

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 980 688**

51 Int. Cl.:

**H04S 7/00**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **02.06.2021** **E 21177505 (1)**

97 Fecha y número de publicación de la concesión europea: **17.04.2024** **EP 3920557**

54 Título: **Control de altavoz**

30 Prioridad:

**05.06.2020 GB 202008547**

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

**02.10.2024**

73 Titular/es:

**AUDIOSCENIC LIMITED (100.0%)  
Suite A, Epsilon House, Enterprise Road  
Southampton Science Park  
Southampton, Hampshire SO16 7NS, GB**

72 Inventor/es:

**FAZI, FILIPPO;  
HAMDAN, ERIC;  
FRANCK, ANDREAS y  
SIMÓN, MARCOS**

74 Agente/Representante:

**GONZÁLEZ PECES, Gustavo Adolfo**

**ES 2 980 688 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

Control de altavoz

Campo

5 La presente descripción se refiere a un método para controlar un grupo de altavoces y un aparato y programa informático correspondientes.

Antecedentes

10 Se pueden usar grupos de altavoces para reproducir una pluralidad de señales de audio diferentes en una pluralidad de puntos de control. Las señales de audio que se aplican al grupo de altavoces se generan utilizando filtros, que pueden diseñarse para evitar la diafonía. Sin embargo, la determinación de los pesos de estos filtros puede resultar costosa computacionalmente, particularmente si los puntos de control se están moviendo y, por lo tanto, los pesos de los filtros deben calcularse en tiempo real. Este puede ser el caso, por ejemplo, si los puntos de control corresponden a las posiciones de los oyentes en un entorno acústico.

En el documento WO 2017/158338 A1 se describe un enfoque anterior para determinar los pesos de filtro para un grupo de altavoces.

15 Xie Bosun, "Capítulo 9 Reproducción binaural a través de altavoces", Función de transferencia relacionada con la cabeza y visualización auditiva virtual, J. Ross Publishing, (20130101), páginas 283 - 326, ISBN 978-1-60427-070-9, direcciones XP055856269, según su introducción, cuestiones relativas a la conversión de señales binaurales para la reproducción de altavoces. Se introducen los conceptos de cancelación de diafonía y procesamiento transaural para la reproducción de altavoces. También se analiza el diseño y la implementación de la cancelación de diafonía y del procesamiento transaural.

20 Resumen

Los aspectos de la presente descripción se definen en las reivindicaciones independientes adjuntas.

Breve descripción de los dibujos

A continuación se explicarán ejemplos de la presente descripción con referencia a los dibujos adjuntos en los que:

25 La Fig. 1 muestra un método para controlar un grupo de altavoces;

La Fig. 2 muestra un aparato para controlar un grupo de altavoces que puede usarse para implementar el método de la Fig. 1;

La Fig. 3a ilustra una aplicación de control de campo de sonido destinada a reproducir audio binaural en 3D realizando cancelación de diafonía y creando haces estrechos dirigidos a los oídos de los oyentes;

30 La Fig. 3b ilustra una aplicación de control de campo de sonido destinada a reproducir diferentes señales de contenido para diferentes oyentes;

La Fig. 3c ilustra una aplicación de control de campo de sonido que tiene como objetivo reproducir audio binaural en 3D realizando cancelación de diafonía y creando haces estrechos dirigidos a una pluralidad de oídos de los oyentes mientras también hace rebotar el sonido en las paredes del entorno para crear otras fuentes de imágenes en 3D;

35 La Fig. 3d ilustra el uso de un sistema de seguimiento de cabeza que estima la posición en 3D en tiempo real de un oyente con respecto a un grupo de altavoces;

La Fig. 4 muestra un diagrama de bloques de procesamiento de señales de un problema de control acústico subyacente para reproducir una pluralidad de señales acústicas en una pluralidad de puntos de control con un grupo de altavoces;

40 La Fig. 5 muestra un diagrama de procesamiento de señales simplificado de un proceso de control de múltiples entradas y múltiples salidas (MIMO) utilizado en el procesamiento de señales de matriz para reproducir M señales de entrada con L altavoces;

La Fig. 6 muestra un diagrama de procesamiento de señales simplificado de un enfoque de filtrado denominado "Tecnología 1" para reproducir M señales de entrada con L altavoces;

45 La Fig. 7 muestra un diagrama de procesamiento de señal ampliado del enfoque de Tecnología 1 que muestra los  $M \times M$  filtros independientes y los  $M \times L$  filtros dependientes;

La figura 8 muestra un diagrama de bloques de procesamiento de señales para un enfoque descrito en el presente documento, denominado "Tecnología 2";

La Fig. 9a ilustra un primer esquema de procesamiento de señales que divide el proceso de Tecnología 2 en múltiples bandas de frecuencia para permitir que los parámetros de procesamiento de señales tomen diferentes valores en diferentes bandas de frecuencia;

5 La Fig. 9b ilustra un segundo esquema de procesamiento de señales que divide el proceso de Tecnología 2 en múltiples bandas de frecuencia;

La Fig. 9c ilustra un tercer esquema de procesamiento de señales que divide el proceso de Tecnología 2 en múltiples bandas de frecuencia;

10 La Fig. 10a muestra los resultados de una simulación de los requisitos de potencia de procesamiento para filtros de grupo adaptativos para el oyente basados en el enfoque de Tecnología 1 en comparación con los enfoques MIMO estáticos y adaptativos para el oyente tradicionales; y

La Fig. 10b muestra una comparación del rendimiento de cancelación de diafonía entre filtros obtenidos utilizando el enfoque de Tecnología 1 y el enfoque de Tecnología 2 descritos en el presente documento.

A lo largo de la descripción y de los dibujos, números de referencia similares se refieren a partes similares.

#### Descripción detallada

15 En términos generales, la presente descripción se refiere a un método para controlar un grupo de altavoces para reproducir una pluralidad de señales de audio de entrada en una pluralidad respectiva de puntos de control de una manera que evita la diafonía, es decir, que reduce la magnitud en la que un audio la señal que se ha de reproducir en un primer punto de control también se reproduce en otros puntos de control. Se aplica un conjunto de filtros a las señales de audio de entrada para obtener la pluralidad de señales de audio de salida que se envían al grupo de altavoces. La  
20 presente descripción se refiere principalmente a modos de determinar esos filtros.

El objeto reivindicado corresponde a los ejemplos denominados en este documento como "Tecnología 2".

En la Fig. 1 se muestra un método para controlar el grupo de altavoces.

En la etapa S100, se reciben una pluralidad de señales de audio de entrada que se han de reproducir, mediante un grupo de altavoces, en una pluralidad respectiva de puntos de control en un entorno acústico.

25 En la etapa S110, la pluralidad de puntos de control se puede recibir usando un sensor de posición. En particular, se puede recibir o determinar la posición de cada uno de la pluralidad de puntos de control.

30 En la etapa S120, se puede determinar un conjunto de filtros. Si se realiza la etapa S110, el conjunto de filtros se puede determinar basándose en la pluralidad determinada de puntos de control. Alternativamente, el conjunto de filtros puede determinarse basándose en una pluralidad predeterminada de puntos de control. La manera en que se determina el conjunto de filtros se describe en detalle a continuación.

En la etapa S130, se determina una señal de audio de salida respectiva para cada uno de los altavoces en el grupo aplicando el conjunto de filtros a la pluralidad de señales de audio de entrada.

35 El conjunto de filtros se puede aplicar en el dominio de la frecuencia. En este caso, se aplica una transformada, tal como una transformada rápida de Fourier (FFT), a las señales de audio de entrada, se aplican los filtros y luego se aplica una transformada inversa para obtener las señales de audio de salida.

En la etapa S140, las señales de audio de salida pueden enviarse al grupo de altavoces.

Las etapas S100 a S140 pueden repetirse con otra pluralidad de señales de audio de entrada. A medida que se repiten las etapas S100 a S140, el conjunto de filtros puede permanecer igual, en cuyo caso no es necesario realizar la etapa S120, o puede cambiar.

40 Como entendería un experto en la técnica, las etapas de la Fig. 1 se pueden realizar con respecto a tramas recibidas sucesivamente de una pluralidad de señales de audio de entrada. Por consiguiente, no es necesario completar todas las etapas S100 a S140 antes de comenzar a repetirse. Por ejemplo, en algunas implementaciones, la etapa S100 se realiza una segunda vez antes de que se haya realizado la etapa S140 por primera vez.

45 En la Fig. 2 se muestra un diagrama de bloques de un aparato ejemplar 200 para implementar cualquiera de los métodos descritos en el presente documento, tal como el método de la Fig. 1. El aparato 200 comprende un procesador 210 (por ejemplo, un procesador de señales digitales) dispuesto para ejecutar instrucciones legibles por ordenador que pueden proporcionarse al aparato 200 a través de una o más de una memoria 220, una interfaz 230 de red o una interfaz 250 de entrada.

50 La memoria 220, por ejemplo una memoria de acceso aleatorio (RAM), está dispuesta para poder recuperar, almacenar y proporcionar al procesador 210, instrucciones y datos que se han almacenado en la memoria 220. La interfaz 230 de

red está dispuesta para permitir que el procesador 210 se comunique con una red de comunicaciones, tal como Internet. La interfaz 250 de entrada está dispuesta para recibir entradas de usuario proporcionadas a través de un dispositivo de entrada (no mostrado) tal como un ratón, un teclado o una pantalla táctil. El procesador 210 puede acoplarse además a un adaptador 240 de visualización, que a su vez está acoplado a un dispositivo de visualización (no mostrado). El procesador 210 puede acoplarse además a una interfaz 260 de audio que puede usarse para emitir señales de audio a uno o más dispositivos de audio, tales como un grupo 300 de altavoces. La interfaz 260 de audio puede comprender un convertidor de digital a analógico (DAC) (no mostrado), por ejemplo, para utilizar con dispositivos de audio con entrada o entradas analógicas.

A continuación se describen varios enfoques para determinar el conjunto de filtros.

## 10 Contexto

Los sistemas de audio 3D con cancelación de diafonía (CTC) basados en adaptación al oyente se basan en múltiples filtros de control para generar el sonido que activa uno o más altavoces. Los parámetros de estos filtros se adaptan en tiempo real según la posición instantánea de uno o más oyentes, que se estima con un dispositivo de seguimiento de oyentes (por ejemplo, una cámara, un dispositivo de sistema de posicionamiento global o un dispositivo portátil). Esta adaptación de los parámetros del filtro requiere costosos recursos computacionales, lo que dificulta el uso de tales enfoques de reproducción de audio para pequeños dispositivos integrados. Parte del consumo de recursos computacionales proviene de la necesidad de múltiples filtros inversos, lo que se deriva del uso de modelos de funciones de transferencia complejos y precisos entre los altavoces del sistema y los oídos de un oyente determinado. Se pueden utilizar funciones de transferencia acústica más simples para reducir la carga computacional, pero esto tiene el coste de una calidad reducida del audio reproducido, especialmente en términos de sus atributos espaciales percibidos. Por lo tanto, es difícil crear un sistema que sea adaptativo, tenga una carga computacional baja y tenga un rendimiento de alta calidad.

Los sistemas CTC adaptables al oyente pueden basarse en disposiciones de altavoces estéreo. Los sistemas adaptables al oyente también pueden utilizar disposiciones de cuatro altavoces para brindarle al oyente la capacidad de girar su cabeza y escuchar sonidos procedentes de un entorno de 360 grados. Estos ejemplos de sistemas CTC adaptables al oyente utilizan enfoques de control de procesamiento de señales que varían en el tiempo para adaptarse a las posiciones del oyente y a las orientaciones de la cabeza que varían en el tiempo. Los filtros de control pueden leerse desde una base de datos o calcularse sobre la marcha con un coste computacional significativo. Si bien estos enfoques de procesamiento de señales pueden implementarse utilizando grandes unidades centrales de procesamiento (CPU), como las disponibles en los ordenadores personales (PC), su procesamiento de señales subyacente resulta un factor limitante en los sistemas integrados cuando se utilizan más de dos altavoces.

Los sistemas de audio en 3D basados en CTC tienen una respuesta mejorada cuando se utilizan más de dos altavoces. Estos se pueden utilizar con un enfoque fijo y adaptable sin oyente. Sin embargo, dicho enfoque puede no ser adecuado para aplicaciones de consumo, ya que suponen que el oyente permanece quieto en una única posición de audición.

Desde el punto de vista del procesamiento de señales, el principal problema con muchos enfoques es que se basan en flujos de señales 'clásicos' de múltiples entradas y múltiples salidas (MIMO) que requieren  $M \times L$  filtros de control, siendo  $M$  el número de puntos de control de presión acústica (normalmente uno para cada uno de los oídos de los oyentes) y  $L$  el número de altavoces del grupo de altavoces. Para un sistema de dos altavoces, sólo se necesitarían cuatro filtros; sin embargo, se necesitaría el doble si el sistema se adaptara al oyente, y si se van a utilizar más altavoces, el costo de procesamiento crece muy rápidamente.

La tecnología descrita en el documento WO 2017/158338 A1, en lo sucesivo denominada "Tecnología 1", permite una reproducción de audio adaptativo al oyente de procesamiento eficiente con grupos de altavoces que utilizan más de dos altavoces. La principal reducción de la sobrecarga (o consumo) de la CPU introducida por la Tecnología 1 resulta de descomponer el flujo de audio del procesamiento de la señal de filtrado en una combinación de filtros dependientes del altavoz (DF) y de filtros independientes del altavoz (IF). En la Tecnología 1, los IF se implementan como un conjunto de filtros de respuesta de impulso finito (FIR) variables en el tiempo, mientras que los DF se implementan como un conjunto de elementos de retardo de ganancia variables en el tiempo. Debido a esta descomposición, sólo se necesitan  $M \times M$  filtros de control y  $M$  líneas de retardo con  $L$  puntos de lectura cada una. Este esquema de procesamiento introduce una gran reducción en la complejidad del procesamiento en comparación con la matriz de  $M \times L$  filtros necesaria para otros enfoques, ya que en la mayoría de las implementaciones  $L$  es mucho mayor que  $M$ .

Sin embargo, los ahorros de procesamiento introducidos por la Tecnología 1 requieren que la función de transferencia acústica entre cada altavoz y los puntos de control de presión acústica sea representable con ganancias lineales independientes de fase y frecuencia, por ejemplo, suponiendo un modelo de propagación puntual-de un solo polo de campo libre. Sin embargo, puede resultar útil utilizar una función de transferencia más compleja que mejoraría significativamente la calidad percibida de las imágenes de sonido virtual y que no puede representarse mediante simples ganancias y retardos.

## Visión general de la Tecnología 1

Los sistemas de control de campo de sonido basados en grupos de altavoces tienen como objetivo reproducir una o más

señales acústicas en uno o más puntos del espacio (puntos de control), al mismo tiempo que eliminan la diafonía acústica (o fuga de sonido) a otros puntos de control. Dicho control acústico conduce a la creación de haces de sonido estrechos que pueden ser controlados o u orientados direccionalmente en el espacio de una manera precisa para facilitar diversas aplicaciones acústicas.

5 Por ejemplo, una aplicación puede controlar con precisión la presión en los oídos de uno o más oyentes 341, 342, 343 para crear "auriculares virtuales" y reproducir sonido en 3D, lo que se conoce como cancelación de diafonía (CTC), como se ilustra en la Fig. 3a. Otra aplicación puede ser para reproducir varios haces 320 de sonido diferentes e independientes a dos o más oyentes, de modo que cada uno de ellos pueda oír un programa de sonido diferente o el mismo programa con un nivel de sonido específico del usuario, como se ilustra en la Fig. 3b. Como los haces 320 de sonido controlan el  
10 campo de sonido alrededor de los oídos, estas técnicas de control son conocidas por la "capacidad de personalizar el sonido alrededor de los oyentes". Además, los haces creados por el grupo 300 de altavoces se pueden controlar para dirigir también el sonido hacia las paredes 330 de la habitación donde se reproduce el sonido. Este sonido rebota en las paredes y llega al oyente u oyentes, creando así una experiencia de inmersión, como se ilustra en la Fig. 3c.

15 Un grupo de altavoces de canal L comprende altavoces ubicados en las posiciones  $y_1, y_2, \dots, y_L \in \mathbb{R}^3$ . Para una frecuencia de reproducción dada  $\omega = 2\pi f$  en radianes por segundo, el objetivo es reproducir un conjunto de M señales de audio  $d(\omega) = [d_1(\omega), \dots, d_M(\omega)]^T$  que son representadas por M haces creados por el grupo de altavoces, en un conjunto de puntos de control  $x_1, \dots, x_M \in \mathbb{R}^3$ . El oyente es libre de moverse alrededor en el espacio de audición y la posición de los puntos de control  $\{x_m\}$  puede variar en el espacio. Para permitir esto, la posición espacial instantánea de los puntos de control  $\{x_m\}$  puede recopilarse mediante un sistema 310 de seguimiento de oyentes (cámara, portátil, láser, basado en  
20 sonido) que proporciona las coordenadas en tiempo real de los oídos de los oyentes con respecto a cada uno de los altavoces del grupo de altavoces, como se muestra en la Fig. 3d.

En la Fig. 4 se muestra un diagrama de bloques del problema de control de la presión acústica reproducido por un grupo de altavoces. El problema de control acústico subyacente se puede expresar en el dominio de frecuencia como

$$P(\omega) = S(\omega)H(\omega)d(\omega), \quad (1)$$

25 donde  $p(\omega) = [p_1(\omega), \dots, p_M(\omega)]^T$  contiene las señales de presión acústica reproducidas en los diferentes puntos de control  $x_m$ ,  $(\cdot)^T$  denota el vector o matriz transpuesta,  $S(\omega) \in \mathbb{C}^{M \times L}$  es la denominada matriz de planta cuyos elementos son las funciones de transferencia acústica entre las L fuentes y los M puntos de control, y  $H(\omega) \in \mathbb{C}^{L \times M}$  es la matriz de filtros de control diseñada para permitir la reproducción de señales  $d(\omega)$  de audio de entrada en los puntos de control, dado  $S(\omega)$ . Cada columna  $h_m$  de H está diseñada para reproducir su correspondiente señal  $d_m$  de audio en el punto de control  $x_m$ , minimizando al mismo tiempo la presión irradiada en los otros puntos de control. La dependencia de  $\omega$  se omitirá en lo que sigue a menos que sea necesario.

El objetivo final del sistema de control de sonido es obtener

$$p = e^{-j\omega T} d, \quad (2)$$

35 donde  $j = \sqrt{-1}$  y  $e^{-j\omega T}$  es un retardo de modelado utilizado para asegurar la causalidad de la solución. Esta condición se cumple si  $SH = e^{-j\omega T} I$ , donde I es la matriz de identidad M x M. Un enfoque que permite que esta condición se cumpla aproximadamente es calcular H como la matriz pseudoinversa regularizada de S, en particular

$$H = e^{-j\omega T} S^H [SS^H + A]^{-1} \quad (3)$$

40 donde A es una matriz de regularización y  $(\cdot)^H$  denota la traspuesta Hermitiana. La ecuación anterior puede denominarse solución pseudoinversa para un sistema indeterminado y, por lo tanto, el conjunto de filtros de control que devuelve pueden denominarse como filtros "inversos". Dicho sistema tendrá M entradas para M señales de audio y L salidas para los L altavoces del grupo, como se muestra en el diagrama de bloques de la Fig. 5. Para el caso de un sistema MIMO como los utilizados en el procesamiento de señales de grupo clásico, se necesitan M x L filtros de control.

En el procesamiento de señales del grupo, los filtros H de control del grupo se calculan para una matriz de planta acústica determinada, S. La matriz de planta es un modelo de las funciones de transferencia electroacústica entre los  
45 altavoces del grupo y los puntos de control donde se ha de controlar la presión acústica. Idealmente, la matriz de planta caracterizará con la mayor precisión posible la función de transferencia física encontrada en un sistema acústico práctico. Sin embargo, esto no siempre es posible en aplicaciones prácticas. Aunque es posible realizar mediciones acústicas y estimar la matriz de planta de un sistema determinado con un grado relativamente alto de precisión, este es un proceso complejo que sólo puede realizarse con precisión en condiciones de laboratorio. Además, la matriz de planta puede  
50 cambiar significativamente incluso con pequeños movimientos del oyente u oyentes, lo que requiere una densa cuadrícula de mediciones para permitir un amplio intervalo de adaptabilidad a los movimientos del oyente. Además, este enfoque da como resultado un conjunto de L x M filtros inversos complejos, lo que provoca una alta complejidad computacional para la reconstrucción. Por lo tanto, es útil utilizar modelos de propagación acústica muy simples pero precisos para representar la matriz de planta S.

55 Un caso particular es cuando la matriz de planta S se aproxima mediante una matriz simple C que se forma suponiendo

un modelo de propagación acústica de fuente puntual de campo libre entre cada uno de los altavoces y los puntos de control de presión acústica. Por lo tanto, la matriz C se define como

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \vdots \\ \mathbf{c}_M \end{bmatrix}, \quad (4)$$

donde cada elemento de esta matriz está formado por un retardo y un elemento de ganancia, por ejemplo,

$$\mathbf{c}_m = \left[ \frac{e^{-jk_r m1}}{r_{m1}}, \dots, \frac{e^{-jk_r mL}}{r_{mL}} \right], \quad (5)$$

donde  $k = \omega/c_0$  es el número de onda y  $c_0$  la velocidad del sonido en el aire y  $r_{ml}$  es un número real independiente de la frecuencia que depende de la distancia entre el emésimo punto de control acústico y el centro acústico del  $l^{\text{ésimo}}$  altavoz. Utilizar dicho modelo de propagación permite que los elementos de la matriz C sean calculados fácilmente una vez que se conocen las posiciones de los puntos de control con respecto al grupo de altavoces, por lo que se requiere un procesamiento modesto para calcular un nuevo conjunto de filtros H de control.

Aunque utilizar un modelo electroacústico simple es útil para reducir la cantidad de cálculos necesarios para obtener un nuevo conjunto de filtros, también es útil para reducir la cantidad de operaciones de bajo nivel necesarias para filtrar una cantidad determinada de contenido de audio digital. Se puede realizar una simplificación adicional analizando la estructura de la ecuación (3), que es la fórmula de la pseudoinversa de un problema de mínimos cuadrados indeterminado. Un análisis cuidadoso muestra que algunos términos (elementos filtrantes) son comunes a algunas de las salidas/altavoces. Estos se denominan como filtros independientes (IF). Otros términos son específicos sólo de algunos de los altavoces y se denominan filtros dependientes (DF). Los términos de la ecuación (3), y por lo tanto la arquitectura de procesamiento de señales resultante, se pueden agrupar de la siguiente manera:

$$\mathbf{H} = \overbrace{e^{-j\omega T_1} \mathbf{C}^H}^{\text{DFs}} \overbrace{e^{-j\omega T_2} [\mathbf{C}\mathbf{C}^H + \mathbf{A}]^{-1}}^{\text{IFs}}. \quad (6)$$

donde  $T_1$  y  $T_2$  son retardos que satisfacen la relación  $T_1 + T_2 = T$ . Esto hace posible dividir el procesamiento de señales en la ecuación (6) en un conjunto de  $M \times M$  IF y un conjunto de  $L \times M$  DF. Esto conduce al esquema de procesamiento de señales que se muestra en la Fig. 6, que se muestra en su forma ampliada en la Fig. 7.

Una de las peculiaridades de este procesamiento de señales de grupo es que es posible implementar los  $M \times M$  IF utilizando filtrado FIR convencional (variable en el tiempo) y los  $M \times L$  DF utilizando M líneas de retardo (variables en el tiempo) con L puntos de acceso cada una. En este punto, los DF actúan como un formador de haz de retardo y suma. En comparación con un enfoque de filtrado MIMO tradicional basado en  $M \times L$  filtros variables, esta implementación introduce una gran reducción en el coste computacional necesario para filtrar una cierta cantidad de audio digital, permitiendo así un número reducido de operaciones de coma flotante por segundo (FLOPS) y para que el procesamiento se integre en dispositivos más pequeños. El único requisito para lograr esta reducción en la complejidad del cálculo es que los elementos de la matriz C incluyan solo ganancias y retardos independientes de la frecuencia.

#### Enfoque de Tecnología 2

Puede ser útil usar modelos de función de transferencia dependientes de la frecuencia, más precisos que los proporcionados por la matriz C introducida anteriormente. Por ejemplo, puede ser deseable utilizar funciones de transferencia relacionadas con la cabeza (HRTF) de esfera rígida o medidas para la cancelación de diafonía para tener en cuenta la difracción de la cabeza de los oyentes y así mejorar la calidad del audio espacial, o puede ser útil compensar la respuesta de frecuencia y la direccionalidad de los altavoces, o compensar la difracción de otros elementos del entorno.

Una forma de lograr esto es sustituir la matriz simple C por una matriz G más compleja que proporcione una mejor aproximación de la matriz S de función de transferencia física. La matriz G podría, por ejemplo, crearse midiendo la función S de transferencia física, en cuyo caso, los elementos de G podrían ser, por ejemplo, funciones de transferencia relacionadas con la cabeza, o mediante el uso de un modelo analítico o numérico de S, tal como una esfera rígida o un modelo de elemento límite de una cabeza humana. Sin embargo, en este caso, los elementos de G no serán simples retardos y ganancias como en el caso de C, sino que se basarán en datos o funciones más complejas dependientes de la frecuencia. Si dicha matriz G tuviera que ser utilizada en la ecuación (6) para el cálculo del filtro digital, esto conduciría, por un lado, a un mejor rendimiento de calidad de audio del sistema, pero, por otro lado, requeriría un rendimiento mucho más complejo de DF, conduciendo así a un aumento significativo de la carga computacional general.

Los inventores han llegado a la conclusión de que la calidad de audio de la Tecnología 1 se puede mejorar significativamente sin aumentar significativamente la carga computacional utilizando tanto una matriz G más precisa,

relativamente compleja y una matriz C menos precisa, relativamente simple.

En primer lugar, se recuerda que, dado que el objetivo de la etapa de diseño del filtro es  $p = e^{j\omega T}d$ , donde  $p = SHd$ , el filtro H debe ser tal que

$$SH \approx e^{-j\omega T}I \quad (7)$$

5 donde I es la matriz de identidad  $M \times M$ .

La ecuación (6) para el cálculo de H se sustituye por (ignorando por el momento la matriz A de regularización)

$$H = \overbrace{e^{-j\omega T_1} C^H}^{\text{DFs}} \overbrace{e^{-j\omega T_2} [GC^H]^{-1}}^{\text{IFs}}. \quad (8)$$

$SC^H[GC^H]^{-1}$  proporciona una aproximación mucho mejor a la matriz I de identidad que lo hace  $SC^H[CC^H]^{-1}$  ya que G es una aproximación mucho mejor a S que lo es C. Esto permite una calidad de audio significativamente mejorada.

10 Sin embargo, el uso de la matriz G, más precisa pero computacionalmente más compleja, se limita a los IF, mientras que los DF son las ganancias y retardos simples contenidos en la matriz  $e^{j\omega T_1}C^H$ . Esto permite un costo computacional mucho menor que el que se requeriría si también se usara la matriz  $G^H$  para los DF.

En este caso, el problema directo de las presiones acústicas ahora viene dado como

$$p = e^{-j\omega T} SC^H [GC^H]^{-1} d. \quad (9)$$

15 También es posible aplicar un esquema de regularización (por ejemplo, la regularización de Tikhonov) al diseño de los IF. En este caso, la ecuación (8) se reescribe como

$$H = e^{-j\omega T} C^H [GC^H + A]^{-1} \quad (10)$$

donde A es una matriz de regularización utilizada para controlar la energía de los filtros de grupo. El diagrama de bloques correspondiente a esta arquitectura de procesamiento de señales digitales (DSP) se muestra en la Fig. 8. Se puede observar cómo los filtros H se pueden dividir en  $M \times M$  filtros independientes IF y  $M \times L$  filtros dependientes DF.

20

Una forma alternativa de calcular los filtros independientes IF es resolver un problema de optimización (convexo)

$$\underset{\text{IFs}}{\operatorname{argmin}} \left\| GC^H \text{IFs} - e^{-j\omega T} I \right\|_{p_1} \quad (11)$$

$$\text{subject to } \left\| C^H \text{IFs} \right\|_{p_2} \leq H_{\max}. \quad (12)$$

Aquí  $\|\cdot\|_{p_1}$  y  $\|\cdot\|_{p_2}$  representan normas matriciales adecuadas, por ejemplo la norma de Frobenius, y  $H_{\max}$  es un límite superior admisible en la norma de la matriz de filtros H en grupo.

25

Vale la pena señalar en este punto que las combinaciones de las matrices G y C ofrecen otras posibilidades para crear filtros de control de grupo que pueden beneficiarse del uso de este enfoque de control híbrido y un modelo de función de transferencia más realista. Por ejemplo, puede ser útil emplear enfoques de control "ponderados" para ajustar la contribución procedente de cualquier altavoz elegido para controlar la presión acústica en cualquiera de los puntos de control, calculando H como

30

$$H = e^{-j\omega T} W_L C^H [GW_L C^H + A]^{-1}, \quad (13)$$

donde en este caso  $W_L$  es una matriz de ponderación diagonal  $L \times L$  que contiene pesos positivos para cada altavoz.

Un enfoque similar puede resultar útil para algunos de los casos de uso donde se desea controlar la presión acústica en cada uno de los puntos de control de una manera diferente. En este caso, se puede utilizar una matriz  $W_M$  con tamaño  $M \times M$  que contiene pesos positivos, donde los filtros de control vienen dados por:

35

$$H = [GW_M C^H + A]^{-1} C^H W_M e^{-j\omega T}. \quad (14)$$

A continuación se definen el siguiente conjunto de términos:

- Los elementos de la matriz G recién introducida, es decir,  $G_{m,l}$ , tienen la forma  $G_{m,l} = G_0(x_m, y_l, \omega) e^{-j\omega \tau(x_m, y_l)}$ , donde  $\tau(x_m, y_l)$  es una posición de retardo dependiente que depende de la posición de cada altavoz y punto de control y  $G_0(x_m, y_l, \omega)$  es una función compleja dependiente de la frecuencia.

40

- Los elementos de C, es decir,  $C_{m,l}$ , están formados por ganancias y retardos de la forma  $C_{m,l} = e^{-j\omega\tau(x_m,y_l)}g_{m,l}$ .

Las ganancias de valor real  $g_{m,l}$  dependen de la posición relativa de los altavoces y los puntos de control.

El término de retardo  $\tau(x_m,y_l)$  incluido en la definición de  $G_{m,l}$  puede ser el mismo retardo que define el elemento correspondiente  $C_{m,l}$  de la matriz C.

- 5 El término de retardo  $\tau(x_m,y_l)$  se puede elegir de tal manera que la fase de los términos en la diagonal de la matriz  $GC^H$  sea lo más cercana posible a cero.

Por lo tanto, una posible elección del retardo es el valor  $\tau(x_m,y_l)$  tal que  $\omega\tau(x_m,y_l)$  es la mejor aproximación lineal (a través de la frecuencia) de la fase de  $G_{m,l}$ .

Otras posibilidades para el diseño de C se basan en el factor de colinealidad

$$10 \quad \gamma_{m,m'} = \frac{|\mathbf{g}_m \cdot \mathbf{c}_{m'}^H|}{\|\mathbf{g}_m\| \|\mathbf{c}_{m'}\|} \quad (15)$$

donde  $\|\cdot\|$  es el operador de norma  $\ell^2$  y  $\mathbf{c}_m$  y  $\mathbf{g}_m$  son la  $m$ ésima fila de la matriz C y la  $m$ ésima fila de la matriz G, respectivamente.

- 15 Una opción es elegir los términos de retardo  $\tau(x_m,y_l)$  y los términos de ganancia  $g_{m,l}$  de tal manera que el factor de colinealidad  $\gamma_{m,m'}$  se maximice (o aumente) para cada combinación de filas con índices  $m = m'$ , sobre un intervalo de frecuencia de interés.

Otra posibilidad es elegir los términos de retardo  $\tau(x_m,y_l)$  y los términos de ganancia  $g_{m,l}$  de tal manera que se logre un equilibrio óptimo entre maximizar (o aumentar) el factor de colinealidad para cada combinación de filas con índices  $m = m'$  y minimizar (o reducir) el factor de colinealidad para filas con índices  $m \neq m'$ , nuevamente sobre un intervalo de frecuencia de interés.

- 20 Como ejemplo, una posible formulación matemática de este problema de optimización es

$$\operatorname{argmax}_{\mathcal{T}, \mathcal{G}} \sum_{k=1}^K \zeta_k \sum_{m=1}^M \left[ \gamma_{m,m}(\omega_k) - \alpha_k \sum_{m' \neq m} \gamma_{m,m'}(\omega_k) \right] \quad (16)$$

donde los parámetros de diseño  $\alpha_k$  y  $\zeta_k$  son números reales no negativos y T y g son los conjuntos de todos los retardos  $\tau(x_m,y_l)$  y ganancias  $g_{m,l}$ , respectivamente.  $\{\omega_k\}_{k=1,\dots,K}$  es un conjunto de frecuencias que abarca el intervalo de frecuencias de interés (obsérvese que  $\gamma_{m,m'}$  es una cantidad dependiente de la frecuencia).

- 25 Una de las ventajas de este enfoque de optimización es que aumenta la estabilidad del sistema. Para el caso en el que  $M = 2$ , esto se demuestra por el hecho de que el valor absoluto de  $\det(GC^H)$ , el determinante de la matriz que se ha de invertir para el cálculo del filtro, es

$$\begin{aligned} |\det(GC^H)| &= |(\mathbf{c}_1^H \mathbf{g}_1)(\mathbf{c}_2^H \mathbf{g}_2) - (\mathbf{c}_1^H \mathbf{g}_2)(\mathbf{c}_2^H \mathbf{g}_1)| \\ &= \|\mathbf{c}_1\| \|\mathbf{c}_2\| \|\mathbf{g}_1\| \|\mathbf{g}_2\| \left| \gamma_{1,1}\gamma_{2,2} - \gamma_{2,1}\gamma_{1,2} e^{j\Phi} \right| \\ &\geq \|\mathbf{c}_1\| \|\mathbf{c}_2\| \|\mathbf{g}_1\| \|\mathbf{g}_2\| |\gamma_{1,1}\gamma_{2,2} - \gamma_{2,1}\gamma_{1,2}| \end{aligned} \quad (17)$$

donde  $\Phi$  es un término de fase.

- 30 Se puede ver que, si no se hace ninguna suposición con respecto a  $\Phi$ , maximizar (o aumentar)  $\gamma_{1,1}$  y  $\gamma_{2,2}$  y minimizar (o reducir)  $\gamma_{1,2}$  y  $\gamma_{2,1}$  maximiza (o aumenta) el valor absoluto del determinante y, por lo tanto, aumenta la estabilidad del sistema.

Los enfoques anteriores utilizan dos conjuntos de funciones de transferencia para calcular filtros de grupo y se denominan "Tecnología 2".

- 35 Implementación del banco de filtros

Para determinadas aplicaciones, puede resultar útil implementar versiones paralelas del mismo algoritmo de procesamiento de señales pero para diferentes bandas de frecuencia. Esto podría ser necesario, por ejemplo, si se utilizan diferentes tipos de activadores acústicos para diferentes intervalos de frecuencia (tweeters y woofers). En este caso, se podría utilizar un número diferente de altavoces  $L_n$  para cada banda diferente. Esto requiere que las matrices C y G se calculen de manera diferente para diferentes bandas de frecuencia, de modo que los elementos de estas matrices puedan tomar diferentes valores para  $n = [1,\dots,N]$  diferentes bandas de frecuencia. A continuación se describen tres

- 40

enfoques diferentes para lograr esto.

La primera arquitectura multibanda se muestra en la Fig. 9a. Se utiliza un conjunto de N filtros de paso de banda  $B_n$  en la entrada y el procesamiento central de Tecnología 2 se duplica N veces. En este caso, los IF y DF son diferentes para cada banda de frecuencia. Los filtros de paso de banda pueden ser alternativamente filtros de paso bajo o filtros de paso alto. En este caso, los IF y DF para la banda de frecuencia enésima se pueden definir como

$$IF_n = [\mathbf{G}_n \mathbf{C}_n^H + \mathbf{A}_n]^{-1} \quad (18)$$

$$DF_n = \mathbf{C}_n \quad (19)$$

donde las matrices  $\mathbf{G}_n, \mathbf{C}_n, \mathbf{A}_n$  son como se han definido anteriormente en este documento, pero con valores de parámetros específicos para la enésima banda de frecuencia. Con estas definiciones de IF y DF, las señales  $\mathbf{q}_n$  de altavoz  $L_n$  correspondientes a la enésima banda de frecuencia vienen dadas, en el dominio de frecuencia, por

$$\mathbf{q}_n = \mathbf{C}_n^H [\mathbf{G}_n \mathbf{C}_n^H + \mathbf{A}_n]^{-1} B_n \mathbf{d} \quad (20)$$

En la figura 9b se muestra una segunda posible arquitectura DSP multibanda. En este caso, los IF tienen en cuenta los distintos retardos en matrices  $\mathbf{C}_n$ , diferentes para cada banda de frecuencia, y la salida de los IF se divide posteriormente en N bandas de frecuencia que se alimentan a N conjuntos de DF con diferentes valores del retardo escalado para cada banda de frecuencia. Este esquema requiere el uso de sólo  $M \times M$  IF, en lugar de tener un conjunto diferente de IF para cada banda de frecuencia. Estos IF se pueden definir como

$$IF = \sum_{n=1}^N W_n [\mathbf{G}_n \mathbf{C}_n^H + \mathbf{A}_n]^{-1} \quad (21)$$

donde  $W_n$  es una función de ponderación de frecuencia que depende principalmente de los filtros  $B_n$  de paso de banda y puede tener valores complejos. Los DF se pueden calcular como en la ecuación (19).

En la figura 9c se muestra una tercera posible arquitectura DSP multibanda. En este caso, el procesamiento multibanda se incluye tanto en los IF como en los DF, de modo que se requiere un único conjunto de  $M \times M$  IF y de  $M \times L$  DF (a diferencia de un conjunto diferente para cada banda de frecuencia). Los IF se pueden definir como en la ecuación (21), mientras que los DF se pueden definir como

$$DF = \sum_{n=1}^N B_n \mathbf{C}_n^H \quad (22)$$

Con este enfoque, los DF ya no son elementos de retardo-ganancia. En este tercer enfoque, las señales relacionadas con las distintas bandas de frecuencia se suman juntas, para cada altavoz determinado. Por lo tanto, este método no es adecuado en casos en los que se utilizan diferentes controladores acústicos para diferentes bandas de frecuencia (tweeter y woofer). Sin embargo, existen otras aplicaciones en las que este enfoque puede resultar útil, por ejemplo cuando los retardos de grupo de los elementos de  $\mathbf{G}$  se aproximan mejor mediante diferentes retardos en diferentes bandas de frecuencia. Con las definiciones de IF y DF anteriores, las señales  $\mathbf{q}$  de altavoz  $L$  están dadas, en el dominio de la frecuencia, por

$$\mathbf{q} = \left[ \sum_{n=1}^N B_n \mathbf{C}_n^H \right] \left[ \sum_{m=1}^N W_m (\mathbf{G}_m \mathbf{C}_m^H + \mathbf{A}_m)^{-1} \right] \mathbf{d} \quad (23)$$

#### Efectos de los enfoques de Tecnología 1 y Tecnología 2

La figura 10a muestra los resultados de una simulación de los requisitos de potencia de procesamiento para filtros de grupo adaptables al oyente basados en el enfoque de Tecnología 1 en comparación con los enfoques MIMO estáticos y adaptables al oyente tradicionales. Específicamente, el número de MFLOPS requeridos en función del número de altavoces  $L$  se muestra para un enfoque MIMO estático 1001, un enfoque MIMO 1002 adaptable al oyente y el enfoque 1003 de Tecnología 1.

Para ilustrar la ventaja que proporciona el enfoque de Tecnología 2, en la Fig. 10b se muestran los resultados de una simulación para un grupo de altavoces con tres altavoces. En esta simulación, se muestra el espectro CTC, que representa la separación de canales de las señales acústicas suministradas a los oídos de un oyente. Idealmente, esta métrica de rendimiento debe ser lo más grande posible para que una matriz que entregue sonido en 3D a través de CTC proporcione una buena inmersión en 3D. Como se observa en la Fig. 10b, el rendimiento de 1004 de la Tecnología 2 es mucho mejor que el de 1005 de la Tecnología 1 en el intervalo de frecuencia de audio, particularmente por encima de 2

kHz, donde los efectos de la difracción de cabeza son grandes.

5 El enfoque de la Tecnología 2 combina la simplicidad y el bajo coste computacional de la Tecnología 1, debido a la presencia de DF simples representados por la matriz  $C^H$ , pero también permite la introducción de una matriz de planta  $G$  más precisa en el cálculo de los IF, sin un aumento significativo del coste computacional general del algoritmo. Esto permite tener en cuenta y compensar fenómenos acústicos complejos (tales como la difracción debida a la cabeza o los reflejos del entorno acústico) y mejorar así la calidad del audio reproducido.

Un efecto de la presente descripción es proporcionar un esquema de cálculo de filtro que permite el uso de modelos de funciones de transferencia complejos mientras utiliza una cantidad limitada de recursos de procesamiento.

Un efecto de la presente descripción es proporcionar un enfoque de filtrado con estabilidad mejorada.

10 Implementaciones alternativas

Se apreciará que los enfoques anteriores, y en particular la Tecnología 1 y la Tecnología 2, pueden implementarse de muchas maneras. A continuación sigue una descripción general de características que pueden ser comunes a muchas implementaciones de los enfoques anteriores. Por supuesto, se entenderá que, a menos que se indique lo contrario, cualquiera de las características de los enfoques anteriores se puede combinar con cualquiera de las características comunes enumeradas a continuación.

15 Se proporciona un método para controlar (o 'controlar') un grupo de altavoces (por ejemplo, un grupo en línea de altavoces  $L$ ).

El método comprende recibir una pluralidad de señales de audio de entrada que se han de reproducir (por ejemplo,  $d$ ), por el grupo, en una pluralidad respectiva de puntos de control (o "posiciones de audición") (por ejemplo,

20 
$$\mathbf{x}_1, \dots, \mathbf{x}_M \in R^3$$

) en un entorno acústico (o "espacio acústico").

Cada una de la pluralidad de señales de audio de entrada puede ser diferente.

Al menos una de la pluralidad de señales de audio de entrada puede ser diferente de al menos otra de la pluralidad de señales de audio de entrada.

25 El método comprende además generar (o "determinar") una señal de audio de salida respectiva (por ejemplo,  $H_d$  o  $q$ ) para cada uno de los altavoces en el grupo aplicando un conjunto de filtros (por ejemplo,  $H$ ) a la pluralidad de señales de audio de entrada (por ejemplo,  $d$ ).

El conjunto de filtros pueden ser filtros digitales. El conjunto de filtros se puede aplicar en el dominio de la frecuencia.

30 El conjunto de filtros se basa en una primera pluralidad de elementos de filtro (por ejemplo,  $C$ ) y una segunda pluralidad de elementos de filtro (por ejemplo,  $G$ ).

La primera pluralidad de elementos de filtro (por ejemplo,  $C$ ) se basa en una primera aproximación de un conjunto de funciones de transferencia (por ejemplo,  $S$ ).

La segunda pluralidad de elementos de filtro (por ejemplo,  $G$ ) se basa en una segunda aproximación del conjunto de funciones de transferencia (por ejemplo,  $S$ ).

35 Cada función de transferencia en el conjunto de funciones de transferencia está entre una señal de audio aplicada a uno respectivo de los altavoces y una señal de audio recibida en uno respectivo de los puntos de control procedente del respectivo de los altavoces.

40 La primera y segunda pluralidades de elementos de filtro se basan en diferentes aproximaciones del conjunto de funciones de transferencia. En particular, las diferentes aproximaciones pueden basarse en diferentes modelos del conjunto de funciones de transferencia.

Un elemento de filtro puede ser el peso de un filtro. Una pluralidad de elementos de filtro puede ser cualquier conjunto de pesos de filtro. Un elemento de filtro puede ser cualquier componente del peso de un filtro. Una pluralidad de elementos de filtro puede ser una pluralidad de componentes de pesos respectivos de un filtro.

45 El conjunto de filtros se puede obtener combinando dos matrices diferentes,  $C$  y  $G$ , que a su vez se calculan utilizando dos aproximaciones diferentes de las funciones de transferencia electroacústicas físicas que constituyen la matriz de planta  $S$  del sistema. La matriz  $G$  (por ejemplo, como se usa en la ecuación 10) puede formarse usando una aproximación precisa y dependiente de la frecuencia de la matriz de planta  $S$ . La matriz  $C$  (por ejemplo, como se usa en la ecuación 10) puede formarse usando ganancias y retardos independientes de la frecuencia o, más generalmente,

elementos que son diferentes de los elementos de G y permiten DF que se pueden calcular con una carga computacional reducida en comparación con los DF que se calculan en función de G.

La primera aproximación (por ejemplo, la utilizada para determinar C) se basa en un modelo de propagación acústica de campo libre y/o un modelo de propagación acústica de fuente puntual.

- 5 La segunda aproximación (por ejemplo, la utilizada para determinar G) tiene en cuenta uno o más de reflexión, refracción, difracción o dispersión del sonido en el entorno acústico. La segunda aproximación puede, alternativa o adicionalmente, tener en cuenta la dispersión de una cabeza de uno o más oyentes. La segunda aproximación puede, alternativa o adicionalmente, tener en cuenta una o más respuestas de frecuencia de cada uno de los altavoces o un patrón de direccionalidad de cada uno de los altavoces.
- 10 El conjunto de filtros (por ejemplo, H) puede comprender:
- un primer subconjunto de filtros (por ejemplo,  $[GC^H]^{-1}$ ) basado en la primera (por ejemplo, C) y segunda (por ejemplo, G) pluralidades de elementos de filtro; y
  - un segundo subconjunto de filtros (por ejemplo,  $C^H$ ) basado en una de la primera o segunda pluralidades de elementos de filtro.
- 15 Generar la señal de audio de salida respectiva para cada uno de los altavoces del conjunto puede comprender:
- generar una señal de audio intermedia respectiva para cada uno de los puntos de control (m) aplicando el o un primer subconjunto de filtros (por ejemplo,  $[GC^H]^{-1}$ ) a las señales de audio de entrada (por ejemplo, d); y
  - generar la señal de audio de salida respectiva para cada uno de los altavoces aplicando el o un segundo subconjunto de filtros (por ejemplo,  $C^H$ ) a las señales de audio intermedias.
- 20 El grupo puede comprender L altavoces y la pluralidad de puntos de control puede comprender M puntos de control, y el primer subconjunto de filtros puede comprender  $M^2$  filtros y el segundo subconjunto de filtros puede comprender L x M filtros.
- El conjunto de filtros o el primer subconjunto de filtros se determina basándose en una inversa de una matriz (por ejemplo,  $[GC^H]$ ) que contiene la primera (por ejemplo, C) y la segunda (por ejemplo, G) pluralidades de elementos de filtro.
- 25 La matriz (por ejemplo,  $[GC^H]$ ) que contiene la primera y segunda pluralidades de elementos de filtro puede regularizarse antes de invertirse (por ejemplo, mediante la matriz A de regularización).
- La matriz (por ejemplo,  $[GC^H]$ ) que contiene la primera y segunda pluralidades de elementos de filtro se determina basándose en:
- 30 en el dominio de la frecuencia, un producto de una matriz (por ejemplo, G) que contiene la segunda pluralidad de elementos de filtro y una matriz (por ejemplo,  $[C^H]$ ) que contiene la primera pluralidad de elementos de filtro; o
- una operación equivalente en el dominio del tiempo.
- El conjunto de filtros puede determinarse basándose en:
- 35 en el dominio de la frecuencia, un producto de una matriz (por ejemplo,  $[C^H]$ ) que contiene la primera pluralidad de elementos de filtro y la inversa de la matriz (por ejemplo,  $[GC^H]$ ) que contiene la primera y segunda pluralidades de elementos de filtro; o
- una operación equivalente en el dominio del tiempo.
- El conjunto de filtros se puede determinar utilizando una técnica de optimización.
- 40 El primer subconjunto de filtros puede determinarse para reducir una diferencia entre una matriz escalar (por ejemplo, una matriz I de identidad) y una matriz que comprende un producto de: una matriz (por ejemplo, G) que comprende la segunda pluralidad de elementos de filtro, una matriz (por ejemplo, C) que comprende la primera pluralidad de elementos de filtro, y una matriz que representa el primer subconjunto de filtros (por ejemplo, IF).
- Cada uno de la primera pluralidad de elementos de filtro (por ejemplo, C) puede ser un elemento de ganancia de retardo independiente de la frecuencia (por ejemplo,  $C_{m,l} = e^{-j\omega t(x_m, y_l)} g_{m,l}$ )
- 45 Cada uno de la primera pluralidad de elementos de filtro puede comprender un término de retardo (por ejemplo,  $e^{-j\omega t(x_m, y_l)}$ ) y/o un término de ganancia (por ejemplo,  $g_{m,l}$ ) que se basa en una posición relativa (por ejemplo,  $x_m$ ) de uno de los puntos de control y uno de los altavoces (p. ej.  $y_l$ ).
- Para cada uno (m) dado de la pluralidad de puntos de control:

un primer vector (por ejemplo,  $c_m$ ) puede contener los elementos de filtro de la primera pluralidad de elementos de filtro (por ejemplo, C) que corresponden al punto de control dado (m), y

un segundo vector (por ejemplo,  $g_m$ ) puede contener los elementos de filtro de la segunda pluralidad de elementos de filtro (por ejemplo, G) que corresponden al punto de control dado (m);

5 y cada uno de la primera pluralidad de elementos de filtro puede comprender un término de retardo y/o un término de ganancia que se determina basándose en una colinealidad (por ejemplo,  $\gamma$ ) entre el primer y segundo vectores.

El término de retardo (por ejemplo,  $e^{-j\omega\tau(x_m, y)}$ ) y/o el término de ganancia (por ejemplo,  $g_{m,i}$ ) pueden determinarse para aumentar (o maximizar), para cada uno (m) dado de la pluralidad de puntos de control, la colinealidad (por ejemplo,  $\gamma_{m,m'}$ ) entre el primer vector (por ejemplo,  $c_m$ ) correspondiente al punto de control dado y el segundo vector (por ejemplo,  $g_m$ ) correspondiente al punto de control dado.

El término de retardo (por ejemplo,  $e^{-j\omega\tau(x_m, y)}$ ) y/o el término de ganancia (por ejemplo,  $g_{m,i}$ ) se pueden determinar de manera para:

15 reducir (o minimizar), para cada par de puntos primero ( $m_1$ ) y segundo ( $m_2$ ) dados diferentes de la pluralidad de puntos de control, la colinealidad (por ejemplo,  $\gamma_{m_1, m_2}$ ) entre el primer vector (por ejemplo,  $c_{m_1}$ ) correspondiente al primer punto de control dado y el segundo vector (por ejemplo,  $g_{m_2}$ ) correspondiente al segundo punto de control dado; y

aumentar (o maximizar), para cada tercero dado ( $m_3$ ) de la pluralidad de puntos de control, la colinealidad (por ejemplo,  $\gamma_{m_3, m_2}$ ) entre el primer vector (por ejemplo,  $c_{m_3}$ ) correspondiente al tercer punto de control dado y el segundo vector (por ejemplo,  $g_{m_3}$ ) correspondiente al tercer punto de control dado.

20 Cada uno de la primera pluralidad de elementos de filtro puede comprender un término de retardo (por ejemplo,  $e^{-j\omega\tau(x_m, y)}$ ) y/o un término de ganancia (por ejemplo,  $g_{m,i}$ ) que se determina para cada fila dada de una primera matriz. (por ejemplo, C) que comprende la primera pluralidad de elementos de filtro, para:

25 aumentar (o maximizar) una colinealidad (por ejemplo,  $\gamma$ ) entre la fila dada de la primera matriz (por ejemplo, C) y una fila correspondiente de una segunda matriz (por ejemplo, G) que comprende la segunda pluralidad de elementos de filtro; y

opcionalmente, reducir (o minimizar) la colinealidad (por ejemplo,  $\gamma$ ) entre la fila dada de la primera matriz (por ejemplo, C) y las filas no correspondientes de la segunda matriz (por ejemplo, G).

30 Cada uno de la primera pluralidad de elementos de filtro puede comprender un término de retardo (por ejemplo,  $e^{-j\omega\tau(x_m, y)}$ ) basado en una aproximación lineal de una fase de uno correspondiente de la segunda pluralidad de elementos de filtro (por ejemplo, G).

La pluralidad de puntos de control (por ejemplo,

$$\mathbf{x}_1, \dots, \mathbf{x}_M \in R^3.$$

) puede comprender ubicaciones de una pluralidad correspondiente de oyentes, por ejemplo, cuando se opera en un modo de "audio personal".

35 La pluralidad de puntos de control (por ejemplo,

$$\mathbf{x}_1, \dots, \mathbf{x}_M \in R^3.$$

) puede comprender ubicaciones de oídos de uno o más oyentes, por ejemplo, cuando se opera en un modo "baural".

40 La segunda aproximación puede basarse en una o más funciones de transferencia relacionadas con la cabeza, HRTF. La una o más HRTF pueden ser HRTF medidas. Una o más HRTF pueden ser HRTF simuladas. Una o más HRTF se pueden determinar usando un modelo de elemento límite de una cabeza.

La segunda pluralidad de elementos de filtro se puede determinar midiendo el conjunto de funciones de transferencia.

El método puede comprender además determinar la pluralidad de puntos de control usando un sensor de posición.

Generar las respectivas señales de audio de salida (por ejemplo, Hd) puede comprender el uso de un banco de filtros para aplicar al menos una parte del conjunto de filtros en una pluralidad de sub-bandas de frecuencia.

45 El primer subconjunto de filtros (por ejemplo,  $[GC^H]^{-1}$ ) y el segundo subconjunto de filtros (por ejemplo,  $C^H$ ) se pueden aplicar en cada una de las sub-bandas de frecuencia (por ejemplo, como se ilustra en la Fig. 9a).

El primer subconjunto de filtros (por ejemplo,  $[GC^H]^{-1}$ ) y el segundo subconjunto de filtros (por ejemplo,  $C^H$ ) se pueden aplicar dentro del banco de filtros (por ejemplo, como se ilustra en la Fig. 9a).

5 El primer subconjunto de filtros (por ejemplo,  $[GC^H]^{-1}$ ) se puede aplicar en banda completa y el segundo subconjunto de filtros (por ejemplo,  $C^H$ ) se puede aplicar en cada una de las sub-bandas de frecuencia (por ejemplo, como se ilustra en la Fig. 9b). En otras palabras, el primer subconjunto de filtros (por ejemplo,  $[GC^H]^{-1}$ ) se puede aplicar fuera del banco de filtros y el segundo subconjunto de filtros (por ejemplo,  $C^H$ ) se puede aplicar dentro del banco de filtros.

Generar una señal de audio de salida respectiva para cada uno de los altavoces en el grupo puede comprender:

generar, para cada uno de un primer subconjunto de altavoces, una señal de audio de salida respectiva en una primera de la pluralidad de sub-bandas de frecuencia; y

10 generar, para cada uno de un segundo subconjunto de los altavoces, una señal de audio de salida respectiva en una segunda de la pluralidad de sub-bandas de frecuencia,

siendo diferentes el primer y segundo subconjunto de los altavoces y siendo diferentes la primera y la segunda de la pluralidad de sub-bandas de frecuencia.

15 La primera pluralidad de elementos de filtro puede comprender un primer subconjunto de primeros elementos de filtro para una primera de la pluralidad de sub-bandas de frecuencia y un segundo subconjunto de primeros elementos de filtro para una segunda de la pluralidad de sub-bandas de frecuencia; y/o la segunda pluralidad de elementos de filtro puede comprender un primer subconjunto de segundos elementos de filtro para la primera de la pluralidad de sub-bandas de frecuencia y un segundo subconjunto de segundos elementos de filtro para la segunda de la pluralidad de sub-bandas de frecuencia.

20 El primer subconjunto de primeros elementos de filtro y el segundo subconjunto de primeros elementos de filtro pueden ser diferentes y/o el primer subconjunto de segundos elementos de filtro y el segundo subconjunto de segundos elementos de filtro pueden ser diferentes.

25 El conjunto de filtros (por ejemplo, H) puede variar en el tiempo. Alternativamente, el conjunto de filtros (por ejemplo, H) puede ser fijo o invariable en el tiempo, por ejemplo, cuando las posiciones del oyente y las orientaciones de la cabeza se consideran relativamente estáticas.

El método puede comprender además enviar las señales de audio de salida (por ejemplo,  $H_d$  o  $q$ ) al grupo de altavoces.

El método puede comprender además recibir el conjunto de filtros (por ejemplo, H), por ejemplo, desde otro dispositivo de procesamiento, o desde un módulo de determinación de filtro. El método puede comprender además determinar el conjunto de filtros (por ejemplo, H).

30 La primera y segunda aproximaciones pueden ser diferentes.

Al menos uno de la primera pluralidad de elementos de filtro (por ejemplo, C) puede ser diferente del correspondiente de la segunda pluralidad de elementos de filtro (por ejemplo, G).

El método puede comprender además determinar cualquiera de las variables enumeradas en el presente documento utilizando cualquiera de las ecuaciones establecidas en el presente documento.

35 El conjunto de filtros se puede determinar usando cualquiera de las ecuaciones establecidas en el presente documento (por ejemplo, ecuaciones 6, 8, 10, 13, 14).

Se proporciona un aparato configurado para realizar cualquiera de los métodos descritos en el presente documento.

El aparato puede comprender un procesador de señales digitales configurado para realizar cualquiera de los métodos descritos en el presente documento.

40 El aparato puede comprender el grupo de altavoces.

El aparato puede estar acoplado, o puede estar configurado para acoplarse, al grupo de altavoces.

Se proporciona un programa informático que comprende instrucciones que, cuando las ejecuta un sistema de procesamiento, hacen que el sistema de procesamiento realice cualquiera de los métodos descritos en el presente documento.

45 Se proporciona un medio (no transitorio) legible por ordenador o una señal portadora de datos que comprende el programa informático.

En algunas implementaciones, los diversos métodos descritos anteriormente se implementan mediante un programa informático. En algunas implementaciones, el programa informático incluye código informático dispuesto para dar instrucciones a un ordenador para que realice las funciones de uno o más de los diversos métodos descritos

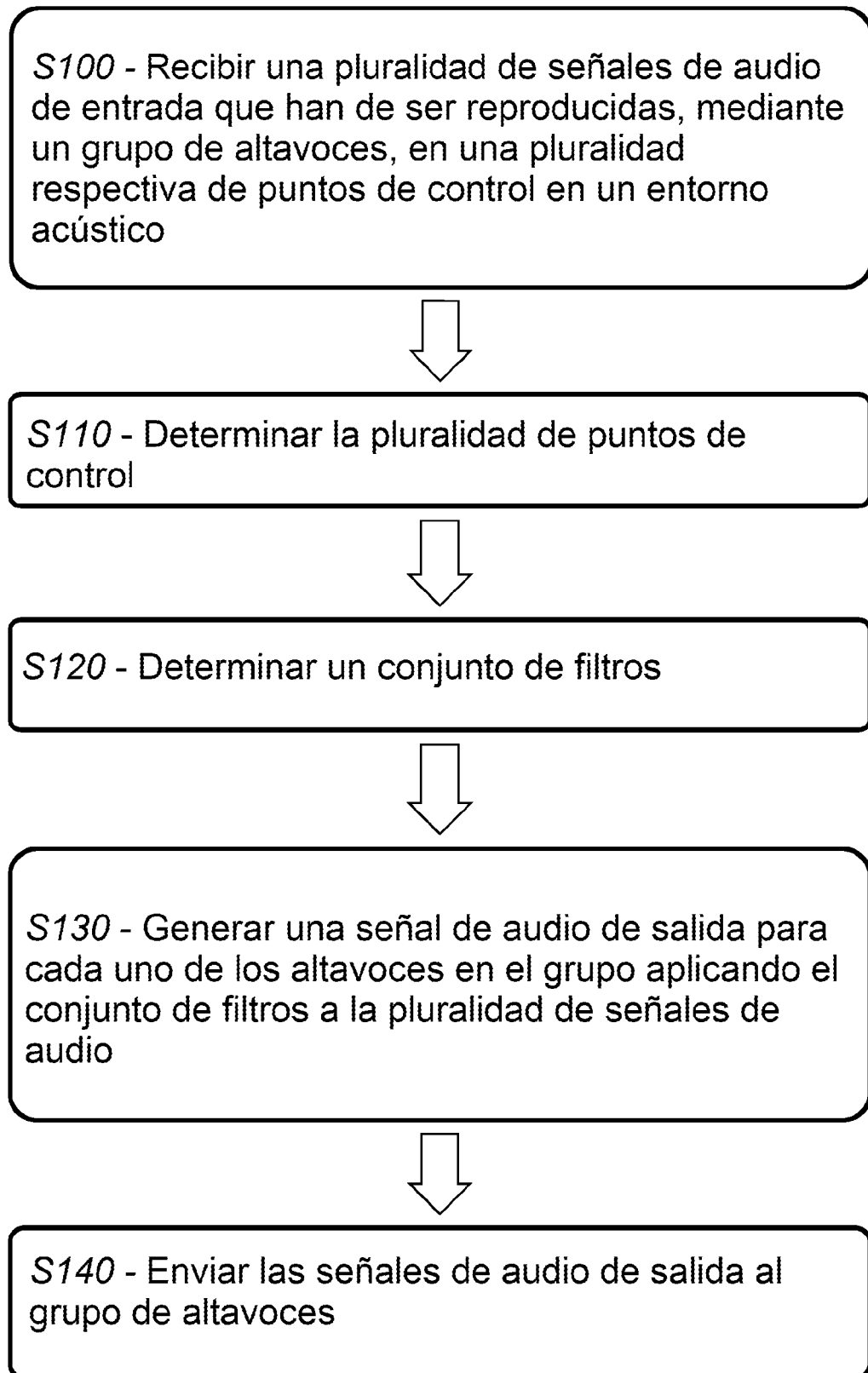
- anteriormente. En algunas implementaciones, el programa informático y/o el código para realizar dichos métodos se proporciona a un aparato, tal como un ordenador, en uno o más medios legibles por computadora o, más generalmente, un producto de programa informático. Los medios legibles por ordenador son transitorios o no transitorios. El uno o más medios legibles por ordenador podrían ser, por ejemplo, un sistema electrónico, magnético, óptico, electromagnético, infrarrojo o semiconductor, o un medio de propagación para la transmisión de datos, por ejemplo para descargar el código a través de Internet. Alternativamente, el uno o más medios legibles por ordenador podrían tomar la forma de uno o más medios físicos legibles por ordenador tales como memoria semiconductora o de estado sólido, cinta magnética, un disquete de ordenador extraíble, una memoria de acceso aleatorio (RAM), una memoria de sólo lectura (ROM), un disco magnético rígido o un disco óptico, tal como un CD-ROM, CD-RW o DVD.
- 5
- 10 En una implementación, los módulos, componentes y otras características descritas en el presente documento se implementan como componentes discretos o se integran en la funcionalidad de componentes de hardware tales como ASICs, FPGA, DSP o dispositivos similares.
- Un 'componente de hardware' es un componente físico tangible (por ejemplo, no transitorio) (por ejemplo, un conjunto de uno o más procesadores) capaz de realizar ciertas operaciones y configurado o dispuesto de una determinada manera física. En algunas implementaciones, un componente de hardware incluye circuitos o lógica dedicados que están configurados permanentemente para realizar determinadas operaciones. En algunas implementaciones, un componente de hardware es o incluye un procesador de propósito especial, como una matriz de puertas programables en campo (FPGA) o un ASIC. En algunas implementaciones, un componente de hardware también incluye lógica o circuitos programables que el software configura temporalmente para realizar determinadas operaciones.
- 15
- 20 Por consiguiente, se debe entender que el término "componente de hardware" abarca una entidad tangible que está construida físicamente, configurada permanentemente (por ejemplo, cableada) o configurada temporalmente (por ejemplo, programada) para operar de cierta manera o para realizar ciertas operaciones descritas en el presente documento.
- Además, en algunas implementaciones, los módulos y componentes se implementan como firmware o circuitos funcionales dentro de dispositivos de hardware. Además, en algunas implementaciones, los módulos y componentes se implementan en cualquier combinación de dispositivos de hardware y componentes de software, o solo en software (por ejemplo, código almacenado o incorporado de otro modo en un medio legible por máquina o en un medio de transmisión).
- 25
- 30 Los expertos en la técnica reconocerán que se puede realizar una amplia variedad de modificaciones, alteraciones y combinaciones con respecto a los ejemplos antes descritos sin apartarse del alcance de los conceptos descritos.
- Se apreciará que, aunque varios de los enfoques anteriores pueden describirse implícita o explícitamente como "óptimos", la ingeniería implica compensaciones y, por lo tanto, un enfoque que es óptimo desde una perspectiva puede no serlo desde otra. Además, pueden resultar útiles enfoques ligeramente sub-óptimos. Como resultado, tanto las soluciones óptimas como las sub-óptimas deben considerarse dentro del alcance de la presente descripción.
- 35 Los expertos en la técnica también reconocerán que el alcance de la invención no está limitado por los ejemplos descritos en el presente documento, sino que en su lugar está definido por las reivindicaciones adjuntas.

REIVINDICACIONES

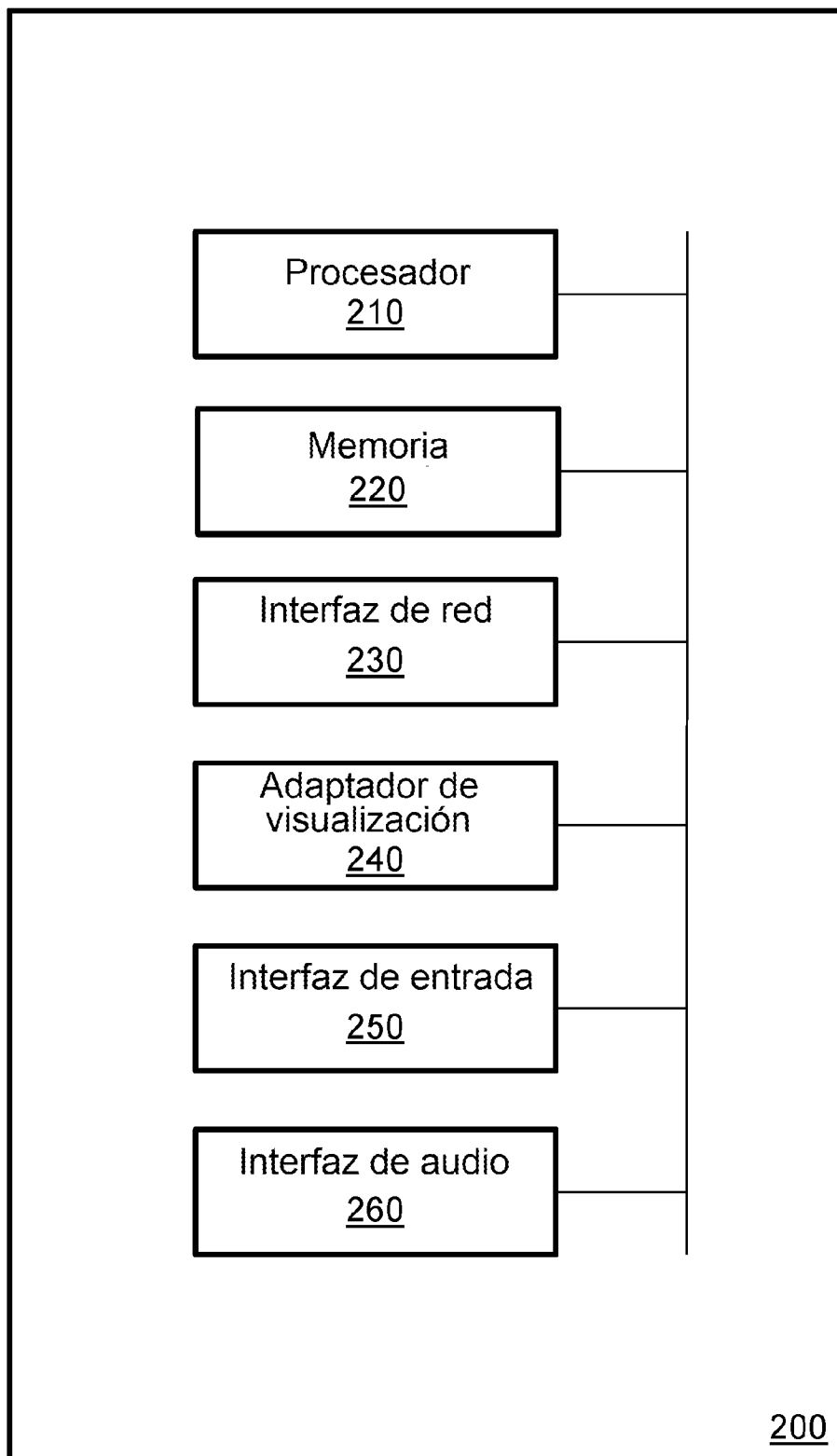
1. Un método para controlar un grupo de altavoces (300), comprendiendo el método:
  - recibir (S100) una pluralidad de señales de audio de entrada ( $d_1, \dots, d_M$ ) que han de ser reproducidas, por el grupo, en una pluralidad respectiva de puntos de control ( $x_1, \dots, x_M$ ) en un entorno acústico; y
  - 5 generar (S130) una señal (q) de audio de salida respectiva para cada uno de los altavoces en el grupo aplicando un conjunto de filtros a la pluralidad de señales de audio de entrada,
    - en donde el conjunto de filtros se basa en:
      - una primera pluralidad de elementos de filtro basada en una primera aproximación de un conjunto de funciones de transferencia, estando cada función de transferencia en el conjunto de funciones de transferencia entre una
      - 10 señal de audio aplicada a uno respectivo de los altavoces y una señal de audio recibida en uno respectivo de los puntos de control procedente del respectivo de los altavoces; y
      - una segunda pluralidad de elementos de filtro basada en una segunda aproximación del conjunto de funciones de transferencia,
    - en donde la primera aproximación se basa en un modelo de propagación acústica de campo libre y/o un modelo de propagación acústica de fuente puntual y la segunda aproximación tiene en cuenta una o más de reflexión, refracción, difracción o dispersión del sonido en el entorno acústico, y
    - 15 en donde el conjunto de filtros se determina basándose en una inversa de una matriz ( $G C^H$ ) que contiene la primera y segunda pluralidades de elementos de filtro, determinándose la matriz basándose:
      - en el dominio de frecuencia, de un producto de una segunda matriz que contiene la segunda pluralidad de
      - 20 elementos (G) de filtro y de una primera matriz que contiene la primera pluralidad de elementos de filtro ( $C^H$ ) o una operación equivalente en el dominio del tiempo.
2. El método de la reivindicación 1, en donde la segunda aproximación tiene en cuenta una o más de una respuesta de frecuencia de cada uno de los altavoces (300) o un patrón de direccionalidad de cada uno de los altavoces.
3. El método de cualquier reivindicación anterior, en donde el conjunto de filtros comprende:
  - 25 un primer subconjunto de filtros basado en la primera y segunda pluralidades de elementos de filtro; y
  - un segundo subconjunto de filtros basado en una de la primera o segunda pluralidades de elementos de filtro.
4. El método de cualquier reivindicación anterior, en el que generar la señal de audio de salida respectiva para cada uno de los altavoces (300) en el grupo comprende:
  - 30 generar una respectiva señal de audio intermedia para cada uno de los puntos de control ( $x_1, \dots, x_M$ ) aplicando el o un primer subconjunto de filtros a las señales de audio de entrada ( $w_1, \dots, w_M$ ) y
  - generar la respectiva señal (q) de audio de salida para cada uno de los altavoces aplicando el segundo subconjunto de filtros a las señales de audio intermedias.
5. El método de cualquier reivindicación anterior, en donde cada uno de la primera pluralidad de elementos de filtro es un elemento de ganancia-retardo independiente de la frecuencia.
- 35 6. El método de cualquier reivindicación anterior, en donde cada uno de la primera pluralidad de elementos de filtro comprende un término de retardo y/o un término de ganancia que se determina para cada fila dada de la primera matriz que comprende la primera pluralidad de elementos de filtro, para:
  - aumentar una colinealidad entre la fila dada de la primera matriz y una fila correspondiente de la segunda matriz que comprende la segunda pluralidad de elementos de filtro; y
  - 40 opcionalmente, reducir la colinealidad entre la fila dada de la primera matriz y las filas no correspondientes de la segunda matriz.
7. El método de cualquier reivindicación anterior, en donde cada uno de la primera pluralidad de elementos de filtro comprende un término de retardo basado en una aproximación lineal de una fase de uno correspondiente de la segunda pluralidad de elementos filtrantes.
- 45 8. El método de cualquier reivindicación anterior, en donde la pluralidad de puntos de control ( $x_1, \dots, x_M$ ) comprende ubicaciones de una pluralidad correspondiente de oyentes o ubicaciones de oídos de uno o más oyentes (341, 342, 343).
9. El método de cualquier reivindicación anterior, en donde la segunda aproximación se basa en una o más funciones de

transferencia relacionadas con la cabeza, HRTF.

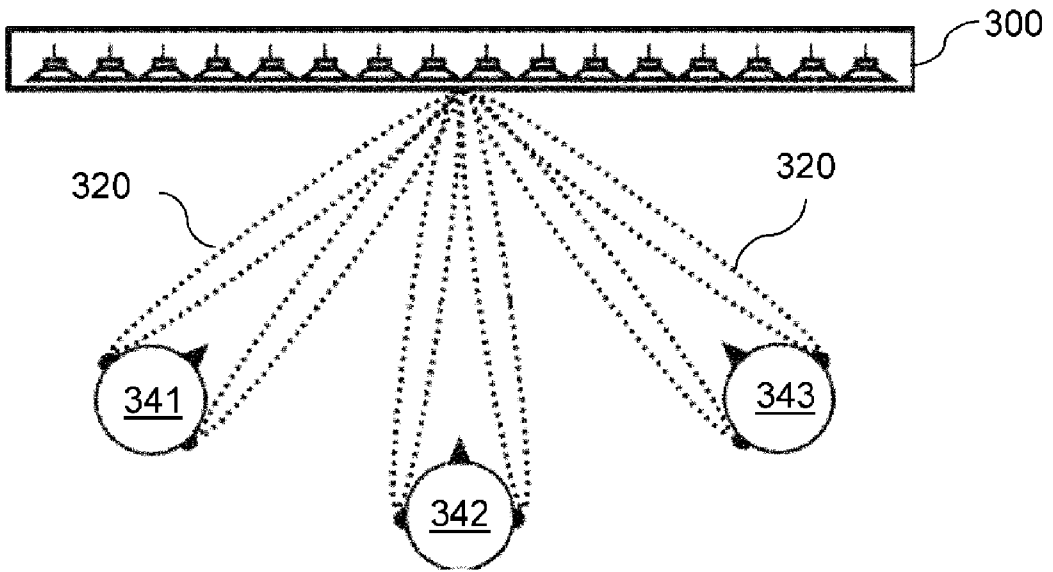
10. El método de cualquier reivindicación anterior, que comprende además determinar (S110) la pluralidad de puntos de control ( $x_1, \dots, x_M$ ) usando un sensor de posición.
11. El método de cualquier reivindicación anterior, en donde el conjunto de filtros varía en el tiempo.
- 5 12. Un aparato (200) que comprende medios configurados para realizar el método de cualquier reivindicación anterior.
13. Un programa informático que comprende instrucciones que, cuando son ejecutadas por un sistema de procesamiento, hacen que el sistema de procesamiento realice el método de cualquiera de las reivindicaciones 1 a 11.
14. Un medio legible por ordenador que comprende instrucciones que, cuando son ejecutadas por un sistema de procesamiento, hacen que el sistema de procesamiento realice el método de cualquiera de las reivindicaciones 1 a 11.
- 10 15. Una señal portadora de datos que comprende instrucciones que, cuando son ejecutadas por un sistema de procesamiento, hacen que el sistema de procesamiento realice el método de cualquiera de las reivindicaciones 1 a 11.



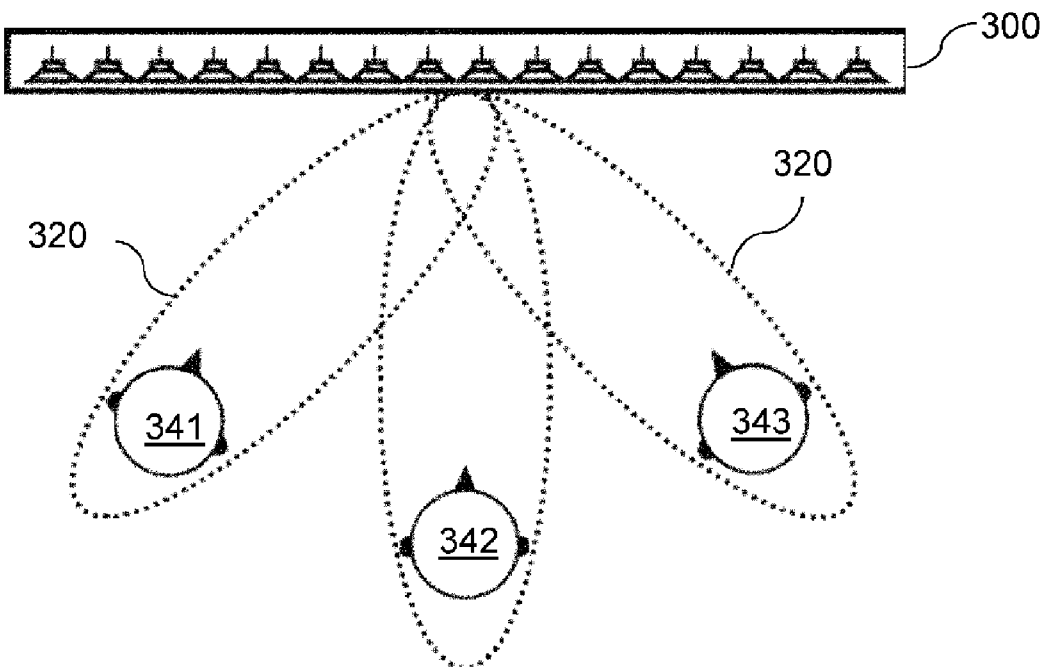
**Fig. 1**



**Fig. 2**



**Fig. 3a**



**Fig. 3b**

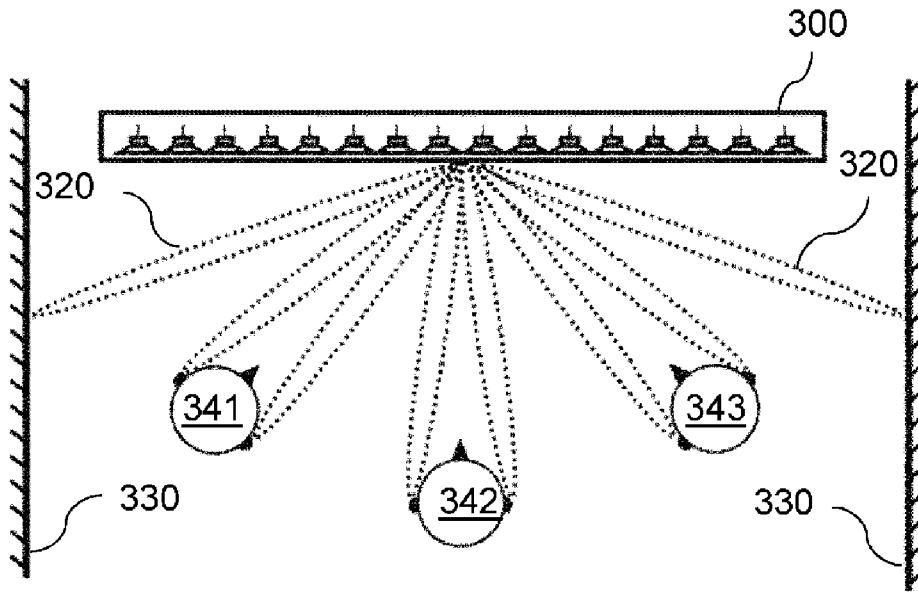


Fig. 3c

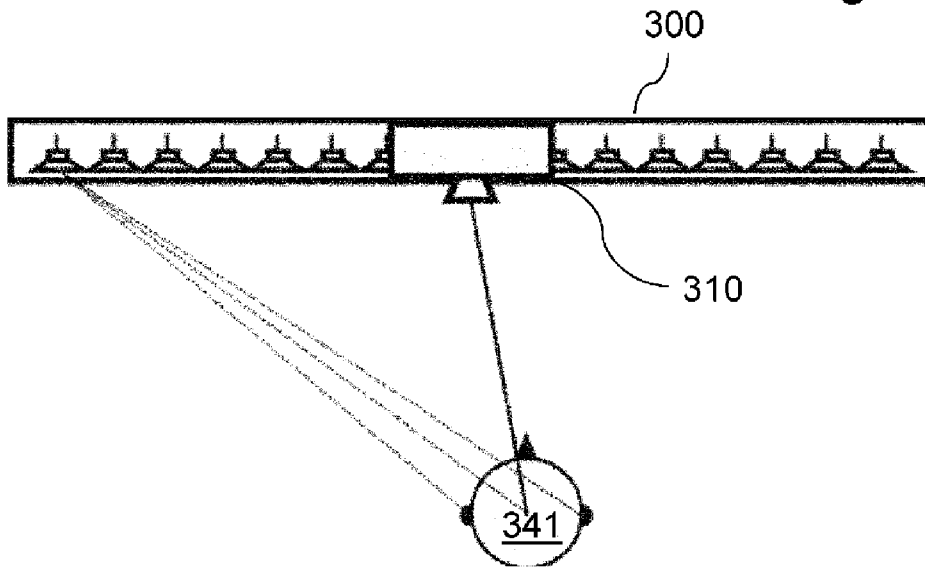


Fig. 3d

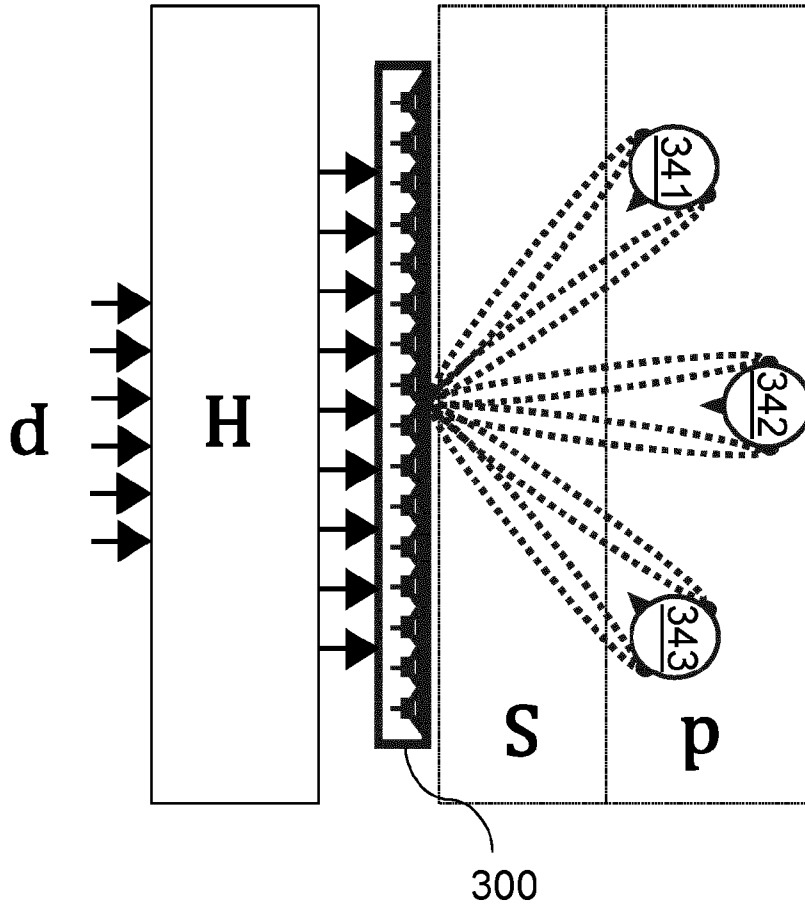


Fig. 4

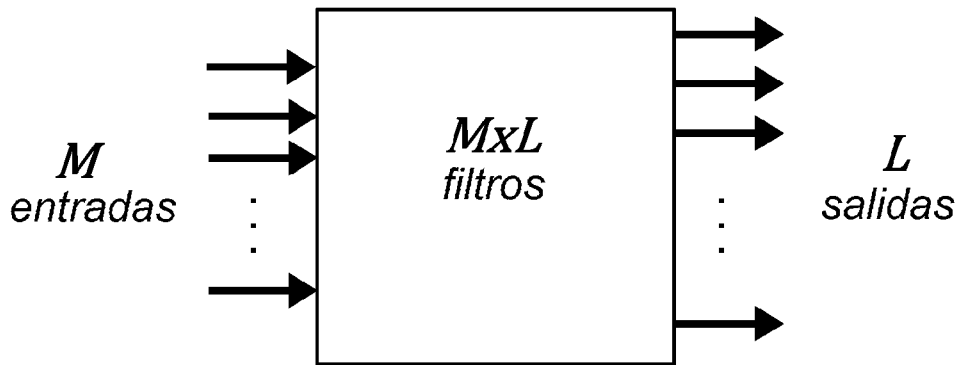


Fig. 5

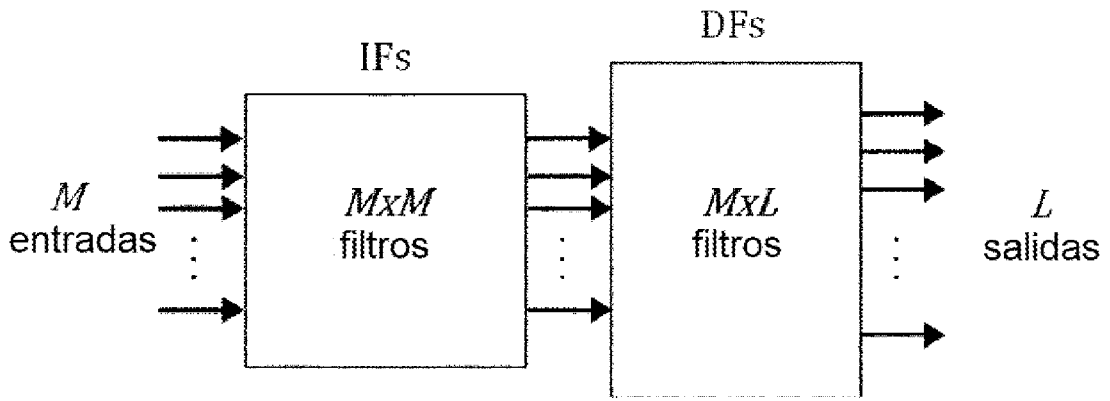


Fig. 6

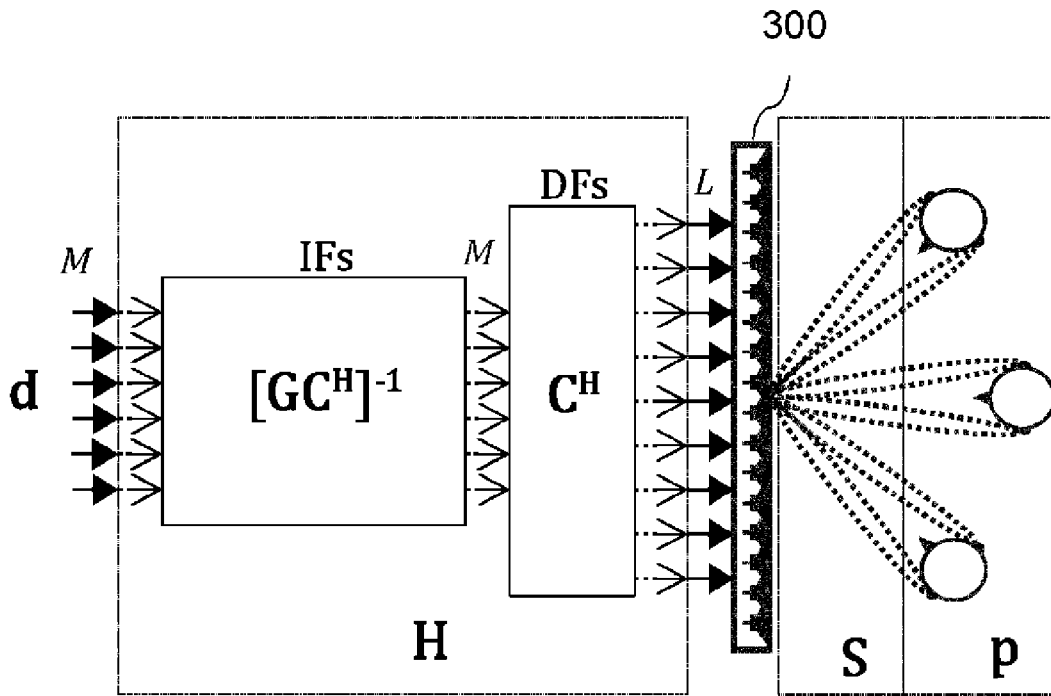


Fig. 8

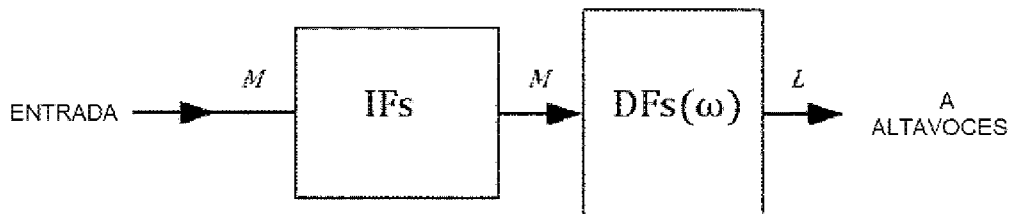


Fig. 9c

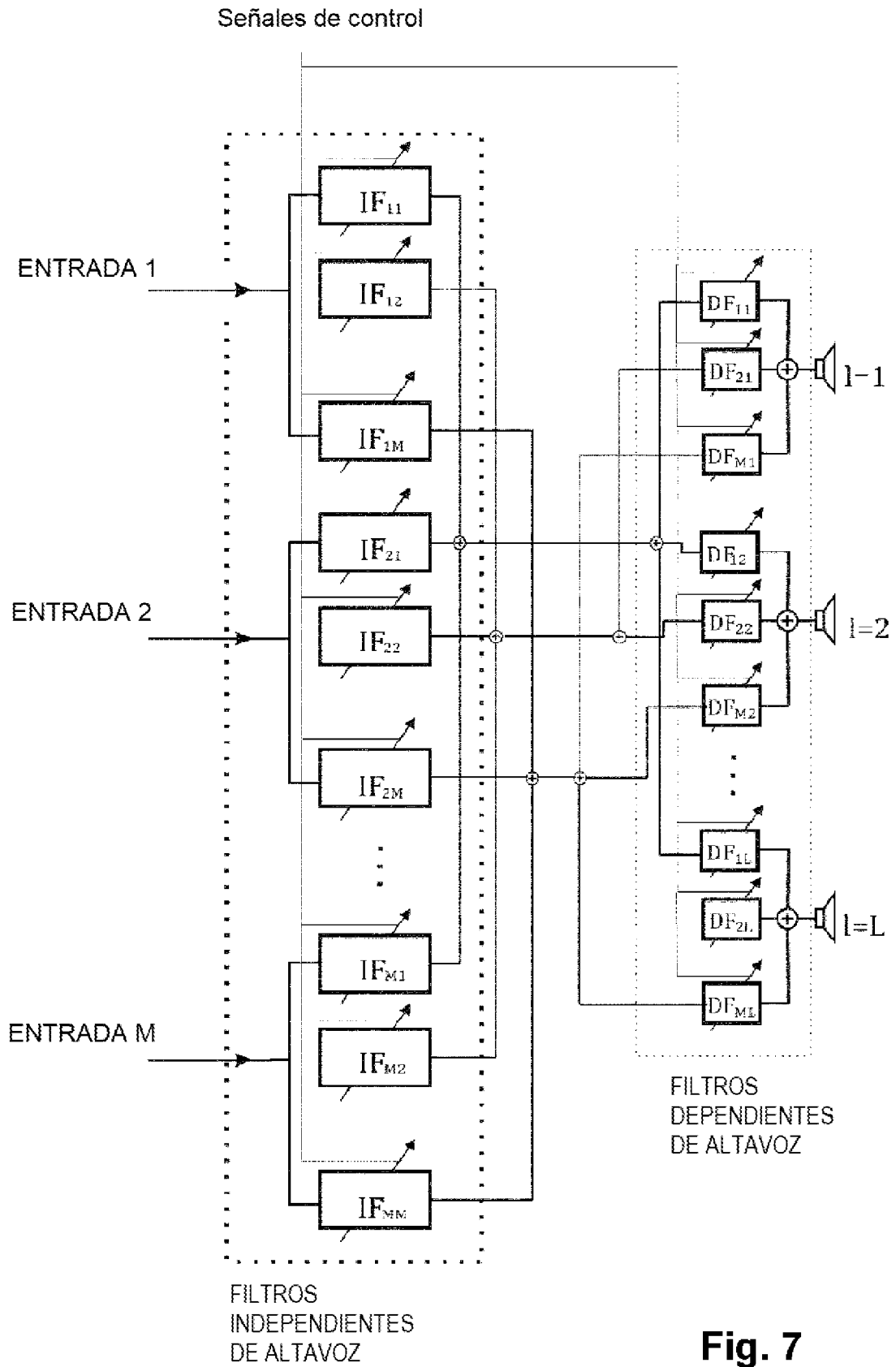


Fig. 7



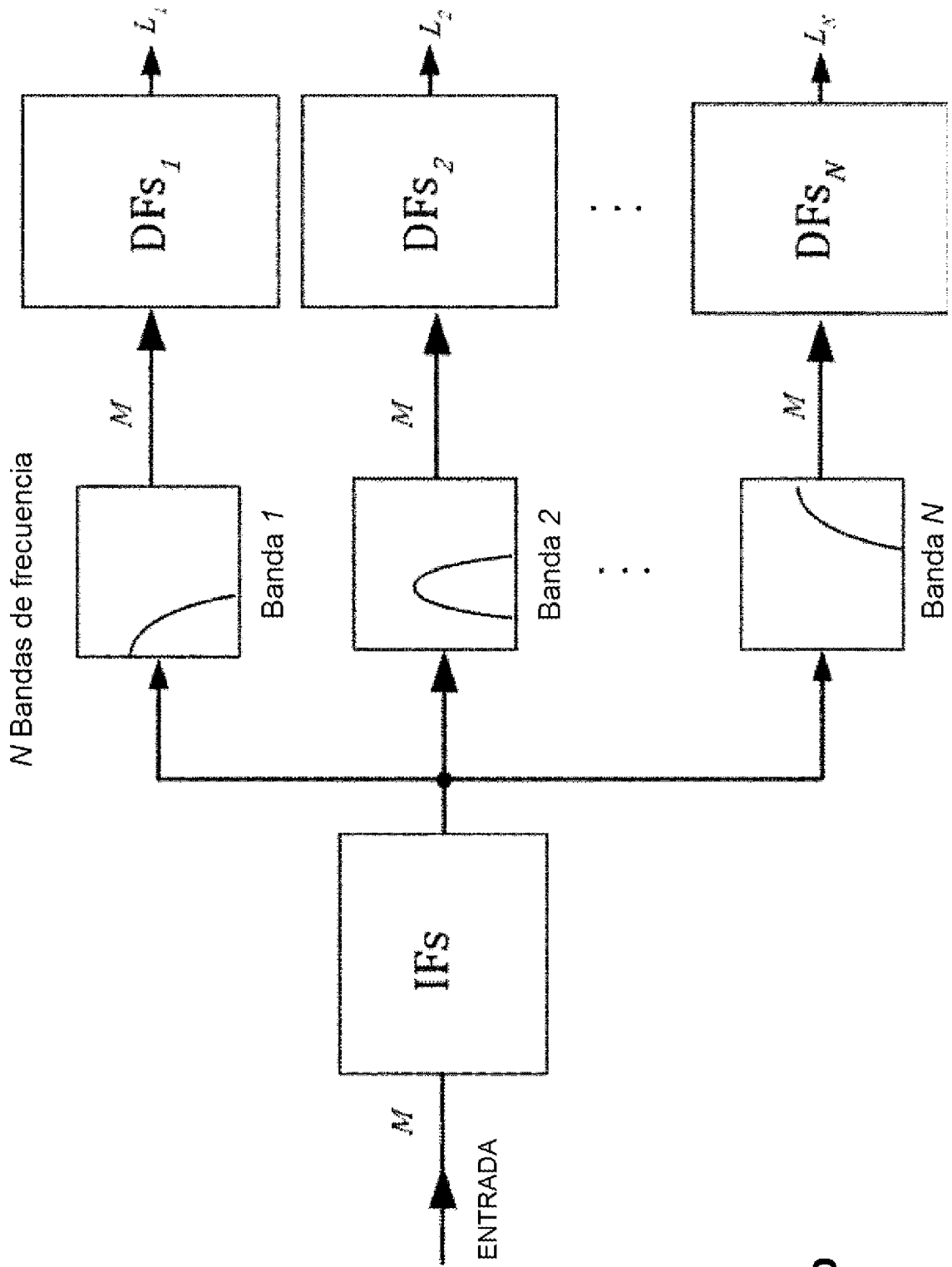


Fig. 9b

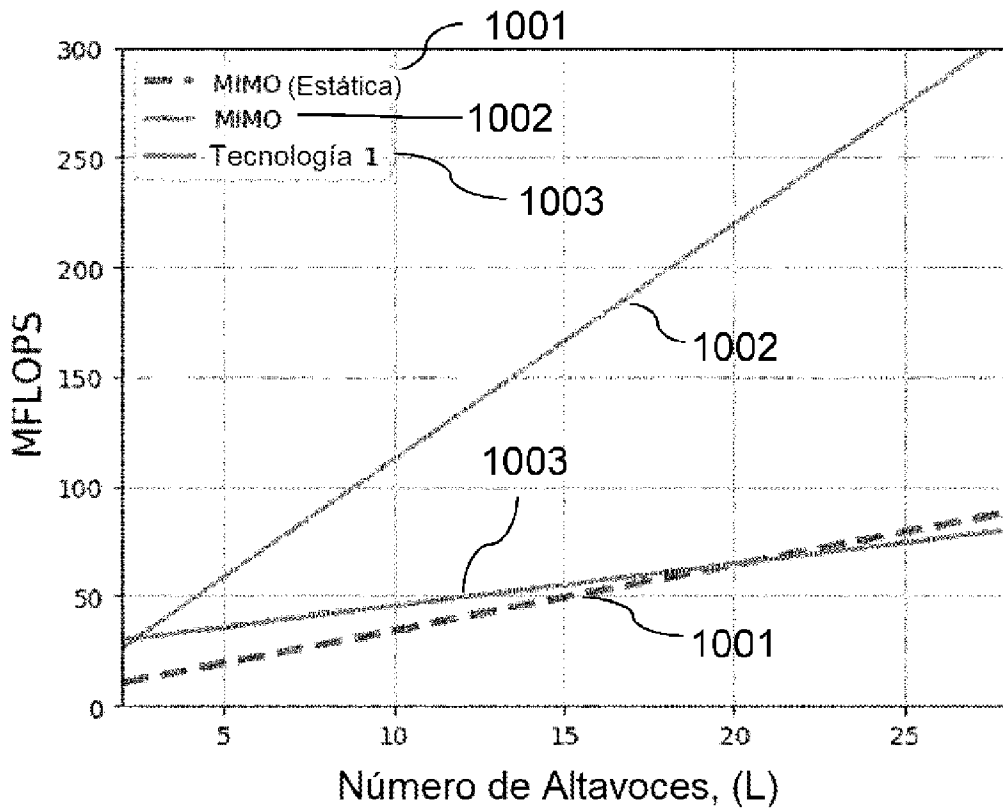


Fig. 10a

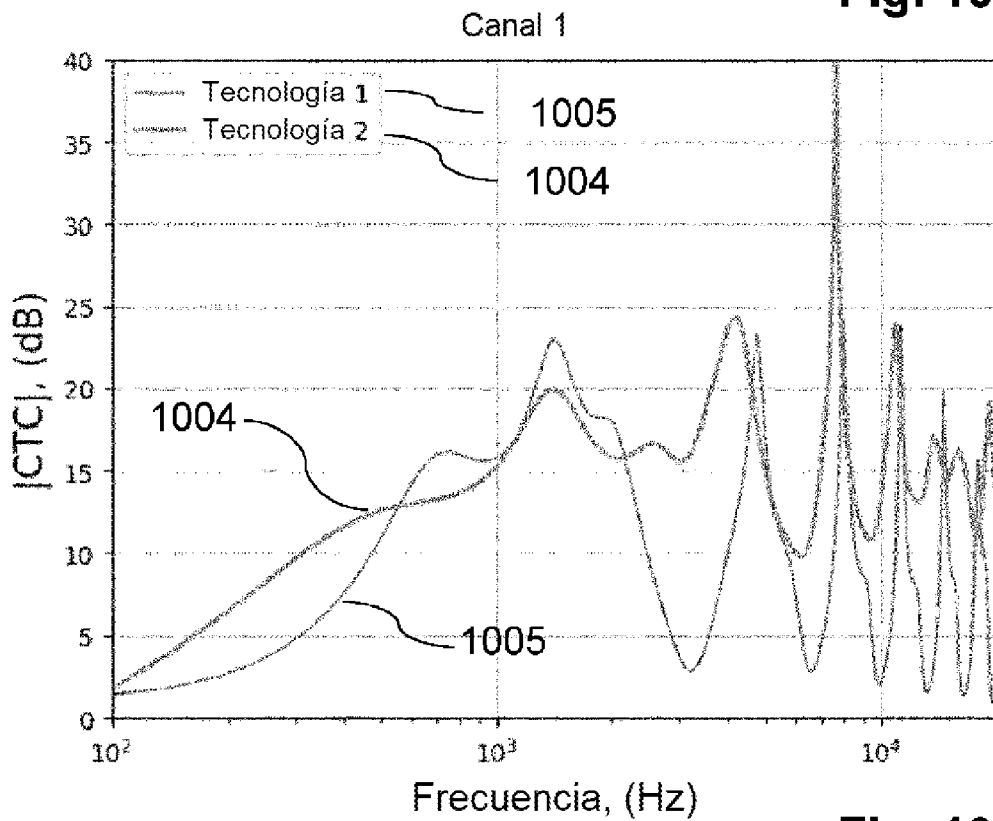


Fig. 10b