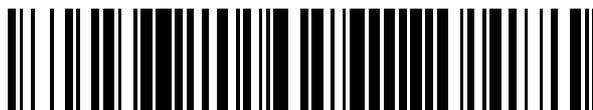


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 592 217**

51 Int. Cl.:

G10L 19/16 (2013.01)
G10L 25/48 (2013.01)
G10L 25/03 (2013.01)
H04S 1/00 (2006.01)
H04S 3/00 (2006.01)
H04S 7/00 (2006.01)
H04N 5/232 (2006.01)
G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **14.12.2010 PCT/EP2010/069669**
- 87 Fecha y número de publicación internacional: **23.06.2011 WO11073210**
- 96 Fecha de presentación y número de la solicitud europea: **14.12.2010 E 10796353 (0)**
- 97 Fecha y número de publicación de la concesión europea: **22.06.2016 EP 2502228**

54 Título: **Un aparato y método para convertir una primera señal de audio paramétrico espacial en una segunda señal de audio paramétrico espacial**

30 Prioridad:

17.12.2009 US 287596 P
11.03.2010 EP 10156263

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.11.2016

73 Titular/es:

FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V. (100.0%)
Hansastraße 27c
80686 München, DE

72 Inventor/es:

SCHULTZ-AMLING, RICHARD;
KÜCH, FABIAN;
KALLINGER, MARKUS;
DEL GALDO, GIOVANNI;
THIERGART, OLIVER;
MAHNE, DIRK y
KUNTZ, ACHIM

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 592 217 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).

Un aparato y método para convertir una primera señal de audio paramétrico espacial en una segunda señal de audio paramétrico espacial

DESCRIPCIÓN

5

La presente invención se refiere al campo del procesamiento de audio, especialmente al campo del procesamiento de audio espacial paramétrico y para la conversión de una primera señal de audio paramétrico espacial en una segunda señal de audio paramétrico espacial.

10

Antecedentes de la invención

La grabación de sonido espacial tiene por objeto capturar un campo de sonido con múltiples micrófonos de tal manera que, en el lado de la reproducción, un oyente perciba la imagen de sonido, tal como estuviera presente en el sitio de la grabación. Los enfoques normales para la grabación de sonido espacial utilizan micrófonos estéreo sencillos o combinaciones más sofisticadas de micrófonos direccionales, por ejemplo, tal como los micrófonos de formato B empleados en Ambisonics y descritos por M.A. Gerzon, "Periphony: Width-Height Sound Reproduction," J. Aud. Eng. Soc., Vol. 21, N.º 1, págs. 2-10, 1973, al que en adelante se hace referencia como [Ambisonics]. Habitualmente, se hace referencia a estos métodos como técnicas de micrófonos coincidentes.

20

Como alternativa, se pueden aplicar métodos basados en una representación paramétrica de campos de sonido, que se denominan como codificadores paramétricos de audio paramétrico espacial. Estos métodos determinan una señal de audio de mezcla descendente junto con la correspondiente información secundaria espacial, que son relevantes para la percepción del sonido espacial. Son ejemplos de esto la Codificación de Audio Direccional (DirAC) como se analiza en Pulkki, V., "Directional audio coding in spatial sound reproduction and stereo upmixing", en Proceedings of The AES 28th International Conference, págs. 251-258, Piteå, Suecia, 30 de junio - 2 de julio de 2006, al que en adelante se hace referencia como [DirAC], o el enfoque de los denominados micrófonos de audio espacial (SAM) propuesta por Faller, C., "Microphone Front-Ends for Spatial Audio Coders", en Proceedings of the AES 125th International Convention, San Francisco, octubre de 2008, al que en adelante se hace referencia como [SAM]. La información de los indicios espaciales básicamente consiste en la dirección de llegada (DOA) del sonido y la difusividad del campo de sonido en subbandas de frecuencia. En una etapa de síntesis, se determinan las señales de altavoz deseadas para la reproducción basándose en la señal de mezcla descendente y la información paramétrica secundaria.

25

30

35

En otras palabras, las señales de mezcla descendente y la correspondiente información espacial secundaria representan la escena de audio de acuerdo con la configuración, por ejemplo, la orientación y/o posición de los micrófonos, en relación con las diferentes fuentes de audio utilizadas en el momento que se grabó la escena de audio.

40

El documento US 2008/0298597 A1 describe métodos, medios legibles por ordenador y aparatos para un procesamiento digital de señales acústicas para crear una reproducción de un entorno de sonido espacial natural o uno artificial. Una realización soporta procesamiento de audio espacial tal como extrayendo un canal central en sonido estéreo de mezcla ascendente para una configuración de altavoz multi-canal o virtualización de auriculares. Una realización soporta escucha direccional en la que las fuentes de sonido en una dirección deseada pueden amplificarse o atenuarse. Se determinan los parámetros de dirección y difusividad para regiones de canales de entrada y se extrae un canal extraído a partir de los canales de entrada de acuerdo con los parámetros de dirección y difusividad. Los estimadores estimados para cada componente de señal que se alimentan en el canal extraído y un canal extraído se sintetizan a partir de una señal de base y la estimación de la ganancia. Los canales de entrada se subdividen en una pluralidad de regiones de tiempo-frecuencia.

50

El documento EP 1 589 754 A1 describe un aparato de procesamiento de información, un aparato de adquisición de imágenes, un método de procesamiento de información y un programa. El aparato de procesamiento de información tiene una función para procesar señales de audio de diversos canales asociadas con una señal de vídeo. El aparato de procesamiento de información incluye una unidad de conversión de audio para generar una pluralidad de señales de audio convirtiendo los niveles de la señal de audio de diversos canales de acuerdo con un parámetro de ajuste definido dependiendo de la información de las condiciones de captura de imágenes con respecto a la señal de vídeo.

55

El documento EP 1 473 971 A2 describe un controlador de campo de sonido. El controlador de campo de sonido incluye una pluralidad de canales de salida para suministrar respectivamente señales a una pluralidad de altavoces. El controlador de campo de sonido incluye una unidad de emisión de coeficiente de ponderación que emite coeficientes de ponderación para determinar cómo se han de distribuir las señales en la pluralidad de canales para reproducir un campo de sonido a los canales de salida, de acuerdo con una relación posicional relativa entre el campo de sonido y un oyente. El controlador de campo de sonido comprende también una unidad de entrada que

60

introduce datos de relación posicional que indican la relación posicional relativa. El controlador de campo de sonido comprende un controlador de distribución que distribuye las señales en la pluralidad de canales a los canales de salida basándose en los coeficientes de ponderación emitidos de acuerdo con los datos de relación posicional introducidos.

5 El artículo "Spatial Sound Reproduction with Directional Audio Coding" de Ville Pulkki (publicado en el Journal of the Audio Engineering Society, volumen 55, n.º 6, junio de 2007) describe codificación de audio direccional. La codificación de audio direccional es un método para representación de sonido espacial, aplicable para diferentes sistemas de reproducción de sonido. En una parte de análisis, se estiman la difusividad y dirección de llegada del sonido en una única localización dependiendo del tiempo y frecuencia. En una parte de síntesis, las señales de
10 micrófono se dividen en primer lugar en partes no difusa y difusa, y a continuación se reproducen usando diferentes estrategias. La codificación de audio direccional se desarrolla a partir de una tecnología existente para reproducción de respuesta de impulso, representación de respuesta de impulso espacial. Se describen implementaciones de codificación de audio direccional para diferentes aplicaciones.

15 El artículo "Virtual Sound Source Positioning Using Vector Base Amplitude Panning" de Ville Pulkki (publicado en el Audio Engineering Society, volumen 45, n.º 6, junio de 1997) describe una reformulación basada en vector de panorámica de amplitud que conduce a ecuaciones sencillas y computacionalmente eficaces para colocación de fuente de sonido virtual. Usando este método, es posible crear campos de sonido bi- o tridimensionales donde puede colocarse cualquier número de altavoces arbitrariamente. El método produce fuentes de sonido virtuales que
20 son tan agudas como sea posible con la configuración de altavoz actual y métodos de panorámica de amplitud. Se describe una herramienta digital que implementa panorámica de amplitud de base de vector bi- y tridimensional con ocho entradas y salidas.

Un objetivo de la presente invención es dar a conocer un concepto para una adaptación flexible de la escena de
25 audio grabada.

Sumario de la invención

30 Este objetivo se resuelve mediante aparatos de acuerdo con las reivindicaciones 1 y 18, métodos de acuerdo con las reivindicaciones 15 y 16 y un programa de computación de acuerdo con la reivindicación 17. Se incluye también una reivindicación 14 del sistema.

35 Todos los métodos antes mencionados tienen en común que tienen por objeto representar el campo de sonido en el lado de la reproducción, como si se percibiera en la posición de la grabación. La posición de grabación, es decir la posición de los micrófonos, también se puede denominar como la posición de escucha de referencia. No se contempla una modificación de la escena de audio grabada en estos métodos conocidos de captura de sonido espacial.

40 Por otro lado, se aplica comúnmente la modificación de la imagen visual, por ejemplo en el contexto de la captura de vídeo. Por ejemplo, en las cámaras de vídeo se usa un zoom óptico para modificar la posición virtual de la cámara, dando así la impresión de que la imagen fue tomada desde un punto de vista diferente. Esto se describe mediante una traslación de la posición de la cámara. Otra sencilla modificación de la imagen es la rotación horizontal o vertical de la cámara alrededor de su propio eje. También se hace referencia a la rotación vertical como panorámica o
45 inclinación.

Las realizaciones de la presente invención proporcionan aparatos y un métodos, que también permiten cambiar virtualmente la posición y/u orientación de escucha de acuerdo con el movimiento visual. En otras palabras, la invención permite alterar la imagen acústica que percibe un oyente durante la reproducción, de manera que
50 corresponda a la grabación obtenida utilizando una configuración de micrófonos ubicada en una posición y/u orientación virtual diferente de la posición física real de los micrófonos. Al hacerlo así, la imagen acústica grabada se puede alinear con la correspondiente imagen de vídeo modificada. Por ejemplo, se puede aplicar el efecto de un zoom de vídeo a una determinada área de una imagen a la imagen de audio espacial grabada de manera constante. De acuerdo con la invención, esto se logra satisfactoriamente modificando los parámetros de indicios espaciales y/o la señal de mezcla descendente en el dominio paramétrico del codificador de audio espacial.

55 Las realizaciones de la presente invención permiten cambiar con flexibilidad la posición y/u orientación de un oyente dentro de una escena de audio espacial dada sin tener que grabar la escena de audio espacial con una configuración diferente de los micrófonos, por ejemplo, una posición y/u orientación diferente de la configuración de micrófonos de grabación con respecto a las fuentes de las señales de audio. En otras palabras, las realizaciones de
60 la presente invención permiten definir una posición de escucha virtual y/o una orientación de escucha virtual que es diferente de la posición de grabación o la posición de escucha en el momento que se grabó la escena de audio espacial.

Ciertas realizaciones de la presente invención solo utilizan una o varias señales de mezcla descendente y/o la información secundaria espacial, por ejemplo, la dirección de llegada y la difusividad para adaptar las señales de mezcla descendente y/o la información secundaria espacial para reflejar la posición y/u orientación de escucha modificada. En otras palabras, estas realizaciones no requieren ninguna información de configuración adicional, por ejemplo, información geométrica de las diferentes fuentes de audio con respecto a la posición de grabación original.

Las realizaciones de la presente invención reciben además señales de audio espacial paramétrico de acuerdo con un determinado formato de audio espacial, por ejemplo, señales mono o estéreo de mezcla descendente con dirección de llegada y difusividad como información secundaria espacial y convierten estos datos de acuerdo con señales de control, por ejemplo, señales de control de zoom o rotación y producen como salida los datos modificados o convertidos en el mismo formato de audio espacial, es decir, en forma de señal de mezcla descendente mono o estéreo con la dirección de llegada y parámetros de difusividad asociados.

En una realización específica, las realizaciones de la presente invención se acoplan a una cámara de vídeo u otras fuentes de vídeo y modifican los datos de audio espacial recibidos u originales en los datos de audio espacial modificados de acuerdo con las señales de control de zoom o control de rotación proporcionadas por la cámara de vídeo para sincronizar, por ejemplo, la experiencia de audio con la experiencia de vídeo y, por ejemplo, para efectuar un zoom acústico en el caso de realizarse un zoom de vídeo y/o realizar una rotación de audio dentro de la escena de audio en caso de que la cámara de vídeo se gire y los micrófonos no giren físicamente con la cámara debido a que no están montados en la cámara.

Breve descripción de las figuras

Las realizaciones de la presente invención se describirán en detalle utilizando las siguientes Figuras.

La Figura 1 ilustra un diagrama de bloques de un codificador de audio espacial paramétrico;

La Figura 2 ilustra el codificador de audio espacial de la Figura 1 junto con una realización del bloque de modificaciones de parámetros espaciales acoplado entre la unidad de análisis de audio espacial y la unidad de síntesis de audio espacial del codificador de audio espacial;

La Figura 3A corresponde a la Figura 2 e ilustra una realización más detallada del bloque de modificaciones de parámetros espaciales;

La Figura 3B corresponde a la Figura 2 e ilustra una realización más detallada del bloque de modificaciones de parámetros espaciales;

La Figura 4 ilustra una vista general geométrica ilustrativa de un zoom acústico;

La Figura 5A ilustra un ejemplo de función de mapeo direccional $f_p(k, n, \varphi, d)$ para el mapeo de dirección de llegada (DOA);

La Figura 5B ilustra un ejemplo de función de mapeo de difusividad $f_d(k, n, \varphi, d)$ para el mapeo de difusividad;

La Figura 6 ilustra diferentes ventanas de ganancia correspondientes al filtro de ponderación $H_1(k, n, \varphi, d)$ del componente de sonido directo dependiendo de un factor de zoom; y

La Figura 7 ilustra una ventana subcardioide para el filtro de ponderación $H_2(k, n, \varphi, d)$ para el componente difuso.

Los elementos iguales o equivalentes con funcionalidad igual o equivalente están indicados en la siguiente descripción de las Figuras por números de referencia iguales o equivalentes.

Descripción detallada de la invención

Para una mejor comprensión de las realizaciones de la presente invención, se describe un codificador de audio espacial típico. La tarea de un codificador de audio espacial paramétrico típico es reproducir la impresión espacial que estaba presente en el lugar donde se realizara la grabación. Por lo tanto, un codificador de audio espacial consiste en una parte de análisis 100 y una parte de síntesis 200, como se ilustra en la Figura 1. En el extremo del frente acústico, hay N micrófonos 102 dispuestos para obtener N señales de entrada de micrófono que se procesan por la unidad de análisis de audio espacial 100 para producir L señales de mezcla descendente 112 con $L \leq N$ junto con la información espacial secundaria 114. En el decodificador, es decir en la unidad de síntesis de audio espacial, se utiliza la señal de mezcla descendente 112 y la información espacial secundaria 114 para calcular M canales de altavoz para M altavoces 202, que reproducen el campo de sonido grabado con la impresión espacial original. Las

líneas gruesas (las líneas entre los micrófonos 102 y la unidad de análisis de audio espacial 100, las L señales de mezcla descendente 112 y las líneas de señal M entre la unidad de síntesis de audio espacial 200 y los M altavoces 202) simbolizan los datos de audio, mientras que las líneas finas 114 entre la unidad de análisis de audio espacial 100 y la unidad de síntesis de audio espacial 200 representan la información espacial secundaria.

5 A continuación, se describen en forma más detallada los pasos básicos incluidos en el cálculo de los parámetros espaciales o, en otras palabras, para el análisis de audio espacial realizado por la unidad de análisis de audio espacial 100. Las señales de micrófono se procesan en una representación adecuada de tiempo/frecuencia, por ejemplo, aplicando una Transformada de Fourier de tiempo reducido (STFT) o cualquier otro banco de filtros. La información espacial secundaria determinada en la etapa de análisis contiene una medida que corresponde a la dirección de llegada (DOA) del sonido y una medida de la difusividad del campo de sonido, que describe la relación entre el sonido directo y difuso del campo de sonido analizado.

15 En DirAC, se ha propuesto determinar la DOA del sonido como dirección opuesta del vector de intensidad activo. La información acústica relevante se deriva de una entrada del denominado micrófono de formato B, que corresponde a la presión de sonido y la velocidad obtenida por la configuración de micrófonos que produce un patrón de captación dipolo, que están alineados con los ejes del sistema de coordenadas cartesianas. En otras palabras, el formato B consiste en cuatro señales, es decir $w(t)$, $x(t)$, $y(t)$ y $z(t)$. La primera corresponde a la presión medida por un micrófono omnidireccional, mientras que las últimas tres son señales de los micrófonos que tienen patrones de captación de figura en ocho dirigidos hacia los tres ejes de un sistema de coordenadas cartesianas. Las señales $x(t)$, $y(t)$ y $z(t)$ son proporcionales a los componentes de los vectores de velocidad de partículas dirigidos hacia x , y y z , respectivamente. Por otro lado, el enfoque presentado en SAM utiliza un conocimiento previo del patrón de direccionalidad de los micrófonos estéreo para determinar la DOA del sonido.

25 Se puede obtener la medida de la difusividad relacionando la intensidad activa del sonido con la energía total del campo de sonido como se propone en DirAC. Por otro lado, el método como se describe en SAM propone la evaluación de la coherencia entre diferentes señales de micrófono. Se debe tener en cuenta que la difusividad también se podría considerar una medida general de la fiabilidad de la DOA estimada. Sin pérdida de generalidad, a continuación se supone que la difusividad radica en el rango de $[1, 0]$, donde un valor de 1 indica un campo de sonido puramente difuso y un valor de 0 corresponde al caso en que solo está presente el sonido directo. En otras realizaciones, se pueden utilizar otros rangos y valores para la difusividad.

35 La señal de mezcla descendente 112, que va acompañada por la información secundaria 114, se calcula a partir de las señales de entrada de micrófono. Puede ser mono o incluir múltiples canales de audio. En el caso de DirAC, por lo general se considera solo una señal mono, que corresponde a la presión del sonido, obtenido por un micrófono omnidireccional. En el caso del enfoque SAM, se utiliza una señal estéreo de dos canales como señal de mezcla descendente.

40 A continuación, se describe en forma más detallada la síntesis de las señales de altavoz empleadas para la reproducción efectuada por la unidad de síntesis de audio espacial 200. La entrada de la síntesis 200 es la señal de mezcla descendente 112 y los parámetros espaciales 114 en su representación de tiempo-frecuencia. A partir de estos datos, se calculan M canales de altavoces de tal manera que la imagen de audio espacial o la impresión de audio espacial se reproduzca correctamente. Siendo $Y_i(k, n)$, donde $i = 1 \dots M$, denota la señal del canal de altavoz físico i -ésimo en la representación de tiempo/frecuencia con los índices de tiempo y frecuencia k y n , respectivamente. El modelo de señal subyacente para la síntesis se proporciona por

$$Y_i(k, n) = g_i(k, n)S(k, n) + D_i\{N(k, n)\} \quad (1)$$

50 donde $S(k, n)$ corresponde al componente de sonido directo y $N(k, n)$ representa el componente de sonido difuso. Obsérvese que para la reproducción correcta del sonido difuso, se aplica una operación de descorrelación $D\{\}$ al componente difuso de cada canal de altavoz. El factor de escala $g_i(k, n)$ depende de la DOA del sonido directo incluido en la información secundaria y de la configuración de altavoces empleados para la reproducción. Una elección adecuada se proporciona por el enfoque de panorámica en amplitud base por vectores propuesta por Pulkki, V., "Virtual sound source positioning using vector base amplitude panning", J. Audio Eng. Soc., Vol. 45, págs. 456-466, junio de 1997, a la que en lo sucesivo se hace referencia como [VBAP].

55 En DirAC, el componente de sonido directo se determina mediante la escala apropiada de la señal de mezcla descendente mono $W(k, n)$, y se obtiene de acuerdo con:

$$60 \quad S(k, n) = W(k, n)\sqrt{1 - \Psi(k, n)} \quad (2)$$

El componente de sonido difuso se obtiene de acuerdo con

$$N(k, n) = \frac{1}{\sqrt{M}} W(k, n) \cdot \sqrt{\Psi(k, n)} \quad (3)$$

5 donde M es el número de altavoces utilizado.

En SAM, se aplica el mismo modelo de señal que en (1), aunque en su lugar se calculan los componentes de sonido directo y difuso basándose en las señales de mezcla descendente estéreo.

10 La Figura 2 ilustra un diagrama de bloques de una realización de la presente invención integrada en el entorno
ejemplar de la Figura 1, es decir, integrado entre una unidad de análisis espacial 100 y una unidad de síntesis de
audio espacial 200. Como se explica basándose en la Figura 1, se graba la escena de audio original con una
configuración de micrófonos de grabación específica que especifica la ubicación y orientación (en el caso de los
micrófonos direccionales) con respecto a las diferentes fuentes de sonido de audio. Los N micrófonos producen
15 señales de micrófonos físicos o señales de canales, que se procesan por la unidad de análisis de audio espacial 100
para generar una o varias señales de mezcla descendente W 112 y la información espacial secundaria 114, por
ejemplo, la dirección de llegada (DOA) φ 114a y la difusividad ψ 114b. A diferencia de la Figura 1, estas señales de
audio espacial 112, 114a, 114b no se proporcionan directamente a la unidad de síntesis de audio espacial 200, sino
que se modifican por un aparato para convertir o modificar una primera señal de audio espacial paramétrico 112,
20 114a, 114b que representa una primera posición de escucha y/o una primera orientación de escucha (en este
ejemplo, la posición de grabación y la orientación de grabación) en una escena de audio espacial en una segunda
señal de audio espacial paramétrico 212, 214a, 214b, es decir una señal de mezcla descendente modificada W_{mod}
212, una señal de dirección de llegada modificada φ_{mod} 214a y/o una señal de difusividad modificada ψ_{mod} 214b que
representa una segunda posición de escucha y/o una segunda orientación de escucha que es diferente de la
25 primera posición de escucha y/o la primera orientación de escucha. También se hace referencia a dirección de
llegada modificada 214a y la difusividad modificada 214b como información de audio espacial modificada 214.
También se hace referencia al aparato 300 como una unidad de modificación de audio espacial o bloque de
modificaciones de señal de audio 300. El aparato 300 de la Figura 3A está adaptado para modificar la primera señal
de audio espacial paramétrico 112, 114 dependiendo de la señal de control d 402 proporcionada por, por ejemplo,
30 una unidad de control externa 400. La señal de control 402 puede ser, por ejemplo una señal de control de zoom
que define o es un factor de zoom d o un parámetro de zoom d , o una señal de control de rotación 402
proporcionada por una unidad de control de zoom y/o de control de rotación 400 de una cámara de vídeo. Se debe
tener en cuenta que un zoom en una determinada dirección y una traslación en la misma dirección son solo dos
maneras diferentes de describir un movimiento virtual en esa dirección determinada (el zoom por un factor de zoom,
35 la traslación por una distancia absoluta o por una distancia relativa con respecto a una distancia de referencia). Por
lo tanto, las explicaciones presentadas en el presente documento con respecto a una señal de control de zoom se
aplican, en correspondencia, a las señales de control de traslación y viceversa, y la señal de control de zoom 402
también se refiere a una señal de control de traslación. El término d puede representar, por un lado, la propia señal
de control 402, y por el otro, la información o parámetro de control contenido en la señal de control. En otras
40 realizaciones, el parámetro de control d ya representa la señal de control 402. El parámetro de control o la
información de control d puede ser una distancia, un factor de zoom y/o un ángulo de rotación y/o una dirección de
rotación.

45 Como se puede apreciar en la Figura 2, el aparato 300 está adaptado para producir señales de audio espacial
paramétrico 212, 214 (señales de mezcla descendente y la información secundaria/parámetros asociados) en el
mismo formato que se recibieron las señales de audio espacial paramétrico 112, 114. Por lo tanto, la unidad de
síntesis de audio espacial 200 tiene la capacidad (sin modificaciones) de procesar la señal de audio espacial
modificada 212, 214 de la misma manera que la señal de audio espacial original o grabada 112, 114 y de
convertirlas a M señales de altavoces físicos 204 para generar la experiencia de sonido en la escena de audio
50 espacial modificada o, en otras palabras, en la posición de escucha modificada y/o la orientación de escucha
modificada dentro de la escena de audio espacial de otro modo inalterada.

En otras palabras, en la Figura 2 se ilustra un diagrama esquemático de bloques de una realización del aparato o
método novedosos. Como se puede apreciar, la salida 112, 114 del codificador de audio espacial 100 se modifica
55 basándose en la información de control externo 402 para obtener una representación de audio espacial 212, 214
que corresponde a una posición de escucha, que es diferente de la utilizada en la ubicación original utilizada para la
captura del sonido. Más precisamente, tanto las señales de mezcla descendente 112 como la información espacial
secundaria 114 se modifican de manera apropiada. La estrategia de modificación está determinada por un control
externo 400, que puede adquirirse directamente de una cámara 400 o de cualquier otra interfaz del usuario 400 que
60 ofrezca información acerca de la posición real de la cámara o zoom. En esta realización, la tarea del algoritmo,
respectivamente de la unidad de modificación 300 es cambiar la impresión espacial de la escena de sonido de la

misma manera que el zoom óptico o la rotación de la cámara cambia el punto de vista del espectador. En otras palabras, la unidad de modificación 300 está adaptada para producir un zoom acústico o experiencia de rotación de audio correspondiente al zoom de vídeo o rotación de vídeo.

- 5 La Figura 3A ilustra un diagrama de bloques o vista general del sistema de una realización del aparato 300 al que se hace referencia como "unidad de zoom acústico". La realización del aparato 300 de la Figura 3A comprende una unidad de modificación de parámetros 301 y una unidad de modificación de mezcla descendente 302. La unidad de modificación de parámetros 301 comprende además una unidad de modificación de dirección de llegada 301a y una unidad de modificación de difusividad 301b. La unidad de modificación de parámetros 301 está adaptada para recibir el parámetro de dirección de llegada 114a y modificar el primer parámetro de dirección de llegada o recibido 114a dependiendo de la señal de control d 402 para obtener el segundo parámetro de dirección de llegada o modificado 214a. La unidad de modificación de parámetros 301 también está adaptada para recibir el primer parámetro de difusividad (u original) 114b y para modificar el parámetro de difusividad 114b mediante la unidad de modificación de difusividad 301b para obtener el segundo parámetro de difusividad o modificado 214b dependiendo de la señal de control 402. La unidad de modificación de mezcla descendente 302 está adaptada para recibir la una o más señales de mezcla descendente 112 y modificar la primera señal de mezcla descendente u original 112 para obtener la segunda señal de mezcla descendente o modificada 212 dependiendo del primer parámetro de dirección de llegada u original 114a, el primer parámetro de difusividad u original 114b y/o la señal de control 402.
- 10
- 15
- 20 Si se controla la cámara independientemente de los micrófonos 102, las realizaciones de la invención ofrecen la posibilidad de sincronizar el cambio de la escena de audio o la percepción de audio de acuerdo con los controles de la cámara 402. Además, se pueden desplazar las direcciones sin modificar las señales de mezcla descendente 112 si solo se hace girar la cámara 400 horizontalmente sin zoom, es decir aplicando solo una señal de control de rotación y no una señal de control de zoom 402. Esto está descrito por el "controlador de rotación" en las Figuras 2 y 3.
- 25

La modificación de rotación se describe en forma más detallada en la sección sobre el remapeo direccional o remapeo de direcciones. Las secciones sobre la modificación de la difusividad y mezcla descendente se relacionan con la aplicación de traslación o zoom.

- 30 Las realizaciones de la invención pueden adaptarse para realizar ambas, una modificación de la rotación y una modificación de la traslación o zoom, por ejemplo efectuando en primer lugar la modificación de la rotación y posteriormente la modificación de la traslación o zoom o viceversa, o ambos al mismo tiempo proporcionando las correspondientes funciones de mapeo direccional.
- 35

Para obtener el efecto de zoom acústico, se cambia virtualmente la posición de escucha, lo que se hace remapeando apropiadamente las direcciones analizadas. Para obtener una impresión general correcta de la escena de sonido modificada, la señal de mezcla descendente es procesada por un filtro, que depende de las direcciones remapeadas. Este filtro modifica las ganancias, como por ejemplo, se incrementa el nivel de los sonidos, que ahora están más cerca, mientras que los sonidos que están fuera de las regiones de interés se pueden atenuar. Además, se escala la difusividad usando los mismos supuestos, como por ejemplo los sonidos que parecen más cercanos a la posición de escucha deben ser reproducidos en forma menos difusa que antes.

40

A continuación, se presenta una descripción más detallada del algoritmo o método ejecutado por el aparato 300. En la Figura 3A se presenta una vista general de la unidad de zoom acústico. En primer lugar, se describe el remapeo de las direcciones (bloque 301a, $f_p(k,n,\varphi,d)$), luego se ilustra el filtro para la modificación de la difusividad (bloque 301b, $f_d(k,n,\varphi,d)$). El bloque 302 describe la modificación de la mezcla descendente, que depende del control de zoom y los parámetros espaciales originales.

- 50 En la siguiente sección, se describe el remapeo de las direcciones, especialmente el remapeo de los parámetros de dirección de llegada realizado, por ejemplo, por el bloque de modificación de la dirección 301a.

El parámetro de dirección de llegada (parámetro de DOA) puede representarse, por ejemplo, por un vector de unidad e. En el caso de un análisis de campo de sonido tridimensional (3D) el vector puede expresarse de la siguiente manera

55

$$\mathbf{e} = \begin{bmatrix} \cos \varphi \cos \theta \\ \sin \varphi \cos \theta \\ \sin \theta \end{bmatrix} \quad (4)$$

donde el ángulo de azimut φ corresponde a la DOA en el plano bidimensional (2D), es decir el plano horizontal. El

ángulo de elevación se proporciona por θ . Este vector se modificará, de acuerdo con la nueva posición virtual del micrófono, como se describe a continuación.

5 Sin pérdida de generalidad, se proporciona un ejemplo de remapeo de DOA para el caso bidimensional para simplificar la presentación (Figura 4). Se puede efectuar el correspondiente remapeo de la DOA tridimensional con consideraciones similares.

10 La Figura 4 ilustra una vista general geométrica de una vista general geométrica ejemplar del zoom acústico. La posición S marca la posición original de grabación del micrófono, es decir la posición de escucha original. A y B marcan posiciones espaciales dentro del plano bidimensional observado. Ahora se supone que la posición de escucha se mueve de S a S2, por ejemplo en la dirección de la primera orientación de escucha. Como se puede apreciar en la Figura 4, el sonido que emerge de la posición espacial A se mantiene en la misma posición angular con respecto al sitio de la grabación, mientras que los sonidos procedentes del área o posición espacial B se mueven al lado. Esto está indicado por un cambio del ángulo analizado α a β . P indica, por lo tanto, la dirección de llegada del sonido que llega de la posición angular de B si el oyente se ha ubicado en S2. En el caso del ejemplo en cuestión, se aumenta el ángulo azimutal de α a β como se ilustra en la Figura 4. Este remapeo de la información de dirección de llegada se puede expresar como una transformación por vectores de acuerdo con

$$\mathbf{e}_{\text{mod}} = \mathbf{f}(\mathbf{e}), \quad (5)$$

20 donde $\mathbf{f}(\)$ indica una función de remapeo y \mathbf{e}_{mod} es el vector de dirección modificado. Esta función es una transformación no lineal, que depende del factor de zoom d y las DOA originales estimadas. La Figura 5A ilustra ejemplos para el mapeo $\mathbf{f}(\)$ de diferentes valores de α que se pueden aplicar en el ejemplo bidimensional expuesto en la Figura 4. En el caso del factor de control de zoom de $d = 1$, es decir, no se aplica zoom, los ángulos son iguales a la DOA original α . Para aumentar los factores de control de zoom, el valor de β se incrementa también. La función puede obtenerse de factores geométricos o, como alternativa, se puede elegir de manera heurística. Por consiguiente, el remapeo de las direcciones significa que cada DOA se modifica de acuerdo con la función $\mathbf{f}(\)$. El mapeo $f_p(k,n,\varphi,d)$ se realiza por cada componente de tiempo y frecuencia (k,n) .

30 Si bien en la Figura 4 se ilustra el parámetro de zoom d en términos de distancia de traslación entre la posición de escucha original S y la posición de escucha modificada S2, como se ha mencionado anteriormente, d también puede ser un factor, por ejemplo un zoom óptico tal como un zoom de 4x u 8x. Especialmente para el control del ancho o filtro, tomando a d como factor, no como distancia, se obtiene una fácil implementación del zoom acústico. En otras palabras, el parámetro de zoom d es, en este caso, una distancia real, o al menos proporcional a una distancia.

35 Se debe indicar asimismo que las realizaciones de la invención se pueden adaptar además para soportar, además del zoom de acercamiento ("zoom-in") como se ha descrito anteriormente, por ejemplo reduciendo la distancia hasta un objeto (por ejemplo hasta el objeto A de la Figura 4 moviéndose de la posición S a la posición S2), también un zoom de alejamiento ("zoom-out"), por ejemplo aumentando una distancia hasta un objeto (por ejemplo hasta el objeto A de la Figura 4 desde la posición S2 a la posición S). En este caso se aplican consideraciones inversas en comparación al zoom de acercamiento descrito, puesto que los objetos situados en un lado del oyente (por ejemplo el objeto B con respecto a la posición S2) se mueven hacia el frente del oyente al desplazarse éste a la posición S. En otras palabras, las magnitudes de los ángulos se reducen (por ejemplo de β a α).

45 El remapeo de la transformación de direcciones o vectores se realiza por la unidad de modificación de dirección de llegada 301a. La Figura 5A ilustra una función de mapeo ilustrativa (que depende del factor de zoom d) para las direcciones de llegada de la situación ilustrada en la Figura 4. El diagrama de la Figura 5A ilustra el factor de zoom en el eje x en el rango de 1 a 2 y el ángulo modificado o mapeado β en el eje y. En el caso de un factor de zoom de 1, $\beta = \alpha$, es decir que el ángulo inicial no se modifica. El número de referencia 512 se refiere a la función de mapeo f_p correspondiente a $\alpha = 10^\circ$, el número de referencia 514 representa la función de mapeo f_p correspondiente a $\alpha = 30^\circ$, el número de referencia 516 representa la función de mapeo $f_p(k,n,\varphi,d)$ correspondiente a $\alpha = 50^\circ$, el número de referencia 518 es la función de mapeo $f_p(k,n,\varphi,d)$ correspondiente a $\alpha = 70^\circ$ y el número de referencia 520 representa la función de mapeo $f_p(k,n,\varphi,d)$ correspondiente a $\alpha = 90^\circ$.

55 Las realizaciones de la invención se pueden adaptar para usar la misma función de mapeo f_p para todos los valores de componentes de tiempo y frecuencia definidas por k y n , o bien pueden emplear diferentes funciones de mapeo para diferentes valores de tiempo y/o componentes de frecuencia.

60 Como resulta evidente a partir de las explicaciones anteriores, la idea detrás del filtro f_d consiste en cambiar la difusividad ψ de tal manera que reduzca la difusividad para las direcciones acercadas ($\varphi < |\gamma|$) y aumenta la difusividad para las direcciones fuera de foco ($\varphi > |\gamma|$).

Para simplificar la determinación del ángulo mapeado β , ciertas realizaciones de la unidad de modificación 301a

están adaptadas para usar solamente la dirección y suponer que todas las fuentes, por ejemplo A y B, que definen la dirección de llegada del sonido tienen la misma distancia hasta la primera posición de escucha, por ejemplo, están dispuestas en un radio unitario.

5 Si se considera una configuración de altavoces que solo reproduce el sonido procedente de direcciones frontales, por ejemplo, una configuración típica de altavoces estéreo, se puede diseñar la función de mapeo $f(\cdot)$ de tal manera que el ángulo máximo respecto del cual se remapean las DOA esté limitado. Por ejemplo, se elige un ángulo máximo de $\pm 60^\circ$, cuando los altavoces están colocados a $\pm 60^\circ$. De esta manera, toda la escena de sonido se mantiene en el frente y solo se amplía cuando se aplica el zoom.

10 En el caso de una rotación de la cámara, los valores azimutales originales simplemente se desplazan de tal manera que la nueva dirección de visualización corresponda a un ángulo de cero. De esa manera, una rotación horizontal de la cámara en 20° daría como resultado $\beta = \alpha - 20^\circ$. Además, no se modifica la mezcla descendente ni la difusividad en este caso especial, a menos que se lleve a cabo simultáneamente la rotación y traslación.

15 Como se puede observar a partir de las explicaciones anteriormente mencionadas, el cambio o diferencia de rotación se obtiene a partir de la primera orientación de escucha, visualizando respectivamente en primer lugar la orientación (por ejemplo la dirección de la "nariz" del oyente o respectivo espectador) que define una primera orientación de referencia o 0° . Cuando cambia la orientación de escucha, la orientación de referencia o 0° cambia de manera correspondiente. Por lo tanto, las realizaciones de la presente invención cambian los ángulos o direcciones de llegada del sonido originales, es decir el primer parámetro direccional de acuerdo con la nueva orientación de referencia o 0° por lo que el segundo parámetro direccional representa la misma "dirección de llegada" en la escena de audio, aunque con respecto a la nueva orientación de referencia o sistema de coordenadas. Se aplican similares consideraciones a la traslación o zoom, respectivamente, donde las direcciones de llegada percibidas cambian debido a la traslación o zoom en la dirección de la primera orientación de escucha (véase la Figura 4).

20 El primer parámetro direccional 114a y el segundo parámetro direccional 214a pueden ser vectores bidimensionales o tridimensionales. Además, el primer parámetro direccional 114a puede ser un vector, donde la señal de control 402 es una señal de control de rotación que define un ángulo de rotación (por ejemplo 20° en el ejemplo anteriormente mencionado) y una dirección de rotación (a la derecha en el ejemplo bidimensional anteriormente mencionado), y donde la unidad de modificación de parámetros 301, 301a está adaptada para hacer rotar el vector en el ángulo de rotación en dirección inversa a la dirección de rotación ($\beta = \alpha - 20^\circ$ en el ejemplo anteriormente mencionado) para obtener el segundo parámetro direccional, es decir el segundo vector o modificado 214a.

35 En la siguiente sección se describe en forma más detallada la escala de difusividad como se realiza, por ejemplo, por la unidad de modificación de difusividad 301b.

40 La difusividad se escala con una ventana dependiente de la DOA. En ciertas realizaciones, los valores de la difusividad $\psi(k, n)$ se reducen para las direcciones acercadas, mientras que se incrementan los valores de difusividad correspondientes a las direcciones sin interés. Esto corresponde a la observación de que las fuentes de sonido se perciben menos difusas si están situadas más cerca de la posición de escucha. Por lo tanto, por ejemplo, para un factor de zoom mínimo (por ejemplo $d = 1$), la difusividad no se modifica. Se puede tomar el rango del ángulo visual cubierto por la imagen de la cámara como un controlador para la escala por la cual se incrementa o reduce el valor de difusividad.

45 Los términos direcciones de zoom de acercamiento o direcciones de interés se refieren a una ventana angular de interés, a la que también se hace referencia como rango central de ángulos, que está dispuesto alrededor de la primera dirección de escucha o dirección original, por ejemplo la dirección de referencia original 0° . La ventana angular o rango central está determinado por los valores angulares γ que definen el borde de la ventana angular. La ventana angular y el ancho de la ventana angular pueden definirse por el ángulo de borde negativo $-\gamma$ y el ángulo de borde positivo γ , donde la magnitud del ángulo de borde negativo puede ser diferente del ángulo de borde positivo. En las realizaciones preferidas, el ángulo de borde negativo y el ángulo de borde positivo tienen la misma magnitud (ventana simétrica o rango central de ángulos centrado alrededor de la primera orientación de escucha). La magnitud del ángulo de borde también se denomina ancho angular y el ancho de la ventana (del ángulo de borde negativo al ángulo de borde positivo) también se denomina ancho angular total.

50 De acuerdo con las realizaciones de la invención, los parámetros de dirección de llegada, los parámetros de difusividad, y/o los componentes directos o difusos se pueden modificar de manera diferente dependiendo de si el parámetro original de dirección de llegada está dentro de la ventana de interés, por ejemplo si el ángulo de DOA o la magnitud del ángulo de DOA con respecto a la primera posición de escucha es menor que la magnitud del ángulo de borde o la amplitud angular γ , o si el parámetro original de dirección de llegada está fuera de la ventana de interés, por ejemplo si el ángulo de DOA o la magnitud del ángulo de DOA con respecto a la primera posición de escucha es mayor que la magnitud del ángulo de borde o la amplitud angular γ . También se hace referencia a esto como

dependiente de la dirección y las correspondientes funciones de filtro como funciones de filtro dependientes de la dirección, donde la amplitud angular o ángulo de borde γ define el ángulo en el cual el filtro correspondiente cambia de aumento del parámetro a reducción del parámetro o viceversa.

5 Haciendo referencia de nuevo, a la unidad de modificación de difusividad 301b, la unidad de modificación de difusividad 301b está adaptada para modificar la difusividad ψ por la función $f_d(k,n,\varphi,d)$ o f_d que depende de los índices de tiempo/frecuencia k,n , la dirección de llegada original φ y el controlador de zoom d . La Figura 5B ilustra una realización de una función de filtro f_d . El filtro f_d se puede implementar en forma de inversión de la función de filtro H_1 , que se explica más adelante, adaptado, sin embargo, para equipararse al rango de difusividad, por ejemplo el rango de $[0..1]$. La Figura 5B ilustra la función de mapeo o filtro f_d , donde el eje x representa la primera difusividad o difusividad original ψ , de la Figura 5B a la que también se refiere como ψ_{in} , donde el rango de 0 a 1, y el eje y representa la segunda difusividad o modificada ψ_{mod} también en el rango de 0 a 1. En caso de no aplicarse zoom ($d = 0$), el filtro f_d no cambia la difusividad en absoluto y se ajusta para desviarse, es decir $\psi_{mod} = \psi_{in}$ respectivamente. El número de referencia 552 ilustra la línea de desviación.

15 Si la dirección de llegada original radica dentro de la amplitud angular γ , la difusividad se reduce. Si la dirección de llegada original está fuera de la amplitud angular γ , la difusividad se incrementa. La Figura 5B ilustra algunas funciones prototipo de f_d , es decir 562, 564, 572 y 574 que dependen del ancho de visión o de la amplitud angular γ . En el ejemplo ilustrado en la Figura 5B, la amplitud angular es menor para γ_2 que para γ_1 , es decir $\gamma_2 < \gamma_1$. Por consiguiente, γ_2 corresponde a un factor de zoom más alto que γ_1 .

20 El área por debajo de la línea de desviación 552 define los valores de difusividad modificados ψ_{mod} en caso de que la dirección de llegada original φ esté dentro de la amplitud angular γ que se refleja por una reducción del valor de difusividad modificado ψ_{mod} en comparación con el valor de difusividad original ψ_{in} o ψ después del mapeo realizado por el filtro f_d . El área por encima de la línea de desviación 552 representa el mapeo de la difusividad original ψ a los valores de difusividad modificada ψ_{mod} en caso de que la dirección de llegada original φ esté fuera de la ventana. En otras palabras, el área por encima de la línea de desviación 552 ilustra el aumento de la difusividad después del mapeo. En las realizaciones preferidas, la amplitud angular γ se reduce con un aumento del factor de zoom d . En otras palabras, cuanto mayor es un factor de zoom d , menor es la amplitud angular γ . Además, las formas de realización se pueden adaptar de tal manera que el factor d o la información de traslación no solo influye sobre la amplitud angular γ de la función de filtro f_d sino que también se incrementa el grado o factor de difusividad en caso de estar dentro de la ventana y se reduce el grado o factor de difusividad ψ en caso de estar fuera de la ventana definida por la amplitud angular γ . Una realización de este tipo se ilustra en la Figura 5B, en la cual la amplitud angular γ_1 corresponde a un factor de zoom d_1 y la amplitud angular γ_2 corresponde a un factor de zoom d_2 , donde d_2 es mayor que d_1 y, por consiguiente, la amplitud angular γ_2 es menor que la amplitud angular γ_1 . Además, la función f_d representada por el número de referencia 564 y que corresponde al factor de zoom mayor d_2 mapea los valores de difusividad originales ψ_{in} con los valores de difusividad modificada más bajos ψ_{mod} que la función de filtro f_d representada por 562 que corresponde al factor de zoom más bajo d_1 . En otras palabras, las realizaciones de la función de filtro puede ser adaptada para reducir la difusividad original más cuanto menor es la amplitud angular γ . Lo siguiente se aplica al área por encima de la línea de desviación 552 de una manera inversa. En otras palabras, las realizaciones de la función de filtro f_d se pueden adaptar para mapear la difusividad original ψ_{in} a la difusividad modificada ψ_{mod} dependiendo del factor de zoom d y la amplitud angular γ , o a mayor factor de zoom d menor amplitud angular γ y/o mayor aumento de la difusividad de la dirección de llegada φ fuera de la ventana.

45 En otras realizaciones, se aplica la misma ventana o función de filtro dependiente de la dirección $f_d(k,n,\varphi,d)$ para todos los factores de zoom. Sin embargo, el uso de una ventana dependiente de la dirección o funciones de filtro diferentes con menores amplitudes angulares para factores de traslación o zoom más elevados se equipara mejor a la experiencia de audio del usuario y ofrece una percepción de audio más realista. La aplicación de diferentes valores de mapeo por diferentes factores de zoom (una mayor reducción de la difusividad con el aumento del factor de zoom para el valor de dirección de llegada φ dentro de la ventana y valores crecientes o de mayor difusividad para factores de zoom más elevados en caso de que el valor de dirección de llegada φ esté fuera de la amplitud ancho angular γ) incluso mejora aún más la percepción de audio realista.

50 A continuación se describen en forma más detallada las realizaciones de la modificación de mezcla descendente, como por ejemplo la ejecutada por la unidad de modificación de mezcla descendente 302.

55 Se utilizan filtros para la señal de mezcla descendente para modificar la ganancia de la parte directa y difusa de la señal de salida. Como consecuencia directa del concepto de codificador de audio espacial, se modifican así las señales de altavoces. El sonido del área acercada se amplifica, mientras que el sonido de las direcciones fuera de la zona de interés se puede atenuar.

60 Como la señal de mezcla descendente 112 puede ser una señal mono o una estéreo para micrófonos de

codificación de audio direccional (DirAC) o de audio espacial (SAM), a continuación se describen dos realizaciones diferentes de la modificación.

5 En primer lugar se describe una realización correspondiente a una modificación de mezcla descendente mono, es decir una realización correspondiente a una modificación de una señal de audio de mezcla descendente mono W 112. Con respecto a las siguientes consideraciones, es ventajoso introducir un modelo de señal de la señal de mezcla descendente mono $W(k,n)$ que es similar al ya aplicado para la síntesis de las señales de altavoces de acuerdo con (1):

$$10 \quad W(k,n) = S(k,n) + N(k,n) \quad (6)$$

15 En este caso, $S(k,n)$ indica el componente de sonido directo de la señal de mezcla descendente, $N(k,n)$ indica los componentes de sonido difuso de la señal de mezcla descendente original y k indica el índice temporal o el instante de tiempo que la señal representa y n representa un componente de frecuencia o canal de frecuencia de la señal en el instante de tiempo dado k .

Supongamos que $W_{\text{mod}}(k,n)$ indica la señal de mezcla descendente mono modificada. Se obtiene procesando la señal de mezcla descendente original de acuerdo con

$$20 \quad W_{\text{mod}}(k,n) = H_1(k,n,\varphi,d)S(k,n) + H_2(k,n,\varphi,d)N(k,n) \quad (7)$$

donde $H_1(k,n,\varphi,d)$ y $H_2(k,n,\varphi,d)$ representan filtros aplicados a los componentes directos y difusos del modelo de señal, φ representa la dirección de llegada original y d el factor de zoom o parámetro de zoom. Los componentes de sonido directos 112a y difusos 112b se pueden calcular de manera similar a (2), (3), es decir según

$$25 \quad S(k,n) = W(k,n)\sqrt{1-\Psi}$$

y

$$30 \quad N(k,n) = W(k,n)\sqrt{\Psi}$$

Ambos filtros son direccionales dependiendo de las funciones de ponderación. Por ejemplo, se puede tomar un patrón de captación de micrófono en forma de cardioide como criterio de diseño para dichas funciones de ponderación.

35 El filtro $H_1(k,n,\varphi,d)$ se puede implementar en forma de ventana de coseno elevada de tal manera que el sonido directo se amplifique para las direcciones del área acercada, mientras que el nivel del sonido procedente de las demás direcciones se atenúa. En general, se pueden aplicar diferentes formas de ventana a los componentes de sonido directo y difuso, respectivamente.

40 El filtro de ganancia implementado por las ventanas puede controlarse por la translación real o el factor de control de zoom d . Por ejemplo, el zoom controla la amplitud de igual ganancia correspondiente a las direcciones enfocadas con la amplitud de la ganancia en general. En la Figura 6 se presentan ejemplos correspondientes a diferentes ventanas de ganancia.

45 La Figura 6 ilustra diferentes ventanas de ganancia para el filtro de ponderación $H_1(k,n,\varphi,d)$. Se ilustran cuatro prototipos diferentes de ganancia:

1. línea plena: no se aplica zoom, la ganancia es 0 dB en todas las direcciones (véase 612).
2. línea de guiones: se aplica un factor de zoom de 1,3, la amplitud de ventana es una amplitud de 210° en el caso de la ganancia máxima y la ganancia máxima es de 2,3 dB (véase 614).
3. línea de puntos: se aplica un factor de zoom de 2,1, la amplitud de ventana para la ganancia máxima se reduce a 140° y la ganancia máxima es de 3 dB, la más baja -2,5 dB (véase 616).
4. línea de guiones y puntos: el factor de zoom es 2,8, la amplitud de ventana es 30° en el caso de la ganancia máxima y la ganancia esté limitada a un máximo de +3 dB y un mínimo -6 dB (véase 618).

55 Como se puede apreciar en la Figura 6, la primera orientación de escucha representada por 0° en la Figura 6, forma el centro de diferentes ventanas dependientes de los factores de zoom dependientes de la dirección, donde el rango central predeterminado o anchura de las ventanas dependientes de la dirección es menor cuanto mayor sea el factor de zoom. Los límites del rango o ventana central están definidos por el ángulo γ en el cual la ganancia es 0 dB. La Figura 6 ilustra ventanas simétricas con límites positivo y negativo que tienen la misma magnitud.

60

La ventana 614 tiene una amplitud de 210° en el caso de la ganancia máxima y una región central predeterminada con una amplitud de 260° con los límites +/- γ_2 a +/- 130°, donde los componentes directos dentro o en el interior de la región central predeterminada se incrementan y los componentes directos fuera de la región central predeterminada se mantienen inalterados (ganancia = 0 dB).

La ventana 616 tiene una amplitud de 140° para la ganancia máxima y una región central predeterminada con una amplitud de 180° con límites o amplitudes angulares +/- γ_3 a +/- 90°, donde los componentes directos dentro o en el interior de la región central predeterminada se incrementan y los componentes directos fuera de la región central predeterminada se reducen (ganancia negativa que desciende a -2,5 dB).

La ventana 618 tiene una amplitud de 30° para la ganancia máxima y una región central predeterminada con una amplitud de 60° con límites o amplitudes angulares +/- γ_4 a +/- 30°, donde los componentes directos dentro o en el interior de la región central predeterminada se incrementan y los componentes directos fuera de la región central predeterminada se reducen (ganancia negativa que desciende a -6 dB).

Por lo tanto, en ciertas realizaciones el factor de zoom d controla la amplitud, es decir los límites negativo y positivo y la amplitud total, y la ganancia de las ventanas prototipo. Por consiguiente, ya se puede diseñar la ventana de tal manera que se aplique la amplitud y la ganancia correctamente a las direcciones de llegada originales φ .

La ganancia máxima debería limitarse para evitar distorsiones en las señales de salida. La amplitud de la ventana, o la forma exacta como se muestra en este punto se debe considerar como un ejemplo ilustrativo de cómo el factor de zoom controla diversos aspectos de una ventana de ganancia. Se puede utilizar otra implementación en diferentes realizaciones.

El filtro $H_2(k, n, \varphi, d)$ se utiliza para modificar la parte difusa 112a de la señal de mezcla descendente de manera análoga a la manera en que se ha modificado la medida de la difusividad $\psi(k, n)$ y puede implementarse como una ventana subcardioide, como se ilustra en la Figura 7. Al aplicar ese tipo de ventanas se atenúa ligeramente la parte difusa de las direcciones fuera de interés, aunque las direcciones acercadas se mantienen sin modificación o casi inalteradas. La Figura 7 ilustra una ventana subcardioide 702 que casi mantiene el componente difuso inalterado en un área entre -30° y +30° de la dirección de llegada original φ y atenúa el componente difuso a mayor desviación, es decir el ángulo que se aleja de la orientación 0°, de la dirección de llegada original φ . En otras palabras, en el caso del área acercada, los componentes de la señal difusa en la señal de mezcla descendente se mantienen inalterados. Esto dará como resultado una reproducción más directa del sonido en la dirección de zoom. Los sonidos que proceden de todas las demás direcciones se representan más difusos, ya que el micrófono se ha colocado virtualmente en un sitio más alejado. De esa manera, esas partes difusas se atenuarán en comparación con las de la señal de mezcla descendente original. Evidentemente, también se puede diseñar el filtro de ganancia deseado utilizando las ventanas de coseno elevado anteriormente descritas. Obsérvese, sin embargo, que la escala será menos pronunciada que en el caso de la modificación del sonido directo. En otras realizaciones, las ventanas pueden depender del factor de zoom, donde la pendiente de la función de ventana 702 es más abrupta cuanto mayor sea el factor de zoom.

A continuación, se describe una realización de una modificación de mezcla descendente estéreo, es decir una modificación de una señal de mezcla descendente estéreo W .

A continuación se describe cómo se ha de realizar la modificación de mezcla descendente en el caso de una mezcla descendente estéreo según se requiere para el enfoque SAM. Para la señal de mezcla descendente estéreo original se introduce un modelo de señal de dos canales de manera análoga al caso mono (6):

$$W_1(k, n) = S(k, n) + N_1(k, n) \tag{8}$$

$$W_2(k, n) = cS(k, n) + N_2(k, n) \tag{9}$$

De nuevo, la señal $S(k, n)$ representa el sonido directo, mientras que N_i indica el sonido difuso correspondiente al i -ésimo micrófono. De manera análoga a (2), (3), los componentes de sonido directo y difuso se pueden determinar a partir de los canales de mezcla descendente basándose en la medida de la difusividad. El factor de ganancia c corresponde a una escala diferente del componente de sonido directo en los diferentes canales estéreo, lo que surge del diferente patrón de direccionalidad asociado a los dos canales de mezcla descendente. Se pueden encontrar más detalles sobre la relación del factor de escalamiento y la DOA del sonido directo en SAM. Dado que este escalamiento depende de la DOA del sonido del campo de sonido observado, su valor tiene que modificarse de acuerdo con el remapeo de la DOA resultante de la ubicación de grabación virtual modificada.

La señal de mezcla descendente estéreo modificada que corresponde a la nueva posición virtual de los micrófonos se puede expresar de la siguiente manera

$$W_{1,\text{mod}}(k, n) = G_{11}(k, n, \varphi, d)S(k, n) + G_{12}(k, n, \varphi, d)N_1(k, n) \quad (10)$$

$$W_{2,\text{mod}}(k, n) = G_{21}(k, n, \varphi, d)c_{\text{mod}}S(k, n) + G_{22}(k, n, \varphi, d)N_2(k, n) \quad (11)$$

El cálculo de los filtros de ganancia $G_{ij}(k, n, \varphi, d)$ se realiza de acuerdo con los correspondientes filtros de ganancia $H_i(k, n, \varphi, d)$ como se analizará con respecto al caso de la mezcla descendente mono. Se determina el nuevo factor de escala estéreo c_{mod} como una función de la DOA modificada de tal manera que corresponde a la nueva ubicación de grabación virtual.

Con referencia, una vez más, a las Figuras 2 y 3A, las realizaciones de la presente invención proporcionan un aparato 300 para convertir una primera señal de audio espacial paramétrico 112, 114 que representa una primera posición de escucha o una primera orientación de escucha en una escena de audio espacial, en una segunda señal de audio espacial paramétrico 212, 214 que representa una segunda posición de escucha o una segunda orientación de escucha, siendo la segunda posición de escucha o segunda orientación de escucha diferente de la primera posición de escucha o primera orientación de escucha. El aparato comprende una unidad de modificación de señales de audio espacial 301, 302 adaptada para modificar la primera señal de audio paramétrico parásita 112, 114 dependiendo de un cambio de la primera posición de escucha o la primera orientación de escucha para obtener la segunda señal de audio paramétrico espacial 212, 214, donde la segunda posición de escucha o la segunda orientación de escucha corresponde a la primera posición de escucha o la primera orientación de escucha alterada por el cambio.

Las realizaciones del aparato 300 se pueden adaptar para convertir solo un único parámetro de información secundaria, por ejemplo, la dirección de llegada 114a o el parámetro de difusividad 114b, o solo la señal de mezcla descendente de audio 112 o parte o la totalidad de las señales y parámetros anteriormente citados.

Como se ha descrito anteriormente, en realizaciones que utilizan la codificación de audio direccional (DirAC), las señales analógicas de micrófono se digitalizan y procesan para generar una representación de tiempo/frecuencia con mezcla descendente $W(k, n)$ de las señales de micrófono, que representan, para cada instante o bloque de tiempo k , una representación de frecuencia, donde cada componente de frecuencia de la representación en frecuencia o espectral está indicada por el índice n . Además de la señal de mezcla descendente 112, la unidad de análisis de audio espacial 100 determina, para cada instante de tiempo k y por cada componente de frecuencia n para el correspondiente instante de tiempo k , un vector unitario \mathbf{e}_{DOA} (consúltese la ecuación (4)) que proporciona, por cada componente de frecuencia n y cada instante de tiempo k , el parámetro o información direccional. Además, la unidad de análisis de audio espacial 100 determina para cada instante de tiempo k y cada componente de frecuencia n , un parámetro de difusividad ψ que define una relación entre los componentes de sonido o audio directo y los componentes de sonido o audio difuso, donde los componentes difusos se provocan, por ejemplo, por dos o más fuentes de audio y/o por reflejos de las señales de audio procedentes de las fuentes de audio.

La DirAC es muy eficiente en su procesamiento y con codificación eficiente en cuanto a la memoria, ya que reduce la información de audio espacial que define la escena de audio, por ejemplo, las fuentes de audio, el reflejo, posición y orientación de los micrófonos y el respectivo oyente (para cada instante de tiempo k y cada componente de frecuencia n) a una información direccional, es decir un vector unitario $\mathbf{e}_{\text{DOA}}(k, n)$ y un valor de difusividad $\psi(k, n)$ de entre 0 y 1, asociado a la correspondiente señal de audio de mezcla descendente (mono) única $W(k, n)$ o a varias señales de audio de mezcla descendente (por ejemplo estéreo) $W_1(k, n)$ y $W_2(k, n)$.

Las realizaciones que utilizan la codificación de audio direccional (DirAC) anteriormente mencionada se adaptan, por lo tanto, para modificar, para cada instante k y cada componente de frecuencia n , el correspondiente valor de mezcla descendente $W(k, n)$ por $W_{\text{mod}}(k, n)$, el correspondiente valor de parámetro de dirección de llegada $\mathbf{e}(k, n)$ por $\mathbf{e}_{\text{mod}}(k, n)$ (en las Figuras 1 a 3 representado por φ , respectivamente φ_{mod}) y/o el valor del parámetro de difusividad $\psi(k, n)$ por $\psi_{\text{mod}}(k, n)$.

La unidad de modificación de señales de audio espacial comprende o está formada, por ejemplo, por la unidad de modificación de parámetros 301 y la unidad de modificación de mezcla descendente 302. De acuerdo con una realización preferida, la unidad de modificación de parámetros 301 está adaptada para procesar el parámetro original 114a para determinar el parámetro direccional modificado 214a, para procesar el parámetro de difusividad ψ dependiendo del parámetro direccional original φ , respectivamente 114a, para dividir la señal de mezcla descendente 112 empleando las ecuaciones (2) y (3) usando el respectivo parámetro de difusividad original ψ 114b, y para aplicar el filtrado dependiente de la dirección $H_1(k, n, \varphi, d)$ y $H_2(k, n, \varphi, d)$ dependiendo del parámetro direccional

original φ 114a, respectivamente. Como se ha explicado anteriormente, estas modificaciones se realizan para cada instante de tiempo k y cada componente de frecuencia n para obtener, para cada instante de tiempo k y cada instante de frecuencia n , las respectivas señales y/o parámetros modificados.

5 De acuerdo con una realización, el aparato 300 está adaptado para modificar solo el primer parámetro direccional 114a de la primera señal de audio espacial paramétrico para obtener un segundo parámetro direccional 214a de la segunda señal de audio espacial paramétrico dependiendo de la señal de control 402, por ejemplo, una señal de control de rotación o una señal de control de zoom. En caso de que el cambio de la posición/orientación de escucha solo comprenda una rotación y no una traslación o zoom, es suficiente una correspondiente modificación o
10 desplazamiento del parámetro direccional $\varphi(k,n)$ 114a. Los correspondientes parámetros de difusividad y los componentes de la señal de mezcla descendente pueden quedar sin modificaciones de manera que la señal de mezcla descendente 212 corresponda a la primera señal de mezcla descendente 112 y el segundo parámetro de difusividad 214b corresponda al primer parámetro de difusividad 114b.

15 En el caso de efectuarse un cambio traslacional, por ejemplo un zoom, una modificación del parámetro direccional $\varphi(k,n)$ 114a de acuerdo con una función de remapeo expuesta en la Figura 5A ya mejora la experiencia de sonido y proporciona una mejor sincronización entre la señal de audio y, por ejemplo, una señal de vídeo en comparación con la señal de audio espacial paramétrico original o sin modificaciones (sin modificar el parámetro de difusividad ni la señal de mezcla descendente).

20 Las dos realizaciones anteriores, que solo comprenden la adaptación o remapeo de la dirección de llegada por el filtro f_p ya producen una impresión favorable del efecto de zoom.

25 De acuerdo con otra realización, el aparato 300 está adaptado para aplicar solo el filtro $H_1(k,n,\varphi,d)$. En otras palabras, esta realización no ejecuta el remapeo de la dirección de llegada ni la modificación de la difusividad. Esta realización está adaptada solamente para determinar, por ejemplo, el componente directo 112a a partir de la señal de mezcla descendente 112 y para aplicar la función de filtro H_1 al componente directo para producir una versión ponderada dependiente de la dirección del componente directo. Dichas realizaciones se pueden adaptar además para usar la versión ponderada dependiente de la dirección del componente directo como señal de mezcla
30 descendente modificada W_{mod} 212, o también para determinar el componente difuso 112b de la señal de mezcla descendente original W 112 y para generar la señal de mezcla descendente modificada W_{mod} 212, añadiendo, o en general combinado, la versión ponderada dependiente de la dirección del componente directo y el componente difuso original o sin modificaciones 112b. Se puede obtener una impresión mejorada del zoom acústico, aunque el efecto de zoom es limitado, ya que no se modifica la dirección de llegada.

35 En una realización adicional, se aplican tanto los filtros $H_1(k,n,\varphi,d)$ como los $H_2(k,n,\varphi,d)$; sin embargo, no se realiza el remapeo de la dirección de llegada ni la modificación de la difusividad. La impresión acústica se mejora en comparación con la señal de audio espacial paramétrico inalterada u original 112, 114. La impresión de zoom también es mejor que solo la aplicación de la función de filtro $H_1(k,n,\varphi,d)$ al componente directo cuando el sonido difuso está presente; sin embargo, sigue aún limitado, porque la dirección de llegada φ no se modifica (mejor que la
40 realización anteriormente mencionada que utiliza solo $H_1(k,n,\varphi,d)$).

En otra realización más, solo se aplica el filtro f_d o, en otras palabras, solo se modifica el componente de difusividad ψ . Se mejora el efecto de zoom en comparación con la señal de audio espacial paramétrico original 112, 114 porque
45 se reduce la difusividad de las áreas acercadas (áreas de interés) y se incrementan los valores de difusividad de las áreas fuera de la zona de interés.

Otras realizaciones están adaptadas para realizar el remapeo de la dirección de llegada φ por la función de filtro f_p en combinación con aplicar el filtro $H_1(k,n,\varphi,d)$ en solitario. En otras palabras, tales realizaciones no realizan una
50 modificación de la difusividad de acuerdo con la función de filtro f_d y no aplican la segunda función de filtro $H_2(k,n,\varphi,d)$ a un componente difuso de la señal de mezcla descendente original W 112. Tales realizaciones proporcionan una muy buena impresión de zoom que es mejor que solo la aplicación del remapeo de la dirección de llegada.

55 Las realizaciones que aplican el remapeo de la dirección de llegada de acuerdo con la función f_p en combinación con una modificación de mezcla descendente utilizando ambas funciones de filtro $H_1(k,n,\varphi,d)$ y $H_2(k,n,\varphi,d)$ producen impresiones de zoom aún mejores que solo la aplicación del remapeo de la dirección de llegada en combinación con la aplicación de la primera función de filtro H_1 solamente.

60 Aplicando el remapeo de la dirección de llegada de acuerdo con la función f_p , la modificación de mezcla descendente utilizando los filtros $H_1(k,n,\varphi,d)$ y $H_2(k,n,\varphi,d)$, y la modificación de la difusividad utilizando la función f_d proporciona la mejor implementación del zoom acústico.

Con referencia, una vez más, a la realización de remapeo solamente de la dirección de llegada, la modificación adicional del parámetro de difusividad 114b mejora aún más la experiencia de audio o, en otras palabras, mejora la adaptación de la experiencia de sonido con respecto a la posición modificada dentro de la escena de audio espacial.

5 Por lo tanto, en otras realizaciones, el aparato 300 puede estar adaptado solo para modificar el parámetro direccional $\varphi(k,n)$ y el parámetro de difusividad $\psi(k,n)$, aunque sin modificar la señal de mezcla descendente $W(k,n)$ 100.

10 Las realizaciones preferidas del aparato 300 como se ha mencionado anteriormente también comprenden la modificación de la señal de mezcla descendente $W(k,n)$ para mejorar aún más la experiencia de audio con respecto a la posición modificada en la escena de audio espacial.

15 Por lo tanto, en las realizaciones en las que el primer parámetro direccional $\varphi(k,n)$ 114a es un vector, la unidad de modificación de parámetros 301 está adaptada para desplazar o modificar el primer parámetro direccional en un ángulo definido por una señal de control de rotación en dirección inversa a una dirección definida por la señal de control de rotación para obtener el segundo parámetro direccional $\varphi_{mod}(k,n)$ 214a.

20 En otras realizaciones, la unidad de modificación de parámetros 301 está adaptada para obtener el segundo parámetro direccional 214a utilizando una función de mapeo no lineal (como, por ejemplo, la ilustrada en la Figura 5A) que define el segundo parámetro direccional 214a dependiendo del primer parámetro direccional $\varphi(k,n)$ y un factor de zoom d definido por una señal de control de zoom 402 u otra información de control de traslación definida por la señal de cambio.

25 Como se ha descrito anteriormente, en otras realizaciones, la unidad de modificación de parámetros 301 puede estar adaptada para modificar el primer parámetro de difusividad $\psi(k,n)$ 114b de la primera señal de audio espacial paramétrico para obtener un segundo parámetro de difusividad $\psi_{mod}(k,n)$ 214b dependiendo del primer parámetro direccional $\varphi(k,n)$ 114a. La unidad de modificación de parámetros puede estar adaptada adicionalmente para obtener el segundo parámetro de difusividad $\psi_{mod}(k,n)$ usando una función dependiente de la dirección adaptada para reducir el primer parámetro de difusividad $\psi(k,n)$ para obtener el segundo parámetro de difusividad $\psi_{mod}(k,n)$ en caso de que el primer parámetro direccional $\varphi(k,n)$ esté dentro de un rango central predeterminado, por ejemplo $\gamma = \pm 30^\circ$ de la orientación de referencia original (véase la Figura 5B), y/o para aumentar el primer parámetro de difusividad $\psi(k,n)$ para obtener el segundo parámetro de difusividad $\psi_{mod}(k,n)$ en caso de que el primer parámetro direccional $\varphi(k,n)$ esté fuera del rango central predeterminado, por ejemplo, en un caso bidimensional fuera del rango central definido por $+\gamma = +30^\circ$ y $-\gamma = -30^\circ$ a partir de la orientación de referencia original 0° .

35 En otras palabras, en ciertas realizaciones la unidad de modificación de parámetros 301, 310b está adaptada para obtener el segundo parámetro de difusividad 214b empleando una función dependiente de la dirección adaptada para reducir el primer parámetro de difusividad 114b para obtener el segundo parámetro de difusividad 214b en caso de que el primer parámetro direccional 114a esté dentro de un rango central predeterminado del segundo parámetro direccional con la segunda orientación de escucha o la modificada, que forma el centro del rango central bidimensional o tridimensional predeterminado y/o para aumentar el primer parámetro de difusividad 114b para obtener el segundo parámetro de difusividad en caso de que el primer parámetro direccional 114a esté fuera del rango central predeterminado. La primera orientación de escucha u original define un centro, por ejemplo 0° , del rango central predeterminado del primer parámetro direccional, donde un límite positivo y otro negativo del rango central predeterminado están definidos por un ángulo positivo y otro negativo γ en un plano bidimensional (por ejemplo horizontal) (por ejemplo $\pm 30^\circ$) independientemente de si la segunda orientación de escucha es un vector bidimensional o tridimensional, o por un correspondiente ángulo γ (por ejemplo 30°) que define un cono circular derecho alrededor de la primera orientación de escucha tridimensional. Otras realizaciones pueden comprender diferentes regiones centrales predeterminadas o ventanas, simétricas y asimétricas, dispuestas o centradas alrededor de la primera orientación de escucha o un vector que define la primera orientación de escucha.

40 En otras realizaciones, la función dependiente de la dirección $f_d(k,n,\varphi,d)$ depende de la señal de cambio, por ejemplo, de la señal de control de zoom, donde el rango central predeterminado, los valores respectivos γ que definen el límite negativo y positivo (o, en general, el límite) del rango central es el menor cuanto mayor sea el cambio de traslación o mayor es el factor de zoom definido por la señal de control de zoom.

45 En otras realizaciones, la unidad de modificación de señales de audio espacial comprende una unidad de modificación de mezcla descendente 302 adaptada para modificar la primera señal de audio de mezcla descendente $W(k,n)$ de la primera señal de audio espacial paramétrico para obtener una segunda señal de mezcla descendente $W_{mod}(k,n)$ de la segunda señal de audio espacial paramétrico dependiendo del primer parámetro direccional $\varphi(k,n)$ y del primer parámetro de difusividad $\psi(k,n)$. Las realizaciones de la unidad de modificación de mezcla descendente 302 pueden adaptarse para dividir la primera señal de audio de mezcla descendente W en un componente directo

$S(k,n)$ 112a y un componente difuso $N(k,n)$ 112b dependiente del primer parámetro de difusividad $\psi(k,n)$, por ejemplo, basándose en las ecuaciones (2) y (3).

5 En otras realizaciones, la unidad de modificación de mezcla descendente 302 está adaptada para aplicar una primera función dependiente de la dirección $H_1(k,n,\varphi,d)$ para obtener una versión ponderada dependiente de la dirección del componente directo y/o para aplicar una segunda función dependiente de la dirección $H_2(k,n,\varphi,d)$ al componente difuso para obtener una versión ponderada dependiente de la dirección del componente difuso. La unidad de modificación de mezcla descendente 302 puede estar adaptada para producir la versión ponderada dependiente de la dirección del componente directo 112a aplicando otra función dependiente de la dirección $H_1(k,n,\varphi,d)$ al componente directo, donde la otra función dependiente de la dirección está adaptada para aumentar el componente directo 112a en caso de que el primer parámetro direccional 114a esté dentro del rango central predeterminado adicional de los primeros parámetros direccionales y/o para reducir el componente directo 112a en caso de que el primer parámetro direccional 114a esté fuera del rango predeterminado adicional de los segundos parámetros direccionales. En otras realizaciones la unidad de modificación de mezcla descendente puede estar adaptada para producir la versión ponderada dependiente de la dirección del componente difuso 112b aplicando una función dependiente de la dirección $H_2(k,n,\varphi,d)$ al componente difuso 112b, donde la función dependiente de la dirección está adaptada para reducir el componente difuso en caso de que el primer parámetro direccional 114a esté dentro de un rango central predeterminado del primer parámetro direccional y/o para aumentar el componente de difusividad 112b en caso de que el primer parámetro direccional 114a esté fuera del rango predeterminado del segundo parámetro direccional.

25 En otras realizaciones, la unidad de modificación de mezcla descendente 302 está adaptada para obtener la segunda señal de mezcla descendente 212 basándose en una combinación, por ejemplo una suma, de una versión ponderada dependiente de la dirección del componente directo 112a y una versión ponderada dependiente de la dirección del componente difuso 112b. Sin embargo, otras realizaciones pueden aplicar otros algoritmos además de sumar los dos componentes para obtener la señal de mezcla descendente modificada 212.

30 Como se ha explicado anteriormente, las realizaciones de la unidad de modificación de mezcla descendente 302 pueden estar adaptadas para dividir la señal de mezcla descendente W en una parte o componente difuso 112b y una parte o componente no difuso o directo 112a por medio de dos multiplicadores, es decir $(\psi)^{1/2}$ y $(1 - \psi)^{1/2}$ y para filtrar la parte no difusa 112a mediante la función de filtro H_1 y para filtrar la parte difusa 112b mediante la función de filtro H_2 . La función de filtro H_1 o $H_1(k,n,\varphi,d)$ puede depender de los índices de tiempo/frecuencia k, n , la dirección de llegada original φ y el parámetro de zoom d . La función de filtro H_1 puede depender, además, de la difusividad ψ . La función de filtro H_2 o $H_2(k,n,\varphi,d)$ puede depender de los índices de tiempo/frecuencia k, n , la dirección de llegada original φ y el parámetro de zoom d . La función de filtro H_2 puede depender, además, de la difusividad ψ . Como se describió anteriormente, la función de filtro H_2 puede implementarse en forma de ventana subcardioide como se ilustra en la Figura 7, o como simple factor de atenuación, independiente de la dirección de llegada φ .

40 Haciendo referencia a las explicaciones precedentes, se puede utilizar el parámetro de zoom d para controlar los filtros H_1, H_2 y los modificadores o funciones f_d y f_p (véase la Figura 3A). En el caso de la función de filtro H_1 y f_d el parámetro de zoom d también puede controlar la amplitud de visión o la amplitud angular γ (a la que también se hace referencia como ángulo límite γ) de las ventanas o regiones centrales aplicadas. La amplitud γ se define, por ejemplo como el ángulo en el cual la función de filtro tiene 0 dB (véase, por ejemplo la línea de 0 dB de la Figura 6). Se puede controlar la amplitud angular γ y/o la ganancia por medio del parámetro de zoom d . En la Figura 6 se ilustra un ejemplo de valores diferentes de γ y ganancias máximas y mínimas diferentes.

50 Si bien se han descrito anteriormente realizaciones del aparato, en las cuales las funciones dependientes de la dirección y la ponderación dependen del primer parámetro direccional u original φ (véase la Figura 3A), otras realizaciones pueden estar adaptadas para determinar la segunda difusividad o modificada ψ_{mod} y/o una o ambas funciones de filtro H_1 y H_2 dependientes del segundo parámetro direccional o modificado φ_{mod} . Como se puede determinar a partir de la Figura 4, donde α corresponde al parámetro direccional original φ y P corresponde al parámetro direccional modificado φ_{mod} (en el caso del zoom de acercamiento), cuanto más elevado es el factor de zoom d , más se desplaza el objeto B de una posición central o frontal a una posición lateral, o incluso (en caso de factores de zoom aún más elevados d al expuesto en la Figura 4) a una posición detrás de la posición virtualmente modificada. En otras palabras, cuanto más elevado sea el factor de zoom d , más aumenta la magnitud de un ángulo inicialmente pequeño que representa una posición en un área frontal del oyente, donde los ángulos mayores representan posiciones en un área lateral del oyente. Esta modificación del parámetro direccional se tiene en cuenta aplicando una función ilustrada en la Figura 5A. Además, las ventanas o funciones dependientes de la dirección correspondientes a los demás parámetros y a los componentes directos y difusos también pueden ser diseñadas para tener en cuenta la modificación del parámetro o ángulo direccional original, mediante la reducción de la amplitud angular γ con el aumento del zoom d , por ejemplo de manera no lineal, correspondiente al parámetro de dirección de llegada o mapeo direccional expuesto en la Figura 5A. Por lo tanto, estas ventanas o funciones

dependientes de la dirección pueden estar adaptadas de tal manera que se pueda utilizar directamente el parámetro direccional original (por ejemplo sin la modificación previa ejecutada por la función f_p) o, como alternativa, se realiza primero el mapeo del parámetro direccional f_p y posteriormente se realiza de manera similar la ponderación dependiente de la dirección f_d , H_1 y/o H_2 del parámetro direccional modificado. Con referencia, una vez más, a la Figura 4, ambas cosas son posibles, las funciones dependientes de la dirección f_d , H_1 y H_2 en referencia directa a α , que representa el parámetro direccional original (para el zoom de acercamiento), o funciones dependientes de la dirección f_d , H_1 y H_2 con referencia a β que representa el parámetro direccional modificado.

Las realizaciones que hacen uso del parámetro direccional modificado pueden emplear, similar al caso de las realizaciones que utilizan el parámetro direccional original, diferentes ventanas con diferentes amplitudes angulares y/o diferentes ganancias para diferentes factores de zoom, o bien ventanas iguales con la misma amplitud angular (puesto que el parámetro direccional ya ha sido mapeado para reflejar los diferentes factores de zoom) y la misma ganancia, o ventanas con las mismas amplitudes angulares aunque con diferentes ganancias, donde un factor de zoom más elevado da como resultado una ganancia mayor (análogo a las ventanas de la Figura 6).

La Figura 3B ilustra otra de las realizaciones del aparato. La unidad de modificación de señales de audio espacial de la Figura 3B comprende o está formada, por ejemplo, por la unidad de modificación de parámetros 301 y la unidad de modificación de mezcla descendente 302. De acuerdo con una realización alternativa, la unidad de modificación de parámetros 301 está adaptada para procesar en primer lugar el parámetro original 114a para determinar el parámetro direccional modificado 214a, para luego procesar el parámetro de difusividad ψ dependiendo del parámetro direccional modificado φ_{mod} , que corresponde a 214a, para dividir la señal de mezcla descendente 112 usando las ecuaciones (2) y (3) y el parámetro de difusividad original ψ respectivo, 114b como se describe con respecto a la Figura 3A, y para aplicar el filtrado dependiente de la dirección H_1 y H_2 dependiendo del parámetro direccional modificado φ_{mod} , que corresponde a 214a. Como se explicara anteriormente, estas modificaciones se ejecutan para cada instante de tiempo k y cada componente de frecuencia n para obtener, para cada instante de tiempo k y cada instante de frecuencia n , las respectivas señales y/o parámetros modificados.

De acuerdo con otra realización alternativa del aparato 300 de acuerdo con la Figura 3B, la unidad de modificación de parámetros 301 está adaptada para procesar el parámetro original 114a para determinar el parámetro direccional modificado 214a, para procesar a continuación el parámetro de difusividad ψ dependiendo del parámetro direccional original φ_{mod} , respectivamente 214a, para dividir la señal de mezcla descendente 112 utilizando las ecuaciones (2) y (3) y el parámetro de difusividad original ψ , respectivamente 114b como se ha descrito basándose en la Figura 3A y para aplicar el filtrado dependiente de la dirección H_1 y H_2 dependiendo del parámetro direccional modificado φ_{mod} , respectivamente 214a. como se ha explicado anteriormente, estas modificaciones se realizan para cada instante de tiempo k y cada componente de frecuencia n para obtener, para cada instante de tiempo k y cada instante de frecuencia n , las respectivas señales y/o parámetros modificados.

De acuerdo con otra realización alternativa del aparato 300 de acuerdo con la Figura 3B, la unidad de modificación de parámetros 301 está adaptada para procesar el parámetro original 114a para determinar el parámetro direccional modificado 214a, para procesar el parámetro de difusividad ψ dependiendo del parámetro direccional original φ o 114a, para determinar el parámetro de difusividad modificado ψ_{mod} o 214b, para dividir la señal de mezcla descendente 112 usando las ecuaciones (2) y (3) y el parámetro de difusividad original ψ o 114b como se ha descrito basándose en la Figura 3A, y para aplicar el filtrado dependiente de la dirección H_1 y H_2 dependiente del parámetro direccional modificado φ_{mod} o 214a.

De acuerdo con una realización, el aparato 300 de acuerdo con la Figura 3B está adaptado solo para modificar el primer parámetro direccional 114a de la primera señal de audio espacial paramétrico para obtener un segundo parámetro direccional 214a de la segunda señal de audio espacial paramétrico dependiendo de la señal de control 402, por ejemplo, una señal de control de rotación o una señal de control de zoom. En el caso en que el cambio de posición/orientación de escucha solo comprenda una rotación y no una traslación o zoom, es suficiente una correspondiente modificación o desplazamiento del parámetro direccional $\varphi(k,n)$ 114a. Los correspondientes parámetros de difusividad y los componentes de la señal de mezcla descendente pueden quedar sin modificaciones de manera que la segunda señal de mezcla descendente 212 corresponda a la primera señal de mezcla descendente 112 y el segundo parámetro de difusividad 214b corresponda al primer parámetro de difusividad 114b.

En el caso de realizarse un cambio por traslación, por ejemplo un zoom, una modificación del parámetro direccional $\varphi(k,n)$ 114a de acuerdo con una función de remapeo ilustrada en la Figura 5A ya mejora la experiencia de sonido y proporciona una mejor sincronización entre la señal de audio y, por ejemplo, una señal de vídeo en comparación con la señal de audio paramétrico original o sin modificar (sin modificar el parámetro de difusividad ni la señal de mezcla descendente).

Modificar el parámetro de difusividad 114b mejora aún más la experiencia de audio o, en otras palabras, mejora la

adaptación de la experiencia de sonido con respecto a la posición cambiada dentro de la escena de audio espacial. Por lo tanto, en otras realizaciones, el aparato 300 puede estar adaptado para modificar solamente el parámetro direccional $\varphi(k,n)$ y el parámetro de difusividad $\psi(k,n)$, este último dependiendo del parámetro direccional modificado $\varphi_{mod}(k,n)$, pero no para modificar la señal de mezcla descendente $W(k,n)$ 100.

5 Las realizaciones preferidas del aparato 300 de acuerdo con la Figura 3B también comprenden la modificación de la señal de mezcla descendente $W(k,n)$ dependiente de la difusividad original $\psi(k,n)$ y del parámetro direccional modificado $\varphi_{mod}(k,n)$ para mejorar aún más la experiencia de audio con respecto a la posición cambiada en la escena de audio espacial.

10 Por lo tanto, en realizaciones en las cuales el primer parámetro direccional $\varphi(k,n)$ 114a es un vector, la unidad de modificación de parámetros 301 está adaptada para desplazar o modificar el primer parámetro direccional en un ángulo definido por una señal de control de rotación en dirección inversa a una dirección definida por la señal de control de rotación para obtener el segundo parámetro direccional $\varphi_{mod}(k,n)$ 214a.

15 En otras realizaciones, la unidad de modificación de parámetros 301 está adaptada para obtener el segundo parámetro direccional 214a usando una función de mapeo no lineal (como, por ejemplo, la expuesta en la Figura 5A) que define el segundo parámetro direccional 214a dependiendo del primer parámetro direccional $\varphi(k,n)$ y de un factor de zoom d definido por una señal de control de zoom 402 u otra información de control de traslación definida por la señal de cambio.

20 Como se ha descrito anteriormente, en otras realizaciones, la unidad de modificación de parámetros 301 puede estar adaptada para modificar el primer parámetro de difusividad $\psi(k,n)$ 114b de la primera señal de audio espacial paramétrico para obtener un segundo parámetro de difusividad $\psi_{mod}(k,n)$ 214b dependiendo del segundo parámetro direccional $\varphi_{mod}(k,n)$ 214a. La unidad de modificación de parámetros puede estar adaptada además para obtener el segundo parámetro de difusividad $\psi_{mod}(k,n)$ usando una función dependiente de la dirección adaptada para reducir el primer parámetro de difusividad $\psi(k,n)$ para obtener el segundo parámetro de difusividad $\psi_{mod}(k,n)$ en caso de que el segundo parámetro direccional $\varphi_{mod}(k,n)$ esté dentro de un rango central predeterminado, por ejemplo $\pm 30^\circ$ de la orientación de referencia original a la que se hace referencia como orientación original a 0° , y/o para aumentar el primer parámetro de difusividad $\psi(k,n)$ para obtener el segundo parámetro de difusividad $\psi_{mod}(k,n)$ en caso de que el segundo parámetro direccional $\varphi_{mod}(k,n)$ esté fuera del rango central predeterminado, por ejemplo, en un caso bidimensional fuera del rango central definido por $+30^\circ$ y -30° de la orientación de referencia original a 0° .

25 En otras palabras, en ciertas realizaciones la unidad de modificación de parámetros 301, 310b está adaptada para obtener el segundo parámetro de difusividad 214b usando una función dependiente de la dirección adaptada para reducir el primer parámetro de difusividad 114b para obtener el segundo parámetro de difusividad 214b en caso de que el segundo parámetro direccional 214a esté dentro de un rango central predeterminado del segundo parámetro direccional, formando la primera orientación de escucha, u orientación de escucha original el centro del rango central bidimensional o tridimensional predeterminado y/o para aumentar el primer parámetro de difusividad 114b para obtener el segundo parámetro de difusividad en caso de que el segundo parámetro direccional 214a esté fuera del rango central predeterminado. La primera orientación de escucha define un centro, por ejemplo 0° , del rango central predeterminado del segundo parámetro direccional, donde un límite positivo y otro negativo del rango central predeterminado está definido por un ángulo positivo y otro negativo en un plano bidimensional (por ejemplo horizontal) (por ejemplo $\pm 30^\circ$) independientemente de si la primera orientación de escucha es un vector bidimensional o tridimensional, o en un ángulo correspondiente (por ejemplo 30°) que define un cono circular derecho alrededor de la segunda orientación de escucha tridimensional. Otras realizaciones pueden comprender diferentes regiones centrales predeterminadas, simétricas y asimétricas, dispuestas alrededor de la primera orientación de escucha o vector que define la primera orientación de escucha.

35 En otras realizaciones, la función dependiente de la dirección $f_d(\psi)$ depende de la señal de cambio, por ejemplo, la señal de control de zoom, donde el rango central predeterminado, respectivamente los valores que definen los límites negativo y positivo (o en general, el límite) del rango central es menor cuanto mayor sea el cambio traslacional o mayor es el factor de zoom definido por la señal de control de zoom.

40 En otras realizaciones, la unidad de modificación de señales de audio espacial comprende una unidad de modificación de mezcla descendente 302 adaptada para modificar la primera señal de audio de mezcla descendente $W(k,n)$ de la primera señal de audio espacial paramétrico para obtener una segunda señal de mezcla descendente $W_{mod}(k,n)$ de la segunda señal de audio espacial paramétrico dependiendo del segundo parámetro direccional $\varphi_{mod}(k,n)$ y del primer parámetro de difusividad $\psi(k,n)$. Ciertas realizaciones de la unidad de modificación de mezcla descendente 302 pueden estar adaptadas para dividir la primera señal de audio de mezcla descendente W en un componente directo $S(k,n)$ 112a y un componente difuso $N(k,n)$ 112b dependiendo del primer parámetro de difusividad $\psi(k,n)$, por ejemplo, basándose en las ecuaciones (2) y (3).

En otras realizaciones, la unidad de modificación de mezcla descendente 302 está adaptada para aplicar una primera función dependiente de la dirección H_1 para obtener una versión ponderada dependiente de la dirección del componente directo y/o para aplicar una segunda función dependiente de la dirección H_2 al componente difuso para obtener una versión ponderada dependiente de la dirección del componente difuso. La unidad de modificación de mezcla descendente 302 puede estar adaptada para producir la versión ponderada dependiente de la dirección del componente directo 112a aplicando otra función dependiente de la dirección H_1 al componente directo, donde la otra función dependiente de la dirección está adaptada para incrementar el componente directo 112a en caso de que el segundo parámetro direccional 214a esté dentro de otro rango central predeterminado del segundo parámetro direccional y/o para reducir el componente directo 112a en caso de que el segundo parámetro direccional 214a esté fuera del otro rango predeterminado del segundo parámetro direccional. En otras realizaciones adicionales, la unidad de modificación de mezcla descendente puede estar adaptada para producir la versión ponderada dependiente de la dirección del componente difuso 112b mediante la aplicación de una función dependiente de la dirección H_2 al componente difuso 112b, donde la función dependiente de la dirección está adaptada para reducir el componente difuso en caso de que el segundo parámetro direccional 214a esté dentro de un rango central predeterminado de los segundos parámetros direccionales y/o para aumentar el componente difuso 112b en caso de que el segundo parámetro direccional 214a esté fuera del rango predeterminado de los segundos parámetros direccionales.

En otras realizaciones, la unidad de modificación de mezcla descendente 302 está adaptada para obtener la segunda señal de mezcla descendente 212 basándose en una combinación, por ejemplo una suma, de una versión ponderada dependiente de la dirección del componente directo 112a y una versión ponderada dependiente de la dirección del componente difuso 112b. Sin embargo, otras realizaciones pueden aplicar otros algoritmos aparte de la suma de los dos componentes para obtener la señal de mezcla descendente modificada 212.

Como se ha explicado anteriormente, las realizaciones de la unidad de modificación de mezcla descendente 302 de acuerdo con la Figura 3B pueden estar adaptadas para dividir la señal de mezcla descendente W en una parte o componente difuso 112b y una parte o componente no difuso o directo 112a mediante dos multiplicadores, es decir $(\psi)^{1/2}$ y $(1 - \psi)^{1/2}$ y para filtrar la parte no difusa 112a mediante la función de filtro H_1 y para filtrar la parte difusa 112b mediante la función de filtro H_2 . La función de filtro H_1 o $H_1(\phi, \psi)$ puede depender de los índices de tiempo/frecuencia k, n , la dirección de llegada modificada y el parámetro de zoom d . La función de filtro H_1 puede depender, además, de la difusividad ψ . La función de filtro H_2 o $H_2(\phi, \psi)$ puede depender de los índices de tiempo/frecuencia k, n , la dirección de llegada original ϕ y el parámetro de zoom d . La función de filtro H_2 o $H_2(\phi, \psi)$ puede depender, además, de la difusividad ψ . Como se ha descrito anteriormente, la función de filtro H_2 puede implementarse en forma de ventana subcardioide de acuerdo con lo expuesto en la Figura 7, o en forma de factor de atenuación simple, independiente de la dirección de llegada modificada ϕ_{mod} .

Haciendo referencia a las explicaciones precedentes, también en las realizaciones de acuerdo con la Figura 3B, se pueden utilizar los parámetros de zoom d para controlar los filtros H_1, H_2 y los modificadores o funciones f_d y f_p . En el caso de la función de filtros H_1 y f_d el parámetro de zoom d también puede controlar la amplitud angular γ (a la que también se hace referencia como ángulo límite γ) de las ventanas o regiones centrales aplicadas. La amplitud γ se define, por ejemplo como ángulo en el cual la función de filtro tiene 0 dB (análogo a la línea de 0 dB de la Figura 6). Se puede controlar la amplitud angular γ y/o la ganancia mediante el parámetro de zoom d . Se debe tener en cuenta que, en general, las explicaciones dadas con respecto a las realizaciones de acuerdo con la Figura 3A se aplican de la misma manera o al menos de una manera análoga a las realizaciones de acuerdo con la Figura 3B.

A continuación se describen aplicaciones ilustrativas en que las realizaciones de la invención llevan a una experiencia mejorada de una reproducción conjunta de vídeo/audio mediante el ajuste de la imagen de audio percibida al control de zoom de una cámara de vídeo.

En la realización de teleconferencias, el estado de la técnica consiste en dirigir automáticamente la cámara hacia el locutor activo. Esto está conectado habitualmente con el zoom del hablante. El sonido no coincide habitualmente con la imagen. Las realizaciones de la presente invención ofrecen la posibilidad de acercar también acústicamente al hablante activo. Esto hace que la impresión global sea más realista para los usuarios remotos, ya que no solo la imagen cambia de foco, sino que el sonido coincide con el cambio de atención pretendido. En síntesis, los indicios acústicos corresponden a los indicios visuales.

Las videocámaras actuales, por ejemplo, para uso de entretenimiento doméstico, tienen la capacidad de grabar sonido envolvente y tienen un teleobjetivo óptico potente. Sin embargo, no hay interacción perceptual equivalente entre el zoom óptico y el sonido grabado, ya que el sonido espacial grabado solo depende de la posición real de la cámara y, por consiguiente, de la posición de los micrófonos instalados en la propia cámara. En el caso de una escena filmada en un modo de primer plano, la invención permite ajustar en consecuencia la imagen de audio. Esto da lugar a una experiencia del consumidor más natural y coherente, ya que el sonido se acerca con la imagen.

Se debe mencionar que la invención se puede aplicar también en una fase de post-procesamiento si las señales de micrófono originales se graban sin modificaciones con el vídeo y no se ha realizado más procesamiento. Si bien puede no conocerse la longitud de zoom original, se puede utilizar la invención en equipos creativos de herramientas de post-procesamiento audiovisual. El usuario puede elegir arbitrariamente una longitud de zoom y dirigir el zoom acústico para que coincida con la imagen. Como alternativa, el usuario puede crear sus propios efectos espaciales preferidos. En todo caso, se modificará la posición de grabación del micrófono original por una posición de grabación virtual definida por el usuario.

Dependiendo de ciertos requisitos de implementación de los métodos de la presente invención, los métodos de la invención se pueden implementar en hardware o en software. La implementación se puede realizar utilizando un medio de almacenamiento digital, en particular un disco, un CD, un DVD o un disco Blu-Ray con una señal de control legible electrónicamente almacenada en el mismo, que coopera con un sistema de computación programable de tal manera que se realice una realización del método de la invención. En términos generales, una realización de la presente invención consiste, por lo tanto, en un programa de computación producido con un código de programa almacenado en un portador legible por una máquina, donde el código de programa es apto para llevar a la práctica el método de la invención al ejecutarse el producto programa de computación en una computadora. En otras palabras, las realizaciones del método de la invención consisten, por lo tanto, en un programa de computación que consta de un código de programa para realizar al menos uno de los métodos de la invención al ejecutarse el programa de computación en una computadora.

Lo anterior se mostró y describió particularmente con referencia a realizaciones particulares de lo mismo. Los expertos en la materia comprenderán que se pueden efectuar otros cambios en forma y detalle sin apartarse del alcance de la misma. Se debe entender, por lo tanto, que se pueden efectuar diversos cambios al adaptar las diferentes realizaciones sin apartarse del concepto más amplio desvelado en el presente documento y cubierto por las siguientes reivindicaciones.

REIVINDICACIONES

1. Aparato (300) para convertir una primera señal de audio espacial paramétrico (112, 114) que representa una primera posición de escucha o una primera orientación de escucha en una escena de audio espacial en una segunda señal de audio espacial paramétrico (212, 214) que representa una segunda posición de escucha o una segunda orientación de escucha, comprendiendo el aparato:
- una unidad de modificación de señales de audio espacial (301, 302) adaptada para modificar al menos una de la señal de mezcla descendente (112) de la primera señal de audio espacial paramétrico (112, 114), el parámetro de dirección de llegada (114a) de la primera señal de audio espacial paramétrico y el parámetro de difusividad (114b) de la primera señal de audio espacial paramétrico dependiendo de un cambio de la primera posición de escucha o la primera orientación de escucha para obtener la segunda señal de audio espacial paramétrico (212, 214), en el que la segunda posición de escucha o la segunda orientación de escucha corresponde a la primera posición de escucha o la primera orientación de escucha modificada por el cambio, en el que la unidad de modificación de señales de audio espacial (301, 302) comprende:
- una unidad de modificación de parámetros (301, 301a) adaptada para modificar un primer parámetro direccional (114a) de la primera señal de audio espacial paramétrico (112, 114) para obtener un segundo parámetro direccional (214a) de la segunda señal de audio espacial paramétrico (212, 214) dependiendo de una señal de control (402) que proporciona información que corresponde al cambio; y en el que la unidad de modificación de parámetros (301, 301b) está adaptada para modificar un primer parámetro de difusividad (114b) de la primera señal de audio espacial paramétrico para obtener un segundo parámetro de difusividad (214b) de la segunda señal de audio espacial paramétrico dependiendo del primer parámetro direccional (114a) o dependiendo del segundo parámetro direccional (214a).
2. Aparato de acuerdo con la reivindicación 1, en el que el primer parámetro direccional (114a) y el segundo parámetro direccional (214a) son vectores bidimensionales o tridimensionales.
3. Aparato de acuerdo con la reivindicación 1 o 2, en el que el primer parámetro direccional (114a) es un vector, en el que la señal de control es una señal de control de rotación que define un ángulo de rotación y una dirección de rotación, y en el que la unidad de modificación de parámetros (301, 301a) está adaptada para rotar el vector en un ángulo de rotación en una dirección inversa a la dirección de rotación para obtener el segundo parámetro direccional (214a).
4. Aparato de acuerdo con una de las reivindicaciones 1 a 3, en el que la señal de control es una señal de control de traslación (402) que define una traslación (d) en la dirección de la primera orientación de escucha, en el que la unidad de modificación de parámetros (301, 301a) está adaptada para obtener el segundo parámetro direccional (214a) usando una función de mapeo no lineal (fp) que define el segundo parámetro direccional dependiendo del primer parámetro direccional (114a) y la traslación (d) definida por la señal de control.
5. Aparato de acuerdo con una de las reivindicaciones 1 a 3, en el que la señal de control es una señal de control de zoom (402) que define un factor de zoom (d) en la dirección de la primera orientación de escucha, donde la unidad de modificación de parámetros (301, 301a) está adaptada para obtener el segundo parámetro direccional (214a) utilizando una función de mapeo no lineal (fp) que define el segundo parámetro direccional dependiendo del primer parámetro direccional (114a) y el factor de zoom (d) definido por la señal de control de zoom.
6. Aparato de acuerdo con la reivindicación 1 a 5, en el que la unidad de modificación de parámetros (301, 301b) está adaptada para obtener el segundo parámetro de difusividad (214b) empleando una función dependiente de la dirección (f_d) adaptada para reducir el primer parámetro de difusividad (114b) para obtener el segundo parámetro de difusividad (214b) en caso de que el primer parámetro direccional (114a) esté dentro de un rango central predeterminado del primer parámetro direccional y/o para incrementar el primer parámetro de difusividad (114b) para obtener el segundo parámetro de difusividad en caso de que el primer parámetro direccional (114a) esté fuera del rango central predeterminado, o en el que la unidad de modificación de parámetros (301, 301b) está adaptada para obtener el segundo parámetro de difusividad (214b) empleando una función dependiente de la dirección (f_d) adaptada para reducir el primer parámetro de difusividad (114b) para obtener el segundo parámetro de difusividad (214b) en caso de que el segundo parámetro direccional (214a) esté dentro de un rango central predeterminado del segundo parámetro direccional y/o para incrementar el primer parámetro de difusividad (114b) para obtener el segundo parámetro de difusividad en caso de que el segundo parámetro direccional (214a) esté fuera del rango central predeterminado.
7. Aparato de acuerdo con la reivindicación 6, en el que la señal de control es una señal de control de traslación (402) que define una traslación (d) en la dirección de la primera orientación de escucha, en el que la función dependiente de la dirección depende de la traslación, y en el que el rango central predeterminado es menor cuanto

mayor sea la traslación definida por la señal de control de traslación; o en el que la señal de control es una señal de control de zoom (402) que define un zoom en la dirección de la primera orientación de escucha, en el que la función dependiente de la dirección depende del zoom y en el que el rango central predeterminado es menor cuanto mayor sea un factor de zoom (d) definido por la señal de control de zoom.

5 8. Aparato de acuerdo con las reivindicaciones 1 a 7, comprendiendo la unidad de modificación de señales de audio espacial (300):

10 una unidad de modificación de mezcla descendente (302) adaptada para modificar una primera señal de audio de mezcla descendente (112) de la primera señal de audio espacial paramétrico para obtener una segunda señal de mezcla descendente (212) de la segunda señal de audio espacial paramétrico dependiendo del primer parámetro direccional (114a) y/o del primer parámetro de difusividad (114b), o
 15 una unidad de modificación de mezcla descendente (302) adaptada para modificar la primera señal de audio de mezcla descendente (112) de la primera señal de audio espacial paramétrico para obtener la segunda señal de mezcla descendente (212) de la segunda señal de audio espacial paramétrico dependiendo del segundo parámetro direccional (214a) y/o del primer parámetro de difusividad (114b).

9. Aparato de acuerdo con la reivindicación 8, en el que la unidad de modificación de mezcla descendente (302) está adaptada para obtener un componente directo (112a) de la primera señal de audio de mezcla descendente (112) y/o un componente difuso (112b) de la primera señal de audio de mezcla descendente (112) dependiendo del primer parámetro de difusividad (114b).

10. Aparato de acuerdo con la reivindicación 9, en el que la unidad de modificación de mezcla descendente (302) está adaptada para determinar el componente directo (112a) según:

25

$$S(k, n) = W(k, n) \sqrt{1 - \Psi}$$

y/o el componente difuso según:

$$N(k, n) = W(k, n) \cdot \sqrt{\Psi}$$

30 en el que k es un índice de tiempo, n es un índice de la componente de frecuencia, $W(k, n)$ se refiere a la primera señal de mezcla descendente, $\Psi(k, n)$ se refiere al primer parámetro de difusividad, $S(k, n)$ se refiere al componente directo y $N(k, n)$ se refiere al componente difuso obtenido de la primera señal de mezcla descendente.

35 11. Aparato de acuerdo con la reivindicación 9 o 10, en el que la unidad de modificación de mezcla descendente (302) está adaptada para obtener la segunda señal de mezcla descendente (212) basándose en una versión ponderada dependiente de la dirección del componente directo (112a), basándose en una versión ponderada dependiente de la dirección del componente difuso (112b) o basándose en una combinación de la versión ponderada dependiente de la dirección del componente directo (112a) y la versión ponderada dependiente de la dirección del componente difuso (112b).

40 12. Aparato de acuerdo con la reivindicación 11, en el que la unidad de modificación de mezcla descendente (302) está adaptada para producir la versión ponderada dependiente de la dirección del componente directo (112a) aplicando una función dependiente de la dirección (H_1) adicional al componente directo, estando adaptada la función dependiente de la dirección adicional para incrementar el componente directo (112a) en caso de que el primer parámetro direccional (114a) esté dentro de un rango central predeterminado adicional de los primeros parámetros direccionales y/o para reducir el componente directo (112a) en caso de que el primer parámetro direccional (114a) esté fuera del rango predeterminado adicional de los primeros parámetros direccionales.

50 13. Aparato de acuerdo con la reivindicación 11 o 12, en el que la unidad de modificación de mezcla descendente está adaptada para producir la versión ponderada dependiente de la dirección del componente difuso (112b) aplicando una función dependiente de la dirección (H_2) al componente difuso (112b), estando adaptada la función dependiente de la dirección para reducir el componente difuso en caso de que el primer parámetro direccional (114a) esté dentro de un rango central predeterminado de los primeros parámetros direccionales y/o para incrementar el componente difuso (112b) en caso de que el primer parámetro direccional (114a) esté fuera del rango predeterminado de los primeros parámetros direccionales, o
 55 estando adaptada la función dependiente de la dirección para reducir el componente difuso en caso de que el segundo parámetro direccional (214a) esté dentro de un rango central predeterminado de los segundos parámetros direccionales y/o para incrementar el componente difuso (112b) en caso de que el segundo parámetro direccional (214a) esté fuera del rango predeterminado de los segundos parámetros direccionales.

60

14. Sistema que comprende:

un aparato de acuerdo con una de las reivindicaciones 1 a 13; y
 una cámara de vídeo, en el que el aparato está acoplado a la cámara de vídeo y está adaptado para recibir una
 5 señal de rotación de vídeo o de zoom de vídeo como una señal de control.

15. Un método para convertir una primera señal de audio espacial paramétrico (112, 114) que representa una
 primera posición de escucha o una primera orientación de escucha en una escena de audio espacial en una
 segunda señal de audio espacial paramétrico (212, 214) que representa una segunda posición de escucha o una
 10 segunda orientación de escucha, comprendiendo el método:

modificar al menos una de la señal de mezcla descendente (112) de la primera señal de audio espacial
 paramétrico (112, 114), el parámetro de dirección de llegada (114a) de la primera señal de audio espacial
 paramétrico y el parámetro de difusividad (114b) de la primera señal de audio espacial paramétrico dependiendo
 15 de un cambio de la primera posición de escucha o la primera orientación de escucha para obtener la segunda
 señal de audio espacial paramétrico (212, 214), en el que la segunda posición de escucha o la segunda
 orientación de escucha corresponde a la primera posición de escucha o la primera orientación de escucha
 modificada por el cambio;
 en el que el método comprende modificar un primer parámetro direccional (114a) de la primera señal de audio
 20 espacial paramétrico (112, 114) para obtener un segundo parámetro direccional (214a) de la segunda señal de
 audio espacial paramétrico (212, 214) dependiendo de una señal de control (402) que proporciona información
 que corresponde al cambio; y
 en el que el método comprende modificar un primer parámetro de difusividad (114b) de la primera señal de audio
 25 espacial paramétrico para obtener un segundo parámetro de difusividad (214b) de la segunda señal de audio
 espacial paramétrico dependiendo del primer parámetro direccional (114a) o dependiendo del segundo
 parámetro direccional (214a).

16. Un método para convertir una primera señal de audio espacial paramétrico (112, 114) que representa una
 primera posición de escucha o una primera orientación de escucha en una escena de audio espacial en una
 30 segunda señal de audio espacial paramétrico (212, 214) que representa una segunda posición de escucha o una
 segunda orientación de escucha; comprendiendo el método:

modificar la primera señal de audio espacial paramétrico dependiendo de un cambio de la primera posición de
 escucha o la primera orientación de escucha para obtener la segunda señal de audio espacial paramétrico, en el
 35 que la segunda posición de escucha o la segunda orientación de escucha corresponde a la primera posición de
 escucha o la primera orientación de escucha modificada por el cambio;
 en el que el método comprende modificar un primer parámetro direccional (114a) de la primera señal de audio
 espacial paramétrico (112, 114) para obtener un segundo parámetro direccional (214a) de la segunda señal de
 audio espacial paramétrico (212, 214) dependiendo de una señal de control (402) que proporciona información
 40 que corresponde al cambio;
 en el que el método comprende modificar una primera señal de audio de mezcla descendente (112) de la primera
 señal de audio espacial paramétrico para obtener una segunda señal de mezcla descendente (212) de la
 segunda señal de audio espacial paramétrico dependiendo del primer parámetro direccional (114a) y/o un primer
 parámetro de difusividad (114b), o
 45 modificar la primera señal de audio de mezcla descendente (112) de la primera señal de audio espacial
 paramétrico para obtener la segunda señal de mezcla descendente (212) de la segunda señal de audio espacial
 paramétrico dependiendo del segundo parámetro direccional (214a) y/o un primer parámetro de difusividad
 (114b);
 en el que un componente directo (112a) se obtiene a partir de la primera señal de audio de mezcla descendente
 50 (112) y un componente difuso (112b) se obtiene a partir de la primera señal de audio de mezcla descendente
 (112) dependiendo del primer parámetro de difusividad (114b);
 en el que la segunda señal de mezcla descendente (212) se obtiene basándose en una combinación de una
 versión ponderada dependiente de la dirección del componente directo (112a) y una versión ponderada
 dependiente de la dirección del componente difuso (112b);
 55 en el que la versión ponderada dependiente de la dirección del componente directo (112a) se produce aplicando
 una primera función dependiente de la dirección (H_1) al componente directo, estando adaptada la primera función
 dependiente de la dirección para incrementar el componente directo (112a) en caso de que el primer parámetro
 direccional (114a) esté dentro de un rango central predeterminado de los primeros parámetros direccionales y/o
 para reducir el componente directo (112a) en caso de que el primer parámetro direccional (114a) esté fuera del
 60 rango predeterminado de los primeros parámetros direccionales; y
 en el que se aplica una segunda función dependiente de la dirección (H_2) al componente difuso para obtener la
 versión ponderada dependiente de la dirección del componente difuso.

17. Un programa de computación que tiene un código de programa para realizar el método de acuerdo con la reivindicación 15 o 16 cuando el programa se ejecuta en una computadora.

18. Un aparato (300) para convertir una primera señal de audio espacial paramétrico (112, 114) que representa una primera posición de escucha o una primera orientación de escucha en una escena de audio espacial en una segunda señal de audio espacial paramétrico (212, 214) que representa una segunda posición de escucha o una segunda orientación de escucha; comprendiendo el aparato:

una unidad de modificación de señal de audio espacial (301, 302) adaptada para modificar la primera señal de audio espacial paramétrica (112, 114) dependiendo de un cambio de la primera posición de escucha o la primera orientación de escucha para obtener la segunda señal de audio espacial paramétrico (212, 214), en el que la segunda posición de escucha o la segunda orientación de escucha corresponde a la primera posición de escucha o la primera orientación de escucha modificada por el cambio;

en el que la unidad de modificación de señal de audio espacial (301, 302) comprende una unidad de modificación de parámetros (301, 301a) adaptada para modificar un primer parámetro direccional (114a) de la primera señal de audio espacial paramétrico (112, 114) para obtener un segundo parámetro direccional (214a) de la segunda señal de audio espacial paramétrico (212, 214) dependiendo de una señal de control (402) que proporciona información correspondiente al cambio;

en el que la unidad de modificación de señal de audio espacial (300) comprende:

una unidad de modificación de mezcla descendente (302) adaptada para modificar una primera señal de audio de mezcla descendente (112) de la primera señal de audio espacial paramétrico para obtener una segunda señal de mezcla descendente (212) de la segunda señal de audio espacial paramétrico dependiendo del primer parámetro direccional (114a) y/o un primer parámetro de difusividad (114b), o

una unidad de modificación de mezcla descendente (302) adaptada para modificar la primera señal de audio de mezcla descendente (112) de la primera señal de audio espacial paramétrico para obtener la segunda señal de mezcla descendente (212) de la segunda señal de audio espacial paramétrico dependiendo del segundo parámetro direccional (214a) y/o un primer parámetro de difusividad (114b);

en el que la unidad de modificación de mezcla descendente (302) está adaptada para obtener un componente directo (112a) de la primera señal de audio de mezcla descendente (112) y un componente difuso (112b) de la primera señal de audio de mezcla descendente (112) dependiendo del primer parámetro de difusividad (114b);

en el que la unidad de modificación de mezcla descendente (302) está adaptada para obtener la segunda señal de mezcla descendente (212) basándose en una combinación de una versión ponderada dependiente de la dirección del componente directo (112a) y una versión ponderada dependiente de la dirección del componente difuso (112b);

en el que la unidad de modificación de mezcla descendente (302) está adaptada para producir la versión ponderada dependiente de la dirección del componente directo (112a) aplicando una primera función dependiente de la dirección (H_1) al componente directo, estando adaptada la primera función dependiente de la dirección para incrementar el componente directo (112a) en caso de que el primer parámetro direccional (114a) esté dentro de un rango central predeterminado de los primeros parámetros direccionales y/o para reducir el componente directo (112a) en caso de que el primer parámetro direccional (114a) esté fuera del rango predeterminado de los primeros parámetros direccionales; y

en el que la unidad de modificación de mezcla descendente (302) está adaptada para aplicar una segunda función dependiente de la dirección (H2) al componente difuso para obtener la versión ponderada dependiente de la dirección del componente difuso.

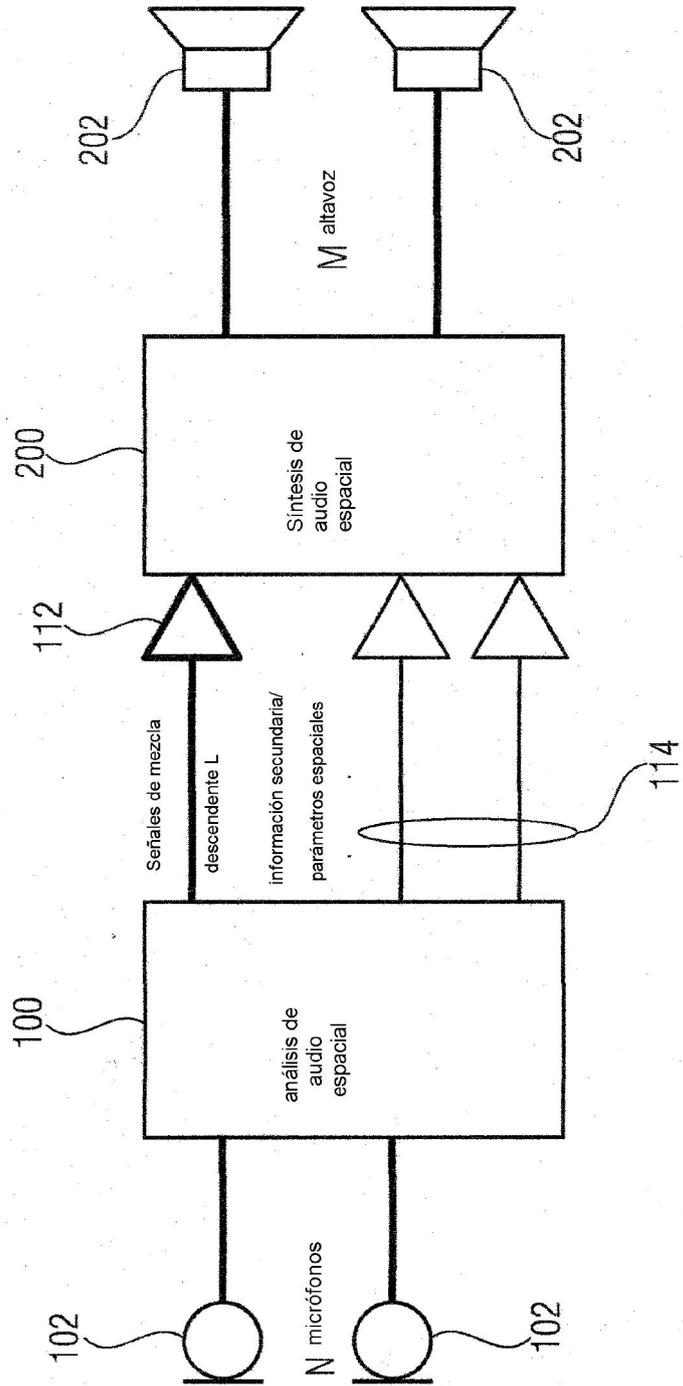


FIG. 1

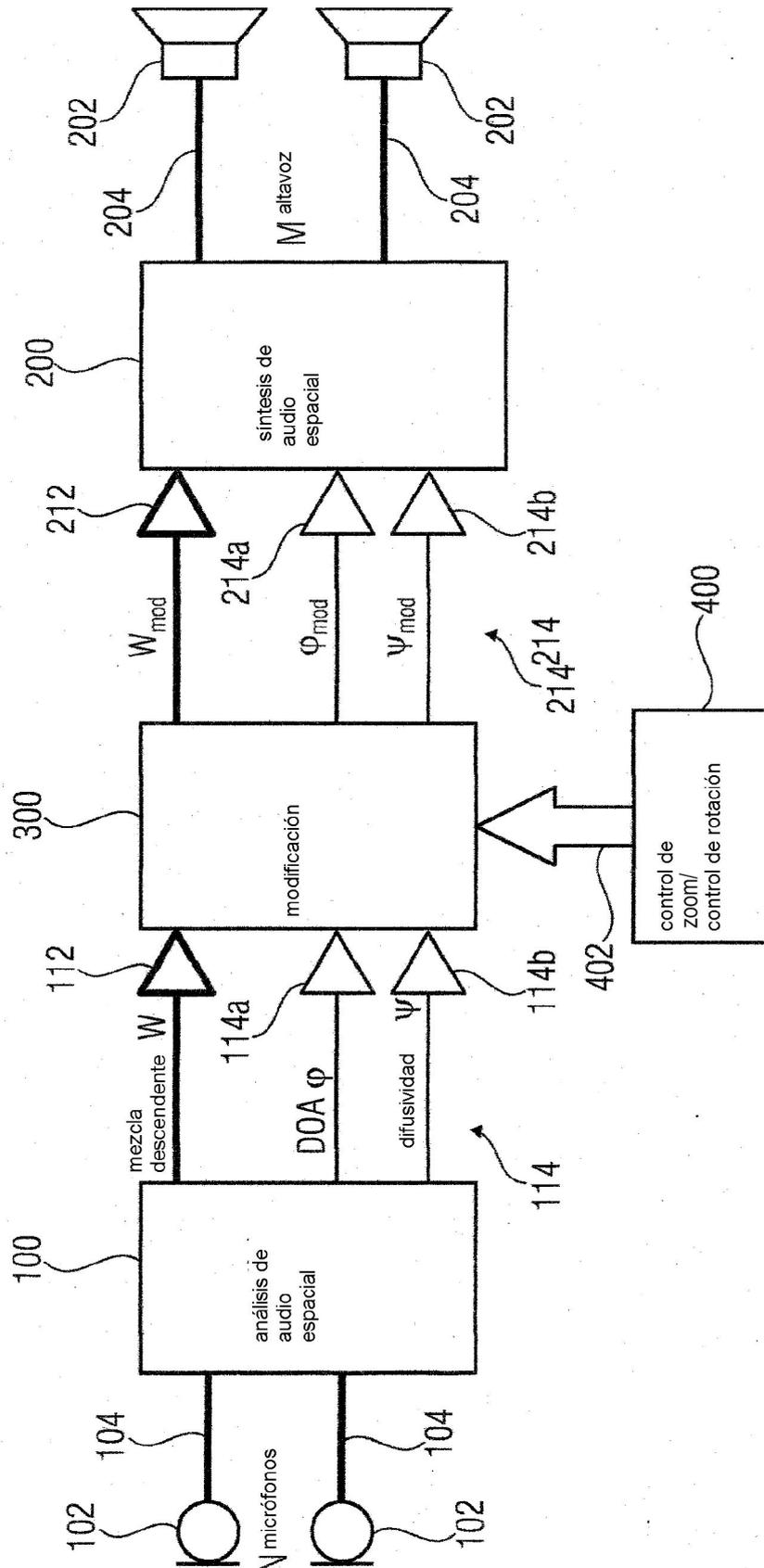


FIG. 2

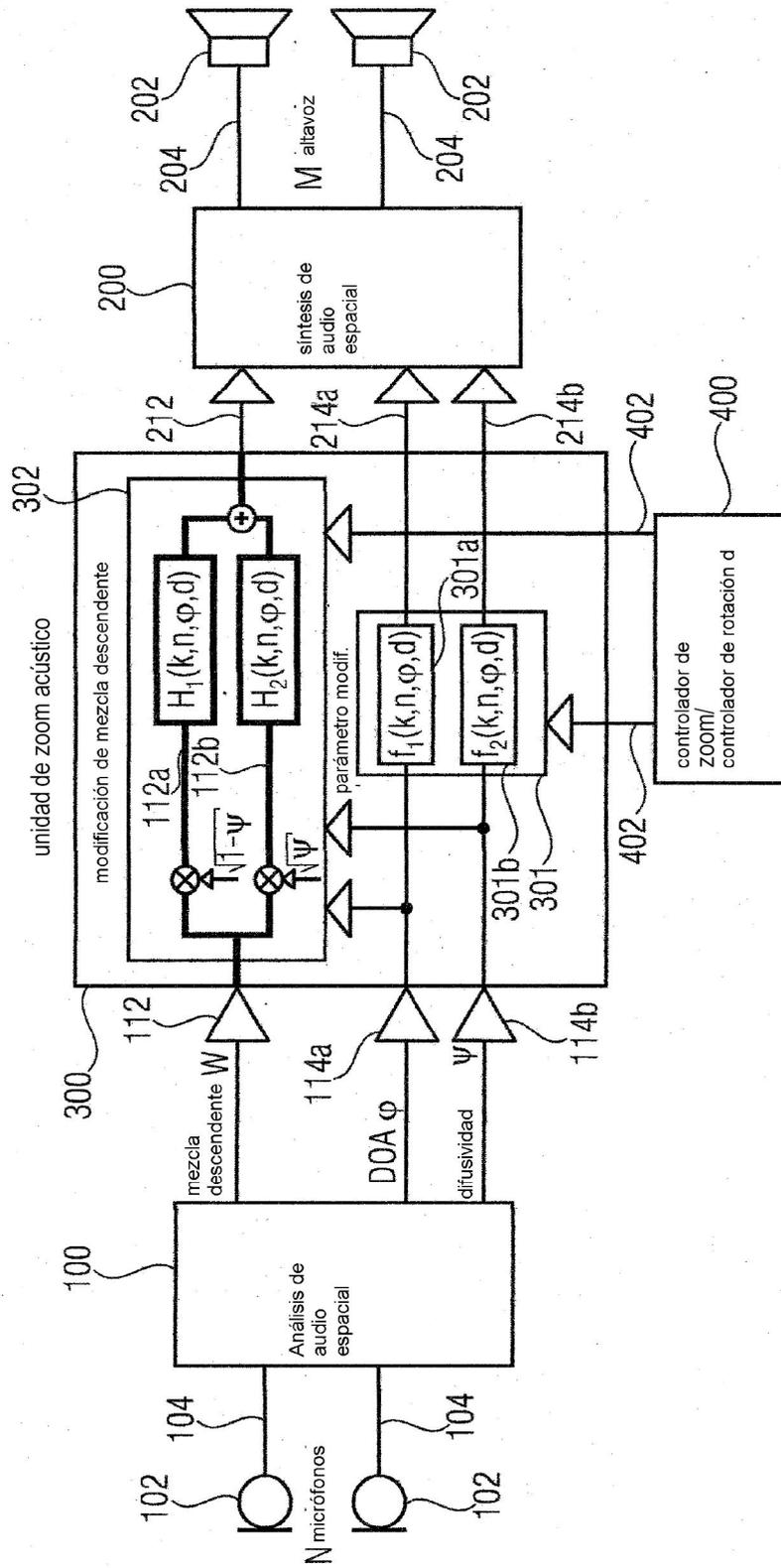


FIG. 3 A

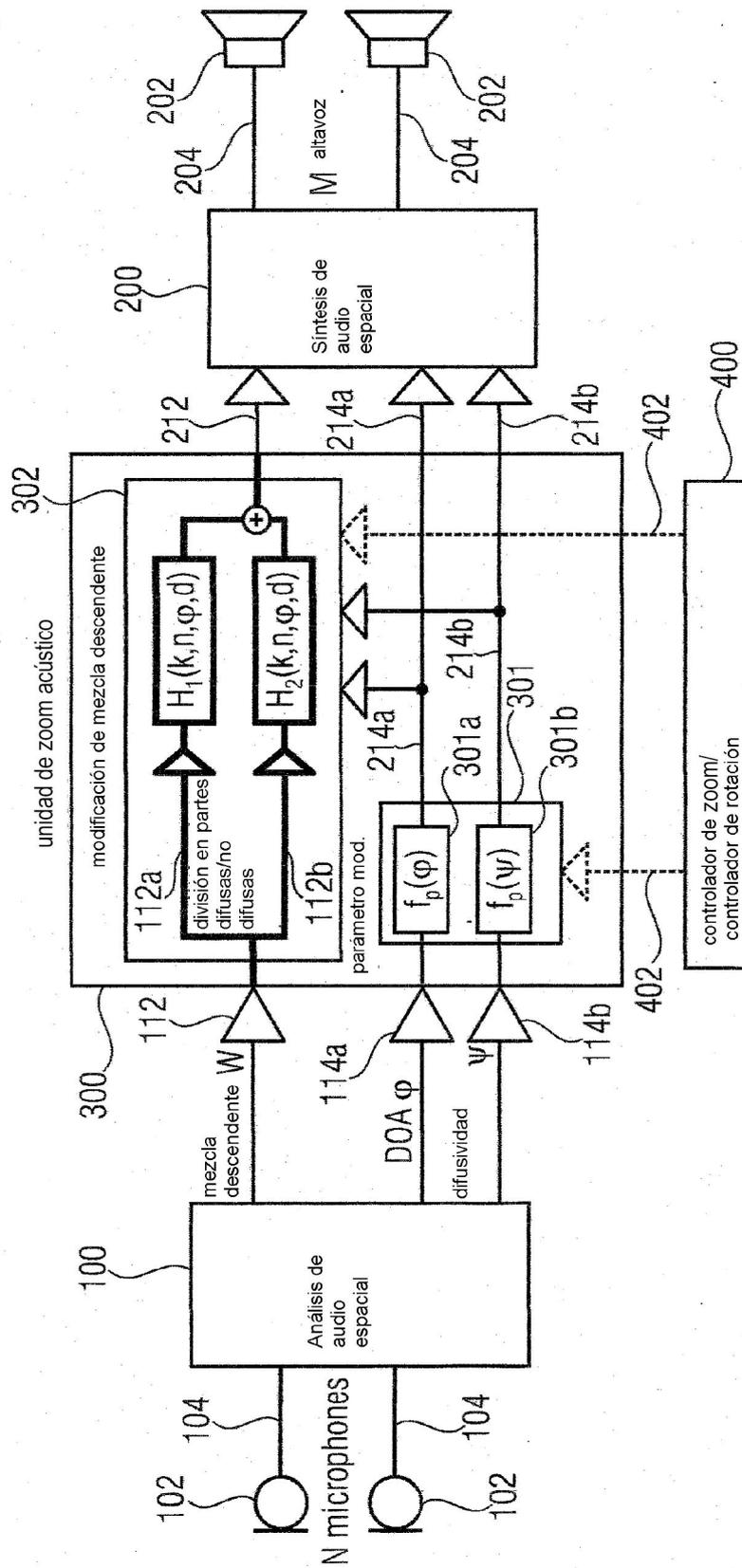


FIG. 3 B

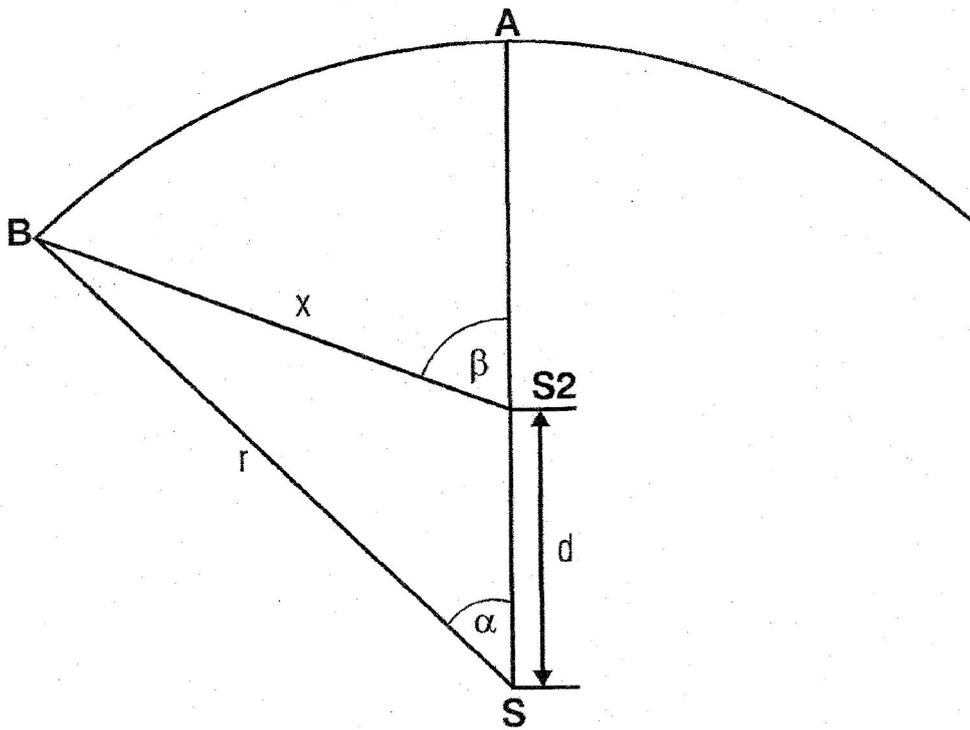


FIG. 4

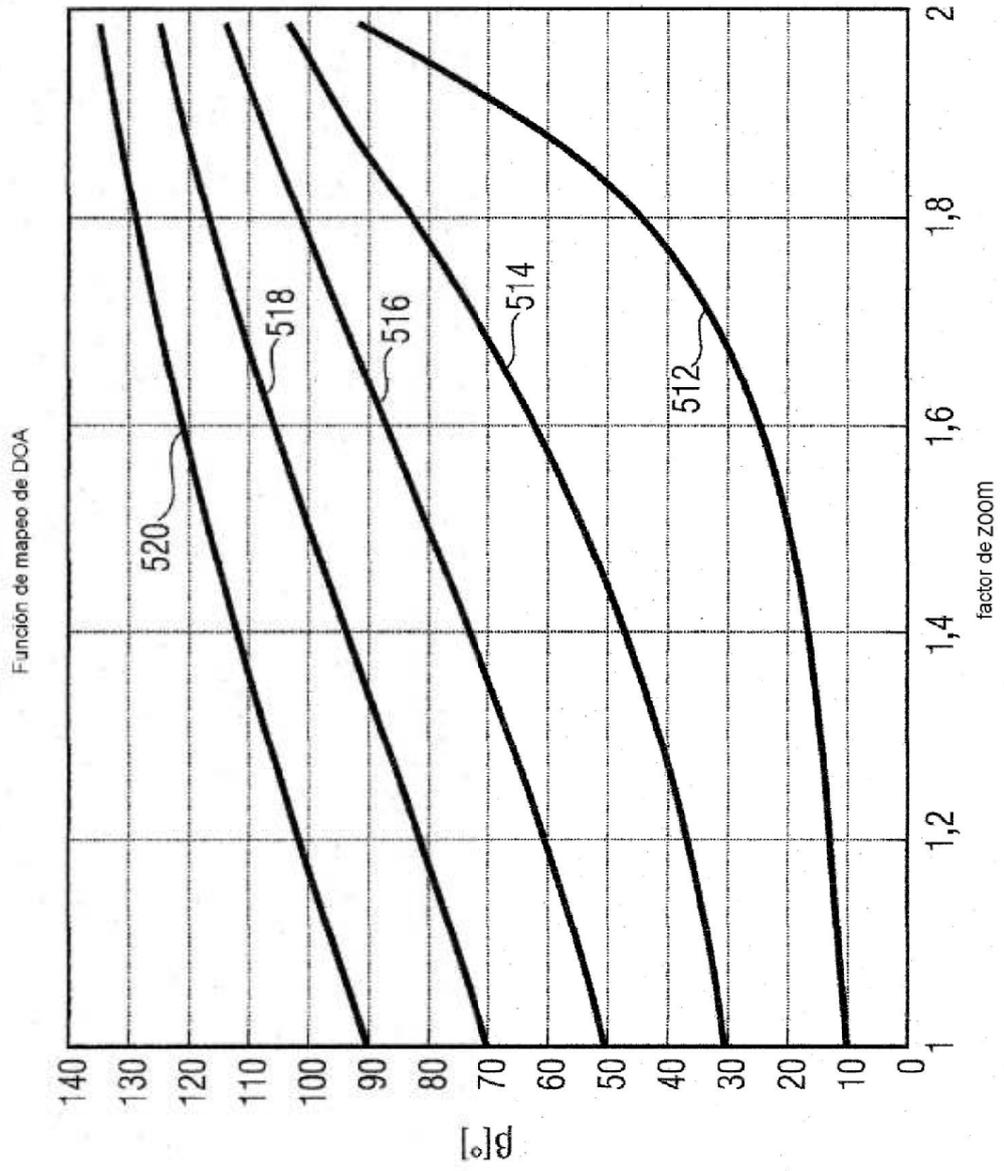


FIG. 5A

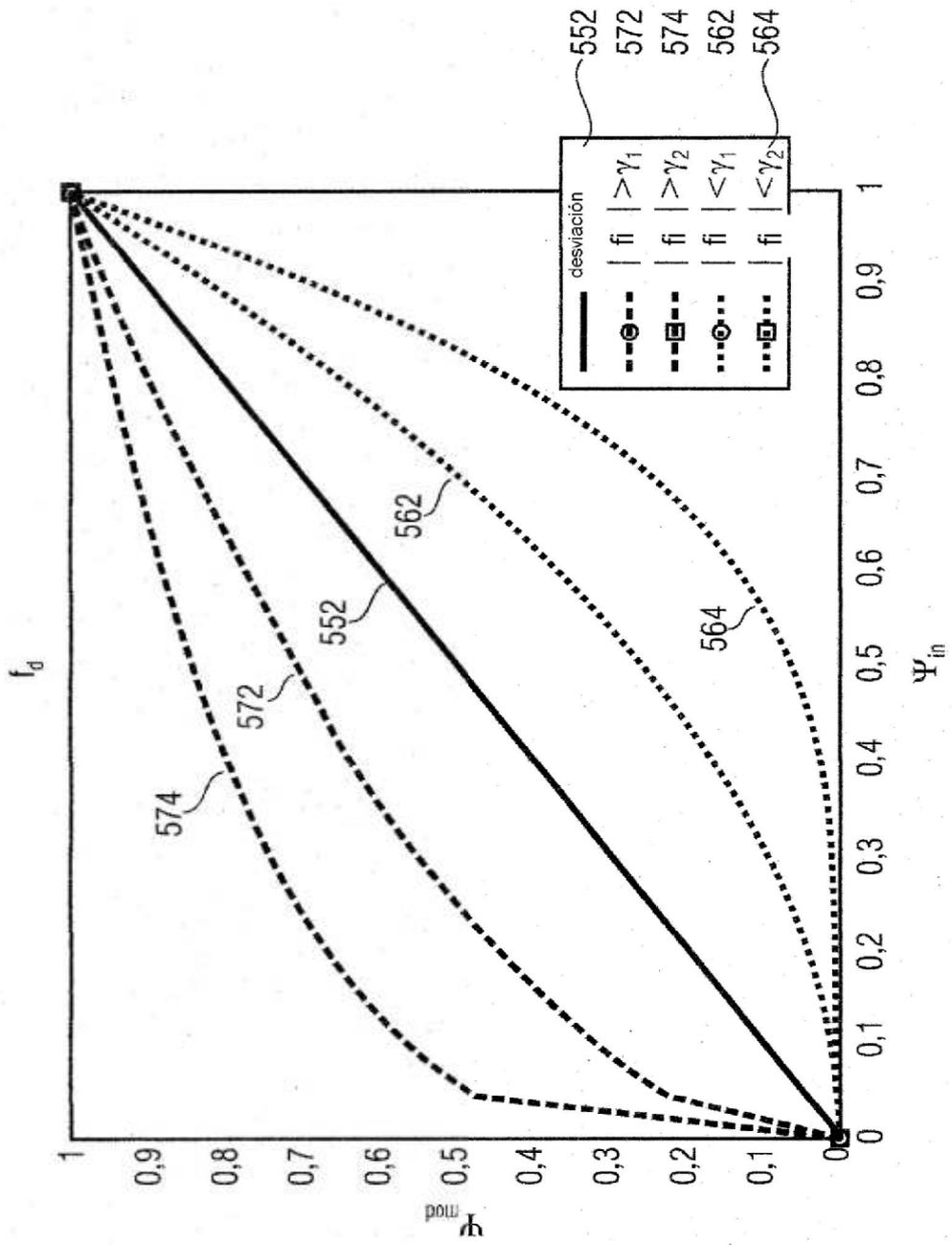


FIG. 5B

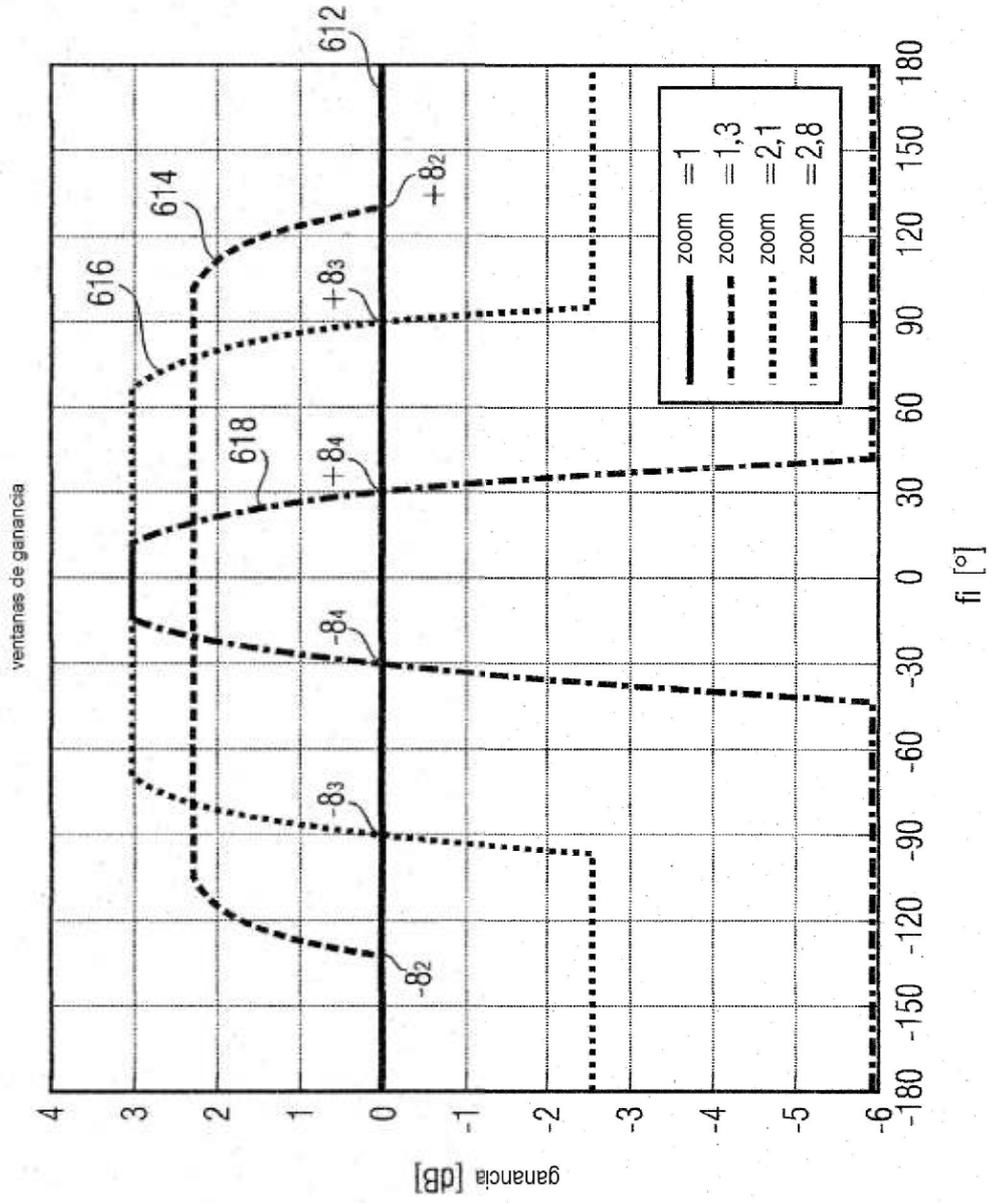


FIG. 6

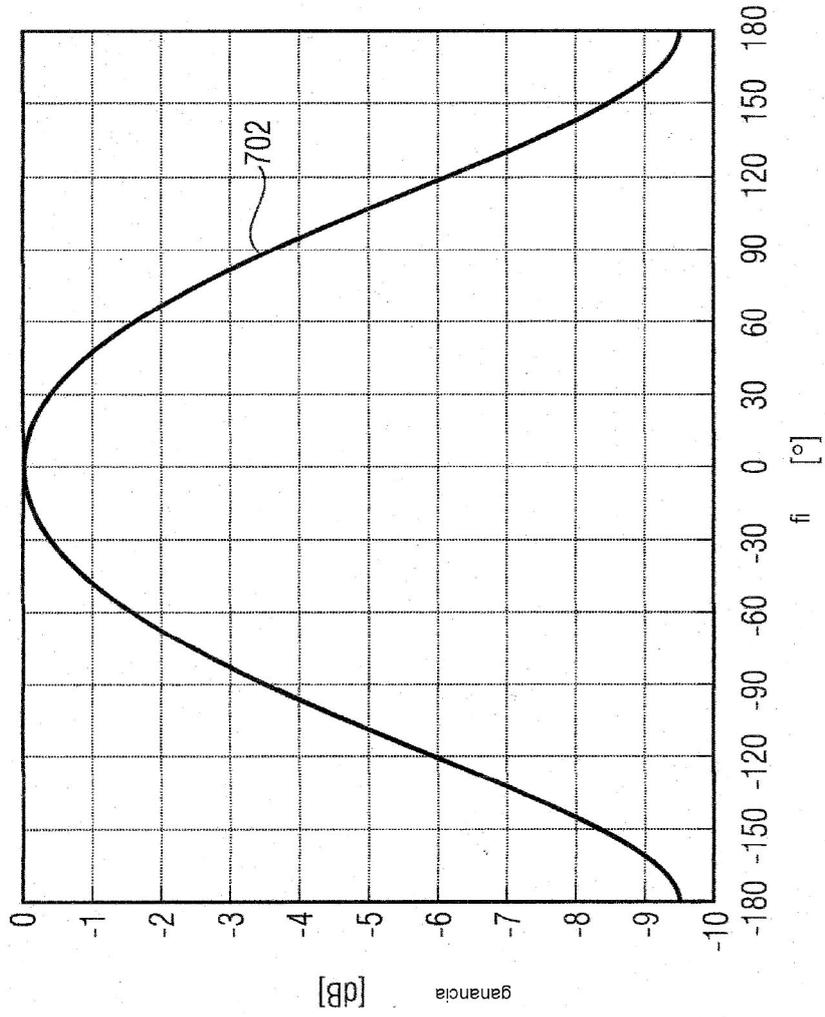


FIG. 7