

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 964 178**

51 Int. Cl.:

<b>H04S 5/00</b>	(2006.01)
<b>H04S 7/00</b>	(2006.01)
<b>H04S 3/02</b>	(2006.01)
<b>H04R 3/12</b>	(2006.01)
<b>G06F 3/16</b>	(2006.01)
<b>G10L 19/008</b>	(2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **12.03.2019 PCT/FI2019/050206**
- 87 Fecha y número de publicación internacional: **03.10.2019 WO19185979**
- 96 Fecha de presentación y número de la solicitud europea: **12.03.2019 E 19776342 (8)**
- 97 Fecha y número de publicación de la concesión europea: **11.10.2023 EP 3777242**

54 Título: **Renderización de sonido espacial**

30 Prioridad:

**29.03.2018 GB 201805215**

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

**04.04.2024**

73 Titular/es:

**NOKIA TECHNOLOGIES OY (100.0%)  
Karakaari 7  
02610 Espoo, FI**

72 Inventor/es:

**ERONEN, ANTTI JOHANNES;  
LAITINEN, MIKKO-VILLE;  
VILKAMO, JUHA TAPIO y  
PIHLAJAKUJA, TAPANI**

74 Agente/Representante:

**DEL VALLE VALIENTE, Sonia**

**ES 2 964 178 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

Renderización de sonido espacial

5 **Campo**

La presente solicitud se refiere a aparatos y a métodos para la renderización de sonido espacial. Esto incluye, aunque no exclusivamente, la renderización de sonido espacial para configuraciones de altavoces multicanal.

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 parámetros de razón que expresan energías relativas de las partes direccionales y no direccionales del sonido captado en 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.

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 proporción de energía sonora que es direccional) 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 de codificación de audio avanzado (AAC). Un decodificador puede decodificar las señales de audio para dar señales de modulación por pulsos codificados (PCM), y procesar el sonido en bandas de frecuencia (usando los metadatos espaciales) para obtener la salida espacial, por ejemplo, una salida binaural.

35 Los formatos de entrada del codificador paramétrico pueden ser uno o varios formatos de entrada.

Un formato de entrada de ejemplo es un formato ambisónico de primer orden (FOA). El análisis de la entrada con FOA para una extracción de metadatos espaciales se documenta en la bibliografía científica relacionada con la codificación de audio direccional (DirAC), como el documento "Application of directional audio coding in audio", 19° International Congress on Acoustics, Madrid 2-7 de septiembre de 2007, y Harmonic planewave expansion (Harpex). Esto se debe a que existen matrices de micrófonos especializadas para proporcionar directamente una señal con FOA (o específicamente una variante, la señal de formato B) y se ha implementado un análisis de una entrada de este tipo.

45 La publicación de solicitud de patente de Estados Unidos 2014/226823 describe un método para generar un flujo de bits representativo del contenido de audio multicanal especificando la renderización de audio que incluye una matriz para renderizar coeficientes armónicos esféricos a un número de alimentaciones de altavoces para generar un contenido de audio multicanal.

50 **Sumario**

Se proporciona un aparato para una decodificación de señales de audio espaciales asociado con una pluralidad de nodos de altavoz colocados dentro de un espacio tridimensional, el aparato configurado para recibir: al menos una señal de audio asociada, la señal de audio asociada en función de una señal de audio de disposición de altavoz definida; metadatos espaciales asociados con la señal de audio asociada; al menos un parámetro que representa una disposición de altavoz definida asociada con la señal de audio de disposición de altavoz definida; y al menos un parámetro que representa una disposición de altavoz de salida; sintetizar, a partir de la al menos una señal de audio asociada, al menos una señal de audio de salida en función de los metadatos espaciales y del al menos un parámetro que representa la disposición de altavoz definida y del al menos un parámetro que representa una disposición de altavoz de salida, en donde el aparato, cuando está configurado para sintetizar, a partir de la al menos una señal de audio asociada, al menos una señal de audio de salida en función de los metadatos espaciales y del al menos un parámetro que representa la disposición de altavoz definida y del al menos un parámetro que representa una disposición de altavoz de salida, está configurado además para: dividir la al menos una señal de audio asociada en una parte directa y una parte difusa en función de los metadatos espaciales; sintetizar una señal de audio directa en función de la parte directa de la al menos una señal de audio asociada; sintetizar una señal de audio difusa en función de la parte difusa de la al menos una señal de audio asociada; y combinar la señal de audio directa y la señal de audio difusa para generar la al menos una señal de audio de salida, en donde al menos una de la síntesis de la señal de

audio directa y/o de la síntesis de la señal de audio difusa se basa en el al menos un parámetro que representa la disposición de altavoz definida y en el al menos un parámetro que representa una disposición de altavoz de salida.

5 El aparato configurado para sintetizar una señal de audio difusa en función de la parte difusa de la al menos una señal de audio asociada puede estar configurado para: decorrelacionar la al menos una señal de audio asociada; generar un conjunto de ganancias para cada disposición de altavoz de salida, las ganancias en función de la coincidencia de una distribución de energía de parte difusa total con la disposición de altavoz de entrada y la disposición de altavoz de salida.

10 El aparato configurado para generar un conjunto de ganancias para cada una de la disposición de altavoz de salida, las ganancias en función de la coincidencia de una distribución de energía de parte difusa total con la disposición de altavoz de entrada y la disposición de altavoz de salida puede estar configurado para: determinar las ubicaciones de disposición de altavoz de salida; determinar sumas absolutas para componentes de eje x de nodos de altavoz en función de las ubicaciones de disposición de altavoz de salida; determinar componentes de eje x normalizados de nodos de altavoz; determinar una corrección de componente de eje x en función de los componentes x normalizados; determinar factores de energía corregidos con el componente de eje x en función de la corrección de componente x; y determinar el conjunto de ganancias para cada uno de los nodos de altavoz de salida.

20 El aparato configurado para generar un conjunto de ganancias para cada una de la disposición de altavoz de salida, las ganancias en función de la coincidencia de una distribución de energía de parte difusa total con la disposición de altavoz de entrada y la disposición de altavoz de salida puede estar configurado para: determinar vectores de disposición de altavoz de salida en función de la disposición de altavoz de salida; determinar vectores de disposición de altavoz de entrada en función de la disposición de altavoz de entrada; establecer un factor de energía inicial en función de la inversa de un número de nodos de altavoz de salida; formar un vector diana en función de un vector de suma ponderado por factor de energía de los vectores de disposición de altavoz de entrada; optimizar los factores de energía de manera que un vector de suma ponderado por factor de energía de los vectores de disposición de altavoz de salida coincide con el vector diana; eliminar cualquier valor negativo de los factores de energía optimizados; normalizar la suma de los factores de energía optimizados restantes a la unidad; y determinar el conjunto de ganancias a partir de los factores de energía optimizados restantes normalizados.

30 El aparato configurado para sintetizar una señal de audio difusa en función de la parte difusa de la al menos una señal de audio asociada puede estar configurado para: determinar y desactivar nodos de altavoz a partir de la disposición de altavoz de salida, de modo que los nodos de altavoz desactivados están ubicados entre otros dos nodos de altavoz de la disposición de altavoz de salida y que la secuencia de tres nodos de altavoz de salida adyacentes se encuentra inclusivamente entre dos nodos de altavoz de entrada adyacentes.

40 El aparato configurado para sintetizar una señal de audio directa en función de la parte directa de la al menos una señal de audio asociada puede estar configurado para: determinar y desactivar nodos de altavoz a partir de la disposición de altavoz de salida, de modo que los nodos de altavoz desactivados están ubicados entre otros dos nodos de altavoz de la disposición de altavoz de salida y que la secuencia de tres nodos de altavoz de salida adyacentes se encuentra inclusivamente entre dos nodos de altavoz de entrada adyacentes.

45 El aparato configurado para determinar y desactivar nodos de altavoz a partir de la disposición de altavoz de salida, de modo que los nodos de altavoz desactivados están ubicados entre otros dos nodos de altavoz de la disposición de altavoz de salida y que la secuencia de tres nodos de altavoz de salida adyacentes se encuentra entre dos nodos de altavoz de entrada adyacentes puede estar configurado para: determinar si hay más altavoces en la disposición de altavoz de salida que en la disposición de altavoz de entrada; clasificar la disposición de altavoz de salida en función de un ángulo de acimut para generar una lista clasificada; agregar un primer y un último altavoz de salida según la clasificación al final de la lista clasificada y al inicio de la lista clasificada respectivamente; iterar a través de la lista clasificada adjunta en bloques de tres altavoces adyacentes con un tamaño de salto de uno; mantener un bloque de tres altavoces de salida adyacentes donde al menos un altavoz de entrada está ubicado entre los dos altavoces externos del bloque; retirar un medio de un bloque de tres altavoces de salida adyacentes donde no hay altavoces de entrada entre los dos altavoces externos del bloque.

55 El aparato configurado para clasificar la disposición de altavoz de salida en función de un ángulo de acimut para generar una lista clasificada puede configurarse para clasificar solamente altavoces de salida de orientación horizontal.

60 El aparato configurado para sintetizar una señal de audio directa en función de la parte directa de la al menos una señal de audio asociada puede estar configurado para: crear una distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada en función de la disposición de altavoz de entrada; crear una distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de salida en función de la disposición de altavoz de salida; examinar cualquier superficie virtual en función de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de salida que coincide con al menos una superficie virtual de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada y retirar todas las superficies virtuales de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de salida que implican un

5 altavoz en el interior de la superficie virtual coincidente; examinar cualquier superficie virtual en función de la disposición de salida y retirar una superficie virtual si está completamente en el interior de una superficie virtual de disposición de entrada; retirar cualquier altavoz donde la superficie virtual de disposición de salida tiene altavoces adicionales en el interior para generar una distribución de superficie virtual de altavoz reducida; generar una panoramización de amplitud de base vectorial en función de las superficies virtuales de base vectorial de disposición de altavoz de salida mantenidas.

10 El aparato configurado para sintetizar una señal de audio directa en función de la parte directa de la al menos una señal de audio asociada puede estar configurado para: crear una distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada en función de la disposición de altavoz de entrada; examinar cualquier nodo de altavoz de disposición de altavoz de salida que coincida con al menos un nodo dentro de la superficie virtual de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada y retirar todos los nodos de altavoz de disposición de altavoz de salida que no forman parte de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada pero están ubicados en el interior de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada; crear una distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de salida en función de los nodos de altavoz de disposición de altavoz de salida restantes; examinar cualquier superficie virtual en función de la disposición de salida y retirar una superficie virtual si está completamente en el interior de una superficie virtual de disposición de entrada; retirar cualquier altavoz donde la superficie virtual de disposición de salida tiene altavoces adicionales en el interior para generar una distribución de superficie virtual de altavoz reducida; generar una panoramización de amplitud de base vectorial en función de las superficies virtuales de base vectorial de disposición de altavoz de salida mantenidas.

25 El aparato configurado para sintetizar una señal de audio difusa en función de la parte difusa de la al menos una señal de audio asociada puede ser producido para: crear una distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada en función de la disposición de altavoz de entrada; examinar cualquier nodo de altavoz de disposición de altavoz de salida que coincida con al menos un nodo dentro de la superficie virtual de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada y retirar todos los nodos de altavoz de disposición de altavoz de salida que no forman parte de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada pero están ubicados en el interior de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada.

35 Según un segundo aspecto, se proporciona un método para una decodificación de señal de audio espacial asociado con una pluralidad de nodos de altavoz colocados dentro de un espacio tridimensional, comprendiendo el método: recibir al menos una señal de audio asociada, la señal de audio asociada en función de una señal de audio de disposición de altavoz definida; metadatos espaciales asociados con la señal de audio asociada; al menos un parámetro que representa una disposición de altavoz definida asociada con la señal de audio de disposición de altavoz definida; y al menos un parámetro que representa una disposición de altavoz de salida; y sintetizar, a partir de la al menos una señal de audio asociada, al menos una señal de audio de salida en función de los metadatos espaciales y del al menos un parámetro que representa la disposición de altavoz definida y del al menos un parámetro que representa una disposición de altavoz de salida, en donde la etapa de sintetizar, a partir de la al menos una señal de audio asociada, al menos una señal de audio de salida en función de los metadatos espaciales y del al menos un parámetro que representa la disposición de altavoz definida y del al menos un parámetro que representa una disposición de altavoz de salida comprende además dividir la al menos una señal de audio asociada en una parte directa y una parte difusa en función de los metadatos espaciales; sintetizar una señal de audio directa en función de la parte directa de la al menos una señal de audio asociada; sintetizar una señal de audio difusa en función de la parte difusa de la al menos una señal de audio asociada; y combinar la señal de audio directa y la señal de audio difusa para generar la al menos una señal de audio de salida, en donde al menos una de la síntesis de la señal de audio directa y/o de la síntesis de la señal de audio difusa se basa en el al menos un parámetro que representa la disposición de altavoz definida y en el al menos un parámetro que representa una disposición de altavoz de salida.

55 También es posible un medio legible por ordenador no transitorio que comprende instrucciones de programa para hacer que un aparato realice el método descrito anteriormente.

También es posible un aparato, configurado para realizar las acciones del método descrito anteriormente.

60 También es posible un programa informático que comprende instrucciones de programa para hacer que un ordenador realice el método descrito anteriormente.

También es posible un producto de programa informático almacenado en un medio para hacer que un aparato realice el método descrito en la presente memoria.

65 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.

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

**5 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:

10 la Figura 1 muestra esquemáticamente un ejemplo de panoramización de amplitud de base vectorial que comprende un triplete de altavoces y un triángulo activo;

la Figura 2 muestra esquemáticamente una triangulación de panoramización de amplitud de base vectorial;

15 la Figura 3a muestra esquemáticamente una captura y un sintetizador espaciales ilustrativos según algunas realizaciones;

la Figura 3b muestra un diagrama de flujo del método para operar una captura y un sintetizador espaciales ilustrativos según algunas realizaciones;

20 la Figura 4a muestra esquemáticamente un sintetizador espacial ilustrativo según algunas realizaciones;

la Figura 4b muestra un diagrama de flujo de un método ilustrativo para operar el sintetizador espacial ilustrativo según algunas realizaciones;

25 la Figura 5 muestra un diagrama de flujo para determinar el método para la corriente difusa según algunas realizaciones;

la Figura 6 muestra un vector de suma ilustrativo para una configuración de altavoces 5.0;

30 la Figura 7 muestra un diagrama de flujo de un primer método para la corriente difusa según algunas realizaciones;

la Figura 8 muestra una notación angular de parlante ilustrativa usada en algunas realizaciones;

35 la Figura 9 muestra un diagrama de flujo de un segundo método para la corriente difusa según algunas realizaciones;

la Figura 10 muestra un diagrama de flujo de un método para modificar la reproducción de corriente directa usando el parámetro de disposición de parlante;

40 la Figura 11 muestra un diagrama de flujo de un método para modificar la reproducción de corriente directa usando el parámetro de disposición de parlante para configuraciones de parlantes en 3D; y

la Figura 12 muestra esquemáticamente un dispositivo ilustrativo adecuado para implementar el aparato mostrado.

**45 Realizaciones de la solicitud**

A continuación se describe con más detalle un aparato adecuado y mecanismos posibles para la provisión de un procesamiento espacial eficiente en función de una serie de formatos de entrada de audio.

50 Los metadatos espaciales que consisten en direcciones y parámetros de relación de energía directa a total (o relación de difusividad) en bandas de frecuencia son particularmente adecuados para expresar las propiedades perceptuales de campos sonoros naturales.

55 Sin embargo, las escenas sonoras, tales como las mezclas de parlantes 5.1, pueden originarse a partir de diversas disposiciones de parlantes diferentes tales como estéreo, 4.0, 5.1, 7.1, 22.1 y similares. Cuando tal escena sonora sintética se codifica con los metadatos espaciales mencionados anteriormente, puede producirse una pérdida de información. El objetivo de las realizaciones como se describe con mayor detalle es que el campo sonoro reproducido intente corresponderse plenamente con lo que un oyente percibiría si estuviera escuchando la mezcla envolvente 5.1 original u otra mezcla envolvente original.

60 En particular, el concepto como se analiza en la siguiente realización intenta reproducir con precisión la distribución espacial de la energía sonora difusa/ambiental en el sonido reproducido en comparación con el sonido espacial original. De manera similar, las realizaciones pueden tener como objetivo reproducir con precisión el efecto espacial percibido de las fuentes cuando se reproduce sobre configuraciones de parlantes distintas de la configuración de entrada.

65

Las realizaciones pueden configurarse adicionalmente para controlar, durante la renderización, si la intención es optimizar la calidad general de reproducción o si la intención es reproducir la experiencia lo más cerca posible de la mezcla envolvente original.

5 Puede tratarse, por ejemplo, de requisitos diferentes. Por ejemplo, por un lado, la representación de audio espacial paramétrica pretende abstraer el efecto de la disposición original de los canales y producir una representación paramétrica genérica que pueda renderizarse a cualquier configuración de parlantes. Por otro lado, puede que se desee preservar la intención y las cualidades artísticas originales en la representación basada en canales.

10 El concepto analizado con más detalle a continuación en el presente documento es añadir un campo de metadatos de disposición de canal en el flujo de bits implementado por el códec y emplear un método para utilizar estos metadatos durante la síntesis y la renderización para permitir la reproducción del audio espacial de modo que represente más estrechamente la mezcla envolvente original.

15 Como tales, las realizaciones descritas a continuación se refieren a la codificación y decodificación de audio usando una parametrización relacionada con un campo sonoro (dirección o direcciones y relación o relaciones en bandas de frecuencia) y donde estas realizaciones tienen como objetivo mejorar la calidad de reproducción de mezclas envolventes de parlantes codificadas con la parametrización mencionada anteriormente. Adicionalmente, estas realizaciones describen dónde se mejora la calidad de las mezclas envolventes de los parlantes al transportar un parámetro de disposición de parlante junto con el o los parámetros direccionales, y reproducir el sonido en función del parámetro o parámetros direccionales y el parámetro de disposición de parlante de entrada, de modo que el parámetro de disposición de parlante de entrada afecta al menos a una síntesis de corriente difusa o síntesis de corriente directa usando la o las direcciones y la o las relaciones en bandas de frecuencia.

25 En particular, las realizaciones analizadas a continuación se configuran para modificar la síntesis de corriente difusa usando el parámetro de disposición de parlante de entrada de manera que la distribución de energía de la mezcla envolvente original se reproduce mejor. Las realizaciones describen además la modificación de la síntesis de corriente directa de manera que las fuentes fantasma presentes en la mezcla envolvente original se reproducen con mayor autenticidad.

30 Con respecto a la Figura 3a, se muestra una captura y un sintetizador ilustrativos espaciales según algunas realizaciones. La captura y el sintetizador espaciales se muestran en este ejemplo recibiendo una señal 300 de audio de formato 5.0 como una entrada.

35 Un determinador 311 de parámetro de disposición de canal está configurado para recibir la señal 300 de audio de formato 5.0 y determinar un parámetro 304 de disposición de canal adecuado. El parámetro de disposición de canal o de disposición de altavoz o de disposición de parlante puede ser una enumeración que identifique posibles configuraciones de altavoces conocidas por el sistema. Cada enumeración puede estar asociada a una configuración de parlantes que comprende el número de parlantes, y los ángulos de acimut y de elevación de parlante.

40 Para diseños de altavoces desconocidos, el parámetro puede definir el número de altavoces y las ubicaciones de altavoces como pares de acimut, de elevación, o como índices para las ranuras cuantificadas de acimut y elevación.

45 Un convertidor 301 ambisónico de primer orden (FOA) (o una variante de formato B) está configurado para recibir la señal 300 de audio de formato 5.0 y convertirla en una corriente 302 con FOA (formato B) adecuada. El convertidor 301 con FOA, en algunas realizaciones, está configurado para generar las señales con FOA (por ejemplo, la variante de formato B) de una mezcla de altavoces en función del conocimiento de las posiciones de los canales en las señales de audio de entrada. En otras palabras, los componentes  $w_i(t)$ ,  $x_i(t)$ ,  $y_i(t)$ ,  $z_i(t)$  de una señal con FOA pueden generarse a partir de una señal de parlante  $s_i(t)$  en  $azi_i$  y  $ele_i$  por

$$50 \quad FOA_i(t) = \begin{bmatrix} w_i(t) \\ x_i(t) \\ y_i(t) \\ z_i(t) \end{bmatrix} = s_i(t) \begin{bmatrix} \cos(azi_i) \cos(ele_i) \\ \sin(azi_i) \cos(ele_i) \\ \sin(ele_i) \end{bmatrix}$$

Las señales  $w$ ,  $x$ ,  $y$ ,  $z$  se generan para cada señal de parlante (u objeto)  $s_i$  que tiene su propia dirección de acimut y elevación.

55 La señal de salida que combina todas estas señales puede calcularse como  $\sum_{i=1}^{NUM\_CH} FOA_i(t)$ . En otras palabras, la combinación de cada altavoz o señal de canal con las señales con FOA totales y  $NUM\_CH$  es el número de señales de altavoz o canal.

60 Un analizador 303 espacial está configurado para recibir la corriente 302 con FOA y generar parámetros espaciales adecuados tales como direcciones 306 y relaciones 308. El analizador 303 espacial puede ser, por ejemplo, un ordenador o un teléfono móvil (que ejecuta software adecuado), o alternativamente un dispositivo específico que

utiliza, por ejemplo, matrices de puertas programables en campo (FPGA) o circuitos integrados específicos de aplicación (ASIC).

5 Un ejemplo de método de análisis espacial adecuado es la codificación de audio direccional (DirAC). Los métodos de DirAC pueden estimar las direcciones y las relaciones de difusividad (información equivalente a un parámetro de relación directa a total) de una señal ambisónica de primer orden (FOA), o su variante, la señal de formato B.

10 En algunas realizaciones, el método de DirAC transforma las señales de  $\sum_{i=1}^{NUM.CH} FOA_i(t)$  en bandas de frecuencia usando una transformada en el dominio del tiempo a una en el dominio de la frecuencia adecuada, por ejemplo, usando una transformada de Fourier de tiempo corto (STFT), lo que da como resultado señales de tiempo-frecuencia  $w(k,n)$ ,  $x(k,n)$ ,  $y(k,n)$ ,  $z(k,n)$  donde  $k$  es el índice de intervalo de frecuencia y  $n$  es el índice de tiempo.

En tales ejemplos, el método de DirAC puede estimar el vector de intensidad mediante

$$15 \quad I(k, n) = -Re \left\{ w^* (k, n) \begin{bmatrix} x(k, n) \\ y(k, n) \\ z(k, n) \end{bmatrix} \right\}$$

donde  $Re$  significa parte real, y asterisco  $*$  significa conjugado complejo. La intensidad expresa la dirección de la energía sonora de propagación y, por lo tanto, el parámetro de dirección puede determinarse por la dirección opuesta del vector de intensidad.

20 El vector de intensidad, en algunas realizaciones, puede promediarse durante varios índices de tiempo y/o frecuencia antes de la determinación del parámetro de dirección.

25 Adicionalmente, en algunas realizaciones, el método de DirAC puede determinar la difusividad basada en componentes con FOA (suponiendo la seminormalización de Schmidt (normalización SN3D)). En la normalización SN3D para el sonido difuso, la suma de las energías de todos los componentes ambisónicos dentro de un orden es igual. Por ejemplo, si el orden cero  $W$  tiene 1 unidad de energía, entonces cada primer orden  $X$   $Y$   $Z$  tiene 1/3 unidades de energía (la suma es 1). Y así sucesivamente para órdenes superiores.

30 Por lo tanto, la difusividad puede determinarse como

$$\Psi(k, n) = 1 - \frac{|E[I(k, n)]|}{E[0,5(w^2(k, n) + x^2(k, n) + y^2(k, n) + z^2(k, n))]}$$

35 La difusividad es un valor de relación que es 1 cuando el sonido es completamente ambiental, y 0 cuando el sonido es completamente direccional. En algunas realizaciones, todos los parámetros en la ecuación se promedian típicamente con el tiempo y/o la frecuencia. El operador de expectativa  $E[ ]$  puede reemplazarse por un operador promedio en algunos sistemas.

40 En algunas realizaciones, el parámetro de dirección y el parámetro de difusividad pueden analizarse a partir de componentes con FOA que se han obtenido de dos formas diferentes. En particular, en esta realización, el parámetro de dirección puede analizarse a partir de las señales  $\sum_{i=1}^{NUM.CH} FOA_i(t)$  como se ha descrito anteriormente. La difusividad puede analizarse a partir de otro conjunto de señales con FOA indicadas como  $\sum_{i=1}^{NUM.CH} \widehat{FOA}_i(t)$ , y descrito con más detalle a continuación. Como un ejemplo particular, considérese la conversión a componentes con FOA a partir de una entrada 5.0 que tiene parlantes en ángulos de acimut 0, +/-30, y +/-110 (todas las elevaciones son cero,  $\cos(e/e_i) = 1$ ,  $\sin(e/e_i) = 0$  para todo  $i$ ). Los componentes con FOA para el análisis del parámetro de dirección se obtienen como anteriormente:

$$FOA_i(t) = \begin{bmatrix} w_i(t) \\ x_i(t) \\ y_i(t) \\ z_i(t) \end{bmatrix} = s_i(t) = \begin{bmatrix} 1 \\ \cos(azi_i) \\ \sin(azi_i) \\ 0 \end{bmatrix}$$

50 La difusividad puede analizarse a partir de otro conjunto de señales con FOA obtenidas como

$$\widehat{FOA}_i(t) = \begin{bmatrix} \widehat{w}_i(t) \\ \widehat{x}_i(t) \\ \widehat{y}_i(t) \\ \widehat{z}_i(t) \end{bmatrix} = s_i(t) = \begin{bmatrix} 1 \\ \cos(\widehat{azi}_i) \\ \sin(\widehat{azi}_i) \\ 0 \end{bmatrix}$$

55 donde  $\widehat{azi}_i$  es una posición de parlante virtual modificada. Las posiciones de parlantes virtuales modificadas para el análisis de difusividad se obtienen de manera que los parlantes virtuales se ubican con una separación uniforme

cuando se crean las señales con FOA. El beneficio de tal posicionamiento uniformemente separado de los altavoces virtuales para el análisis de difusividad es que el sonido incoherente llega de manera uniforme desde diferentes direcciones alrededor del micrófono virtual y el promedio temporal del vector de intensidad suma hasta valores cercanos a cero. En el caso de 5.0, las posiciones de parlantes virtuales modificadas son 0, +/-72, +/-144 grados. Por lo tanto, los parlantes virtuales tienen una separación constante de 72 grados.

Se pueden crear posiciones de parlantes virtuales modificadas similares para otras configuraciones de parlantes para garantizar una separación constante entre altavoces adyacentes. En una realización de la invención, la separación de parlantes virtuales modificada se obtiene dividiendo los 360 grados completos con el número de parlantes en el plano horizontal. Las posiciones de parlantes virtuales modificadas se obtienen entonces colocando los parlantes virtuales con la separación obtenida comenzando desde el altavoz central u otro altavoz de inicio adecuado.

En algunas realizaciones puede determinarse un parámetro de relación alternativo, por ejemplo, una relación de energía directa a total, que puede obtenerse como

$$r(k, n) = 1 - \Psi(k, n)$$

Cuando se promedian, los parámetros de difusividad (y dirección) pueden determinarse en bandas de frecuencia que combinan varios intervalos de frecuencia  $k$ , por ejemplo, aproximando la resolución de frecuencia de Bark.

DirAC, como se ha descrito anteriormente, es una posible opción de método de análisis espacial para determinar los metadatos direccionales y de relación. Los parámetros de audio espacial también llamados metadatos espaciales o metadatos pueden determinarse según cualquier método adecuado. Por ejemplo, simulando una matriz de micrófonos y usando un algoritmo de captura de audio espacial (SPAC). Adicionalmente, los metadatos espaciales pueden incluir (pero no se limitan a): Dirección y relación de energía directa a total; Dirección y difusividad; Diferencia de nivel entre canales, diferencia de fase entre canales y coherencia entre canales. En algunas realizaciones, estos parámetros se determinan en el dominio de tiempo-frecuencia. Cabe señalar que también pueden usarse otras parametrizaciones que las presentadas anteriormente. En general, típicamente las parametrizaciones de audio espacial describen cómo se distribuye el sonido en el espacio, ya sea generalmente (por ejemplo, usando direcciones) o relativamente (por ejemplo, como diferencias de nivel entre ciertos canales).

Un generador 305 de señales de transporte se configura además para recibir la corriente 302 con FOA/ formato B y generar señales 310 de audio de transporte adecuadas. Las señales de audio de transporte también se pueden conocer como señales de audio asociadas y se pueden basar en las señales de audio espaciales que contienen información direccional de un campo sonoro y que se introducen en el sistema. El generador 305 de señales de transporte puede configurarse para generar cualquier número adecuado de señales (o canales) de audio de transporte y/o formato de señal. Por ejemplo, en algunas realizaciones, el generador de señales de transporte está configurado para generar dos señales de audio de transporte. En algunas realizaciones, el preprocesador está configurado además para codificar las señales de audio. En algunas realizaciones, las señales de transporte están en cualquier formato adecuado. Por ejemplo, en algunas realizaciones, las señales de transporte son señales con FOA o canales de entrada no procesados (originales). En estas realizaciones, el analizador espacial puede implementarse en el lado del decodificador de la línea discontinua con el sintetizador 309.

Por ejemplo, en algunas realizaciones, las señales de audio pueden codificarse usando una codificación de audio avanzada (AAC) o servicios de voz potenciada (EVS). En algunas realizaciones, el generador 305 de señales de transporte puede configurarse para ecualizar las señales de audio, aplicar control automático de ruido, procesamiento dinámico o cualquier otro procesamiento adecuado.

El parámetro 304 de disposición de canal, metadatos espaciales (las direcciones 306 y relaciones 308) y las señales 310 de audio de transporte pueden transmitirse o almacenarse, por ejemplo, dentro de algún almacenamiento 307, tal como memoria, o alternativamente procesarse directamente en el mismo dispositivo. En algunas realizaciones, el parámetro 304 de disposición de canal, metadatos espaciales 306, 308 y señales 310 de audio de transporte pueden codificarse o cuantificarse o combinarse o multiplexarse en un flujo de datos individual mediante una operación de codificación y/o multiplexación adecuada. En algunas realizaciones, la señal de audio codificada se agrupa con un flujo de vídeo (por ejemplo, vídeo de 360 grados) en un contenedor multimedia tal como un contenedor de Expert Picture Experts Group MPEG-4 Parte 14 o el contenedor mp4, para transmitirse a un receptor adecuado.

El sintetizador 309 está configurado para recibir el parámetro 304 de disposición de canal, señales 310 de audio de transporte, los parámetros espaciales tales como las direcciones 306 y las relaciones 308 y el parámetro de disposición de canal y generar las señales 312 de audio de parlante.

El sintetizador 309, por ejemplo, puede estar configurado para generar las señales de audio de parlante empleando una reproducción de sonido espacial donde el sonido en el espacio 3D se coloca en direcciones arbitrarias. El sintetizador 309 puede ser, por ejemplo, un ordenador o un teléfono móvil (que ejecuta software adecuado), o alternativamente un dispositivo específico que utiliza, por ejemplo, matrices de puertas programables en campo (FPGA) o circuitos integrados específicos de aplicación (ASIC). En función del flujo de datos (las señales de audio de

transporte y los metadatos). El sintetizador 309 puede estar configurado para producir señales de audio de salida, que se muestran en la Figura 3a como señales 312 de altavoz. Las señales de salida pueden ser, en algunas realizaciones, cualquier señal de audio de formato adecuada y pueden ser, por ejemplo, señales binaurales para la escucha de auriculares, señales de parlantes de múltiples canales o señales ambisónicas. En algunas realizaciones, el analizador espacial y el sintetizador (y otros componentes como se describe en la presente descripción) pueden implementarse dentro del mismo dispositivo, y pueden ser también parte del mismo software.

Un método ilustrativo para implementar tal posicionamiento es el uso de una panoramización de amplitud de base vectorial (VBAP).

La VBAP se basa en

- 1) triangular automáticamente la configuración de parlantes,
- 2) seleccionar un triángulo apropiado en función de la dirección (de modo que para una dirección dada se seleccionan tres parlantes que forman un triángulo donde se encuentra la dirección), y
- 3) calcular ganancias para los tres parlantes que forman el triángulo particular.

En algunas realizaciones, las ganancias de VBAP (para cada acimut y elevación) y los tripletes de parlantes u otros números adecuados de parlantes o nodos de altavoz (para cada acimut y elevación) pueden formularse previamente en una tabla de consulta almacenada en la memoria. En algunas realizaciones, un método en tiempo real realiza la panoramización de amplitud hallando a partir de la memoria el triplete (o número) de altavoces apropiado para la dirección de panoramización deseada, y las ganancias para estos parlantes correspondientes a la dirección de panoramización deseada.

La primera etapa de VBAP es la división de la configuración de altavoces en 3D para dar triángulos. No hay una solución única a la generación de la triangulación y la configuración de altavoces puede triangularse de muchas maneras. En algunas realizaciones, se intenta encontrar triángulos o polígonos de tamaño mínimo (sin parlantes dentro de los triángulos y lados que tienen una longitud lo más igual posible). En un caso general, este es un enfoque válido, ya que trata los objetos auditivos en cualquier dirección por igual, e intenta minimizar las distancias hasta los altavoces que están usándose para crear el objeto auditivo en esa dirección.

Otro método computacionalmente rápido para la triangulación o distribución de superficie virtual es generar un casco convexo en función de los puntos de datos determinados por los ángulos de parlante. Este también es un enfoque genérico que trata todas las direcciones y puntos de datos por igual.

En la Figura 2 se muestra una triangulación de ejemplo. La Figura 2 muestra una serie de superficies virtuales en forma de triángulos que tienen esquinas ubicadas en la posición de los nodos 201, 203, 205, 207, 209, 211, 221, 223, 225, 227, 231 de parlantes. Esta triangulación ilustrativa ejemplifica un caso con parlantes solamente en el plano horizontal y elevaciones positivas, y no se realizó ninguna generación de triángulos por debajo de la horizontal. En tal caso, podría ser que cualquier dirección reproducida prevista por debajo del plano horizontal se reproduzca usando los parlantes horizontales.

La siguiente etapa o segunda etapa es formular ganancias de panoramización correspondientes a las direcciones de panoramización.

La Figura 1 muestra, por ejemplo, tres parlantes, canal 1 101 ubicado en la dirección del vector unitario  $\mathbf{l}_1$ , canal 2 102 ubicado en la dirección del vector unitario  $\mathbf{l}_2$  y canal 3 101 ubicado en la dirección del vector unitario  $\mathbf{l}_3$ . Estos vectores se definen con respecto al oyente 100 en un punto de origen y muestran el triángulo activo 105 definido, definido por los tres altavoces. Además, se muestra una fuente 104 virtual ubicada en la dirección del vector unitario  $\mathbf{p}$  con respecto al oyente 100 dentro del triángulo activo 105.

La siguiente etapa es formular ganancias de panoramización correspondientes a las direcciones de panoramización.

La panoramización de amplitud de base vectorial se refiere al método donde los tres vectores unitarios  $\mathbf{l}_1$ ,  $\mathbf{l}_2$ ,  $\mathbf{l}_3$  (la base vectorial) se suponen desde el punto de origen hasta las posiciones de los tres parlantes que forman el triángulo donde se encuentra la dirección de panoramización tal como se muestra en la Figura 1.

Las ganancias de panoramización para los tres altavoces se determinan de manera que estos tres vectores unitarios se ponderan de manera que su vector de suma ponderado apunta hacia la dirección de panoramización de amplitud deseada. Esto puede resolverse de la siguiente manera. Se formula un vector unitario de columna  $\mathbf{p}$  que apunta hacia la dirección de panoramización de amplitud deseada, y un vector  $\mathbf{g}$  que contiene las ganancias de panoramización de amplitud puede resolverse mediante una multiplicación de matriz

$$\mathbf{g}^T = \mathbf{p}^T \begin{bmatrix} l_1^T \\ l_2^T \\ l_3^T \end{bmatrix}^{-1}$$

donde  $^{-1}$  designa la matriz inversa. Después de formular las ganancias  $\mathbf{g}$ , se normaliza su nivel global de modo que para las ganancias finales la suma de energía  $\mathbf{g}^T \mathbf{g} = 1$ .

5 Con respecto a la Figura 3b, se muestra un resumen ilustrativo de las operaciones del aparato mostrado en la Figura 3a.

10 La operación inicial es recibir señales de audio de parlante (por ejemplo, formato 5.0) como se muestra en la Figura 3b, en la etapa 351.

15 La determinación de un parámetro 304 de disposición de canal en función de las señales de audio de disposición de altavoz recibidas (o recibidas junto con las señales de audio de disposición de altavoz recibidas) se muestra en la Figura 3b, en la etapa 353.

Las señales de audio de disposición de altavoz recibidas pueden convertirse en una corriente de formato con FOA (por ejemplo, la variante de formato B) como se muestra en la Figura 3b, en la etapa 355.

20 Las señales de audio con FOA convertidas pueden acto seguido procesarse para generar señales de audio de transporte como se muestra en la Figura 3b, en la etapa 357.

Asimismo, las señales de audio con FOA convertidas pueden analizarse para generar los metadatos (por ejemplo, las direcciones y/o relaciones de energía) como se muestra en la Figura 3b, en la etapa 359.

25 El parámetro de disposición de canal, las señales de audio de transporte y los metadatos pueden combinarse opcionalmente para formar un flujo de datos como se muestra en la Figura 3b, en la etapa 361.

30 El parámetro de disposición de canal, las señales de audio de transporte y los metadatos (o el flujo de datos combinado) pueden transmitirse y recibirse (o almacenarse y recuperarse) como se muestra en la Figura 3b, en la etapa 363.

35 Habiendo recibido o recuperado el parámetro de disposición de canal, las señales de audio de transporte y los metadatos (o flujo de datos), las señales de audio de salida pueden sintetizarse en función de al menos en el parámetro de disposición de canal, las señales de audio de transporte y los metadatos como se muestra en la Figura 3b, en la etapa 365.

Las señales de salida de señales de audio sintetizadas pueden entonces emitirse a una salida adecuada.

40 Con respecto a la Figura 4a, se muestra un sintetizador 309 ilustrativo según algunas realizaciones.

45 Las entradas en el sintetizador 309 pueden, en algunas realizaciones, ser la o las dirección o direcciones 306, la o las relación o relaciones 308, los metadatos espaciales, la corriente 310 de señal de audio de transporte y el parámetro 304 de disposición de parlante de entrada. Otras entradas en el sistema pueden ser un parámetro 452 de disposición de parlante de salida y una entrada 450 de activación/desactivación.

50 Un generador 401 de señales de salida prototipo o cualquier otro medio para generar una señal de salida prototipo de las señales de transporte puede configurarse para recibir las señales 310 de audio de transporte y generar una "señal de salida prototipo". La corriente 310 de señal de audio de transporte puede estar en el dominio del tiempo y convertirse en un dominio de tiempo-frecuencia antes de generar la señal de salida prototipo. Una generación ilustrativa de una señal prototipo a partir de dos señales de transporte puede realizarse mediante el ajuste del canal o de los canales de salida prototipo del lado izquierdo para que sean copias del canal de transporte izquierdo, estableciendo el canal o canales de salida prototipo del lado derecho para que sean copias de los canales de transporte derecho, y los canales prototipo centrales (o medios) para que sean una mezcla de los canales de transporte izquierdo y derecho.

55 Un ejemplo de la señal de salida prototipo es una señal de micrófono virtual que intenta regenerar una señal de micrófono virtual cuando la señal de transporte es en realidad una señal con FOA.

60 Un procesador 403 de raíz cuadrada (relación) puede recibir la o las relaciones 308 que, en los siguientes ejemplos, está en forma de relación de energía directa a total, y generar una raíz cuadrada del valor.

Una primera etapa 411 de ganancia (un generador de señal directa) puede recibir la raíz cuadrada de la o las relaciones y aplicar esto a las señales de audio de salida prototipo para generar la parte de señal de audio directa.

5 Un altavoz activo para el determinador 409 de parte directa está configurado para recibir la entrada 450 de activación/desactivación, la disposición 452 de parlante de salida y el parámetro 304 de disposición de parlante de entrada y configurado para hacer que el sonido de corriente directa sea más similar en la mezcla envolvente original modificando el conjunto activo de parlantes para una reproducción de sonido directo.

10 Por lo tanto, en algunas realizaciones, los altavoces activos para el determinador 409 de parte directa están configurados para desactivar de manera selectiva los altavoces de la configuración de salida, de modo que los altavoces desactivados están entre otros dos altavoces de la configuración de salida y la secuencia de tres altavoces de salida adyacentes se encuentra entre dos altavoces de entrada adyacentes.

15 Esto, en algunas realizaciones, puede producir el efecto de que los sonidos directos que se reproducen originalmente como fuentes fantasma con una panoramización de amplitud entre dos parlantes también se reproducen como fuentes fantasma en la disposición de altavoz de salida más a menudo, ya que no se usa un altavoz físico adicional en la disposición de altavoz de salida entre dos altavoces de disposición de altavoz de entrada.

20 Esto puede hacer que la calidad tímbrica y la extensión espacial de la fuente estén más cerca del original. Sin embargo, los altavoces activos para el determinador 409 de parte directa pueden emplearse para garantizar que la configuración de altavoz de salida no sea demasiado dispersa, ya que solamente se eliminan el o los altavoces medios de un bloque de tres altavoces de salida adyacentes que se encuentran entre un conjunto de dos altavoces de entrada.

25 Se configura una VBAP 407 para recibir la o las direcciones 306 y la salida de los parlantes activos para la salida 409 del determinador de parte directa y generar ganancias de VBAP adecuadas de una manera similar a la descrita anteriormente.

La etapa 415 de ganancia de parte directa está configurada para aplicar las ganancias de VBAP a las señales de audio de parte directa para generar una parte directa procesada espacialmente.

30 Un procesador 405 de raíz cuadrada (1 relación) puede recibir la o las relaciones 308 que, en los siguientes ejemplos, está en forma de una relación de energía directa a total, y generar una raíz cuadrada del valor de 1 relación.

Una segunda etapa 413 de ganancia (un generador de señal difusa) puede recibir la raíz cuadrada de la o las 1 relaciones y aplicar esto a las señales de audio de salida prototipo para generar la parte de señal de audio difusa.

35 Un determinador 423 de ganancia de parte difusa está configurado para recibir la entrada 450 de activación/desactivación, la disposición 452 de parlante de salida, la o las relaciones 308 y el parámetro 304 de disposición de parlante de entrada.

40 Para las señales de audio de parte difusa, el determinador 423 de ganancia de parte difusa puede configurarse para distribuir de manera selectiva (en función de las entradas) la energía de forma no uniforme a diferentes direcciones donde se determina que la configuración de parlantes original no era uniforme. Por lo tanto, la distribución de energía en la reproducción difusa puede estar más cerca de la mezcla envolvente original. Además, el determinador 423 de ganancia de parte difusa puede distribuir de manera selectiva la energía de manera uniforme a todos los parlantes, pero no sonará tan parecido a la mezcla envolvente original si la configuración de parlantes original de la mezcla envolvente tenía un equilibrio espacial diferente de los parlantes.

Una etapa 417 de ganancia difusa puede configurarse para recibir las ganancias de parte difusa y aplicarlas a la parte de señal de audio difusa.

50 A continuación, un combinador 421 puede configurarse para combinar la parte de señal de audio difusa procesada y la parte de señal directa procesada y generar señales de audio de salida adecuadas. En algunas realizaciones, estas señales de audio combinadas pueden convertirse adicionalmente en una forma en el dominio del tiempo antes de la salida a un dispositivo de salida adecuado.

55 Con respecto a la Figura 4b, se muestra un diagrama de flujo de las operaciones del sintetizador 309 que se muestra en la Figura 4a.

60 El método puede comprender recibir las señales de audio con FOA, los metadatos, el parámetro de parlante de entrada, el parámetro de parlante de salida y parámetro de activación/desactivación como se muestra en la Figura 4b, en la etapa 471.

El método también puede comprender generar las señales de audio de salida prototipo en función de las señales de audio con FOA como se muestra en la Figura 4b, en la etapa 475.

65 El método también puede comprender determinar la parte directa de las señales de audio de salida prototipo y los metadatos de relación como se muestra en la Figura 4b, en la etapa 479.

El método también puede comprender determinar la parte difusa de las señales de audio de salida prototipo y los metadatos de relación como se muestra en la Figura 4b, en la etapa 481.

5 Los altavoces de salida activos para el procesamiento de VBAP de parte directa pueden determinarse en función del parámetro de parlante de entrada, el parámetro de parlante de salida y el parámetro de activación/desactivación se muestra en la Figura 4b, en la etapa 473.

10 La aplicación de VBAP a la parte directa y en función de los altavoces activos determinados se muestra en la Figura 4b, en la etapa 483.

15 El método también puede comprender determinar ganancias de parte difusa en función del parámetro de parlante de entrada, el parámetro de parlante de salida y parámetro de activación/desactivación como se muestra en la Figura 4b, en la etapa 477.

El método puede comprender además aplicar las ganancias de parte difusa a la parte difusa determinada como se muestra en la Figura 4b, en la etapa 485.

20 Las partes directas y difusas procesadas pueden combinarse acto seguido para generar las señales de audio de salida como se muestra en la Figura 4b, en la etapa 487.

Entonces pueden emitirse las señales de audio de salida combinadas como se muestra en la Figura 4b, en la etapa 489.

25 Con respecto a la Figura 5, un diagrama de flujo que muestra el funcionamiento de un determinador 423 de ganancia de parte difusa de ejemplo según algunas realizaciones.

30 El determinador 423 de ganancia de parte difusa de ejemplo puede configurarse para recibir/obtener la disposición de parlante de salida como se muestra en la Figura 5, en la etapa 501.

En algunas realizaciones, el determinador 423 de ganancia de parte difusa de ejemplo puede configurarse entonces para determinar si la disposición de parlante de salida es uniforme de izquierda a derecha como se muestra en la Figura 5, en la etapa 503.

35 Cuando la disposición de parlante de salida es uniforme de izquierda a derecha, entonces el determinador 423 de ganancia de parte difusa de ejemplo está configurado para determinar si se necesita una corrección de elevación como se muestra en la Figura 5, en la etapa 505.

40 Cuando la disposición de parlante de salida es uniforme de derecha a izquierda y no se necesita ninguna corrección de elevación, entonces se puede emplear un primer método de determinación de ganancia como se describe a continuación en la presente descripción como se muestra en la Figura 5, en la etapa 509.

45 Cuando la disposición de parlante de salida no es uniforme de izquierda a derecha (o en otras palabras, asimétrica de derecha a izquierda) o la disposición de parlante de salida es uniforme de izquierda a derecha y la corrección de elevación es necesaria, entonces se necesita un segundo método de determinación de ganancia como se describe a continuación, como se muestra en la Figura 5, en la etapa 507.

50 El determinador 423 de ganancia de parte difusa generalmente está configurado para intentar igualar o hacer coincidir la distribución de energía sonora ambiental total con la disposición de parlante de entrada y la disposición de parlante de salida.

55 En algunas realizaciones, en el determinador 423 de ganancia de parte difusa, el método intenta encontrar ganancias  $g_m$  y factores de energía  $e_m (g_m = \sqrt{e_m})$  para los  $M$  altavoces de salida con  $m = 1, \dots, M$  de manera que, después de obtener nuevas ganancias,

1) La suma de vectores ponderados por factor de energía desde la posición central a los parlantes de entrada es igual o coincide con la suma de vectores ponderados por factor de energía desde la posición central hasta los parlantes de salida.

60 2) 
$$\sum_{m=1}^M g_m^2 = 1$$

3) Los factores de energía  $e_m$  no se desvían excesivamente del promedio de  $\frac{1}{M}$ .

Con respecto a la Figura 6, se muestra un ejemplo de disposición de parlante de salida que comprende el altavoz 600 central a 0 grados, el altavoz 604 derecho a -30 grados, el altavoz 602 izquierdo a 30 grados, el altavoz 606 trasero izquierdo a 110 grados y el altavoz 608 trasero derecho a -110 grados. La Figura 6 muestra además el vector 601 de suma formado a partir de los vectores 610 de parlante para la configuración de parlantes 5.0 que se muestra.

Con respecto a la Figura 7, se muestra el primer método de determinación de ganancia que se emplea por el determinador de ganancia de parte difusa.

El primer método de determinación de ganancia intenta fijar el equilibrio de energía frontal-posterior y proporciona un enfoque simple desde el punto de vista computacional y algorítmico que no requiere la inversión de matrices.

La Figura 8 muestra las notaciones usadas para los ángulos de altavoces. El ángulo de acimut  $\phi_m$  es el ángulo de acimut del altavoz de orden  $m$  en radianes. Los ángulos positivos se ejecutan en sentido contrario a las agujas del reloj hasta  $\pi$  y los ángulos negativos en el sentido de las agujas del reloj hasta  $-\pi$ . Los componentes del eje x de los vectores de parlante de salida se indican con  $x_m = \cos(\phi_m)$ .

La primera operación es la obtención o recepción de la disposición de altavoz de salida y la creación de coordenadas x de altavoz de salida como se muestra en la Figura 7, en la etapa 701.

A continuación, se forman las sumas absolutas de las coordenadas x, para coordenadas x positivas y coordenadas x negativas por separado. La suma de coordenadas x positivas se almacena en a y la suma de los valores absolutos de las coordenadas x negativas se almacena en b como se muestra en la Figura 7, en la etapa 703. En otras palabras, definir  $a = \sum_k x_k$ , donde k pertenece al conjunto de índices donde  $x_k > 0$  y  $b = \sum_l |x_l|$ , donde l pertenece al conjunto de índices donde  $x_l < 0$ .

A continuación, los componentes de eje x normalizados se determinan como se muestra en la Figura 7, en la etapa 705. Los componentes normalizados pueden determinarse, por ejemplo, como

$$y_m = \begin{cases} \frac{x_m}{a}, & \text{if } x_m \geq 0 \\ \frac{x_m}{b} & \text{if } x_m < 0. \end{cases}$$

Entonces, la cantidad de corrección de eje x necesaria puede calcularse como se muestra en la Figura 7, en la etapa 707. Esta corrección del eje x se puede determinar usando

$$\Delta x = \frac{1}{N} \sum_{n=1}^N \cos(\theta_n) - \frac{1}{M} \sum_{m=1}^M \cos(\phi_m),$$

donde  $\theta_n$  es el acimut del parlante con configuración de entrada, y  $\phi_m$  es el acimut del parlante con configuración de salida.

Entonces, los factores de energía corregidos por eje x pueden determinarse como se muestra en la Figura 7, en la etapa 709. Los factores de energía corregidos pueden ser

$$\bar{e}_m = \frac{1}{M} + \frac{\Delta x \cdot y_m}{\sum_{k=1}^M (\cos(\phi_k) \cdot y_k)}$$

Para garantizar que los factores de energía son positivos, cualquier factor de energía negativa se fija en cero.

Los factores de energía se normalizan adicionalmente para sumar a la unidad:

$$\hat{e}_m = e_m / \sum_{j=1}^M e_j.$$

Además, en algunas realizaciones, los factores de energía pueden estar limitados de modo que no puedan desviarse del factor de energía promedio de  $\frac{1}{M}$  en más de 6 dB. Después de la limitación, los factores de energía pueden normalizarse de nuevo para que sumen a la unidad, denotando los factores tras la limitación y la segunda normalización  $\tilde{e}_m$ .

Finalmente, las ganancias se calculan como los valores de raíz cuadrada de los factores de energía como se muestra en la Figura 7, en la etapa 711. Esto puede derivarse matemáticamente como:  $g_m = \sqrt{\tilde{e}_m}$ .

El sistema puede entonces, como se ha descrito anteriormente, aplicar las ganancias obtenidas a la corriente difusa durante la síntesis para hacer coincidir la distribución de energía del sonido difuso a la distribución de energía en la mezcla envolvente original.

5 Con respecto a la Figura 9, se muestra el segundo método de determinación de ganancia que se emplea por el determinador de ganancia de parte difusa. El segundo método de determinación de ganancia puede emplearse para obtener las ganancias para configuraciones de parlantes arbitrarias (por ejemplo, asimétricas de izquierda a derecha), incluyendo altavoces elevados. En el ejemplo que se muestra a continuación, la elevación (coordenada z) no se observa para simplificar la presentación, pero la adición de la coordenada z es sencilla.

10 Una operación es la obtención o recepción de la disposición de altavoz de entrada y la creación de vectores de altavoz de entrada como se muestra en la Figura 9, en la etapa 901.

15 El vector de parlante de orden n para la disposición de altavoz de entrada puede definirse como  $I_n = [\cos(\theta_n), \sin(\theta_n)]^T$ ,  $n = 1, \dots, N$ , y  $\theta_n$  es el ángulo de acimut de parlante en radianes.

Otra operación es la obtención o recepción de la disposición de altavoz de salida y la creación de vectores de altavoz de salida como se muestra en la Figura 9, en la etapa 903.

20 El vector de parlante de orden m para la disposición de altavoz de salida se indica con  $O_m = [\cos(\phi_m), \sin(\phi_m)]^T$ , con  $m = 1, \dots, M$  y  $\phi_m$  el acimut correspondiente en radianes.

25 El factor de energía normalizado inicial para cada parlante se establece en la inversa del número de parlantes, en otras palabras  $e_n = \frac{1}{N}$ , como se muestra en la Figura 9, en la etapa 905. En algunas realizaciones, pueden aplicarse otras ponderaciones.

A partir de esto, se forma un vector de suma ponderado por factor de energía de la disposición de parlante de entrada  $\frac{1}{N} \sum_{n=1}^N I_n$ .

30 Un vector de limitación  $t$  se forma al agregar uno al vector de suma para limitar la energía para sumar a la unidad:

$$t = \left[ \frac{1}{N} \sum_{n=1}^N I_n \right]$$

35 Este vector tiene un tamaño 3 x 1 en nuestro ejemplo de configuración de envolvente horizontal, y sería 4 x 1 en un caso de configuraciones en 3D.

Una diana 3 mediante  $M$  matriz de parlantes de salida con 1 añadido similar se forma como se muestra en la Figura 9, en la etapa 907 y se muestra como

40 
$$L = \begin{bmatrix} O_1 & \dots & O_M \\ 1 & \dots & 1 \end{bmatrix}$$

Por lo tanto, L es una matriz 3xM en una configuración en 2D y 4xM para una configuración en 3D.

45 Los factores de energía  $e_m$  pueden representarse como una M por 1 vector  $e$ . El objetivo es encontrar el vector de factor de energía  $e$  de manera que  $t = Le$  como se muestra en la Figura 9, en la etapa 909.

Esto se puede lograr minimizando la norma  $\|Le - t\|$ . La solución se puede encontrar como:

50 
$$e = L^+ t$$

donde  $L^+$  es la pseudoinversa de  $L$ .

55 Después de encontrar el vector de factor de energía, el método comprende además eliminar factores de energía negativos y normalizar su suma a la unidad como se muestra en la Figura 9, en la etapa 911.

La siguiente operación es obtener las ganancias a través de la raíz cuadrada como se describe en relación con el primer método como se muestra en la Figura 9, en la etapa 913.

60 En algunas realizaciones, el segundo método comprende además limitar los factores de energía que no se desvían del factor de energía promedio en más de 6 dB de la manera descrita anteriormente.

En algunas realizaciones, las ganancias de corrección para diferentes pares de disposición de altavoces de entrada-salida pueden precalcularse en una tabla de consulta y luego obtenerse de ahí durante el tiempo de ejecución sin tener que realizar la inversión de matriz en un sistema con recursos limitados.

5 Como se analiza en algunas realizaciones, la operación de determinar los altavoces de salida activos para la parte directa se realiza para intentar hacer que el sonido de corriente directa sea más similar en la mezcla envolvente original modificando el conjunto activo de parlantes para una reproducción de sonido directo.

10 El método desactiva los parlantes desde la configuración de salida, de modo que los altavoces desactivados están entre otros dos altavoces de la configuración de salida y la secuencia de tres altavoces de salida adyacentes se encuentra entre dos altavoces de entrada adyacentes.

15 Esto tiene el efecto de que los sonidos directos que se reproducen originalmente como fuentes fantasma con una panoramización de amplitud entre dos parlantes también se reproducen como fuentes fantasma en la disposición de altavoz de salida más a menudo, ya que no se usa un altavoz físico adicional en la disposición de altavoz de salida entre dos altavoces de disposición de altavoz de entrada. Esto provoca que la calidad tímbrica y la extensión espacial de la fuente estén más cerca del original. Sin embargo, el método también asegura que la configuración de parlantes de salida no sea demasiado dispersa, ya que solamente se eliminan los altavoces medios de un bloque de tres altavoces de salida adyacentes que se encuentran entre un conjunto de dos altavoces de entrada.

20 El siguiente procedimiento se utiliza para obtener el conjunto de altavoces activos para una reproducción de sonido directo según el método:

25 Si en la disposición de altavoz de salida hay un número menor o igual de altavoces que en la disposición de altavoz de entrada, el procedimiento se detiene y no se realiza ninguna modificación en la síntesis de corriente directa.

30 Si hay más altavoces en la disposición de altavoz de salida que en la disposición de altavoz de entrada, las siguientes etapas se ejecutan como se muestra con respecto a la Figura 10, que muestra un diagrama de flujo ilustrativo de la operación de determinar los altavoces de salida activos para la parte directa.

Una operación es la obtención o recepción de la disposición de altavoz de entrada como se muestra en la Figura 10, en la etapa 1001.

35 Otra operación es la obtención o recepción del altavoz de salida como se muestra en la Figura 10, en la etapa 1003.

La disposición de altavoz de salida se clasifica en función del ángulo de acimut en orden ascendente como se muestra en la Figura 10, en la etapa 1005.

40 Los altavoces envolventes de 360 grados se añaden al altavoz de salida primero y último (acimut de último altavoz - 360, acimut de primer altavoz + 360) como se muestra en la Figura 10, en la etapa 1007.

El método entonces itera a través de la secuencia de altavoz de salida creada en bloques de tres altavoces adyacentes con un tamaño de salto de uno como se muestra en la Figura 10, en la etapa 1009.

45 Un bloque de tres altavoces de salida adyacentes está marcado como inadecuado para la eliminación de altavoces si al menos un altavoz de entrada está entre los dos altavoces externos del bloque como se muestra en la Figura 10, en la etapa 1011. En este caso, la definición de entre es entre inclusivo.

50 Para bloques de tres altavoces de salida adyacentes sin ningún altavoz de entrada entre ellos (según lo determinado por la etapa anterior), el método está configurado para retirar el altavoz medio del bloque de tres altavoces de salida adyacentes como se muestra en la Figura 10, en la etapa 1013.

55 Por ejemplo, cuando los altavoces de entrada tienen 30, -30, 110, -110 grados y los altavoces de salida tienen 30, -30, 0, 110, -110 grados, entonces se puede eliminar un altavoz en un ángulo de 0.

Las ganancias de VBAP pueden inicializarse entonces con los altavoces de salida restantes como se muestra en la Figura 1015.

60 En otras palabras, el principio del método es iterar a través de la salida en conjuntos de tres altavoces adyacentes. Si hay al menos un altavoz de la disposición de altavoz de entrada en el centro de los tres altavoces de salida, entonces el bloque de tres altavoces de salida no puede residir entre dos altavoces de entrada. Por lo tanto, no se pueden eliminar los altavoces.

65 El método puede extenderse a configuraciones de parlantes en 3D. En algunas realizaciones, cuando se aplican a una configuración de parlantes en 3D, el método está configurado para procesar solamente el círculo de parlantes horizontal con el método anterior y dejar los altavoces elevados sin procesar.

- 5 En algunas realizaciones, una versión completa para disposiciones de parlantes en 3D puede ser una donde los altavoces de salida candidatos para la retirada se determinan iterando a través de altavoces de entrada en pares de dos, para pares que están conectados en la triangulación de VBAP de entrada, encontrando los altavoces de salida que se encuentran entre ellos, y si hay al menos tres altavoces entre ellos, quitar el central del conjunto de tres.
- En algunas realizaciones, las ganancias de VBAP con disposición de altavoz modificada para la reproducción de sonido directo se utilizan para la panoramización de los componentes de sonido directo.
- 10 En algunas realizaciones para eliminar la necesidad de calcular productos de puntos en el decodificador, se pueden calcular previamente máscaras binarias para que indiquen, por combinaciones conocidas de entrada/salida, el conjunto de altavoces de salida activos.
- 15 En algunas realizaciones, la disposición de altavoz modificada para la reproducción de sonido directo se usa como la disposición de altavoz de salida para la parte difusa.
- En algunas realizaciones, se puede implementar una versión de reproducción de sonido directo para configuraciones de parlantes en 3D junto con la etapa de triangulación de VBAP. La triangulación de VBAP es el proceso de encontrar para cada dirección de acimut y elevación el conjunto de tres parlantes que se usarán para el cálculo de ganancia.
- 20 Con respecto a la Figura 11, el método de la reproducción de sonido directo combinado para configuraciones de parlantes en 3D con la operación de triangulación de VBAP puede comprender el siguiente método.
- 25 Una operación es la obtención o recepción del diseño de altavoz de entrada como se muestra en la Figura 11, en la etapa 1101.
- Otra operación es la obtención o recepción del diseño de altavoz de salida como se muestra en la Figura 11, en la etapa 1103.
- 30 La creación de una triangulación de VBAP en función de la disposición de parlante de entrada se muestra en la Figura 11, en la etapa 1105. Aunque en la presente memoria se describe la triangulación de VBAP, puede analizarse más generalmente la creación de una distribución de superficie virtual o de polígono adecuada.
- 35 A continuación, la creación de una triangulación de VBAP (o distribución de superficie virtual) para la disposición de altavoz de salida usando el siguiente esquema modificado se muestra en la Figura 11, en la etapa 1107. Para ello, antes de realizar cualquier otra reducción de triángulos (o polígonos), compruebe si cada triángulo (o polígono) coincide exactamente con la disposición de altavoz de entrada o dentro de un margen de tolerancia y establezca que se mantengan los triángulos que coincidan con los triángulos de la disposición de altavoz de entrada y, a continuación,
- 40 elimine todos los parlantes de la disposición de altavoz de salida que estén dentro del triángulo mantenido, es decir, elimine todos los triángulos que incluyan un parlante dentro del triángulo mantenido.
- La eliminación de triángulos continúa para cualquier triángulo que tenga bordes cruzados, como se muestra en la Figura 11, en la etapa 1109.
- 45 En algunas realizaciones, un método adicional puede ser para:
- 1) Comprobar triángulos de entrada, si las esquinas de un triángulo coinciden bien con tres parlantes (dentro de la tolerancia, por ejemplo, un grado) en la disposición de altavoz de salida
  - 50 2) En caso afirmativo, hacer un triángulo de estos parlantes y comprobar si hay parlantes de salida dentro de este triángulo, y eliminarlos (a partir de la inicialización de VBAP)
  - 3) Cuando haya terminado con todos los triángulos de entrada, utilice los parlantes de salida restantes para
- 55 inicializar VBAP, es decir, realizar la triangulación mediante cualquier método adecuado. Esto puede considerarse una alternativa a las etapas 1107 y 1109.
- 60 A continuación, el método puede examinar todos los triángulos de la disposición de altavoz de salida y comprobar si algún triángulo de la disposición de altavoz de salida está completamente dentro de un triángulo de la disposición de altavoz de entrada. Si el triángulo de disposición de altavoz de salida tiene más altavoces de disposición de altavoz de salida en su interior, elimine estos altavoces adicionales y también los triángulos de disposición de altavoz de salida que incluyen este altavoz adicional, como se muestra en la Figura 11, en la etapa 1111.
- 65 Los triángulos restantes se pueden usar para inicializar las ganancias de VBAP como se muestra en la Figura 11, en la etapa 1113.

En algunas realizaciones, los nodos de altavoz de disposición de altavoz pueden desactivarse a partir de la determinación de ganancia de parte difusa usando un método adicional:

1. Crear triángulos de disposición de altavoz de entrada
2. Comprobar los triángulos de entrada, si las esquinas de un triángulo coinciden bien con tres parlantes (dentro de la tolerancia, por ejemplo, un grado) en la disposición de altavoz de salida
3. En caso afirmativo, hacer un triángulo de estos parlantes y comprobar si hay parlantes de salida dentro de este triángulo, y eliminarlos de la determinación de ganancia de parte difusa

Con respecto a la Figura 12, 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.

5 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.

15 Tal y como se emplea en esta solicitud, el término «circuitería» puede referirse a uno o a más o a todos de los siguientes conceptos:

- (a) implementaciones de circuitos solo en hardware (tales como implementaciones en conjuntos de circuitos solo analógicos y/o digitales) y
- 20 (b) combinaciones de circuitos físicos y software, tales como (según corresponda):
  - (i) una combinación de circuito(s) de hardware analógico y/o digital con software/firmware y
  - (ii) cualquier porción de procesador o procesadores de hardware con software (incluyendo procesador o procesadores de señales digitales), software y memoria o memorias que funcionan juntos para hacer que un aparato, tal como teléfono móvil o servidor, realice diversas funciones) y
- 25 (c) un(os) circuito(s) de hardware y/o procesador(es), tales como un(os) microprocesador(es) o parte de un(os) microprocesador(es), que requiere(n) un software (p. ej., firmware) para su funcionamiento, pero el software puede no estar presente cuando no sea necesario para el funcionamiento.
- 30

Esta definición de circuitería aplica a todos los usos de este término en esta solicitud, incluyendo en cualquier reivindicación. Como ejemplo adicional, tal y como se emplea en esta solicitud, el término circuitería también cubre una implementación de tan solo un circuito físico o un procesador (o múltiples procesadores) o de una parte de un circuito físico o de un procesador y de su software y/o firmware acompañante. Por ejemplo, y si fuese aplicable a un elemento de reivindicación particular, el término circuitería también abarca un circuito integrado de banda base o circuito integrado de procesador para un teléfono móvil o un circuito integrado similar en un servidor, un dispositivo de red celular u otro dispositivo informático o de red.

40 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 etapas del programa, o circuitos lógicos interconectados, bloques y funciones, o una combinación de etapas 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.

50 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.

60 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.

65 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.

5 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, la materia para la cual se busca protección es única y definida por las reivindicaciones adjuntas.

10

REIVINDICACIONES

1. Un aparato (1400) para una decodificación de señal de audio espacial asociado con una pluralidad de nodos de altavoz colocados dentro de un espacio tridimensional, el aparato configurado para:

5 recibir al menos una señal (300) de audio asociada, la señal de audio asociada en función de una señal de audio de disposición de altavoz definida; metadatos espaciales asociados con la señal (306) (308) de audio asociada; al menos un parámetro que representa una disposición de altavoz definida asociada con la señal (304) de audio de disposición de altavoz definida; y al menos un parámetro que representa una disposición (452) de altavoz de salida;

10 sintetizar (309), a partir de la al menos una señal de audio asociada, al menos una señal de audio de salida en función de los metadatos espaciales y del al menos un parámetro que representa la disposición de altavoz definida y del al menos un parámetro que representa una disposición de altavoz de salida, en donde el aparato, cuando está configurado para sintetizar, a partir de la al menos una señal de audio asociada, al menos una señal de audio de salida en función de los metadatos espaciales y del al menos un parámetro que representa la disposición de altavoz definida y del al menos un parámetro que representa una disposición de altavoz de salida, está configurado además para:

20 dividir la al menos una señal de audio asociada en una parte directa y una parte difusa en función de los metadatos espaciales;

sintetizar una señal de audio directa en función de la parte directa de la al menos una señal (411) (403) (415) (407) de audio asociada;

25 sintetizar una señal de audio difusa en función de la parte difusa de la al menos una señal (405) (413) (419) (417) de audio asociada; y

combinar la señal de audio directa y la señal de audio difusa para generar la al menos una señal (421) de audio de salida, en donde al menos una de la síntesis de la señal de audio directa y/o de la síntesis de la señal de audio difusa se basa en el al menos un parámetro que representa la disposición (304) de altavoz definida y en el al menos un parámetro que representa una disposición (452) de altavoz de salida.

30
  
2. El aparato según la reivindicación 1, en donde el aparato configurado para sintetizar una señal de audio difusa en función de la parte difusa de la al menos una señal de audio asociada está configurado para:

35 decorrelacionar (419) la al menos una señal de audio asociada;

generar un conjunto de ganancias para cada disposición (423) de altavoz de salida, las ganancias en función de la coincidencia de una distribución de energía de parte difusa total con la disposición de altavoz de entrada y la disposición de altavoz de salida.
  
3. El aparato según la reivindicación 2, en donde el aparato configurado para generar un conjunto de ganancias (423) para cada una de la disposición (452) de altavoz de salida, las ganancias en función de la coincidencia de una distribución de energía de parte difusa total con la disposición de altavoz de entrada y la disposición de altavoz de salida, está configurado para:

45 determinar las ubicaciones (701) de disposición de altavoz de salida;

determinar sumas absolutas para componentes de eje x de nodos de altavoz en función de las ubicaciones (703) de disposición de altavoz de salida;

determinar componentes de eje x normalizados de nodos (705) de altavoz;

40 determinar una corrección de componente de eje x en función de los componentes (707) x normalizados;

50 determinar factores de energía corregidos con el componente de eje x en función de la corrección (709) de componente x; y

determinar el conjunto de ganancias para cada uno de los nodos (711) de altavoz de salida.
  
4. El aparato según la reivindicación 2, en donde el aparato configurado para generar un conjunto de ganancias (423) para cada una de la disposición (452) de altavoz de salida, las ganancias en función de la coincidencia de una distribución de energía de parte difusa total con la disposición de altavoz de entrada y la disposición de altavoz de salida, está configurado para:

55 determinar vectores de disposición de altavoz de salida en función de la disposición (903) de altavoz de salida;

60 determinar vectores de disposición de altavoz de entrada en función de la disposición (901) de altavoz de entrada;

establecer un factor de energía inicial en función de la inversa de un número de nodos (905) de altavoz de salida;

65

- 5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60  
 65
- formar un vector diana en función de un vector de suma ponderado por factor de energía de los vectores (907) de disposición de altavoz de entrada;  
 optimizar los factores de energía de manera que un vector de suma ponderado por factor de energía de los vectores de disposición de altavoz de salida coincide con el vector (909) diana;  
 eliminar cualquier valor negativo de los factores (911) de energía optimizados;  
 normalizar la suma de los factores de energía optimizados restantes a la unidad (911); y  
 determinar el conjunto de ganancias a partir de los factores (913) de energía optimizados restantes normalizados.
5. El aparato según cualquiera de las reivindicaciones 1 a 4, en donde el aparato configurado para sintetizar (405) (413) (419) (417) una señal de audio difusa en función de la parte difusa de la al menos una señal de audio asociada está configurado para:  
 determinar y desactivar nodos de altavoz a partir de la disposición de altavoz de salida, de modo que los nodos de altavoz desactivados están ubicados entre otros dos nodos de altavoz de la disposición de altavoz de salida y que la secuencia de tres nodos de altavoz de salida adyacentes se encuentra inclusivamente entre dos nodos de altavoz de entrada adyacentes.
6. El aparato según cualquiera de las reivindicaciones 1 a 5, en donde el aparato configurado para sintetizar (411) (403) (415) (407) una señal de audio directa en función de la parte (409) directa de la al menos una señal de audio asociada está configurado para:  
 determinar y desactivar nodos de altavoz a partir de la disposición de altavoz de salida, de modo que los nodos de altavoz desactivados están ubicados entre otros dos nodos de altavoz de la disposición de altavoz de salida y que la secuencia de tres nodos de altavoz de salida adyacentes se encuentra inclusivamente entre dos nodos de altavoz de entrada adyacentes.
7. El aparato según cualquiera de las reivindicaciones 5 y 6, en donde el aparato configurado para determinar y desactivar nodos de altavoz a partir de la disposición de altavoz de salida, de modo que los nodos de altavoz desactivados están ubicados entre otros dos nodos de altavoz de la disposición de altavoz de salida y que la secuencia de tres nodos de altavoz de salida adyacentes se encuentra entre dos nodos de altavoz de entrada adyacentes está configurado para:  
 determinar si hay más altavoces en la disposición de altavoz de salida que en la disposición de altavoz de entrada;  
 clasificar la disposición de altavoz de salida en función de un ángulo de acimut para generar una lista (1005) clasificada;  
 agregar un primer y un último altavoz de salida según la clasificación al final de la lista clasificada y al inicio de la lista (1007) clasificada respectivamente;  
 iterar a través de la lista clasificada adjunta en bloques de tres altavoces adyacentes con un tamaño de salto de uno (1009);  
 mantener un bloque de tres altavoces de salida adyacentes donde al menos un altavoz de entrada está ubicado entre los dos altavoces externos del bloque (1011); y  
 retirar un medio de un bloque de tres altavoces de salida adyacentes donde no hay altavoces de entrada entre los dos altavoces externos del bloque (1013).
8. El aparato según la reivindicación 7, en donde el aparato configurado para clasificar la disposición de altavoz de salida en función de un ángulo de acimut para generar una lista clasificada se hace para clasificar solamente altavoces de salida de orientación horizontal.
9. El aparato según cualquiera de las reivindicaciones 1 a 4, en donde el aparato configurado para sintetizar (411) (403) (415) (407) una señal de audio directa en función de la parte directa de la al menos una señal de audio asociada está configurado para:  
 crear una distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz en función de la disposición (1105) de altavoz de entrada;  
 crear una distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de salida en función de la disposición (1107) de altavoz de salida;  
 examinar cualquier superficie virtual en función de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de salida que coincide con al menos una superficie virtual de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada y retirar todas las superficies virtuales de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de salida que implican un altavoz en el interior de la superficie (1109) virtual coincidente;  
 examinar cualquier superficie virtual en función de la disposición de salida y retirar una superficie virtual si está completamente en el interior de una superficie (1111) virtual de disposición de entrada;

retirar cualquier altavoz donde la superficie virtual de disposición de salida tiene altavoces adicionales en el interior para generar una distribución de superficie virtual de altavoz reducida; y generar una panoramización de amplitud de base vectorial en función de las superficies (1113) virtuales de base vectorial de disposición de altavoz de salida mantenidas.

5  
10. El aparato según cualquiera de las reivindicaciones 1 a 4, en donde el aparato configurado para sintetizar (411) (403) (415) (407) una señal de audio directa en función de la parte directa de la al menos una señal de audio asociada está configurado para:

10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65

crear una distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada en función de la disposición de altavoz de entrada; examinar cualquier nodo de altavoz de disposición de altavoz de salida que coincida con al menos un nodo dentro de la superficie virtual de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada y retirar todos los nodos de altavoz de disposición de altavoz de salida que no forman parte de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada pero están ubicados en el interior de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada; crear una distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de salida en función de los nodos de altavoz de disposición de altavoz de salida restantes; examinar cualquier superficie virtual en función de la disposición de salida y retirar una superficie virtual si está completamente en el interior de una superficie virtual de disposición de entrada; retirar cualquier altavoz donde la superficie virtual de disposición de salida tiene altavoces adicionales en el interior para generar una distribución de superficie virtual de altavoz reducida; y generar una panoramización de amplitud de base vectorial en función de las superficies virtuales de base vectorial de disposición de altavoz de salida mantenidas.

11. El aparato según cualquiera de las reivindicaciones 1 a 4, en donde el aparato configurado para sintetizar (405) (413) (419) (417) una señal de audio difusa en función de la parte difusa de la al menos una señal de audio asociada está configurado para:

35  
40  
45  
50  
55  
60  
65

crear una distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada en función de la disposición de altavoz de entrada; examinar cualquier nodo de altavoz de disposición de altavoz de salida que coincida con al menos un nodo dentro de la superficie virtual de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada y retirar todos los nodos de altavoz de disposición de altavoz de salida que no forman parte de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada pero están ubicados en el interior de la distribución de superficie virtual de panoramización de amplitud de base vectorial de disposición de altavoz de entrada.

12. Un método para una decodificación (365) de señal de audio espacial asociado con una pluralidad de nodos de altavoz colocados dentro de un espacio tridimensional, comprendiendo el método:

45  
50  
55  
60  
65

recibir (471) al menos una señal de audio asociada, la señal de audio asociada en función de una señal de audio de disposición de altavoz definida; metadatos espaciales asociados con la señal de audio asociada; al menos un parámetro que representa una disposición de altavoz definida asociada con la señal de audio de disposición de altavoz definida; y al menos un parámetro que representa una disposición de altavoz de salida; y sintetizar, a partir de la al menos una señal de audio asociada, al menos una señal de audio de salida en función de los metadatos espaciales y del al menos un parámetro que representa la disposición de altavoz definida y del al menos un parámetro que representa una disposición de altavoz de salida, en donde la etapa de sintetizar, a partir de la al menos una señal de audio asociada, al menos una señal de audio de salida en función de los metadatos espaciales y del al menos un parámetro que representa la disposición de altavoz definida y del al menos un parámetro que representa una disposición de altavoz de salida comprende además dividir la al menos una señal de audio asociada en una parte directa y una parte difusa en función de los metadatos espaciales; sintetizar una señal de audio directa en función de la parte directa de la al menos una señal (473) (479) (483) de audio asociada; sintetizar una señal de audio difusa en función de la parte difusa de la al menos una señal (477) (481) (485) de audio asociada; y combinar la señal de audio directa y la señal de audio difusa para generar la al menos una señal de audio de salida, en donde al menos una de la síntesis de la señal de audio directa y/o de la síntesis de la señal de audio difusa se basa en el al menos un parámetro que representa la disposición de

altavoz definida y en el al menos un parámetro que representa una disposición (487) de altavoz de salida.

5 13. El método según la reivindicación 12, en donde sintetizar una señal de audio difusa en función de la parte difusa de la al menos una señal de audio asociada comprende:

10 decorrelacionar la al menos una señal (481) de audio asociada;  
generar un conjunto de ganancias para cada disposición de altavoz de salida, las ganancias en función de la coincidencia de una distribución de energía de parte difusa total con la disposición de altavoz de entrada y la disposición (477) de altavoz de salida.

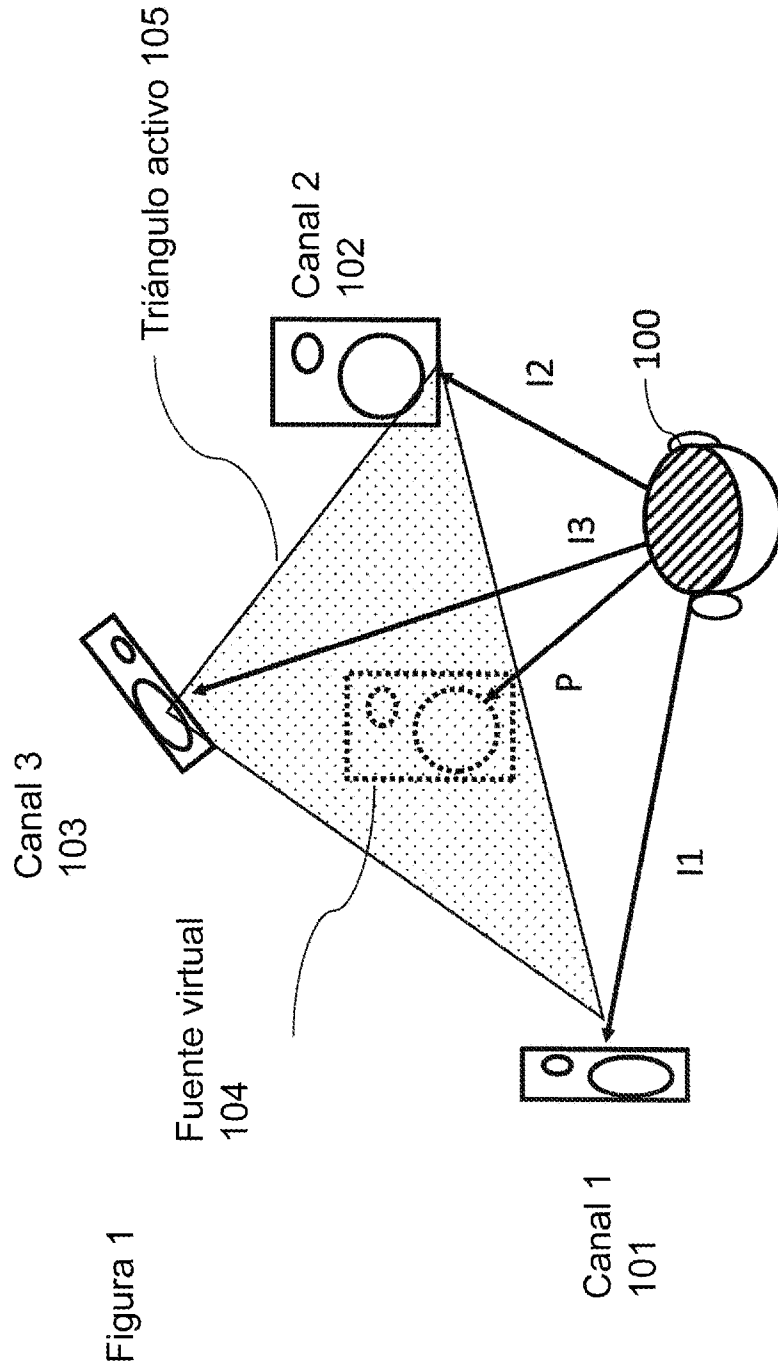
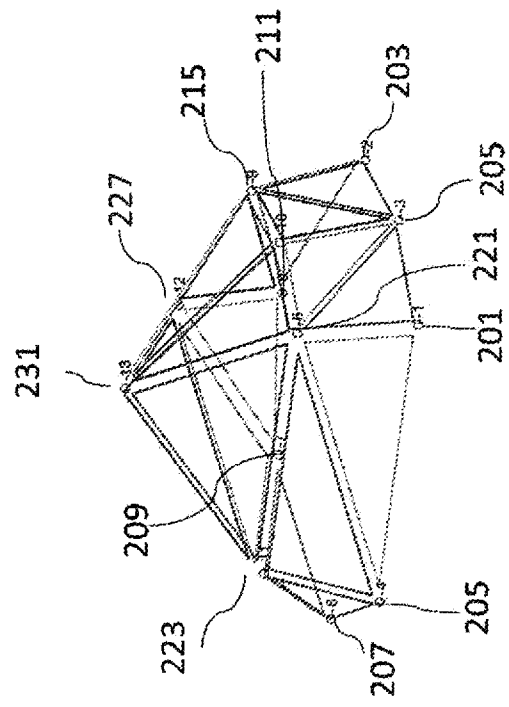


Figura 2



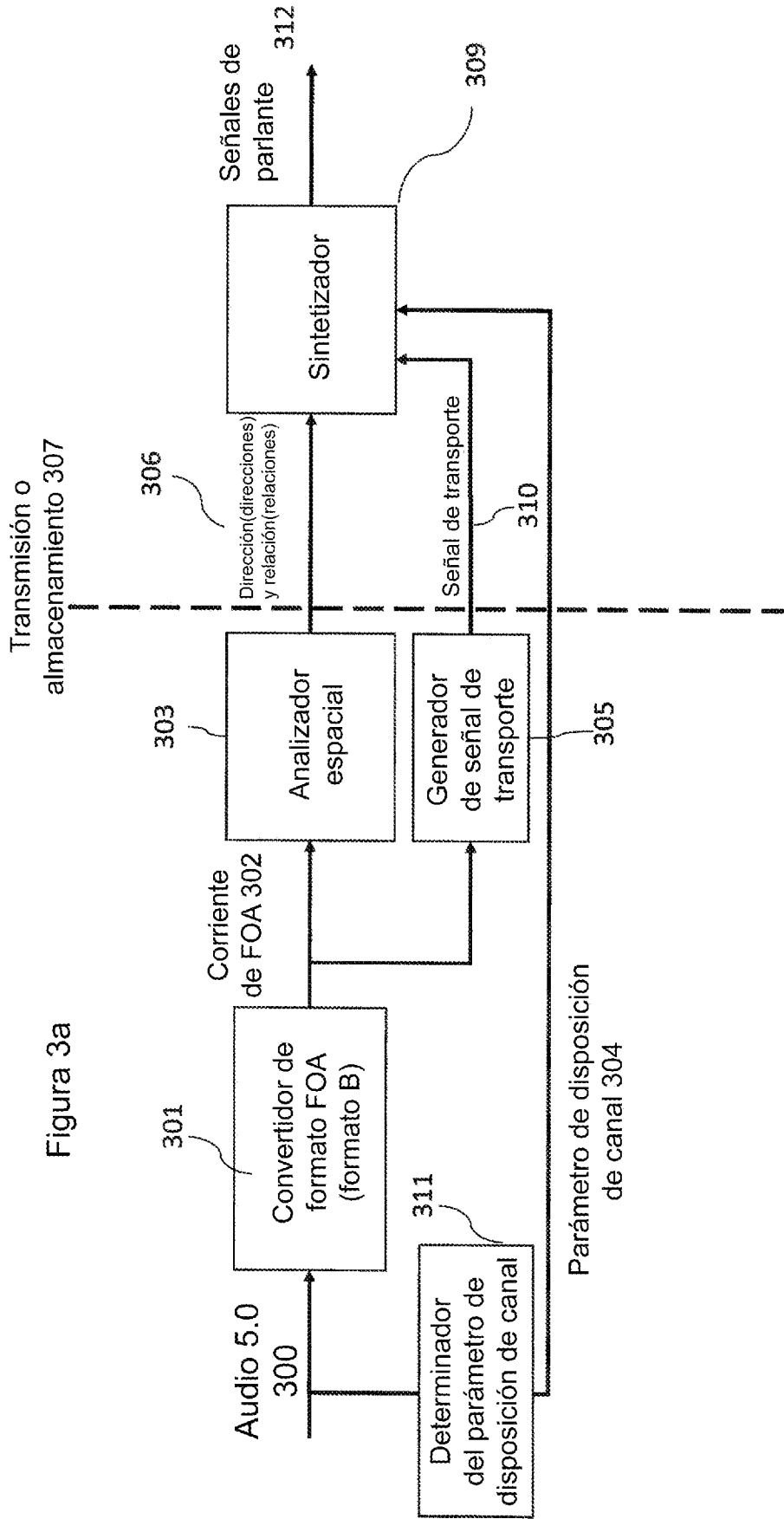
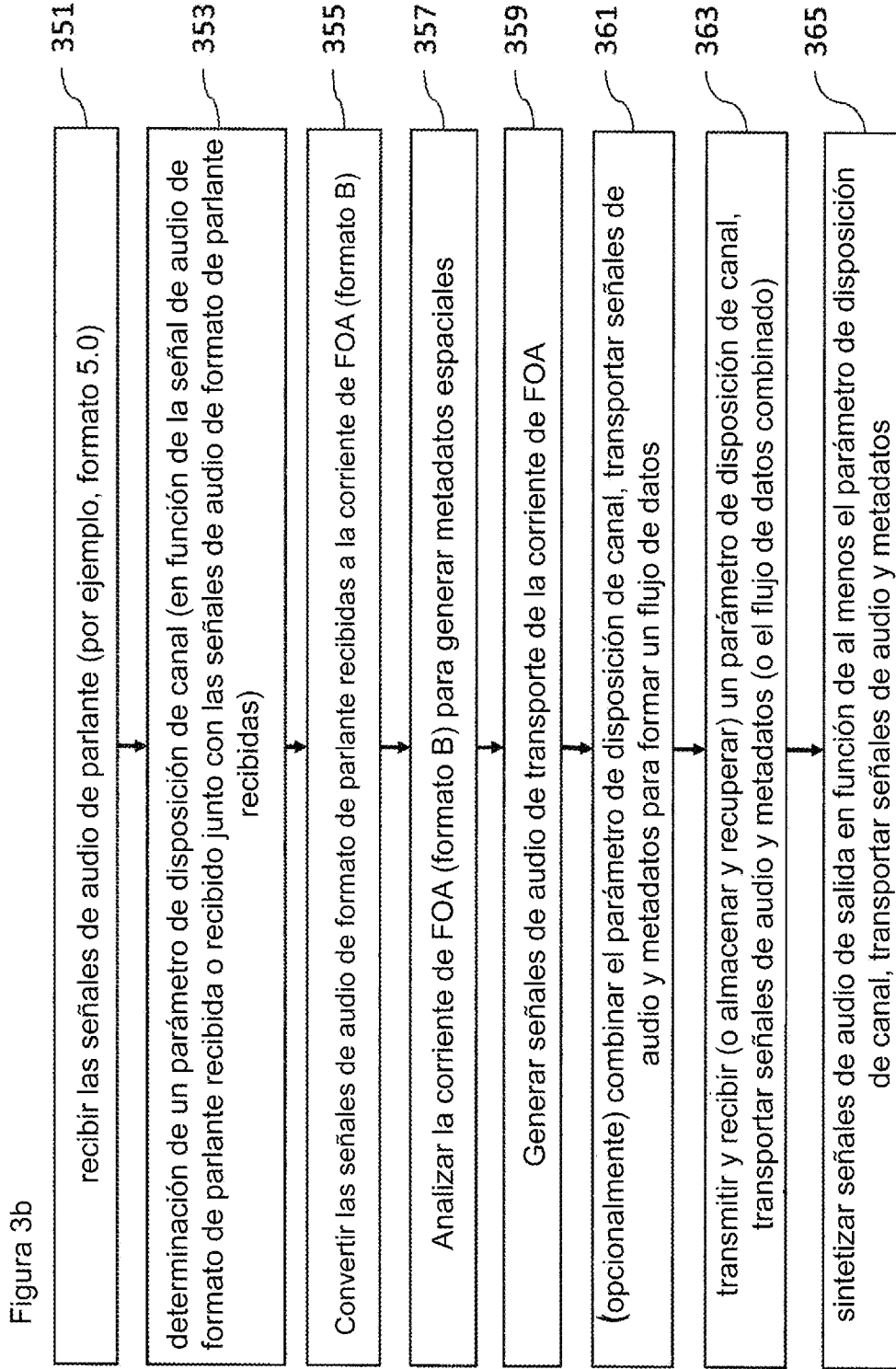
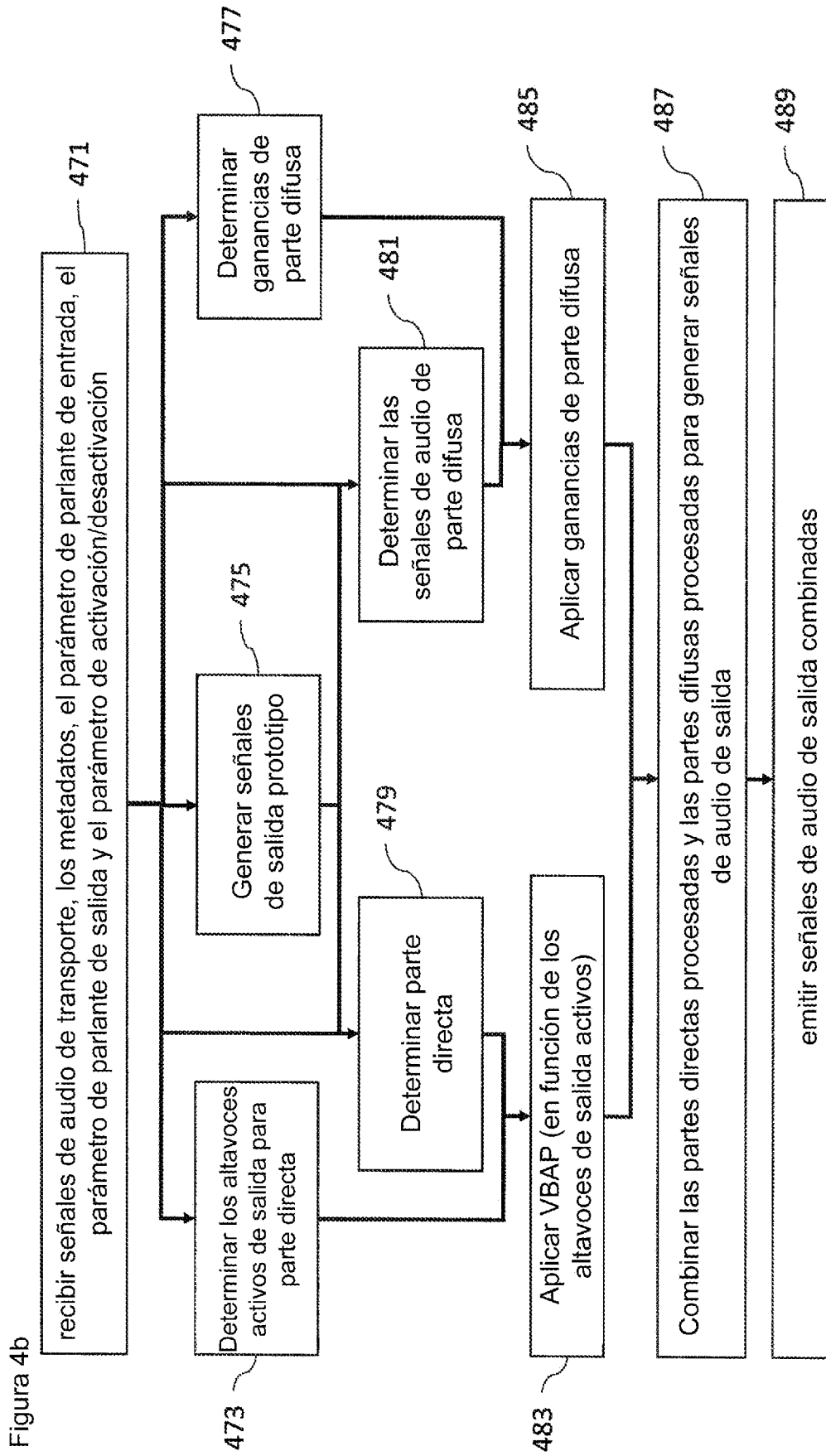


Figura 3a







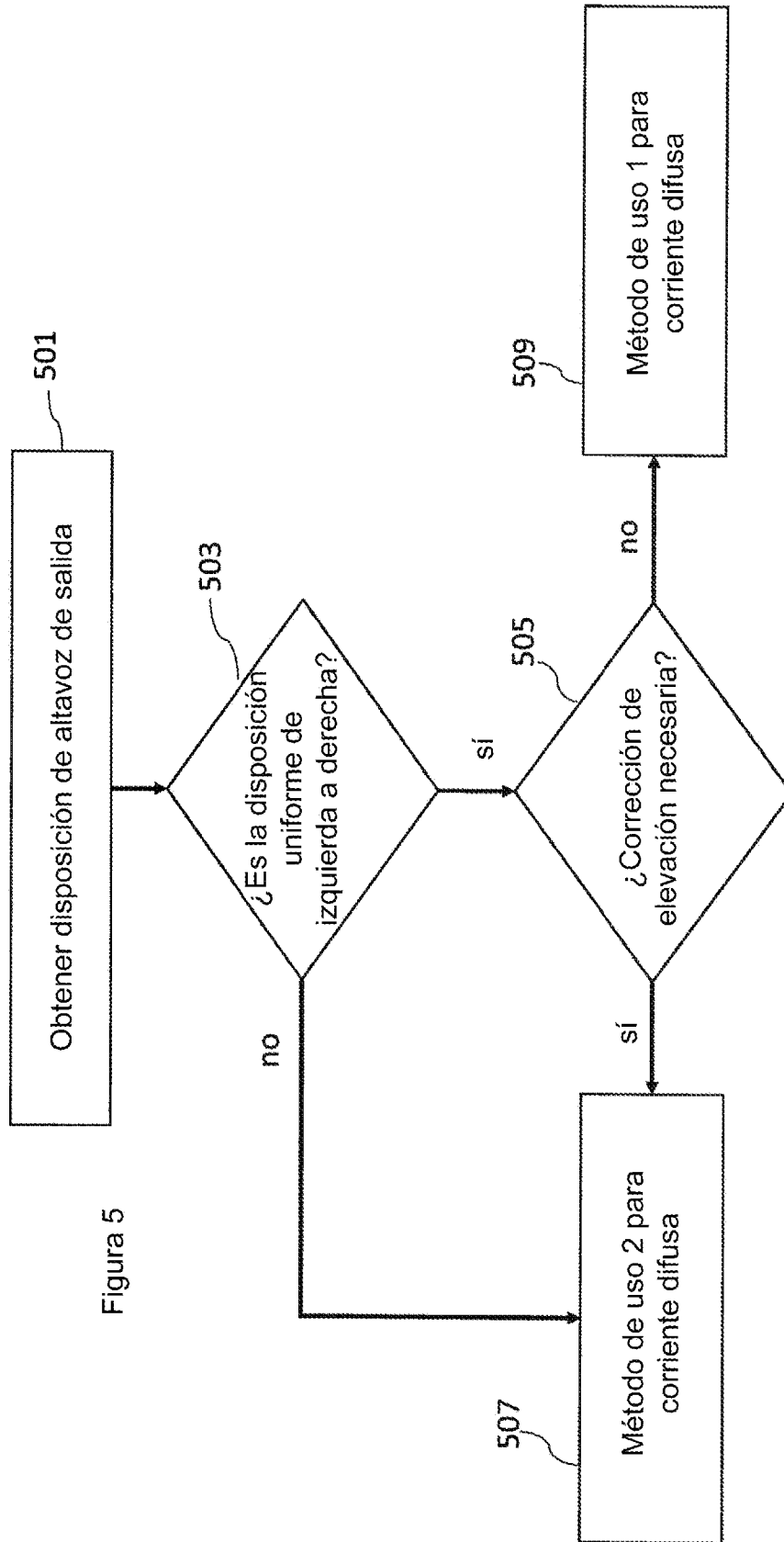


Figura 5

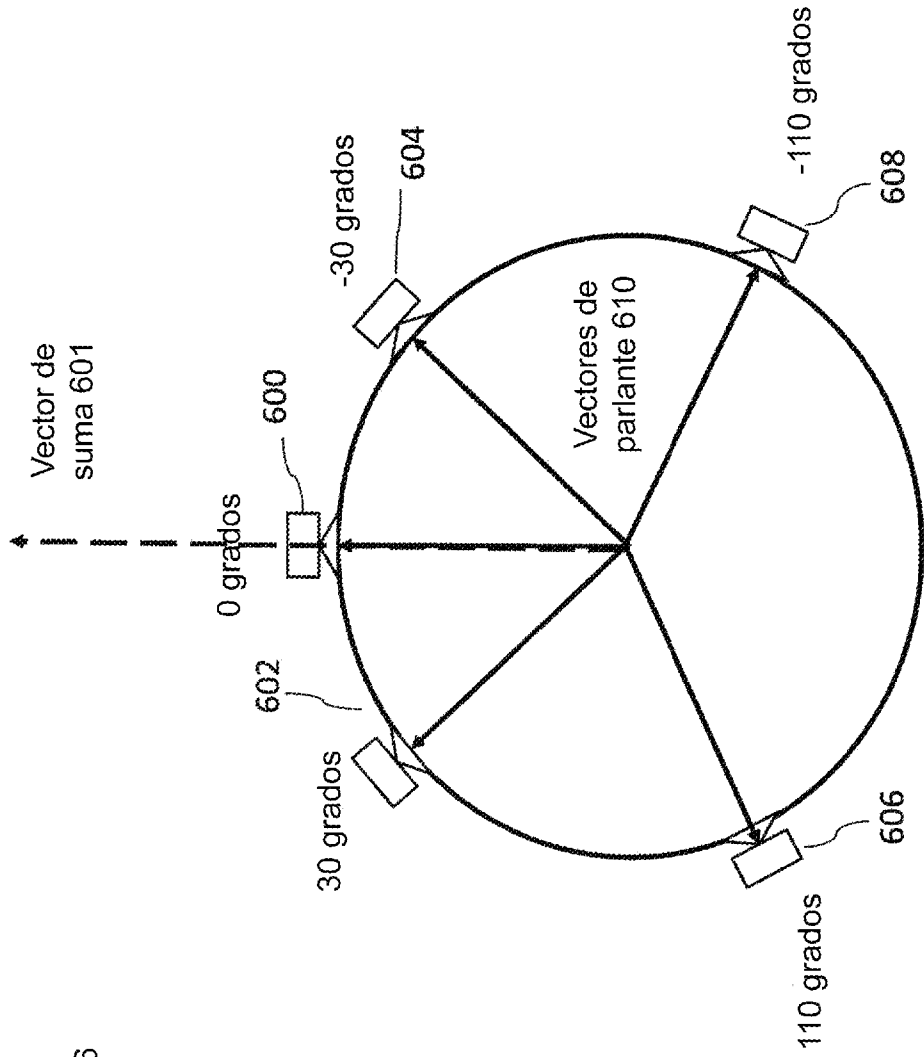
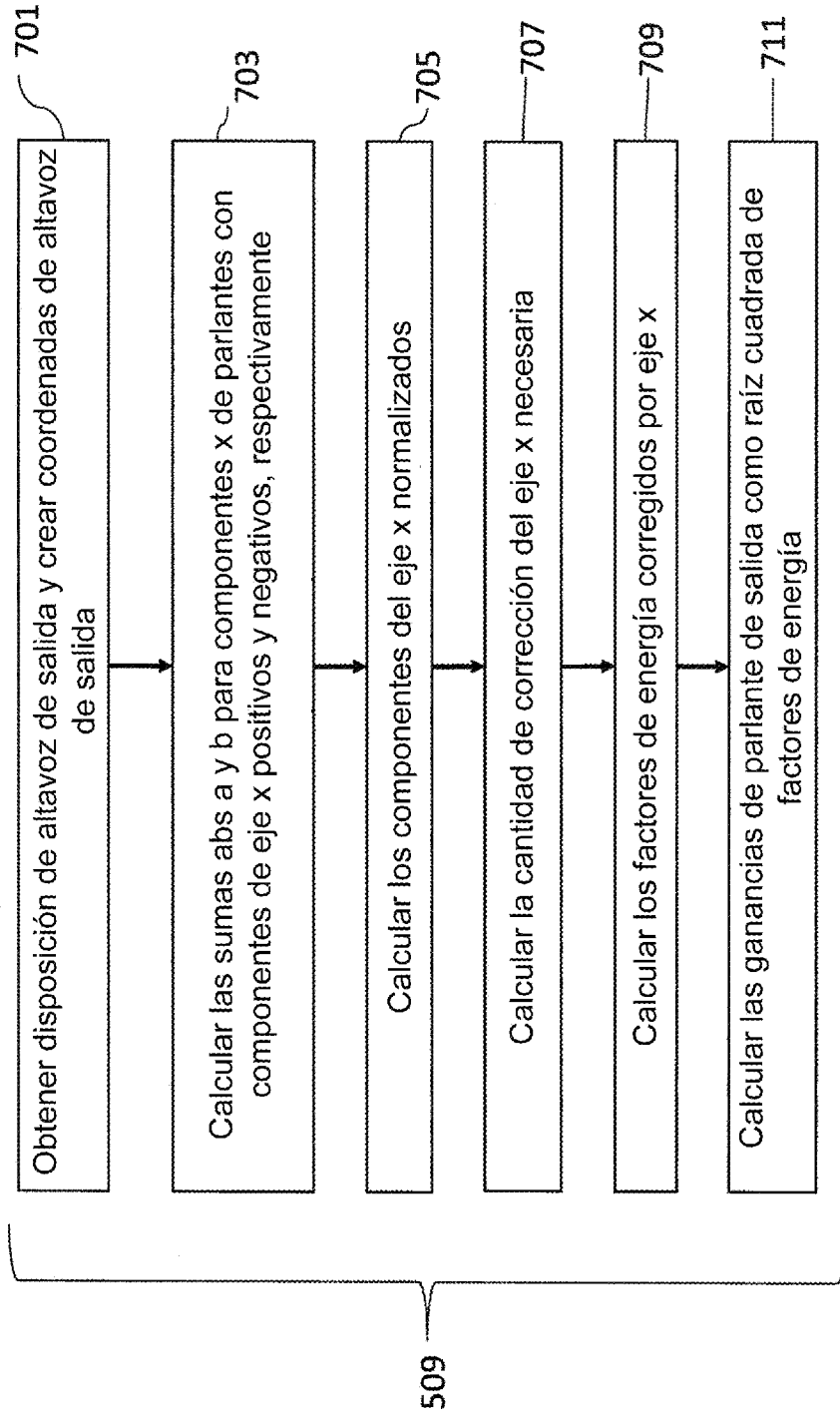


Figura 6

Figura 7



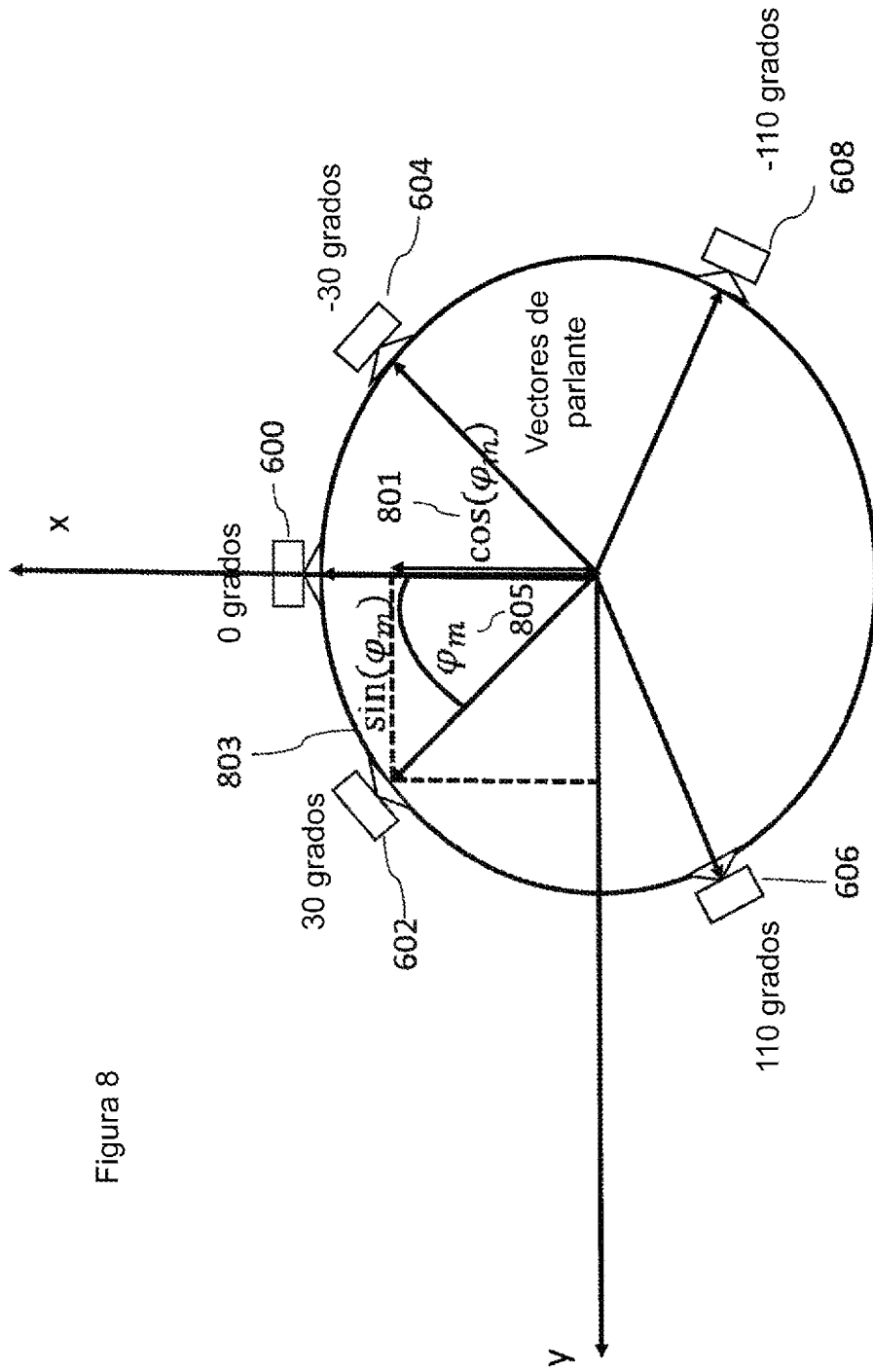


Figura 8

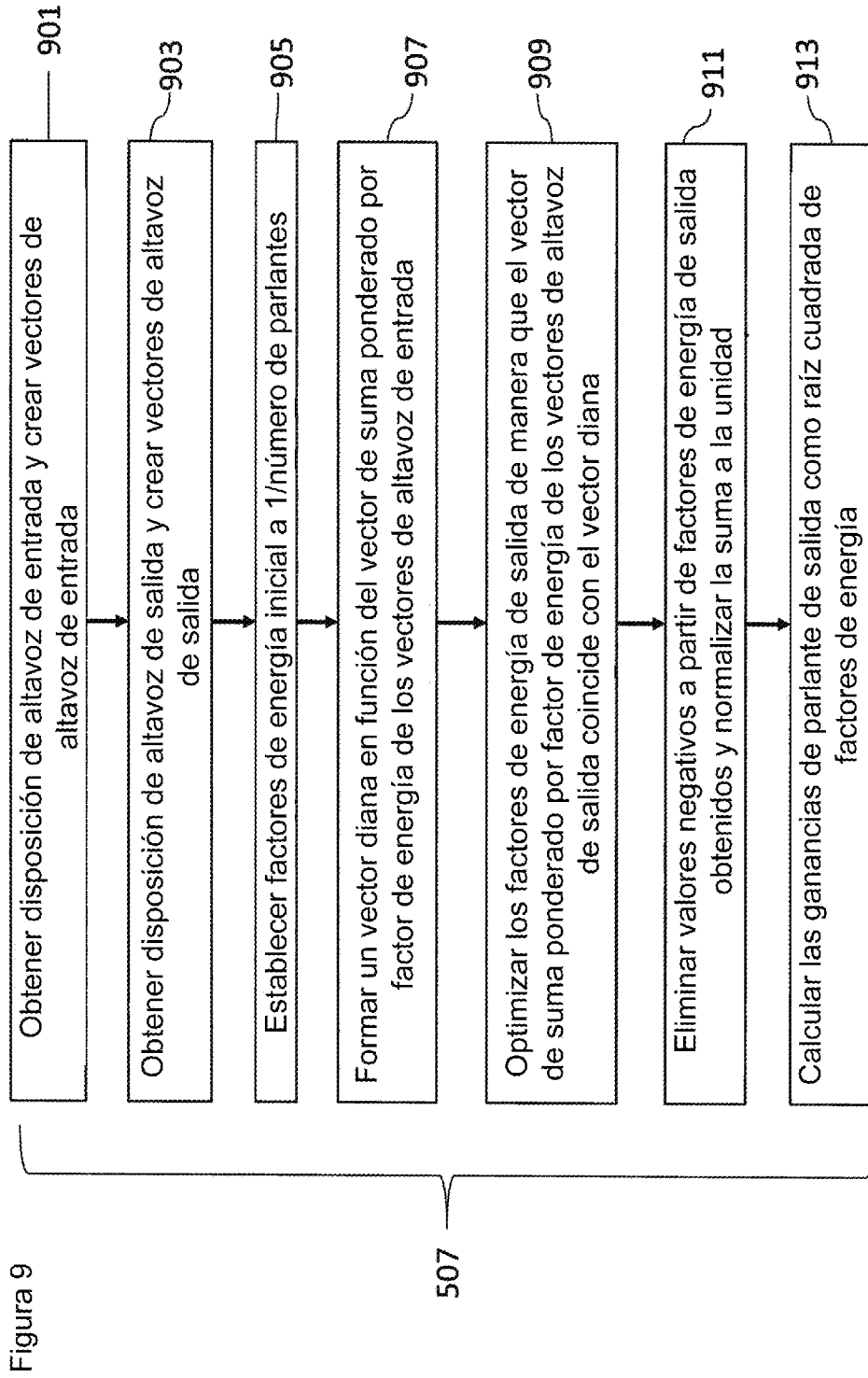


Figura 10

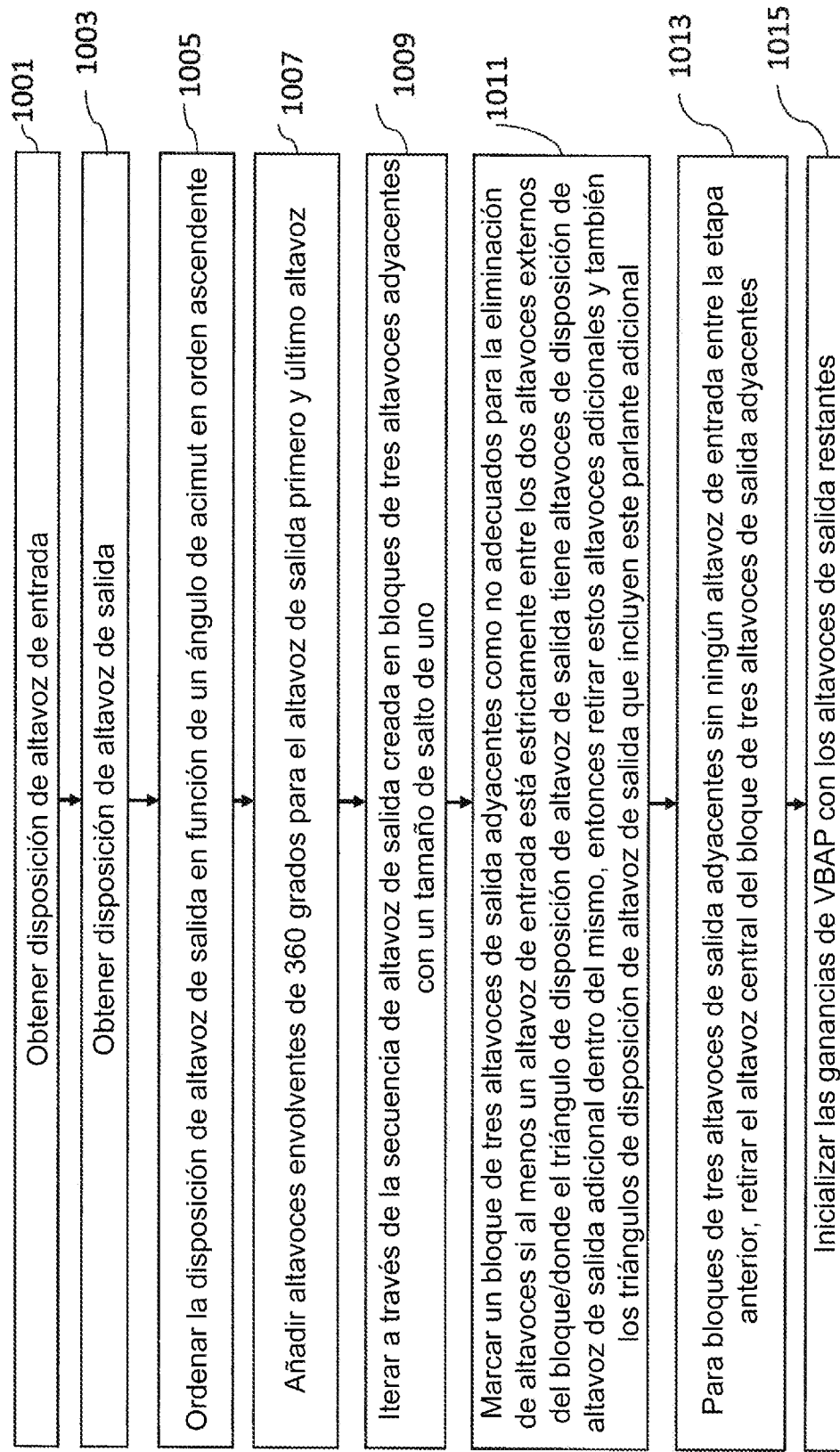
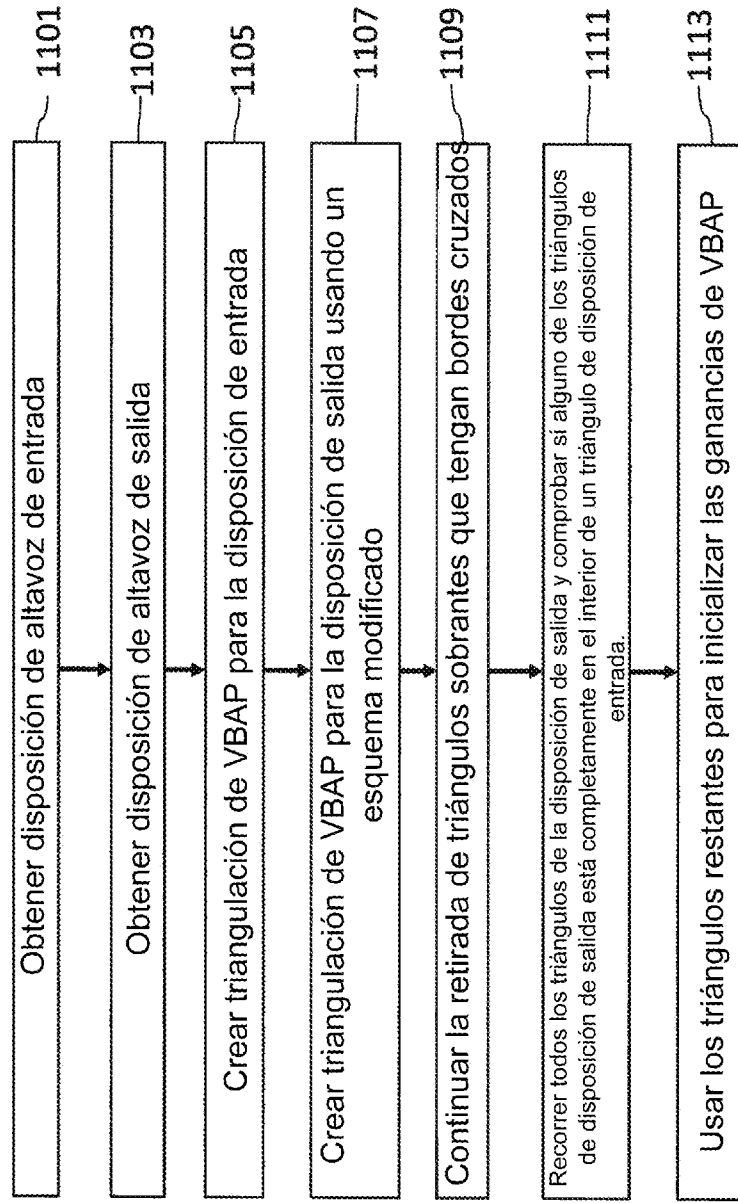


Figura 11



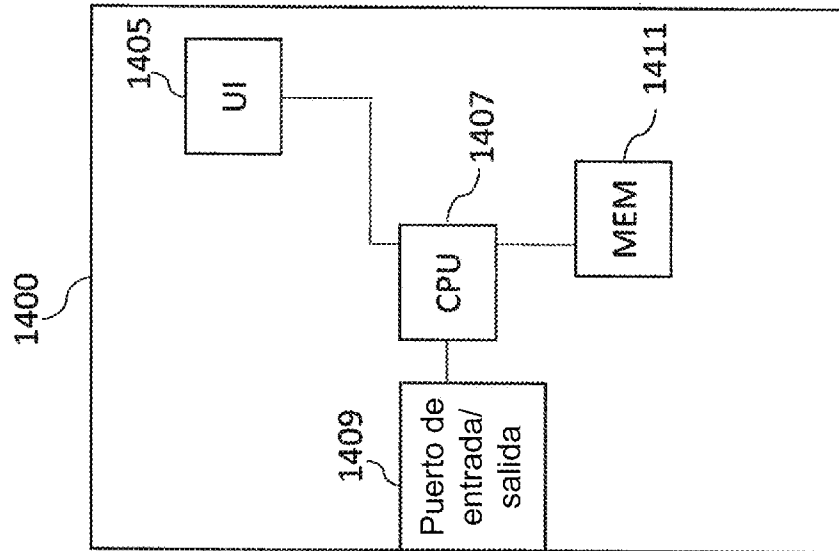


Figura 12