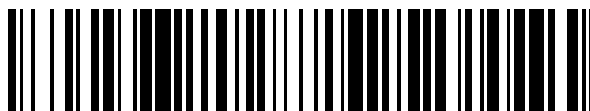


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 609 054**

51 Int. Cl.:

H04S 7/00 (2006.01)

G10L 19/08 (2013.01)

G10L 19/008 (2013.01)

H04S 3/00 (2006.01)

H04S 5/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **12.11.2013 PCT/EP2013/073574**

87 Fecha y número de publicación internacional: **22.05.2014 WO14076058**

96 Fecha de presentación y número de la solicitud europea: **12.11.2013 E 13789558 (7)**

97 Fecha y número de publicación de la concesión europea: **28.09.2016 EP 2904818**

54 Título: **Aparato y método para generar una pluralidad de transmisiones de audio paramétricas y aparato y método para generar una pluralidad de señales de altavoz**

30 Prioridad:

15.11.2012 US 201261726887 P

15.03.2013 EP 13159421

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

18.04.2017

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (50.0%)**

Hansastrasse 27c

80686 München, DE y

TECHNISCHE UNIVERSITÄT ILMENAU (50.0%)

72 Inventor/es:

KÜCH, FABIAN;

DEL GALDO, GIOVANNI;

KUNTZ, ACHIM;

PULKKI, VILLE y

POLITIS, ARCHONTIS

74 Agente/Representante:

SALVA FERRER, Joan

ES 2 609 054 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

Aparato y método para generar una pluralidad de transmisiones de audio paramétricas y aparato y método para generar una pluralidad de señales de altavoz

5

Campo Técnico

La presente invención generalmente se relaciona con un procesamiento de audio espacial paramétrico, y en particular con un aparato y método para generar una pluralidad de transmisiones de audio paramétricas y un aparato y método para generar una pluralidad de señales de altavoz. Otras realizaciones de la presente invención se relacionan con un procesamiento de audio espacial paramétrico basado en sectores.

10

Antecedentes de la Invención

Al escuchar por múltiples canales, el oyente se encuentra rodeado de múltiples altavoces. Existen varios métodos conocidos de captura de audio para dichas configuraciones. Consideremos primero los sistemas de altavoz y la impresión espacial que puede ser creada con ellos. Sin técnicas especiales, las configuraciones comunes estereofónicas de dos canales sólo pueden crear eventos auditivos en la línea que conecta los altavoces. El sonido que emana de otras direcciones no puede ser producido. Lógicamente, al usar más altavoces alrededor del oyente, se podrán cubrir más direcciones y crear una impresión espacial más natural. El sistema y diseño de altavoces con múltiples canales más conocido es el estándar 5.1 ("ITU-R 775-1"), que consiste en cinco altavoces en ángulos azimutales de 0°, 30° y 110° con respecto a la posición de escucha. También se conocen otros sistemas con varios altavoces situados en diferentes direcciones.

20

En la técnica, se han diseñado varios métodos de grabación para los sistemas de altavoz mencionados antes, para reproducir la impresión espacial en la situación de escucha como se percibiría en el ambiente de grabación. La manera ideal de grabar el sonido espacial para un sistema de altavoz con múltiples canales elegido utilizará el mismo número de micrófonos como altavoces existan. En dicho caso, los patrones de direccionalidad de los micrófonos deberán corresponder a la disposición del altavoz de modo que el sonido que emane de una dirección será grabado sólo con uno, dos o tres micrófonos. Cuantos más altavoces se utilicen, más estrechos los patrones de direccionalidad necesarios. Sin embargo, los micrófonos de direccionalidad estrechos son relativamente onerosos, y tienen típicamente una respuesta de frecuencia no plana, no deseada. Asimismo, al utilizar varios micrófonos con patrones de direccionalidad muy amplios como entrada de la reproducción por múltiples canales da por resultado una percepción auditiva colorida e imprecisa, debido a que el sonido que emana de una dirección siempre se reproduce con más altavoces de los necesarios. Por ende, los micrófonos actuales se adaptan mejor al sistema de grabación y reproducción de dos canales sin tener como objetivo una impresión espacial circundante.

30

35

Otro método conocido de grabación de sonido espacial es grabar una gran cantidad de micrófonos que se distribuyen sobre un área espacial amplia. Por ejemplo, al grabar una orquesta en un escenario, los instrumentos individuales podrán ser levantados por los denominados micrófonos de punto, que se colocan cerca de las fuentes de sonido. La distribución espacial del escenario de sonido delantero podrá ser, por ejemplo, capturada por micrófonos estéreos convencionales. Los componentes de campo de sonido que corresponden a la reverberación tardía podrán ser capturados por varios micrófonos colocados a una distancia relativamente lejana al escenario. Un ingeniero de sonido podrá mezclar la salida por múltiples canales deseada utilizando una combinación de todos los canales de micrófonos disponibles. Sin embargo, esta técnica de grabación implica una gran configuración de grabación y mezcla artesanal de los canales grabados, no siempre posible en la práctica.

40

45

Los sistemas convencionales para la grabación y reproducción de audio espacial basados en la codificación de audio direccional (DirAC, por su sigla en inglés), como se describe en T. Lokki, J. Merimaa, V. Pulkki: Método para Reproducir Impresión espacial Natural o Modificada en Escuchas por Múltiples canales, Patente de EEUU 7.787.638 B2, 31 de Agosto de 2010 y V. Pulkki: Reproducción de Sonido Espacial con Codificación de audio direccional. J. Soc. Ing. Audio, Vol. 55, No. 6, pp. 503-516, 2007, recae en un simple modelo global para el campo de sonido. En consecuencia, sufren de inconvenientes sistemáticos, que limitan la calidad de sonido a lograr y la experiencia en la práctica.

50

55

El documento US2011/0081024 muestra un aparato para generar una pluralidad de transmisiones de audio desde una señal de audio de entrada, donde cada transmisión de audio depende de un segmento correspondiente del espacio de grabación.

60

Un problema general de las soluciones conocidas radica en su relativa complejidad y degradación de la calidad de sonido espacial.

En consecuencia, un objetivo de la presente invención es proveer un mejor concepto del procesamiento espacial de audio paramétrico que permite una mejor calidad, grabación y reproducción de sonido espacial más realista y configuraciones de micrófonos simples y compactas.

65

5 Descripción de la Invención

Este objetivo se logra con un aparato de acuerdo con la reivindicación 1, un aparato de acuerdo con la reivindicación 10, un método de acuerdo con la reivindicación 11, un método de acuerdo con la reivindicación 12, un programa de computación de acuerdo con la reivindicación 13 o un programa de computación de acuerdo con la reivindicación 14.

De acuerdo con una realización de la presente invención, un aparato para generar una pluralidad de transmisiones de audio paramétricas para una señal espacial de audio de entrada que se obtiene de una grabación en un espacio de grabación comprende un segmentador y un generador. El segmentador está configurado para proveer al menos dos señales de audio segmentales de entrada de la señal espacial de audio de entrada. En este caso, las dos al menos señales segmentales de entrada se asocian con los correspondientes segmentos del espacio de grabación. El generador está configurado para generar la transmisión de audio paramétrica para cada de al menos dos señales de audio segmentales de entrada y obtener la pluralidad de transmisiones de audio paramétricas.

La idea básica de la presente invención radica en que el procesamiento espacial de audio paramétrico podrá lograrse si al menos dos señales de audio segmentales de entrada son emanadas de la señal espacial de audio de entrada, donde las al menos dos señales de audio segmentales de entrada se asocian con correspondientes segmentos del espacio de grabación y si una transmisión de audio paramétrica se genera para cada una de las al menos dos señales de audio segmentales de entrada para obtener la pluralidad de transmisiones de audio paramétricas. Se logra así una mayor calidad, una grabación y reproducción de sonido especial más realista utilizando configuraciones de micrófonos simples y compactas.

De acuerdo con otra realización, el segmentador está configurado para utilizar un patrón de direccionalidad para cada segmento del espacio de grabación. En este caso, el patrón de direccionalidad indica una dirección de las al menos dos señales de audio segmentales de audio. Al utilizar el patrón de direcciones, es posible obtener una mayor correspondencia de modelo del campo de sonido observado, especialmente en escenas de sonido complejas.

De acuerdo con otra realización, el generador está configurado para obtener la pluralidad de transmisiones de audio paramétrica, donde la pluralidad de transmisiones de audio paramétrica comprende un componente de las al menos dos señales de audio segmentales de entrada y una información espacial paramétrica correspondiente. Por ejemplo, la información espacial paramétrica de cada transmisión de audio paramétrica comprende un parámetro de dirección de llegada (DOA, por sus siglas en inglés) y/o un parámetro de difusión. Al proveer los parámetros DOA y/o parámetros de difusión, resulta posible describir el campo de sonido observado en un dominio de representación de señal paramétrica.

De acuerdo con otra realización, un aparato para generar una pluralidad de señales de altavoz desde una pluralidad de transmisiones de audio paramétricas derivadas de una señal espacial de audio de entrada grabada en un espacio de grabación comprende un suministrador y un combinador. El suministrador se configura para proveer una pluralidad de señales de altavoz segmentales de entrada desde una pluralidad de transmisiones de audio paramétricas. En este caso, las señales de altavoz segmentales de entrada se asocian con correspondientes segmentos del espacio de grabación. El combinador está configurado para combinar las señales de altavoz segmentales de entrada para obtener la pluralidad de señales de altavoz.

Otras realizaciones de la presente invención proveen métodos para generar una pluralidad de transmisiones de audio paramétricas y generar una pluralidad de señales de altavoz.

Breve Descripción de las Figuras

A continuación se explican las realizaciones de la presente invención con referencia a los dibujos adjuntos, donde:

La Fig. 1 muestra un diagrama en bloque de una realización de un aparato para generar una pluralidad de transmisiones de audio paramétricas desde una grabación de señal espacial de audio de entrada en un espacio de grabación con un segmentador y un generador;

La Fig. 2 muestra una ilustración esquemática del segmentador de la realización del aparato de acuerdo con la Fig. 1 basado en una operación de mezcla o matizaje;

La Fig. 3 muestra una ilustración esquemática del segmentador de la realización del aparato de acuerdo con la Fig. 1 que usa un patrón de direccionalidad;

La Fig. 4 muestra una ilustración esquemática del generador de la realización del aparato de acuerdo con la Fig. 1 basado en un análisis espacial paramétrico;

5 La Fig. 5 muestra un diagrama en bloque de una realización de un aparato para generar una pluralidad de señales de altavoz desde una pluralidad de transmisiones de audio paramétricas con un suministrador y un combinador;

10 La Fig. 6 muestra una ilustración esquemática de ejemplos de segmentos de un espacio de grabación, cada uno representa un sub grupo de direcciones dentro de un plano en dos dimensiones (2D) o dentro de un espacio en tres dimensiones (3D);

Fig. 7 muestra una ilustración esquemática de un ejemplo de computación de señal de altavoz para dos segmentos o sectores de un espacio de grabación;

15 La Fig. 8 muestra una ilustración esquemática de un ejemplo computación de señal de altavoz para dos segmentos o sectores de un espacio de grabación utilizando señales de entrada con formato B se segundo orden;

20 La Fig. 9 muestra una ilustración esquemática de un ejemplo computación de señal de altavoz para dos segmentos o sectores de un espacio de grabación que incluye una modificación de señal en un dominio de representación de señal paramétrica;

La Fig. 10 muestra una ilustración esquemática de un ejemplo de patrones polares de señales de audio segmentales de entrada que provee el segmentador de la realización del aparato de acuerdo con la Fig. 1;

25 La Fig.11 muestra una ilustración esquemática de un ejemplo de configuración de micrófono para realizar un campo de sonido grabación; y

30 La Fig. 12 muestra una ilustración esquemática de un ejemplo de serie circular de micrófonos omnidireccionales para obtener señales de micrófono de mayor orden.

Descripción Detallada de las Realizaciones

35 Antes de analizar la presente invención en detalle teniendo en cuenta los dibujos, cabe señalar que en las figuras los elementos idénticos, los elementos con igual función o mismo efecto poseen los mismos números de referencia para que la descripción de los elementos and y su funcionalidad ilustrada en las diferentes realizaciones pueda ser intercambiada mutuamente o aplicada entre sí en diferentes realizaciones.

40 La Fig. 1 muestra un diagrama en bloque de una realización de un aparato 100 para generar una pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i) desde una señal espacial de audio de entrada 105 que se obtiene de una grabación en un espacio de grabación con un segmentador 110 y un generador 120. Por ejemplo, la señal espacial de audio de entrada 105 comprende una señal omnidireccional W y una pluralidad de diferentes señales direccionales X, Y, Z, U, V (o X, Y, U, V). En la Fig. 1, el aparato 100 comprende un segmentador 110 y un generador 120. Por ejemplo, el segmentador 110 está configurado para proveer al menos dos señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i) de la señal omnidireccional W y la pluralidad de diferentes señales direccionales X, Y, Z, U, V de la señal espacial de audio de entrada 105, donde las al menos dos señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i) se asocian con los correspondientes segmentos Seg_i del espacio de grabación. Asimismo, el generador 120 podrá generar una transmisión de audio paramétrica para cada una de las al menos dos señales de audio del segmentador de entrada 115 (W_i, X_i, Y_i, Z_i) y obtener la pluralidad de transmisiones de audio paramétrica 125 (θ_i, Ψ_i, W_i).

50 Con el aparato 100 para generar la pluralidad de transmisiones de audio paramétricas 125, es posible evitar una degradación de la calidad de sonido espacial y evitar las configuraciones relativamente complejas del micrófono. En consecuencia, la realización del aparato 100 de acuerdo con la Fig. 1 permite una mayor calidad, una grabación de sonido espacial más realista utilizando configuraciones del micrófono relativamente simples y compactas.

55 En realizaciones, los segmentos Seg_i del espacio de grabación representan un sub grupo de direcciones dentro de un plano en dos dimensiones (2D) o dentro de un espacio en tres dimensiones (3D).

60 En realizaciones, los segmentos Seg_i del espacio de grabación se caracterizan por una medida de direccionalidad asociada.

De acuerdo con realizaciones, el aparato 100 está configurado para realizar una grabación de campo de sonido y obtener la señal espacial de audio de entrada 105. Por ejemplo, el segmentador 110 está configurado para dividir un rango de interés de ángulo completo en los segmentos Seg_i del espacio de grabación. Asimismo, los segmentos

Seg_i del espacio de grabación podrán cubrir un rango de ángulo reducido comparado con el rango de interés de ángulo completo.

5 La Fig. 2 muestra una ilustración esquemática del segmentador 110 de la realización del aparato 100 de acuerdo con la Fig. 1 basada en una operación de mezcla (o matrizaje). En la Fig. 2, el segmentador 110 está configurado para generar las al menos dos señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i) desde la señal omnidireccional W y la pluralidad de diferentes señales direccionales X, Y, Z, U, V utilizando a operación de mezcla o matrizaje que depende de los segmentos Seg_i del espacio de grabación. Con el segmentador 110 de la Fig. 2, es posible mapear la señal omnidireccional W y la pluralidad de diferentes señales direccionales X, Y, Z, U, V que forman la señal espacial de audio de entrada 105 a las al menos dos señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i) utilizando una operación de mezcla o matrizaje predefinida. Esta predefinida operación de mezcla o matrizaje depende de los segmentos Seg_i del espacio de grabación y sustancialmente podrá usarse para bifurcar las al menos dos señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i) de la señal espacial de audio de entrada 105. La bifurcación de las al menos dos señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i) por el segmentador 110 basada en la operación de mezcla o matrizaje sustancialmente permite lograr las ventajas anteriores en oposición a un modelo global simple para el campo de sonido.

20 La Fig. 3 muestra una ilustración esquemática del segmentador 110 de la realización del aparato 100 de acuerdo con la Fig. 1 utilizando un patrón de direccionalidad (deseado o predeterminado) 305, q_i(ϑ). En la Fig. 3, el segmentador 110 está configurado para usar un patrón de direccionalidad 305, q_i(ϑ) para cada segmento Seg_i del espacio de grabación. Asimismo, el patrón de direccionalidad 305, q_i(ϑ), puede indicar una dirección de las al menos dos señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i).

25 En realizaciones, el patrón de direccionalidad 305, q_i(ϑ), está dado por:

$$q_i(\vartheta) = a + b \cos(\vartheta + \Theta_i) \quad (1)$$

30 donde a y b denotan multiplicadores que pueden modificarse para obtener un deseado patrón de direcciones y donde ϑ denota un ángulo azimutal y Θ_i indica una dirección preferida del segmento de orden i del espacio de grabación. Por ejemplo, a yace en un rango de 0 a 1 y b en un rango de f -1 a 1.

Una opción útil de multiplicadores a, b puede ser a=0,5 y b=0,5, resultando en el siguiente patrón de direccionalidad:

$$35 \quad q_i(\vartheta) = 0,5 + 0,5 \cos(\vartheta + \Theta_i) \quad (1a)$$

40 Con el segmentador 110 de la Fig. 3, es posible obtener las al menos dos señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i) asociadas con los correspondientes segmentos Seg_i del espacio de grabación con un predeterminado patrón de direccionalidad 305, q_i(ϑ), respectivamente. Cabe señalar en este caso que el uso del patrón de direccionalidad 305, q_i(ϑ), para cada segmento Seg_i del espacio de grabación permite mejorar la calidad de sonido espacial que se obtiene con el aparato 100.

45 La Fig. 4 muestra una ilustración esquemática del generador 120 de la realización del aparato 100 de acuerdo con la Fig. 1 basado en un análisis espacial paramétrico. En el ejemplo de la Fig. 4, el generador 120 está configurado para obtener la pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i). Asimismo, la pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i) podrán comprender un componente W_i de las al menos dos señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i) y una información espacial paramétrica correspondiente θ_i, Ψ_i.

50 En realizaciones, el generador 120 podrá realizar un análisis espacial paramétrico para cada una de las al menos dos señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i) y obtener la correspondiente información espacial paramétrica θ_i, Ψ_i.

55 En realizaciones, la información espacial paramétrica θ_i, Ψ_i de cada transmisión de audio paramétrica 125 (θ_i, Ψ_i, W_i) comprende un parámetro de dirección de llegada (DOA) θ_i y/o un parámetro de difusión Ψ_i.

60 En realizaciones, el parámetro de dirección de llegada (DOA) θ_i y el un parámetro de difusión Ψ_i del generador 120 de la Fig. 4 podrán formar parámetros DirAC para un procesamiento de señal espacial de audio paramétrica. Por ejemplo, el generador 120 está configurado para generar los parámetros DirAC (ej. el parámetro DOA θ_i y el parámetros de difusión Ψ_i) utilizando una representación de frecuencia de tiempo de las al menos dos señales de audio segmentales de entrada 115.

La Fig. 5 muestra un diagrama en bloque de una realización de un aparato 500 para generar una pluralidad de señales de altavoz 525 (L₁, L₂, ...) de una pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i) con un

5 suministrador 510 y un combinador 520. En la realización de la Fig. 5, la pluralidad de transmisiones de audio paramétrica 125 (θ_i, Ψ_i, W_i) podrá derivar de una señal espacial de audio de entrada (ej. señal espacial de audio de entrada 105 de la realización de la Fig. 1) grabada en un espacio de grabación. En la Fig. 5, el aparato 500 comprende un suministrador 510 y un combinador 520. Por ejemplo, el suministrador 510 está configurado para proveer una pluralidad de señales de altavoz segmentales de entrada 515 de la pluralidad de transmisiones de audio paramétrica 125 (θ_i, Ψ_i, W_i), donde las señales de altavoz segmentales de entrada 515 se asocian con los correspondientes segmentos (Seg_i) del espacio de grabación. Asimismo, el combinador 520 está configurado para combinar las señales de altavoz segmentales de entrada 515 para obtener la pluralidad de señales de altavoz 525 (L_1, L_2, \dots).

10 Con el aparato 500 de la Fig. 5, es posible generar la pluralidad de señales de altavoz 525 (L_1, L_2, \dots) de la pluralidad de transmisiones de audio paramétrica 125 (θ_i, Ψ_i, W_i), donde la transmisión de audio paramétrica 125 (θ_i, Ψ_i, W_i) podrá transmitirse del aparato 100 de la Fig. 1. Asimismo, el aparato 500 de la Fig. 5 permite lograr una mayor calidad, una reproducción de sonido espacial más realista utilizando transmisiones de audio paramétricas derivadas de configuraciones de micrófonos relativamente simples y compactas.

15 En realizaciones, el suministrador 510 está configurado para recibir la pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i). Por ejemplo, la pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i) comprende un componente de audio segmental W_i y una información espacial paramétrica correspondiente θ_i, Ψ_i . Asimismo, el suministrador 510 podrá suministrar cada componente de audio segmental W_i utilizando la información espacial paramétrica correspondiente 505 (θ_i, Ψ_i) para obtener la pluralidad de señales de altavoz segmentales de entrada 515.

20 La Fig. 6 muestra una ilustración esquemática 600 de ejemplos de segmentos Seg_i ($i = 1, 2, 3, 4$) 610, 620, 630, 640 de un espacio de grabación. En la ilustración esquemática 600 de la Fig. 6, los ejemplos de segmentos 610, 620, 630, 640 del espacio de grabación representan un sub grupo de direcciones dentro de un plano de dos dimensiones (2D). Además, los segmentos Seg_i del espacio de grabación representan un sub grupo de direcciones dentro de un espacio de tres dimensiones (3D). Por ejemplo, los segmentos Seg_i representan los sub grupos de direcciones dentro del espacio en tres dimensiones (3D) podrán ser similares a los segmentos 610, 620, 630, 640 de la Fig. 6. De acuerdo con la ilustración esquemática 600 de la Fig. 6, se muestran cuatro ejemplos de segmentos 610, 620, 630, 640 del 100 de la Fig. 1. Sin embargo, es posible usar un diferente número de segmentos Seg_i ($i = 1, 2, \dots, n$, donde i es un índice entero, y n denota el número de segmentos). Los ejemplos de segmentos 610, 620, 630, 640 podrán estar representados en un sistema de coordenadas polares (ver por ej. Fig. 6). Para el espacio en tres dimensiones (3D), los segmentos Seg_i podrán de modo similar estar representados en un sistema de coordenadas esféricas.

25 En realizaciones, el segmentador 110 de la Fig. 1 utiliza los segmentos Seg_i (ej. ejemplos de segmentos 610, 620, 630, 640 de la Fig. 6) para proveer las al menos dos señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i). Al utilizar los segmentos (o sectores), es posible realizar un modelo paramétrico basado en segmentos (o basado en sectores) del campo de sonido. De esta manera se logra una mejor calidad de grabación y reproducción espacial de sonido con una configuración de micrófono relativamente compacta.

30 La Fig. 7 muestra una ilustración esquemática 700 de un ejemplo de computación de señal de altavoz para dos segmentos o sectores de un espacio de grabación. En la ilustración esquemática 700 de la Fig. 7, se representan ejemplos de la realización del aparato 100 para generar la pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i) y la realización del aparato 500 para generar la pluralidad de señales de altavoz 525 (L_1, L_2, \dots). En la ilustración esquemática 700 de la Fig. 7, el segmentador 110 podrá recibir la señal espacial de audio de entrada 105 (ej. señal de micrófono). Asimismo, el segmentador 110 podrá configurarse para proveer las al menos dos señales de audio segmentales de entrada 115 (ej. señales de micrófono segmentales 715-1 de un primer segmento y señales de micrófono segmentales 715-2 de un segundo segmento). El generador 120 podrá comprender un primer bloque de análisis especial paramétrico 720-1 y un segundo bloque de análisis especial paramétrico 720-2. Asimismo, el generador 120 podrá generar la transmisión de audio paramétrica para cada una de las al menos dos señales de audio segmentales de entrada 115. En la salida de la realización del aparato 100, se obtiene la pluralidad de transmisiones de audio paramétricas 125. Por ejemplo, el primer bloque de análisis especial paramétrico 720-1 emitirá una primera transmisión de audio paramétrica 725-1 de un primer segmento, y el segundo bloque de análisis especial paramétrico 720-2 emitirá una segunda transmisión de audio paramétrica 725-2 de un segundo segmento. Asimismo, la primera transmisión de audio paramétrica 725-1 del primer bloque de análisis especial paramétrico 720-1 podrá comprender información espacial paramétrica (ej. θ_1, Ψ_1) de un primer segmento y una o más señales de audio segmentales (ej. W_1) del primer segmento, y la segunda transmisión de audio paramétrica 725-2 del segundo bloque de análisis especial paramétrico 720-2 podrá comprender información espacial paramétrica (ej. θ_2, Ψ_2) de un segundo segmento y una o más señales de audio segmentales (ej. W_2) del segundo segmento. La realización del aparato 100 podrá transmitir la pluralidad de transmisiones de audio paramétricas 125. En la ilustración esquemática 700 de la Fig. 7, la realización del aparato 500 podrá recibir la pluralidad de transmisiones de audio paramétricas 125 de la realización del aparato 100. El suministrador 510 podrá comprender una primera unidad de suministro 730-1 y una segunda unidad de suministro 730-2. Asimismo, el suministrador 510 podrá

proveer la pluralidad of señales de altavoz segmentales de entrada 515 de la pluralidad de transmisiones de audio paramétricas recibidas 125. Por ejemplo, la primer unidad de suministro 730-1 podrá proveer señales de altavoz segmentales de entrada 735-1 de un primer segmento de la primera transmisión de audio paramétrica 725-1 del primer segmento, y la segunda unidad de suministro 730-2 podrá proveer señales de altavoz segmentales de entrada 735-2 de un segundo segmento de la segunda transmisión de audio paramétrica 725-2 del segundo segmento. Asimismo, el combinador 520 podrá combinar las señales de altavoz segmentales de entrada 515 para obtener la pluralidad de señales de altavoz 525 (ej. L_1, L_2, \dots).

La realización de la Fig. 7 esencialmente representa un concepto de grabación y reproducción especial de audio de mayor calidad utilizando un modelo paramétrico basado en segmentos (o basado en sectores) del campo de sonido, que permite grabar además escenas espaciales de audio complejas con x audio espacial con una configuración de micrófono relativamente compacta.

La Fig. 8 muestra una ilustración esquemática 800 de un ejemplo de computación de señal de altavoz para dos segmentos o sectores de un espacio de grabación utilizando señales de entrada de formato B de segundo orden 105. El ejemplo de computación de señal de altavoz que se ilustra en forma esquemática en la Fig. 8 esencialmente corresponde al ejemplo de computación de la señal de altavoz que se ilustra en forma esquemática en la Fig. 7. En la ilustración esquemática de la Fig. 8, se representan ejemplos de la realización del aparato 100 para generar la pluralidad de transmisiones de audio paramétricas 125 y la realización del aparato 500 para generar la pluralidad de señales de altavoz 525. En la Fig. 8, la realización del aparato 100 podrá recibir la señal espacial de audio de entrada 105 (ej. canales de micrófono con formato B como $[W, X, Y, U, V]$). En este caso, cabe notar que las U, V en la Fig. 8 son componentes con formato B de segundo orden. El segmentador 110 denotado como "matrizador" podrá generar las al menos dos señales de audio segmentales de entrada 115 de la señal omnidireccional y la pluralidad de diferentes señales de direccionalidad utilizando una operación de mezcla o matrizaje que depende de los segmentos Seg_i del espacio de grabación. Por ejemplo, las al menos dos señales de audio segmentales de entrada 115 podrán comprender la señal de micrófono segmental 715-1 de un primer segmento (ej. $[W_1, X_1, Y_1]$) y la señal de micrófono segmental 715-2 de un segundo segmento (ej. $[W_2, X_2, Y_2]$). Asimismo, el generador 120 podrá comprender un primer bloque de análisis de direccionalidad y difusión 720-1 y un segundo bloque de análisis de direccionalidad y difusión 720-2. El primer y segundo bloque de análisis de direccionalidad y difusión 720-1, 720-2 de la Fig. 8 esencialmente corresponden al primer y segundo bloque de análisis espacial paramétrico 720-1, 720-2 de la Fig. 7. El generador 120 podrá generar una transmisión de audio paramétrica para las al menos dos señales de audio segmentales de entrada 115 y obtener la pluralidad de transmisiones de audio paramétricas 125. Por ejemplo, el generador 120 podrá realizar un análisis espacial en las señales de micrófono segmentales 715-1 del primer segmento utilizando el primer bloque de análisis de direccionalidad y difusión 720-1 y para extraer un primer componente (ej. señal de audio segmental W_1) de las señales de micrófono segmentales 715-1 del primer segmento para obtener la primera transmisión de audio paramétrica 725-1 del primer segmento. Asimismo, el generador 120 podrá realizar un análisis espacial en las señales de micrófono segmentales 715-2 del segundo segmento y para extraer un segundo componente (ej. señal de audio segmental W_2) de las señales de micrófono segmentales 715-2 del segundo segmento utilizando el segundo bloque de análisis de direccionalidad y difusión 720-2 para obtener la segunda transmisión de audio paramétrica 725-2 del segundo segmento. Por ejemplo, la primera transmisión de audio paramétrica 725-1 del primer segmento podrá comprender información espacial paramétrica del primer segmento que comprende un primer parámetro de dirección de llegada (DOA) θ_1 y un primer parámetro de difusión Ψ_1 y un primer componente extraído W_1 , y la segunda transmisión de audio paramétrica 725-2 del segundo segmento podrá comprender información espacial paramétrica del segundo segmento que comprende un segundo parámetro de dirección de llegada (DOA) θ_2 y un segundo parámetro de difusión Ψ_2 y un segundo componente extraído W_2 . La realización del aparato 100 podrá transmitir la pluralidad de transmisiones de audio paramétricas 125.

En la ilustración esquemática 800 de la Fig. 8, la realización del aparato 500 para generar la pluralidad de señales de altavoz 525 podrá recibir la pluralidad de transmisiones de audio paramétricas 125 transmitidas de la realización del aparato 100. En la ilustración esquemática 800 de la Fig. 8, el suministrador 510 comprende la primera unidad de suministro 730-1 y la segunda unidad de suministro 730-2. Por ejemplo, la primera unidad de suministro 730-1 comprende un primer multiplicador 802 y un segundo multiplicador 804. El primer multiplicador 802 de la primera unidad de suministro 730-1 podrá aplicar un primer factor de ponderación 803 (ej. $\sqrt{1-\Psi}$) a la señal de audio segmental W_1 de la primera transmisión de audio paramétrica 725-1 del primer segmento para obtener una sub transmisión de sonido directa 810 por la primera unidad de suministro 730-1, y el segundo multiplicador 804 de la primera unidad de suministro 730-1 podrá aplicar un segundo factor de ponderación 805 (ej. $\sqrt{\Psi}$) a la señal de audio segmental W_1 de la primera transmisión de audio paramétrica 725-1 del primer segmento para obtener una sub transmisión difusa 812 por la primera unidad de suministro 730-1. Asimismo, la segunda unidad de suministro 730-2 podrá comprender un primer multiplicador 806 y un segundo multiplicador 808. Por ejemplo, el primer multiplicador 806 de la segunda unidad de suministro 730-2 podrá aplicar un primer factor de ponderación 807 (ej. $\sqrt{1-\Psi}$) a la señal de audio segmental W_2 de la segunda transmisión de audio paramétrica 725-2 del segundo segmento para obtener una transmisión de sonido directa 814 por la segunda unidad de suministro 730-2, y el

segundo multiplicador 808 de la segunda unidad de suministro 730-2 podrá aplicar un segundo factor de ponderación 809 (ej. $\sqrt{\Psi}$) a la señal de audio segmental W_2 de la segunda transmisión de audio paramétrica 725-2 del segundo segmento para obtener una sub transmisión difusa 816 por la segunda unidad de suministro 730-2. En realizaciones, el primer y segundo factor de ponderación 803, 805, 807, 809 de la primera y segunda unidad de suministro 730-1, 730-2 derivan de los parámetros de difusión correspondientes Ψ_i . De acuerdo con realizaciones, la primera unidad de suministro 730-1 podrá comprender factor de ganancia del multiplicador 811, bloques de procesamiento de decorrelación 813 y unidades de combinación 832, y la segunda unidad de suministro 730-2 podrá comprender factor de ganancia del multiplicador 815, bloques de procesamiento de decorrelación 817 y unidades de combinación 834. Por ejemplo, el factor de ganancia del multiplicador 811 de la primera unidad de suministro 730-1 podrá aplicar factores de ganancia que se obtienen de una operación de paneo en amplitud basada en un vector (VBAP, por su sigla en inglés) operación mediante bloques 822 a la salida de sub transmisión de sonido directo 810 por el primer multiplicador 802 de la primera unidad de suministro 730-1. Asimismo, los bloques de procesamiento de decorrelación 813 de la primera unidad de suministro 730-1 podrán aplicar una operación de decorrelación/ganancia a la sub transmisión difusa 812 en la salida del segundo multiplicador 804 de la primera unidad de suministro 730-1. Además, las unidades de combinación 832 de la primera unidad de suministro 730-1 podrán combinar las señales que se obtienen del factor de ganancia del multiplicador 811 y los bloques de procesamiento de decorrelación 813 para obtener las señales de altavoz segmentales 735-1 del primer segmento. Por ejemplo, el factor de ganancia del multiplicador 815 de la segunda unidad de suministro 730-2 podrá aplicar factores de ganancia que se obtienen de una operación de paneo en amplitud basada en un vector (VBAP) mediante bloques 824 a la salida de sub transmisión de sonido directo 814 mediante el primer multiplicador 806 de la segunda unidad de suministro 730-2. Asimismo, los bloques de procesamiento de decorrelación 817 de la segunda unidad de suministro 730-2 podrán aplicar una operación de decorrelación/ganancia a la sub transmisión difusa 816 en la salida del segundo multiplicador 808 de la segunda unidad de suministro 730-2. Además, las unidades de combinación 834 de la segunda unidad de suministro 730-2 podrán combinar señales obtenidas del factor de ganancia del multiplicador 815 y bloques de procesamiento de decorrelación 817 para obtener las señales de altavoz segmentales 735-2 del segundo segmento.

En realizaciones, la operación de paneo en amplitud basada en un vector (VBAP) mediante bloques 822, 824 de la primera y segunda unidad de suministro 730-1, 730-2 depende de los correspondientes parámetros de dirección de llegada (DOA, por su sigla en inglés) θ . En la Fig. 8, el combinador 520 podrá combinar señales de altavoz segmentales de entrada 515 para obtener la pluralidad de señales de altavoz 525 (ej. L_1, L_2, \dots). En la Fig. 8, el combinador 520 podrá comprender una primera unidad de síntesis 842 y una segunda unidad de síntesis 844. Por ejemplo, la primera unidad de síntesis 842 sintetiza una primera señal de altavoz segmental 735-1 del primer segmento y una primera señal de altavoz segmental 735-2 del segundo segmento para obtener una primera señal de altavoz 843. Además, la segunda unidad de síntesis 844 podrá sintetizar una segunda señal de altavoz segmental 735-1 del primer segmento y una segunda señal de altavoz segmental 735-2 del segundo segmento para obtener una segunda señal de altavoz 845. La primera y segunda señal de altavoz 843, 845 podrán formar la pluralidad de señales de altavoz 525. En la realización de la Fig. 8, debe notarse que para cada segmento, podrán generarse señales de altavoz potenciales para todos los altavoces de la reproducción.

La Fig. 9 muestra una ilustración esquemática 900 de un ejemplo de computación de señal de altavoz para dos segmentos o sectores de un espacio de grabación que incluye una modificación de señal en un dominio de representación de señal paramétrica. El de computación de señal de altavoz en la ilustración esquemática 900 de la Fig. 9 esencialmente corresponde al ejemplo de computación de señal de altavoz en la ilustración esquemática 700 de la Fig. 7. Sin embargo, el ejemplo de computación de señal de altavoz en la ilustración esquemática 900 de la Fig. 9 incluye otra modificación de señal.

En la ilustración esquemática 900 de la Fig. 9, el aparato 100 comprende el segmentador 110 y el generador 120 para obtener la pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i). Asimismo, el aparato 500 comprende el suministrador 510 y combinador 520 para obtener la pluralidad de señales de altavoz 525.

Por ejemplo, el aparato 100 podrá comprender un modificador 910 para modificar la pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i) en un dominio de representación de señal paramétrica. Asimismo, el modificador 910 podrá modificar al menos la transmisión de audio paramétrica 125 (θ_i, Ψ_i, W_i) utilizando el correspondiente parámetro de control de modificación 905. De este modo, se obtiene una primera transmisión de audio paramétrica modificada 916 de un primer segmento y una segunda transmisión de audio paramétrica modificada 918 de un segundo segmento. La primera y segunda transmisión de audio paramétrica modificada 916, 918 podrán formar una pluralidad de transmisiones de audio paramétricas modificadas 915. En realizaciones, el aparato 100 podrá transmitir la pluralidad de transmisiones de audio paramétricas modificadas 915. Además, el aparato 500 podrá recibir la pluralidad de transmisiones de audio paramétricas modificadas 915 transmitidas desde el aparato 100.

Con el ejemplo de computación de señal de altavoz de acuerdo con la Fig. 9, es posible lograr un esquema de grabación y reproducción espacial de audio más flexible. En particular, es posible obtener mayor calidad en las señales de salida al aplicar modificaciones en el dominio paramétrico. Al segmentar las señales de entrada antes

de generar la pluralidad de representaciones de audio paramétricas (transmisiones), se obtiene una selección espacial mayor que permite tratar diferentes componentes del campo de sonido capturado en forma diferente.

5 La Fig. 10 muestra una ilustración esquemática 1000 de ejemplos de patrones polares de señales de audio segmentales de entrada 115 (ej. W_i , X_i , Y_i) del segmentador 110 de la realización del aparato 100 para generar la pluralidad de transmisiones de audio paramétricas 125 (θ_i , Ψ_i , W_i) de acuerdo con la Fig. 1. En la ilustración esquemática 1000 de la Fig. 10, el ejemplo de señales de audio segmentales de entrada 115 se visualiza en un sistema de coordenadas polares respectivas para el plano en dos dimensiones (2D). De modo similar, el ejemplo de señales de audio segmentales de entrada 115 puede ser visualizado en un sistema de coordenadas esférico respectivo para el espacio en tres dimensiones (3D). La ilustración esquemática 1000 de la Fig. 10 representa la primera respuesta direccional 1010 para una primera señal de audio segmental de entrada (ej. señal omnidireccional W_i), una segunda respuesta direccional 1020 de una segunda señal de audio segmental de entrada (ej. primera señal direccional X_i) y una tercera respuesta direccional 1030 de una tercera señal de audio segmental de entrada (ej. segunda señal direccional Y_i). Asimismo, una cuarta respuesta direccional 1022 con signo opuesto en comparación con la segunda respuesta direccional 1020 y una quinta respuesta direccional 1032 con signo opuesto en comparación con la tercera respuesta direccional 1030 se representan en la ilustración esquemática 1000 de la Fig. 10. De este modo, diferentes respuestas direccionales 1010, 1020, 1030, 1022, 1032 (patrones polares) podrán usarse para las señales de audio segmentales de entrada 115 por el segmentador 110. Cabe señalar que las señales de audio segmentales de entrada 115 podrán depender del tiempo y frecuencia, es decir $W_i = W_i(m, k)$, $X_i = X_i(m, k)$, y $Y_i = Y_i(m, k)$, donde (m, k) son índices que indican un mosaico de frecuencia de tiempo en una representación de señal espacial de audio.

En este contexto, cabe notar que la Fig. 10 representa los diagramas polares para un solo grupo de señales de entrada, es decir las señales 115 para un solo sector i (ej. $[W_i, X_i, Y_i]$). Asimismo, las partes positivas y negativas del diagrama polar representan juntas el diagrama polar de una señal, respectivamente (por ejemplo, las partes 1020 y 1022 juntas muestran el diagrama polar de la señal X_i , y las partes 1030 y 1032 muestran el diagrama polar de la señal Y_i).

La Fig. 11 muestra una ilustración esquemática 1100 de un ejemplo de configuración de micrófono 1110 para realizar una grabación de campo de sonido. En la ilustración esquemática 1100 de la Fig. 11, la configuración de micrófono 1110 podrá comprender múltiples series lineales de micrófonos direccionales 1112, 1114, 1116. La ilustración esquemática 1100 de la Fig. 11 representa como un espacio de observación en dos dimensiones (2D) puede dividirse en diferentes segmentos o sectores 1101, 1102, 1103 (ej. Seg_i , $i = 1, 2, 3$) del espacio de grabación. En este caso, los segmentos 1101, 1102, 1103 de la Fig. 11 corresponden a los segmentos Seg_i que se representan en la Fig. 6. Similarmente, el ejemplo de configuración de micrófono 1110 podrá usarse además en un espacio de observación en tres dimensiones (3D), donde el espacio de observación en tres dimensiones (3D) puede dividirse en segmentos o sectores para la configuración de micrófono dada. En realizaciones, el ejemplo de configuración de micrófono 1110 en la ilustración esquemática 1100 de la Fig. 11 podrá usarse para proveer la señal espacial de audio de entrada 105 para la realización del aparato 100 de acuerdo con la Fig. 1. Por ejemplo, las múltiples series lineales de micrófonos direccionales 1112, 1114, 1116 de la configuración de micrófono 1110 podrá proveer diferentes señales direccionales para la señal espacial de audio de entrada 105. Con el uso del ejemplo de configuración de micrófono 1110 de la Fig. 11, es posible optimizar la calidad de grabación espacial de audio utilizando el modelo paramétrico basado en segmentos (o basado en sectores) del campo de sonido.

En las previas realizaciones, el aparato 100 y aparato 500 podrán operar en dominio de frecuencia de tiempo.

En síntesis, las realizaciones de la presente invención se relacionan con el campo de la grabación y reproducción espacial de audio de mayor calidad. El uso de un modelo paramétrico basado en segmentos o basado en sectores permite grabar escenas espaciales de audio complejas con una configuración de micrófono relativamente compacta. En oposición a un simple modelo global del campo de sonido asumido por los métodos de la técnica actual, la información paramétrica podrá determinarse para un número de segmentos donde todo el espacio de observación está dividido. En consecuencia, el suministro de una configuración de altavoz casi arbitraria podrá realizarse teniendo en cuenta la información paramétrica que se muestra con los canales de audio grabados.

De acuerdo con realizaciones, para una grabación de campo de sonido plano en dos dimensiones (2D), todo el rango de interés del ángulo azimutal podrá dividirse en múltiples sectores o segmentos que cubren un reducido rango de ángulos azimutales. En forma análoga, en el caso 3D todo el rango del ángulo sólido (azimutal y elevación) podrá dividirse en sectores o segmentos que cubren un rango menor del ángulo. Los diferentes sectores o segmentos podrán superponerse en forma parcial.

De acuerdo con realizaciones, cada sector o segmento se caracteriza por una medida direccional asociada, que puede usarse para especificar o referirse al correspondiente sector o segmento. La medida direccional podrá, por ejemplo, ser un vector que señala a (o desde) el centro del sector o segmento, o un ángulo azimutal en el caso 2D, o un grupo de un ángulo azimutal y en elevación en el caso 3D. El segmento o sector podrán denominarse como

subgrupo de direcciones dentro de un plano 2D o dentro de un espacio 3D. Para una simple presentación, los previos ejemplos se describieron para los casos 2D; sin embargo la extensión a configuraciones 3D es directa.

5 En la Fig. 6, la mediada direccional podrá definirse como un vector que, para el segmento Seg_3 , señala desde el origen, es decir, el centro con la coordenada (0, 0), a la derecha, es decir, hacia la coordenada (1, 0) en el diagrama polar, o el ángulo azimutal de 0° si, en la Fig. 6, los ángulos se cuentan desde (o en referencia a) el eje x (eje horizontal).

10 En la realización de la Fig. 1, el aparato 100 podrá recibir un número de señales de micrófono como entrada (señal espacial de audio de entrada 105). Estas señales can de micrófono, por ejemplo, provienen de una grabación real o son artificialmente generadas por una grabación estimulada en un ambiente virtual. De estas señales de micrófono, podrán determinarse correspondiente señales de micrófono segmentales (señales de audio segmentales de entrada 115), asociadas a los correspondientes segmentos (Seg_i). Las señales de micrófono segmentales presentan características especiales. Su patrón de levante direccional podrá mostrar una sensibilidad significativamente
15 aumentada dentro del sector angular asociado comparado con la sensibilidad fuera de ese sector. Un ejemplo de segmentación de un rango total azimutal 360° y los patrones de levante de las señales de micrófono segmentales asociadas se ilustran en la Fig. 6. En el ejemplo de la Fig. 6, la dirección de los micrófonos asociados con los sectores exhiben patrones cardioides que giran de acuerdo al rango angular cubierto por los correspondientes sectores. Por ejemplo, la dirección de los micrófonos asociados al sector 3 (Seg_3) que señalan a 0° señala además
20 a 0° . En este caso, cabe notar que en el diagrama polar de la Fig. 6, la dirección de la sensibilidad máxima es la dirección donde el radio de la curva representada comprende el máximo. De este modo, Seg_3 posee la mayor sensibilidad para los componentes de sonido que provienen de la derecha. En otras palabras, el segmento Seg_3 posee su dirección preferida en el ángulo azimutal de 0° (asumiendo que los ángulos se cuentan desde el eje x).

25 De acuerdo con realizaciones, para cada sector, se determina un parámetro DOA (θ_i) con un parámetro de difusión a base de un sector (Ψ_i). En una simple realización, el parámetro de difusión (Ψ_i) puede ser el mismo para todos los sectores. En principio, se puede aplicar un algoritmo de estimación preferido DOA (ej. por el generador 120). Por ejemplo, el parámetro DOA (θ_i) puede interpretarse como que refleja la dirección opuesta donde viaja la mayor parte de la energía del sonido dentro del sector considerado. En consecuencia, la difusión basada en el sector se
30 relaciona con la relación de la energía de sonido difusa y la energía de sonido total dentro del sector considerado. Debe observarse que la estimación del parámetro (realizada por el generador 120) debe realizarse con variación de tiempo y en forma individual para cada banda de frecuencia.

35 De acuerdo con realizaciones, para cada sector, podrá componerse una transmisión de audio direccional (transmisión de audio paramétrica) incluyendo la señal de micrófono segmental (W_i) y el DOA basado en sectores y los parámetros de difusión (θ_i , Ψ_i) que describen en forma predominante las propiedades espaciales de audio del campo de sonido dentro del rango angular representado por ese sector. Por ejemplo, las señales de altavoz 525 para la reproducción podrán determinarse utilizando la información direccional paramétrica (θ_i , Ψ_i) y una o más
40 señales de micrófono segmentales 125 (ej. W_i). Por ende, se puede determinar un grupo de señales de altavoz segmentales 515 para cada segmento que pueden combinarse como por el combinador 520 (ej. sintetizadas o mezcladas) para formar las señales de altavoz finales 525 de reproducción. Los componentes de sonido directo dentro de un sector podrán, por ejemplo, suministrarse como fuentes tipo punto aplicando un ejemplo de paneo en amplitud basado en vectores (como describe V. Pulkki: Posición de fuente de sonido virtual utilizando paneo en amplitud basado en vectores. J. Audio Eng. Soc., Vol. 45, pp. 456-466, 1997), mientras que el sonido difundido
45 podrá reproducirse desde varios altavoces al mismo tiempo.

El diagrama en bloque en la Fig. 7 ilustra la computación de las señales de altavoz 525 como se describe antes para el caso de dos sectores. En la Fig. 7, las flechas gruesas representan señales de audio, y las flechas delgadas representan señales paramétricas o de control. En la Fig. 7, la generación de señales de micrófono segmentales 115 por el segmentador 110, la aplicación del análisis de la señal paramétrica espacial (bloques 720-1, 720-1) para cada sector (ej. por el generador 120), la generación de las señales de altavoz segmentales 515 por el suministrador 510 y la combinación de las señales de altavoz segmentales 515 por el combinador 520 se ilustran en forma esquemática.

55 En realizaciones, el segmentador 110 podrá realizar la generación de señales de micrófono segmentales 115 desde un grupo de señales de entrada de micrófono 105. Asimismo, el generador 120 podrá realizar la aplicación del análisis de la señal paramétrica espacial para cada sector de modo que las transmisiones de audio paramétricas 725-1, 725-2 para cada sector puedan obtenerse. Por ejemplo, cada transmisión de audio paramétrica 725-1, 725-2 podrá tener al menos una señal de audio segmental (ej. W_1 , W_2 , respectivamente) e información paramétrica asociada (ej. Parámetro DOAs θ_1 , θ_2 y parámetros de difusión Ψ_1 , Ψ_2 , respectivamente). El suministrador 510 podrá
60 generar las señales de altavoz segmentales 515 para cada sector basado en las transmisiones de audio paramétricas 725-1, 725-2 generadas de los sectores particulares. El combinador 520 podrá combinar las señales de altavoz segmentales 515 para obtener las señales de altavoz finales 525.

El diagrama en bloque en la Fig. 8 ilustra la computación de señales de altavoz 525 para el ejemplo del caso de dos sectores como ejemplo de aplicación de señal de micrófono con formato B de segundo orden. En la realización de la Fig. 8, dos (grupo de) señales de micrófono segmentales 715-1 (ej. $[W_1, X_1, Y_1]$) y 715-2 (ej. $[W_2, X_2, Y_2]$) podrán generarse de un grupo de señales de micrófono d entrada 105 mediante una operación de mezcla o matrizaje (ej. por el bloque 110) como se describe antes. Para cada una de las dos señales de micrófono segmentales, se realiza un análisis de audio direccional (ej. por bloques 720-1, 720-2), produciendo las transmisiones de audio direccionales 725-1 (ej. θ_1, Ψ_1, W_1) y 725-2 (ej. θ_2, Ψ_2, W_2) para el primer sector y segundo sector, respectivamente.

En la Fig. 8, las señales de altavoz segmentales 515 podrán generarse por separado para cada sector de la siguiente manera. El componente de audio segmental W_i podrá dividirse en dos sub transmisiones complementarias 810, 812, 814, 816 ponderando con los multiplicadores 803, 805, 807, 809 derivados del parámetro de difusión Ψ_i . Una sub transmisión podrá llevar predominantemente componentes de sonido directo, y la otra sub transmisión podrá llevar predominantemente componentes de sonido difundido. Las sub transmisiones de sonido directo 810, 814 podrán suministrarse utilizando ganancias en paneo 811, 815 determinadas por el Parámetro DOA θ_i , y las sub transmisiones difusas 812, 816 podrán suministrarse en forma incoherente utilizando bloques de procesamiento de decorrelación 813, 817.

Como ejemplo de último paso, las señales de altavoz segmentales 515 podrán combinarse (ej. por el bloque 520) para obtener las señales de salida finales 525 para la reproducción del altavoz.

En la realización de la Fig. 9, debe mencionarse que los parámetros estimados (dentro de las transmisiones de audio paramétricas 125) podrán modificarse (ej. por el modificador 910) antes que las señales de altavoz reales 525 para la reproducción sean determinadas. Por ejemplo, el parámetro DOA θ_i podrá ser re mapeado para lograr una manipulación de la escena de sonido. En otros casos, las señales de audio (ej. W_i) de ciertos sectores podrá ser atenuadas antes de computar las señales de altavoz 525 si el sonido que proviene de cierta o todas las direcciones incluyendo esos sectores no son deseados. En forma análoga, los componentes de sonido difundido podrán atenuarse si el sonido principal o sólo el sonido directo debe ser suministrado. Este proceso incluyendo una modificación 910 de la transmisión de audio paramétrica 125 se ilustra en la Fig. 9 para el ejemplo de segmentación en dos segmentos.

Una realización de estimación de parámetro basado en sectores en el ejemplo del caso 2D de las realizaciones previas se describe a continuación. Se supone que las señales de micrófono usadas para la captura pueden convertirse en las denominadas señales de formato B de segundo orden. Las señales de formato B de segundo orden pueden describirse por la forma de patrón de direcciones de los correspondientes micrófonos:

$$b_w(\vartheta) = 1 \quad (2)$$

$$b_x(\vartheta) = \cos(\vartheta) \quad (3)$$

$$b_y(\vartheta) = \sin(\vartheta) \quad (4)$$

$$b_u(\vartheta) = \cos(2\vartheta) \quad (5)$$

$$b_v(\vartheta) = \sin(2\vartheta) \quad (6)$$

donde ϑ denota el ángulo azimutal. Las correspondientes señales con formato B (ej. entrada 105 de la Fig. 8) se denotan por $W(m, k)$, $X(m, k)$, $Y(m, k)$, $U(m, k)$ y $V(m, k)$, donde m y k representan un índice de tiempo y frecuencia, respectivamente. Se asume que las señal de micrófono asociada don el sector de orden i posee un patrón de direccionalidad $q_i(\vartheta)$. Se podrá determinar (ej. por el bloque 110) las otras señales de micrófono 115, $W_i(m, k)$, $X_i(m, k)$, $Y_i(m, k)$ con un patrón de direccionalidad que se expresa por

$$b_{W_i}(\vartheta) = q_i(\vartheta) \quad (7)$$

$$b_{X_i}(\vartheta) = q_i(\vartheta)\cos(\vartheta) \quad (8)$$

$$b_{Y_i}(\vartheta) = q_i(\vartheta)\sin(\vartheta) \quad (9)$$

Algunos ejemplos para el patrón de direcciones de las señales de micrófono que se describen en caso del ejemplo del patrón cardioide $q_i(\vartheta) = 0,5 + 0,5 \cos(\vartheta + \Theta_i)$ se muestran en la Fig. 10. La dirección preferida del sector de orden i depende de un ángulo azimutal Θ_i . En la Fig. 10, las líneas de guiones indican las respuestas direccionales 1022, 1032 (patrones polares) con signo opuesto comparado con las respuestas direccionales 1020, 1030 representadas con líneas sólidas.

Obsérvese que para el ejemplo del caso de $\Theta_i = 0$, las señales $W_i(m, k)$, $X_i(m, k)$, $Y_i(m, k)$ podrán determinarse de las señales con formato B de segundo orden al mezclar los componentes de entrada W, X, Y, U, V de acuerdo con

$$5 \quad W_i(m, k) = 0.5W(m, k) + 0.5X(m, k) \quad (10)$$

$$X_i(m, k) = 0.25W(m, k) + 0.5X(m, k) + 0.25U(m, k) \quad (11)$$

$$Y_i(m, k) = 0.5Y(m, k) + 0.25V(m, k) \quad (12)$$

10 Esta operación de mezcla se realiza por ej. en la Fig. 2 en el bloque de construcción 110. Obsérvese que una opción diferente de $q_i(\vartheta)$ lleva a una regla de mezcla diferente para obtener los componentes W_i, X_i, Y_i de señales con formato B de segundo orden.

15 De las señales de micrófono segmentales 115, $W_i(m, k)$, $X_i(m, k)$, $Y_i(m, k)$, se determina (ej. por el bloque 120) que el parámetro DOA θ_i asociado al sector de orden i al computar el vector de intensidad activo basado en sectores

$$\mathbf{I}_{a_i}(m, k) = -\frac{1}{2\rho_0 c} \operatorname{Re} \left\{ W_i^*(m, k) \cdot \begin{bmatrix} X_i(m, k) \\ Y_i(m, k) \end{bmatrix} \right\} \quad (13)$$

20 donde $\operatorname{Re}\{A\}$ denota la parte real del número complejo A y $*$ denota un conjugado complejo. Asimismo, ρ_0 es la densidad del aire y c es la velocidad del sonido. El cálculo de DOA deseado $\theta_i(m, k)$, por ejemplo representado por la vector de unidad $\mathbf{e}_i(m, k)$, puede obtenerse por

$$\mathbf{e}_i(m, k) = -\frac{\mathbf{I}_{a_i}(m, k)}{\|\mathbf{I}_{a_i}(m, k)\|} \quad (14)$$

25 Se puede determinar además la cantidad relacionada de energía del campo de sonido basado en sectores

$$E_i(m, k) = \frac{1}{4\rho_0 c^2} \left(|W_i(m, k)|^2 + |X_i(m, k)|^2 + |Y_i(m, k)|^2 \right) \quad (15)$$

30 El parámetro de difusión deseado $\Psi_i(m, k)$ del sector de orden i podrá determinarse por

$$\Psi_i(m, k) = g \left(1 - \frac{\|\mathbf{E}\{\mathbf{I}_{a_i}(m, k)\}\|}{c E_i(m, k)} \right) \quad (16)$$

35 donde g denota un adecuado factor de escala, $\mathbf{E}\{\}$ es el operador de expectación y $\|\cdot\|$ denota la norma del vector. Puede observarse que el parámetro de difusión $\Psi_i(m, k)$ es cero si sólo se encuentra una onda de plano presente y toma un valor positivo menor o igual a uno en caso de campos de sonidos puramente difundidos. En general, se puede definir una alternativa función de mapeo para la difusión que exhibe una conducta similar, es decir dar 0 para sonido directo sólo, y acercar 1 para un campo de sonido completamente difundido.

40 En la realización de la Fig. 11, una alternativa de realización para estimar el parámetro podrá usarse para las diferentes configuraciones de micrófonos. En la Fig. 11, se podrán usar múltiples series lineales 1112, 1114, 1116 de micrófonos direccionales. La Fig. 11 muestra un ejemplo de como dividir el espacio de observación 2D en sectores 1101, 1102, 1103 para la configuración de micrófono dada. Las señales de micrófono segmentales 115 pueden determinarse por medio de técnicas que forman haces como formación de haces de filtro y suma aplicados a cada serie de micrófono lineal 1112, 1114, 1116. La formación de haces podrá omitirse, es decir los patrones de direccionalidad de los micrófonos direccionales podrán usarse como únicos medios para obtener señales de micrófono segmentales 115 que muestran la deseada selección espacial para cada sector (Seg). El parámetro DOA θ_i dentro de cada sector puede estimarse utilizando técnicas de estimación comunes como algoritmo "ESPRIT" (se describe en R. Roy y T. Kailath: estimación ESPRIT de parámetros de señal por medio de técnicas de invariancia rotativas, IEEE Transacciones sobre Acústicas, Discurso y Procesamiento de Señal, vol. 37, nº 7, pp. 984995, Julio de 1989). El parámetro de difusión Ψ_i para cada sector podrá, por ejemplo, determinarse al evaluar la variación

temporal de los cálculos de DOA (como se describe en J. Ahonen, V. Pulkki: Estimación de difusión utilizando variación temporal de vectores de intensidad, IEEE Taller en Aplicaciones del Procesamiento de Señal en Audio y Acústica, 2009. WAS-PAA '09. , pp. 285-288, 18-21 Oct. 2009). Alternativamente, pueden emplearse relaciones conocidas de la coherencia entre diferentes micrófonos y la de sonido directo a difundido (como se describe en O. Thierngart, G. Del Galdo, E.A.P. Habets,: Estimación de relación de señal a reverberante basada en la coherencia espacial compleja entre micrófonos omnidireccionales, IEEE Conferencia Internacional en Acústica, Discurso y Procesamiento de Señal (ICASSP), 2012, pp. 309-312, 25-30 Marzo de 2012).

La Fig. 12 muestra una ilustración esquemática 1200 de ejemplo de serie circular de micrófonos omnidireccionales 1210 para obtener señales de micrófono de mayor orden (ej. señal espacial de audio de entrada 105). En la ilustración esquemática 1200 de la Fig. 12, la serie circular de micrófonos omnidireccionales micrófonos 1210 comprende, por ejemplo, 5 micrófonos equidistantes a lo largo de un círculo (línea de puntos) en un diagrama polar. En realizaciones, la serie circular de micrófonos omnidireccionales 1210 podrá usarse para obtener señales de micrófono de mayor orden (HO, por su sigla en inglés), que se describen abajo. Para computar el ejemplo de señales de micrófono de segundo orden U y V de las señales de micrófono omnidireccionales (de los micrófonos omnidireccionales 1210), al menos 5 señales de micrófono independientes deberán usarse. Esto se logra en forma elegante, ej. utilizando una Serie Circular Uniforme (UCA, por su sigla en inglés) como el ejemplo de la Fig. 12. El vector que se obtiene de señales de micrófono en cierto tiempo y frecuencia podrá, por ejemplo, transformarse con DFT (Transformada Discreta de Fourier). Las señales de micrófono W, X, Y, U y V (es decir la señal espacial de audio de entrada 105) pueden obtenerse mediante combinación lineal de coeficientes DFT. Obsérvese que los coeficientes DFT representan coeficientes de series Fourier calculados desde el vector de señales de micrófono.

Dejemos que Υ_m denote el orden generalizado de la señal de micrófono m-th, definida por el patrón de direcciones

$$\begin{aligned} \Upsilon_m^{(\cos)} &\Rightarrow \text{pattern} : \cos(m\vartheta) \\ \Upsilon_m^{(\sin)} &\Rightarrow \text{pattern} : \sin(m\vartheta) \end{aligned} \quad (17)$$

donde ϑ denota un ángulo azimutal para que

$$\begin{aligned} X &= \Upsilon_1^{(\cos)} \\ Y &= \Upsilon_1^{(\sin)} \\ U &= \Upsilon_2^{(\cos)} \\ V &= \Upsilon_2^{(\sin)} \end{aligned} \quad (18)$$

Entonces, puede probarse que

$$\Upsilon_m^{(\cos)} = \frac{A_m}{2j^m}$$

$$\Upsilon_m^{(\sin)} = \frac{B_m}{2j^m}$$

donde

$$A_m = \frac{1}{J_m(kr)} \left(\overset{\circ}{P}_m + \overset{\circ}{P}_{-m} \right)$$

$$B_m = j \cdot \frac{1}{J_m(kr)} \left(\overset{\circ}{P}_m - \overset{\circ}{P}_{-m} \right)$$

$$P(\varphi, r) = \sum_{m=-\infty}^{\infty} \overset{\circ}{P}_m e^{jm\varphi} \quad (19)$$

donde j es la unidad imaginaria, k es el número de onda, r y φ son radio y ángulo azimutal que definen un sistema de coordenadas polar, $J_m(\cdot)$ es la función Bessel de orden m del primer tipo, y $\overset{\circ}{P}_m$ son coeficientes de las series Fourier de la señal de presión medida en las coordenadas polares (r, φ).

Obsérvese que debe ponerse cuidado en el diseño e implementación de las series de cálculo de las señales con formato B (mayor orden) para evitar excesiva amplificación de ruido por las propiedades numéricas de la función Bessel.

5 Los antecedentes matemáticos y derivaciones de la transformación de señal que se describe pueden hallarse en, ej. en A. Kuntz, *Análisis de campo de onda utilizando serie de micrófonos circular virtual*, Dr. Hut, 2009, ISBN: 978-3-86853-006-3.

10 Otras realizaciones de la presente invención se relacionan con un método para generar una pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i) de una señal espacial de audio de entrada 105 obtenida de la grabación en un espacio de grabación. Por ejemplo, la señal espacial de audio de entrada 105 comprende una señal omnidireccional W y una pluralidad de diferentes señales direccionales X, Y, Z, U, V . El método comprende proveer al menos dos señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i) de la señal espacial de audio de entrada 105 (ej. señal omnidireccional W y pluralidad de diferentes señales direccionales X, Y, Z, U, V), donde al menos dos
15 señales de audio segmentales de entrada 115 (W_i, X_i, Y_i, Z_i) se asocian a correspondientes segmentos Seg_i del espacio de grabación. Asimismo, el método comprende generar una transmisión de audio paramétrica para cada señales de audio segmental de entrada 115 (W_i, X_i, Y_i, Z_i) para obtener la pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i).

20 Otras realizaciones de la presente invención se relacionan con un método para generar una pluralidad de señales de altavoz 525 (L_1, L_2, \dots) de una pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i) derivadas de señal espacial de audio de entrada 105 grabada en un espacio de grabación. El método comprende proveer una pluralidad de señales de altavoz segmentales de entrada 515 de la pluralidad de transmisiones de audio paramétricas 125 (θ_i, Ψ_i, W_i), donde las señales de altavoz segmentales de entrada 515 se asocian a correspondientes segmentos Seg_i
25 del espacio de grabación. Asimismo, el método comprende combinar señales de altavoz segmentales de entrada 515 para obtener la pluralidad de señales de altavoz 525 (L_1, L_2, \dots).

30 Aunque la presente invención se describe en contexto de diagramas en bloque donde los representan componentes, la presente invención podrá implementarse mediante un método que implementa computadora. En el último caso, los bloques representan correspondientes pasos de método donde los pasos representan funciones que realizan los correspondientes bloques de hardware lógicos o físicos.

35 Las realizaciones que se describen son solo ilustrativas a los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y detalles que se describen en la presente resultan evidentes a los expertos en la técnica. Por ende, se intenta quedar limitado solo por el alcance de las reivindicaciones de la patente debajo y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en la presente.

40 Aunque algunos aspectos se describen en el contexto de un aparato, resulta claro que esos aspectos representan una descripción de los correspondientes métodos, donde un bloque o dispositivo corresponde a un paso del método o rasgo de un método. En forma análoga, los aspectos que se describen en el contexto de un paso del método representan una descripción de un correspondiente bloque o artículo o rasgo de un correspondiente aparato. Algunos o todos los pasos del método podrán ejecutarse (o utilizando) mediante un aparato de hardware como, por
45 ejemplo, un microprocesador, computadora programable o circuito electrónico. En algunas realizaciones, algunos o los más importantes pasos del método podrán ejecutarse mediante dicho aparato.

50 La transmisión de audio paramétrica 125 (θ_i, Ψ_i, W_i) podrá almacenarse en un medio de almacenamiento digital o transmitirse en un medio de transmisión como medio de transmisión inalámbrico o medio de transmisión cableado como internet.

55 Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención podrán implementarse en hardware o en software. La implementación podrá llevarse a cabo utilizando un medio de almacenamiento digital, por ejemplo un disco blando, DVD, Blu-Ray, CD, ROM, EPROM, EEPROM o memoria FLASH, con señal de control legible en forma electrónica en ellos, que cooperan (pueden cooperar) con un sistema de computación programable a medida que se aplica el respectivo método. Por lo tanto, el medio de almacenamiento digital podrá ser legible por computadora.

60 Algunas realizaciones de acuerdo con la invención comprenden un portador de datos con señales de control legibles en forma electrónica, que cooperan con un sistema de computación programable a medida que se aplica uno de los métodos de la presente.

65 Generalmente, las realizaciones de la presente invención se implementan como producto de programa de computación con un código de programa, el código de programa aplica un método cuando el producto de programa de computación funciona en una computadora. El código de programa podrá por ejemplo almacenarse en un portador legible por máquina.

Otras realizaciones comprenden el programa de computación para aplicar un método de la presente, almacenado en un portador legible por máquina.

5 En otras palabras, una realización del método de la invención, es por lo tanto, un programa de computación con un código de programa para aplicar un método que se describe en la presente, cuando el programa de computación opera en una computadora.

10 Otra realización del método de la invención es por ende un portador de datos (o medio de almacenamiento digital, o medio legible por computadora) que comprende el programa de computación que aplica un método de la presente grabado en el. El portador de datos, el medio de almacenamiento digital o medio grabado son típicamente tangibles y/o no de transición.

15 Otra realización del método de la invención es por lo tanto una transmisión de datos o secuencia de señales que representan el programa de computación para aplicar uno de los métodos de la presente. La transmisión de datos o secuencia de señales podrá, por ejemplo, transferirse por conexión de comunicación de datos, por ejemplo por internet.

20 Otra realización comprende un medio de procesamiento, por ejemplo una computadora o dispositivo lógico programable configurado o adaptado para aplicar uno de los métodos de la presente.

Otra realización comprende una computadora con un programa de computación instalado en ella para aplicar uno de los métodos de la presente.

25 Otra realización de acuerdo con la invención comprende un aparato o sistema configurado para transferir (por ejemplo, en forma electrónica u óptica) un programa de computación para aplicar uno de los métodos de la presente a un receptor. El receptor podrá, por ejemplo, ser una computadora, dispositivo móvil, dispositivo de memoria o similar. El aparato o sistema podrá, por ejemplo, comprender un servidor de archivos para transferir el programa de computación al receptor.

30 En algunas realizaciones, podrá emplearse un dispositivo lógico programable (por ejemplo una matriz de puertas programables de campo) para aplicar algunas o todas las funciones de los métodos de la presente. En algunas realizaciones, una a matriz de puertas programables de campo podrá operar con un microprocesador para aplicar uno de los métodos de la presente. Generalmente, los métodos se aplican mediante aparato de hardware.

35 Las realizaciones de la presente invención proveen una grabación and reproducción de sonido espacial realista de alta calidad utilizando una simple y compacta configuración de micrófonos.

40 Las realizaciones de la presente invención se basan en la codificación de audio direccional (DirAC, por su sigla en inglés) (T. Lokki, J. Merimaa, V. Pulkki: Método para Reproducir Impresión espacial Natural o Modificada en Escuchas por Múltiples canales, Patente de EEUU 7.787.638 B2, 31 de Agosto 31 de 2010 y V. Pulkki: Reproducción de Sonido Espacial con codificación de audio direccional. J. Audio Eng. Soc., Vol. 55, No. 6, pp. 503-516, 2007), que se usad con diferentes sistemas de micrófonos, y configuraciones arbitrarias de altavoz. El beneficio de DirAC es reproducir la impresión espacial de un ambiente acústico existente lo más preciso posible utilizando un sistema de altavoz con múltiples canales. Dentro del ambiente elegido, las respuestas (sonido continuo o respuestas de impulsos) pueden medirse con un micrófono omnidireccional (W) y grupo de micrófonos que miden la dirección de llegada (DOA) del sonido y difusión del sonido. Un posible método es aplicar tres micrófonos con figura de ocho (X, Y, Z) alineados con los ejes coordenados Cartesianos correspondientes. Una forma de hacer esto es utilizar un micrófono "Campo de sonido", que directamente produce las respuestas deseadas. Es interesante observar que la señal del micrófono omnidireccional representa la presión de sonido, donde las señales de dipolos son proporcionales a los correspondientes elementos del vector de velocidad de partícula.

50 De estas señales, los parámetros DirAC, es decir DOA de sonido y difusión del campo de sonido observado pueden medirse en una trama de tiempo/frecuencia adecuada con resolución correspondiente al sistema auditivo humano. Las señales de altavoz reales podrán determinarse de la señal del micrófono omnidireccional basada en los parámetros DirAC (V. Pulkki: Reproducción de Sonido Espacial con Codificación de Audio Direccional J. Audio Eng. Soc., Vol. 55, No. 6, pp. 503-516, 2007). Los componentes del Sonido Directo podrán reproducirse por un número menor de altavoces (ej. uno o dos) utilizando técnicas de paneo, y los componentes de sonido difundido podrán reproducirse desde todos los altavoces al mismo tiempo.

60 Las realizaciones de la presente invención basadas en DirAC representan un simple acercamiento a la grabación de sonido espacial con configuración de micrófonos compacta. En particular, la presente invención previene inconvenientes sistemáticos que limitan la calidad de sonido logrado y experiencia en práctica en técnicas previas.

5 En oposición a DirAC convencional, las realizaciones de la presente invención proveen un procesamiento espacial de audio paramétrico de mayor calidad. El DirAC convencional se basa en un modelo simple global para el campo de sonido, que emplea sólo un DOA y un parámetro de difusión para todo el espacio de observación. Se basa en la suposición que el campo de sonido puede representarse por sólo un componente de sonido directo, como onda de plano, y parámetro de difusión global para cada mosaico de tiempo/frecuencia. En la práctica, sin embargo, esta simplificada suposición sobre el campo de sonido no se sostiene. Esto es verdadero en el caso de acústica compleja, del mundo real, ej. donde múltiples fuentes de sonido como parlantes o instrumento se encuentran activos al mismo tiempo. Por otro lado, las realizaciones de la presente invención no provocan discordancia de modelo del campo de sonido observado, y los correspondientes cálculos de parámetro son más correctos. Se evita además que haya discordancia de modelo, especialmente en casos donde los componentes de sonido directo sean difundidos y no se perciba dirección al escuchar las salidas del altavoz. En realizaciones, los decorrelacionadores podrán generar sonido difundido no correlacionado reproducido desde todos los altavoces (V. Pulkki: Reproducción de sonido espacial con codificación de audio Direccional. J. Audio Eng. Soc., Vol. 55, No. 6, pp. 503-516, 2007). En oposición a técnicas previas, donde los decorrelacionadores introducen un efecto de ambiente agregado no deseado, es posible con la presente invención reproducir en forma más correcta Fuentes de sonido con alcance espacial (opuesto al caso de modelo simple de campo de sonido de DirAC que no captura en forma precisa dichas fuentes de sonido).

20 Las realizaciones de la presente invención proveen un mayor número de grados de libertad en el supuesto modelo de señal, permitiendo mejor concordancia de modelo en escenas de sonido complejo.

25 Asimismo, en caso de utilizar micrófonos direccionales y generar sectores (u otro medio lineal que sin variación en el tiempo, ej. físico), se obtiene una dirección inherente aumentada de micrófonos. Por lo tanto, hay menor necesidad de aplicar ganancias sin variación en el tiempo para evitar direcciones vagas, interferencias, y coloración. De este modo se obtiene un menor procesamiento no lineal en el recorrido de la señal de audio, obteniendo una mejor calidad.

30 En general, se pueden suministrar más componentes de sonido directo como fuentes de sonido directo (fuentes de punto/fuentes de onda de plano). En consecuencia, hay menos artefactos de decorrelación, se perciben más eventos localizables (correctamente), y se logra una reproducción espacial más exacta.

35 Las realizaciones de la presente invención proveen un mayor desempeño de manipulación en el dominio paramétrico, ej. filtro direccional (M. Kallinger, H. Ochsenfeld, G. Del Galdo, F. Kuech, D. Mahne, R. Schultz-Amling, y O. Thiergart: Método de Filtro Espacial para Codificación de Audio Direccional, 126th AES Convención, Doc 7653, Munich, Alemania 2009), comparado con el modelo simple global, ya que una mayor fracción de la energía de señal total se atribuye a eventos de sonido directo con un correcto DOA asociado al mismo, y se dispone de una mayor cantidad de información. La provisión de más (paramétrica) información permite, por ejemplo, separar múltiples componentes de sonido directo o dirigir los componentes de sonido directo de reflexiones tempranas que impactan de diferentes direcciones.

40 Específicamente, las realizaciones proveen los siguientes aspectos. En el caso 2D, todo el rango de ángulo azimutal podrá dividirse en sectores que cubren rangos de ángulo azimutal reducidos. En el caso 3D, todo el rango de ángulo sólido podrá dividirse en sectores que cubren rangos de ángulo sólido reducidos. Cada sector puede asociarse a un rango de ángulo referido. Para cada sector, las señales de micrófono segmentales podrán determinarse de las señales de micrófono recibidas, que predominantemente consisten en sonido que llega de direcciones asignadas a/cubiertas por el particular sector. Las señales de micrófono podrán determinarse en forma artificial mediante grabaciones virtuales simuladas. Para cada sector, un análisis de campo de sonido paramétrico podrá realizarse para determinar parámetros como DOA y difusión. Para cada sector, la información de dirección paramétrica (DOA y difusión) predominantemente describe las propiedades espaciales del rango angular del campo de sonido asociado al particular sector. En caso de reproducción, para cada sector, las señales de altavoz pueden determinarse teniendo en cuenta los parámetros de direccionalidad y señales de micrófono segmentales. Toda la salida se obtiene al combinar salidas de todos los sectores. En caso de manipulación, antes de computar señales de altavoz para la reproducción, los parámetros y/o señales de audio segmentales estimados podrán modificarse y lograr una manipulación de la escena de sonido.

55

REIVINDICACIONES

1. Un aparato (100) para generar una pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i) desde una señal espacial de audio de entrada (105) obtenida de una grabación en un espacio de grabación, el aparato (100) comprende:
- un segmentador (110) para generar al menos dos señales de audio segmentales de entrada (115) (W_i, X_i, Y_i, Z_i) desde la señal espacial de audio de entrada (105), donde el segmentador (110) está configurado para generar las al menos dos señales de audio segmentales de entrada (115) (W_i, X_i, Y_i, Z_i) dependiendo de unos correspondientes segmentos (Seg_i) del espacio de grabación, donde cada uno de los segmentos (Seg_i) del espacio de grabación representa un sub grupo de direcciones dentro de un plano en dos dimensiones (2D) o dentro de un espacio en tres dimensiones (3D), y donde los segmentos (Seg_i) son diferentes entre sí, y
- un generador (120) para generar una transmisión de audio paramétrica para cada una de las al menos dos señales de audio segmentales de entrada (115) (W_i, X_i, Y_i, Z_i) para obtener la pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i), de manera que cada una de la pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i) comprende un componente (W_i) de las al menos dos señales de audio segmentales de entrada (115) (W_i, X_i, Y_i, Z_i) y correspondiente información espacial paramétrica (θ_i, Ψ_i), donde la información espacial paramétrica (θ_i, Ψ_i) de cada transmisión de audio paramétrica (125) (θ_i, Ψ_i, W_i) comprende un parámetro (θ_i) de dirección de llegada (DOA) y/o un parámetro de difusión (Ψ_i).
2. El aparato (100) de acuerdo con la reivindicación 1,
- donde cada uno de los segmentos (Seg_i) del espacio de grabación se caracteriza por una medida direccional asociada.
3. El aparato (100) de acuerdo con la reivindicación 1 o 2,
- donde el aparato (100) está configurado para realizar una grabación de campo de sonido para obtener la señal espacial de audio de entrada (105);
- donde el segmentador (110) está configurado para dividir un rango de interés de ángulo total en los segmentos (Seg_i) del espacio de grabación;
- donde los segmentos (Seg_i) del espacio de grabación cubren un rango de ángulo reducido comparado con el rango de interés de ángulo total,
4. El aparato (100) de acuerdo con una de las reivindicaciones 1 a 3,
- donde la señal espacial de audio de entrada (105) comprende una señal omnidireccional (W) y una pluralidad de diferentes señales direccionales (X, Y, Z, U, V).
5. El aparato (100) de acuerdo con una de las reivindicaciones 1 a 4,
- donde el segmentador (110) está configurado para generar las al menos dos señales de audio segmentales de entrada (115) (W_i, X_i, Y_i, Z_i) a partir de la señal omnidireccional (W) y de la pluralidad de diferentes señales direccionales (X, Y, Z, U, V) utilizando una operación de mezcla que depende de los segmentos (Seg_i) del espacio de grabación.
6. El aparato (100) de acuerdo con una de las reivindicaciones 1 a 5,
- donde el segmentador (110) está configurado para usar un patrón de direccionalidad (305) ($q_i(\vartheta)$) para cada segmento (Seg_i) del espacio de grabación;
- donde el patrón de direccionalidad (305) ($q_i(\vartheta)$) indica una direccionalidad de las al menos dos señales de audio segmentales de entrada (115) (W_i, X_i, Y_i, Z_i).
7. El aparato (100) de acuerdo con la reivindicación 6,
- donde el patrón de direccionalidad (305) ($q_i(\vartheta)$) se da por
- $$q_i(\vartheta) = a + b \cos(\vartheta + \Theta_i),$$
- donde a y b denotan multiplicadores modificados para obtener un patrón de direccionalidad deseado (305) ($q_i(\vartheta)$);

donde ϑ denota un ángulo azimutal y Θ_i indica una dirección preferida del segmento de orden i del espacio de grabación.

5 8. El aparato (100) de acuerdo con una de las reivindicaciones 1 a 7,

donde el generador (120) está configurado para realizar un análisis espacial paramétrico para cada una de las al menos dos señales de audio segmentales de entrada (115) (W_i, X_i, Y_i, Z_i) para obtener la correspondiente información espacial paramétrica (θ_i, Ψ_i).

10

9. El aparato (100) de acuerdo con una de las reivindicaciones 1 a 8, que además comprende:

un modificador (910) para modificar la pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i) en un dominio de representación de señal paramétrica;

15

donde el modificador (910) está configurado para modificar al menos una de las transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i) utilizando un parámetro de control correspondiente de modificación (905).

20

10. Un aparato (500) para generar una pluralidad de señales de altavoz (525) (L_1, L_2, \dots) desde una pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i); donde cada una de la pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i) comprende un componente de audio segmental (W_i) y correspondiente información espacial paramétrica (θ_i, Ψ_i); donde la información espacial paramétrica (θ_i, Ψ_i) de cada transmisión de audio paramétrica (125) (θ_i, Ψ_i, W_i) comprende un parámetro (θ_i) de dirección de llegada (DOA) y/o un parámetro de difusión (Ψ_i); donde el aparato (500) comprende:

25

un suministrador (510) para proveer una pluralidad de señales de altavoz segmentales de entrada (515) desde la pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i), de manera que las señales de altavoz segmentales de entrada (515) dependen de unos correspondientes segmentos (Seg_i) del espacio de grabación, donde cada uno de los segmentos (Seg_i) del espacio de grabación representa un sub grupo de direcciones dentro de un plano en dos dimensiones (2D) o dentro de un espacio en tres dimensiones (3D), y donde los segmentos (Seg_i) son diferentes entre sí, donde el suministrador (510) está configurado para suministrar cada una de las componentes de audio segmentales (W_i) utilizando la correspondiente información espacial paramétrica (505) (θ_i, Ψ_i) para obtener la pluralidad de señales de altavoz segmentales de entrada (515); y

30

35

un combinador (520) para combinar la señales de altavoz segmentales de entrada (515) para obtener la pluralidad de señales de altavoz (525) (L_1, L_2, \dots).

40

11. Un método para generar una pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i) desde una señal espacial de audio de entrada (105) que se obtiene de una grabación en un espacio de grabación, el método comprende:

45

generar al menos dos señales de audio segmentales de entrada (115) (W_i, X_i, Y_i, Z_i) de la señal espacial de audio de entrada (105), donde la generación de las al menos dos señales de audio segmentales de entrada (115) (W_i, X_i, Y_i, Z_i) se lleva a cabo dependiendo de unos correspondientes segmentos (Seg_i) del espacio de grabación, donde cada uno de los segmentos (Seg_i) del espacio de grabación representa un sub grupo de direcciones dentro de un plano en dos dimensiones (2D) o dentro de un espacio en tres dimensiones (3D), y donde los segmentos (Seg_i) son diferentes entre sí, y

50

generar una transmisión de audio paramétrica para cada una de las al menos dos señales de audio segmentales de entrada (115) (W_i, X_i, Y_i, Z_i) para obtener la pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i), de manera que cada una de la pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i) comprende un componente (W_i) de las al menos dos señales de audio segmentales de entrada (115) (W_i, X_i, Y_i, Z_i) y correspondiente información espacial paramétrica (θ_i, Ψ_i), donde la información espacial paramétrica (θ_i, Ψ_i) de cada una de las transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i) comprende un parámetro (θ_i) de dirección de llegada (DOA) y/o un parámetro de difusión (Ψ_i).

55

60

12. Un método para generar una pluralidad de señales de altavoz (525) (L_1, L_2, \dots) de una pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i); donde cada una de la pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i) comprende un componente (W_i) de audio segmental y correspondiente información espacial paramétrica (θ_i, Ψ_i); donde la información espacial paramétrica (θ_i, Ψ_i) de cada transmisión de audio paramétrica (125) (θ_i, Ψ_i, W_i) comprende un parámetro (θ_i) de dirección de llegada (DOA) y/o un parámetro de difusión (Ψ_i); donde el método comprende:

65

proveer una pluralidad de señales de altavoz segmentales de entrada (515) de la pluralidad de transmisiones de audio paramétricas (125) (θ_i, Ψ_i, W_i), de manera que las señales de altavoz segmentales de entrada (515) dependen

- 5 de unos correspondientes segmentos (Seg_i) del espacio de grabación, donde cada uno de los segmentos (Seg_i) del espacio de grabación representa un sub grupo de direcciones dentro de un plano en dos dimensiones (2D) o dentro de un espacio en tres dimensiones (3D), y donde los segmentos (Seg_i) son diferentes entre sí, donde la provisión de la pluralidad de señales de altavoz segmentales de entrada (515) se lleva a cabo mediante el suministro de cada una de las componentes de audio segmentales (W_i) utilizando la correspondiente información espacial paramétrica (505) (θ_i, Ψ_i) para obtener la pluralidad de señales de altavoz segmentales de entrada (515); y
- 10 combinar la señales de altavoz segmentales de entrada (515) para obtener la pluralidad of señales de altavoz (525) (L_1, L_2, \dots).
13. Un programa de computación con un código de programa para aplicar el método de acuerdo con la reivindicación 11 cuando el programa de computación se ejecuta en una computadora.
- 15 14. Un programa de computación con un código de programa para aplicar el método de acuerdo con la reivindicación 12 cuando el programa de computación se ejecuta en una computadora.

100

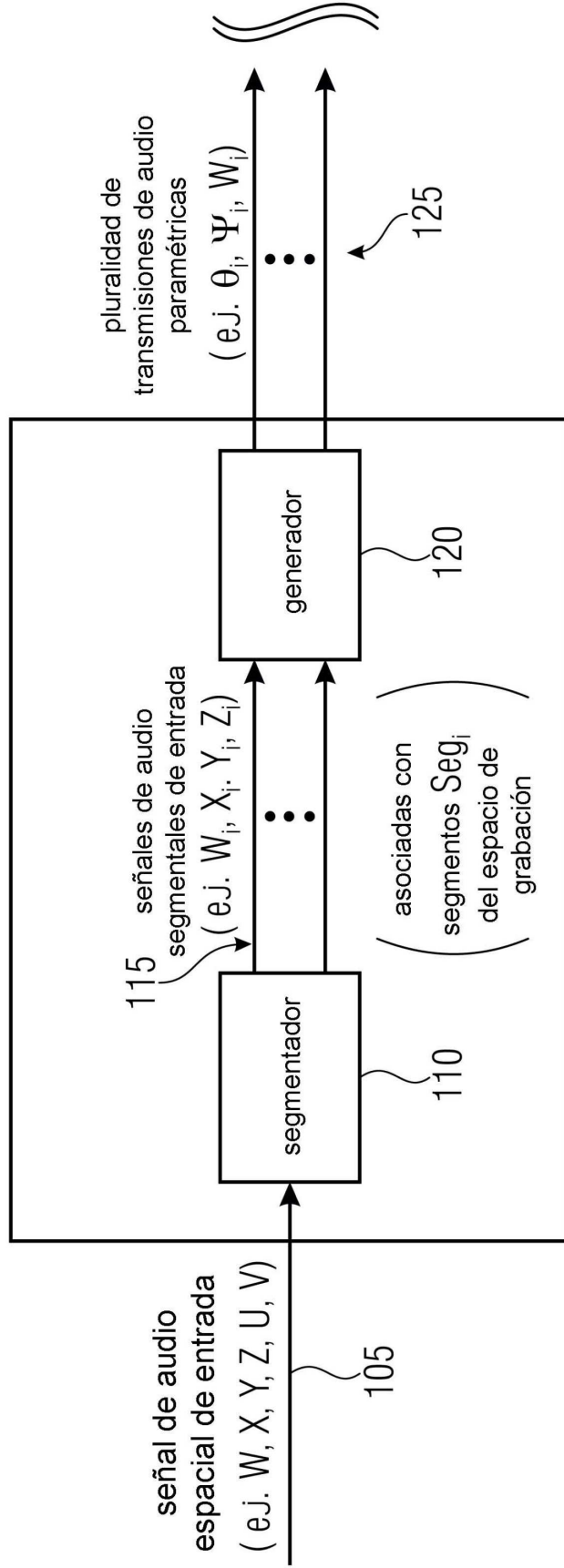


FIGURA 1

100

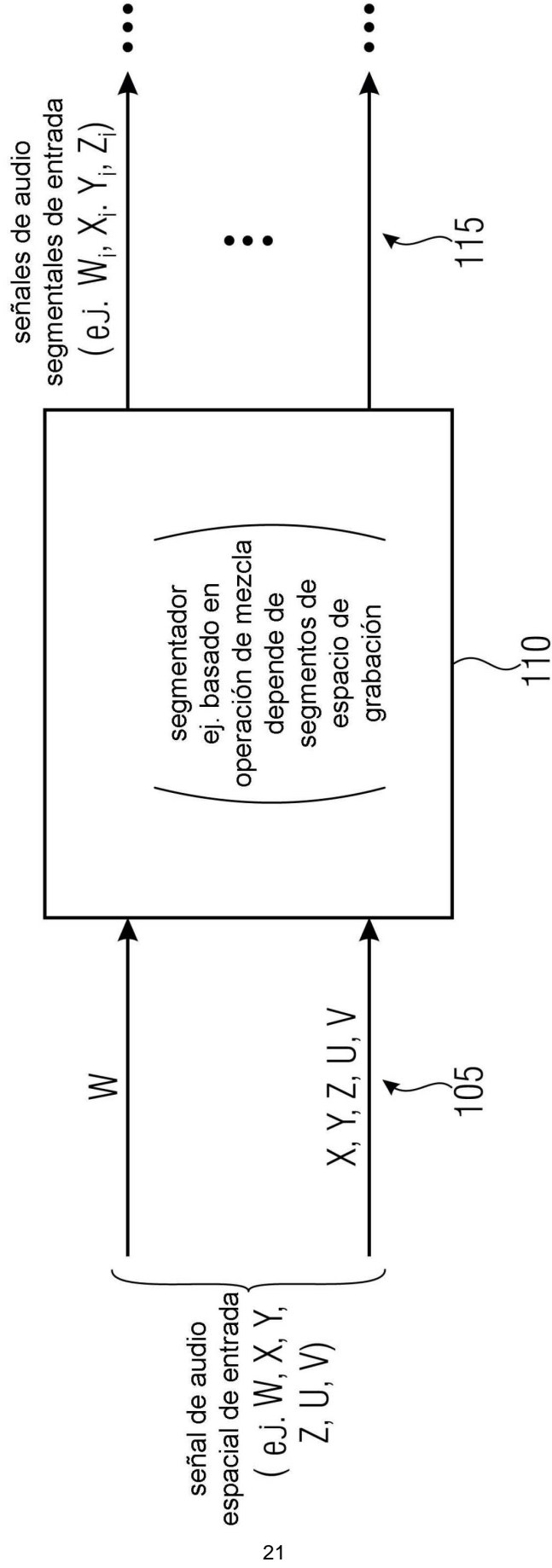


FIGURA 2

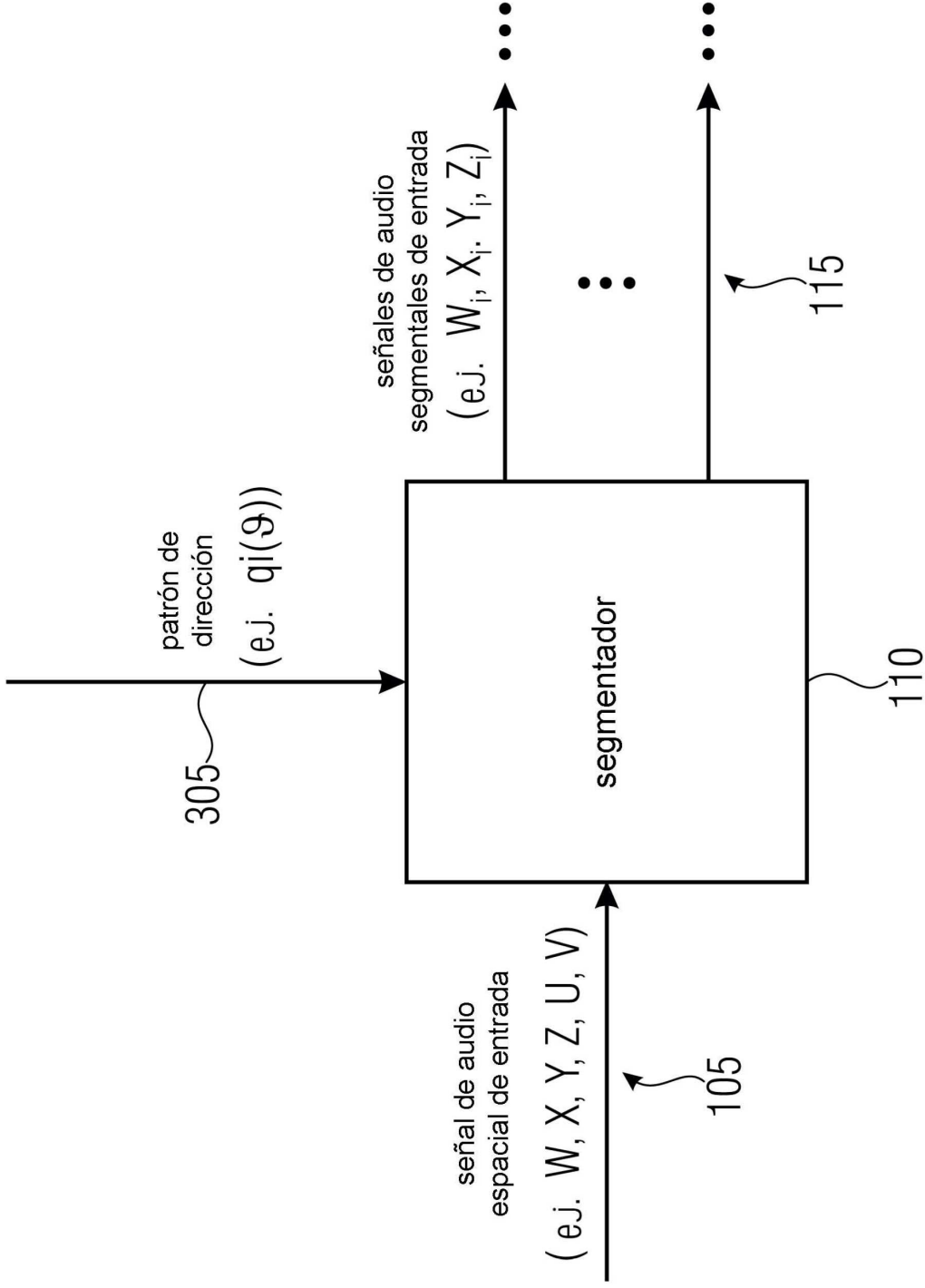


FIGURA 3

100

100

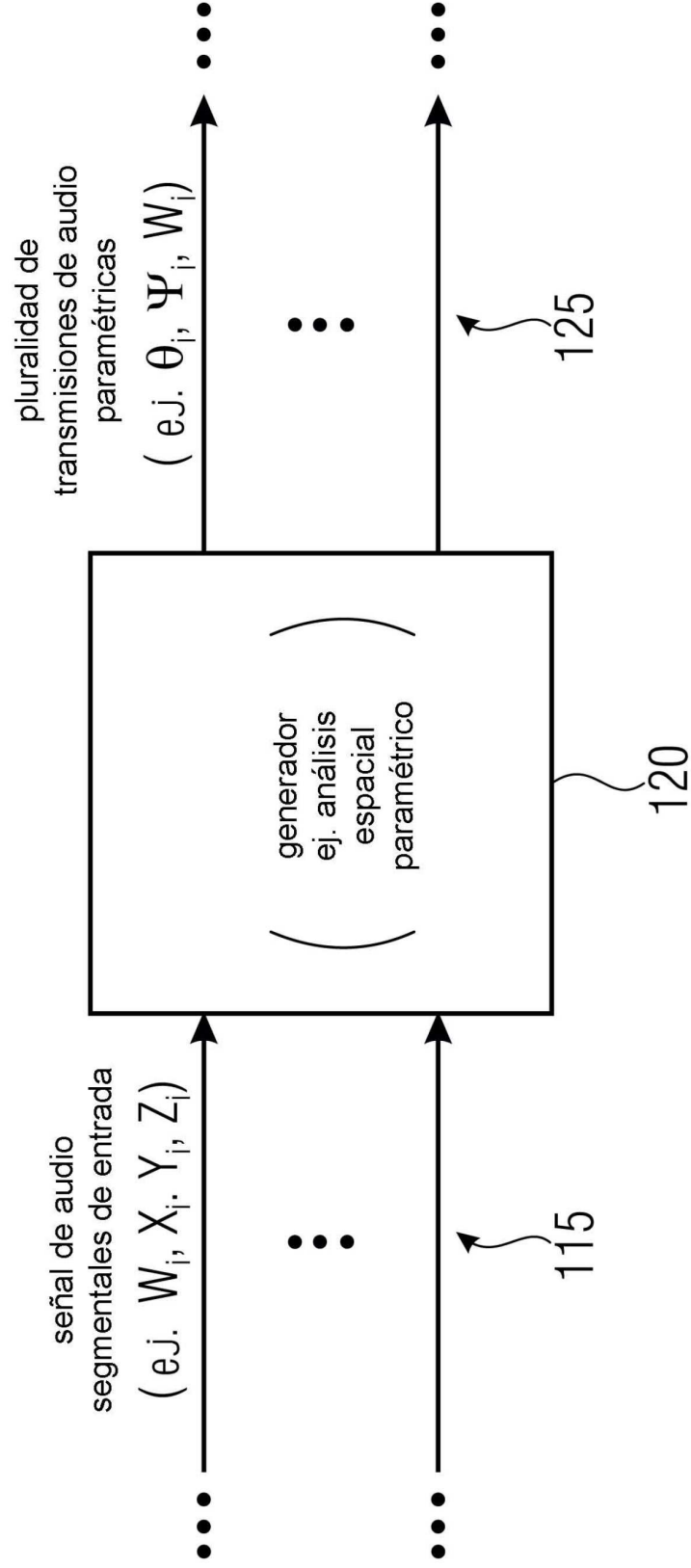


FIGURA 4

500

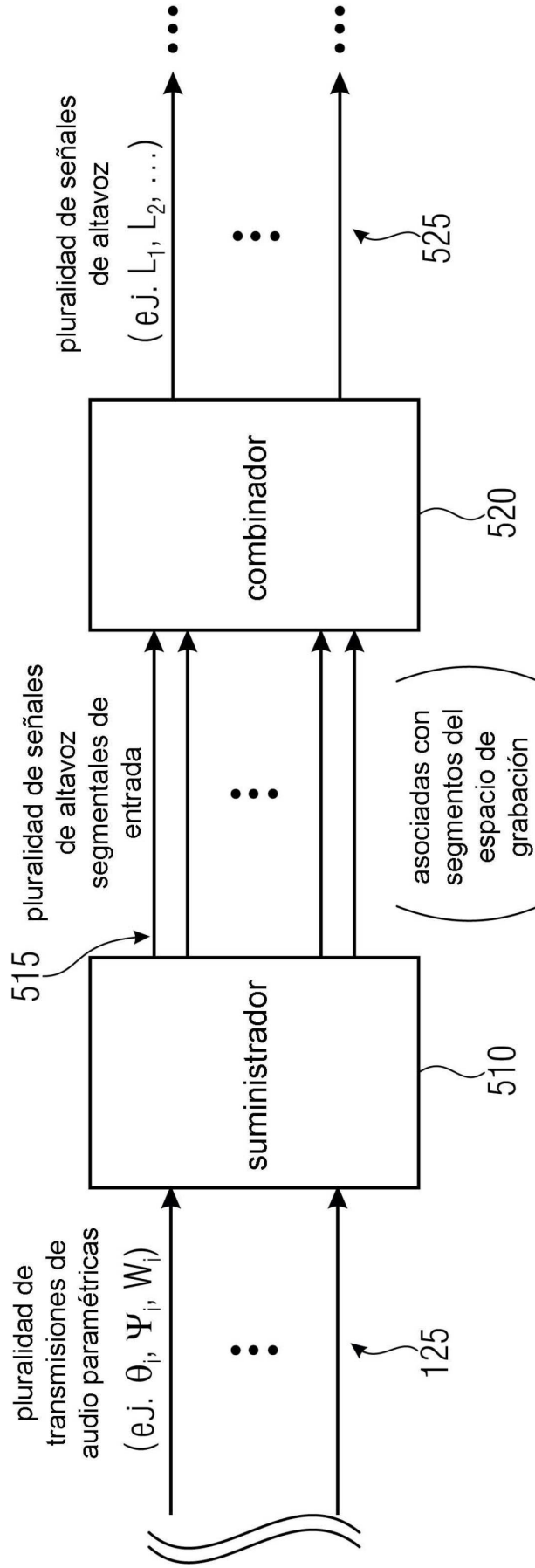


FIGURA 5

600

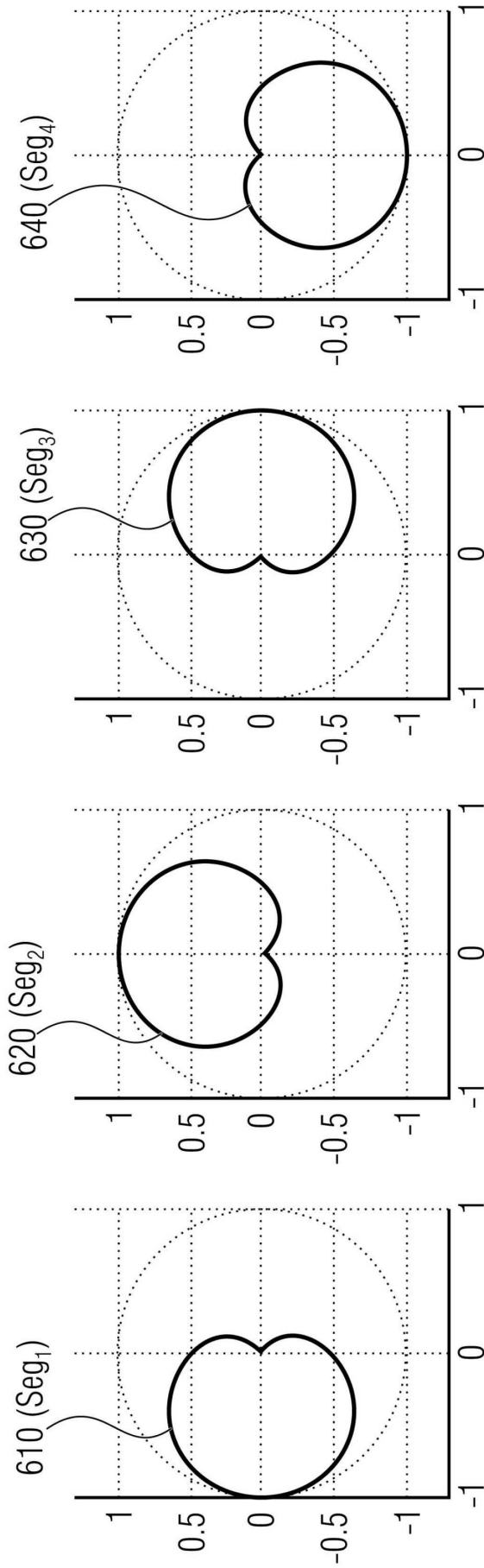


FIGURA 6

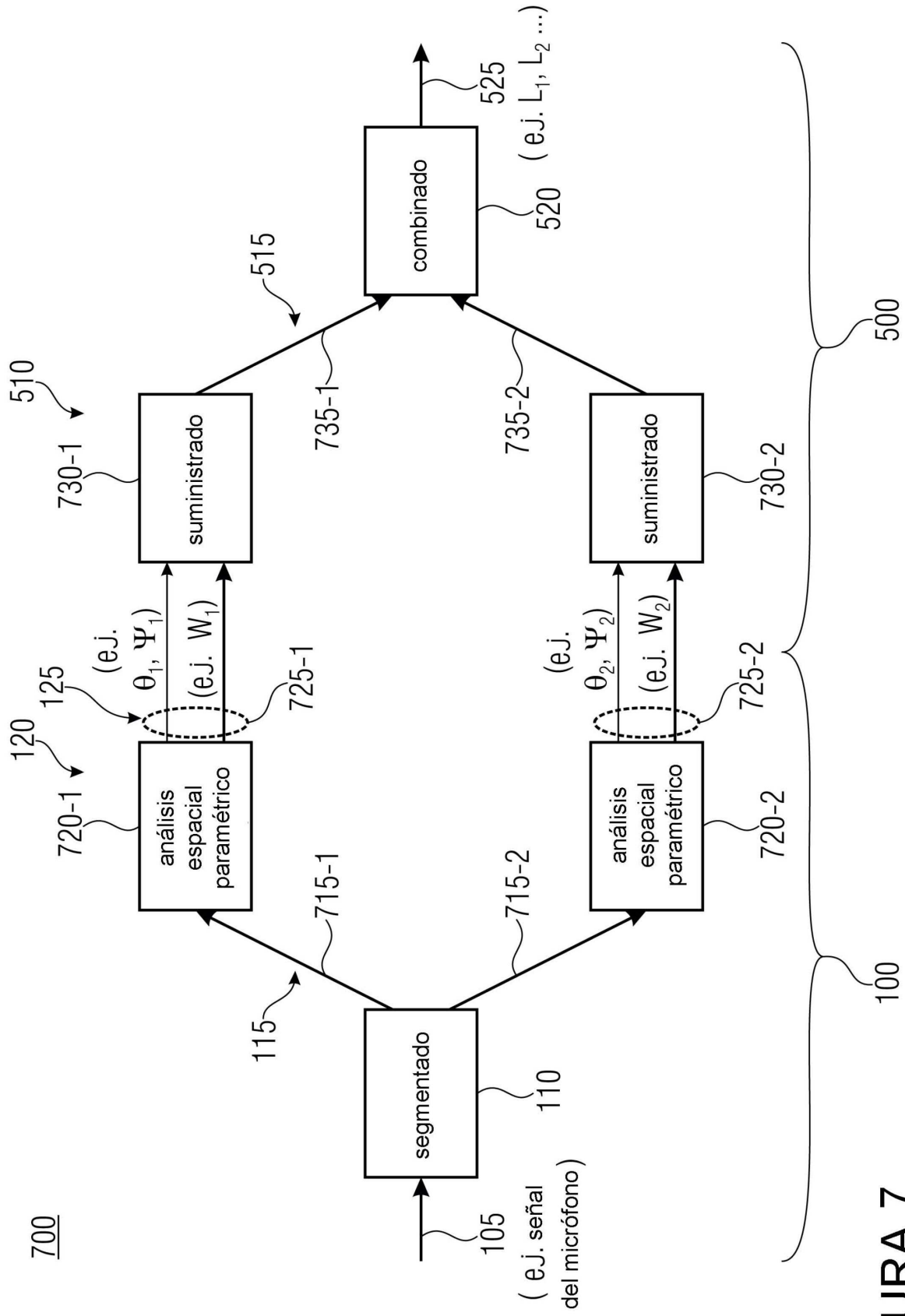


FIGURA 7

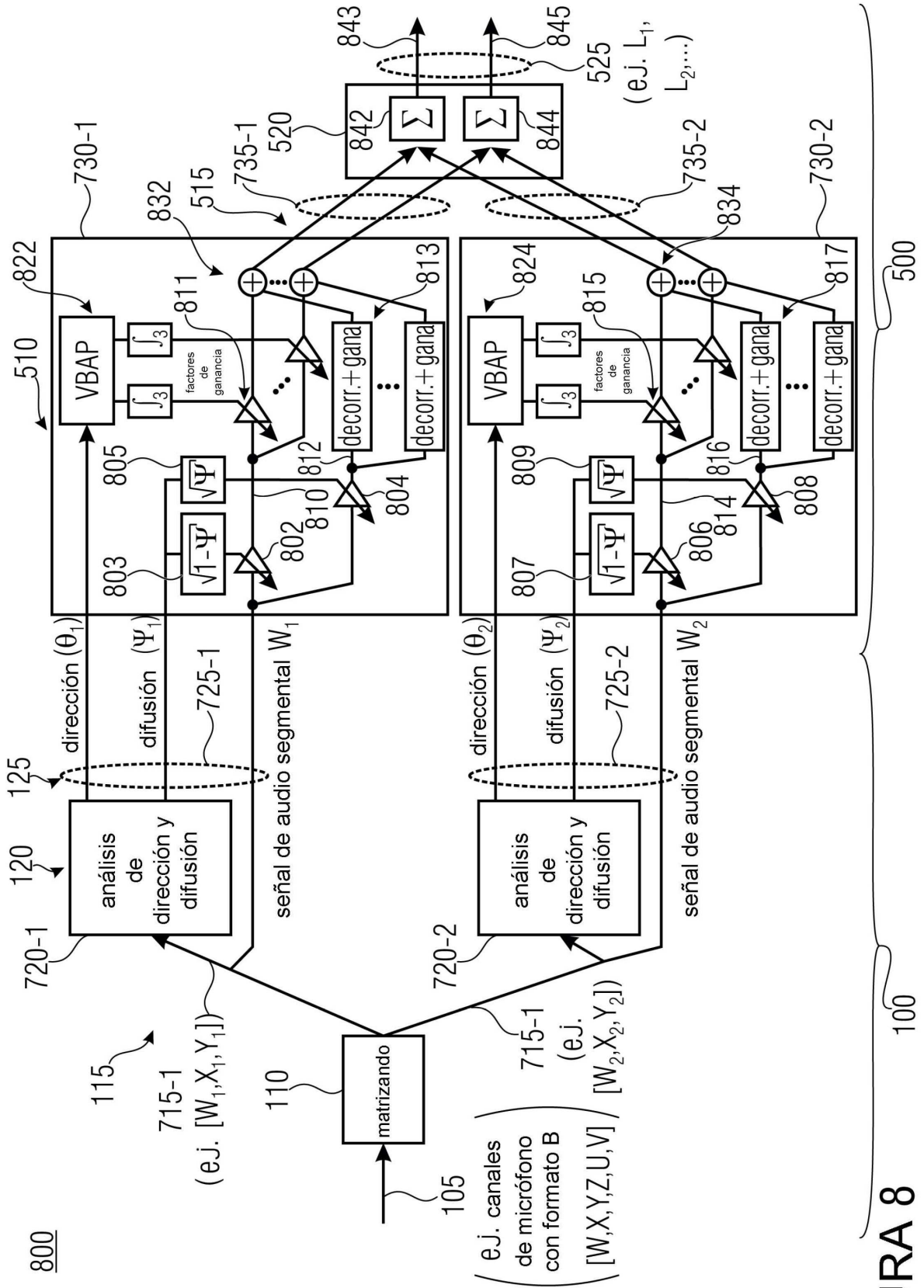


FIGURA 8

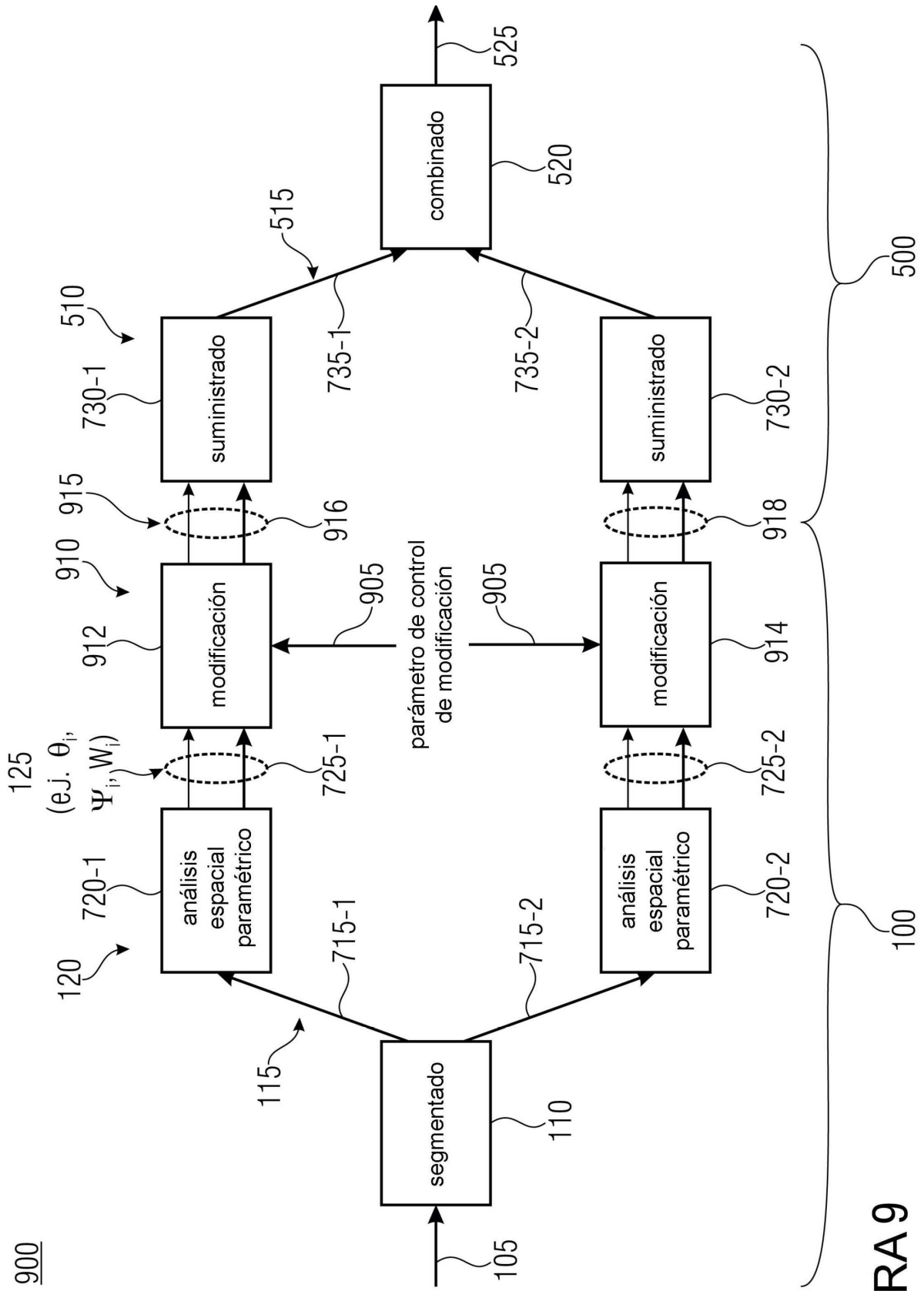


FIGURA 9

115

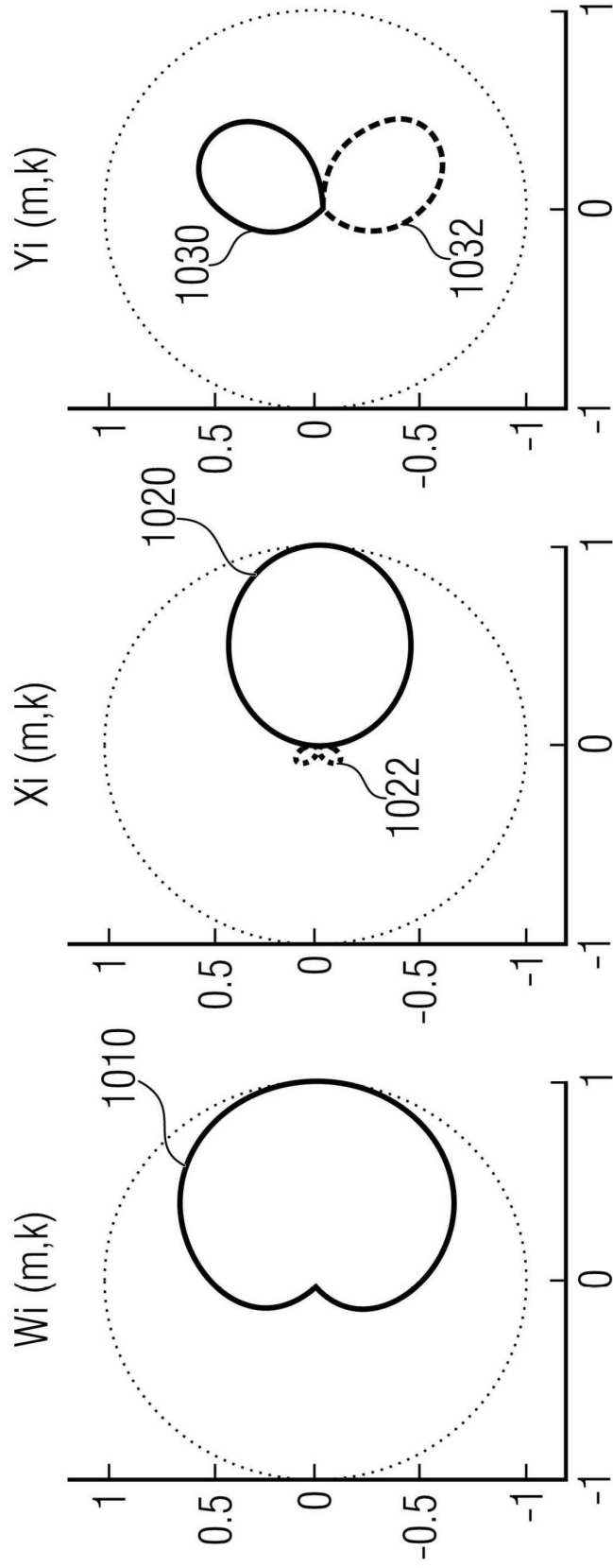


FIGURA 10

1000

1100

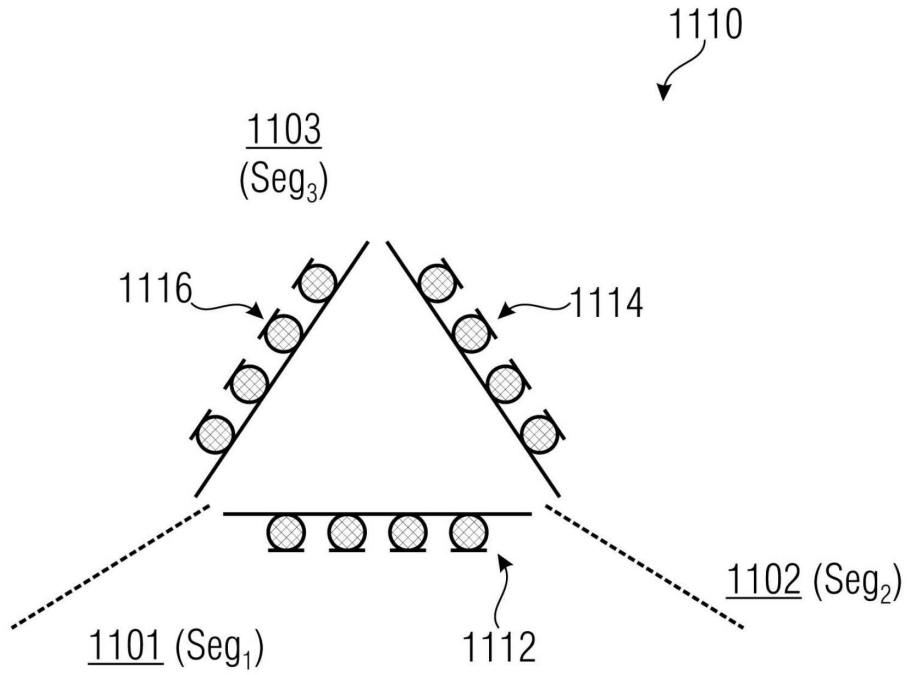


FIGURA 11

1200

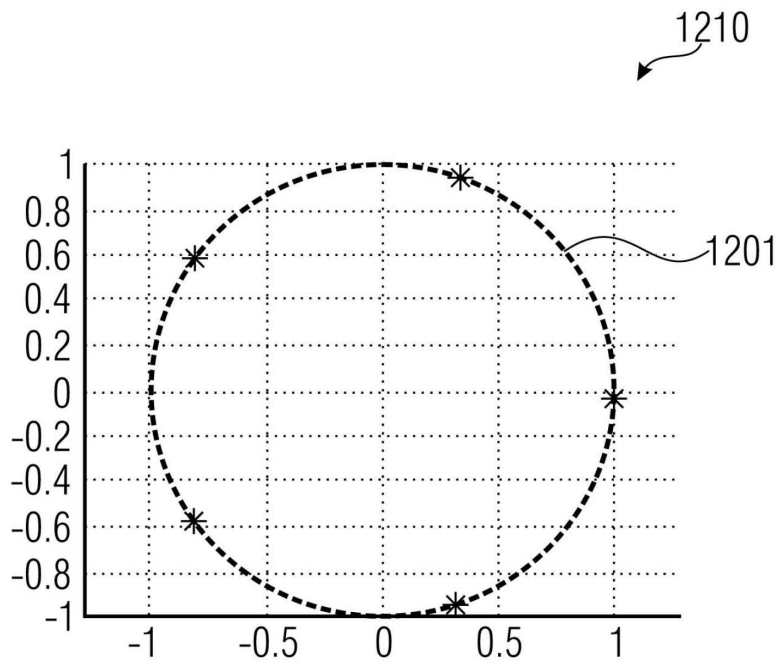


FIGURA 12