Esta adjudicación dinámica de tasas de bits puede facilitar un mejor uso de la tasa de bits objetivo global, asignando tasas de bits más altas a aquellos de los SHC 27 transformados que describen partes más relevantes del campo sonoro mientras asigna tasas de bits más bajas (en

comparación con las tasas de bits más altas) a aquellos de los SHC 27 transformados que describen partes comparativamente menos relevantes (o, en otras palabras, partes ambientales o de fondo) del campo sonoro.

[0048] Para ilustrar, considérese una vez más el ejemplo de la FIG. 2. El dispositivo de generación de flujo de bits 36 puede, basado en la función ventana, asignar una tasa de bits a cada suborden de los coeficientes armónicos esféricos transformados, donde para el cuarto (4) orden, el dispositivo de generación de flujo de bits 36 identifica nueve (de menos cuatro a más cuatro) subconjuntos diferentes de los coeficientes armónicos esféricos transformados. Por ejemplo, el dispositivo de generación de flujo de bits 36 puede, basado en la función ventana, asignar una primera tasa de bits para expresar los coeficientes armónicos esféricos transformados de suborden 0, una segunda tasa de bits para expresar los coeficientes armónicos esféricos transformados de suborden -1/+1, una tercera tasa de bits para expresar los coeficientes armónicos esféricos transformados de suborden -2/+2, una cuarta tasa de bits para expresar los coeficientes armónicos esféricos transformados de suborden -3/+3 y una quinta tasa de bits para expresar los coeficientes armónicos esféricos transformados de suborden -4/+4.

[0049] En algunos casos, el dispositivo de generación de flujo de bits 36 puede asignar tasas de bits de una manera incluso más granular, donde la tasa de bits no varía sólo por suborden, sino también por orden. Dado que las funciones de base esférica de orden superior tienen lóbulos más pequeños, estas funciones de base esférica de orden superior no son tan importantes al representar partes de alta energía del campo sonoro. Como resultado, el dispositivo de generación de flujo de bits 36 puede asignar una tasa de bits más baja a los coeficientes armónicos esféricos transformados de orden superior relativa a esta tasa de bits asignada a los coeficientes armónicos esféricos transformados de orden inferior. De nuevo, el dispositivo de generación de flujo de bits 36 puede asignar estas tasas de bits específicas de orden basado en una función ventana de una manera similar a la descrita anteriormente con respecto a la asignación de las tasas de bits específicas de suborden.

[0050] En este sentido, el dispositivo de generación de flujo de bits 36 puede asignar una tasa de bits a al menos un subconjunto de coeficientes armónicos esféricos transformados basado en uno o más de un orden y un suborden de una función de base esférica a la que el subconjunto de los coeficientes armónicos esféricos transformados corresponde, habiéndose transformado los coeficientes armónicos esféricos de acuerdo con una operación de transformada que transforma un campo sonoro.

[0051] En algunos casos, la operación de transformación comprende una operación de rotación que gira el campo sonoro.

[0052] En algunos casos, el dispositivo de generación de flujo de bits 36 puede identificar uno o más ángulos mediante los que girar el campo sonoro de modo que una parte del campo sonoro que tiene la energía más alta se alinea con un eje, donde la operación de transformación puede comprender una operación de rotación que gira el campo sonoro mediante el uno o más ángulos identificados a fin de generar los coeficientes armónicos esféricos transformados.

[0053] En algunos casos, el dispositivo de generación de flujo de bits 36 puede identificar uno o más ángulos mediante los que girar el campo sonoro de modo que una parte del campo sonoro que tiene la energía más alta se alinea con un eje Z, donde la operación de transformación puede comprender una operación de rotación que gira el campo sonoro mediante el uno o más ángulos identificados a fin de generar los coeficientes armónicos esféricos transformados.

[0054] En algunos casos, el dispositivo de generación de flujo de bits 36 puede realizar un análisis espacial con respecto al campo sonoro para identificar uno o más ángulos mediante los que girar el campo sonoro, donde la operación de transformación puede comprender una operación de rotación que gira el campo sonoro mediante el uno o más ángulos identificados a fin de generar los coeficientes armónicos esféricos transformados.

[0055] En algunos casos, el dispositivo de generación de flujo de bits 36, cuando asigna la tasa de bits, puede asignar dinámicamente, de acuerdo con una función ventana, diferentes tasas de bits a diferentes subconjuntos de los coeficientes armónicos esféricos transformados basado en uno o más del orden y el suborden de la función de base esférica a la que corresponde cada uno de los coeficientes armónicos esféricos transformados. La función ventana puede comprender una o más de una función ventana de Hanning, una función ventana de Hamming, una función ventana rectangular y una función ventana triangular.

[0056] En algunos casos, el dispositivo de generación de flujo de bits 36, cuando asigna la tasa de bits, puede asignar una primera tasa de bits a un primer subconjunto de los coeficientes armónicos esféricos transformados correspondientes al subconjunto de las funciones de base esférica que tienen un suborden de cero, y asigna una segunda tasa de bits a un segundo subconjunto de los coeficientes armónicos esféricos transformados correspondientes al subconjunto de las funciones de base esférica que tienen un suborden de uno positivo o bien negativo, siendo la primera tasa de bits mayor que la segunda tasa de bits. En este sentido, las técnicas pueden proporcionar la asignación dinámica de tasas de bits basada en el suborden de las funciones de base esférica a las que corresponde el SHC 27.

[0057] En algunos casos, el dispositivo de generación de flujo de bits 36, cuando asigna la tasa de bits, puede asignar una primera tasa de bits a un primer subconjunto de los coeficientes armónicos esféricos transformados

correspondientes al subconjunto de la función de base esférica que tiene un orden de uno, y asignar una segunda tasa de bits a un segundo subconjunto de los coeficientes armónicos esféricos transformados correspondientes al subconjunto de las funciones de base esférica que tienen un orden de dos, siendo la primera tasa de bits mayor que la segunda tasa de bits. De esta manera, las técnicas pueden proporcionar la asignación dinámica de tasas de bits basada en el orden de las funciones de base esférica a las que corresponden los SHC 27.

[0058] En algunos casos, el dispositivo de generación de flujo de bits 36 puede generar un flujo de bits que especifica el primer subconjunto de los coeficientes armónicos esféricos transformados usando la primera tasa de bits y el segundo subconjunto de los coeficientes armónicos esféricos transformados usando la segunda tasa de bits.

[0059] En algunos casos, el dispositivo de generación de flujo de bits 36, cuando asigna la tasa de bits, puede asignar dinámicamente tasas de bits progresivamente decrecientes a medida que el suborden de las funciones de base esférica a las que corresponden los coeficientes armónicos esféricos transformados se aleja de cero.

[0060] En algunos casos, el dispositivo de generación de flujo de bits 36, cuando asigna la tasa de bits, puede asignar dinámicamente tasas de bits progresivamente decrecientes a medida que el orden de las funciones de base esférica a las que corresponden los coeficientes armónicos esféricos transformados aumenta.

[0061] En algunos casos, el dispositivo de generación de flujo de bits 36, cuando asigna la tasa de bits, puede asignar dinámicamente diferentes tasas de bits a diferentes subconjuntos de coeficientes armónicos esféricos transformados basado en uno o más del orden y el suborden de la función de base esférica a la que corresponde el subconjunto de los coeficientes armónicos esféricos transformados.

[0062] Dentro del consumidor de contenido 24, el dispositivo de extracción 38 puede a continuación realizar un procedimiento de procesamiento del flujo de bits 31 representativo del contenido de audio de acuerdo con aspectos de las técnicas recíprocas a los descritos anteriormente con respecto al dispositivo de generación de flujo de bits 36. El dispositivo de extracción 38 puede determinar, a partir del flujo de bits 31, el subconjunto de los SHC 27' que describen un campo sonoro que se incluyen en el flujo de bits 31, y analizar sintácticamente el flujo de bits 31 para determinar el subconjunto identificado de los SHC 27'.

[0063] En algunos casos, el dispositivo de extracción 38 puede, cuando determina el subconjunto de los SHC 27' que se incluyen en el flujo de bits 31, el dispositivo de extracción 38 puede analizar sintácticamente el flujo de bits 31 para determinar un campo que tiene una pluralidad de bits identificando cada uno de la pluralidad de bits si uno de los SHC 27' correspondientes se incluye en el flujo de bits 31.

[0064] En algunos casos, el dispositivo de extracción 38, cuando determina el subconjunto de los SHC 27' que se incluyen en el flujo de bits 31, puede especificar un campo que tiene una pluralidad de bits igual a $(n+1)^2$ bits, donde de nuevo n denota un orden del conjunto jerárquico de elementos que describen el campo sonoro. De nuevo, cada uno de la pluralidad de bits identifica si se incluye uno de los SHC 27' correspondientes en el flujo de bits 31.

[0065] En algunos casos, el dispositivo de extracción 38, cuando determina el subconjunto de los SHC 27' que se incluyen en el flujo de bits 31, puede analizar sintácticamente el flujo de bits 31 para identificar un campo en el flujo de bits 31 que tenga una pluralidad de bits con uno diferente de la pluralidad de bits que identifica si uno de los SHC 27' correspondientes se incluye en el flujo de bits 31. El dispositivo de extracción 38 cuando analiza sintácticamente el flujo de bits 31 para determinar el subconjunto identificado de los SHC 27', puede analizar sintácticamente el flujo de bits 31 para determinar el subconjunto identificado de los SHC 27' directamente a partir del flujo de bits 31 después del campo que tiene la pluralidad de bits.

[0066] En algunos casos, el dispositivo de extracción 38 puede analizar sintácticamente el flujo de bits 31 para determinar la información de ajuste que describe cómo se ajustó el campo sonoro para reducir un número de los SHC 27' que proporcionan información pertinente al describir el campo sonoro. El dispositivo de extracción 38 puede proporcionar esta información al sistema de reproducción de audio 32, que cuando reproduce el campo sonoro basado en el subconjunto de los SHC 27' que proporcionan información pertinente al describir el campo sonoro, ajusta el campo sonoro basado en la información de ajuste para invertir el ajuste realizado para reducir el número de la pluralidad de elementos jerárquicos.

[0067] En algunos casos, el dispositivo de extracción 38, como alternativa a o junto con los aspectos de las técnicas descritos anteriormente, puede analizar sintácticamente el flujo de bits 31 para determinar la información de rotación que describe cómo se giró el campo sonoro para reducir un número de los SHC 27' que proporcionan información pertinente al describir el campo sonoro. El dispositivo de extracción 38 puede proporcionar esta información al sistema de reproducción de audio 32, que cuando reproduce el campo sonoro basado en el subconjunto de los SHC 27' que proporcionan información pertinente al describir el campo sonoro, gira el campo sonoro basado en la información de rotación para invertir la rotación realizada para reducir el número de la pluralidad de elementos jerárquicos.

[0068] En algunos casos, el dispositivo de extracción 38, como alternativa a o junto con los aspectos de las técnicas descritos anteriormente, puede analizar sintácticamente el flujo de bits 31 para determinar la información de

transformación que describe cómo el campo sonoro se transformó para reducir un número de los SHC 27' que proporcionan información pertinente al describir el campo sonoro. El dispositivo de extracción 38 puede proporcionar esta información al sistema de reproducción de audio 32, que cuando reproduce el campo sonoro basado en el subconjunto de los SHC 27' que proporciona información pertinente al describir el campo sonoro, transforma el campo sonoro basado en la información de ajuste para invertir la transformación realizada para reducir el número de la pluralidad de elementos jerárquicos.

[0069] En algunos casos, el dispositivo de extracción 38, como alternativa a o junto con los aspectos de las técnicas descritos anteriormente, puede analizar sintácticamente el flujo de bits 31 para determinar la información de ajuste que describe cómo se ajustó el campo sonoro para reducir un número de los SHC 27' que tienen valores distintos de cero. El dispositivo de extracción 38 puede proporcionar esta información al sistema de reproducción de audio 32, que cuando reproduce el campo sonoro basado en el subconjunto de los SHC 27' que tienen valores distintos de cero, ajusta el campo sonoro basado en la información de ajuste para invertir el ajuste realizado para reducir el número de la pluralidad de elementos jerárquicos.

[0070] En algunos casos, el dispositivo de extracción 38, como alternativa a o junto con los aspectos de las técnicas descritos anteriormente, puede analizar sintácticamente el flujo de bits 31 para determinar la información de rotación que describe cómo se giró el campo sonoro para reducir un número de los SHC 27' que tienen valores distintos de cero. El dispositivo de extracción 38 puede proporcionar esta información al sistema de reproducción de audio 32, que cuando reproduce el campo sonoro basado en el subconjunto de los SHC 27' que tienen valores distintos de cero, gira el campo sonoro basado en la información de rotación para invertir la rotación realizada para reducir el número de la pluralidad de elementos jerárquicos.

[0071] En algunos casos, el dispositivo de extracción 38, como alternativa a o junto con los aspectos de las técnicas descritos anteriormente, puede analizar sintácticamente el flujo de bits 31 para determinar la información de transformación que describe cómo el campo sonoro se transformó para reducir un número de los SHC 27' que tienen valores distintos de cero. El dispositivo de extracción 38 puede proporcionar esta información al sistema de reproducción de audio 32, que cuando reproduce el campo sonoro basado en los de los SHC 27' que tienen valores distintos de cero, transforma el campo sonoro basado en la información de transformación para invertir la transformación realizada para reducir el número de la pluralidad de elementos jerárquicos.

[0072] En este sentido, diversos aspectos de las técnicas pueden hacer posible la señalización, en un flujo de bits, de los de una pluralidad de elementos jerárquicos, tales como los coeficientes ambisónicos de orden superior (HOA) (que también se pueden denominar coeficientes armónicos esféricos), que se incluyen en el flujo de bits (donde los que se van a incluir en el flujo de bits se pueden denominar "subconjunto de la pluralidad de los SHC"). Dado que algunos de los coeficientes HOA pueden no proporcionar información pertinente al describir un campo sonoro, el codificador de audio puede reducir la pluralidad de coeficientes HOA a un subconjunto de los coeficientes HOA que proporcionan información pertinente al describir el campo sonoro, aumentando de este modo la eficacia de codificación. Como resultado, diversos aspectos de las técnicas pueden hacer posible especificar en el flujo de bits que incluye los coeficientes HOA y/o las versiones codificadas de los mismos, los de los coeficientes HOA que se incluyen realmente en el flujo de bits (por ejemplo, el subconjunto distinto de cero de los coeficientes HOA que incluye al menos uno de los coeficientes HOA pero no la totalidad de los coeficientes). La información que identifica el subconjunto de los coeficientes HOA se puede especificar en el flujo de bits como se indica anteriormente, o en algunos casos, en la información de canal lateral.

[0073] Las FIGS. 4A y 4B son diagramas de bloques que ilustran una implementación de ejemplo del dispositivo de generación de flujo de bits 36. Como se ilustra en el ejemplo de la FIG. 4A, la primera implementación del dispositivo de generación de flujo de bits 36, indicado como dispositivo de generación de flujo de bits 36A, incluye una unidad de análisis espacial 150, una unidad de rotación 154, un motor de codificación 160 y un multiplexor (MUX) 164.

[0074] El ancho de banda, en términos de bits/segundo, requerido para representar datos de audio 3D en forma de SHC puede hacerlo prohibitivo en términos de uso del consumidor. Por ejemplo, cuando se usa una frecuencia de muestreo de 48 kHz y con 32 bits/misma resolución, una representación de SHC de cuarto orden representa un ancho de banda de 36 Mbits/segundo ($25 \times 48000 \times 32$ 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. 5 pueden reducir el ancho de banda de las representaciones de audio 3D.

[0075] La unidad de análisis espacial 150 y la unidad de rotación 154 pueden recibir SHC 27. Como se describe en otra parte de la presente divulgación, los SHC 27 pueden ser representativos de un campo sonoro. En el ejemplo de la FIG. 4A, la unidad de análisis espacial 150 y la unidad de rotación 154 pueden recibir muestras de veinticinco SHC para una representación de cuarto orden ($N = 4$) del campo sonoro. Típicamente, una trama de datos de audio incluye 1028 muestras, aunque las técnicas se pueden realizar con respecto a una trama que tenga cualquier número de muestras. La unidad de análisis espacial 150 y la unidad de rotación 154 pueden funcionar de la manera descrita a continuación con respecto a una trama de los datos de audio. Si bien se describe como funcionando en una trama de

datos de audio, las técnicas se pueden realizar con respecto a cualquier cantidad de datos de audio, incluyendo una muestra única y hasta la totalidad de los datos de audio.

[0076] La unidad de análisis espacial 150 puede analizar el campo sonoro representado por los SHC 27 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 de otro modo son distintos de los componentes de fondo o difusos del campo sonoro. Por ejemplo, el sonido generado por un instrumento musical individual se puede percibir 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. En algunos casos, los distintos componentes también se pueden denominar "componentes relevantes" o "componentes de primer plano", mientras que los componentes difusos se pueden denominar "componentes ambientales" o "componentes de fondo".

[0077] Típicamente, estos distintos componentes tienen alta energía en una ubicación identificable del campo sonoro. La unidad de análisis espacial 150 puede identificar estas ubicaciones de "alta energía" del campo sonoro, analizando cada ubicación de alta energía para determinar una ubicación en el campo sonoro que tenga la energía más alta. A continuación, la unidad de análisis espacial 150 puede determinar un ángulo óptimo mediante el que girar el campo sonoro para alinear los de los distintos componentes que tienen la mayor energía con un eje (en relación con un presunto micrófono que registró este campo sonoro), tal como el eje Z. La unidad de análisis espacial 150 puede identificar este ángulo óptimo de manera que el campo sonoro se pueda girar de modo que estos distintos componentes se alineen mejor con las funciones de base esféricas subyacentes mostradas en los ejemplos de las FIGS. 1 y 2.

[0078] En algunos ejemplos, la unidad de análisis espacial 150 puede representar una unidad configurada para realizar una forma de análisis de difusión para identificar un porcentaje del campo sonoro representado por los SHC 27 que incluye sonidos difusos (que se pueden referir a sonidos que tienen bajos niveles de dirección o SHC de orden inferior, es decir, los de SHC 27 que tienen un orden menor o igual a uno). Como un ejemplo, la unidad de análisis espacial 150 puede realizar un análisis de difusión de una manera similar a la descrita en un documento de Ville Pulkki, titulado "Reproducción de sonido espacial con codificación de audio direccional", publicado en J. Audio Eng. Soc., vol. 55, n.º 6, con fecha de junio de 2007. En algunos casos, la unidad de análisis espacial 150 sólo puede analizar un subconjunto distinto de cero de los coeficientes SHC 27, tales como los de orden cero y de primer orden de los SHC 27, al realizar el análisis de difusión para determinar el porcentaje de difusión.

[0079] La unidad de rotación 154 puede realizar una operación de rotación de los SHC 27 basado en el ángulo óptimo identificado (o ángulos como sea el caso). Como se analiza en otra parte de la presente divulgación (por ejemplo, con respecto a la FIG. 5A y 5B), realizar la operación de rotación puede reducir el número de bits requeridos para representar los SHC 27. La unidad de rotación 154 puede emitir coeficientes armónicos esféricos transformados 155 ("SHC transformados 155") al motor de codificación 160.

[0080] El motor de codificación 160 puede representar una unidad configurada para comprimir el ancho de banda de los SHC transformados 155. El motor de codificación 160 puede asignar diferentes tasas de bits a diferentes subconjuntos de los SHC transformados 155 de acuerdo con las técnicas descritas en la presente divulgación. Como se muestra en el ejemplo de la FIG. 4A, el motor de codificación 160 incluye una función ventana 161 y unidades de codificación AAC 163. El motor de codificación 160 puede aplicar la función ventana 161 a una tasa de bits objetivo para asignar tasas de bits a una o más de las unidades de codificación AAC 163. Las funciones ventana 161 pueden identificar tasas de bits diferentes para cada orden y/o suborden de las funciones de base esférica a las que corresponden los SHC transformados 155. A continuación, el motor de codificación 160 puede configurar la unidad de codificación AAC 163 con las tasas de bits identificadas, con lo cual el motor de codificación 160 puede dividir los SHC transformados 155 en diferentes subconjuntos y pasar estos subconjuntos diferentes a uno correspondiente de las unidades de codificación AAC 163. Es decir, si se configura una tasa de bits en una de las unidades de codificación AAC 163 para los de los SHC transformados 155 correspondientes a las funciones de base esférica de suborden cero, el motor de codificación 160 pasa los de los SHC transformados 127 que corresponden a las funciones de base esférica de suborden cero a la de las unidades de codificación AAC 163. A continuación, las unidades de codificación AAC 163 pueden realizar AAC con respecto a los subconjuntos de los SHC transformados 155, emitiendo versiones comprimidas del subconjunto diferente de los SHC transformados 155 al multiplexor 164. A continuación, el multiplexor 164 puede multiplexar estos subconjuntos conjuntamente con el ángulo óptimo para generar el flujo de bits 31.

[0081] Como se ilustra en el ejemplo de la FIG. 4B, el dispositivo de generación de flujo de bits 36B incluye una unidad de análisis espacial 150, una unidad de análisis de características de contenido 152, una unidad de rotación 154, una unidad de componentes coherentes de extracción 156, una unidad de componentes difusos de extracción 158, motores de codificación 160 y un multiplexor (MUX) 164. Aunque es similar al dispositivo de generación de flujo de bits 36A, el dispositivo de generación de flujo de bits 36B incluye unidades adicionales 152, 156 y 158.

[0082] La unidad de análisis de características de contenido 152 puede determinar basada, al menos en parte, en los SHC 27, si los SHC 27 se generaron por medio 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 152 puede determinar a continuación, basada al menos en parte en si los SHC 27 se generaron por medio de una grabación real de un campo sonoro o a partir de un objeto de audio artificial, el número total de canales a incluir en el flujo de bits 31. Por ejemplo, la unidad de análisis de características de contenido 152 puede determinar, basada al menos en parte en si los SHC 27 se generaron a partir de una grabación de un campo sonoro real o a partir de un objeto de audio artificial, que el flujo de bits 31 va a incluir dieciséis canales. Cada uno de los canales puede ser un canal mono. La unidad de análisis de características de contenido 152 puede realizar además la determinación del número total de canales a incluir en el flujo de bits 31 basada en una tasa de bits de salida del flujo de bits 31, por ejemplo, de 1,2 Mbps.

[0083] Además, la unidad de análisis de características de contenido 152 puede determinar, basada al menos en parte en si los SHC 27 se generaron a partir de una grabación de un campo sonoro real o a partir de un objeto de audio artificial, cuántos de los canales adjudicar a componentes coherentes o, en otras palabras, distintos del campo sonoro y cuántos de los canales adjudicar a componentes difusos o, en otras palabras, de fondo del campo sonoro. Por ejemplo, cuando los SHC 27 se generaron a partir de una grabación de un campo sonoro real usando, como un ejemplo, un Eigenmic, la unidad de análisis de características de contenido 152 puede adjudicar tres de los canales a componentes coherentes del campo sonoro y puede adjudicar los canales restantes a componentes difusos del campo sonoro. En este ejemplo, cuando los SHC 27 se generaron a partir de un objeto de audio artificial, la unidad de análisis de características de contenido 152 puede adjudicar cinco de los canales a componentes coherentes del campo sonoro y puede adjudicar 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 152) 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.

[0084] La tasa de bits objetivo puede influir en el número de componentes y la tasa de bits de los motores de codificación AAC individuales (por ejemplo, los motores de codificación 160). En otras palabras, la unidad de análisis de características de contenido 152 puede realizar además la determinación de cuántos canales adjudicar a componentes coherentes y cuántos canales adjudicar a componentes difusos basado en una tasa de bits de salida del flujo de bits 31, por ejemplo, de 1,2 Mbps.

[0085] En algunos ejemplos, los canales adjudicados a los componentes coherentes del campo sonoro pueden tener mayores tasas de bits que los canales adjudicados a los componentes difusos del campo sonoro. Por ejemplo, una tasa de bits máxima del flujo de bits 31 puede ser de 1,2 Mb/s. En este ejemplo, puede haber cuatro canales adjudicados a componentes coherentes y 16 canales adjudicados a componentes difusos. Además, en este ejemplo, cada uno de los canales adjudicados a los componentes coherentes puede tener una tasa de bits máxima de 64 kb/s. En este ejemplo, cada uno de los canales adjudicados a los componentes difusos puede tener una velocidad de bits máxima de 48 kb/s.

[0086] Como se indica anteriormente, la unidad de análisis de características de contenido 152 puede determinar si los SHC 27 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 152 puede realizar esta determinación de diversas maneras. Por ejemplo, el dispositivo de generación de flujo de bits 36 puede usar SHC de 4.º orden. En este ejemplo, la unidad de análisis de características de contenido 152 puede codificar 24 canales y predecir un 25.º canal (que se puede representar como un vector). La unidad de análisis de características de contenido 152 puede aplicar escalares a, al menos, algunos de los 24 canales y añadir los valores resultantes para determinar el 25.º vector. Además, en este ejemplo, la unidad de análisis de características de contenido 152 puede determinar una exactitud del 25.º canal predicho. En este ejemplo, si la exactitud del 25.º canal predicho es relativamente alta (por ejemplo, la exactitud excede un umbral particular), es probable que los SHC 27 se generen a partir de un objeto de audio sintético. Por el contrario, si la exactitud del 25.º canal predicho es relativamente baja (por ejemplo, la exactitud está por debajo del umbral particular), es más probable que los SHC 27 representen un campo sonoro grabado. Por ejemplo, en este ejemplo, si una relación señal-ruido (SNR) del 25.º canal predicho es superior a 100 decibelios (db), es más probable que los SHC 27 representen un campo sonoro generado a partir de un objeto de audio sintético. Por el contrario, la SNR de un campo sonoro grabado usando un Eigenmike puede ser de 5 a 20 db. Por tanto, puede existir una demarcación aparente en las relaciones SNR entre el campo sonoro representado con los SHC 27 generados a partir de una grabación directa real y de un objeto de audio sintético.

[0087] Además, la unidad de análisis de características de contenido 152 puede seleccionar, basada al menos en parte en si los SHC 27 se generaron a partir de una grabación de un campo sonoro real o a partir de un objeto de audio artificial, libros de códigos para cuantificar el vector V. En otras palabras, la unidad de análisis de características de contenido 152 puede seleccionar diferentes libros de códigos para su uso al cuantificar el vector V, dependiendo de si el campo sonoro representado por los coeficientes HOA es grabado o sintético.

[0088] En algunos ejemplos, la unidad de análisis de características de contenido 152 puede determinar, en una base recurrente, si los SHC 27 se generaron a partir de una grabación de un campo sonoro real o a partir de un objeto de audio artificial. En algunos de dichos ejemplos, la base recurrente puede ser cada trama. En otros ejemplos, la unidad de análisis de características de contenido 152 puede realizar esta determinación una vez. Además, la unidad de análisis de características de contenido 152 puede determinar, en una base recurrente, el número total de canales y

la adjudicación de canales de componentes coherentes y canales de componentes difusos. En algunos de dichos ejemplos, la base recurrente puede ser cada trama. En otros ejemplos, la unidad de análisis de características de contenido 152 puede realizar esta determinación una vez. En algunos ejemplos, la unidad de análisis de características de contenido 152 puede seleccionar, en una base recurrente, libros de códigos para su uso en la cuantificación del vector V. En algunos de dichos ejemplos, la base recurrente puede ser cada trama. En otros ejemplos, la unidad de análisis de características de contenido 152 puede realizar esta determinación una vez.

[0089] La unidad de rotación 154 puede realizar una operación de rotación de los coeficientes HOA. Como se analiza en otra parte de la presente divulgación (por ejemplo, con respecto a la FIG. 5A y 5B), realizar la operación de rotación puede reducir el número de bits requeridos para representar los SHC 27. En algunos ejemplos, el análisis de rotación realizado por la unidad de rotación 152 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 transformada de Karhunen-Loeve (KLT) son técnicas relacionadas que pueden ser aplicables.

[0090] En este sentido, las técnicas pueden proporcionar un procedimiento de generación de un flujo de bits comprendido por una pluralidad de elementos jerárquicos que describen un campo sonoro, donde, en un primer ejemplo, el procedimiento comprende transformar la pluralidad de elementos jerárquicos representativos de un campo sonoro a desde un dominio de armónicos esféricos a otro dominio a fin de reducir un número de la pluralidad de elementos jerárquicos, y especificar información de transformación en el flujo de bits que describe cómo se transformó el campo sonoro.

[0091] En un segundo ejemplo, el procedimiento del primer ejemplo, en el que transformar la pluralidad de elementos jerárquicos comprende realizar una transformación basada en vectores con respecto a la pluralidad de elementos jerárquicos.

[0092] En un tercer ejemplo, el procedimiento del segundo ejemplo, en el que realizar la transformación basada en vectores comprende realizar una o más de una descomposición en valores singulares (SVD), un análisis de componentes principales (PCA) y una transformada de Karhunen-Loeve (KLT) con respecto a la pluralidad de elementos jerárquicos.

[0093] En un cuarto ejemplo, un dispositivo comprende uno o más procesadores configurados para transformar una pluralidad de elementos jerárquicos representativos de un campo sonoro desde un dominio de armónicos esféricos a otro dominio a fin de reducir un número de la pluralidad de elementos jerárquicos, y especificar la información de transformación en un flujo de bits que describe cómo se transformó el campo sonoro.

[0094] En un quinto ejemplo, el dispositivo del cuarto ejemplo, en el que el uno o más procesadores se configuran para, cuando se transforma la pluralidad de elementos jerárquicos, realizar una transformación basada en vectores con respecto a la pluralidad de elementos jerárquicos.

[0095] En un sexto ejemplo, el dispositivo del quinto ejemplo, en el que el uno o más procesadores se configuran para, cuando se realiza la transformación basada en vectores, realizar una o más de una descomposición en valores singulares (SVD), un análisis de componentes principales (PCA) y una transformada de Karhunen-Loeve (KLT) con respecto a la pluralidad de elementos jerárquicos.

[0096] En un séptimo ejemplo, un dispositivo comprende medios para transformar una pluralidad de elementos jerárquicos representativos de un campo sonoro desde un dominio de armónicos esféricos a otro dominio a fin de reducir un número de la pluralidad de elementos jerárquicos, y medios para especificar la información de transformación en un flujo de bits que describe cómo se transformó el campo sonoro.

[0097] En un octavo ejemplo, el dispositivo del séptimo ejemplo, en el que los medios para transformar la pluralidad de elementos jerárquicos comprenden medios para realizar una transformación basada en vectores con respecto a la pluralidad de elementos jerárquicos.

[0098] En un noveno ejemplo, el dispositivo del octavo ejemplo, en el que los medios para realizar la transformación basada en vectores comprenden medios para realizar una o más de una descomposición en valores singulares (SVD), un análisis de componentes principales (PCA) y una transformada de Karhunen-Loeve (KLT) con respecto a la pluralidad de elementos jerárquicos.

[0099] En un décimo ejemplo, un medio de almacenamiento legible por ordenador no transitorio ha almacenado en el mismo instrucciones que, cuando se ejecutan, hacen que uno o más procesadores transformen una pluralidad de elementos jerárquicos representativos de un campo sonoro desde un dominio de armónicos esféricos a otro dominio a fin de reducir un número de la pluralidad de elementos jerárquicos, y especificar la información de transformación en un flujo de bits que describe cómo se transformó el campo sonoro.

[0100] En un undécimo ejemplo, un procedimiento comprende analizar sintácticamente un flujo de bits para determinar la información de traslación que describe cómo se transformó una pluralidad de elementos jerárquicos que describen

un campo sonoro desde un dominio de armónicos esféricos a otro dominio para reducir un número de la pluralidad de elementos jerárquicos, y reconstruir, cuando se reproduce el campo sonoro basado en la pluralidad de elementos jerárquicos, la pluralidad de elementos jerárquicos basados en la pluralidad de elementos jerárquicos transformados.

[0101] En un duodécimo ejemplo, el procedimiento del undécimo ejemplo, en el que la información de transformación describe cómo se transformó la pluralidad de elementos jerárquicos usando descomposición basada en vectores para reducir el número de la pluralidad de elementos jerárquicos, y en el que transformar el campo sonoro comprende, cuando se reproduce el campo sonoro basado en la pluralidad de elementos jerárquicos, reconstruir la pluralidad de elementos jerárquicos basada en la pluralidad de elementos jerárquicos descompuestos basados en vectores.

[0102] En un decimotercero ejemplo, el procedimiento del duodécimo ejemplo, en el que la descomposición basada en vectores comprende una o más de una descomposición en valores singulares (SVD), un análisis de componentes principales (PCA) y una transformada de Karhunen-Loeve (KLT).

[0103] En un decimocuarto ejemplo, un dispositivo comprende uno o más procesadores configurados para analizar sintácticamente un flujo de bits para determinar la información de traslación que describe cómo se transformó una pluralidad de elementos jerárquicos que describen un campo sonoro desde un dominio de armónicos esféricos a otro dominio para reducir un número de la pluralidad de elementos jerárquicos, y reconstruir, cuando se reproduce el campo sonoro basado en la pluralidad de elementos jerárquicos, la pluralidad de elementos jerárquicos basada en la pluralidad de elementos jerárquicos transformados.

[0104] En un decimoquinto ejemplo, el dispositivo del decimocuarto ejemplo, en el que la información de transformación describe cómo se transformó la pluralidad de elementos jerárquicos usando descomposición basada en vectores para reducir el número de la pluralidad de elementos jerárquicos, y en el que el uno o más procesadores se configuran para, cuando transforman el campo sonoro, reconstruir, cuando se reproduce el campo sonoro basado en la pluralidad de elementos jerárquicos, reconstruyendo la pluralidad de elementos jerárquicos basada en la pluralidad de elementos jerárquicos descompuestos basados en vectores.

[0105] En un decimosexto ejemplo, el dispositivo del decimoquinto ejemplo, en el que la descomposición basada en vectores comprende una o más de una descomposición en valores singulares (SVD), un análisis de componentes principales (PCA) y una transformada de Karhunen-Loeve (KLT).

[0106] En un decimoséptimo ejemplo, un dispositivo comprende medios para analizar sintácticamente un flujo de bits para determinar la información de traslación que describe cómo se transformó una pluralidad de elementos jerárquicos que describen un campo sonoro desde un dominio de armónicos esféricos a otro dominio para reducir un número de la pluralidad de los elementos jerárquicos, y medios para reconstruir, cuando se reproduce el campo sonoro basado en la pluralidad de elementos jerárquicos, la pluralidad de elementos jerárquicos basada en la pluralidad de elementos jerárquicos transformados.

[0107] En un decimoctavo ejemplo, el dispositivo del decimoséptimo ejemplo, en el que la información de transformación describe cómo se transformó la pluralidad de elementos jerárquicos usando descomposición basada en vectores para reducir el número de la pluralidad de elementos jerárquicos, y en el que los medios para transformar el campo sonoro comprenden medios para reconstruir, cuando se reproduce el campo sonoro basado en la pluralidad de elementos jerárquicos, la pluralidad de elementos jerárquicos basada en la pluralidad de elementos jerárquicos descompuestos basados en vectores.

[0108] En un decimonoveno ejemplo, el dispositivo del decimoctavo ejemplo, en el que la descomposición basada en vectores comprende una o más de una descomposición en valores singulares (SVD), un análisis de componentes principales (PCA) y una transformada de Karhunen-Loeve (KLT).

[0109] En un vigésimo ejemplo, un medio de almacenamiento legible por ordenador no transitorio que tiene almacenado en el mismo instrucciones que, cuando se ejecutan, hacen que uno o más procesadores analicen sintácticamente un flujo de bits para determinar la información de traslación que describe cómo una pluralidad de elementos jerárquicos que describen un campo sonoro se transformó desde un dominio de armónicos esféricos a otro dominio para reducir un número de la pluralidad de elementos jerárquicos, y reconstruir, cuando se reproduce el campo sonoro basado en la pluralidad de elementos jerárquicos, la pluralidad de elementos jerárquicos basada en la pluralidad de elementos jerárquicos transformados.

[0110] En el ejemplo de la FIG. 4B, la unidad de componentes coherentes de extracción 156 recibe SHC girados 27 desde la unidad de rotación 154. Además, la unidad de componentes coherentes de extracción 156 extrae, a partir de los SHC girados 27, los de los SHC girados 27 asociados con los componentes coherentes del campo sonoro.

[0111] Además, la unidad de componentes coherentes de extracción 156 genera uno o más canales de componentes coherentes. Cada uno de los canales de componentes coherentes puede incluir un subconjunto diferente de los SHC girados 27 asociados con los coeficientes coherentes del campo sonoro. En el ejemplo de la FIG. 4B, la unidad de componentes coherentes de extracción 156 puede generar desde uno a 16 canales de componentes coherentes. El

número de canales de componentes coherentes generados por la unidad de componentes coherentes de extracción 156 se puede determinar por el número de canales adjudicados por la unidad de análisis de características de contenido 152 a los componentes coherentes del campo sonoro. Las tasas de bits de los canales de componentes coherentes generados por la unidad de componentes coherentes de extracción 156 se pueden determinar mediante la unidad de análisis de características de contenido 152.

[0112] De forma similar, en el ejemplo de la FIG. 4B, la unidad de componentes difusos de extracción 158 recibe SHC girados 27 desde la unidad de rotación 154. Además, la unidad de componentes difusos de extracción 158 extrae, a partir de los SHC girados 27, los de los SHC girados 27 asociados con los componentes difusos del campo sonoro.

[0113] Además, la unidad de componentes difusos de extracción 158 genera uno o más canales de componentes difusos. Cada uno de los canales de componentes difusos puede incluir un subconjunto diferente de los SHC girados 27 asociados con los coeficientes difusos del campo sonoro. En el ejemplo de la FIG. 4B, la unidad de componentes difusos de extracción 158 puede generar desde uno a 9 canales de componentes difusos. El número de canales de componentes difusos generados mediante la unidad de componentes difusos de extracción 158 se puede determinar por el número de canales adjudicados por la unidad de análisis de características de contenido 152 a los componentes difusos del campo sonoro. Las tasas de bits de los canales de componentes difusos generados por la unidad de componentes difusos de extracción 158 pueden ser las determinadas por la unidad de análisis de características de contenido 152.

[0114] En el ejemplo de la FIG. 4B, el motor de codificación 160 puede funcionar como se describe anteriormente con respecto al ejemplo de la FIG. 4A, solo que esta vez con respecto a los componentes difusos y coherentes. El multiplexor 164 ("MUX 164") 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 150), para generar el flujo de bits 31.

[0115] Las FIGS. 5A y 5B son diagramas que ilustran un ejemplo de realización de diversos aspectos de las técnicas descritas en la presente divulgación para girar un campo sonoro 40. La FIG. 5A es un diagrama que ilustra el campo sonoro 40 antes de la rotación de acuerdo con los diversos aspectos de las técnicas descritas en la presente divulgación. En el ejemplo de la FIG. 5A, el campo sonoro 40 incluye dos ubicaciones de alta presión, indicadas como ubicación 42A y 42B. Estas ubicaciones 42A y 42B ("ubicaciones 42") residen a lo largo de una línea 44 que tiene una pendiente no infinita (que es otra forma de denominar una línea que no es vertical, ya que las líneas verticales tienen una pendiente infinita). Dado que las ubicaciones 42 tienen una coordenada z, además de las coordenadas x e y, se pueden requerir funciones de base esférica de orden superior para representar correctamente este campo sonoro 40 (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 40 directamente a los SHC 27, el dispositivo de generación de flujo de bits 36 puede girar el campo sonoro 40 hasta que la línea 44 que conecta las ubicaciones 42 sea vertical.

[0116] La FIG. 5B es un diagrama que ilustra el campo sonoro 40 después de girarse hasta que la línea 44 que conecta las ubicaciones 42 es vertical. Como resultado de la rotación del campo sonoro 40 de esta manera, los SHC 27 se pueden derivar de modo que los de suborden distinto de cero de SHC 27 se especifiquen como ceros, dado que el campo sonoro girado 40 ya no tiene ninguna ubicación de presión (o energía) a lo largo del eje no vertical (por ejemplo, el eje X y/o el eje Y). De esta manera, el dispositivo de generación de flujo de bits 36 puede girar, transformar o, más en general, ajustar el campo sonoro 40 para reducir el número de SHC girados 27 que tienen valores distintos de cero. A continuación, el dispositivo de generación de flujo de bits 36 puede adjudicar tasas de bits más bajas a los de suborden distintos de cero de los SHC girados 27 en relación con los de suborden cero de los SHC girados 27, como se describe anteriormente. El dispositivo de generación de flujo de bits 36 también puede especificar información de rotación en el flujo de bits 31 que indica cómo se giró el campo sonoro 40, a menudo mediante la expresión de un acimut y una elevación de la manera descrita anteriormente.

[0117] De forma alternativa o adicionalmente, el dispositivo de generación de flujo de bits 36 puede a continuación, en lugar de señalar un número con signo de 32 bits que identifica que estos de orden superior de SHC 27 tienen valores de cero, señalar en un campo del flujo de bits 31 que estos de orden superior de SHC 27 no están señalizados. El dispositivo de extracción 38, en estos casos, puede implicar que estos no señalizados de los SHC girados 27 tienen un valor de cero y, cuando se reproduce el campo sonoro 40 basado en SHC 27, realizar la rotación para girar el campo sonoro 40 de manera que el campo sonoro 40 se asemeje al campo sonoro 40 mostrado en el ejemplo de la FIG. 5A. De esta manera, el dispositivo de generación de flujo de bits 36 puede reducir el número de SHC 27 que se requiere que se especifiquen en el flujo de bits 31 o de otro modo reducir la tasa de bits asociada con los de suborden distinto de cero de los SHC girados 27.

[0118] Se puede usar un algoritmo de 'compactación espacial' para determinar la rotación óptima del campo sonoro. En un modo de realización, el dispositivo de generación de flujo de bits 36 puede realizar el algoritmo para iterar a través de la totalidad de las posibles combinaciones de acimut y elevación (es decir, 1024 x 512 combinaciones en el ejemplo anterior), girando el campo sonoro para cada combinación, y calculando el número de SHC 27 que están por encima del valor umbral. La combinación candidata de acimut/elevación que produce el menor número de SHC 27 por encima del valor umbral se puede considerar que es lo que se puede denominar la "rotación óptima". En esta forma

girada, el campo sonoro puede requerir el menor número de SHC 27 para representar el campo sonoro y a continuación se puede considerar 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).

[0119] En algunos casos, en lugar de solo especificar el ángulo de acimut y el ángulo de elevación, el dispositivo de generación de flujo de bits 36 puede especificar ángulos adicionales en forma de ángulos de Euler, como un ejemplo. Los ángulos de Euler especifican el ángulo de rotación en torno al eje Z, el eje X anterior y el eje Z anterior. Si bien se describen en la presente divulgación con respecto a combinaciones de ángulos de acimut y de elevación, las técnicas de la presente divulgación no se deben limitar a especificar solo los ángulos de acimut y de elevación, sino que pueden incluir especificar cualquier número de ángulos, incluyendo los tres ángulos de Euler indicados anteriormente. En este sentido, el dispositivo de generación de flujo de bits 36 puede girar el campo sonoro para reducir un número de la pluralidad de elementos jerárquicos que proporcionan información pertinente al describir el campo sonoro y especificar los ángulos de Euler como información de rotación en el flujo de bits. Los ángulos de Euler, como se indica anteriormente, pueden describir cómo se giró el campo sonoro. Cuando se usan los ángulos de Euler, el dispositivo de extracción de flujo de bits 38 puede analizar sintácticamente el flujo de bits para determinar la información de rotación que incluye los ángulos de Euler y, cuando se reproduce el campo sonoro basado en los de la pluralidad de elementos jerárquicos que proporcionan información pertinente al describir el campo sonoro, girar el campo sonoro basado en los ángulos de Euler.

[0120] Asimismo, en algunos casos, en lugar de especificar explícitamente estos ángulos en el flujo de bits 31, el dispositivo de generación de flujo de bits 36 puede especificar un índice (que se puede denominar "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 usar en relación con una tabla de rotación. Es decir, el dispositivo de generación de flujo de bits 36 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.

[0121] De forma alternativa, la tabla de rotación puede incluir una entrada para cada transformada de matriz representativa de cada combinación del ángulo de acimut y el ángulo de elevación. Es decir, el dispositivo de generación de flujo de bits 36 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 de elevación. Típicamente, el dispositivo de generación de flujo de bits 36 recibe los SHC 27 y deriva los SHC 27', cuando se realiza la rotación, de acuerdo con la siguiente ecuación:

$$\begin{bmatrix} SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (25 \times 32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32 \times 25) \end{bmatrix} \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

En la ecuación anterior, los SHC 27' se computan como una función de una matriz de codificación para codificar un campo sonoro en términos de una segunda trama de referencia ($EncMat_2$), una matriz de inversión para revertir los SHC 27 de vuelta a un campo sonoro en términos de una primera trama de referencia ($InvMat_1$) y SHC 27. $EncMat_2$ tiene un tamaño de 25×32 , mientras que $InvMat_1$ tiene un tamaño de 32×25 . Tanto SHC 27' como SHC 27 tienen un tamaño de 25, donde SHC 27' se pueden reducir adicionalmente debido a la eliminación de los 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 estática con respecto a cada combinación de ángulo de acimut y de elevación. La tabla de rotación puede incluir una entrada que almacena el resultado de multiplicar cada $EncMat_2$ diferente a $InvMat_1$.

[0122] La FIG. 6 es un diagrama que ilustra un campo sonoro de ejemplo capturado de acuerdo con una primera trama de referencia que a continuación se gira de acuerdo con las técnicas descritas en la presente divulgación para expresar el campo sonoro en términos de una segunda trama de referencia. En el ejemplo de la FIG. 6, el campo sonoro que rodea a un micrófono Eigen 46 se captura suponiendo una primera trama de referencia, que se indica por los ejes X_1 , Y_1 y Z_1 en el ejemplo de la FIG. 6. Los SHC 27 describen el campo sonoro en términos de esta primera trama de referencia. La $InvMat_1$ transforma los SHC 27 de vuelta al campo sonoro, haciendo posible que el campo sonoro se gire a la segunda trama de referencia indicada por los ejes X_2 , Y_2 y Z_2 en el ejemplo de la FIG. 6. La $EncMat_2$ descrita anteriormente puede girar el campo sonoro y generar SHC 27' que describe este campo sonoro girado en términos de la segunda trama de referencia.

[0123] En cualquier caso, la ecuación anterior se puede derivar como sigue. Dado que el campo sonoro se graba con un determinado sistema de coordenadas, de modo que el frente se considera la dirección del eje X, las 32 posiciones de micrófono de un Eigenmike (u otras configuraciones de micrófono) se definen a partir de este sistema de coordenadas de referencia. La rotación del campo sonoro se puede considerar entonces como una rotación de esta trama de referencia. Para la supuesta trama de referencia, los SHC 27 se pueden calcular como sigue:

$$\begin{bmatrix} SHC \\ 27 \end{bmatrix} = \begin{bmatrix} Y_0^0(Pos_1) & Y_0^0(Pos_2) & \cdots & Y_0^0(Pos_{32}) \\ Y_1^{-1}(Pos_1) & \cdots & & Y_1^{-1}(Pos_{32}) \\ \vdots & \ddots & & \\ Y_4^4(Pos_1) & & & Y_4^4(Pos_{32}) \end{bmatrix} \begin{bmatrix} mic_1(t) \\ mic_2(t) \\ \vdots \\ mic_{32}(t) \end{bmatrix}$$

En la ecuación anterior, la Y_n^m representa 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 del 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 anterior a la rotación en este ejemplo).

[0124] La ecuación anterior se puede expresar de forma alternativa en términos de las expresiones matemáticas indicadas anteriormente como:

$$[SHC_27] = [E_s(\theta, \phi)] [m_i(t)].$$

[0125] 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 las señales de micrófono originales estén presentes, el campo sonoro se puede girar 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 ($mic_i(t)$) a partir de los SHC 27. Si se usa un diseño en T (como en un Eigenmike de 32 micrófonos), la solución a este problema se puede lograr resolviendo la siguiente ecuación:

$$\begin{bmatrix} mic_1(t) \\ mic_2(t) \\ \vdots \\ mic_{32}(t) \end{bmatrix} = [InvMat_1] \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

Este $InvMat_1$ puede especificar las funciones de base armónica esférica computadas de acuerdo con la posición de los micrófonos como se especifica en relación con la primera trama de referencia. Esta ecuación también se puede expresar como $[m_i(t)] = [E_s(\theta, \phi)]^{-1} [SHC]$, como se indica anteriormente.

[0126] Aunque se denominan anteriormente "señales de micrófono", las señales de micrófono se pueden referir a una representación de dominio espacial que usa el diseño en T de la posición de cápsula de 32 micrófonos en lugar de las "señales de micrófono" *per se*. Asimismo, si bien se describe con respecto a posiciones de cápsula de 32 micrófonos, las técnicas se pueden realizar con respecto a cualquier número de posiciones de cápsula de micrófonos, incluyendo 16, 64 o cualquier otro número (incluyendo aquellos que no son un factor de dos).

[0127] Una vez que las señales del micrófono ($mic_i(t)$) se recuperan de acuerdo con la ecuación anterior, las señales de micrófono ($mic_i(t)$) que describen el campo sonoro se pueden girar para computar los SHC 27' correspondientes a la segunda trama de referencia, dando como resultado la siguiente ecuación:

$$\begin{bmatrix} SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (25 \times 32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32 \times 25) \end{bmatrix} \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

[0128] El $EncMat_2$ especifica las funciones de base armónica esférica desde una posición girada (Pos_i'). De esta manera, el $EncMat_2$ puede especificar eficazmente una combinación del ángulo de acimut y de elevación. Por tanto,

cuando la tabla de rotación almacena el resultado de $\begin{bmatrix} EncMat_2 \\ (25 \times 32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32 \times 25) \end{bmatrix}$ para cada combinación de los ángulos de acimut y de elevación, la tabla de rotación especifica eficazmente cada combinación de los ángulos de acimut y de elevación. La ecuación anterior se puede expresar también como:

$$[SHC\ 27'] = [E_s(\theta_2, \phi_2)] [E_s(\theta_1, \phi_1)]^{-1} [SHC\ 27],$$

donde θ_2, ϕ_2 representan un segundo ángulo de acimut y un segundo ángulo de elevación diferentes del primer ángulo de acimut y ángulo de elevación representados por θ_1, ϕ_1 . El θ_1, ϕ_1 corresponden a la primera trama de referencia mientras que el θ_2, ϕ_2 corresponden a la segunda trama de referencia. Por lo tanto, el $InvMat_1$ puede corresponder a $[E_s(\theta_1, \phi_1)]^{-1}$, mientras que el $EncMat_2$ puede corresponder a $[E_s(\theta_2, \phi_2)]$.

[0129] Lo anterior puede representar una versión más simplificada del cómputo que no considera la operación de filtrado, representada anteriormente en diversas ecuaciones que denotan la derivación de SHC 27 en el dominio de frecuencia por la función $j_n(\cdot)$, que se refiere a la función de Bessel esférica de orden n . En el dominio de tiempo, esta función $j_n(\cdot)$ representa una operación de filtrado que es específica para un orden particular, n . Con el filtrado, la rotación se puede realizar por orden. Para ilustrar, considérense las siguientes ecuaciones:

$$a_n^k(t) \square b_n(t) * \left[Y_n^m \right] [m_i(t)]$$

$$a_n^k(t) \square \left[Y_n^m \right] [b_n(t) * m_i(t)]$$

[0130] Si bien se describe con respecto a dichas operaciones de filtrado, en diversos ejemplos, las técnicas se pueden realizar sin estas operaciones de filtrado. En otras palabras, se pueden realizar diversas formas de rotación sin realizar o de otro modo aplicar las operaciones de filtrado a los SHC 27, como se indica anteriormente. Debido a que diferentes 'n' SHC no interactúan entre sí en esta operación, no se requieren filtros dado que los filtros solo dependen de 'n' y no de 'm'. Por ejemplo, se puede aplicar una matriz-d de Wigner a los SHC 27 para realizar la rotación, donde la aplicación de esta matriz-d de Wigner puede no requerir la aplicación de las operaciones de filtrado. Como resultado de no transformar los SHC 27 de vuelta a señales de micrófono, se pueden requerir las operaciones de filtrado en esta transformada. Asimismo, considerando que 'n' solo entra en 'n', la rotación se hace en bloques de $2m+1$ de los SHC 27 y el resto pueden ser ceros. Para una adjudicación de memoria más eficaz (posiblemente en el software), la rotación se puede hacer por orden como se describe en la presente divulgación. Además, debido a que solo existe un SHC 27 en $n=0$, siempre es el mismo. Diversas implementaciones de las técnicas pueden hacer uso de este único de los SHC 27 en $n=0$ para proporcionar eficacia (en términos de cálculos y/o consumo de memoria).

[0131] A partir de estas ecuaciones, los SHC 27' girados para órdenes se hacen por separado ya que los $b_n(t)$ son diferentes para cada orden. Como resultado, la ecuación anterior se puede modificar como sigue para computar los de primer orden de los SHC 27' girados:

$$\begin{bmatrix} \text{SHC} \\ 27' \\ \text{de 1.}^{\text{er}} \\ \text{orden} \end{bmatrix} = \begin{bmatrix} \text{EncMat}_2 \\ (3 \times 32) \end{bmatrix} \begin{bmatrix} \text{InvMat}_1 \\ (32 \times 3) \end{bmatrix} \begin{bmatrix} \text{SHC} \\ 27 \\ \text{de 1.}^{\text{er}} \\ \text{orden} \end{bmatrix}$$

Dado que existen tres de primer orden de SHC 27, cada uno de los vectores SHC 27' y 27 tiene un tamaño de tres en la ecuación anterior. Igualmente, para el segundo orden se puede aplicar la siguiente ecuación:

$$\begin{bmatrix} \text{SHC} \\ 27' \\ \text{de 2.}^{\text{o}} \\ \text{orden} \end{bmatrix} = \begin{bmatrix} \text{EncMat}_2 \\ (5 \times 32) \end{bmatrix} \begin{bmatrix} \text{InvMat}_1 \\ (32 \times 5) \end{bmatrix} \begin{bmatrix} \text{SHC} \\ 27 \\ \text{de 2.}^{\text{o}} \\ \text{orden} \end{bmatrix}$$

De nuevo, dado que existen cinco de segundo orden de SHC 27, cada uno de los vectores SHC 27' y 27 tiene un tamaño de cinco en la ecuación anterior. Las ecuaciones restantes para los demás órdenes, es decir, el tercer y cuarto órdenes, pueden ser similares a las descritas anteriormente, siguiendo el mismo patrón con respecto a los tamaños de las matrices (ya que el número de filas de EncMat_2 , el número de columnas de InvMat_1 y los tamaños de los vectores SHC 27 y SHC 27' de tercer y cuarto orden son iguales al número de subórdenes (m por dos más 1) de cada una de las funciones de base armónica esférica de tercer y cuarto orden. Aunque se describe como que es una representación de cuarto orden, las técnicas se pueden aplicar a cualquier orden y no se deben limitar al cuarto orden.

[0132] El dispositivo de generación de flujo de bits 36 por lo tanto 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 generación de flujo de bits 36 puede, después de realizar esta operación de rotación, computar el número de SHC 27' por encima del valor umbral. En algunos casos, el dispositivo de generación de flujo de bits 36 puede realizar esta rotación para derivar una serie de SHC 27' que representan el campo sonoro durante una duración de tiempo, tal como una trama de audio. Al realizar esta rotación para derivar la serie de los SHC 27' que representan el campo sonoro durante esta duración de tiempo, el dispositivo de generación de flujo de bits 36 puede reducir el número de operaciones de rotación que se tienen que realizar en comparación con hacer esto para cada conjunto de los SHC 27 que describen el campo sonoro para duraciones de tiempo menores que una trama u otra longitud. En cualquier caso, el dispositivo de generación de flujo de bits 36 puede ahorrar, a lo largo de este procedimiento, los SHC 27' que tienen el menor número de los SHC 27' mayores que el valor umbral.

[0133] Sin embargo, la realización de esta operación de rotación con respecto a cada combinación de ángulo de acimut y de elevación puede ser intensiva en el procesador o requerir mucho tiempo. Como resultado, el dispositivo de generación de flujo de bits 36 puede no realizar lo que se puede caracterizar como esta implementación de "fuerza bruta" del algoritmo de rotación. En cambio, el dispositivo de generación de flujo de bits 36 puede realizar rotaciones con respecto a un subconjunto de combinaciones posiblemente conocidas (estadísticamente) de ángulo de acimut y de elevación que ofrecen una compactación en general buena, realizando rotaciones adicionales con respecto a combinaciones alrededor de las de este subconjunto que proporcionan mejor compactación en comparación con otras combinaciones en el subconjunto.

[0134] Como otra alternativa, el dispositivo de generación de flujo de bits 36 puede realizar esta rotación con respecto a sólo el subconjunto de combinaciones conocido. Como otra alternativa, el dispositivo de generación de flujo de bits 36 puede seguir una trayectoria (espacialmente) de combinaciones, realizando las rotaciones con respecto a esta trayectoria de combinaciones. Como otra alternativa, el dispositivo de generación de flujo de bits 36 puede especificar un umbral de compactación que define un número máximo de SHC 27' que tienen valores distintos de cero por encima del valor umbral. Este umbral de compactación puede establecer eficazmente un punto de parada para la búsqueda, de modo que, cuando el dispositivo de generación de flujo de bits 36 realiza una rotación y determina que el número de SHC 27' que tienen 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 generación de flujo de bits 36 deja de realizar cualquier operación de rotación adicional con respecto a las combinaciones restantes. Aún como otra alternativa, el dispositivo de generación de flujo de bits 36 puede recorrer 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 27' que tienen un valor distinto de cero mayor que el valor umbral.

[0135] 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 27' con un valor distinto de cero mayor que el valor umbral. En consecuencia, el dispositivo de generación de flujo de bits 36 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 pertinente al describir el campo sonoro. El dispositivo de generación de flujo de bits 36 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 pertinente al describir el campo sonoro. Además, el dispositivo de generación de flujo de bits 36 puede seleccionar la primera operación de rotación o la segunda operación de rotación basada en una comparación del primer número de la pluralidad de elementos jerárquicos y el segundo número de la pluralidad de elementos jerárquicos.

[0136] 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 basado en la información de rotación pasada determinada cuando se gira el campo sonoro durante una duración de tiempo previa. Por ejemplo, el dispositivo de generación de flujo de bits 36 puede girar el campo sonoro durante una primera duración de tiempo, por ejemplo, una trama de audio, para identificar SHC 27' durante esta primera duración de tiempo. El dispositivo de generación de flujo de bits 36 puede especificar la información de rotación y los SHC 27' en el flujo de bits 31 de cualquiera de las maneras descritas anteriormente. Esta información de rotación se puede denominar primera información de rotación ya que describe la rotación del campo sonoro durante la primera duración de tiempo. A continuación, el dispositivo de generación de flujo de bits 31 puede, basado en esta primera información de rotación, girar el campo sonoro durante una segunda duración de tiempo, por ejemplo, una segunda trama de audio, para identificar SHC 27' durante esta segunda duración de tiempo. El dispositivo de generación de flujo de bits 36 puede utilizar esta primera información de rotación cuando realiza la segunda operación de rotación durante la segunda duración de tiempo para inicializar una búsqueda de la combinación "óptima" de ángulos de acimut y de elevación, como un ejemplo. A continuación, el dispositivo de generación de flujo de bits 36 puede especificar los SHC 27' 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 31.

[0137] Si bien se describen anteriormente con respecto a un número de maneras diferentes mediante las que implementar el algoritmo de rotación para reducir el tiempo de procesamiento y/o el consumo, las técnicas se pueden realizar con respecto a cualquier algoritmo que pueda reducir o de otro modo acelerar la identificación de lo que se puede denominar la "rotación óptima". Asimismo, las técnicas se pueden realizar 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.

[0138] Las FIGS. 7A-7E son, cada uno, un diagrama que ilustra los flujos de bits 31A-31E formados de acuerdo con las técnicas descritas en la presente divulgación. En el ejemplo de la FIG. 7A, el flujo de bits 31A puede representar un ejemplo del flujo de bits 31 mostrado en la FIG. 3 anterior. El flujo de bits 31A incluye un campo presente de SHC 50 y un campo que almacena SHC 27' (donde el campo se indica como "SHC 27'"). El campo presente de SHC 50 puede incluir un bit correspondiente a cada uno de los SHC 27. Los SHC 27' pueden representar los de SHC 27 que se especifican en el flujo de bits, que pueden ser menores en número que el número de los SHC 27. Típicamente, cada uno de los SHC 27' son los de SHC 27 que tienen valores distintos de cero. Como se indica 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, que se pueden adjudicar para expresar otras partes del campo sonoro con más detalle o eliminarse de otro modo para facilitar la utilización del ancho de banda eficaz.

[0139] En el ejemplo de la FIG. 7B, el flujo de bits 31B puede representar un ejemplo del flujo de bits 31 mostrado en la FIG. 3 anteriormente. El flujo de bits 31B incluye un campo de información de transformación 52 ("información de transformación 52") y un campo que almacena SHC 27' (donde el campo se indica como "SHC 27'"). La información de transformación 52, como se indica anteriormente, puede comprender información de transformación, 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 52 también puede especificar un orden más alto de SHC 27 que se especifican en el flujo de bits 31B como SHC 27'. Es decir, la información de transformación 52 puede indicar un orden de tres, que el dispositivo de extracción 38 puede entender como que indica que el SHC 27' incluye los de SHC 27 hasta e incluyendo los de SHC 27 que tienen un orden de tres. A continuación, el dispositivo de extracción 38 se puede configurar para establecer que SHC 27 tenga un orden de cuatro o superior a cero, eliminando de este modo potencialmente la señalización explícita de SHC 27 de orden cuatro o superior en el flujo de bits.

[0140] En el ejemplo de la FIG. 7C, el flujo de bits 31C puede representar un ejemplo del flujo de bits 31 mostrado en la FIG. 3 anteriormente. El flujo de bits 31C incluye el campo de información de transformación 52 ("información de transformación 52"), el campo presente de SHC 50 y un campo que almacena SHC 27' (donde el campo se indica como "SHC 27'"). En lugar de estar configurado para comprender qué orden de SHC 27 no se señalizan como se describe anteriormente con respecto a la FIG. 7B, el campo presente de SHC 50 puede señalar explícitamente cuáles de los SHC 27 se especifican en el flujo de bits 31C como SHC 27'.

[0141] En el ejemplo de la FIG. 7D, el flujo de bits 31D puede representar un ejemplo del flujo de bits 31 mostrado en la FIG. 3 anteriormente. El flujo de bits 31D incluye un campo de orden 60 ("orden 60"), el campo presente de SHC 50, un indicador de acimut 62 ("AZF 62"), un indicador de elevación 64 ("ELF 64"), un campo de ángulo de acimut 66 ("acimut 66"), un campo de ángulo de elevación 68 ("elevación 68") y un campo que almacena SHC 27' (donde, de nuevo, el campo se indica como "SHC 27'"). El campo de orden 60 especifica el orden de SHC 27', es decir, el orden indicado por n anteriormente para el orden más alto de la función de base esférica usada para representar el campo sonoro. El campo de orden 60 se muestra como que es un campo de 8 bits, pero puede ser de otros tamaños de bits diversos, tal como tres (que es el número de bits requerido para especificar el cuarto orden). El campo presente de SHC 50 se muestra como un campo de 25 bits. De nuevo, sin embargo, el campo presente de SHC 50 puede ser de otros tamaños de bits diversos. El campo presente de SHC 50 se muestra como de 25 bits para indicar que el campo presente de SHC 50 puede incluir un bit para cada uno de los coeficientes armónicos esféricos correspondientes a una representación de cuarto orden del campo sonoro.

[0142] El indicador de acimut 62 representa un indicador de un bit que especifica si el campo de acimut 66 está presente en el flujo de bits 31D. Cuando el indicador de acimut 62 se establece en uno, el campo de acimut 66 para SHC 27' está presente en el flujo de bits 31D. Cuando el indicador de acimut 62 se establece en cero, el campo de acimut 66 para SHC 27' no está presente o especificado de otro modo en el flujo de bits 31D. Igualmente, el indicador de elevación 64 representa un indicador de un bit que especifica si el campo de elevación 68 está presente en el flujo de bits 31D. Cuando el indicador de elevación 64 se establece en uno, el campo de elevación 68 para SHC 27' está presente en el flujo de bits 31D. Cuando el indicador de elevación 64 se establece en cero, el campo de elevación 68 para SHC 27' no está presente o especificado de otro modo en el flujo de bits 31D. Si bien se describe como que uno señala que el campo correspondiente está presente y que cero señala que el campo correspondiente no está presente, la convención se puede invertir de modo que un cero especifica que el campo correspondiente se especifica en el flujo de bits 31D y un uno especifica que el campo correspondiente no se especifica en el flujo de bits 31D. Por lo tanto, las técnicas descritas en la presente divulgación no se deben limitar en este sentido.

[0143] El campo de acimut 66 representa un campo de 10 bits que especifica, cuando está presente en el flujo de bits 31D, el ángulo de acimut. Si bien se muestra como un campo de 10 bits, el campo de acimut 66 puede tener otros tamaños de bits. El campo de elevación 68 representa un campo de 9 bits que especifica, cuando está presente en el flujo de bits 31D, el ángulo de elevación. El ángulo de acimut y el ángulo de elevación especificados en los campos 66 y 68, respectivamente, pueden, junto con los indicadores 62 y 64, representar la información de rotación descrita anteriormente. Esta información de rotación se puede usar para girar el campo sonoro a fin de recuperar SHC 27 en la trama de referencia original.

[0144] El campo de SHC 27' se muestra como un campo variable que tiene un tamaño de X. El campo de SHC 27' puede variar debido al número de SHC 27' especificado en el flujo de bits como se indica por el campo presente de SHC 50. El tamaño de X se puede derivar en función del número de unidades en el campo presente de SHC 50 por 32 bits (que es el tamaño de cada SHC 27').

[0145] En el ejemplo de la FIG. 7E, el flujo de bits 31E puede representar otro ejemplo del flujo de bits 31 mostrado en la FIG. 3 anteriormente. El flujo de bits 31E incluye un campo de orden 60 ("orden 60"), un campo presente de SHC 50 y un campo de índice de rotación 70, y un campo que almacena SHC 27' (donde, de nuevo, el campo se indica como "SHC 27"). El campo de orden 60, el campo presente de SHC 50 y el campo de SHC 27' pueden ser sustancialmente similares a los descritos anteriormente. El campo de índice de rotación 70 puede representar un campo de 20 bits usado para especificar una de las combinaciones de 1024x512 (o, en otras palabras, 524288) de los ángulos de elevación y de acimut. En algunos casos, sólo se pueden usar 19 bits para especificar este campo de índice de rotación 70, y el dispositivo de generación de flujo de bits 36 puede especificar un indicador adicional en el flujo de bits para indicar si se realizó una operación de rotación (y, por lo tanto, si el campo de índice de rotación 70 está presente en el flujo de bits). Este campo de índice de rotación 70 especifica el índice de rotación indicado anteriormente, que se puede referir a una entrada en una tabla de rotación común a tanto el dispositivo de generación de flujo de bits 36 como al dispositivo de extracción de flujo de bits 38. Esta tabla de rotación puede, en algunos casos, almacenar las diferentes combinaciones de ángulos de acimut y de elevación. De forma alternativa, la tabla de rotación puede almacenar la matriz descrita anteriormente, que almacena eficazmente las diferentes combinaciones de los ángulos de acimut y de elevación en forma de matriz.

[0146] La FIG. 8 es un diagrama de flujo que ilustra el funcionamiento de ejemplo del dispositivo de generación de flujo de bits 36 mostrado en el ejemplo de la FIG. 3 al implementar los aspectos de rotación de las técnicas descritas en la presente divulgación. Inicialmente, el dispositivo de generación de flujo de bits 36 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 (80). A continuación, el dispositivo de generación de flujo de bits 36 puede girar el campo sonoro de acuerdo con el ángulo de acimut y de elevación seleccionados (82). Como se describe anteriormente, el dispositivo de generación de flujo de bits 36 en primer lugar puede derivar el campo sonoro a partir de SHC 27 usando la $InvMat_1$ indicada anteriormente. El dispositivo de generación de flujo de bits 36 también puede determinar SHC 27' que representan el campo sonoro girado (84). Si bien se describe como que son etapas u operaciones separadas, el dispositivo de generación de flujo de bits 36 puede aplicar una transformada (que puede representar el resultado de $[EncMat_2][InvMat_1]$) que representa la selección de la combinación del ángulo de acimut y del ángulo de elevación, derivando la campo sonoro a partir de los SHC 27, girando el campo sonoro y determinando los SHC 27' que representan el campo sonoro girado.

[0147] En cualquier caso, el dispositivo de generación de flujo de bits 36 puede a continuación computar un número de los SHC 27' determinados que sean mayores que un valor umbral, comparando este número con un número computado para una iteración previa con respecto a una combinación de ángulo de acimut y ángulo de elevación (86, 88) previa. 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 previo predefinido (que se puede establecer en cero). En cualquier caso, si el número determinado de los SHC 27' es menor que el número previo ("Sí" 88), el dispositivo de generación de flujo de bits 36 almacena los SHC 27', el ángulo de acimut y el ángulo de elevación, reemplazando a menudo los SHC 27', el ángulo de acimut y el ángulo de elevación previos almacenados a partir de una iteración previa del algoritmo de rotación (90).

[0148] Si el número determinado de los SHC 27' no es menor que el número previo ("NO" 88) o después de almacenar los SHC 27', el ángulo de acimut y el ángulo de elevación en lugar de los SHC 27', el ángulo de acimut y el ángulo de elevación previamente almacenados, el dispositivo de generación de flujo de bits 36 puede determinar si el algoritmo de rotación ha terminado (92). Es decir, el dispositivo de generación de flujo de bits 36 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 generación de flujo de bits 36 puede determinar si se cumplen otros criterios (tales como que se ha realizado la totalidad de un subconjunto definido de combinación, si se ha recorrido una trayectoria dada, si se ha recorrido un árbol jerárquico para un nodo de hoja, etc.) de modo que el dispositivo de generación de flujo de bits 36 haya terminado de realizar el algoritmo de rotación. Si no ha terminado ("NO" 92), el dispositivo de generación de flujo de bits 36 puede realizar el procedimiento anterior con respecto a otra combinación seleccionada (80-92). Si ha terminado ("Sí" 92), el dispositivo de generación de flujo de bits 36 puede especificar los SHC 27', el ángulo de acimut y el ángulo de elevación almacenados en el flujo de bits 31 de una de las diversas maneras descritas anteriormente (94).

[0149] La FIG. 9 es un diagrama de flujo que ilustra el funcionamiento de ejemplo del dispositivo de generación de flujo de bits 36 mostrado en el ejemplo de la FIG. 4 al realizar los aspectos de transformación de las técnicas descritas en la presente divulgación. Inicialmente, el dispositivo de generación de flujo de bits 36 puede seleccionar una matriz que representa una transformada invertible lineal (100). Un ejemplo de una matriz que representa una transformada invertible lineal puede ser la matriz mostrada anteriormente que es el resultado de $[EncMat_1][IncMat_1]$. A continuación, el dispositivo de generación de flujo de bits 36 puede aplicar la matriz al campo sonoro para transformar el campo sonoro (102). El dispositivo de generación de flujo de bits 36 también puede determinar SHC 27' que representan el

campo sonoro girado (104). Si bien se describe como que son etapas u operaciones separadas, el dispositivo de generación de flujo de bits 36 puede aplicar una transformada (que puede representar el resultado de $[EncMat_2][InvMat_1]$), derivando el campo sonoro a partir de los SHC 27, transformar el campo sonoro y determinando los SHC 27' que representan el campo sonoro de transformada.

[0150] En cualquier caso, el dispositivo de generación de flujo de bits 36 puede a continuación computar un número de los SHC 27' determinados que sean mayores que un valor umbral, comparando este número con un número computado para una iteración previa con respecto a una aplicación previa de una matriz de transformada (106, 108). Si el número determinado de los SHC 27' es menor que el número previo ("Sí" 108), el dispositivo de generación de flujo de bits 36 almacena los SHC 27' y la matriz (o algún derivado de la misma, tal como un índice asociado con la matriz), reemplazando a menudo los SHC 27' previos y la matriz (o un derivado de la misma) almacenados de una iteración previa del algoritmo de rotación (110).

[0151] Si el número determinado de los SHC 27' no es menor que el número previo ("NO" 108) o después de almacenar los SHC 27' y la matriz en lugar de los SHC 27' y la matriz previamente almacenados, el dispositivo de generación de flujo de bits 36 puede determinar si el algoritmo de transformada ha terminado (112). Es decir, el dispositivo de generación de flujo de bits 36 puede, como un ejemplo, determinar si se han evaluado todas las matrices de transformada disponibles. En otros ejemplos, el dispositivo de generación de flujo de bits 36 puede determinar si se cumplen otros criterios (tales como que se ha realizado la totalidad de un subconjunto definido de las matrices de transformada disponibles, si se ha recorrido una trayectoria dada, si se ha recorrido un árbol jerárquico para un nodo de hoja, etc.) de modo que el dispositivo de generación de flujo de bits 36 haya terminado de realizar el algoritmo de transformada. Si no ha terminado ("NO" 112), el dispositivo de generación de flujo de bits 36 puede realizar el procedimiento anterior con respecto a otra matriz de transformada seleccionada (100-112). Si ha terminado ("Sí" 112), el dispositivo de generación de flujo de bits 36 puede a continuación, como se indica anteriormente, identificar diferentes tasas de bits para los diferentes subconjuntos transformados de los SHC 27' (114). A continuación, el dispositivo de generación de flujo de bits 36 puede codificar los diferentes subconjuntos usando las tasas de bits identificadas para generar el flujo de bits 31 (116).

[0152] En algunos ejemplos, el algoritmo de transformada puede realizar una única iteración, evaluando una única matriz de transformada. Es decir, la matriz de transformada puede comprender cualquier matriz que represente una transformada invertible lineal. En algunos casos, la transformada invertible lineal puede transformar el campo sonoro del dominio espacial al dominio de frecuencia. Los ejemplos de dicha transformada invertible lineal pueden incluir una transformada de Fourier discreta (DFT). La aplicación de la DFT sólo puede implicar una única iteración y, por lo tanto, no necesariamente incluirá etapas para determinar si el algoritmo de transformada ha terminado. En consecuencia, las técnicas no se deben limitar al ejemplo de la FIG. 9.

[0153] En otras palabras, un ejemplo de una transformada invertible lineal es una transformada de Fourier discreta (DFT). Se pueden realizar operaciones en los veinticinco SHC 27' por la DFT para formar un conjunto de veinticinco coeficientes complejos. El dispositivo de generación de flujo de bits 36 también puede rellenar con ceros los veinticinco SHC 27' para que sean un múltiplo entero de 2, a fin de aumentar potencialmente la resolución del tamaño de intervalo de la DFT, y potencialmente tienen una implementación más eficaz de la DFT, por ejemplo, a través de la aplicación de 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 transformada, el dispositivo de generación de flujo de bits 36 puede aplicar un umbral para determinar si existe alguna energía espectral en un intervalo particular. A continuación, el dispositivo de generación de flujo de bits 36 puede, en este contexto, descartar o llevar a cero energía de coeficiente espectral que esté por debajo de este umbral, y el dispositivo de generación de flujo de bits 36 puede aplicar una transformada inversa para recuperar SHC 27' que tengan uno o más de los SHC 27' descartados o llevados a cero. 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.

[0154] Otra transformada invertible lineal puede comprender una matriz que realiza lo que se denomina "descomposición en valores singulares". Si bien se describe con respecto a la SVD, las técnicas se pueden realizar con respecto a cualquier transformación o descomposición similar que proporcione conjuntos de datos linealmente no correlacionados. Además, la referencia a "conjuntos" o "subconjuntos" en la presente divulgación está en general concebida para referirse a conjuntos o subconjuntos "distintos de cero" a menos que se mencione específicamente lo contrario y no está concebida para referirse a la definición matemática clásica de conjuntos que incluye el denominado "conjunto vacío".

[0155] Las transformaciones alternativas pueden incluir un análisis de componentes principales, que a menudo se abrevia por las siglas PCA. Un PCA se refiere a un procedimiento matemático que emplea una transformación ortogonal para convertir un conjunto de observaciones de variables posiblemente correlacionadas en un conjunto de variables no correlacionadas linealmente, denominadas componentes principales. Las variables no correlacionadas linealmente representan variables que no tienen una relación (o dependencia) estadística lineal entre sí. Estos componentes principales se pueden describir 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. Típicamente, 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 (que se puede reexpresar como no correlacionado) con los componentes precedentes. Un PCA puede realizar una forma de reducción de orden, que en términos de los SHC puede dar como resultado la compresión de los SHC. Dependiendo del contexto, un PCA se puede denominar mediante un número de nombres diferentes, tales como la transformada de Karhunen-Loeve discreta, la transformada de Hotelling, la descomposición ortogonal propia (POD) y la descomposición en autovalores (EVD), para nombrar unos pocos ejemplos.

[0156] En cualquier caso, una SVD representa un procedimiento que se aplica a los SHC para transformar los SHC en dos o más conjuntos de coeficientes armónicos esféricos transformados. El dispositivo de generación de flujo de bits 36 puede realizar una SVD con respecto a los SHC 27 para generar una denominada matriz V, una matriz S y una matriz U. Una SVD, en álgebra lineal, puede representar una factorización de una matriz X real o compleja de m por n (donde X puede representar datos de audio multicanal, tales como los SHC 11A) en la siguiente forma:

$$X = USV^*$$

[0157] U puede representar una matriz unitaria real o compleja de m por m, donde las m columnas de U se conocen comúnmente como los vectores singulares izquierdos de los datos de audio multicanal. S puede representar una matriz diagonal rectangular de m por n 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 de n por n, donde las n columnas de V* se conocen comúnmente como los vectores singulares derechos de los datos de audio multicanal.

[0158] Si bien se describe en la presente divulgación como que se aplican a datos de audio multicanal que comprenden coeficientes armónicos esféricos 27, las técnicas se pueden aplicar a cualquier forma de datos de audio multicanal. De esta manera, el dispositivo de generación de flujo de bits 36 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.

[0159] En general, la matriz V* en la expresión matemática de SVD mencionada anteriormente se indica como la 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*) se puede considerar igual a la matriz V. A continuación se supone, para propósitos de facilidad de ilustración, que los SHC 11A comprenden números reales con el resultado de que la matriz V se emite a través de SVD en lugar de la matriz V*. Si bien se supone que es la matriz V, las técnicas se pueden aplicar de una manera similar a los SHC 11A que tienen coeficientes complejos, donde la salida de la SVD es la matriz V*. En consecuencia, las técnicas no se deben limitar en este sentido a proporcionar solo la aplicación de SVD para generar una matriz V, sino que pueden incluir la aplicación de SVD a los SHC 11A que tienen componentes complejos para generar una matriz V*.

[0160] En el contexto de la SVD, el dispositivo de generación de flujo de bits 36 puede especificar la información de transformación en el flujo de bits como un indicador definido por uno o más bits que indican si se aplicó SVD (o más en general, una transformación basada en vectores) a los SHC 27 o si se aplicaron otras transformaciones o esquemas de codificación variables.

[0161] En consecuencia, en un campo sonoro tridimensional, las direcciones en las que se origina una fuente de sonido se pueden considerar las más importantes. Como se describe anteriormente, se proporciona una metodología para girar el campo sonoro calculando la dirección en la que está presente la energía principal. A continuación, el campo sonoro se puede girar de manera que esta energía, o la ubicación espacial más importante, se gire a continuación para estar en los coeficientes armónicos esféricos de an0. El motivo de esto es simple, de manera que cuando se eliminan los armónicos esféricos innecesarios (es decir, por debajo de un umbral dado) probablemente habrá la menor cantidad de coeficientes armónicos esféricos necesarios para cualquier orden N dado, que es N armónicos esféricos. Debido al gran ancho de banda requerido para almacenar incluso estos coeficientes HOA reducidos, a continuación se puede requerir una forma de compresión de datos. Si se usa la misma tasa de bits a través de todos los armónicos esféricos, entonces algunos de los coeficientes están usando potencialmente más bits de los necesarios para producir una codificación perceptualmente transparente, mientras que otros coeficientes de armónicos esféricos no usan potencialmente una tasa de bits lo suficientemente grande como para hacer que el coeficiente sea perceptualmente transparente. Por consiguiente, se puede requerir un procedimiento para adjudicar la tasa de bits de manera inteligente a través de los coeficientes HOA.

[0162] Las técnicas descritas en la presente divulgación pueden proporcionar que, para la compresión de velocidad de transferencia de datos de audio de armónicos esféricos, el campo sonoro se gira en primer lugar de manera que,

como un ejemplo, la dirección donde se origina la energía más grande se sitúa en el eje Z. Con esta rotación, el coeficiente armónico esférico de a_{n0} puede tener la mayor energía, ya que las funciones base armónicas esféricas Y_{n0} tienen lóbulos máximos y mínimos que apuntan en el eje Z (eje arriba-abajo). Debido a la naturaleza de las funciones base armónicas esféricas, la distribución de energía probablemente residirá en gran medida en el coeficiente a_{n0} , mientras que la energía mínima estará en $a_{n+/-n}$ de base horizontal y la energía en otros coeficientes de valor m $-n < m < n$ aumentará entre $m = -n$ y $m=0$ y a continuación disminuirá de nuevo entre $m = 0$ y $m = n$. A continuación, las técnicas pueden asignar una tasa de bits mayor a los coeficientes a_{n0} y la cantidad mínima a los coeficientes $a_{n+/-n}$. En este sentido, las técnicas pueden proporcionar una adjudicación de la tasa de bits dinámica que varía por orden y/o suborden. Los coeficientes intermedios para un orden dado probablemente tienen tasas de bits intermedias. Para calcular las tasas, se puede usar una función ventana (WIN) que puede tener p número de puntos para cada orden de HOA incluido en la señal de HOA. Las tasas se podrían aplicar, como un ejemplo, usando el factor WIN de la diferencia entre las tasas de bits altas y bajas. Las tasas de bits altas y bajas se pueden definir por orden de los órdenes incluidos dentro de la señal de HOA. La ventana resultante en tres dimensiones se asemejaría a un tipo de 'carpa principal' de carpa de circo que apunta hacia arriba en el eje Z y otra como su reflejo que apunta hacia abajo en el eje Z, donde se reflejan en el plano horizontal.

[0163] La FIG. 10 es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de extracción, tal como el dispositivo de extracción 38 mostrado en el ejemplo de la FIG. 3, al realizar diversos aspectos de las técnicas descritas en la presente divulgación. Inicialmente, el dispositivo de extracción 38 puede determinar la información de transformación 52 (120), que se puede especificar en el flujo de bits 31 como se muestra en los ejemplos de las FIGS. 7A-7E. A continuación, el dispositivo de extracción 38 puede determinar los SHC 27 transformados, como se describe anteriormente (122). A continuación, el dispositivo de extracción 38 puede transformar los SHC 27 transformados basado en la información de transformación 52 determinada para generar los SHC 27'. En algunos ejemplos, el dispositivo de extracción 38 puede seleccionar un reproductor que realice eficazmente esta transformación basada en la información de transformación 52. Es decir, el dispositivo de extracción 38 puede funcionar de acuerdo con la siguiente ecuación para generar los SHC 27':

$$\begin{bmatrix} SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (25 \times 32) \end{bmatrix} \begin{bmatrix} Reprodutor \\ (32 \times 25) \end{bmatrix} \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

En la ecuación anterior, la $[EncMat_2][Reprodutor]$ se puede usar para transformar el reproductor en la misma cantidad de manera que ambas direcciones frontales coincidan y, de este modo, deshacer o contrarrestar la rotación realizada en el dispositivo de generación de flujo de bits.

[0164] La FIG. 11 es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de generación de flujo de bits, tal como el dispositivo de generación de flujo de bits 36 mostrado en el ejemplo de la FIG. 3, y un dispositivo de extracción, tal como el dispositivo de extracción 38 también mostrado en el ejemplo de la FIG. 3, al realizar diversos aspectos de las técnicas descritas en la presente divulgación. Inicialmente, el dispositivo de generación de flujo de bits 36 puede identificar un subconjunto de SHC 27 que se van a incluir en el flujo de bits 31 de cualquiera de las diversas maneras descritas anteriormente y mostradas con respecto a las FIGS. 7A-7E (140). A continuación, el dispositivo de generación de flujo de bits 36 puede especificar el subconjunto identificado de los SHC 27 en el flujo de bits 31 (142). A continuación, el dispositivo de extracción 38 puede obtener el flujo de bits 31, determinar el subconjunto de los SHC 27 especificados en el flujo de bits 31 y analizar sintácticamente el subconjunto de los SHC 27 determinados a partir del flujo de bits.

[0165] En algunos ejemplos, el dispositivo de generación de flujo de bits 36 y el dispositivo de extracción 38 pueden realizar otros aspectos diversos de las técnicas junto con este subconjunto de SHC que señalizan aspectos de las técnicas. Es decir, el dispositivo de generación de flujo de bits 36 puede realizar una transformación con respecto a los SHC 27 para reducir el número de SHC 27 que se van a especificar en el flujo de bits 31. A continuación, el dispositivo de generación de flujo de bits 36 puede identificar el subconjunto de los SHC 27 restantes después de realizar esta transformación en el flujo de bits 31 y especificar estos SHC 27 transformados en el flujo de bits 31, mientras que también especifica la información de transformación 52 en el flujo de bits 31. A continuación, el dispositivo de extracción 38 puede obtener el flujo de bits 31, determinar el subconjunto de los SHC 27 transformados y analizar sintácticamente el subconjunto de los SHC 27 transformados determinado a partir del flujo de bits 31. A continuación, el dispositivo de extracción 38 puede recuperar los SHC 27 (que se muestran como SHC 27') transformando los SHC 27 transformados basado en la información de transformación para generar los SHC 27'. Por tanto, si bien se muestran por separado unos de otros, diversos aspectos de las técnicas se pueden realizar junto con unos y otros.

[0166] Se debe entender que, dependiendo del ejemplo, determinados actos o acontecimientos de cualquiera de los procedimientos descritos en el presente documento se pueden realizar en una secuencia diferente, se pueden añadir, fundir u omitir por completo (por ejemplo, no todos los actos o acontecimientos descritos son necesarios para la puesta en práctica del procedimiento). Asimismo, en determinados ejemplos, los actos o acontecimientos se pueden realizar simultáneamente, por ejemplo, a través de procesamiento de múltiples hilos, procesamiento de interrupciones o múltiples procesadores, en lugar de secuencialmente. Además, si bien determinados aspectos de la presente divulgación se describen como que se realizan mediante un único dispositivo, módulo o unidad para propósitos de

claridad, se debe entender que las técnicas de la presente divulgación se pueden realizar mediante una combinación de dispositivos, unidades o módulos.

[0167] En uno o más ejemplos, las funciones descritas se pueden implementar en hardware, software, firmware o cualquier combinación de los mismos. Si se implementan en software, las funciones se pueden almacenar en, o transmitir por un medio legible por ordenador, 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 correspondan a un medio tangible tal como medios de almacenamiento de datos, o medios de comunicación que incluyan cualquier medio que facilite la transferencia de un programa informático desde un lugar a otro, por ejemplo, de acuerdo a un protocolo de comunicación.

[0168] De esta manera, los medios legibles por ordenador pueden corresponder en general a (1) medios de almacenamiento legibles por ordenador tangibles 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 pueda acceder mediante uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en la presente divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

[0169] A modo de ejemplo, y no de limitación, 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 se pueda usar para almacenar un código de programa deseado en forma de instrucciones o estructuras de datos y al que se pueda acceder mediante un ordenador. Además, cualquier conexión recibe apropiadamente la denominación de medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde una sede de la Red, un servidor u otro origen remoto usando 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.

[0170] Sin embargo, se debe entender 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 dirigen a medios de almacenamiento tangibles no transitorios. Los discos, como se usan en el presente documento, incluyen el disco compacto (CD), el disco láser, el disco óptico, el disco versátil digital (DVD), el disco flexible y el disco Blu-ray, donde algunos discos reproducen normalmente los datos magnéticamente, mientras que otros discos reproducen los datos ópticamente con láseres. Las combinaciones de lo anterior también se deben incluir dentro del alcance de los medios legibles por ordenador.

[0171] Las instrucciones se pueden ejecutar 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. En consecuencia, el término "procesador", como se usa en el presente documento, se puede referir a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento se puede proporcionar dentro de módulos de hardware y/o software dedicados, configurados para la codificación y la decodificación, o incorporados en un códec combinado. Además, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.

[0172] Las técnicas de la presente divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, incluyendo un teléfono inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). Diversos componentes, módulos o unidades se describen en la presente divulgación para enfatizar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no requieren necesariamente su realización mediante diferentes unidades de hardware. En cambio, como se describe anteriormente, diversas unidades se pueden combinar en una unidad de hardware de códec, o proporcionarse mediante un grupo de unidades de hardware interoperativas, incluyendo uno o más procesadores, como se describe anteriormente, junto con software y/o firmware adecuados.

[0173] Se han descrito diversos modos de realización de las técnicas. Estos y otros modos de realización están dentro del alcance de las siguientes reivindicaciones.

REIVINDICACIONES

1. Un procedimiento de generación de un flujo de bits (31) representativo de contenido de audio, comprendiendo el procedimiento:

determinar que uno o más de una pluralidad de elementos jerárquicos tienen información pertinente al describir un campo sonoro en tres dimensiones, en el que la pluralidad de elementos jerárquicos comprende coeficientes armónicos esféricos (27) representativos del campo sonoro, comprendiendo la determinación determinar si un coeficiente armónico esférico está por encima de un valor umbral, siendo el valor umbral mayor que cero;

especificar (140), en el flujo de bits y basado en la determinación, cuáles de los coeficientes armónicos esféricos se incluyen en el flujo de bits; y

especificar (142), en el flujo de bits, el uno o más de la pluralidad de coeficientes armónicos esféricos determinados.

2. El procedimiento de la reivindicación 1, en el que especificar cuáles de los coeficientes armónicos esféricos se incluyen en el flujo de bits comprende especificar un campo que tiene una pluralidad de bits identificando uno diferente de la pluralidad de bits si uno correspondiente de la pluralidad de coeficientes armónicos esféricos se incluye en el flujo de bits.

3. El procedimiento de la reivindicación 1, en el que especificar cuáles de los coeficientes armónicos esféricos se incluyen en el flujo de bits comprende especificar un campo que tiene una pluralidad de bits igual a $(1+n)^2$ bits,

en el que n indica un orden de la pluralidad de coeficientes armónicos esféricos que describen el campo sonoro, y

en el que cada uno de la pluralidad de bits identifica si se incluye uno correspondiente de la pluralidad de coeficientes armónicos esféricos en el flujo de bits.

4. El procedimiento de la reivindicación 1,

en el que especificar cuáles de los coeficientes armónicos esféricos se incluyen en el flujo de bits comprende especificar un campo en el flujo de bits que tiene una pluralidad de bits identificando uno diferente de la pluralidad de bits si se incluye uno correspondiente de la pluralidad de coeficientes armónicos esféricos en el flujo de bits, y

en el que especificar la pluralidad de coeficientes armónicos esféricos identificados comprende especificar, en el flujo de bits, el uno o más de la pluralidad de coeficientes armónicos esféricos identificados directamente después del campo que tiene la pluralidad de bits.

5. El procedimiento de la reivindicación 1, que comprende además

especificar, en el flujo de bits, que los restantes de la pluralidad de coeficientes armónicos esféricos que tienen información no pertinente al describir el campo sonoro no se incluyen en el flujo de bits.

6. Un dispositivo configurado para generar un flujo de bits representativo de contenido de audio, comprendiendo el dispositivo:

medios para determinar que uno o más de una pluralidad de elementos jerárquicos tienen información pertinente al describir un campo sonoro en tres dimensiones, en el que la pluralidad de elementos jerárquicos son coeficientes armónicos esféricos representativos del campo sonoro, comprendiendo la determinación determinar si un coeficiente armónico esférico está por encima de un valor umbral, siendo el valor umbral mayor que cero;

medios para especificar (140), en el flujo de bits y basado en la determinación, cuáles de los coeficientes armónicos esféricos se incluyen en el flujo de bits; y

medios para especificar (142), en el flujo de bits, el uno o más de la pluralidad de coeficientes armónicos esféricos determinados.

7. El dispositivo de la reivindicación 6 que comprende medios para llevar a cabo un procedimiento de acuerdo con cualquiera de las reivindicaciones 2 a 5.

8. Un procedimiento de procesamiento de un flujo de bits (31) representativo de contenido de audio, comprendiendo el procedimiento:
- 5 identificar (146), a partir del flujo de bits, que uno o más de una pluralidad de coeficientes armónicos esféricos (27') se incluyen en el flujo de bits, en el que se ha determinado que el uno o más de la pluralidad de coeficientes armónicos esféricos proporcionan información pertinente al describir un campo sonoro en tres dimensiones al estar por encima de un valor umbral, siendo el valor umbral mayor que cero, y en el que la pluralidad de coeficientes armónicos esféricos son representativos del campo sonoro; y
- 10 analizar sintácticamente (148) el flujo de bits para determinar el uno o más de la pluralidad de coeficientes armónicos esféricos identificados.
9. El procedimiento de la reivindicación 8, en el que identificar que el uno o más de la pluralidad de coeficientes armónicos esféricos se incluyen en el flujo de bits comprende analizar sintácticamente el flujo de bits para
- 15 identificar un campo que tiene una pluralidad de bits identificando cada uno de la pluralidad de bits si uno correspondiente de la pluralidad de coeficientes armónicos esféricos se incluye en el flujo de bits.
10. El procedimiento de la reivindicación 8, en el que identificar que el uno o más de la pluralidad de coeficientes armónicos esféricos se incluyen en el flujo de bits comprende analizar sintácticamente un campo que tiene una
- 20 pluralidad de bits igual a $(1+n)^2$ bits,
- en el que n indica un orden de la pluralidad de coeficientes armónicos esféricos que describen el campo sonoro, y
- 25 en el que cada uno de la pluralidad de bits identifica si se incluye uno correspondiente de la pluralidad de coeficientes armónicos esféricos en el flujo de bits.
11. El procedimiento de la reivindicación 8,
- 30 en el que identificar que el uno o más de la pluralidad de coeficientes armónicos esféricos se incluyen en el flujo de bits comprende analizar sintácticamente un campo en el flujo de bits que tiene una pluralidad de bits identificando uno diferente de la pluralidad de bits si uno correspondiente de la pluralidad de coeficientes armónicos esféricos se incluye en el flujo de bits, y
- 35 en el que analizar sintácticamente el flujo de bits para determinar el uno o más de la pluralidad de coeficientes armónicos esféricos identificados comprende analizar sintácticamente el flujo de bits para determinar la pluralidad de coeficientes armónicos esféricos identificados directamente a partir del flujo de bits después del campo que tiene la pluralidad de bits.
12. El procedimiento de la reivindicación 8, que comprende además identificar, a partir del flujo de bits, que los
- 40 restantes de la pluralidad de coeficientes armónicos esféricos que tienen información no pertinente al describir el campo sonoro no se incluyen en el flujo de bits.
13. Un dispositivo configurado para procesar un flujo de bits (31) representativo de contenido de audio, comprendiendo el dispositivo:
- 45 medios para identificar (146), a partir del flujo de bits, que uno o más de una pluralidad de coeficientes armónicos esféricos (27') se incluyen en el flujo de bits, en el que se ha determinado que el uno o más de la pluralidad de coeficientes armónicos esféricos proporcionan información pertinente al describir un campo sonoro en tres dimensiones al estar por encima de un valor umbral, siendo el valor umbral mayor que cero, y en el que la pluralidad de coeficientes armónicos esféricos son representativos del campo sonoro; y
- 50 medios para analizar sintácticamente (148) el flujo de bits para determinar la pluralidad de coeficientes armónicos esféricos identificados.
14. El dispositivo de la reivindicación 13 que comprende medios para llevar a cabo un procedimiento de acuerdo con cualquiera de las reivindicaciones 9 a 12.
- 55 15. Un medio de almacenamiento legible por ordenador no transitorio que tiene almacenadas en el mismo instrucciones que, cuando se ejecutan, hacen que uno o más procesadores lleven a cabo un procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 5 y 8 a 12.
- 60

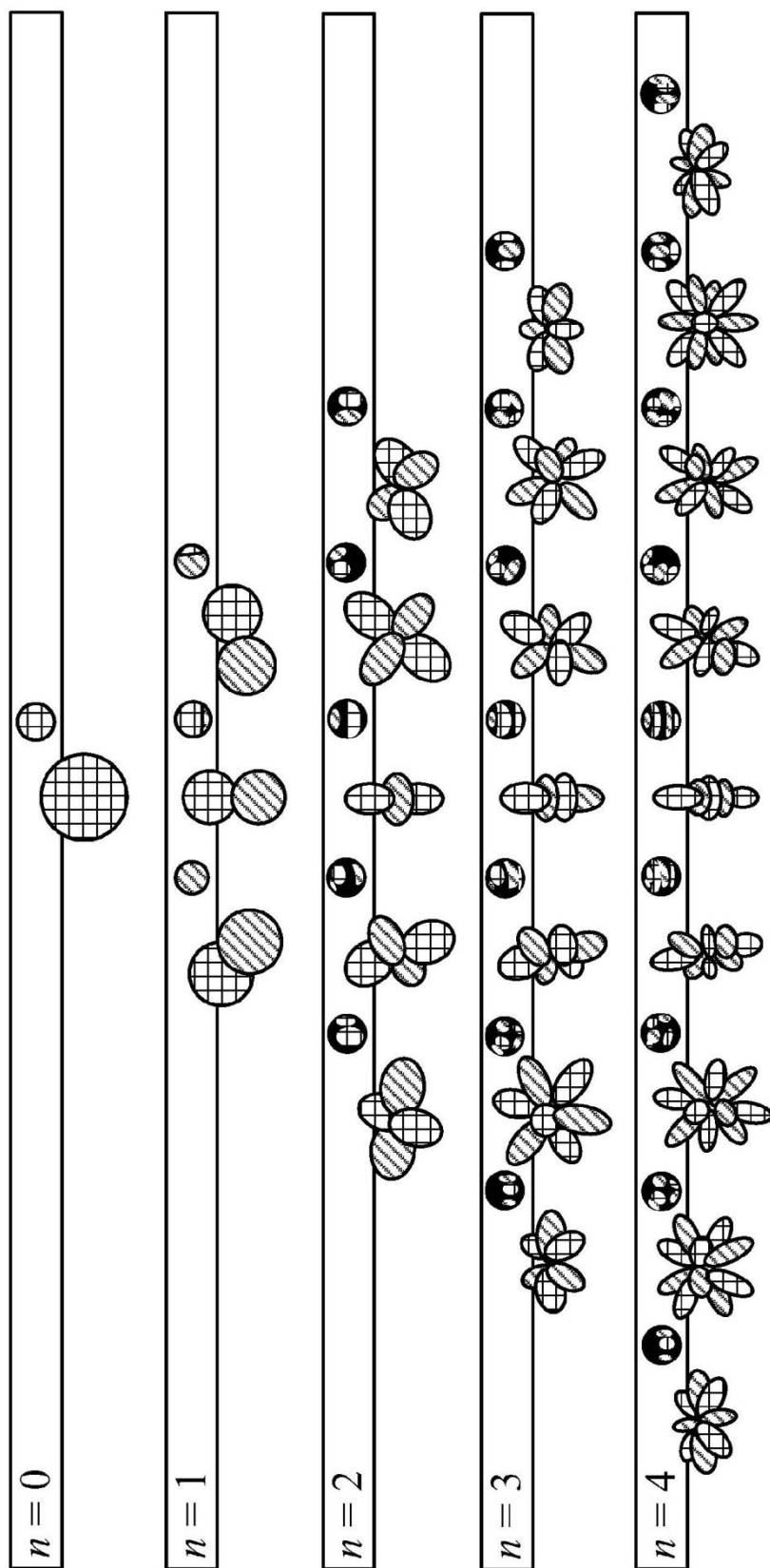


FIG. 1

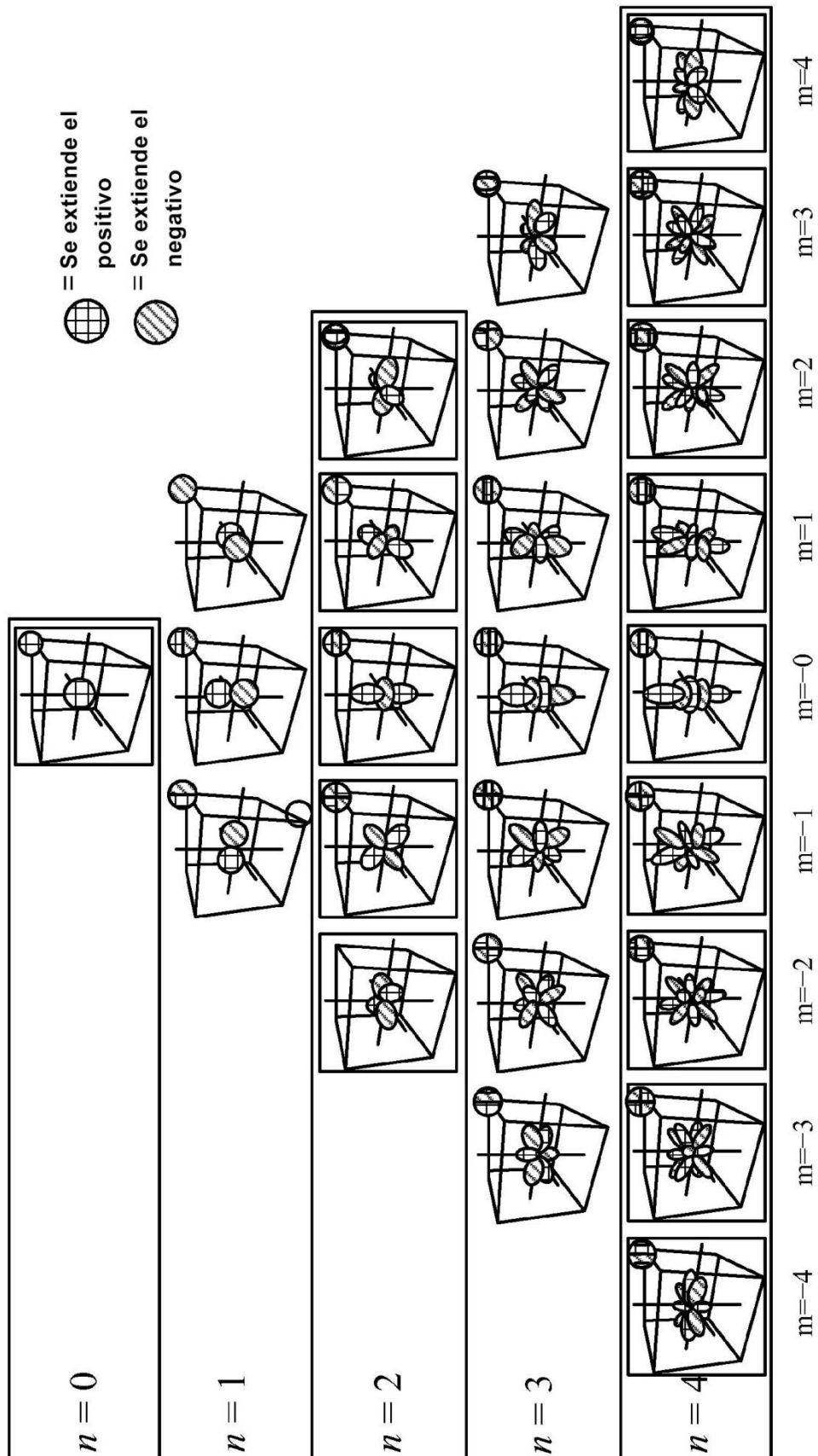


FIG. 2

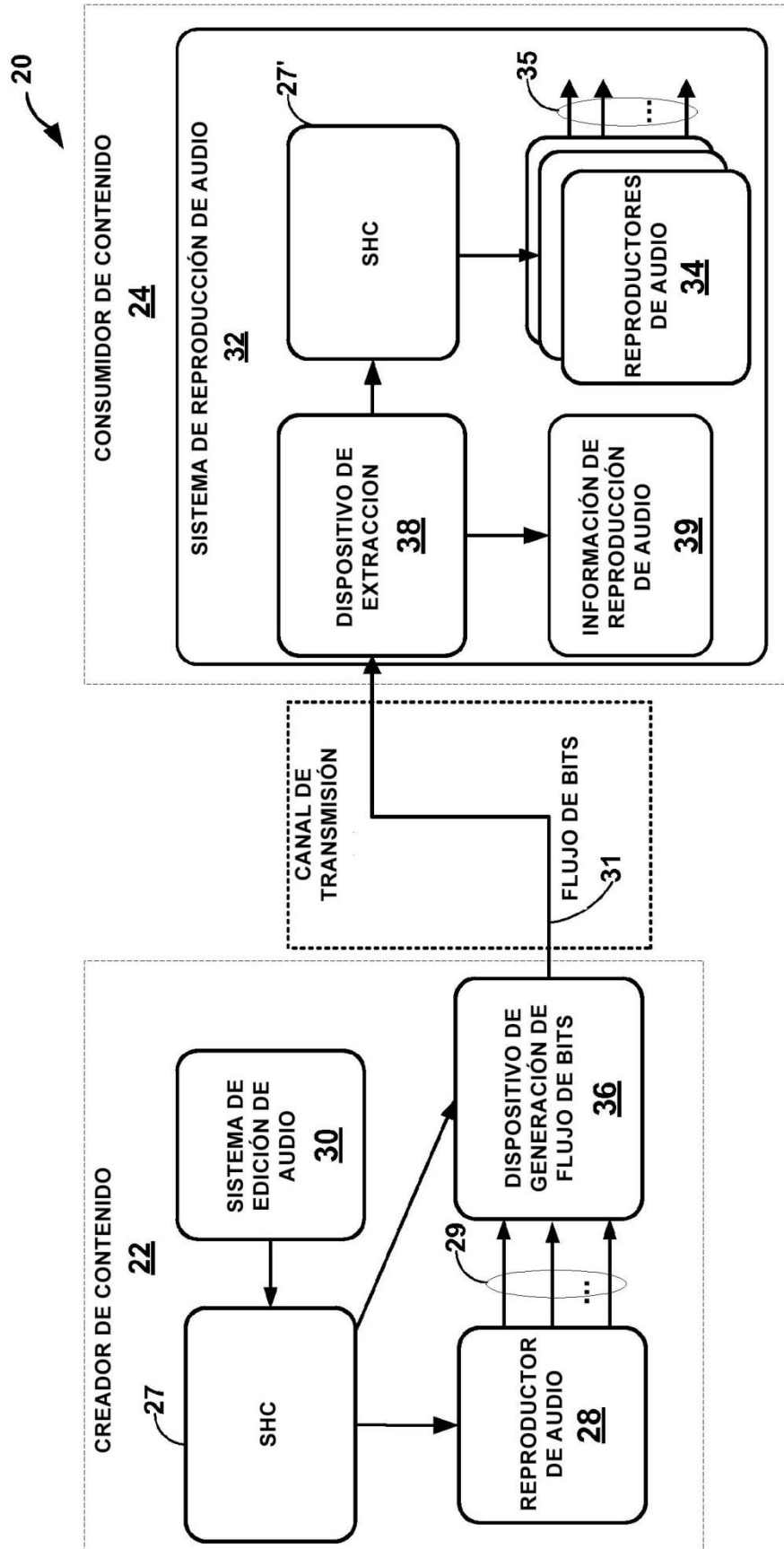


FIG. 3

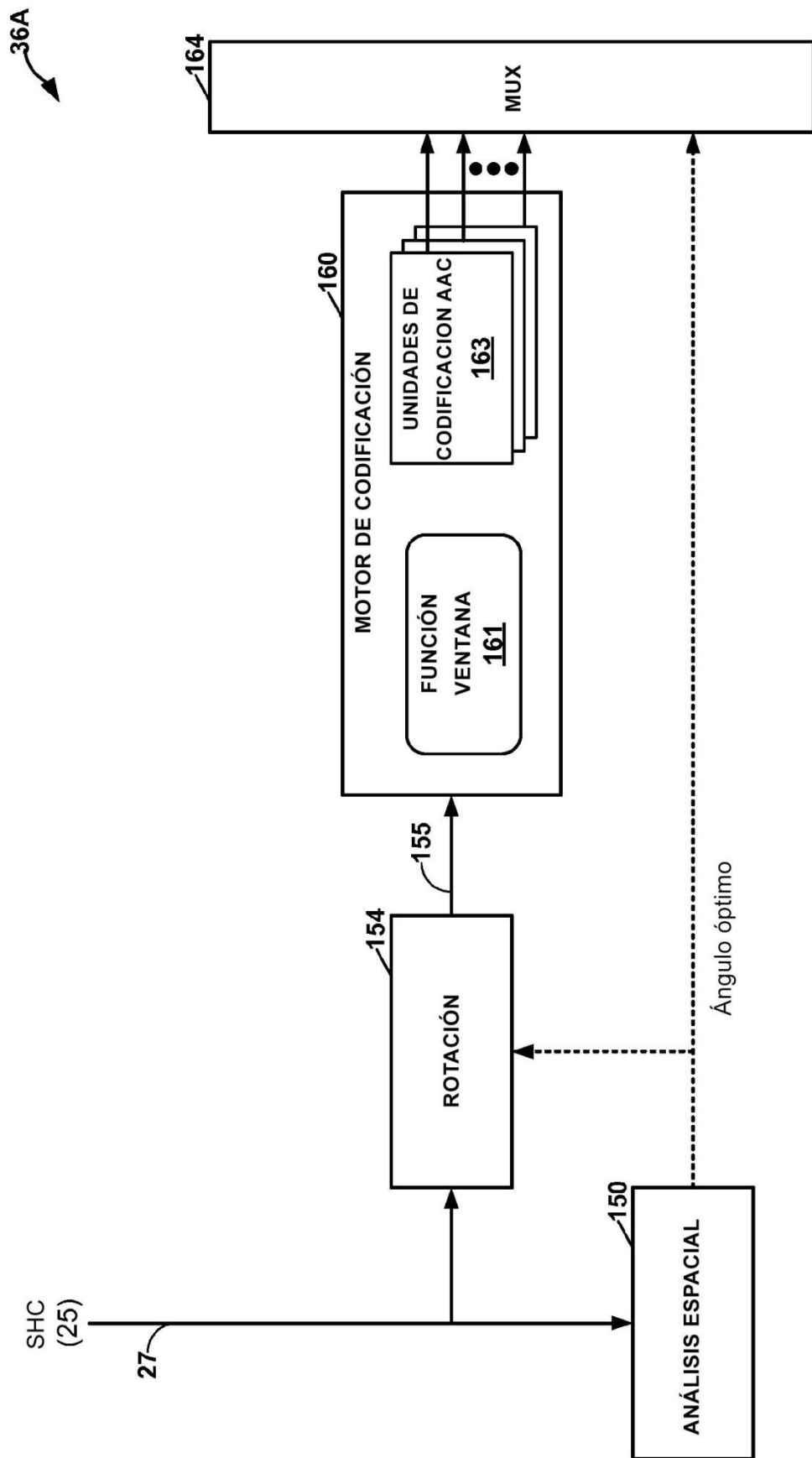


FIG. 4A

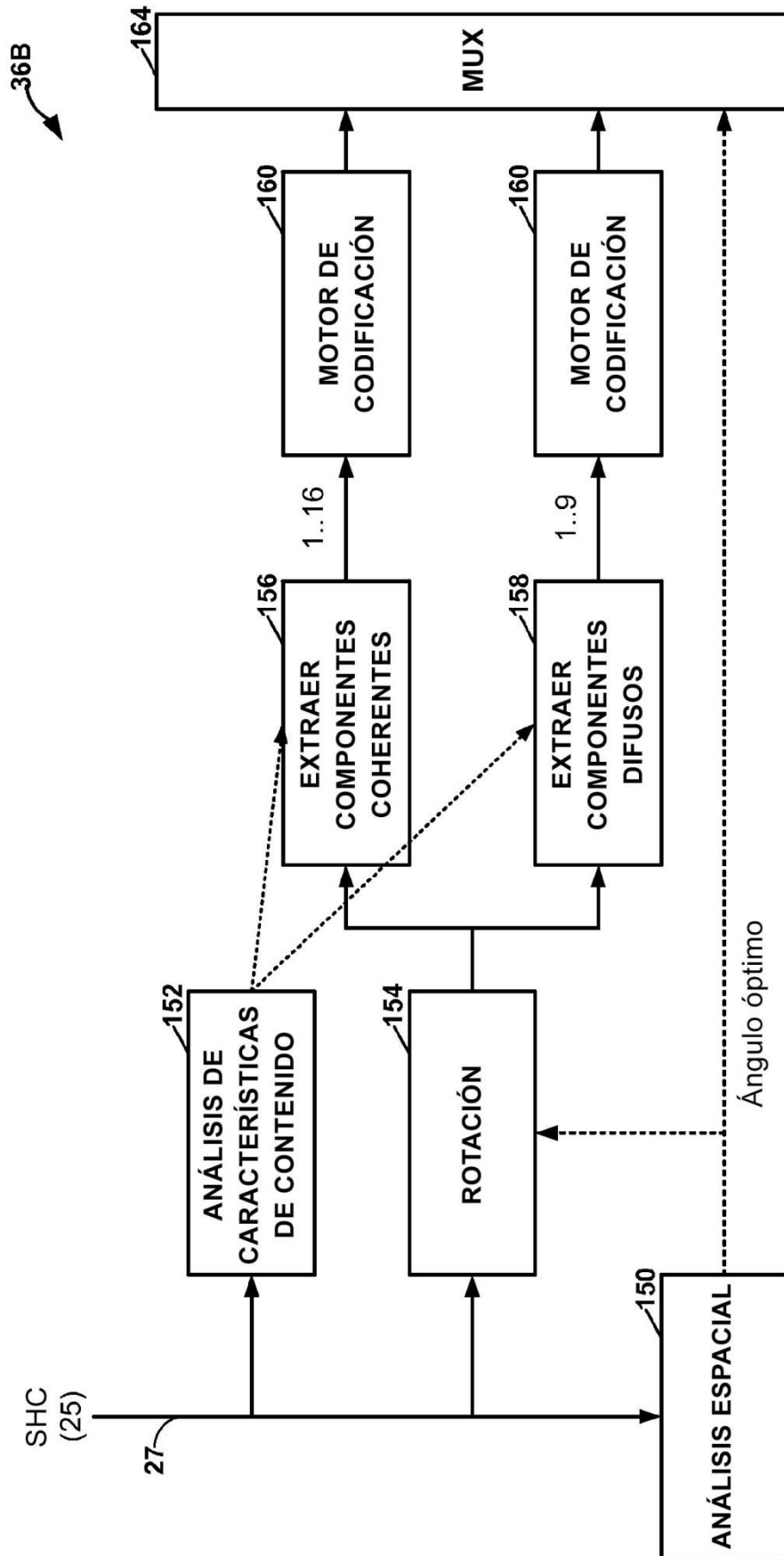


FIG. 4B

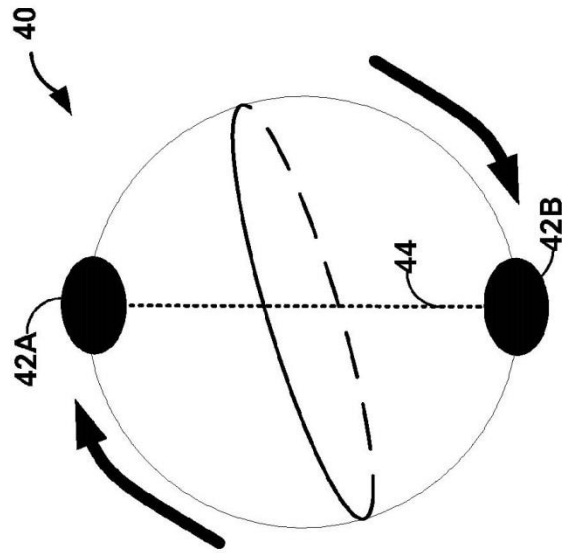


FIG. 5B

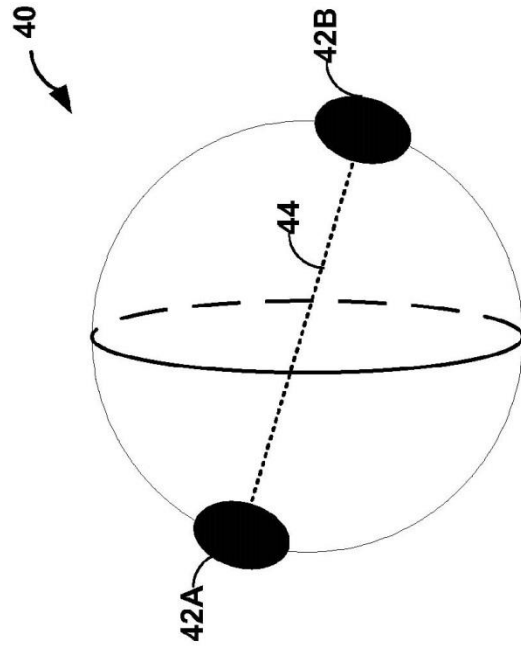


FIG. 5A

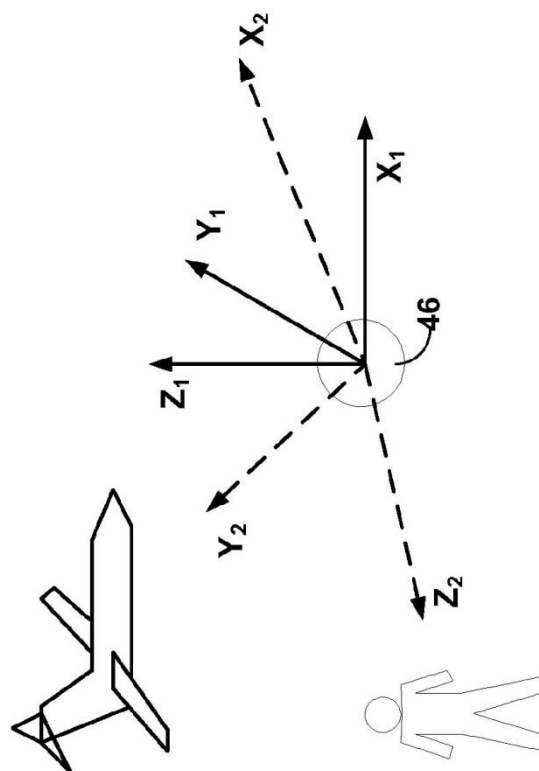


FIG. 6

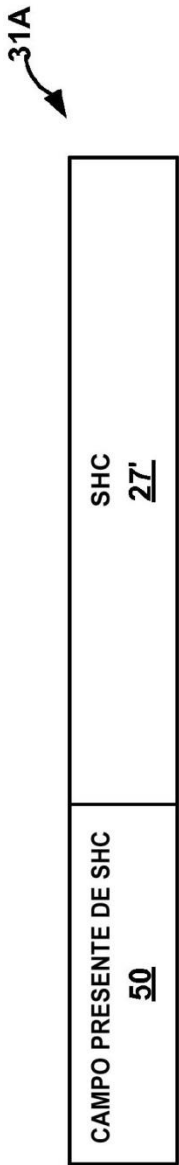


FIG. 7A

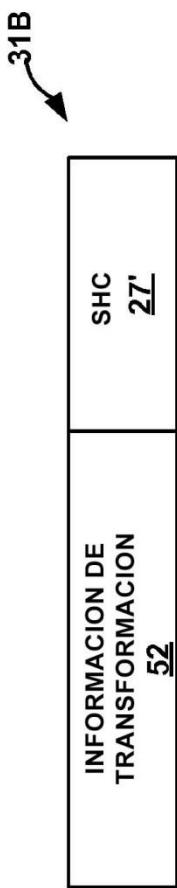


FIG. 7B

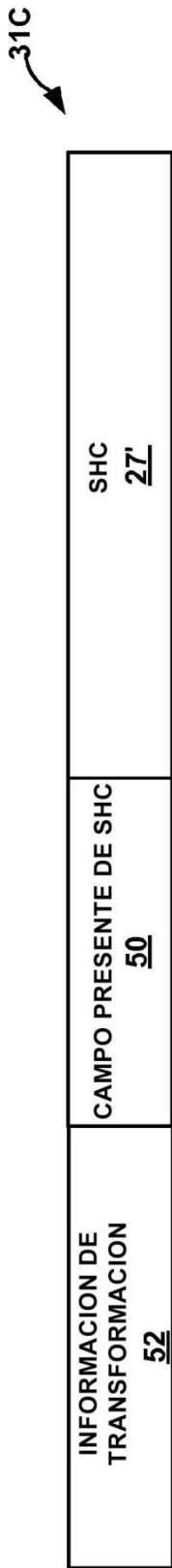


FIG. 7C

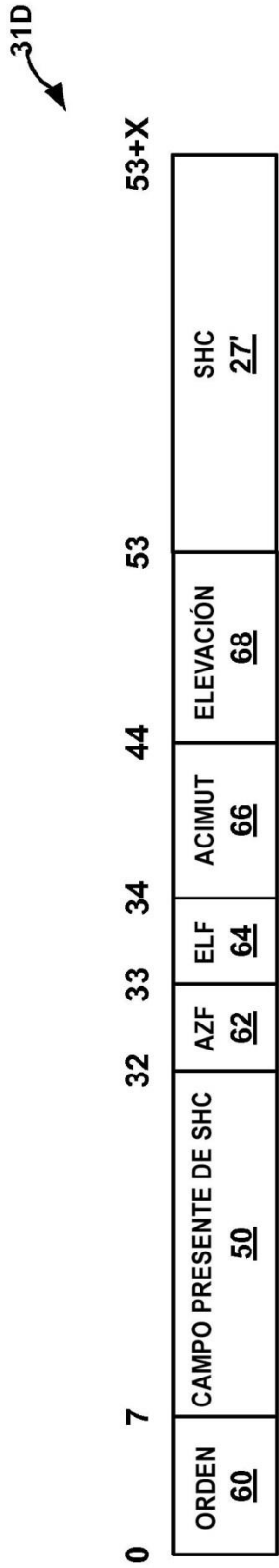


FIG. 7D

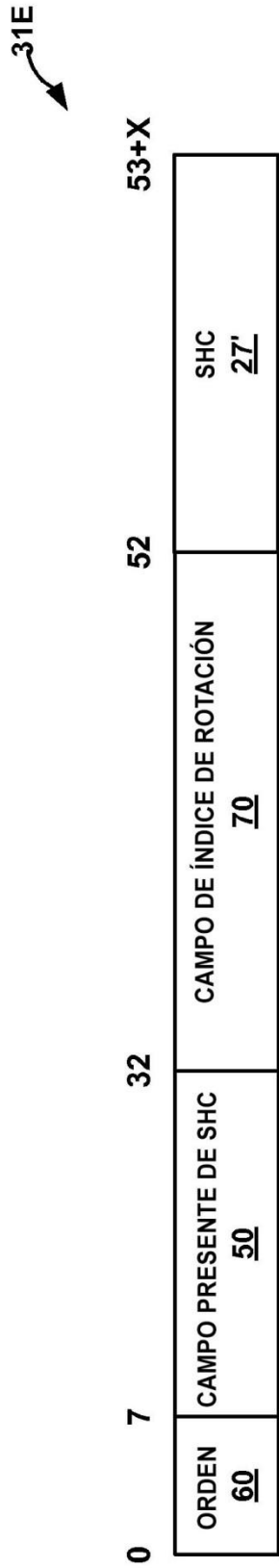


FIG. 7E

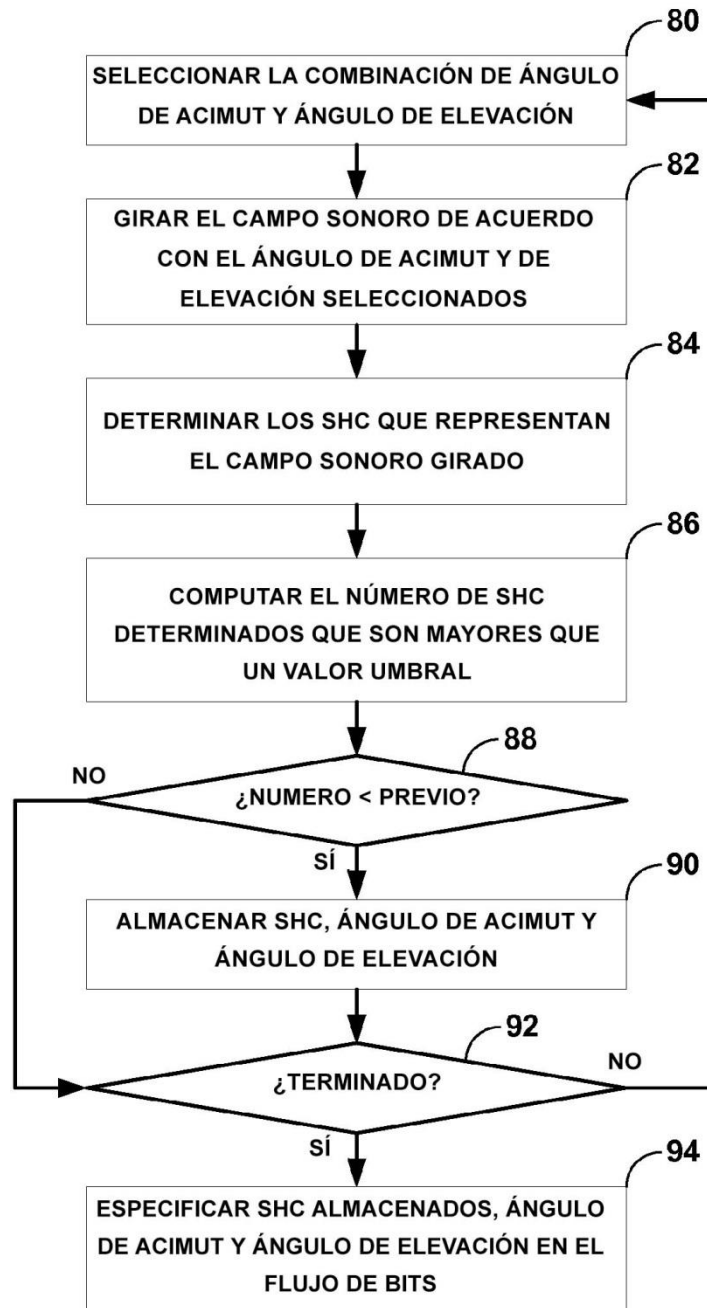


FIG. 8

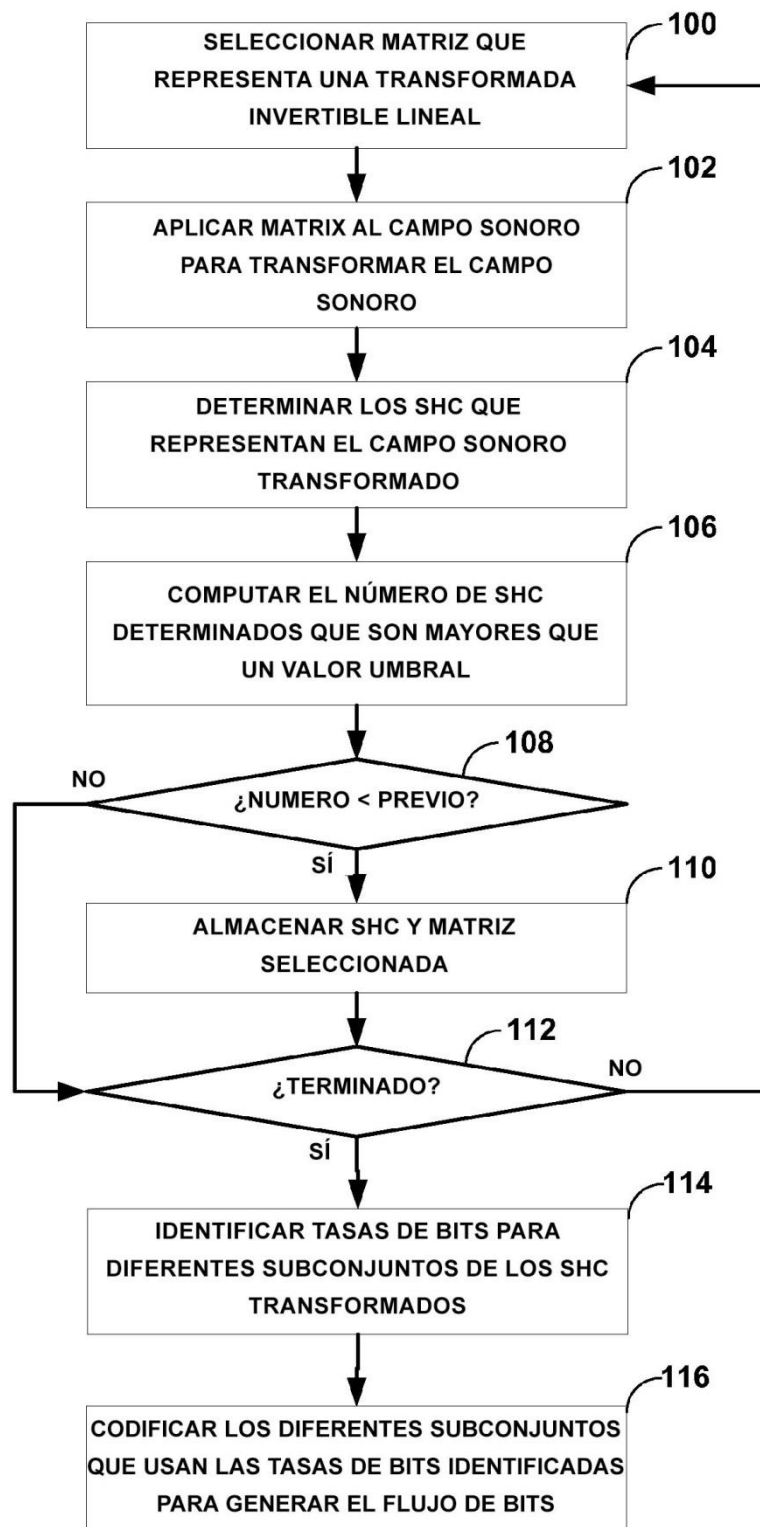


FIG. 9

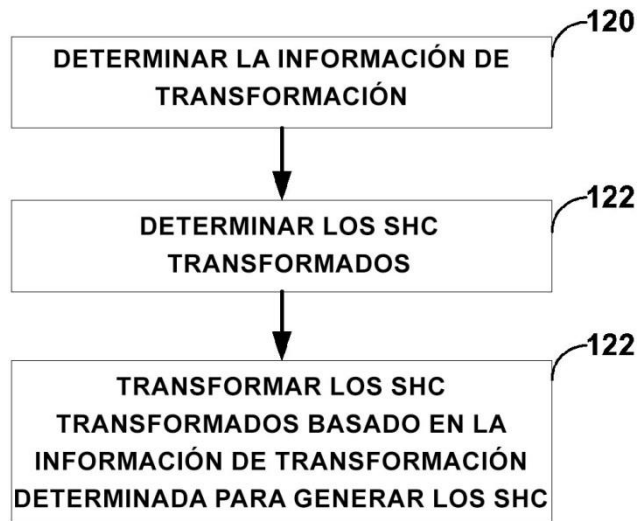


FIG. 10

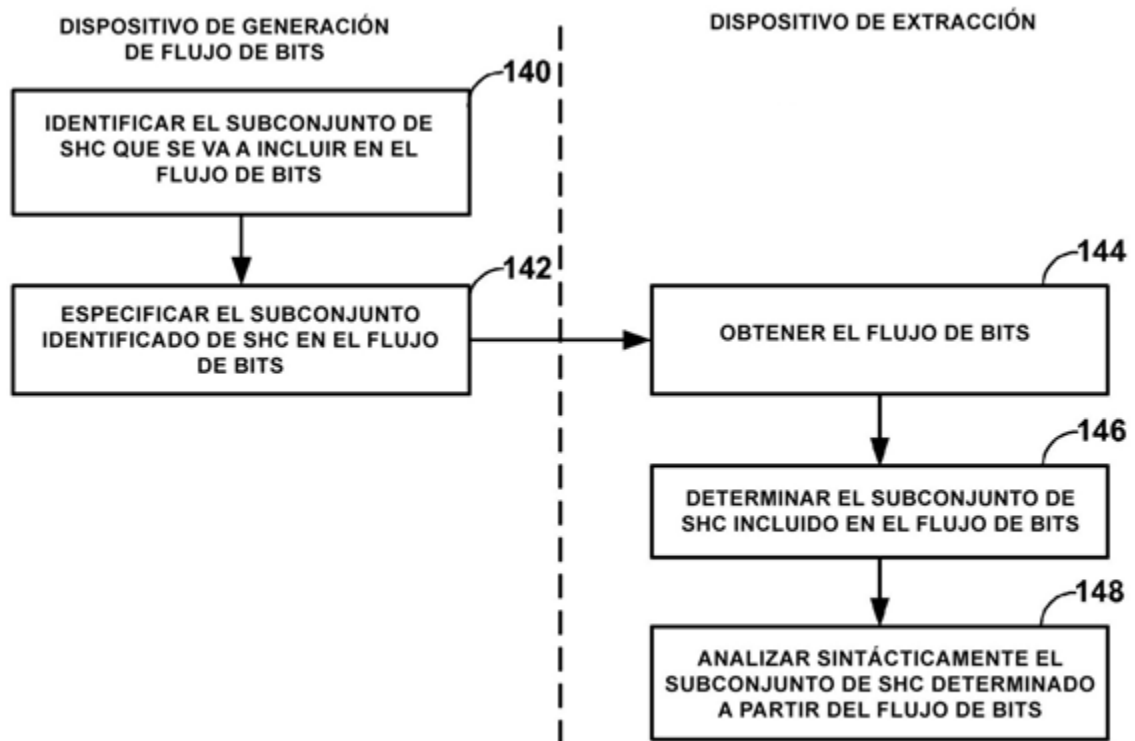


FIG. 11