

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 965 395**

51 Int. Cl.:

G10L 19/008 (2013.01)

H04S 3/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.12.2017** **PCT/EP2017/084748**

87 Fecha y número de publicación internacional: **04.07.2019** **WO19129350**

96 Fecha de presentación y número de la solicitud europea: **28.12.2017** **E 17822336 (8)**

97 Fecha y número de publicación de la concesión europea: **15.11.2023** **EP 3732678**

54 Título: **Determinación de codificación de parámetros de audio espacial y decodificación asociada**

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
15.04.2024

73 Titular/es:
NOKIA TECHNOLOGIES OY (100.0%)
Karakaari 7
02610 Espoo, FI

72 Inventor/es:
LAAKSONEN, LASSE JUHANI;
RÄMÖ, ANSSI SAKARI;
VASILACHE, ADRIANA;
TAMMI, MIKKO y
VILERMO, MIIKKA

74 Agente/Representante:
DEL VALLE VALIENTE, Sonia

ES 2 965 395 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

Determinación de codificación de parámetros de audio espacial y decodificación asociada

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.

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.

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.

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.

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.

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, acimut (y difusión) de una dirección resultante, para cada subbanda de tiempo/frecuencia considerada. La cuantificación de estos componentes direccionales es un tema de investigación actual.

Li Gang et al. The perceptual Lossless Quantization of Spatial Parameter for 3D Audio Signals. Network and parallel Computing. Lecture Notes in Computer Science. Springer International Publishing, Cham, páginas 381 - 392. Disclose a Perceptual Lossless Quantization of Spatial Parameter (PLQSP) method, in which the azimuth and elevation quantization step sizes of spatial audio parameters are combined with Just Noticeable Difference (JND) characteristics of human auditory system to reduce spatial perceptual redundancy.

Yang Cheng et al. 3D Audio Coding Approach Based on Spatial Perception Features. China Communications. China Institute of Communications Piscataway, NJ, USA. Vol. 14, No. 11, 1st November 2017, páginas 126-140. Describir un enfoque de codificación de audio (3D) tridimensional que mejora la calidad perceptual espacial del sonido 3D. El enfoque se denomina método de cuantificación espacial de esferas concéntricas (CSSQ), donde se cuantifica la información lateral de distancia y se propone una cuantificación perceptual no uniforme basándose en las características de percepción espacial del sistema auditivo humano.

Resumen

Se proporciona según un primer aspecto un aparato para la codificación de señal de audio espacial configurada para: determinar, para dos o más señales de audio, al menos un parámetro de audio espacial para proporcionar reproducción de audio espacial, el al menos un parámetro de audio espacial comprende un parámetro de dirección con una elevación y un componente de acimut; definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas, disponiéndose las esferas más pequeñas en círculos de esferas donde un primer círculo de esferas comprende una de las esferas más pequeñas ubicadas con un centro a una elevación de 90 grados con respecto a una dirección de referencia de la esfera; y convertir la elevación y el componente de acimut del parámetro de dirección en un valor de índice basado en la retícula esférica definida.

El aparato configurado para definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas, las esferas más pequeñas dispuestas en círculos de esferas pueden configurarse además para seleccionar un primer número determinado de las esferas más pequeñas para un círculo adicional de la esfera, el círculo adicional definido por un diámetro de la una de las esferas más pequeñas ubicada con el centro a una elevación de 90 grados con respecto a la dirección de referencia de la esfera.

El círculo adicional puede ubicarse paralelo a un ecuador de la esfera.

El aparato configurado para definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas puede configurarse además para definir un orden de índice de círculo asociado con el primer círculo y el círculo adicional.

El aparato configurado para definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas, el aparato puede configurarse además para separar las esferas más pequeñas sobre la esfera aproximadamente de forma equidistante entre sí.

Se puede determinar un número de las esferas más pequeñas basándose en un valor de cuantificación de entrada.

El aparato configurado para convertir la elevación y el componente de acimut del parámetro de dirección a un valor de índice basado en la retícula esférica definida puede configurarse además para: determinar un valor de índice de círculo en base a un orden definido del primer círculo y en base al componente de elevación del parámetro de dirección; determinar un valor de índice intracírculo basado en el componente de acimut del parámetro de dirección; y generar el valor de índice basándose en la combinación del valor de índice intracírculo y un valor de compensación basándose en el valor de índice de círculo.

El aparato puede configurarse además para determinar al menos una dirección de referencia en base a un análisis de las dos o más señales de audio.

El aparato configurado para determinar la al menos una dirección de referencia en base a un análisis de las dos o más señales de audio puede configurarse para determinar la al menos una dirección de referencia en base a un parámetro direccional asociado con al menos una subbanda con un valor de energía de subbanda más alto.

El aparato configurado para definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas puede configurarse además para definir los círculos de esferas, de modo que los círculos de esferas son coplanares con la dirección de referencia y tienen diámetros que se definen basándose en una elevación desde la dirección de referencia de modo que un círculo más cercano a la dirección de referencia tiene un diámetro más grande.

El aparato configurado para definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas puede configurarse además para definir para el primer círculo la esfera más pequeña que tiene un primer diámetro y para el círculo adicional las esferas más pequeñas que tienen un segundo diámetro.

Según un segundo aspecto, se proporciona un aparato para la decodificación de señal de audio espacial, configurada para: determinar, al menos un índice de dirección asociado con dos o más señales de audio para proporcionar reproducción de audio espacial, representando al menos un índice de dirección un parámetro espacial con una elevación y un componente de acimut; definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas, disponiéndose las esferas más pequeñas en círculos de esferas donde un primer círculo de esferas comprende una de las esferas más pequeñas ubicadas con un centro a una elevación de 90 grados con respecto a una dirección de referencia de la esfera; y convertir al menos un índice de dirección a una elevación cuantificada y una representación de acimut cuantificada de la elevación y el componente de acimut del parámetro de dirección a un valor de índice basado en la retícula esférica definida.

El aparato producido para definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas, las esferas más pequeñas dispuestas en círculos de esferas pueden configurarse además para: seleccionar un primer número determinado de las esferas más pequeñas para un círculo adicional de las esferas, el círculo adicional definido por un diámetro de la una de las esferas más pequeñas ubicada con el centro a una elevación de 90 grados con respecto a la dirección de referencia de la esfera.

El círculo adicional puede ubicarse paralelo a un ecuador de la esfera.

5 El aparato configurado para definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas puede configurarse además para definir un orden de índice de círculo asociado con el primer círculo y el círculo adicional.

10 El aparato configurado para definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas puede configurarse además para separar las esferas más pequeñas sobre la esfera aproximadamente de forma equidistante entre sí.

Se puede determinar un número de las esferas más pequeñas basándose en un valor de cuantificación de entrada.

15 El aparato configurado para convertir al menos un índice de dirección a una elevación cuantificada y una representación de acimut cuantificada de la elevación y el componente de acimut del parámetro de dirección a un valor de índice basado en la retícula esférica definida puede configurarse además para: determinar un valor de índice de círculo en base al valor de índice; determinar la representación de elevación cuantificada del componente de elevación basándose en el valor de índice de círculo; y generar la representación de acimut cuantificada del componente de acimut basándose en un valor de índice restante después de eliminar un desplazamiento asociado con el valor de índice de círculo a partir del valor de índice.

El aparato puede configurarse además para determinar la al menos una dirección de referencia basándose en al menos uno de: un valor de dirección de referencia recibido; y un análisis basado en las dos o más señales de audio.

25 El aparato configurado para determinar la al menos una dirección de referencia en base a un análisis basado en las dos o más señales de audio puede configurarse para determinar la al menos una dirección de referencia en base a un parámetro direccional asociado con al menos una subbanda con un valor de energía de subbanda más alto.

30 El aparato configurado para definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas puede configurarse además para definir los círculos de esferas, de modo que los círculos de esferas son coplanares con la dirección de referencia y tienen diámetros definidos en función de una elevación desde la dirección de referencia de modo que un círculo más cercano a la dirección de referencia tiene un diámetro más grande.

35 El aparato configurado para definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas puede configurarse además para definir para el primer círculo la esfera más pequeña que tiene un primer diámetro y para el círculo adicional las esferas más pequeñas que tienen un segundo diámetro.

Un aparato que integra medios para realizar las acciones del método descrito anteriormente.

40 Un aparato configurado para realizar las acciones del método descrito anteriormente.

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

45 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.

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

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

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

55 Resumen de las figuras

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

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

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

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

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

5 la Figura 3c muestra esquemáticamente una configuración de ubicación de esfera de ejemplo 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;

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

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;

15 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 una operación de ejemplo de conversión de un índice de dirección a partir de un parámetro de dirección con más detalle;

20 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;

25 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

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

Realizaciones de la solicitud

30 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.

40 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, 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.

55 Por lo tanto, el concepto es un intento de determinar un parámetro de dirección cuantificado para metadatos espaciales y para indexar el parámetro basándose en una esfera práctica que cubre la distribución de las direcciones para definir una distribución de direcciones más uniforme. En particular, las realizaciones como se analiza con más detalle a continuación intentan producir una cuantificación y/o codificación que implementa una granularidad uniforme a lo largo del acimut y los componentes de elevación por separado (cuando estos dos parámetros se añaden por separado a los metadatos) y además tienen como objetivo producir una distribución uniforme de 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.

Por lo tanto, el concepto puede implementarse de tal manera que una retícula esférica, usada para la cuantificación de los parámetros de dirección, se define comenzando desde una dirección de referencia, de modo que dentro de la trama la información correspondiente a la dirección es relativa a la dirección más importante y la cantidad de información que debe codificarse es mínima. En la práctica, significa que el índice de dirección se envía primero para la subbanda que tiene, por ejemplo, la relación de energía más alta. Los índices de direcciones de las siguientes subbandas están formados por una retícula construida alrededor de la dirección principal y donde se determina que la dirección principal o de referencia tiene una elevación de ± 90 , en otras palabras, la dirección del polo "norte" o "sur" con respecto a la posición de referencia.

El índice de metadatos propuesto puede usarse a continuación junto a una señal de mezcla descendente ("canales"), para definir un formato inmersivo paramétrico que puede utilizarse, por ejemplo, para el códec 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, la relación de energía y los parámetros 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 secundario NAS 109 que está configurado para recibir las señales de mezcla descendente (o de otro modo) 104 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 alternativamente 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 se puede implementar 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 la corriente codificada de audio a un extractor de mezcla descendente 135 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 alternativamente un dispositivo específico que utiliza, por ejemplo, FPGA o ASIC.

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

La parte de "síntesis" 100 del sistema muestra además un procesador de síntesis 139 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 como las señales binaural o ambisónica, 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.

Después de esto, el sistema puede almacenar/transmitir las submezclas y los metadatos codificados como se muestra en la Figura 4 mediante el paso 409.

El sistema puede recuperar/recibir las submezclas y los 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 fotograma 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 del tiempo. Estos binarios de frecuencia pueden agruparse en subbandas que forman uno o más de los

binarios en una subbanda de un índice de banda $K = 0, \dots, K-1$. Cada subbanda k tiene un contenedor b más bajo k , bajo y un depósito más alto $b_{k, \text{alto}}$, y la subbanda contiene todos los intervalos de $b_{k, \text{bajo}}$ a $b_{k, \text{alto}}$. 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 de dirección 203 puede configurarse para recibir las señales de frecuencia de tiempo 202 y basándose en estas señales estima los parámetros de dirección 108. 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 de dirección 203 puede configurarse para proporcionar un acimut para cada banda de frecuencia y un marco temporal, denotado como acimut $\phi(k, n)$ y elevación $\theta(k, n)$. El parámetro de dirección 108 también se puede pasar a un analizador de señal 205. El analizador de dirección 203 también puede configurarse para determinar un parámetro de relación de energía 110. 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, por ejemplo, usando una medida de estabilidad de la estimación direccional, o usando cualquier medida de correlación, o cualquier otro método adecuado para obtener un parámetro de relación.

La dirección estimada 108 y la relación de energía 110 se emiten (y pasan a un codificador).

En algunas realizaciones, el procesador de análisis 105 comprende un analizador de señal 205. El analizador de señal 205 está configurado para recibir parámetros de dirección (tales como el acimut $\phi(k, n)$ y la elevación $\theta(k, n)$ 108 y los parámetros de relación de energía 110 desde el analizador de dirección 203. El analizador de señal 205 se configura además para recibir las señales de tiempo-frecuencia $s_i(b, n)$ 202 desde el transformador de dominio de tiempo-frecuencia 201. 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 consiste potencialmente en varios binarios 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 se combinan durante 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 consiste en varios binarios de frecuencia b . Lo mismo se aplica para todos los parámetros espaciales analizados en la presente memoria.

El analizador de señal 205 se configura para producir un número de parámetros de señal tales como coherencia y difusión, ambos analizados en el dominio de tiempo-frecuencia. Además, en algunas realizaciones, el analizador de señal 205 puede configurarse para modificar las relaciones de energía estimadas $r(k, n)$. 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 de metadatos de dirección 300 en algunas realizaciones comprende una entrada de cuantificación 302. La entrada de cuantificación, que también puede ser conocida 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 determina el parámetro de dirección. En algunas realizaciones, la entrada de cuantificación es un valor predefinido o fijo. Además, la entrada de cuantificación 302 puede definir otros aspectos o entradas que permiten la configuración de las operaciones de cuantificación esférica. Por ejemplo, la entrada de cuantificación 302 comprende una dirección de referencia (por ejemplo, en relación con una dirección absoluta tal como el norte magnético). La dirección de referencia se determina o define en base a un análisis de las señales de entrada. Por ejemplo, la dirección de referencia se determina en base a la dirección de la subbanda con el mayor valor de energía o relación de energía.

El codificador de metadatos de dirección 300 comprende un posicionador de esfera 303. 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 definen una retícula de direcciones casi equidistantes.

El concepto como se muestra en la presente memoria es uno en el que una esfera se define en relación con la ubicación de referencia y una dirección 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 la Figura 3c. Por ejemplo, la Figura 3c muestra una configuración de dirección de referencia "polar" ilustrativa que muestra una primera esfera 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 círculos) 381, 391, 393, 395, 397 y 399 ubicadas de modo que cada esfera más pequeña tiene una circunferencia que en un punto toca la circunferencia de esfera principal y al menos un punto adicional que toca al menos una circunferencia de esfera más pequeña adicional. Por lo tanto, como se muestra en la Figura 3c, la esfera más pequeña 381 toca la esfera principal 370 y las esferas más pequeñas 391, 393, 395, 397 y 399. Además, la esfera 381 más pequeña está ubicada de modo que el centro de la esfera más pequeña está ubicado en la línea de elevación de +/- 90 grados (el eje z) que se extiende a través del centro de la esfera 370 principal.

Las esferas 391, 393, 395, 397 y 399 más pequeñas están situadas de tal manera que cada una toca la esfera principal 370, la esfera más pequeña 381 y adicionalmente un par de esferas más pequeñas adyacentes. Por ejemplo, la esfera más pequeña 391 toca adicionalmente las esferas adyacentes 399 y 393 más pequeñas, la esfera más pequeña 393 toca adicionalmente las esferas más pequeñas adyacentes 391 y 395, la esfera más pequeña 395 toca adicionalmente esferas más pequeñas adyacentes 393 y 397, la esfera más pequeña 397 toca adicionalmente esferas más pequeñas adyacentes 399 y 391, y la esfera más pequeña 399 toca adicionalmente esferas más pequeñas 397 y 391 adyacentes.

Por lo tanto, la esfera más pequeña 381 define un cono 380 o un ángulo sólido alrededor de la línea de elevación de +90 grados y las esferas más pequeñas 391, 393, 395, 397 y 399 definen un cono adicional 390 o un ángulo sólido alrededor de la línea de elevación de 90 grados, en donde el cono adicional es un ángulo de sólido más grande que el cono.

En otras palabras, se puede considerar que la esfera 381 más pequeña (que define un primer círculo de esferas) se ubica en una primera elevación (con el centro de esfera más pequeño + 90 grados), y las esferas más pequeñas 391, 393, 395, 397 y 399 (que definen un segundo círculo de esferas) se consideran ubicadas una segunda elevación (con los centros de esfera más pequeños <90 grados) con respecto a la esfera principal y con una elevación inferior al círculo anterior.

Esta disposición puede repetirse adicionalmente con círculos adicionales de esferas de contacto ubicadas en elevaciones adicionales con respecto a la esfera principal y con una elevación menor que los círculos anteriores.

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

Entrada: resolución angular para elevación, $\Delta \theta$ (idealmente de tal manera que $\frac{\pi}{2\Delta \theta}$ es un entero)

Salida: número de círculos, N_c y número de puntos en cada círculo, $n(i)$, $i = 0, N_c-1$

1. $n(0) = 1$

2. $M = \left\lceil \frac{\pi}{2 \Delta \theta} \right\rceil$

3. Para $i = 1; M-1$

a. $n(i) = \pi \sin(\Delta \theta \cdot i) / \sin \frac{\Delta \theta}{2}$

b. $\theta(i) = \frac{\pi}{2} - i \cdot \Delta \theta$ (elevación)

c. $\Delta \phi(i) = 2 \pi / n(i)$

d. Si i es impar

i. $\phi_i(0) = 0$

e. De lo contrario

i. $\phi_i(0) = \frac{\Delta \phi(i)}{2}$ (primer valor de acimut en el círculo i)

f. Finalizar si

4. Finalizar para

Por lo tanto, según lo anterior, 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 bajo el equipo (el plano definido por los ejes X-Y).

Además, como se discutió anteriormente, 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.

En otras palabras, los círculos se ordenan a partir del "Polo Norte" hacia abajo.

En otra realización, el número de puntos a lo largo de los círculos paralelos al ecuador $n(i) = \pi \sin(\Delta \theta \cdot i) / \sin \frac{\Delta \theta}{2}$ también se puede obtener mediante $n(i) = \pi \sin(\Delta \theta \cdot i) / \lambda_i \sin \frac{\Delta \theta}{2}$, donde $\lambda_i \geq 1$, $\lambda_i \leq \lambda_{i+1}$. En otras palabras, las esferas a lo largo de los círculos paralelos al ecuador tienen radios más grandes ya que están más lejos del polo norte, es decir, están más lejos del polo norte de la dirección principal.

Para una trama de metadatos, hay al menos información de dirección e información de relación de energía para cada subbanda. La dirección principal en algunas realizaciones puede decidirse como la dirección dada por los datos en la subbanda que tiene al menos la relación de energía más grande. La información de dirección principal se proporciona por tanto por los datos direccionales correspondientes a la subbanda con al menos la relación de energía más alta. Si, basándose en los valores de relación de energía, hay más de una dirección principal (es decir, los valores de relación de energía más grandes están muy próximos entre sí) la dirección principal se puede obtener como una combinación ponderada de estas direcciones.

Para una trama, si la información de dirección principal (ϕ_D, θ_D) se envía primero, la información de dirección correspondiente a las siguientes subbandas se envía en relación con la dirección principal. Esto significa que los valores para la elevación y el acimut para las subbandas posteriores son $\phi^{(i)} - \phi_D, \theta^{(i)} - \theta_D$ y se cuantifican e indexan en la red propuesta por el algoritmo A.

El posicionador de esfera ha determinado el número de círculos y el número de círculos, N_c , 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 de índice de dirección 305 de acimut a acimut está configurado para recibir la entrada de parámetro de dirección 108 y la información de posicionador de esfera y convertir el valor de acimut de elevación de la entrada 108 de parámetro de dirección a un índice de dirección que va a emitirse.

El convertidor de índice de dirección 305 de acimut a dirección (EA-DI) se configura para realizar esta conversión según el siguiente algoritmo:

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

Salida: I_d

La granularidad $\Delta \theta$ a lo largo de la elevación se conoce. Los valores θ son de un conjunto discreto de valores, correspondientes a las direcciones indexadas. Se conoce el número de puntos en cada círculo y los desplazamientos correspondientes, según el orden de los círculos considerado, *apagado* (i).

1. Encuentre el índice de círculo $i = (-\theta + \pi/2)/\Delta \theta$

2. Encuentre el índice del acimut dentro del círculo $j = \left\lfloor \frac{\varphi}{\Delta \Phi(i)} \right\rfloor$

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

Puede generarse el índice de dirección I_d 306.

Con respecto a la Figura 6, se muestra un método de ejemplo 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 procedimiento determina el posicionamiento de esfera basándose en la entrada de cuantificación como se muestra en la Figura 6 mediante el paso 603.

Además, el método comprende 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 en un índice de dirección basado 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.

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

El método comienza encontrando un índice de círculo i del valor de elevación θ como se muestra en la Figura 7 mediante el paso 701.

Habiendo determinado el índice de círculo, se encuentra el índice del acimut basado en el valor de acimut φ como se muestra en la Figura 7 mediante el paso 703.

Habiendo determinado el índice de círculo i y el índice del acimut, se determina la dirección añadiendo el valor del índice del acimut al desplazamiento asociado con el índice circular como se muestra en la Figura 7 mediante el paso 705.

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

El extractor de metadatos de dirección 350 comprende una entrada de cuantificación 352. Esto 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. Además, la entrada de cuantificación define además la configuración de las esferas, por ejemplo, la orientación de la dirección de referencia (en relación con una dirección absoluta tal como el norte magnético).

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

5 El extractor de metadatos de dirección 350 comprende un posicionador de esfera 353. 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. La entrada de cuantificación y el posicionador de esfera 353 es opcional y la disposición de la información de esferas se pasa desde el codificador en lugar de ser generada en el extractor.

10 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. La conversión se puede realizar según el siguiente algoritmo.

15 Entrada: I_d

Salida: (θ, ϕ)

1. Encuentre el índice de círculo i , de modo que $apagado(i) \leq I_d \leq apagado(i+1)$

2.

$$\theta = i \cdot \Delta\theta$$

3.

$$\phi = \begin{cases} \frac{I_d - off(i)}{\Delta\phi(i)}, & \text{si } i \text{ es impar} \\ \frac{I_d - off(i)}{\Delta\phi(i)} + \frac{\Delta\phi(i)}{2}, & \text{si } i \text{ es par} \end{cases}$$

Para el caso bidimensional, donde solo el acimut describe la dirección, después de decidir en la dirección principal dada por ϕ_M , la indexación de las direcciones viene dada por el siguiente orden: $\phi_M, \phi_M + \Delta\phi, \phi_M - \Delta\phi, \phi_M + 2\Delta\phi, \phi_M - 2\Delta\phi$.

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

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

Entonces el procedimiento determina el posicionamiento de esfera basándose en la entrada de cuantificación como se muestra en la Figura 8 mediante el paso 803.

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

Habiendo recibido el índice de dirección y la información de posicionamiento de esfera, el método comprende convertir el índice de dirección en 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 genera después el parámetro de dirección cuantificado como puede verse en la Figura 8 mediante el paso 807.

Con respecto a la Figura 9, se muestra un método de ejemplo para convertir el índice de dirección en 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 $apagado(i) \leq I_d \leq apagado(i+1)$ 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 puede usar un protocolo de universal mobile telecommunications system (sistema universal de telecomunicaciones móviles - UMTS), un protocolo de wireless local area network (red de área local inalámbrica - WLAN) tal como por ejemplo IEEE 802.X, un protocolo de comunicación de radiofrecuencia de corto alcance adecuado tal como Bluetooth, o infrared data communication pathway (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, tales como los proporcionados por Synopsys, Inc. de Mountain View, California y Cadence Design, de San Jose, California, enrutan automáticamente conductores y localizan componentes en un chip de semiconductor usando 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.

REIVINDICACIONES

1. Un aparato configurado para:
 - 5 determinar, para dos o más señales de audio, al menos un parámetro de audio espacial para proporcionar reproducción de audio espacial, el al menos un parámetro de audio espacial comprende un parámetro de dirección con una elevación y un componente de acimut; definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas, en donde las esferas más pequeñas son cada una más pequeñas que la esfera, en donde las esferas más pequeñas están dispuestas en círculos de esferas, en donde un primer círculo de esferas comprende una de las esferas más pequeñas ubicadas con un centro a una elevación de 90 grados con respecto a una dirección de referencia de la esfera; y
 - 10 convertir la elevación y el componente de acimut del parámetro de dirección con un valor de índice basado en la retícula esférica definida.
- 15 2. El aparato según la reivindicación 1, en donde el aparato configurado para definir la retícula esférica generada cubriendo la esfera con las esferas más pequeñas, está configurado además para seleccionar un primer número determinado de las esferas más pequeñas para un círculo adicional de la esfera, en donde el círculo adicional está definido por un diámetro de una de las esferas más pequeñas ubicadas con el centro a la elevación de 90 grados con respecto a la dirección de referencia de la esfera.
- 20 3. El aparato según la reivindicación 2, en donde el círculo adicional está ubicado paralelo a un ecuador de la esfera.
- 25 4. El aparato según cualquiera de las reivindicaciones 2 a 3, en donde el aparato configurado para definir una retícula esférica generada cubriendo la esfera con las esferas más pequeñas, está configurado además para definir un orden de índice de círculo asociado con el primer círculo y el círculo adicional.
- 30 5. El aparato según cualquiera de las reivindicaciones 2 a 4, en donde el aparato configurado para definir la retícula esférica generada cubriendo la esfera con las esferas más pequeñas, está configurado además para separar las esferas más pequeñas sobre la esfera aproximadamente equidistantemente entre sí.
- 35 6. El aparato según cualquiera de las reivindicaciones 2 a 5, en donde el aparato configurado para definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas está configurada para definir un número de esferas más pequeñas basándose en un valor de cuantificación de entrada.
- 40 7. El aparato según cualquiera de las reivindicaciones 1 a 6, en donde el aparato configurado para convertir la elevación y el componente de acimut del parámetro de dirección con respecto al valor de índice basado en la retícula esférica definida se configura además para:
 - 45 determinar un valor de índice de círculo basándose en un orden definido del primer círculo y en base al componente de elevación del parámetro de dirección;
 - determinar un valor de índice intracírculo basado en el componente de acimut del parámetro de dirección; y
 - generar el valor de índice basándose en la combinación del valor de índice intracírculo y un valor de compensación basándose en el valor de índice de círculo.
- 50 8. El aparato según cualquiera de las reivindicaciones 1 a 7, en donde el aparato está configurado además para determinar la al menos una dirección de referencia en base a un análisis de las dos o más señales de audio.
- 55 9. El aparato según la reivindicación 8, en donde el aparato configurado para determinar la al menos una dirección de referencia en base al análisis de las dos o más señales de audio está configurado para determinar la al menos una dirección de referencia en base a un parámetro direccional asociado con al menos una subbanda con un valor de energía de subbanda más alto.
- 60 10. El aparato según cualquiera de las reivindicaciones 1 a 9, en donde el aparato configurado para definir la retícula esférica generada cubriendo la esfera con esferas más pequeñas está configurado además para definir los círculos de esferas de modo que los círculos de esferas son coplanares con la dirección de referencia y tienen diámetros que se definen basándose en una elevación desde la dirección de referencia de modo que un círculo más cercano a la dirección de referencia tiene un diámetro más grande.
- 65 11. El aparato según cualquiera de las reivindicaciones 1 a 10, en donde el aparato configurado para definir la retícula esférica generada cubriendo la esfera con esferas más pequeñas, está configurado además para definir, para el primer círculo, la esfera más pequeña que tiene un primer diámetro y para el círculo adicional las esferas más pequeñas que tienen un segundo diámetro.

12. Un aparato configurado para:
 - determinar, al menos un índice de dirección asociado con dos o más señales de audio para proporcionar reproducción de audio espacial, representando el al menos un índice de dirección un parámetro espacial con una elevación y un componente de acimut;
 - definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas, en donde las esferas más pequeñas son cada una más pequeñas que la esfera, en donde las esferas más pequeñas están dispuestas en círculos de esferas, en donde un primer círculo de esferas comprende una de las esferas más pequeñas ubicadas con un centro a una elevación de 90 grados con respecto a una dirección de referencia de la esfera; y
 - convertir el al menos un índice de dirección a una elevación cuantificada y una representación de acimut cuantificada de la elevación y el componente de acimut del parámetro de dirección con un valor de índice basado en la retícula esférica definida.
13. El aparato según la reivindicación 12, en donde el aparato configurado para definir la retícula esférica generada cubriendo la esfera con esferas más pequeñas, está configurado además para seleccionar un primer número determinado de las esferas más pequeñas para un círculo adicional de las esferas, el círculo adicional está definido por un diámetro de la una de las esferas más pequeñas ubicadas con el centro a una elevación de 90 grados con respecto a la dirección de referencia de la esfera.
14. El aparato según la reivindicación 13, en donde el círculo adicional está ubicado paralelo a un ecuador de la esfera.
15. El aparato según cualquiera de las reivindicaciones 13 y 14, en donde el aparato configurado para definir la red esférica generada cubriendo la esfera con las esferas más pequeñas está configurado además para definir un orden de índice de círculo asociado con el primer círculo y el círculo adicional.
16. El aparato según cualquiera de las reivindicaciones 13 a 15, en donde el aparato configurado para definir la retícula esférica generada cubriendo la esfera con las esferas más pequeñas, está configurado además para separar las esferas más pequeñas sobre la esfera aproximadamente equidistantemente entre sí.
17. El aparato según cualquiera de las reivindicaciones 13 a 16, en donde el aparato configurado para definir una retícula esférica generada cubriendo una esfera con esferas más pequeñas está configurada para definir un número de esferas más pequeñas basándose en un valor de cuantificación de entrada.
18. El aparato según cualquiera de las reivindicaciones 12 a 17, en donde el aparato configurado para convertir el al menos un índice de dirección a la elevación cuantificada y la representación de acimut cuantificada de la elevación y el componente de acimut del parámetro de dirección al valor de índice basándose en la retícula esférica definida se configura además para:
 - determinar un valor de índice de círculo en base al valor de índice;
 - determinar la representación de elevación cuantificada del componente de elevación basándose en el valor de índice de círculo; y
 - generar la representación de acimut cuantificada del componente de acimut basándose en un valor de índice de resto después de eliminar un desplazamiento asociado con el valor de índice de círculo a partir del valor de índice.
19. El aparato según cualquiera de las reivindicaciones 12 a 18, en donde el aparato está configurado además para determinar la al menos una dirección de referencia basándose en al menos uno de:
 - un valor de dirección de referencia recibido; o
 - un análisis basado en las dos o más señales de audio.
20. El aparato según la reivindicación 19, en donde el aparato configurado para determinar la al menos una dirección de referencia en base a un análisis basado en las dos o más señales de audio está configurado para determinar la al menos una dirección de referencia en base a un parámetro direccional asociado con al menos una subbanda con un valor de energía de subbanda más alto.
21. El aparato según cualquiera de las reivindicaciones 12 a 20, en donde el aparato configurado para definir la retícula esférica generada cubriendo la esfera con las esferas más pequeñas, está configurado además para definir los círculos de esferas de modo que los círculos de esferas son coplanares con la dirección de referencia y tienen diámetros definidos en base a una elevación desde la dirección de referencia de modo que un círculo más cercano a la dirección de referencia tiene un diámetro más grande.
22. El aparato según cualquiera de las reivindicaciones 12 a 21, en donde el aparato configurado para definir la retícula esférica generada cubriendo la esfera con las esferas más pequeñas está configurado además para

definir para el primer círculo la esfera más pequeña que tiene un primer diámetro y para el círculo adicional las esferas más pequeñas que tienen un segundo diámetro.

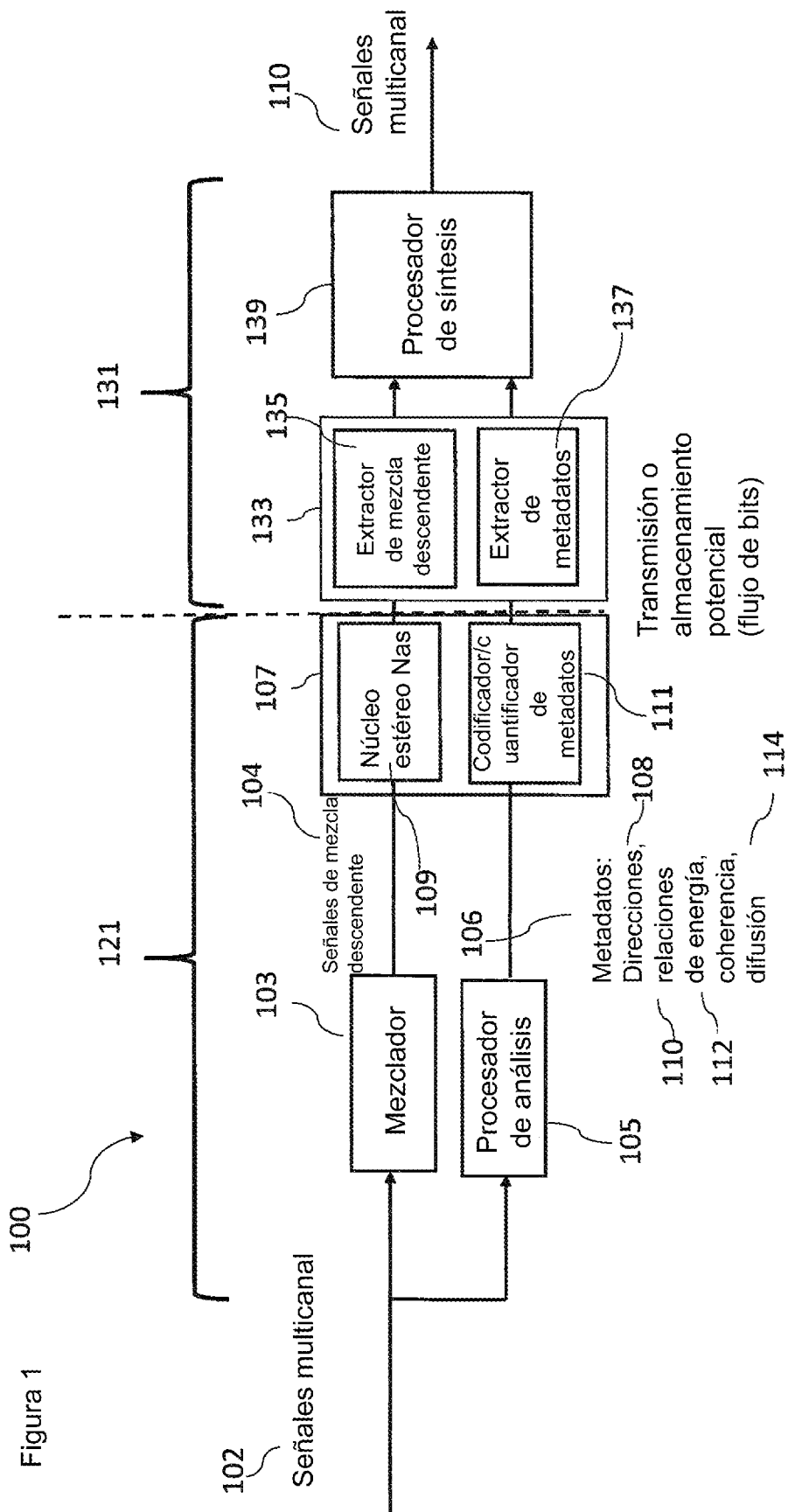
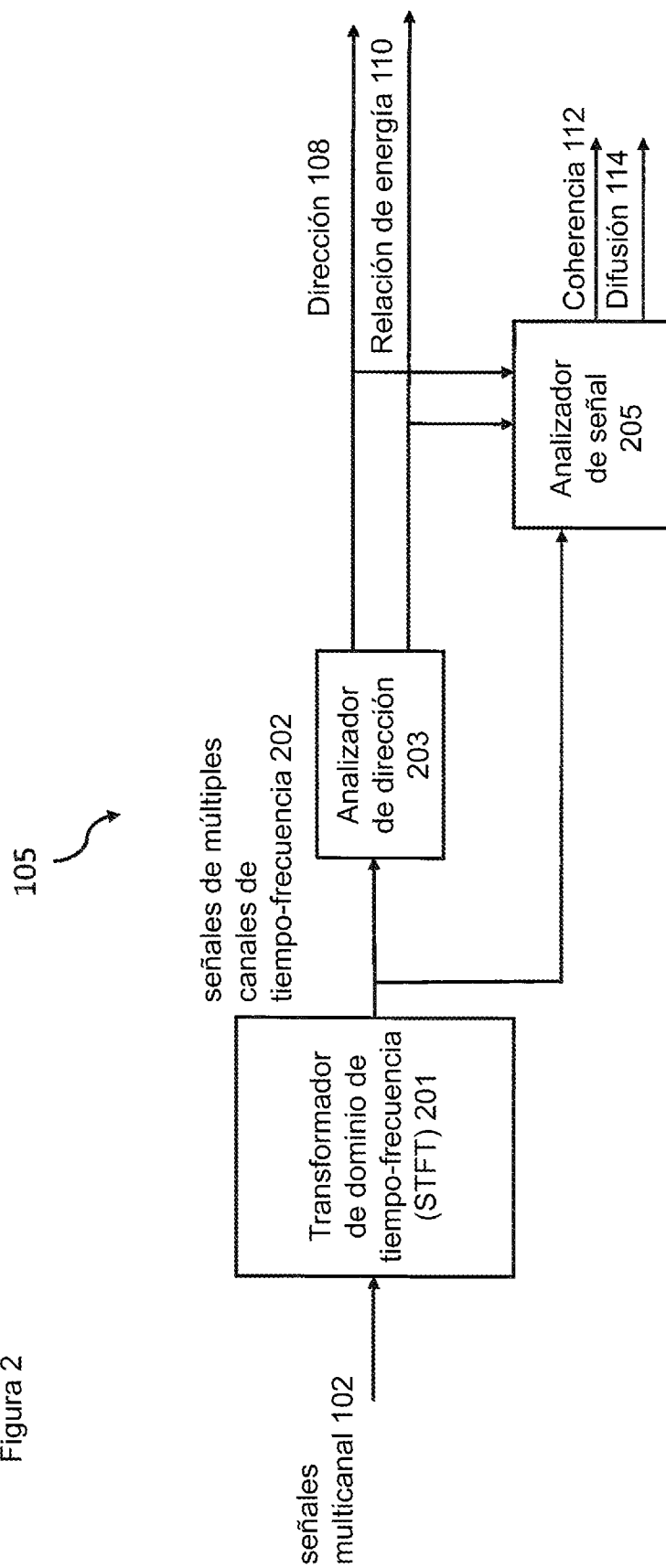


Figura 2



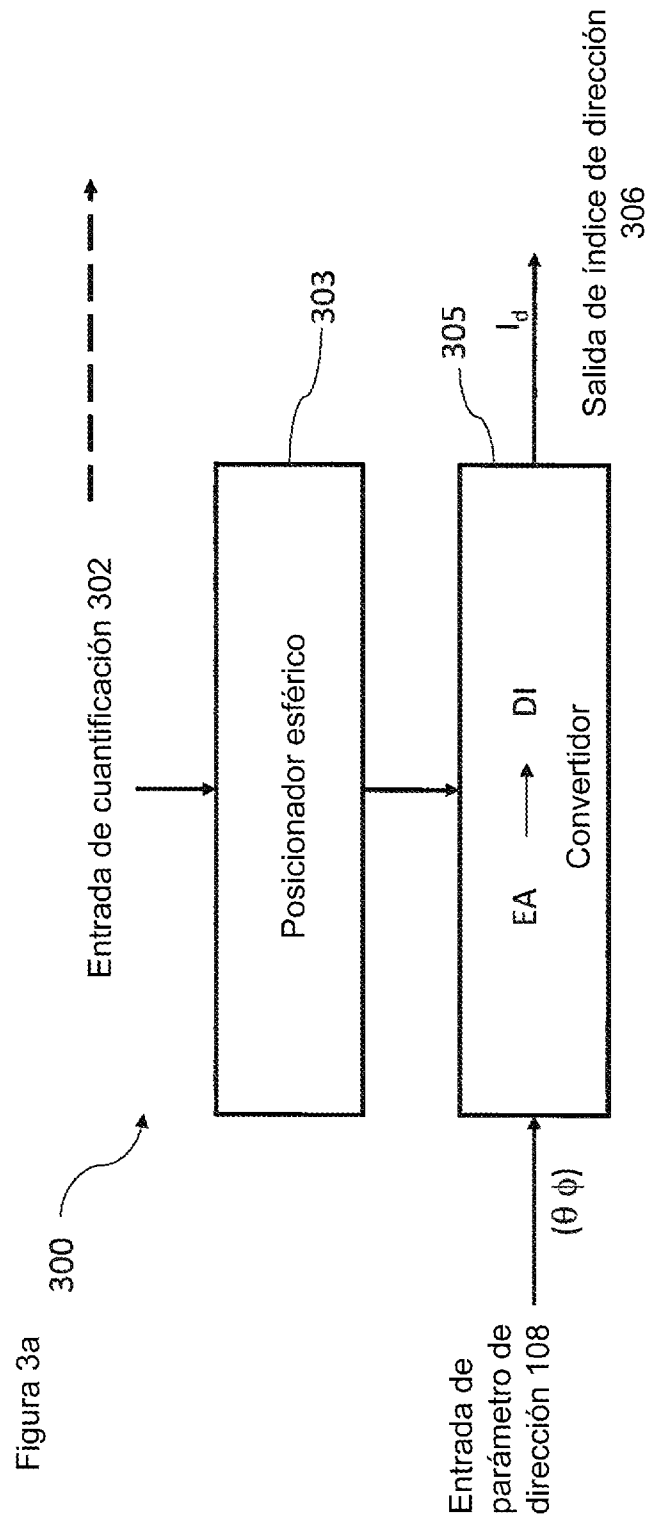
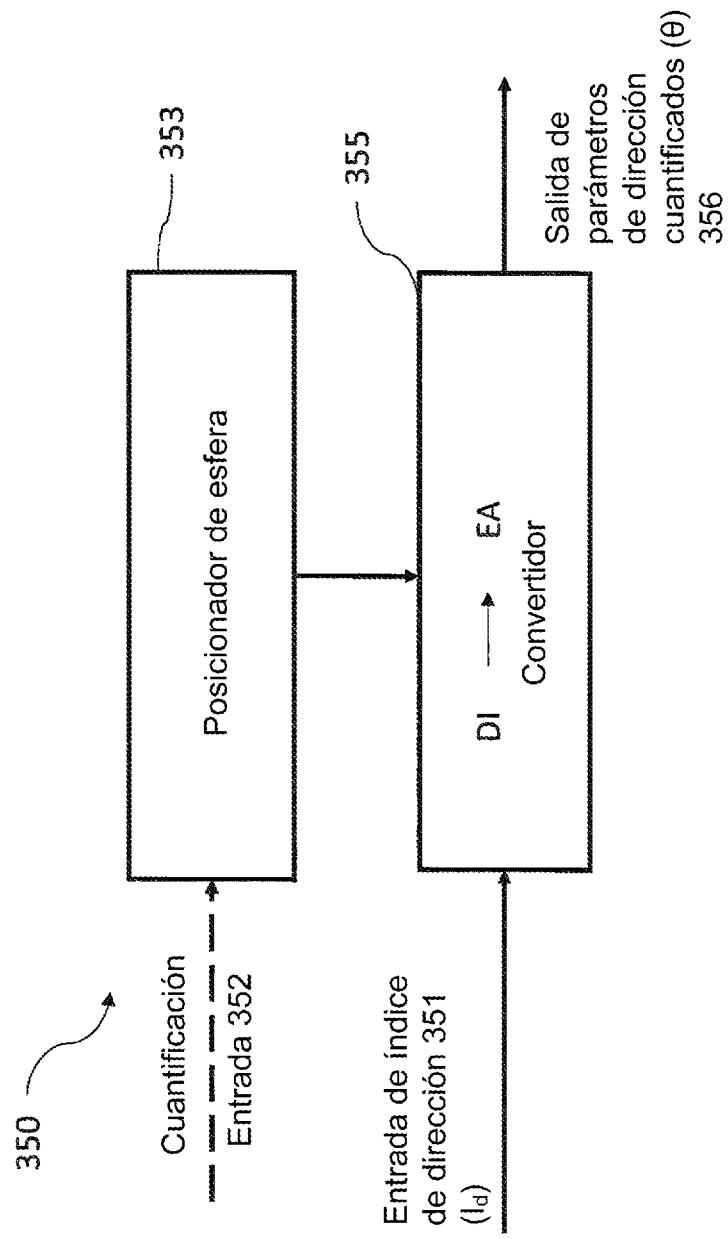


Figura 3b



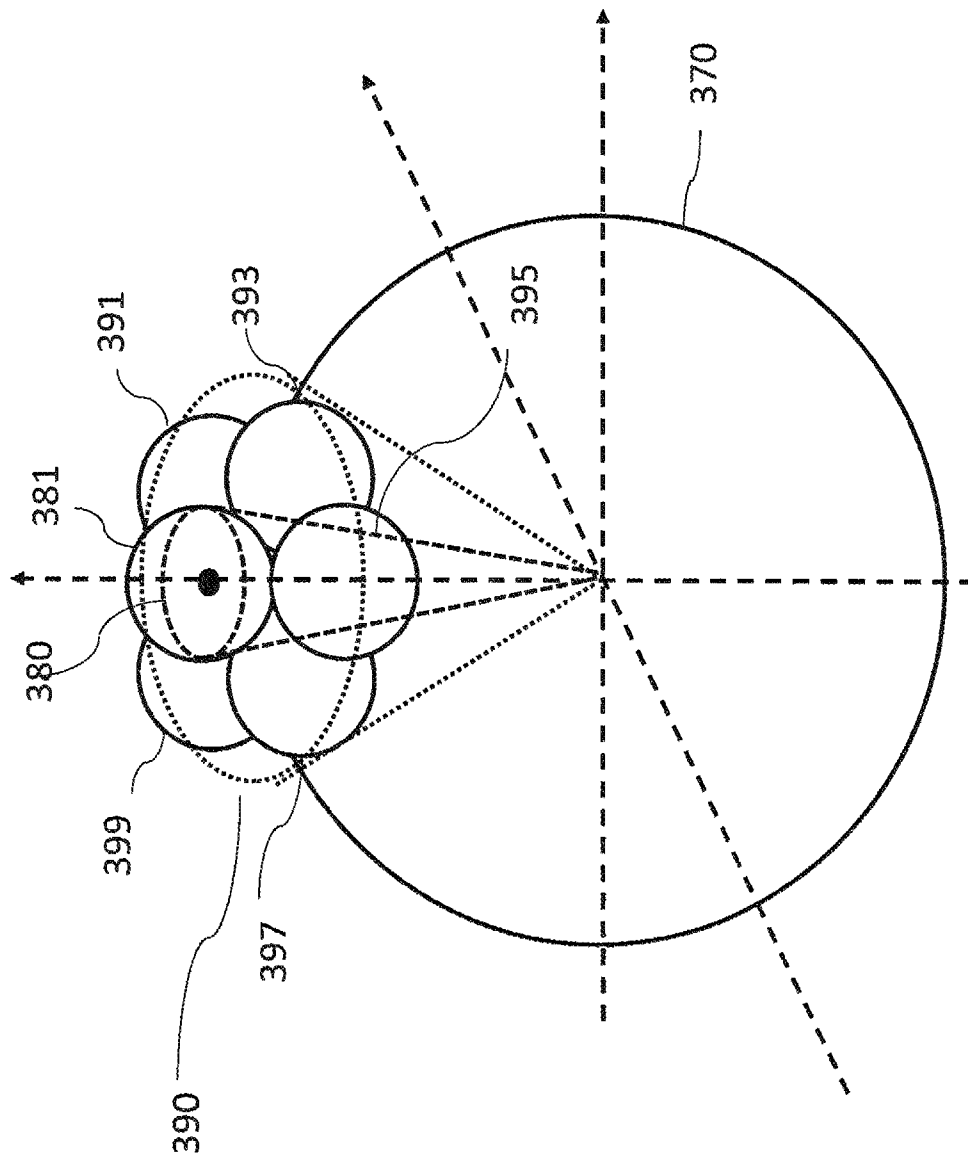


Figure 3c

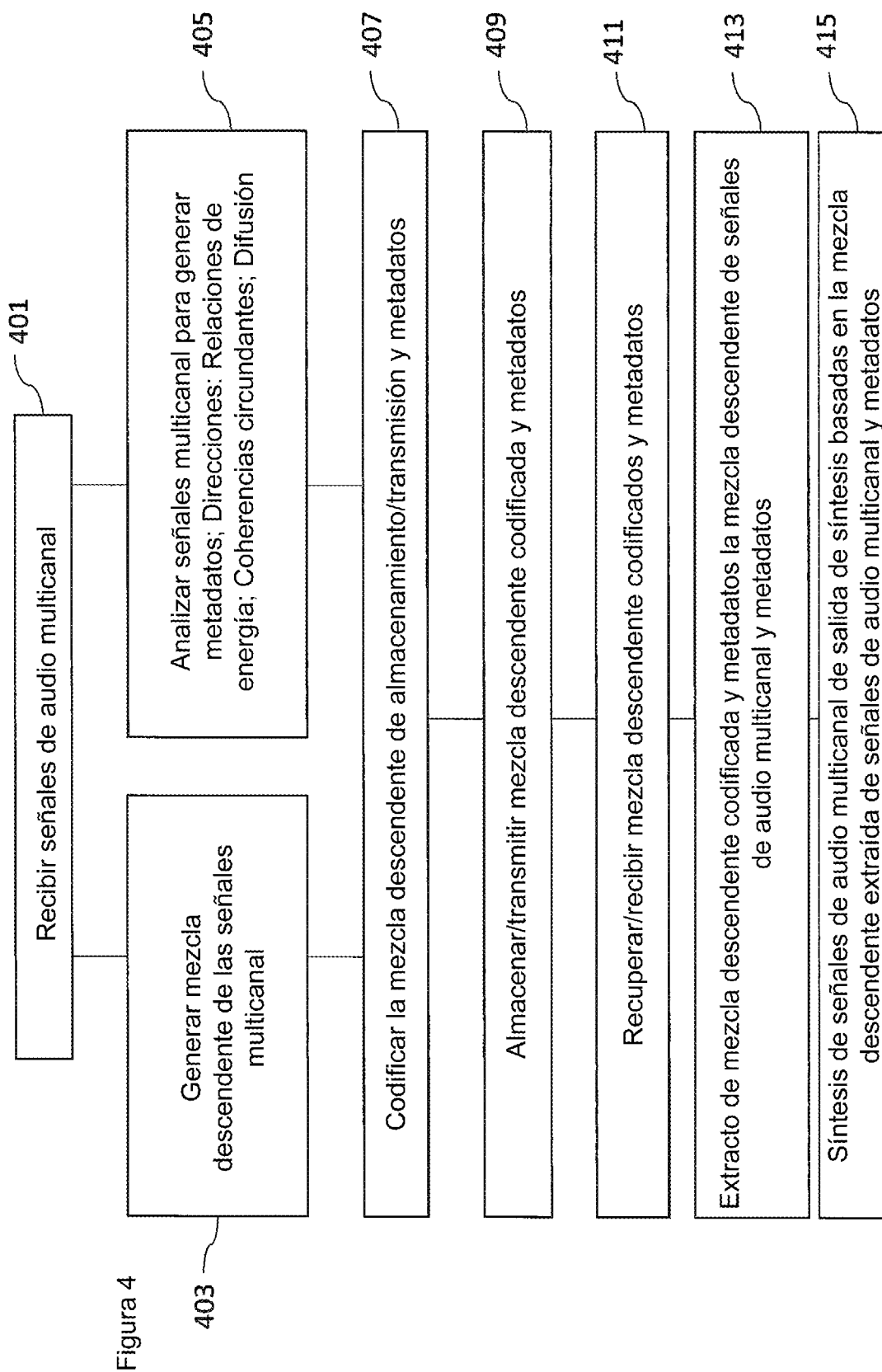
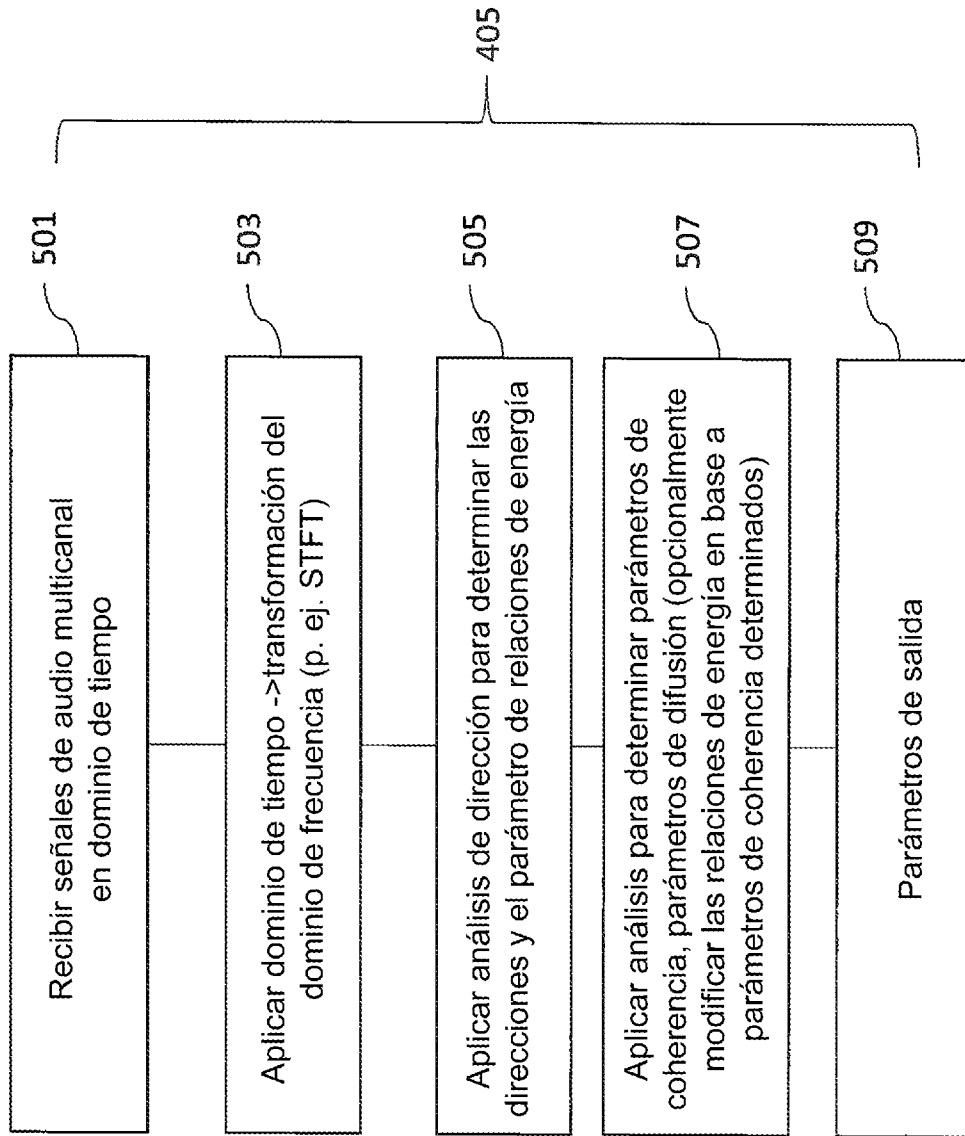


Figura 5



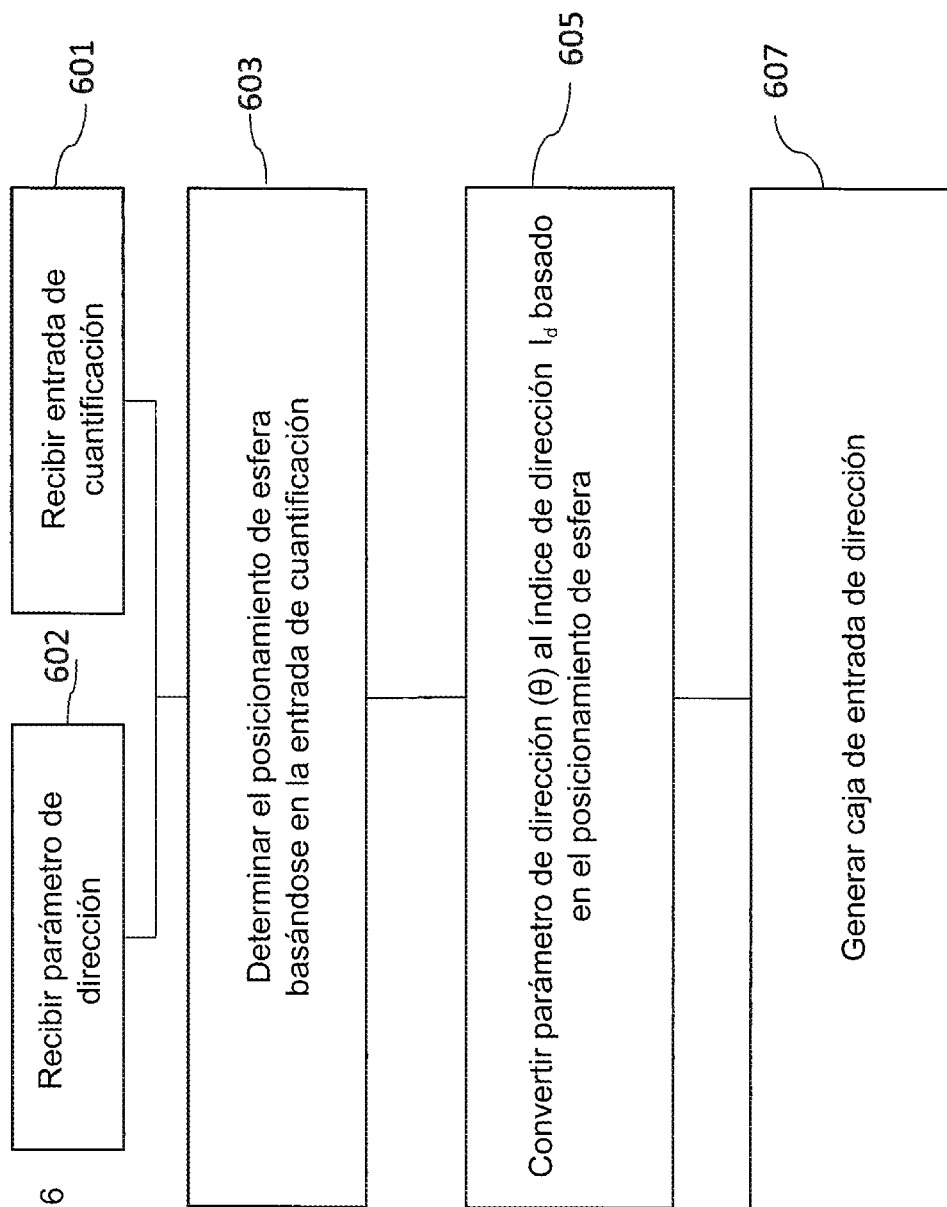
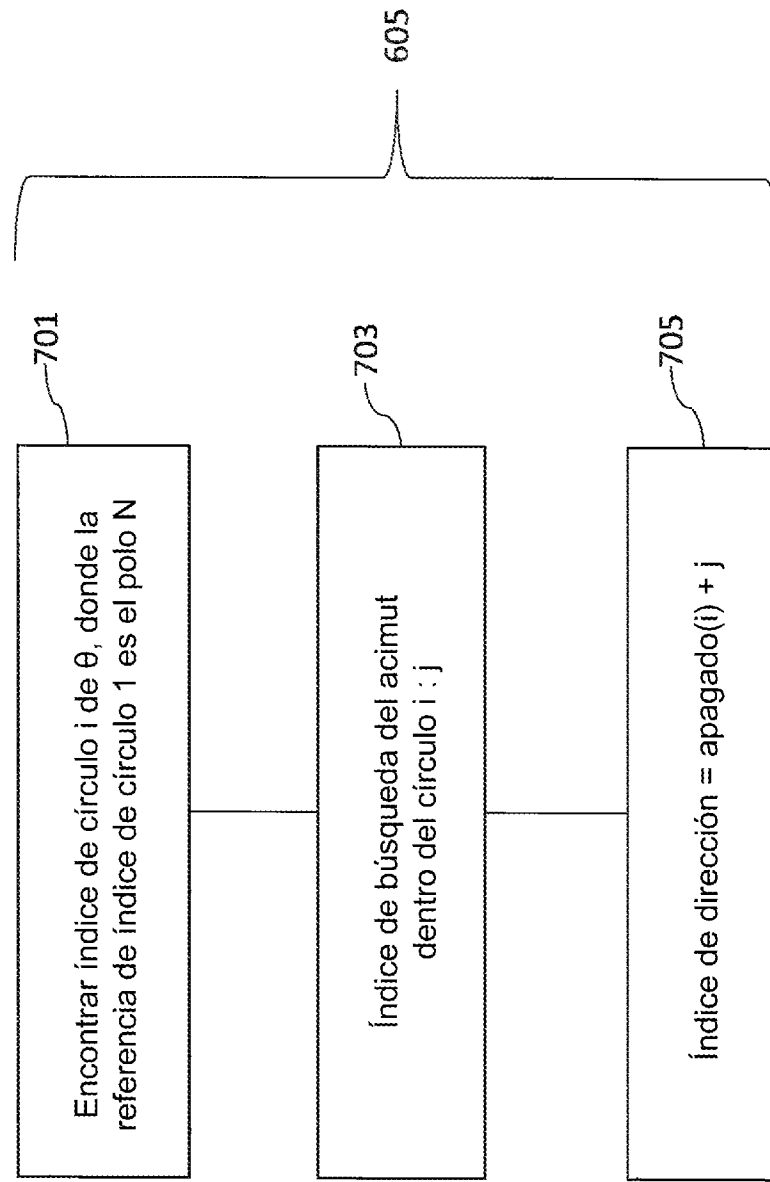
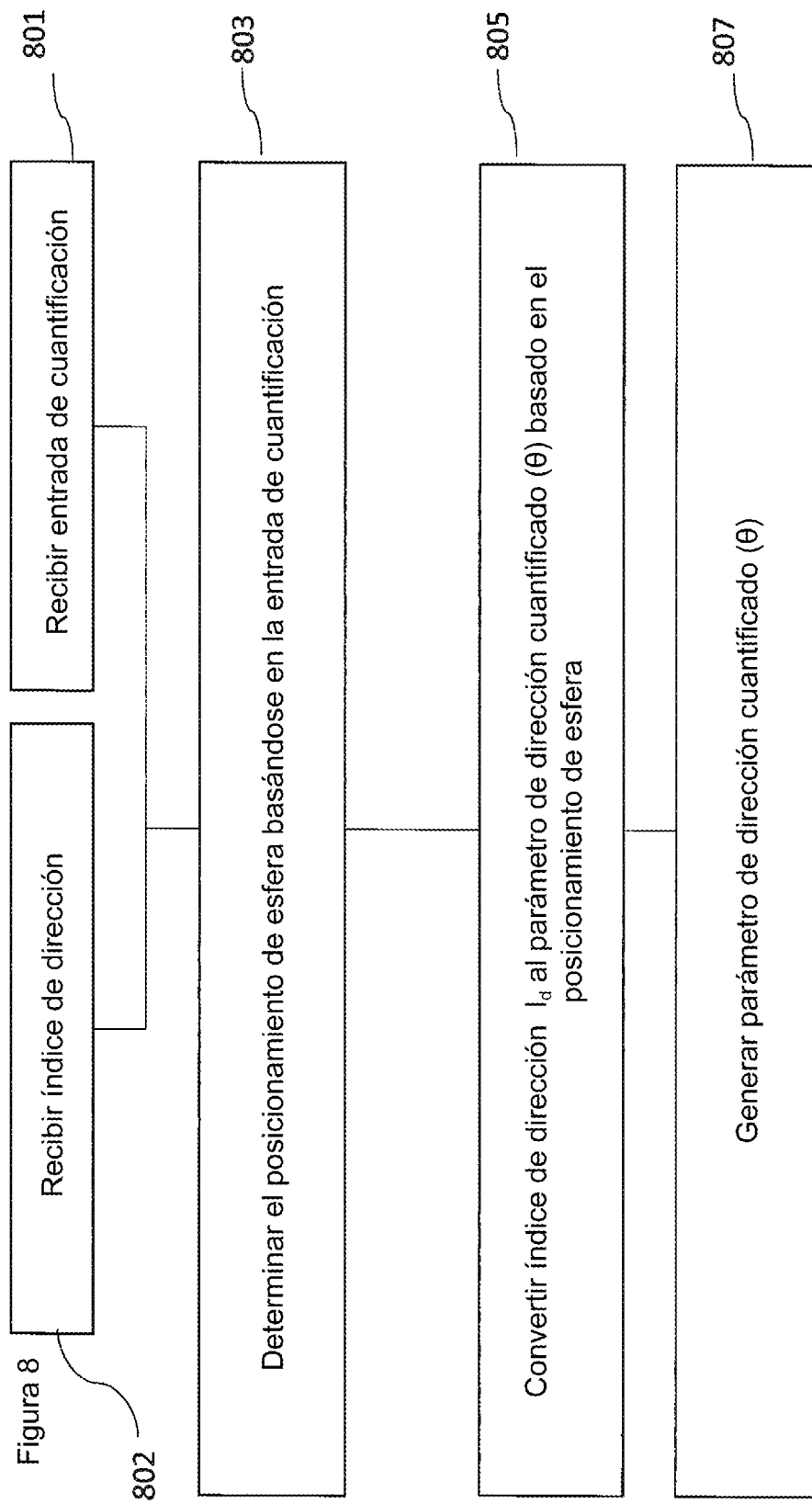
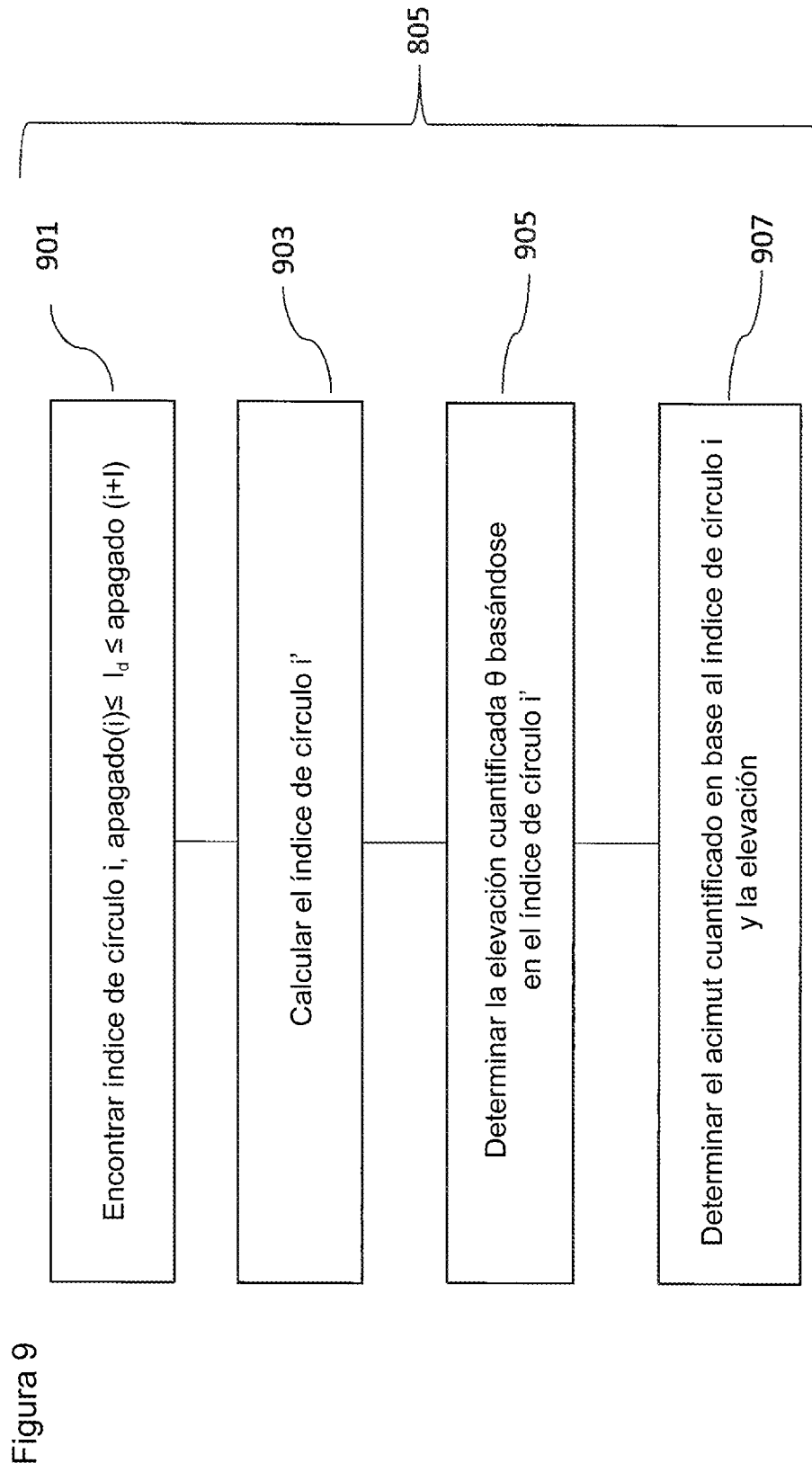


Figura 6

Figura 7







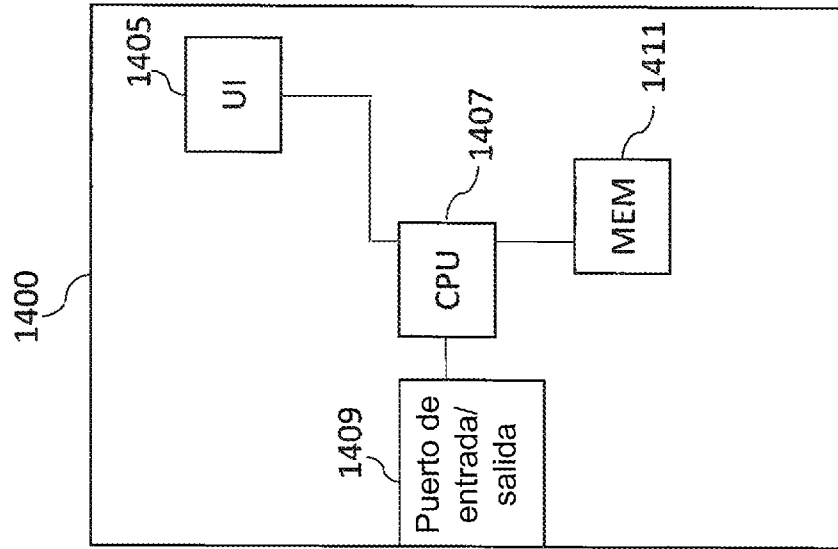


Figura 10