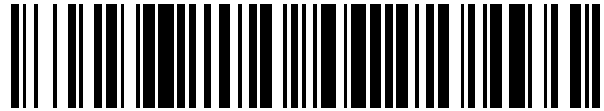


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 656 815**

51 Int. Cl.:

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **07.10.2010 E 10186808 (1)**

97 Fecha y número de publicación de la concesión europea: **22.11.2017 EP 2375410**

54 Título: **Procesador de audio espacial y procedimiento para proporcionar parámetros espaciales en base a una señal de entrada acústica**

30 Prioridad:

29.03.2010 US 318689 P

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

28.02.2018

73 Titular/es:

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

72 Inventor/es:

**THIERGART, OLIVER;
KUECH, FABIAN;
SCHULTZ-AMLING, RICHARD;
KALLINGER, MARKUS;
DEL GALDO, GIOVANNI;
KUNTZ, ACHIM;
MAHNE, DIRK;
PULKKI, VILLE y
LAITINEN, MIKKO-VILLE**

74 Agente/Representante:

SALVA FERRER, Joan

ES 2 656 815 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Procesador de audio espacial y procedimiento para proporcionar parámetros espaciales en base a una señal de entrada acústica

5

CAMPO DE TÉCNICO

[0001] Las realizaciones de la presente invención crean un procesador de audio espacial para proporcionar parámetros espaciales en base a una señal de entrada acústica. Realizaciones adicionales de la presente invención crean un procedimiento para proporcionar parámetros espaciales en base a una señal de entrada acústica. Las realizaciones de la presente invención pueden estar relacionadas con el campo de análisis acústica, descripción paramétrica y reproducción de sonido espacial, por ejemplo basada en grabaciones de micrófonos.

ANTECEDENTES DE LA INVENCION

15

[0002] La grabación de sonido espacial tiene como objetivo capturar un campo de sonido con múltiples micrófonos, de tal manera que en el sitio de reproducción, un oyente percibe la imagen de sonido como si estuviera presente en el sitio de la grabación. Estrategias estándar para la grabación de sonido espacial usan micrófonos estereofónicos simples o combinaciones más sofisticadas de micrófonos direccionales, por ejemplo tales como micrófonos de formato B usados en Ambisonics. Comúnmente, estos procedimientos son denominados como técnicas de micrófono coincidentes.

[0003] Alternativamente, se pueden aplicar procedimientos basados en una representación paramétrica de campos de sonido, que son denominados como procesadores de audio espacial paramétricos. Recientemente, se han propuesto varias técnicas para el análisis, descripción paramétrica y reproducción de audio espacial. Cada sistema tiene ventajas y desventajas únicas con respecto al tipo de la descripción paramétrica, el tipo de las señales requeridas, la dependencia e independencia de un montaje de altavoz específico, etc.

[0004] Un ejemplo para una descripción paramétrica eficiente de sonido espacial es dado por la Codificación de Audio Direccional (DirAC) (V. Pulkki: Spatial Sound Reproduction with Directional Audio Coding, Journal of the AES, Vol. 55, No. 6, 2007). DirAC representa una estrategia al análisis acústico y descripción paramétrica del sonido espacial (análisis de DirAC), también como su reproducción (síntesis de DirAC). El análisis de DirAC toma múltiples señales de micrófono como entrada. La descripción de sonido espacial es provista para un número de sub-bandas de frecuencia en términos de una o varias señales de audio de mezcla descendente e información lateral paramétrica que contiene dirección del sonido y difusividad. El último parámetro describe cómo de difuso está el campo de sonido grabado. Además, la difusividad puede ser usada como una medida de fiabilidad para el valor estimativo de la dirección. Otra aplicación consiste en el procesamiento dependiente de la dirección de la señal de audio espacial (M. Kallinger et al.: A Spatial Filtering Approach for Directional Audio Coding, 126th AES Convention, Munich, May 2009). En base a la representación paramétrica, el audio espacial puede ser reproducido con montajes de altavoces arbitrarios. Además, el análisis de DirAC puede ser considerado como un extremo frontal acústico para el sistema de codificación paramétrica que son capaces de codificar, transmitir y reproducir audio espacial de multicanal, por ejemplo surround de MPEG.

[0005] Otra estrategia para el análisis de campo de sonido espacial es representada por el llamado Micrófono de Audio Espacial (SAM) (C. Faller: Microphone Front-Ends for Spatial Audio Coders, in Proceedings of the AES 125th International Convention, San Francisco, Oct. 2008). SAM toma las señales de micrófonos direccionales coincidentes como entrada. Similar a DirAC, SAM determina la DOA (DOA-dirección de llegada) del sonido para una descripción paramétrica del campo de sonido, junto con un valor estimativo de los componentes de sonido difusos.

[0006] Técnicas paramétricas para la grabación y análisis de audio espacial, tales como DirAC y SAM, dependen de valores estimativos de parámetros de campo de sonido específicos. El rendimiento de estas estrategias son así fuertemente dependientes del rendimiento de estimación de los parámetros de pista espacial tales como la dirección de llegada del sonido o la difusividad del campo de sonido.

[0007] En general, cuando se estiman parámetros de pista espacial, se pueden hacer suposiciones específicas en cuanto a las señales de entrada acústicas (por ejemplo, en lo estacionario o en la tonalidad) con el fin de emplear el mejor algoritmo (es decir, el más eficiente o más exacto) para el procesamiento de audio. Tradicionalmente, un solo modelo de señal no variante en el tiempo puede ser definido para este propósito. Sin embargo, un problema que surge comúnmente es que diferentes señales de audio pueden exhibir una varianza

temporal significativa de tal manera que un modelo invariante en el tiempo general que describe la entrada de audio es frecuentemente inapropiado. En particular, cuando se considera un modelo de señal invariante en el tiempo individual para el procesamiento de audio, se pueden presentar desajustes de modelo que degradan el rendimiento del algoritmo aplicado.

5

[0008] Kallinger et al. describe en "A Spatial Filtering Approach for Directional Audio Coding" (una estrategia de filtrado espacial para la codificación de audio direccional) una estructura de filtrado espacial que funciona en la representación de señal paramétrica de codificación de audio direccional. En un procedimiento de filtrado espacial presentado, los parámetros de codificación de audio direccional básico sirven como una entrada a un bloque de procesamiento de señal de filtrado espacial. El procedimiento propuesto modifica la difusividad de los parámetros de codificación de audio direccional basados en el ángulo acimut ϕ y la difusividad ψ .

10

[0009] Es un objetivo de las realizaciones de la presente invención proporcionar parámetros espaciales para una señal de entrada acústica con desajustes de modelo más bajas provocadas por una varianza temporal o un no estacionario temporal de la señal de entrada acústica.

15

BREVE DESCRIPCIÓN DE LA INVENCION

[0010] Este objetivo es resuelto por un procesador de audio espacial según la reivindicación 1, un procedimiento para proporcionar parámetros espaciales en base a una señal de entrada acústica según la reivindicación 12 y un programa informático según la reivindicación 13.

20

[0011] Realizaciones de la presente invención crean un procesador de audio espacial para proporcionar parámetros espaciales en base a una señal de entrada acústica. El procesador de audio espacial comprende un determinador de características de señal y un estimador de parámetros controlables. El determinador de características de señal está configurado para determinar una característica de la señal de la señal de entrada acústica. El estimador de parámetros controlable está configurado para calcular los parámetros espaciales para la señal de entrada acústica según una regla de cálculo de parámetros espaciales variable. El estimador de parámetros está configurado además para modificar la regla de cálculo del parámetro espacial variable según la característica de señal determinada.

25

30

[0012] Es una idea de las realizaciones de la presente invención que un procesador de audio espacial para proporcionar parámetros espaciales basados en una señal de entrada acústica, que reduce desajustes de modelo provocados por una varianza temporal de la señal de entrada acústica, pueden ser creados cuando una regla de cálculo para calcular el parámetro espacial es modificada en base a una característica de señal de la señal de entrada acústica. Se ha encontrado que los desajustes de modelo pueden ser reducidos cuando una característica de señal de la señal de entrada acústica es determinada y en base a esta característica de señal determinada, los parámetros espaciales para la señal de entrada acústica son calculados.

35

[0013] En otras palabras, las realizaciones de la presente invención pueden gestionar el problema de desajustes de modelo provocados por una varianza temporal de la señal de entrada acústica al determinar características (características de señal) de las señales de entrada acústicas, por ejemplo en una etapa de pre-procesamiento (en el determinador de característica de señal) y, a continuación, identificar el modelo de señal (por ejemplo, una regla de cálculo de parámetros espaciales o parámetros del arreglo de cálculo de parámetros espaciales) que se ajustan mejor a la situación actual (las características de señal actuales). Esta información puede ser implementada al estimador de parámetros que pueden seleccionar a continuación la mejor estrategia de estimación de parámetros (con respecto a la varianza temporal de la señal de entrada acústica) para calcular los parámetros espaciales. Por consiguiente, es una ventaja de las realizaciones de la presente invención que una descripción de campo paramétrica (los parámetros espaciales) con un desajuste de modelo significativamente reducido pueda ser obtenida.

40

45

50

[0014] La señal de entrada acústica puede por ejemplo una señal medida con uno o más micrófono(s), por ejemplo con arreglos de micrófonos o con un micrófono de formato B. Diferentes micrófonos pueden tener diferentes directividades. Las señales de entrada acústicas pueden ser por ejemplo una presión de sonido "P" o una velocidad particular "U", por ejemplo en un dominio de tiempo o en un dominio de frecuencia (por ejemplo, en un dominio STFT-, STFT= transformada de Fourier de tiempo corto) o en otras palabras, ya sea en una representación de tiempo o en una representación de frecuencia. La señal de entrada acústica puede comprender por ejemplo componentes en tres direcciones diferentes (por ejemplo, ortogonales) (por ejemplo, un componente X, un componente Y, y un componente Z) y de un componente omnidireccional (por ejemplo, un componente W). Además,

55

las señales de entrada acústicas pueden contener únicamente componentes de las tres direcciones y ningún componente omnidireccional. Además, la señal de entrada acústica puede comprender dos componentes direccionales (por ejemplo, el componente X, el componente Y, el componente X y el componente Z o el componente Y, y el componente Z) y el componente omnidireccional o ningún componente omnidireccional.

5 Además, la señal de entrada acústica puede comprender solamente un componente direccional (por ejemplo, el componente X, un componente Y, y un componente Z) y el componente omnidireccional o ningún componente omnidireccional.

10 **[0015]** La característica de señal determinada por el determinador de características de señal de la señal de entrada acústica, por ejemplo de señales de micrófono puede ser por ejemplo: intervalos estacionarios con respecto al tiempo, frecuencia, espacio; presencia de doble habla o múltiples fuentes de sonido; presencia de tonalidad o transitorios; una proporción de señal a ruido de la señal de entrada acústica o presencia de señales semejantes a aplausos.

15 **[0016]** Las señales semejantes a aplausos son definidas en esta invención como señales que comprenden una secuencia temporal rápida de transitorios, por ejemplo diferentes direcciones.

20 **[0017]** La información reunida por el determinador de características de señal puede ser usada para controlar el estimador de parámetros controlables, por ejemplo en la codificación de audio direccional (DirAC) o micrófono de audio espacial (SAM), por ejemplo para seleccionar la estrategia del estimador o los ajustes del estimador (o en otras palabras para modificar la regla de cálculo del parámetro espacial variable) que se ajuste mejor a la situación actual (la característica de señal actual de la señal de entrada acústica).

25 **[0018]** Las realizaciones de la presente invención pueden ser aplicadas de manera similar a ambos sistemas, micrófono de audio espacial (SAM) y codificación de audio direccional (DirAC) o a cualquier otro sistema paramétrico. En lo sucesivo, un enfoque principal caerá en el análisis de codificación de audio direccional.

30 **[0019]** Según algunas realizaciones de la presente invención, el estimador de parámetros controlable puede estar configurado para calcular los parámetros espaciales como parámetros de codificación de audio direccional que comprenden un parámetro de difusividad para un segmento del tiempo y una sub-banda de frecuencia y/o un parámetro de dirección de llegada para un segmento de tiempo y una sub-banda de frecuencia o como parámetros de micrófono de audio espacial.

35 **[0020]** En lo sucesivo, la codificación de audio de dirección y micrófono de audio espacial son considerados como extremos frontales acústicos para sistemas que operan en parámetros espaciales tales como, por ejemplo, la dirección de llegada y la difusividad del sonido. Se debe destacar que es directo aplicar el concepto de la presente invención a otros extremos frontales acústicos también. Tanto la codificación de audio direccional como el micrófono de audio espacial proporcionan parámetros específicos (espaciales) obtenidos de señales de entrada acústicas para describir el sonido tradicional. Tradicionalmente, cuando se procesa audio espacial con extremos frontales acústicos, tales como codificación de audio de dirección y micrófono de audio espacial, un solo modelo general para las señales de entrada acústico es definido de tal manera que estimadores de parámetros óptimos (o casi óptimos) puedan ser derivados. Los estimadores se desempeñan como se desee siempre y cuando se satisfagan las suposiciones fundamentales tomadas en cuenta por el modelo. Como se ha mencionado antes, si este no es el caso, surgen desajustes de modelo, que usualmente conducen a errores severos en los valores estimativos. Tales desajustes de modelo representan un problema recurrente puesto que las señales de entrada acústicas son de forma usual altamente variantes en el tiempo.

BREVE DESCRIPCIÓN DE LAS FIGURAS

50 **[0021]** Realizaciones según la presente invención serán descritas con referencia a las figuras adjuntas, en las cuales:

La Figura 1 muestra un diagrama esquemático por bloques de un procesador de audio espacial según una realización de la presente invención;

55 La Figura 2 muestra un diagrama de bloques esquemático de un codificador de audio direccional como un ejemplo de referencia;

La Figura 3 muestra un diagrama de bloques esquemático de un procesador de audio espacial según una realización adicional de la presente invención;

La Figura 4 muestra un diagrama de bloques esquemático de un procesador de audio espacial según una

realización adicional de la presente invención;

La Figura 5 muestra un diagrama de bloques esquemático de un procesador de audio espacial según una realización adicional de la presente invención;

La Figura 6 muestra un diagrama de bloques esquemático de un procesador de audio espacial según una realización adicional de la presente invención;

La Figura 7a muestra un diagrama de bloques esquemático de un estimador de parámetros que puede ser usado en un procesador de audio espacial según una realización de la presente invención;

La Figura 7b muestra un diagrama de bloques esquemático de un estimador de parámetros que puede ser usado en un procesador de audio espacial según una realización de la presente invención;

La Figura 8 muestra un diagrama de bloques esquemático de un procesador de audio espacial según una realización adicional de la presente invención;

La Figura 9 muestra un diagrama de bloques esquemático de un procesador de audio espacial según una realización adicional de la presente invención y

La Figura 10 muestra un diagrama de flujo de un procedimiento según una realización adicional de la presente invención.

DESCRIPCIÓN DETALLADA DE REALIZACIONES DE LA PRESENTE INVENCIÓN

[0022] Antes de que las realizaciones de la presente invención sean explicadas con mayor detalle utilizando las figuras adjuntas, se debe indicar que los mismos o elementos funcionalmente iguales son proporcionados con los mismos números de referencia y que se omitirá una descripción repetida de estos elementos. Las descripciones de elementos proporcionados con los mismos números de referencia son, por consiguiente, mutuamente intercambiables.

25 Procesador de audio espacial según la Figura 1

[0023] En lo sucesivo, un procesador de audio espacial 100 será descrito con referencia a la Figura 1 que muestra un diagrama de bloques esquemático de tal procesador de audio espacial. El procesador de audio espacial 100 para proporcionar parámetros espaciales 102 o valores estimativos de parámetros espaciales 102 en base a una señal de entrada acústica 104 (o en una pluralidad de señales de entrada acústicas 104) comprende un estimador de parámetros controlables 106 y un determinador de características de señal 108. El determinador de características de señal 108 está configurado para determinar una característica de señal 110 de la señal de entrada acústica 104. El estimador de parámetros controlables 106 está configurado para calcular los parámetros espaciales 102 para la señal de entrada acústica 104 según una regla de cálculo de parámetros espaciales variable. El estimador de parámetros controlables 106 está configurado además para modificar la regla de cálculo de parámetros espaciales variables según las características de señal determinadas 110.

[0024] En otras palabras, el estimador de parámetros controlables 106 es controlado dependiendo de las características de las señales de entrada acústica o la entrada de señal acústica 104.

[0025] La señal de entrada acústica 104 puede comprender, como se ha descrito anteriormente, componentes direccionales y/o componentes omnidireccionales. Una característica de señal apropiada 110, como ya se ha mencionado puede ser, por ejemplo, intervalos estacionarios con respecto al tiempo, frecuencia, espacio de la señal de entrada acústica 104, una presencia de doble habla o múltiples fuentes de sonido en la señal de entrada acústica 104, una presencia de tonalidad o transitorios al interior de la señal de entrada acústica 104, una presencia de aplausos o una proporción de señal a ruido de la señal de entrada acústica 104. Esta enumeración de características de señal apropiadas es solo un ejemplo de características de señal que el determinador de características de señal 108 puede determinar. Según realizaciones adicionales de la presente invención, el determinador de características de señal 108 puede determinar también otras características de señal (no mencionadas) de la señal de entrada acústica 104 y el estimador de parámetros controlables 106 puede modificar la regla de cálculo del parámetro espacial variable en base a estas otras características de señal de la señal de entrada acústica 104.

[0026] El estimador de parámetros controlables 106 puede estar configurado para calcular los parámetros espaciales 102 como parámetros de codificación de audio direccionales que comprenden un parámetro de difusividad $\Psi(k,n)$ para un segmento de tiempo n y una sub-banda de frecuencia k y/o un parámetro de dirección de llegada $\phi(k,n)$ para un segmento de tiempo n y una sub-banda de frecuencia k o como parámetros de micrófono de audio espacial, por ejemplo para un segmento de tiempo n y una sub-banda de frecuencia k .

[0027] El estimador de parámetros controlable 106 puede estar configurado además para calcular los parámetros espaciales 102 utilizando otro concepto de DirAC o SAM. El cálculo de los parámetros de DirAC y parámetros de SAM solamente serán comprendidos como ejemplos. El estimador de parámetros controlable puede estar configurado por ejemplo para calcular los parámetros espaciales 102 de tal manera que los parámetros espaciales comprendan una dirección del sonido, una difusividad de sonido o una medida estadística de la dirección del sonido.

[0028] La señal de entrada acústica 104 puede ser proporcionada por ejemplo en un dominio de tiempo o un dominio de frecuencia (tiempo corto), por ejemplo en el dominio de STFT.

[0029] Por ejemplo, la señal acústica 104, en la que se proporciona en el dominio de tiempo, puede comprender una pluralidad de corrientes de audio acústicas $x_1(t)$ a $x_N(t)$ comprendiendo cada una, una pluralidad de muestras de entrada acústica con el paso del tiempo. Cada una de las corrientes de entrada acústica puede ser proporcionada por ejemplo desde un micrófono diferente y pueden corresponder con una dirección de observación diferente. Por ejemplo, una primera corriente de entrada acústica $x_1(t)$ puede comprender con una primera dirección (por ejemplo, con una dirección x), una segunda corriente de entrada acústica $x_2(t)$ puede corresponder con una segunda dirección, que puede ser ortogonal a la primera dirección (por ejemplo, una dirección y), una tercera corriente de entrada acústica $x_3(t)$ puede corresponder con una tercera dirección, que puede ser ortogonal a la primera dirección y a la segunda dirección (por ejemplo, una dirección z) y una cuarta corriente de entrada acústica $x_4(t)$ puede ser un componente omnidireccional. Estas diferentes corrientes de audio acústicas pueden ser grabadas desde diferentes micrófonos, por ejemplo en una orientación ortogonal y pueden ser digitalizadas utilizando un convertidor análogo a digital.

[0030] Según realizaciones adicionales de la presente invención, la señal de entrada acústica 104 puede comprender corrientes de entrada acústicas en una representación de frecuencia, por ejemplo en un dominio de frecuencia de tiempo, tal como el dominio STFT. Por ejemplo, la señal de entrada acústica 104 puede ser proporcionada en el formato B que comprende un vector de velocidad particular $U(k,n)$ y un vector de presión de sonido $P(k,n)$ en el que k denota una sub-banda de frecuencia y n denota un espacio de tiempo. El vector de velocidad particular $U(k,n)$ es un componente direccional de la señal de entrada acústica 104, en el que la presión de sonido $P(k,n)$ representa un componente omnidireccional de la señal de entrada acústica 104.

[0031] Como se ha mencionado anteriormente, el estimador de parámetros controlables 106 puede ser configurado para proporcionar los parámetros espaciales 102 como parámetros de codificación de audio direccionales o como parámetros de micrófonos de audio espacial. En lo sucesivo, un codificador de audio direccional convencional será presentado como un ejemplo de referencia. Un diagrama de bloques esquemático de tal codificador de audio direccional convencional se muestra en la Figura 2.

Audio Direccional Convencional según la Figura 2

[0032] La Figura 2 muestra un diagrama de bloques esquemático de un codificador de audio direccional 200. El codificador de audio direccional 200 comprende un estimador de formato B 202. El estimador de formato B 202 comprende un banco de filtros. El codificador de audio direccional 200 comprende además un estimador de parámetros de codificación de audio direccional 204. El estimador de parámetros de codificación de audio direccional 204 comprende un analizador energético 206 para efectuar un análisis energético.

[0033] Además, el estimador de parámetros de codificación de audio direccional 204 comprende un estimador de dirección 208 y un estimador de difusividad 210.

[0034] Directional Audio Coding (DirAC) (V. Pulkki: Spatial Sound Reproduction with Directional Audio Coding, Journal of the AES, Vol. 55, No. 6, 2007) representa una estrategia motivada perceptualmente, eficiente al análisis y reproducción de sonido espacial. El análisis de DirAC proporciona una descripción paramétrica del campo de sonido en términos de una señal de audio de mezcla descendente e información lateral adicional, por ejemplo dirección de llegada (DOA) del sonido y difusividad del campo de sonido. DirAC toma en cuenta los elementos que son relevantes para la audición humana. Por ejemplo, supone que diferencias de tiempo interaurales (ITD) y diferencias de nivel interaurales (ILD) puede ser descrita por la DOA del sonido. Correspondientemente, se supone que la coherencia interaural (IC) puede ser representada por la difusividad del campo de sonido. De la salida del análisis de DirAC un sistema de reproducción de sonido puede generar elementos para reproducir el sonido con la impresión espacial original con el conjunto arbitrario de altavoces. Se debe destacar que la difusividad puede ser considerada también como una medida de continuidad para las DOA estimadas. Mientras más alta es la difusividad,

más baja es la fiabilidad de la DOA y viceversa. Esta información puede ser usada por muchas herramientas basadas en DirAC tales como la localización de fuente (O. Thiérgart et al.: Localization of Sound Sources in Reverberant Environment Base don Directional Audio COding Parameters, 127th AES Convention, NY, Octubre 2009). Realizaciones de la presente invención se enfocan en la parte de análisis de DirAC en lugar de sobre la reproducción de sonido.

[0035] En el análisis de DirAC, los parámetros son estimados vía un análisis energético efectuado por el analizador energético 206 del campo de sonido, en base a las señales de formato B proporcionadas por el estimador de formato B 202. Las señales de formato B consisten en una señal omnidireccional, correspondiente a la presión de sonido $P(k,n)$ y una, dos o tres señales de dipolo alineadas con la dirección x , y , y z de un sistema de coordenadas cartesiano. Las señales de dipolo corresponden a los elementos del vector de velocidad de partículas $U(k,n)$. El análisis de DirAC se representa en la Figura 2. Las señales de micrófono en el dominio de tiempo, es decir $x_1(t), x_2(t), \dots, x_N(t)$, son proporcionadas al estimador de formato B 202. Estas señales de micrófono de dominio de tiempo pueden ser denominadas como “señales de entrada acústica en el dominio de tiempo” en lo sucesivo. El estimador de formato B 202, que contiene una transformada de Fourier de tiempo corto (STFT) u otro banco de filtros (FB), calcula la señales de formato B en el dominio de frecuencia de tiempo corto, es decir, la presión de sonido $P(k,n)$ y el vector de velocidad de partículas $U(k,n)$, en donde k y n denotan el índice de frecuencia (una sub-banda de frecuencia) y el índice de bloque de tiempo (un segmento de tiempo), respectivamente. Las señales $P(k,n)$ y $U(k,n)$ pueden ser denominadas como “señales de entrada acústica en el dominio d frecuencia de tiempo corto” en lo sucesivo. Las señales de formato B pueden ser obtenidas a partir de mediciones con arreglos de micrófono como se explica en R. Schultz-Amling et al.: Planar Microphone Array Processing for the Analysis and Reproduction of Spatial Audio using Directional Audio Coding, 124th AES COvention, Amsterdam, The Netherlands, Mayo 2008 o directamente al usar por ejemplo un micrófono de formato B. En el análisis energético, el vector de intensidad de sonido activo $I_a(k,n)$ puede ser estimado de forma independiente para diferentes bandas de frecuencia utilizando

$$I_a(k,n) = \text{Re}\{P(k,n)U^*(k,n)\}, \quad (1)$$

en donde $\text{Re}(\cdot)$ produce la parte real y $U^*(k,n)$ denota el conjugado complejo del vector de velocidad de partículas $U(k,n)$.

[0036] En lo sucesivo, el vector de intensidad de sonido activo también será llamado parámetro de intensidad.

[0037] Utilizando la representación de dominio de STFT en la ecuación 1, la DOA del sonido $\varphi(k,n)$ puede ser determinada en el estimador de dirección 208 para cada k y n como la dirección opuesta del vector de intensidad de sonido activo $I_a(k,n)$. En el estimador de difusividad 210, la difusividad del campo de sonido $\Psi(k,n)$ puede ser calculada en base a fluctuaciones de la intensidad activa según:

$$\tilde{\Psi}(k,n) = 1 - \frac{|E(I_a(k,n))|}{E(|I_a(k,n)|)}, \quad (2)$$

en donde $|\cdot|$ denota la norma del vector y $E(\cdot)$ devuelve la esperanza. En la aplicación práctica, la esperanza $E(\cdot)$ puede ser aproximada mediante un promedio obtenido a lo largo de una o más dimensiones específicas, por ejemplo a lo largo del tiempo, frecuencia o espacio.

[0038] Se ha encontrado que la esperanza $E(\cdot)$ en la ecuación 2 puede ser aproximada mediante un promedio a lo largo de una dimensión específica. Para esta cuestión, el promedio puede ser llevado a cabo a lo largo del tiempo (promedio temporal), frecuencia (promedio espectral) o espacio (promedio espacial). El promedio espacial significa por ejemplo que el vector de intensidad del sonido activo $I_a(k,n)$ de la ecuación 2 es estimado con múltiples arreglos de micrófono colocados en puntos diferentes. Por ejemplo, se pueden colocar cuatro arreglos diferentes (micrófonos) en cuatro puntos diferentes en el interior de la habitación. Como resultado, se tienen entonces para

cada tiempo de punto de tiempo-frecuencia(k,n), cuatro vectores de intensidad $I_a(k,n)$ que pueden ser promediados (de la misma manera, por ejemplo el promedio espacial) para obtener una aproximación del operador de esperanza $E(\cdot)$.

- 5 **[0039]** Por ejemplo, cuando se usa un promedio temporal sobre varios n, se obtiene un valor estimativo $\Psi(k,n)$ para el parámetro de difusividad dado por

$$\Psi(k,n) = 1 - \frac{|\langle I_a(k,n) \rangle_n|}{\langle |I_a(k,n)| \rangle_n}. \quad (3)$$

- 10 **[0040]** Existen procedimientos comunes para realizar un promedio temporal según se requiera en (3). Un procedimiento es promedio por bloques (promedio de intervalos) sobre un número específico N de instancia de tiempo n, dados por

$$\langle y(k,n) \rangle_n = \frac{1}{N} \sum_{m=0}^{N-1} y(k,n-m), \quad (4)$$

- 15 donde $y(k,n)$ es la cantidad que se va a promediar, por ejemplo $I_a(k,n)$ o $|I_a(k,n)|$. Un segundo procedimiento para calcular promedios temporales, que es usado comúnmente en DirAC debido a su eficacia es aplicar filtros de respuesta de impulso infinita (IIR). Por ejemplo, cuando se usa un filtro de paso bajo de orden con coeficiente de filtro $\alpha \in [0,1]$ un promedio temporal de una cierta señal $y(k,n)$ a lo largo de n puede ser obtenido con

20

$$\langle y(k,n) \rangle_n = \bar{y}(k,n) = \alpha \cdot y(k,n) + (1-\alpha) \cdot \bar{y}(k,n-1) \quad (5)$$

- 25 donde $\bar{y}(k,n)$ denota el resultado de promedio real y $\bar{y}(k,n-1)$ es el resultado de promedio pasado, es decir, el resultado de promedio para la instancia de tiempo (n-1). Un promedio temporal más largo es obtenido para α más pequeña, mientras que una α de mayor tamaño produce resultados más instantáneos en donde el resultado pasado $\bar{y}(k,n-1)$ cuenta menos. Un valor típico para α usado en DirAC es $\alpha=0.1$.

- [0041]** Se ha encontrado que además de usar el promedio temporal, el operador de expectativa en la ecuación 2 puede ser aproximado también mediante promedio espectral a lo largo de varias o todas las sub-bandas de frecuencia k. Este procedimiento es aplicable solo si ningún valor estimativo de difusividad independiente es necesario para la diferentes sub-bandas de frecuencia en el último procesamiento, por ejemplo cuando únicamente una sola fuente de sonido está presente. Por consiguiente, la manera más apropiada usualmente para calcular la difusividad en la práctica puede ser emplear promedio temporal.

- 35 **[0042]** En general, cuando se aproxima un operador de expectativa como en la ecuación 2 mediante un proceso de promedio, se supone estacionario de la señal considerada con respecto a la cantidad que se va a promediar. Cuanto más largo sea el promedio, es decir, más muestras son tomadas en cuenta, más precisos son los resultados de forma habitual.

- 40 **[0043]** En lo sucesivo, el análisis de micrófono de audio espacial (SAM) también se explicará en breve.

Análisis de Micrófono de Audio Espacial (SAM)

[0044] Similar a DirAC el análisis de SAM (C. Faller: Microphone Fronts-Ends for Spatial Audio Coders, in Proceedings of the AES 125th International Convention, San Francisco, Oct. 2008) proporciona una descripción paramétrica del sonido espacial. La representación de campo de sonido está basada en una señal de audio de mezcla descendente e información lateral paramétrica, es decir la DOA del sonido y valores estimativos de los niveles de los componentes de sonido directos y difusos. La entrada al análisis de SAM son las señales de medidas con múltiples micrófonos direccionales coincidentes, por ejemplo dos sensores cardiodes colocados en el mismo punto. La base para el análisis de SAM son las densidades espectrales de potencia (PSD) y las densidades espectrales cruzadas (CSD) de las señales de entrada.

[0045] Por ejemplo, $X_1(k,n)$ y $X_2(k,n)$ deberían ser las señales en el dominio de tiempo-frecuencia medidas por dos micrófonos direccionales coincidentes. Las PSD de ambas señales de entrada pueden ser determinadas con

$$\begin{aligned} \text{PSD}_1(k,n) &= E\{X_1(k,n)X_1^*(k,n)\} \\ \text{PSD}_2(k,n) &= E\{X_2(k,n)X_2^*(k,n)\}. \end{aligned} \quad (5a)$$

[0046] La CSD entre ambas entradas es dada por la correlación

$$\text{CSD}(k,n) = E\{X_1(k,n)X_2^*(k,n)\}. \quad (5b)$$

[0047] La SAM supone que las señales de entrada medidas $X_1(k,n)$ y $X_2(k,n)$ representan una superposición del sonido directo y sonido difuso, mientras que el sonido directo y el sonido difuso no están correlacionados. En base a esta suposición, se muestra en C. Faller: Microphone Fronts-Ends for Spatial Audio Coders, in Proceedings of the AES 125th International Convention, San Francisco, Oct. 2008, que es posible derivar de las ecuaciones 5a y 5b para cada sensor la PSD del sonido directo medido y el sonido difuso medido. A partir de la proporción entre las PSD del sonido directo es entonces posible determinar la DOA $\varphi(k,n)$ del sonido con un conocimiento a priori de la respuestas direccionales de los micrófonos.

[0048] Se ha encontrado que en una aplicación práctica, las expectativas $E\{\cdot\}$ en la ecuación 5a y 5b pueden ser aproximadas mediante operaciones de promedio temporal y/o espectral. Es decir similar al cálculo de difusividad en DirAC descrito en la sección anterior. Similarmente, el promedio puede ser llevado a cabo utilizando por ejemplo la ecuación 4 ó 5. Para dar un ejemplo, la estimación de la CSD puede ser efectuada en base al promedio temporal recursivo según

$$\text{CDS}(k,n) \approx \alpha \cdot X_1(k,n)X_2^*(k,n) + (1 - \alpha) \cdot \text{CDS}(k,n - 1). \quad (5c)$$

[0049] Como se ha tratado en la sección anterior, cuando se aproxima un operador de expectativa como en las ecuaciones 5a y 5b mediante un proceso de promedio, lo estacionario de la señal considerada con respecto a la cantidad que se va a promediar, puede tener que ser supuesta.

[0050] En lo sucesivo, se explicará una realización de la presente invención que efectúa una estimación de parámetros variante en el tiempo dependiendo de un intervalo estacionario.

Procesador de Audio Espacial según la Figura 3

[0051] La Figura 3 muestra un procesador de audio espacial 300 según una realización de la presente invención. En una funcionalidad del procesador de audio espacial 300 puede ser similar a una funcionalidad del procesador de audio de audio espacial 100 según la Figura 1. El procesador de audio espacial 300 puede

comprender los elementos adicionales mostrados en la Figura 3. El procesador de audio espacial 300 comprende un estimador de parámetros controlable 306, una funcionalidad del cual puede ser similar a una funcionalidad de estimador de parámetros controlables 106 según la Figura 1 y que puede comprender los elementos adicionales descritos en lo sucesivo. El procesador de audio espacial 300 comprende además un determinador de características de señal 308, una funcionalidad en el cual puede ser similar a una funcionalidad del determinador de características de señal 108 según la Figura 1 y que puede comprender los elementos adicionales descritos en lo sucesivo.

[0052] El determinador de características de señal 308 puede estar configurado para determinar un intervalo estacionario de la señal de entrada acústica 104 que constituye la característica de señal determinada 110, por ejemplo utilizando un determinador de intervalo estacionario 310. El estimador de parámetro 306 puede estar configurado para modificar la regla de cálculo del parámetro variable según las características de señal determinada 110, es decir el intervalo estacionario determinado. El estimador de parámetros 306 puede estar configurado para modificar la regla de cálculo del parámetro variable, de tal manera que un periodo de promedio o longitud de promedio para calcular los parámetros espaciales 102 es comparativamente más largo (más alto) para un intervalo estacionario comparativamente más largo y es comparativamente más corto (más bajo) para un intervalo estacionario comparativamente más corto. La longitud de promedio puede ser por ejemplo igual al intervalo estacionario.

[0053] En otras palabras, el procesador de audio espacial 300 crea un concepto para mejorar la estimación de difusividad en la codificación de audio de dirección al considerar el intervalo variable estacionario de la señal de entrada acústica 104 o las señales de entrada acústicas.

[0054] El intervalo estacionario de la señal de entrada acústica 104 puede definir por ejemplo un periodo en el tiempo en el cual no se presenta ningún movimiento (o solamente uno insignificamente pequeño) de una fuente de sonido de la entrada de la señal acústica 104. En general, lo estacionario de la señal de entrada acústica 104 puede definir un periodo de tiempo en el cual una cierta característica de la señal de la señal de entrada acústica 104 permanece constante a lo largo del tiempo. Las características de señal pueden ser por ejemplo una energía de señal, una difusividad espacial, una tonalidad, una proporción de señal a ruido y/o u otros. Al tomar en cuenta el intervalo estacionario de la señal de entrada acústica 104 para calcular los parámetros espaciales 102, una longitud de promedio para calcular los parámetros espaciales 102 puede ser modificada, de tal manera que la precisión de los parámetros espaciales 102 que representan la entrada de señal acústica 104 pueden ser mejorados. Por ejemplo, para un intervalo estacionario más largo, lo que significa que la fuente de sonido de la señal de entrada acústica 104 no se ha movido durante un intervalo más largo, un promedio temporal más largo (o tiempo) puede ser aplicado entonces para un intervalo estacionario más corto. Por consiguiente, una estimación de parámetros espaciales casi óptima (o en algunos casos incluso una óptima) puede ser efectuada (siempre) por el estimador de parámetros controlables 306 dependiendo del intervalo estacionario de la señal de entrada acústica 104.

[0055] El estimador de parámetros controlables 306 puede ser configurado por ejemplo para proporcionar un parámetro de difusividad $\Psi(k,n)$ por ejemplo en un dominio de STFT para una sub-banda de frecuencia k y un segmento de tiempo o bloque de tiempo n . El estimador de parámetros controlables 306 puede comprender un estimador de difusividad 312 para calcular el parámetro de difusividad $\Psi(k,n)$, por ejemplo en base a un promedio temporal de un parámetro de intensidad de una entrada de señal acústica 104 en un dominio STFT. Además, el estimador de parámetros controlables 306 puede comprender un analizador energético 314 para efectuar un análisis energético de la señal de entrada acústica 104 para determinar el parámetro de intensidad $I_a(k,n)$. El parámetro de intensidad $I_a(k,n)$ puede ser designado también como vector de intensidad de salida activo y puede ser calculado por el analizador energético 314 según la ecuación 1.

[0056] Por consiguiente, la señal de entrada acústica 104 puede ser proporcionada también en el dominio STFT por ejemplo en el formato B que comprende una presión de sonido $P(k,n)$ y un vector de velocidad particular $U(k,n)$ para una sub-banda de frecuencia k y un segmento de tiempo n .

[0057] El estimador de difusividad 312 puede calcular el parámetro de difusividad $\Psi(k,n)$ en base a un promedio temporal de los parámetros de intensidad $I_a(k,n)$ de la señal acústica 104, por ejemplo de la misma sub-banda de frecuencia k . El estimador de difusividad 312 puede calcular el parámetro de difusividad $\Psi(k,n)$ según la ecuación 3, en el que un numero de parámetros de intensidad y por consiguiente la longitud de promedio pueden ser variados por el estimador de difusividad 312 en dependencia del intervalo estacionario determinado.

[0058] Como un ejemplo numérico, si un intervalo estacionario comparativamente largo es determinado por el

determinador de intervalo estacionario 310, el estimador de difusividad 312 puede efectuar el promedio temporal de los parámetros de intensidad $I_a(k,n)$ con respecto a los parámetros de intensidad $I_a(k,n-10)$ a $I_a(k,n-1)$. Para un intervalo estacionario comparativamente corto determinado por el determinador de intervalo estacionario 310, el estimador de difusividad 312 puede efectuar el promedio temporal de los parámetros de intensidad $I_a(k,n)$ para los parámetros de intensidad $I_a(k,n-1)$.

[0059] Como se puede ver, la longitud de promedio del promedio temporal aplicado por el estimador de difusividad 312 se corresponde con el número de parámetros de intensidad $I_a(k,n-1)$ usado por el promedio temporal.

10

[0060] En otras palabras, la estimación de difusividad de codificación de audio direccional es mejorada al considerar el intervalo estacionario invariante en el tiempo (también llamado tiempo de coherencia) de las señales de entrada acústica con la señal de entrada acústica 104. Como se explica anteriormente, la manera común en la práctica para estimar el parámetro de difusividad $\Psi(k,n)$ es usar la ecuación 3 que comprende un promedio temporal del vector de intensidad activo $I_a(k,n)$. Se ha encontrado que la longitud de promedio óptima depende del estacionario temporal de las señales de entrada acústica o la entrada de señal de entrada acústica 104. Se ha encontrado que los resultados más exactos pueden ser obtenidos cuando la longitud de promedio es escogida para que sea igual al intervalo estacionario.

15

[0061] Tradicionalmente, como se muestra con el codificador de audio direccional convencional 200, un modelo invariante en el tiempo en general para la señal de entrada acústica es definido a partir del cual la estrategia de estimación de parámetros óptimos es definida a continuación, que en este caso significa la longitud de promedio temporal óptima. Para la estimación de difusividad, se supone comúnmente que la señal de entrada acústica posee el estacionario del tiempo dentro de un cierto intervalo de tiempo, por ejemplo 20 ms. En otras palabras, el intervalo estacionario considerado es ajustado a un valor constante que es típico para varias señales de entrada. Del intervalo estacionario supuesto, la estrategia de promedio temporal óptima es derivada a continuación, por ejemplo, el mejor valor para alfa cuando se usa un promedio de IIR como se muestra en la ecuación 5 o en la mejor N cuando se usa promedio por bloques como se muestra en la ecuación 4.

20

25

[0062] Sin embargo, se ha encontrado que diferentes señales de entrada acústicas están caracterizadas usualmente por diferentes intervalos estacionarios. De este modo, el procedimiento tradicional para suponer un modelo invariante en el tiempo para la señal de entrada acústica no se mantiene. En otras palabras, cuando la señal de entrada exhibe intervalos estacionarios que son diferentes del supuesto por el estimador, se puede ejecutar en un desajuste de modelo que puede dar como resultado valores estimativos de parámetros escasos.

30

35

[0063] Por consiguiente, la nueva estrategia propuesta (por ejemplo realizada en el procesador de audio espacial 300) adapta la estrategia de estimación de parámetros (la regla de cálculo de parámetros espaciales variables) dependiendo de la característica de señal real, tal como es visualizada en la Figura 3 para la estimación de difusividad: el intervalo estacionario de la señal de entrada acústica 104, es decir, de la señal de formato B, es determinado en una etapa de pre-procesamiento (por el determinador de características de señal 308). A partir de esta información (del intervalo estacionario determinado) la mejor longitud de promedio temporal (o en algunos casos casi la mejor), el mejor valor (o en algunos casos casi el mejor valor) para α o para N es escogido y, a continuación, el cálculo de parámetros (espacial) es llevado a cabo con el estimador de difusividad 312.

40

[0064] Se debería mencionar que además de una estimación de difusividad adaptable de señal en DirAC, es posible mejorar la estimación de dirección en SAM de una manera muy similar. En efecto, el cálculo de la PSD y la CSD de las señales de entrada acústicas en las ecuaciones 5a y 5b también requieren aproximar los operadores de expectativa mediante un proceso de promedio temporal (por ejemplo, al usar las ecuaciones 4 ó 5). Como se explica anteriormente, los resultados más exactos pueden ser obtenidos cuando la longitud de promedio corresponde al intervalo estacionario de las señales de entrada acústicas. Esto significa que el análisis de SAM puede ser mejorado al determinar primero el intervalo estacionario de las señales de entrada acústicas y escoger a continuación de esta información la mejor longitud de promedio. El intervalo estacionario de las señales de entrada acústicas y el filtro de promedio óptimo correspondiente pueden ser determinados como se explica en lo sucesivo.

45

50

[0065] En lo sucesivo, se presentará un procedimiento ejemplar que determina el intervalo estacionario de la señal de entrada acústica 104. A partir de esta información, se escoge entonces la longitud de promedio temporal óptima para el cálculo de difusividad mostrado en la Figura 3.

55

Determinación del intervalo estacionario

[0066] En lo sucesivo, se describe una manera posible para determinar el intervalo estacionario de una señal de entrada acústica (por ejemplo, la entrada de señal acústica 104) así como el coeficiente de filtro IIR óptimo α (por ejemplo usado en la ecuación 5) que produce un promedio temporal correspondiente. La determinación del intervalo estacionario descrita en lo sucesivo puede ser efectuada por el determinador de intervalo estacionario 310 del determinador de características de señal 308. El procedimiento presentado permite usar la ecuación 3 para estimar exactamente la difusividad (parámetro) $\Psi(k,n)$ dependiendo del intervalo estacionario de la señal de entrada acústica 104. La presión de sonido del dominio de frecuencia $P(k,n)$ que es parte de la señal de formato B, puede ser considerada como la señal de entrada acústica 104. En otras palabras, la señal de entrada acústica 104 puede comprender al menos un componente correspondiente a la presión de sonido $P(k,n)$.

[0067] Las señales de entrada acústica exhiben en general un intervalo estacionario corto si la energía de señal varía fuertemente dentro de un intervalo de tiempo corto y viceversa. Ejemplos típicos para los cuales el intervalo estacionario es corto son transitorios, inicios de habla y "desplazamientos", es decir cuando un usuario para de hablar. El último caso está caracterizado por energía de señal decreciente fuertemente (ganancia negativa) dentro de un tiempo corto, mientras que en los dos primeros casos, la energía se incrementa fuertemente (ganancia positiva).

[0068] El algoritmo deseado, que tiene como objetivo encontrar el coeficiente de filtro óptimo α , tiene que proporcionar valores cercanos a $\alpha = 1$ (correspondiente a un promedio temporal corto) para señales no estacionarias altas y valores cerca de $\alpha = \alpha'$ en el caso de estacionario. El símbolo α' denota un coeficiente de filtro independiente de señal apropiado para promediar señales estacionarias. Expresado en términos matemáticos, un algoritmo es dado por

$$\alpha^+(k,n) = \frac{\alpha' \cdot W(k,n)}{\alpha' \cdot W(k,n) + (1 - \alpha') \cdot \bar{W}(k,n)}, \quad (7)$$

25

en donde $\alpha^+(k,n)$ es el coeficiente de filtro óptimo para cada bandeja de tiempo-frecuencia, $W(k,n) = |P(k,n)|^2$ es el valor absoluto de la energía de señal instantánea de $P(k,n)$ y $\bar{W}(k,n)$ es un promedio temporal de $W(k,n)$. Para señales estacionarias, la energía instantánea $W(k,n)$ es igual al promedio temporal $\bar{W}(k,n)$ que produce $\alpha^+ = \alpha'$ según se desee. En el caso de señales altamente no estacionarias debido a ganancias de energía positivas, el denominador de la ecuación 7 se convierte en $\alpha' \cdot W(k,n)$, ya que $W(k,n)$ es mayor en comparación con $\bar{W}(k,n)$. Así, se obtiene $\alpha^+ \approx 1$ según se desee. En caso de no estacionario debido a ganancias de energía negativas, se obtiene el resultado indeseable $\alpha^+ \approx 0$, puesto que $\bar{W}(k,n)$ se vuelve mayor en comparación con $W(k,n)$. Por consiguiente, un candidato alternativo para el coeficiente de filtro óptimo α , es decir

35

$$\alpha^-(k,n) = \frac{\alpha' \cdot \bar{W}(k,n)}{(1 - \alpha') \cdot W(k,n) + \alpha' \cdot \bar{W}(k,n)}, \quad (8)$$

es introducido que es similar a la ecuación 7, pero exhibe el comportamiento inverso en el caso de no estacionario. Esto significa que en el caso de no estacionario debido a ganancia de energía positiva, se tiene $\alpha^- \approx 0$, mientras que para ganancias de energía negativa se obtiene $\alpha^- \approx 1$. De aquí, tomando el máximo de la ecuación 7 y ecuación 8 es decir

40

$$\alpha = \max(\alpha^+, \alpha^-), \quad (9)$$

45 produce el valor óptimo deseado para el coeficiente de promedio recursivo α , conduciendo a un promedio temporal que corresponde al intervalo estacionario de las señales de entrada acústicas.

[0069] En otras palabras, el determinador de características de señal 308 está configurado para determinar el parámetro de ponderación α en base a una proporción entre la energía de señal actual (instantánea) de al menos un componente (omnidireccional) (por ejemplo, la presión de sonido $P(k,n)$) de la señal de entrada acústica 104 y un promedio temporal sobre un segmento de tiempo dado (anterior) de energía de señal de al menos un componente (omnidireccional) de la señal de entrada acústica 104. El segmento de tiempo dado puede corresponder por ejemplo a un número dado de coeficiente de energía de señal para diferentes segmentos de tiempo (anterior).

[0070] En el caso de un análisis de SAM, la señal de energía $W(k,n)$ puede estar compuesta de las de energías de las dos señales de micrófono $X_1(k,n)$ y $X_2(k,n)$, por ejemplo $W(k,n) = |X_1(k,n)|^2 + |X_2(k,n)|^2$. El coeficiente α para la estimación recursiva de las correlaciones en la ecuación 5a o ecuación 5b, según la ecuación 5c, puede ser escogido apropiadamente utilizando el criterio de la ecuación 9 descrita anteriormente.

[0071] Como se puede ver a partir de lo anterior, el estimador de parámetros controlable 306 puede ser configurado para aplicar el promedio temporal de los parámetros de $I_a(k,n)$ de la señal acústica 104 utilizando un filtro de paso de bajos (por ejemplo, el filtro de respuesta de impulso infinita) (IIR) mencionado o un filtro de respuesta de impulso finita (FIR). Además, el estimador de parámetros controlable 306 puede ser configurado para ajustar la ponderación entre un parámetro de intensidad actual de la señal de audio acústica 104 y parámetros de intensidad previos de la señal de entrada acústica 104 en base al parámetro de ponderación alfa. En un caso especial del filtro de IIR de primer orden como se muestra con la ecuación 5, una ponderación entre el parámetro de intensidad actual y el parámetro de intensidad anterior puede ser ajustada. Cuanto más alto es el factor de ponderación alfa, más corta es la longitud de promedio temporal y, por consiguiente, más alto es el peso del parámetro de intensidad actual en comparación con el peso de los parámetros de intensidad previos. En otras palabras, la longitud de promedio temporal está basada en el parámetro de ponderación α .

[0072] El estimador de parámetros controlable 306 puede estar configurado por ejemplo de tal manera que el peso del parámetro de intensidad actual en comparación con el peso de los parámetros de intensidad anteriores sea comparativamente más alto para un intervalo estacionario comparativamente más corto y de tal manera que el peso del parámetro de intensidad actual en comparación con el peso de los parámetros de intensidad anteriores sea comparativamente más bajo para un intervalo estacionario comparativamente más largo. Por consiguiente, la longitud de promedio temporal es comparativamente más corta para un intervalo estacionario comparativamente más corto y es comparativamente más larga para un intervalo estacionario comparativamente más largo.

[0073] Según realizaciones adicionales de la presente invención, un estimador de parámetros controlable de un procesador de audio espacial según una realización de la presente invención puede estar configurado para seleccionar una regla de cálculo de parámetros espaciales de una pluralidad de reglas de cálculo de parámetros espaciales para calcular los parámetros espaciales en dependencia de la característica de señal determinada. Una pluralidad de reglas de cálculo de parámetros espaciales puede diferir por ejemplo en parámetros de cálculo o pueden ser incluso completamente diferentes entre sí. Como se muestra con las ecuaciones 4 y 5, un promedio temporal puede ser calculado utilizando un promedio por bloques como se muestra en la ecuación 4 o un filtro de paso de bajos como se muestra en la ecuación 5. Una primera regla de cálculo de parámetros espaciales puede corresponder por ejemplo al promedio por bloques según la ecuación 4 y una segunda regla de cálculo de parámetros puede corresponder por ejemplo al promedio que usa el filtro de paso de bajo según la ecuación 5. El estimador de parámetros controlable puede escoger las reglas de cálculo de la pluralidad de las reglas de cálculo que proporciona la estimación más precisa de los parámetros espaciales, en base a la característica de señal determinada.

[0074] Según realizaciones adicionales de la presente invención, el estimador de parámetros controlable puede ser configurado de tal manera que una primera regla de cálculo de parámetros espaciales de la pluralidad de reglas de cálculo de parámetros espaciales es diferente de una segunda regla de cálculo de parámetros espaciales de la pluralidad de reglas de cálculo de parámetros espaciales. La primera regla de cálculo de parámetros espaciales y la segunda regla de cálculo de parámetros espaciales pueden ser seleccionadas de un grupo que está constituido por:

promedio en el tiempo sobre una pluralidad de segmentos de tiempo en una sub-banda de frecuencia (por ejemplo, como se muestra en la ecuación 3), promedio de frecuencia sobre una pluralidad de sub-bandas de frecuencia en un segmento de tiempo, promedio de tiempo de frecuencia, promedio espacial y sin promedio.

[0075] En lo sucesivo, este concepto de escoger una regla de cálculo de parámetros espaciales de una pluralidad de reglas de cálculo de parámetros espaciales por un estimador de parámetros controlable será descrito

utilizando dos realizaciones ejemplares de la presente invención mostradas en las Figura 4 y 5.

Estimación de la dirección de llegada y difusividad variantes en el tiempo dependiendo de doble habla utilizando un codificador espacial según la Figura 4

5

[0076] La Figura 4 muestra un diagrama de bloques esquemático de un procesador de audio espacial 400 según una realización de la presente invención. La funcionalidad del procesador de audio espacial 400 puede ser similar a la funcionalidad del procesador de audio espacial 100 según la Figura 1. El procesador de audio espacial 400 puede comprender los elementos adicionales descritos en lo sucesivo. El procesador de audio espacial 400 comprende un estimador de parámetros controlable 406, la funcionalidad del cual puede ser similar a la funcionalidad del estimador de parámetros controlables 106 de la Figura 1 y que puede comprender los elementos adicionales descritos en lo sucesivo. El procesador de audio espacial 400 comprende además un determinador de características de señal 408, la funcionalidad del cual puede ser similar a la funcionalidad del determinador de características de señal según la Figura 1 y que puede comprender los elementos adicionales descritos en lo sucesivo.

[0077] El estimador de parámetros controlable 406 está configurado para seleccionar una regla de cálculo de parámetros espaciales de una pluralidad de reglas de cálculo de parámetros espaciales para calcular parámetros espaciales 102, en dependencia de una característica de señal determinada 110, que es determinada por el determinador de características de señal 408. En la realización ejemplar mostrada en la Figura 4, el determinador de características de señal está configurado para determinar si una señal de entrada acústica 104 comprende componentes de diferentes fuentes de sonido o solamente comprende componentes de una fuente de sonido. En base a esta determinación, el estimador de parámetros controlable 406 puede escoger una primera regla de cálculo de parámetros espaciales 410 para calcular los parámetros espaciales 102 si la señal de entrada acústica 104 solamente comprende componentes de una fuente de sonido y puede escoger una segunda regla de cálculo de parámetros espaciales 412 para calcular los parámetros espaciales 102 si la señal de entrada acústica 104 comprende componentes de una más de una fuente de sonido. La primera regla de cálculo de parámetros espaciales 410 puede comprender por ejemplo un promedio espectral o promedio de frecuencia sobre una pluralidad de sub-bandas de frecuencia y la segunda regla de cálculo de parámetros espaciales 412 puede no comprender promedio espectral o promedio de frecuencia.

[0078] La determinación de si la señal de entrada acústica 104 comprende componentes de más de una fuente de sonido o no puede ser efectuada por un detector de doble habla 414 del determinador de características de señal 408. El estimador de parámetros controlable 406 puede estar configurado por ejemplo para proporcionar un parámetro de difusividad $\Psi(k, n)$ de la señal de entrada acústica 104 en el dominio de STFT para una sub-banda de frecuencia k y un bloque de tiempo n .

[0079] En otras palabras, el procesador de audio espacial 400 muestra un concepto para mejorar la estimación de difusividad en la codificación de audio direccional al tomar en cuenta situaciones de doble habla.

40

[0080] De otra manera, en otras palabras, el determinador de características de señal 408 está configurado para determinar si la señal de entrada acústica 104 comprende componentes de diferentes fuentes de sonido al mismo tiempo. EL estimador de parámetros controlable 406 está configurado para seleccionar según el resultado de la determinación de características de señal una regla de cálculo de parámetros espaciales (por ejemplo, la primera regla de cálculo de parámetros espaciales 410 o la segunda regla de cálculo de parámetros espaciales 412) de la pluralidad de reglas de cálculo de parámetros espaciales, para calcular los parámetros espaciales 102 (por ejemplo, para calcular el parámetro de difusividad $\Psi(k, n)$). La primera regla de cálculo de parámetros espaciales 410 es escogida cuando la señal de entrada acústica 104 comprende componentes como máximo de una fuente de sonido y la segunda regla de cálculo de parámetros espaciales de la pluralidad de reglas de cálculo de parámetros espaciales es escogida cuando la señal de entrada acústica 104 comprende componentes de más de una fuente de sonido al mismo tiempo. La primera regla de cálculo de parámetros espaciales 410 incluye un promedio de frecuencia (por ejemplo, de parámetro de intensidad $I_a(k, n)$) de la señal de entrada acústica 104 sobre una pluralidad de sub-bandas de frecuencia. La segunda regla de cálculo de parámetros espaciales 412 no incluye promedio de frecuencia.

50

[0081] En el ejemplo mostrado en la Figura 4, la estimación del parámetro de difusividad $\Psi(k, n)$ y/o parámetro de dirección (de llegada) $\phi(k, n)$ en el análisis de codificación de audio direccional es mejorada al ajustar los estimadores correspondientes dependiendo de situaciones de doble habla. Se ha encontrado que el cálculo de difusividad de la ecuación 2 puede ser realizado en la práctica mediante promedio del vector de intensidad activo I_a

55

(k, n) sobre sub-bandas de frecuencia k o al combinar un promedio temporal y promedio espectral. Sin embargo, el promedio espectral no es apropiado si se requieren valores estimativos de difusividad independientes para las diferentes sub-bandas de frecuencia, como es el caso en una llamada situación de doble habla, en donde múltiples fuentes de sonido (por ejemplo, usuarios que hablan) están activos al mismo tiempo. Por consiguiente, tradicionalmente (como en el codificador de audio direccional mostrado en la Figura 2) el promedio espectral no es empleado, ya que el modelo general de las señales de entrada acústicas siempre supone situaciones de doble habla. Se ha encontrado que esta suposición de modelo no es óptima en el caso de situaciones de una sola habla, debido a que se ha encontrado que en situaciones de una sola habla, el promedio espectral puede mejorar la exactitud de estimación de parámetros.

10

[0082] La nueva estrategia propuesta, como se muestra en la Figura 4, escoge la estrategia de estimación de parámetros óptima (la regla de cálculo de parámetros espaciales óptima) al seleccionar el modelo básico para la señal de entrada acústica 104 o para las señales de entrada acústicas. En otras palabras, la Figura 4 muestra una aplicación de una realización de la presente invención para mejorar la estimación de difusividad dependiendo de situaciones de doble habla: primero se emplea el detector de doble habla 414 que determina a partir de la señal de entrada acústica 104 o las señales de entrada acústicas si está presente doble habla en la situación actual o no. Si no se decide para un estimador de parámetros (o en otras palabras, el estimador de parámetros controlable 406 escoge una regla de cálculo de parámetros espaciales) que calcula la difusividad (parámetros) $\Psi(k, n)$ al aproximar la ecuación 2 al utilizar el promedio espectral (frecuencia) y temporal del vector de intensidad activo $I_a(k, n)$, es decir:

15

20

$$\Psi(k, n) = \Psi(n) = 1 - \frac{|\langle\langle I_a(k, n) \rangle_n \rangle_k|}{\langle\langle |I_a(k, n)| \rangle_n \rangle_k} \quad (10)$$

25

[0083] De otra manera, si existe doble habla, un estimador es escogido (o en otras palabras, el estimador de parámetros controlable 406 escoge una regla de cálculo de parámetros espaciales) que usa promedio temporal solamente, como en la ecuación 3. Una idea similar puede ser aplicada a la estimación de dirección: en el caso de situaciones de una sola habla, pero solamente en este caso, la estimación de dirección $\varphi(k, n)$ puede ser mejorada mediante un promedio espectral de los resultados sobre varias o todas la sub-bandas de frecuencia k, es decir:

30

$$\bar{\varphi}(k, n) = \bar{\varphi}(n) = \langle \varphi(k, n) \rangle_k \quad (11)$$

[0084] Según algunas realizaciones de la presente invención, es también concebible aplicar el promedio (espectral) sobre partes del espectro y no todo el ancho de banda necesariamente.

35

[0085] Para efectuar el promedio temporal y espectral, el estimador de parámetros controlable 406 puede determinar el vector de intensidad activo $I_a(k, n)$, por ejemplo en el dominio de STFR para cada sub-banda k y cada segmento de tiempo n, por ejemplo utilizando un análisis energético, por ejemplo al emplear un analizador energético 416 del estimador de parámetros controlable 406.

40

[0086] En otras palabras, el estimador de parámetros 406 puede estar configurado para determinar un parámetro de difusividad actual $\Psi(k, n)$ para una sub-banda de frecuencia k y un segmento de tiempo actual n de la señal de entrada acústica 104 en base al promedio espectral y temporal de los parámetros de intensidad activos determinados $I_a(k, n)$ de la señal de entrada acústica 104 incluida en la primera regla de cálculo de parámetros espaciales 410 o basada solamente en el promedio temporal de los vectores de intensidad activos determinados $I_a(k, n)$ en dependencia de la característica de señal determinada.

45

[0087] En lo sucesivo, se describirá otra realización ejemplar de la presente invención que está basada también en el concepto de escoger una regla de cálculo de parámetros espaciales de ajuste para mejorar el cálculo de los parámetros espaciales de la señal de entrada acústica utilizando un procesador de audio espacial 500 mostrado en la Figura 5, basado en una tonalidad de la señal de entrada acústica.

50

Estimación de parámetros dependiente de la tonalidad utilizando un procesador de audio espacial según la Figura 5

[0088] La Figura 5 muestra un diagrama de bloques esquemático de un procesador de audio espacial 500 según una realización de la presente invención. Una funcionalidad del procesador de audio espacial 500 puede ser similar a la funcionalidad del procesador de audio espacial 100 según la Figura 1. El procesador de audio espacial 500 puede comprender además los elementos adicionales descritos en lo sucesivo. El procesador de audio espacial 500 comprende un estimador de parámetros controlables 506 y un determinador de características de señal 508. Una funcionalidad del estimador de parámetros controlable 506 puede ser similar a la funcionalidad del estimador de parámetros controlable 106 según la Figura 1, el estimador de parámetros controlable 506 puede comprender los elementos adicionales descritos en lo sucesivo. Una funcionalidad del determinador de características de señal 508 puede ser similar a la funcionalidad del determinador de características de señal 108 según la Figura 1. El determinador de características de señal 508 puede comprender los elementos adicionales descritos en lo sucesivo.

[0089] El procesador de audio espacial 500 difiere del procesador de audio espacial 400 en el hecho de que el cálculo de los parámetros espaciales 102 es modificado en base a la tonalidad determinada de la señal de entrada acústica 104. El determinador de características de señal 508 puede determinar la tonalidad de señal de entrada acústica 104 y el estimador de parámetros controlable 506 puede escoger, en base a la tonalidad determinada de la señal de entrada acústica 104, una regla de cálculo de parámetros espaciales de una pluralidad de reglas de cálculo de parámetros espaciales para calcular los parámetros espaciales 102.

[0090] En otras palabras, el procesador de audio espacial 500 muestra un concepto para mejorar la estimación en parámetros de codificación de audio direccional al considerar la tonalidad de la señal de entrada acústica 104 o de las señales de entrada acústicas.

[0091] El determinador de características de señal 508 puede determinar la tonalidad de la señal de entrada acústica utilizando una estimación de tonalidad, por ejemplo, utilizando el estimador de tonalidad 510 del determinador de características de señal 508. El determinador de características de señal 508 puede proporcionar por consiguiente la tonalidad de la señal de entrada acústica 104 o una información correspondiente a la tonalidad de la señal de entrada acústica 104 como la característica de señal determinada 110 de la señal de entrada acústica 104.

[0092] El estimador de parámetros controlable 506 puede estar configurado para seleccionar, según un resultado de la determinación de características de señal (de la estimación de tonalidad) una regla de cálculo de parámetros espaciales de la pluralidad de reglas de cálculo de parámetros espaciales para calcular los parámetros espaciales 102, de tal manera que una primera regla de cálculo de parámetros espaciales de la pluralidad de reglas de cálculo de parámetros espaciales es escogida cuando la tonalidad de la señal de entrada acústica 104 está debajo de un nivel de umbral de tonalidad dado y de tal manera que una segunda regla de cálculo de parámetros espaciales de la pluralidad de reglas de cálculo de parámetros espaciales es escogida cuando la tonalidad de la señal de entrada acústica 104 está por encima de un nivel de umbral de tonalidad dado. Similar al estimador de parámetros controlables 406 según la Figura 4, la primera regla de cálculo de parámetros espaciales puede incluir un promedio de frecuencia y la segunda regla de cálculo de parámetros espaciales puede no incluir promedio de frecuencia.

[0093] En general, la tonalidad de una señal acústica proporciona información de si la señal tiene o no un espectro de banda ancha. Una alta tonalidad indica que el espectro de señal contiene solo unas pocas frecuencias con alta energía. En contraste, la baja tonalidad indica señales de banda amplia, es decir, señales en donde energía similar está presente en un intervalo de frecuencia grande.

[0094] Esta información en la tonalidad de una señal de entrada acústica (de la tonalidad de la señal de entrada acústica 104) puede ser aprovechada para mejorar, por ejemplo la estimación de parámetros de codificación de audio direccional. Con referencia al diagrama de bloques esquemático mostrado en la Figura 5, de la señal de entrada acústica 104 o las señales de entrada acústicas, primero la tonalidad es determinada (por ejemplo, como se explica en S. Molla y B. Torresani: Determining Local Transientness of Audio Signals, IEEE Signal Processing Letters, Vol. 11, No. 7, Julio 2007) de la entrada utilizando el detector de tonalidad o estimador de tonalidad 510. La información en cuanto a la tonalidad (la característica de señal determinada 110) controla la estimación de los parámetros de codificación de audio direccionales (de los parámetros de audio espaciales 102). Una salida del estimador de parámetros controlables 506 son los parámetros espaciales 102 con exactitud incrementada en comparación con el procedimiento tradicional con el codificador de audio direccional según la Figura 2.

[0095] La estimación de la difusividad $\Psi(k, n)$ puede servirse del conocimiento de la tonalidad de señal de entrada del siguiente modo: el cálculo de la difusividad $\Psi(k, n)$ requiere un proceso de promedio como se muestra

en la ecuación 3. Este promedio es llevado a cabo tradicionalmente solo a lo largo del tiempo n. Particularmente en campos de sonido difuso, una estimación exacta de la difusividad es solamente posible cuando el promedio es suficientemente largo. Un promedio temporal largo sin embargo no es posible usualmente debido al intervalo estacionario corto de las señales de entrada acústicas. Para mejorar la estimación de difusividad, se puede
5 combinar el promedio temporal con un promedio espectral sobre las bandas de frecuencia k, es decir,

$$\Psi(k, n) = 1 - \frac{|\langle\langle I_a(k, n) \rangle_n \rangle_k|}{\langle\langle |I_a(k, n)| \rangle_n \rangle_k}, \quad (12)$$

[0096] Sin embargo, este procedimiento puede requerir señales de banda amplia en donde la difusividad es similar para diferentes bandas de frecuencia. En el caso de señales tonales, en donde solo pocas frecuencias poseen energía significativa, la verdadera difusividad del campo de sonido puede variar fuertemente a lo largo de las bandas de frecuencia k. Esto significa que cuando el detector de tonalidad (el estimador de tonalidad 510 del determinador de características de señal 508) indica una alta tonalidad de la señal acústica 104, entonces se evita el promedio espectral.
10

[0097] En otras palabras, el estimador de parámetros controlables 506 está configurado para derivar los parámetros espaciales 102, por ejemplo un parámetro de difusividad $\Psi(k, n)$ por ejemplo en el dominio de STFT para una sub-banda de frecuencia k y un segmento de tiempo n en base a un promedio temporal y espectral de los parámetros de intensidad $I_a(k, n)$ de la señal de entrada acústica 104 si la tonalidad determinada de la señal acústica 104 es comparativamente pequeña y proporcionar los parámetros espaciales 102, por ejemplo el parámetro de difusividad $\Psi(k, n)$ basado solamente en un promedio temporal y no promedio espectral de los parámetros de intensidad $I_a(k, n)$ de la señal de entrada acústica 104 si la tonalidad determinada de la señal de entrada acústica 104 es comparativamente alta.
15
20

[0098] La misma idea puede ser aplicada a la estimación del parámetro de dirección (de llegada) $\phi(k, n)$ para mejorar la proporción de la señal a ruido de los resultados (de los parámetros espaciales determinados 102). En otras palabras, el estimador de parámetros controlable 506 puede estar configurado para determinar el parámetro de dirección de llegada $\phi(k, n)$ en base al promedio espectral si la tonalidad determinada de la señal de entrada acústica 104 es comparativamente pequeña y para derivar el parámetro de dirección de llegada $\phi(k, n)$ sin efectuar un promedio espectral si la tonalidad ya es comparativamente alta.
25
30

[0099] Esta idea de mejorar la proporción de la señal a ruido mediante promedio espectral del parámetro de dirección de llegada $\phi(k, n)$ será descrita en lo sucesivo en más detalle utilizando otra realización de la presente invención. El promedio espectral puede ser aplicado a la señal de entrada acústica 104 o las señales de entrada acústicas a la intensidad de señal activa o directamente al parámetro de dirección (de llegada) $\phi(k, n)$.
35

[0100] Para experto en la técnica queda claro que el procesador de audio espacial 500 puede ser aplicado también al análisis de micrófono de audio espacial de manera similar con la diferencia de que ahora los operadores de expectativa en la ecuación 5a y 5b son aproximados al considerar un promedio espectral en el caso en el que ninguna doble habla está presente o en el caso de una baja tonalidad.
40

[0101] En lo sucesivo, se explicarán otras dos realizaciones de la presente invención, que efectúan una estimación de dirección dependiente de la proporción de señal a ruido para mejorar el cálculo de los parámetros espaciales.
45

Estimación de dirección dependiente de la proporción de señal a ruido utilizando un procesador de audio espacial según la Figura 6

[0102] La Figura 6 muestra un diagrama de bloques esquemático del procesador de audio espacial 600. El procesador de audio espacial 600 está configurado para efectuar la estimación de dirección dependiente de la producción de ruido mencionada anteriormente.
50

[0103] Una funcionalidad del procesador de audio espacial 600 puede ser similar a la funcionalidad del procesador de audio espacial 100 según la Figura 1. El procesador de audio espacial 600 puede comprender los elementos adicionales descritos en lo sucesivo. El procesador de audio espacial 600 comprende un estimador de parámetros controlable 606 y un determinador de características de señal 608. La funcionalidad del estimador de
55

parámetros controlable 606 puede ser similar a la funcionalidad del estimador de parámetros controlable 106 según la Figura 1, el estimador de parámetros controlable 606 puede comprender los elementos adicionales descritos en lo sucesivo. Una funcionalidad del determinador de características de señal 608 puede ser similar a la funcionalidad del determinador de características de señal 108 según la Figura 1 y el determinador de características de señal 608 puede comprender los elementos adicionales descritos en lo sucesivo.

[0104] El determinador de características de señal 608 puede estar configurado para determinar la proporción de señal a ruido (SNR) de una señal de entrada acústica 104 como una característica de señal 110 de la señal de entrada acústica 104. El estimador de parámetros controlable 606 puede estar configurado para proporcionar una regla de cálculo espacial variable para calcular parámetros espaciales 102 de la señal de entrada acústica 104 en base a la proporción de señal a ruido de la señal de entrada acústica 104.

[0105] El estimador de parámetros controlable 606 puede efectuar por ejemplo un promedio temporal para determinar los parámetros espaciales 102 y puede hacer variar la longitud de promedio del promedio temporal (o un número de elementos usados para el promedio temporal) en dependencia de la proporción de señal a ruido determinado de la señal de entrada acústica 104. Por ejemplo, el estimador de parámetros 606 puede estar configurado para hacer variar la longitud de promedio del promedio temporal, de tal manera que la longitud de promedio es comparativamente alta para una proporción de señal a ruido comparativamente baja de la señal de entrada acústica 104 y de tal manera que la longitud de promedio es comparativamente baja para una proporción de señal a ruido comparativamente alta de la señal de entrada acústica 104.

[0106] El estimador de parámetros 606 puede estar configurado para proporcionar el parámetro de dirección de llegada $\varphi(k, n)$ como parámetro espacial 602 en base al promedio temporal mencionado. Como se ha mencionado anteriormente, el parámetro de dirección de llegada $\varphi(k, n)$ puede ser determinado en el estimador de parámetros controlables 606 (por ejemplo, en un estimador de dirección 610 del estimador de parámetros 606) para cada sub-banda de frecuencia k y segmento de tiempo como la dirección opuesta del vector de intensidad de sonido activo $I_a(k, n)$. El estimador de parámetros 606 puede comprender por consiguiente un analizador energético 612 para efectuar un análisis energético en la señal de entrada acústica 104 para determinar el vector de intensidad de sonido activo $I_a(k, n)$ por cada sub-banda de frecuencia k y cada segmento de tiempo n . El estimador de dirección 610 puede efectuar el promedio temporal, por ejemplo en el vector de intensidad activo determinado para una sub-banda de frecuencia k sobre una pluralidad de segmentos de tiempo n . En otras palabras, el estimador de dirección 610 puede efectuar un promedio temporal de los parámetros de intensidad $I_a(k, n)$ para una sub-banda de frecuencia k y una pluralidad de segmentos de tiempo (anterior) para calcular el parámetro de dirección de llegada $\varphi(k, n)$ para una sub-banda de frecuencia k y un segmento de tiempo n . Según realizaciones adicionales de la presente invención, el estimador de dirección 610 puede también (por ejemplo, en lugar de un promedio temporal de los parámetros de intensidad $I_a(k, n)$) efectuar el promedio temporal en una pluralidad de parámetros de dirección de llegada determinados $\varphi(k, n)$ para una sub-banda de frecuencia k y una pluralidad de segmentos de tiempo (anteriores). La longitud de promedio del promedio temporal corresponde por consiguiente al número de parámetros de intensidad o el número de parámetros de dirección de llegada usados para efectuar el promedio temporal. En otras palabras, el estimador de parámetros 606 puede estar configurado para aplicar el promedio temporal a un subconjunto de parámetros de intensidad $I_a(k, n)$ para una pluralidad de segmentos de tiempo y una sub-banda de frecuencia k o a un subconjunto de parámetros de dirección de llegada $\varphi(k, n)$ para una pluralidad de segmentos de tiempo y una sub-banda de frecuencia k . El número de parámetros de intensidad en el subconjunto de parámetros de intensidad o el número de parámetros de dirección de llegada en el subconjunto de parámetros de dirección de llegada usados para el promedio temporal corresponde con la longitud de promedio del promedio temporal. El estimador de parámetros controlable 606 está configurado para ajustar el número de parámetros de intensidad o el número de parámetros de dirección de llegada en el subconjunto usado para calcular el promedio temporal, de tal manera que el número de parámetros de intensidad en el subconjunto de parámetros de intensidad o el número de parámetros de dirección de llegada en el subconjunto de parámetros de dirección de llegada es comparativamente bajo para una proporción de señal a ruido comparativamente alta de la señal de entrada acústica 104 y de tal manera que el número de parámetros de intensidad o el número de parámetros de dirección de llegada es comparativamente alto para una proporción de señal a ruido comparativamente baja de la señal de entrada acústica 104.

[0107] En otras palabras, la realización de la presente invención proporciona una estimación de dirección de codificación de audio direccional que está basada en la proporción de señal a ruido de las señales de entrada acústicas o de la señal de entrada acústica 104.

[0108] En general, la exactitud de la dirección estimada $\varphi(k, n)$ (o del parámetro de dirección de llegada $\varphi(k, n)$)

n)) del sonido, definido según el codificador de audio direccional 200 según la Figura 2 es influenciada por el ruido que está siempre presente dentro de las señales de entrada acústicas.

[0109] El impacto del ruido sobre la exactitud de estimación depende de la SNR, es decir, de la proporción entre la energía de señal del sonido que llega en el arreglo (micrófono) y la energía del ruido. Una SNR pequeña reduce significativamente la exactitud de estimación de la dirección $\varphi(k, n)$. La señal de ruido es usualmente introducida por el equipo de medición, es decir, los micrófonos y el amplificador del micrófono y conduce a errores en $\varphi(k, n)$. Se ha encontrado que la dirección $\varphi(k, n)$ está con igual probabilidad ya sea subestimada o sobreestimada, pero la expectativa de $\varphi(k, n)$ es todavía correcta.

[0110] Se ha encontrado que al tener varias estimaciones independientes del parámetro de dirección de llegada $\varphi(k, n)$, por ejemplo, al repetir la medición varias veces, la influencia del ruido puede ser reducida y así la exactitud de la estimación de dirección puede ser incrementada mediante promedio del parámetro de dirección de llegada $\varphi(k, n)$ sobre las varias instancias de medición. Efectivamente, el proceso de promedio incrementa la proporción de señal a ruido del estimador. Cuanto más pequeña es la proporción a ruido en los micrófonos o en general en los dispositivos de grabación o cuanto más alta es la proporción de señal a ruido objetivo deseada en el estimador, más alto es el número de instancias de medición que pueden ser requeridas en el proceso de promedio.

[0111] El codificador espacial 600 mostrado en la Figura 6 efectúa este proceso de promedio en dependencia de la proporción de señal a ruido de la señal de entrada acústica 104. De otra manera, en otras palabras, el procesador de audio espacial 600 muestra un concepto para mejorar la estimación de dirección en codificación de audio direccional al tomar en cuenta la SNR en la entrada acústica o de la señal de entrada acústica 104.

[0112] Antes de estimar la dirección $\varphi(k, n)$ con el estimador de dirección 610, la proporción de señal a ruido de la señal de entrada acústica 104 o de las señales de entrada acústicas es determinada con el estimador de proporción de señal a ruido 614 del determinador de características de señal 608. La proporción de señal a ruido puede ser estimada para cada bloque de tiempo n y una banda de frecuencia k , por ejemplo en el dominio de STFT. La información en cuanto a la proporción de señal a ruido real de la señal de entrada acústica 104 se proporciona como la característica de señal determinada 110 del estimador de proporción de señal a ruido 614 al estimador de dirección 610 que incluye un promedio temporal dependiente de la frecuencia y el tiempo de señales de codificación de audio direccional específicas para mejorar la proporción de señal a ruido. Además, una proporción de señal a ruido objetivo deseada se puede hacer pasar al estimador de dirección 610. La proporción de señal a ruido objetivo deseada puede ser definida externamente, por ejemplo por un usuario. El estimador de dirección 610 puede ajustar la longitud de promedio del promedio temporal, de tal manera que una proporción de señal a ruido obtenida de la señal de entrada acústica 104 o en la salida del estimador de parámetros controlable 606 (después del promedio) coincide con la proporción de señal a ruido deseada. De otra manera, en otras palabras, el promedio (en el estimador de dirección 610) se lleva a cabo hasta que se obtiene la proporción de señal a ruido objetivo deseada.

[0113] El estimador de dirección 610 puede comparar continuamente la proporción de señal a ruido obtenida de la señal de entrada acústica 104 con la proporción de señal a ruido objetivo y puede efectuar el promedio hasta que se obtiene la proporción de señal a ruido deseada. Utilizando este concepto, la proporción de señal a ruido obtenida, la señal de entrada acústica 104 es monitoreada continuamente y el promedio se finaliza, cuando la proporción de señal a ruido obtenida de la señal de entrada acústica 104 coincide con la proporción de señal a ruido objetivo, así, no hay necesidad de calcular la longitud de promedio por adelantado.

[0114] Además, el estimador de dirección 610 puede determinar, en base a la proporción de señal a ruido de la señal de entrada acústica 104 en la entrada del estimador de parámetros controlables 606, la longitud de promedio para el promedio de la proporción de señal a ruido de la señal de entrada acústica 104, de tal manera que la proporción de señal a ruido obtenida de la señal de entrada acústica 104 en la salida de estimador de parámetros controlables 606 coincide con la proporción de señal a ruido objetivo. Así, mediante el uso de este concepto, la proporción de señal a ruido obtenida de la señal de entrada acústica 104 no es monitoreada continuamente.

[0115] Un resultado generado por los dos conceptos para el estimador de dirección 610 descrito anteriormente es el mismo: durante la estimación de los parámetros espaciales 102, se puede obtener la precisión de los parámetros espaciales 102, como si la señal de entrada acústica 104 tuviera la proporción de señal a ruido objetivo, aunque la proporción de señal a ruido actual de la señal de entrada acústica 104 (en la entrada del estimador de parámetros controlables 606) es peor.

[0116] Cuanto más pequeña es la proporción de señal a ruido de la señal de entrada acústica 104 en

comparación con la proporción de señal a ruido objetivo, más largo es el promedio temporal. Una salida del estimador de dirección 610, por ejemplo es un valor estimativo $\varphi(k, n)$ es decir, en parámetros de dirección de llegada $\varphi(k, n)$ con exactitud incrementada. Como se menciona anteriormente, diferentes posibilidades para el promedio de la señales de codificación de audio direccionales existen: promedio del vector de intensidad del sonido activo $I_a(k, n)$ para una sub-banda de frecuencia k y una pluralidad de segmentos de tiempo proporcionados en la ecuación 1 o promedio directamente de la dirección estimada $\varphi(k, n)$ (el parámetro de dirección de llegada $\varphi(k, n)$) definido ya antes como la dirección opuesta del vector de intensidad de sonidos activo $I_a(k, n)$ a lo largo del tiempo.

[0117] El procesador de audio espacial 600 puede ser aplicado también al análisis de dirección de micrófono de audio espacial de manera similar. La exactitud de la estimación de dirección puede ser incrementada al promediar resultados sobre varias instancias de medición. Esto significa que similar a DirAC en la Figura 6, el estimador de SAM es mejorado al terminar primero la SNR de la(s) señal(es) de entrada acústica(s) 104. La información en cuanto a la SNR real y la SNR objetivo deseada se hace pasar al estimador de dirección de SAM que incluye un promedio temporal dependiente de la frecuencia y el tiempo de señales de SAM específicas para mejorar la SNR. El promedio es llevado a cabo hasta que se obtiene la SNR objetivo deseada. En efecto, dos señales de SAM pueden ser promediadas, es decir, la dirección estimada $\varphi(k, n)$ o las PSD y CSD definidas en la ecuación 5a y ecuación 5b. La última estimación significa posiblemente que los operadores de expectativa son aproximados por un proceso de promedio cuya longitud depende de la SNR real y la SNR (objetivo) deseada. El promedio de la dirección estimada $\varphi(k, n)$ es explicado para DirAC según la Figura 7b, pero sigue siendo de la misma manera para SAM.

[0118] Según una realización adicional de la presente invención, que será explicada más tarde utilizando la Figura 8, en lugar de promediar explícitamente las cantidades físicas con estos dos procedimientos, es posible cambiar a un banco de filtros utilizado, ya que el banco de filtros puede contener un promedio inherente de las señales de entrada. En lo sucesivo, los dos procedimientos mencionados para promediar las señales de codificación de audio direccional serán explicados en más detalle utilizando las Figuras 7a y 7b. El procedimiento alternativo de cambio del banco de filtros con un procesador de audio espacial se muestra en la Figura 8.

Promedio del vector de densidad de sonido activo en codificación de audio direccional según la Figura 7a

[0119] La Figura 7a muestra un diagrama de bloques esquemático de una primera realización posible del estimador de dirección dependiente de la proporción de señal a ruido 610 de la Figura 6. La realización que se muestra en la Figura 7a está basada en el promedio temporal de la intensidad de sonido acústico o de los parámetros de intensidad de sonido $I_a(k, n)$ por un estimador de dirección 610a. La funcionalidad del estimador de dirección 610a puede ser similar a la funcionalidad del estimador de dirección 610 de la Figura 6, en donde el estimador de dirección 610a puede comprender los elementos adicionales descritos en lo sucesivo.

[0120] El estimador de dirección 610a está configurado para efectuar un promedio y una estimación de dirección. El estimador de dirección 610a es conectado al analizador energético 612 de la Figura 6, el estimador de dirección 610 con el analizador energético 612 puede constituir un estimador de parámetros controlable 606a, una funcionalidad de la cual es similar a la funcionalidad del estimador de parámetros controlables 606 mostrado en la Figura 6. El estimador de parámetros controlables 606a determina en primer lugar a partir de la señal de entrada acústica 104 o las señales de entrada acústica un vector de intensidad de sonido activo 706 ($I_a(k, n)$) en el análisis energético utilizando el analizador energético 612 utilizando la ecuación 1 como se explica anteriormente. En un bloque de promedio 702 del estimador de dirección 610a que efectúa el promedio de este vector (el vector de intensidad de sonido 706) es promediado a lo largo del tiempo n , independientemente para todas (o al menos parte de todas) las bandas de frecuencia o sub-bandas de frecuencia k , que conduce a un vector de intensidad acústica promediado 708 ($I_{avg}(k, n)$) según la siguiente ecuación.

$$I_{avg}(k,n) = \langle I_a(k,n) \rangle_n . \quad (13)$$

[0121] Para llevar a cabo el promedio, el estimador de dirección 610a considera los valores estimativos de intensidad del pasado. Una entrada al bloque de promedio 702 es la proporción de señal a ruido 710 de la entrada acústica 104 o de la señal de entrada acústica 104, que es determinada con el estimador de proporción de señal a ruido 714 mostrado en la Figura 6. La proporción de señal a ruido real 710 de la señal de entrada acústica 104 constituye la característica de señal determinada 110 de la señal de entrada acústica 104. La proporción de señal a ruido es determinada para cada sub-banda de frecuencia k y cada segmento de tiempo n en el dominio de

frecuencia de tiempo corto. Una segunda entrada al bloque de promedio 702 es una proporción de señal a ruido deseada o una proporción de señal a ruido objetivo 712, que debe ser obtenida en una salida del estimador de parámetros controlables 606a, es decir, la proporción de señal a ruido objetivo. La proporción de señal a ruido objetivo 712 es una entrada externa, dada por ejemplo por un usuario. El bloque de promedio 702 promedia el vector de intensidad 706 ($I_a(k, n)$) hasta que se obtiene la proporción de señal a ruido objetivo 712. En base al vector de intensidad promediado (acústico) 708 ($I_{avg}(k, n)$) finalmente la dirección $\varphi(k, n)$ del sonido puede ser calculada utilizando un bloque de estimación de dirección 704 del estimador de dirección 610a que efectúa la estimación de dirección, como se explica anteriormente. El parámetro de dirección de llegada $\varphi(k, n)$ constituye un parámetro espacial 702 determinado por el estimador de parámetros controlables 606a. El estimador de dirección 610a puede determinar el parámetro de dirección de llegada $\varphi(k, n)$ para cada sub-banda de frecuencia k y segmento de tiempo como la dirección opuesta del vector de intensidad de sonido promediado 708 ($I_{avg}(k, n)$) de la sub-banda de frecuencia correspondiente k y el segmento de tiempo correspondiente n .

[0122] Dependiendo de la proporción de señal a ruido objetivo deseada 712, el estimador de parámetros controlable 610a puede hacer variar la longitud de promedio para el promedio de los parámetros de intensidad de sonido 706 ($I_a(k, n)$) de tal manera que la proporción de señal a ruido en la salida del estimador de parámetros controlables 606a coincide (o es igual a) la proporción de señal a ruido objetivo 712. Típicamente, el estimador de parámetros controlable 610a puede escoger una longitud de promedio comparativamente larga para una diferencia comparativamente alta entre la proporción de señal a ruido real 710 de la señal de entrada acústica 104 y la proporción de señal a ruido objetivo 712. Para una diferencia comparativamente baja entre la proporción de señal a ruido real 710 de la señal de entrada acústica 104 y la proporción de señal a ruido objetivo 712, el estimador de parámetros controlable 610a escogerá una longitud de promedio comparativamente corta.

[0123] En otras palabras, el estimador de dirección 606a está basado en el promedio de la intensidad acústica de los parámetros de intensidad acústica.

Promedio del parámetro de dirección de codificación de audio direccional directamente según la Figura 7b

[0124] La Figura 7b muestra un diagrama de bloques esquemático de un estimador de parámetros controlables 606b, la funcionalidad del cual puede ser similar a la funcionalidad del estimador de parámetros controlables 606 mostrado en la Figura 6. El estimador de parámetros controlable 606b comprende el analizador energético 612 y un estimador de dirección 610b configurado para efectuar una estimación de dirección y un promedio. El estimador de dirección 610b difiere del estimador de dirección 610a en que en primer lugar efectúa una estimación de dirección para determinar el parámetro de dirección de llegada 718 ($\varphi(k, n)$) para cada sub-banda de frecuencia k y cada segmento de tiempo n y en segundo lugar efectúa el promedio del parámetro de dirección de llegada determinado 718 para determinar un parámetro de dirección de llegada $\varphi_{prom}(k, n)$ promediado para cada sub-banda de frecuencia k y cada segmento de tiempo n . El parámetro de dirección de llegada promediado ($\varphi_{prom}(k, n)$) constituye un parámetro espacial 102 determinado por el estimador de parámetros controlable 606b.

[0125] En otras palabras, la Figura 7b muestra otra realización posible del estimador de dirección dependiente de la proporción de señal a ruido 610, que es mostrado en la Figura 6. La realización, que es mostrada en la Figura 7b, está basada en el promedio temporal de la dirección estimada (el parámetro de dirección de llegada 718 ($\varphi(k, n)$)) que puede ser obtenida con un procedimiento de codificación de audio convencional, por ejemplo para cada sub-banda de frecuencia k y cada segmento de tiempo n como la dirección opuesta del vector de intensidad del sonido activo 706 ($I_a(k, n)$).

[0126] A partir de la entrada acústica o la señal de entrada acústica 104 el análisis energético se efectúa utilizando el analizador energético 612 y, a continuación, la dirección de sonido (el parámetro de dirección de llegada 718 ($\varphi(k, n)$)) es determinada en un bloque de estimación de dirección 714 del estimador de dirección 710b que efectúa la estimación de dirección, por ejemplo con un procedimiento de codificación de audio direccional convencional explicado anteriormente. A continuación, en un bloque de promedio 716 del estimador de dirección 710b, un promedio temporal es aplicado sobre esta dirección (sobre el parámetro de dirección de llegada 718 ($\varphi(k, n)$)). Como se explica anteriormente, el promedio es llevado a cabo a lo largo del tiempo y para todas (o para al menos parte de todas) las bandas de frecuencia o sub-bandas de frecuencia k , que produce la dirección promediada ($\varphi_{prom}(k, n)$):

$$\varphi_{prom}(k, n) = \langle \varphi(k, n) \rangle_n. \quad (14)$$

[0127] La dirección promediada $\varphi_{prom}(k, n)$ para cada sub-banda de frecuencia k y cada segmento de tiempo

n constituye un parámetro espacial 102 determinado por el estimador de parámetros comparables 606b.

[0128] Como se ha descrito anteriormente, las entradas al bloque de promedio 716 son la proporción de señal a ruido real 710 de la entrada acústica o de la señal de entrada acústica 104, así como la proporción de señal a ruido objetivo 712, que será obtenida en la salida del estimador de parámetros controlable 606b. La proporción de señal a ruido real 710 es determinada para cada sub-banda de frecuencia y cada segmento de tiempo n, por ejemplo en el dominio de SFTF. El promedio 716 es llevado a cabo en un número suficiente de bloques de tiempo (o segmentos de tiempo) hasta que se obtiene la proporción de señal a ruido 712. El resultado final es la dirección promediada temporal $\varphi_{\text{prom}}(k, n)$ con exactitud incrementada.

[0129] Para resumir en breve, el determinador de características de señal 708 está configurado para proporcionar la proporción de señal a ruido 710 de la señal de entrada acústica 104 como una pluralidad de parámetros de proporción de señal a ruido para una sub-banda de frecuencia k y un segmento de tiempo n, de la señal de entrada acústica 104. Los estimadores de parámetros controlables 606a, 606b están configurados para recibir la proporción de señal a ruido objetivo 712 como una pluralidad de parámetros de proporción de señal a ruido objetivo para una sub-banda de frecuencia k y un segmento de tiempo n. Los estimadores de parámetros controlables 606a, 606b están configurados además para derivar la longitud de promedio del promedio temporal según un parámetro de proporción de señal a ruido actual de la señal de entrada acústica de tal manera que un parámetro de proporción de señal a ruido actual del parámetro de dirección de llegada actual (promediado $\varphi_{\text{prom}}(k, n)$) coincide con el parámetro de proporción de señal a ruido objetivo actual.

[0130] Los estimadores de parámetros controlables 606a, 606b están configurados para derivar los parámetros de intensidad $I_a(k, n)$ para cada sub-banda de frecuencia k y cada segmento de tiempo n de la señal de entrada acústica 104. Además, los estimadores de parámetros controlables 606a, 606b están configurados para derivar los parámetros de dirección de llegada $\varphi(k, n)$ para cada sub-banda de frecuencia y cada segmento de tiempo n de la señal de entrada acústica 104 en base a los parámetros de intensidad $I_a(k, n)$ de la señal de audio acústica determinado por los estimadores de parámetros controlables 606a, 606b. Los estimadores de parámetros controlables 606a, 606b están configurados además para derivar el parámetro de dirección de llegada actual $\varphi(k, n)$ para una sub-banda de frecuencia actual y un segmento de tiempo actual en base al promedio temporal de al menos un subconjunto de parámetros de intensidad derivados de la señal de entrada acústica 104 o basados en el promedio temporal de al menos un subconjunto de parámetros de dirección de llegada derivados.

[0131] Los estimadores de parámetros controlables 606a, 606b están configurados para derivar los parámetros de intensidad $I_a(k, n)$ para cada sub-banda de frecuencia k y cada segmento de tiempo n, en el dominio de STFT, además, los estimadores de parámetros controlables 606a, 606b están configurados para derivar el parámetro de dirección de llegada $\varphi(k, n)$ para cada sub-banda de frecuencia k y cada segmento de tiempo n, por ejemplo en el dominio de STFT. El estimador de parámetros controlables 606a está configurado para escoger el subconjunto de parámetros de intensidad para efectuar el promedio temporal, de tal manera que un subcanal de frecuencia asociado con todos los parámetros de intensidad del subconjunto de parámetros de intensidad es igual a una sub-banda de frecuencia actual asociada con el parámetro de dirección de llegada actual. El parámetro controlable 606b está configurado para escoger el subconjunto de parámetros de dirección de llegada para efectuar el promedio temporal 716, de tal manera que un subcanal de frecuencia asociado con todos los parámetros de dirección de llegada del subconjunto de parámetros de dirección de llegada es igual al subcanal de frecuencia actual asociado con el parámetro de dirección de llegada actual.

[0132] Además, el estimador de parámetros controlable 606a está configurado para escoger el subconjunto de parámetros de intensidad, de tal manera que los segmentos de tiempos asociados con los parámetros de intensidad del subconjunto de parámetros de intensidad son adyacentes en el tiempo. El estimador de parámetros controlables 606b está configurado para escoger el subconjunto de parámetros de dirección de llegada, de tal manera que los segmentos de tiempo asociados con los parámetros de dirección de llegada del subconjunto de parámetros de dirección de llegada son adyacentes en el tiempo. El número de parámetros de intensidad en el subconjunto de parámetros de intensidad o el número de parámetros de dirección de llegada en el subconjunto de parámetros de dirección de llegada corresponde a la longitud de promedio del promedio temporal. El estimador de parámetros controlable 606a está configurado para derivar el número de parámetros de intensidad en el subconjunto de parámetros de intensidad para efectuar el promedio temporal en dependencia de la diferencia entre la proporción de señal a ruido actual de la señal acústica 104 y la proporción de señal a ruido objetivo actual. El estimador de parámetros controlable 606b está configurado para derivar el número de parámetros de dirección de llegada en el subconjunto de parámetros de dirección de llegada para efectuar el promedio temporal en base a la diferencia entre la proporción de señal a ruido actual de la señal de entrada acústica 104 y la proporción de señal a ruido objetivo

actual.

[0133] En otras palabras el estimador de dirección 606b está basado en el promedio de la dirección 718 $\varphi(k, n)$ obtenida con una estrategia de codificación de audio direccional convencional.

5

[0134] En lo sucesivo, otra realización de un procesador de audio espacial será descrita, que también efectúa una estimación de parámetros dependiente de la proporción de señal a ruido.

Uso de un banco de filtros con una resolución espectro-temporal apropiada en codificación de audio direccional utilizando un codificador de audio según la Figura 8

[0135] La Figura 8 muestra un procesador de audio espacial 800 que comprende un estimador de parámetros controlable 806 y un determinador de características de señal 808. La funcionalidad del codificador de audio direccional 800 puede ser similar a la funcionalidad del codificador de audio direccional 100. El codificador de audio direccional 800 puede comprender los elementos adicionales descritos en lo sucesivo. La funcionalidad del estimador de parámetros controlable 806 puede ser similar a la funcionalidad del estimador de parámetros controlable 106 y la funcionalidad del determinador de características de señal 808 puede ser similar a la funcionalidad del determinador de características de señal 108. El estimador de parámetros controlable 806 y el determinador de características de señal 808 pueden comprender los elementos adicionales descritos en lo sucesivo.

[0136] El determinador de características de señal 808 difiere del determinador de características de señal 608 en que determina la proporción de señal a ruido 810 de la señal de entrada acústica 104, que es también denotada como proporción de señal a ruido de entrada, en el dominio de tiempo y no en el dominio de STFT. La proporción de señal a ruido 810 de la señal de entrada acústica 104 constituye una característica de señal determinada por el determinador de características de señal 808. El estimador de parámetros controlable 806 difiere del estimador de parámetros controlable 606 mostrado en la Figura 6 en que comprende un estimador de formato B 812 que comprende un banco de filtros 814 y un bloque de cálculo de formato B 816, que está configurado para transformar la señal de entrada acústica 104 en el dominio de tiempo a la representación de formato B, por ejemplo en el dominio de STFT.

[0137] Además, el estimador de formato B 812 está configurado para hacer variar la determinación de formato B de la señal de entrada acústica 104 en base a las características de señal determinadas por el determinador de características de señal 808 o en otras palabras en dependencia de la proporción de señal a ruido 810 de la señal de entrada acústica 104 en el dominio de tiempo.

[0138] Una salida del estimador de formato B 812 es una representación de formato 818 de la señal de entrada acústica 104. La representación de formato B 818 comprende un componente omnidireccional, por ejemplo el vector de presión de sonido mencionado anteriormente $P(k, n)$ y un componente direccional, por ejemplo, el vector de velocidad de sonido mencionado anteriormente $U(k, n)$ para cada sub-banda de frecuencia k y cada segmento de tiempo n .

[0139] Un estimador de dirección 820 del estimador de parámetros controlable 806 deriva un parámetro de dirección de llegada $\varphi(k, n)$ de la señal de entrada acústica 104 para cada sub-banda de frecuencia k y cada segmento de tiempo n . El parámetro de dirección de llegada $\varphi(k, n)$ constituye un parámetro espacial 102 determinado por el estimador de parámetros controlable 806. El estimador de dirección 820 puede efectuar la estimación de dirección al determinar un parámetro de intensidad activo $I_a(k, n)$ para cada sub-banda de frecuencia k y cada segmento de tiempo n y al derivar los parámetros de dirección de llegada $\varphi(k, n)$ en base a los parámetros de intensidad activos $I_a(k, n)$.

50

[0140] El banco de filtros 814 del estimador de formato B 812 está configurado para recibir la proporción de señal a ruido real 810 de la señal de entrada acústica 104 y para recibir una proporción de señal a ruido objetivo 822. El estimador de parámetros controlable 806 está configurado para hacer variar la longitud del bloque del banco de filtro 814 dependiendo de la diferencia entre la proporción de señal a ruido real 810 de la señal de entrada acústica 104 y la proporción de señal a ruido objetivo 822. Una salida de banco de filtros 814 es la representación de frecuencia (por ejemplo, en el dominio de STFT) de la señal de entrada acústica 104, en base a la cual el bloque de cálculo de formato B 816 calcula la representación de formato 818 de la señal de entrada acústica 104. En otras palabras, la conversión de la señal de entrada acústica 104 del dominio de tiempo a la representación de frecuencia puede ser efectuada por el banco de filtros 814 en dependencia de la proporción de señal a ruido real determinada

55

810 de la señal de entrada acústica 104 y en dependencia de la proporción de señal a ruido objetivo 822. En breve, el cálculo de formato B puede ser efectuado por el bloque de formato B 816 en dependencia de la proporción de señal a ruido real determinado 810 y la proporción de señal a ruido objetivo 822.

5 **[0141]** En otras palabras, el determinador de características de señal 808 está configurado para determinar la proporción de señal a ruido 810 de la señal de entrada acústica 104 en el dominio de tiempo. El estimador de parámetros controlable 806 comprende el banco de filtros 814 para convertir la señal de entrada acústica 104 del dominio de tiempo a la representación de frecuencia. El estimador de parámetros controlable 806 está configurado para hacer variar la longitud de bloques del banco de filtros 814 según la proporción de señal a ruido determinada
10 810 de la señal de entrada acústica 104. El estimador de parámetros controlable 806 está configurado para recibir la proporción de señal a ruido objetivo 812 y para hacer variar la longitud del bloque del banco de filtros 814, de tal manera que la proporción de señal a ruido de la señal de entrada acústica 104 en el dominio de frecuencia coincide con la proporción de señal a ruido objetivo 824 o en otras palabras, de tal manera que la proporción de señal a ruido de la representación de frecuencia 824 de la señal de entrada acústica 104 coincide con la proporción de señal a
15 ruido objetivo 822.

[0142] El estimador de parámetros controlable 806 mostrado en la Figura 8 puede ser entendido también como otra realización del estimador de dirección dependiente de la proporción de señal a ruido 610 mostrado en la Figura 6. La realización que se muestra en la Figura 8 está basada en escoger una resolución temporal espectral
20 asociada del banco de filtros 814. Como se ha explicado anteriormente, la codificación de audio direccional opera en el dominio de STFT. Así, las señales de entrada acústicas o la señal de entrada acústica 104 en el dominio de tiempo, por ejemplo medidas con micrófonos son transformadas utilizando por ejemplo una transformación de Fourier de tiempo corto o cualquier otro banco de filtros. El estimado de formato B 812 proporciona a continuación la representación de frecuencia de tiempo corto 818 de la señal de entrada acústica 104 o en otras palabras
25 proporciona la señal de formato B como se denota por la presión de sonido $P(k, n)$ y el vector de velocidad particular $U(k, n)$, respectivamente. La aplicación del banco de filtros 814 sobre las señales de entrada del dominio de tiempo acústicas (sobre la señal de entrada acústica 104 en el dominio de tiempo) promedia inherentemente la señal transformada (la representación de frecuencia de tiempo corto 824 de la señal de entrada acústica 104), mientras que la longitud de promedio corresponde a la longitud de transformada (o longitud de bloques) del banco
30 de filtros 814. El procedimiento de promedio descrito en conjunción con el procesador de audio espacial 800 aprovecha este promedio temporal inherente de la señales de entrada.

[0143] La entrada acústica o la señal de entrada acústica 104 que puede ser medida con los micrófonos, es transformada al dominio de frecuencia de tiempo corto utilizando el banco de filtros 814. La longitud de transformada
35 o longitud de filtro o longitud de bloque es controlada por la proporción de señal a ruido de entrada real 810 de la señal de entrada acústica 104 o de las señales de entrada acústicas y la proporción de señal a ruido objetivo deseadas 822, que debe ser obtenida mediante el proceso de promedio. En otras palabras, se desea efectuar el promedio en el banco de filtros 814, de tal manera que la proporción de señal a ruido de la representación de tiempo-frecuencia 824 de la señal de entrada acústica 104 coincide o es igual con la proporción de señal a ruido
40 objetivo 822. La proporción de señal a ruido es determinada a partir de la señal de entrada acústica 104 o las señales de entrada acústicas en el dominio de tiempo. En el caso de una alta proporción de señal a ruido de entrada 810 una longitud de transformada más corta es escogida y viceversa una para proporción de señal a ruido de entrada baja 810, se escoge una longitud de transformada más larga. Como se explica en la sección anterior, la proporción de señal a ruido de entrada 810 de la señal de entrada acústica 104 es proporcionada por un estimador
45 de proporción de señal a ruido del determinador de características de señal 808, mientras que la proporción de señal a ruido objetivo 822 puede ser controlada externamente, por ejemplo por un usuario. La salida del banco de filtros 814 y el cálculo de formato B subsecuente efectuado por el bloque de cálculo de formato B 816 son la señales de entrada acústicas 818, por ejemplo en el dominio de STFT, es decir $P(k, n)$ o $U(k, n)$. Estas señales (la señal de entrada acústica 818 en el dominio de STFT) son procesadas adicionalmente, por ejemplo, con el procesamiento de
50 audio direccional convencional en el estimador de dirección 820 para obtener la dirección x para cada sub-banda de frecuencia k y cada segmento de tiempo n .

[0144] En otras palabras, el procesador de audio espacial 800 o el estimador de dirección están basados en escoger un banco de filtro apropiado para la señal de entrada acústica 104 o para las señales de entrada acústica.
55

[0145] En breve, el determinador de características de señal 808 está configurado para determinar la proporción de señal a ruido 810 de la señal de entrada acústica 104 en el dominio de tiempo. El estimador de parámetros controlable 806 comprende el banco de filtros 814 configurado para convertir la señal de entrada acústica 104 desde el dominio de tiempo a la representación de frecuencia. El estimador de parámetros controlable

806 está configurado para hacer variar la longitud de bloque del banco de filtros 814, según la proporción de señal a ruido determinada 810 de la señal de entrada acústica 104. Además, el estimador de parámetros controlable 806 está configurado para recibir la proporción de señal a ruido objetivo 822 y para hacer variar la longitud de bloques del banco de filtros 814, de tal manera que la proporción de señal a ruido de la señal de entrada acústica 824 en la representación de frecuencia coincide con la proporción de señal a ruido objetivo 822.

[0146] La estimación de la proporción de señal a ruido efectuada por el estimador de características de señales 608, 808 es un problema bien conocido. En lo sucesivo, se describirá una implementación posible de un estimador de la proporción de señal a ruido.

10

Implementación posible de un estimador de SNR

[0147] En lo sucesivo, se describirá una implementación posible del estimador de proporción de señal a ruido de entrada 614 en la Figura 600. El estimador de la proporción de señal a ruido descrito en lo sucesivo puede ser usado por el estimador de parámetros controlable 606a y el estimador de parámetros controlables 606b mostrado en la Figura 7a y 7b. El estimador de la proporción de señal a ruido estima la proporción de señal a ruido de la señal de entrada acústica 104, por ejemplo en el dominio de STFT. Una implementación en dominio de tiempo (por ejemplo, implementado en el determinador de características de señal 808) puede ser realizada de manera similar.

[0148] El estimador de SNR puede estimar la SNR de las señales de entrada acústicas, por ejemplo en el dominio de STFT para cada bloque de tiempo n y banda de frecuencia k o para señal de dominio de tiempo. La SNR es estimada al calcular la energía de la señal para la bandeja de tiempo-frecuencia considerada. Sea x (k, n) la señal de entrada acústica. La energía de la señal S (k, n) puede ser determinada con

$$S(k,n) = |x(k,n)|^2 \quad (15)$$

25

[0149] Para obtener la SNR, la proporción entre la energía de la señal y la energía de ruido N (k) es calculada, es decir

$$SNR = S(k,n) / N(k).$$

30

[0150] Dado que S (k, n) ya contiene ruido, un estimador de SNR más exacto en el caso de SNR baja es dado por:

$$SNR = (S(k,n) - N(k)) / N(k). \quad (16)$$

35

[0151] Se supone que la señal de energía de ruido N (k) es constante a lo largo del tiempo n. Se puede determinar para cada k de la entrada acústica. En efecto, es igual a la potencia media de la señal de entrada acústica en el caso de que ningún sonido este presente, es decir durante el silencio. Expresado en términos matemáticos

40

$$N (k) = \langle |x(k, n)|^2 \rangle_n, \quad x (k, n) \text{ medido durante el silencio} \quad (17)$$

[0152] En otras palabras, según algunas realizaciones de la presente invención, un determinador de características de señal está configurado para medir una señal de ruido durante una fase silente de la señal de entrada acústica 104 y para calcular la energía N (k) de la señal de ruido. El determinador de características de señal puede estar configurado además para medir una señal activa durante una fase no silente de la señal de entrada acústica 104 y para calcular la energía S (k, n) de la señal activa. El determinador de características de señal puede estar configurado además para calcular el estimador de la proporción de señal a ruido de la entrada señal acústica 104 en base a la potencia calculada N (k) de la señal de ruido y la potencia calculada S(k) de la señal activa.

50

[0153] Este esquema puede ser aplicado también al determinador de características de señal 808 con la

diferencia de que el determinador de características de señal 808 determina la energía $S(t)$ de la señal activa en el dominio de tiempo y determina la energía $N(t)$ de la señal de ruido en el dominio de tiempo, para obtener el estimador de la proporción de señal a ruido real de la señal de entrada acústica 104 en el dominio de tiempo.

5 **[0154]** En otras palabras, los determinadores de características de señal 608, 808 están configurados para medir una señal de ruido durante una fase silente durante la fase de señal de entrada acústica 104 y para calcular la energía $N(k)$ de señal de ruido. Los determinadores de características de señal 606, 808 están configurados para medir una señal activa durante una fase no silente de la señal de entrada acústica 104 y para calcular la energía de la señal activa ($S(k, n)$). Además, los determinadores de características de señal 608, 808 están configurados para
10 determinar la proporción de señal a ruido de la señal de entrada acústica 104 en base a la energía calculada $N(k)$ en base a la señal de ruido y la energía calculada $S(k)$ de la señal activa.

[0155] En lo sucesivo, otra realización de la presente invención será descrita que efectúa una estimación de parámetros dependiente de aplausos.

15

Estimación de parámetros dependiente de aplauso utilizando un procesador de audio espacial según la Figura 9

[0156] La Figura 9 muestra un diagrama de bloques esquemático de un procesador de audio espacial 900 según una realización de la presente invención. Una funcionalidad del procesador de audio espacial 900 puede ser similar a la funcionalidad del procesador de audio espacial 100 y el procesador de audio espacial 900 puede comprender los elementos adicionales descritos en lo sucesivo. El procesador de audio espacial 900 comprende un estimador de parámetros controlables 906 y un determinador de características de señal 908. Una funcionalidad del estimador de parámetros controlables 906 puede ser similar a la funcionalidad del estimador de parámetros controlables 106 y el estimador de parámetros controlables 906 puede comprender los elementos adicionales descritos en lo sucesivo. Una funcionalidad del determinador de características de señal 908 puede ser similar a la funcionalidad del determinador de características de señal 108 al determinar de características de señal 908 puede comprender los elementos adicionales descritos en lo sucesivo.

30 **[0157]** El determinador de características de señal 908 está configurado para determinar si la señal de entrada acústica 104 comprende componentes transitorios que corresponden a señales semejantes a aplauso utilizando un detector de aplauso 910.

[0158] Las señales semejantes a aplausos definidas en la presente como señales que comprenden una
35 secuencia temporal rápida de transitorios, por ejemplo con diferentes direcciones.

[0159] El estimador de parámetros controlable 906 comprende un banco de filtros 912 que está configurado para convertir la señal de entrada acústica 104 del dominio de tiempo a una representación de frecuencia (a un dominio de STFT) en base a una regla de cálculo de conversión. El estimador de parámetros controlable 906 está
40 configurado para escoger la regla de cálculo de conversión para convertir la señal de entrada acústica 104 del dominio de tiempo a la representación de frecuencia de una pluralidad de reglas de cálculo de conversión según el resultado de una determinación de características de señal efectuada por el determinador de características de señal 908. El resultado de la determinación de características de señal constituye la característica de señal determina 110 del determinador de características de señal 908. El estimador de parámetros controlables 906
45 escoge la regla de cálculo de conversión de una pluralidad de reglas de cálculo de conversión, de tal manera que una primera regla de cálculo de conversión de la pluralidad de reglas de cálculo de conversión es escogida para convertir la señal de entrada acústica 104 del dominio de tiempo a la representación de frecuencia cuando la señal de entrada acústica comprende componentes correspondientes a aplausos y de tal manera que una segunda regla de cálculo de conversión de la pluralidad de reglas de cálculo de conversión es escogida para convertir la señal de
50 entrada acústica 104 del dominio de tiempo a la representación de frecuencia cuando la señal de entrada acústica 104 no comprende componentes correspondientes a aplausos.

[0160] En otras palabras, el estimador de parámetros controlable 906 está configurado para escoger una
55 regla de cálculo de conversión apropiada para convertir la señal de entrada acústica 104 del dominio de tiempo a la representación de frecuencia en dependencia de la detección de aplauso.

[0161] En breve, el procesador de audio espacial 900 es mostrado como una realización ejemplar de la invención, en donde la descripción paramétrica del campo de sonido es determinada dependiendo de las características de las señales de entrada acústica o la señal de entrada acústica 104. En el caso de micrófonos que

capturan el aplauso o la señal de entrada acústica 104 comprende componentes correspondientes a señales semejantes a aplauso, se usa un procesamiento especial con el fin de incrementar la exactitud de la estimación del parámetro.

5 **[0162]** El aplauso está caracterizado usualmente por una variación rápida de la dirección de llegada de sonido dentro de un periodo de tiempo muy corto. Además, las señales de sonido capturadas contienen principalmente transitorios. Se ha encontrado que para un análisis exacto del sonido es ventajoso un sistema que pueda resolver la variación temporal rápida de la dirección de llegada y que pueda conservar el carácter transitorio de los componentes de señal.

10

[0163] Estos objetivos pueden ser obtenidos al utilizar un banco de filtro con alta resolución temporal (por ejemplo, una STFT con transformada corta o longitud de bloque corta) para transformar las señales de entrada de dominio de tiempo acústicas. Cuando se usa tal banco de filtros, la resolución espectral del sistema será reducida. Esto no es problemático para señales de aplauso ya que la DOA del sonido no varía mucho a lo largo de la frecuencia debido a las características transitorias del sonido, sin embargo, se ha encontrado que una resolución espectral pequeña es problemática para otras señales tales como habla en un escenario de doble habla, en donde una cierta resolución espectral es requerida para ser capaces de distinguir entre los interlocutores individuales. Se ha encontrado que una estimación de parámetros exacta puede requerir un cambio dependiente de la señal del banco de filtro (o de la transformada correspondiente o longitud correspondiente del banco de filtro) dependiendo de las características de las señales de entrada acústicas o de la señal de entrada acústica 104.

[0164] El codificador espacial 900 mostrado en la Figura 9 representa una realización posible para efectuar la conmutación dependiente de la señal del banco de filtro 912 o de escoger la regla de cálculo de conversión de banco de filtro 912. Antes de transformar las señales de entrada o la señal de entrada acústica 104 a la representación de frecuencia (por ejemplo, al dominio de STFT) con el banco de filtro 912, las señales de entrada o la señal de entrada acústica 104 se hace pasar al detector de aplausos 910 del determinador de características de señal 908. La señal de entrada acústica 104 se hace pasar al detector de aplausos 910 en el dominio de tiempo. El detector de aplausos 910 del determinador de características de señal 908 controla el banco de filtro 912 en base a la característica de señal determinada 110 (que en este caso señala así la señal de entrada acústica 104 contiene componentes correspondientes a señales semejantes a aplauso o no). Si el aplauso es detectado en las señales de entrada acústica o en la señal de entrada acústica 104, el estimador de parámetros controlables 900 cambia a un banco de filtros o en otras palabras una regla de cálculo de conversión es escogida en el banco de filtros que es apropiada para el análisis de aplauso. En el caso de que ningún aplauso esté presente, un banco de filtros convencional o en otras palabras una regla de cálculo de conversión convencional que puede ser por ejemplo conocido del codificador de audio direccional 200 es usada. Después de transformar la señal de entrada acústica 104 al dominio de STFT (u otra representación de frecuencia), un procesamiento de codificación de audio direccional convencional puede ser llevado a cabo (utilizando un bloque de cálculo de formato B 914 y un bloque de estimación de parámetros 916 del estimador de parámetros controlable 906). En otras palabras, la determinación de los parámetros de codificación de audio direccionales, que constituyen los parámetros espaciales 102, que son determinados por el procesador de audio espacial 900, puede ser llevado a cabo utilizando el bloque de cálculo de formato B 914 y el bloque de estimación de parámetros 916 como se describe según el codificador de audio direccional 200 mostrado en la Figura 2. Los resultados son por ejemplo los parámetros de codificación de audio direccional, es decir, dirección $\phi(k, n)$ y difusividad $\Psi(k, m)$.

45 **[0165]** En otras palabras, el procesador de audio espacial 900 proporciona un concepto en el cual la estimación de los parámetros de codificación de audio direccionales es mejorada al cambiar el banco de filtros en el caso de señales de aplauso o señales semejantes a aplauso.

[0166] En breve, el estimador de parámetros controlable 906 está configurado de tal manera que la primera regla de cálculo de conversión corresponde a una resolución temporal más alta de la señal de entrada acústica en la representación de frecuencia que la segunda regla de cálculo de conversión y de tal manera que la segunda regla de cálculo de conversión corresponde a una resolución espectral más alta de la señal de entrada acústica en la representación de frecuencia que la primera regla de cálculo de conversión.

55 **[0167]** El detector de aplausos 910 del determinador de características de señal 908 puede determinar por ejemplo si la señal de entrada acústica 104 comprende señales semejantes a aplauso en base a metadatos, por ejemplo generados por un usuario.

[0168] El procesador de audio espacial 900 mostrado en la Figura 9 puede ser aplicado también al análisis de

SAM de manera similar con la diferencia de que ahora el banco de filtros del SAM es controlado por el detector de aplausos 910 del determinador de características de señal 908.

[0169] En una realización adicional de la presente invención, el estimador de parámetros controlable puede determinar los parámetros espaciales utilizando diferentes estrategias de estimación de parámetros independientes de la característica de señal determinada, de tal manera que para cada estrategia de estimación de parámetros, el estimador de parámetros controlable determina un conjunto de parámetros espaciales de la señal de entrada acústica. El estimador de parámetros controlable puede estar configurado además para seleccionar un conjunto de parámetros espaciales de los conjuntos determinados de parámetros espaciales como el parámetro espacial de la señal de entrada acústica y por consiguiente como el resultado del proceso de estimación en dependencia de la característica de señal determinada. Por ejemplo, una primera regla de cálculo de parámetros espaciales variables puede comprender: determinar parámetros espaciales de la señal de entrada acústica para cada estrategia de estimación de parámetros y seleccionar el conjunto de parámetros espaciales determinados con una primera estrategia de estimación de parámetros. Una segunda regla de cálculo de parámetros espaciales variables puede comprender: determinar parámetros espaciales de la señal de entrada acústica para cada estrategia de estimación de parámetros y seleccionar el conjunto de parámetros espaciales determinado con una segunda estrategia de estimación de parámetros.

[0170] La Figura 10 muestra un diagrama de flujo de un procedimiento 1000 según una realización de la presente invención.

[0171] El procedimiento 1000 para proporcionar parámetros espaciales en base a una señal de entrada acústica comprende una etapa 1010 para determinar una característica de señal de la señal de entrada acústica.

[0172] El procedimiento 1000 comprende además una etapa 1020 de modificación de una regla de cálculo de parámetros espaciales variables según la característica de señal determinada.

[0173] El procedimiento 1000 comprende además una etapa 1030 para calcular parámetros espaciales de la señal de entrada acústica según la regla de cálculo de parámetros espaciales variable.

[0174] Realizaciones de la presente invención se refieren al procedimiento que controla estrategias de estimación de parámetros en sistemas para representación de sonido espacial basados en características de señales de entrada acústicas, es decir, señales de micrófono.

[0175] En lo sucesivo, algunos aspectos de realizaciones de la presente invención serán resumidos.

[0176] Al menos algunas realizaciones de la presente invención están configuradas para recibir señales de audio de multicanal acústicas, es decir señales de micrófono. A partir de las señales de entrada acústica, realizaciones de la presente invención pueden determinar las características de señal específicas. En base a las características de señal, realizaciones de la presente invención pueden escoger el modelo de señal de mejor ajuste. El modelo de señal puede controlar a continuación la estrategia de estimación de parámetros. En base a la estrategia de estimación de parámetros controlada o seleccionada, realizaciones de la presente invención pueden estimar parámetros espaciales de mejora ajuste para la señal de entrada acústica dada.

[0177] La estimación de descripción de campo de sonido paramétrico depende de suposiciones específicas en las señales de entrada acústicas. Sin embargo, esta entrada puede exhibir una varianza temporal significativa y así un modelo invariante en el tiempo es frecuentemente inapropiado. En la codificación paramétrica, este problema puede ser resuelto al identificar a priori las características de señal y escoger entonces la estrategia de codificación mejor de una manera variante en el tiempo. Realizaciones de la presente invención determinan las características de señal de las señales de entrada acústicas no a priori sino continuamente, por ejemplo por bloques, por ejemplo para una sub-banda de frecuencia y un segmento de tiempo o para un subconjunto de sub-bandas de frecuencia y/o un subconjunto de segmento de tiempo. Realizaciones de la presente invención pueden aplicar esta estrategia a extremos frontales acústicos para el procesamiento de audio espacial paramétrico y/o codificación de audio espacial tal como codificación de audio direccional (DirAC) o micrófono de audio espacial (SAM).

[0178] Es una idea de realizaciones de la presente invención usar estrategias de procesamiento de datos dependientes de la señal dependientes en el tiempo para la estimación de parámetros en codificación de audio espacial paramétrica basada en señales de micrófono o en otras señales de entrada acústicas.

[0179] Realizaciones de la presente invención han sido descritas con un enfoque principal sobre la estimación de parámetros en codificación de audio direccional, sin embargo el concepto presentado puede ser aplicado también a otras estrategias paramétricas, tal como un micrófono de audio espacial.

5 **[0180]** Realizaciones de la presente invención proporcionan una estimación de parámetros adaptables de señal para sonido espacial basado en señales de entrada acústicas.

[0181] Diferentes realizaciones de la presente invención han sido descritas. Algunas realizaciones de la presente invención efectúan una estimación de parámetros dependiendo de un intervalo estacionario de las señales de entrada. Realizaciones adicionales de la presente invención efectúan una estimación de parámetros dependiendo de situaciones de doble habla. Realizaciones adicionales de la presente invención efectúan una estimación de parámetros dependiente de una proporción de señal a ruido de las señales de entrada. Realizaciones adicionales de la presente invención efectúan una estimación de parámetros basada en el promedio del vector de intensidad de sonido dependiendo de la proporción de señal a ruido de entrada. Realizaciones adicionales de la presente invención efectúan la estimación de parámetros basada en el promedio del parámetro de dirección estimado dependiendo de la proporción de señal a ruido de entrada. Realizaciones adicionales de la presente invención efectúan la estimación de parámetros al escoger un banco de filtros apropiado o una regla de cálculo de conversión apropiada dependiendo de la proporción de señal a ruido de entrada. Realizaciones adicionales de la presente invención efectúan la estimación de parámetros dependiendo de la tonalidad de las señales de entrada acústicas. Realizaciones adicionales de la presente invención efectúan la estimación de parámetros dependiente de las señales semejantes a aplauso.

[0182] Un procesador de audio espacial puede ser en general un aparato que procesa audio espacial y genera o procesa información paramétrica.

25 **Alternativas de implementación**

[0183] Aunque algunos aspectos han sido descritos en el contexto de un aparato, está claro que estos aspectos también representan una descripción del procedimiento correspondiente en donde un bloque o dispositivo corresponde a una etapa de procedimiento o un elemento de una etapa de procedimiento. Análogamente, aspectos descritos en el contexto de una etapa de procedimiento también representan una descripción de un bloque o ítem o elemento correspondiente de un aparato correspondiente. Algunas o todas las etapas del procedimiento pueden ser ejecutadas al (o usando) un aparato de elementos físicos, por ejemplo un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, una o más de las etapas de procedimiento más importantes pueden ser ejecutadas mediante tal aparato.

[0184] Dependiendo de ciertos requisitos de implementación, realizaciones de la invención pueden ser implementadas en hardware o en software. La implementación puede ser efectuada utilizando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un Blue-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control que se pueden leer electrónicamente almacenadas en la misma que cooperan (o son capaces de cooperar) con un sistema informático programable, de tal manera que el procedimiento respectivo es efectuado. Por consiguiente, el medio de almacenamiento digital puede ser legible por ordenador.

45 **[0185]** Algunas realizaciones según la invención comprenden un portador de datos que tiene señales de control que se pueden leer electrónicamente que son capaces de cooperar con un sistema informático programable, de tal manera que se efectúa uno de los procedimientos descritos en esta invención.

[0186] En general, realizaciones de la presente invención pueden ser implementadas como un producto de programa informático con un código de programa, el código de programa es operativo para efectuar uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. Los códigos de programa pueden ser almacenados por ejemplo en un soporte legible por máquina.

[0187] Otras realizaciones comprenden el programa informático para efectuar uno de los procedimientos descritos en esta invención, almacenados en un soporte legible por máquina.

[0188] En otras palabras, una realización del procedimiento de la invención es por consiguiente un programa informático que tiene códigos de programa para efectuar uno de los procedimientos descritos en esta invención, cuando el programa informático se ejecuta en un ordenador.

[0189] Una realización adicional de los procedimientos de la invención es por consiguiente un soporte de datos (o medio de almacenamiento digital o medio legible por ordenador) que comprende, registrado sobre el mismo, el programa informático para efectuar uno de los procedimientos descritos en esta invención.

5

[0190] Una realización adicional del procedimiento de la invención es, por consiguiente, una corriente de datos o una secuencia de señales que representan el programa informático para efectuar uno de los procedimientos descritos en esta invención. La corriente de datos o la secuencia de señales puede por ejemplo estar configurada para ser transferida vía una conexión de comunicación de datos, por ejemplo vía Internet.

10

[0191] Una realización adicional comprende un medio de procesamiento, por ejemplo un ordenador o un dispositivo lógico programable configurado para o adaptado para efectuar uno de los procedimientos descritos en esta invención.

15 **[0192]** Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para efectuar uno de los procedimientos descritos en esta invención.

[0193] En algunas realizaciones, un dispositivo lógico programable (por ejemplo, un arreglo de compuerta programable en el campo) puede ser usado para efectuar algunas o todas las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, un arreglo de compuerta programable en el campo puede cooperar con un microprocesador con el fin de efectuar uno de los procedimientos descritos en esta invención. En general, los procedimientos son efectuados preferiblemente por cualquier aparato de hardware.

20

[0194] Las realizaciones descritas anteriormente son meramente ilustrativas para los principios de la presente invención. Se comprenderá que modificaciones y variaciones de los arreglos y los detalles descritos en esta invención se harán evidentes para otros expertos en la técnica. Se pretende, por consiguiente, estar limitados solo por el alcance de las reivindicaciones de patentes pendientes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones de esta invención.

25

REIVINDICACIONES

1. Un procesador de audio espacial para proporcionar parámetros espaciales (102, $\phi(k, n)$, $\Psi(k, n)$) en base a una señal de entrada acústica, el procesador de audio espacial esta caracterizado porque comprende:
- 5 un determinador de características de señal (108, 308, 408, 508, 608, 808, 908) configurado para determinar una característica de señal (110, 710, 810) de la señal de entrada acústica (104), y un estimador de parámetros controlable (106, 306, 406, 506, 606, 606a, 606b, 806, 906) para calcular los parámetros espaciales (102, $\phi(k, n)$, $\Psi(k, n)$) para la señal de entrada acústica (104) según una regla de cálculo de parámetros espaciales variable;
- 10 en el que el estimador de parámetros controlable (106, 306, 406, 506, 606, 606a, 606b, 806, 906) está configurado para modificar la regla de cálculo de parámetros espaciales según la característica de señal determinada (110, 710, 810);
- 15 en el que el determinador de características de señal (308) está configurado para determinar un intervalo estacionario de la señal de entrada acústica (104) y el estimador de parámetros controlable (306) está configurado para modificar la regla de cálculo de parámetros espaciales variables según el intervalo estacionario determinado, de forma que un período de promedio para el cálculo de los parámetros espaciales (102, $\Psi(k, n)$, $\phi(k, n)$) es comparativamente más largo para un intervalo estacionario comparativamente más largo y es comparativamente más corto para un intervalo estacionario comparativamente más corto; o
- 20 en el que el estimador de parámetros controlable (406, 506, 906) está configurado para seleccionar una regla de cálculo de parámetros espaciales (410, 412) fuera de una pluralidad de reglas de cálculo de parámetros espaciales (410, 412) para el cálculo de los parámetros espaciales (102, $\Psi(k, n)$, $\phi(k, n)$), en dependencia de la característica de señal determinada (110).
- 25 2. El procesador de audio espacial según la reivindicación 1, en el que los parámetros espaciales (102) comprenden una dirección del sonido y/o una difusividad del sonido y/o una medida estadística de la dirección del sonido.
3. El procesador de audio espacial según la reivindicación 1 ó 2, en el que el estimador de parámetros controlable (106, 306, 406, 506, 606, 606a, 606b, 806, 906) está configurado para calcular los parámetros espaciales (102, $\phi(k, n)$, $\Psi(k, n)$) como parámetros de codificación de audio direccional que comprende un parámetro de difusividad ($\Psi(k, n)$) para un segmento de tiempo (n) y para un sub-banda de frecuencia (k) y/o un parámetro de dirección de llegada ($\phi(k, n)$) para un segmento de tiempo (n) y para una sub-banda de frecuencia (k) o como parámetros de micrófono de audio espacial.
- 35 4. El procesador de audio espacial según una de las reivindicaciones 1 a 3, en el que el estimador de parámetros controlable (306) está configurado para calcular los parámetros espaciales (102, $\psi(k, n)$) a partir de la señal de entrada acústica (104) para un segmento de tiempo (n) y una sub-banda de frecuencia (k) en base al menos a un promedio en tiempo de los parámetros de señal ($I_a(k, n)$) de la señal de entrada acústica (104); y
- 40 en el que el estimador de parámetros controlable (306) está configurado para hacer variar un período de promedio del promedio en el tiempo de los parámetros de señal ($I_a(k, n)$) de la señal de entrada acústica (104) según el intervalo estacionario determinado.
- 45 5. El procesador de audio espacial según la reivindicación 4, en el que el estimador de parámetros controlable (306) está configurado para aplicar el promedio en tiempo de los parámetros de señal ($I_a(k, n)$) de la señal de entrada acústica utilizando un filtro de paso de bajo;
- 50 en el que el estimador de parámetros controlable (306) está configurado para ajustar la ponderación entre un parámetro de señal actual de la señal de entrada acústica (104) y parámetros de señal previos de la señal de entrada acústica (104) en base a un parámetro de ponderación (α), de tal manera que el período de promedio está basado en el parámetro de ponderación (α), de tal manera que un peso del parámetro de señal actual comparado con el peso de los parámetros de señal previos es comparativamente alto para un intervalo estacionario comparativamente corto y de tal manera que el peso del parámetro de señal actual comparado con el peso de los
- 55 parámetros de señal previos es comparativamente bajo para un intervalo estacionario comparativamente largo.
6. El procesador de audio espacial según una de las reivindicaciones 1 a 5, en el que el estimador de parámetros controlable (406, 506) está configurado se tal forma que una primera regla de cálculo de parámetros espaciales (410) de una pluralidad de reglas de cálculo de parámetros espaciales (410, 412) es diferente a una

segunda regla de cálculo de parámetros espaciales (410) de la pluralidad de reglas de cálculo de parámetros espaciales (410, 412) y en el que la primera regla de cálculo de parámetros espaciales (410) y la segunda regla de parámetros espaciales (412) son seleccionadas de un grupo reconstituido por: promedio en el tiempo sobre una pluralidad de segmentos de tiempo en una sub-banda de frecuencia, promedio de frecuencia sobre una pluralidad de sub-bandas de frecuencia en un segmento en el tiempo, promedio en el tiempo y promedio de frecuencia y sin promedio.

7. El procesador de audio espacial según una de las reivindicaciones 1 a 6, en el que el determinador de características de señal (408) está configurado para determinar si la señal de entrada acústica (104) comprende componentes de diferentes fuentes de sonido al mismo tiempo o en el que el determinador de características de señal está configurado para determinar una tonalidad de la señal de entrada acústica (104);

en el que el estimador de parámetros controlables (406, 506) está configurado para seleccionar, según un resultado de la determinación de características de señal, una regla de cálculo de parámetros espaciales (410, 412) de una pluralidad de reglas de cálculo de parámetros espaciales (410, 412) para calcular los parámetros espaciales ($102, \Psi(k, n), \phi(k, n)$), de tal manera que una primera regla de cálculo de parámetros espaciales (410) de la pluralidad de reglas de cálculo de parámetros espaciales (410, 412) es escogida cuando la señal de entrada acústica (104) comprende componentes como máximo de una fuente de sonido o cuando la tonalidad de la señal de entrada acústica (104) está por debajo de un nivel de umbral de tonalidad dada y de tal manera que una segunda regla de cálculo de parámetros espaciales (412) de la pluralidad de reglas de cálculo de parámetros espaciales (410, 412) es escogida cuando la señal de entrada acústica (104) comprende componentes de más de una fuente de sonido al mismo tiempo o cuando la tonalidad de la señal de entrada acústica (104) está por encima de un nivel de umbral de tonalidad dado;

en el que la primera regla de cálculo de parámetros espaciales (410) incluye un promedio de frecuencia sobre un primer número de sub-bandas de frecuencia (k) y la segunda regla de cálculo de parámetros espaciales (412) incluye un promedio de frecuencia sobre un segundo número de sub-bandas de frecuencia (k) y no incluye un promedio de frecuencia y en el que el primer número es mayor que el segundo número.

8. El procesador de audio espacial según una de las reivindicaciones 1 a 7, en el que el determinador de características de señal (608) está configurado para determinar la proporción de señal a ruido (110, 710) de la señal de entrada acústica (104);

en el que el estimador de parámetros controlable (606, 606a, 606b) está configurado para aplicar un promedio en el tiempo sobre una pluralidad de segmentos de tiempo en una sub-banda de frecuencia (k), un promedio de frecuencia sobre una pluralidad de sub-bandas de frecuencia (k) en un segmento de tiempo (n), un promedio espacial o una combinación de los mismos; y

en el que el estimador de parámetros controlable (606, 606a, 606b) está configurado para hacer variar un período de promedio del promedio en el tiempo, del promedio de frecuencia, del promedio espacial o de la combinación de los mismos, según la proporción de señal a ruido determinada (110, 710), de tal manera que el período de promedio es comparativamente más largo para una proporción de señal a ruido comparativamente más bajo (110, 710) de la señal de entrada acústica y de tal manera que el período de promedio es comparativamente más corto para una proporción de señal a ruido comparativamente más alta (110, 710) de la señal de entrada acústica (104).

9. El procesador de audio espacial según la reivindicación 8, en el que el estimador de parámetros controlable (606a, 606b) está configurado para aplicar el promedio en el tiempo a un subconjunto de parámetros de intensidad ($I_a(k, n)$) sobre una pluralidad de segmentos de tiempo y una sub-banda de frecuencia (k) o a un subconjunto de parámetros de dirección de llegada ($\phi(k, n)$) sobre una pluralidad de segmentos de tiempo y una sub-banda de frecuencia (k); y

en el que un número de parámetros de intensidad ($I_a(k, n)$) en el subconjunto de parámetros de intensidad ($I_a(k, n)$) o un número de parámetros de dirección de llegada ($\phi(k, n)$) en el subconjunto de parámetros de dirección de llegada ($\phi(k, n)$) corresponde al período de promedio del promedio en el tiempo, de tal manera que el número de parámetros de intensidad ($I_a(k, n)$) en el subconjunto de parámetros de intensidad ($I_a(k, n)$) o el número de parámetros de dirección de llegada ($\phi(k, n)$) en el subconjunto de parámetros de dirección de llegada ($\phi(k, n)$) es comparativamente más bajo para una proporción de señal a ruido comparativamente más alta (110, 710) de la señal de entrada acústica (104) y de tal manera que el número de parámetros de intensidad ($I_a(k, n)$) en el subconjunto de parámetros de intensidad ($I_a(k, n)$) o el número de parámetros de dirección de llegada ($\phi(k, n)$) en el subconjunto de parámetros de dirección de llegada ($\phi(k, n)$) es comparativamente más alto para una proporción de señal a ruido

comparativamente más baja (110, 710) de la señal de entrada acústica (104).

10. El procesador de audio espacial según una de las reivindicaciones 8 a 9, en el que el determinador de características de señal (608) está configurado para proporcionar la proporción de señal a ruido (110, 710) de la señal de entrada acústica (104) como una pluralidad de parámetros de proporción de señal a ruido de la señal de entrada acústica (104), cada parámetro de proporción de señal a ruido de la señal de entrada acústica (104) está asociado con una sub-banda de frecuencia y un segmento de tiempo, en el que el estimador de parámetros controlable (606a, 606b) está configurado para recibir una proporción de señal a ruido objetivo (712) como una pluralidad de parámetros de proporción de señal a ruido objetivo, cada parámetro de proporción de señal a ruido objetivo está asociado con una sub-banda de frecuencia y un segmento de tiempo; y

en el que el estimador de parámetros controlable (606a, 606b) está configurado para hacer variar el período de promedio del promedio en el tiempo según un parámetro de proporción de señal a ruido actual en la señal de entrada acústica, de tal manera que un parámetro de proporción de señal a ruido actual (102) intenta coincidir con un parámetro de proporción de señal a ruido objetivo actual.

11. El procesador de audio espacial según una de las reivindicaciones 1 a 10, en el que el determinador de características de señal (908) está configurado para determinar si la señal de entrada acústica (104) comprende componentes transitorios que corresponden a señales semejantes a aplauso;

en el que el estimador de parámetros controlable (906) comprende un banco de filtros (912) que está configurado para convertir la señal de entrada acústica (104) de un dominio de tiempo a una representación de frecuencia basada en una regla de cálculo de conversión; y

en el que el estimador de parámetros controlable (906) está configurado para escoger la regla de cálculo de conversión para convertir la señal de entrada acústica (104) del dominio de tiempo a la representación de frecuencia de una pluralidad de reglas de cálculo de conversión según el resultado de la determinación de características de señal, de tal manera que una primera regla de cálculo de conversión de la pluralidad de las reglas de cálculo de conversión es escogida para convertir la señal de entrada acústica (104) del dominio de tiempo a la representación de frecuencia cuando la señal de entrada acústica comprende componentes correspondientes a señales semejantes a aplauso y de tal manera que una segunda regla de conversión de la pluralidad de reglas de cálculo de conversión es escogida para convertir la entrada de señal acústica (104) del dominio de tiempo a la representación de frecuencia cuando la señal de entrada acústica no comprende componentes correspondientes a señales semejantes a aplauso.

12. Un procedimiento para proporcionar parámetros espaciales en base a una señal de entrada acústica, comprendiendo el procedimiento:

la determinación (1010) de una característica de señal de la señal de entrada acústica;
la modificación (1020) de una regla de cálculo de parámetros espaciales variables según la característica de señal determinada;

el cálculo (1030) de los parámetros espaciales de la señal de entrada acústica según la regla de cálculo de parámetros espaciales variables; y

la determinación de un intervalo estacionario de la señal de entrada acústica y la modificación de la regla de cálculo de parámetros espaciales variable según el intervalo estacionario determinado, de forma que un período de promedio para el cálculo de los parámetros espaciales sea comparativamente más largo para un intervalo estacionario comparativamente más largo y sea comparativamente más corto para un intervalo estacionario comparativamente más corto; o

la selección de una regla de cálculo de parámetro espacial fuera de una pluralidad de reglas de cálculo de parámetros espaciales para el cálculo de los parámetros espaciales en dependencia de la característica de señal determinada.

13. Un programa informático que tiene un código de programa adaptado para efectuar, cuando se ejecuta en un ordenador el procedimiento según la reivindicación 12.

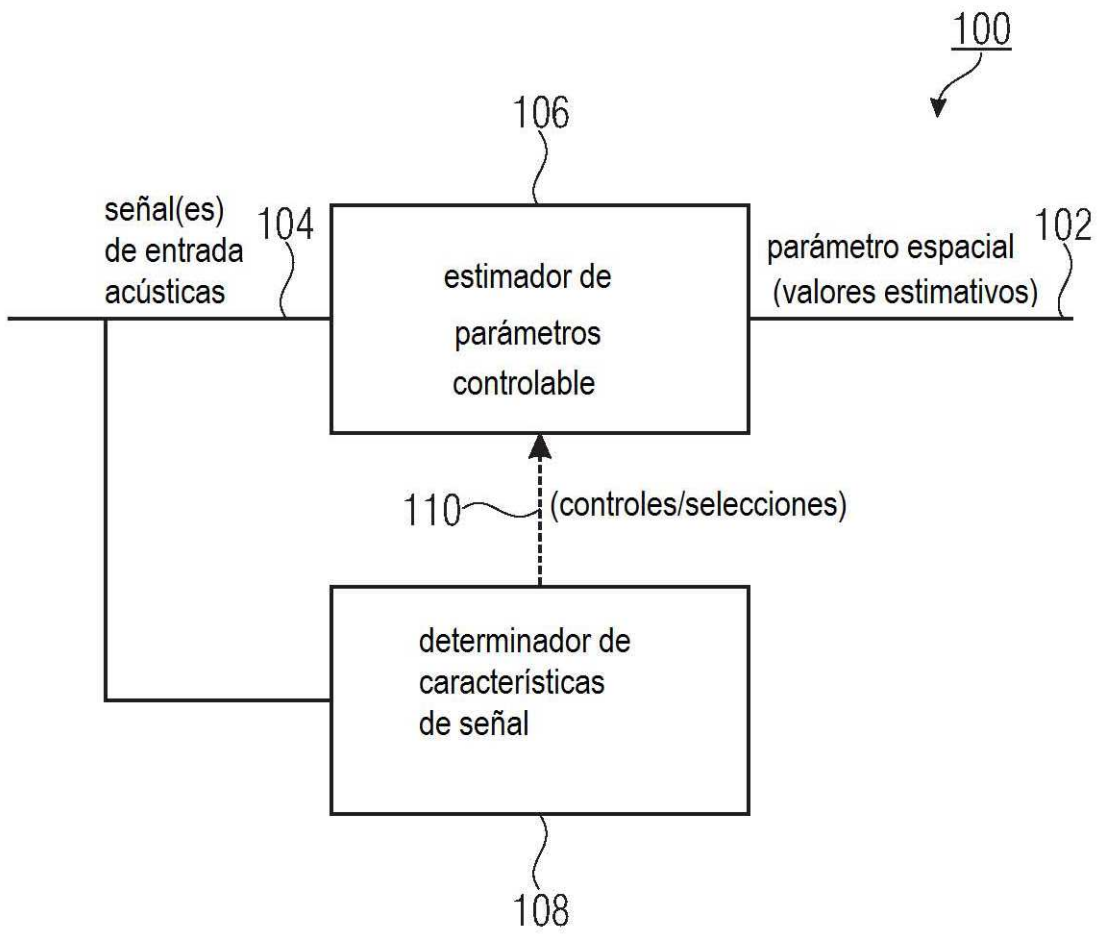


FIG 1

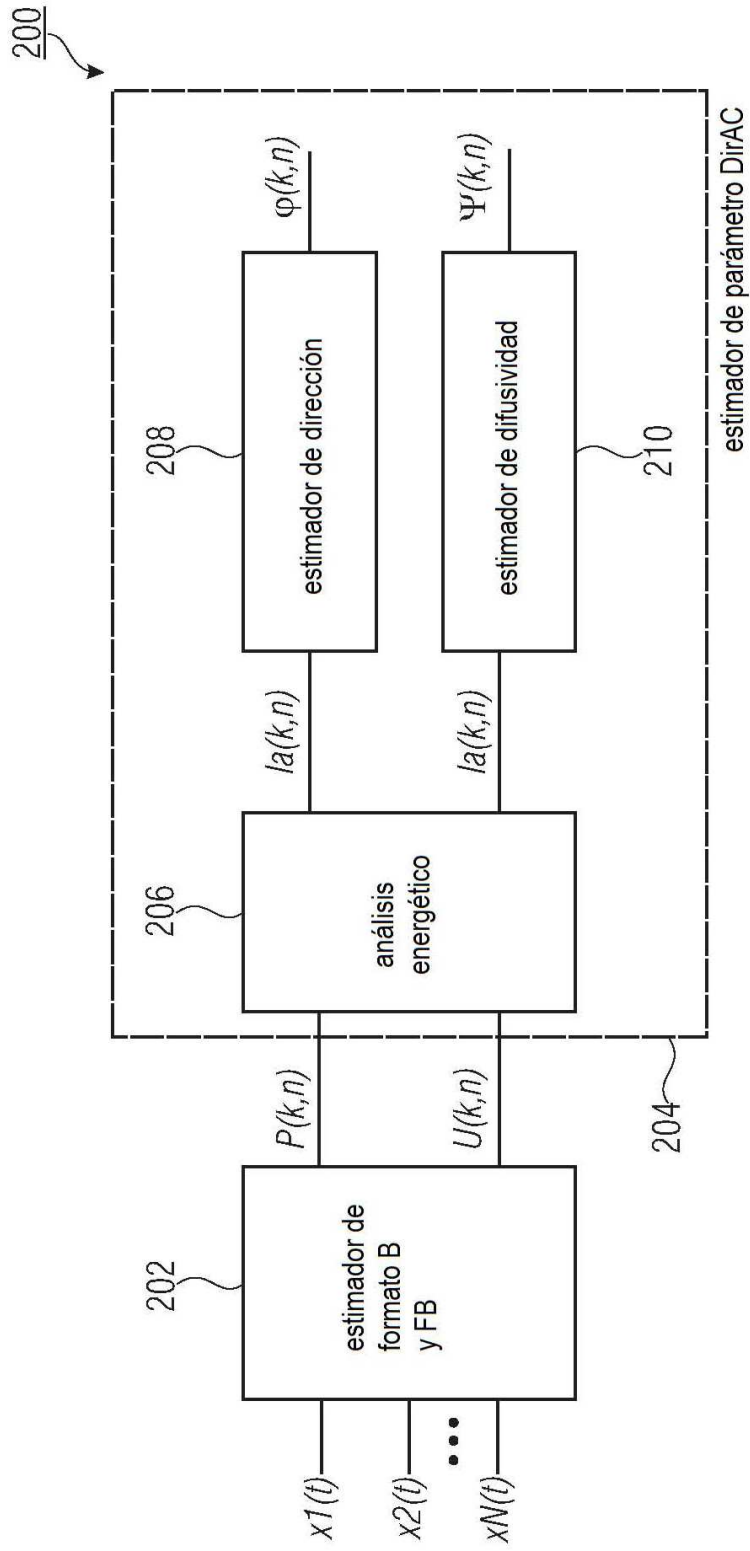


FIG 2

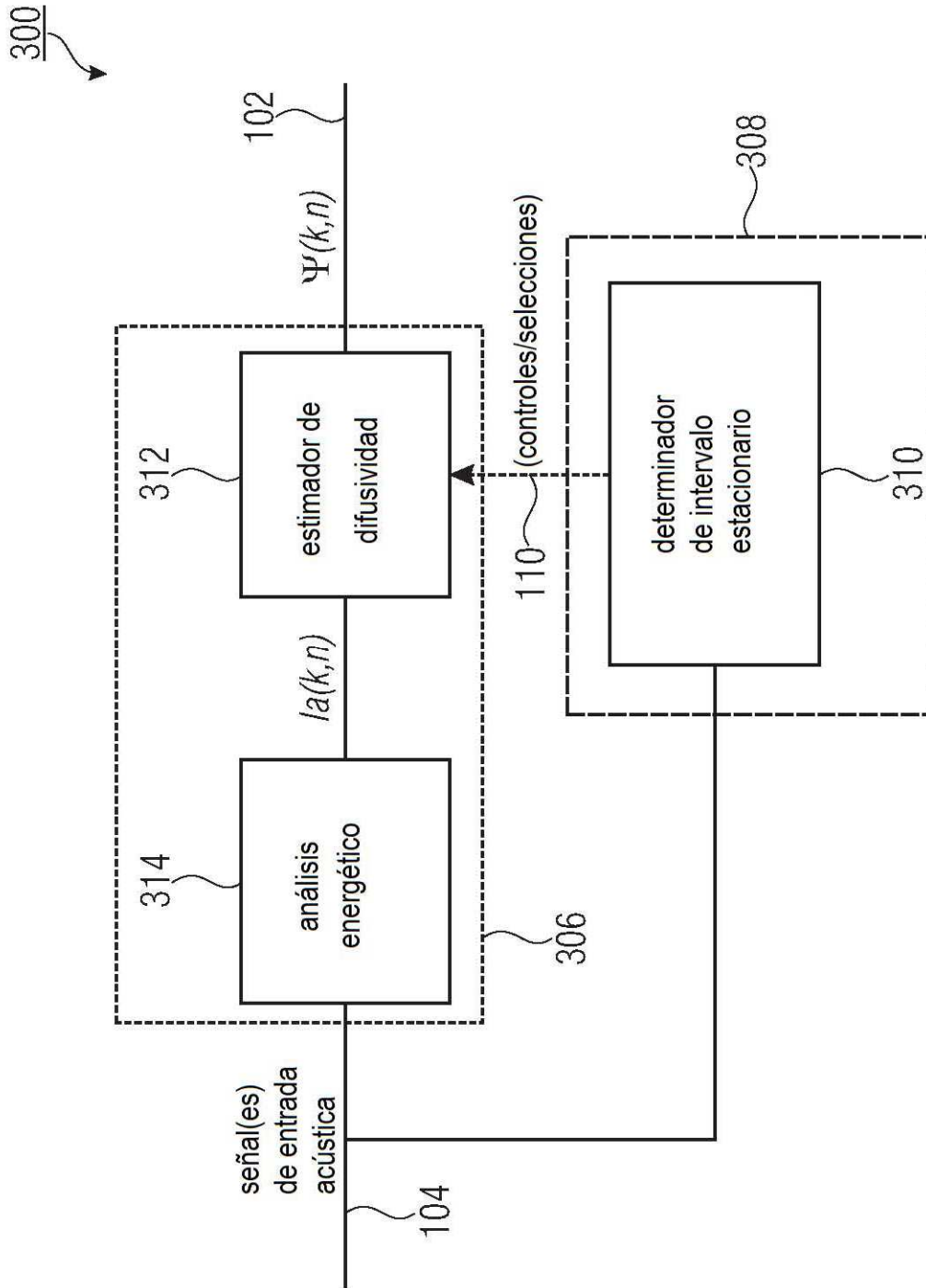


FIG 3

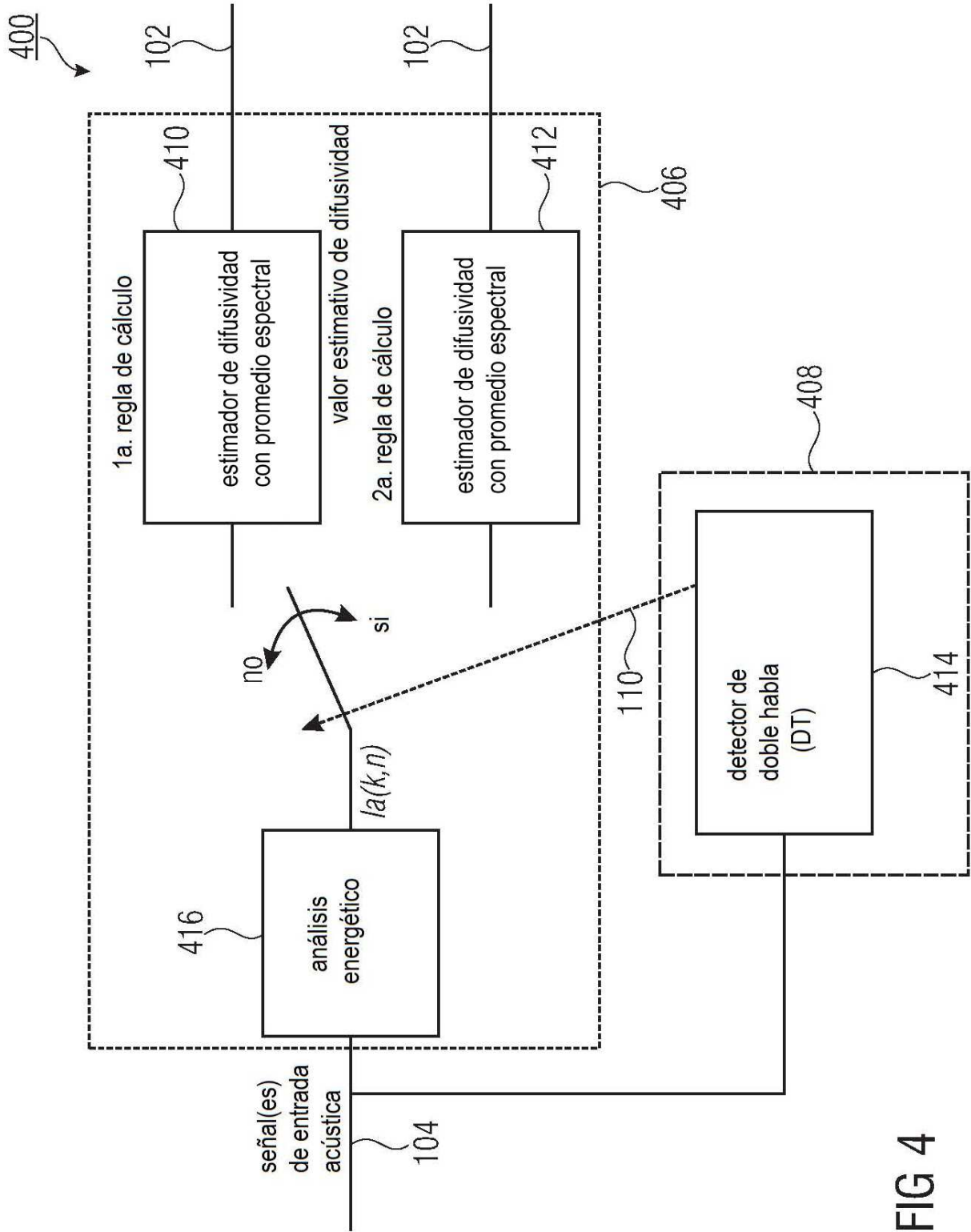


FIG 4

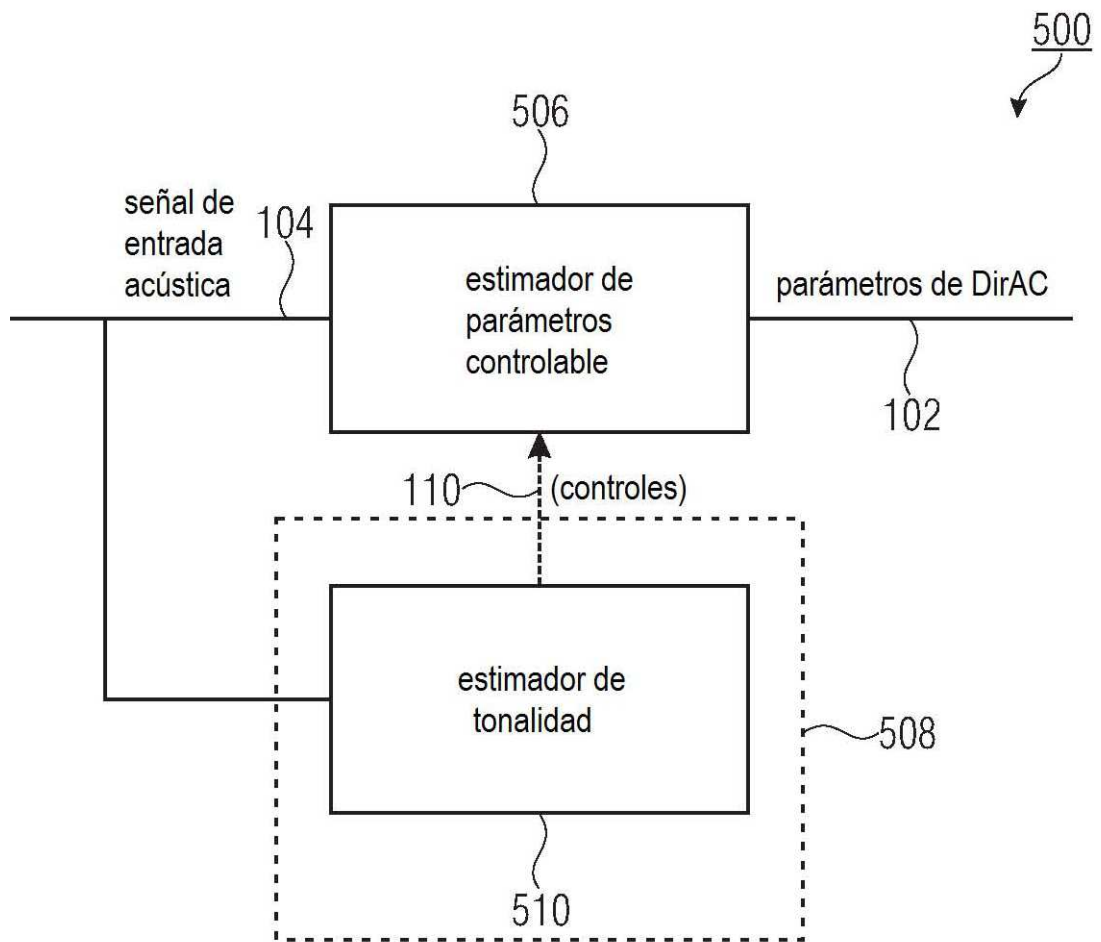


FIG 5

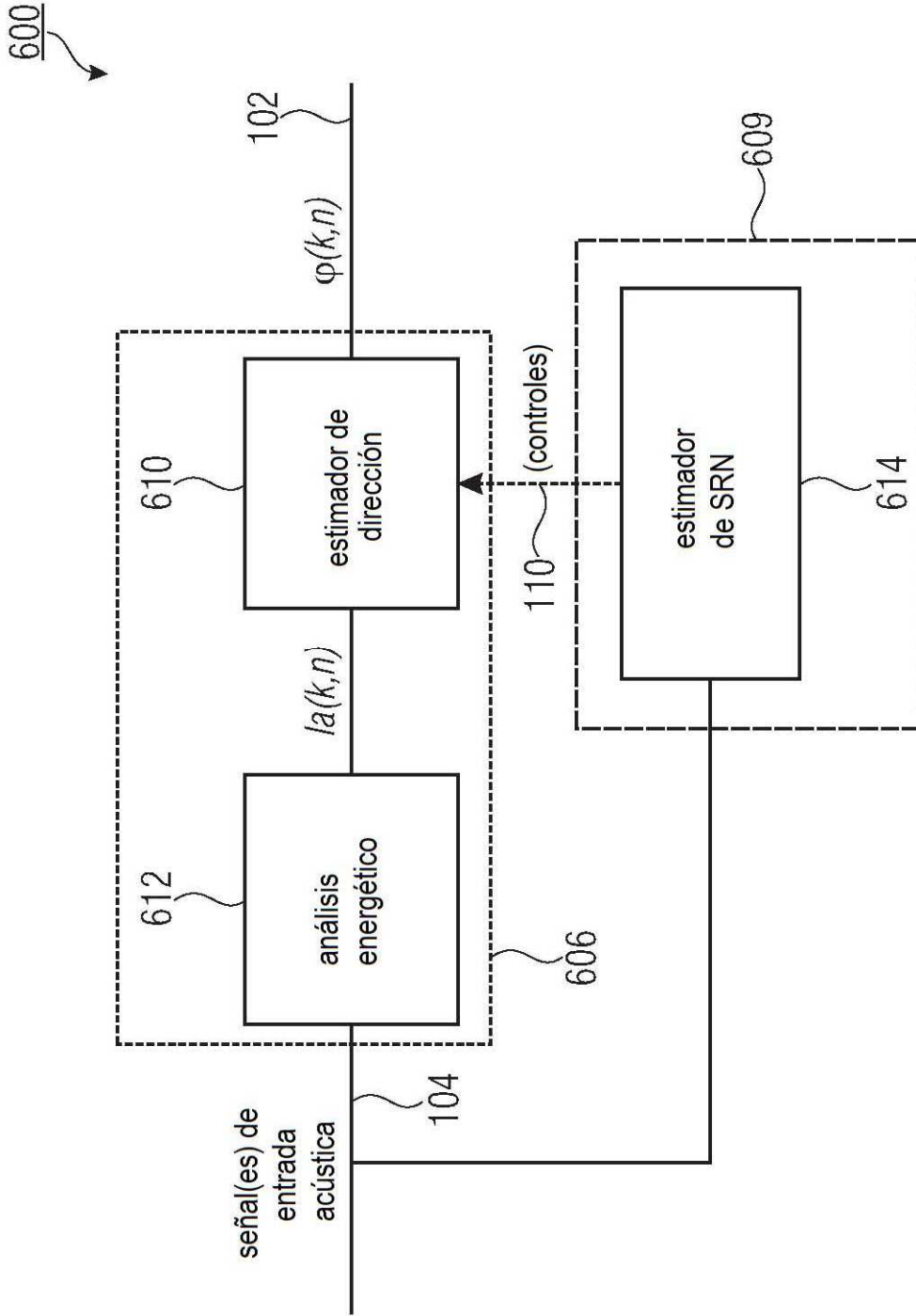


FIG 6

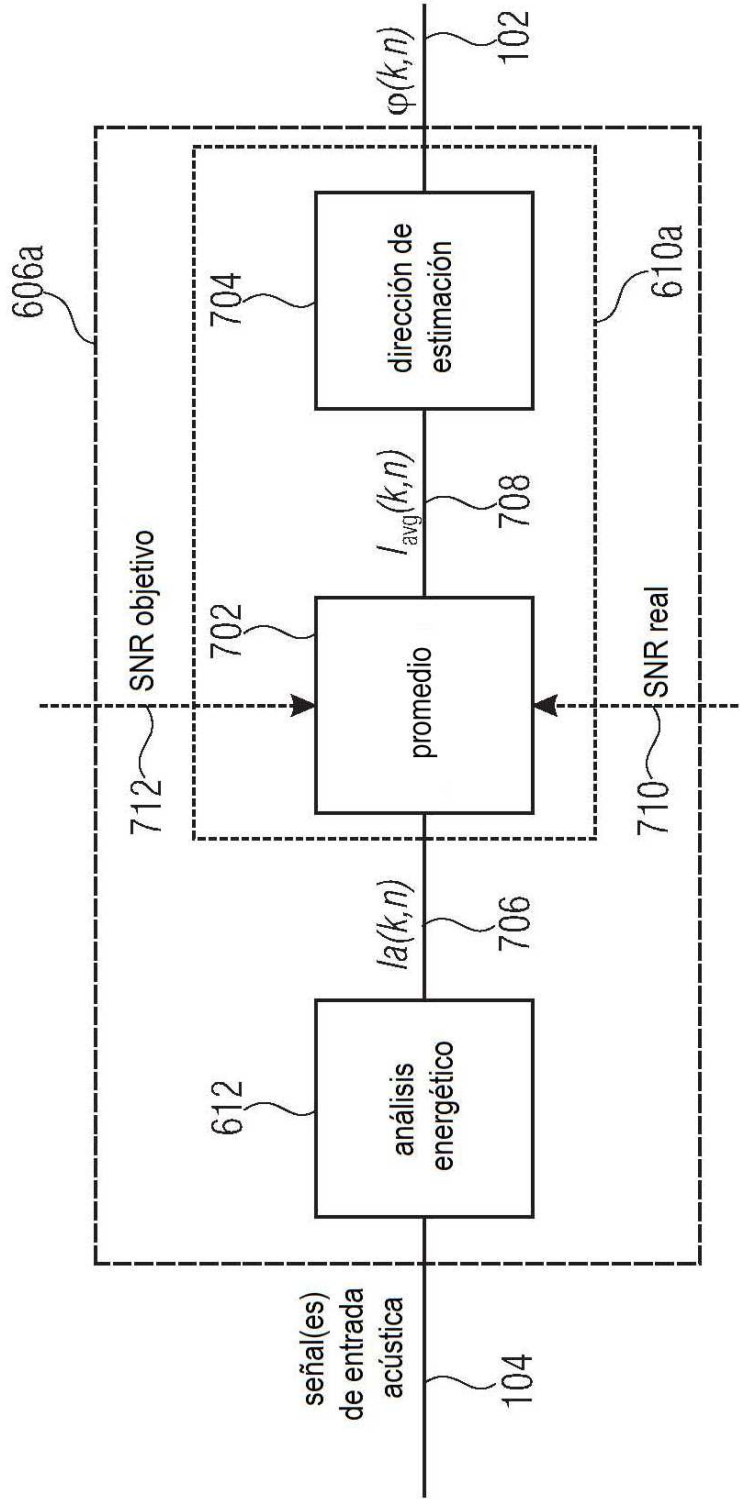


FIG 7A

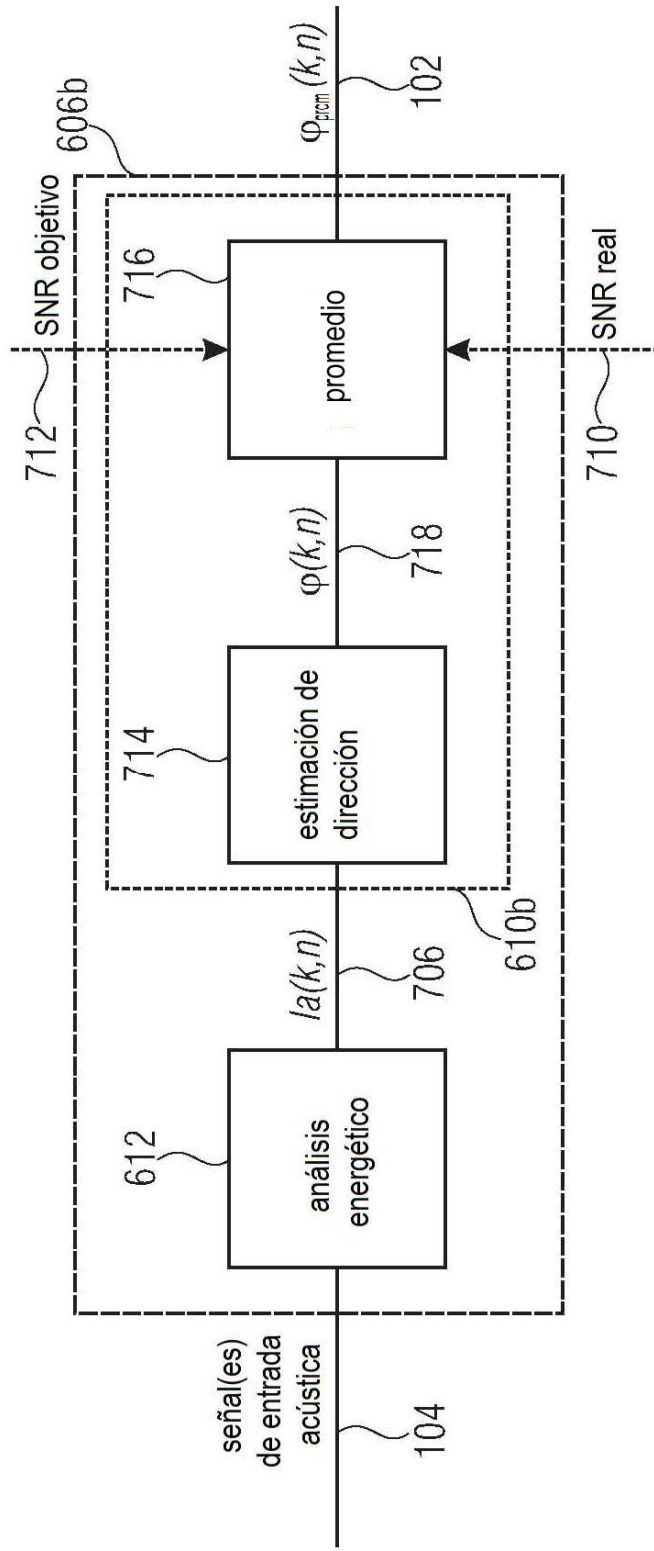
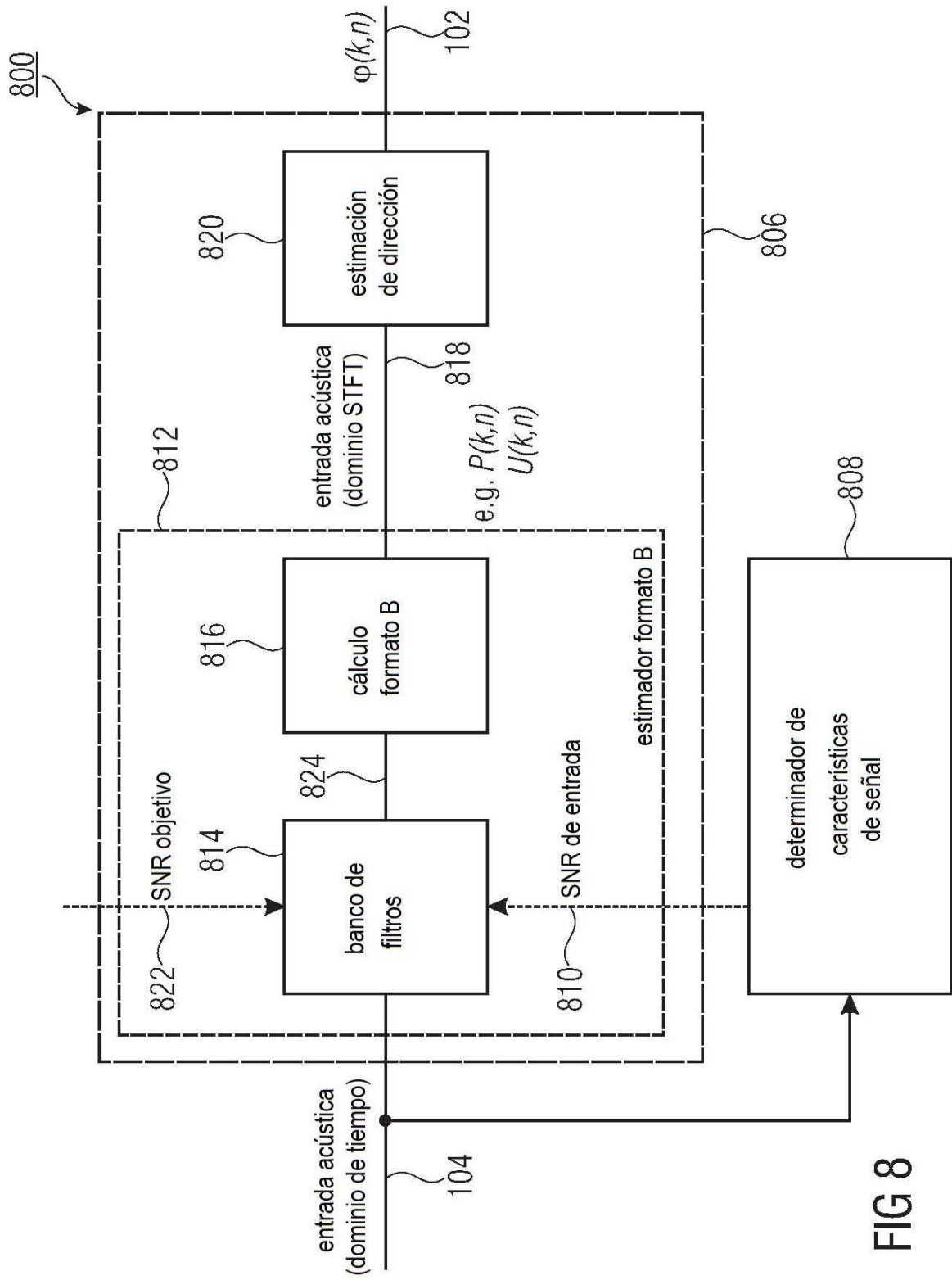


FIG 7B



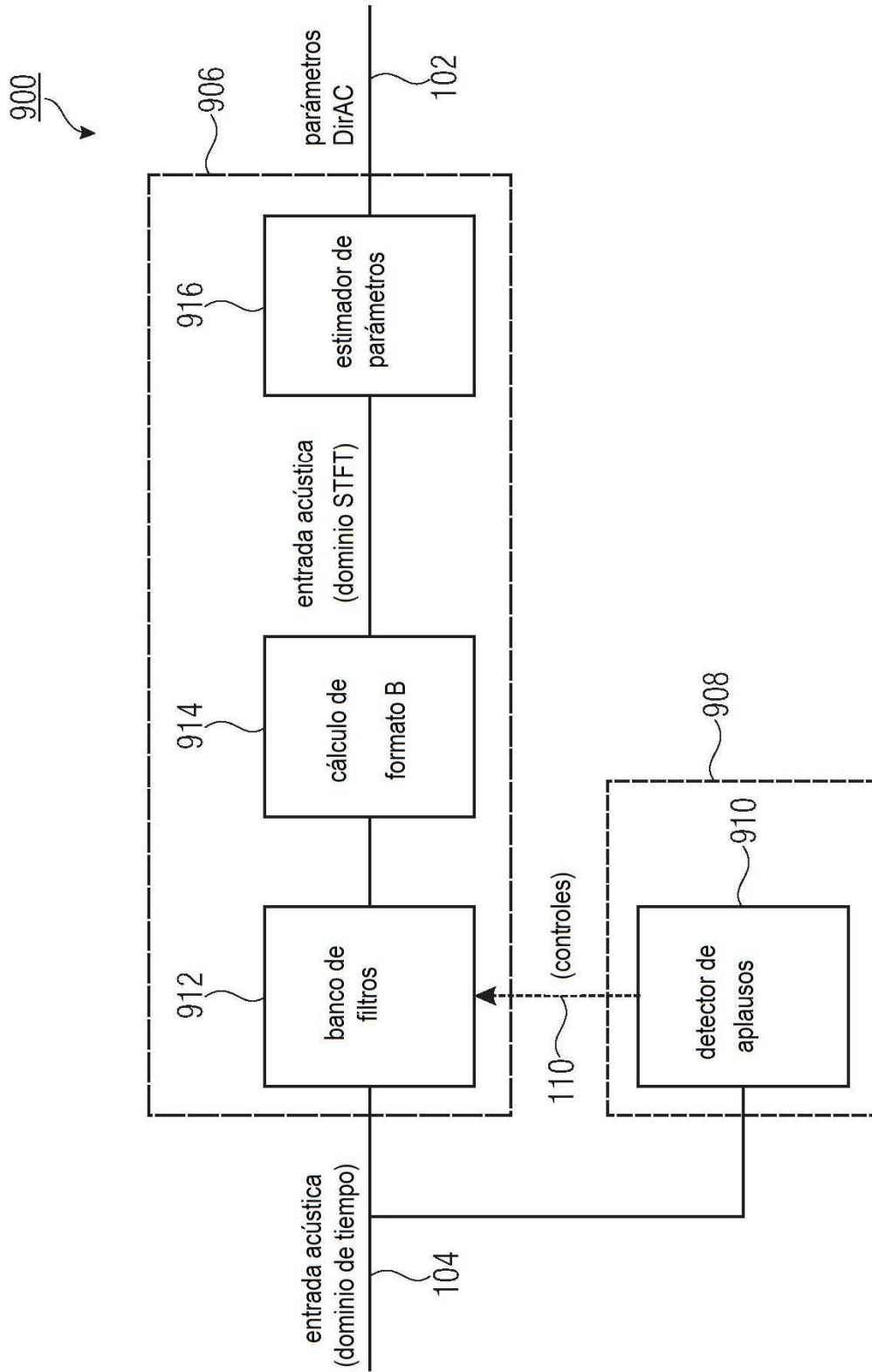


FIG 9

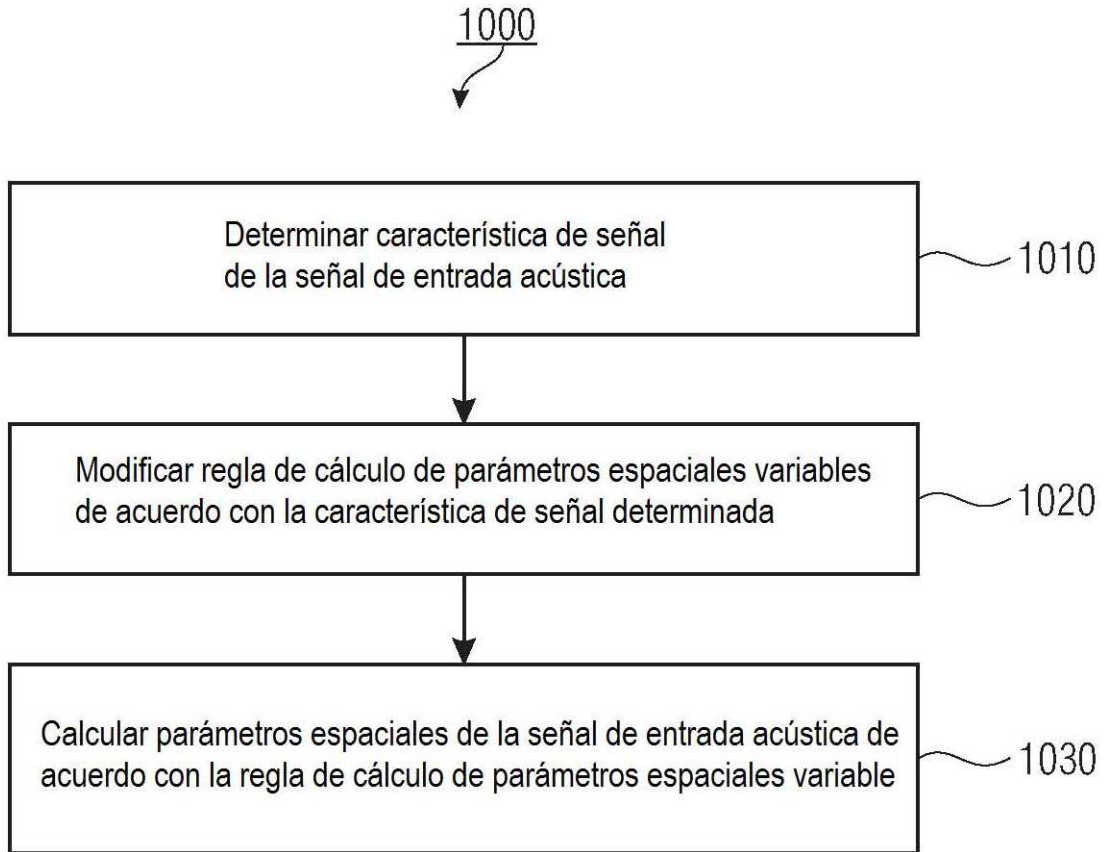


FIG 10