

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 637 922**

51 Int. Cl.:

H04S 3/02 (2006.01)

H04S 7/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.10.2014 PCT/EP2014/072411**

87 Fecha y número de publicación internacional: **30.04.2015 WO15059081**

96 Fecha de presentación y número de la solicitud europea: **20.10.2014 E 14786876 (4)**

97 Fecha y número de publicación de la concesión europea: **12.07.2017 EP 3061270**

54 Título: **Método y dispositivo para decodificar una representación de campo sonoro de audio ambisónico para reproducir audio utilizando configuraciones 2D**

30 Prioridad:

23.10.2013 EP 13290255

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

17.10.2017

73 Titular/es:

**DOLBY INTERNATIONAL AB (100.0%)
Apollo Building, 3E, Herikerbergweg 1-35
1101 CN Amsterdam Zuidoost, NL**

72 Inventor/es:

**KEILER, FLORIAN y
BOEHM, JOHANNES**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 637 922 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

Método y dispositivo para decodificar una representación de campo sonoro de audio ambisónico para reproducir audio utilizando configuraciones 2D

Campo de la invención

5 Esta invención se refiere a un método y a un aparato para decodificar una representación de campo sonoro de audio y, en particular, una representación de audio formateado ambisónico, para la reproducción de audio utilizando una configuración 2D o cercana a 2D.

Antecedentes

10 La localización precisa es un objetivo clave para cualquier sistema de reproducción de audio espacial. Tales sistemas de reproducción son muy aplicables para sistemas de conferencias, juegos u otros entornos virtuales que se benefician del sonido 3D. Se pueden sintetizar o capturar escenas de sonido en 3D como un campo sonoro natural. Señales de campo sonoro, tal como p. ej. el ambisónico, llevan una representación de un campo sonoro deseado. Se requiere un proceso de decodificación para obtener las señales de los altavoces individuales de una representación del campo sonoro. Decodificar una señal formateada ambisónica también se conoce como "representación". Con el fin de sintetizar escenas de audio, se requieren funciones de paneo que se refieren a la disposición de altavoces espacial para la obtención de una localización espacial de la fuente de sonido dada. Para la grabación de un campo sonoro natural, se requieren conjuntos de micrófonos para capturar la información espacial. El enfoque ambisónico es una herramienta muy adecuada para lograr esto. Las señales formateadas ambisónicas llevan una representación del campo sonoro deseado, basada en la descomposición armónica esférica del campo sonoro. Mientras que el formato ambisónico básico o formato-B utiliza armónicos esféricos de orden cero y uno, el llamado ambisónico de orden superior (HOA) utiliza también armónicos esféricos adicionales de al menos 2º orden. La disposición espacial de los altavoces se conoce como configuración de altavoces. Para el proceso de decodificación, se requiere una matriz de decodificación (también llamada matriz de representación), la cual es específica para una configuración de altavoces dada y la cual se genera utilizando las posiciones de altavoces conocidas.

15 Las configuraciones de altavoces utilizadas habitualmente son la configuración estéreo que utiliza dos altavoces, la configuración envolvente estándar que utiliza cinco altavoces y las extensiones de la configuración envolvente que utilizan más de cinco altavoces. Sin embargo, estas configuraciones bien conocidas se limitan a dos dimensiones (2D), p. ej., no se reproduce información de la altura. Representar para configuraciones de altavoces conocidas que pueden reproducir información de la altura tiene desventajas en la localización del sonido y la coloración: bien se perciben platillos verticales espaciales con sonoridad muy desigual, o señales del altavoz tienen fuertes lóbulos laterales, los cuales son desventajosos especialmente para posiciones de escucha descentradas. Por lo tanto, se prefiere un denominado diseño de representación de conservación de energía cuando se representa una descripción del campo sonoro HOA a los altavoces. Esto significa que la representación de una sola fuente de sonido resulta en señales de altavoz de energía constante, independientemente de la dirección de la fuente. En otras palabras, la energía de entrada llevada por la representación ambisónica es conservada por el representador del altavoz. La publicación de patente internacional WO2014/012945A1 [1] de los presentes inventores describe un diseño de representación de HOA con buenas propiedades de conservación de energía y de localización para configuraciones de altavoces 3D. Sin embargo, aunque este enfoque funciona bastante bien para configuraciones de altavoces 3D que cubren todas las direcciones, algunas direcciones de las fuentes son atenuadas para configuraciones de altavoces 2D (como p. ej. sonido envolvente 5.1). Esto aplica especialmente para las direcciones donde no hay altavoces colocados, p. ej., desde la parte superior.

20 En F. Zotter y M. Frank, "All-Round Ambisonic Panning and Decoding" [2], se añade un altavoz "imaginario" si hay un agujero en la envoltura convexa creada por los altavoces. Sin embargo, se omite la señal resultante para ese altavoz imaginario para su reproducción en el altavoz real. Por lo tanto, todavía se atenúa una señal de fuente de esa dirección (es decir, una dirección en donde no está colocado un altavoz real). Además, ese documento muestra el uso del altavoz imaginario para su uso sólo con VBAP (paneo vectorial basado en amplitud).

Sumario de la invención

5 Por lo tanto, es un problema pendiente diseñar presentadores ambisónicos de conservación de energía para configuraciones de altavoces 2D (2-dimensional), en donde las fuentes de sonido de direcciones donde no están colocados altavoces son menos atenuadas, o no atenuadas en absoluto. Las configuraciones de altavoces 2D se pueden clasificar como aquellas en las que los ángulos de elevación de los altavoces están dentro de un pequeño rango definido (p. ej., $< 10^\circ$), por lo que están cerca del plano horizontal.

10 La presente memoria describe una solución para representar/decodificar una representación de campo sonoro de audio formateado ambisónico para distribuciones de altavoces espaciales regulares o no regulares, en donde la representación/decodificación proporciona propiedades de localización y de coloración altamente mejoradas y conserva la energía, y en donde incluso se representa el sonido de direcciones en las que no está disponible un altavoz. Ventajosamente, el sonido de direcciones en las que no está disponible un altavoz se representa sustancialmente con la misma energía y sonoridad percibida que tendría si estuviese disponible un altavoz en la dirección respectiva. Por supuesto, no es posible una localización exacta de estas fuentes de sonido ya que no está disponible un altavoz en su dirección.

15 En particular, al menos algunas realizaciones descritas proporcionan una nueva manera de obtener la matriz de decodificación para la decodificación de los datos de campo sonoro en formato HOA. Puesto que al menos el formato HOA describe un campo sonoro que no está directamente relacionado con las posiciones de los altavoces y ya que las señales de los altavoces que se obtienen son necesariamente en un formato de audio basado en canales, la decodificación de señales HOA está siempre estrechamente relacionada con la representación de la señal de audio. En principio, lo mismo aplica también para otros formatos de campo sonoro de audio. Por lo tanto, la presente divulgación se refiere a la decodificación y representación de formatos de audio relacionados con el campo sonoro. Los términos matriz de decodificación y matriz de representación se utilizan como sinónimos.

25 Para obtener una matriz de decodificación para una configuración dada con buenas propiedades de conservación de energía, se añaden uno o más altavoces virtuales en posiciones donde no está disponible un altavoz. Por ejemplo, para la obtención de una matriz de decodificación mejorada para una configuración 2D, se añaden dos altavoces virtuales en la parte superior e inferior (correspondientes a ángulos de elevación $+ 90^\circ$ y $- 90^\circ$, con los altavoces 2D colocados aproximadamente a una elevación de 0°). Para esta configuración de altavoz 3D virtual, está diseñada una matriz de decodificación que satisface la propiedad de conservación de energía. Finalmente, los factores de ponderación de la matriz de decodificación de los altavoces virtuales se mezclan con ganancias constantes a los altavoces reales de la configuración 2D.

30 De acuerdo con una realización, una matriz de decodificación (o matriz de representación) para representar o decodificar una señal de audio en formato ambisónico a un conjunto de altavoces dado, se genera mediante la generación de una primera matriz de decodificación preliminar utilizando un método convencional y utilizando posiciones de altavoces modificadas, en donde las posiciones de altavoces modificadas incluyen las posiciones de altavoces de un conjunto de altavoces dado y al menos una posición de altavoz virtual adicional, y la mezcla de la primera matriz de decodificación preliminar, en donde los coeficientes relativos al por lo menos un altavoz virtual adicional se retiran y se distribuyen a los coeficientes relativos a los altavoces del conjunto de altavoces dado. En una realización, sigue un paso subsiguiente de normalización de la matriz de decodificación. La matriz de decodificación resultante es adecuada para la representación o la decodificación de la señal ambisónica al conjunto de altavoces dado, en donde incluso el sonido de posiciones en las que no está presente un altavoz se reproduce con energía de señal correcta. Esto es debido a la construcción de la matriz de decodificación mejorada. Preferiblemente, la primera matriz de decodificación preliminar es de conservación de energía.

35 En una realización, la matriz de decodificación tiene filas L y O_{3D} columnas. El número de filas corresponde al número de altavoces en la configuración de altavoces 2D y el número de columnas corresponde al número de coeficientes ambisónicos O_{3D} , el cual depende del HOA de orden N de acuerdo con $O_{3D} = (N + 1)^2$. Cada uno de los coeficientes de la matriz de decodificación para una configuración de altavoces 2D, es una suma de al menos un primer coeficiente intermedio y un segundo coeficiente intermedio. El primer coeficiente intermedio se obtiene por un método de diseño de la matriz 3D de conservación de energía para la posición de altavoz actual de la configuración de altavoces 2D, en donde el método de diseño de la matriz 3D de conservación de energía utiliza al menos una posición de altavoz virtual. El segundo coeficiente intermedio se obtiene por un coeficiente que se obtiene a partir de dicho método de diseño de la matriz 3D de conservación de energía para la al menos una posición de altavoz virtual, multiplicada con un factor de ponderación g . En una realización, el factor de ponderación g se calcula de acuerdo con $g = \frac{1}{\sqrt{L}}$, en donde L es el número de altavoces en la configuración de altavoces 2D.

En una realización, la invención se refiere a un medio de almacenamiento legible por ordenador que tiene almacenado en el mismo instrucciones ejecutables para hacer que un ordenador realice un método que comprende los pasos del método divulgado anteriormente o en las reivindicaciones.

En la reivindicación 9 se describe un aparato que utiliza el método.

- 5 Las realizaciones ventajosas se describen en las reivindicaciones dependientes, la siguiente descripción y las figuras.

Breve descripción de los dibujos

Se describen ejemplos de realización de la invención con referencia a los dibujos que se acompañan, los cuales muestran en

- 10 la Fig.1 un diagrama de flujo de un método de acuerdo con una realización;
 la Fig.2 una construcción ejemplar de una matriz de decodificación HOA mezclada;
 la Fig.3 un diagrama de flujo para obtener y modificar las posiciones de los altavoces;
 la Fig.4 un diagrama de bloques de un aparato de acuerdo con una realización;
 la Fig.5 la distribución de energía resultante de una matriz de decodificación convencional;
 15 la Fig.6 la distribución de energía resultante de una matriz de decodificación de acuerdo con las realizaciones; y
 la Fig.7 el uso de matrices de decodificación optimizadas por separado para diferentes bandas de frecuencia.

Descripción detallada de realizaciones

- 20 La Fig.1 muestra un diagrama de flujo de un método para decodificar una señal de audio, en particular, una señal de campo sonoro, de acuerdo con una realización. La decodificación de las señales de campo sonoro requiere generalmente las posiciones de los altavoces a los cuales será representada la señal de audio. Tales posiciones de altavoces $\hat{\Omega}_1 \dots \hat{\Omega}_L$ para L altavoces son la entrada i10 para el proceso. Tenga en cuenta que cuando se mencionan las posiciones, en realidad en este documento se entienden direcciones espaciales, es decir, las posiciones de los
 25 altavoces se definen por sus ángulos θ_i de inclinación y los ángulos azimutales φ_i , los cuales se combinan en un vector $\hat{\Omega}_i = [\theta_i, \varphi_i]^T$. Después, se añade 10 al menos una posición de un altavoz virtual. En una realización, todas las posiciones de los altavoces que se introducen en el proceso i10 están sustancialmente en el mismo plano, de forma que constituyen una configuración 2D, y el al menos un altavoz virtual que se añade está fuera de este plano. En una realización particularmente ventajosa, todas las posiciones de los altavoces que se introducen en el proceso i10
 30 están sustancialmente en el mismo plano y se añaden las posiciones de dos altavoces virtuales en el paso 10. Las posiciones ventajosas de los dos altavoces virtuales se describen a continuación. En una realización, la adición se realiza de acuerdo con la Ec. (6) de más abajo. El paso de adición 10 da como resultado un conjunto modificado de ángulos de altavoz $\hat{\Omega}'_1 \dots \hat{\Omega}'_{L+L_{virt}}$ en q10. L_{virt} es el número de altavoces virtuales. El conjunto modificado de los ángulos de altavoces se utiliza en un paso 11 de diseño de la matriz de decodificación 3D. También el HOA de orden N
 35 (generalmente del orden de los coeficientes de la señal de campo sonoro) necesita ser proporcionado i11 al paso 11.

- El paso 11 de diseño de la matriz de decodificación 3D realiza cualquier método conocido para generar una matriz de decodificación 3D. Preferiblemente, la matriz de decodificación 3D es adecuada para un tipo de decodificación/representación de conservación de energía. Por ejemplo, se puede utilizar el método descrito en el
 40 documento PCT/EP2013/065034. El paso 11 de diseño de matriz de decodificación 3D da como resultado una matriz de decodificación, o matriz de representación D' que es adecuada para la representación de las señales de

altavoces $L' = L + L_{\text{virt}}$, con L_{virt} siendo el número de posiciones de altavoces virtuales que se añadieron en el paso 10 "adición de posiciones de altavoz virtual".

5 Dado que sólo están físicamente disponibles L altavoces, la matriz de decodificación \mathbf{D}' que resulta del paso 11 de de diseño de la matriz de decodificación 3D, necesita ser adaptada a los L altavoces en un paso 12 de mezcla. Este paso realiza la mezcla de la matriz de decodificación \mathbf{D}' , en donde son ponderados los coeficientes relativos a los altavoces virtuales y distribuidos a los coeficientes relativos a los altavoces existentes. Preferiblemente, los coeficientes de cualquier de orden HOA particular (es decir, columna de la matriz de decodificación \mathbf{D}') se ponderan y se añaden a los coeficientes del mismo orden HOA (es decir, la misma columna de la matriz de decodificación \mathbf{D}'). Un ejemplo es una mezcla de acuerdo con la Ec. (8) de más abajo. El paso 12 de mezcla da como resultado la matriz de decodificación 3D mezclada $\tilde{\mathbf{D}}$ que tiene L filas, es decir, menos filas que la matriz de decodificación \mathbf{D}' , pero tiene el mismo número de columnas que la matriz de decodificación \mathbf{D}' . En otras palabras, la dimensión de la matriz de decodificación \mathbf{D}' es $(L + L_{\text{virt}}) \times O_{3D}$, y la dimensión de la matriz de decodificación 3D mezclada $\tilde{\mathbf{D}}$ es $L \times O_{3D}$.

15 La Fig.2 muestra una construcción a modo de ejemplo de una matriz de decodificación HOA mezclada $\tilde{\mathbf{D}}$ a partir de una matriz de decodificación HOA \mathbf{D}' . La matriz de decodificación HOA \mathbf{D}' tiene $L + 2$ filas, lo cual significa que se han añadido dos posiciones de altavoces virtuales a las L posiciones de altavoces disponibles, y O_{3D} columnas, con $O_{3D} = (N + 1)^2$ y siendo N el orden HOA. En el paso 12 de mezcla, se ponderan los coeficientes de filas $L + 1$ y $L + 2$ de la matriz de decodificación HOA \mathbf{D}' y se distribuyen a los coeficientes de su respectiva columna, y se eliminan las filas $L + 1$ y $L + 2$. Por ejemplo, se ponderan los primeros coeficientes $d'_{L+1,1}$ y $d'_{L+2,1}$ de cada una de las filas $L + 1$ y $L + 2$ y se suman a los primeros coeficientes de cada una de las filas restantes, tales como $d'_{1,1}$. El coeficiente resultante $\tilde{d}_{1,1}$ de la matriz de decodificación HOA mezcla $\tilde{\mathbf{D}}$ es una función de $d'_{1,1}$, $d'_{L+1,1}$, $d'_{L+2,1}$ y del factor de ponderación g . De la misma manera, p. ej., el coeficiente $\tilde{d}_{2,1}$ resultante de la matriz de decodificación HOA mezclada $\tilde{\mathbf{D}}$ es una función de $d'_{2,1}$, $d'_{L+1,1}$, $d'_{L+2,1}$ y del factor de ponderación g , y el coeficiente resultante $\tilde{d}_{1,2}$ de la matriz de decodificación HOA mezclada $\tilde{\mathbf{D}}$ es una función de $d'_{1,2}$, $d'_{L+1,2}$, $d'_{L+2,2}$ y del factor de ponderación g .

25 Por lo general, la matriz de decodificación HOA mezclada $\tilde{\mathbf{D}}$ se normalizará en un paso de normalización 13. Sin embargo, este paso 13 es opcional ya que también podría ser utilizada una matriz de decodificación no normalizada para la decodificación de una señal de campo sonoro. En una realización, la matriz de decodificación HOA mezclada $\tilde{\mathbf{D}}$ se normaliza de acuerdo con la Ec. (9) de más abajo. El paso 13 de normalización da como resultado una la matriz de decodificación HOA mezclada normalizada \mathbf{D} , la cual tiene la misma dimensión $L \times O_{3D}$ que la matriz de decodificación HOA mezclada $\tilde{\mathbf{D}}$.

35 La matriz de decodificación HOA mezclada normalizada \mathbf{D} , se puede entonces utilizar en un paso 14 de decodificación de campo sonoro, donde se decodifica una señal i_{14} de campo sonoro de entrada a señales q_{14} de L altavoces. Por lo general, la matriz de decodificación HOA mezclada normalizada \mathbf{D} , no necesita ser modificada hasta que se modifica la configuración de los altavoces. Por lo tanto, en una realización la matriz de decodificación HOA mezclada normalizada \mathbf{D} se almacena en un almacenamiento de matriz de decodificación HOA.

La Fig.3 muestra detalles de cómo en una realización, se obtienen y modifican las posiciones de los altavoces. Esta realización comprende los pasos para determinar 101 las posiciones $\hat{\Omega}_1 \dots \hat{\Omega}_L$ de los L altavoces y un orden N de coeficientes de la señal de campo sonoro, determinar 102 a partir de las posiciones que los L altavoces están sustancialmente en un plano 2D, y generar 103 al menos una posición virtual $\hat{\Omega}'_{L+1}$ de un altavoz virtual.

40 En una realización, la al menos una posición virtual $\hat{\Omega}'_{L+1}$ es una de $\hat{\Omega}'_{L+1} = [0,0]^T$ y $\hat{\Omega}'_{L+1} = [\pi,0]^T$.

En una realización, se generan 103 dos posiciones virtuales $\hat{\Omega}'_{L+1}$ y $\hat{\Omega}'_{L+2}$ correspondientes a dos altavoces virtuales, con $\hat{\Omega}'_{L+1} = [0,0]^T$ y $\hat{\Omega}'_{L+2} = [\pi, 0]^T$.

45 De acuerdo con una realización, un método para decodificar una señal de audio codificada para L altavoces en posiciones conocidas comprende los pasos para determinar 101 las posiciones $\hat{\Omega}_1 \dots \hat{\Omega}_L$ de los L altavoces y un orden N de coeficientes de la señal de campo sonoro, determinar 102 a partir de las posiciones que los L altavoces que están sustancialmente en un plano 2D, generar 103 al menos una posición virtual $\hat{\Omega}'_{L+1}$ de un altavoz virtual, generar 11 una matriz de decodificación 3D \mathbf{D}' , en donde se utilizan las posiciones determinadas $\hat{\Omega}_1 \dots \hat{\Omega}_L$ de los L altavoces y la al menos una posición virtual $\hat{\Omega}'_{L+1}$, y la matriz de decodificación 3D \mathbf{D}' tiene coeficientes para dichas posiciones determinadas y de altavoces virtuales, mezclar 12 la matriz de decodificación 3D \mathbf{D}' , en donde se ponderan los

5 coeficientes para las posiciones de los altavoces virtuales y se distribuyen a los coeficientes relativos a las posiciones de altavoces determinados, y en donde se obtiene una matriz de decodificación 3D adaptada \hat{D} que tienen los coeficientes para las posiciones de los altavoces determinadas, y decodificar la señal de audio codificada i14 usando la matriz de decodificación 3D adaptada \hat{D} , en donde se obtiene una pluralidad de señales de altavoz q14 decodificadas.

En una realización, la señal de audio codificada es una señal de campo sonoro, p. ej., en formato HOA. En una realización, la al menos una posición virtual $\hat{\Omega}_{L+1}$ de un altavoz virtual es una de $\hat{\Omega}'_{L+1} = [0,0]^T$ y $\hat{\Omega}'_{L+2} = [\pi, 0]^T$. En una realización, los coeficientes para las posiciones de los altavoces virtuales se ponderan con un factor de ponderación $g = \frac{1}{\sqrt{L}}$.

10 En una realización, el método tiene un paso adicional de normalización de la matriz de decodificación 3D adaptada \hat{D} , en donde se obtiene la matriz de decodificación 3D adaptada normalizada D , y el paso de decodificación 14 de la señal de audio codificada i14 usa la matriz de decodificación 3D adaptada normalizada D . En una realización, el método tiene un paso adicional de almacenamiento de la matriz de decodificación 3D adaptada \hat{D} o de la matriz de decodificación HOA mezclada normalizada D , en un almacenamiento de la matriz de decodificación.

15 De acuerdo con una realización, una matriz de decodificación para la representación o la decodificación de una señal de campo sonoro para un conjunto de altavoces dado, se genera mediante la generación de una primera matriz de decodificación preliminar utilizando un método convencional y utilizando las posiciones de los altavoces modificadas, en donde las posiciones de los altavoces modificadas incluyen posiciones de altavoces del conjunto de altavoces dado y al menos una posición de altavoz virtual adicional, y se mezcla la primera matriz de decodificación preliminar, en donde los coeficientes relativos al por lo menos un altavoz virtual adicional se retiran y se distribuyen a los coeficientes relativos a los altavoces del conjunto de altavoces dado. En una realización, sigue un paso subsiguiente de normalización de la matriz de decodificación. La matriz de decodificación resultante es adecuada para la representación o la decodificación de la señal de campo sonoro para el conjunto de altavoces dado, en donde incluso el sonido de las posiciones en las que no está presente un altavoz se reproduce con energía de señal correcta. Esto es debido a la construcción de la matriz de decodificación mejorada. Preferiblemente, la primera matriz de decodificación preliminar es de conservación de energía.

La Fig.4 a) muestra un diagrama de bloques de un aparato de acuerdo con una realización. El aparato 400 para decodificar una señal de audio codificada en formato de campo sonoro para L altavoces en posiciones conocidas, comprende una unidad sumadora 410 para la adición de al menos una posición de al menos un altavoz virtual a las posiciones de los L altavoces, una unidad generadora de matrices de decodificación 411 para la generación de una matriz de decodificación 3D D' , en donde se utilizan las posiciones $\hat{\Omega}_1 \dots \hat{\Omega}_L$ de los L altavoces y la al menos una posición virtual $\hat{\Omega}_{L+1}$ y la matriz de decodificación 3D D' tiene los coeficientes para dichas posiciones determinadas y de altavoces virtuales, una unidad de mezcla de matrices 412 para mezclar la matriz de decodificación 3D D' , en donde se ponderan los coeficientes para las posiciones de los altavoces virtuales y se distribuyen a los coeficientes relativos a las posiciones de los altavoces determinados, y en donde se obtiene una matriz de decodificación 3D mezclada \hat{D} que tiene los coeficientes para las posiciones de los altavoces determinadas y la unidad de decodificación 414 para decodificar la señal de audio codificada usando la matriz de decodificación 3D mezclada \hat{D} , en donde se obtiene una pluralidad de señales de altavoces decodificadas.

40 En una realización, el aparato comprende además una unidad de normalización 413 para normalizar la matriz de decodificación 3D mezclada \hat{D} , en donde se obtiene una matriz de decodificación 3D mezclada normalizada D y la unidad de decodificación 414 utiliza la matriz de decodificación 3D mezclada normalizada D .

En una realización mostrada en la Fig.4 b), el aparato comprende además una primera unidad de determinación 4101, para determinar las posiciones (Ω_L) de los L altavoces y un orden N de coeficientes de la señal de campo sonoro, una segunda unidad de determinación 4102, para determinar a partir de las posiciones que los L altavoces están sustancialmente en un plano 2D, y una unidad generadora de posiciones de altavoces virtuales 4103, para generar al menos una posición virtual ($\hat{\Omega}_{L+1}$) de un altavoz virtual.

En una realización, el aparato comprende además una pluralidad de filtros de paso de banda 715b para separar la señal de audio codificada en una pluralidad de bandas de frecuencia, en donde se genera una pluralidad de matrices de decodificación 3D separadas D_b' 711b, una para cada una de las bandas de frecuencia, y se mezcla cada una de las matrices de decodificación 3D D_b' 712b y, opcionalmente, se normaliza por separado, y en donde la unidad de decodificación 714b decodifica cada una de las bandas de frecuencia por separado. En esta realización, el aparato comprende además una pluralidad de unidades sumadoras 716b, una para cada uno de los altavoces. Cada una de las unidades sumadoras añade las bandas de frecuencia relativas al altavoz respectivo.

5 Cada una de las unidades sumadoras 410, las unidades generadoras de matrices de decodificación 411, las unidades de mezcla de matrices 412, las unidades de normalización 413, las unidades de decodificación 414, las primeras unidades de determinación 4101, las segundas unidades de determinación 4102 y las unidades de generación de posiciones de altavoces virtuales 4103, pueden ser implementadas por uno o más procesadores y cada una de estas unidades puede compartir el mismo procesador con cualquier otra de éstas u otras unidades.

10 La Fig.7 muestra una realización que utiliza matrices de decodificación optimizadas separadas para las diferentes bandas de frecuencia de la señal de entrada. En esta realización, el método de decodificación comprende un paso de separación de la señal de audio codificada en una pluralidad de bandas de frecuencias utilizando filtros de paso de banda. Se genera una pluralidad de matrices de decodificación 3D \mathbf{D}_b separadas 711b, una para cada una de las bandas de frecuencia, y se mezcla 712b cada una de las matrices de decodificación 3D \mathbf{D}_b y, opcionalmente, se normaliza por separado. La decodificación 714b de la señal de audio codificada se realiza para cada una de las bandas de frecuencias por separado. Esto tiene la ventaja de que pueden ser tomadas en consideración las diferencias dependientes de la frecuencia en la percepción humana y puede dar lugar a diferentes matrices de decodificación para diferentes bandas de frecuencia. En una realización, se genera sólo una o más (pero no todas) de las matrices de decodificación mediante la adición de posiciones de altavoces virtuales y luego la ponderación y la distribución de sus coeficientes a los coeficientes de las posiciones de altavoces existentes como descrito anteriormente. En otra realización, cada una de las matrices de decodificación se genera mediante la adición de posiciones de altavoces virtuales y luego la ponderación y la distribución de sus coeficientes a los coeficientes de las posiciones de altavoces existentes como descrito anteriormente. Por último, se suman todas las bandas de frecuencia relativas al mismo altavoz en una unidad sumadora de banda de frecuencia 716b por cada uno de los altavoces, en una operación inversa a la división de banda de frecuencia.

25 Cada una de las unidades sumadoras 410, las unidades generadoras de matrices de decodificación 711b, las unidades de mezcla de matrices 712b, las unidades de normalización 713b, las unidades de decodificación 714b, las unidades sumadoras de banda de frecuencia 716b y las unidades de filtro de paso de banda 715b, pueden ser implementadas por uno o más procesadores y cada una de estas unidades puede compartir el mismo procesador con cualquier otra de éstas u otras unidades.

30 Un aspecto de la presente divulgación es obtener una matriz de representación para una configuración 2D con buenas propiedades de conservación de energía. En una realización, se añaden dos altavoces virtuales en la parte superior e inferior (ángulos de elevación + 90° y - 90° con los altavoces 2D colocados aproximadamente a una elevación de 0°). Para esta configuración de altavoz virtual 3D, se diseña una matriz de representación que satisface la propiedad de conservación de energía. Finalmente, se mezclan los factores de ponderación de la matriz de representación de los altavoces virtuales con ganancias constantes a los altavoces reales de la configuración 2D. A continuación, se describe la representación ambisónica (en particular HOA)

35 La representación ambisónica es el proceso de cálculo de las señales de altavoces de una descripción de campo sonoro ambisónico. A veces, también se le llama decodificación ambisónica. Se considera una representación de campo sonoro ambisónico 3D de orden N, donde el número de coeficientes es

$$O_{3D} = (N + 1)^2 \quad (1)$$

Los coeficientes para la muestra de tiempo t están representados por el vector $\mathbf{b}(t) \in \mathbb{C}^{O_{3D} \times 1}$ con O_{3D} elementos.

40 Con la matriz de representación $\mathbf{D} \in \mathbb{C}^{L \times O_{3D}}$ las señales de los altavoces para la muestra de tiempo t se calculan por

$$\mathbf{w}(t) = \mathbf{D} \mathbf{b}(t) \quad (2)$$

con $\mathbf{D} \in \mathbb{C}^{L \times O_{3D}}$ y $\mathbf{w} \in \mathbb{R}^{L \times 1}$ y siendo L el número de altavoces.

45 Las posiciones de los altavoces se definen por sus ángulos de inclinación θ_l y los ángulos azimutales ϕ_l , los cuales se combinan en un vector $\hat{\mathbf{q}}_l = [\theta_l, \phi_l]^T$ para $l = 1, \dots, L$. Las diferentes distancias de los altavoces de la posición de escucha se compensan mediante el uso de retardos individuales para los canales de los altavoces.

La energía de señal en el dominio HOA está dada por

$$E = \mathbf{b}^H \mathbf{b} \quad (3)$$

donde H indica la transpuesta (compleja conjugada). La energía correspondiente de las señales del altavoz se calcula por

50
$$\hat{E} = \mathbf{w}^H \mathbf{w} = \mathbf{b}^H \mathbf{D}^H \mathbf{D} \mathbf{b}. \quad (4)$$

La relación \hat{E}/E para una matriz de decodificación/procesamiento de conservación de energía, debe ser constante a fin de lograr la decodificación/representación de conservación de energía.

En principio, se propone la siguiente extensión para la representación 2D mejorada: Para el diseño de la representación de matrices para configuraciones de altavoces 2D, se añaden uno o más altavoces virtuales. Las configuraciones 2D se entienden como aquellas en las que los ángulos de elevación de los altavoces están dentro de un pequeño intervalo definido, de modo que están cerca del plano horizontal. Esto puede ser expresado por

$$5 \quad \left| \theta_l - \frac{\pi}{2} \right| \leq \theta_{thres2d}; \quad l = 1, \dots, L \quad (5)$$

En una realización, el valor umbral $\theta_{thres2d}$ se elige normalmente para corresponder a un valor en el intervalo de 5° a 10°.

Para el diseño de representación, se define un conjunto modificado de ángulos $\hat{\Omega}'_l$ de altavoces. Las últimas (en este ejemplo dos) posiciones de altavoces son las de dos altavoces virtuales en los polos norte y sur (en dirección vertical, es decir, superior e inferior) del sistema de coordenadas polares:

$$\hat{\Omega}'_l = \hat{\Omega}_l; \quad l = 1, \dots, L$$

$$\hat{\Omega}'_{L+1} = [0, 0]^T \quad (6)$$

$$\hat{\Omega}'_{L+2} = [\pi, 0]^T$$

Así, el nuevo número de altavoces usado para el diseño de representación es $L' = L + 2$. A partir de estas posiciones

de altavoces modificadas, se diseña una matriz de representación $\mathbf{D}' \in \mathbb{C}^{(L+2) \times O_{3D}}$ con un enfoque de conservación de energía. Por ejemplo, se puede utilizar el método de diseño descrito en [1]. Ahora, se deriva de \mathbf{D}' la matriz de representación final para la configuración de altavoces original. Una idea es mezclar los factores de ponderación para el altavoz virtual como se define en la matriz \mathbf{D}' con los altavoces reales. Se utiliza un factor de ganancia fijo que se elige como

$$g = \frac{1}{\sqrt{L'}} \quad (7)$$

Los coeficientes de la matriz intermedia $\tilde{\mathbf{D}} \in \mathbb{C}^{L \times O_{3D}}$ (también llamada matriz de decodificación 3D adaptada en este documento) se definen por

$$\tilde{d}_{l,q} = d'_{l,q} + g \cdot d'_{L+1,q} + g \cdot d'_{L+2,q} \quad \text{para } l = 1, \dots, L \quad \text{y } q = 1, \dots, O_{3D} \quad (8)$$

donde $\tilde{d}_{l,q}$ es el elemento de la matriz $\tilde{\mathbf{D}}$ en la fila l -ésima y la columna q -ésima. En un paso final opcional, la matriz intermedia (matriz de decodificación 3D adaptada) se normaliza usando la norma de Frobenius:

$$\mathbf{D} = \frac{\tilde{\mathbf{D}}}{\sqrt{\sum_{l=1}^L \sum_{q=1}^{O_{3D}} |\tilde{d}_{l,q}|^2}} \quad (9)$$

25 Las Figs. 5 y 6 muestran las distribuciones de energía para una configuración de altavoces de sonido envolvente 5.0. En ambas figuras, los valores de energía se muestran como escalas de grises y los círculos indican las posiciones de los altavoces. Con el método descrito, especialmente, se reduce claramente la atenuación en la parte superior (y también en la parte inferior, no mostrada aquí).

30 La Fig.5 muestra la distribución de energía resultante de una matriz de decodificación convencional. Los pequeños círculos alrededor del plano $z = 0$ representan las posiciones de los altavoces. Como puede verse, está cubierto un intervalo de energía de [-3,9, ..., 2,1] dB, lo que resulta en diferencias de energía de 6 dB. Además, las señales de la parte superior (y de la parte inferior, no visible) de la n-esfera se reproducen con energía muy baja, es decir, no audible, ya que aquí no están disponibles altavoces.

35 La Fig.6 muestra la distribución de energía resultante de una matriz de decodificación de acuerdo con una o más realizaciones, con la misma cantidad de altavoces que están en las mismas posiciones que en la Fig.5. Al menos se proporcionan las siguientes ventajas: primero, se cubre un intervalo de energía menor de [-1,6, ..., 0,8] dB, lo que resulta en diferencias de energía más pequeñas de únicamente 2,4 dB. Segundo, se reproducen las señales de todas las direcciones de la n-esfera con su energía correcta, incluso si aquí no están disponibles altavoces. Puesto que estas señales se reproducen a través de los altavoces disponibles, su localización no es correcta, pero las

40 señales son audibles con sonoridad correcta. En este ejemplo, las señales de la parte superior y en la parte inferior (no visible) se hacen audibles debido a la decodificación con la matriz de decodificación mejorada.

En una realización, un método para decodificar una señal de audio codificada en formato ambisónico para L altavoces en posiciones conocidas, comprende pasos de adición de al menos una posición de al menos un altavoz virtual a las posiciones de los L altavoces, generando una matriz de decodificación 3D \mathbf{D}' , en donde se utilizan las

5 posiciones $\hat{\Omega}_1, \dots, \hat{\Omega}_L$ de los L altavoces y la al menos una posición virtual $\hat{\Omega}'_{L+1}$, y la matriz de decodificación 3D \mathbf{D}' tiene coeficientes para dichas posiciones determinadas y de altavoces virtuales, mezclar la matriz de decodificación 3D \mathbf{D}' , en donde se ponderan los coeficientes para las posiciones de altavoces virtuales y se distribuyen a los coeficientes relativos a las posiciones de altavoces determinadas, y en donde se obtiene una matriz de decodificación 3D adaptada $\tilde{\mathbf{D}}$ que tiene los coeficientes para las posiciones de los altavoces determinadas, y la decodificación de la señal de audio codificada usando la matriz de decodificación 3D adaptada $\tilde{\mathbf{D}}$, en donde se obtiene una pluralidad de señales de altavoz decodificadas.

10 En otra realización, un aparato para decodificar una señal de audio codificada en formato ambisónico para L altavoces en posiciones conocidas, comprende una unidad sumadora 410 para la adición de al menos una posición de al menos un altavoz virtual a las posiciones de los L altavoces, una unidad generadora de matrices de decodificación 411 para generar una matriz de decodificación 3D \mathbf{D}' , en donde se utilizan las posiciones $\hat{\Omega}_1, \dots, \hat{\Omega}_L$ de los L altavoces y la al menos una posición virtual $\hat{\Omega}'_{L+1}$, y la matriz de decodificación 3D \mathbf{D}' tiene los coeficientes para dichas posiciones determinadas y de altavoces virtuales, una unidad de mezcla de matrices 412 para mezclar la matriz de decodificación 3D \mathbf{D}' , en donde se ponderan los coeficientes para las posiciones de los altavoces virtuales y se distribuyen a los coeficientes relativos a las posiciones de altavoces determinadas, y en donde se obtiene una matriz de decodificación 3D adaptada $\tilde{\mathbf{D}}$ que tiene los coeficientes para las posiciones de los altavoces determinadas, y una unidad de decodificación 414 para decodificar la señal de audio codificada usando la matriz de decodificación 3D adaptada $\tilde{\mathbf{D}}$, en donde se obtiene una pluralidad de señales de altavoz decodificadas.

20 En otra realización, un aparato para decodificar una señal de audio codificada en formato ambisónico para L altavoces en posiciones conocidas, comprende al menos un procesador y al menos una memoria, la memoria tiene almacenadas instrucciones que cuando se ejecutan en el procesador, implementan una unidad sumadora 410 para la adición de al menos una posición de al menos un altavoz virtual a las posiciones de los L altavoces, una unidad generadora de matrices de decodificación 411 para generar una matriz de decodificación 3D \mathbf{D}' , en donde se utilizan las posiciones $\hat{\Omega}_1, \dots, \hat{\Omega}_L$ de los L altavoces y la al menos una posición virtual $\hat{\Omega}'_{L+1}$, y la matriz de decodificación 3D \mathbf{D}' tiene coeficientes para dichas posiciones determinadas y de altavoces virtuales, una unidad de mezcla de matrices 412 para mezclar la matriz de decodificación 3D \mathbf{D}' , en donde se ponderan los coeficientes para las posiciones de los altavoces virtuales y se distribuyen a los coeficientes relativos a las posiciones de altavoces determinadas, y en donde se obtiene una matriz de decodificación 3D adaptada $\tilde{\mathbf{D}}$ que tiene los coeficientes para las posiciones de altavoces determinadas, y una unidad de decodificación 414 para decodificar la señal de audio codificada usando la matriz de decodificación 3D adaptada $\tilde{\mathbf{D}}$, en donde se obtiene una pluralidad de señales de altavoz decodificadas.

35 En otra realización, un medio de almacenamiento legible por ordenador que tiene almacenado en el mismo instrucciones ejecutables para hacer que un ordenador realice un método para decodificar una señal de audio codificada en formato ambisónico para L altavoces en posiciones conocidas, en donde el método comprende pasos de adición de al menos una posición de al menos un altavoz virtual a las posiciones de los L altavoces, generando una matriz de decodificación 3D \mathbf{D}' , en donde se utilizan las posiciones $\hat{\Omega}_1, \dots, \hat{\Omega}_L$ de los L altavoces y la al menos una posición virtual $\hat{\Omega}'_{L+1}$, y la matriz de decodificación 3D \mathbf{D}' tiene los coeficientes para dichas posiciones determinadas y de altavoces virtuales, mezclar la matriz de decodificación 3D \mathbf{D}' , en donde se ponderan los coeficientes para las posiciones de altavoces virtuales y se distribuyen a los coeficientes relativos a las posiciones de altavoces determinadas, y en donde se obtiene una matriz de decodificación 3D adaptada $\tilde{\mathbf{D}}$ que tiene los coeficientes para las posiciones de altavoces determinadas, y decodificar la señal de audio codificada usando la matriz de decodificación 3D adaptada $\tilde{\mathbf{D}}$, en donde se obtiene una pluralidad de señales de altavoz decodificadas. Otras realizaciones de medios de almacenamiento legibles por ordenador pueden incluir cualquier característica descrita anteriormente en las características particulares divulgadas en las reivindicaciones dependientes que se refieran a la reivindicación 1.

45 Se entenderá que la presente invención se ha descrito puramente a modo de ejemplo y se pueden hacer modificaciones de detalle sin apartarse del alcance de la invención. Por ejemplo, aunque se ha descrito sólo con respecto a HOA, la invención también se puede aplicar para otros formatos de audio de campo sonoro.

50 Cada una de las características descritas en la descripción y (donde sea apropiado) las reivindicaciones y los dibujos, puede ser proporcionada independientemente o en cualquier combinación apropiada. Las características pueden, en su caso, ser implementadas en hardware, software o una combinación de los dos. Los números de referencia que aparecen en las reivindicaciones son sólo a modo de ilustración y no tendrán efecto limitativo en el alcance de las reivindicaciones.

Las siguientes referencias se han citado anteriormente.

[1] Publicación de Patente Internacional n.º WO2014/012945A1 (PD120032)

[2] F. Zotter y M. Frank, " All-Round Ambisonic Panning and Decoding", J. Audio Eng. Soc., 2012, vol. 60, pp. 807-820

REIVINDICACIONES

1. Un método para decodificar una señal de audio codificada en formato ambisónico para L altavoces en posiciones determinadas conocidas, que comprende los pasos de
- añadir (10) al menos una posición de al menos un altavoz virtual a las posiciones de los L altavoces;
- 5 - generar (11) una matriz de decodificación 3D (\mathbf{D}^*), en donde se utilizan las posiciones ($\hat{\Omega}_1, \dots, \hat{\Omega}_L$) de los L altavoces y la al menos una posición virtual ($\hat{\Omega}_{L+1}$) y la matriz de decodificación 3D (\mathbf{D}^*) tiene los coeficientes para dichas posiciones determinadas y de altavoces virtuales;
- mezclar (12) la matriz de decodificación 3D (\mathbf{D}^*), en donde se ponderan los coeficientes para las posiciones de altavoces virtuales y se distribuyen a los coeficientes relativos a las posiciones de altavoces determinadas, y en donde se obtiene una matriz de decodificación 3D adaptada ($\tilde{\mathbf{D}}$) que tienen los coeficientes para las posiciones de altavoces determinadas; y
- 10 - decodificar (14) la señal de audio codificada (i14) usando la matriz de decodificación 3D adaptada ($\tilde{\mathbf{D}}$), en donde se obtiene una pluralidad de señales de altavoz decodificadas (q14).
2. El método de acuerdo con la reivindicación 1, en donde se ponderan los coeficientes para las posiciones de los
- 15 altavoces virtuales con un factor de ponderación $g = \frac{1}{\sqrt{L}}$, en donde L es el número de altavoces.
3. El método de acuerdo con la reivindicación 1 o 2, en donde la al menos una posición virtual ($\hat{\Omega}_{L+1}$) de un altavoz virtual es una de $\hat{\Omega}'_{L+1} = [0,0]^T$ y $\hat{\Omega}_{L+1} = [\pi,0]^T$.
4. El método de acuerdo con cualquiera de las reivindicaciones 1-3, que comprende además un paso de
- 20 normalización (13) de la matriz de decodificación 3D adaptada ($\tilde{\mathbf{D}}$) utilizando una norma de Frobenius, en donde se obtiene una matriz de decodificación 3D adaptada normalizada (\mathbf{D}), y el paso de decodificación (14) de la señal de audio codificada utiliza la matriz de decodificación 3D adaptada normalizada (\mathbf{D}).
5. El método de acuerdo con la reivindicación 4, en donde la normalización se realiza de acuerdo con
- $$\mathbf{D} = \frac{\tilde{\mathbf{D}}}{\sqrt{\sum_{l=1}^L \sum_{q=1}^{O_{3D}} |\tilde{d}_{l,q}|^2}}.$$
6. El método de acuerdo con cualquiera de las reivindicaciones 1-5, que comprende además los pasos de
- 25 - determinar (101) posiciones ($\hat{\Omega}_1, \dots, \hat{\Omega}_L$) de los L altavoces y un orden N de coeficientes de la señal de campo sonoro;
- determinar (102) a partir de las posiciones que los L altavoces están sustancialmente en un plano 2D; y
 - generar (103) al menos una posición virtual ($\hat{\Omega}_{L+1}$) de un altavoz virtual.
7. El método de acuerdo con cualquiera de las reivindicaciones 1-6, que comprende además un paso de separación
- 30 de la señal de audio codificada en una pluralidad de bandas de frecuencia utilizando filtros de paso de banda, en donde se genera una pluralidad de matrices de decodificación 3D separadas (\mathbf{D}_b^*) (711 b), una para cada una de las bandas de frecuencia, y cada una de las matrices de decodificación 3D (\mathbf{D}_b^*) se mezcla (712b) y opcionalmente se normaliza por separado (713b), y donde el paso de decodificación (714b) de la señal de audio codificada (i14) se realiza para cada una de las bandas de frecuencias por separado.
- 35 8. El método de acuerdo con cualquiera de las reivindicaciones 1-7, en donde las posiciones de los L altavoces conocidas están sustancialmente dentro de un plano 2D, con elevaciones de no más de 10°.
9. Un aparato para decodificar una señal de audio codificada en formato ambisónico para L altavoces en posiciones determinadas conocidas, que comprende
- la unidad sumadora (410) para añadir al menos una posición de al menos un altavoz virtual a las
- 40 posiciones de los L altavoces;
- la unidad generadora de matrices de decodificación (411) para generar una matriz de decodificación 3D (\mathbf{D}^*), en donde se utilizan las posiciones ($\hat{\Omega}_1, \dots, \hat{\Omega}_L$) de los L altavoces y la al menos una posición virtual ($\hat{\Omega}_{L+1}$) y la

- matriz de decodificación 3D (\mathbf{D}') que tiene los coeficientes para dichas posiciones determinadas y de altavoces virtuales;
- la unidad de mezcla de matrices (412) para mezclar la matriz de decodificación 3D (\mathbf{D}'), en donde se ponderan los coeficientes para las posiciones de los altavoces virtuales y se distribuyen a los coeficientes relativos a las posiciones de altavoces determinadas, y en donde se obtiene una matriz de decodificación 3D adaptada ($\tilde{\mathbf{D}}$) que tiene los coeficientes para las posiciones de altavoces determinadas; y
 - la unidad de decodificación (414) para decodificar la señal de audio codificada (i14) usando la matriz de decodificación 3D adaptada ($\tilde{\mathbf{D}}$), en donde se obtiene una pluralidad de señales de altavoz decodificadas (q14).
10. El aparato de acuerdo con la reivindicación 9, que comprende además una unidad de normalización (413) para normalizar la matriz de decodificación 3D adaptada ($\tilde{\mathbf{D}}$) utilizando una norma de Frobenius, en donde se obtiene una matriz de decodificación 3D adaptada normalizada (\mathbf{D}), y la unidad de decodificación (414) utiliza la matriz de decodificación 3D adaptada normalizada (\mathbf{D}).
11. El aparato de acuerdo con la reivindicación 9 o 10, que comprende además
- la primera unidad de determinación (101) para determinar las posiciones ($\hat{\Omega}_1, \dots, \hat{\Omega}_L$) de los L altavoces y un orden N de los coeficientes de la señal de campo sonoro;
 - la segunda unidad de determinación (102) para determinar a partir de las posiciones que los L altavoces están sustancialmente en un plano 2D; y
 - la unidad generadora de posiciones de altavoces virtuales (103) para generar al menos una posición virtual ($\hat{\Omega}_{L+1}$) de un altavoz virtual.
12. El aparato de acuerdo con una de las reivindicaciones 9-11, que comprende además una pluralidad de filtros de paso de banda (715b) para separar la señal de audio codificada en una pluralidad de bandas de frecuencia, en donde se genera una pluralidad de matrices de decodificación 3D separadas (\mathbf{D}_b') (711 b), una para cada una de las bandas de frecuencia, y se mezcla cada una de las matrices de decodificación 3D (\mathbf{D}_b') (712b) y opcionalmente se normaliza por separado, y en donde la unidad de decodificación (714b) decodifica cada una de las bandas de frecuencias por separado.
13. Un medio de almacenamiento legible por ordenador que tiene almacenado en el mismo instrucciones ejecutables para hacer que un ordenador realice un método para decodificar una señal de audio codificada en formato ambisónico para L altavoces en posiciones determinadas conocidas, el método comprende los pasos de
- añadir (10) al menos una posición de al menos un altavoz virtual a las posiciones de los L altavoces;
 - generar (11) una matriz de decodificación 3D (\mathbf{D}'), en donde se utilizan las posiciones ($\hat{\Omega}_1, \dots, \hat{\Omega}_L$) de los L altavoces y la al menos una posición virtual ($\hat{\Omega}_{L+1}$) y la matriz de decodificación 3D (\mathbf{D}') tiene los coeficientes para dichas posiciones determinadas y de altavoces virtuales;
 - mezclar (12) la matriz de decodificación 3D (\mathbf{D}'), en donde se ponderan los coeficientes para las posiciones de los altavoces virtuales y se distribuyen a los coeficientes relativos a las posiciones de altavoces determinadas, y en donde se obtiene una matriz de decodificación 3D adaptada ($\tilde{\mathbf{D}}$) que tienen los coeficientes para las posiciones de altavoces determinadas; y
 - decodificar (14) la señal de audio codificada (i14) usando la matriz de decodificación 3D adaptada ($\tilde{\mathbf{D}}$), en donde se obtiene una pluralidad de señales de altavoz decodificadas (q14).
14. El medio de almacenamiento legible por ordenador de acuerdo con la reivindicación 13, en donde se ponderan los coeficientes para las posiciones de altavoces virtuales con un factor de ponderación $g = \frac{1}{\sqrt{L}}$, en donde L es el número de altavoces.
15. El medio de almacenamiento legible por ordenador de acuerdo con la reivindicación 13 o 14, en donde la al menos una posición virtual ($\hat{\Omega}_{L+1}$) de un altavoz virtual es una de $\hat{\Omega}'_{L+1} = [0,0]^T$ y $\hat{\Omega}''_{L+1} = [\pi,0]^T$.

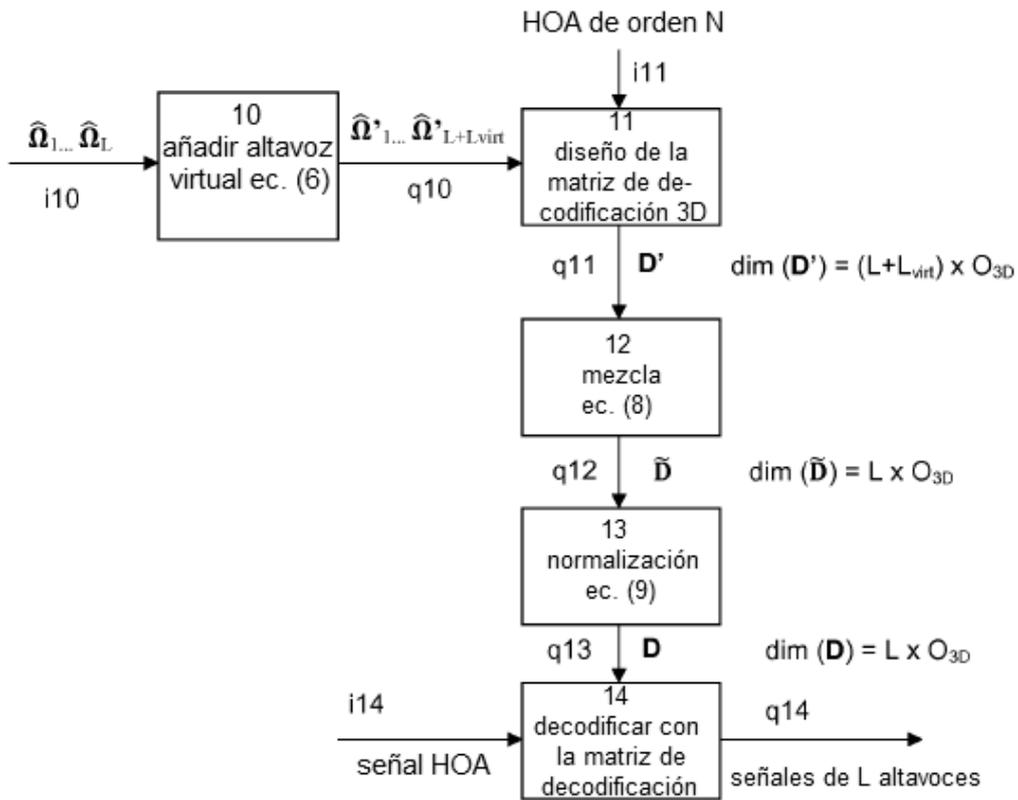


Fig.1

$$\mathbf{D}' = \begin{bmatrix} d'_{1,1} & d'_{1,2} & \dots & d'_{1,O3D} \\ \dots & \dots & \dots & \dots \\ d'_{L,1} & d'_{L,2} & \dots & d'_{L,O3D} \\ d'_{L+1,1} & d'_{L+1,2} & \dots & d'_{L+1,O3D} \\ d'_{L+2,1} & d'_{L+2,2} & \dots & d'_{L+2,O3D} \end{bmatrix} \rightarrow \tilde{\mathbf{D}} = \begin{bmatrix} \tilde{d}_{1,1} & \tilde{d}_{1,2} & \dots & \tilde{d}_{1,O3D} \\ \dots & \dots & \dots & \dots \\ \tilde{d}_{L,1} & \tilde{d}_{L,2} & \dots & \tilde{d}_{L,O3D} \end{bmatrix}$$

$$\mathbf{D}' = \begin{bmatrix} d'_{1,1} & d'_{1,2} & \dots & d'_{1,O3D} \\ \dots & \dots & \dots & \dots \\ d'_{L,1} & d'_{L,2} & \dots & d'_{L,O3D} \\ d'_{L+1,1} & d'_{L+1,2} & \dots & d'_{L+1,O3D} \\ d'_{L+2,1} & d'_{L+2,2} & \dots & d'_{L+2,O3D} \end{bmatrix} \rightarrow \tilde{\mathbf{D}} = \begin{bmatrix} \tilde{d}_{1,1} & \tilde{d}_{1,2} & \dots & \tilde{d}_{1,O3D} \\ \dots & \dots & \dots & \dots \\ \tilde{d}_{L,1} & \tilde{d}_{L,2} & \dots & \tilde{d}_{L,O3D} \end{bmatrix}$$

Fig.2

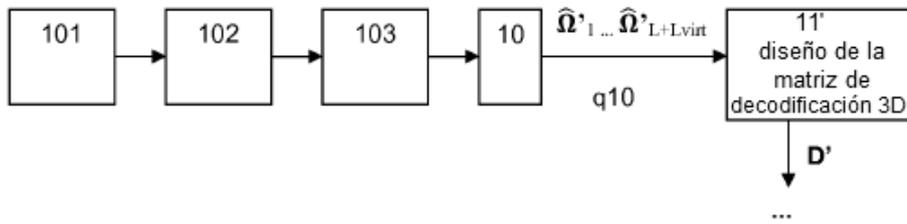
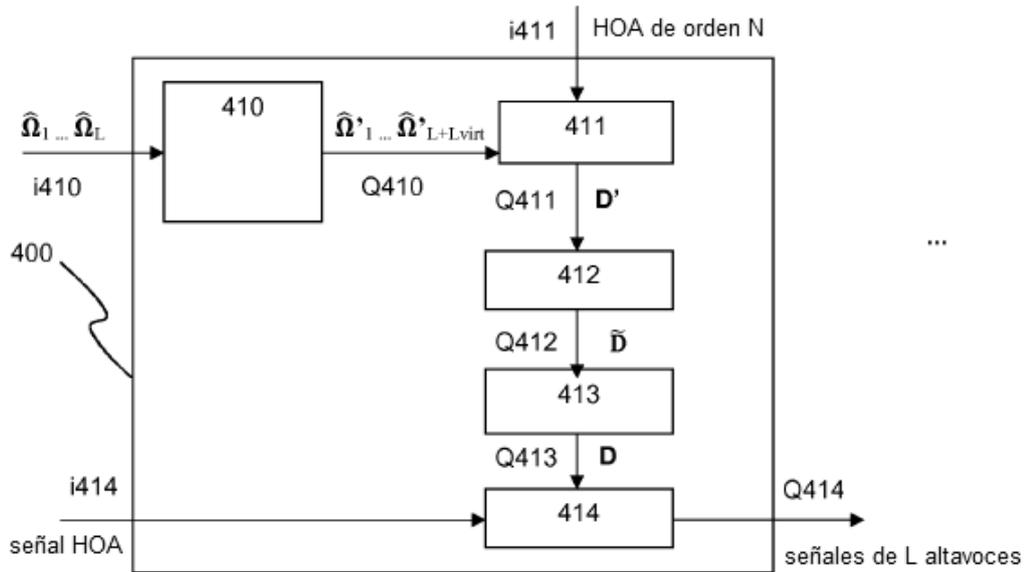
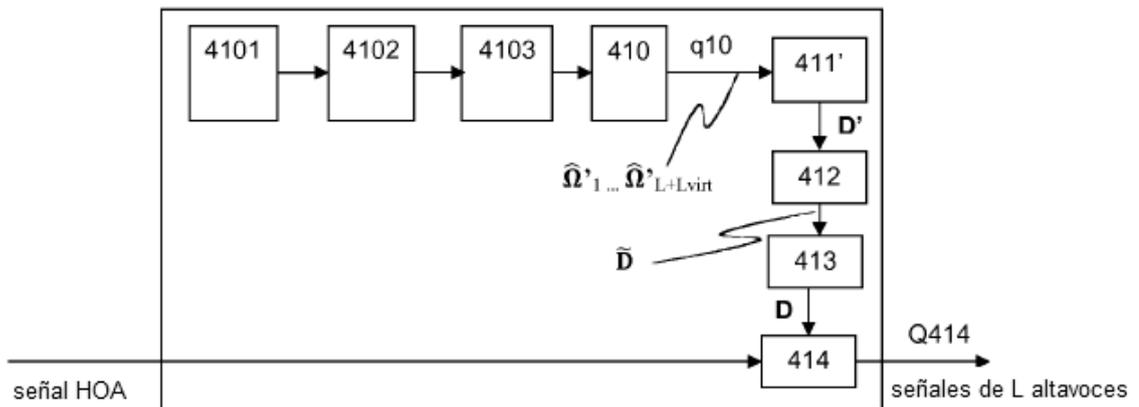


Fig.3



a)



b)

Fig.4

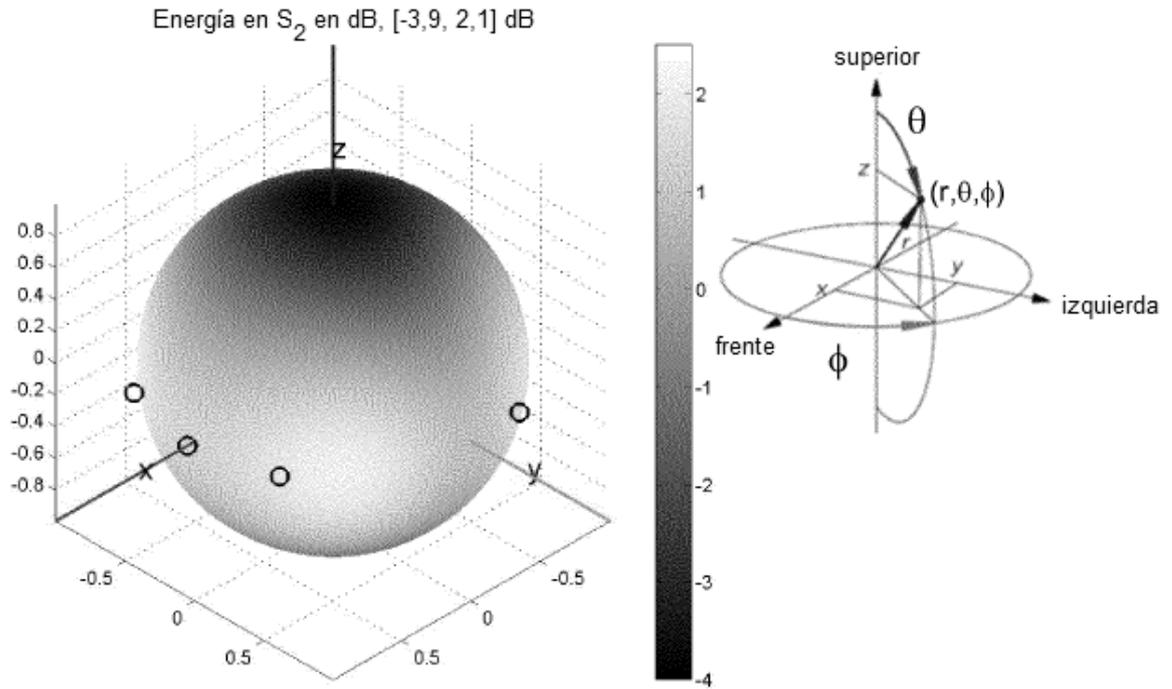


Fig.5

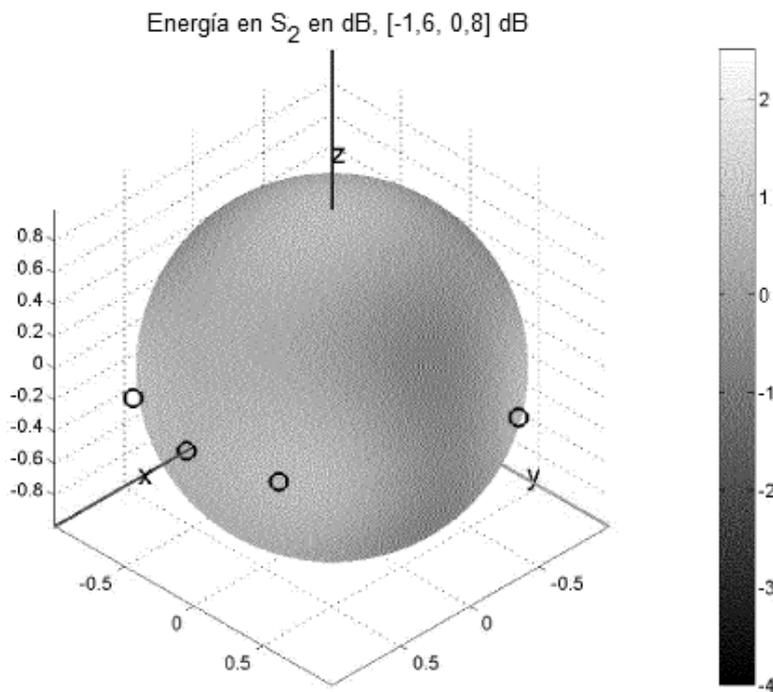


Fig.6

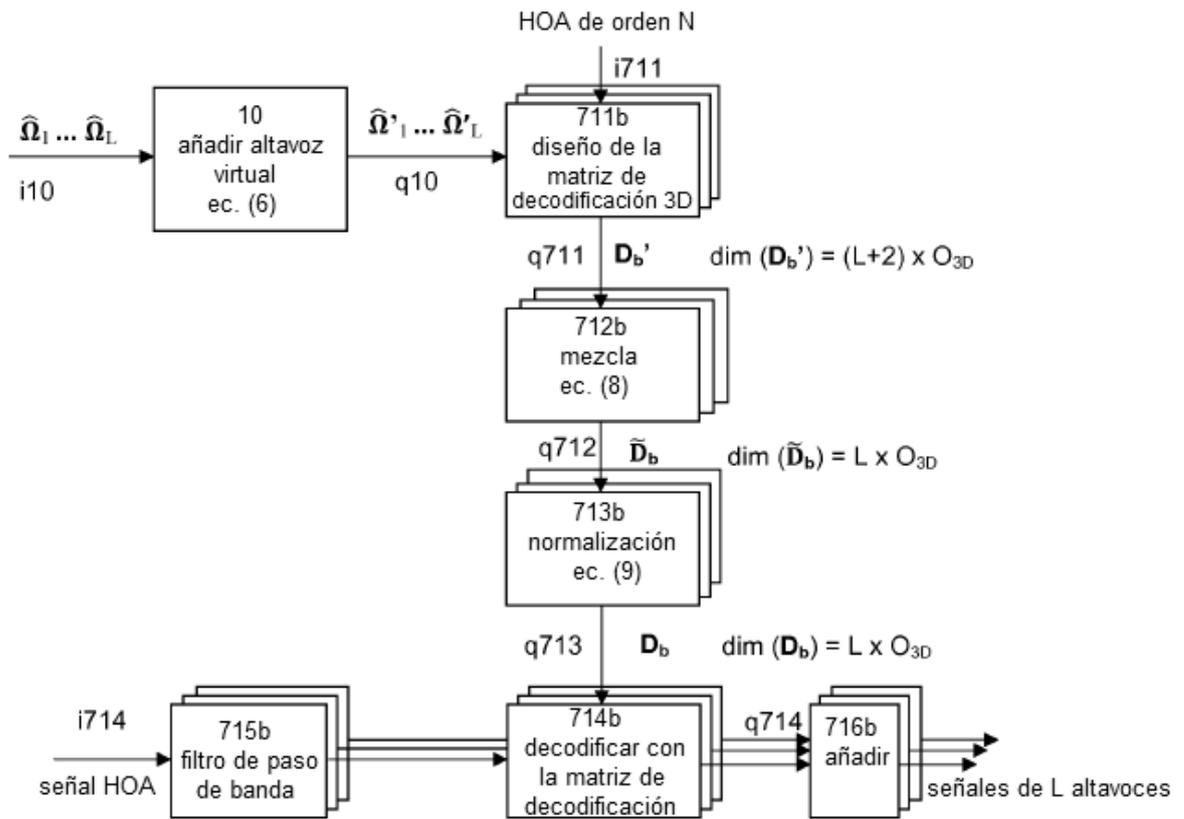


Fig.7