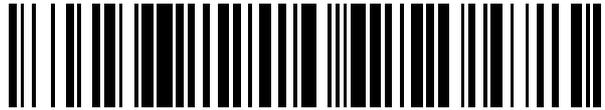


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 729 624**

51 Int. Cl.:

H04S 5/00 (2006.01)

G10L 19/008 (2013.01)

H04S 3/00 (2006.01)

H04R 5/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.07.2015 PCT/US2015/038943**

87 Fecha y número de publicación internacional: **07.01.2016 WO16004277**

96 Fecha de presentación y número de la solicitud europea: **02.07.2015 E 15741701 (5)**

97 Fecha y número de publicación de la concesión europea: **06.03.2019 EP 3165001**

54 Título: **Reducción de la correlación entre canales de fondo ambisónicos de orden superior (HOA)**

30 Prioridad:

02.07.2014 US 201462020348 P

06.10.2014 US 201462060512 P

01.07.2015 US 201514789961

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

05.11.2019

73 Titular/es:

QUALCOMM INCORPORATED (100.0%)

5775 Morehouse Drive

San Diego, CA 92121-1714, US

72 Inventor/es:

PETERS, NILS GÜNTHER;

SEN, DIPANJAN y

MORRELL, MARTIN JAMES

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 729 624 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

Reducción de la correlación entre canales de fondo ambisónicos de orden superior (HOA)

5 **CAMPO TÉCNICO**

[0001] La presente divulgación se refiere a datos de audio y, más específicamente, a la codificación y descodificación de datos de audio ambisónico de orden superior.

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 el campo sonoro de una manera que sea independiente de la geometría del altavoz local usado para reproducir una señal de audio multicanal reproducida a partir de la señal SHC. La señal SHC también puede facilitar la retrocompatibilidad, ya que la 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 SHC puede, por lo tanto, hacer posible una mejor representación de un campo sonoro que también asimila la retrocompatibilidad.

20 **SUMARIO**

[0003] En general, se describen técnicas para codificar y descodificar datos de audio ambisónico de orden superior. Los datos de audio ambisónico de orden superior pueden comprender al menos un coeficiente ambisónico de orden superior (HOA) correspondiente a una función de base armónica esférica que tiene un orden mayor que uno. Se describen técnicas para reducir la correlación entre canales de fondo ambisónicos de orden superior (HOA).

[0004] En un aspecto, un procedimiento incluye obtener una representación descorrelacionada de coeficientes ambisónicos ambientales que representan al menos una señal izquierda y una señal derecha, habiéndose extraído los coeficientes ambisónicos ambientales de una pluralidad de coeficientes ambisónicos de orden superior y representativos de un componente de fondo de un campo sonoro descrito por la pluralidad de coeficientes ambisónicos de orden superior, habiéndose descorrelacionado la representación descorrelacionada de los coeficientes ambisónicos ambientales usando una transformada basada en fase, en los que al menos uno de la pluralidad de coeficientes ambisónicos de orden superior está asociado con una función de base esférica que tiene un orden de uno o cero; aplicar una transformada de recorrelación a la representación descorrelacionada de los coeficientes ambisónicos ambientales para obtener una pluralidad de coeficientes ambisónicos ambientales correlacionados; y generar una señal de altavoz en base a la pluralidad de coeficientes ambisónicos ambientales correlacionados obtenidos de la representación descorrelacionada de los coeficientes ambisónicos ambientales.

[0005] En otro aspecto, un dispositivo para procesar datos de audio incluye una memoria configurada para almacenar al menos una porción de los datos de audio que se van a procesar; y uno o más procesadores configurados para obtener una representación descorrelacionada de coeficientes ambisónicos ambientales que tienen al menos una señal izquierda y una señal derecha, habiéndose extraído los coeficientes ambisónicos ambientales de una pluralidad de coeficientes ambisónicos de orden superior y representativos de un componente de fondo de un campo sonoro descrito por la pluralidad de coeficientes ambisónicos de orden superior, en los que al menos uno de la pluralidad de coeficientes ambisónicos de orden superior está asociado con una función de base esférica que tiene un orden de uno o cero; aplicar una transformada de recorrelación a la representación descorrelacionada de los coeficientes ambisónicos para obtener una pluralidad de coeficientes ambisónicos ambientales correlacionados; y generar una señal de altavoz en base a la representación descorrelacionada de los coeficientes ambisónicos ambientales.

[0006] En otro aspecto, un dispositivo para comprimir datos de audio incluye una memoria configurada para almacenar al menos una porción de los datos de audio que se van a comprimir; y uno o más procesadores configurados para aplicar una transformada de descorrelación a los coeficientes ambisónicos ambientales para obtener una representación descorrelacionada de los coeficientes ambisónicos ambientales, habiéndose extraído los coeficientes ambisónicos ambientales de una pluralidad de coeficientes ambisónicos de orden superior y representativos de un componente de fondo de un campo sonoro descrito por la pluralidad de coeficientes ambisónicos de orden superior, en los que al menos uno de la pluralidad de coeficientes ambisónicos de orden superior está asociado con una función de base esférica que tiene un orden de uno o cero.

[0007] 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 las técnicas resultarán evidentes a partir de la descripción y de los dibujos y a partir de las reivindicaciones.

65 **BREVE DESCRIPCIÓN DE LOS DIBUJOS**

[0008]

La FIG. 1 es un diagrama que ilustra funciones de base armónica esférica de diversos órdenes y subórdenes.

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

La FIG. 3 es un diagrama de bloques que ilustra, con mayor detalle, un ejemplo del dispositivo de codificación de audio mostrado en el ejemplo de la FIG. 2 que puede realizar diversos aspectos de las técnicas descritas en la presente divulgación.

La FIG. 4 es un diagrama de bloques que ilustra el dispositivo de descodificación de audio de la FIG. 2 con mayor detalle.

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

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

La FIG. 6B es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de codificación de audio y un dispositivo de descodificación de audio en la realización de diversos aspectos de las técnicas de codificación descritas en la presente divulgación.

DESCRIPCIÓN DETALLADA

[0009] 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 para el consumidor son en su mayoría "canales" basados en que, implícitamente, especifican las señales a los altavoces en determinadas coordenadas geométricas. Los formatos de sonido envolvente para el consumidor incluyen el popular formato 5.1 (que incluye los siguientes seis canales: delantero izquierdo (FL), delantero derecho (FR), central o delantero central, posterior izquierdo o envolvente izquierdo, posterior derecho o envolvente derecho, y los efectos de baja frecuencia (LFE)), el formato creciente 7.1, diversos formatos que incluyen altavoces de altura tales como el formato 7.1.4 y el formato 22.2 (por ejemplo, para su uso con la norma de Televisión de Ultra Alta Definición). Los formatos que no son para el consumidor pueden abarcar cualquier número de altavoces (en geometrías simétricas y no simétricas) a menudo denominados "formaciones envolventes". Un ejemplo de una formación de este tipo incluye 32 altavoces situados en coordenadas en las esquinas de un icosaedro truncado.

[0010] La entrada a un futuro codificador de MPEG es opcionalmente uno de tres formatos posibles: (i) audio tradicional basado en canales (como se analiza anteriormente), que está destinado para reproducirse a través de altavoces en posiciones preestablecidas; (ii) audio basado en objetos, que implica datos discretos de modulación por código de pulso (PCM) para objetos de audio individuales con metadatos asociados que contienen sus coordenadas de ubicación (entre otra información); y (iii) audio basado en la escena, que implica la representación del campo sonoro usando coeficientes de funciones de base armónica esférica (también llamados "coeficientes armónicos esféricos" o SHC, "ambisónicos de orden superior" o HOA y "coeficientes HOA"). El futuro codificador de MPEG se puede describir con mayor detalle en un documento titulado "[Call for Proposals for 3D Audio] Convocatoria de propuestas para audio 3D", por la Organización Internacional de Normalización/Comisión Electrotécnica Internacional (ISO)/(IEC) JTC1/SC29/WG11/N13411, publicado en enero de 2013 en Ginebra, Suiza, y disponible en <http://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/w13411.zip>.

[0011] Existen diversos formatos basados en canales con "sonido envolvente" en el mercado. Varían, por ejemplo, desde el sistema de cine en casa 5.1 (que es el que ha tenido el mayor éxito en cuanto a irrumpir en las salas de estar, más allá del estéreo) hasta el sistema 22.2 desarrollado por NHK (Nippon Hoso Kyokai o Corporación Difusora de Japón). Los creadores de contenido (por ejemplo, los estudios de Hollywood) desearían producir la banda sonora de una película una vez y no dedicar esfuerzo en remezclarla para cada configuración de altavoces. Recientemente, las Organizaciones de Desarrollo de Normas han estado considerando maneras de proporcionar una codificación en un flujo de bits estandarizado y una descodificación subsiguiente que sea adaptable e indiferente con respecto a la geometría (y número) de los altavoces y a las condiciones acústicas en la ubicación de la reproducción (que implica un reproductor).

[0012] Para proporcionar dicha flexibilidad a los creadores de contenido, se puede usar un conjunto jerárquico de elementos para representar un campo sonoro. El conjunto jerárquico de elementos puede hacer referencia a un conjunto de elementos en el que los elementos están ordenados 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 torna más detallada, aumentando la resolución.

[0013] Un ejemplo de un conjunto jerárquico de elementos es un conjunto de coeficientes armónicos esféricos (SHC). La siguiente expresión demuestra una descripción o representación de un campo sonoro usando SHC:

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

5 **[0014]** La expresión muestra que la presión p_i en cualquier punto $\{r_r, \theta_r, \phi_r\}$ del campo sonoro, en el tiempo t , se puede representar de forma única por los SHC,

$A_n^m(k)$. Aquí, $k = \frac{\omega}{c}$, c es la velocidad del sonido (~ 343 m/s), $\{r_r, \theta_r, \phi_r\}$ es un punto de referencia (o punto de observación), $j_n(\cdot)$ es la función de Bessel esférica de orden n , y $Y_n^m(\theta_r, \varphi_r)$ 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. Las señales de ambisónicos de orden superior se procesan truncando los órdenes superiores para que solo queden el primer orden y el cero. Normalmente se hace una compensación de energía de las señales restantes debido a la pérdida de energía en el coeficiente de orden superior.

20 **[0015]** Diversos aspectos de la presente divulgación están dirigidos a reducir la correlación entre las señales de fondo. Por ejemplo, las técnicas de la presente divulgación pueden reducir o posiblemente eliminar la correlación entre las señales de fondo expresadas en el dominio HOA. Una ventaja potencial de reducir la correlación entre las señales de HOA de fondo es la mitigación del desenmascaramiento de ruido. Como se usa en el presente documento, la expresión "desenmascaramiento de ruido" puede hacer referencia a la atribución de objetos de audio a ubicaciones que no corresponden al objeto de audio en el dominio espacial. Además de mitigar los problemas potenciales relacionados con el desenmascaramiento de ruido, las técnicas de codificación descritas en el presente documento pueden generar señales de salida que representan señales de audio izquierda y derecha, tales como señales que juntas forman una salida en estéreo. A su vez, un dispositivo de descodificación puede descodificar las señales de audio izquierda y derecha para obtener una salida en estéreo, o puede mezclar las señales izquierda y derecha para obtener una salida mono. Adicionalmente, en escenarios donde un flujo de bits codificado representa una disposición puramente horizontal, un dispositivo de descodificación puede implementar diversas técnicas de la presente divulgación para descodificar solo señales de fondo de HOA descorrelacionadas de componentes horizontales. Al limitar el proceso de descodificación a las señales de fondo de HOA descorrelacionadas de componentes horizontales, el descodificador puede implementar las técnicas para conservar los recursos informáticos y reducir el consumo de ancho de banda.

35 **[0016]** 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 ver, para cada orden, existe una expansión de los subórdenes m que se muestran pero no se señalan explícitamente en el ejemplo de la FIG. 1, para propósitos de facilidad de ilustración.

40 **[0017]** Los SHC $A_n^m(k)$ se pueden adquirir físicamente (por ejemplo, grabarse) mediante diversas configuraciones de formaciones de micrófonos o bien, de forma alternativa, pueden provenir de descripciones basadas en canales o basadas en objetos del campo sonoro. Los SHC representan el audio basado en la escena, donde los SHC se pueden introducir en un codificador de audio para obtener SHC codificados que pueden fomentar una transmisión o almacenamiento más eficaz. Por ejemplo, se puede usar una representación de cuarto orden que implique coeficientes $(1+4)^2$ (25 y, de ahí, de cuarto orden).

50 **[0018]** Como se indica anteriormente, los SHC pueden provenir de una grabación con micrófono usando una formación de micrófonos. En el artículo de Poletti, M., "Three-Dimensional Surround Sound Systems Based on Spherical Harmonics [Sistemas tridimensionales de sonido envolvente basados en armónicos esféricos]", J. Audio Eng. Soc., vol. 53, n.º 11, noviembre de 2005, pp. 1004-1025, se describen diversos ejemplos de cómo se pueden obtener los SHC de formaciones de micrófonos.

55 **[0019]** Para ilustrar cómo los SHC se pueden obtener 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:

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

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

5 que lo anterior es una descomposición lineal y ortogonal) que los coeficientes $A_n^m(k)$ para cada objeto son aditivos. De esta manera, una multitud de objetos de PCM 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, los coeficientes contienen información sobre el campo sonoro (la presión como una función de coordenadas 3D), y lo anterior representa la transformación de los objetos individuales a una representación del campo sonoro global en las

10 proximidades del punto de observación $\{r_r, \theta_r, \phi_r\}$. Las cifras restantes se describen a continuación en el contexto de la codificación de audio basada en objetos y basada en SHC.

[0020] La FIG. 2 es un diagrama que ilustra un sistema 10 que puede realizar diversos aspectos de las técnicas descritas en la presente divulgación. Como se muestra en el ejemplo de la FIG. 2, el sistema 10 incluye un dispositivo creador de contenido 12 y un dispositivo consumidor de contenido 14. Si bien se describen en el contexto del dispositivo creador de contenido 12 y del dispositivo consumidor de contenido 14, las técnicas se pueden implementar en cualquier contexto en el que los SHC (que también se pueden denominar coeficientes HOA), o cualquier otra representación jerárquica de un campo sonoro, se codifican para formar un flujo de bits representativo de los datos de audio. Además, el dispositivo creador de contenido 12 puede representar cualquier forma de dispositivo informático que pueda implementar las técnicas descritas en la presente divulgación, incluyendo un equipo manual (o teléfono celular), un ordenador de tableta, un teléfono inteligente o un ordenador de escritorio, para proporcionar unos pocos ejemplos. Asimismo, el dispositivo consumidor de contenido 14 puede representar cualquier forma de dispositivo informático que pueda implementar las técnicas descritas en la presente divulgación, incluyendo un equipo manual (o teléfono celular), un ordenador de tableta, un teléfono inteligente, un descodificador o un ordenador de escritorio, para proporcionar unos pocos ejemplos.

[0021] El dispositivo creador de contenido 12 se puede hacer funcionar por un estudio de cine u otra entidad que pueda generar contenido de audio multicanal para el consumo por operadores de dispositivos consumidores de contenido, tales como el dispositivo consumidor de contenido 14. En algunos ejemplos, el dispositivo creador de contenido 12 se puede hacer funcionar por un usuario individual que desea comprimir los coeficientes HOA 11. A menudo, el creador de contenido genera contenido de audio junto con el contenido de vídeo. El dispositivo consumidor de contenido 14 se puede hacer funcionar por un individuo. El dispositivo consumidor de contenido 14 puede incluir un sistema de reproducción de audio 16, que puede hacer referencia a cualquier forma de sistema de reproducción de audio que pueda reproducir los SHC para su reproducción como contenido de audio multicanal.

[0022] El dispositivo creador de contenido 12 incluye un sistema de edición de audio 18. El dispositivo creador de contenido 12 obtiene grabaciones en vivo 7 en diversos formatos (incluyendo directamente como coeficientes HOA) y objetos de audio 9, que el dispositivo creador de contenido 12 puede editar usando el sistema de edición de audio 18. Un micrófono 5 puede capturar las grabaciones en vivo 7. El creador de contenido puede, durante el proceso de edición, reproducir los coeficientes HOA 11 de los objetos de audio 9, escuchando las señales de los altavoces reproducidas en un intento de identificar diversos aspectos del campo sonoro que requieran una edición adicional. El dispositivo creador de contenido 12 puede a continuación editar los coeficientes HOA 11 (potencialmente de manera indirecta a través de la manipulación de diferentes de los objetos de audio 9 a partir de los cuales se pueden obtener los coeficientes HOA fuente, de la manera descrita anteriormente). El dispositivo creador de contenido 12 puede emplear el sistema de edición de audio 18 para generar los coeficientes HOA 11. El sistema de edición de audio 18 representa cualquier sistema que pueda editar datos de audio y emitir los datos de audio como uno o más coeficientes armónicos esféricos fuente.

[0023] Cuando se completa el proceso de edición, el dispositivo creador de contenido 12 puede generar un flujo de bits 21 en base a los coeficientes HOA 11. Es decir, el dispositivo creador de contenido 12 incluye un dispositivo de codificación de audio 20 que representa un dispositivo configurado para codificar o comprimir de otro modo los coeficientes HOA 11 de acuerdo con diversos aspectos de las técnicas descritas en la presente divulgación para generar el flujo de bits 21. El dispositivo de codificación de audio 20 puede generar el flujo de bits 21 para su transmisión, como un ejemplo, a través de un canal de transmisión, que puede ser un canal cableado o inalámbrico, un dispositivo de almacenamiento de datos o similar. El flujo de bits 21 puede representar una versión codificada de los coeficientes HOA 11 y puede incluir un flujo de bits primario y otro flujo de bits lateral, que se puede denominar información de canal lateral.

[0024] Si bien se muestra en la FIG. 2 como transmitido directamente al dispositivo consumidor de contenido 14, el dispositivo creador de contenido 12 puede emitir el flujo de bits 21 a un dispositivo intermedio situado entre el dispositivo creador de contenido 12 y el dispositivo consumidor de contenido 14. Este dispositivo intermedio puede almacenar el flujo de bits 21 para su posterior suministro al dispositivo consumidor de contenido 14, que puede solicitar el flujo de bits. El dispositivo intermedio puede comprender un servidor de archivos, un servidor de la Red, un

ordenador de escritorio, un ordenador portátil, un ordenador de tableta, un teléfono móvil, un teléfono inteligente o cualquier otro dispositivo que pueda almacenar el flujo de bits 21 para su posterior recuperación mediante un decodificador de audio. Este dispositivo intermedio puede residir en una red de suministro de contenido, que puede transmitir en flujo continuo el flujo de bits 21 (y posiblemente junto con la transmisión de un correspondiente flujo de bits de datos de vídeo) a los abonados, tales como el dispositivo consumidor de contenido 14 que solicita el flujo de bits 21.

[0025] De forma alternativa, el dispositivo creador de contenido 12 puede almacenar el flujo de bits 21 en un medio de almacenamiento, tal como un disco compacto, un disco de vídeo digital, un disco de vídeo de alta definición u otros medios de almacenamiento, la mayoría de los cuales se pueden leer mediante un ordenador y, por lo tanto, se pueden denominar medios de almacenamiento legibles por ordenador o medios de almacenamiento no transitorio legibles por ordenador. En este contexto, el canal de transmisión puede hacer referencia a los canales por los cuales se transmiten contenidos almacenados en los medios (y pueden incluir tiendas minoristas y otros mecanismos de suministro basados en tiendas). En cualquier caso, las técnicas de la presente divulgación no se deberían, por lo tanto, limitar, a este respecto, al ejemplo de la FIG. 2.

[0026] Como se muestra además en el ejemplo de la FIG. 2, el dispositivo consumidor de contenido 14 incluye el sistema de reproducción de audio 16. El sistema de reproducción de audio 16 puede representar cualquier sistema de reproducción de audio que pueda reproducir datos de audio multicanal. El sistema de reproducción de audio 16 puede incluir un número de diferentes reproductores 22. Los reproductores 22 pueden proporcionar cada uno una forma diferente de reproducción 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. Como se usa en el presente documento, "A y/o B" significa "A o B" o ambos "A y B".

[0027] El sistema de reproducción de audio 16 puede incluir además un dispositivo de decodificación de audio 24. El dispositivo de decodificación de audio 24 puede representar un dispositivo configurado para decodificar los coeficientes HOA 11' desde el flujo de bits 21, donde los coeficientes HOA 11' pueden ser similares a los coeficientes HOA 11, pero difieren debido a operaciones con pérdidas (por ejemplo, cuantificación) y/o transmisión por medio el canal de transmisión. El sistema de reproducción de audio 16 puede, después de decodificar el flujo de bits 21, obtener los coeficientes HOA 11' y reproducir los coeficientes HOA 11' para emitir las señales de los altavoces 25. Las señales de los altavoces 25 pueden accionar uno o más altavoces (que no se muestran en el ejemplo de la FIG. 2, para propósitos de facilidad de ilustración).

[0028] Para seleccionar el reproductor apropiado o, en algunos casos, generar un reproductor apropiado, el sistema de reproducción de audio 16 puede obtener información de los altavoces 13, indicativa de un número de altavoces y/o una geometría espacial de los altavoces. En algunos casos, el sistema de reproducción de audio 16 puede obtener la información de los altavoces 13 usando un micrófono de referencia y accionando los altavoces de tal manera como para determinar dinámicamente la información de los altavoces 13. En otros casos, o junto con la determinación dinámica de la información de los altavoces 13, el sistema de reproducción de audio 16 puede inducir a un usuario a interactuar con el sistema de reproducción de audio 16 e introducir la información de los altavoces 13.

[0029] El sistema de reproducción de audio 16 puede a continuación seleccionar uno de los reproductores de audio 22 en base a la información de los altavoces 13. En algunos casos, el sistema de reproducción de audio 16 puede, cuando ninguno de los reproductores de audio 22 está dentro de ninguna medición de similitud de umbral (en cuanto a geometría de altavoces) con la geometría de los altavoces especificada en la información de los altavoces 13, generar el uno de los reproductores de audio 22 en base a la información de los altavoces 13. El sistema de reproducción de audio 16 puede, en algunos casos, generar uno de los reproductores de audio 22 en base a la información de los altavoces 13 sin intentar primero seleccionar uno existente de los reproductores de audio 22. Uno o más altavoces 3 pueden a continuación reproducir las señales de los altavoces reproducidas 25.

[0030] La FIG. 3 es un diagrama de bloques que ilustra, con mayor detalle, un ejemplo del dispositivo de codificación de audio 20 mostrado en el ejemplo de la FIG. 2 que puede realizar diversos aspectos de las técnicas descritas en la presente divulgación. El dispositivo de codificación de audio 20 incluye una unidad de análisis de contenido 26, una unidad de metodología de síntesis basada en vectores 27, una unidad de metodología de síntesis basada en la dirección 28 y una unidad de descorrelación 40'. Aunque se describe brevemente a continuación, más información con respecto al dispositivo de codificación de audio 20 y los diversos aspectos de compresión o de otro modo codificación de los coeficientes HOA está disponible en la publicación de solicitud de patente internacional n.º WO 2014/194099, titulada "INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD [INTERPOLACIÓN PARA REPRESENTACIONES DESCOMPUESTAS DE UN CAMPO SONORO]", presentada el 29 de mayo de 2014.

[0031] La unidad de análisis de contenido 26 representa una unidad configurada para analizar el contenido de los coeficientes HOA 11, para identificar si los coeficientes HOA 11 representan contenido generado a partir de una grabación en vivo o un objeto de audio. La unidad de análisis de contenido 26 puede determinar si los coeficientes HOA 11 se generaron a partir de una grabación de un campo sonoro real o de un objeto de audio artificial. En algunos casos, cuando los coeficientes HOA en tramas 11 se generaron a partir de una grabación, la unidad de análisis de

contenido 26 pasa los coeficientes HOA 11 a la unidad de descomposición basada en vectores 27. En algunos casos, cuando los coeficientes HOA en tramas 11 se generaron a partir de un objeto de audio sintético, la unidad de análisis de contenido 26 pasa los coeficientes HOA 11 a la unidad de síntesis basada en la dirección 28. La unidad de síntesis basada en la dirección 28 puede representar una unidad configurada para realizar una síntesis basada en la dirección de los coeficientes HOA 11 para generar un flujo de bits basado en la dirección 21.

[0032] Como se muestra en el ejemplo de la FIG. 3, la unidad de descomposición basada en vectores 27 puede incluir una unidad de transformada invertible lineal (LIT) 30, una unidad de cálculo de parámetros 32, una unidad de reordenamiento 34, una unidad de selección de primer plano 36, una unidad de compensación de energía 38, una unidad codificadora de audio psicoacústico 40, una unidad de generación de flujo de bits 42, una unidad de análisis de campo sonoro 44, una unidad de reducción de coeficientes 46, una unidad de selección de fondo (BG) 48, una unidad de interpolación espacio-temporal 50 y una unidad de cuantificación 52.

[0033] La unidad de transformada invertible lineal (LIT) 30 recibe los coeficientes HOA 11 en forma de canales HOA, siendo cada canal representativo de un bloque o trama de un coeficiente asociado con un orden dado, suborden de las funciones de base esférica (que se puede indicar como $HOA[k]$, donde k puede indicar la trama o bloque actual de muestras). La matriz de coeficientes HOA 11 puede tener las dimensiones $D: M \times (N+1)^2$.

[0034] La unidad de LIT 30 puede representar una unidad configurada para realizar una forma de análisis denominada descomposición en valores singulares. Si bien se describen con respecto a la SVD, las técnicas descritas en la presente divulgación se pueden realizar con respecto a cualquier transformación o descomposición similar que proporcione conjuntos de salida compactada de energía, no correlacionada linealmente. También, la referencia a "conjuntos" en la presente divulgación en general pretende hacer referencia a conjuntos no nulos, a menos que se exprese específicamente lo contrario y no pretende hacer referencia a la definición matemática clásica de conjuntos que incluye el denominado "conjunto vacío". Una transformación alternativa puede comprender un análisis de componentes principales, que se denomina a menudo "PCA". Dependiendo del contexto, el PCA se puede denominar mediante un número de nombres diferentes, tales como transformada discreta de Karhunen-Loeve, la transformada de Hotelling, descomposición ortogonal propia (POD) y descomposición en autovalores (EVD), para nombrar unos pocos ejemplos. Las propiedades de dichas operaciones que conducen a la meta subyacente de compresión de datos de audio son la "compactación energética" y la "descorrelación" de los datos de audio multicanal.

[0035] En cualquier caso, suponiendo que la unidad de LIT 30 realice una descomposición en valores singulares (que, de nuevo, se puede denominar "SVD") para propósitos de ejemplo, la unidad de LIT 30 puede transformar los coeficientes HOA 11 en dos o más conjuntos de coeficientes HOA transformados. Los "conjuntos" de coeficientes HOA transformados pueden incluir vectores de coeficientes HOA transformados. En el ejemplo de la FIG. 3, la unidad de LIT 30 puede realizar la SVD con respecto a los coeficientes HOA 11 para generar una denominada matriz V , una matriz S y una matriz U . La SVD, en álgebra lineal, puede representar una factorización de una matriz X real o compleja de y por z (donde X puede representar datos de audio multicanal, tales como los coeficientes HOA 11) de la siguiente forma:

$$X = USV^*$$

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

[0036] En algunos ejemplos, la matriz V^* en la expresión matemática de SVD a la que se hace referencia anteriormente se indica como la traspuesta conjugada de la matriz V para reflejar que la SVD se puede aplicar a matrices que comprenden números complejos. Cuando se aplica a matrices que comprenden solo números reales, la conjugada compleja de la matriz V (o, en otras palabras, la matriz V^*) se puede considerar que es la traspuesta de la matriz V . A continuación se supone, para propósitos de facilidad de ilustración, que los coeficientes HOA 11 comprenden números reales, con el resultado de que la matriz V se emite a través de la SVD, en lugar de la matriz V^* . Además, si bien se indica como la matriz V en la presente divulgación, la referencia a la matriz V se debería entender para hacer referencia a la traspuesta de la matriz V cuando corresponda. Si bien se supone que es la matriz V , las técnicas se pueden aplicar de forma similar a los coeficientes HOA 11 que tienen coeficientes complejos, donde la salida de la SVD es la matriz V^* . En consecuencia, las técnicas no se deberían limitar a este respecto para proporcionar solo la aplicación de la SVD para generar una matriz V , sino que pueden incluir la aplicación de la SVD a coeficientes HOA 11 que tengan componentes complejos para generar una matriz V^* .

[0037] De esta manera, la unidad de LIT 30 puede realizar una SVD con respecto a los coeficientes HOA 11 para emitir los vectores $US[k]$ 33 (que pueden representar una versión combinada de los vectores S y los vectores U) que tienen las dimensiones $D: M \times (N+1)^2$, y los vectores $V[k]$ 35 que tienen las dimensiones $D: (N+1)^2 \times (N+1)^2$. Los

elementos vectoriales individuales en la matriz $US[k]$ también se pueden denominar $X_{PS}(k)$, mientras que los vectores individuales de la matriz $V[k]$ también se pueden denominar $v(k)$.

5 **[0038]** Un análisis de las matrices U, S y V puede revelar que las matrices llevan o representan características espaciales y temporales del campo sonoro subyacente representado anteriormente por X. Cada uno de los N vectores en U (de muestras de longitud M) puede representar señales de audio separadas normalizadas como una función del tiempo (para el período de tiempo representado por las M muestras), que son ortogonales entre sí y que se han desacoplado de cualquier característica espacial (que también se puede denominar información direccional). Las características espaciales, que representan la forma y posición espaciales (r, theta, phi), se pueden representar en cambio por vectores ^{ésimos} individuales, $v^{\theta}(k)$, en la matriz V (cada uno de longitud $(N+1)^2$). Los elementos individuales de cada uno de los vectores $v^{\theta}(k)$ pueden representar un coeficiente HOA que describe la forma (incluyendo la anchura) y posición del campo sonoro para un objeto de audio asociado. Tanto los vectores de la matriz U como los de la matriz V se normalizan de modo que sus energías de media cuadrática son iguales a la unidad. La energía de las señales de audio en U, por tanto, se representa por los elementos diagonales en S. La multiplicación de U por S para formar $US[k]$ (con los elementos vectoriales individuales $X_{PS}(k)$), por tanto, representa la señal de audio con energías. La capacidad de la descomposición SVD para desacoplar las señales de tiempo de audio (en U), sus energías (en S) y sus características espaciales (en V) puede admitir diversos aspectos de las técnicas descritas en la presente divulgación. Además, el modelo de síntesis de los coeficientes HOA[k] subyacentes, X, por una multiplicación de vectores de $US[k]$ y $V[k]$ da lugar al término "descomposición basada en vectores" que se usa a lo largo de todo el presente documento.

25 **[0039]** Aunque se describe como realizada directamente con respecto a los coeficientes HOA 11, la unidad de LIT 30 puede aplicar la transformada invertible lineal a los derivados de los coeficientes HOA 11. Por ejemplo, la unidad de LIT 30 puede aplicar una SVD con respecto a una matriz de densidad espectral de potencia que proviene de los coeficientes HOA 11. Al realizar la SVD con respecto a la densidad espectral de potencia (PSD) de los coeficientes HOA, en lugar de los coeficientes en sí, la unidad de LIT 30 puede reducir potencialmente la complejidad de cálculo para realizar la SVD en términos de uno o más ciclos de procesador y espacio de almacenamiento, mientras que logra la misma eficacia de codificación de audio fuente que si la SVD se aplicara directamente a los coeficientes HOA.

30 **[0040]** La unidad de cálculo de parámetros 32 representa una unidad configurada para calcular diversos parámetros, tales como un parámetro de correlación (R), parámetros de propiedades direccionales (θ , ϕ , r), y una propiedad de energía (e). Cada uno de los parámetros para la trama actual se puede indicar como $R[k]$, $\theta[k]$, $\phi[k]$, $r[k]$ y $e[k]$. La unidad de cálculo de parámetros 32 puede realizar un análisis de energía y/o correlación (o la denominada correlación cruzada) con respecto a los vectores $US[k]$ 33 para identificar los parámetros. La unidad de cálculo de parámetros 32 también puede determinar los parámetros para la trama previa, donde los parámetros de la trama previa se pueden indicar como $R[k-1]$, $\theta[k-1]$, $\phi[k-1]$, $r[k-1]$ y $e[k-1]$, en base a la trama previa del vector $US[k-1]$ y los vectores $V[k-1]$. La unidad de cálculo de parámetros 32 puede emitir los parámetros actuales 37 y los parámetros previos 39 a la unidad de reordenamiento 34.

40 **[0041]** Los parámetros calculados por la unidad de cálculo de parámetros 32 se pueden usar por la unidad de reordenamiento 34 para reordenar los objetos de audio para representar su evaluación natural o su continuidad en el tiempo. La unidad de reordenamiento 34 puede comparar cada uno de los parámetros 37 de los primeros vectores $US[k]$ 33, por turnos, con cada uno de los parámetros 39 para los segundos vectores $US[k-1]$ 33. La unidad de reordenamiento 34 puede reordenar (usando, como ejemplo, un algoritmo húngaro) los diversos vectores dentro de la matriz $US[k]$ 33 y la matriz $V[k]$ 35 en base a los parámetros actuales 37 y los parámetros previos 39 para emitir una matriz $US[k]$ reordenada 33' (que se puede indicar matemáticamente como $US[k]$) y una matriz $V[k]$ reordenada 35' (que se puede indicar matemáticamente como $V[k]$) a una unidad de selección de sonido de primer plano 36 (o sonido predominante - PS) ("unidad de selección de primer plano 36") y una unidad de compensación de energía 38.

50 **[0042]** La unidad de análisis de campo sonoro 44 puede representar una unidad configurada para realizar un análisis de campo sonoro con respecto a los coeficientes HOA 11 para lograr potencialmente una velocidad de bits objetivo 41. La unidad de análisis de campo sonoro 44 puede, en base al análisis y/o a una velocidad de bits objetivo 41 recibida, determinar el número total de instanciaciones de codificador psicoacústico (que puede ser una función del número total de canales ambientales o de fondo (BG_{TOT}) y el número de canales de primer plano o, en otras palabras, canales predominantes. El número total de instanciaciones de codificador psicoacústico se puede indicar como numHOATransportChannels.

60 **[0043]** La unidad de análisis de campo sonoro 44 también puede determinar, de nuevo para lograr potencialmente la velocidad de bits objetivo 41, el número total de canales de primer plano (nFG) 45, el orden mínimo del campo sonoro de fondo (o, en otras palabras, ambiental) (N_{BG} o, de forma alternativa, MinAmbHOAorder), el número correspondiente de canales reales representativos del orden mínimo de campo sonoro de fondo ($nBGa = (\text{MinAmbHOAorder} + 1)^2$), y los índices (i) de los canales HOA de BG adicionales para enviar (que se pueden indicar conjuntamente como información de canal de fondo 43 en el ejemplo de la FIG. 3). La información de canal de fondo 42 también se puede denominar información de canal ambiental 43. Cada uno de los canales que queda de numHOATransportChannels-nBGa puede ser un "canal de fondo/ambiental adicional", un "canal predominante basado en un vector activo", una "señal predominante basada en la dirección activa" o bien "completamente inactivo". En un aspecto, los tipos de

canales se pueden indicar como elemento sintáctico (como un "ChannelType") con dos bits (por ejemplo, 00: señal basada en la dirección; 01: señal predominante basada en vectores; 10: señal ambiental adicional; 11: señal inactiva). El número total de señales de fondo o ambientales, nBGa, puede estar dado por $(\text{MinAmbHOAorder} + 1)^2$ + el número de veces que el índice 10 (en el ejemplo anterior) aparece como un tipo de canal en el flujo de bits para esa trama.

[0044] La unidad de análisis de campo sonoro 44 puede seleccionar el número de canales de fondo (o, en otras palabras, ambientales) y el número de canales de primer plano (o, en otras palabras, predominantes) en base a la velocidad de bits objetivo 41, seleccionando más canales de fondo y/o de primer plano cuando la velocidad de bits objetivo 41 es relativamente mayor (por ejemplo, cuando la velocidad de bits objetivo 41 es igual o mayor que 512 Kbps). En un aspecto, el numHOATransportChannels se puede fijar en 8 mientras que el MinAmbHOAorder se puede fijar en 1 en la sección de cabecera del flujo de bits. En este escenario, en cada trama, se pueden dedicar cuatro canales para representar la porción de fondo o ambiental del campo sonoro, mientras que los otros 4 canales pueden variar, de trama a trama, según el tipo de canal, por ejemplo, usado como un canal de fondo/ambiental adicional o bien un canal de primer plano/predominante. Las señales de primer plano/predominantes pueden ser una de señales basadas en vectores o bien basadas en la dirección, como se describe anteriormente.

[0045] En algunos casos, el número total de señales predominantes basadas en vectores para una trama puede estar dado por el número de veces que el índice ChannelType es 01 en el flujo de bits de esa trama. En el aspecto anterior, para cada canal de fondo/ambiental adicional (por ejemplo, correspondiente a un ChannelType de 10), se puede representar en ese canal la información correspondiente de cuál de los posibles coeficientes HOA (más allá de los cuatro primeros). La información, para contenido HOA de cuarto orden, puede ser un índice para indicar los coeficientes HOA 5-25. Los cuatro primeros coeficientes HOA ambientales 1-4 se pueden enviar todo el tiempo cuando minAmbHOAorder se fija en 1, de ahí que el dispositivo de codificación de audio solo pueda necesitar indicar uno de los coeficientes HOA ambientales adicionales que tengan un índice de 5-25. La información, por tanto, se podría enviar usando un elemento sintáctico de 5 bits (para el contenido de 4.º orden), que se puede indicar como "CodedAmbCoeffIdx". En cualquier caso, la unidad de análisis de campo sonoro 44 emite la información de canal de fondo 43 y los coeficientes HOA 11 a la unidad de selección de fondo (BG) 36, la información de canal de fondo 43 a la unidad de reducción de coeficientes 46 y a la unidad de generación de flujo de bits 42, y el nFG 45 a una unidad de selección de primer plano 36.

[0046] La unidad de selección de fondo 48 puede representar una unidad configurada para determinar los coeficientes HOA de fondo o ambientales 47 en base a la información de canal de fondo (por ejemplo, el campo sonoro de fondo (N_{BG}) y el número (nBGa) y los índices (i) de los canales HOA de BG adicionales a enviar). Por ejemplo, cuando N_{BG} es igual a uno, la unidad de selección de fondo 48 puede seleccionar los coeficientes HOA 11 para cada muestra de la trama de audio que tenga un orden igual o inferior a uno. La unidad de selección de fondo 48 puede, en este ejemplo, seleccionar a continuación los coeficientes HOA 11 que tengan un índice identificado por uno de los índices (i) como coeficientes HOA de BG adicionales, donde el nBGa se proporcione a la unidad de generación de flujo de bits 42 para que se especifique en el flujo de bits 21 para hacer posible que el dispositivo de decodificación de audio, tal como el dispositivo de decodificación de audio 24, mostrado en el ejemplo de las FIGS. 2 y 4, analice sintácticamente los coeficientes HOA de fondo 47 del flujo de bits 21. La unidad de selección de fondo 48 puede emitir a continuación los coeficientes HOA ambientales 47 a la unidad de compensación de energía 38. Los coeficientes HOA ambientales 47 pueden tener las dimensiones D: $M \times [(N_{BG}+1)^2 + nBGa]$. Los coeficientes HOA ambientales 47 también se pueden denominar "coeficientes HOA ambientales 47", donde cada uno de los coeficientes HOA ambientales 47 corresponde a un canal HOA ambiental independiente 47, que se va a codificar por la unidad codificadora de audio psicoacústico 40.

[0047] La unidad de selección de primer plano 36 puede representar una unidad configurada para seleccionar la matriz $US[k]$ reordenada 33' y la matriz $V[k]$ reordenada 35', que representan componentes de primer plano, o distintos, del campo sonoro en base a nFG 45 (que pueden representar uno o más índices que identifican los vectores de primer plano). La unidad de selección de primer plano 36 puede emitir señales de nFG 49 (que se pueden indicar como una $US[k]_i$ reordenada, ..., nFG 49, FG_1 , ..., nFG[k] 49, o

$$X_{PS}^{(1..nFG)}(k)$$

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

[0048] La unidad de compensación de energía 38 puede representar una unidad configurada para realizar compensación de energía con respecto a los coeficientes HOA ambientales 47, para compensar la pérdida de energía debido a la eliminación de unos diversos de los canales HOA por la unidad de selección de fondo 48. La unidad de compensación de energía 38 puede realizar un análisis de energía con respecto a una o más de la matriz $US[k]$

reordenada 33', la matriz $V[k]$ reordenada 35', las señales de nFG 49, los vectores $V[k]$ de primer plano 51_k y los coeficientes HOA ambientales 47, y a continuación realizar una compensación de energía en base al análisis de energía para generar los coeficientes HOA ambientales compensados en energía 47'. La unidad de compensación de energía 38 puede emitir los coeficientes HOA ambientales compensados en energía 47' a la unidad de descorrelación 40'. A su vez, la unidad de descorrelación 40' puede implementar las técnicas de la presente divulgación para reducir o eliminar la correlación entre las señales de fondo de los coeficientes HOA 47' para formar uno o más coeficientes HOA descorrelacionados 47". La unidad de descorrelación 40' puede emitir los coeficientes HOA descorrelacionados 47" a la unidad codificadora de audio psicoacústico 40.

[0049] La unidad de interpolación espacio-temporal 50 puede representar una unidad configurada para recibir los vectores $V[k]$ de primer plano 51_k para la $k^{\text{ésima}}$ trama y los vectores $V[k-1]$ de primer plano 51_{k-1} para la trama previa (de ahí la notación $k-1$) y realizar una interpolación espacio-temporal para generar los vectores $V[k]$ de primer plano interpolados. La unidad de interpolación espacio-temporal 50 puede recombinar las señales de nFG 49 con los vectores $V[k]$ de primer plano 51_k para recuperar los coeficientes HOA de primer plano reordenados. La unidad de interpolación espacio-temporal 50 puede a continuación dividir los coeficientes HOA de primer plano reordenados por los vectores $V[k]$ interpolados para generar señales de nFG interpoladas 49'. La unidad de interpolación espacio-temporal 50 también puede emitir los vectores $V[k]$ de primer plano 51_k que se usaron para generar los vectores $V[k]$ de primer plano interpolados, para que un dispositivo de descodificación de audio, tal como el dispositivo de descodificación de audio 24, pueda generar los vectores $V[k]$ de primer plano interpolados y recuperar de este modo los vectores $V[k]$ de primer plano 51_k . Los vectores $V[k]$ de primer plano 51_k usados para generar los vectores $V[k]$ de primer plano interpolados se indican como los vectores $V[k]$ de primer plano restantes 53. Para asegurar que los mismos $V[k]$ y $V[k-1]$ se usen en el codificador y el descodificador (para crear los vectores $V[k]$ interpolados) se pueden usar versiones cuantificadas/descuantificadas de los vectores en el codificador y el descodificador. La unidad de interpolación espacio-temporal 50 puede emitir las señales de nFG interpoladas 49' a la unidad codificadora de audio psicoacústico 46 y los vectores $V[k]$ de primer plano interpolados 51_k a la unidad de reducción de coeficientes 46.

[0050] La unidad de reducción de coeficientes 46 puede representar una unidad configurada para realizar una reducción de coeficientes con respecto a los restantes vectores $V[k]$ de primer plano 53, en base a la información de canal de fondo 43 para emitir los vectores $V[k]$ de primer plano reducidos 55 a la unidad de cuantificación 52. Los vectores $V[k]$ de primer plano reducidos 55 pueden tener las dimensiones D: $[(N+1)^2 - (N_{BG}+1)^2 - BG_{TOT}] \times nFG$. La unidad de reducción de coeficientes 46 puede, a este respecto, representar una unidad configurada para reducir el número de coeficientes de los restantes vectores $V[k]$ de primer plano 53. En otras palabras, la unidad de reducción de coeficientes 46 puede representar una unidad configurada para eliminar los coeficientes en los vectores $V[k]$ de primer plano (que forman los restantes vectores $V[k]$ de primer plano 53) que tienen poca a ninguna información direccional. En algunos ejemplos, los coeficientes de los vectores $V[k]$ distintos o, en otras palabras, de primer plano correspondientes a funciones de base de primer orden y de orden cero (que se pueden indicar como N_{BG}) proporcionan poca información direccional y por lo tanto se pueden eliminar de los vectores V de primer plano (a través de un proceso que se puede denominar "reducción de coeficientes"). En este ejemplo, se puede proporcionar mayor flexibilidad para identificar no solo los coeficientes que corresponden a N_{BG} , sino para identificar canales HOA adicionales (que se pueden indicar con la variable $TotalOfAddAmbHOAChan$) a partir del conjunto de $[(N_{BG} + 1)^2 + 1, (N+1)^2]$.

[0051] La unidad de cuantificación 52 puede representar una unidad configurada para realizar cualquier forma de cuantificación para comprimir los vectores $V[k]$ de primer plano reducidos 55 para generar vectores $V[k]$ de primer plano codificados 57, emitiendo los vectores $V[k]$ de primer plano codificados 57 a la unidad de generación de flujo de bits 42. En funcionamiento, la unidad de cuantificación 52 puede representar una unidad configurada para comprimir un componente espacial del campo sonoro, es decir, uno o más de los vectores $V[k]$ de primer plano reducidos 55 en este ejemplo. La unidad de cuantificación 52 puede realizar uno cualquiera de los siguientes 12 modos de cuantificación, como se indica mediante un elemento sintáctico de modo de cuantificación indicado como "NbitsQ":

Valor de NbitsQ	Tipo de modo de cuantificación
0-3:	Reservado
4:	Cuantificación de vector
5:	Cuantificación escalar sin codificación Huffman
6:	Cuantificación escalar de 6 bits con codificación Huffman
7:	Cuantificación escalar de 7 bits con codificación Huffman
8:	Cuantificación escalar de 8 bits con codificación Huffman
...	...
16:	Cuantificación escalar de 16 bits con codificación Huffman

La unidad de cuantificación 52 también puede realizar versiones predichas de cualquiera de los tipos anteriores de modos de cuantificación, donde se determina una diferencia entre un elemento de (o una ponderación cuando se

realiza la cuantificación del vector) del vector V de una trama previa y el elemento (o ponderación cuando se realiza la cuantificación del vector) del vector V de una trama actual. La unidad de cuantificación 52 puede a continuación cuantificar la diferencia entre los elementos o ponderaciones de la trama actual y de la trama previa en lugar del valor del elemento del vector V de la propia trama actual.

5
 [0052] La unidad de cuantificación 52 puede realizar múltiples formas de cuantificación con respecto a cada uno de los vectores $V[k]$ de primer plano reducidos 55 para obtener múltiples versiones codificadas de los vectores $V[k]$ de primer plano reducidos 55. La unidad de cuantificación 52 puede seleccionar una de las versiones codificadas de los vectores $V[k]$ de primer plano reducidos 55 como el vector $V[k]$ de primer plano codificado 57. La unidad de
 10
 cuantificación 52 puede, en otras palabras, seleccionar uno del vector V cuantificado en vector no predicho, el vector V cuantificado en vector predicho, el vector V cuantificado escalar sin codificación Huffman y el vector V cuantificado escalar con codificación Huffman para usar como el vector V cuantificado conmutado de salida en base a cualquier combinación de los criterios analizados en la presente divulgación. En algunos ejemplos, la unidad de cuantificación 52 puede seleccionar un modo de cuantificación de un conjunto de modos de cuantificación que incluye un modo de
 15
 cuantificación vectorial y uno o más modos de cuantificación escalar, y cuantificar un vector V de entrada en base a (o de acuerdo con) el modo seleccionado. La unidad de cuantificación 52 puede proporcionar a continuación uno del vector V cuantificado en vector no predicho (por ejemplo, en términos de valores de ponderación o bits indicativos del mismo), el vector V cuantificado en vector predicho (por ejemplo, en términos de valores de error o bits indicativos del mismo), el vector V cuantificado escalar sin codificación Huffman y el vector V cuantificado escalar con codificación
 20
 Huffman a la unidad de generación de flujo de bits 52 como los vectores $V[k]$ de primer plano codificados 57. La unidad de cuantificación 52 también puede proporcionar los elementos sintácticos indicativos del modo de cuantificación (por ejemplo, el elemento sintáctico NbitsQ) y cualquier otro elemento sintáctico usado para descuantificar o reconstruir de otro modo el vector V.

25
 [0053] La unidad de descorrelación 40' incluida dentro del dispositivo de codificación de audio 20 puede representar casos individuales o múltiples de una unidad configurada para aplicar una o más transformadas de descorrelación a los coeficientes HOA 47', para obtener los coeficientes HOA descorrelacionados 47". En algunos ejemplos, la unidad de descorrelación 40' puede aplicar una matriz UHJ a los coeficientes HOA 47'. En diversos casos de la presente divulgación, la matriz UHJ también se puede denominar una "transformada basada en fase". La aplicación de la
 30
 transformada basada en fase también se puede denominar en el presente documento "descorrelación de desplazamiento de fase".

[0054] El formato UHJ ambisónico es un desarrollo del sistema de sonido envolvente ambisónico diseñado para que sea compatible con los medios mono y estéreo. El formato UHJ incluye una jerarquía de sistemas en los que el campo sonoro grabado se reproducirá con un grado de exactitud que varía de acuerdo con los canales disponibles. En
 35
 diversos casos, el UHJ también se denomina "formato C". Las iniciales indican algunas de las fuentes incorporadas en el sistema: U de universal (UD-4); H de matriz H; y J del sistema 45J.

[0055] El UHJ es un sistema jerárquico de codificación y decodificación de información de sonido direccional dentro de la tecnología ambisónica. Dependiendo del número de canales disponibles, un sistema puede transportar más o
 40
 menos información. El UHJ es totalmente compatible con estéreo y mono. Se pueden usar hasta cuatro canales (L, R, T, Q).

[0056] En una forma, se puede transportar información envolvente horizontal (o "plana") en UHJ de 2 canales (L, R) por canales de señal en estéreo normales, CD, FM o radio digital, etc., que se puede recuperar usando un
 45
 decodificador de UHJ en el extremo de escucha. La suma de los dos canales puede proporcionar una señal mono compatible, que puede ser una representación más exacta de la versión de dos canales que la suma de una fuente "mono de olla panorámica" convencional. Si está disponible un tercer canal (T), el tercer canal se puede usar para proporcionar una exactitud de localización mejorada para el efecto envolvente plano cuando se descodifica por medio de un decodificador UHJ de 3 canales. Es posible que no se requiera que el tercer canal tenga un ancho de banda de audio completo para este propósito, dando lugar a la posibilidad de los denominados sistemas de "2½ canales", donde el tercer canal tiene un ancho de banda limitado. En un ejemplo, el límite puede ser de 5 kHz. El tercer canal se puede transmitir por medio de radio FM, por ejemplo, mediante modulación de fase en cuadratura. Añadir un cuarto canal (Q) al sistema UHJ puede permitir la codificación de sonido envolvente completo con altura, a veces denominado
 50
 n como perifonía, con un nivel de exactitud idéntico al formato B de 4 canales.

[0057] El UHJ de 2 canales es un formato comúnmente usado para la distribución de grabaciones ambisónicas. Las grabaciones de UHJ de 2 canales se pueden transmitir por medio de todos los canales estéreo normales y cualquiera de los medios de 2 canales normales se puede usar sin alteración. El UHJ es compatible con estéreo por que, sin
 55
 descodificar, el oyente puede percibir una imagen en estéreo, pero que es significativamente más amplia que el estéreo convencional (por ejemplo, el denominado "súper estéreo"). Los canales izquierdo y derecho también se pueden sumar para un muy alto grado de compatibilidad mono. Se puede revelar la capacidad envolvente reproducida por medio de un decodificador UHJ.

[0058] Una representación matemática de ejemplo de la unidad de descorrelación 40' aplicando la matriz UHJ (o transformada basada en fase) es como sigue:

Codificación UHJ:

$$S = (0,9397 * W) + (0,1856 * X);$$

$$D = \text{imag}(\text{hilbert}((-0,3420 * W) + (0,5099 * X))) + (0,6555 * Y);$$

$$T = \text{imag}(\text{hilbert}((-0,1432 * W) + (0,6512 * X))) - (0,7071 * Y);$$

$$Q = 0,9772 * Z;$$

5

conversión de S y D en izquierda y derecha:

$$\text{Izquierda} = (S+D)/2$$

$$\text{Derecha} = (S-D)/2$$

10 **[0059]** De acuerdo con algunas implementaciones de los cálculos anteriores, los supuestos con respecto a los cálculos anteriores pueden incluir lo siguiente: Los canales de fondo de HOA son ambisónicos de 1.er orden, FuMa normalizado, en el orden de numeración de canales ambisónicos W (a00), X(a11), Y(a11-), Z(a10).

15 **[0060]** En los cálculos mencionados anteriormente, la unidad de descorrelación 40' puede realizar una multiplicación escalar de diversas matrices por valores constantes. Por ejemplo, para obtener la señal S, la unidad de descorrelación 40' puede realizar la multiplicación escalar de una matriz W por el valor constante de 0,9397 (por ejemplo, por multiplicación escalar), y de una matriz X por el valor constante de 0,1856. Como también se ilustra en los cálculos mencionados anteriormente, la unidad de descorrelación 40' puede aplicar una transformada de Hilbert (indicada por la función "Hilbert ()" en la codificación UHJ anterior) para obtener cada una de las señales D y T. La función "imag()" en la codificación UHJ anterior indica que se obtiene el imaginario (en el sentido matemático) del resultado de la transformada de Hilbert.

20 **[0061]** Otra representación matemática de ejemplo de la unidad de descorrelación 40' aplicando la matriz UHJ (o transformada basada en fase) es como sigue:

25

Codificación UHJ:

$$S = (0,9396926 * W) + (0,151520536509082 * X);$$

$$D = \text{imag}(\text{hilbert}((-0,3420201 * W) + (0,416299273350443 * X))) + (0,535173990363608 * Y);$$

$$T = 0,940604061228740 * (\text{imag}(\text{hilbert}((-0,1432 * W) + (0,531702573500135 * X))) - (0,577350269189626 * Y));$$

$$Q=Z;$$

30 conversión de S y D en izquierda y derecha:

$$\text{Izquierda} = (S+D)/2$$

$$\text{Derecha} = (S-D)/2$$

35 **[0062]** En algunas implementaciones de ejemplo de los cálculos anteriores, los supuestos con respecto a los cálculos anteriores pueden incluir lo siguiente: Los canales de fondo de HOA son ambisónicos de 1.er orden, N3D (o "3D completo") normalizado, en el orden de numeración de canales ambisónicos W (a00), X(a11), Y(a11-), Z(a10). Aunque se describe en el presente documento con respecto a la normalización en N3D, se apreciará que los cálculos de ejemplo también se pueden aplicar a canales de fondo de HOA que están normalizados en SN3D (o "seminormalizados de Schmidt). La normalización en N3D y SN3D puede diferir en términos de los factores de escala usados. Una representación de ejemplo de la normalización en N3D, relativa a la normalización en SN3D, se expresa a continuación:

40

$$N_{l,m}^{N3D} = N_{l,m}^{SN3D} \sqrt{2l+1}$$

[0063] Un ejemplo de coeficientes de ponderación usados en la normalización en SN3D se expresa a continuación:

$$N_{l,m}^{SN3D} = \sqrt{\frac{2 - \delta_m (l - |m|)!}{4\pi (l + |m|)!}}, \delta_m \begin{cases} 1 & \text{if } m = 0 \\ 0 & \text{if } m \neq 0 \end{cases}$$

5

[0064] En los cálculos mencionados anteriormente, la unidad de descorrelación 40' puede realizar una multiplicación escalar de diversas matrices por valores constantes. Por ejemplo, para obtener la señal S, la unidad de descorrelación 40' puede realizar la multiplicación escalar de una matriz W por el valor constante de 0,9396926 (por ejemplo, por multiplicación escalar), y de una matriz X por el valor constante de 0,151520536509082. Como también se ilustra en los cálculos mencionados anteriormente, la unidad de descorrelación 40' puede aplicar una transformada de Hilbert (indicada por la función "Hilbert ()" en la codificación UHJ anterior o descorrelación de desplazamiento de fase) para obtener cada una de las señales D y T. La función "imag()" en la codificación UHJ anterior indica que se obtiene el imaginario (en el sentido matemático) del resultado de la transformada de Hilbert.

10

15

[0065] La unidad de descorrelación 40' puede realizar los cálculos mencionados anteriormente, de modo que las señales S y D resultantes representen señales de audio izquierda y derecha (o en otras palabras, señales de audio estéreo). En algunos escenarios de este tipo, la unidad de descorrelación 40' puede emitir las señales T y Q como parte de los coeficientes HOA descorrelacionados 47", pero un dispositivo de descodificación que recibe el flujo de bits 21 no puede procesar las señales T y Q cuando se reproducen en una geometría de altavoz estéreo (o, en otras palabras, configuración de altavoz estéreo). En los ejemplos, los coeficientes HOA 47' pueden representar un campo sonoro que se va a reproducir en un sistema de reproducción de audio mono. La unidad de descorrelación 40' puede emitir las señales S y D como parte de los coeficientes HOA descorrelacionados 47", y un dispositivo de descodificación que recibe el flujo de bits 21 puede combinar (o "mezclar") las señales S y D para formar una señal de audio que se va a reproducir y/o emitir en formato de audio mono. En estos ejemplos, el dispositivo de descodificación y/o el dispositivo de reproducción pueden recuperar la señal de audio mono de diversas maneras. Un ejemplo es mezclando las señales izquierda y derecha (representadas por las señales S y D). Otro ejemplo es aplicando una matriz UHJ (o transformada basada en fase) para descodificar una señal W (analizada con mayor detalle a continuación, con respecto a la FIG. 5). Al producir una señal izquierda natural y una señal derecha natural en forma de las señales S y D aplicando la matriz UHJ (o transformada basada en fase), la unidad de descorrelación 40' puede implementar las técnicas de la presente divulgación para proporcionar ventajas potenciales y/o mejoras potenciales sobre las técnicas que aplican otras transformadas de descorrelación (tales como una matriz de modo descrita en la norma MPEG-H).

20

25

30

[0066] En diversos ejemplos, la unidad de descorrelación 40' puede aplicar diferentes transformadas de descorrelación, en base a una velocidad de bits de los coeficientes HOA 47' recibidos. Por ejemplo, la unidad de descorrelación 40' puede aplicar la matriz UHJ (o transformada basada en fase) descrita anteriormente en escenarios donde los coeficientes HOA 47' representan una entrada de cuatro canales. Más específicamente, en base a los coeficientes HOA 47' que representan una entrada de cuatro canales, la unidad de descorrelación 40' puede aplicar una matriz UHJ de 4 x 4 (o transformada basada en fase). Por ejemplo, la matriz de 4 x 4 puede ser ortogonal a la entrada de cuatro canales de los coeficientes HOA 47'. En otras palabras, en los casos donde los coeficientes HOA 47' representen un número menor de canales (por ejemplo, cuatro), la unidad de descorrelación 40' puede aplicar la matriz UHJ como la transformada de descorrelación seleccionada, para descorrelacionar las señales de fondo de las señales de HOA 47' para obtener los coeficientes HOA descorrelacionados 47".

35

40

45

[0067] De acuerdo con este ejemplo, si los coeficientes HOA 47' representan un mayor número de canales (por ejemplo, nueve), la unidad de descorrelación 40' puede aplicar una transformada de descorrelación diferente de la matriz UHJ (o transformada basada en fase). Por ejemplo, en un escenario donde que los coeficientes HOA 47' representan una entrada de nueve canales, la unidad de descorrelación 40' puede aplicar una matriz de modo (por ejemplo, como se describe en la norma MPEG-H), para descorrelacionar los coeficientes HOA 47'. En ejemplos donde los coeficientes HOA 47' representan una entrada de nueve canales, la unidad de descorrelación 40' puede aplicar una matriz de modo de 9 x 9 para obtener los coeficientes HOA descorrelacionados 47".

50

[0068] A su vez, diversos componentes del dispositivo de codificación de audio 20 (tal como el codificador de audio psicoacústico 40) pueden codificar de forma perceptual los coeficientes HOA descorrelacionados 47" de acuerdo con AAC o USAC. La unidad de descorrelación 40' puede aplicar la transformada de descorrelación de desplazamiento de fase (por ejemplo, la matriz UHJ o transformada basada en fase en el caso de una entrada de cuatro canales), para optimizar la codificación AAC/USAC para HOA. En los ejemplos donde los coeficientes HOA 47' (y, de este modo, los coeficientes HOA descorrelacionados 47") representan datos de audio que se van a reproducir en un sistema de reproducción estéreo, la unidad de descorrelación 40' puede aplicar las técnicas de la presente divulgación para

55

60

mejorar u optimizar la compresión, en base a AAC y USAC que están relativamente orientadas (u optimizadas para) datos de audio estéreo.

5 **[0069]** Se entenderá que la unidad de descorrelación 40' puede aplicar las técnicas descritas en el presente documento en situaciones donde los coeficientes HOA 47' compensados en energía incluyen canales de primer plano, así como en situaciones donde los coeficientes HOA compensados en energía 47' no incluyen ningún canal de primer plano. Como un ejemplo, la unidad de descorrelación 40' puede aplicar las técnicas y/o cálculos descritos anteriormente, en un escenario donde los coeficientes HOA compensados en energía 47' incluyen cero (0) canales de primer plano y

10 cuatro (4) canales de fondo (por ejemplo, un escenario de una velocidad de bits más baja/menor).
[0070] En algunos ejemplos, la unidad de descorrelación 40' puede hacer que la unidad de generación de flujo de bits 42 señalice, como parte del flujo de bits basado en vectores 21, uno o más elementos sintácticos que indiquen que la unidad de descorrelación 40' aplicó una transformada de descorrelación a los coeficientes HOA 47'. Al proporcionar una indicación de este tipo a un dispositivo de descodificación, la unidad de descorrelación 40' puede hacer posible
 15 que el dispositivo de descodificación realice las transformadas de descorrelación recíprocas en los datos de audio en el dominio HOA. En algunos ejemplos, la unidad de descorrelación 40' puede hacer que la unidad de generación de flujo de bits 42 señalice elementos sintácticos que indiquen qué transformada de descorrelación se aplicó, tal como la matriz UHJ (u otra transformada basada en fase) o la matriz de modo.

20 **[0071]** La unidad de descorrelación 40' puede aplicar una transformada basada en fase al coeficiente HOA ambiental compensado en energía 47'. La transformada basada en fase para las primeras secuencias de coeficientes HOA O_{MIN} de $C_{AMB}(k-1)$ se define por

$$\begin{bmatrix} x_{AMB,BAJA,1}(k-2) \\ x_{AMB,BAJA,2}(k-2) \\ x_{AMB,BAJA,3}(k-2) \\ x_{AMB,BAJA,4}(k-2) \end{bmatrix} = \begin{bmatrix} d(9) \cdot (S(k-2) + M(k-2)) \\ d(9) \cdot (M(k-2) - S(k-2)) \\ d(8) \cdot (B_{+90}(k-2) + d(5) \cdot c_{AMB,2}(k-2)) \\ c_{AMB,3}(k-2) \end{bmatrix},$$

25 con los coeficientes d como se define en la tabla 1, definiéndose las tramas de señales $S(k-2)$ y $M(k-2)$ por

$$S(k-2) = A_{+90}(k-2) + d(6) \cdot c_{AMB,2}(k-2)$$

$$30 \quad M(k-2) = d(4) \cdot c_{AMB,1}(k-2) + d(5) \cdot c_{AMB,4}(k-2)$$

y $A_{+90}(k-2)$ y $B_{+90}(k-2)$ son las tramas de las señales A y B desplazadas en fase de +90 grados definidas por

$$A(k-2) = d(0) \cdot c_{AMB,BAJA,1}(k-2) + d(1) \cdot c_{AMB,4}(k-2)$$

$$35 \quad B(k-2) = d(2) \cdot c_{AMB,BAJA,1}(k-2) + d(3) \cdot c_{AMB,4}(k-2)$$

La transformada basada en fase para las primeras secuencias de coeficientes HOA O_{MIN} de $C_{P,AMB}(k-1)$ se define en consecuencia. La transformada descrita puede introducir un retraso de una trama.

40 **[0072]** En lo anterior, la $x_{AMB,BAJA,1}(k-2)$ hasta la $x_{AMB,BAJA,4}(k-2)$ pueden corresponder a los coeficientes HOA ambientales descorrelacionados 47". En la ecuación anterior, la variable $C_{AMB,1}(k)$ indica los coeficientes HOA para la k ésima trama correspondiente a las funciones de base esférica que tienen un (orden:suborden) de (0:0), que también se puede denominar el canal o componente "W". La variable $C_{AMB,2}(k)$ variable indica los coeficientes HOA para la k ésima trama correspondiente a las funciones de base esférica que tienen un (orden:suborden) de (1:-1), que también se puede denominar el canal o componente "Y". La variable $C_{AMB,3}(k)$ variable indica los coeficientes HOA para la k ésima
 45 trama correspondiente a las funciones de base esférica que tienen un (orden:suborden) de (1:0), que también se puede denominar el canal o componente "Z". La variable $C_{AMB,4}(k)$ variable indica los coeficientes HOA para la k ésima trama correspondiente a las funciones de base esférica que tienen un (orden:suborden) de (1:1), que también se puede denominar el canal o componente "X". La $C_{AMB,1}(k)$ hasta la $C_{AMB,3}(k)$ pueden corresponder a los coeficientes HOA ambientales 47'.

50 **[0073]** La tabla 1 a continuación ilustra un ejemplo de coeficientes que la unidad de descorrelación 40 puede usar para realizar una transformada basada en fase.

Tabla 1 Coeficientes para la transformada basada en fase

55

n	d(n)
0	0,34202009999999999
1	0,41629927335044281
2	0,14319999999999999
3	0,53170257350013528
4	0,93969259999999999
5	0,15152053650908184
6	0,53517399036360758
7	0,57735026918962584
8	0,94060406122874030
9	0,500000000000000

5 **[0074]** En algunos ejemplos, los diversos componentes del dispositivo de codificación de audio 20 (tales como la unidad de generación de flujo de bits 42) se pueden configurar para transmitir solo representaciones HOA de primer orden para velocidades de bits objetivo más bajas (por ejemplo, una velocidad de bits objetivo de 128K o 256K). De acuerdo con algunos ejemplos de este tipo, el dispositivo de codificación de audio 20 (o los componentes del mismo, tales como la unidad de generación de flujo de bits 42) se puede configurar para descartar coeficientes HOA de orden superior (por ejemplo, coeficientes con un orden mayor que el primer orden, o en otras palabras, $N > 1$). Sin embargo, en los ejemplos donde el dispositivo de codificación de audio 20 determina que la velocidad de bits objetivo es relativamente alta, el dispositivo de codificación de audio 20 (por ejemplo, la unidad de generación de flujo de bits 42) puede separar los canales de primer plano y de fondo, y puede asignar bits (por ejemplo, en mayores cantidades) a los canales de primer plano.

15 **[0075]** La unidad codificadora de audio psicoacústico 40 incluida dentro del dispositivo de codificación de audio 20 puede representar múltiples casos de un codificador de audio psicoacústico, cada uno de los cuales se usa para codificar un objeto de audio diferente o canal HOA de cada uno de los coeficientes HOA descorrelacionados 47" y las señales de nFG interpoladas 49' para generar los coeficientes HOA ambientales codificados 59 y las señales de nFG codificadas 61. La unidad codificadora de audio psicoacústico 40 puede emitir los coeficientes HOA ambientales codificados 59 y las señales de nFG codificadas 61 a la unidad de generación de flujo de bits 42.

20 **[0076]** La unidad de generación de flujo de bits 42 incluida dentro del dispositivo de codificación de audio 20 representa una unidad que formatea datos para adaptarse a un formato conocido (que puede hacer referencia a un formato conocido por un dispositivo de descodificación), generando de este modo el flujo de bits basado en vectores 21. El flujo de bits 21 puede, en otras palabras, representar datos de audio codificados, que se han codificado de la manera descrita anteriormente. La unidad de generación de flujo de bits 42 puede representar un multiplexor en algunos ejemplos, que puede recibir los vectores $V[k]$ de primer plano codificados 57, los coeficientes HOA ambientales codificados 59, las señales de nFG codificadas 61 y la información de canal de fondo 43. La unidad de generación de flujo de bits 42 puede generar a continuación un flujo de bits 21 en base a los vectores $V[k]$ de primer plano codificados 57, los coeficientes HOA ambientales codificados 59, las señales de nFG codificadas 61 y la información de canal de fondo 43. De esta manera, la unidad de generación de flujo de bits 42 puede especificar de este modo los vectores 57 en el flujo de bits 21 para obtener el flujo de bits 21. El flujo de bits 21 puede incluir un flujo de bits primario o principal y uno o más flujos de bits de canal lateral.

35 **[0077]** Aunque no se muestra en el ejemplo de la FIG. 3, el dispositivo de codificación de audio 20 también puede incluir una unidad de salida de flujo de bits que conmuta la salida del flujo de bits desde el dispositivo de codificación de audio 20 (por ejemplo, entre el flujo de bits 21 basado en la dirección y el flujo de bits 21 basado en vectores), en base a si una trama actual va a ser codificada usando la síntesis basada en la dirección o la síntesis basada en vectores. La unidad de salida de flujo de bits puede realizar la conmutación en base al elemento sintáctico emitido por la unidad de análisis de contenido 26, indicando si se realizó una síntesis basada en la dirección (como resultado de detectar que los coeficientes HOA 11 se generaron a partir de un objeto de audio sintético) o si se realizó una síntesis basada en vectores (como resultado de detectar que se registraron los coeficientes HOA). La unidad de salida de flujo de bits puede especificar la sintaxis de cabecera correcta para indicar la conmutación o codificación actual usada para la trama actual, junto con la respectiva de los flujos de bits 21.

45 **[0078]** Además, como se indica anteriormente, la unidad de análisis del campo sonoro 44 puede identificar coeficientes HOA ambientales de BG_{TOT} 47, que pueden cambiar de trama a trama (aunque a veces BG_{TOT} puede permanecer constante o igual a través de dos o más tramas adyacentes (en el tiempo)). El cambio en BG_{TOT} puede dar como resultado cambios en los coeficientes expresados en los vectores $V[k]$ de primer plano reducidos 55. El cambio en BG_{TOT} puede dar como resultado los coeficientes HOA de fondo (que también se pueden denominar "coeficientes HOA ambientales") que cambian de trama a trama (aunque, de nuevo, a veces BG_{TOT} puede permanecer constante o

igual a través de dos o más tramas adyacentes (en el tiempo)). Los cambios a menudo dan como resultado un cambio de energía para los aspectos del campo sonoro representados por la adición o eliminación de los coeficientes HOA ambientales adicionales y la correspondiente eliminación de coeficientes de, o adición de coeficientes a, los vectores $V[k]$ de primer plano reducidos 55.

5 **[0079]** Como resultado, la unidad de análisis del campo sonoro 44 puede determinar además cuándo los coeficientes HOA ambientales cambian de trama a trama, y generar un indicador u otro elemento sintáctico indicativo del cambio para el coeficiente HOA ambiental, en términos de usarse para representar los componentes ambientales del campo sonoro (donde el cambio también se puede denominar una "transición" del coeficiente HOA ambiental o una "transición" del coeficiente HOA ambiental). En particular, la unidad de reducción de coeficientes 46 puede generar el indicador (que se puede indicar como un indicador AmbCoeffTransition o un indicador AmbCoeffIdxTransition), proporcionando el indicador a la unidad de generación de flujo de bits 42 para que el indicador se pueda incluir en el flujo de bits 21 (posiblemente como parte de la información de canal lateral).

15 **[0080]** La unidad de reducción de coeficientes 46 puede, además de especificar el indicador de transición de coeficientes ambientales, también modificar cómo se generan los vectores $V[k]$ de primer plano reducidos 55. En un ejemplo, tras determinar que uno de los coeficientes HOA ambientales está en transición durante la trama actual, la unidad de reducción de coeficientes 46 puede especificar un coeficiente vectorial (que también se puede denominar "elemento vectorial" o "elemento") para cada uno de los vectores V de los vectores $V[k]$ de primer plano reducidos 55, que corresponde al coeficiente HOA ambiental en transición. De nuevo, el coeficiente HOA ambiental en transición puede añadir o eliminar del número total BG_{TOT} de coeficientes de fondo. Por lo tanto, el cambio resultante en el número total de coeficientes de fondo afecta si el coeficiente HOA ambiental se incluye o no en el flujo de bits, y si el elemento correspondiente de los vectores V se incluye para los vectores V especificados en el flujo de bits en los segundo y tercer modos de configuración, descritos anteriormente. Se proporciona más información con respecto a cómo la unidad de reducción de coeficientes 46 puede especificar los vectores $V[k]$ de primer plano reducidos 55 para superar los cambios en la energía en la solicitud de EE. UU. con n.º de serie 14/594,533, titulada "TRANSITIONING OF AMBIENT HIGHER-ORDER AMBISONIC COEFFICIENTS [TRANSICIÓN DE COEFICIENTES AMBISÓNICOS DE ORDEN SUPERIOR AMBIENTALES]" presentada el 12 de enero de 2015.

30 **[0081]** Por tanto, el dispositivo de codificación de audio 20 puede representar un ejemplo de un dispositivo para comprimir audio configurado para aplicar una transformada de descorrelación a los coeficientes ambisónicos ambientales para obtener una representación descorrelacionada de los coeficientes ambisónicos ambientales, habiéndose extraído los coeficientes HOA ambientales de una pluralidad de coeficientes ambisónicos de orden superior y representativos de un componente de fondo de un campo sonoro descrito por la pluralidad de coeficientes ambisónicos de orden superior, en el que al menos uno de la pluralidad de coeficientes ambisónicos de orden superior está asociado con una función de base esférica que tiene un orden mayor que uno. En algunos ejemplos, para aplicar la transformada de descorrelación, el dispositivo se configura para aplicar una matriz UHJ a los coeficientes ambisónicos ambientales.

40 **[0082]** En algunos ejemplos, el dispositivo se configura además para normalizar la matriz UHJ de acuerdo con la normalización en N3D (3D completa). En algunos ejemplos, el dispositivo se configura además para normalizar la matriz UHJ de acuerdo con la normalización en SN3D (seminormalización de Schmidt). En algunos ejemplos, los coeficientes ambisónicos ambientales están asociados con funciones de base esférica que tienen un orden de cero o un orden de uno, y para aplicar la matriz UHJ a los coeficientes ambisónicos ambientales, el dispositivo se configura para realizar una multiplicación escalar de la matriz UHJ con respecto a al menos un subconjunto de los coeficientes ambisónicos ambientales. En algunos ejemplos, para aplicar la transformada de descorrelación, el dispositivo se configura para aplicar una matriz de modo a los coeficientes ambisónicos ambientales.

50 **[0083]** De acuerdo con algunos ejemplos, para aplicar la transformada de descorrelación, el dispositivo se configura para obtener una señal izquierda y una señal derecha de los coeficientes ambisónicos ambientales descorrelacionados. De acuerdo con algunos ejemplos, el dispositivo se configura además para señalar los coeficientes ambisónicos ambientales descorrelacionados junto con uno o más canales en primer plano. De acuerdo con algunos ejemplos, para señalar los coeficientes ambisónicos ambientales descorrelacionados junto con uno o más canales de primer plano, el dispositivo se configura para señalar los coeficientes ambisónicos ambientales descorrelacionados junto con uno o más canales de primer plano en respuesta a una determinación de que una velocidad de datos objetivo cumple o supera un umbral predeterminado.

60 **[0084]** En algunos ejemplos, el dispositivo se configura además para señalar los coeficientes ambisónicos ambientales descorrelacionados sin señalar ningún canal de primer plano. En algunos ejemplos, para señalar los coeficientes ambisónicos ambientales descorrelacionados sin señalar ningún canal de primer plano, el dispositivo se configura para señalar los coeficientes ambisónicos ambientales descorrelacionados sin señalar ningún canal de primer plano en respuesta a una determinación de que una velocidad de bits objetivo está por debajo de un umbral predeterminado. En algunos ejemplos, el dispositivo se configura además para señalar una indicación de que se ha aplicado la transformada de descorrelación a los coeficientes ambisónicos ambientales. En algunos ejemplos, el dispositivo incluye además una formación de micrófonos configurada para capturar los datos de audio que se van a comprimir.

[0085] La FIG. 4 es un diagrama de bloques que ilustra el dispositivo de descodificación de audio 24 de la FIG. 2 con mayor detalle. Como se muestra en el ejemplo de la FIG. 4, el dispositivo de descodificación de audio 24 puede incluir una unidad de extracción 72, una unidad de reconstrucción basada en la direccionalidad 90, una unidad de reconstrucción basada en vectores 92 y una unidad de recorrelación 81.

[0086] . Aunque se describe a continuación, más información con respecto al dispositivo de descodificación de audio 24 y los diversos aspectos de descompresión o de otro modo descodificación de los coeficientes HOA está disponible en la publicación de solicitud de patente internacional n.º WO 2014/194099, titulada "INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD [INTERPOLACIÓN PARA REPRESENTACIONES DESCOMPUESTAS DE UN CAMPO SONORO]", presentada el 29 de mayo de 2014.

[0087] La unidad de extracción 72 puede representar una unidad configurada para recibir el flujo de bits 21 y extraer las diversas versiones codificadas (por ejemplo, una versión codificada basada en la dirección o una versión codificada basada en vectores) de los coeficientes HOA 11. La unidad de extracción 72 puede determinar a partir del elemento sintáctico mencionado anteriormente indicaciones de si los coeficientes HOA 11 se codificaron por medio de las diversas versiones basadas en la dirección o basadas en vectores. Cuando se realizó una codificación basada en la dirección, la unidad de extracción 72 puede extraer la versión basada en la dirección de los coeficientes HOA 11 y los elementos sintácticos asociados con la versión codificada (que se indica como información basada en la dirección 91 en el ejemplo de la FIG. 4), pasando la información basada en la dirección 91 a la unidad de reconstrucción basada en la dirección 90. La unidad de reconstrucción basada en la dirección 90 puede representar una unidad configurada para reconstruir los coeficientes HOA en forma de coeficientes HOA 11', en base a la información basada en la dirección 91. A continuación se describe el flujo de bits y la disposición de los elementos sintácticos dentro del flujo de bits.

[0088] Cuando el elemento sintáctico indica que los coeficientes HOA 11 se codificaron usando una síntesis basada en vectores, la unidad de extracción 72 puede extraer los vectores $V[k]$ de primer plano codificados 57 (que pueden incluir ponderaciones codificadas 57 y/o índices 63 o vectores V con cuantificación escalar), los coeficientes HOA ambientales codificados 59 y los correspondientes objetos de audio 61 (que también se pueden denominar las señales de nFG codificadas 61). Los objetos de audio 61 corresponden cada uno a uno de los vectores 57. La unidad de extracción 72 puede pasar los vectores $V[k]$ de primer plano codificados 57 a la unidad de reconstrucción de vector V 74 y los coeficientes HOA ambientales codificados 59 junto con las señales de nFG codificadas 61 a la unidad de descodificación psicoacústica 80.

[0089] La unidad de reconstrucción de vector V 74 puede representar una unidad configurada para reconstruir los vectores V a partir de los vectores $V[k]$ de primer plano codificados 57. La unidad de reconstrucción de vector V 74 puede funcionar de una manera recíproca a la de la unidad de cuantificación 52.

[0090] La unidad de descodificación psicoacústica 80 puede funcionar de una manera recíproca a la unidad codificadora de audio psicoacústico 40 mostrada en el ejemplo de la FIG. 3 para descodificar los coeficientes HOA ambientales codificados 59 y las señales de nFG codificadas 61 y generar de este modo coeficientes HOA ambientales compensados en energía 47' y las señales de nFG interpoladas 49' (que también se pueden denominar objetos de audio de nFG interpolados 49'). La unidad de descodificación psicoacústica 80 puede pasar los coeficientes HOA ambientales compensados en energía 47' a la unidad de recorrelación 81 y las señales de nFG 49' a la unidad de formulación de primer plano 78. A su vez, la unidad de recorrelación 81 puede aplicar una o más transformadas de recorrelación a los coeficientes HOA ambientales compensados en energía 47' para obtener uno o más coeficientes HOA recorrelacionados 47" (o coeficientes HOA correlacionados 47") y puede pasar los coeficientes HOA correlacionados 47" a la unidad de formulación de coeficientes HOA 82 (opcionalmente, a través de la unidad de fundido 770).

[0091] De forma similar a las descripciones anteriores, con respecto a la unidad de descorrelación 40' del dispositivo de codificación de audio 20, la unidad de recorrelación 81 puede implementar las técnicas de la presente divulgación para reducir la correlación entre los canales de fondo de los coeficientes HOA ambientales compensados en energía 47' para reducir o mitigar el desenmascaramiento del ruido. En los ejemplos donde la unidad de recorrelación 81 aplica una matriz UHJ (por ejemplo, una matriz UHJ inversa) como la transformada de recorrelación seleccionada, la unidad de recorrelación 81 puede mejorar las tasas de compresión y conservar los recursos informáticos al reducir las operaciones de procesamiento de datos. En algunos ejemplos, el flujo de bits basado en vectores 21 puede incluir uno o más elementos sintácticos que indican que se aplicó una transformada de descorrelación durante la codificación. La inclusión de dichos elementos sintácticos en el flujo de bits basado en vectores 21 puede hacer posible que la unidad de recorrelación 81 realice transformadas de descorrelación recíprocas (por ejemplo, correlación o recorrelación) sobre los coeficientes HOA compensados en energía 47'. En algunos ejemplos, los elementos sintácticos de señal pueden indicar qué transformada de descorrelación se aplicó, tal como la matriz UHJ o la matriz de modo, haciendo posible de este modo que la unidad de recorrelación 81 seleccione la transformada de recorrelación apropiada para aplicar a los coeficientes HOA compensados en energía 47'.

[0092] En los ejemplos donde la unidad de reconstrucción basada en vectores 92 emite los coeficientes HOA 11' a un sistema de reproducción que comprende un sistema estéreo, la unidad de recorrelación 81 puede procesar las señales S y D (por ejemplo, una señal izquierda natural y una señal derecha natural) para producir los coeficientes HOA recorrelacionados 47". Por ejemplo, debido a que las señales S y D representan una señal izquierda natural y una señal derecha natural, el sistema de reproducción puede usar las señales S y D como los dos flujos de salida estéreo. En los ejemplos donde la unidad de reconstrucción 92 emite los coeficientes HOA 11' a un sistema de reproducción que comprende un sistema de audio mono, el sistema de reproducción puede combinar o mezclar las señales S y D (como se representa en los coeficientes HOA 11') para obtener la salida de audio mono para la reproducción. En el ejemplo de un sistema de audio mono, el sistema de reproducción puede añadir la salida de audio mono mixta a uno o más canales de primer plano (si hay algún canal de primer plano) para generar la salida de audio.

[0093] Con respecto a algunos codificadores capaces de UHJ existentes, las señales se procesan en una matriz de amplitud de fase para recuperar un conjunto de señales que se asemeja al formato B. En la mayoría de los casos, la señal será realmente en formato B, pero en el caso de UHJ de 2 canales, no hay suficiente información disponible para que se pueda reconstruir una verdadera señal en formato B, sino más bien, una señal que presenta características similares a una señal en formato B. A continuación, la información se pasa a una matriz de amplitud que desarrolla las señales de los altavoces, por medio de un conjunto de filtros limitadores, que mejoran la exactitud y rendimiento del descodificador en entornos de escucha más pequeños (se pueden omitir en aplicaciones de mayor escala). La ambisónica se diseñó para adaptarse a habitaciones reales (por ejemplo, salas de estar) y a posiciones de altavoces prácticas: muchas habitaciones de este tipo son rectangulares y, como resultado, el sistema básico se diseñó para descodificar a cuatro altavoces en un rectángulo, con lados entre 1:2 (anchura dos veces la longitud) y 2:1 (longitud dos veces la anchura) de longitud, adaptándose por tanto a la mayoría de dichas habitaciones. En general se proporciona un control de disposición para permitir que el descodificador se configure para las posiciones de los altavoces. El control de disposición es un aspecto de la reproducción ambisónica que difiere de otros sistemas de sonido envolvente: el descodificador se puede configurar específicamente para el tamaño y la disposición de la formación de altavoces. El control de disposición puede adoptar la forma de un botón giratorio, un interruptor de 2 vías (1:2,2:1) o de 3 vías (1:2,1:1,2:1). Cuatro altavoces es el mínimo requerido para la descodificación envolvente horizontal, y si bien una disposición de cuatro altavoces puede ser adecuada para varios entornos de escucha, los espacios más grandes pueden requerir más altavoces para dar una localización envolvente completa.

[0094] Un ejemplo de cálculos que la unidad de recorrelación 81 puede realizar con respecto a la aplicación de una matriz UHJ (por ejemplo, una matriz de UHJ inversa o transformada basada en fase inversa) como una transformada de recorrelación se enumeran a continuación:

Descodificación UHJ:

conversión de izquierda y derecha en S y D:

$$S = \text{izquierda} + \text{derecha}$$

$$D = \text{izquierda} - \text{derecha}$$

$$W = (0,982*S) + 0,197.*\text{imag}(\text{hilbert}((0,828*D) + (0,768*T)));$$

$$X = (0,419*S) - \text{imag}(\text{hilbert}((0,828*D) + (0,768*T)));$$

$$Y = (0,796*D) - 0,676*T + \text{imag}(\text{hilbert}(0,187*S));$$

$$Z = (1,023*Q);$$

[0095] En algunas implementaciones de ejemplo de los cálculos anteriores, los supuestos con respecto a los cálculos anteriores pueden incluir lo siguiente: Los canales de fondo de HOA son ambisónicos de 1.er orden, FuMa normalizado, en el orden de numeración de canales ambisónicos W (a00), X(a11), Y(a11-), Z(a10).

[0096] Un ejemplo de cálculos que la unidad de recorrelación 81 puede realizar con respecto a la aplicación de una matriz UHJ (o transformada basada en fase inversa) como una transformada de recorrelación se enumeran a continuación:

Descodificación UHJ:

conversión de izquierda y derecha en S y D:

conversión de izquierda y derecha en S y D:

ES 2 729 624 T3

$$S = \text{izquierda} + \text{derecha};$$

$$D = \text{izquierda} - \text{derecha};$$

$$h1 = \text{imag}(\text{hilbert}(1,014088753512236 * D + T));$$

$$h2 = \text{imag}(\text{hilbert}(0,229027290950227 * S));$$

$$W = 0,982 * S + 0,160849826442762 * h1;$$

$$X = 0,513168101113076 * S - h1;$$

$$Y = 0,974896917627705 * D - 0,880208333333333 * T + h2;$$

$$Z = Q;$$

5 **[0097]** En algunas implementaciones de los cálculos anteriores, los supuestos con respecto a los cálculos anteriores pueden incluir lo siguiente: Los canales de fondo de HOA son ambisónicos de 1.er orden, N3D (o "tres-D completo") normalizado, en el orden de numeración de canales ambisónicos W (a00), X(a11), Y(a11-), Z(a10). Aunque se describe en el presente documento con respecto a la normalización en N3D, se apreciará que los cálculos de ejemplo también se pueden aplicar a canales de fondo de HOA que están normalizados en SN3D (o "seminormalizados de Schmidt").
10 Como se describe anteriormente con respecto a la FIG. 4, la normalización en N3D y SN3D puede diferir en términos de los factores de escala usados. Una representación de ejemplo de los factores de escala usados en la normalización en N3D se describe anteriormente con respecto a la FIG. 4. Una representación de ejemplo de los coeficientes de ponderación usados en la normalización en SN3D se describe anteriormente con respecto a la FIG. 4.

15 **[0098]** En algunos ejemplos, los coeficientes HOA compensados en energía 47' pueden representar una disposición solo horizontal, tal como datos de audio que no incluyen ningún canal vertical. En estos ejemplos, la unidad de recorrelación 81 puede no realizar los cálculos con respecto a la señal Z anterior, porque la señal Z representa datos de audio direccionales verticales. En cambio, en estos ejemplos, la unidad de recorrelación 81 solo puede realizar los cálculos anteriores con respecto a las señales W, X e Y, porque las señales W, X e Y representan datos direccionales horizontales. En algunos ejemplos donde los coeficientes HOA compensados en energía 47' representan datos de audio que se van a reproducir en un sistema de reproducción de audio mono, la unidad de recorrelación 81 solo puede obtener la señal W de los cálculos anteriores. Más específicamente, debido a que la señal W resultante representa los datos de audio mono, la señal W puede proporcionar todos los datos necesarios donde los coeficientes HOA compensados en energía 47' representan datos que se van a reproducir en formato de audio mono, o donde el sistema de reproducción comprende un sistema de audio mono.
20
25

[0099] De forma similar a como se describe anteriormente con respecto a la unidad de descorrelación 40' del dispositivo de codificación de audio 20, la unidad de recorrelación 81 puede, en los ejemplos, aplicar la matriz UHJ (o una matriz UHJ inversa o transformada basada en fase inversa) en los escenarios donde los coeficientes HOA compensados en energía 47' incluyen un número menor de canales de fondo, pero pueden aplicar una matriz de modo o matriz de modo inversa (por ejemplo, como se describe en la norma MPEG-H) en escenarios donde los coeficientes HOA compensados en energía 47' incluyen un mayor número de canales de fondo.
30

[0100] Se entenderá que la unidad de recorrelación 81 puede aplicar las técnicas descritas en el presente documento en situaciones donde los coeficientes HOA compensados en energía 47' incluyen canales de primer plano, así como en situaciones donde los coeficientes HOA compensados en energía 47' no incluyen ningún canal de primer plano. Como un ejemplo, la unidad de recorrelación 81 puede aplicar las técnicas y/o cálculos descritos anteriormente, en un escenario donde los coeficientes HOA compensados en energía 47' incluyen cero (0) canales de primer plano y ocho (8) canales de fondo (por ejemplo, un escenario de una velocidad de bits más baja/menor).
35
40

[0101] Diversos componentes del dispositivo de decodificación de audio 24, tales como la unidad de recorrelación 81, pueden ser un elemento sintáctico, tal como un indicador UsePhaseShiftDecorr, para determinar cuál de los dos procedimientos de procesamiento se aplicó para la descorrelación. En los casos donde la unidad de descorrelación 40' usó una transformada espacial para la descorrelación, la unidad de recorrelación 81 puede determinar que el indicador UsePhaseShiftDecorr se fija en un valor de cero.
45

[0102] En los casos donde la unidad de recorrelación 81 determina que el indicador UsePhaseShiftDecorr se fija en un valor de uno, la unidad de recorrelación 81 puede determinar que la recorrelación se va a realizar usando una transformada basada en fase. Si el indicador UsePhaseShiftDecorr es de valor 1, se aplica el siguiente procesamiento para reconstruir las primeras cuatro secuencias de coeficientes del componente HOA ambiental mediante
50

$$\begin{bmatrix} \mathbf{c}_{AMB,1}(k) \\ \mathbf{c}_{AMB,2}(k) \\ \mathbf{c}_{AMB,3}(k) \\ \mathbf{c}_{AMB,4}(k) \end{bmatrix} = \begin{bmatrix} c(3) \cdot A_{+90}(k) + c(2) \cdot [\mathbf{c}_{I,AMB,1}(k) + \mathbf{c}_{I,AMB,2}(k)] \\ B_{+90}(k) + c(5) \cdot [\mathbf{c}_{I,AMB,1}(k) - \mathbf{c}_{I,AMB,2}(k)] + c(6) \cdot \mathbf{c}_{I,AMB,3}(k) \\ \mathbf{c}_{I,AMB,4}(k) \\ c(4) \cdot [\mathbf{c}_{I,AMB,1}(k) + \mathbf{c}_{I,AMB,2}(k)] - A_{+90}(k) \end{bmatrix}$$

con los coeficientes c como se define en la tabla 1 a continuación, y $A_{+90}(k)$ y $B_{+90}(k)$ son las tramas de las señales A y B desplazadas en fase de +90 grados definidas por

5

$$A(k) = c(0) \cdot [\mathbf{c}_{I,AMB,1}(k) - \mathbf{c}_{I,AMB,2}(k)],$$

$$B(k) = c(1) \cdot [\mathbf{c}_{I,AMB,1}(k) + \mathbf{c}_{I,AMB,2}(k)].$$

10 **[0103]** La tabla 2 a continuación ilustra los coeficientes de ejemplo que la unidad de descorrelación 40' puede usar para implementar una transformada basada en fase.

Tabla 2. Coeficientes para la transformada basada en fase

n	c(n)
0	1,0140887535122356
1	0,22902729095022714
2	0,9819999999999998
3	0,16084982644276205
4	0,51316810111307576
5	0,97489691762770481
6	-0,8802083333333337

15

[0104] En la ecuación anterior, la variable $C_{AMB,1}(k)$ variable indica los coeficientes HOA para la $k^{\text{ésima}}$ trama correspondiente a las funciones de base esférica que tienen un (orden:suborden) de (0:0), que también se puede denominar el canal o componente "W". La variable $C_{AMB,2}(k)$ variable indica los coeficientes HOA para la $k^{\text{ésima}}$ trama correspondiente a las funciones de base esférica que tienen un (orden:suborden) de (1:-1), que también se puede denominar el canal o componente "Y". La variable $C_{AMB,3}(k)$ variable indica los coeficientes HOA para la $k^{\text{ésima}}$ trama correspondiente a las funciones de base esférica que tienen un (orden:suborden) de (1:0), que también se puede denominar el canal o componente "Z". La variable $C_{AMB,4}(k)$ variable indica los coeficientes HOA para la $k^{\text{ésima}}$ trama correspondiente a las funciones de base esférica que tienen un (orden:suborden) de (1:1), que también se puede denominar el canal o componente "X". La $C_{AMB,1}(k)$ hasta la $C_{AMB,3}(k)$ pueden corresponder a los coeficientes HOA ambientales 47'.

20

25

[0105] La notación $[C_{I,AMB,1}(k) + C_{I,AMB,2}(k)]$ anterior indica lo que de forma alternativa se denomina "S", que es equivalente al canal izquierdo más el canal derecho. La variable $C_{I,AMB,1}(k)$ indica el canal izquierdo generado como resultado de la codificación UHJ, mientras que la variable $C_{I,AMB,2}(k)$ indica el canal derecho generado como resultado de la codificación UHJ. La notación "I" en el subíndice indica que el canal correspondiente se ha descorrelacionado (por ejemplo, a través de la aplicación de la matriz UHJ o transformada basada en fase) de los otros canales ambientales. La notación $[C_{I,AMB,1}(k) - C_{I,AMB,2}(k)]$ indica lo que se denomina "D" a lo largo de toda la presente divulgación, que es representativa del canal izquierdo menos el canal derecho. La variable $C_{I,AMB,3}(k)$ indica lo que se denomina la variable "T" a lo largo de toda la presente divulgación. La variable $C_{I,AMB,4}(k)$ indica lo que se denomina la variable "Q" a lo largo de toda la presente divulgación.

30

35

[0106] La notación $A_{+90}(k)$ indica un desplazamiento de fase de 90 grados positivo de $c(0)$ multiplicado por S (que también se indica por la variable "h1" a lo largo de toda la presente divulgación). La notación $B_{+90}(k)$ indica un desplazamiento de fase de 90 grados positivo de $c(1)$ multiplicado por D (que también se indica por la variable "h2" a lo largo de toda la presente divulgación).

40

[0107] La unidad de interpolación espacio-temporal 76 puede funcionar de una manera similar a la descrita anteriormente con respecto a la unidad de interpolación espacio-temporal 50. La unidad de interpolación espacio-temporal 76 puede recibir los vectores $V[k]$ reducidos de primer plano 55_k y realizar la interpolación espacio-temporal con respecto a los vectores $V[k]$ de primer plano 55_k y los vectores $V[k-1]$ reducidos de primer plano 55_{k-1} para generar vectores $V[k]$ de primer plano interpolados 55_k ". La unidad de interpolación espacio-temporal 76 puede reenviar los vectores $V[k]$ de primer plano interpolados 55_k " a la unidad de fundido 770.

45

[0108] La unidad de extracción 72 también puede emitir una señal 757 indicativa de cuándo uno de los coeficientes HOA ambientales está en transición hacia la unidad de fundido 770, que puede a continuación determinar cuál de los SHC_{BG} 47' (donde los SHC_{BG}47' también se pueden indicar como "canales HOA ambientales 47'" o "coeficientes HOA ambientales 47'") y los elementos de los vectores $V[k]$ de primer plano interpolados 55_k van a efectuar un fundido de entrada o bien un fundido de salida. En algunos ejemplos, la unidad de fundido 770 puede funcionar opuesta con respecto a cada uno de los coeficientes HOA ambientales 47' y los elementos de los vectores $V[k]$ de primer plano interpolados 55_k . Es decir, la unidad de fundido 770 puede realizar un fundido de entrada o un fundido de salida, o tanto un fundido de entrada como un fundido de salida, con respecto al correspondiente de los coeficientes HOA ambientales 47', mientras realiza un fundido de entrada o un fundido de salida, o tanto un fundido de entrada como un fundido de salida, con respecto al correspondiente de los elementos de los vectores $V[k]$ de primer plano interpolados 55_k . La unidad de fundido 770 puede emitir coeficientes HOA ambientales ajustados 47' a la unidad de formulación de coeficientes HOA 82 y vectores $V[k]$ de primer plano ajustados 55_k a la unidad de formulación de primer plano 78. A este respecto, la unidad de fundido 770 representa una unidad configurada para realizar una operación de fundido con respecto a diversos aspectos de los coeficientes HOA o los derivados de los mismos, por ejemplo, en forma de los coeficientes HOA ambientales 47' y los elementos de los vectores $V[k]$ de primer plano interpolados 55_k .

[0109] La unidad de formulación de primer plano 78 puede representar una unidad configurada para realizar la multiplicación matricial con respecto a los vectores $V[k]$ de primer plano ajustados 55_k y las señales de nFG interpoladas 49' para generar los coeficientes HOA de primer plano 65. A este respecto, la unidad de formulación de primer plano 78 puede combinar los objetos de audio 49' (que es otra manera mediante la cual indicar las señales de nFG interpoladas 49') con los vectores 55_k para reconstruir el primer plano o, en otras palabras, los aspectos predominantes de los coeficientes HOA 11'. La unidad de formulación de primer plano 78 puede realizar una multiplicación matricial de las señales de nFG interpoladas 49' por los vectores $V[k]$ de primer plano ajustados 55_k .

[0110] La unidad de formulación de coeficientes HOA 82 puede representar una unidad configurada para combinar los coeficientes HOA de primer plano 65 con los coeficientes HOA ambientales ajustados 47' para obtener los coeficientes HOA 11'. La notación principal refleja que los coeficientes HOA 11' pueden ser similares pero no iguales a los coeficientes HOA 11. Las diferencias entre los coeficientes HOA 11 y 11' pueden ser el resultado de pérdidas debidas a la transmisión por un medio de transmisión con pérdidas, cuantificación u otras operaciones con pérdidas.

[0111] El UHJ es un procedimiento de transformada de matriz que se ha usado para crear un flujo estéreo de 2 canales a partir de contenido de ambisónicos de primer orden. El UHJ se ha usado en el pasado para transmitir contenido envolvente estéreo o solo horizontal por medio de un transmisor de FM. Sin embargo, se apreciará que el UHJ no se limita al uso en transmisores de FM. En el esquema de codificación de HOA de MPEG-H, los canales de fondo de HOA se pueden preprocesar con una matriz de modo para convertir los canales de fondo de HOA en puntos ortogonales en el dominio espacial. A continuación, los canales transformados se codifican de forma perceptual por medio de USAC o AAC.

[0112] Las técnicas de la presente divulgación están dirigidas en general al uso de la transformada de UHJ (o transformada basada en fase) en la aplicación de la codificación de los canales de fondo de HOA en lugar de usar esta matriz de modo. Ambos métodos ((1) transformación en dominio espacial por medio de una matriz de modo (2) transformada de UHJ) están dirigidos en general a reducir la correlación entre los canales de fondo de HOA, lo que puede dar como resultado (el potencialmente no deseado) efecto de desenmascaramiento del ruido dentro del campo sonoro descodificado.

[0113] Por tanto, el dispositivo de descodificación de audio 24 puede, en los ejemplos, representar un dispositivo configurado para obtener una representación descorrelacionada de los coeficientes ambisónicos ambientales que tienen al menos una señal izquierda y una señal derecha, habiéndose extraído los coeficientes ambisónicos ambientales de una pluralidad de coeficientes ambisónicos de orden superior y representativos de un componente de fondo de un campo sonoro descrito por la pluralidad de coeficientes ambisónicos de orden superior, en los que al menos uno de la pluralidad de coeficientes ambisónicos de orden superior está asociado con una función de base esférica que tiene un orden mayor que uno, y para generar una señal de altavoz en base a la representación descorrelacionada de los coeficientes ambisónicos ambientales. En algunos ejemplos, el dispositivo se configura además para aplicar una transformada de recorrelación a la representación descorrelacionada de los coeficientes ambisónicos ambientales para obtener una pluralidad de coeficientes ambisónicos ambientales correlacionados.

[0114] En algunos ejemplos, para aplicar la transformada de recorrelación, el dispositivo se configura para aplicar una matriz de UHJ inversa (o transformada basada en fase) a los coeficientes ambisónicos ambientales. De acuerdo con algunos ejemplos, la matriz de UHJ inversa (o transformada basada en fase inversa) se ha normalizado de acuerdo con la normalización en N3D (3D completa). De acuerdo con algunos ejemplos, la matriz de UHJ inversa (o transformada basada en fase inversa) se ha normalizado de acuerdo con la normalización en SN3D (seminormalización de Schmidt).

[0115] De acuerdo con algunos ejemplos, los coeficientes ambisónicos ambientales están asociados con funciones de base esférica que tienen un orden de cero o un orden de uno, y para aplicar la matriz de UHJ inversa (o

transformada basada en fase inversa), el dispositivo se configura para realizar una multiplicación escalar de la matriz de UHJ con respecto a la representación descorrelacionada de los coeficientes ambisónicos ambientales. En algunos ejemplos, para aplicar la transformada de recorrelación, el dispositivo se configura para aplicar una matriz de modo inverso a la representación descorrelacionada de los coeficientes ambisónicos ambientales. En algunos ejemplos, para generar la señal de altavoz, el dispositivo se configura para generar, para su emisión por un sistema de reproducción estéreo, una señal de altavoz izquierdo en base a la señal izquierda y una señal de altavoz derecho en base a la señal derecha.

[0116] En algunos ejemplos, para generar la señal de altavoz, el dispositivo se configura para usar la señal izquierda como una señal de altavoz izquierdo y la señal derecha como una señal de altavoz derecho sin aplicar una transformada de recorrelación a las señales derecha e izquierda. De acuerdo con algunos ejemplos, para generar la señal de altavoz, el dispositivo se configura para mezclar la señal izquierda y la señal derecha para su emisión por un sistema de audio mono. De acuerdo con algunos ejemplos, para generar la señal de altavoz, el dispositivo se configura para combinar los coeficientes ambisónicos ambientales correlacionados con uno o más canales de primer plano.

[0117] De acuerdo con algunos ejemplos, el dispositivo se configura además para determinar que no haya canales de primer plano disponibles con los que combinar los coeficientes ambisónicos ambientales correlacionados. En algunos ejemplos, el dispositivo se configura además para determinar que el campo sonoro se va a emitir por medio de un sistema de reproducción de audio mono, y para descodificar al menos un subconjunto de los coeficientes ambisónicos de orden superior descorrelacionados que incluyen datos para su emisión por el sistema de reproducción de audio mono. En algunos ejemplos, el dispositivo se configura además para obtener una indicación de que la representación descorrelacionada de los coeficientes ambisónicos ambientales se descorrelacionó con una transformada de descorrelación. De acuerdo con algunos ejemplos, el dispositivo incluye además una formación de altavoces configurados para emitir la señal de altavoz generada en base a la representación descorrelacionada de los coeficientes ambisónicos ambientales.

[0118] La FIG. 5 es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de codificación de audio, tal como el dispositivo de codificación de audio 20 mostrado en el ejemplo de la FIG. 3, en la realización de diversos aspectos de las técnicas de síntesis basadas en vectores descritas en la presente divulgación. Inicialmente, el dispositivo de codificación de audio 20 recibe los coeficientes HOA 11 (106). El dispositivo de codificación de audio 20 puede invocar la unidad de LIT 30, que puede aplicar una LIT con respecto a los coeficientes HOA para emitir los coeficientes HOA transformados (por ejemplo, en el caso de la SVD, los coeficientes HOA transformados pueden comprender los vectores $US[k]$ 33 y los vectores $V[k]$ 35) (107).

[0119] El dispositivo de codificación de audio 20 puede invocar, a continuación, la unidad de cálculo de parámetros 32 para realizar el análisis descrito anteriormente con respecto a cualquier combinación de los vectores $US[k]$ 33, vectores $US[k-1]$ 33, los vectores $V[k]$ y/o $V[k-1]$ 35, para identificar diversos parámetros de la manera descrita anteriormente. Es decir, la unidad de cálculo de parámetros 32 puede determinar al menos un parámetro en base a un análisis de los coeficientes HOA transformados 33/35 (108).

[0120] El dispositivo de codificación de audio 20 puede a continuación invocar la unidad de reordenamiento 34, que puede reordenar los coeficientes HOA transformados (que, de nuevo en el contexto de la SVD, puede hacer referencia a los vectores $US[k]$ 33 y a los vectores $V[k]$ 35) en base al parámetro para generar los coeficientes HOA transformados reordenados 33'/35' (o, en otras palabras, los vectores $US[k]$ 33' y los vectores $V[k]$ 35'), como se describe anteriormente (109). El dispositivo de codificación de audio 20 puede, durante cualquiera de las operaciones anteriores u operaciones subsiguientes, invocar también la unidad de análisis del campo sonoro 44. La unidad de análisis del campo sonoro 44 puede, como se describe anteriormente, realizar un análisis de campo sonoro con respecto a los coeficientes HOA 11 y/o los coeficientes HOA transformados 33/35 para determinar el número total de canales de primer plano (nFG) 45, el orden del campo sonoro de fondo (N_{BG}) y el número ($nBGa$) e índices (i) de canales HOA de BG adicionales a enviar (que se pueden indicar conjuntamente como información de canal de fondo 43 en el ejemplo de la FIG. 3) (109).

[0121] El dispositivo de codificación de audio 20 también puede invocar la unidad de selección de fondo 48. La unidad de selección de fondo 48 puede determinar los coeficientes HOA de fondo o ambientales 47 en base a la información de canal de fondo 43 (110). El dispositivo de codificación de audio 20 puede invocar además la unidad de selección de primer plano 36, que puede seleccionar los vectores $US[k]$ reordenados 33' y los vectores $V[k]$ reordenados 35' que representan componentes de primer plano o distintos del campo sonoro en base a nFG 45 (que puede representar uno o más índices que identifican los vectores de primer plano) (112).

[0122] El dispositivo de codificación de audio 20 puede invocar la unidad de compensación de energía 38. La unidad de compensación de energía 38 puede realizar compensación de energía con respecto a los coeficientes HOA ambientales 47 para compensar la pérdida de energía debida a la eliminación de diversos de los coeficientes HOA por la unidad de selección de fondo 48 (114) y generar de este modo coeficientes HOA ambientales compensados en energía 47'.

[0123] El dispositivo de codificación de audio 20 también puede invocar la unidad de interpolación espacio-temporal 50. La unidad de interpolación espacio-temporal 50 puede realizar interpolación espacio-temporal con respecto a los coeficientes HOA transformados reordenados 33'/35' para obtener las señales de primer plano interpoladas 49' (que también se pueden denominar las "señales de nFG interpoladas 49'") y la información direccional de primer plano restante 53 (que también se puede denominar los vectores "V[k] 53") (116). A continuación, el dispositivo de codificación de audio 20 puede invocar la unidad de reducción de coeficientes 46. La unidad de reducción de coeficientes 46 puede realizar reducción de coeficientes con respecto a los vectores V[k] de primer plano restantes 53, en base a la información de canal de fondo 43, para obtener información direccional de primer plano reducida 55 (que también se puede denominar los vectores V[k] de primer plano reducidos 55) (118).

[0124] El dispositivo de codificación de audio 20 puede a continuación invocar la unidad de cuantificación 52 para comprimir, de la manera descrita anteriormente, los vectores V[k] de primer plano reducidos 55 y generar los vectores V[k] de primer plano codificados 57 (120). El dispositivo de codificación de audio 20 también puede invocar la unidad de descorrelación 40' para aplicar la descorrelación de desplazamiento de fase para reducir o eliminar la correlación entre las señales de fondo de los coeficientes HOA 47' para formar uno o más coeficientes HOA descorrelacionados 47" (121).

[0125] El dispositivo de codificación de audio 20 también puede invocar la unidad codificadora de audio psicoacústico 40. La unidad codificadora de audio psicoacústico 40 puede codificar psicoacústicamente cada vector de los coeficientes HOA ambientales compensados en energía 47' y las señales de nFG interpoladas 49' para generar coeficientes HOA ambientales codificados 59 y señales de nFG codificadas 61. A continuación, el dispositivo de codificación de audio puede invocar la unidad de generación de flujo de bits 42. La unidad de generación de flujo de bits 42 puede generar el flujo de bits 21 en base a la información direccional de primer plano codificada 57, los coeficientes HOA ambientales codificados 59, las señales de nFG codificadas 61 y la información de canal de fondo 43.

[0126] La FIG. 6 es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de descodificación de audio, tal como el dispositivo de descodificación de audio 24 mostrado en la FIG. 4, en la realización de diversos aspectos de las técnicas descritas en la presente divulgación. Inicialmente, el dispositivo de descodificación de audio 24 puede recibir el flujo de bits 21 (130). Tras recibir el flujo de bits, el dispositivo de descodificación de audio 24 puede invocar la unidad de extracción 72. Suponiendo, para propósitos de análisis, que el flujo de bits 21 indica que se va a realizar una reconstrucción basada en vectores, la unidad de extracción 72 puede analizar sintácticamente el flujo de bits para recuperar la información mencionada anteriormente, pasando la información a la unidad de reconstrucción basada en vectores 92.

[0127] En otras palabras, la unidad de extracción 72 puede extraer la información direccional de primer plano codificada 57 (que, de nuevo, también se puede denominar los vectores V[k] de primer plano codificados 57), los coeficientes HOA ambientales codificados 59 y las señales de primer plano codificadas (que también se pueden denominar las señales de nFG de primer plano codificadas 59 o los objetos de audio de primer plano codificados 59) del flujo de bits 21, de la manera descrita anteriormente (132).

[0128] El dispositivo de descodificación de audio 24 puede invocar además la unidad de descuantificación 74. La unidad de descuantificación 74 puede descodificar por entropía y descuantificar la información direccional de primer plano codificada 57 para obtener información direccional de primer plano reducida 55_k (136). El dispositivo de descodificación de audio 24 puede invocar la unidad de recorrelación 81. La unidad de recorrelación 81 puede aplicar una o más transformadas de recorrelación a los coeficientes HOA ambientales compensados en energía 47' para obtener uno o más coeficientes HOA recorrelacionados 47" (o coeficientes HOA correlacionados 47") y puede pasar los coeficientes HOA correlacionados 47" a la unidad de formulación de coeficientes HOA 82 (opcionalmente, a través de la unidad de fundido 770) (137). El dispositivo de descodificación de audio 24 también puede invocar la unidad de descodificación psicoacústica 80. La unidad de descodificación de audio psicoacústico 80 puede descodificar los coeficientes HOA ambientales codificados 59 y las señales de primer plano codificadas 61 para obtener coeficientes HOA ambientales compensados en energía 47' y las señales de primer plano interpoladas 49' (138). La unidad de descodificación psicoacústica 80 puede pasar los coeficientes HOA ambientales compensados en energía 47' a la unidad de fundido 770 y las señales de nFG 49' a la unidad de formulación de primer plano 78.

[0129] El dispositivo de descodificación de audio 24 puede invocar a continuación la unidad de interpolación espacio-temporal 76. La unidad de interpolación espacio-temporal 76 puede recibir la información direccional de primer plano reordenada 55_{k'} y realizar la interpolación espacio-temporal con respecto a la información direccional de primer plano reducida 55_k/55_{k-1} para generar la información direccional de primer plano interpolada 55_{k''} (140). La unidad de interpolación espacio-temporal 76 puede reenviar los vectores V[k] de primer plano interpolados 55_{k''} a la unidad de fundido 770.

[0130] El dispositivo de descodificación de audio 24 puede invocar la unidad de fundido 770. La unidad de fundido 770 puede recibir u obtener de otro modo elementos sintácticos (por ejemplo, desde la unidad de extracción 72) indicativos de cuándo los coeficientes HOA ambientales compensados en energía 47' están en transición (por ejemplo, el elemento sintáctico AmbCoeffTransition). La unidad de fundido 770 puede, en base a los elementos sintácticos de

transición y la información de estado de transición mantenida, efectuar un fundido de entrada o un fundido de salida de los coeficientes HOA ambientales compensados en energía 47', emitiendo coeficientes HOA ambientales ajustados 47" a la unidad de formulación de coeficientes HOA 82. La unidad de fundido 770 también puede, en base a los elementos sintácticos y la información de estado de transición mantenida, efectuar un fundido de entrada o un fundido de salida de los correspondientes uno o más elementos de los vectores $V[k]$ de primer plano interpolados $55_k''$, emitiendo los vectores $V[k]$ de primer plano ajustados $55_k''$ a la unidad de formulación de primer plano 78 (142).

[0131] El dispositivo de descodificación de audio 24 puede invocar la unidad de formulación de primer plano 78. La unidad de formulación de primer plano 78 puede realizar la multiplicación matricial de las señales de nFG 49' por la información direccional de primer plano ajustada $55_k'''$ para obtener los coeficientes HOA de primer plano 65 (144). El dispositivo de descodificación de audio 24 también puede invocar la unidad de formulación de coeficientes HOA 82. La unidad de formulación de coeficientes HOA 82 puede añadir los coeficientes HOA de primer plano 65 a los coeficientes HOA ambientales ajustados 47" para obtener los coeficientes HOA 11' (146).

[0132] La FIG. 6B es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de codificación de audio y un dispositivo de descodificación de audio en la realización de las técnicas de codificación descritas en la presente divulgación. La FIG. 6B es un diagrama de flujo que ilustra un proceso de codificación y descodificación de ejemplo 160, de acuerdo con uno o más aspectos de la presente divulgación. Aunque el proceso 160 se puede realizar por una variedad de dispositivos, para facilidad de análisis, el proceso 160 se describe en el presente documento con respecto al dispositivo de codificación de audio 20 y al dispositivo de descodificación de audio 24 descrito anteriormente. Las secciones de codificación y descodificación del proceso 160 se delimitan usando una línea discontinua en la FIG. 6B. El proceso 160 puede comenzar con uno o más componentes del dispositivo de codificación de audio 20 (por ejemplo, la unidad de selección de primer plano 36 y la unidad de selección de fondo 48) generando los canales de primer plano 164 y los canales de fondo de HOA de primer orden 166 desde una entrada de HOA usando codificación espacial de HOA (162). A su vez, la unidad de descorrelación 40' puede aplicar una transformada de descorrelación (por ejemplo, en la forma de una transformada o matriz de descorrelación basada en fase) a los coeficientes HOA ambientales compensados en energía 47'. Más específicamente, el dispositivo de codificación de audio 20 puede aplicar una matriz de UHJ o transformada de descorrelación basada en fase (por ejemplo, por multiplicación escalar) a los coeficientes HOA ambientales compensados en energía 47' (168).

[0133] En algunos ejemplos, la unidad de descorrelación 40' puede aplicar la matriz de UHJ (o transformada basada en fase) si la unidad de descorrelación 40', en los casos donde la unidad de descorrelación 40' determina que los canales de fondo de HOA incluyen un menor número de canales (por ejemplo, cuatro). A la inversa, en estos ejemplos, si la unidad de descorrelación 40' determina que los canales de fondo de HOA incluyen un mayor número de canales (por ejemplo, nueve), el dispositivo de codificación de audio 20 puede seleccionar y aplicar una transformada de descorrelación diferente de la matriz de UHJ (tal como una matriz de modo descrita en la norma MPEG-H) a los canales de fondo de HOA. Al aplicar la transformada de descorrelación (por ejemplo, la matriz de UHJ) a los canales de fondo de HOA, el dispositivo de codificación de audio 20 puede obtener canales de fondo de HOA descorrelacionados.

[0134] Como se muestra en la FIG. 6B, el dispositivo de codificación de audio 20 (por ejemplo, al invocar la unidad codificadora de audio psicoacústico 40) puede aplicar codificación temporal (por ejemplo, al aplicar AAC y/o USAC) a las señales de fondo de HOA descorrelacionadas (170) y a cualquier canal de primer plano (166). Se apreciará que, en algunos escenarios, la unidad codificadora de audio psicoacústico 40 puede determinar que el número de canales de primer plano puede ser cero (es decir, en estos escenarios, la unidad codificadora de audio psicoacústico 40 puede no obtener ningún canal de primer plano a partir de la entrada de HOA). Como es posible que AAC y/o USAC no estén optimizadas o de otro modo no sean adecuadas para los datos de audio estéreo, la unidad de descorrelación 40' puede aplicar la matriz de descorrelación para reducir o eliminar la correlación entre los canales de fondo de HOA. La correlación reducida mostrada en los canales de fondo de HOA descorrelacionados proporciona la ventaja potencial de mitigar o eliminar el desmascaramiento de ruido en la fase de codificación temporal de AAC/USAC, ya que es posible que AAC y USAC no estén optimizadas para datos de audio estéreo.

[0135] A su vez, el dispositivo de descodificación de audio 24 puede realizar la descodificación temporal de la salida de flujo de bits codificados por el dispositivo de codificación de audio 20. En el ejemplo del proceso 160, uno o más componentes del dispositivo de descodificación de audio 24 (por ejemplo, la unidad de descodificación psicoacústica 80) pueden realizar la descodificación temporal por separado con respecto a los canales de primer plano (si se incluye algún canal de primer plano en el flujo de bits) (172) y los canales de fondo (174). Adicionalmente, la unidad de recorrelación 81 puede aplicar una transformada de recorrelación a los canales de fondo de HOA descodificados temporalmente. Como un ejemplo, la unidad de recorrelación 81 puede aplicar la transformada de descorrelación de manera recíproca a la unidad de descorrelación 40'. Por ejemplo, como se describe en el ejemplo específico del proceso 160, la unidad de recorrelación 81 puede aplicar una matriz de UHJ o una transformada basada en fase a las señales de fondo de HOA descodificadas temporalmente (176).

[0136] En algunos ejemplos, la unidad de recorrelación 81 puede aplicar la matriz de UHJ o transformada basada en fase, si la unidad de recorrelación 81 determina que los canales de fondo de HOA descodificados temporalmente incluyen un menor número de canales (por ejemplo, cuatro). A la inversa, en estos ejemplos, si la unidad de recorrelación 81

determina que los canales de fondo de HOA decodificados temporalmente incluyen un mayor número de canales (por ejemplo, nueve), la unidad de correlación 81 puede seleccionar y aplicar una transformada de descorrelación diferente de la matriz de UHJ (tal como la matriz de modo descrita en la norma MPEG-H) a los canales de fondo de HOA.

5 **[0137]** Adicionalmente, la unidad de formulación de coeficientes HOA 82 puede realizar decodificación espacial de HOA de los canales de fondo de HOA correlacionados, y cualquier canal de primer plano decodificado disponible (178). A su vez, la unidad de formulación de coeficientes HOA 82 puede reproducir las señales de audio decodificadas a uno o más dispositivos de salida (180), tal como altavoces y/o auriculares (incluyendo, pero no limitado a, dispositivos de salida con estéreo o capacidades de sonido envolvente).

15 **[0138]** Las técnicas anteriores se pueden realizar con respecto a cualquier número de contextos y ecosistemas de audio diferentes. A continuación se describen una serie de contextos de ejemplo, aunque las técnicas se deberían limitar a los contextos de ejemplo. Un ecosistema de audio de ejemplo puede incluir contenido de audio, estudios de cine, estudios de música, estudios de audio de juegos, contenido de audio basado en canales, motores de codificación, derivaciones de audio de juegos, motores de codificación/reproducción de audio de juegos y sistemas de suministro.

20 **[0139]** Los estudios de cine, los estudios de música y los estudios de audio de juegos pueden recibir contenido de audio. En algunos ejemplos, el contenido de audio puede representar la salida de una adquisición. Los estudios de cine pueden emitir contenido de audio basado en canales (por ejemplo, en 2.0, 5.1 y 7.1), tal como usando una estación de trabajo de audio digital (DAW). Los estudios de música pueden emitir contenido de audio basado en canales (por ejemplo, en 2.0 y 5.1), tal como usando una DAW. En cualquier caso, los motores de codificación pueden recibir y codificar el contenido de audio basado en canales en base a uno o más códecs (por ejemplo, AAC, AC3, Dolby True HD®, Dolby Digital Plus® y DTS Master Audio®) para su emisión mediante los sistemas de suministro. Los estudios de audio de juegos pueden emitir una o más derivaciones de audio de juegos, tal como usando una DAW. Los motores de codificación/reproducción de audio de juegos pueden codificar y/o reproducir las derivaciones de audio como contenido de audio basado en canales para su emisión por los sistemas de suministro. Otro contexto de ejemplo en el que se pueden realizar las técnicas comprende un ecosistema de audio que puede incluir objetos de audio de grabación difundida, sistemas de audio profesionales, captura de consumidores en el dispositivo, formato de audio de HOA, reproducción en el dispositivo, audio, televisión y accesorios para el consumidor, y sistemas de audio para automóviles.

35 **[0140]** Los objetos de audio de grabación difundida, los sistemas de audio profesionales y la captura de consumidores en el dispositivo pueden codificar todos su salida usando el formato de audio de HOA. De esta manera, el contenido de audio se puede codificar usando el formato de audio de HOA en una representación única que se puede reproducir usando la reproducción en el dispositivo, el audio, televisión y accesorios para el consumidor, y los sistemas de audio para automóviles. En otras palabras, la representación única del contenido de audio se puede reproducir en un sistema genérico de reproducción de audio (es decir, en lugar de requerir una configuración particular tal como 5.1, 7.1, etc.), tal como un sistema de reproducción de audio 16.

40 **[0141]** Otros ejemplos de contexto en los que se pueden realizar las técnicas incluyen un ecosistema de audio que puede incluir elementos de adquisición y elementos de reproducción. Los elementos de adquisición pueden incluir dispositivos de adquisición cableados y/o inalámbricos (por ejemplo, micrófonos Eigen), captura de sonido envolvente en el dispositivo y dispositivos móviles (por ejemplo, teléfonos inteligentes y tabletas). En algunos ejemplos, los dispositivos de adquisición cableados y/o inalámbricos se pueden acoplar al dispositivo móvil por medio de un canal(es) de comunicación cableado(s) y/o inalámbrico(s).

50 **[0142]** De acuerdo con una o más técnicas de la presente divulgación, el dispositivo móvil se puede usar para adquirir un campo sonoro. Por ejemplo, el dispositivo móvil puede adquirir un campo sonoro por medio de dispositivos de adquisición cableados y/o inalámbricos y/o la captura de sonido envolvente en el dispositivo (por ejemplo, una pluralidad de micrófonos integrados en el dispositivo móvil). A continuación, el dispositivo móvil puede codificar el campo sonoro adquirido en los coeficientes HOA para la reproducción por uno o más de los elementos de reproducción. Por ejemplo, un usuario del dispositivo móvil puede grabar (adquirir un campo sonoro de) un suceso en vivo (por ejemplo, una reunión, una conferencia, una obra de teatro, un concierto, etc.) y codificar la grabación en coeficientes HOA.

60 **[0143]** El dispositivo móvil también puede utilizar uno o más de los elementos de reproducción para reproducir el campo sonoro codificado por HOA. Por ejemplo, el dispositivo móvil puede decodificar el campo sonoro codificado por HOA y emitir una señal a uno o más de los elementos de reproducción, que hace que los uno o más de los elementos de reproducción recreen el campo sonoro. Como un ejemplo, el dispositivo móvil puede utilizar los canales de comunicación inalámbricos y/o inalámbricos para emitir la señal a uno o más altavoces (por ejemplo, formaciones de altavoces, barras de sonido, etc.). Como otro ejemplo, el dispositivo móvil puede utilizar soluciones de acoplamiento para emitir la señal a una o más estaciones de acoplamiento y/o uno o más altavoces acoplados (por ejemplo, sistemas de sonido en coches y/u hogares inteligentes). Como otro ejemplo, el dispositivo móvil puede utilizar la reproducción de auriculares para emitir la señal a un conjunto de auriculares, por ejemplo, para crear un sonido binaural realista.

[0144] En algunos ejemplos, un dispositivo móvil particular puede tanto adquirir un campo sonoro 3D como reproducir el mismo campo sonoro 3D en un momento posterior. En algunos ejemplos, el dispositivo móvil puede adquirir un campo sonoro 3D, codificar el campo sonoro 3D en HOA y transmitir el campo sonoro 3D codificado a uno o más de otros dispositivos (por ejemplo, otros dispositivos móviles y/u otros dispositivos no móviles) para su reproducción.

[0145] Aún otro contexto en el que las técnicas se pueden realizar incluye un ecosistema de audio que puede incluir contenido de audio, estudios de juegos, contenido de audio codificado, motores de reproducción y sistemas de suministro. En algunos ejemplos, los estudios de juego pueden incluir una o más DAW que pueden admitir la edición de señales de HOA. Por ejemplo, las una o más DAW pueden incluir acoples HOA y/o herramientas que se pueden configurar para funcionar con (por ejemplo, trabajar con) uno o más sistemas de audio de juegos. En algunos ejemplos, los estudios de juego pueden emitir nuevos formatos de derivación que admitan HOA. En cualquier caso, los estudios de juego pueden emitir contenido de audio codificado a los motores de reproducción que pueden reproducir un campo sonoro para su reproducción mediante los sistemas de suministro.

[0146] Las técnicas también se pueden realizar con respecto a los dispositivos de adquisición de audio ejemplares. Por ejemplo, las técnicas se pueden realizar con respecto a un micrófono Eigen que puede incluir una pluralidad de micrófonos que están configurados conjuntamente para grabar un campo sonoro 3D. En algunos ejemplos, la pluralidad de micrófonos del micrófono Eigen puede estar ubicada en la superficie de una bola sustancialmente esférica con un radio de aproximadamente 4 cm. En algunos ejemplos, el dispositivo de codificación de audio 20 se puede integrar en el micrófono Eigen para emitir un flujo de bits 21 directamente desde el micrófono.

[0147] Otro contexto de adquisición de audio ejemplar puede incluir un camión de producción que se puede configurar para recibir una señal desde uno o más micrófonos, tal como uno o más micrófonos Eigen. El camión de producción también puede incluir un codificador de audio, tal como el codificador de audio 20 de la FIG. 3.

[0148] El dispositivo móvil puede incluir también, en algunos casos, una pluralidad de micrófonos que se configuran conjuntamente para grabar un campo sonoro 3D. En otras palabras, la pluralidad de micrófonos puede tener diversidad X, Y, Z. En algunos ejemplos, el dispositivo móvil puede incluir un micrófono que se puede girar para proporcionar diversidad X, Y, Z con respecto a uno o más de otros micrófonos del dispositivo móvil. El dispositivo móvil también puede incluir un codificador de audio, tal como el codificador de audio 20 de la FIG. 3.

[0149] Un dispositivo de captura de vídeo robustecido se puede configurar además para grabar un campo sonoro 3D. En algunos ejemplos, el dispositivo de captura de vídeo robustecido se puede unir a un casco de un usuario que participa en una actividad. Por ejemplo, el dispositivo de captura de vídeo robustecido se puede unir a un casco de un usuario que realiza descenso de ríos. De esta manera, el dispositivo de captura de vídeo robustecido puede capturar un campo sonoro 3D que representa la acción alrededor del usuario (por ejemplo, agua que se estrella detrás del usuario, otro navegante que habla delante del usuario, etc.).

[0150] Las técnicas también se pueden realizar con respecto a un dispositivo móvil potenciado accesorio, que se puede configurar para grabar un campo sonoro 3D. En algunos ejemplos, el dispositivo móvil puede ser similar a los dispositivos móviles analizados anteriormente, con la adición de uno o más accesorios. Por ejemplo, un micrófono Eigen se puede unir al dispositivo móvil mencionado anteriormente para formar un dispositivo móvil potenciado accesorio. De esta manera, el dispositivo móvil potenciado accesorio puede capturar una versión de calidad superior del campo sonoro 3D que usando únicamente componentes de captura de sonido integrados en el dispositivo móvil potenciado accesorio.

[0151] Los dispositivos de reproducción de audio de ejemplo que pueden realizar diversos aspectos de las técnicas descritas en la presente divulgación se analizan adicionalmente a continuación. De acuerdo con una o más técnicas de la presente divulgación, los altavoces y/o las barras de sonido se pueden disponer en cualquier configuración arbitraria mientras se sigue reproduciendo un campo sonoro 3D. Además, en algunos ejemplos, los dispositivos de reproducción de auriculares se pueden acoplar a un decodificador 24 por medio de una conexión cableada o bien inalámbrica. De acuerdo con una o más técnicas de la presente divulgación, se puede utilizar una única representación genérica de un campo sonoro para reproducir el campo sonoro sobre cualquier combinación de altavoces, barras de sonido y dispositivos de reproducción de auriculares.

[0152] Un número de diferentes entornos de reproducción de audio de ejemplo también pueden ser adecuados para realizar diversos aspectos de las técnicas descritas en la presente divulgación. Por ejemplo, un entorno de reproducción de altavoces 5.1, un entorno de reproducción de altavoces 2.0 (por ejemplo, estéreo), un entorno de reproducción de altavoces 9.1 con altavoces frontales de altura completa, un entorno de reproducción de altavoces 22.2, un entorno de reproducción de altavoces 16.0, un entorno de reproducción de altavoces automotrices y un dispositivo móvil con entorno de reproducción por audífono pueden ser entornos adecuados para realizar diversos aspectos de las técnicas descritas en la presente divulgación.

[0153] De acuerdo con una o más técnicas de la presente divulgación, se puede utilizar una única representación genérica de un campo sonoro para reproducir el campo sonoro en cualquiera de los entornos de reproducción anteriores. Adicionalmente, las técnicas de la presente divulgación hacen posible que un reproductor reproduzca un

campo sonoro a partir de una representación genérica para la reproducción en entornos de reproducción distintos al descrito anteriormente. Por ejemplo, si las consideraciones de diseño prohíben la colocación apropiada de los altavoces de acuerdo con un entorno de reproducción de altavoces 7.1 (por ejemplo, si no es posible colocar un altavoz envolvente derecho), las técnicas de la presente divulgación hacen posible que un reproductor compense con los otros 6 altavoces de modo que la reproducción se pueda lograr en un entorno de reproducción de altavoces 6.1.

[0154] Además, un usuario puede ver un juego deportivo mientras utiliza los auriculares. De acuerdo con una o más técnicas de la presente divulgación, el campo sonoro 3D del juego deportivo se puede adquirir (por ejemplo, uno o más micrófonos Eigen se pueden colocar en y/o alrededor del estadio de béisbol), los coeficientes HOA correspondientes al campo sonoro 3D se pueden obtener y transmitir a un descodificador, el descodificador puede reconstruir el campo sonoro 3D en base a los coeficientes HOA y emitir el campo sonoro 3D reconstruido a un reproductor, el reproductor puede obtener una indicación en cuanto al tipo de entorno de reproducción (por ejemplo, los auriculares) y reproducir el campo sonoro 3D reconstruido en señales que hacen que los auriculares emitan una representación del campo sonoro 3D del juego deportivo.

[0155] En cada uno de los diversos casos descritos anteriormente, se debería entender que el dispositivo de codificación de audio 20 puede realizar un procedimiento o comprender de otro modo medios para realizar cada etapa del procedimiento para el cual se configura el dispositivo de codificación de audio 20 para realizar. En algunos casos, los medios pueden comprender uno o más procesadores. En algunos casos, los uno o más procesadores pueden representar un procesador de propósito especial configurado mediante instrucciones almacenadas en un medio de almacenamiento no transitorio legible por ordenador. En otras palabras, diversos aspectos de las técnicas en cada uno de los conjuntos de ejemplos de codificación pueden proporcionar un medio de almacenamiento no transitorio legible por ordenador que tiene almacenadas en el mismo instrucciones que, cuando se ejecutan, hacen que los uno o más procesadores realicen el procedimiento para el cual se ha configurado el dispositivo de codificación de audio 20 para realizar.

[0156] 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. Los medios de almacenamiento de datos pueden ser cualquier medio disponible al que se pueda acceder desde uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en la presente divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

[0157] Asimismo, en cada uno de los diversos casos descritos anteriormente, se debería entender que el dispositivo de descodificación de audio 24 puede realizar un procedimiento o comprender de otro modo medios para realizar cada etapa del procedimiento para el cual se configura el dispositivo de descodificación de audio 24 para realizar. En algunos casos, los medios pueden comprender uno o más procesadores. En algunos casos, los uno o más procesadores pueden representar un procesador de propósito especial configurado mediante instrucciones almacenadas en un medio de almacenamiento no transitorio legible por ordenador. En otras palabras, diversos aspectos de las técnicas en cada uno de los conjuntos de ejemplos de codificación pueden proporcionar un medio de almacenamiento no transitorio legible por ordenador que tiene almacenadas en el mismo instrucciones que, cuando se ejecutan, hacen que los uno o más procesadores realicen el procedimiento para el cual se ha configurado el dispositivo de descodificación de audio 24 para realizar.

[0158] 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. Sin embargo, se debería 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, están dirigidos a medios de almacenamiento tangibles no transitorios. El término disco, como se usa en el presente documento, incluye disco compacto (CD), disco láser, disco óptico, disco versátil digital (DVD), disco flexible y disco Blu-ray, donde algunos discos normalmente reproducen datos magnéticamente, mientras que otros discos emiten datos ópticamente con láseres. Las combinaciones de lo anterior también se deberían incluir dentro del alcance de los medios legibles por ordenador.

[0159] 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

descodificación, o incorporados en un códec combinado. Además, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.

5 **[0160]** 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 proporcionar por un grupo de unidades de hardware interoperativas, incluyendo uno o más procesadores, como se describe anteriormente, junto con software y/o firmware adecuados.

10 **[0161]** Se han descrito diversos aspectos de las técnicas. Estos y otros aspectos de las técnicas están dentro del alcance de las reivindicaciones siguientes.

15

REIVINDICACIONES

1. Un procedimiento que comprende:

5 obtener una representación descorrelacionada de los coeficientes ambisónicos ambientales que representan al menos una señal izquierda y una derecha, habiéndose extraído los coeficientes ambisónicos ambientales de una pluralidad de coeficientes ambisónicos de orden superior y representativos de un componente de fondo de un campo sonoro descrito por la pluralidad de coeficientes ambisónicos de orden superior, habiéndose descorrelacionado la representación descorrelacionada de los coeficientes ambisónicos ambientales usando una transformada basada en fase, en los que al menos uno de la pluralidad de coeficientes ambisónicos de orden superior está asociado con una función de base esférica que tiene un orden de uno o cero;

15 aplicar una transformada de recorrelación a la representación descorrelacionada de los coeficientes ambisónicos ambientales para obtener una pluralidad de coeficientes ambisónicos ambientales correlacionados; y

20 generar una señal de altavoz en base a la pluralidad de coeficientes ambisónicos ambientales correlacionados obtenidos a partir de la representación descorrelacionada de los coeficientes ambisónicos ambientales.

2. El procedimiento de la reivindicación 1, en el que aplicar la transformada de recorrelación comprende aplicar una transformada basada en fase inversa a los coeficientes ambisónicos ambientales.

25 3. El procedimiento de la reivindicación 2, en el que los coeficientes ambisónicos ambientales están asociados con funciones de base esférica que tienen un orden de cero o un orden de uno, y en el que aplicar la transformada basada en fase inversa comprende realizar una multiplicación escalar de la transformada basada en fase con respecto a la representación descorrelacionada de los coeficientes ambisónicos ambientales.

30 4. El procedimiento de la reivindicación 1, que comprende además obtener una indicación de que la representación descorrelacionada de los coeficientes ambisónicos ambientales se descorrelacionó con una transformada de descorrelación.

35 5. El procedimiento de la reivindicación 1, que comprende además obtener uno o más componentes espaciales que definen características espaciales de componentes de primer plano del campo sonoro, definidos los componentes espaciales en un dominio armónico esférico y generados al realizar una descomposición con respecto a la pluralidad de coeficientes ambisónicos de orden superior,

40 en el que generar la señal de altavoz comprende combinar los coeficientes ambisónicos ambientales correlacionados con uno o más canales de primer plano obtenidos en base al uno o más componentes espaciales.

6. Un dispositivo para procesar datos de audio, comprendiendo el dispositivo:

45 una memoria configurada para almacenar al menos una porción de los datos de audio que se van a procesar; y

uno o más procesadores configurados para:

50 obtener una representación descorrelacionada de coeficientes ambisónicos ambientales que representan al menos una señal izquierda y una derecha, habiéndose extraído los coeficientes ambisónicos ambientales de una pluralidad de coeficientes ambisónicos de orden superior y representativos de un componente de fondo de un campo sonoro descrito por la pluralidad de coeficientes ambisónicos de orden superior, habiéndose descorrelacionado la representación descorrelacionada de los coeficientes ambisónicos ambientales usando una transformada basada en fase, en los que al menos uno de la pluralidad de coeficientes ambisónicos de orden superior está asociado con una función de base esférica que tiene un orden de uno o cero;

55 aplicar una transformada de recorrelación a la representación descorrelacionada de los coeficientes ambisónicos ambientales para obtener una pluralidad de coeficientes ambisónicos ambientales correlacionados; y

60 generar una señal de altavoz en base a la representación descorrelacionada de los coeficientes ambisónicos ambientales.

65

7. El dispositivo de la reivindicación 6, en el que aplicar la transformada de recorrelación comprende aplicar una transformada basada en fase inversa a los coeficientes ambisónicos ambientales, en el que la transformada basada en fase inversa se ha normalizado de acuerdo con una normalización en N3D (3D completa).
- 5 8. El dispositivo de la reivindicación 6, en el que aplicar la transformada de recorrelación comprende aplicar una transformada basada en fase inversa a los coeficientes ambisónicos ambientales, en el que la transformada basada en fase inversa se ha normalizado de acuerdo con la normalización en SN3D (seminormalización de Schmidt).
- 10 9. El dispositivo de la reivindicación 6, en el que, para generar la señal de altavoz, el uno o más procesadores se configuran para generar, para su emisión por un sistema de reproducción estéreo, una señal de altavoz izquierdo en base a la señal izquierda y una señal de altavoz derecho en base a la señal derecha.
- 15 10. El dispositivo de la reivindicación 6, en el que, para generar la señal de altavoz, el uno o más procesadores se configuran para usar la señal izquierda como una señal de altavoz izquierdo y la señal derecha como una señal de altavoz derecho sin aplicar la transformada de recorrelación a las señales derecha e izquierda.
- 20 11. El dispositivo de la reivindicación 6, en el que, para generar la señal de altavoz, el uno o más procesadores se configuran para mezclar la señal izquierda y la señal derecha para su emisión por un sistema de audio mono.
- 25 12. El dispositivo de la reivindicación 6, en el que, para generar la señal de altavoz, el uno o más procesadores se configuran para combinar los coeficientes ambisónicos ambientales correlacionados con uno o más canales de primer plano.
- 30 13. El dispositivo de la reivindicación 6, en el que el uno o más procesadores se configuran además para determinar que no hay canales de primer plano disponibles con los que combinar los coeficientes ambisónicos ambientales correlacionados.
- 35 14. El dispositivo de cualquiera de las reivindicaciones 6 a 13, que comprende además un altavoz configurado para emitir la señal de altavoz generada en base a la representación descorrelacionada de los coeficientes ambisónicos ambientales.
- 40 15. Un dispositivo para comprimir datos de audio, comprendiendo el dispositivo:
una memoria configurada para almacenar al menos una porción de los datos de audio que se van a comprimir; y
uno o más procesadores configurados para:
45 aplicar una transformada de descorrelación basada en fase a los coeficientes ambisónicos ambientales que representan al menos una señal izquierda y una derecha para obtener una representación descorrelacionada de los coeficientes ambisónicos ambientales, habiéndose extraído los coeficientes ambisónicos ambientales de una pluralidad de coeficientes ambisónicos de orden superior y representativos de un componente de fondo de un campo sonoro descrito por la pluralidad de coeficientes ambisónicos de orden superior, en los que al menos uno de la pluralidad de coeficientes ambisónicos de orden superior está asociado con una función de base esférica que tiene un orden de uno o cero.
16. El dispositivo de la reivindicación 15, que comprende además una formación de micrófonos configurada para capturar los datos de audio que se van a comprimir.

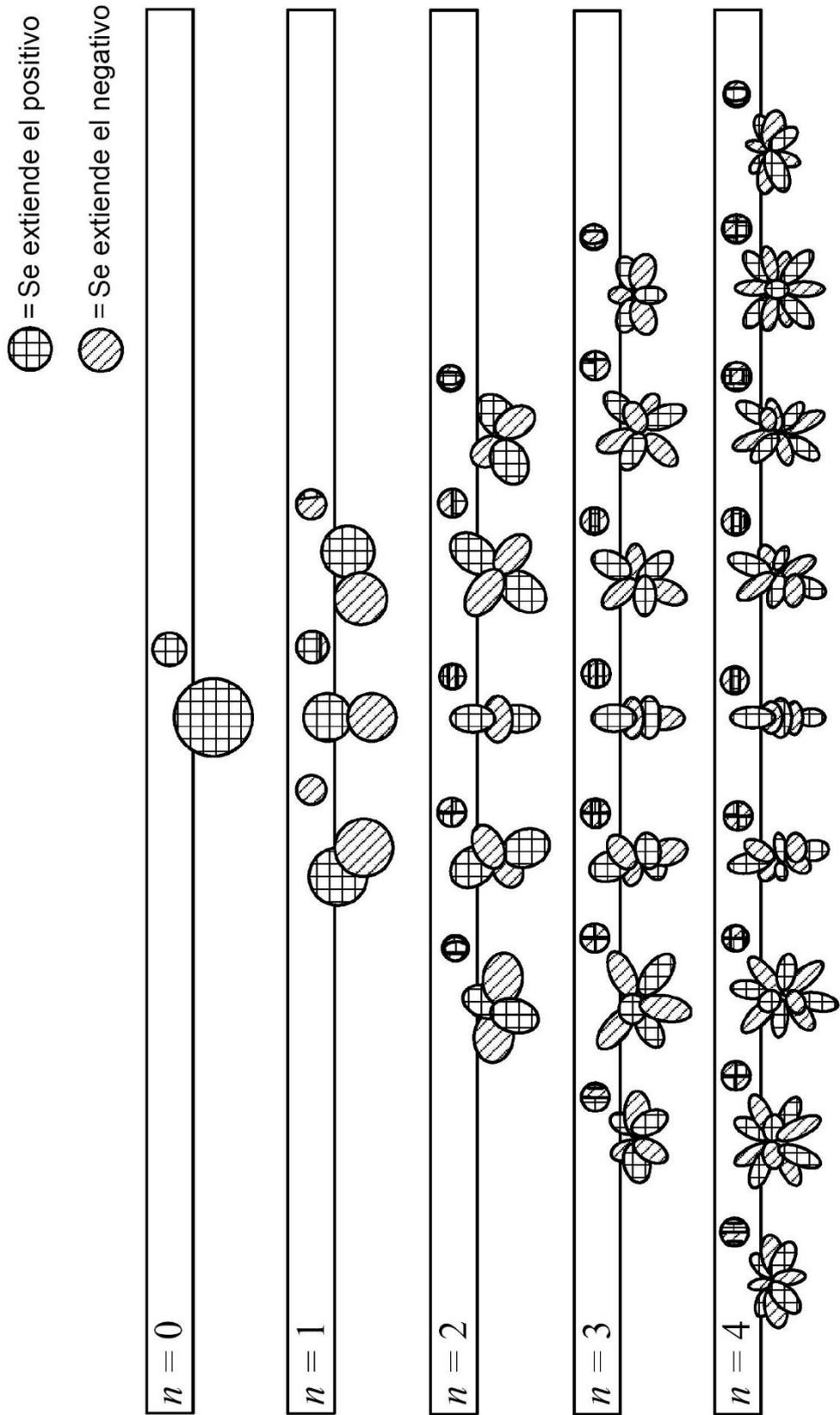


FIG. 1

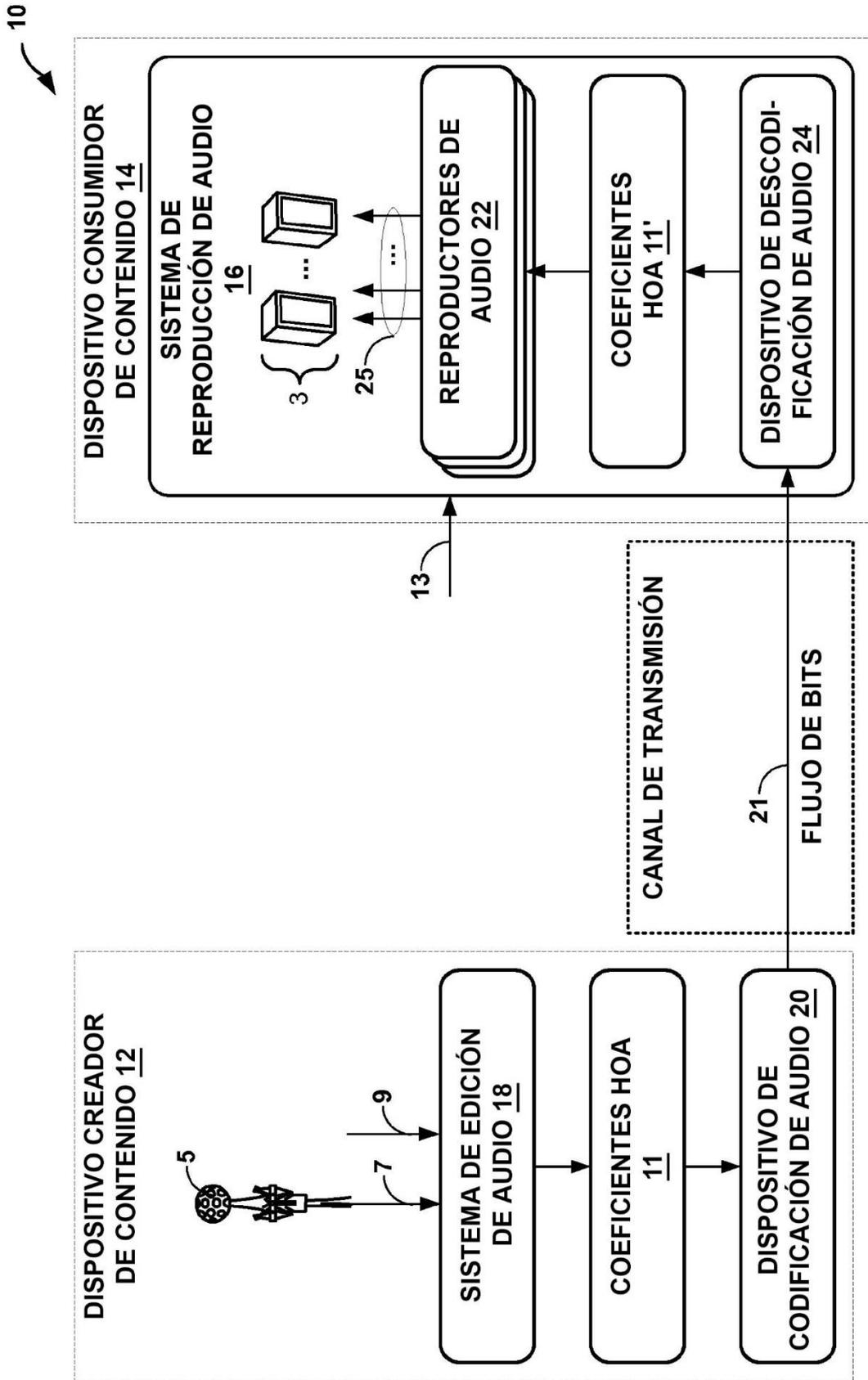


FIG. 2

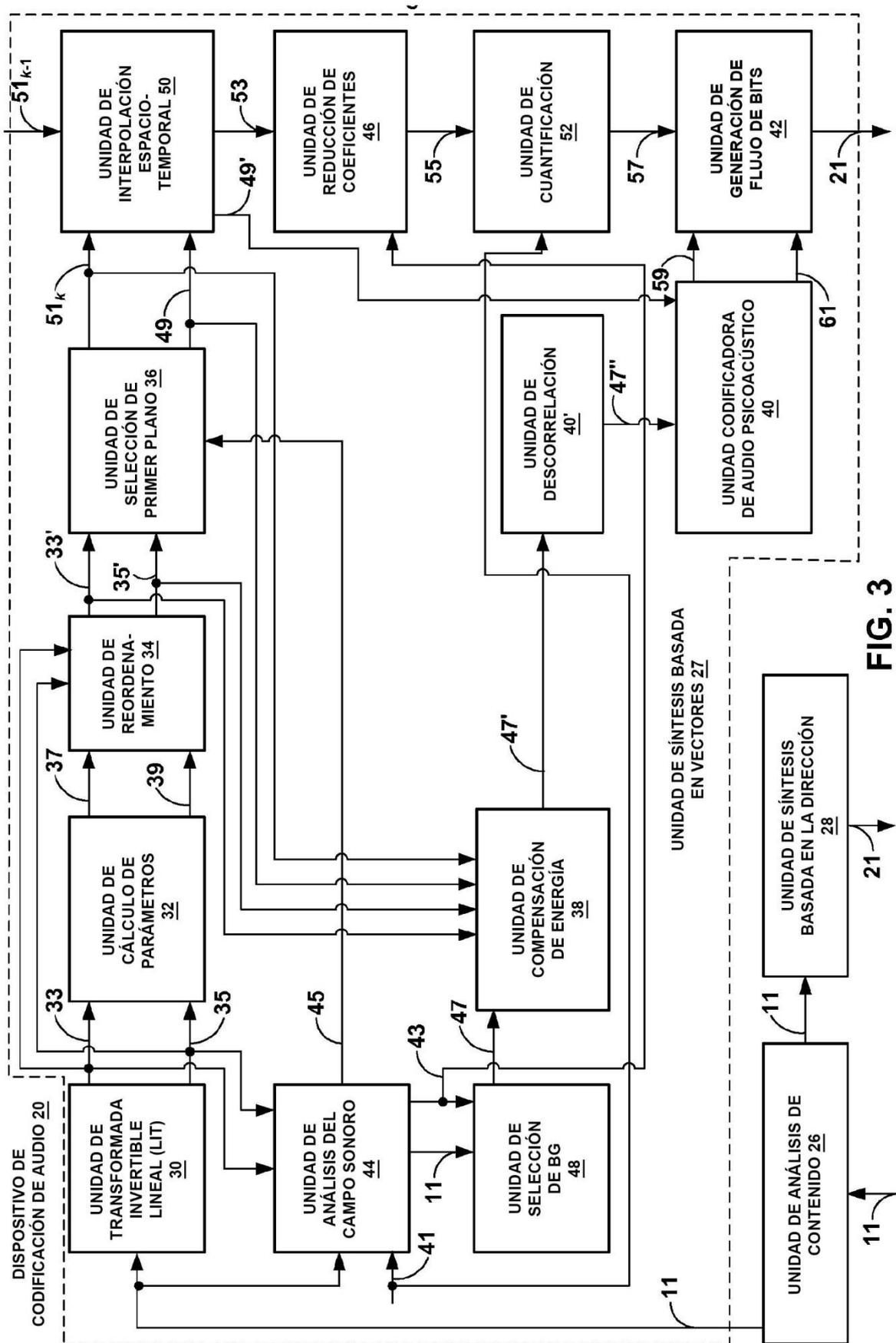


FIG. 3

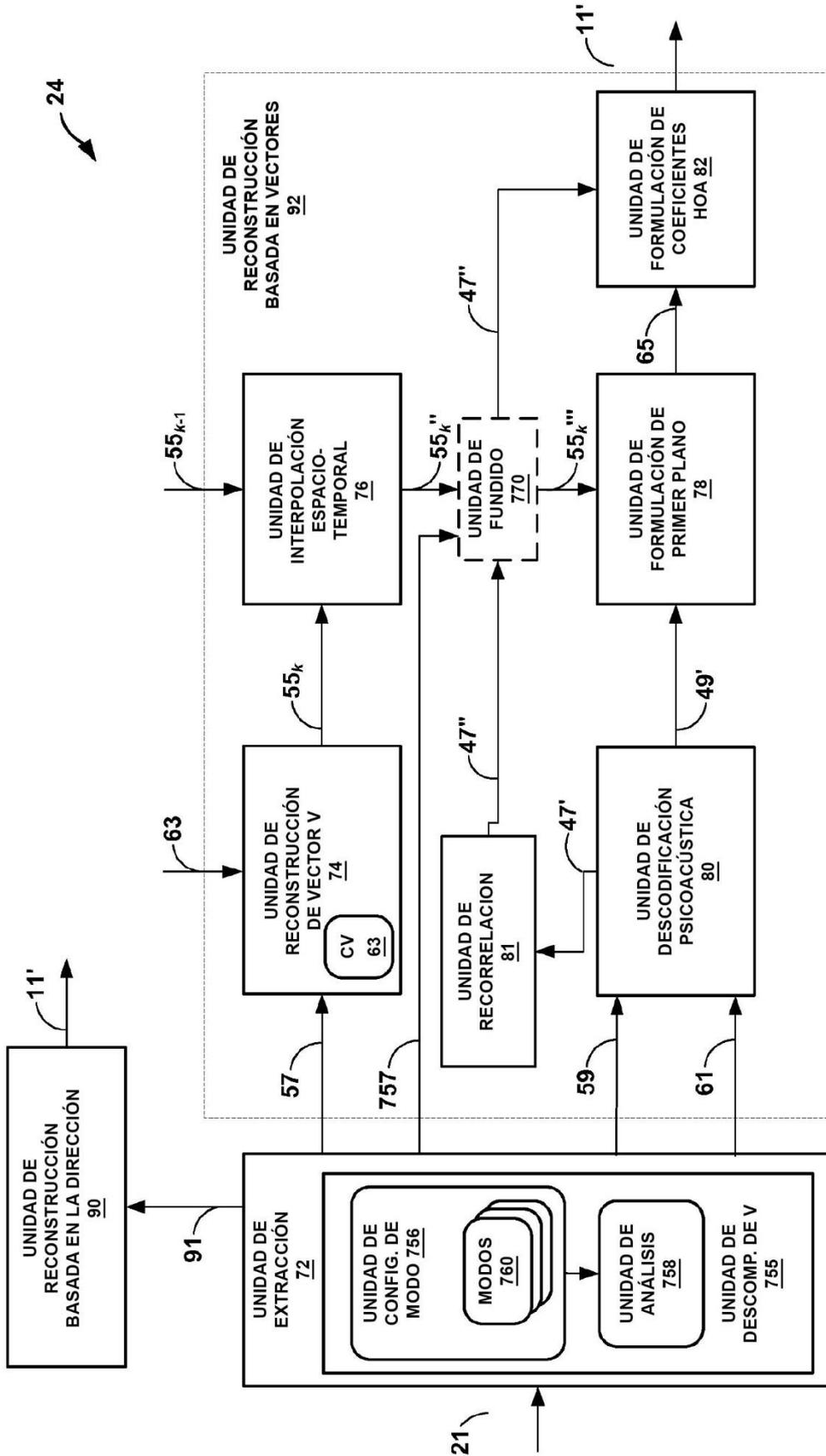


FIG. 4

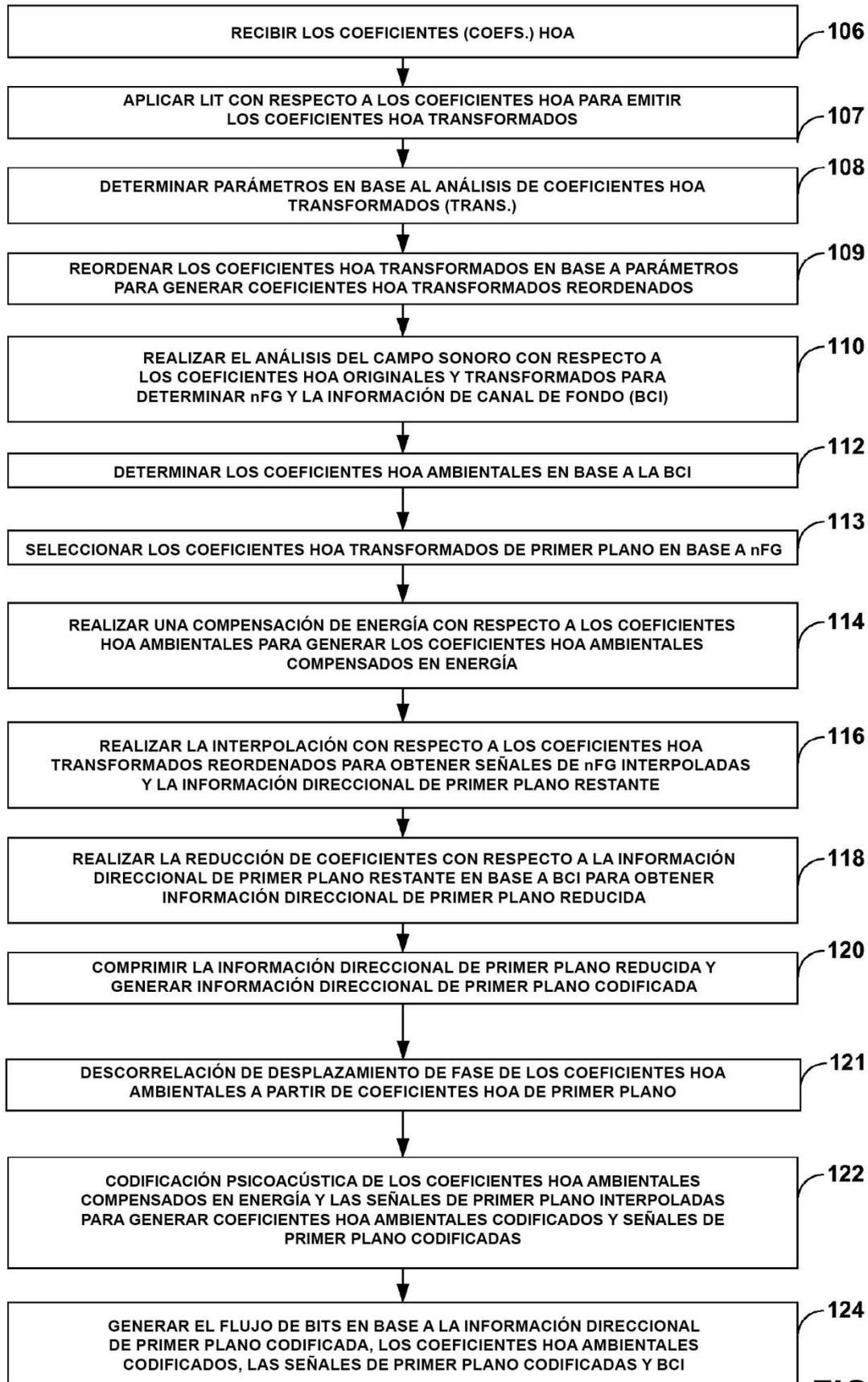


FIG. 5

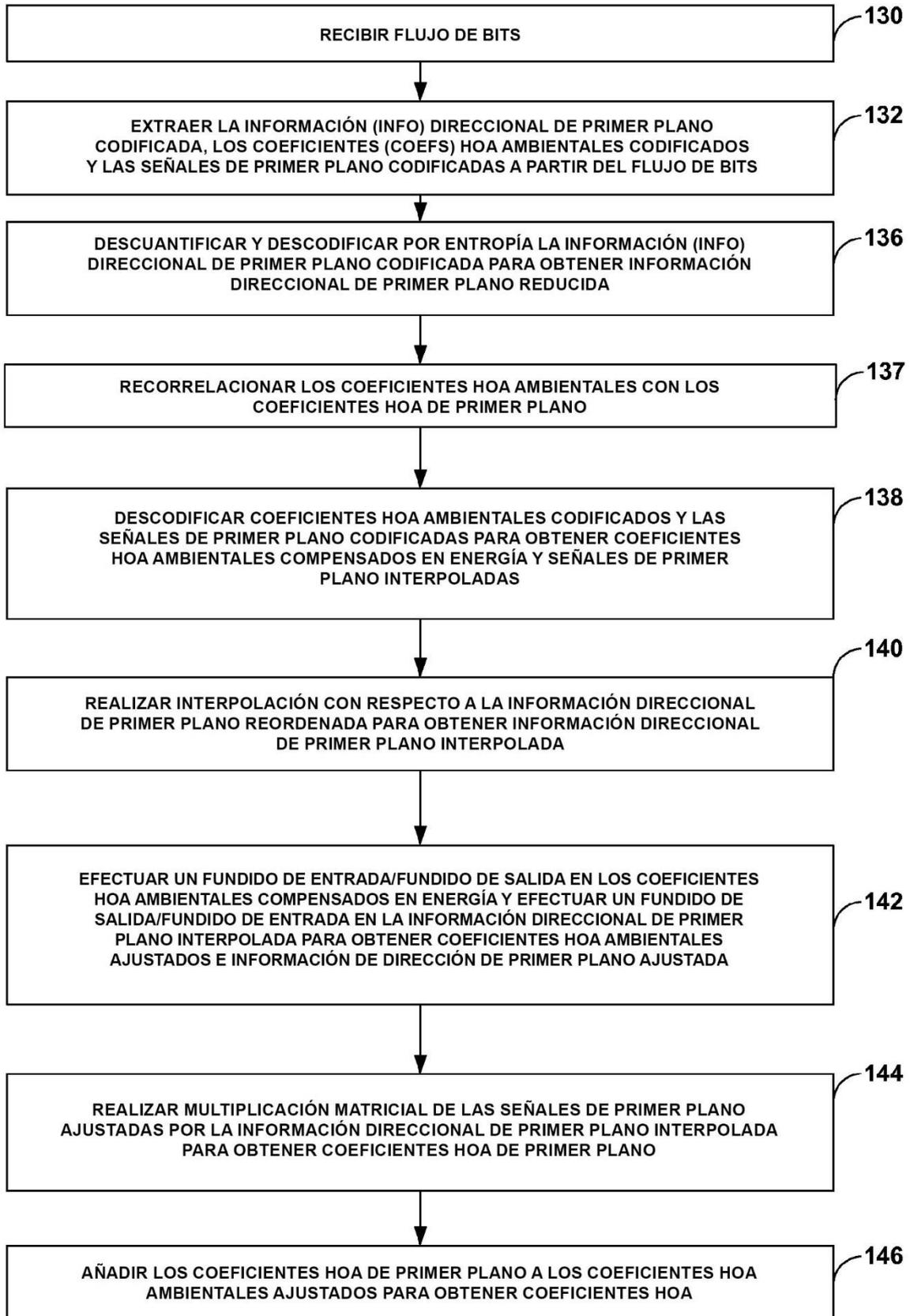


FIG. 6A

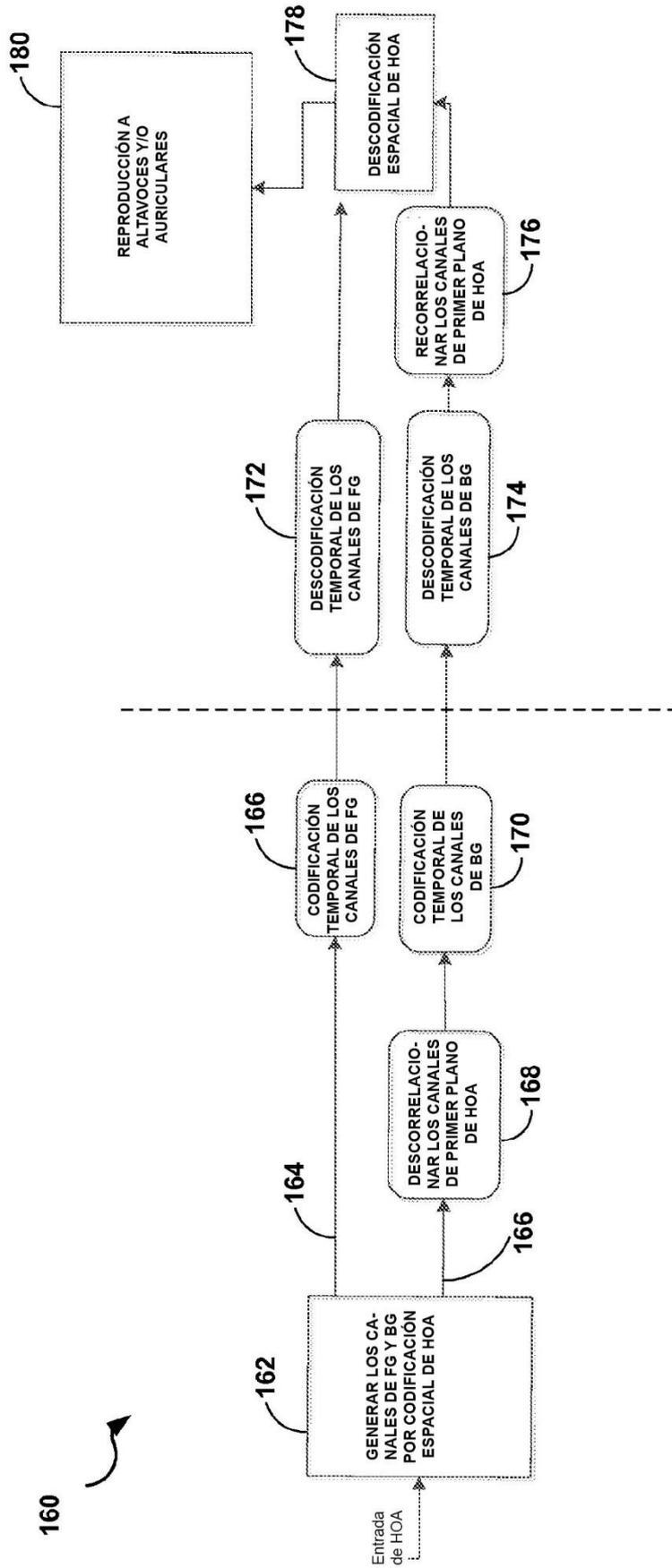


FIG. 6B