

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 382 986**

51 Int. Cl.:  
**H04S 3/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **09806392 .8**  
96 Fecha de presentación: **11.08.2009**  
97 Número de publicación de la solicitud: **2324645**  
97 Fecha de publicación de la solicitud: **25.05.2011**

54 Título: **Aparato para fusionar corrientes de audio espacial**

30 Prioridad:  
**13.08.2008 US 88520 P**  
**02.02.2009 EP 09001397**

45 Fecha de publicación de la mención BOPI:  
**15.06.2012**

45 Fecha de la publicación del folleto de la patente:  
**15.06.2012**

73 Titular/es:  
**Fraunhofer-Gesellschaft zur Förderung der  
angewandten Forschung e.V.**  
**Hansastraße 27c**  
**80686 München, DE**

72 Inventor/es:  
**DEL GALDO, Giovanni;**  
**KUECH, Fabian;**  
**KALLINGER, Markus;**  
**PULKKI, Ville;**  
**LAITINEN, Mikko-Ville y**  
**SCHULTZ-AMLING, Richard**

74 Agente/Representante:  
**Arizti Acha, Monica**

ES 2 382 986 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).

## DESCRIPCION

Aparato para fusionar corrientes de audio espacial

- 5 La presente invención se refiere al campo de procesamiento de audio, especialmente procesamiento de audio espacial, y a la fusión de múltiples corrientes de audio espacial.

10 DirAC (DirAC = *Directional Audio Coding*, codificación de audio direccional), cf. V. Pulkki y C. Faller, *Directional audio coding in spatial sound reproduction and stereo upmixing*, en la 28ª Conferencia Internacional de la AES, Pitea, Suecia, junio de 2006 y V. Pulkki, *A method for reproducing natural or modified spatial impression in Multichannel listening*, patente WO 2004/077884 A1, septiembre de 2004, es un enfoque eficaz para el análisis y reproducción de sonido espacial. DirAC utiliza una representación paramétrica de campos de sonido basada en las características que son  
15 relevantes para la percepción de sonido espacial, es decir la dirección de llegada (DOA = *Direction Of Arrival*) y la difusividad del campo de sonido en subbandas de frecuencia. En efecto, DirAC asume que las diferencias de tiempo interaural (ITD = *Interaural Time Differences*) y las diferencias de nivel interaural (ILD = *Interaural Level Differences*) son percibidas correctamente cuando la DOA de un campo de sonido se reproduce correctamente, mientras que la coherencia interaural (IC = *Interaural Coherence*) es percibida correctamente, si la difusividad se reproduce de manera precisa.

20 Estos parámetros, es decir DOA y difusividad, representan información secundaria que acompaña a una señal mono en lo que se denomina como una corriente de DirAC mono. Los parámetros de DirAC se obtienen a partir de una representación de tiempo-frecuencia de las señales de micrófono. Por consiguiente, los parámetros son dependientes del tiempo y de la frecuencia. En el lado de la reproducción, esta información permite una presentación espacial precisa. Para recrear el sonido espacial en una posición de escucha deseada se requiere un montaje de múltiples altavoces. Sin embargo, su geometría es arbitraria. En efecto, las señales para los altavoces se determinan en función de los parámetros de DirAC.

25 Hay diferencias substanciales entre DirAC y codificación de audio multicanal paramétrica tal como MPEG Surround aunque comparten estructuras de procesamiento muy similares, cf. Lars Villemoes, Juergen Herre, Jeroen Breebaart, Gerard Hotho, Sascha Disch, Heiko Purnhagen, y Kristofer Kjrlingm, *MPEG Surround: The forthcoming ISO Standard for spatial audio coding*, en la 28ª Conferencia Internacional de la AES, Pitea, Suecia, junio de 2006. Mientras que MPEG Surround está basado en un análisis de tiempo-frecuencia de los diferentes canales de altavoz, DirAC toma como  
30 entrada los canales de micrófonos coincidentes, que describen eficazmente el campo de sonido en un punto. Así, DirAC también representa una técnica de grabación eficaz para audio espacial.

Otro sistema convencional que trata de audio espacial es SAOC (SAOC = *Spatial Audio Object Coding*, codificación de objeto de audio espacial), cf. Jonas Engdegard, Barbara Resch, Cornelio Falch, Oliver Hellmuth, Johannes Hilpert, Andreas Hoelzer, Leonid Ternetiev, Jeroen Breebaart, Jeroen Koppens, Erik Schuijjer y Werner Oomen, *Spatial audio object coding (SAOC) the upcoming MPEG Standard on parametric object based audio coding*, en la 124ª Convención de la AES, 17-20 de mayo de 2008, Ámsterdam, Países Bajos, 2008, actualmente en proceso de normalización en ISO/MPEG.

40 Se integra en el motor de presentación de MPEG Surround y trata diferentes fuentes de sonido como objetos. Esta codificación de audio ofrece una eficacia muy alta en términos de tasa de transmisión de bits y da una libertad de interacción sin precedentes en el lado de reproducción. Este enfoque promete nuevas características convincentes y funcionalidad en sistemas heredados, así como varias otras nuevas aplicaciones.

Es el objeto de la presente invención proporcionar un concepto aprobado para la fusión de señales de audio espacial.

El objeto se consigue mediante un aparato para fusionar según la reivindicación 1 y un método para fusionar según la reivindicación 13.

45 Obsérvese que la fusión sería trivial en el caso de una corriente de DirAC multicanal, es decir si los 4 canales de audio de formato B estuvieran disponibles. En efecto, las señales de diferentes fuentes pueden sumarse directamente para obtener las señales de formato B de la corriente fusionada. Sin embargo, si estos canales no están disponibles, la fusión directa es problemática.

50 La presente invención está basada en el descubrimiento de que las señales de audio espacial pueden representarse por la suma de una representación de onda, por ejemplo una representación de onda plana, y una representación de campo difuso. A la primera se le puede asignar una dirección. Cuando se fusionan varias corrientes de audio, algunas realizaciones pueden permitir obtener la información secundaria de la corriente fusionada, por ejemplo en términos de una difusividad y una dirección. Algunas realizaciones pueden obtener esta información a partir de las representaciones de onda así como de las corrientes de audio de entrada. Cuando se fusionan varias corrientes de audio, que pueden  
55 todas estar modeladas por una parte o representación de onda y una parte o representación difusa, las partes o componentes de onda y las partes o componentes difusas pueden fusionarse por separado. La fusión de la parte de onda produce una parte de onda fusionada, para la cual se puede obtener una dirección fusionada basándose en las

direcciones de las representaciones de la parte de onda. Además, las partes difusas pueden también fusionarse por separado, a partir de la parte difusa fusionada puede derivarse un parámetro de difusividad global.

5 Algunas realizaciones pueden proporcionar un método para fusionar dos o más señales de audio espacial codificadas como corrientes de DirAC mono. La señal fusionada resultante puede representarse como una corriente de DirAC mono también. En realizaciones, la codificación de DirAC mono puede ser una manera compacta de describir audio espacial, ya que solamente tiene que transmitirse un canal de audio junto con información secundaria.

10 En realizaciones, un escenario posible puede ser una aplicación de teleconferencia con más de dos partes. Por ejemplo, sea el usuario A que se comunica con los usuarios B y C, que generan dos corrientes de DirAC mono separadas. En la ubicación de A, la realización puede permitir que las corrientes del usuario B y C se fusionen en una sola corriente de DirAC mono, que puede reproducirse con la técnica de síntesis de DirAC convencional. En una realización que utiliza una topología de red que observa la presencia de una unidad de control multipunto (MCU = *multipoint control unit*), la operación de fusión se realizaría por la propia MCU, de tal manera que el usuario A recibiría una única corriente de DirAC mono que ya contiene el habla tanto de B como de C. Claramente, las corrientes de DirAC que van a fusionarse pueden también generarse sintéticamente, lo que significa que puede añadirse información secundaria apropiada a una señal de audio mono. En el ejemplo recién mencionado, el usuario A podría recibir dos corrientes de audio de B y C sin ninguna información secundaria. Es posible entonces asignar a cada corriente una determinada dirección y difusividad, añadiendo así la información secundaria necesaria para construir las corrientes de DirAC, que pueden después fusionarse mediante una realización.

20 Otro escenario posible en realizaciones se puede encontrar en juegos en línea de múltiples jugadores y aplicaciones de realidad virtual. En estos casos, se generan varias corrientes ya sea de jugadores u objetos virtuales. Cada corriente está caracterizada por una determinada dirección de llegada en relación con el oyente y puede por consiguiente expresarse por una corriente de DirAC. La realización puede usarse para fusionar las diferentes corrientes en una única corriente de DirAC, que se reproduce luego en la posición del oyente.

Realizaciones de la presente invención se detallarán utilizando las figuras adjuntas, en las que:

25 la figura 1a muestra una realización de un aparato para fusionar;

la figura 1b muestra la presión y componentes de un vector de velocidad de partícula en un plano gaussiano para una onda plana;

la figura 2 muestra una realización de un codificador de DirAC;

la figura 3 ilustra una fusión ideal de corrientes de audio;

30 la figura 4 muestra las entradas y salidas de una realización de un bloque de procesamiento de fusión de DirAC general;

la figura 5 muestra un diagrama de bloques de una realización; y

la figura 6 muestra un diagrama de flujo de una realización de un método para fusionar.

35 La figura 1a ilustra una realización de un aparato 100 para fusionar una primera corriente de audio espacial con una segunda corriente de audio espacial para obtener una corriente de audio fusionada. La realización ilustrada en la figura 1a ilustra la fusión de dos corrientes de audio, sin embargo no debe limitarse a dos corrientes de audio, de manera similar, múltiples corrientes de audio espacial pueden fusionarse. La primera corriente de audio espacial y la segunda corriente de audio espacial pueden por ejemplo corresponder a corrientes de DirAC mono y la corriente de audio fusionada puede también corresponder a una única corriente de audio de DirAC mono. Como se detallará posteriormente, una corriente de DirAC mono puede comprender una señal de presión, por ejemplo capturada por un micrófono omnidireccional, e información secundaria. Esta última puede comprender medidas dependientes del tiempo-frecuencia de difusividad y dirección de llegada del sonido.

45 La figura 1a muestra una realización de un aparato 100 para fusionar una primera corriente de audio espacial con una segunda corriente de audio espacial para obtener una corriente de audio fusionada, que comprende un estimador 120 para estimar una primera representación de onda que comprende una primera medida de dirección de onda y una primera medida de campo de onda para la primera corriente de audio espacial, teniendo la primera corriente de audio espacial una primera representación de audio y una primera dirección de llegada, y para estimar una segunda representación de onda que comprende una segunda medida de dirección de onda y una segunda medida de campo de onda para la segunda corriente de audio espacial, teniendo la segunda corriente de audio espacial una segunda representación de audio y una segunda dirección de llegada. En realizaciones, la primera y/o segunda representación de onda puede corresponder a una representación de onda plana.

50 En la realización mostrada en la figura 1a, el aparato 100 comprende además un procesador 130 para procesar la primera representación de onda y la segunda representación de onda para obtener una representación de onda fusionada que comprende una medida de campo fusionada y una medida de dirección de llegada fusionada y para procesar la primera representación de audio y la segunda representación de audio para obtener una representación de

audio fusionada, el procesador 130 está adaptado además para proporcionar la corriente de audio fusionada que comprende la representación de audio fusionada y la medida de dirección de llegada fusionada.

- 5 El estimador 120 puede estar adaptado para estimar la primera medida de campo de onda en términos de una primera amplitud de campo de onda, para estimar la segunda medida de campo de onda en términos de una segunda amplitud de campo de onda y para estimar una diferencia de fase entre la primera medida de campo de onda y la segunda medida de campo de onda. En realizaciones, el estimador puede estar adaptado para estimar una primera fase de campo de onda y una segunda fase de campo de onda. En realizaciones, el estimador 120 puede estimar solamente un desplazamiento o diferencia de fase entre las representaciones de onda primera y segunda, las medidas de campo de onda primera y segunda, respectivamente. El procesador 130 puede estar adaptado entonces, por consiguiente, para procesar la primera representación de onda y la segunda representación de onda para obtener una representación de onda fusionada que comprende una medida de campo de onda fusionada, que puede comprender una amplitud de campo de onda fusionada, una fase de campo de onda fusionada y una medida de dirección de llegada fusionada, y para procesar la primera representación de audio y la segunda representación de audio para obtener una representación de audio fusionada.
- 10 En realizaciones, el procesador 130 puede estar adaptado adicionalmente para procesar la primera representación de onda y la segunda representación de onda para obtener la representación de onda fusionada que comprende la medida de campo de onda fusionada, la medida de dirección de llegada fusionada y un parámetro de difusividad fusionado, y para proporcionar la corriente de audio fusionada que comprende la representación de audio fusionada, la medida de dirección de llegada fusionada y el parámetro de difusividad fusionado.
- 15 En otras palabras, en realizaciones, un parámetro de difusividad puede determinarse basándose en las representaciones de onda para la corriente de audio fusionada. El parámetro de difusividad puede establecer una medida de una difusividad espacial de una corriente de audio, es decir una medida de una distribución espacial como por ejemplo una distribución angular alrededor de una cierta dirección. En una realización, un escenario posible podría ser la fusión de dos señales sintéticas mono con sólo información direccional.
- 20 El procesador 130 puede estar adaptado para procesar la primera representación de onda y la segunda representación de onda para obtener la representación de onda fusionada, en el que el parámetro de difusividad fusionado está basado en la primera medida de dirección de onda y en la segunda medida de dirección de onda. En realizaciones, las representaciones de onda primera y segunda pueden tener direcciones de llegada diferentes y la dirección de llegada fusionada puede situarse entre las mismas. En esta realización, aunque las corrientes de audio espacial primera y segunda pueden no proporcionar ningún parámetro de difusividad, el parámetro de difusividad fusionado puede determinarse a partir de las representaciones de onda primera y segunda, es decir basándose en la primera medida de dirección de onda y en la segunda medida de dirección de onda. Por ejemplo, si dos ondas planas chocan desde diferentes direcciones, es decir la primera medida de dirección de onda difiere de la segunda medida de dirección de onda, la representación de audio fusionada puede comprender una dirección de llegada fusionada combinada con un parámetro de difusividad fusionado que no se desvanece, con el fin de tener en cuenta la primera medida de dirección de onda y la segunda medida de dirección de onda. En otras palabras, mientras que dos corrientes de audio espacial enfocadas pueden no tener o proporcionar ninguna difusividad, la corriente de audio fusionada puede tener una difusividad que no se desvanece, ya que está basada en la distribución angular establecida por las corrientes de audio primera y segunda.
- 25 Algunas realizaciones pueden estimar un parámetro de difusividad, por ejemplo, para una corriente de DirAC fusionada. En general, algunas realizaciones pueden establecer o asumir entonces los parámetros de difusividad de las corrientes individuales a un valor fijo, por ejemplo 0 ó 0,1, o a un valor variable derivado de un análisis de las representaciones de audio y/o representaciones de dirección.
- 30 En otras realizaciones, el aparato 100 para fusionar la primera corriente de audio espacial con la segunda corriente de audio espacial para obtener una corriente de audio fusionada puede comprender el estimador 120 para estimar la primera representación de onda que comprende una primera medida de dirección de onda y una primera medida de campo de onda para la primera corriente de audio espacial, teniendo la primera corriente de audio espacial la primera representación de audio, la primera dirección de llegada y un primer parámetro de difusividad. En otras palabras, la primera representación de audio puede corresponder a una señal de audio con un determinado ancho espacial o ser difusa en cierta medida. En una realización, esto puede corresponder a un escenario en un juego de ordenador. Un primer jugador puede estar en un escenario, en el que la primera representación de audio representa una fuente de audio como por ejemplo un tren que pasa, creando un campo de sonido difuso en cierta medida. En tal realización, los sonidos evocados por el propio tren pueden ser difusos, un sonido producido por la bocina del tren, es decir las componentes de frecuencia correspondientes, pueden no ser difusos.
- 35 El estimador 120 puede además estar adaptado para estimar la segunda representación de onda que comprende la segunda medida de dirección de onda y la segunda medida de campo de onda para la segunda corriente de audio espacial, teniendo la segunda corriente de audio espacial la segunda representación de audio, la segunda dirección de llegada y un segundo parámetro de difusividad. En otras palabras, la segunda representación de audio puede corresponder a una señal de audio con un determinado ancho espacial o ser difusa en cierta medida. Una vez más, esto

puede corresponder al escenario en el juego de ordenador, en el que una segunda fuente de sonido puede representarse por la segunda corriente de audio, por ejemplo ruido de fondo de otro tren que pasa por otra vía. Para el primer jugador en el juego de ordenador, ambas fuentes de sonido pueden ser difusas ya que se encuentra en la estación de tren.

- 5 En realizaciones, el procesador 130 puede estar adaptado para procesar la primera representación de onda y la segunda representación de onda para obtener la representación de onda fusionada que comprende la medida de campo de onda fusionada y la medida de dirección de llegada fusionada, y para procesar la primera representación de audio y la segunda representación de audio para obtener la representación de audio fusionada, y para proporcionar la corriente de audio fusionada que comprende la representación de audio fusionada y la medida de dirección de llegada fusionada.
- 10 En otras palabras, el procesador 130 puede no determinar un parámetro de difusividad fusionado. Esto puede corresponder al campo de sonido experimentado por un segundo jugador en el juego de ordenador descrito anteriormente. El segundo jugador puede encontrarse más alejado de la estación de tren, de tal manera que las dos fuentes de sonido pueden no experimentarse como difusas por el segundo jugador, sino más bien representar fuentes de sonido enfocadas, debido a la mayor distancia.
- 15 En realizaciones, el aparato 100 puede comprender además medios 110 para determinar, para la primera corriente de audio espacial, la primera representación de audio y la primera dirección de llegada y para determinar, para la segunda corriente de audio espacial, la segunda representación de audio y la segunda dirección de llegada. En realizaciones, puede proporcionarse a los medios 110 para determinar una corriente de audio directa, es decir la determinación puede solamente referirse a la lectura de la representación de audio en términos de por ejemplo una señal de presión y una DOA y opcionalmente también parámetros de difusividad en términos de la información secundaria.

20 El estimador 120 puede estar adaptado para estimar la primera representación de onda a partir de la primera corriente de audio espacial que tiene además un primer parámetro de difusividad y/o para estimar la segunda representación de onda a partir de la segunda corriente de audio espacial que tiene además un segundo parámetro de difusividad, el procesador 130 puede estar adaptado para procesar la medida de campo de onda fusionada, las representaciones de audio primera y segunda y los parámetros de difusividad primero y segundo para obtener el parámetro de difusividad fusionado para la corriente de audio fusionada, y el procesador 130 puede estar adaptado adicionalmente para proporcionar la corriente de audio que comprende el parámetro de difusividad fusionado. Los medios 110 para determinar pueden estar adaptados para determinar el primer parámetro de difusividad para la primera corriente de audio espacial y el segundo parámetro de difusividad para la segunda corriente de audio espacial.

25 El procesador 130 puede estar adaptado para procesar las corrientes de audio espacial, las representaciones de audio, la DOA y/o los parámetros de difusividad bloque por bloque, es decir en términos de segmentos de muestras o valores. En algunas realizaciones, un segmento puede comprender un número predeterminado de muestras correspondientes a una representación de frecuencia de una determinada banda de frecuencia en un determinado momento de una corriente de audio espacial. Tal segmento puede corresponder a una representación mono y tener asociada una DOA y un parámetro de difusividad.

30 En realizaciones, los medios 110 para determinar pueden estar adaptados para determinar la primera y segunda representación de audio, la primera y segunda dirección de llegada y los parámetros de difusividad primero y segundo de una manera dependiente del tiempo-frecuencia y/o el procesador 130 puede estar adaptado para procesar las representaciones de onda, parámetros de difusividad y/o medidas de DOA primera y segunda y/o para determinar la representación de audio fusionada, la medida de dirección de llegada fusionada y/o el parámetro de difusividad fusionado de una manera dependiente del tiempo-frecuencia.

35 En realizaciones, la primera representación de audio puede corresponder a una primera representación mono y la segunda representación de audio puede corresponder a una segunda representación mono y la representación de audio fusionada puede corresponder a una representación mono fusionada. En otras palabras, las representaciones de audio pueden corresponder a un único canal de audio.

40 En realizaciones, los medios 110 para determinar pueden estar adaptados para determinar y/o el procesador puede estar adaptado para procesar la primera y segunda representación mono, la primera y la segunda DOA y un primer y un segundo parámetro de difusividad y el procesador 130 puede proporcionar la representación mono fusionada, la medida de DOA fusionada y/o el parámetro de difusividad fusionado de una manera dependiente del tiempo-frecuencia. En realizaciones, la primera corriente de audio espacial puede estar ya proporcionada en términos de, por ejemplo, una representación de DirAC, los medios 110 para determinar pueden estar adaptados para determinar la primera y segunda representación mono, la primera y segunda DOA y los parámetros de difusividad primero y segundo simplemente mediante extracción a partir de las corrientes de audio primera y segunda, por ejemplo a partir de la información secundaria de DirAC.

45 A continuación se expondrá una realización en detalle, debiendo introducirse en primer lugar la notación y el modelo de datos. En realizaciones, los medios 110 para determinar pueden estar adaptados para determinar las representaciones de audio primera y segunda y/o el procesador 130 puede estar adaptado para proporcionar una representación mono fusionada en términos de una señal de presión  $p(t)$  o una señal de presión transformada en tiempo-frecuencia  $P(k,n)$ , donde  $k$  denota un índice de frecuencia y  $n$  denota un índice de tiempo.

5 En realizaciones, las medidas de dirección de onda primera y segunda, así como la medida de dirección de llegada fusionada pueden corresponder a cualquier cantidad direccional, tal como por ejemplo un vector, un ángulo, una dirección, etc. y pueden derivarse de cualquier medida direccional que representa una componente de audio, tal como por ejemplo un vector de intensidad, un vector de velocidad de partícula, etc. Las medidas de campo de onda primera y segunda, así como la medida de campo de onda fusionada pueden corresponder a cualquier cantidad física que describe una componente de audio, que puede ser de valor real o complejo, corresponder a una señal de presión, una amplitud o magnitud de velocidad de partícula, intensidad de sonido, etc. Además, las medidas pueden considerarse en el dominio de tiempo y/o de frecuencia.

10 Algunas realizaciones pueden estar basadas en la estimación de una representación de onda plana para las medidas de campo de onda de las representaciones de onda de las corrientes de entrada, lo que puede llevarse a cabo por el estimador 120 en la figura 1a. En otras palabras, la medida de campo de onda puede modelarse utilizando una representación de onda plana. En general, existen varias descripciones exhaustivas (es decir, completas) equivalentes de una onda u ondas planas en general. A continuación se introducirá una descripción matemática para calcular parámetros de difusividad y direcciones de llegadas o medidas de dirección para diferentes componentes. Aunque solamente algunas descripciones se refieren directamente a cantidades físicas, como por ejemplo presión, velocidad de partícula, etc., existe potencialmente un número infinito de diferentes maneras para describir representaciones de onda, de las cuales se presentará una como un ejemplo posteriormente, sin que se pretenda sin embargo limitarse de ninguna manera a las realizaciones de la presente invención.

20 Con el fin de detallar adicionalmente diferentes descripciones potenciales se consideran dos números reales a y b. La información contenida en a y b puede transferirse al enviar c y d, cuando

$$\begin{bmatrix} c \\ d \end{bmatrix} = \Omega \begin{bmatrix} a \\ b \end{bmatrix},$$

25 donde  $\Omega$  es una matriz de 2x2 conocida. El ejemplo considera solamente combinaciones lineales, en general cualquier combinación, es decir también una combinación no lineal es concebible.

A continuación, los escalares se representan por letras minúsculas, a, b, c, mientras que los vectores de columna se representan por letras minúsculas en negrita **a**, **b**, **c**. El superíndice  $(\cdot)^\dagger$  denota la transpuesta, respectivamente, mientras que  $(\cdot)^{\bar{}}$  y  $(\cdot)^*$  denotan conjugación compleja. La notación de fasor complejo se distingue del temporal. Por ejemplo, la presión  $p(t)$  que es un número real y a partir del cual puede derivarse una medida de campo de onda posible, puede expresarse por medio del fasor P, que es un número complejo y a partir del cual puede derivarse otra medida de campo de onda posible, mediante

$$p(t) = \text{Re}\{P e^{j\omega t}\}$$

35 donde  $\text{Re}\{\cdot\}$  denota la parte real y  $\omega = 2\pi f$  es la frecuencia angular. Además, letras mayúsculas usadas para cantidades físicas representan fasores a continuación. Para el siguiente ejemplo introductorio y para evitar confusión, por favor obsérvese que todas las cantidades con el subíndice "PW" consideradas a continuación se refieren a ondas planas.

Para una onda plana monocromática ideal, el vector de velocidad de partícula  $\mathbf{U}_{PW}$  puede indicarse como

40

$$\mathbf{U}_{PW} = \frac{P_{PW}}{\rho_0 c} \mathbf{e}_d = \begin{bmatrix} U_x \\ U_y \\ U_z \end{bmatrix},$$

donde el vector unitario  $\mathbf{e}_d$  apunta hacia la dirección de propagación de la onda, por ejemplo correspondiente a una medida de dirección. Puede demostrarse que

$$\begin{aligned} \mathbf{I}_a &= \frac{1}{2\rho_0 c} |P_{PW}|^2 \mathbf{e}_d \\ E &= \frac{1}{2\rho_0 c^2} |P_{PW}|^2 \\ \Psi &= 0 \end{aligned} \quad , \text{ (a)}$$

5

donde  $\mathbf{I}_a$  denota la intensidad activa,  $\rho_0$  denota la densidad del aire,  $c$  denota la velocidad del sonido,  $E$  denota la energía del campo de sonido y  $\Psi$  denota la difusividad.

Es interesante observar que, puesto que todas las componentes de  $\mathbf{e}_d$  son números reales, las componentes de  $\mathbf{U}_{PW}$  están todas en fase con  $P_{PW}$ . La figura 1b ilustra un  $\mathbf{U}_{PW}$  y  $P_{PW}$  ejemplar en el plano gaussiano. Tal como acaba de

10 mencionarse, todas las componentes de  $\mathbf{U}_{PW}$  comparten la misma fase que  $P_{PW}$ , es decir  $\theta$ . Sus magnitudes, por otra parte, están limitadas a

$$\frac{|P_{PW}|}{c} = \sqrt{|U_x|^2 + |U_y|^2 + |U_z|^2} = \|\mathbf{U}_{PW}\|.$$

15 Aun cuando múltiples fuentes de sonido estén presentes, la presión y velocidad de partícula pueden todavía expresarse como una suma de componentes individuales. Sin pérdida de generalidad, puede exponerse el caso de dos fuentes de sonido. En efecto, la extensión a un mayor número de fuentes es directa. Sean  $P^{(1)}$  y  $P^{(2)}$  las presiones que se habrían registrado para la primera y la segunda fuente, respectivamente, por ejemplo que representan las medidas de campo de onda primera y segunda.

20 De manera similar, sean  $\mathbf{U}^{(1)}$  y  $\mathbf{U}^{(2)}$  los vectores de velocidad de partícula complejos. Dada la linealidad del fenómeno de propagación, cuando las fuentes se reproducen conjuntamente, la presión observada  $\mathbf{P}$  y la velocidad de partícula  $\mathbf{U}$  son

$$\begin{aligned} P &= P^{(1)} + P^{(2)} \\ \mathbf{U} &= \mathbf{U}^{(1)} + \mathbf{U}^{(2)} \end{aligned}$$

25

Por consiguiente, las intensidades activas son

$$\begin{aligned} \mathbf{I}_a^{(1)} &= \frac{1}{2} \operatorname{Re}\{P^{(1)} \cdot \overline{\mathbf{U}^{(1)}}\} \\ \mathbf{I}_a^{(2)} &= \frac{1}{2} \operatorname{Re}\{P^{(2)} \cdot \overline{\mathbf{U}^{(2)}}\} \end{aligned}$$

30

Así,

$$\mathbf{I}_a = \mathbf{I}_a^{(1)} + \mathbf{I}_a^{(2)} + \frac{1}{2} \operatorname{Re}\{P^{(1)} \cdot \overline{\mathbf{U}^{(2)}} + P^{(2)} \cdot \overline{\mathbf{U}^{(1)}}\}$$

35 Obsérvese que, aparte de casos especiales,

$$\mathbf{I}_a \neq \mathbf{I}_a^{(1)} + \mathbf{I}_a^{(2)}$$

Cuando las dos ondas, por ejemplo planas, están exactamente en fase (aunque desplazándose en direcciones diferentes),

$$P^{(2)} = \gamma \cdot P^{(1)}$$

donde  $\gamma$  es un número real. Se desprende que:

$$\begin{aligned} 5 \quad \mathbf{I}_a^{(1)} &= \frac{1}{2} \operatorname{Re}\{P^{(1)} \cdot \overline{\mathbf{U}^{(1)}}\} \\ \mathbf{I}_a^{(2)} &= \frac{1}{2} \operatorname{Re}\{P^{(2)} \cdot \overline{\mathbf{U}^{(2)}}\} \\ \|\mathbf{I}_a^{(2)}\| &= |\gamma|^2 \|\mathbf{I}_a^{(1)}\| \end{aligned}$$

y

$$10 \quad \mathbf{I}_a = (1 + \gamma)\mathbf{I}_a^{(1)} + (1 + \frac{1}{\gamma})\mathbf{I}_a^{(2)}.$$

Cuando las ondas están en fase y se desplazan en la misma dirección pueden ser interpretadas claramente como una onda.

15 Para  $\gamma = -1$  y cualquier dirección, la presión se desvanece y no puede haber ningún flujo de energía, es decir  $\|\mathbf{I}_a\| = 0$ .

Cuando las ondas están perfectamente en cuadratura, entonces

$$\begin{aligned} 20 \quad P^{(2)} &= \gamma \cdot e^{j\pi/2} P^{(1)} \\ \mathbf{U}^{(2)} &= \gamma \cdot e^{j\pi/2} \mathbf{U}^{(1)} \\ \mathbf{U}_x^{(2)} &= \gamma \cdot e^{j\pi/2} \mathbf{U}_x^{(1)} \\ \mathbf{U}_y^{(2)} &= \gamma \cdot e^{j\pi/2} \mathbf{U}_y^{(1)} \\ \mathbf{U}_z^{(2)} &= \gamma \cdot e^{j\pi/2} \mathbf{U}_z^{(1)} \end{aligned}$$

donde  $\gamma$  es un número real. A partir de esto, se desprende que:

$$\begin{aligned} 25 \quad \mathbf{I}_a^{(1)} &= \frac{1}{2} \operatorname{Re}\{P^{(1)} \cdot \overline{\mathbf{U}^{(1)}}\} \\ \mathbf{I}_a^{(2)} &= \frac{1}{2} \operatorname{Re}\{P^{(2)} \cdot \overline{\mathbf{U}^{(2)}}\} \\ \|\mathbf{I}_a^{(2)}\| &= |\gamma|^2 \|\mathbf{I}_a^{(1)}\| \end{aligned}$$

30 Y

$$\mathbf{I}_a = \mathbf{I}_a^{(1)} + \mathbf{I}_a^{(2)}.$$

Utilizando las ecuaciones anteriores, se puede demostrar fácilmente que, para una onda plana, cada una de las cantidades ejemplares  $\mathbf{U}$ ,  $\mathbf{P}$  y  $\mathbf{e}_d$  o  $\mathbf{P}$  y  $\mathbf{I}_a$  pueden representar una descripción equivalente y exhaustiva, ya que todas las demás cantidades físicas pueden derivarse de ellas, es decir cualquier combinación de las mismas puede usarse en algunas realizaciones en lugar de la medida de campo de onda o medida de dirección de onda. Por ejemplo, en realizaciones, la norma dos del vector de intensidad activa puede usarse como medida de campo de onda.

Una descripción mínima puede identificarse para efectuar la fusión tal como se especifica por las realizaciones. Los vectores de presión y velocidad de partícula para la  $i$ -ésima onda pueden expresarse como

$$\begin{aligned}
 P^{(i)} &= |P^{(i)}| e^{j\angle P^{(i)}} \\
 \mathbf{U}^{(i)} &= \frac{|P^{(i)}|}{\rho_0 c} \mathbf{e}_d^{(i)} e^{j\angle P^{(i)}}
 \end{aligned}$$

donde  $\angle P^{(i)}$  representa la fase de  $P^{(i)}$ . Expresando el vector de intensidad fusionado, es decir la medida de campo de onda fusionada y la medida de dirección de llegada fusionada, con respecto a estas variables, se desprende que:

$$\begin{aligned}
 \mathbf{I}_a &= \frac{1}{2\rho_0 c} |P^{(1)}|^2 \mathbf{e}_d^{(1)} + \frac{1}{2\rho_0 c} |P^{(2)}|^2 \mathbf{e}_d^{(2)} + \\
 &+ \frac{1}{2} \operatorname{Re} \left\{ |P^{(1)}| e^{j\angle P^{(1)}} \frac{|P^{(2)}|}{\rho_0 c} \mathbf{e}_d^{(2)} e^{-j\angle P^{(2)}} \right\} + \\
 &+ \frac{1}{2} \operatorname{Re} \left\{ |P^{(2)}| e^{j\angle P^{(2)}} \frac{|P^{(1)}|}{\rho_0 c} \mathbf{e}_d^{(1)} e^{-j\angle P^{(1)}} \right\}.
 \end{aligned}$$

Obsérvese que los primeros dos sumandos son  $\mathbf{I}_a^{(1)}$  y  $\mathbf{I}_a^{(2)}$ . La ecuación puede simplificarse adicionalmente a:

$$\begin{aligned}
 \mathbf{I}_a &= \frac{1}{2\rho_0 c} |P^{(1)}|^2 \mathbf{e}_d^{(1)} + \frac{1}{2\rho_0 c} |P^{(2)}|^2 \mathbf{e}_d^{(2)} + \\
 &+ \frac{1}{2\rho_0 c} |P^{(1)}| \cdot |P^{(2)}| \mathbf{e}_d^{(2)} \cdot \cos(\angle P^{(1)} - \angle P^{(2)}) + \\
 &+ \frac{1}{2\rho_0 c} |P^{(2)}| \cdot |P^{(1)}| \mathbf{e}_d^{(1)} \cdot \cos(\angle P^{(2)} - \angle P^{(1)}) .
 \end{aligned}$$

Introduciendo

$$\Delta^{(1,2)} = |\angle P^{(2)} - \angle P^{(1)}|$$

se obtiene

$$\mathbf{I}_a = \frac{1}{2\rho_0 c} \left\{ |P^{(1)}|^2 \mathbf{e}_d^{(1)} + |P^{(2)}|^2 \mathbf{e}_d^{(2)} + |P^{(1)}| \cdot |P^{(2)}| \cos(\Delta^{(1,2)}) \cdot (\mathbf{e}_d^{(1)} + \mathbf{e}_d^{(2)}) \right\} . \quad (\text{b})$$

Esta ecuación muestra que la información requerida para calcular  $\mathbf{I}_a$  puede reducirse a  $|P^{(i)}|$ ,  $\mathbf{e}_d^{(i)}$ ,  $|\angle P^{(2)} - \angle P^{(1)}|$ . En otras palabras, la representación para cada onda, por ejemplo plana, puede reducirse a la amplitud de la onda y la dirección de propagación. Además, la diferencia de fase relativa entre las ondas puede considerarse también. Cuando van a fusionarse más de dos ondas, las diferencias de fase entre todos los pares de ondas pueden considerarse. Claramente, existen otras diversas descripciones que contienen la misma información. Por ejemplo, conocer los vectores de intensidad y la diferencia de fase sería equivalente.

En general, una descripción energética de las ondas planas puede no ser suficiente para llevar a cabo la fusión correctamente. La fusión podría aproximarse suponiendo que las ondas están en cuadratura. Un descriptor exhaustivo de las ondas (es decir todas las cantidades físicas de la onda son conocidas) puede ser suficiente para la fusión, sin

embargo puede no ser necesario en todas las realizaciones. En realizaciones que llevan a cabo una fusión correcta, la amplitud de cada onda, la dirección de propagación de cada onda y la diferencia de fase relativa entre cada par de ondas que van a fusionarse, pueden tenerse en cuenta.

5 Los medios 110 para determinar pueden estar adaptados para proporcionar y/o el procesador 130 puede estar adaptado para procesar las direcciones de llegada primera y segunda y/o para proporcionar la medida de dirección de llegada fusionada en términos de un vector unitario  $\mathbf{e}_{DOA}(k, n)$ , con  $\mathbf{e}_{DOA}(k, n) = -\mathbf{e}_I(k, n)$  y  $\mathbf{I}_a(k, n) = \|\mathbf{I}_a(k, n)\| \cdot \mathbf{e}_I(k, n)$ , con

$$\mathbf{I}_a(k, n) = \frac{1}{2} \text{Re}\{P(k, n) \cdot \mathbf{U}^*(k, n)\}$$

y

10 
$$\mathbf{U}(k, n) = [U_x(k, n), U_y(k, n), U_z(k, n)]^T$$

que denotan el vector de velocidad de partícula  $\mathbf{u}(t) = [u_x(t), u_y(t), u_z(t)]^T$  transformado en tiempo-frecuencia. En otras palabras, sea  $p(t)$  y  $\mathbf{u}(t) = [u_x(t), u_y(t), u_z(t)]^T$  el vector de presión y de velocidad de partícula, respectivamente, para un punto específico en el espacio, donde  $[\cdot]^T$  denota la transpuesta. Estas señales pueden transformarse a un dominio de tiempo-frecuencia por medio de un banco de filtros apropiado, por ejemplo, una transformada de Fourier de tiempo corto (STFT) como se sugiere por ejemplo por V. Pulkki y C. Faller, Directional audio coding: Filterbank and STFT-based design, en la 120ª Convención de la AES, 20-23 de mayo de 2006, París, Francia, mayo de 2006.

15

Sea  $P(k, n)$  y  $\mathbf{U}(k, n) = [U_x(k, n), U_y(k, n), U_z(k, n)]^T$  que denotan las señales transformadas, donde  $k$  y  $n$  son índices de frecuencia (o banda de frecuencia) y tiempo, respectivamente. El vector de intensidad activo  $\mathbf{I}_a(k, n)$  puede definirse como

20

$$\mathbf{I}_a(k, n) = \frac{1}{2} \text{Re}\{P(k, n) \cdot \mathbf{U}^*(k, n)\} \quad (1)$$

donde  $(\cdot)^*$  denota la conjugación compleja y  $\text{Re}\{\cdot\}$  extrae la parte real. El vector de intensidad activo expresa el flujo de energía neto que caracteriza el campo de sonido, por ejemplo cf. F.J. Fahy, Sound Intensity, Essex: Elsevier Science Publishers Ltd., 1989, y puede usarse por tanto como una medida de campo de onda.

25

Sea  $c$  que denota la velocidad del sonido en el medio considerado y  $E$  la energía del campo de sonido definida por F.J. Fahy

$$E(k, n) = \frac{\rho_0}{4} \|\mathbf{U}(k, n)\|^2 + \frac{1}{4\rho_0 c^2} |P(k, n)|^2 \quad (2)$$

30 donde  $\|\cdot\|$  calcula la norma 2. A continuación se detallará el contenido de una corriente de DirAC mono.

La corriente de DirAC mono puede consistir en la señal mono  $p(t)$  y en información secundaria. Esta información secundaria puede comprender la dirección de llegada dependiente del tipo-frecuencia y una medida de difusividad dependiente del tiempo-frecuencia. La primera puede denotarse con  $\mathbf{e}_{DOA}(k, n)$  que es un vector unitario que apunta hacia la dirección desde la cual llega el sonido. La última, la difusividad, se denota por

35

$$\Psi(k, n).$$

En realizaciones, los medios 110 y/o el procesador 130 pueden estar adaptados para proporcionar/procesar las DOA primera y segunda y/o la DOA fusionada en términos de un vector unitario  $\mathbf{e}_{DOA}(k, n)$ . La dirección de llegada puede obtenerse como:

$$\mathbf{e}_{DOA}(k, n) = -\mathbf{e}_I(k, n)$$

donde el vector unitario  $\mathbf{e}_I(k, n)$  indica la dirección hacia la cual apunta la intensidad activa, es decir

$$\begin{aligned} \mathbf{I}_a(k, n) &= \|\mathbf{I}_a(k, n)\| \cdot \mathbf{e}_I(k, n) \\ \mathbf{e}_I(k, n) &= \mathbf{I}_a(k, n) / \|\mathbf{I}_a(k, n)\| \end{aligned} \quad (3)$$

Alternativamente, en realizaciones, la DOA puede expresarse en términos de ángulos de azimut y de elevación en un sistema de coordenadas esférico. Por ejemplo, si  $\varphi$  y  $\vartheta$  son los ángulos de azimut y de elevación, respectivamente, entonces

$$\mathbf{e}_{DOA}(k, n) = [\cos(\varphi) \cdot \cos(\vartheta), \sin(\varphi) \cdot \cos(\vartheta), \sin(\vartheta)]^T \quad (4)$$

En realizaciones, los medios 110 para determinar y/o el procesador 130 pueden estar adaptados para proporcionar/procesar los parámetros de difusividad primero y segundo y/o el parámetro de difusividad fusionado mediante  $\Psi(k, n)$  de manera dependiente del tiempo-frecuencia. Los medios 110 para determinar pueden estar adaptados para proporcionar los parámetros de difusividad primero y/o segundo y/o el procesador 130 puede estar adaptado para proporcionar un parámetro de difusividad fusionado en términos de:

$$\Psi(k, n) = 1 - \frac{\|\langle \mathbf{I}_a(k, n) \rangle_t\|}{c \langle E(k, n) \rangle_t} \quad (5)$$

donde  $\langle \cdot \rangle_t$  indica un promedio temporal.

Existen diferentes estrategias para obtener  $P(k, n)$  y  $\mathbf{U}(k, n)$  en la práctica. Una posibilidad es utilizar un micrófono de formato B, que entrega cuatro señales, es decir  $w(t)$ ,  $x(t)$ ,  $y(t)$  y  $z(t)$ . La primera,  $w(t)$ , corresponde a la lectura de presión de un micrófono omnidireccional. Las últimas tres son lecturas de presión de micrófonos que tienen patrones de captación en ocho dirigidos hacia los tres ejes de un sistema de coordenadas cartesiano. Estas señales son también proporcionales a la velocidad de partícula. Por consiguiente, en algunas realizaciones:

$$\begin{aligned} P(k, n) &= W(k, n) \\ \mathbf{U}(k, n) &= -\frac{1}{\sqrt{2}\rho_0 c} [X(k, n), Y(k, n), Z(k, n)]^T \end{aligned} \quad (6)$$

donde  $W(k, n)$ ,  $X(k, n)$ ,  $Y(k, n)$  y  $Z(k, n)$  son las señales de formato B transformadas. Obsérvese que el factor  $\sqrt{2}$  en la ecuación (6) viene de la convención usada en la definición de las señales de formato \*B, cf. Michael Gerzon, Surround sound psychoacoustics, In *WirelessWorld*, volumen 80, páginas 483-486, diciembre de 1974.

Alternativamente,  $P(k, n)$  y  $\mathbf{U}(k, n)$  pueden estimarse por medio de una disposición de micrófonos omnidireccionales tal como se sugiere en J. Merimaa, Applications of a 3-D microphone array, en la 112ª Convención de la AES, Documento 5501, Múnich, mayo de 2002. Las etapas de procesamiento descritas anteriormente se ilustran también en la figura 2.

La figura 2 muestra un codificador 200 de DirAC, que está adaptado para calcular un canal de audio mono e información secundaria a partir de señales de entrada apropiadas, por ejemplo, señales de micrófono. En otras palabras, la figura 2 ilustra un codificador 200 de DirAC para determinar la difusividad y la dirección de llegada a partir de señales de micrófono apropiadas. La figura 2 muestra un codificador 200 de DirAC que comprende una unidad 210 de estimación de P/U. La unidad de estimación de P/U recibe las señales del micrófono como información de entrada, en la que se basa la estimación de P/U. Puesto que toda la información está disponible, la estimación de P/U es directa según las ecuaciones anteriores. Una fase 220 de análisis energético permite la estimación de la dirección de llegada y el parámetro de difusividad de la corriente fusionada.

En realizaciones, pueden fusionarse otras corrientes de audio distintas de las corrientes de audio de DirAC mono. En otras palabras, en algunas realizaciones, los medios 110 para determinar pueden estar adaptados para convertir

5 cualquier otra corriente de audio en las corrientes de audio primera y segunda, tal como por ejemplo datos de audio estéreo o *surround*. En caso de que algunas realizaciones fusionen corrientes de DirAC distintas de mono, pueden distinguir entre diferentes casos. Si la corriente de DirAC transporta señales de formato B como señales de audio, entonces los vectores de velocidad de partícula serían conocidos y una fusión sería trivial, tal como se detallará posteriormente. Cuando la corriente de DirAC lleva señales de audio distintas de señales de formato B o una señal omnidireccional mono, los medios 110 para determinar pueden estar adaptados para la conversión a dos corrientes de DirAC mono en primer lugar, y una realización puede luego fusionar las corrientes convertidas de correspondiente. En realizaciones, las corrientes de audio espacial primera y segunda pueden por tanto representar corrientes de DirAC mono convertidas.

10 Algunas realizaciones pueden combinar los canales de audio disponibles para aproximar un patrón de captación omnidireccional. Por ejemplo, en el caso de una corriente de DirAC estéreo, esto puede obtenerse sumando el canal izquierdo L y el canal derecho R.

15 A continuación se expondrá la física en un campo generado por múltiples fuentes sonido. Cuando múltiples fuentes de sonido están presentes, todavía es posible expresar la presión y velocidad de partícula como una suma de componentes individuales.

Sea  $P^{(i)}(k, n)$  y  $U^{(i)}(k, n)$  la presión y la velocidad de partícula que se habrían registrado para la i-ésima fuente, si fuera a reproducirse sola. Suponiendo linealidad del fenómeno de propagación, cuando N fuentes se reproducen conjuntamente, la presión  $P(k, n)$  y velocidad de partícula  $U(k, n)$  observadas son

20 
$$P(k, n) = \sum_{i=1}^N P^{(i)}(k, n) \quad (7)$$

y

$$U(k, n) = \sum_{i=1}^N U^{(i)}(k, n) \quad (8)$$

25 Las ecuaciones previas muestran que si tanto la presión como la velocidad de partícula fueran conocidas, la obtención de la corriente de DirAC mono fusionada sería directa. Tal situación se ilustra en la figura 3. La figura 3 ilustra una realización que efectúa una fusión optimizada o posiblemente ideal de múltiples corrientes de audio. La figura 3 supone que todos los vectores de presión y velocidad de partícula son conocidos. Desafortunadamente, tal fusión trivial no es posible para corrientes de DirAC mono, para las que la velocidad de partícula  $U^{(i)}(k, n)$  no es conocida.

30 La figura 3 ilustra N corrientes, para cada una de las cuales se lleva a cabo una estimación de P/U en los bloques 301, 302-30N. El resultado de los bloques de estimación de P/U son las representaciones de tiempo-frecuencia correspondientes de las señales  $P^{(i)}(k, n)$  y  $U^{(i)}(k, n)$  individuales, que pueden luego combinarse según las ecuaciones anteriores (7) y (8), ilustrada por los dos sumadores 310 y 311. Una vez obtenidas  $P(k, n)$  y  $U(k, n)$  combinadas, una fase 320 de análisis energético puede determinar el parámetro de difusividad  $\Psi(k, n)$  y la dirección de

35 llegada  $e_{DOA}(k, n)$  de una manera directa.

La figura 4 ilustra una realización para fusionar múltiples corrientes de DirAC mono. Según la descripción anterior, N corrientes van a fusionarse mediante la realización de un aparato 100 ilustrado en la figura 4. Como se ilustra en la figura 4, cada una de las N corrientes de entrada pueden representarse por una representación mono dependiente del tiempo-frecuencia  $P^{(i)}(k, n)$ , una dirección de llegada  $e_{DOA}^{(i)}(k, n)$  y  $\Psi^{(i)}(k, n)$ , donde  $^{(1)}$  representa la primera corriente. Una representación correspondiente se ilustra también en la figura 4 para la corriente fusionada.

40 La tarea de fusionar dos o más corrientes de DirAC mono se ilustra en la figura 4. Dado que la presión  $P(k, n)$  puede obtenerse simplemente sumando las cantidades conocidas  $P^{(i)}(k, n)$  como en (7), el problema de fusionar dos o más corrientes de DirAC mono se reduce a la determinación de  $e_{DOA}(k, n)$  y  $\Psi(k, n)$ . La siguiente realización está basada en la suposición de que el campo de cada fuente consiste en una onda plana sumada a un campo difuso. Por

45 consiguiente, la presión y la velocidad de partícula para la i-ésima fuente pueden expresarse como

$$P^{(i)}(k, n) = P_{PW}^{(i)}(k, n) + P_{diff}^{(i)}(k, n) \quad (9)$$

$$U^{(i)}(k, n) = U_{PW}^{(i)}(k, n) + U_{diff}^{(i)}(k, n) , \quad (10)$$

5 donde los subíndices "PW" y "diff" denotan la onda plana y el campo difuso, respectivamente. A continuación se presenta una realización que tiene una estrategia para estimar la dirección de llegada del sonido y la difusividad. Las etapas de procesamiento correspondientes se ilustran en la figura 5.

La figura 5 ilustra otro aparato 500 para fusionar múltiples corrientes de audio que se detallará a continuación. La figura 5 ejemplifica el procesamiento de la primera corriente de audio espacial en términos de una primera representación

10 mono  $P^{(1)}$ , una primera dirección de llegada  $\hat{e}_{DOA}^{(1)}$  y un primer parámetro de difusividad  $\Psi^{(1)}$ . Según la figura 5, la primera corriente de audio espacial se descompone en una representación de onda plana aproximada  $\hat{P}_{PW}^{(1)}(k, n)$  así como la segunda corriente de audio espacial y, potencialmente, otras corrientes de audio espacial de manera correspondiente en  $\hat{P}_{PW}^{(2)}(k, n) \dots \hat{P}_{PW}^{(N)}(k, n)$ . Los valores estimativos están indicados por el sombrero encima de la representación de fórmula respectiva.

15 El estimador 120 puede estar adaptado para estimar una pluralidad de N representaciones de onda  $\hat{P}_{PW}^{(i)}(k, n)$  y representaciones de campo difuso  $\hat{P}_{diff}^{(i)}(k, n)$  como aproximaciones  $\hat{P}^{(i)}(k, n)$  para una pluralidad de N corrientes de audio espacial, con  $1 \leq i \leq N$ . El procesador 130 puede estar adaptado para determinar la dirección de llegada fusionada basándose en un valor estimativo,

20

$$\hat{e}_{DOA}(k, n) = -\frac{\hat{I}_z(k, n)}{\|\hat{I}_z(k, n)\|} , \text{ con}$$

con

$$\hat{I}_a(k, n) = \frac{1}{2} \text{Re}\{\hat{P}_{PW}(k, n) \cdot \hat{U}_{PW}^*(k, n)\}$$

25

$$\hat{P}_{PW}(k, n) = \sum_{i=1}^N \hat{P}_{PW}^{(i)}(k, n)$$

$$\hat{P}_{PW}^{(i)}(k, n) = \alpha^{(i)}(k, n) \cdot P^{(i)}(k, n)$$

30

$$\hat{U}_{PW}(k, n) = \sum_{i=1}^N \hat{U}_{PW}^{(i)}(k, n)$$

$$\hat{U}_{PW}^{(i)}(k, n) = -\frac{1}{\rho_0 c} \beta^{(i)}(k, n) \cdot P^{(i)}(k, n) \cdot \hat{e}_{DOA}^{(i)}(k, n)$$

35 Con los números reales  $\alpha^{(i)}(k, n), \beta^{(i)}(k, n) \in \{0 \dots 1\}$ .

La figura 5 muestra en líneas de puntos el estimador 120 y el procesador 130. En la realización mostrada en la figura 5, los medios 110 para determinar no están presentes, ya que se supone que la primera corriente de audio así como potencialmente otras corrientes de audio, se proporcionan en la representación de DirAC mono, es decir las representaciones mono, la DOA y los parámetros de difusividad simplemente se separan de la corriente. Como se

muestra en la figura 5, el procesador 130 puede estar adaptado para determinar la DOA fusionada basándose en un valor estimativo.

La dirección de llegada del sonido, es decir las medidas de dirección, puede estimarse por  $\hat{e}_{DOA}(k, n)$ , que se calcula como

5

$$\hat{e}_{DOA}(k, n) = -\frac{\hat{I}_a(k, n)}{\|\hat{I}_a(k, n)\|} \quad (11)$$

donde  $\hat{I}_a(k, n)$  es el valor estimativo de la intensidad activa para la corriente fusionada. Puede obtenerse como sigue:

10

$$\hat{I}_a(k, n) = \frac{1}{2} \text{Re} \left\{ \hat{P}_{PW}(k, n) \cdot \hat{U}_{PW}^*(k, n) \right\} \quad (12)$$

donde  $\hat{P}_{PW}(k, n)$  y  $\hat{U}_{PW}^*(k, n)$  son los valores estimativos de la presión y la velocidad de partícula correspondientes a las ondas planas, por ejemplo como medidas de campo de onda, solamente. Pueden definirse como:

15

$$\hat{P}_{PW}(k, n) = \sum_{i=1}^N \hat{P}_{PW}^{(i)}(k, n) \quad (13)$$

$$\hat{P}_{PW}^{(i)}(k, n) = \alpha^{(i)}(k, n) \cdot P^{(i)}(k, n) \quad (14)$$

20

$$\hat{U}_{PW}(k, n) = \sum_{i=1}^N \hat{U}_{PW}^{(i)}(k, n) \quad (15)$$

$$\hat{U}_{PW}^{(i)}(k, n) = -\frac{1}{\rho_0 c} \beta^{(i)}(k, n) \cdot P^{(i)}(k, n) \cdot e_{DOA}^{(i)}(k, n) \quad (16)$$

Los factores  $\alpha^{(i)}(k, n)$  y  $\beta^{(i)}(k, n)$  son en general dependientes de la frecuencia y pueden exhibir una proporcionalidad inversa a la difusividad  $\Psi^{(i)}(k, n)$ . En efecto, cuando la difusividad  $\Psi^{(i)}(k, n)$  está próxima a 0, se puede suponer que el campo está compuesto por una única onda plana, de tal manera que

25

$$\hat{P}_{PW}^{(i)}(k, n) \approx P(k, n) \quad \text{y} \quad (17)$$

30

$$\hat{U}_{PW}^{(i)}(k, n) \approx -\frac{1}{\rho_0 c} P^{(i)}(k, n) \cdot e_{DOA}^{(i)}(k, n) \quad (18)$$

y

lo que implica que  $\alpha^{(i)}(k, n) = \beta^{(i)}(k, n) = 1$ .

A continuación se presentarán dos realizaciones que determinan  $\alpha^{(i)}(k, n)$  y  $\beta^{(i)}(k, n)$ . En primer lugar se consideran consideraciones energéticas de los campos difusos. En realizaciones, el estimador 120 puede estar

35

adaptado para determinar los factores  $\alpha^{(i)}(k, n)$  y  $\beta^{(i)}(k, n)$  basándose en los campos difusos. Algunas

realizaciones pueden suponer que el campo está compuesto por una onda plana sumada a un campo difuso ideal. En realizaciones, el estimador 120 puede estar adaptado para determinar  $\alpha^{(i)}(k, n)$  y  $\beta^{(i)}(k, n)$  según

$$\begin{aligned} \alpha^{(i)}(k, n) &= \beta^{(i)}(k, n) \\ \beta^{(i)}(k, n) &= \sqrt{1 - \Psi^{(i)}(k, n)} \end{aligned} \quad (19)$$

estableciendo la densidad del aire  $\rho_0$  igual a 1 y abandonando la dependencia funcional  $(k, n)$  por motivos de simplicidad, se puede escribir:

$$\Psi^{(i)} = 1 - \frac{\langle |P_{PW}^{(i)}|^2 \rangle_t}{\langle |P_{PW}^{(i)}|^2 \rangle_t + 2c^2 \langle E_{diff} \rangle_t} \quad (20)$$

En realizaciones, el procesador 130 puede estar adaptado para aproximar los campos difusos basándose en sus propiedades estadísticas, una aproximación puede obtenerse por:

$$\langle |P_{PW}^{(i)}|^2 \rangle_t + 2c^2 \langle E_{diff} \rangle_t \approx \langle |P^{(i)}|^2 \rangle_t \quad (21)$$

donde  $E_{diff}$  es la energía del campo difuso. Algunas realizaciones pueden por tanto estimar:

$$\langle |P_{PW}^{(i)}| \rangle_t \approx \langle |\hat{P}_{PW}^{(i)}| \rangle_t = \sqrt{1 - \Psi^{(i)}} \langle |P^{(i)}| \rangle_t \quad (22)$$

Para calcular valores estimativos instantáneos (es decir para cada loseta de tiempo-frecuencia), algunas realizaciones pueden eliminar los operadores de esperanza, obteniendo:

$$\hat{P}_{PW}^{(i)}(k, n) = \sqrt{1 - \Psi^{(i)}(k, n)} P^{(i)}(k, n) \quad (23)$$

Aprovechando la suposición de onda plana, el valor estimativo para la velocidad de partícula puede derivarse directamente

$$\hat{U}_{PW}^{(i)}(k, n) = \frac{1}{c\rho_0} \hat{P}_{PW}^{(i)}(k, n) \cdot \mathbf{e}_I^{(i)}(k, n) \quad (24)$$

En realizaciones, puede aplicarse un modelado simplificado de la velocidad de partícula. En realizaciones, el estimador 120 puede estar adaptado para aproximar los factores  $\alpha^{(i)}(k, n)$  y  $\beta^{(i)}(k, n)$  basándose en el modelado simplificado. Algunas realizaciones pueden utilizar una solución alternativa, que puede derivarse introduciendo un modelado simplificado de la velocidad de partícula:

A continuación se proporciona una derivación. La velocidad de partícula  $U^{(i)}(k, n)$  se modela como

$$\alpha^{(i)}(k, n) = 1$$

$$\beta^{(i)}(k, n) = \frac{1 - \sqrt{1 - (1 - \Psi^{(i)}(k, n))^2}}{1 - \Psi^{(i)}(k, n)} \quad (25)$$

$$U^{(i)}(k, n) = \beta^{(i)}(k, n) \cdot \frac{P^{(i)}}{\rho_0 c} \cdot e_I^{(i)}(k, n) \quad (26)$$

El factor  $\beta^{(i)}(k, n)$  puede obtenerse sustituyendo (26) en (5), lo que lleva a

$$\Psi^{(i)}(k, n) = 1 - \frac{\frac{1}{\rho_0 c} \left\| \left\langle \left| \beta^{(i)}(k, n) \cdot P^{(i)}(k, n) \right|^2 \cdot e_I^{(i)}(k, n) \right\rangle_t \right\|}{c < \frac{1}{2\rho_0 c^2} \left| P^{(i)}(k, n) \right|^2 \cdot (\beta^{(i)^2}(k, n) + 1) >_t} \quad (27)$$

Para obtener valores instantáneos, los operadores de esperanza pueden eliminarse y resolverse para  $\beta^{(i)}(k, n)$ , obteniendo

$$\beta^{(i)}(k, n) = \frac{1 - \sqrt{1 - (1 - \Psi^{(i)}(k, n))^2}}{1 - \Psi^{(i)}(k, n)} \quad (28)$$

Obsérvese que este procedimiento conduce a direcciones de llegada de sonido similares como la dada en (19), sin embargo, con una complejidad computacional más baja dado que el factor  $\alpha^{(i)}(k, n)$  es la unidad.

En realizaciones, el procesador 130 puede estar adaptado para estimar la difusividad, es decir para estimar el parámetro de difusividad fusionado. La difusividad de la corriente fusionada, denotada por  $\Psi(k, n)$ , puede estimarse directamente a partir de las cantidades conocidas  $\Psi^{(i)}(k, n)$  y  $P^{(i)}(k, n)$  y del valor estimativo  $\hat{\mathbf{I}}_a(k, n)$  obtenido tal como se describió anteriormente. Siguiendo las consideraciones energéticas introducidas en la sección previa, algunas realizaciones pueden usar el estimador:

$$\Psi(k, n) = 1 - \frac{\left\| \left\langle \hat{\mathbf{I}}_a(k, n) \right\rangle_t \right\|}{\left\| \hat{\mathbf{I}}_a(k, n) \right\| + \frac{1}{2c} \sum_{i=1}^2 \Psi^{(i)}(k, n) \cdot \left| P^{(i)}(k, n) \right|^2} >_t \quad (29)$$

El conocimiento de  $\hat{P}_{PW}^{(i)}$  y  $\hat{U}_{PW}^{(i)}$  permite el uso de las representaciones alternativas dadas en la ecuación (b) en algunas realizaciones. En efecto, la dirección de la onda puede obtenerse por  $\hat{U}_{PW}^{(i)}$  mientras que  $\hat{P}_{PW}^{(i)}$  da la amplitud y fase de la  $i$ -ésima onda. A partir de esta última, todas las diferencias de fase  $\Delta^{(i,j)}$  pueden calcularse fácilmente. Los parámetros de DirAC de la corriente fusionada pueden calcularse entonces sustituyendo la ecuación (b) en la ecuación (a), (3) y (5).

La figura 6 ilustra una realización de un método para fusionar dos o más corrientes de DirAC. Algunas realizaciones pueden proporcionar un método para fusionar una primera corriente de audio espacial con una segunda corriente de audio espacial para obtener una corriente de audio fusionada. En realizaciones, el método puede comprender una etapa de determinar, para la primera corriente de audio espacial, una primera representación de audio y una primera DOA, así como para la segunda corriente de audio espacial, una segunda representación de audio y una segunda DOA. En realizaciones, las representaciones de DirAC de las corrientes de audio espacial pueden estar disponibles, la etapa de determinación lee entonces simplemente las representaciones correspondientes a partir de las corrientes de audio. En

la figura 6, se supone que las dos o más corrientes de DirAC pueden obtenerse simplemente a partir de las corrientes de audio según la etapa 610.

- 5 En realizaciones, el método puede comprender una etapa de estimar una primera representación de onda que comprende una primera medida de dirección de onda y una primera medida de campo de onda para la primera corriente de audio espacial basándose en la primera representación de audio, la primera DOA y opcionalmente un primer parámetro de difusividad. Por consiguiente, el método puede comprender una etapa de estimar una segunda representación de onda que comprende una segunda medida de dirección de onda y una segunda medida de campo de onda para la segunda corriente de audio espacial basada en la segunda representación de audio, la segunda DOA y opcionalmente un segundo parámetro de difusividad.
- 10 El método puede comprender además una etapa de combinar la primera representación de onda y la segunda representación de onda para obtener una representación de onda fusionada que comprende una medida de campo fusionada y una medida de DOA fusionada y una etapa de combinar la primera representación de audio y la segunda representación de audio para obtener una representación de audio fusionada, que se indica en la figura 6 por la etapa 620 para canales de audio mono. La realización ilustrada en la figura 6 comprende una etapa de calcular  $\alpha^{(i)}(k, n)$  y
- 15  $\beta^{(i)}(k, n)$  según (19) y (25) permitiendo la estimación de los vectores de presión y de velocidad de partícula para las representaciones de onda plana en la etapa 640. En otras palabras, las etapas de estimar las representaciones de onda plana primera y segunda se llevan a cabo en las etapas 630 y 640 en la figura 6 en términos de representaciones de onda plana.
- 20 La etapa de combinar las representaciones de onda plana primera y segunda se lleva a cabo en la etapa 650, en la que los vectores de presión y de velocidad de partícula de todas las corrientes pueden sumarse.
- En la etapa 660 de la figura 6, el cálculo del vector de intensidad activa y la estimación de la DOA se llevan a cabo basándose en la representación de onda plana fusionada.
- 25 Algunas realizaciones pueden comprender una etapa de combinar o procesar la medida de campo fusionada, las representaciones mono primera y segunda y los parámetros de difusividad primero y segundo para obtener un parámetro de difusividad fusionado. En la realización ilustrada en la figura 6, el cálculo de la difusividad se lleva a cabo en la etapa 670, por ejemplo basándose en (29).
- Algunas realizaciones pueden proporcionar la ventaja de que la fusión de las corrientes de audio espacial se puede efectuar con alta calidad y complejidad moderada.
- 30 Dependiendo de determinados requisitos de implementación de los métodos de la invención, los métodos de la invención pueden implementarse en hardware o software. La implementación puede realizarse utilizando un medio de almacenamiento digital y particularmente una memoria *flash*, un disco, un DVD o un CD con señales de control legibles electrónicamente almacenadas en los mismos, que actúan conjuntamente con un sistema informático programable de tal manera que se llevan a cabo los métodos de la invención. En general, la presente invención es, por consiguiente, un código de programa informático con un código de programa almacenado en un soporte legible por máquina, siendo el
- 35 código de programa operativo para llevar a cabo los métodos de la invención cuando el programa informático se ejecuta en un ordenador o procesador. En otras palabras, los métodos de la invención son, por consiguiente, un programa informático que tiene un código de programa para llevar a cabo al menos uno de los métodos de la invención, cuando el programa informático se ejecuta en un ordenador.

REIVINDICACIONES

1. Aparato (100) para fusionar una primera corriente de audio espacial con una segunda corriente de audio espacial para obtener una corriente de audio fusionada, que comprende:  
 5 un estimador (120) para estimar una primera representación de onda que comprende una primera medida de dirección de onda  $(\hat{\mathbf{U}}_{PW}^{(1)})$  que es una cantidad direccional de una primera onda y una primera medida de campo de onda  $(\hat{P}_{PW}^{(1)})$  que está relacionada con una magnitud de la primera onda para la primera corriente de audio espacial, teniendo la primera corriente de audio espacial una primera representación de audio que comprende una medida de una presión o una magnitud de una primera señal de audio ( $P^{(1)}$ ) y una primera  
 10 dirección de llegada  $(\hat{\mathbf{e}}_{DOA}^{(1)})$ , y para estimar una segunda representación de onda que comprende una segunda medida de dirección de onda que es una cantidad direccional de una segunda onda  $(\hat{\mathbf{U}}_{PW}^{(2)})$  y una segunda medida de campo de onda  $(\hat{P}_{PW}^{(2)})$  que está relacionada con una magnitud de la segunda onda para la segunda corriente de audio espacial, teniendo la segunda corriente de audio espacial una segunda representación de audio que comprende una medida de una presión o una magnitud de una segunda señal de  
 15 audio ( $P^{(2)}$ ) y una segunda dirección de llegada  $(\hat{\mathbf{e}}_{DOA}^{(2)})$ ; y  
 un procesador (130) para procesar la primera representación de onda y la segunda representación de onda para obtener una representación de onda fusionada que comprende una medida de campo de onda fusionada  $(\hat{I}_a)$ , una medida de dirección de llegada fusionada  $(\hat{\mathbf{e}}_{DOA})$  y un parámetro de difusividad fusionado  $(\hat{\Psi})$ ,  
 en el que el parámetro de difusividad fusionado está basado en la medida de campo de onda fusionada  $(\hat{I}_a)$ , la  
 20 primera representación de audio ( $P^{(1)}$ ) y la segunda representación de audio ( $P^{(2)}$ ), y  
 en el que la medida de campo de onda fusionada  $(\hat{I}_a)$  está basada en la primera medida de campo de onda, la segunda medida de campo de onda, la primera medida de dirección de onda  $(\hat{\mathbf{U}}_{PW}^{(1)})$ , y la segunda medida de  
 25 dirección de onda  $(\hat{\mathbf{U}}_{PW}^{(2)})$ , y  
 en el que el procesador (130) está configurado para procesar la primera representación de audio ( $P^{(1)}$ ) y la segunda representación de audio ( $P^{(2)}$ ) para obtener una representación de audio fusionada ( $P$ ), y para proporcionar la corriente de audio fusionada que comprende la representación de audio fusionada ( $P$ ), la medida de dirección de llegada fusionada  $(\hat{\mathbf{e}}_{DOA})$  y el parámetro de difusividad fusionado  $(\hat{\Psi})$ .
2. Aparato (100) según la reivindicación 1, en el que el estimador (120) está adaptado para estimar la primera medida de campo de onda en términos de una primera amplitud de campo de onda y para estimar la segunda medida de campo de onda en términos de una segunda amplitud de campo de onda, y para estimar una diferencia de fase entre la primera medida de campo de onda y la segunda medida de campo de onda, y/o para estimar una primera fase de campo de onda y una segunda fase de campo de onda.
3. Aparato según una de las reivindicaciones 1 a 2, que comprende medios (110) para determinar, para la primera corriente de audio espacial, la primera representación de audio, la primera medida de dirección de llegada y el primer parámetro de difusividad y para determinar, para la segunda corriente de audio espacial, la segunda representación de audio, la segunda medida de dirección de llegada y el segundo parámetro de difusividad.
4. Aparato según una de las reivindicaciones 1 a 3, en el que el procesador (130) está adaptado para determinar la representación de audio fusionada, la medida de dirección de llegada fusionada y el parámetro de difusividad fusionado de una manera dependiente del tiempo-frecuencia.
5. Aparato (100) según una de las reivindicaciones 1 a 4, en el que el estimador (120) está adaptado para estimar las representaciones de onda primera y/o segunda, y en el que el procesador (130) está adaptado para proporcionar la representación de audio fusionada en términos de una señal de presión  $p(t)$  o una señal de presión transformada en tiempo-frecuencia  $P(k,n)$ , donde  $k$  denota un índice de frecuencia y  $n$  denota un índice de tiempo.
6. Aparato (100) según la reivindicación 5, en el que el procesador (130) está adaptado para procesar las medidas de direcciones de llegada primera y segunda y/o para proporcionar la medida de dirección de llegada

fusionada en términos de un vector unitario  $\mathbf{e}_{DOA}(k, n)$ , con

$$\mathbf{e}_{DOA}(k, n) = -\mathbf{e}_l(k, n)$$

y

$$\mathbf{I}_a(k, n) = \|\mathbf{I}_a(k, n)\| \cdot \mathbf{e}_l(k, n),$$

con

$$\mathbf{I}_a(k, n) = \frac{1}{2} \operatorname{Re}\{P(k, n) \cdot \mathbf{U}^*(k, n)\}$$

en el que  $P(k, n)$  es la presión de la corriente fusionada y  $\mathbf{U}(k, n) = [U_x(k, n), U_y(k, n), U_z(k, n)]^T$  denota el vector de velocidad de partícula  $\mathbf{u}(t) = [u_x(t), u_y(t), u_z(t)]^T$  transformado en tiempo-frecuencia de la corriente de audio fusionada, donde  $\operatorname{Re}\{\cdot\}$  denota la parte real.

7. Aparato (100) según una de la reivindicación 6, en el que el procesador (130) está adaptado para procesar los parámetros de difusividad primero y/o segundo y/o para proporcionar el parámetro de difusividad fusionado en términos de

$$\Psi(k, n) = 1 - \frac{\|\langle \mathbf{I}_a(k, n) \rangle_t\|}{c \langle E(k, n) \rangle_t}, \quad (k, n)$$

$$\mathbf{I}_a(k, n) = \frac{1}{2} \operatorname{Re}\{P(k, n) \cdot \mathbf{U}^*(k, n)\}$$

y  $\mathbf{U}(k, n) = [U_x(k, n), U_y(k, n), U_z(k, n)]^T$  denota un vector de velocidad de partícula  $\mathbf{u}(t) = [u_x(t), u_y(t), u_z(t)]^T$  transformado en tiempo-frecuencia,  $\operatorname{Re}\{\cdot\}$  denota la parte real,  $P(k, n)$  denota una señal de presión  $p(t)$  transformada en tiempo-frecuencia, donde  $k$  denota un índice de frecuencia y  $n$  denota un índice de tiempo,  $c$  es la velocidad del sonido y

$$E(k, n) = \frac{\rho_0}{4} \|\mathbf{U}(k, n)\|^2 + \frac{1}{4\rho_0 c^2} |P(k, n)|^2$$

denota la energía de campo de sonido, donde  $\rho_0$  denota la densidad del aire y  $\langle \cdot \rangle_t$  denota un promedio temporal.

8. Aparato (100) según la reivindicación 7, en el que el estimador (120) está adaptado para estimar una pluralidad de  $N$  representaciones de onda  $\hat{P}_{PW}^{(i)}(k, n)$  y representaciones de campo difuso  $\hat{P}_{diff}^{(i)}(k, n)$  como aproximaciones para una pluralidad de  $N$  corrientes de audio espacial  $\hat{P}_{diff}^{(i)}(k, n)$ , con  $1 \leq i \leq N$ , y en el que el procesador (130) está adaptado para determinar la medida de dirección de llegada fusionada basándose en un valor estimativo,

$$\hat{\mathbf{e}}_{DOA}(k, n) = -\frac{\hat{\mathbf{I}}_a(k, n)}{\|\hat{\mathbf{I}}_a(k, n)\|}$$

5

$$\hat{\mathbf{I}}_a(k, n) = \frac{1}{2} \operatorname{Re} \left\{ \hat{P}_{PW}(k, n) \cdot \hat{\mathbf{U}}_{PW}^*(k, n) \right\}$$

10

$$\hat{P}_{PW}(k, n) = \sum_{i=1}^N \hat{P}_{PW}^{(i)}(k, n)$$

15

$$\hat{P}_{PW}^{(i)}(k, n) = \alpha^{(i)}(k, n) \cdot P^{(i)}(k, n)$$

20

$$\hat{\mathbf{U}}_{PW}(k, n) = \sum_{i=1}^N \hat{\mathbf{U}}_{PW}^{(i)}(k, n)$$

25

$$\hat{\mathbf{U}}_{PW}^{(i)}(k, n) = -\frac{1}{\rho_0 c} \beta^{(i)}(k, n) \cdot P^{(i)}(k, n) \cdot \mathbf{e}_{DOA}^{(i)}(k, n)$$

30

denotando los números reales  $\alpha^{(i)}(k, n), \beta^{(i)}(k, n) \in \{0 \dots 1\}$  y  $\mathbf{U}(k, n) = [U_x(k, n), U_y(k, n), U_z(k, n)]^T$  un vector de velocidad de partícula  $\mathbf{u}(t) = [u_x(t), u_y(t), u_z(t)]^T$  transformado en tiempo-frecuencia,  $\operatorname{Re}\{\cdot\}$  denota la parte real,  $P^{(i)}(k, n)$  denota una señal de presión transformada en tiempo-frecuencia  $p^{(i)}(t)$ , donde  $k$  denota un índice de frecuencia y  $n$  denota un índice de tiempo,  $N$  el número de corrientes de audio espacial,  $c$  es la velocidad del sonido y  $\rho_0$  denota la densidad del aire.

35

9. Aparato (100) según la reivindicación 8, en el que el estimador (120) está adaptado para determinar  $\alpha^{(i)}(k, n)$  y  $\beta^{(i)}(k, n)$  según

$$\begin{aligned} \alpha^{(i)}(k, n) &= \beta^{(i)}(k, n) \\ \beta^{(i)}(k, n) &= \sqrt{1 - \Psi^{(i)}(k, n)} \end{aligned}$$

40

10. Aparato (100) según la reivindicación 8, en el que el procesador (130) está adaptado para determinar  $\alpha^{(i)}(k, n)$  y  $\beta^{(i)}(k, n)$  mediante

$$\begin{aligned} \alpha^{(i)}(k, n) &= 1 \\ \beta^{(i)}(k, n) &= \frac{1 - \sqrt{1 - (1 - \Psi^{(i)}(k, n))^2}}{1 - \Psi^{(i)}(k, n)} \end{aligned}$$

45

50

11. Aparato (100) según una de las reivindicaciones 9 a 10, en el que el procesador (130) está adaptado para determinar el parámetro de difusividad fusionado mediante

55

$$\hat{\Psi}(k, n) = 1 - \frac{\| \langle \hat{\mathbf{I}}_a(k, n) \rangle_t \|}{\| \hat{\mathbf{I}}_a(k, n) \| + \frac{1}{2c} \sum_{i=1}^2 \Psi^{(i)}(k, n) \cdot |P^{(i)}(k, n)|^2}_{>_t}$$

12. Aparato según la reivindicación 1, en el que la primera corriente de audio espacial comprende

adicionalmente un primer parámetro de difusividad ( $\hat{\Psi}^{(1)}$ ), en el que la segunda corriente de audio espacial comprende adicionalmente un segundo parámetro de difusividad ( $\hat{\Psi}^{(2)}$ ), y

en el que el procesador (130) está configurado para calcular el parámetro de difusividad fusionado ( $\hat{\Psi}$ ) basándose adicionalmente en el primer parámetro de difusividad ( $\Psi^{(1)}$ ) y el segundo parámetro de difusividad ( $\Psi^{(2)}$ ).

13. Método para fusionar una primera corriente de audio espacial con una segunda corriente de audio espacial para obtener una corriente de audio fusionada, que comprende:

estimar una primera representación de onda que comprende una primera medida de dirección de onda ( $\hat{U}_{PW}^{(1)}$ ) que es una cantidad direccional de una primera onda y una primera medida de campo de onda ( $\hat{P}_{PW}^{(1)}$ ) que está relacionada con una magnitud de la primera onda para la primera corriente de audio espacial, teniendo la primera corriente de audio espacial una primera representación de audio que comprende una medida de una presión o una magnitud de una primera señal de audio ( $P^{(1)}$ ) y una primera dirección de llegada ( $e_{DOA}^{(1)}$ );

estimar una segunda representación de onda que comprende una segunda medida de dirección de onda ( $\hat{U}_{PW}^{(2)}$ ) que es una cantidad direccional de una segunda onda y una segunda medida de campo de onda ( $\hat{P}_{PW}^{(2)}$ ) que está relacionada con una magnitud de la segunda onda para la segunda corriente de audio espacial, teniendo la segunda corriente de audio espacial una segunda representación de audio que comprende una medida de una presión o una magnitud de una segunda señal de audio ( $P^{(2)}$ ) y una segunda dirección de llegada ( $e_{DOA}^{(2)}$ );

procesar la primera representación de onda y la segunda representación de onda para obtener una representación de onda fusionada que tiene una medida de campo de onda fusionada ( $\hat{I}_a$ ), una medida de dirección de llegada fusionada ( $\hat{e}_{DOA}$ ) y un parámetro de difusividad fusionado ( $\hat{\Psi}$ ), en el que el parámetro de difusividad fusionado ( $\hat{\Psi}$ ) está basado en la medida de campo de onda fusionada ( $\hat{I}_a$ ), la primera representación de audio ( $P^{(1)}$ ) y la segunda representación de audio ( $P^{(2)}$ ), y en el que la medida de campo de onda fusionada ( $\hat{I}_a$ ) está basada en la primera medida de campo de onda, la segunda medida de campo de onda, la primera medida de dirección de onda ( $\hat{U}_{PW}^{(1)}$ ), y la segunda medida de dirección de onda ( $\hat{U}_{PW}^{(2)}$ );

procesar la primera representación de audio ( $P^{(1)}$ ) y la segunda representación de audio ( $P^{(2)}$ ) para obtener una representación de audio fusionada (P); y proporcionar la corriente de audio fusionada que comprende la representación de audio fusionada (P), una medida de dirección de llegada fusionada ( $\hat{e}_{DOA}$ ) y el parámetro de difusividad fusionado ( $\hat{\Psi}$ ).

14. Método según la reivindicación 13, en el que la primera corriente de audio espacial comprende adicionalmente un primer parámetro de difusividad ( $\Psi^{(1)}$ ), en el que la segunda corriente de audio espacial comprende adicionalmente un segundo parámetro de difusividad ( $\Psi^{(2)}$ ), y en el que el parámetro de difusividad fusionado ( $\hat{\Psi}$ ) se calcula en la etapa de procesamiento basándose adicionalmente en el primer parámetro de difusividad ( $\Psi^{(1)}$ ) y el segundo parámetro de difusividad ( $\Psi^{(2)}$ ).

15. Programa informático que tiene un código de programa para realizar el método de la reivindicación 14, cuando el código de programa se ejecuta en un ordenador o un procesador.

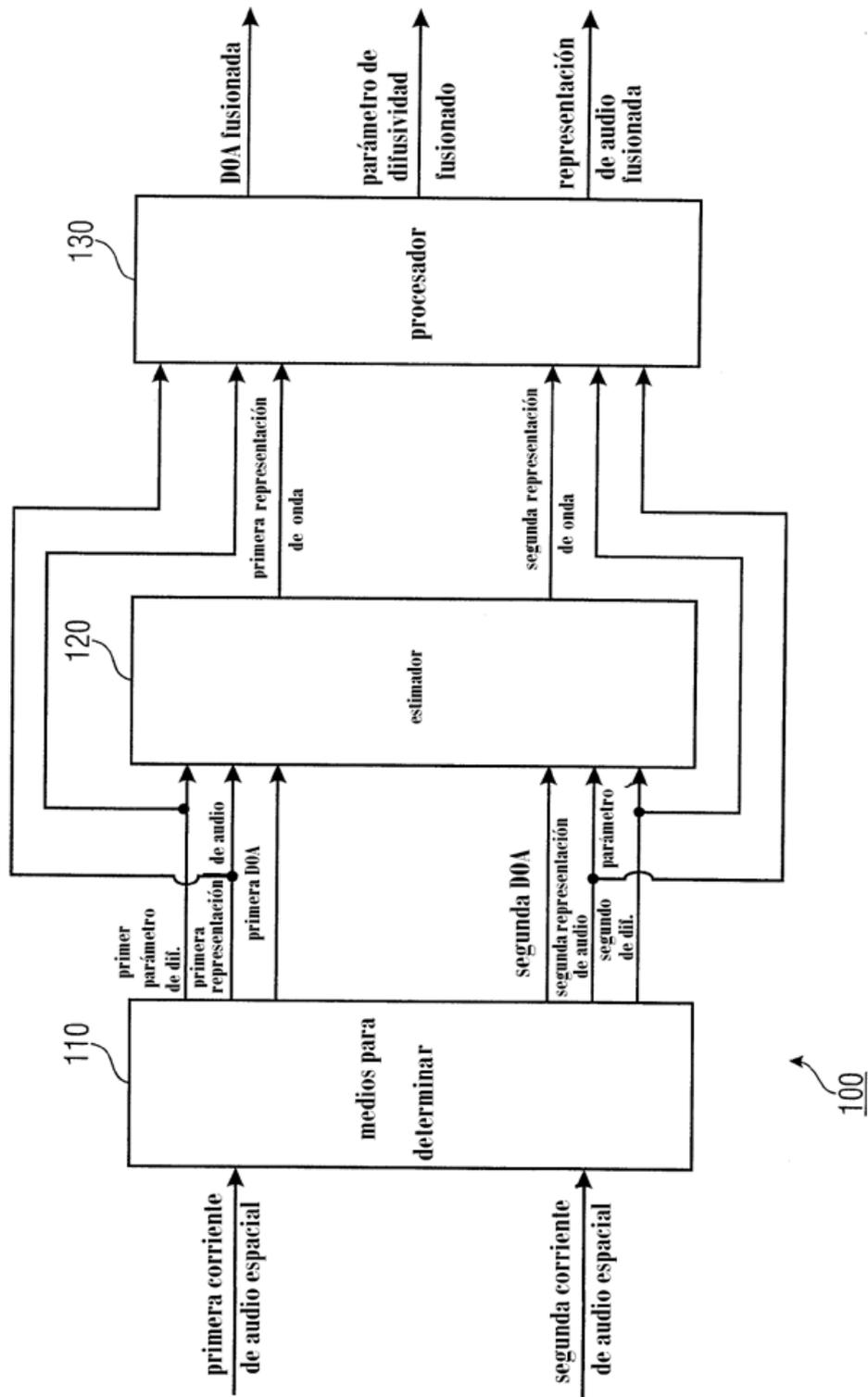


FIGURA 1A

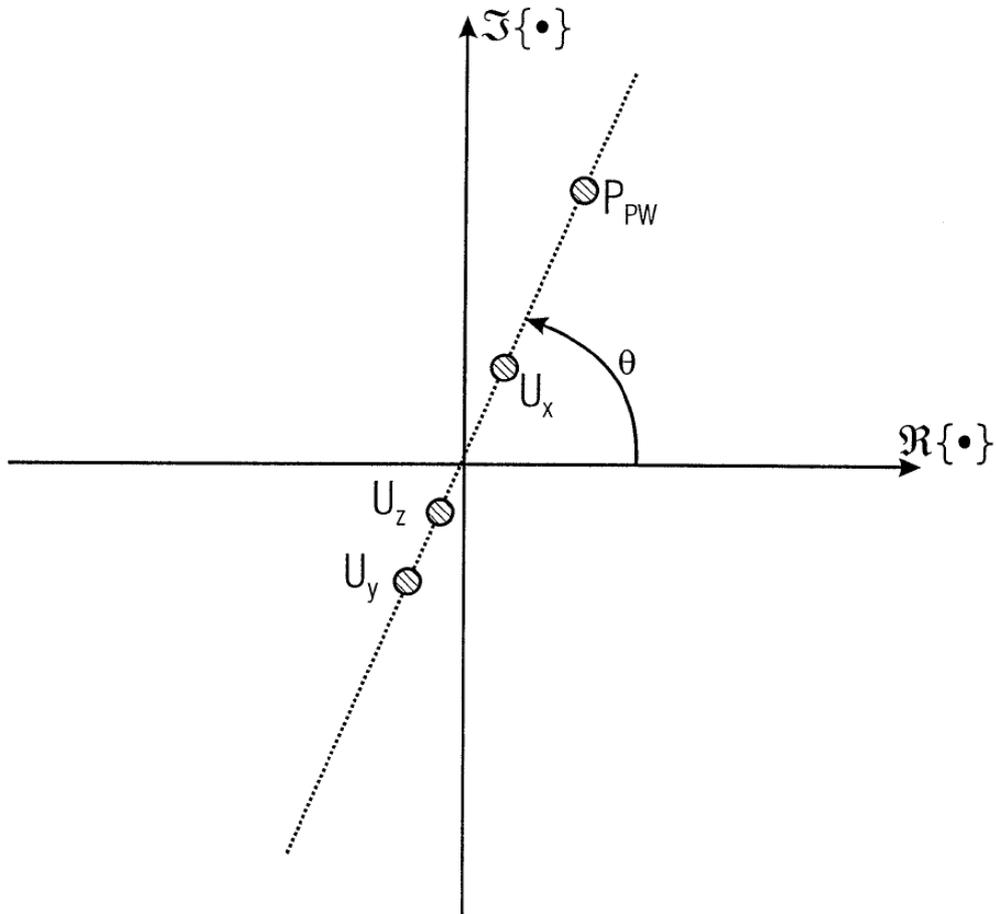


FIGURA 1B

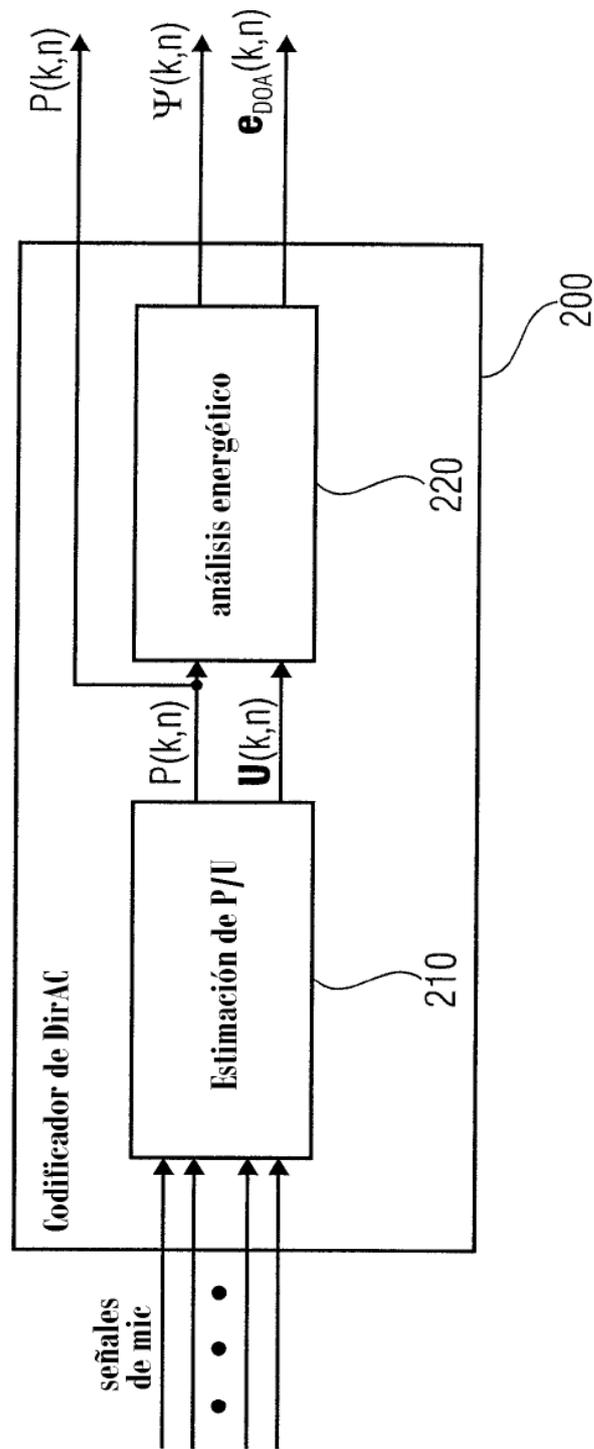


FIGURA 2

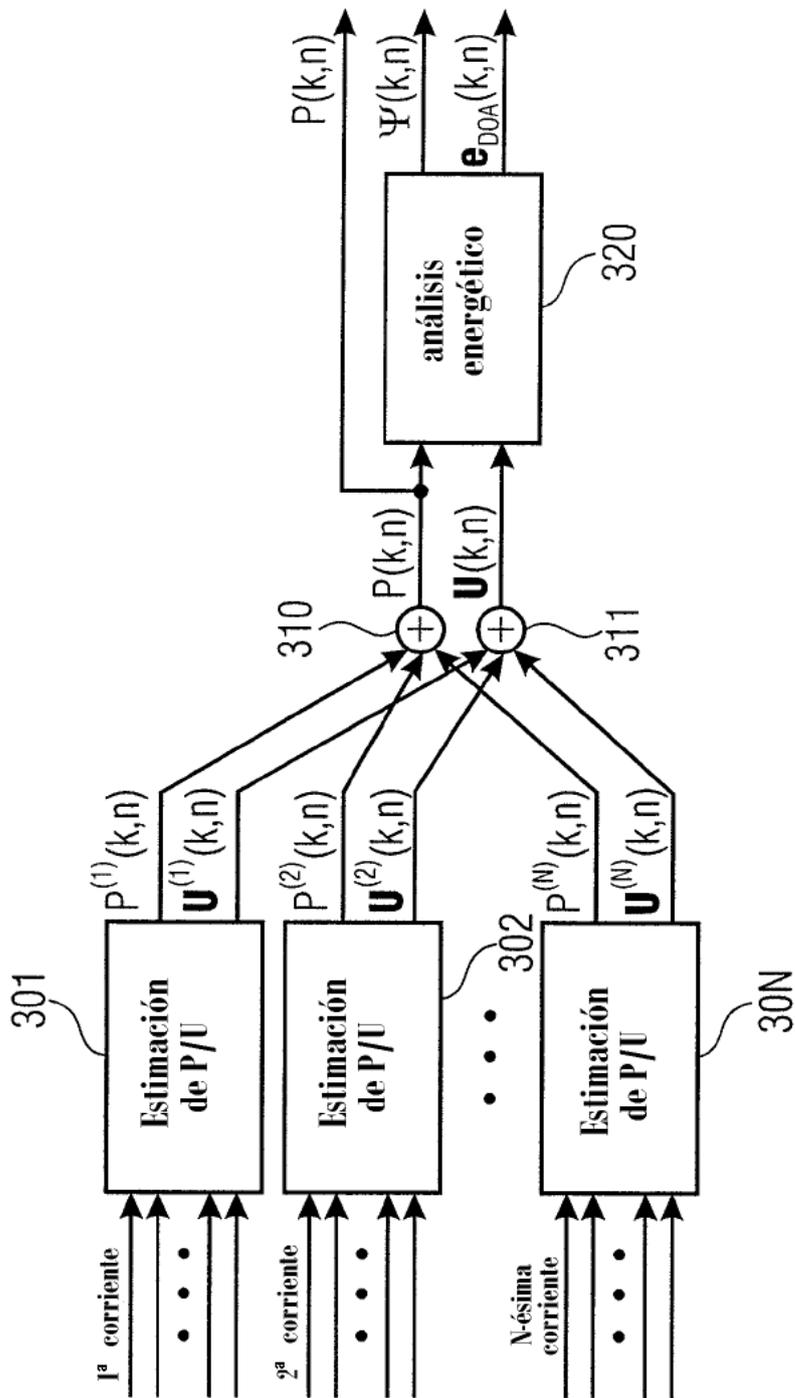


FIGURA 3

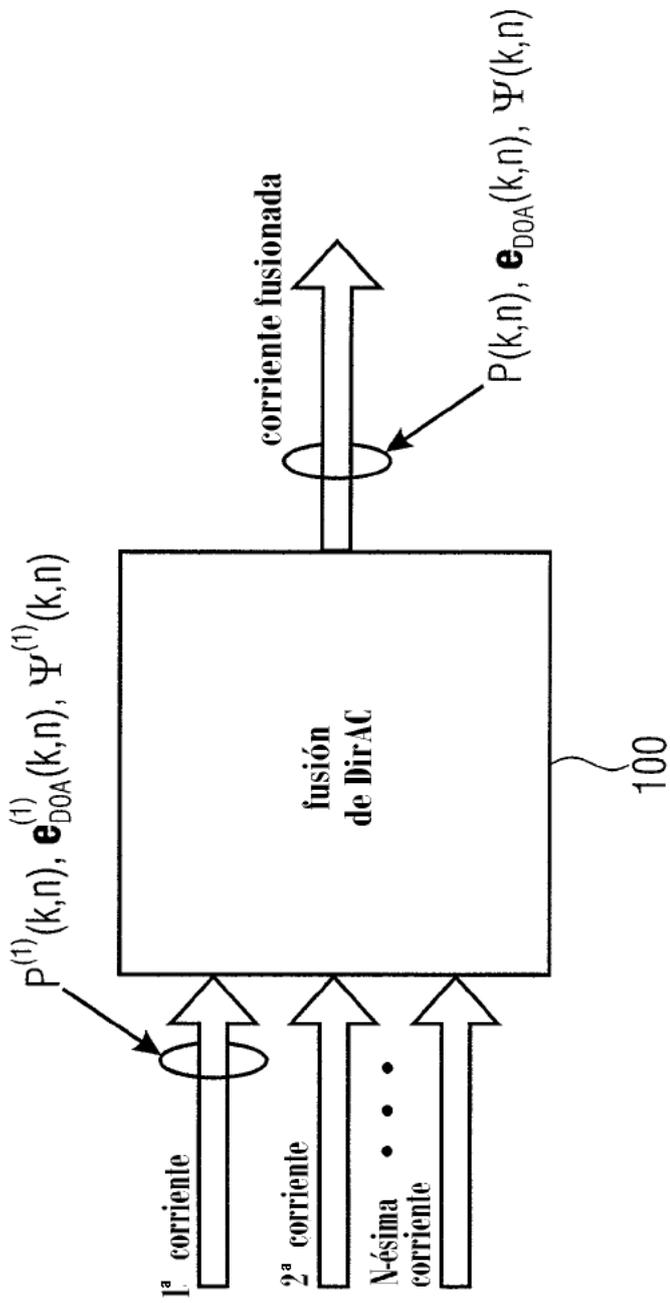


FIGURA 4

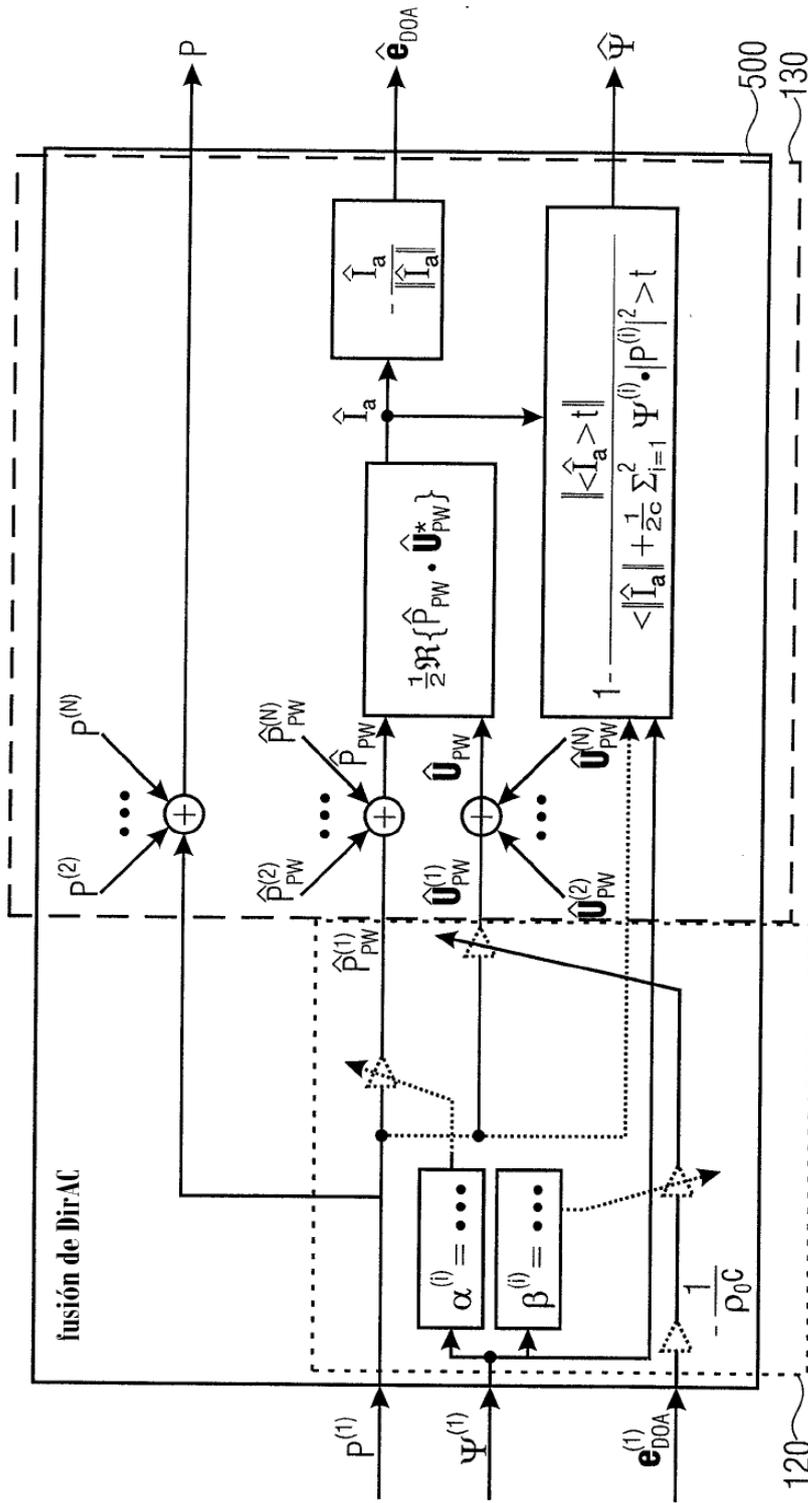


FIGURA 5

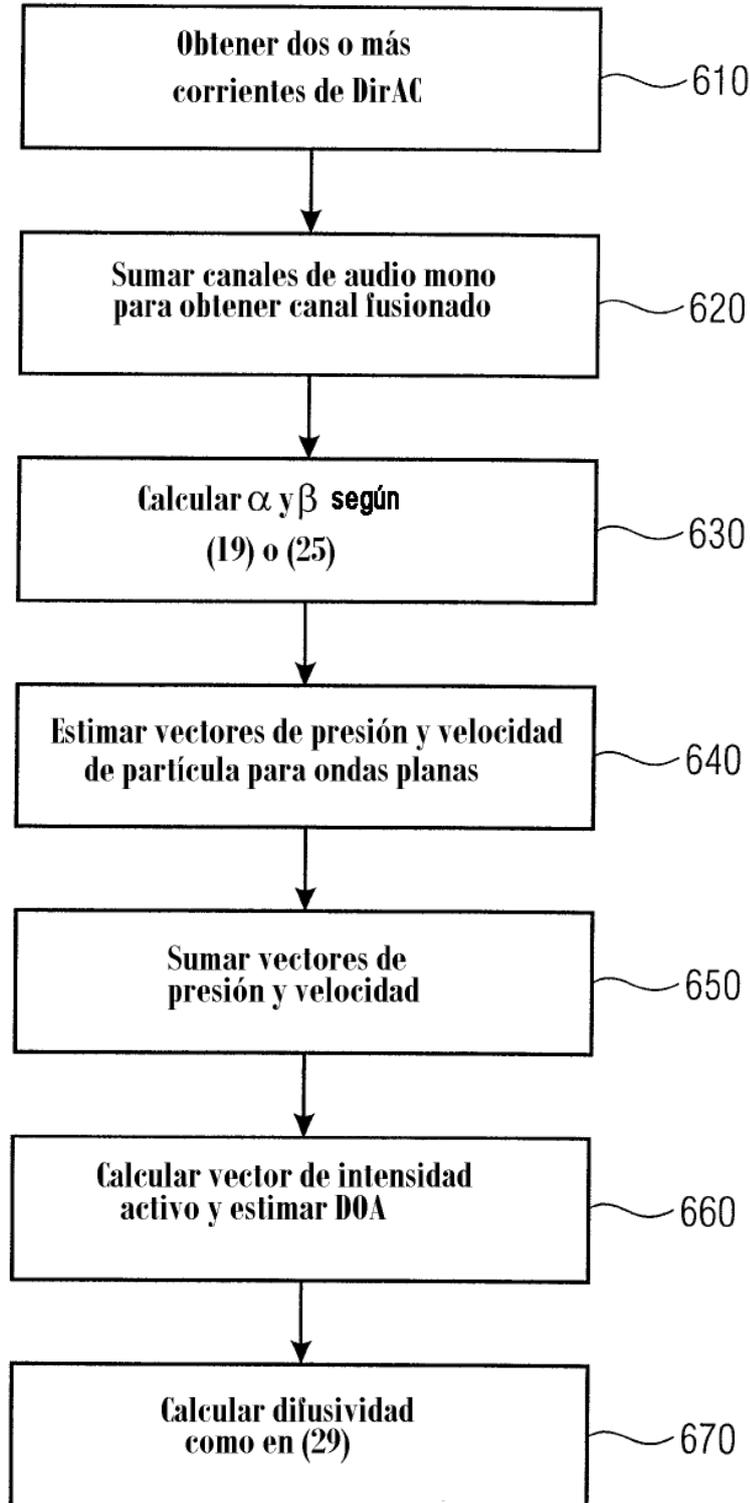


FIGURA 6