

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 987 561**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/02 (2013.01)

H04S 3/02 (2006.01)

H04S 3/00 (2006.01)

H04S 7/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.04.2019 PCT/FI2019/050264**

87 Fecha y número de publicación internacional: **17.10.2019 WO19197713**

96 Fecha de presentación y número de la solicitud europea: **02.04.2019 E 19785175 (1)**

97 Fecha y número de publicación de la concesión europea: **17.04.2024 EP 3776545**

54 Título: **Cuantificación de parámetros de audio espacial**

30 Prioridad:

09.04.2018 GB 201805833

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
15.11.2024

73 Titular/es:

NOKIA TECHNOLOGIES OY (100.0%)

Karakaari 7

02610 Espoo, FI

72 Inventor/es:

VASILACHE, ADRIANA

74 Agente/Representante:

DEL VALLE VALIENTE, Sonia

ES 2 987 561 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

Cuantificación de parámetros de audio espacial

5 **Campo**

La presente solicitud se refiere a aparatos y métodos para codificación de parámetros relacionados con el campo de sonido, pero no exclusivamente para parámetros relacionados con la dirección de dominio de tiempo-frecuencia que codifican un codificador y decodificador de audio.

10

Antecedentes

El procesamiento de audio espacial paramétrico es un campo de procesamiento de señales de audio donde el aspecto espacial del sonido se describe usando un conjunto de parámetros. Por ejemplo, en la captura de audio espacial paramétrica a partir de matrices de micrófonos, una elección típica y efectiva es estimar, a partir de las señales de matriz de micrófonos, un conjunto de parámetros tales como direcciones del sonido en bandas de frecuencia, y los coeficientes entre las partes direccionales y no direccionales del sonido capturado en las bandas de frecuencia. Se sabe que estos parámetros describen bien las propiedades espaciales perceptuales del sonido captado en la posición de la matriz de micrófonos. Estos parámetros pueden usarse en la síntesis del sonido espacial, por consiguiente, para auriculares de manera binaural, para altavoces o para otros formatos, tales como ambisónico.

20

Por tanto, las direcciones y las razones de energía directa con respecto a total en bandas de frecuencia son de este modo una parametrización que es particularmente eficaz para la captación de audio espacial.

25

También puede usarse un conjunto de parámetros que consiste en un parámetro de dirección en bandas de frecuencia y un parámetro de razón de energía en bandas de frecuencia (que indica la direccionalidad del sonido) como metadatos espaciales para un códec de audio. Por ejemplo, estos parámetros pueden estimarse a partir de señales de audio captadas en matriz de micrófonos y, por ejemplo, se puede generar una señal estéreo a partir de las señales de matriz de micrófonos que se transportan con los metadatos espaciales. La señal estéreo puede codificarse, por ejemplo, con un codificador AAC. Un decodificador puede decodificar las señales de audio para dar señales de PCM, y procesar el sonido en bandas de frecuencia (usando los metadatos espaciales) para obtener la salida espacial, por ejemplo, una salida binaural.

30

La solución mencionada anteriormente es particularmente adecuada para codificar sonido espacial capturado a partir de matrices de micrófonos (por ejemplo, en teléfonos móviles, cámaras de RV, matrices de micrófono independientes). Sin embargo, puede ser deseable que dicho codificador tenga también otros tipos de entrada que las señales capturadas de matriz de micrófonos, por ejemplo, señales de altavoces, señales de objeto de audio o señales ambisónicas.

35

El análisis de entradas ambisónicas (FOA) de primer orden para la extracción de metadatos espaciales se ha documentado completamente en la literatura científica relacionada con la codificación de audio direccional (DirAC) y la expansión de onda de plano armónico (Harpex). Esto es dado que existen matrices de micrófono que proporcionan directamente una señal FOA (de forma más precisa: su variante, la señal de formato B) y el análisis de dicha entrada ha sido por lo tanto un punto de estudio en el campo. Por ejemplo, la solicitud de patente internacional WO 2017/062.159, "Quantization of Spatial Vectors" de Kim Moo Young, 13 de abril de 2017, describe un dispositivo de codificación y decodificación de audio que comprende la cuantificación de vectores espaciales que representan señales de audio que se representan en un dominio ambisónico de orden superior (HOA), en donde cada par de valores de elevación y azimut se almacenan en forma de tabla y se referencian mediante un índice único, que se cuantifica en un solo paso.

45

50

Una entrada adicional para el codificador también es la entrada de altavoz multicanal, como las entradas envolventes del canal 5.1 o 7.1.

Sin embargo, con respecto a los componentes direccionales de los metadatos, que pueden comprender una elevación, un acimut (y una difusión) de una dirección resultante, para cada subbanda de tiempo/frecuencia considerada, una cuantificación y/o codificación que implemente una granularidad uniforme a lo largo de los componentes de acimut y elevación por separado (cuando estos dos parámetros se añadan por separado a los metadatos), puede resultar en una distribución desigual de los estados de cuantificación y codificación. Por ejemplo, un enfoque uniforme para ambos resultados por separado en un esquema de codificación con una mayor densidad más cercana a los "polos" de la esfera de dirección, en otras palabras directamente por encima o por debajo del lugar o ubicación de referencia.

55

60

Resumen

Según un primer aspecto, hay un aparato para la codificación de señal de audio espacial, que comprende: medios para determinar, para dos o más señales de audio, al menos un parámetro de audio espacial para proporcionar reproducción de audio espacial, comprendiendo el al menos un parámetro de audio espacial un parámetro de dirección

65

con un componente de elevación y un componente de acimut; medios para cuantificar escalarmente el componente de elevación para dar un primer valor de elevación indexado de un libro de códigos que comprende valores de elevación indexados y valores de acimut indexados, en donde el primer valor de elevación indexado está asociado a una pluralidad de valores de acimut indexados, y medios para cuantificar escalarmente el componente de acimut para dar un primer valor de acimut indexado, en donde el primer valor de acimut indexado se extrae de la pluralidad de valores de acimut indexados asociados al primer valor de elevación indexado; medios para cuantificar escalarmente el componente de elevación para dar un segundo valor de elevación indexado, en donde la cuantificación escalar del componente de elevación se limita a un valor de elevación indexado cuyo índice está próximo al índice del primer valor de elevación indexado del libro de códigos, en donde el segundo valor de elevación indexado está asociado a una pluralidad adicional de valores de acimut indexados, y medios para cuantificar escalarmente el componente de acimut para dar un segundo índice valor acimutal, en donde el segundo valor acimutal indexado se extrae de la pluralidad adicional de valores de acimut indexados; medios para determinar una primera distancia entre un punto en una esfera dado por el componente de elevación y el componente de acimut, y un punto en la esfera dado por el primer valor de elevación indexado y el primer valor de acimut indexado; medios para determinar una segunda distancia entre un punto en la esfera dado por la elevación y el componente de acimut, y un punto en la esfera dado por el segundo valor de elevación indexado y el segundo valor acimutal indexado; y medios para seleccionar uno del primer valor de elevación indexado y el primer valor de acimut indexado, o el segundo valor de elevación indexado y el segundo valor de acimut indexado, como un componente de elevación indexado y un componente de acimut indexado, dependiendo de un mínimo de la primera distancia y la segunda distancia.

La primera distancia y la segunda distancia pueden ser, cada una, una distancia estándar L2 medida entre dos puntos en la esfera.

El componente de elevación y el componente de acimut pueden indexarse hasta un punto en una retícula esférica, en donde la retícula esférica comprende una serie de puntos dispuestos en forma de esfera.

La retícula esférica puede formarse cubriendo la esfera con esferas más pequeñas, en donde las esferas más pequeñas definen los puntos de la retícula esférica.

El componente de elevación y el componente de acimut pueden ser componentes de un vector de acimut/elevación, en cuyo caso los medios para convertir el componente de elevación y el componente de acimut del parámetro de dirección a un valor de índice pueden comprender, entonces: medios para determinar una primera medida de distancia entre el vector de acimut/elevación y un primer vector de código de acimut/elevación de un libro de códigos vectoriales que comprende una pluralidad de vectores de códigos de acimut/elevación, y determinar al menos una segunda medida de distancia entre el vector de acimut/elevación y al menos un segundo vector de código de acimut/elevación del libro de códigos vectoriales; medios para seleccionar una medida de distancia mínima a partir de la primera medida de distancia y la al menos segunda medida de distancia; y medios para asignar el valor del índice como el índice del vector de código de acimut/elevación asociado a la medida de distancia mínima.

Como antes, la primera distancia y la segunda distancia pueden ser, cada una, una distancia estándar L2 medida entre dos puntos en una esfera.

La esfera puede ser una esfera unitaria, y la distancia estándar L2 entre los dos puntos de la esfera unitaria puede ser de la forma $-(\sin(\theta_1) \sin(\theta_2) + \cos(\theta_1) \cos(\theta_2) \cos(\varphi_1 - \varphi_2))$, en donde θ_1 y θ_2 son valores de elevación para un primer punto y un segundo punto de los dos puntos, en donde φ_1 y φ_2 son valores de acimut para el primer y el segundo punto de los dos puntos.

Se proporciona según un primer aspecto, un método para la codificación de señal de audio espacial, que comprende: determinar, para dos o más señales de audio, al menos un parámetro de audio espacial para proporcionar reproducción de audio espacial, comprendiendo el al menos un parámetro de audio espacial un parámetro de dirección con un componente de elevación y un componente de acimut; cuantificar escalarmente el componente de elevación para dar un primer valor de elevación indexado de un libro de códigos que comprende valores de elevación indexados y valores de acimut indexados, en donde el primer valor de elevación indexado está asociado a una pluralidad de valores de acimut indexados, y cuantificar escalarmente el componente de acimut para dar un primer valor de acimut indexado, en donde el primer valor de acimut indexado se extrae de la pluralidad de valores de acimut indexados asociados al primer valor de elevación indexado; cuantificar escalarmente el componente de elevación para dar un segundo valor de elevación indexado, en donde la cuantificación escalar del componente de elevación se limita a un valor de elevación indexado cuyo índice está cerca del índice del primer valor de elevación indexado del libro de códigos, en donde el segundo valor de elevación indexado está asociado a una pluralidad adicional de valores de acimut indexados, y cuantificar escalarmente el componente de acimut para dar un segundo valor de acimut indexado, en donde el segundo valor de acimut indexado se extrae de la pluralidad adicional de valores de acimut indexados; determinar una primera distancia entre un punto en una esfera dado por el componente de elevación y el componente de acimut, y un punto en la esfera dado por el primer valor de elevación indexado y el primer valor de acimut indexado; determinar una segunda distancia entre un punto en la esfera dado por la elevación y el componente de acimut, y un punto en la esfera dado por el segundo valor de elevación indexado y el segundo valor de acimut indexado; y seleccionar uno del primer valor de elevación indexado y el primer valor de acimut indexado, o el segundo valor de

elevación indexado y el segundo valor de acimut indexado, como un componente de elevación indexado y un componente de acimut indexado, dependiendo de un mínimo de la primera distancia y la segunda distancia.

5 Un programa informático que comprende instrucciones de programa para hacer que un ordenador realice el método descrito anteriormente.

Un producto de programa informático almacenado en un medio puede hacer que un aparato realice el método tal como se describe en la presente memoria.

10 Un dispositivo electrónico puede comprender un aparato tal como se describe en la presente memoria.

Un conjunto de chips puede comprender un aparato tal como se describe en la presente memoria.

15 Las realizaciones de la presente solicitud tienen como objetivo abordar problemas asociados con el estado de la técnica.

Resumen de las figuras

20 Para una mejor comprensión de la presente solicitud, ahora se hará referencia a modo de ejemplo a los dibujos adjuntos, en los que:

la Figura 1 muestra esquemáticamente un sistema de aparato adecuado para implementar algunas realizaciones;

25 la Figura 2 muestra esquemáticamente el procesador de análisis como se muestra en la Figura 1 según algunas realizaciones;

la Figura 3a muestra esquemáticamente un cuantificador/codificador de metadatos como se muestra en la Figura 1 según algunas realizaciones;

30 la Figura 3b muestra esquemáticamente el extractor de metadatos como se muestra en la Figura 1 según algunas realizaciones;

35 las Figuras 3c a 3e muestran esquemáticamente ejemplos de configuraciones de ubicaciones de esferas como se usa en el codificador/cuantificador de metadatos y el extractor de metadatos, como se muestra en las Figuras 3a y 3b según algunas realizaciones;

la Figura 4 muestra un diagrama de la operación del sistema como se muestra en la Figura 1 según algunas realizaciones;

40 la Figura 5 muestra un diagrama de flujo de la operación del procesador de análisis como se muestra en la Figura 2 según algunas realizaciones;

45 la Figura 6 muestra un diagrama de flujo de generación de un índice de dirección basado en un parámetro de dirección de entrada con detalles adicionales;

la Figura 7 muestra un diagrama de flujo de un ejemplo de una operación de cuantificación del parámetro de dirección para obtener un índice de dirección;

50 la Figura 8 muestra un diagrama de flujo de generación de un parámetro de dirección cuantificado basado en un índice de dirección de entrada con más detalle;

la Figura 9 muestra un diagrama de flujo de una operación de ejemplo de conversión de un parámetro de dirección cuantificado a partir de un índice de dirección con más detalle; y

55 la Figura 10 muestra esquemáticamente un dispositivo ilustrativo adecuado para implementar el aparato mostrado.

Realizaciones de la solicitud

60 A continuación se describe con más detalle un aparato adecuado y mecanismos posibles para la provisión de parámetros de metadatos derivados de análisis espacial efectivos para señales de audio de formato de entrada multicanal. En la siguiente discusión, el sistema multicanal se analiza con respecto a una implementación de micrófono multicanal. Sin embargo, como se discutió anteriormente, el formato de entrada puede ser cualquier formato de entrada adecuado, tal como un altavoz multicanal, ambisónica (FOA/HOA), etc. Se entiende que en algunas realizaciones la ubicación del canal se basa en una ubicación del micrófono o es una ubicación o dirección virtual. Además, la salida del sistema de ejemplo es una disposición de altavoz multicanal. Sin embargo, se entiende que la salida puede hacerse

al usuario mediante medios distintos a los altavoces. Además, las señales del altavoz multicanal pueden generalizarse para ser dos o más señales de audio de reproducción.

Como se discutió anteriormente, los parámetros de metadatos espaciales tales como la dirección y la relación de energía directa a total (o la relación de difusión, energías absolutas, o cualquier expresión adecuada que indique la direccionalidad/no direccionalidad del sonido en el intervalo de tiempo-frecuencia dado) en las bandas de frecuencia son particularmente adecuadas para expresar las propiedades perceptuales de campos de sonido natural. Las escenas de sonido sintéticas, tales como las mezclas de altavoces 5.1 comúnmente utilizan efectos de audio y métodos de panoramización de amplitud que proporcionan sonido espacial que difiere de los sonidos que ocurren en campos de sonido natural. En particular, una mezcla 5.1 o 7.1 puede configurarse de modo que contenga sonidos coherentes reproducidos desde múltiples direcciones. Por ejemplo, es común que algunos sonidos de una mezcla 5.1 percibida directamente en la parte delantera no sean producidos por un altavoz central (canal), pero, por ejemplo, coherentemente desde los altavoces frontal (canales) izquierdo y derecho, y potencialmente también del altavoz central (canal). Los parámetros de metadatos espaciales tales como la(s) dirección(s) y la (s) relación (s) de energía no expresan (n) tales características espacialmente coherentes con precisión. Como tales otros parámetros de metadatos, tales como los parámetros de coherencia, pueden determinarse a partir del análisis de las señales de audio para expresar las relaciones de señal de audio entre los canales.

Como se expresó anteriormente, un ejemplo de la incorporación de la información de dirección en los metadatos, es usar valores de acimut y elevación determinados. Sin embargo, el muestreo uniforme convencional de acimut y elevación produce una distribución de dirección no uniforme.

Por lo tanto, el concepto es un intento de determinar un parámetro de dirección para metadatos espaciales y para indexar el parámetro basándose en una distribución de cobertura de esferas práctica de las direcciones, con el fin de definir una distribución de direcciones más uniforme.

El índice de metadatos propuesto puede entonces utilizarse junto a una señal de mezcla descendente ("canales"), para definir un formato inmersivo paramétrico que pueda utilizarse, p. ej., para el códec de Servicio de audio y de voz inmersivos (IVAS). Alternativamente y además, el formato de retícula esférica se puede usar en el códec para cuantificar las direcciones.

El concepto describe además la decodificación de dichos parámetros de dirección indexados para producir parámetros direccionales cuantificados que pueden usarse en la síntesis de audio espacial basado en parametrización relacionada con campo de sonido (dirección o direcciones y relación o relaciones en bandas de frecuencia).

Con respecto a la Figura 1, se muestran un aparato y un sistema de ejemplo para implementar realizaciones de la solicitud. El sistema 100 se muestra con una parte de "análisis" 121 y una parte de "síntesis" 131. La parte de "análisis" 121 es la parte desde la recepción de las señales de altavoces de múltiples canales hasta la codificación de los metadatos y la señal de mezcla descendente; y la parte de "síntesis" 131 es la parte de una decodificación de los metadatos codificados y la señal de mezcla descendente a la presentación de la señal regenerada (por ejemplo, en forma de altavoces de múltiples canales).

La entrada al sistema 100 y la parte de análisis 121 son las señales multicanal 102. En los siguientes ejemplos, se describe una entrada de señal de canal de micrófono, aunque puede implementarse cualquier formato de entrada (o multicanal sintético) adecuado en otras realizaciones.

Las señales multicanal se pasan a un mezclador descendente 103 y a un procesador de análisis 105.

En algunas realizaciones, el mezclador descendente 103 está configurado para recibir las señales multicanal y la mezcla descendente de las señales a un número determinado de canales y emitir las señales de mezcla descendente 104. Por ejemplo, el mezclador descendente 103 puede configurarse para generar una mezcla de canales de audio 2 de las señales multicanal. El número determinado de canales puede ser cualquier número adecuado de canales. En algunas realizaciones, el mezclador descendente 103 es opcional y las señales multicanal se pasan sin procesar a un codificador 107 de la misma manera que la señal de mezcla descendente en este ejemplo.

En algunas realizaciones, el procesador de análisis 105 también está configurado para recibir las señales multicanal y analizar las señales para producir metadatos 106 asociados con las señales multicanal y, por lo tanto, asociarse con las señales de mezcla descendente 104. El procesador de análisis 105 puede configurarse para generar los metadatos que pueden comprender, para cada intervalo de análisis de tiempo-frecuencia, un parámetro de dirección 108, un parámetro de relación de energía 110, un parámetro de coherencia 112 y un parámetro de difusión 114. La dirección, los parámetros de relación de energía y de difusión pueden considerarse, en algunas realizaciones, como parámetros de audio espacial. En otras palabras, los parámetros de audio espacial comprenden parámetros que tienen como objetivo caracterizar el campo sonoro creado por las señales multicanal (o dos o más señales de audio de reproducción en general). Los parámetros de coherencia pueden considerarse parámetros de audio de relación de señal que tienen como objetivo caracterizar la relación entre las señales multicanal.

En algunas realizaciones, los parámetros generados pueden diferir entre las bandas de frecuencia. Por lo tanto, por ejemplo, en la banda X todos los parámetros se generan y transmiten, mientras que en la banda Y solo se genera y transmite uno de los parámetros, y además en la banda Z no se generan o transmiten parámetros. Un ejemplo práctico de esto es que para algunas bandas de frecuencia, tales como la banda más alta, no se requieren algunos de los parámetros por motivos perceptuales. Las señales de mezcla descendente 104 y los metadatos 106 pueden pasarse a un codificador 107.

El codificador 107 puede comprender un núcleo 109 estéreo IVAS que se configure para recibir las señales 104 de mezcla descendente (o de otro tipo), y generar una codificación adecuada de estas señales de audio. En algunas realizaciones, el codificador 107 puede ser un ordenador (que ejecuta un software adecuado almacenado en la memoria y en al menos un procesador), o alternatively un dispositivo específico que utiliza, por ejemplo, FPGAs o ASIC. La codificación puede implementarse usando cualquier esquema adecuado. El codificador 107 puede comprender además un codificador de metadatos o cuantificador 109 que está configurado para recibir los metadatos y emitir una forma codificada o comprimida de la información. En algunas realizaciones, el codificador 107 puede además intercalar, multiplexar a un flujo de datos individual o incorporar los metadatos dentro de las señales de mezcla descendente codificadas antes de la transmisión o el almacenamiento mostrado en la Figura 1 por la línea discontinua. La multiplexación puede implementarse usando cualquier esquema adecuado.

En el lado del decodificador, los datos recibidos o recuperados (flujo) pueden ser recibidos por un decodificador/demultiplexor 133. El decodificador/demultiplexor 133 puede demultiplexar los flujos codificados y pasar el flujo codificado de audio a un extractor 135 de mezcla descendente que está configurado para decodificar las señales de audio para obtener las señales de mezcla descendente. De manera similar, el decodificador/demultiplexor 133 puede comprender un extractor de metadatos 137 que está configurado para recibir los metadatos codificados y generar metadatos. En algunas realizaciones, el decodificador/demultiplexor 133 puede ser un ordenador (que ejecuta un software adecuado almacenado en la memoria y en al menos un procesador), o alternatively un dispositivo específico que utiliza, por ejemplo, FPGA o ASIC.

Los metadatos decodificados y las señales de audio de mezcla descendente pueden pasarse a un procesador 139 de síntesis.

La parte 131 de "síntesis" del sistema 100 muestra, además, un procesador 139 de síntesis configurado para recibir la mezcla descendente y los metadatos, y vuelve a crear, en cualquier formato adecuado, un audio espacial sintetizado en forma de señales multicanal 110 (pueden ser un formato de altavoces multicanal o, en algunas realizaciones, cualquier formato de salida adecuado, tales como las señales binaurales o Ambisonics, dependiendo del caso de uso) en base a las señales de mezcla descendente y los metadatos.

Con respecto a la Figura 4, se muestra un diagrama de flujo ilustrativo de la descripción general mostrada en la Figura 1.

En primer lugar, el sistema (parte de análisis) está configurado para recibir señales de audio multicanal como se muestra en la Figura 4 mediante el paso 401.

Entonces el sistema (parte de análisis) está configurado para generar una mezcla descendente de las señales multicanal como se muestra en la Figura 4 mediante el paso 403.

También el sistema (parte de análisis) está configurado para analizar señales para generar metadatos tales como parámetros de direccionamiento; parámetros de relación de energía; parámetros de difusión y parámetros de coherencia como se muestra en la Figura 4 mediante el paso 405.

El sistema se configura entonces para codificar para su almacenamiento/transmisión la señal de mezcla descendente y metadatos como se muestra en la Figura 4 mediante el paso 407.

Tras esto, el sistema puede almacenar/transmitir la mezcla descendente y metadatos codificados, como se muestra en la Figura 4 mediante el paso 409.

El sistema puede recuperar/recibir la mezcla descendente y metadatos codificados, como se muestra en la Figura 4 mediante el paso 411.

Posteriormente, el sistema se configura para extraer la mezcla descendente y los metadatos de parámetros codificados de mezcla descendente y metadatos, por ejemplo demultiplexar y decodificar los parámetros codificados de mezcla descendente y metadatos, como se muestra en la Figura 4 mediante el paso 413.

El sistema (parte de síntesis) está configurado para sintetizar una señal de audio multicanal de salida basándose en la mezcla descendente extraída de señales de audio multicanal y los metadatos con parámetros de coherencia como se muestra en la Figura 4 mediante el paso 415.

Con respecto a la Figura 2, se describe con más detalle un procesador de análisis de ejemplo 105 (como puede verse en la Figura 1) según algunas realizaciones con más detalle. El procesador de análisis 105 en algunas realizaciones comprende un transformador de dominio de tiempo-frecuencia 201.

En algunas realizaciones, el transformador de dominio de tiempo-frecuencia 201 está configurado para recibir las señales multicanal 102 y aplicar una transformación de dominio de tiempo a frecuencia adecuada tal como una Transformación de Fourier de Tiempo Corto (STFT) para convertir las señales de dominio del tiempo de entrada en una señal de tiempo-frecuencia adecuada. Estas señales de tiempo-frecuencia pueden pasar a un analizador de dirección 203 y a un analizador de señal 205.

Por lo tanto, por ejemplo, las señales de tiempo-frecuencia 202 pueden representarse en la representación del dominio de tiempo-frecuencia en

$$s_i(b, n),$$

donde b es el índice del intervalo de frecuencia y n es el índice de trama, e i es el índice de canal. En otra expresión, n puede considerarse un índice de tiempo con una tasa de muestreo más baja que la de las señales originales del dominio de tiempo. Estos intervalos de frecuencia pueden agruparse en subbandas que agrupen uno o más de los intervalos en un índice de banda $k = 0, \dots, K-1$. Cada subbanda k tiene un intervalo más bajo $b_{k,low}$ y un intervalo más alto $b_{k,high}$, y la subbanda contiene todos los intervalos de $b_{k,low}$ a $b_{k,high}$. Las anchuras de las subbandas pueden aproximarse a cualquier distribución adecuada. Por ejemplo, la escala de ancho de banda rectangular equivalente (ERB) o la escala Bark.

En algunas realizaciones, el procesador de análisis 105 comprende un analizador de dirección 203. El analizador 203 de dirección puede configurarse para recibir las señales 202 de frecuencia de tiempo, y basándose en estas señales, estimar los parámetros 108 de dirección. Los parámetros de dirección se pueden determinar basándose en cualquier determinación de "dirección" basada en audio.

Por ejemplo, en algunas realizaciones, el analizador de dirección 203 está configurado para estimar la dirección con dos o más entradas de señal. Esto representa la configuración más simple para estimar una "dirección" y un procesamiento más complejo puede realizarse con incluso más señales.

Por lo tanto, el analizador 203 de dirección puede configurarse para proporcionar un acimut para cada banda de frecuencia y trama temporal, denotados como acimut $\phi(k, n)$ y elevación $\theta(k, n)$. El parámetro 108 de dirección también se puede pasar a un analizador 205 de señal.

En algunas realizaciones, además del parámetro de dirección, el analizador 203 de dirección se configura para determinar un parámetro 110 de relación de energía. Se puede considerar que la relación de energía es una determinación de la energía de la señal de audio que puede considerarse que llega desde una dirección. La relación de energía directa a total $r(k, n)$ puede estimarse, p. ej., utilizando una medida de estabilidad de la estimación direccional, o utilizando cualquier medida de correlación, o cualquier otro método adecuado para obtener un parámetro de relación.

Los parámetros 108 de dirección estimados pueden emitirse (y pasarse a un codificador). Los parámetros 110 de relación de energía estimados pueden pasarse a un analizador 205 de señal.

En algunas realizaciones, el procesador de análisis 105 comprende un analizador de señal 205. El analizador 205 de señal está configurado para recibir parámetros (tales como el de acimut $\phi(k, n)$ y el de elevación $\theta(k, n)$ 108, y los de las relaciones 110 de energía directa a total $r(k, n)$) desde el analizador 203 de dirección. El analizador 205 de señal puede configurarse, además, para recibir las señales 202 de tiempo-frecuencia ($s_i(b, n)$) desde el transformador 201 de dominio de tiempo-frecuencia. Todos estos están en el dominio de tiempo-frecuencia; b es el índice de intervalo de frecuencia, k es el índice de banda de frecuencia (cada banda potencialmente consiste en varios intervalos b), n es el índice de tiempo, e i es el canal.

Aunque aquí se expresan direcciones y relaciones para cada índice de tiempo n , en algunas realizaciones los parámetros pueden combinarse a lo largo de varios índices de tiempo. Lo mismo se aplica para el eje de frecuencia, como se ha expresado, la dirección de varios intervalos de frecuencia b , podría expresarse por un parámetro de dirección en la banda k , que consista en varios intervalos de frecuencia b . Lo mismo se aplica para todos los parámetros espaciales analizados en la presente memoria.

El analizador 205 de señal está configurado para producir una serie de parámetros de señal. En la siguiente descripción se encuentran los dos parámetros: coherencia y difusión, ambos analizados en el dominio de tiempo-frecuencia. Además, en algunas realizaciones, el analizador 205 de señal se configura para modificar las relaciones de energía ($r(k, n)$) estimadas. El analizador de señal 205 se configura para generar los parámetros de coherencia y difusión en base a cualquier método conocido adecuado.

Con respecto a la Figura 5, se muestra un diagrama de flujo que resume las operaciones del procesador de análisis 105.

La primera operación es una recepción de señales de audio de múltiples canales de dominio de tiempo (altavoz) como se muestra en la Figura 5 mediante el paso 501.

Después, se aplica una transformación de dominio de tiempo a dominio de frecuencia (p.ej. STFT) para generar señales de dominio de tiempo-frecuencia adecuadas para su análisis como se muestra en la Figura 5 mediante el paso 503.

Después, la aplicación de los parámetros de dirección y relación de energía se muestran en la Figura 5 mediante el paso 505.

Posteriormente, se aplica un análisis para determinar los parámetros de coherencia (tales como parámetros de coherencia circundante y/o de propagación) y los parámetros de difusión se muestra en la Figura 5 mediante el paso 507. En algunas realizaciones, la relación de energía también puede modificarse basándose en los parámetros de coherencia determinados en este paso.

La operación de determinación de los parámetros estéreo se muestra en la Figura 5 mediante el paso 509.

Con respecto a la Figura 3a, se muestra un codificador de metadatos de ejemplo y específicamente el codificador de metadatos de dirección 300 según algunas realizaciones.

El codificador 300 de metadatos de dirección comprende una entrada 302 de cuantificación. La entrada de cuantificación, que también se conoce como una entrada de codificación, está configurada para definir la granularidad de esferas dispuestas alrededor de una ubicación o posición de referencia a partir de la cual se determine el parámetro de dirección. En algunas realizaciones, la entrada de cuantificación es un valor predefinido o fijo.

El codificador 300 de metadatos de dirección comprende un posicionador 303 de esfera. El posicionador de esfera está configurado para configurar la disposición de esferas basándose en el valor de entrada de cuantificación. La retícula esférica propuesta utiliza la idea de cubrir una esfera con esferas más pequeñas, y considerar los centros de las esferas más pequeñas como puntos que definan una retícula de direcciones casi equidistantes, de manera que la retícula esférica comprenda un número de puntos dispuestos en una forma de esfera.

El concepto, como se muestra en la presente memoria, es uno en el que una esfera se defina en relación con la ubicación de referencia. La esfera se puede visualizar como una serie de círculos (o intersecciones) y para cada intersección de círculo se encuentran en la circunferencia del círculo un número definido de esferas (más pequeñas). Esto se muestra, por ejemplo, con respecto a las Figuras 3c a 3e. Por ejemplo, la Figura 3c muestra un ejemplo de una “sección transversal ecuatorial” o un primer círculo principal 370 que tiene un radio definido como el “radio de esfera principal”. También se muestran en la Figura 3c las esferas más pequeñas (mostradas como secciones transversales circulares de círculos) 371, 373, 375, 377 y 379, ubicadas de modo que cada esfera más pequeña tenga una circunferencia que en un punto toque la circunferencia de esfera principal, y al menos un punto adicional que toque al menos una circunferencia de esfera más pequeña adicional. Por lo tanto, como se muestra en la Figura 3c, la esfera más pequeña 371 toca la esfera principal 370 y la esfera más pequeña 373, la esfera más pequeña 373 toca la esfera principal 370 y las esferas más pequeñas 371 y 375, la esfera más pequeña 375 toca la esfera principal 370 y las esferas más pequeñas 373 y 377, la esfera más pequeña 377 toca la esfera principal 370 y las esferas más pequeñas 375 y 379, y la esfera más pequeña 379 toca la esfera principal 370 y la esfera más pequeña 377.

La Figura 3d muestra un ejemplo de “sección transversal tropical” o círculo principal 380 adicional y las esferas más pequeñas (mostradas como secciones transversales de círculos) 381, 383, 385, ubicadas de manera que cada esfera más pequeña tenga una circunferencia que en un punto toque la circunferencia de la esfera (círculo) principal y al menos un punto adicional que toque al menos una circunferencia de esfera más pequeña adicional. Por lo tanto, como se muestra en la Figura 3d, la esfera más pequeña 381 toca la esfera principal 380 y la esfera más pequeña 383, la esfera más pequeña 383 toca la esfera principal 380 y las esferas más pequeñas 381 y 385, la esfera más pequeña 385 toca la esfera principal 380 y la esfera más pequeña 383.

La Figura 3e muestra un ejemplo de esfera y las secciones transversales 370, 380 y las esferas más pequeñas (secciones transversales) 371 asociadas a la sección transversal 370, la esfera más pequeña 381 asociada a la sección transversal 380 y otras esferas más pequeñas 392, 393, 394, 395, 397, 398. En este ejemplo, solo se dibujan los círculos con un valor de acimut inicial en 0.

Por lo tanto, el posicionador 303 de esfera está configurado para realizar las siguientes operaciones para definir las direcciones correspondientes a las esferas de cobertura:

Entrada: Entrada de cuantificación (número de puntos en el «ecuador», $n(0) = M$)

Salida: número de círculos, N_c , y número de puntos en cada círculo, $n(i)$, i

$= 0, N_c - 1$

1. $n(0) = M$

2. $\alpha = \frac{2\pi}{n(0)}$

3. $R(0) = 1$ (radio del círculo en el ecuador)

4. $\theta(0) = 0$ (elevación)

5. $r = 2 R(0) \sin\left(\frac{\alpha}{4}\right)$ (radio de las esferas más pequeñas)

6. $\phi(0) = 0$

7. $p = \arcsin\left(r \frac{\sqrt{3}}{R(0)}\right)$

8. $R(1) = R(0) \cos p$

9. $i = 1$

10. Mientras $n(i - 1) > 1$

a. $n(i) = \left\lfloor \frac{\pi R(i)}{r} \right\rfloor$ (esto es válido cuando $r \ll R(0)$)

b. $\theta(i) = p \cdot i$

c. $\Delta\phi(i) = \frac{\pi}{n(i-1)}$ (granularidad del acimut en el círculo i)

d. $R(i + 1) = R(i) \cos((i + 1) \cdot p)$

e. Si i es impar

i. $\phi_0(i) = \frac{\Delta\phi(i)}{2}$ (primer valor de acimut en el círculo i)

f. De lo contrario

i. $\phi_0(i) = 0$

g. Finalizar si

h. $i = i + 1$

11. Finalizar mientras

12. $N_c = i + 1$

$$r = 2 R(0) \sin\left(\frac{\alpha}{k}\right)$$

El paso 5 también se puede reemplazar por $r = 2 R(0) \sin\left(\frac{\alpha}{k}\right)$, donde el factor k controla la distribución de puntos a lo largo de la elevación. Para $k = 4$, la resolución de elevación es de aproximadamente 1 grado. Para k más pequeños, la resolución es correspondientemente menor. La elevación para cada punto del círculo i viene dada por los valores en $\theta(i)$. Para cada círculo por encima del ecuador hay un círculo correspondiente por debajo del ecuador.

Cada punto de dirección en un círculo puede indexarse en orden creciente con respecto al valor de acimut. El índice del primer punto en cada círculo viene dado por un desplazamiento que puede deducirse del número de puntos en cada círculo, $n(i)$. Para obtener los desplazamientos, para un orden considerado de los círculos, los desplazamientos se calculan como el número acumulado de puntos en los círculos para el orden dado, comenzando con el valor 0 como primer desplazamiento.

Un posible orden de los círculos podría ser comenzar con el ecuador seguido del primer círculo por encima del ecuador, luego el primero por debajo del ecuador, el segundo por encima del ecuador, y así sucesivamente.

Otra opción es comenzar con el ecuador, luego el círculo por encima del ecuador que está a una elevación aproximada de 45 grados, y luego el círculo correspondiente por debajo del ecuador, y luego los círculos restantes en orden alternativo. De esta forma, para un posicionamiento más sencillo de los altavoces, solo se utilizan los primeros círculos, lo que reduce la cantidad de bits para enviar la información.

También es posible otro orden de los círculos en otras realizaciones.

En algunas realizaciones, la retícula esférica también se puede generar considerando el meridiano 0 en lugar del ecuador, o cualquier otro meridiano.

El posicionador de esfera ha determinado el número de círculos y el número de puntos en cada círculo, $n(i)$, $i = 0, N_c - 1$ y el orden de indexación pueden configurarse para pasar esta información a un convertidor 305 de EA a DI.

El codificador de metadatos de dirección 300 comprende una entrada de parámetro de dirección 108. La entrada del parámetro de dirección define una elevación y un valor de acimut $D = (\theta, \phi)$.

Los procedimientos de transformación de (elevación/acimut) (EA) al índice de dirección (DI) y viceversa se presentan en los siguientes párrafos. El orden alternativo de los círculos se considera aquí.

El codificador de metadatos de dirección 300 comprende un convertidor de índice de dirección 305 de acimut a dirección (EA-DI). El convertidor 305 de elevación-acimut a índice de dirección está configurado para recibir la entrada 108 de parámetro de dirección y la información de posicionador de esfera, y convertir el valor de elevación-acimut de la entrada 108, de parámetro de dirección a un índice de dirección, mediante la cuantificación del valor de elevación-acimut.

Con respecto a la Figura 6, se muestra un método para generar el índice de dirección según algunas realizaciones.

La recepción de la entrada de cuantificación se muestra en la Figura 6 mediante el paso 601.

Entonces, el método puede determinar el posicionamiento de esfera basándose en la entrada de cuantificación, como se muestra en la Figura 6 mediante el paso 603.

Además, el método puede comprender la recepción del parámetro de dirección, como se muestra en la Figura 6 mediante el paso 602.

Habiendo recibido el parámetro de dirección y la información de posicionamiento de esfera, el método comprende convertir el parámetro de dirección a un índice de dirección basándose en la información de posicionamiento de esfera, como se muestra en la Figura 6 mediante el paso 605.

El método luego emite el índice de dirección como se muestra en la Figura 6 mediante el paso 607. El convertidor de índice de dirección 305 de acimut a dirección (EA-DI) se configura para realizar esta conversión según el siguiente algoritmo:

$$\text{Entrada: } (\theta, \phi), \theta \in S_\theta \subset \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \phi \in S_\phi \subset [0, 2\pi]$$

$$\text{Salida: } I_d$$

S_θ puede adoptar la forma de un libro de códigos indexado con N entradas discretas, correspondiendo cada entrada θ_i a un valor de elevación para $i = 0: N - 1$. Además, el libro de códigos también comprende para cada valor discreto de elevación θ_i un conjunto de valores discretos de acimut ϕ_j , donde el número de valores de acimut en el conjunto depende de la elevación θ_i . En otras palabras, para cada entrada de elevación θ_i puede haber diferentes números de valores discretos de acimut ϕ_j , para $j = 0: f(\theta_i)$, donde $f(\theta_i)$ indica que el número de valores de acimut del conjunto de valores de acimut asociados al valor de elevación θ_i , es una función del valor de elevación θ_i .

En otras realizaciones, la indexación del valor de elevación-acimut puede no usar una estructura de libro de códigos como lo anterior, para almacenar valores discretos de elevación y acimut. En su lugar, se puede usar la cuantificación lineal donde el número de círculos N_c y la granularidad p proporcionados por el posicionador de esferas, se puedan usar para dividir uniformemente el rango de elevación desde

$-\frac{\pi}{2}$ a $\frac{\pi}{2}$. La posición en el rango de elevación proporciona un índice de círculo, y se conocen el número de puntos discretos de acimut y el correspondiente desplazamiento, $off(i)$.

$$i = \begin{cases} 2\frac{\theta}{p} - 1, & \text{if } \theta > 0 \\ 0, & \text{if } \theta = 0 \\ -2\frac{\theta}{p}, & \text{if } \theta < 0 \end{cases}$$

1. Encontrar el índice de círculo

2. Encontrar el índice del acimut dentro del círculo i: $j = \left\lfloor \frac{\phi}{\Delta\phi(i)} \right\rfloor$ donde $i' = \left\lfloor \frac{\theta}{p} \right\rfloor$

3. El índice de dirección es $l_d = \text{off}(i) + j$

La explicación del método de indexación EA-DI se deberá realizar sobre la base de la estructura de libro de códigos anterior, tal como se explicó anteriormente. Sin embargo, debe entenderse que el siguiente método se puede aplicar igualmente a una estructura cuantificadora uniforme, en la que el rango de elevación se divida a lo largo de una escala $-\frac{\pi}{2}$ a $\frac{\pi}{2}$.

Con respecto a la Figura 7, se muestra un método ilustrativo del paso 605 en la Figura 6, para convertir elevación-acimut a índice de dirección (EA-DI).

El primer paso de cuantificación de un valor de elevación-acimut comprende cuantificar escalarmente el valor de elevación θ encontrando la entrada θ_i del libro de códigos más cercana, para dar un primer valor de elevación cuantificado $\hat{\theta} = \theta_i$. El valor de elevación θ puede nuevamente cuantificarse escalarmente encontrando la siguiente entrada del libro de códigos más próxima. Esto se da como una de las entradas θ_{i+1} o θ_{i-1} del libro de códigos, dependiendo de cuál esté más cerca de θ , lo que produce un segundo valor de elevación cuantificado $\hat{\theta}'$.

Los pasos de procesamiento de la cuantificación escalar del valor de elevación θ al valor de elevación indexado θ_i más cercano y adicionalmente al siguiente valor de elevación indexado θ_{i+1} o θ_{i-1} más cercano, se muestran como los pasos 701 y 703 de procesamiento, respectivamente.

Para cada valor de elevación cuantificado $\hat{\theta}$ y $\hat{\theta}'$, se puede encontrar el valor de acimut cuantificado escalarmente correspondiente. En otras palabras, un primer valor de acimut cuantificado escalarmente correspondiente a $\hat{\theta}$, se determina encontrando el valor de acimut más cercano del conjunto de valores de acimut asociados al valor de elevación indexado θ_i para el primer valor de elevación cuantificado $\hat{\theta}$. El primer valor de acimut cuantificado escalarmente correspondiente al primer valor de elevación cuantificado $\hat{\theta}$, se expresa como $\hat{\phi}$. De manera similar, un segundo valor de acimut cuantificado escalarmente correspondiente a $\hat{\theta}'$, también se determina y se expresa como $\hat{\phi}'$. Sin embargo, esto se puede realizar volviendo a cuantificar el valor de acimut ϕ , pero esta vez utilizando el conjunto de valores de acimut asociados al índice del segundo valor de elevación cuantificado escalarmente $\hat{\theta}'$.

Los pasos de procesamiento de la cuantificación escalar del valor de acimut ϕ correspondiente al valor de elevación indexado θ_i más cercano, y adicionalmente la cuantificación escalar del valor de acimut correspondiente al siguiente valor de elevación indexado θ_{i+1} o θ_{i-1} más cercano, se muestran como los pasos 705 y 707 de procesamiento, respectivamente.

Una vez se determinen el primer par de valores cuantificados escalarmente de elevación-acimut y el segundo par de valores cuantificados escalarmente de elevación-acimut, se calcula una medida de distancia en una esfera unitaria para cada par. La medida de la distancia se puede considerar tomando la distancia estándar L2 entre dos puntos en la esfera unitaria, de modo que para el primer par de elevación-acimut cuantificado escalarmente $(\hat{\theta}, \hat{\phi})$, la distancia d se calcula como la distancia entre el primer par de elevación-acimut cuantificado escalarmente $(\hat{\theta}, \hat{\phi})$ y el par de elevación-acimut no cuantificado (θ, ϕ) en la esfera unitaria.

De manera similar, para el segundo par de elevación-acimut cuantificado escalarmente $(\hat{\theta}', \hat{\phi}')$, la distancia d' se calcula como la distancia entre el segundo par de elevación-acimut cuantificado escalarmente $(\hat{\theta}', \hat{\phi}')$ y el par de elevación-acimut no cuantificado (θ, ϕ) en la esfera unitaria.

Se debe apreciar en las realizaciones, que la distancia estándar L2 entre dos puntos x e y en una esfera unitaria, se considera a partir de $\|x - y\|^2$, donde x e y son coordenadas esféricas en el espacio tridimensional. En términos del par de elevación-acimut (θ, ϕ) , las coordenadas esféricas se expresan como $x = (r\cos(\theta)\cos(\phi), r\cos(\theta)\sin(\phi), r\sin(\theta))$,

y para el par de elevación-acimut $(\hat{\theta}, \hat{\phi})$ las coordenadas esféricas corresponden a $y = (r \cos(\hat{\theta}) \cos(\hat{\phi}), r \cos(\hat{\theta}) \sin(\hat{\phi}), r \sin(\hat{\theta}))$. Al considerar una esfera unitaria el radio $r = 1$, y la distancia d se puede reducir al cálculo $d = -(\sin(\theta) \sin(\hat{\theta}) + \cos(\theta) \cos(\hat{\theta}) \cos(\varphi - \hat{\phi}))$, donde se puede ver que la distancia d depende únicamente de los valores de los ángulos.

Del mismo modo, la distancia d' entre el segundo par de elevación-acimut cuantificado escalarmemente $(\hat{\theta}', \hat{\phi}')$ y el par de elevación-acimut no cuantificado (θ, φ) en la esfera unitaria, se puede expresar como $d' = -(\sin(\theta) \sin(\hat{\theta}') + \cos(\theta) \cos(\hat{\theta}') \cos(\varphi - \hat{\phi}'))$.

El paso de procesamiento para encontrar la distancia entre el primer par de elevación-acimut cuantificado escalarmemente $(\hat{\theta}, \hat{\phi})$ y el par de elevación-acimut no cuantificado (θ, φ) , se muestra como 709 en la Figura 7.

El paso de procesamiento para encontrar la distancia entre el segundo par de elevación-acimut cuantificado escalarmemente $(\hat{\theta}', \hat{\phi}')$ y el par de elevación-acimut no cuantificado (θ, φ) , se muestra como 711 en la Figura 7.

Finalmente, el par de elevación-acimut cuantificado escalarmemente que tiene una medida de distancia mínima se selecciona como los valores cuantificados de elevación-acimut para la elevación-acimut (θ, φ) . Los índices correspondientes asociados al par cuantificado de elevación y acimut seleccionado, pasan entonces a formar el índice de dirección I_d .

El paso de procesamiento para encontrar la distancia mínima se muestra en la Figura 7 como 713.

El paso de procesamiento de selección entre los índices de $(\hat{\theta}, \hat{\phi})$ y $(\hat{\theta}', \hat{\phi}')$, tiene los índices de elevación-acimut cuantificados θ, φ según la distancia mínima, se muestra como 715 en la Figura 7.

Debe apreciarse que, aunque el esquema de cuantificación esférica anterior se ha definido en términos de una esfera unitaria, otras realizaciones pueden implementar el esquema de cuantificación anterior basándose en una esfera general cuyo radio no sea igual a uno. En tales realizaciones, el paso anterior de encontrar la distancia mínima, aún se mantiene, ya que el cálculo de la distancia mínima correspondiente tanto al primer par de elevación-acimut cuantificado escalarmemente como al segundo par de elevación cuantificado escalarmemente, es independiente del radio r .

Cabe señalar que los pasos de procesamiento anteriores para encontrar el índice de dirección I_d correspondiente a los valores cuantificados seleccionados para el par de elevación-acimut $(\hat{\theta}, \hat{\phi})$, se han descrito en términos de una realización preferida, utilizando un primer y un segundo valores cuantificados escalarmemente para el valor de elevación θ . Sin embargo, otras realizaciones pueden partir de la posición de utilizar más de dos valores cuantificados escalarmemente diferentes, para la elevación θ . En tales realizaciones, las entradas del libro de códigos más cercanas al valor de elevación θ se pueden seleccionar como los valores de elevación cuantificados en los que se basen los cálculos en curso. Por consiguiente, en tales realizaciones habrá una comparación final sobre un número correspondiente de medidas de distancia con el par de elevación-acimut cuantificado escalarmemente que produzca la distancia mínima que se selecciona como el valor cuantificado de elevación-acimut para la elevación-acimut (θ, φ) . Debe apreciarse que la operación de indexación descrita anteriormente, se refiere al proceso de asignar un valor cuantificado a un valor específico de un parámetro y, a continuación, utilizar un índice para hacer referencia al o referenciar el valor del parámetro cuantificado. El índice que representa el valor del parámetro cuantificado o se almacena o se transmite como parte de un flujo codificado de bits.

Debe apreciarse, además, que la medida de distorsión anterior, basada en la distancia entre dos puntos en una esfera unitaria, se puede utilizar en cualquier esquema de cuantificación de componentes direccionales que utilice los valores de elevación y acimut como componentes vectoriales. Con este fin, en otras realizaciones, un vector de elevación/acimut puede cuantificarse utilizando un libro de códigos vectoriales que comprenda una pluralidad de vectores de códigos de elevación/acimut. El libro de códigos vectoriales se puede buscar utilizando la rutina de búsqueda del vecino más cercano, en la que el vector de código de elevación/acimut con la medida de distancia mínima se selecciona como el vector de código más cercano al vector de elevación/acimut. En esta realización, el índice de dirección se determina entonces como el más cercano índice del vector de código de elevación/acimut. Debe observarse que la medida de distancia utilizada para la rutina de búsqueda del vecino más cercano del libro de códigos vectoriales, puede ser la misma que la distancia estándar L2 entre dos puntos x e y en una esfera, tal como se describió anteriormente para el caso de la retícula esférica. Es decir, para el par de elevación-acimut (θ, φ) las coordenadas de esfera se pueden expresar como $x = (r \cos(\theta) \cos(\varphi), r \cos(\theta) \sin(\varphi), r \sin(\theta))$, y para un vector de códigos de elevación/acimut $(\hat{\theta}, \hat{\phi})$ las coordenadas de esfera corresponden a $y = (r \cos(\hat{\theta}) \cos(\hat{\phi}), r \cos(\hat{\theta}) \sin(\hat{\phi}), r \sin(\hat{\theta}))$. Por lo tanto, la medida de distancia entre el vector de elevación/acimut y un vector de códigos de elevación/acimut del libro de códigos vectoriales, también se puede reducir al cálculo $d = -(\sin(\hat{\theta}) \sin(\hat{\theta}) + \cos(\theta) \cos(\hat{\theta}) \cos(\varphi - \hat{\phi}))$ para una esfera unitaria.

Puede emitirse el índice 306 de dirección l_d .

Sigue una descripción de un decodificador adecuado que no forma parte de la invención.

- 5 Con respecto a la Figura 3b, se muestra, según algunas realizaciones, un ejemplo de un extractor 137 de metadatos, y específicamente un extractor 350 de metadatos de dirección.

10 El extractor 350 de metadatos de dirección, en algunas realizaciones comprende una entrada 352 de cuantificación. Esto en algunas realizaciones se pasa desde el codificador de metadatos, o se acuerda de otro modo con el codificador. La entrada de cuantificación está configurada para definir la granularidad de las esferas dispuestas alrededor de una ubicación o posición de referencia.

15 El extractor 350 de metadatos de dirección en algunas realizaciones comprende una entrada 351 de índice de dirección. Esto se recibe desde el codificador, o se recupera por cualquier medio adecuado.

20 El extractor 350 de metadatos de dirección en algunas realizaciones comprende un posicionador 353 de esfera. El posicionador de esfera 353 está configurado para recibir como entrada la entrada de cuantificación y generar la disposición de esfera de la misma manera que se genera en el codificador. En algunas realizaciones, la entrada de cuantificación y el posicionador 353 de esfera son opcionales, y la disposición de la información de esferas se pasa desde el codificador, en lugar de generarse en el extractor.

25 El extractor de metadatos de dirección 350 comprende un convertidor de dirección de elevación-acimut (DI-EA) 355. El índice de dirección al convertidor de acimut de elevación 355 está configurado para recibir el índice de dirección y, además, la información de posición de esfera y generar una salida de elevación-acimut aproximada o cuantificada. En algunas realizaciones, la conversión se realiza según el siguiente algoritmo.

Entrada: l_d

Salida: (θ, ϕ)

1. Encontrar el índice de círculo i , de modo que $off(i) \leq l_d \leq off(i+1)$

35 2. Calcular el índice de círculo en el hemisferio: $i' = \begin{cases} \frac{i+1}{2} & \text{si } i \text{ es impar} \\ -\frac{i}{2} & \text{si } i \text{ es par} \end{cases}$

40 3. $\hat{\theta} = \frac{i'}{p}$

45 4. $\hat{\phi} = \begin{cases} \frac{l_d - off(i)}{\Delta\phi(i')}, & \text{si } i' \text{ es impar} \\ \frac{l_d - off(i)}{\Delta\phi(i')} + \frac{\Delta\phi(i')}{2}, & \text{si } i' \text{ es par} \end{cases}$

50 En otras realizaciones, el índice de dirección l_d puede decodificarse en un índice de elevación y un índice de acimut, con el fin de obtener el valor de elevación cuantificado $\hat{\theta}$ y el valor de acimut cuantificado $\hat{\phi}$ de los libros de códigos respectivos.

Con respecto a la Figura 8, se muestra un método ilustrativo para extraer los parámetros de dirección (o generar parámetros de dirección cuantificados).

55 La recepción de la entrada de cuantificación se muestra en la Figura 8 mediante el paso 801.

Entonces, el método puede determinar el posicionamiento de esfera basándose en la entrada de cuantificación, como se muestra en la Figura 8 mediante el paso 803.

60 Además, el método puede comprender recibir el índice de dirección, como se muestra en la Figura 8 mediante el paso 802.

65 Habiendo recibido el índice de dirección y la información de posicionamiento de esfera, el método puede comprender convertir el índice de dirección a un parámetro de dirección en forma de un parámetro de dirección cuantificado basado en la información de posicionamiento de esfera, como se muestra en la Figura 8 mediante el paso 805.

El método puede entonces generar el parámetro de dirección cuantificado, como se muestra en la Figura 8 mediante el paso 807.

Con respecto a la Figura 9, se muestra, según algunas realizaciones, un método ilustrativo para convertir el índice de dirección a un parámetro cuantificado de elevación-acimut (DI-EA), como se muestra en la Figura 8 mediante el paso 805.

En algunas realizaciones, el método comprende encontrar el valor de índice de círculo i , de modo que $off(i) \leq l_d \leq off(i + 1)$, como se muestra en la Figura 9 mediante el paso 901.

Habiendo determinado el índice de círculo, la siguiente operación es calcular el índice de círculo en la semiesfera a partir de la información de posicionamiento de esfera como se muestra en la Figura 9 mediante el paso 903.

Luego, una elevación cuantificada se determina basándose en el índice de círculo como se muestra en la Figura 9 mediante el paso 905.

Habiendo determinado la elevación cuantificada, el acimut cuantificado se determina basándose en el índice de círculo y la información de elevación como se muestra en la Figura 9 mediante el paso 907.

Aunque no se repite a lo largo del documento, debe entenderse que el procesamiento de audio espacial, tanto típicamente como en este contexto, tiene lugar en bandas de frecuencia. Esas bandas podrían ser, por ejemplo, las cajas de frecuencia de la transformación de tiempo-frecuencia, o bandas de frecuencia que combinan varios binarios. La combinación podría ser tal que se aproxime a las propiedades de la audición humana, tal como la resolución de frecuencia de Bark. En otras palabras, en algunos casos, podría medir y procesar el audio en áreas de tiempo-frecuencia que combinan varios de los intervalos de frecuencia b y/o índices de tiempo n . Por simplicidad, estos aspectos no fueron expresados por todas las ecuaciones anteriores. En caso de que se combinen muchas muestras de tiempo-frecuencia, típicamente un conjunto de parámetros tal como una dirección se estima para esa área de tiempo-frecuencia, y todas las muestras de tiempo-frecuencia dentro de esa área se sintetizan según ese conjunto de parámetros, tal como ese parámetro de dirección.

El uso de una resolución de frecuencia para el análisis de parámetros que es diferente de la resolución de frecuencia del banco de filtros aplicado es un enfoque típico en los sistemas de procesamiento de audio espacial.

Con respecto a la Figura 10, se muestra un dispositivo electrónico de ejemplo que puede usarse como dispositivo de análisis o síntesis. El dispositivo puede ser cualquier dispositivo o aparato electrónico adecuado. Por ejemplo, en algunas realizaciones, el dispositivo 1400 es un dispositivo móvil, un equipo de usuario, un ordenador de tipo tableta, un ordenador, un aparato de reproducción de audio, etc.

En algunas realizaciones, el dispositivo 1400 comprende al menos un procesador o unidad 1407 de procesamiento central. El procesador 1407 puede estar configurado para ejecutar diversos códigos de programa tales como los métodos tales como los descritos en la presente memoria.

En algunas realizaciones, el dispositivo 1400 comprende una memoria 1411. En algunas realizaciones, el al menos un procesador 1407 está acoplado a la memoria 1411. La memoria 1411 puede ser cualquier medio de almacenamiento adecuado. En algunas realizaciones, la memoria 1411 comprende una sección de código de programa para almacenar códigos de programa que pueden implementarse en el procesador 1407. Además, en algunas realizaciones, la memoria 1411 puede comprender además una sección de datos almacenados para almacenar datos, por ejemplo datos que se han procesado o que van a procesarse según las realizaciones como se describe en la presente memoria. El código de programa implementado almacenado dentro de la sección de código de programa y los datos almacenados dentro de la sección de datos almacenados pueden recuperarse por el procesador 1407 siempre que sea necesario a través del acoplamiento de memoria-procesador.

En algunas realizaciones, el dispositivo 1400 comprende una interfaz 1405 de usuario. En algunas realizaciones, la interfaz 1405 de usuario puede estar acoplada al procesador 1407. En algunas realizaciones, el procesador 1407 puede controlar el funcionamiento de la interfaz 1405 de usuario y recibir entradas de la interfaz 1405 de usuario. En algunas realizaciones, la interfaz 1405 de usuario puede permitir que un usuario introduzca comandos al dispositivo 1400, por ejemplo, a través de un teclado. En algunas realizaciones, la interfaz 1405 de usuario puede permitir que el usuario obtenga información a partir del dispositivo 1400. Por ejemplo, la interfaz 1405 de usuario puede comprender un elemento de visualización configurado para visualizar información del dispositivo 1400 al usuario. En algunas realizaciones, la interfaz 1405 de usuario puede comprender una pantalla táctil o una interfaz táctil que puede tanto permitir introducir información al dispositivo 1400 como visualizar adicionalmente información al usuario del dispositivo 1400. En algunas realizaciones, la interfaz 1405 de usuario puede ser la interfaz de usuario para comunicarse con el determinador de posición como se describe en la presente memoria.

En algunas realizaciones, el dispositivo 1400 comprende un puerto 1409 de entrada/salida. En algunas realizaciones, el puerto 1409 de entrada/salida comprende un transceptor. En tales realizaciones, el transceptor puede estar

acoplado al procesador 1407 y configurado para permitir una comunicación con otros aparatos o dispositivos electrónicos, por ejemplo, a través de una red de comunicación inalámbrica. En algunas realizaciones, el transceptor o cualquier transceptor o medios de transmisor y/o receptor adecuado puede estar configurado para comunicarse con otros dispositivos o aparatos electrónicos a través de un cable o acoplamiento por cable.

El transceptor puede comunicarse con un aparato adicional mediante cualquier protocolo de comunicaciones conocido adecuado. Por ejemplo, en algunas realizaciones, el transceptor o medios de transceptor, pueden utilizar un protocolo adecuado del sistema universal de telecomunicaciones móviles (UMTS), un protocolo de red de área local inalámbrica (WLAN), tal como, por ejemplo, IEEE 802.X, un protocolo adecuado de comunicación por radiofrecuencia de corto alcance, tal como Bluetooth, o una ruta de comunicación de datos por infrarrojos (IRDA).

El puerto 1409 de entrada/salida del transceptor puede estar configurado para recibir las señales y, en algunas realizaciones, determinar los parámetros como se describe en la presente memoria mediante el uso del procesador 1407 que ejecuta un código adecuado. Además, el dispositivo puede generar una salida de señal de mezcla descendente y de parámetros adecuada que va a transmitirse al dispositivo de síntesis.

En algunas realizaciones, el dispositivo 1400 puede emplearse como al menos parte del dispositivo de síntesis. Como tal, el puerto 1409 de entrada/salida puede estar configurado para recibir las señales de mezcla descendente y, en algunas realizaciones, los parámetros determinados en el dispositivo de captación o dispositivo de procesamiento como se describe en la presente memoria, y generar una salida de formato de señal de audio adecuada mediante el uso del procesador 1407 que ejecuta un código adecuado.

El puerto 1409 de entrada/salida puede acoplarse a cualquier salida de audio adecuada, por ejemplo, a un sistema de altavoces de múltiples canales y/o auriculares o similar.

En general, las diversas realizaciones de la invención pueden implementarse en hardware o circuitos de propósito especial, software, lógica o cualquier combinación de los mismos. Por ejemplo, algunos aspectos pueden implementarse en hardware, mientras que otros aspectos pueden implementarse en firmware o software que puede ejecutarse por un controlador, microprocesador u otro dispositivo informático, aunque la invención no se limita a los mismos. Aunque diversos aspectos de la invención pueden ilustrarse y describirse como diagramas de bloques, diagramas de flujo, o usando alguna otra representación gráfica, se entiende que estos bloques, aparatos, sistemas, técnicas o métodos descritos en la presente memoria pueden implementarse, como ejemplos no limitativos, en hardware, software, firmware, circuitos o lógica de propósito especial, hardware de propósito general o controlador u otros dispositivos informáticos, o alguna combinación de los mismos.

Las realizaciones de esta invención pueden implementarse mediante software informático ejecutable por un procesador de datos del dispositivo móvil, tal como en la entidad de procesador, o mediante hardware, o mediante una combinación de software y hardware. Además, en este sentido, cabe señalar que cualesquiera bloques del flujo lógico como en las Figuras puede representar pasos del programa, o circuitos lógicos interconectados, bloques y funciones, o una combinación de pasos de programa y circuitos lógicos, bloques y funciones. El software puede almacenarse en medios físicos tales como chips de memoria, o bloques de memoria implementados dentro del procesador, medios magnéticos tales como disco duro o disquetes, y medios ópticos tales como, por ejemplo, DVD y sus variantes de datos, CD.

La memoria puede ser de cualquier tipo adecuado para el entorno técnico local y puede implementarse usando cualquier tecnología de almacenamiento de datos adecuada, tal como dispositivos de memoria basados en semiconductores, dispositivos y sistemas de memoria magnéticos, dispositivos y sistemas de memoria ópticos, memoria fija y memoria extraíble. Los procesadores de datos pueden ser de cualquier tipo adecuado para el entorno técnico local, y pueden incluir uno o más de ordenadores de propósito general, ordenadores de propósito especial, microprocesadores, digital signal processors (procesadores de señales digitales - DSP), application specific integrated circuits (circuitos integrados específicos de aplicación - ASIC), circuitos de nivel de puerta y procesadores basados en arquitectura de procesador de múltiples núcleos, como ejemplos no limitativos.

Las realizaciones de las invenciones pueden ponerse en práctica en diversos componentes tales como módulos de circuito integrado. El diseño de circuitos integrados es, en gran medida, un proceso altamente automatizado. Hay herramientas de software complejas y potentes disponibles para convertir un diseño de nivel lógico en un diseño de circuito de semiconductores listo para grabarse y formarse en un sustrato semiconductor.

Los programas pueden enrutar automáticamente conductores y ubicar componentes en un chip semiconductor, utilizando reglas de diseño bien establecidas, así como bibliotecas de módulos de diseño previamente almacenados. Una vez que se ha completado el diseño para un circuito de semiconductor, el diseño resultante, en un formato electrónico normalizado (p. ej., Opus, GDSII o similar) puede transmitirse a una instalación de fabricación de semiconductores o "fab" para su fabricación.

La descripción anterior ha proporcionado, a modo de ejemplo y como ejemplos no limitativos, una descripción completa e informativa de la realización ilustrativa de esta invención. Sin embargo, diversas modificaciones y adaptaciones

pueden resultar evidentes para los expertos en las técnicas relevantes a la vista de la descripción anterior, cuando se lee junto con los dibujos adjuntos y las reivindicaciones adjuntas. Sin embargo, el alcance de esta invención es como se define en las reivindicaciones adjuntas.

5

10

15

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Un aparato para la codificación de señal de audio espacial, que comprende:

5 medios para determinar, para dos o más señales de audio, al menos un parámetro de audio espacial para proporcionar reproducción de audio espacial, comprendiendo el al menos un parámetro de audio espacial un parámetro de dirección con un componente de elevación y un componente de acimut;

10 medios para cuantificar escalarmente el componente de elevación para dar un primer valor de elevación indexado de un libro de códigos que comprende valores de elevación indexados y valores de acimut indexados, en donde el primer valor de elevación indexado está asociado a una pluralidad de valores de acimut indexados, y medios para cuantificar escalarmente el componente de acimut para dar un primer valor de acimut indexado, en donde el primer valor de acimut indexado se extrae de la pluralidad de valores de acimut indexados asociados al primer valor de elevación indexado;

15 medios para cuantificar escalarmente el componente de elevación para dar un segundo valor de elevación indexado, en donde la cuantificación escalar del componente de elevación se limita a un valor de elevación indexado cuyo índice está cerca del índice del primer valor de elevación indexado del libro de códigos, en donde el segundo valor de elevación indexado está asociado a una pluralidad adicional de valores de acimut indexados, y medios para cuantificar escalarmente el componente de acimut para dar un segundo valor del acimut indexado, en donde el segundo valor del acimut indexado se extrae de la pluralidad adicional de valores de acimut indexados;

20 medios para determinar una primera distancia entre un punto en una esfera dado por el componente de elevación y el componente de acimut, y un punto en la esfera dado por el primer valor de elevación indexado y el primer valor de acimut indexado;

25 medios para determinar una segunda distancia entre un punto en la esfera dado por la elevación y el componente de acimut, y un punto en la esfera dado por el segundo valor de elevación indexado y el segundo valor acimutal indexado; y

30 medios para seleccionar uno del primer valor de elevación indexado y el primer valor de acimut indexado, o el segundo valor de elevación indexado y el segundo valor de acimut indexado, como un componente de elevación indexado y un componente de acimut indexado, dependiendo de un mínimo de la primera distancia y la segunda distancia.

2. El aparato según lo reivindicado en la reivindicación 1, en donde la primera distancia y la segunda distancia son, cada una, una distancia estándar L2 medida entre dos puntos en la esfera.

3. El aparato según lo reivindicado en las reivindicaciones 1 y 2, en donde el componente de elevación y el componente de acimut están indexados con respecto a un punto en una retícula esférica, en donde la retícula esférica comprende una serie de puntos dispuestos en forma de esfera.

4. El aparato según lo reivindicado en la reivindicación 3, en donde la retícula esférica se forma cubriendo la esfera con esferas más pequeñas, en donde las esferas más pequeñas definen los puntos de la retícula esférica.

5. El aparato según lo reivindicado en la reivindicación 1, en donde el componente de elevación y el componente de acimut son componentes de un vector de acimut/elevación, en donde los medios para convertir el componente de elevación y el componente de acimut del parámetro de dirección a un valor de índice, comprenden:

50 medios para determinar una primera medida de distancia entre el vector de acimut/elevación y un primer vector de códigos de acimut/elevación de un libro de códigos vectoriales que comprende una pluralidad de vectores de códigos de acimut/elevación, y determinar al menos una segunda medida de distancia entre el vector de acimut/elevación y al menos un segundo vector de códigos de acimut/elevación del libro de códigos vectoriales;

55 medios para seleccionar una medida de distancia mínima a partir de la primera medida de distancia y la al menos segunda medida de distancia; y

medios para asignar el valor del índice como el índice del vector de código de acimut/elevación asociado a la medida de distancia mínima.

6. El aparato según lo reivindicado en la reivindicación 5, en donde la primera distancia y la segunda distancia son, cada una, una distancia estándar L2 medida entre dos puntos en una esfera.

7. El aparato según lo reivindicado en las reivindicaciones 2 a 4 y 6, en donde la esfera es una esfera unitaria, y en donde la distancia estándar L2 entre los dos puntos en la esfera unitaria, es de la forma $-(\sin(\theta_1) \sin(\theta_2) + \cos(\theta_1) \cos(\theta_2) \cos(\varphi_1 - \varphi_2))$, donde θ_1 y θ_2 son valores de elevación para un primer punto y un segundo punto de los dos puntos, en donde φ_1 y φ_2 son valores de acimut para el primer y el segundo puntos de los dos puntos.

8. Un método para la codificación de señal de audio espacial, que comprende:

determinar, para dos o más señales de audio, al menos un parámetro de audio espacial para proporcionar reproducción de audio espacial, comprendiendo el al menos un parámetro de audio espacial un parámetro de dirección con un componente de elevación y un componente de acimut; cuantificar escalarmente el componente de elevación para dar un primer valor de elevación indexado de un libro de códigos que comprende valores de elevación indexados y valores de acimut indexados, en donde el primer valor de elevación indexado está asociado a una pluralidad de valores de acimut indexados, y cuantificar escalarmente el componente de acimut para dar un primer valor de acimut indexado, en donde el primer valor de acimut indexado se extrae de la pluralidad de valores de acimut indexados asociados al primer valor de elevación indexado; cuantificar escalarmente el componente de elevación para dar un segundo valor de elevación indexado, en donde la cuantificación escalar del componente de elevación se limita a un valor de elevación indexado cuyo índice está cerca del índice del primer valor de elevación indexado del libro de códigos, en donde el segundo valor de elevación indexado está asociado a una pluralidad adicional de valores de acimut indexados, y cuantificar escalarmente el componente de acimut para dar un segundo valor de acimut indexado, en donde el segundo valor de acimut indexado se extrae de la pluralidad adicional de valores de acimut indexados; determinar una primera distancia entre un punto en una esfera dado por el componente de elevación y el componente de acimut, y un punto en la esfera dado por el primer valor de elevación indexado y el primer valor de acimut indexado; determinar una segunda distancia entre un punto en la esfera dado por la elevación y el componente de acimut, y un punto en la esfera dado por el segundo valor de elevación indexado y el segundo valor de acimut indexado; y seleccionar uno del primer valor de elevación indexado y el primer valor de acimut indexado, o el segundo valor de elevación indexado y el segundo valor de acimut indexado, como un componente de elevación indexado y un componente de acimut indexado, dependiendo de un mínimo de la primera distancia y la segunda distancia.

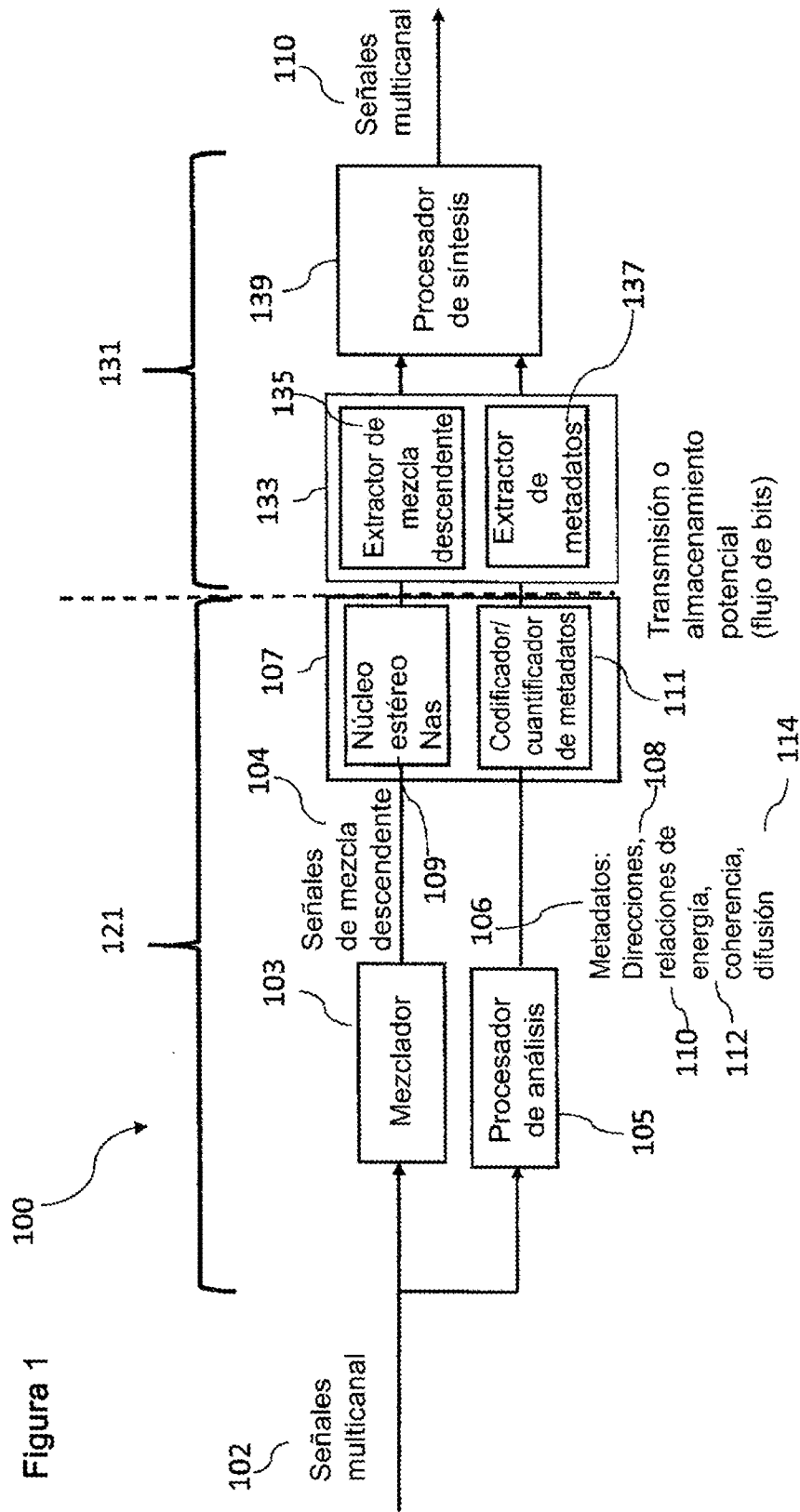
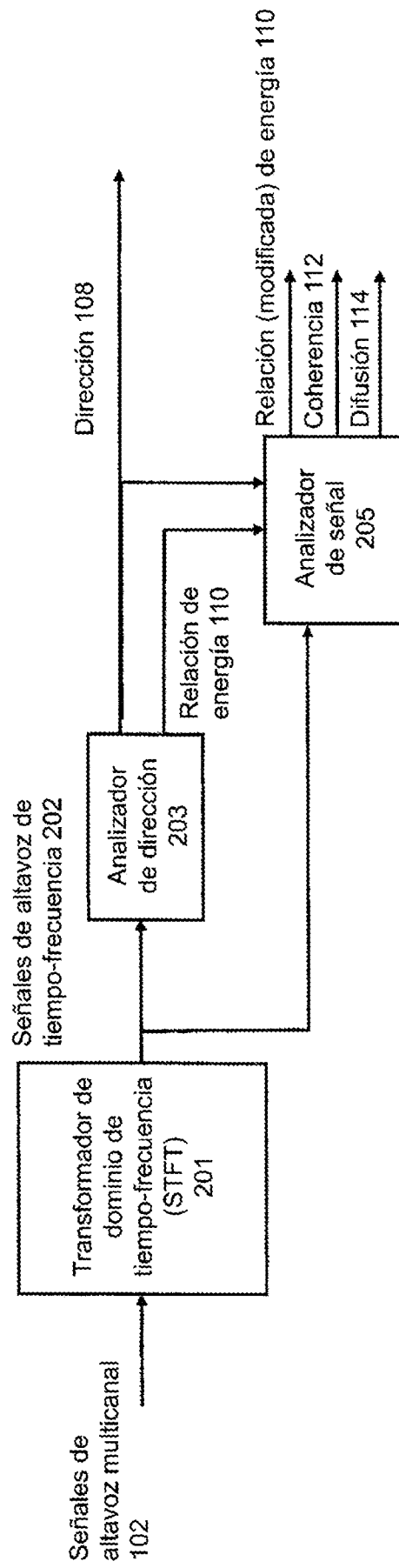
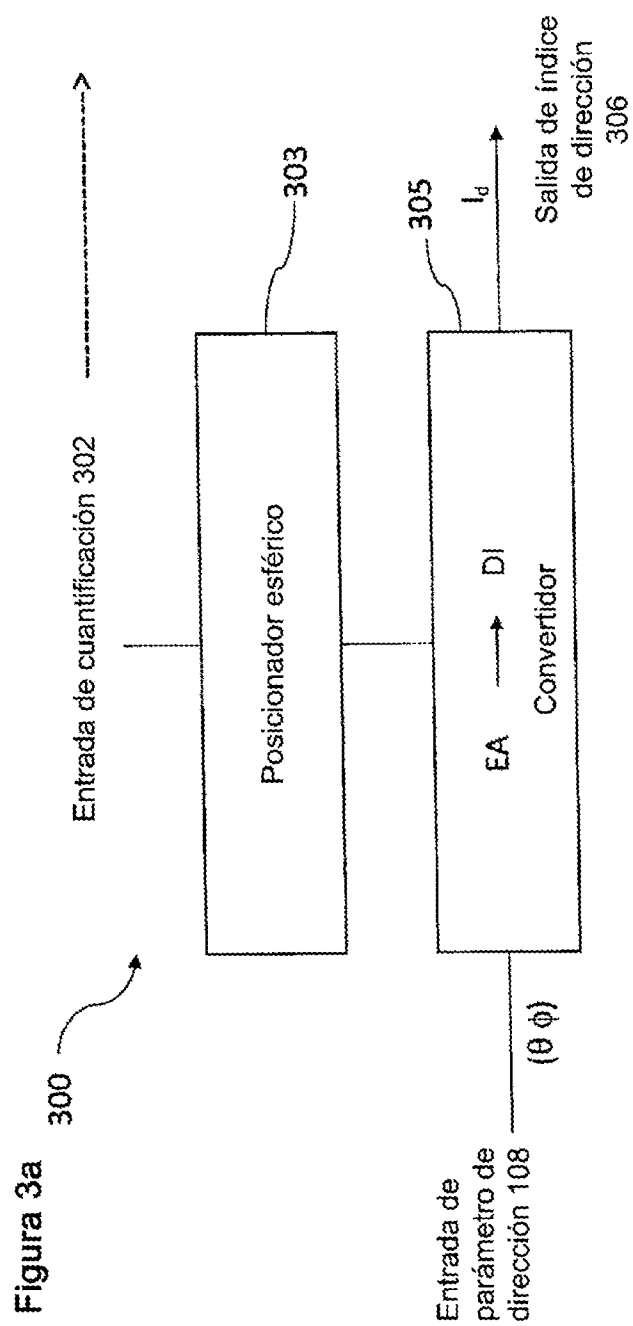


Figura 2

105





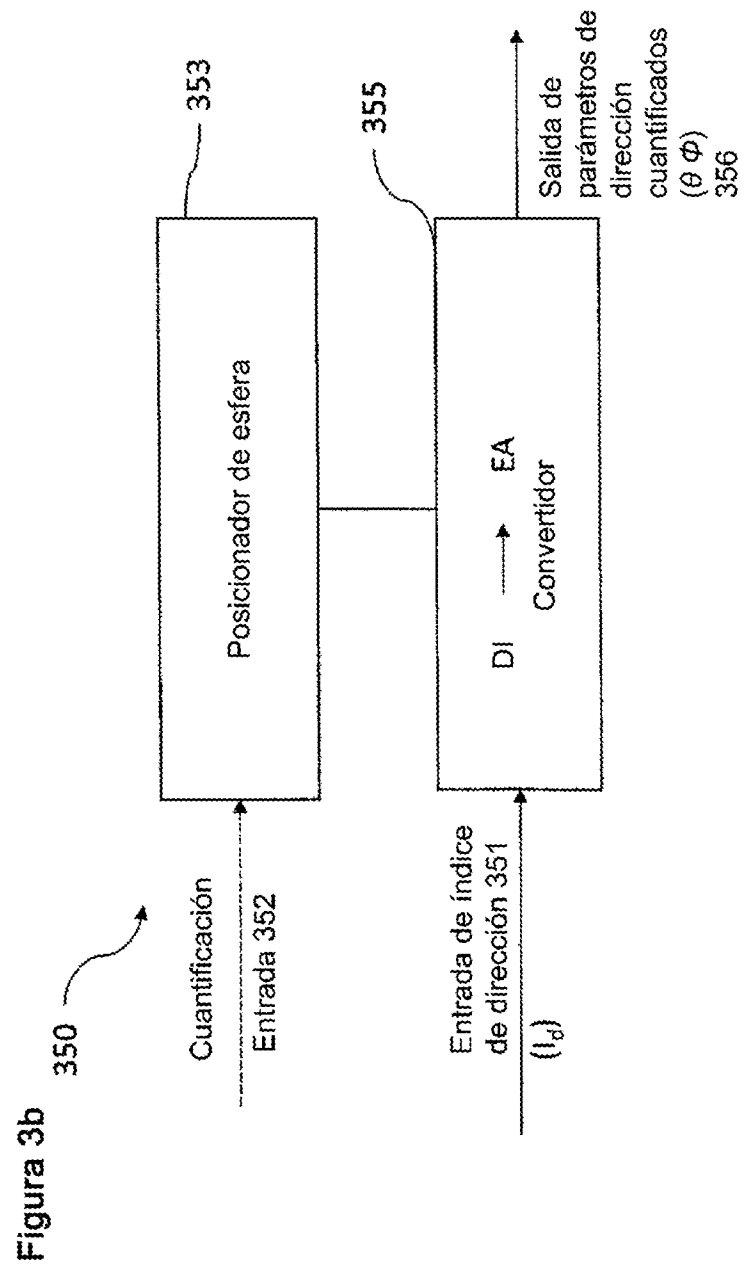


Figura 3d

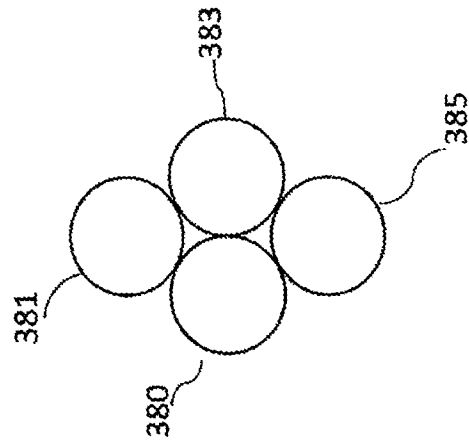


Figura 3c

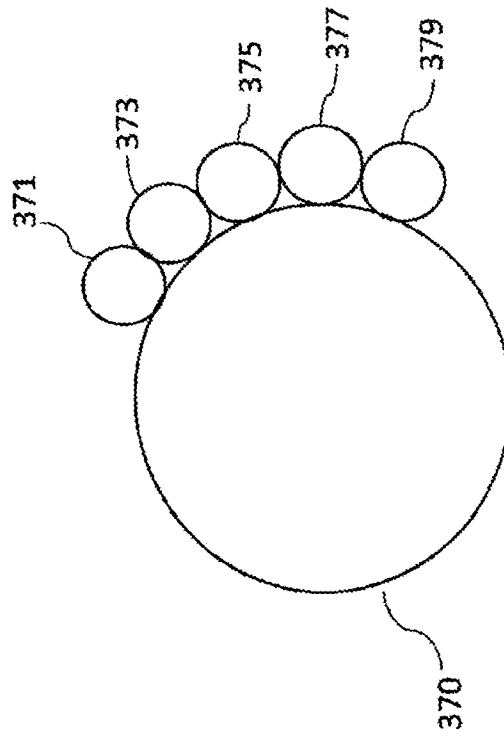
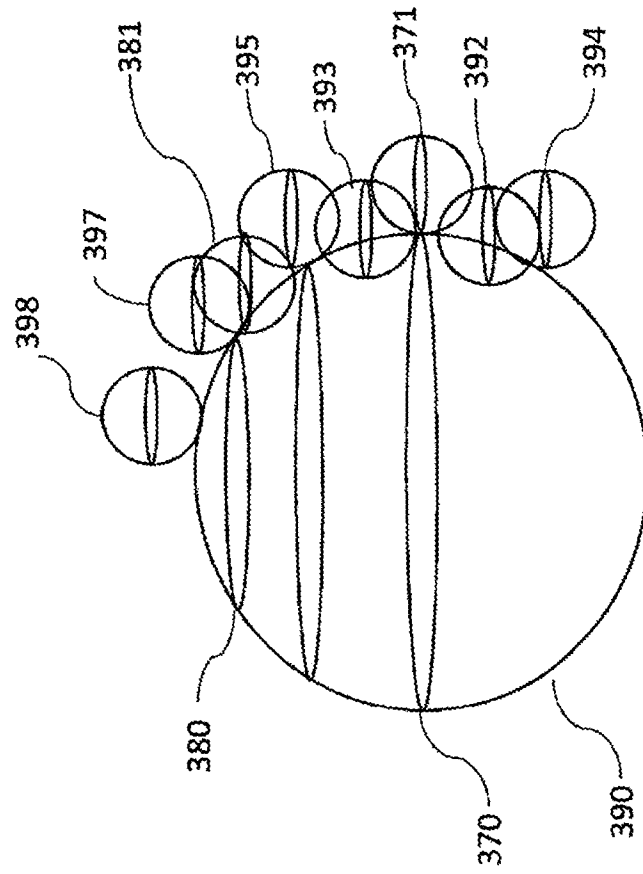


Figura 3e



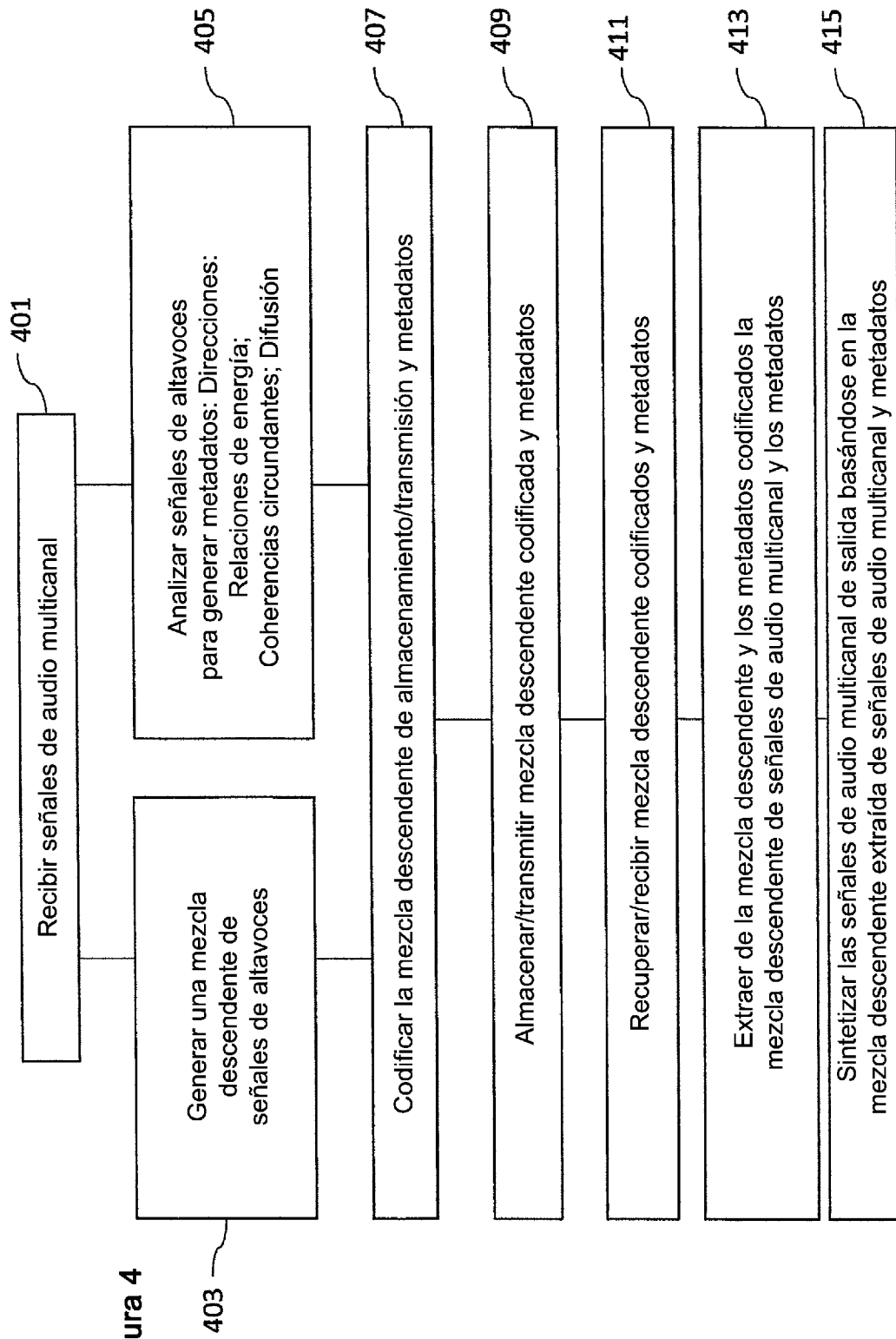
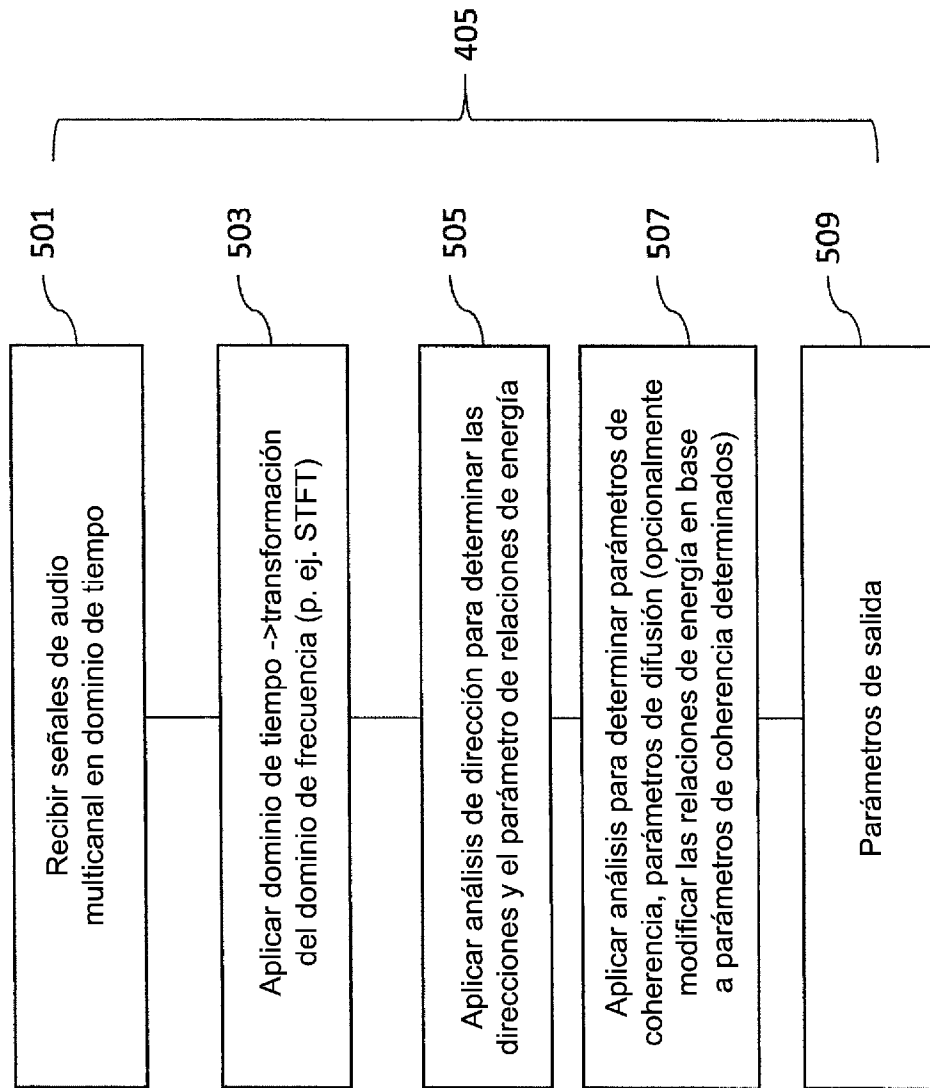


Figura 5



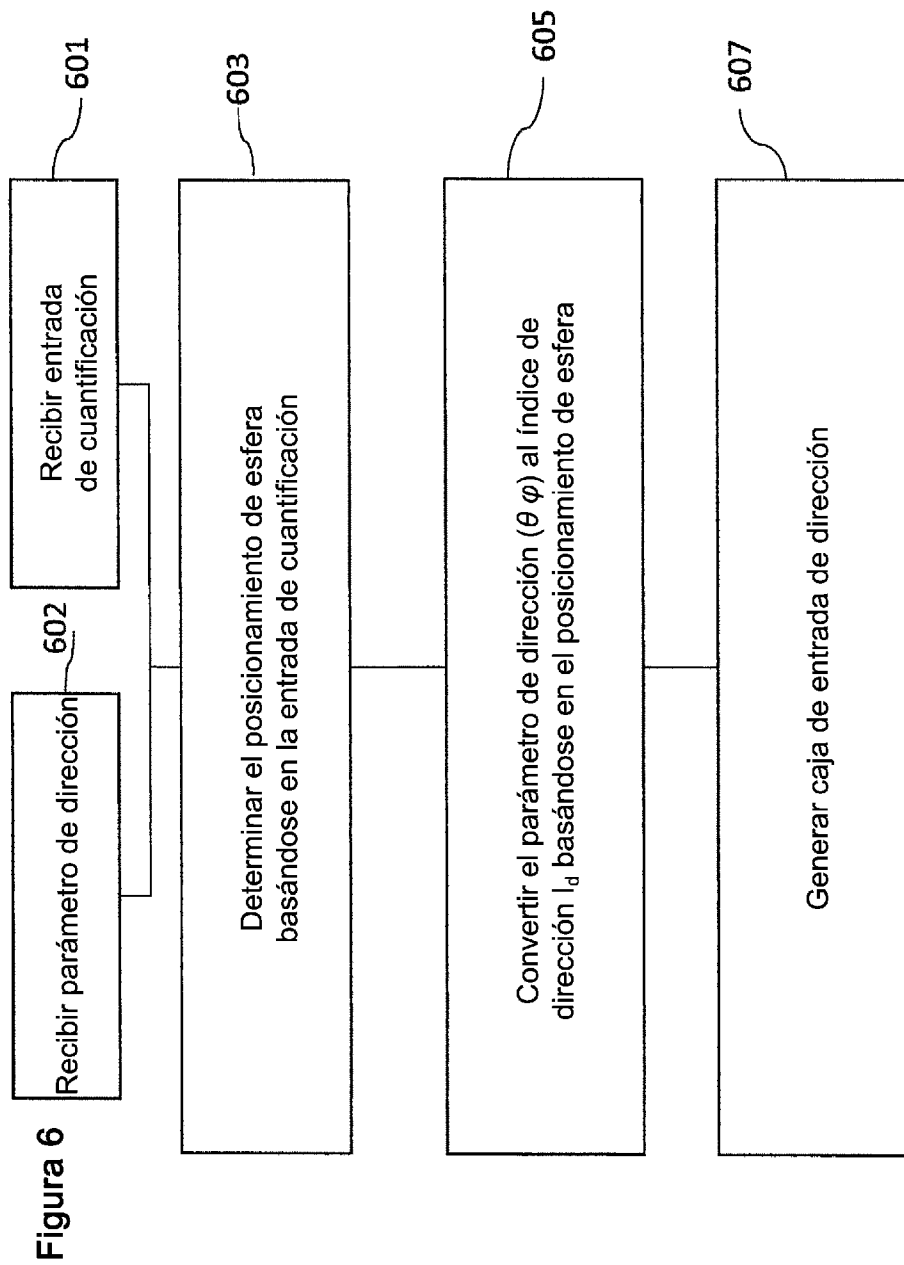
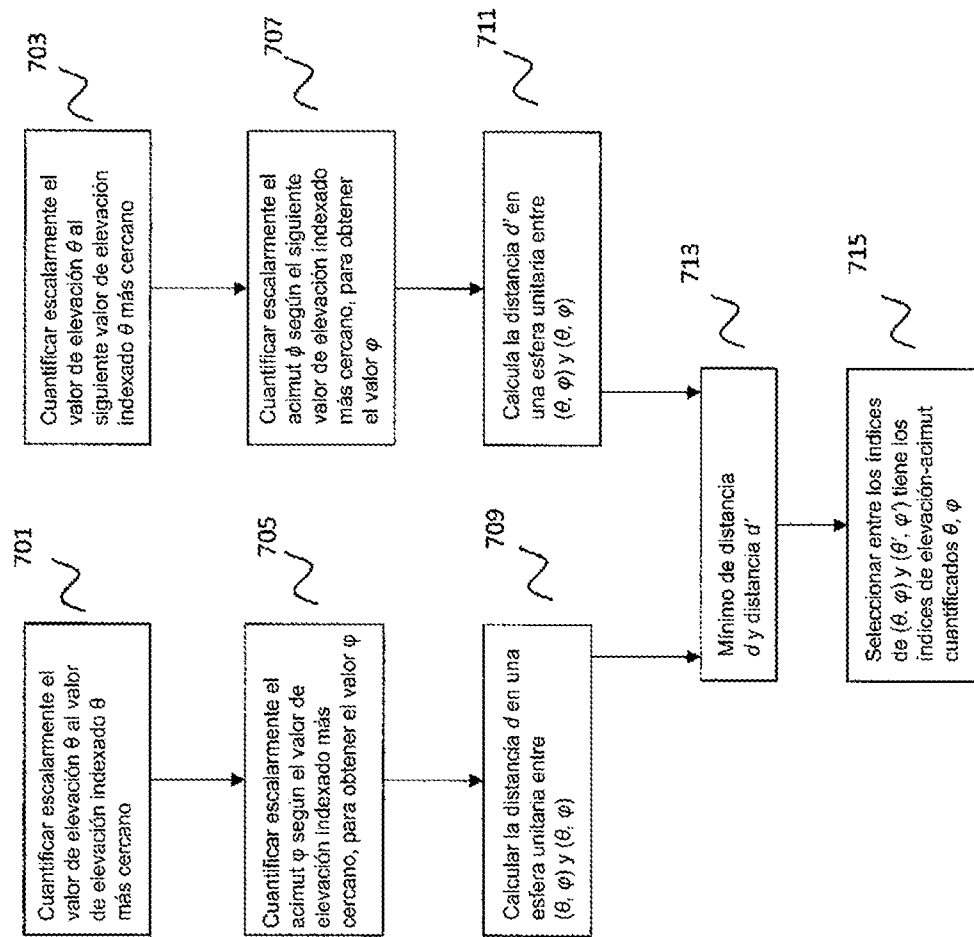
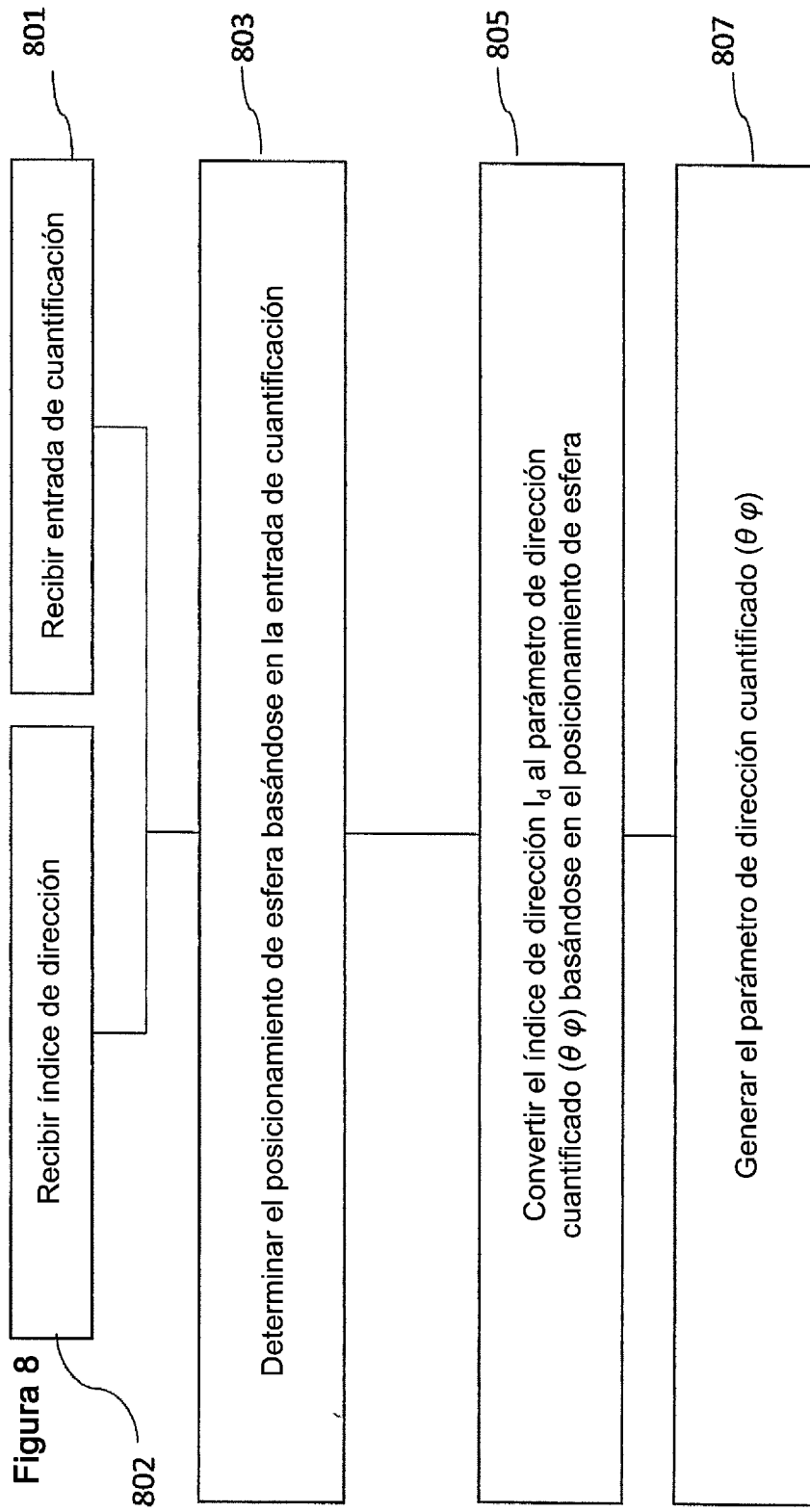


Figura 7





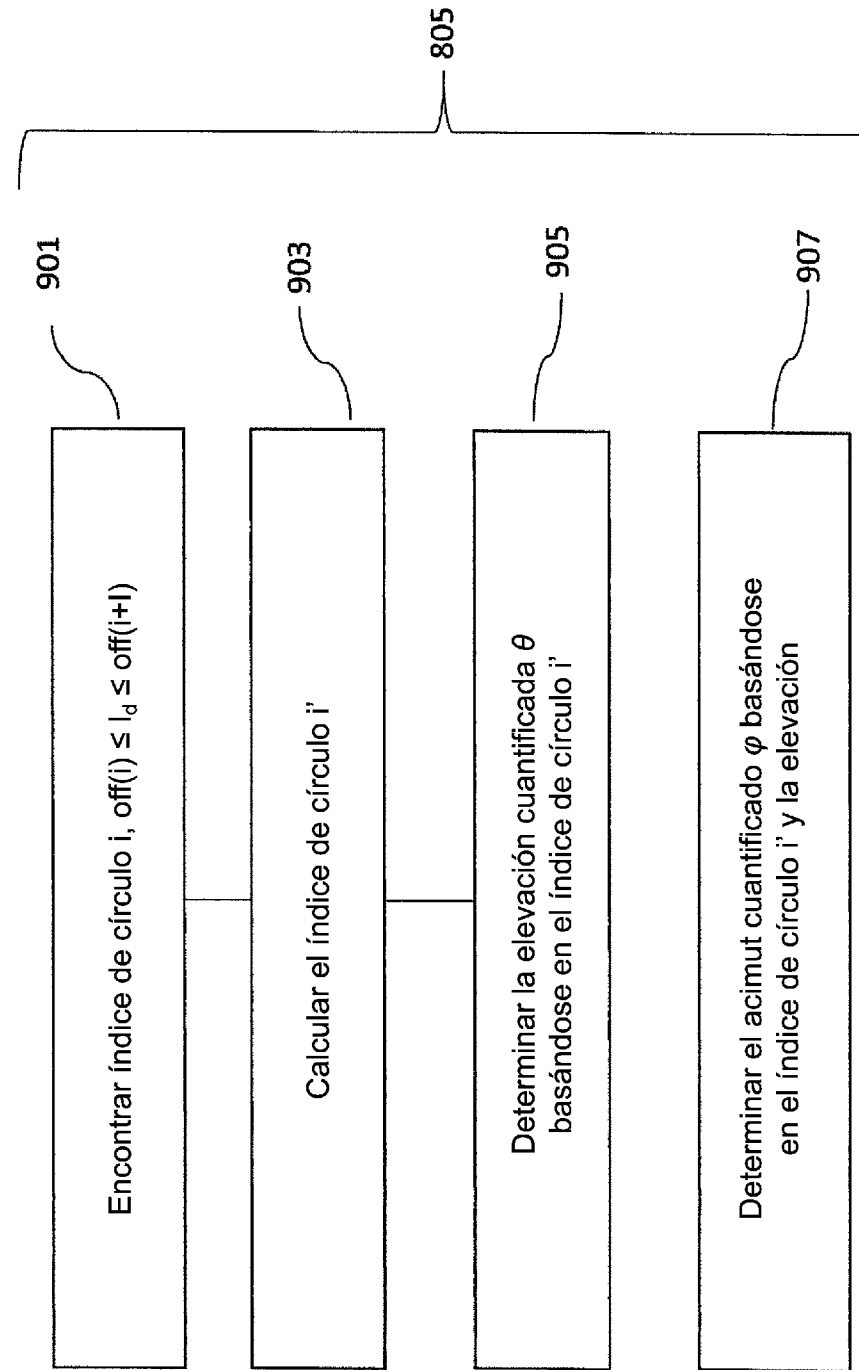


Figura 9

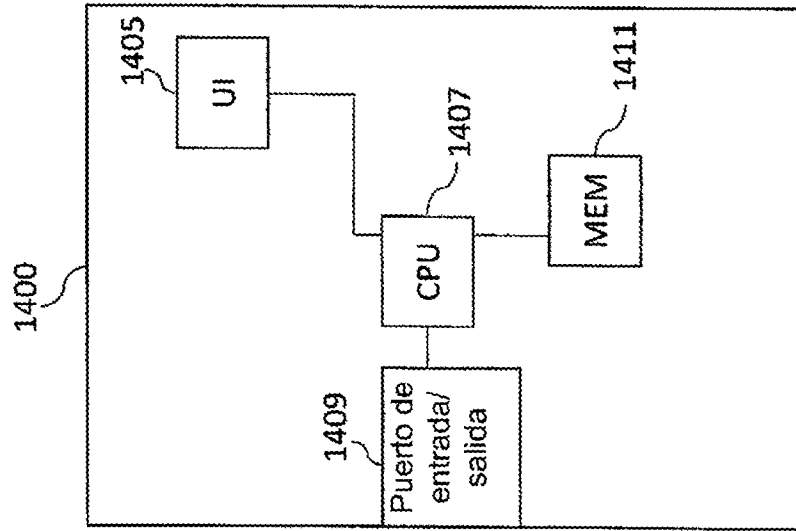


Figura 10