

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 605 248**

51 Int. Cl.:

**G10L 19/008** (2013.01)

**G10L 21/02** (2013.01)

**H04R 5/00** (2006.01)

**G10L 19/26** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **15.02.2011 PCT/EP2011/052246**

87 Fecha y número de publicación internacional: **01.09.2011 WO11104146**

96 Fecha de presentación y número de la solicitud europea: **15.02.2011 E 11703882 (8)**

97 Fecha y número de publicación de la concesión europea: **24.08.2016 EP 2539889**

54 Título: **Aparato para generar señal de mezcla descendente mejorada, método para generar señal de mezcla descendente mejorada y programa de ordenador**

30 Prioridad:

**24.02.2010 US 307553 P**

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

**13.03.2017**

73 Titular/es:

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

72 Inventor/es:

**KÜCH, FABIAN;  
HERRE, JÜRGEN;  
FALLER, CHRISTOF y  
TOURNERY, CHRISTOPHE**

74 Agente/Representante:

**PONTI SALES, Adelaida**

**ES 2 605 248 T3**

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

## DESCRIPCIÓN

Aparato para generar señal de mezcla descendente mejorada, método para generar señal de mezcla descendente mejorada y programa de ordenador

5

Las realizaciones de acuerdo con la invención son concernientes con un aparato para generar una señal de mezcla descendente mejorada con un método para generar una señal de mezcla descendente mejorada y con un programa de ordenador para generar una señal de mezcla descendente mejorada.

10

Una realización de acuerdo con la invención es concerniente con el cálculo de mezcla descendente mejorada para micrófonos de audio espaciales.

Antecedentes de la invención

15

La grabación del sonido envolvente ("surround") con una configuración de micrófono pequeño sigue siendo un reto. Una de tal configuración más ampliamente conocida es un micrófono de campo de sonido y decodificadores envolvente correspondientes (véase, por ejemplo referencia [3]), que filtran y combinan sus cuatro señales de cápsula de micrófono casi coincidentes para generar los canales de salida de sonido envolvente. En tanto que la alta fidelidad de señal de un solo canal es mantenida, la debilidad de este procedimiento es su separación de canal limitada concerniente con la directividad limitada de respuestas direccionales de micrófono de primer orden.

20

Alternativamente, se pueden aplicar técnicas basadas en una representación paramétrica del campo de sonido observado. En la referencia [2], se ha propuesto un método utilizando pares de micrófono estéreo coincidentes convencionales para grabar sonido envolvente. Se demostró como estimar las proporciones de sonido directo a sonido difuso de los parámetros de pista espacial y direcciones de llegada del sonido a partir de estas señales de micrófono direccionales y como aplicar esta información para conducir una síntesis de codificación de audio espacial para generar sonidos envolvente. En la referencia [2] también se ha discutido como la información paramétrica, esto es dirección de llegada (DOA) del sonido y la proporción de sonido difuso (DSR) del campo de sonido puede ser usada para calcular directamente los parámetros espaciales específicos que son usados en el esquema de codificación envolvente de MPEG (MPS) (véase, por ejemplo referencia [6]).

25

30

El envolvente de MPEG es una representación paramétrica de señales de audio de multicanal que representa un procedimiento eficiente a la codificación de audio espacial de alta calidad. MPS aprovecha el hecho de que, desde un punto de vista perceptual, las señales de audio de multicanal contienen redundancia significativa con respecto a los diferentes canales de altavoz. El codificador de MPS toma múltiples señales de altavoz como entrada, con la configuración espacial correspondiente de los altavoces que tiene que ser conocida de antemano. En base a estas señales de entrada, el codificador de MPS calcula los parámetros espaciales en sub-bandas de frecuencia, tales como diferencias de nivel de canal (CLD) entre dos canales y correlación de intercanal (ICC) entre dos canales. Luego, la información lateral de MPS real es derivada a partir de estos parámetros espaciales. Además, el codificador calcula una señal de mezcla descendente, que podría consistir de uno o más canales de audio.

35

40

Se ha encontrado que las señales de entrada de micrófono estero son apropiadas para estimar los parámetros de pista espacial. Sin embargo, también se ha encontrado que la señal de entrada de micrófono esterero sin procesar es en general no apropiada para ser usada directamente como la señal de mezcla descendente envolvente de MPEG correspondiente. Se ha encontrado que en muchos casos, la diafonía entre canales izquierdo y derecho es demasiado alta, dando como resultado una separación de canal escasa en las señales decodificadas envolvente de MPEG.

45

En vista de esta situación, hay la necesidad de un concepto para generar una señal de mezcla descendente mejorada en base a una señal de micrófono de multicanal, de tal manera que las señales de mezcla descendente mejoradas conducen a una calidad de audio espacial suficientemente buena y propiedad de localización después de la decodificación envolvente de MPEG.

50

Un ejemplo más de un sistema de síntesis conocido se describe en la referencia [9].

55

Breve descripción de la invención

Este objetivo es obtenido por el aparato reivindicado para generar una señal de mezcla descendente mejorada, mediante el método reivindicado para generar una señal de mezcla descendente mejora y por el programa de ordenador reivindicado para mejorar una señal de mezcla descendente mejorada.

60

Una realización de acuerdo con la invención crea un aparato para generar una señal de mezcla descendente mejorada en base a una señal de micrófono de multicanal. El aparato comprende un analizador espacial configurado para calcular un conjunto de parámetros de pista espacial que comprenden información de dirección que describe la dirección de llegada de sonido directo, información de potencia de sonido directo e información de potencia de sonido difuso en base a la señal de micrófono de multicanal. El aparato también comprende un calculador de filtro para calcular parámetros de filtro mejorados en dependencia de la información de dirección que describe la dirección de llegada de sonido directo, en dependencia de la información de potencia de sonido directo y en dependencia de la información de potencia de sonido difuso. El aparato también comprende un filtro para filtrar la señal de micrófono o una señal derivada de la misma, utilizando los parámetros de filtro mejorados, para obtener la señal de mezcla descendente mejorada.

Esta realización de acuerdo con la invención está basada en el hallazgo de que una mezcla de señal descendente mejorada, que es apropiada más que la señal de micrófono de multicanal de entrada, puede ser derivada de la señal de micrófono de multicanal de entrada por una operación de filtración y que los parámetros de filtro para tal operación de filtración mejorada de señal puede ser derivada eficientemente de los parámetros de pista espacial.

Así, es posible reutilizar la misma información, es decir los parámetros de pista espacial, que es también apropiada para la derivación de los parámetros de envolvente de MPEG, para el cálculo de los parámetros de filtro mejorados. Así, un sistema altamente eficiente puede ser creado utilizando el concepto descrito anteriormente.

Además, es posible derivar una señal de mezcla descendente que permite una buena separación de canal cuando es procesada en un descodificador envolvente de MPEG aun si las señales de canal de la señal de micrófono de multicanal solamente comprenden una baja separación espacial. Así, la señal de mezcla descendente mejorada puede conducir a una calidad de audio espacial significativamente mejorada y propiedad de localización después de la descodificación envolvente de MPEG en comparación con los sistemas convencionales.

Para resumir, la realización descrita anteriormente de acuerdo con la invención permite proveer una señal de mezcla descendente mejorada que tiene buenas propiedades de separación espacial en un esfuerzo computacional moderado.

En una realización preferida, el calculador de filtro está calculado para calcular los parámetros de filtro mejorados, de tal manera que la señal de mezcla descendente mejorada se aproxima a una señal de mezcla descendente deseada. Utilizando este procedimiento, se puede asegurar que los parámetros de filtro mejorados estén bien adaptados a un resultado deseado del filtrado. Por ejemplo, parámetros de filtro mejorados pueden ser calculados de tal manera que una o más propiedades estadísticas de la señal de mezcla descendente mejorada se aproximan a las propiedades estadísticas deseadas de la señal de mezcla descendente. Así, se puede alcanzar que la señal de mezcla descendente mejorada es bien adaptada a las expectativas, en donde las expectativas pueden ser definidas numéricamente en términos de valores de correlación deseados.

En una realización preferida, el calculador de filtro está configurado para calcular valores de correlación deseados entre la señal de micrófono de multicanal (o más precisamente, señales de canal del mismo) y señales de canal deseadas de la señal de mezcla descendente en dependencia de los parámetros de pista espacial. En este caso, el calculador de filtro está configurado preferiblemente para calcular los parámetros de filtro mejorados en dependencia de los valores de correlación cruzada deseados. Se ha encontrado que dichos valores de correlación cruzada son una buena medida de si las señales de canal de la señal de mezcla descendente exhiben características de separación de canal suficientemente buenas. También, se ha encontrado que los valores de correlación deseados pueden ser calculados con esfuerzo computacional moderado en base a los parámetros de pista espacial.

En una realización preferida, el calculador de filtro está configurado para calcular los valores de correlación cruzada deseados en dependencia de los factores de ganancia dependientes de la dirección, que describen contribuciones deseadas de un componente de sonido directo de un componente de sonido directo de la señal de micrófono de multicanal a una pluralidad de señales de altavoz y en dependencia de uno o más valores de matriz de mezcla descendente que describen contribuciones deseadas de una pluralidad de canales de audio (por ejemplo, señales de altavoz) a uno o más canales de la señal de mezcla descendente mejoradas. Se ha encontrado que tanto los factores de ganancia dependientes de la dirección como los valores de matriz de mezcla descendente son apropiados para calcular los valores de correlación cruzada deseados y que dichos factores de ganancia dependientes de la dirección y dichos valores de matriz de mezcla descendente son fácilmente obtenibles. Además, se ha encontrado que los valores de correlación cruzada deseados son fácilmente obtenibles en base a dicha información.

En una realización preferida, el calculador de filtro está configurado para mapear la información de dirección sobre un conjunto de factores de ganancia dependientes de la dirección. Se ha encontrado que se puede usar una ley de toma panorámica de amplitud de multicanal para determinar los factores de ganancia con el esfuerzo moderado en dependencia de la información de dirección. Se ha encontrado que la información de dirección de llegada es apropiada para determinar

los factores de ganancia dependientes de la dirección, que pueden describir, por ejemplo cuales altavoces deben presentar el componente de sonido directo. Es fácilmente comprensible que el componente de sonido directo este distribuido a diferentes señales de altavoz en dependencia de la información de dirección de llegada (designada brevemente como información de dirección) y que es relativamente simple determinar los factores de ganancia que describen cuál de los altavoces deben presentar el componente de sonido directo. Por ejemplo, la regla de mapeo que es usada para mapear la información de dirección sobre conjunto de factores de ganancia dependientes de la dirección, puede simplemente determinar que aquellos altavoces que están asociados con la dirección de llegada podrían presentar (o principalmente presentar) el componente de sonido directo, mientras que los otros altavoces que están asociados con otras direcciones deben solamente presentar una pequeña porción del componente de sonido directo o deben aun suprimir el componente de sonido directo.

En una realización preferida, el calculador de filtro está configurado para considerar la información de potencia de sonido directo y la información de potencia de sonido difuso para calcular los valores de correlación cruzada deseados. Se ha encontrado que la consideración de las potencias de ambos de dichos componentes de sonido (componente de sonido directo y componente de sonido difuso) da como resultado una impresión de audición particularmente buena, debido a que tanto el componente de sonido directo como el componente de sonido difuso pueden ser asignados apropiadamente a las señales de canal de la señal de mezcla descendente (comúnmente de multicanal).

En una realización preferida, el calculador de filtro está configurado para ponderar la información de potencia de sonido directo en dependencia de la información de dirección y para aplicar una ponderación predeterminada que es independiente de la información de dirección a la información de potencia de sonido difuso con el fin de calcular los valores de correlación cruzada deseados. Así, se puede distinguir entre los componentes de sonido directo y los componentes de sonido difuso lo que da como resultado una estimación particularmente real de los valores de correlación cruzada deseados.

En una realización preferida, el calculador de filtro está configurado para evaluar una ecuación de Wiener-Hopf para derivar los parámetros de filtro mejorados. En este caso, la ecuación de Wiener-Hopf describe una relación entre valores de correlación que describen una correlación entre diferentes pares de canal de la señal de micrófono de multicanal, parámetros de filtro mejorados y valores de correlación cruzada deseados entre señales de canal desde las señales de micrófono de multicanal y señales de canal deseados de la señal de mezcla descendente. Se ha encontrado que la evaluación de tal ecuación de Wiener-Hopf da como resultados parámetros de filtro mejorados que son bien adaptados a las características de correlación deseadas de las señales de canal de la señal de mezcla descendente.

En una realización preferida, el calculador de filtro está configurado para calcular los parámetros de filtro mejorados en dependencia de un modelo de canales de mezcla descendente deseados. Al modelar los canales de mezcla descendente deseados, los parámetros de filtro mejorados pueden ser calculados de tal manera que producen una señal de mezcla descendente que permite una buena reconstrucción de las señales de altavoz de multicanal deseadas en un decodificador de multicanal.

En algunas realizaciones, el modelo de los canales de mezcla descendente deseados puede comprender un modelo de una mezcla descendente ideal, que sería efectuada si las señales de canal (por ejemplo, señales de altavoz) estuvieran disponibles individualmente. Además, el modelado puede incluir un modelo de como las señales de canal individual podrían ser obtenidas a partir de la señal de micrófono de multicanal, aun si la señal de micrófono de multicanal comprende señales de canal que tienen solamente una separación espacial limitada. Así, un modelo global de los canales de mezcla descendente deseados puede ser obtenido, por ejemplo al combinar un modelado de cómo obtener señales de canal individuales (por ejemplo, señales de altavoz) y como derivar canales de mezcla descendente deseados a partir de dichas señales de canal individuales. Así, es una referencia suficientemente buena para el cálculo de los parámetros de filtro mejorados obtenibles con un esfuerzo computacional relativamente pequeño.

En una realización preferida, el calculador de filtro está configurado para efectuar selectivamente un filtrado de un solo canal, en el cual un primer canal de la señal de mezcla descendente es derivado mediante un filtrado de un primer canal de la señal de micrófono de multicanal y en el cual un segundo canal la señal de mezcla descendente es derivado mediante un filtrado de un segundo canal de la señal de micrófono de multicanal mientras que evita la diafonía del primer canal de la señal de micrófono de multicanal al segundo canal de señal de mezcla descendente y del segundo canal de la señal de micrófono de multicanal al primer canal de la señal de mezcla descendente o un filtrado de dos canales, en el cual un primer canal de la señal de mezcla descendente es derivado mediante filtración de un primero y un segundo canal de la señal de micrófono de multicanal y en el cual un segundo canal de la señal de mezcla descendente es derivada mediante filtración de un primero y un segundo canal de la señal de micrófono de multicanal. La selección de el filtrado de un solo canal y de el filtrado de dos canales se hace en dependencia de un valor de correlación que describe una correlación entre el primer canal de la señal de micrófono de multicanal y el segundo canal de la señal de micrófono de multicanal. Al seleccionar entre el filtrado de un solo canal y el filtrado de dos canales, se pueden evitar errores numéricos que pueden algunas veces aparecer si se usa filtración de dos canales en una situación en la cual el canal izquierdo y derecho están altamente

correlacionados. Así, una señal de mezcla descendente de buena calidad puede ser obtenida sin consideración de si las señales de canal de la señal micrófono de multicanal están altamente correlacionadas o no.

Otra realización de acuerdo con la invención crea un método para generar una señal de mezcla descendente mejorada.

Otra realización de acuerdo con la invención crea un programa de ordenador para efectuar dichos métodos para generar una señal de mezcla descendente mejorada.

El método y programas de ordenador están basados en los mismos hallazgos como el aparato y pueden ser complementados por cualquiera de los elementos y funcionalidades discutidas con respecto al aparato.

Breve descripción de las figuras

Realizaciones de acuerdo con la invención serán subsecuentemente descritas con referencia a las figuras adjuntas en las cuales:

La Figura 1 muestra un diagrama esquemático en bloques de un aparato para generar una señal de mezcla descendente mejorada de acuerdo con una realización de la invención;

la Figura 2 muestra una ilustración grafica del procesamiento de micrófono de audio espacial de acuerdo con una realización de la invención;

la Figura 3 muestra una ilustración grafica del cálculo de mezcla descendente mejorada de acuerdo con una realización de la invención;

la Figura 4 muestra una ilustración grafica del mapeo de canal para el cálculo de las señales de mezcla descendente deseada  $Y_1$  e  $Y_2$ , que pueden ser usadas en realizaciones de acuerdo con la invención;

la Figura 5 muestra una ilustración grafica de un cálculo de mezcla descendente mejorada en base a señales de micrófono pre procesadas de acuerdo con una realización de la invención;

la Figura 6 muestra una representación esquemática de cálculos para derivar los parámetros de filtro mejorados a partir de la señal de micrófono de multicanal de acuerdo con una realización de la invención y

la Figura 7 muestra una representación esquemática de cálculos para derivar los parámetros de filtro mejorados a partir de la señal de micrófono de multicanal de acuerdo con otra realización de la invención.

Descripción detallada de las realizaciones

1. Aparato para generar una señal de mezcla descendente mejorada de acuerdo con la Figura 1

La Figura 1 muestra un diagrama de bloque esquemático de un aparato para generar una señal de mezcla descendente mejorada en base a una señal de micrófono de multicanal. El aparato 100 está configurada para recibir una señal de micrófono de multicanal 100 y para proveer, en base a la misma, una señal de mezcla descendente mejorada 112. El aparato 100 comprende un analizador espacial 120 configurado para calcular un conjunto de parámetros de pista espaciales 122 en base a la señal de micrófono de multicanal 110. Los parámetros de pista espacial comprenden comúnmente una información de dirección que describe una dirección de llegada de sonido directo (tal sonido directo está incluido en la señal de micrófono de multicanal), información de potencia de sonido directo e información de potencia de sonido difuso. El aparato 100 también comprende un calculador de filtro 130 para calcular parámetros de filtro mejorados 132 en dependencia de los parámetros de pista espacial 122, esto es, en dependencia de la información de dirección que describe la dirección de llegada de sonido directo, en dependencia de la información de potencia de sonido directo y en dependencia de la información de potencia de sonido difuso. El aparato 100 también comprende un filtro 140 para filtrar la señal de micrófono 110 o una señal 110' derivada de la misma, utilizando los parámetros de filtro mejorados 132, para obtener la señal de mezcla descendente 112. La señal 110' puede opcionalmente ser derivada de la señal de micrófono de multicanal 110 utilizando un pre-procesamiento opcional 150.

Con respecto a la funcionalidad del aparato 100, se puede notar que la señal de mezcla descendente mejorada 112 es provista comúnmente de tal manera que la señal de mezcla descendente mejorada 112 permite una calidad de audio espacial mejorada después de la descodificación envolvente de MPEG cuando se compara con la señal de micrófono multicanal 110, debido a que los parámetros de filtro mejorados 132, son provistos comúnmente por el calculador de filtro 130 con el fin de obtener este objetivo. La provisión de los parámetros de filtro mejorados 130 está basada en los

parámetros de pista espacial 122 provistos por el analizador espacial, de tal manera que los parámetros de filtro mejorados 130 son provistos de acuerdo con una característica espacial de la señal de micrófono del multicanal 110 y con el fin de enfatizar la característica espacial de la señal de micrófono del multicanal 110. Así, el filtrado efectuado por el filtro 140 permite una mejora señal-adaptable de la característica espacial de la señal de mezcla descendente mejorada 112 cuando se compara con la señal de micrófono del multicanal de entrada 110.

Detalles con respecto al análisis espacial efectuado por el administrador espacial 120 con respecto al cálculo de parámetros de filtro efectuados por el calculador de filtro 130 y con respecto a el filtrado efectuada por el filtro 140 serán subsecuentemente descritos en más detalle.

2. Aparato para generar una señal de mezcla descendente mejorada de acuerdo con la Figura 2

La Figura 2 muestra un diagrama de bloque esquemático de un aparato 200 para generar una señal de mezcla descendente mejorada (que puede tomar la forma de una señal de audio de dos canales) y un conjunto de pistas espaciales asociadas con una señal de mezcla ascendente que tiene más de dos canales. El aparato 200 comprende una disposición de micrófono 205 configurado para proveer una señal de micrófono de dos canales que comprende una señal de primer canal 210a y una señal del segundo canal 210b.

El aparato 200 comprende además un procesador 210 para proveer un conjunto de pistas espaciales asociadas con una señal de mezcla ascendente que tiene más de dos canales en base a la señal de micrófono de dos canales. El procesador 216 también está configurado para proveer parámetros de filtro mejorados 232. El procesador 216 está configurado para recibir señales de entrada, la señal del primer canal 210a y la señal del segundo canal 210b proporcionada por la disposición de micrófono 205. El aparato 216 está configurado para proveer los parámetros de filtro mejorados 232 y para también proveer una información de pista espacial 262. EL aparato 200 comprende además un proveedor de señal de audio de dos canales 240 que está configurado para recibir la señal del primer canal 210a y la señal del segundo canal 210b provista por la disposición de micrófono 205 y para proveer versiones procesadas de la señal de micrófono del primer canal 210a y la señal de micrófono del segundo canal 210b como la señal de audio de dos canales 212 que comprende las señales de canal 212a, 212b.

La disposición de micrófono 205 comprende un primer micrófono direccional 206 y un segundo micrófono direccional 208. El primer micrófono direccional 206 y el segundo micrófono direccional 208 están preferiblemente espaciados por no más de 30 cm. Así, la señales recibidas por el primer micrófono direccional 206 y el segundo micrófono direccional 208 están correlacionadas fuertemente, lo que se ha encontrado que es benéfico para el cálculo de la información del componente de energía (o información de energía componente) 122a e información de dirección 122b por el analizador de señal 220. Sin embargo, el primer micrófono direccional 206 y el segundo micrófono direccional 208 están orientados de tal manera que una característica direccional 209 del segundo micrófono direccional 208 es una versión girada de una característica direccional 207 del primer micrófono direccional 206. Así, la señal de micrófono del primer canal 210a y la señal de micrófono del segundo canal 210b están fuertemente correlacionadas (debido a la proximidad espacial de los micrófonos 206, 208) y todavía diferentes (debido a las diferentes características direccional 207, 209 de los micrófonos direccionales 206, 208). En particular, una señal direccional incidente sobre la disposición de micrófono 205 desde una dirección aproximadamente constante provoca que componentes de señal fuertemente correlacionados de la señal de micrófono del primer canal 210a y la señal de micrófono del segundo canal 210b que tienen una proporción de amplitud dependiente de la dirección temporalmente constante (o proporción de intensidad). Una señal de audio ambiental incidente sobre la disposición de micrófono 205 desde direcciones que varían temporalmente provoca que los componentes de señal de la señal de micrófono del primer canal 210a y la señal de micrófono del segundo canal 210b que tienen una correlación significativa pero proporciones de amplitud que fluctúan temporalmente (o proporciones de intensidad). Así, la disposición de micrófono 205 provee una señal de micrófono de dos canales, 210a, 210b que permite que el analizador de señal 220 del procesador 216 distinga que entre sonidos directos y sonidos difusos aunque los micrófonos 206, 208 estén espaciados estrechamente. Así, el aparato 200 constituye un proveedor de señal de audio que puede ser implementado de forma espacialmente compacta y que es no obstante apto de proveer pistas espaciales asociadas con una señal de mezcla ascendente que tiene más de dos canales.

Las pistas espaciales 262 pueden ser usadas en combinación con la señal de audio de dos canales provista 212a, 212b por un descodificador de audio espacial para proveer una señal de salida de sonido envolvente.

En lo siguiente, se darán algunas explicaciones adicionales con respecto al aparato 200. El aparato 200 comprende opcionalmente una disposición de micrófono 205 que provee la señal del primer canal 210a y la señal del segundo canal 210b. La señal del primer canal 210a es también designada con  $x_1(t)$  y la señal del segundo canal 210b es también designada con  $x_2(t)$ . También se debe notar que la señal del primer canal 210a y la señal del segundo canal 210b pueden representar la señal de micrófono de multicanal 110, que es introducida al aparato 100 de acuerdo con la Figura 1.

El proveedor de señal de audio de dos canales 240 provee la señal del primer canal 210a y la señal del segundo canal 210b y comúnmente también recibe la información del parámetro de filtro mejorado 232. El proveedor de señal de audio de dos canales 240 puede por ejemplo efectuar la funcionalidad del pre-procesamiento opcional 150 y del filtro 240 para proveer la señal de audio de dos canales 212 que es representada por una señal de primer canal 212a y una señal del segundo canal 212b. La señal de audio de dos canales 212 puede ser equivalente a la señal de mezcla descendente mejorada 112 emitida por el aparato 100 de la Figura 1.

El analizador de señal 220 puede ser configurado para recibir la señal del primer canal 219a y la señal del segundo canal 210b. También, el administrador de señal 220 puede estar configurado para obtener información de energía componente 112a e información de dirección 122b en base a la señal de micrófono de dos canales 210 esto es, en base a la señal del primer canal 210a y la señal del segundo canal 210b. Preferiblemente, el analizador de señal 220 está configurado para obtener información de energía componente 122a y la información de dirección 122b de tal manera que la información de energía componente 122a describe valores estimativos de energía (o equivalentemente de potencia) de un componente de sonido directo de la señal de micrófono de dos canales y de un componente de sonido difuso de la señal de micrófono de dos canales y de tal manera que la información de dirección 122 describe un valor estimativo de la dirección desde la cual el componente de sonido directa de la señal de micrófono de dos canales 210a y 210b se origina. Así, el analizador de señal 220 puede tomar la funcionalidad del analizador espacial 120 y la información de energía componente 122a y la información de dirección 122b puede ser equivalente a los parámetros de pista espacial 122. La información de energía componente 122a puede ser equivalente a la información de potencia de sonido directa y la información de potencia de sonido difuso. El procesador 216 también comprende el generador de información lateral espacial 260 que recibe la información de energía componente 122a y la información de dirección 122b del analizador de señal 220. El generador de información lateral espacial 260 está configurado para proveer en base al mismo, la información de pista espacial 262. Preferiblemente, el generador de información lateral espacial 260 está configurado para mapear la información de energía componente 122a de la señal de micrófono de dos canales 210a, 210b y la información de dirección 122b de la señal de micrófono de dos canales 210a, 210b sobre la información de pista espacial 262. Así, la información lateral espacial 262 es obtenida de tal manera que la información de pista espacial 262 describe un conjunto de pistas espaciales asociadas con una señal de audio de mezcla ascendente que tiene más de dos canales.

El procesador 216 permite un cálculo computacionalmente muy eficiente de la información de pista espacial 262 que esta asociada con una señal de audio de mezcla ascendente que tiene más de dos canales en base a una señal de micrófono de dos canales 210a, 210b. El analizador de señal 220 es apto de extraer una gran cantidad de información a partir de la señal de micrófono de dos canales, es decir la información de energía componente 122a que describe tanto un valor estimativo de la energía de un componente de sonido directo como un valor estimativo de la energía de un componente de sonido difuso y la información de dirección 122b que describe un valor estimativo de la dirección desde la cual el componente de sonido directo de la señal de micrófono de dos canales se origina. Se ha encontrado que esta información que puede ser obtenida por el analizador de señal 220 en base a la señal de micrófono de dos canales 210a, 210b es suficiente para derivar la información de pista espacial 262 a un para una señal de audio de mezcla ascendente que tiene más de dos canales. Importantemente se ha encontrado que la información de energía componente 122a y la información de dirección 122b son suficientes para determinar directamente la información de pista espacial 262 sin usar realmente los canales de audio de mezcla ascendente como una cantidad intermedia.

Además, el procesador 216 comprende un calculador de filtro 230 que está configurado para recibir la información de energía componente 122a y la información de dirección 122b y para proveer, en base a las mismas, la información de parámetro de filtro mejorado 232. Así, el calculador de filtro 230 puede tomar la funcionalidad del calculador de filtro 130.

Para resumir lo anterior, el aparato 200 es apto de determinar eficientemente tanto la señal de mezcla descendente mejorada 212 como la información de pista espacial 262 de manera eficiente, utilizando la misma información intermedia 122a, 122b en ambos casos. También, se debe notar que el aparato 200 es apto de usar una disposición de micrófono espacialmente pequeño 205 con el fin de obtener tanto la señal de mezcla descendente (mejorada) 212 como la información de pista espacial 262. La señal de mezcla descendente 212 comprende una característica de separación espacial particularmente buena a pesar del uso de la disposición de micrófono pequeño 205 (que puede ser parte del aparato 200 o que puede ser externo al aparato 200 pero conectado al aparato 200) debido al cálculo de los parámetros de filtro mejorados 232 por el calculador de filtro 230. Así, la señal de mezcla descendente (mejorada) 212 puede ser apropiada para una presentación espacial (por ejemplo, utilizando un descodificador envolvente de MPEG) cuando es tomado en combinación con la información de pista espacial 262.

Para resumir, la Figura 2 muestra un diagrama de bloque esquemático de un procedimiento de micrófono de audio espacial. Como se puede ver, las señales de entrada de micrófono estéreo 210a (también designada con  $x_1(t)$ ) y 210b (también designada con  $x_2(t)$ ) son usadas en el bloque 216 para calcular el conjunto de información de pista espacial 262 asociada con la señal de mezcla ascendente de multicanal (por ejemplo, la señal de audio de dos canales 212). Además, se provee una señal de mezcla descendente de dos canales 212.

En las siguientes secciones, se requerirán las etapas requeridas para determinar la información de pista espacial 262 en base al análisis de las señales de micrófono estéreo. Así, se hará referencia a la presentación en referencia [2].

5 3. Análisis de señal estéreo

En lo siguiente, un análisis de señal estéreo será descrito que puede ser efectuado por el analizador espacial 120 o por el analizador de señal 220. Se debe notar que en algunas realizaciones en las cuales hay más de dos micrófonos usados y en las cuales hay más de dos señales de canal de una señal de micrófono de multicanal, se puede usar un análisis de señal mejorado.

10 El análisis de señal estéreo descrito en la presente puede ser usado para proveer los parámetros de pista espacial 122, que pueden tomar la forma de la información de energía del componente 122a e información de dirección 122b. Se debe notar que el análisis de señal estéreo puede ser efectuado en un dominio de tiempo-frecuencia. Así, las señales de canal 210a, 210b de la señal de micrófono de multicanal 110, 210 pueden ser transformadas a una representación de dominio de tiempo-frecuencia por el propósito del análisis adicional.

15 La representación de tiempo-frecuencia de las señales de micrófono son  $x_1(t)$  y  $x_2(t)$  y  $X_1(k, i)$  y  $X_2(k, i)$ , en donde  $k$  e  $i$  son índices de tiempo y frecuencia. Se supone que  $X_1(k, i)$  y  $X_2(k, i)$  pueden ser modelados como

$$\begin{aligned} X_1(k, i) &= S(k, i) + N_1(k, i) \\ X_2(k, i) &= a(k, i)S(k, i) + N_2(k, i), \end{aligned} \tag{1}$$

20 en donde  $a(k,i)$  es un factor de ganancia,  $S(k,i)$  es el sonido directo en el canal izquierdo y  $N_1(k,i)$  y  $N_2(k,i)$  representan sonido difuso.

25 La señal de mezcla descendente de codificación de audio espacial (SAC) 112, 212 e información lateral 262 son calculadas como función de  $E\{SS^*\}$ ,  $E\{N_1N_1^*\}$  y  $E\{N_2N_2^*\}$ , en donde  $E\{\cdot\}$  es una operación de promedio de tiempo corto y en donde  $*$  denota conjugado complejo. Estos valores son derivados en lo siguiente.

De (1) se sigue que

$$\begin{aligned} E\{X_1X_1^*\} &= E\{SS^*\} + E\{N_1N_1^*\} \\ E\{X_2X_2^*\} &= a^2E\{SS^*\} + E\{N_2N_2^*\} \\ E\{X_1X_2^*\} &= aE\{SS^*\} + E\{N_1N_2^*\}. \end{aligned} \tag{2}$$

40 Se debe notar aquí que  $E\{SS^*\}$  puede ser considerado como información de potencia de sonido directo o equivalentemente, información de energía de sonido directo y que  $E\{N_1N_1^*\}$  y  $E\{N_2N_2^*\}$  pueden ser considerados como información de potencia de sonido difuso o información de energía de sonido difuso.  $E\{SS^*\}$  y  $E\{N_1N_1^*\}$  pueden ser considerados como información de energía del componente.  $a$  puede ser considerada como información de dirección.

45 Se supone que la cantidad de sonido difuso en ambas señales de micrófono es la misma, esto es  $E\{N_1N_1^*\} = E\{N_2N_2^*\} = E\{NN^*\}$  y que el coeficiente de relación cruzada normalizado entre  $N_1$  y  $N_2$  es  $\Phi_{diff}$  esto es

$$\Phi_{diff} = \frac{E\{N_1N_2^*\}}{\sqrt{E\{N_1N_1^*\}E\{N_2N_2^*\}}}. \tag{3}$$

50  $\Phi_{diff}$  puede por ejemplo tomar un valor predeterminado o puede ser calculado de acuerdo con algún algoritmo.

Dadas estas suposiciones, (2) puede ser escrito como



$$\begin{aligned}
 E\{X_1 X_1^*\} &= E\{SS^*\} + E\{NN^*\} \\
 E\{X_2 X_2^*\} &= a^2 E\{SS^*\} + E\{NN^*\} \\
 E\{X_1 X_2^*\} &= a E\{SS^*\} + \Phi_{\text{diff}} E\{NN^*\}.
 \end{aligned}
 \tag{4}$$

La (2) produce la ecuación cuadrática

eliminación de  $\{SS^*\}$  y a en

$$AE\{NN^*\}^2 + BE\{NN^*\} + C = 0
 \tag{5}$$

Con

$$\begin{aligned}
 A &= 1 - \Phi_{\text{diff}}^2. \\
 B &= 2\Phi_{\text{diff}} E\{X_1 X_2^*\} - E\{X_1 X_1^*\} - E\{X_2 X_2^*\}. \\
 C &= E\{X_1 X_1^*\} E\{X_2 X_2^*\} - E\{X_1 X_2^*\}^2.
 \end{aligned}
 \tag{6}$$

Entonces  $\{NN^*\}$  es una de las dos soluciones de (5) la físicamente posible, esto es

$$E\{NN^*\} = \frac{-B - \sqrt{B^2 - 4AC}}{2A}.
 \tag{7}$$

La otra solución de (5) produce una potencia de salida difuso más grande que la potencia de señal del micrófono, lo que es físicamente imposible.

Dado (7), es fácil calcular a y  $E\{SS^*\}$ :

$$\begin{aligned}
 a &= \sqrt{\frac{E\{X_2 X_2^*\} - E\{NN^*\}}{E\{X_1 X_1^*\} - E\{NN^*\}}} \\
 E\{SS^*\} &= E\{X_1 X_1^*\} - E\{NN^*\} \\
 a^2 E\{SS^*\} &= E\{X_2 X_2^*\} - E\{NN^*\}.
 \end{aligned}
 \tag{8}$$

Como se discute en la referencia [2], la dirección de llegada de  $a(k,i)$  de sonido directo puede ser determinada como función de la proporción de amplitud estimada  $a(k,i)$

$$\alpha(k, i) = f(a(k, i)).
 \tag{9}$$

El mapeo específico depende de las características direccionales de micrófonos estéreo usados para la grabación del

sonido.

4. Generación de información lateral espacial

5 En lo siguiente, se describirá la generación de la información de pista espacial 262, que puede ser provista por el generador de información lateral espacial 260. Sin embargo, se debe notar que la generación de información lateral espacial en forma de la información de pista espacial 262 no es un elemento necesario de realizaciones de la presente invención. Así, se debe notar que la generación de la información lateral espacial puede ser omitida en algunas realizaciones. También, se debe notar que se pueden usar diferentes métodos para obtener la información de pista espacial 262 o cualquier otra información lateral espacial.

No obstante, también se debe notar que la generación de la información lateral espacial que es discutida en lo siguiente puede ser considerada como un concepto preferido para generar una información de pista espacial.

15 Dados los resultados de análisis de la señal estéreo 122a, 122b, esto es los parámetros  $\alpha$  de acuerdo con la ecuación (9)  $E\{SS^*}$  y  $E\{NN^*}$ , parámetros espaciales compatibles con el decodificador de SAC son generados, por ejemplo por el generador de información lateral espacial 260. Se ha encontrado que una manera eficiente de hacer esto es considerando un modelo de señal de multicanal. Como un ejemplo, se considera la configuración de altavoces como se muestra en la Figura 4 en lo siguiente, que implican:

20

$$\begin{aligned}
 L(k, i) &= g_1(k, i)\tilde{S}(k, i) + h_1(k, i)\tilde{N}_1(k, i) \\
 R(k, i) &= g_2(k, i)\tilde{S}(k, i) + h_2(k, i)\tilde{N}_2(k, i) \\
 C(k, i) &= g_3(k, i)\tilde{S}(k, i) + h_3(k, i)\tilde{N}_3(k, i) \\
 L_s(k, i) &= g_4(k, i)\tilde{S}(k, i) + h_4(k, i)\tilde{N}_4(k, i) \\
 R_s(k, i) &= g_5(k, i)\tilde{S}(k, i) + h_5(k, i)\tilde{N}_5(k, i),
 \end{aligned}
 \tag{10}$$

35 donde  $\tilde{S}(k, i)$  es la señal de sonido directo y  $\tilde{N}_1$  a  $\tilde{N}_5$  son señales difusas (independientes de inter-canal).  $\tilde{S}$  corresponde a la cantidad total ganancia-compensada de sonido directo en la señal de micrófono estéreo, esto es

$$\tilde{S}(k, i) = 10^{\frac{g(\alpha)}{20}} \sqrt{1 + a^2} S(k, i), \tag{11}$$

40 y la señales de sonido difuso,  $\tilde{N}_1$  a  $\tilde{N}_5$ , tienen todas la misma potencia igual a  $E\{NN^*}$ . Se debe notar que esta definición de potencia de sonido difuso es arbitraria, puesto que finalmente las ganancias  $h_1$  a  $h_5$  determinan la cantidad de sonido difuso.

Se debe notar que  $L(k, i)$ ,  $R(k, i)$ ,  $C(k, i)$ ,  $L_s(k, i)$  y  $R_s(k, i)$  pueden ser por ejemplo señales de canal deseados o señales de altavoz deseadas.

45

En una primera etapa, como función de la dirección de llegada de sonido directo  $\alpha(k, i)$ , se aplica una ley de toma panorámica de amplitud de multicanal (véanse, por ejemplo referencias [7] y [4]) para determinar los factores de ganancia  $g_1$  a  $g_5$ . Luego, se usa un procedimiento heurístico para determinar las ganancias de sonido difuso  $h_1$  a  $h_5$ . Los valores constantes  $h_1= 1.0$ ,  $h_2= 1.0$ ,  $h_3= 0$ ,  $h_4= 1.0$  y  $h_5= 1.0$  son una elección razonable, esto es, el ambiente es distribuido igualmente a la parte frontal y parte posterior, mientras que el canal central es generado por como una señal seca. Sin embargo, una elección diferente de  $h_1$  a  $h_5$  es posible.

50

El sonido directo de la parte lateral y parte posterior es atenuado en relación con el sonido que llega desde direcciones delanteras. El sonido directo contenido en las señales de micrófono es preferiblemente compensado en ganancia por un factor  $g(\alpha)$  que depende del patrón de directividad de los micrófonos.

55

Dado, el modelo de señal envolvente (10), el análisis de pista espacial del SAC específico usado es aplicado al modelo de señal para obtener las pistas espaciales para envolvente de MPEG.

5 Los espectros de potencia de las señales definidas en (10) son

$$\begin{aligned}
 P_L(k, i) &= g_1^2 E \{ \tilde{S} \tilde{S}^* \} + h_1^2 E \{ NN^* \} \\
 P_R(k, i) &= g_2^2 E \{ \tilde{S} \tilde{S}^* \} + h_2^2 E \{ NN^* \} \\
 P_C(k, i) &= g_3^2 E \{ \tilde{S} \tilde{S}^* \} + h_3^2 E \{ NN^* \} \\
 P_{L_s}(k, i) &= g_4^2 E \{ \tilde{S} \tilde{S}^* \} + h_4^2 E \{ NN^* \} \\
 P_{R_s}(k, i) &= g_5^2 E \{ \tilde{S} \tilde{S}^* \} + h_5^2 E \{ NN^* \} ,
 \end{aligned}
 \tag{12}$$

donde

$$E \{ \tilde{S} \tilde{S}^* \} = 10^{\frac{g(\alpha)}{10}} (1 + a^2) E \{ SS^* \} .
 \tag{13}$$

30 Los espectros cruzados, usados en lo siguiente son

$$\begin{aligned}
 P_{LL_s}(k, i) &= g_1 g_4 10^{\frac{g(\alpha)}{10}} (1 + a^2) E \{ SS^* \} \\
 P_{RR_s}(k, i) &= g_2 g_5 10^{\frac{g(\alpha)}{10}} (1 + a^2) E \{ SS^* \} .
 \end{aligned}
 \tag{14}$$

40 Envoltente de MPEG aplica una ganancia de -3dB ( $g_s 1/\sqrt{2}$ ) a los canales envolvente antes de procesarlos adicionalmente. Esto puede ser considerado para generar mezcla descendente e información lateral espacial.

El primer bloque de dos a uno (TTO) del envolvente MPEG utiliza diferencia de nivel de inter-canal (ICLD) y coherencia de inter-canal (ICC) entre L y  $L_s$ . En base a (10) y compensado por el pre-escalamiento de los canales envolvente estas pistas son:

$$\begin{aligned}
 ICLD_{LL_s} &= 10 \log_{10} \frac{P_L(k, i)}{g_s^2 P_{L_s}(k, i)} \\
 ICC_{LL_s} &= \frac{P_{LL_s}(k, i)}{\sqrt{P_L(k, i) P_{L_s}(k, i)}} .
 \end{aligned}
 \tag{15}$$

55 Similarmente, la ICLD e ICC del segundo bloque de TTO para R  $R_s$  son calculados:

60

$$\begin{aligned}
 \text{ICLD}_{RR_s} &= 10 \log_{10} \frac{P_R(k, i)}{g_s^2 P_{R_s}(k, i)} \\
 \text{ICC}_{RR_s} &= \frac{P_{RR_s}(k, i)}{\sqrt{P_R(k, i) P_{R_s}(k, i)}}.
 \end{aligned} \tag{16}$$

El bloque de tres a dos (TTT) de envolvente de MPEG es usado en “modo de energía”, véase, por ejemplo referencia [1]. Nótese que el bloque de TTT escala descendientemente el canal central por  $\sqrt{1/2}$  antes de calcular las mezclas descendentes e información lateral espacial. Tomando en cuenta el pre-escalamiento de los canales envolvente, los dos parámetros de ICLD usados por el bloque de TTT son

$$\begin{aligned}
 \text{ICLD}_1 &= 10 \log_{10} \frac{P_L + g_s^2 P_{L_s} + P_R + g_s^2 P_{R_s}}{\frac{1}{2} P_c} \\
 \text{ICLD}_2 &= 10 \log_{10} \frac{P_L + g_s^2 P_{L_s}}{P_R + g_s^2 P_{R_s}}.
 \end{aligned} \tag{17}$$

Nótese que los índices  $i$  y  $k$  han sido dejados otra vez por brevedad de notación.

Así, una información de pista espacial que comprende las pistas  $\text{ICLD}_{LL_s}$ ,  $\text{ICC}_{LL_s}$ ,  $\text{ICLD}_{RR_s}$ ,  $\text{ICC}_{RR_s}$ ,  $\text{ICLD}_1$  y  $\text{ICLD}_2$  es obtenida por el generador de información lateral espacial 260 en base a los parámetros de pista espacial 122, 122a, 122b esto es, en base a la información de energía de componente 122a y la información de dirección 122b.

### 5. Descodificación envolvente de MPEG

En lo siguiente, se describirá una descodificación de envolvente de MPEG posible, que puede ser usada para derivar señales de canal múltiple, por ejemplo señales de altavoz múltiples a partir de una señal de mezcla descendente (por ejemplo, a partir de la señal de mezcla descendente mejorada 112 o la señal de mezclada descendente mejorada 212)

Utilizando la información de pista espacial 262 (o cualquier otra información de pista espacial apropiada).

En el descodificador envolvente de MPEG, la señal de mezcla descendente recibida 112, 212 es expandida a más de dos canales utilizando la información lateral espacial recibida 262. Esta mezcla ascendente es efectuada mediante la representación en cascada apropiadamente de los llamados bloques de una a dos inverso (R-OTT) y tres a dos inverso (R-TTT), respectivamente (véase, por ejemplo referencia [6]). Mientras que el bloque de R-OTT emite dos canales de audio en base a una entrada de audio mono e información lateral, el bloque de R-TTT determina tres canales de audio en base a una entrada de audio de dos canales y la información lateral asociada. En otras palabras, los bloques inversos efectúan el procesamiento inverso como los bloques de TTT y OTT correspondientes descritos anteriormente.

Análogamente al modelo de señal de multicanal en el codificador, el descodificador supone una configuración de altavoces específica para reproducir correctamente el sonido envolvente original. Adicionalmente, el descodificador supone que el codificador de MPS (codificador envolvente de MPEG) efectúa una mezcla específica de los múltiples canales de entrada para calcular de señal de mezcla descendente correcta.

El cálculo de la mezcla descendente estéreo envolvente de MPEG es presentado en la siguiente sección.

### 6. Generación de señal de mezcla descendente estéreo envolvente de MPEG

En lo siguiente, se describirá como se genera la señal de mezcla descendente estéreo envolvente de MPEG.

En realizaciones preferidas, la mezcla descendente es determinada de tal manera que no hay diafonía entre canales de altavoz correspondientes al hemisferio izquierdo y derecho. Esto tiene la ventaja de que no hay fugas indeseables de

energía de sonido del hemisferio izquierdo al derecho, lo que incrementa significativamente la separación izquierdo/derecho después de la decodificación de la corriente envolvente de MPEG. Además, el mismo razonamiento se aplica para las fugas de señal de los canales derecho a izquierdo.

5 Cuando se usa envolvente de MPEG para la codificación de señales de audio envolvente de 5.1 convencionales, la mezcla descendente estéreo que es usada es

$$[ Y_1 \ Y_2 ]^T = M [ L \ R \ C \ L_s \ R_s ]^T , \quad (18)$$

10 donde la matriz de mezcla descendentes es

$$M = \begin{bmatrix} 1 & 0 & \sqrt{\frac{1}{2}} & g_s & 0 \\ 0 & 1 & \sqrt{\frac{1}{2}} & 0 & g_s \end{bmatrix} , \quad (19)$$

15 donde  $g_s$  es la pre-ganancia mencionada previamente dada al canal envolvente.

20 El calculo de mezcla descendente de acuerdo a (18), (19) puede ser considerado como un mapeo de las áreas de reproducción, cubiertas por posiciones de altavoz correspondientes, a los dos canales de mezcla descendente. Este mapeo es ilustrado en la Figura 4 para casos específicos del cálculo de mezcla descendente convencional (18), (19).

## 25 7. Calculo de mezcla descendente mejorada

### 7.1 Vista general con respecto al cálculo de mezcla descendente mejorada

30 En lo siguiente, se describirán detalles con respecto al cálculo de mezcla descendente mejorada. Con el fin de facilitar el entendimiento de las ventajas del concepto presente, se dará en la presente una comparación con algunos sistemas convencionales.

35 En el caso del micrófono de audio espacial como se describe en la sección 2, la señal de mezcla descendente correspondería básicamente a las señales grabadas del micrófono estéreo (por ejemplo, de la disposición de micrófono 205) en ausencia del cálculo de mezcla descendente mejorada descrito en lo siguiente. Se ha encontrado que los micrófonos estéreo prácticos no proveen la separación deseada de los componentes de señal izquierdos y derechos debido a sus patrones de directividad específicos. También se ha encontrado que consecuentemente, la diafonía entre los canales izquierdo y derecho (por ejemplo, señales de canal 210a, 210b) es demasiado alta, dando como resultado una separación de canal escasa en la señal decodificada envolvente de MPEG.

40 Las realizaciones de acuerdo con la invención crean un procedimiento para calcular una señal de mezcla descendente mejorada 112, 212 que se aproxima a las señales de mezcla descendente de SAC deseadas (por ejemplo, la señales  $Y_1$ ,  $Y_2$ ), esto es, exhibe un nivel deseado de diafonía entre los diferentes canales, que es diferente del nivel de diafonía incluido en la entrada estéreo original 110, 210. Esto da como resultado una calidad de sonido mejorada después de la decodificación de audio espacial utilizando la información lateral espacial asociada 262.

45 Los diagramas de bloques esquemáticos mostrados en las Figuras 1, 2, 3 y 5 ilustran el procedimiento propuesto. Como se puede ver, las señales de micrófono originales 110, 210, 310 son procesadas por una unidad de mejora de mezcla descendente 140, 240, 340 para obtener canales de mezcla descendente mejorada 112, 212, 312. La modificación de las señales de micrófono 110, 210, 310 es controlada por una unidad de control 120, 130, 216, 316. La unidad de control toma en cuenta el modelo de señal de multicanal para la reproducción de altavoz y los parámetros de pista espacial estimados 122, 122a, 122b, 322. A partir de esta información, la unidad de control determina un objetivo para la mejora, esto es, el modelo de la señal de mezcla descendente deseada (por ejemplo, señales de mezcla descendente  $Y_1$ ,  $Y_2$ ). Los detalles de la invención serán discutidos en lo siguiente.

### 55 7.2 Modelo para la señal de mezcla descendente estéreo deseada

60 En esta sección, se discute un modelo de la señal de mezcla descendente estéreo deseada, que también presenta el objetivo para el cálculo de mezcla descendente mejorada propuesta.

Si se aplican las ecuaciones (18) y (19) al modelo de señal envolvente supuesta de acuerdo con la ecuación (10) que tiene un modelo de señal de mezcla descendente deseada de acuerdo con

$$\begin{aligned} Y_1 &= (g_1 + \frac{1}{\sqrt{2}}g_3 + g_s g_4)\tilde{S} + \bar{N}_1 \\ Y_2 &= (g_2 + \frac{1}{\sqrt{2}}g_3 + g_s g_5)\tilde{S} + \bar{N}_2, \end{aligned} \quad (20)$$

donde las dos señales de sonido difuso  $\bar{N}_1$  y  $\bar{N}_2$  son

$$\begin{aligned} \bar{N}_1 &= h_1\tilde{N}_1 + \frac{1}{\sqrt{2}}\tilde{N}_3 + g_s h_4\tilde{N}_4 \\ \bar{N}_2 &= h_2\tilde{N}_2 + \frac{1}{\sqrt{2}}\tilde{N}_3 + g_s h_5\tilde{N}_5. \end{aligned} \quad (21)$$

El sonido difuso en la señal de micrófono izquierda y derecha es  $N_1$  y  $N_2$ . Así, la mezcla descendente debe estar basada en el sonido difuso relacionado con  $N_1$  y  $N_2$ . Puesto que como se define previamente, la potencia de  $N_1$ ,  $N_2$  y  $\tilde{N}_1$  y  $\tilde{N}_5$  son las mismas, las señales difusas basadas en  $N_1$  y  $N_2$  con la misma potencia como  $\tilde{N}_1$  a  $\tilde{N}_2$  (21) son

$$\begin{aligned} \bar{N}_1 &= \sqrt{h_1^2 + \frac{1}{2}h_3^2 + g_s^2 h_4^2} N_1 \\ \bar{N}_2 &= \sqrt{h_2^2 + \frac{1}{2}h_3^2 + g_s^2 h_5^2} N_2. \end{aligned} \quad (22)$$

Así, el modelo de señal de mezcla descendente estéreo deseada permite expresar las señales de canal  $Y_1$ ,  $Y_2$  de la señal de mezcla descendente estéreo deseada como función de valores de ganancia  $g_1$ ,  $g_2$ ,  $g_3$ ,  $g_4$ ,  $g_5$ ,  $g_s$ ,  $h_1$ ,  $h_2$ ,  $h_3$ ,  $h_4$ ,  $h_5$  y también en dependencia de la cantidad total compensada en ganancia  $\tilde{S}$  de sonido directo en la señal de micrófono estéreo y la señal difusa  $N_1$ ,  $N_2$ .

### 7.3 Filtración de un solo canal

En lo siguiente, se describirá un procedimiento en el cual un primer canal de la señal de mezcla descendente mejorada es derivada a partir de una señal del primer canal de la señal de micrófono de multicanal y en el cual un segundo canal de la señal de mezcla descendente mejorada es derivado a partir de una señal del segundo canal de la señal del micrófono del multicanal. Se debe notar que el filtrado descrito en lo siguiente puede ser efectuada por el filtro 40 o por el proveedor de señal de audio de dos canales 240 o por el mejorador de mezcla descendente 340. También se debe notar que los parámetros de filtro mejorados  $H_1$ ,  $H_2$  pueden ser provistos por el calculador de filtro 130, por el calculador de filtro 230 o por el control 316.

Un procedimiento posible para determinar las señales de mezcla descendente deseadas  $Y_1(k, i)$  y  $Y_2(k, i)$  de acuerdo con (20) es aplicar un filtro mejorado a la entrada de micrófono estéreo original  $X_1(k, i)$  y  $X_2(k, i)$ , esto es

$$\begin{aligned} \hat{Y}_1(k, i) &= H_1(k, i)X_1(k, i) \\ \hat{Y}_2(k, i) &= H_2(k, i)X_2(k, i). \end{aligned} \quad (23)$$

Estos filtros son escogidos de tal manera que  $\hat{Y}_1(k, i)$  y  $\hat{Y}_2(k, i)$  (esto es, las señales de mezcla descendente reales obtenidas mediante filtración de las señales de canal de la señal de micrófono de multicanal) se aproximan a las señales de mezcla descendente deseadas  $Y_1(k, i)$  y  $Y_2(k, i)$ , respectivamente. Una aproximación apropiada es que  $\hat{Y}_1(k, i)$  y  $\hat{Y}_2(k, i)$  comparten la

misma distribución de energía con respecto a las energías del modelo de señal de altavoz de multicanal como es dado en las señales de mezcla descendente objetivo  $Y_1(k, i)$  y  $Y_2(k, i)$ , respectivamente. En otras palabras, los filtros son escogidos de tal manera que las señales de mezcla descendente reales obtenidas mediante filtración de las señales de canal de la señal de micrófono de multicanal se aproximan a las señales de mezcla descendente deseadas con respecto a algunas propiedades estadísticas como por ejemplo características de energía o características de correlación cruzada.

En el caso en que los filtros mejorados corresponden a filtros de Wiener (véase, por ejemplo referencia [5]),  $H_1(k, i)$  y  $H_2(k, i)$  pueden ser determinados de acuerdo con

$$\begin{aligned} H_1 &= \frac{E\{X_1 Y_1^*\}}{E\{X_1 X_1^*\}} \\ H_2 &= \frac{E\{X_2 Y_2^*\}}{E\{X_2 X_2^*\}}. \end{aligned} \quad (24)$$

Sustituyendo (20) con (22) en (24), producen

$$\begin{aligned} H_1 &= \frac{w_1 E\{SS^*\} + w_3 E\{NN^*\}}{E\{SS^*\} + E\{NN^*\}} \\ H_2 &= \frac{w_2 E\{SS^*\} + w_4 E\{NN^*\}}{a^2 E\{SS^*\} + E\{NN^*\}}, \end{aligned} \quad (25)$$

Con

$$w_1 = 10^{\frac{g(\alpha)}{20}} \sqrt{1 + a^2} (g_1 + \frac{1}{\sqrt{2}} g_3 + g_s g_4) \quad (26)$$

$$w_2 = 10^{\frac{g(\alpha)}{20}} a \sqrt{1 + a^2} (g_2 + \frac{1}{\sqrt{2}} g_3 + g_s g_5) \quad (27)$$

$$w_3 = \sqrt{h_1^2 + \frac{1}{2} h_3^2 + g_s^2 h_4^2} \quad (28)$$

$$w_4 = \sqrt{h_2^2 + \frac{1}{2} h_3^2 + g_s^2 h_5^2}. \quad (29)$$

como se puede notar, los filtros mejorados dependen directamente de los diferentes componentes del modelo de señal multicanal (10). Puesto que estos componentes son estimados en base a los parámetros de pista espacial, se puede concluir que los filtros  $H_1(k, i)$  y  $H_2(k, i)$  para el cálculo de mezcla descendente mejorada dependen de estos parámetros de pista espacial también. En otras palabras, el cálculo de los filtros mejorados puede ser controlado por los parámetros de pista espacial estimados, como también se ilustra en la Figura 3.

#### 7. 4 Filtración de dos canales

En esta sección, se presenta un método alternativo al procedimiento de un solo canal discutido en la sección titulada "filtración de un solo canal". En este caso, cada canal de mezcla descendente mejorado  $\hat{Y}_1, \hat{Y}_2$  es determinado a partir de versiones filtradas de ambas señales de entrada de micrófono  $X_1, X_2$ . Ya que este procedimiento es apto de combinar ambos canales de micrófono de manera óptima, se puede esperar un desempeño mejorado en comparación con el método de filtración de un solo canal.

La señal de mezcla descendente real puede ser obtenida de acuerdo con



$$\hat{Y}_1(k, i) = [H_{1,1} \quad H_{1,2}] \begin{bmatrix} X_1(k, i) \\ X_2(k, i) \end{bmatrix} \quad (30)$$

$$\hat{Y}_2(k, i) = [H_{2,1} \quad H_{2,2}] \begin{bmatrix} X_1(k, i) \\ X_2(k, i) \end{bmatrix} \quad (31)$$

En lo siguiente, se muestra el ejemplo para estimar los filtros mejorados en base a filtros de Wiener de dos canales. Por simplicidad de presentación, se dejan los índices (k, i) en lo siguiente. La ecuación de Wiener-Hopf para el primer canal de mezcla descendente es  $\hat{Y}_1(k, i)$  :

$$\begin{bmatrix} E\{X_1X_1^*\} & E\{X_1X_2^*\} \\ E\{X_2X_1^*\} & E\{X_2X_2^*\} \end{bmatrix} \begin{bmatrix} H_{1,1} \\ H_{1,2} \end{bmatrix} = \begin{bmatrix} E\{X_1Y_1^*\} \\ E\{X_2Y_1^*\} \end{bmatrix} \quad (32)$$

Los filtros son por consiguiente obtenidos como

$$\begin{bmatrix} H_{1,1} \\ H_{1,2} \end{bmatrix} = \frac{1}{d} \begin{bmatrix} E\{X_2X_2^*\} & -E\{X_1X_2^*\} \\ -E\{X_2X_1^*\} & E\{X_1X_1^*\} \end{bmatrix} \begin{bmatrix} E\{X_1Y_1^*\} \\ E\{X_2Y_1^*\} \end{bmatrix}$$

$$\begin{bmatrix} H_{2,1} \\ H_{2,2} \end{bmatrix} = \frac{1}{d} \begin{bmatrix} E\{X_2X_2^*\} & -E\{X_1X_2^*\} \\ -E\{X_2X_1^*\} & E\{X_1X_1^*\} \end{bmatrix} \begin{bmatrix} E\{X_1Y_2^*\} \\ E\{X_2Y_2^*\} \end{bmatrix} \quad (33)$$

donde

$$d = E\{X_1X_1^*\}E\{X_2X_2^*\} - E\{X_1X_2^*\}E\{X_2X_1^*\}. \quad (34)$$

La correlación cruzada entre las señales de entrada de micrófono  $X_1, X_2$  y los canales de mezcla descendente  $Y_1, Y_2$  pueden ser expresadas por

$$\begin{aligned} E\{X_1Y_1^*\} &= w_1E\{SS^*\} + w_3E\{NN^*\} \\ E\{X_2Y_1^*\} &= aw_1E\{SS^*\} + w_3\Phi_{\text{diff}}E\{NN^*\} \\ E\{X_1Y_2^*\} &= \frac{w_2}{a}E\{SS^*\} + w_4\Phi_{\text{diff}}E\{NN^*\} \\ E\{X_2Y_2^*\} &= w_2E\{SS^*\} + w_4E\{NN^*\} \end{aligned} \quad (35)$$

donde los pesos  $w_i$  han sido introducidos en (26)-(29).

### 7.5 Selección entre filtración de un canal y filtración de dos canales

En lo siguiente, se describirá un concepto que permite una selección señal-adaptable entre filtración de un canal y filtración de dos canales.

El filtrado de dos canales, como se describe hasta ahora tiene el problema de que en la práctica algunas veces (o aun frecuentemente) produce filtros que introducen artefactos de audio. Siempre que el canal izquierdo y derecho están altamente correlacionados, la matriz de co-varianza en la ecuación de Wiener-Hopf es malamente acondicionada. La sensibilidad numérica resultante da como resultado entonces filtros que son irrazonables y provocan artefactos de audio. Para impedir esto, se usa el filtrado de un solo canal, siempre que los dos canales exceden un cierto grado de correlación. Esto puede ser implementado al calcular los filtros como



$$\begin{aligned}
 H_{1,1} &= H_1 \\
 H_{1,2} &= 0 \\
 H_{2,1} &= 0 \\
 H_{2,2} &= H_2,
 \end{aligned} \tag{36}$$

donde

$$\frac{|E \{X_1 X_2^*\}|}{\sqrt{E \{X_1 X_1^*\} E \{X_2 X_2^*\}}} > T, \tag{37}$$

donde el umbral de coherencia/correlación T determina el grado de correlación en que el filtrado de un solo canal es usada. Un valor de T = 0.9 produce buenos resultados.

En otras palabras, es posible cambiar selectivamente entre filtración de un canal y filtración de dos canales dependiendo del grado de correlación entre cualesquier señales de canal de la señal de micrófono de multicanal. Si la correlación es mayor que un valor de correlación predeterminada, se puede usar filtración de un canal en lugar de filtración de dos canales.

#### 7.6 Caso general de multicanal

En lo siguiente, se generalizara el cálculo mejorada de las señales de mezcla descendente estéreo envolvente de MPEG en base a un modelo de señal de multicanal de acuerdo con (10) a configuraciones de canal más generales. Análogamente a (10) el modelo de señal de multicanal generalizado que supone K de canales de altavoz es dado por

$$Z_l(k, i) = g_l(k, i)\tilde{S}(k, i) + h_l(k, i)\tilde{N}_l(k, i), \tag{38}$$

con l = 1, 2 ..., K. Los factores de ganancia  $g_l(k, i)$  dependen de la DOA de sonido directo y la posición del lth altavoz dentro de la configuración de reproducción. Los factores de ganancia  $h_l$  pueden ser predeterminados y usados como se explica anteriormente.  $Z_l$  representa señales de canal deseadas de una pluralidad de canales con l = 1,2, ...K.

El calculo de la señal  $Y_j(k, i)$  de un canal de mezcla descendente deseado j es obtenido por una operación de mezcla apropiada de acuerdo con

$$Y_j(k, i) = \sum_{l=0}^{K-1} m_{j,l} Z_l(k, i). \tag{39}$$

Los pesos de mezcla  $m_j$  representan una división o mapeo espacial especifica de áreas de reproducción, que están asociadas con la posición de lth altavoz al j-th de mezcla descendente.

Para dar un ejemplo: En el caso que de que un canal de altavoz L esto es a una cierta de reproducción, no contribuyera a la jth de mezcla descendente, el peso de la mezcla correspondiente  $m_j$  es ajustado a cero.

Análogamente a (23), (30) y (30), respectivamente, los canales de entrada de micrófono originales  $X_j(k, i)$  son modificados por filtros mejorados escogidos apropiadamente para aproximar los canales de mezcla descendentes deseados  $m_j$ .

En el caso de un filtro de un solo canal, se tiene

$$\hat{Y}_j(k, i) = H_l(k, i)X_j(k, i). \tag{40}$$

Aquí,  $\hat{Y}_j$  designa señales de canal reales de la señal de mezcla descendente de multicanal.

5 Nótese que (40) puede también ser aplicada en el caso de que haya más de dos señales de micrófono de entrada disponibles. Los filtros resultantes también de los parámetros de pista espacial estimados. Así, sin embargo, no se discute la estimación de los parámetros de pista espacial en base a más de dos canales de entrada de micrófono, ya que esto no es una parte esencial de la invención.

10 Es posible derivar las ecuaciones requeridas para los filtros mejorados de mezcla descendente de multicanal generales análogamente a (30), (30). Suponiendo M señales de entrada de micrófono, el jth canal de mezcla descendente deseado  $Y_j(k, i)$  es aproximado al aplicar N filtros mejorados a las señales de micrófonos mejorados  $X_m(k, i)$ :

15 
$$\hat{Y}_j(k, i) = \mathbf{H}_j^T(k, i)\mathbf{X}(k, i), \quad (41)$$

$$\mathbf{X}(k, i) = [X_1(k, i), X_2(k, i), \dots, X_M(k, i)]^T, \quad (42)$$

20 
$$\mathbf{H}_j(k, i) = [H_{j,1}(k, i), H_{j,2}(k, i), \dots, H_{j,M}(k, i)]^T. \quad (43)$$

El canal de mezcla descendente deseado correspondiente  $Y_j(k, i)$  puede ser obtenido de (39) utilizando el modelo de señal generalizado (38).

25 Los elementos de la matriz mejorada de multicanal  $H_j(k, i)$  pueden ser obtenidos al resolver la ecuación de Wiener-Hopf correspondiente

$$E \{ \mathbf{X}(k, i)\mathbf{X}^H(k, i) \} \mathbf{H}_j(k, i) = E \{ \mathbf{X}(k, i)Y_j^*(k, i) \}, \quad (44)$$

30 en donde H denota el ermitaño de un operando.

Se debe mencionar que el método descrito anteriormente puede ser considerado como un supresor de diafonía de micrófono general basado en la información de pista espacial si el numero de altavoces K en el modelo de señal de multicanal (38) es escogido grande. En este caso, la posición del altavoz puede ser directamente ser considerado como una DOA correspondiente de sonido directo. Aplicando la invención, un supresor de diafonía flexible puede ser implementado utilizando uno o más filtros de supresión.

40 8. Pre-procesamiento de las señales de micrófono

Hasta ahora, solamente se consideró el caso en donde las señales  $X_j(k, i)$  representan las señales de salida de micrófonos. El nuevo concepto o método propuesto puede ser alternativamente ser aplicado a señales de micrófono pre-procesados en lugar de esto. El procedimiento correspondiente es ilustrado en la Figura 5.

45 El pre-procesamiento puede ser implementado al aplicar formación de las no variantes en el tiempo fija (véase, por ejemplo referencia [8] en base a las señales de entrada de micrófono originales). Como resultado del pre-procesamiento, alguna parte de las fugas de señal indeseables a ciertas señales de micrófono pueden ya ser mitigadas, antes de la aplicación de los filtros mejorados.

50 Los filtros mejorados basados en los canales de entrada pre-procesados pueden ser derivados análogamente a los filtros discutidos anteriormente, al reemplazar  $X_j(k, i)$  por la señales de salida de la etapa de pre-procesamiento  $X_{j,mod}(k, i)$ .

9. Aparato de acuerdo con la Figura 3

55 La Figura 3 muestra un diagrama de bloques esquemático de un aparato 300 para generar una señal de mezcla descendente mejorada en base a una señal de micrófono de multicanal de acuerdo con otra realización de la invención.

EL aparato 300 comprende dos micrófonos 306, 308 que proveen una señal de micrófono de dos canales 310, que comprende una señal del primer canal, que es representada por una representación de dominio de tiempo-frecuencia  $X_1(k,$

i) y una señal del segundo canal que es representada por una segunda representación de tiempo-frecuencia  $X_2(k, i)$ . El aparato 300 también comprende un análisis espacial 320 que recibe la señal de micrófono de dos canales 310 y provee, en base a la misma, parámetros de pista espacial 322. El análisis espacial 320 puede tomar la funcionalidad del analizador espacial 320 o del analizado de señal 220, de tal manera que los parámetros de pista espacial 322 pueden ser equivalentes a los parámetros de pista espacial 122 o a la información de energía compuesta 122a y la información de dirección 122b. El aparato 300 también comprende un dispositivo de control 316 que recibe los parámetros de pista espacial 322 y que también recibe la señal de micrófono de dos canales 310. La unidad de control 316 también recibe un modelo de señal de multicanal 318 o comprende parámetros de tal modelo de señal de multicanal 318. El dispositivo de control 316 provee parámetros de filtro mejorados 332 al dispositivo de mejora de mezcla descendente 340. El dispositivo de control 316 puede por ejemplo tomar la funcionalidad del calculador de filtro 130 o del calculador de filtro 230, de tal manera que los parámetros de filtro mejorados 332 pueden ser equivalentes a los parámetros de filtro mejorados 132 o los parámetros de filtro mejorados 232. El dispositivo de mejora de mezcla descendente 340 recibe la señal de micrófono de dos canales 310 y también los parámetros de filtro mejorados 332 y provee, en base a los mismos, la señal de mezcla descendente de multicanal mejorada 312. Una señal del primer canal de la señal de mezcla descendente mejorada de multicanal 312 es representada por una representación de tiempo-frecuencia  $\hat{Y}_1(k, i)$  y una señal de segundo canal de la señal de mezcla descendente de multicanal mejorada 312 es representada por una representación de tiempo-frecuencia  $\hat{Y}_2(k, i)$ . Se debe notar que el dispositivo de mejora de mezcla descendente 340 puede tomar la funcionalidad del filtro 140 o del proveedor de señal de audio de dos canales 240.

#### 10. Aparato de acuerdo con la Figura 5

La Figura 5 muestra un diagrama de bloques esquemático de un aparato 500 para generar una señal de mezcla descendente mejorada en base a la señal de micrófono de multicanal. El aparato 500 de acuerdo con la Figura 5 es muy similar al aparato 300 de acuerdo con la Figura 3, de tal manera que los medios y señales idénticos son designados con números de referencia iguales y no serán explicados otra vez. Sin embargo, además de los bloques funcionales del aparato 300, el aparato 500 también comprende un pre-procesamiento 580 que recibe la señal de micrófono de multicanal 310 y provee en base a la misma una versión pre procesada 310' de la señal de micrófono de multicanal. En este caso, la mejora de mezcla de descendente 340 recibe la versión procesada 310' de la señal de micrófono de multicanal 210, en lugar de la señal de micrófono de multicanal 310 misma. También, el dispositivo de control 316 recibe la versión procesada 310' de la señal de micrófono de multicanal, en lugar de la señal de micrófono de multicanal 310 misma. Sin embargo, la funcionalidad de la mejora de mezcla descendente 340 y del dispositivo de control 316 no es afectada sustancialmente por esta modificación.

#### 11. Asignación de señales de canal a señales de mezcla descendente de acuerdo con la Figura 4

Como se discute anteriormente, el modelado de la mezcla descendente, que es usado para derivar los canales de mezcla descendente deseados  $Y_1, Y_2$  o algunas de las características estadísticas de los mismos, comprende un mapeo de un componente de sonido directo (por ejemplo,  $\tilde{S}(k, i)$ ) y de componente de sonido difuso (por ejemplo,  $\tilde{N}_l(k, i)$ ) sobre señales de canal (por ejemplo,  $L(k, i), R(k, i), C(k, i), L_s(k, i), R_s(k, i)$  o  $Z_l(k, i)$ ) y un mapeo de señales de canal de altavoz sobre las señales de canal de mezcla descendente.

Con respecto al primer mapeo del componente de sonido directo y el componente de sonido difuso sobre las señales de canal de altavoz, un mapeo dependiente de la dirección puede ser usado, que es descrito por los factores de ganancia  $g_1$ . Sin embargo, con respecto al mapeo de las señales de canal de altavoz sobre las señales de canal de mezcla descendente, se pueden usar suposiciones fijas que pueden ser descritas por una matriz de mezcla descendente. Como se ilustra en la Figura 4 se puede suponer que solamente las señales de canal de altavoz  $C, L$  y  $L_s$  deben contribuir a la primera señal de canal de mezcla descendente  $Y_1$  y que solamente las señales de canal de altavoz  $C, R$  y  $R_s$  deben contribuir a la señal de canal de mezcla descendente  $Y_2$ .

Esto es ilustrado en la Figura 4.

#### 12. Flujo de procesamiento de señal de acuerdo con la Figura 6

En lo siguiente, se describirá el flujo del procesamiento de señal en una realización de acuerdo con la invención con referencia a la Figura 6. La Figura 6 muestra una representación esquemática del flujo de procesamiento de señal para derivar los parámetros de filtro mejorados  $H$  a partir de la señal de micrófono de multicanal representada por ejemplo, por representaciones de tiempo-frecuencia  $X_1$  y  $X_2$ .

El flujo de procesamiento 600 comprende por ejemplo como una primera etapa, un análisis espacial 610 que puede tomar la funcionalidad de un cálculo de parámetro de pista espacial. Así, una información de potencia de sonido directo (o

información de energía de sonido directo)  $E\{SS^*\}$ , una información de potencia de sonido difuso (o información de energía de sonido difuso)  $E\{NN^*\}$  y una información de dirección  $\alpha$ , a puede ser obtenida en base a la señales de micrófono de multicanal. Detalles con respecto a la derivación de la información de potencia de sonido directo (o información de energía de sonido directo) de la información de potencia de sonido difuso (o información de energía de sonido difuso) y la información de dirección han sido discutidos anteriormente.

EL flujo de procesamiento 600 también comprende un mapeo del factor de ganancia 320, en el cual la información de dirección es mapeada sobre una pluralidad de factores de ganancia (por ejemplo, factores de ganancia  $g_1$  a  $g_5$ ). El mapeo de factor de ganancia 620 puede por ejemplo ser efectuado utilizando una ley de toma panorámica de amplitud de multicanal como se describe anteriormente.

El flujo de procesamiento 600 también comprende un cálculo del parámetro de filtro 630, en el cual los parámetros de filtro mejorados H son derivados a partir de la información de potencia de sonido directo, la información de potencia de sonido difuso, la información de dirección y los factores de ganancia. El cálculo del parámetro de filtro 630 puede usar adicionalmente uno o más parámetros constantes que describen por ejemplo un mapeo deseado de canales de altavoz sobre señales de canal de mezcla descendente. También, se pueden aplicar parámetros predeterminados que describen un mapeo del componente de sonido difuso sobre las señales de altavoz.

El cálculo del parámetro de filtro comprende, por ejemplo un mapeo W632. En el mapeo w, que puede ser efectuado de acuerdo con las ecuaciones 26 a 29 se pueden obtener valores  $w_1$  a  $w_4$  que pueden servir como cantidades intermedias. El cálculo del parámetro de filtro 630 comprende además un mapeo H 634 que puede ser efectuado por ejemplo de acuerdo con la ecuación 25. En el mapeo H 634, los parámetros de filtro mejorados H pueden ser determinados. Para el mapeo H, valores de correlación cruzada deseados  $E\{X_1, Y_1^*\}$ ,  $E\{X_2, Y_2^*\}$  entre canales de la señal de micrófono y los canales de la señal de mezcla descendente pueden ser usados. Estos valores de correlación cruzada deseados pueden ser obtenidos en base a la información de potencia de sonido directo  $E\{SS^*\}$  y  $E\{NN^*\}$ , como se puede ver en el numerador de las ecuaciones (25), que es idéntico al numerador de las ecuaciones (24).

Para concluir, el flujo de procesamiento de la Figura 6 puede ser aplicado para derivar los parámetros de filtro mejorados H a partir de la señal de micrófono de multicanal representada por las señales de canal  $X_1, X_2$ .

### 13. Flujo de procesamiento de señal de acuerdo con la Figura 7

La Figura 7 muestra una representación esquemática de un flujo de procesamiento de señal 700 de acuerdo con otra realización de la invención. El flujo de procesamiento de señal 700 puede ser usado para derivar parámetros de filtro mejorado H a partir de una señal de micrófono de multicanal.

El flujo de procesamiento de señal 700 comprende un análisis espacial 710 que puede ser idéntico al análisis espacial 610. También, el flujo de procesamiento de señal 700 comprende un mapeo del factor de ganancia 720, que puede ser idéntico al mapeo de facto de ganancia 620.

El flujo de procesamiento de señal 700 también comprende un cálculo del parámetro de filtro 730. El cálculo del parámetro de filtro 730 puede comprender un mapeo w 732 que puede ser idéntico al mapeo w 632 en algunos casos. Sin embargo, se pueden usar diferentes mapeos w si se aprecia que es apropiado.

El cálculo del parámetro de filtro 730 también comprende un cálculo de correlación cruzada deseada 734, en el curso del cual se calcula una correlación cruzada deseada entre canales de la señal de micrófono de multicanal y canales de la señal de mezcla descendente deseada. Este cálculo puede ser efectuado por ejemplo de acuerdo con la ecuación 35. Se debe notar que se puede aplicar un modelo de una señal de mezcla descendente deseada en el cálculo de correlación cruzada deseada 734. Por ejemplo, suposiciones en como el componente de sonido directo de la señal de micrófono de multicanal debe ser mapeado a una pluralidad de señales de altavoz en dependencia de la información de dirección pueden ser aplicadas en el cálculo de correlación cruzada deseada 734. Además, suposiciones de como los componentes de sonido difuso de la señal de micrófono de multicanal deben ser reflejados en las señales de altavoz pueden también ser evaluadas en el cálculo de correlación cruzada deseada 734. Además, suposiciones con respecto a un mapeo deseado de canales de altavoz múltiples sobre la señal de mezcla descendente pueden también ser aplicadas al cálculo de correlación cruzada deseada 734. Así, una correlación cruzada deseada  $E\{X_i, Y_j^*\}$  entre canales de la señal de micrófono y canales de la señal de mezcla descendente (deseada) puede ser obtenida en base a la información de potencia de sonido directo, la información de potencia de sonido difuso, la dirección de información y factores de ganancia dependientes de la dirección (en donde la última información puede ser combinada para obtener valores de w intermedios).

El cálculo del parámetro de filtro 730 también comprende una solución de la ecuación de Wiener-Hopf 736, que puede por ejemplo ser efectuado de acuerdo con las ecuaciones 33 y 34. Para este propósito, la ecuación de Wiener-Hopf puede ser

ajustada en dependencia de la información de potencia de sonido directo, la información de potencia de sonido difuso y la correlación cruzada deseada entre canales de la señal de micrófono de multicanal y canales de la señal de mezcla descendente (deseada). Como una solución de la ecuación de Wiener-Hopf (por ejemplo, la ecuación 32) se tienen parámetros de filtro mejorados H.

5

Para resumir lo anterior, la determinación de los parámetros de filtro mejorado H puede comprender etapas separadas de cálculo de una correlación cruzada deseada y de establecimiento y solución de una ecuación de Wiener-Hopf (etapa 736) en algunas realizaciones.

10 14. Conclusiones

Para resumir lo anterior, las realizaciones de acuerdo con la invención crean un concepto y métodos mejorados para calcular una señal de c deseada de codificadores de audio espacial paramétricos en base a señales de entrada de micrófono. Un ejemplo importante es dado por la conversión de una señal de micrófono estéreo a una mezcla de descendente envolvente de MPEG correspondiente a los parámetros de MPS calculados. La señal de mezcla descendente mejorada conduce a una calidad de audio espacial significativamente mejorada y propiedad de localización después de la descodificación de MPS en comparación con el caso del estado del arte propuesto en la referencia [2]. Una realización simple de acuerdo con la invención comprende las siguientes etapas 1 a 4:

20 (1) recibir señales de entrada de micrófono;

(2) calcular parámetros de pista espacial;

25 (3) determinar filtros mejorados de mezcla descendente en base al modelo de los canales de mezcla descendente deseados, un modelo de señal de altavoz de multicanal para la salida del descodificador y parámetros de pista espacial y

(4) aplicar los filtros mejorados a las señales de entrada de micrófono para obtener señales de mezcla descendente mejoradas para uso con micrófonos de audio espacial.

30 Otra realización simple de acuerdo con la invención crea un aparato, un método o un programa de computador para generar una señal de mezcla descendente, el aparato, método o programa de ordenador comprende un calculador de filtro para calcular parámetros de filtro mejorados en base a información de señal de micrófono o en base a información en un montaje de reproducción propuesto y el aparato, método o programa de ordenador comprende una disposición de filtro (o etapa de filtración) para filtrar señales de micrófono utilizando los parámetros de filtro mejorados para obtener la señal de mezcla descendente mejorada.

35

Este aparato, método o programa de ordenador puede opcionalmente puede ser mejorado en que el calculador de filtro está configurado para calcular los parámetros de filtro mejorados en base a un modelo de los canales de mezcla descendentes deseados, modelo de señal de altavoz de multicanal para la salida del descodificador o parámetros de pista espacial.

40

15. Alternativas de implementación

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

50

La señal de audio codificada de la invención puede ser almacenada en un medio de almacenamiento digital o puede ser transmitida en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión cableado tal como internet.

55 Dependiendo de ciertos requerimientos de implementación, realizaciones de la invención pueden ser implementados en elementos físicos o en elementos de programación. La implementación puede ser efectuada utilizando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un blue-ray, un cd, un ROM, un PROM, un EPROM, un EEPROM o una memoria instantáneas, que tiene señales de control que se pueden leer electrónicamente almacenadas en el mismo que cooperan (o son aptas de cooperar) con un sistema de ordenador programable de tal manera que el método respectivo es efectuado. Por consiguiente, el medio de almacenamiento digital se puede leer por ordenador.

60

Algunas realizaciones de acuerdo con la invención comprenden un portador de datos que tiene señales de control que se pueden leer electrónicamente que son aptas de cooperar un sistema de ordenador programable, de tal manera que uno de los métodos descritos en la presente es efectuado.

5 En general, las realizaciones de la presente invención pueden ser implementadas como un producto de programa de ordenador con códigos de programa, los códigos de programa son operativos para efectuar uno de los métodos cuando el producto de programa de ordenador se ejecuta en un ordenador. Los códigos de programa pueden por ejemplo ser almacenados en un portador que se puede leer por la máquina.

10 Otras realizaciones comprenden el programa de ordenador para efectuar uno de los métodos descritos en la presente, almacenados en un portador que se puede leer por máquina.

15 En otras palabras, una realización del método de la invención es por consiguiente un programa de ordenador que tiene códigos de programa para efectuar uno de los métodos descritos en la presente, cuando el programa de ordenador se ejecuta en un ordenador.

20 Una realización adicional de los métodos de la invención es por consiguiente un portador de datos (o medio de almacenamiento digital o medio que se puede leer por ordenador) que comprende registrarlos en el mismo, el programa de ordenador para efectuar uno de los métodos descritos en la presente. El portador de datos, medios de almacenamiento digital o el medio grabado son comúnmente tangibles y/o sin transición.

25 Una realización adicional del método de la invención es por consiguiente una corriente de datos o una secuencia de señales que representan el programa de ordenador para efectuar uno de los métodos descritos en la presente. La corriente de datos o la secuencia de señales pueden por ejemplo estar configuradas para ser transferidas vía una colección de comunicación de datos, por ejemplo vía internet.

Una realización adicional comprende medios de procesamiento, por ejemplo un ordenador o un dispositivo lógico programable configurado para o apto para efectuar uno de los métodos descritos en la presente.

30 Una realización adicional comprende un ordenador que tiene instalado en la misma el programa de ordenador para efectuar uno de los métodos descritos en la presente.

35 Una realización adicional de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa de ordenador para efectuar uno de los métodos descritos en la presente a un receptor. El receptor puede ser por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria o los semejantes. El aparato o sistema puede comprender por ejemplo un servidor de archivos para transferir programas de ordenador al receptor.

40 En algunas realizaciones, se puede usar un dispositivo lógico programable (por ejemplo, una disposición de compuerta programable en el campo) para efectuar algunas o todas las funcionalidades de los métodos descritos en la presente. En algunas realizaciones, una disposición de compuerta programable al campo puede cooperar con un microprocesador con el fin de efectuar uno de los métodos descritos en la presente. En general, los métodos son preferiblemente efectuados mediante cualquier aparato de elementos físicos.

45 Las realizaciones descritas anteriormente son solamente ilustrativas para los principios de la presente invención. Se entenderá que modificaciones y variaciones de las disposiciones y los detalles descritos en la presente serán evidentes para otros experimentados en el arte. Es la intención por consiguiente estar limitado solamente por el alcance de las reivindicaciones de patente pendientes y no por los detalles específicos presentados a manera de descripción y explicación de las realizaciones en el presente documento.

50

## References

55 [1] ISO/IEC 23003-1:2007. Information technology - MPEG Audio technologies - Part 1: MPEG Surround. International Standards Organization, Geneva, Switzerland, 2007.

[2] C. Faller. Microphone front-ends for spatial audio coders. In 125th AES Convention, Paper 7508, San Francisco, Oct. 2008.

60 [3] M. A. Gerzon. Periphony: Width-Height Sound Reproduction. J. Aud. Eng. Soc., 21(1):2-10, 1973.

- [4] D. Griesinger. Stereo and surround panning in practice. In Preprint 112th Conv. Aud. Eng. Soc., May 2002.
- [5] S. Haykin. Adaptive Filter Theory (third edition). Prentice Hall, 1996.
- 5 [6] J. Herre, K. Kjörling, J. Breebaart, C. Faller, S. Disch, H. Purnhagen, J. Koppens, J. Hilpert, J. R-od'en, W. Oomen, K. Linzmeier, and K. S. Chong. Mpeg surround-the *iso/mpeg* standard for efficient and compatible multi-channel audio coding. In Preprint 122th Conv. Aud. Eng. Soc., May 2007.
- 10 [7] V. Pulkki. Virtual sound source positioning using Vector Base Amplitude Panning. J. Audio Eng. Soc., 45:456-466, June 1997.
- [8] B. D. Van Veen and K. M. Buckley. Beamforming: A versatile approach to spatial filtering. IEEE ASSP Magazine, 5(2):4-24, April 1988.
- 15 [9] European Patent Application EP 1 565 036 A2, AGERE SYSTEM INC: Late reverberation-based synthesis of auditory scenes, Published on 17 August 2005 .

**REIVINDICACIONES**

1. Un aparato (100; 200; 300; 500) para generar una señal de mezcla descendente mejorada (112; 212; 312) en base a una señal de micrófono de multicanal (110; 210; 310), donde el aparato comprende:
- 5 un analizador espacial (120; 220; 320) configurado para calcular un conjunto de parámetros de pista espacial ( $E\{NN^*\}$ ,  $E\{SS^*\}$ ,  $\alpha$ ,  $\alpha$ ) que comprenden una información de dirección ( $\alpha$ ,  $\alpha$ ) que describe una dirección de llegada de sonido directo, una información de potencia de sonido directo ( $E\{SS^*\}$ ), y una información de potencia de sonido difuso ( $E\{NN^*\}$ , en base a la señal de micrófono de multicanal;
- 10 un calculador de filtro (130; 230; 316) para calcular parámetros de filtro mejorados (132; 232; 332) en dependencia de la información de dirección ( $\alpha$ ,  $\alpha$ ) que describe la dirección de llegada de sonido directo, en dependencia de la información de potencia de sonido directo ( $E\{SS^*\}$ ) y en dependencia de la información de potencia de sonido difuso ( $E\{NN^*\}$ ); y
- 15 un filtro (140; 240; 340) para filtrar la señal de micrófono (110; 210; 310) o una señal derivada de la misma, utilizando los parámetros de filtro mejorados (132; 232; 332) para obtener la señal de mezcla descendente mejorada (112; 212; 312);
- 20 en donde el calculador de filtro está configurado para calcular los parámetros de filtro mejorados ( $H_1, H_2; H_{1,1}, H_{1,2}, H_{2,1}, H_{2,2}$ ) en dependencia de los factores de ganancia dependientes de la dirección ( $g_1, g_2, g_3, g_4, g_5$ ) que describen contribuciones deseadas de un componente de sonido directo (S) de la señal de micrófono de multicanal a una pluralidad de señales de altavoz (L, R, C,  $L_s, R_s; Z_i$ ) y en dependencia de uno o más valores de matriz de mezcla descendente ( $g_s; m_{i,j}$ ) que describen contribuciones deseadas de una pluralidad de canales de audio (L, R, C,  $L_s, R_s; Z_i$ ) a uno o más canales de la señal de mezcla descendente mejorada.
- 25 2. El aparato de acuerdo con la reivindicación 1, en el que el calculador de filtro (130; 230; 316) está configurado para calcular los parámetros de filtro mejorados (132; 232; 332;  $H_1, H_2; H_{1,1}, H_{1,2}, H_{2,1}, H_{2,2}$ ) de tal manera que la señal de mezcla descendente mejorada (112; 212; 312;  $\hat{Y}_1; \hat{Y}_2$ ) se aproxima a una señal de mezcla descendente deseada ( $Y_1, Y_2$ ).
- 30 3. El aparato de acuerdo con la reivindicación 1 o la reivindicación 2, en el que el calculador de filtro (130; 230; 316) está configurado para calcular valores de correlación cruzada deseados ( $E\{X_1Y_1^*\}, E\{X_2Y_1^*\}, E\{X_1Y_2^*\}, E\{X_2Y_2^*\}$ ) entre señales de canal ( $X_1, X_2$ ) de la señal de micrófono de multicanal (110; 210; 310) y señales de canal deseadas ( $Y_1, Y_2$ ) de la señal de mezcla descendente en dependencia de los parámetros de pista espacial y
- 35 en donde el calculador de filtro está configurado para calcular los parámetros de filtro mejorados ( $H_1, H_2; H_{1,1}, H_{1,2}, H_{2,1}, H_{2,2}$ ) en dependencia de los valores de correlación cruzada deseados.
- 40 4. El aparato de acuerdo con la reivindicación 3, en el que el calculador de filtro está configurado para calcular los valores de correlación cruzada deseados en dependencia de los factores de ganancia dependientes de la dirección ( $g_1, g_2, g_3, g_4, g_5$ ) que describen contribuciones deseadas de un componente de sonido directo (S) de la señal de micrófono de multicanal a una pluralidad de señales de altavoz (L, R, C,  $L_s, R_s; Z_i$ ).
- 45 5. El aparato de acuerdo con la reivindicación 4, en el que el calculador de filtro (130; 230; 316) está configurado para mapear la información de dirección ( $\alpha$ ,  $\alpha$ ) sobre un conjunto de factores de ganancia dependiente de la dirección ( $g_1, g_2, g_3, g_4, g_5$ ).
- 50 6. El aparato de acuerdo con cualquiera de las reivindicaciones 3 a 5, en el que el calculador de filtro (130; 230; 316) está configurado para considerar la información de potencia de sonido directo ( $E\{SS^*\}$ ) y la información de potencia de sonido difuso ( $E\{NN^*\}$ ) para calcular los valores de correlación cruzada deseados ( $E\{X_1Y_1^*\}, E\{X_2Y_1^*\}, E\{X_1Y_2^*\}, E\{X_2Y_2^*\}$ ).
- 55 7. El aparato de acuerdo con la reivindicación 6, en el que el calculador de filtro (130; 230; 316) está configurado para ponderar la información de potencia de sonido directo ( $E\{SS^*\}$ ) en dependencia de la información de dirección ( $\alpha$ ,  $\alpha$ ), y para aplicar una ponderación predeterminada, que es independiente de la información de dirección, a la información de potencia de sonido difuso ( $E\{NN^*\}$ ) con el fin de calcular los valores de correlación cruzada deseados ( $E\{X_1Y_1^*\}, E\{X_2Y_1^*\}, E\{X_1Y_2^*\}, E\{X_2Y_2^*\}$ ).
- 60 8. El aparato de acuerdo con cualquiera de las reivindicaciones 1 a 7, en el que el calculador de filtro (130; 230; 316) está configurado para calcular los coeficientes de filtro  $H_1, H_2$  de acuerdo con



$$H_1 = \frac{w_1 E\{SS^*\} + w_3 E\{NN^*\}}{E\{SS^*\} + E\{NN^*\}}$$

$$H_2 = \frac{w_2 E\{SS^*\} + w_4 E\{NN^*\}}{a^2 E\{SS^*\} + E\{NN^*\}}$$

10 donde  $E\{SS^*\}$  es información de potencia de sonido directo,

donde  $E\{NN^*\}$  es información de potencia de sonido difuso,

15 donde  $w_1$  y  $w_2$  son coeficientes que son dependientes de la información de dirección ( $a$ ,  $\alpha$ ), y

donde  $w_3$  y  $w_4$  son coeficientes determinados por las ganancias de sonido difuso ( $h_1, h_2, h_3, h_4, h_5$ ); y

donde el filtro (140; 240; 340) está configurado para determinar una señal del primer canal  $\hat{Y}_1(k,i)$  y una señal del segundo canal  $\hat{Y}_2(k,i)$  de la señal de mezcla descendente mejorada (112; 212; 312) en dependencia de una señal del primer canal  $X_1(k,i)$  y de una señal del segundo canal  $X_2(k,i)$  de la señal de micrófono de multicanal de acuerdo con

$$\hat{Y}_1(k,i) = H_1(k,i)X_1(k,i)$$

$$\hat{Y}_2(k,i) = H_2(k,i)X_2(k,i)$$

25

9. El aparato de acuerdo con cualquiera de las reivindicaciones 1 a 7, en el que el calculador de filtro (130; 230; 316) está configurado para calcular coeficientes de filtro ( $H_{1,1}, H_{1,2}, H_{2,1}$  y  $H_{2,2}$ ) de acuerdo con

30

$$\begin{bmatrix} H_{1,1} \\ H_{1,2} \end{bmatrix} = \frac{1}{d} \begin{bmatrix} E\{X_2X_2^*\} & -E\{X_1X_2^*\} \\ -E\{X_2X_1^*\} & E\{X_1X_1^*\} \end{bmatrix} \begin{bmatrix} E\{X_1Y_1^*\} \\ E\{X_2Y_1^*\} \end{bmatrix}$$

$$\begin{bmatrix} H_{2,1} \\ H_{2,2} \end{bmatrix} = \frac{1}{d} \begin{bmatrix} E\{X_2X_2^*\} & -E\{X_1X_2^*\} \\ -E\{X_2X_1^*\} & E\{X_1X_1^*\} \end{bmatrix} \begin{bmatrix} E\{X_1Y_2^*\} \\ E\{X_2Y_2^*\} \end{bmatrix}$$

donde

$$d = E\{X_1X_1^*\}E\{X_2X_2^*\} - E\{X_1X_2^*\}E\{X_2X_1^*\}.$$

45

donde

$X_1$  designa una señal del primer canal de la señal de micrófono de multicanal,

50  $X_2$  designa una señal del segundo canal de la señal de micrófono de multicanal,

$E\{\cdot\}$  designa una operación de promedio de tiempo corto,

55 \* designa una operación conjugada compleja,

$E\{X_1Y_1^*\}$ ,  $E\{X_2Y_1^*\}$ ,  $E\{X_1Y_2^*\}$  y  $E\{X_2Y_2^*\}$  designan valores de correlación cruzada entre señales de canal  $X_1, X_2$  de la señal de micrófono de multicanal y señales de canal deseados  $Y_1, Y_2$  de la señal de mezcla descendente mejorada.

10. El aparato de acuerdo con la cualquiera de las reivindicaciones 1 a 9, en el que el calculador de filtro (130; 230; 316) está configurado para calcular los parámetros de filtro mejorados  $H_{j,1}(k,i)$  a  $H_{j,M}(k,i)$  de tal manera que las señales de canal  $\hat{Y}_j(k,i)$  de la señal de mezcla descendente mejorada (112; 212; 312) obtenida mediante filtración de señales de canal ( $X_1$ ,  $X_2$ ) de la señal de micrófono de multicanal de acuerdo con los parámetros de filtro mejorados se aproxima, con respecto a una medida estadística de similaridad, a señales de canal  $Y_j(k,i)$  deseadas definidas como

$$Y_j(k,i) = \sum_{l=0}^{K-1} m_{j,l} Z_l(k,i).$$

con

$$Z_l(k,i) = g_l(k,i) \tilde{S}(k,i) + h_l(k,i) \tilde{N}_l(k,i).$$

donde  $g_l$  son factores de ganancia que son dependientes de la información de dirección ( $\alpha$ ) y que representan contribuciones deseadas de un componente de sonido directo ( $\tilde{S}$ ) de la señal de micrófono de multicanal (110; 210; 310) a una pluralidad de señales de altavoz ( $Z$ );

donde  $h_l$  son valores predeterminados que describen contribuciones deseadas del componente de sonido difuso ( $\tilde{N}$ ) de la señal de micrófono de multicanal (110; 210; 310) a una pluralidad de señales de altavoz.

11. El aparato de acuerdo con cualquiera de las reivindicaciones 1 a 10, en el que el calculador de filtro (130; 230; 316) está configurado para evaluar una ecuación de Wiener-Hopf para derivar los parámetros de filtro mejorados (132; 232; 332;  $H_1$ ,  $H_2$ ;  $H_{1,1}$ ,  $H_{1,2}$ ,  $H_{2,1}$ ,  $H_{2,2}$ ),

donde la ecuación de Wiener-Hopf describe una relación entre los valores de correlación  $E\{X_1 X_1^*\}$ ,  $E\{X_1 X_2^*\}$ ,  $E\{X_2 X_1^*\}$ ,  $E\{X_2 X_2^*\}$ , tales valores de correlación describen una relación entre diferentes pares de canal de la señal de micrófono de multicanal, parámetros de filtro mejorados ( $H_{1,1}$ ,  $H_{1,2}$ ,  $H_{2,1}$ ,  $H_{2,2}$ ) y valores de correlación cruzada deseados ( $E\{X_1 Y_1^*\}$ ,  $E\{X_2 Y_1^*\}$ ,  $E\{X_1 Y_2^*\}$ ,  $E\{X_2 Y_2^*\}$ ) entre señales de canal ( $X_1$ ,  $X_2$ ) de señal de micrófono de multicanal (110; 210; 310) y señales de canal deseadas ( $Y_1, Y_2$ ) de la señal de mezcla descendente.

12. El aparato de acuerdo con cualquiera de las reivindicaciones 1 a 11, en el que el calculador de filtro (130; 230; 316) está configurado para calcular los parámetros de filtro mejorados (132; 232; 332) en dependencia de un modelo de canales de mezcla descendentes deseados.

13. El aparato de acuerdo con cualquiera de las reivindicaciones 1 a 12, en el que el calculador de filtro (130; 230; 316) está configurado para efectuar selectivamente un filtrado de un solo canal, en el cual un primer canal ( $\hat{Y}_1$ ) de la señal de mezcla descendente mejorada (112; 212; 312) es derivado mediante un filtrado de un primer canal ( $X_1$ ) de la señal de micrófono de multicanal (110; 210; 310) y en el cual un segundo canal ( $\hat{Y}_2$ ) de la señal de mezcla descendente mejorada es derivado mediante un filtrado de un segundo canal ( $X_2$ ) de la señal de micrófono de multicanal mientras que se evita la diafonía del primer canal de la señal de micrófono de multicanal al segundo canal de la señal de mezcla descendente mejorada y del segundo canal de la señal de micrófono de multicanal al primer canal de la señal de mezcla descendente mejorada,

o un filtrado de dos canales en el cual un primer canal ( $\hat{Y}_1$ ) de la señal de mezcla descendente mejorada es derivado mediante filtrado de un primero y un segundo canal ( $X_1, X_2$ ) de la señal de micrófono de multicanal, y en el cual un segundo canal ( $\hat{Y}_2$ ) de la señal de mezcla descendente mejorada es derivado mediante filtrado de un primero y segundo canal ( $X_1, X_2$ ) de la señal de micrófono de multicanal,

en dependencia de un valor de correlación que describe una correlación entre el primer canal ( $X_1$ ) de la señal de micrófono de multicanal y el segundo canal ( $X_2$ ) de la señal de micrófono de multicanal.

14. Un método para generar una señal de mezcla descendente mejorada en base a una señal de micrófono de multicanal, donde el método comprende:

calcular un conjunto de parámetros de pista espacial que comprenden una información de dirección que describe una

- dirección de llegada de un sonido directo, una información de potencia de sonido directo e información de potencia de sonido difuso en base a la señal de micrófono de multicanal,
- 5 calcular parámetros de filtro mejorados en dependencia de la información de dirección que describe la dirección de llegada de sonido directo, en dependencia de la información de potencia de sonido directo y en dependencia de la información de potencia de sonido difuso; y
- 10 filtrar la señal de micrófono o una señal derivada de la misma, utilizando los parámetros de filtro mejorados, para obtener la señal de mezcla descendente mejorada;
- 15 donde los parámetros de filtro mejorados ( $H_1, H_2; H_{1,1}, H_{1,2}, H_{2,1}, H_{2,2}$ ) son calculados en dependencia de los factores de ganancia dependientes de la dirección ( $g_1, g_2, g_3, g_4, g_5$ ) que describen contribuciones deseadas de un componente de sonido directo (S) de la señal de micrófono de multicanal a una pluralidad de señales de altavoz (L, R, C,  $L_s, R_s; Z_i$ ) y en dependencia de uno o más valores de matriz de mezcla descendente ( $g_s; m_{j,i}$ ) que describen contribuciones deseadas de una pluralidad de canales de audio (L, R, C,  $L_s, R_s; Z_i$ ) a uno o más canales de la señal de mezcla descendente mejorada.
15. Un aparato (100; 200; 300; 500) para generar una señal de mezcla descendente mejorada (112; 212; 312) en base a una señal de micrófono de multicanal (110; 210; 310), donde el aparato comprende:
- 20 un analizador espacial (120; 220; 320) configurado para calcular un conjunto de parámetros de pista espacial, ( $E\{NN^*\}, E\{SS^*\}, a, \alpha$ ) que comprende información de dirección ( $a, \alpha$ ) que describe una dirección de llegada de sonido directo, una información de potencia de sonido directo ( $E\{SS^*\}$ ) y una información de potencia de sonido difuso ( $E\{NN^*\}$ ), en base a la señal de micrófono de multicanal;
- 25 un calculador de filtro (130; 230; 316) para calcular parámetros de filtro mejorados (132; 232; 332) en dependencia de la información de dirección ( $a, \alpha$ ) que describe la dirección de llegada del sonido directo, en dependencia de la información de potencia de sonido directo ( $E\{SS^*\}$ ) y en dependencia de la información de potencia de sonido difuso ( $E\{NN^*\}$ ); y
- 30 un filtro (140; 240; 340) para filtrar la señal de micrófono (110; 210; 310), o una señal derivada de la misma, utilizando los parámetros de filtro mejorados (132; 232; 332), para obtener la señal de mezcla descendente mejorada (112; 212; 312);
- en donde el calculador de filtro (130; 230; 316) está configurado para efectuar selectivamente un filtrado de un solo canal, en el cual un primer canal ( $\hat{Y}_1$ ) de la señal de mezcla descendente mejorada (112; 212; 312) es derivado mediante un filtrado
- 35 de un primer canal ( $X_1$ ) de la señal de micrófono de multicanal (110; 210; 310) y en el cual un segundo canal ( $\hat{Y}_2$ ) de la señal de mezcla descendente mejorada es derivado mediante un filtrado de un segundo canal ( $X_2$ ) de la señal de micrófono de multicanal mientras que se evita la diafonía del primer canal de la señal de micrófono de multicanal al segundo canal de la señal de mezcla descendente mejorada y del segundo canal de la señal de micrófono de multicanal al primer canal de la señal de mezcla descendente mejorada,
- 40 o un filtrado de dos canales en el cual un primer canal ( $\hat{Y}_1$ ) de la señal de mezcla descendente mejorada es derivado mediante filtrado de un primero y un segundo canal ( $X_1, X_2$ ) de la señal de micrófono de multicanal, y en el cual un segundo canal ( $\hat{Y}_2$ ) de la señal de mezcla descendente mejorada es derivado mediante filtrado de un primero y un segundo canal ( $X_1, X_2$ ) de la señal del micrófono de multicanal,
- 45 en dependencia de un valor de correlación que describe una correlación entre el primer canal ( $X_1$ ) de la señal de micrófono del multicanal y el segundo canal ( $X_2$ ) de la señal de micrófono de multicanal.
16. Un método para generar una señal de mezcla descendente mejorada en base a una señal de micrófono de multicanal, donde el método comprende:
- 50 calcular un conjunto de parámetros de pista espacial que comprende información de dirección que describe una dirección de llegada de sonido directo, información de potencia de sonido directo e información de potencia de sonido difuso en base a la señal de micrófono de multicanal;
- 55 calcular parámetros de filtro mejorados en dependencia de la información de dirección que describe la dirección de llegada de sonido directo, en dependencia de la información de potencia de sonido directo y en dependencia de la información de potencia de sonido difuso y

filtrar la señal de micrófono, o una señal derivada de la misma, utilizando los parámetros de filtro mejorados, para obtener la señal de mezcla descendente mejorada;

5 donde el método comprende efectuar selectivamente un filtrado de un solo canal, en el cual un primer canal ( $\hat{Y}_1$ ) de la señal de mezcla descendente mejorada (112; 212; 312) es derivado mediante filtrado de un primer canal (X1) de la señal de micrófono de multicanal (110; 210; 310) y en el cual un segundo canal ( $\hat{Y}_2$ ) de la señal de mezcla descendente mejorada es derivado mediante un filtrado de un segundo canal (X2) de la señal de micrófono de multicanal mientras que se evita la diafonía del primer canal de la señal de micrófono de multicanal al segundo canal de la señal de mezcla descendente mejorada y del segundo canal de la señal de micrófono de multicanal al primer canal de la señal de mezcla descendente mejorada,

10

o un filtrado de dos canales en el cual un primer canal ( $\hat{Y}_1$ ) de la señal de mezcla descendente mejorada es derivado mediante filtrado de un primero y un segundo canal (X1, X2) de la señal de micrófono de multicanal, y en el cual un segundo canal ( $\hat{Y}_2$ ) de la señal de mezcla descendente mejorada es derivado mediante filtrado de un primero y segundo canal (X1, X2) de la señal de micrófono de multicanal,

15

en dependencia de un valor de correlación que describe una correlación entre el primer canal (X1) de la señal de micrófono de multicanal y el segundo canal (X2) de la señal de micrófono de multicanal.

20 17. Un programa de ordenador adaptado para efectuar un método de acuerdo con la reivindicación 14 o la reivindicación 16, cuando el programa de ordenador se ejecuta en un ordenador.

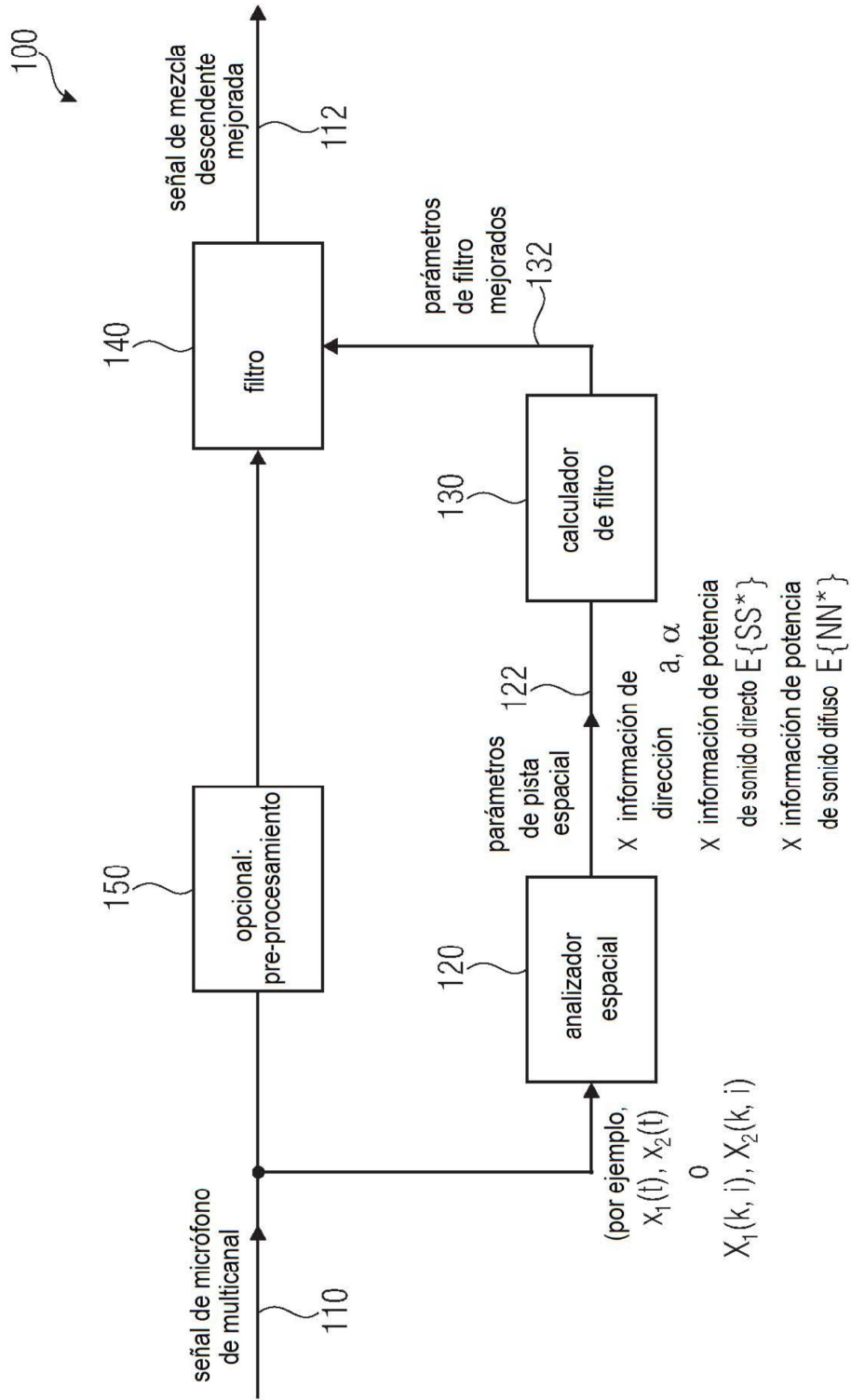


FIG 1

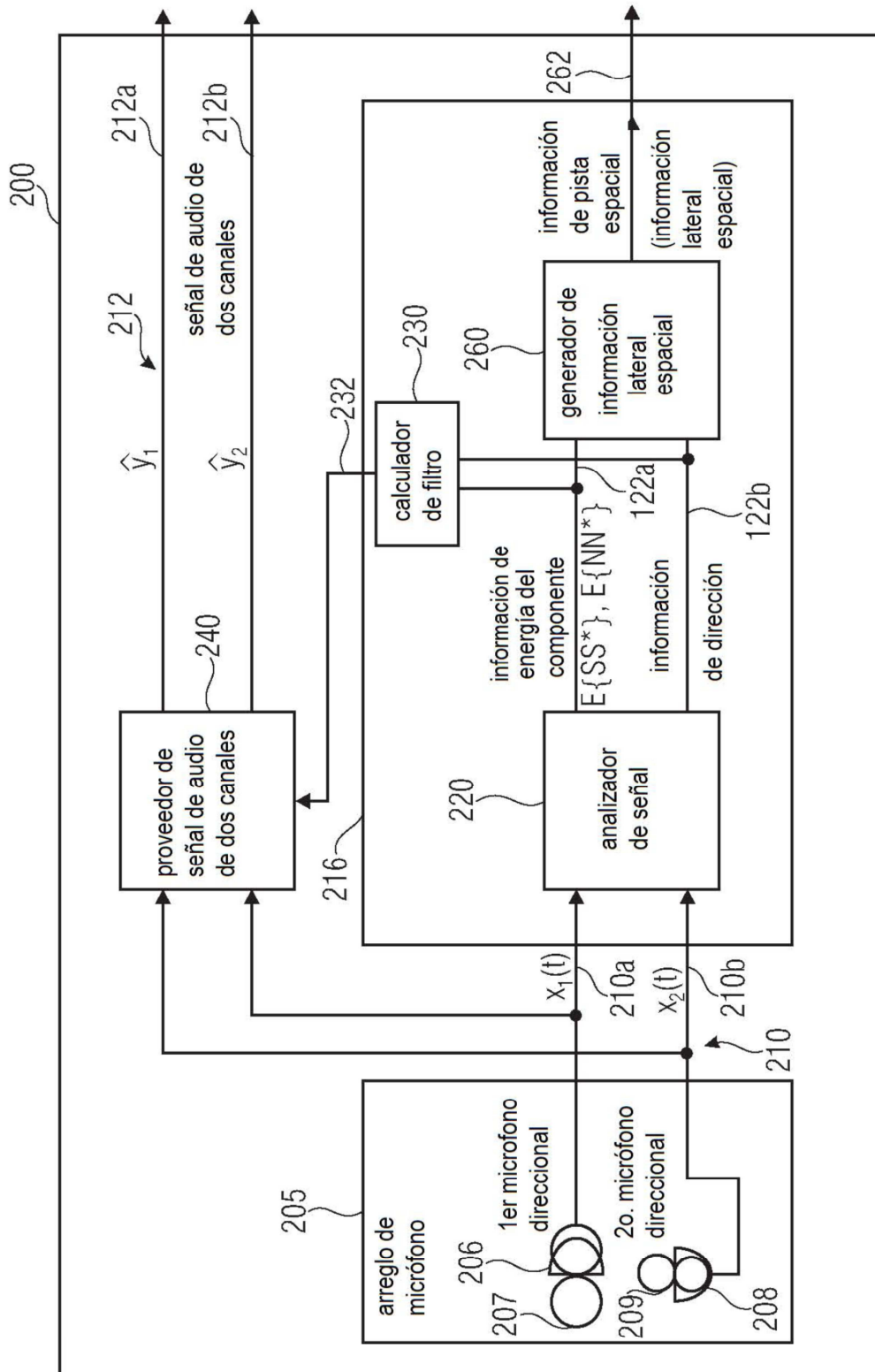


FIG 2

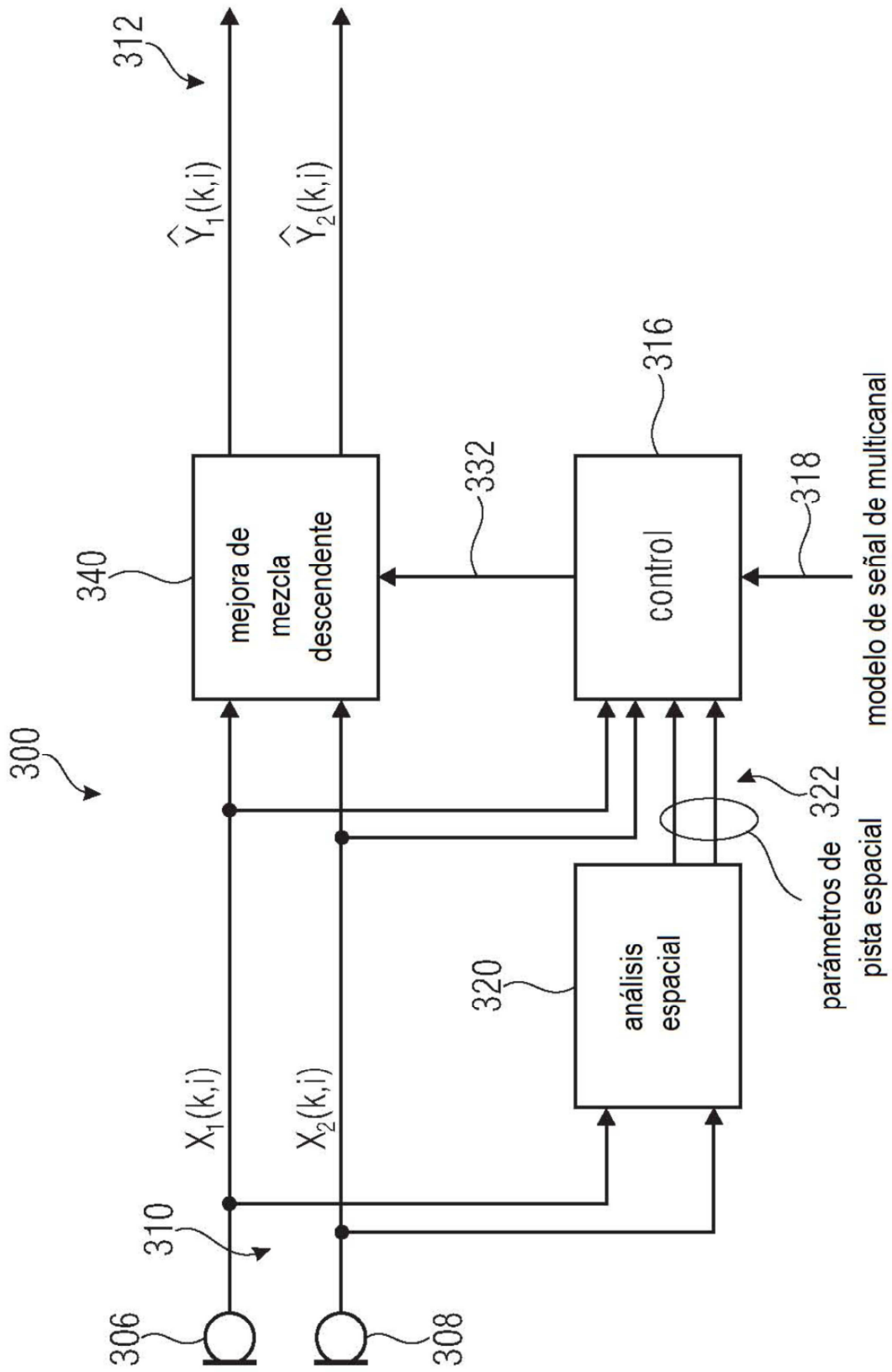


FIG 3

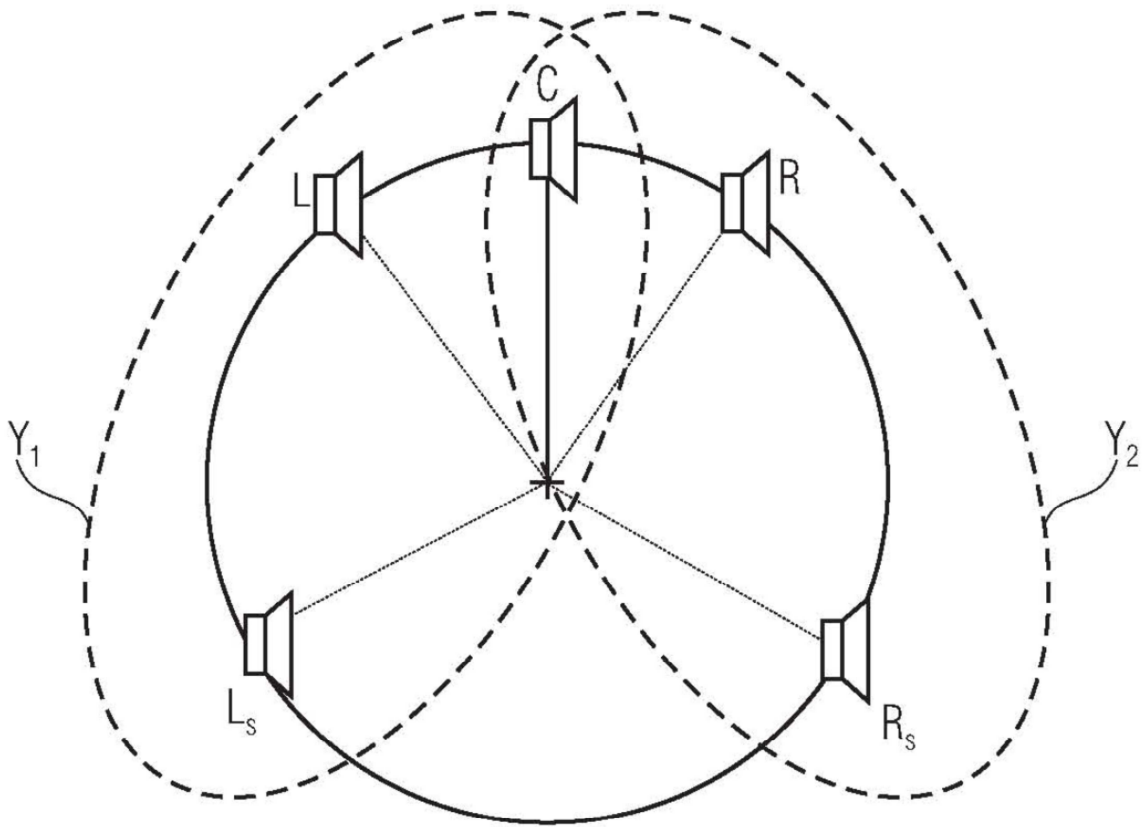


FIG 4



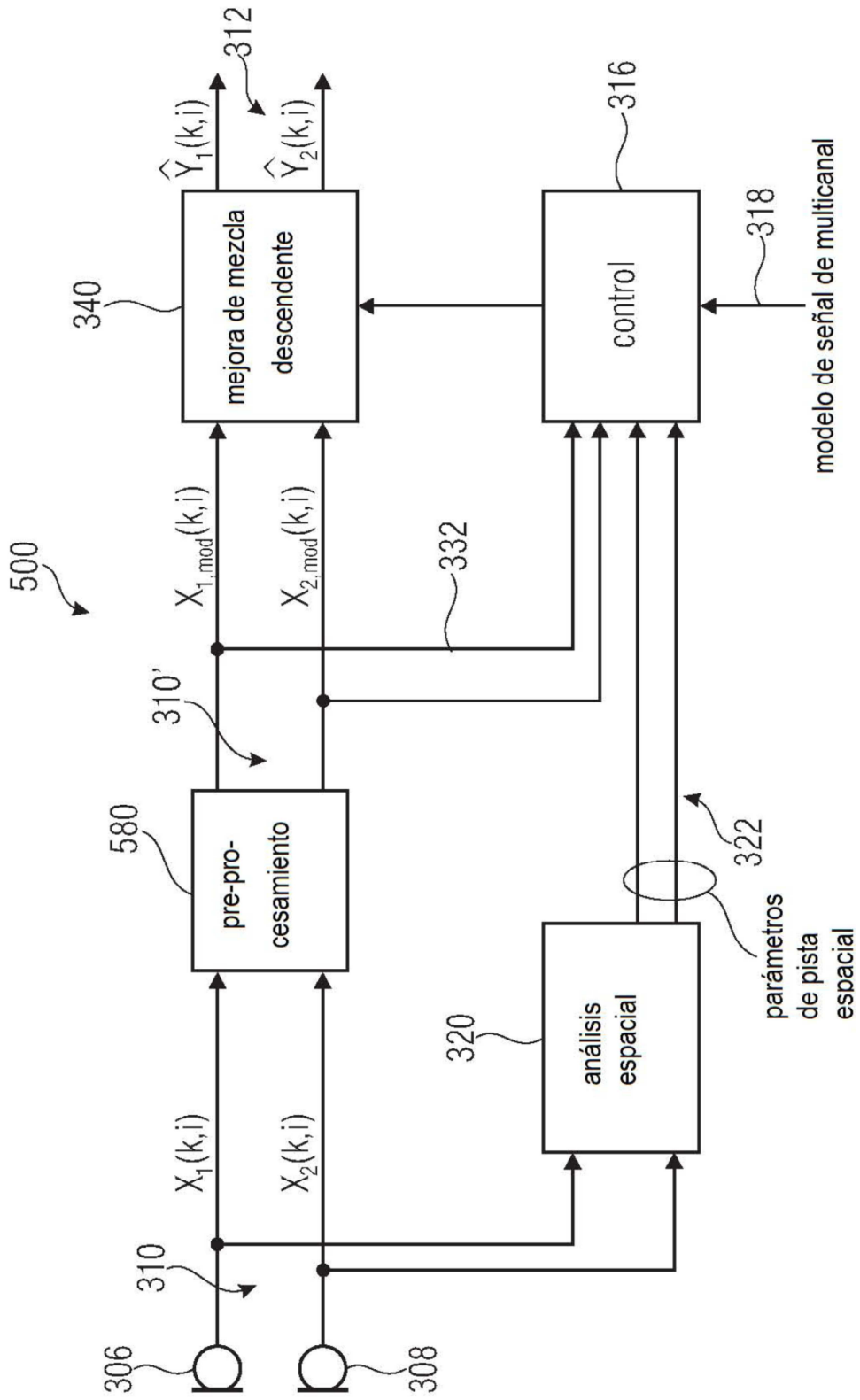


FIG 5

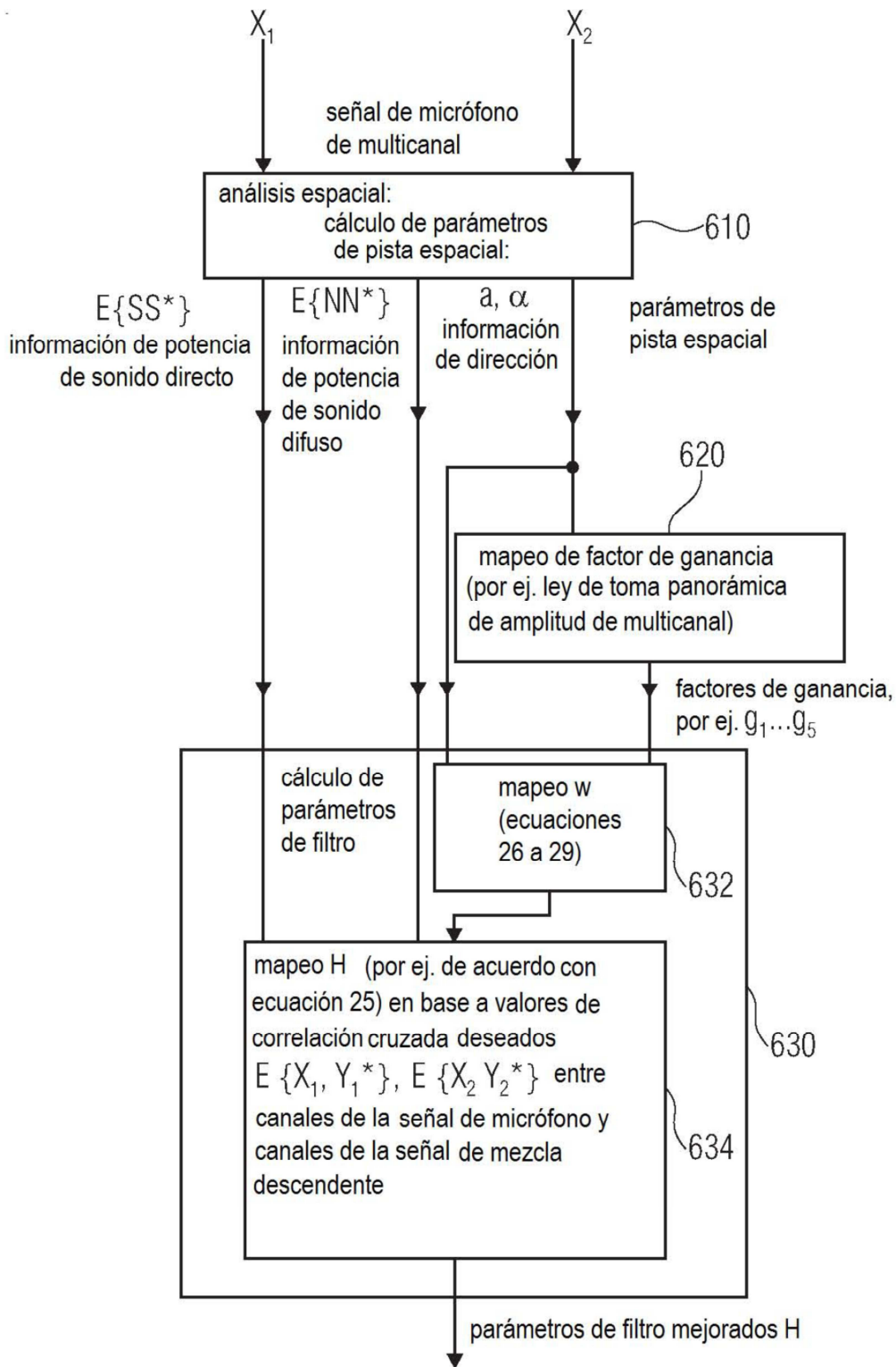


FIG 6

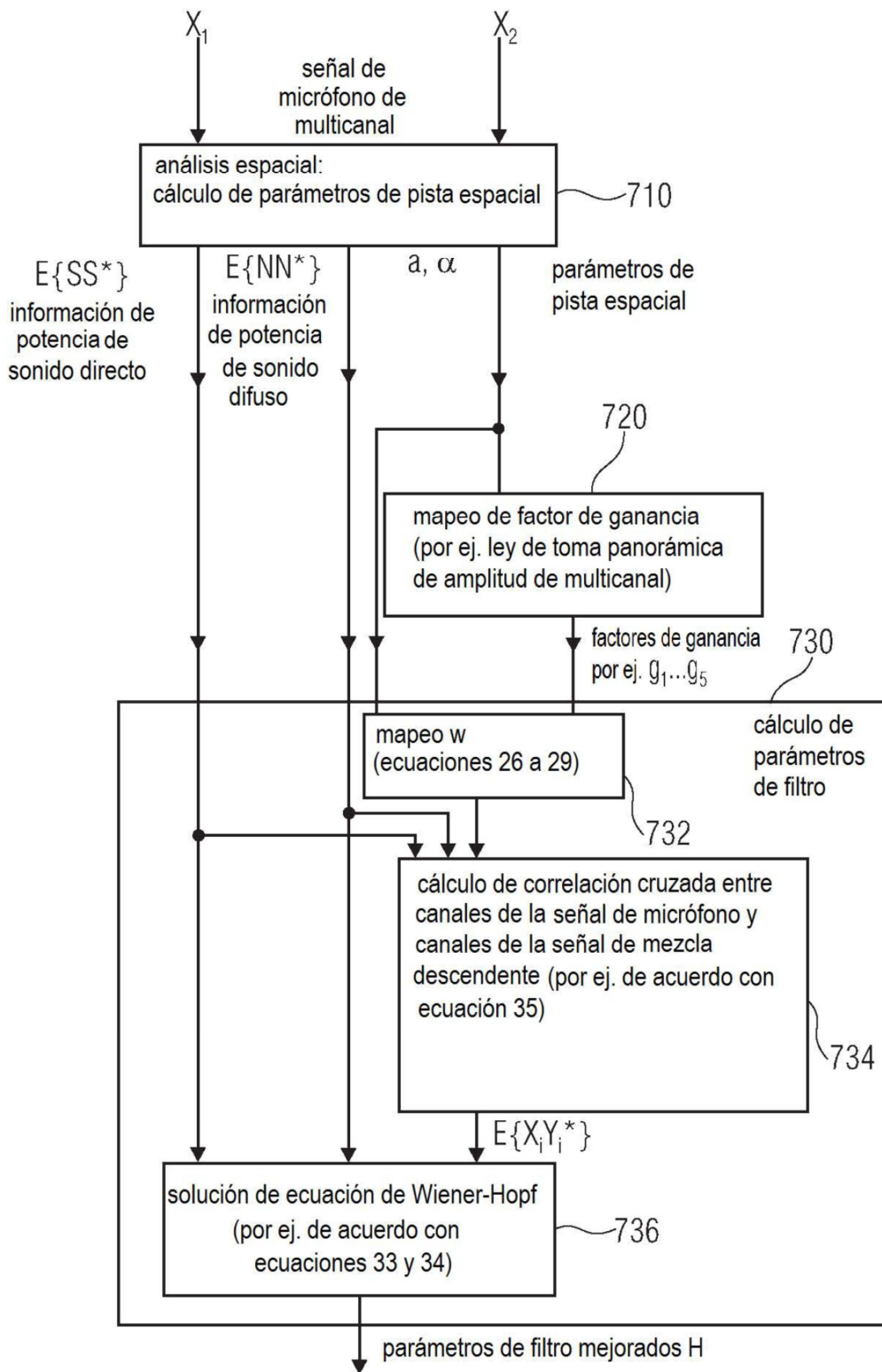


FIG 7