

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 010 378**

51 Int. Cl.:

**H04S 7/00** (2006.01)

**H04S 3/00** (2006.01)

**G10L 19/008** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.07.2014 E 22170897 (7)**

97 Fecha y número de publicación de la concesión europea: **25.12.2024 EP 4061020**

54 Título: **Aparato y procedimiento para la correspondencia de un primer y un segundo canal de entrada de audio con un primer y segundo canal de salida de audio**

30 Prioridad:

**22.07.2013 EP 13177360**

**18.10.2013 EP 13189243**

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

**02.04.2025**

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR  
FÖRDERUNG DER ANGEWANDTEN  
FORSCHUNG E.V. (100.00%)**

**Hansastraße 27c  
80686 München, DE**

72 Inventor/es:

**HERRE, JÜRGEN;  
KÜCH, FABIAN;  
KRATSCHMER, MICHAEL;  
KUNTZ, ACHIM y  
FALLER, CHRISTOF**

74 Agente/Representante:

**PONTI & PARTNERS, S.L.P.**

**ES 3 010 378 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 y procedimiento para la correspondencia de un primer y un segundo canal de entrada de audio con un primer y segundo canal de salida de audio

5

**[0001]** La presente solicitud se refiere a un aparato y un procedimiento para la correspondencia de un primer y un segundo canal de entrada con un primer y segundo canal de salida y, en particular, a un aparato y un procedimiento adecuados para su uso en una conversión de formato entre diferentes configuraciones de canales de altavoz.

10 **[0002]** Las herramientas de codificación de audio espacial son muy conocidas en la técnica y se han normalizado, por ejemplo, en la norma MPEG-ambiente. La codificación de audio espacial se inicia con una pluralidad de entradas originales, por ejemplo, cinco o siete canales de entrada, que se identifican por su ubicación en una configuración de reproducción, por ejemplo, como un canal izquierdo, un canal central, un canal derecho, un canal ambiente izquierdo, un canal ambiente derecho y un canal de intensificación de efectos de baja frecuencia (LFE). Un codificador de audio espacial puede deducir uno o más canales de mezcla descendente de los canales originales y, además, puede deducir datos paramétricos relacionados con indicios espaciales tales como diferencias de nivel entre canales en los valores de coherencia de canales, diferencias de fase entre canales, diferencias de tiempo entre canales, etc. Dichos uno o más canales de mezcla descendente se transmiten junto con la información lateral paramétrica que indica los indicios espaciales a un decodificador de audio espacial para decodificar los canales de mezcla descendente y los datos paramétricos asociados con el fin de obtener, en última instancia, canales de salida que son una versión aproximada de los canales de entrada originales. La ubicación de los canales en la configuración de salida puede ser fija, por ejemplo, un formato 5.1, un formato 7.1, etc.

25 **[0003]** Además, las herramientas de codificación de audio espacial son bien conocidas en la técnica y se han normalizado, por ejemplo, en la norma MPEG SAOC (SAOC = codificación de objetos de audio espacial). A diferencia de la codificación de audio espacial que parte de los canales originales, la codificación de objetos de audio espacial parte de objetos de audio que no están dedicados automáticamente a una determinada configuración de presentación y reproducción. Por el contrario, la ubicación de los objetos de audio en la escena de reproducción es flexible y puede ser ajustada por un usuario, por ejemplo, introduciendo cierta información de presentación en un decodificador para la codificación de objetos de audio espacial. De forma adicional o alternativa, se puede transmitir información de presentación en forma de información lateral adicional o metadatos; la información de presentación puede incluir información sobre en la posición de la configuración de reproducción en que se deben colocar ciertos objetos de audio (por ejemplo, según tiempo). Para obtener una determinada compresión de audio, se codifica un número de objetos de audio usando un codificador SAOC que calcula, a partir de los objetos de entrada, uno o más canales de transporte mediante la mezcla descendente de los objetos según cierta información de mezcla descendente. Además, el codificador SAOC calcula información lateral paramétrica que representa indicios entre objetos tales como diferencias de nivel de los objetos (OLD), valores de coherencia de los objetos, etc. En lo que respecta a SAC (SAC = Codificación de Audio Espacial), se calculan los datos paramétricos entre objetos con respecto a mosaicos individuales en tiempo/frecuencia. Para una trama determinada (por ejemplo, 1.024 o 2.048 muestras) de la señal de audio se tiene en cuenta una pluralidad de bandas de frecuencia (por ejemplo, 24, 32 o 64 bandas) de manera que se suministran datos paramétricos por cada trama y cada banda de frecuencia. Por ejemplo, cuando una pieza de audio tiene 20 tramas y cuando cada trama está subdividida en 32 bandas de frecuencia, el número de mosaicos de tiempo/frecuencia es 640.

45 **[0004]** Un formato de reproducción deseado, es decir, una configuración de los canales de salida (configuración de altavoz de salida), puede diferir de una configuración de los canales de entrada, donde el número de canales de salida es generalmente diferente del número de canales de entrada. Por consiguiente, puede ser necesaria la conversión de formato para la correspondencia de los canales de entrada de la configuración de los canales de entrada con los canales de salida de la configuración de los canales de salida.

50

**[0005]** El documento WO 2011/152044 A1 describe un dispositivo de generación de sonido proporcionado con un procesador de señal que emite una pluralidad de señales de sonido de entrada de canal 7.1 a una pluralidad de altavoces prescritos como los destinos de salida de cada canal 7.1 de la pluralidad de señales de sonido de entrada y que genera altavoces virtuales en posiciones diferentes de las de cada uno de la pluralidad de altavoces prescritos y emite al menos una de la pluralidad de señales de sonido de canal 7.1 desde los altavoces virtuales generados. Los altavoces virtuales se generan en el mismo plano horizontal que los altavoces reales o en un plano diferente. Los aparatos y procedimientos para generar señales de sonido para los altavoces se describen además en los documentos WO 2010/012478 A2 y WO 2012/109019 A1. Jens Blauert: "Ein neuartiges Präsenzfilter", Fernseh- und Kinotechnik 1970, n.º 3, 1 de octubre de 1970, páginas 75-78 describe distorsiones lineales de señales de sonido que utilizan relaciones psicofisiológicas implicadas en la audición direccional en el plano medio.

60

**[0006]** El objeto subyacente de la invención es proporcionar un aparato y un procedimiento que permitan una reproducción mejorada del sonido, en particular en el caso de una conversión de formato entre diferentes configuraciones de canales de altavoz.

65

**[0007]** Este objeto se consigue mediante un aparato según la reivindicación 1, un procedimiento según la reivindicación 2 y un programa informático según la reivindicación 3.

**[0008]** La invención se basa en el hallazgo de que se puede obtener una reproducción de audio mejorada incluso en el caso de un proceso de mezcla descendente de un número de canales de entrada con un número más pequeño de canales de salida si se usa una estrategia destinada a intentar conservar la diversidad espacial de al menos dos canales de entrada que se ponen en correspondencia con al menos un canal de salida. Según las realizaciones de la invención, esto se consigue mediante el procesamiento de uno de los canales de entrada puestos en correspondencia con el mismo canal de salida mediante la aplicación de un filtro de ecualización.

**[0009]** En la invención, se aplica un filtro de ecualización al segundo canal de entrada y está configurado para incrementar la ganancia de una porción espectral del segundo canal de entrada, conocida por transmitir al oyente la impresión de que el sonido proviene de una posición correspondiente a la posición del segundo canal de entrada. Un ángulo de elevación del segundo canal de entrada es mayor que un ángulo de elevación del canal de salida con los cuales se pone en correspondencia el canal de entrada. Por ejemplo, un altavoz asociado al segundo canal de entrada puede estar en una posición por encima de un plano horizontal del oyente, mientras que los altavoces asociados al canal de salida pueden estar en una posición en el plano horizontal del oyente. El filtro de ecualización puede estar configurado para incrementar la ganancia de una porción espectral del segundo canal en un intervalo de frecuencia de entre 7 kHz y 10 kHz. Procesando de esta manera la segunda señal de entrada, se puede transmitir al oyente la impresión de que el sonido proviene de una posición elevada, incluso si en realidad no procede de una posición elevada.

**[0010]** El segundo canal de entrada se puede procesar mediante la aplicación de un filtro de ecualización configurado para procesar el segundo canal de entrada con el fin de compensar las diferencias de timbre causadas por las diferentes posiciones del segundo canal de entrada y el canal de salida con el cual se pone en correspondencia el segundo canal de entrada. Por consiguiente, el timbre del segundo canal de entrada, que es reproducido por un altavoz en una posición errónea se puede manipular de manera que un usuario tenga la impresión de que el sonido proviene de otra posición más cercana a la posición original, es decir, la posición del segundo canal de entrada.

**[0011]** Los coeficientes del filtro de ecualización se pueden determinar basándose en una respuesta binaural al impulso (BRIR) medida de un recinto de escucha específico o se pueden establecer basándose en el conocimiento empírico sobre la acústica del recinto (que también puede tener en cuenta una sala de escucha específica). Por consiguiente, el procesamiento respectivo para tener en cuenta la diversidad espacial de los canales de entrada se puede adaptar a la situación específica, como por ejemplo, la sala de escucha específica en la cual se debe reproducir la señal por medio de la configuración de los canales de salida.

**[0012]** A continuación se explican la invención y los sistemas útiles para el entendimiento de la invención con referencia a las figuras adjuntas, donde:

La Fig. 1 muestra una visión general de un codificador de audio 3D de un sistema de audio 3D;  
 La Fig. 2 muestra una visión general de un decodificador de audio 3D de un sistema de audio 3D;  
 La Fig. 3 muestra un ejemplo para implementar un convertidor de formato que se puede implementar en el decodificador de audio 3D de la Fig. 2;  
 La Fig. 4 muestra una vista superior esquemática de una configuración de altavoz;  
 La Fig. 5 muestra una vista posterior esquemática de otra configuración de altavoz;  
 Las Fig. 6a y 6b muestran vistas esquemáticas de un aparato para la correspondencia de los canales de entrada primero y segundo con un canal de salida, que no forma parte de la invención;  
 Las Fig. 7a y 7b muestran vistas esquemáticas de un aparato para la correspondencia de los canales de entrada primero y segundo con varios canales de salida, que no forman parte de la invención;  
 La Fig. 8 muestra una vista esquemática de un aparato para la correspondencia de los canales primero y segundo con un canal de salida, que no forma parte de la invención;  
 La Fig. 9 muestra una vista esquemática de un aparato para la correspondencia de los canales de entrada primero y segundo con diferentes canales de salida según una realización de la invención;  
 La Fig. 10 muestra un diagrama de bloques de una unidad de procesamiento de señales para la correspondencia de canales de entrada de una configuración de canales de entrada con canales de salida de una configuración de canales de salida;  
 La Fig. 11 muestra una unidad de procesamiento de señales; y  
 La Fig. 12 es un diagrama que muestra las denominadas bandas de Blauert.

**[0013]** Antes de describir en detalle la estrategia de la invención, se presenta una visión general de un sistema de códec de audio 3D en el cual se puede implementar la estrategia de la invención.

**[0014]** Las Figs. 1 y 2 muestran los bloques algorítmicos de un sistema de audio 3D. Más específicamente, la Fig. 1 muestra una visión general de un codificador de audio 3D 100. El codificador de audio 100 recibe en un circuito de pre-presentación/mezcla 102, que se puede incluir opcionalmente, señales de entrada, más específicamente una

pluralidad de canales de entrada que envían al codificador de audio 100 una pluralidad de señales de canales 104, una pluralidad de señales de objeto 106 y sus correspondientes metadatos de objeto 108. Las señales de objeto 106 procesadas por el pre-presentador/mezclador 102 (véanse las señales 110) se pueden enviar a un codificador SAOC 112 (SAOC = Codificación de Objetos de Audio Espacial). El codificador SAOC 112 genera los canales de transporte de SAOC 114 proporcionados a las entradas de un codificador USAC 116 (USAC = Codificación Unificada de Voz y Audio). Además, la SAOC-SI de señal 118 (SAOC-SI = información lateral de SAOC) se envía también a las entradas del codificador USAC 116. El codificador USAC 116 recibe a su vez señales de objeto 120 directamente del pre-presentador/mezclador, así como las señales de canales y señales de objeto pre-presentadas 122. La información de metadatos de objeto 108 se aplica a un codificador de OAM 124 (OAM = metadatos de objeto) que proporciona la información comprimida de metadatos de objeto 126 al codificador USAC. El codificador USAC 116, sobre la base de las señales de entrada antes mencionadas, genera una señal de salida comprimida MP4, tal como se indica en 128.

**[0015]** La Fig. 2 muestra una visión general de un decodificador de audio 3D 200 del sistema de audio 3D. La señal codificada 128 (MP4) generada por el codificador de audio 100 de la Fig. 1 es recibida en el decodificador de audio 200, más específicamente en un decodificador de USAC 202. El decodificador USAC 202 decodifica la señal recibida 128 en las señales de canales 204, las señales de objeto pre-presentadas 206, las señales de objeto 208 y las señales de canales de transporte de SAOC 210. Asimismo, la información comprimida de metadatos de objeto 212 y la SAOC-SI de señal 214 es emitida por el decodificador USAC. Las señales de objeto 208 se envían a un presentador de objetos 216 que emite las señales de objeto presentadas 218. Las señales de canales de transporte de SAOC 210 se suministran al decodificador SAOC 220 que emite las señales de objeto presentadas 222. La metainformación de objeto comprimida 212 se envía a un decodificador OAM 224 que envía las señales de control respectivas al presentador de objetos 216 y al decodificador SAOC 220 para generar las señales de objeto presentadas 218 y las señales de objeto presentadas 222. El decodificador comprende además un mezclador 226 que recibe, tal como se muestra en la Fig. 2, las señales de entrada 204, 206, 218 y 222 para emitir las señales de canales 228. Las señales de canales se pueden enviar directamente a un altavoz, por ejemplo, un altavoz de 32 canales, tal como se indica en 230. Alternativamente, las señales 228 se pueden enviar a un circuito de conversión de formato 232 que recibe, como entrada de control, una señal de distribución de la reproducción que indica la forma en que se deben convertir las señales de canales 228. En el sistema descrito en la Fig. 2, se supone que la conversión se debe realizar de tal manera que se puedan enviar las señales a un sistema de altavoz 5.1 tal como se indica en 234. Asimismo, las señales de canal 228 se envían a un presentador binaural 236 que genera dos señales de salida, por ejemplo, para un auricular, tal como se indica en 238.

**[0016]** El sistema de codificación/decodificación ilustrado en las Figs. 1 y 2 se puede basar en el códec MPEG-D USAC para la codificación de señales de canal y de objeto (véanse las señales 104 y 106). Para aumentar la eficiencia en la codificación de una gran cantidad de objetos, se puede emplear la tecnología MPEG SAOC. Tres tipos de presentadores pueden ejecutar las tareas de presentación de objetos a canales, presentación de canales a auriculares o presentación de canales a una configuración diferente de altavoz (véase la Fig. 2, números de referencia 230, 234 y 238). Cuando las señales de objeto se transmiten o codifican explícitamente de forma paramétrica usando SAOC, la información de metadatos de objeto 108 correspondiente se comprime (véase la señal 126) y se multiplexa en el flujo de bits de audio 3D 128.

**[0017]** Las Figs. 1 y 2 muestran los bloques algorítmicos correspondientes a la totalidad del sistema de audio 3D que se describe a continuación con más detalle.

**[0018]** Se puede incluir opcionalmente el pre-presentador/mezclador 102 para convertir un canal más una escena de entrada de objeto en una escena de canal antes de la codificación. Funcionalmente, es idéntico al presentador/mezclador de objetos que se describe más adelante en detalle. La pre-presentación de objetos puede ser ventajosa para garantizar una entropía de señal determinista a la entrada del codificador que es básicamente independiente del número de señales de objeto activas simultáneamente. Con la pre-presentación de objetos, no es necesaria la transmisión de metadatos de objeto. Se presentan señales de objeto discretas a la distribución de canales que el codificador está configurado para usar. Los pesos de los objetos correspondientes a cada canal se obtienen de los metadatos de objeto (OAM) asociados.

**[0019]** El codificador USAC 116 es el códec de núcleo para las señales de altavoz-canales, señales de objeto discretas, señales de mezcla descendente de objetos y señales pre-presentadas. Se basa en la tecnología MPEG-D USAC. Se encarga de la codificación de las señales enumeradas generando información de canales y objetos basada en la información geométrica y semántica de la asignación de canales de entrada y objetos. Esta información de correspondencia describe cómo se ponen en correspondencia los canales de entrada y objetos con elementos de canales de USAC, como elementos de pares de canales (CPE), elementos de canales individuales (SCE), efectos de baja frecuencia (LFE) y elementos de cuatro canales (QCE) y CPE, SCE y LFE, y la información correspondiente se transmite al decodificador. Todas las cargas útiles adicionales como los datos de SAOC 114, 118 o los metadatos de objeto 126 se tienen en cuenta en el control de tasa del codificador. La codificación de objetos es posible de maneras diferentes, dependiendo de los requisitos de tasa/distorsión y los requisitos de interactividad impuestos al presentador. Son posibles las siguientes variantes de codificación de objetos:

65

- **Objetos pre-presentados:** Las señales de objeto son pre-presentadas y mezcladas con las señales de canales 22.2 antes de la codificación. La cadena de codificación subsiguiente ve señales de canales 22.2.

- **Formas de onda discretas de objetos:** Los objetos se suministran al codificador a modo de formas de onda monofónicas. El codificador usa elementos de canal individual (SCE) para transmitir los objetos además de las señales de canales. Los objetos decodificados son presentados y mezclados en el lado del receptor. Se transmite información comprimida de metadatos de objeto al receptor/presentador.

- **Formas de onda paramétricas de objetos:** Las propiedades de los objetos y su relación mutua se describen por medio de parámetros de SAOC. La mezcla descendente de las señales de objeto se codifica con la USAC. La información paramétrica se transmite conjuntamente. El número de canales de mezcla descendente se elige dependiendo del número de objetos y la tasa de datos en general. Se transmite información comprimida de metadatos de objeto al presentador de SAOC.

**[0020]** El codificador SAOC 112 y el decodificador SAOC 220 para señales de objeto se pueden basar en la tecnología MPEG SAOC. El sistema tiene capacidad para recrear, modificar y presentar un número de objetos de audio basados en un número más pequeño de canales transmitidos y datos paramétricos adicionales, tales como OLD, IOC (Coherencia Entre Objetos), DMG (Ganancias de Mezcla Descendente). Los datos paramétricos adicionales muestran una tasa de datos significativamente más baja que la necesaria para transmitir individualmente todos los objetos, lo que aporta gran eficiencia a la codificación. El codificador SAOC 112 toma como entrada las señales de objeto/canales como formas de onda monofónicas y emite como salida información paramétrica (que está incluida en el flujo de bits de audio 3D 128) y los canales de transporte de SAOC (que se codifican usando elementos de canal único y se transmiten). El decodificador SAOC 220 reconstruye las señales de objeto/canales procedentes de los canales de transporte de SAOC decodificados 210 y la información paramétrica 214, y genera la escena de audio de salida sobre la base del formato de reproducción, la información de metadatos de objeto descomprimida y, opcionalmente, sobre la base de la información de interacción con el usuario.

**[0021]** El códec de metadatos de objeto (véanse el codificador de OAM 124 y el decodificador de OAM 224) se incluye para que, por cada objeto, los metadatos asociados que especifican la posición geométrica y el volumen de los objetos en el espacio 3D sean codificados de manera eficiente mediante la cuantificación de las propiedades de los objetos en tiempo y espacio. Los metadatos de objeto comprimidos cOAM 126 se transmiten al receptor 200 en forma de información lateral.

**[0022]** El presentador de objetos 216 usa los metadatos de objeto comprimidos para generar formas de onda de objetos según el formato de reproducción dado. Cada objeto se presenta a un determinado canal de salida 218 según sus metadatos. La salida de este bloque es el resultado de la suma de los resultados parciales. Si se decodifica tanto el contenido basado en los canales como los objetos discretos/paramétricos, las formas de onda basadas en los canales y las formas de onda de objetos presentados son mezcladas por el mezclador 226 antes de emitir las formas de onda obtenidas 228 o antes de suministrarlas a un módulo postprocesador como el módulo presentador binaural 236 o el módulo presentador de altavoz 232.

**[0023]** El módulo presentador binaural 236 produce una mezcla descendente binaural del material de audio multicanal de tal manera que cada canal de entrada esté representado por una fuente de sonido virtual. El procesamiento se lleva a cabo trama por trama en el dominio de QMF (Banco de Filtros de Cuadratura en Espejo), y la binauralización se basa en respuestas a los impulsos binaurales del recinto medidas.

**[0024]** El presentador de altavoz 232 realiza la conversión entre la configuración de canales transmitida 228 y el formato de reproducción deseado. También se le puede denominar "convertidor de formato". El convertidor de formato realiza las conversiones a números menores de canales de salida, es decir, crea mezclas descendentes.

**[0025]** En la Fig. 3 se muestra una implementación posible de un convertidor de formato 232. La unidad de procesamiento de señal puede ser tal convertidor de formato. El convertidor de formato 232, al que también se hace referencia como presentador de altavoz, convierte entre la configuración de canales transmisores y el formato de reproducción deseado mediante la correspondencia de los canales transmisores (entrada) de la configuración de canales de transmisión (entrada) con los canales (salida) del formato de reproducción deseado (configuración de canales de salida). El convertidor de formato 232 ejecuta en general las conversiones a un número más bajo de canales de salida, es decir, ejecuta un proceso de mezcla descendente (DMX) 240. El dispositivo de mezcla descendente 240, que opera preferentemente en el dominio QMF, recibe las señales de salida del mezclador 228 y emite las señales del altavoz 234. Se puede incluir un configurador 242, al que también se hace referencia como controlador, que recibe, como entrada de control, una señal 246 indicativa de la distribución de salida del mezclador (configuración de canales de entrada), es decir, la distribución de cuyos datos representados por la señal de salida del mezclador 228 se determina, y la señal 248 indicativa de la distribución de reproducción deseada (configuración de canales de salida). Basándose en esta información, el controlador 242 genera, preferentemente de manera automática, matrices de mezcla descendente correspondientes a la combinación dada de formatos de entrada y salida y aplica estas matrices al dispositivo de mezcla descendente 240. El convertidor de formato 232 permite las configuraciones estándar de altavoz, así como también configuraciones aleatorias con posiciones no estándar de altavoz.

**[0026]** Las realizaciones de la presente invención se refieren a la implementación del presentador de altavoz 232, es decir, con aparato y procedimientos para implementar parte de la funcionalidad del presentador de altavoz 232.

5

**[0027]** A continuación, se hace referencia a las Figs. 4 y 5. La Fig. 4 muestra una configuración de altavoces que representa un formato 5.1 que comprende seis altavoces que representan un canal izquierdo LC, un canal central CC, un canal derecho RC, un canal izquierdo de ambiente LSC, un canal derecho de envolvente LRC y un canal de intensificación de baja frecuencia LFC. La Fig. 5 muestra otra configuración de altavoces que comprende altavoces que representan un canal izquierdo LC, un canal central CC, un canal derecho RC y un canal central elevado ECC.

10

**[0028]** En lo sucesivo, no se hace referencia al canal de intensificación de baja frecuencia dado que la posición exacta del altavoz (sub-bajo) asociado al canal de intensificación de baja frecuencia no es importante.

15 **[0029]**

Los canales están dispuestos en direcciones específicas con respecto a una posición central del oyente P. La dirección de cada canal está definida por un ángulo azimutal  $\alpha$  y un ángulo de elevación  $\beta$ , véase la Fig. 5. El ángulo azimutal representa el ángulo del canal en un plano horizontal del oyente 300 y puede representar la dirección del canal respectivo con respecto a una dirección central anterior 302. Como se puede ver en la Fig. 4, la dirección anterior central 302 se puede definir como la dirección supuesta de visualización de un oyente situado en la posición central del oyente P. Una dirección posterior central 304 comprende un ángulo azimutal de  $180^\circ$  con respecto a la dirección anterior central 300. Todos los ángulos azimutales a la izquierda de la dirección anterior central y todos los ángulos azimutales a la derecha de la dirección anterior central entre la dirección anterior central y la dirección posterior central están a la derecha de la dirección anterior central. Los altavoces situados delante de una línea virtual 306, que es ortogonal a la dirección anterior central 302 y pasa por la posición central del oyente P, son altavoces delanteros y los altavoces situados detrás de la línea virtual 306 son los altavoces traseros. En el formato 5.1, el ángulo azimutal  $\alpha$  del canal LC está  $30^\circ$  a la izquierda,  $\alpha$  de CC está a  $0^\circ$ ,  $\alpha$  de RC está  $30^\circ$  a la derecha,  $\alpha$  de LSC está  $110^\circ$  a la izquierda y  $\alpha$  de RSC está  $110^\circ$  a la derecha.

20

25

30 **[0030]**

El ángulo de elevación  $\beta$  de un canal define el ángulo entre el plano horizontal del oyente 300 y la dirección de una línea de conexión virtual entre la posición central del oyente y el altavoz asociado al canal. En la configuración mostrada en la Fig. 4, todos los altavoces están dispuestos dentro del plano horizontal del oyente 300 y, por lo tanto, todos los ángulos de elevación son iguales a cero. En la Fig. 5, los ángulos de elevación  $\beta$  del canal ECC pueden ser de  $30^\circ$ . Un altavoz situado exactamente por encima de la posición central del oyente tendría un ángulo de elevación de  $90^\circ$ . Los altavoces dispuestos por debajo del plano horizontal del oyente 300 tienen ángulos de elevación negativos. En la Fig. 5, LC tiene una dirección  $x_1$ , CC tiene una dirección  $x_2$ , RC tiene una dirección  $x_3$  y ECC tiene una dirección  $x_4$ .

35

40 **[0031]**

La posición de un canal específico en el espacio, es decir, la posición del altavoz asociado al canal específico, está dada por el ángulo azimutal, el ángulo de elevación y la distancia del altavoz desde la posición central del oyente. Cabe señalar que el término "posición de un altavoz" es descrito con frecuencia por los expertos en la materia para referirse al ángulo azimutal y al ángulo de elevación.

45

**[0032]** Por lo general, se realiza una conversión de formato entre diferentes configuraciones de canales de altavoz en forma de proceso de mezcla descendente que pone en correspondencia un número de canales de entrada con un número de canales de salida, donde el número de canales de salida es generalmente menor que el número de canales de entrada, y donde las posiciones de los canales de salida pueden ser diferentes de las posiciones de los canales de entrada. Uno o más canales de entrada se pueden mezclar entre sí hacia el mismo canal de salida. Al mismo tiempo, uno o más canales de entrada se pueden presentar a través de más de un canal de salida. Esta correspondencia de los canales de entrada con el canal de salida se determina por lo general por una serie de coeficientes de mezcla descendente, o alternativamente se formula en forma de matriz de mezcla descendente. La elección de los coeficientes de mezcla descendente afecta significativamente a la calidad del sonido de salida con mezcla descendente que se puede obtener. Las elecciones inadecuadas pueden conducir a una mezcla desequilibrada o a una reproducción espacial deficiente de la escena de sonido de entrada.

50

55 **[0033]**

Cada canal tiene asociada una señal de audio que debe ser reproducida por el altavoz asociado. La enseñanza de que se procesa un canal específico (como, por ejemplo, mediante la aplicación de un coeficiente, mediante la aplicación de un filtro de ecualización o mediante la aplicación de un filtro de decorrelación) significa que se procesa la correspondiente señal de audio asociada a este canal. En el contexto de la presente solicitud, el término "filtro de ecualización" debe comprender cualquier medio para aplicar una ecualización a la señal, de tal manera que se obtenga una ponderación dependiente de la frecuencia de porciones de la señal. Por ejemplo, un filtro de ecualización puede estar configurado para aplicar coeficientes de ganancia dependientes de la frecuencia a las bandas de frecuencia de la señal.

60

**[0034]** La correspondencia de un canal de entrada con uno o más canales de salida incluye la aplicación de al menos un coeficiente que se ha de aplicar al canal de entrada por cada canal de salida con el cual se pone en correspondencia el canal de entrada. Dicho al menos un coeficiente puede incluir un coeficiente de ganancia, es decir, un valor de ganancia, que se ha de aplicar a la señal de entrada asociada al canal de entrada, y/o un coeficiente de retardo, es decir, un valor de retardo que se ha de aplicar a la señal de entrada asociada al canal de entrada. La correspondencia puede incluir la aplicación de coeficientes selectivos de la frecuencia, es decir, coeficientes diferentes para diferentes bandas de frecuencia de los canales de entrada. La correspondencia de los canales de entrada con los canales de salida puede incluir generar una o más matrices de coeficientes con los coeficientes. Cada matriz define un coeficiente que se ha de aplicar a cada canal de entrada de la configuración de canales de entrada por cada canal de salida de la configuración de los canales de salida. En el caso de los canales de salida, con los cuales no se pone en correspondencia el canal de entrada, el coeficiente respectivo de la matriz de coeficientes ha de ser cero. Se pueden generar matrices de coeficientes separadas para los coeficientes de ganancia y los coeficientes de retardo. Se puede generar una matriz de coeficientes por cada banda de frecuencia en el caso en que los coeficientes sean selectivos de la frecuencia. La correspondencia puede incluir además la aplicación de coeficientes derivados a las señales de entrada asociadas a los canales de entrada.

**[0035]** Para obtener buenos coeficientes de mezcla descendente, un experto (por ejemplo, un ingeniero de sonido) puede afinar los coeficientes a mano, teniendo en cuenta su conocimiento experto. Otra posibilidad consiste en deducir automáticamente los coeficientes de mezcla descendente correspondientes a una combinación dada de configuraciones de entrada y salida tratando cada canal de entrada como fuentes de sonido virtual cuya posición en el espacio está dada por la posición en el espacio asociada al canal específico, es decir, la posición del altavoz asociado al canal de entrada específico. Cada fuente virtual puede ser reproducida por un algoritmo de ajuste panorámico genérico como el ajuste panorámico de la ley de tangentes en 2D o el ajuste panorámico vectorial basado en amplitud (VBAP) en 3D, véase V. Pulkki: "Virtual Sound Source Positioning Using Vector Base Amplitude Panning", Journal of the Audio Engineering Society, vol. 45, págs. 456-466, 1997. Otra propuesta para la deducción matemática, es decir automática, de coeficientes de mezcla descendente para una combinación dada de configuraciones de entrada y salida es la ofrecida por A. Ando: "Conversion of Multichannel Sound Signal Maintaining Physical Properties of Sound in Reproduced Sound Field", IEEE Transactions on Audio, Speech, and Language Processing, vol. 19, n.º 6, agosto de 2011.

**[0036]** En consecuencia, las técnicas de mezcla descendente existentes se basan principalmente en tres estrategias para la deducción de coeficientes de mezcla descendente. La primera estrategia es una correspondencia directa de los canales de entrada descartados con canales de salida en una posición azimutal igual o comparable. No se consideran los desplazamientos de elevación. Por ejemplo, una práctica común consiste en presentar los canales de altura directamente con los canales horizontales en una posición azimutal igual o comparable, si la capa de altura no está presente en la configuración de los canales de salida. Una segunda estrategia consiste en el uso de algoritmos de ajuste panorámico genérico, que trata los canales de entrada como fuentes virtuales de sonido y conserva la información de azimut mediante la introducción de fuentes fantasma en la posición de los canales de entrada desechados. No se consideran los desplazamientos de elevación. En los procedimientos de la técnica actual el ajuste panorámico se usa sólo si no existe ningún altavoz de salida disponible en la posición de salida deseada, por ejemplo, en el ángulo azimutal deseado. Una tercera estrategia consiste en la incorporación de conocimiento experto para la deducción de coeficientes de mezcla descendente óptimos en el sentido empírico, artístico o psicoacústico. Se puede emplear una aplicación separada o combinada de diferentes estrategias.

**[0037]** Los ejemplos proporcionan una solución técnica que permite mejorar u optimizar un proceso de mezcla descendente de tal manera que se puedan obtener señales de salida con mezcla descendente de mayor calidad que sin usar esta solución. La solución puede mejorar la calidad de la mezcla descendente en los casos en que la diversidad espacial inherente a la configuración de los canales de entrada se perdería durante la mezcla descendente sin aplicar la solución propuesta.

**[0038]** Para este fin, los ejemplos permiten conservar la diversidad espacial que es inherente a la configuración de canales de entrada y que no se conserva en la estrategia de mezcla descendente (DMX) estricta. En las situaciones de mezcla descendente donde el número de canales acústicos es reducido, los ejemplos persiguen principalmente reducir la pérdida de diversidad y envoltura, que tiene lugar implícitamente cuando se pone en correspondencia un número mayor con un número menor de canales.

**[0039]** Los autores de la invención han reconocido que, dependiendo de la configuración específica, con frecuencia la diversidad espacial inherente y la envoltura espacial de una configuración de canales de entrada se reducen considerablemente o se pierden por completo en la configuración de los canales de salida. Más aún, si se reproducen simultáneamente eventos auditivos de varios oradores en la configuración de entrada, estos pueden ser más coherentes, condensados y focalizados en la configuración de salida. Esto puede llevar a una impresión espacial perceptualmente más exigente, que con frecuencia parece menos agradable que la configuración de canales de entrada. Los ejemplos persiguen conservar por primera vez la diversidad espacial en la configuración de los canales de salida. Los ejemplos de la invención persiguen conservar la ubicación percibida de un evento auditivo lo más cerca posible en comparación con el caso en que se usa la configuración original de altavoz de los canales de entrada.

**[0040]** En consecuencia, la invención proporciona una técnica específica para la correspondencia de un primer canal de entrada y un segundo canal de entrada, que están asociados a diferentes posiciones de altavoz de una configuración de canales de entrada y que por lo tanto comprenden una diversidad espacial, con un canal de salida.

5 En la invención, los canales de entrada primero y segundo están en elevaciones diferentes con respecto a un plano horizontal del oyente. Por consiguiente, se pueden tener en cuenta los desplazamientos de elevación entre el primer canal de entrada y el segundo canal de entrada para mejorar la reproducción de sonido usando los altavoces de la configuración de los canales de salida.

10 **[0041]** En el contexto de la presente solicitud, la diversidad se puede describir de la siguiente manera. Los diferentes altavoces de una configuración de canales de entrada dan origen a diferentes canales acústicos desde los altavoces a los oídos, por ejemplo, los oídos del oyente en la posición P. Existe un número de trayectorias acústicas directas y un número de trayectorias acústicas indirectas, también conocidas como reflexiones o reverberación, que surgen de una excitación ambiental de escucha diversa y que añaden cambios adicionales de decorrelación y timbre

15 a las señales percibidas desde los altavoces en diferentes posiciones. Los canales acústicos pueden ser completamente modelados por BRIR, que son característicos por cada recinto de escucha. La experiencia de escucha de una configuración de canales de entrada depende intensamente de una combinación característica de diferentes canales de entrada y diversos BRIR, que corresponden a las posiciones específicas de los altavoces. Por consiguiente, la diversidad y la envoltura surgen de diversas modificaciones de las señales, que son aplicadas de modo inherente a

20 todas las señales de altavoz por el recinto de acústica.

**[0042]** A continuación se presenta una fundamentación razonada de la necesidad de estrategias de mezcla descendente que conserven la diversidad espacial de una configuración de canales de entrada. Una configuración de canales de entrada puede usar más altavoces que una configuración de los canales de salida o puede usar al menos

25 un altavoz que no esté presente en la configuración de altavoces de salida. Meramente con fines ilustrativos, una configuración de canales de entrada puede usar los altavoces LC, CC, RC, ECC tal como se muestra en la Fig. 5, mientras que una configuración de los canales de salida puede usar sólo los altavoces LC, CC y RC, es decir, no hace uso del altavoz ECC. Por consiguiente, la configuración de canales de entrada puede usar un número mayor de capas de reproducción que la configuración de canales de salida. Por ejemplo, la configuración de canales de entrada puede

30 presentar tanto altavoces horizontales (LC, CC, RC) como de altura (ECC), mientras que la configuración de salida puede incluir sólo altavoces horizontales (LC, CC, RC). De esa manera, el número de canales acústicos de altavoz a oídos se reduce con la configuración de los canales de salida en situaciones con mezcla descendente. Específicamente, se ven afectadas sobre todo las mezclas descendentes 3D (por ejemplo, 22.2) a 2D (por ejemplo, 5.1) (DMX) debido a la falta de diferentes capas de reproducción en la configuración de los canales de salida. Los

35 grados de libertad para obtener una experiencia auditiva similar con la configuración de los canales de salida con respecto a la diversidad y envoltura son reducidos y, por lo tanto, limitados. Los ejemplos proporcionan estrategias de mezcla descendente que mejoran la conservación de la diversidad espacial de una configuración de canales de entrada, donde los aparatos y procedimientos descritos no se limitan a ningún tipo específico de estrategia de mezcla descendente y se pueden aplicar en diversos contextos y aplicaciones.

40

**[0043]** En lo sucesivo, se describen ejemplos con referencia a la situación específica expuesta en la Fig. 5. Sin embargo, los problemas y soluciones descritos pueden adaptarse fácilmente a otras situaciones con condiciones similares. Sin perder la generalidad, se suponen las siguientes configuraciones de canales de entrada y salida:

Configuración de canales de entrada: cuatro altavoces LC, CC, RC y ECC en las posiciones  $x_1 = (\alpha_1, \beta_1)$ ,  $x_2 = (\alpha_2, \beta_1)$ ,

45  $x_3 = (\alpha_3, \beta_1)$  y  $x_4 = (\alpha_4, \beta_2)$ , donde  $\alpha_2 \approx \alpha_4$  o  $\alpha_2 = \alpha_4$ .

**[0044]** Configuración de canales de salida: tres altavoces en la posición  $x_1 = (\alpha_1, \beta_1)$ ,  $x_2 = (\alpha_2, \beta_1)$  y  $x_3 = (\alpha_3, \beta_1)$ , es decir, el altavoz en la posición  $x_4$  se descarta en la mezcla descendente,  $\alpha$  representa el ángulo azimutal y  $\beta$  representa el ángulo de elevación.

50

**[0045]** Tal como se explica en lo anterior, una técnica de DMX simple daría prioridad a la conservación de la información de azimut direccional y sólo omitiría el desplazamiento de elevación. Por consiguiente, las señales del altavoz ECC en la posición  $x_4$  pasarían simplemente al altavoz CC en la posición  $x_2$ . Sin embargo, al hacerlo se perderían algunas características. En primer lugar se pierden las diferencias de timbre debido a diferentes BRIR, que

55 se aplican de modo inherente en las posiciones de reproducción  $x_2$  y  $x_4$ . En segundo lugar se pierde la diversidad espacial de las señales de entrada, que se reproducen en diferentes posiciones  $x_2$  y  $x_4$ . En tercer lugar se pierde una decorrelación inherente de las señales de entrada debido a las diferentes trayectorias de propagación acústica desde las posiciones  $x_2$  y  $x_4$  a los oídos del oyente.

60 **[0046]** Los ejemplos persiguen la conservación o emulación de una o más de las características descritas mediante la aplicación de las estrategias explicadas en la presente memoria de forma separada o en combinación para el proceso de mezcla descendente.

**[0047]** Las Figs. 6a y 6b muestran vistas esquemáticas que explican un aparato 10 según la invención para implementar una estrategia en la cual se ponen en correspondencia un primer canal de entrada 12 y un segundo canal de entrada 14 con el mismo canal de salida 16, donde el procesamiento del segundo canal de entrada se lleva a cabo mediante la aplicación de un filtro de ecualización al segundo canal de entrada. Este procesamiento está indicado en la Fig. 6a por el bloque 18.

**[0048]** Para los expertos en la materia es evidente que los aparatos explicados y descritos en la presente solicitud se pueden implementar por medio de ordenadores o procesadores configurados y/o programados para obtener la funcionalidad descrita. Alternativamente, los aparatos se pueden implementar en forma de otras estructuras de hardware programadas tales como matrices de puertas programables de campo y similares.

**[0049]** El primer canal de entrada 12 de la Fig. 6a puede estar asociado al altavoz central CC en la dirección  $x_2$  y el segundo canal de entrada 14 puede estar asociado al altavoz central elevado ECC en la posición  $x_4$  (en la configuración de canales de entrada, respectivamente). El canal de salida 16 puede estar asociado al altavoz central ECC en la posición  $x_2$  (en la configuración de los canales de salida). La Fig. 6b muestra que el canal 14 asociado al altavoz en la posición  $x_4$  se pone en correspondencia con el primer canal de salida 16 asociado al altavoz CC en la posición  $x_2$  y que esta correspondencia comprende el procesamiento 18 del segundo canal de entrada 14, es decir, el procesamiento de la señal de audio asociada al segundo canal de entrada 14. El procesamiento del segundo canal de entrada comprende la aplicación de un filtro de ecualización al segundo canal de entrada con el fin de conservar las diferentes características entre los canales de entrada primero y segundo en la configuración de canales de entrada. En las realizaciones, el filtro de ecualización puede estar configurado de manera que conserve las características referentes a las diferencias de timbre debido a los diferentes BRIR, que se aplican de manera inherente en las diferentes posiciones de altavoz  $x_2$  y  $x_4$  asociadas a los canales de entrada primero y segundo. En las realizaciones, el filtro de ecualización está configurado de manera que conserve la diversidad espacial de las señales de entrada, que se reproducen en posiciones diferentes de tal manera que la diversidad espacial de los canales de entrada primero y segundo se mantenga perceptible pese al hecho de que los canales de entrada primero y segundo se ponen en correspondencia con el mismo canal de salida.

**[0050]** Se aplica un filtro de ecualización al segundo canal de entrada, es decir, la señal de audio asociada al segundo canal de entrada en la posición  $x_4$ , si se somete a mezcla descendente hacia el altavoz CC en la posición  $x_2$ . El filtro de ecualización compensa los cambios de timbre de los diferentes canales acústicos y se puede deducir basándose en el conocimiento experto y/o en datos de BRIR medidos o similares. Por ejemplo, se supone que la configuración de canales de entrada incluye un canal de tipo Voz de Dios (VoG) en una elevación de  $90^\circ$ . Si la configuración de los canales de salida sólo incluye altavoz en una capa y el canal VoG es desechado como, por ejemplo, con una configuración de salida 5.1, una técnica sencilla y directa consiste en distribuir el canal VoG a todos los altavoces de salida para conservar la información direccional del canal VoG al menos en el punto dulce. Sin embargo, el altavoz VoG original se percibe de manera bastante diferente debido a una BRIR diferente. Mediante la aplicación de un filtro de ecualización especializado al canal VoG antes de la distribución a todos los altavoces de salida, se puede compensar la diferencia de timbre.

**[0051]** El filtro de ecualización puede estar configurado para ejecutar una ponderación dependiente de la frecuencia del canal de entrada correspondiente para tener en cuenta hallazgos psicoacústicos sobre la percepción direccional de las señales de audio. Un ejemplo de dichos hallazgos son las denominadas bandas de Blauert, que representan bandas determinantes de la dirección. La Fig. 12 muestra tres gráficos 20, 22 y 24 que representan la probabilidad de que se reconozca una dirección específica de las señales de audio. Como se puede apreciar en el gráfico 20, las señales de audio de la parte superior pueden ser reconocidas con alta probabilidad en una banda de frecuencia 1200 entre 7 kHz y 10 kHz. Tal como se puede apreciar en el gráfico 22, las señales de audio procedentes de la parte posterior pueden ser reconocidas con alta probabilidad en una banda de frecuencia 1202 de entre aproximadamente 0,7 kHz y aproximadamente 2 kHz y en una banda de frecuencia 1204 de entre aproximadamente 10 kHz y aproximadamente 12,5 kHz. Como se puede apreciar en el gráfico 24, las señales de audio procedentes de la parte delantera pueden ser reconocidas con alta probabilidad en una banda de frecuencia 1206 de entre aproximadamente 0,3 kHz y 0,6 kHz y en una banda de frecuencia 1208 de entre aproximadamente 2,5 y aproximadamente 5,5 kHz.

**[0052]** El filtro de ecualización se configura aprovechando este reconocimiento. Dicho de otro modo, el filtro de ecualización puede estar configurado de manera que aplique coeficientes de ganancia más elevados (incremento de ganancia) a bandas de frecuencia conocidas por transmitir a un usuario la impresión de que el sonido llega de diferentes direcciones, en comparación con las otras bandas de frecuencia. En términos más específicos, si se pone en correspondencia un canal de entrada con un canal de salida más bajo, se puede incrementar la ganancia de una porción espectral del canal de entrada en la banda de frecuencia 1200 en el intervalo de entre 7 kHz y 10 kHz en comparación con otras porciones espectrales del segundo canal de entrada de tal manera que el oyente pueda tener la impresión de que la señal correspondiente proviene de una posición elevada. Del mismo modo, el filtro de ecualización puede estar configurado de manera que incremente la ganancia de otras porciones espectrales del segundo canal de entrada tal como se muestra en la Fig. 12. Por ejemplo, en el caso en que se pone en correspondencia un canal de entrada con un canal de salida dispuesto en una posición más anterior las ganancias de

las bandas 1206 y 1208 se pueden incrementar y en el caso de un canal de entrada que se pone en correspondencia con un canal de salida dispuesto en la posición más posterior las ganancias de las bandas 1202 y 1204 se pueden incrementar.

5 **[0053]** La Fig. 8 muestra una vista esquemática de un aparato 30, que comprende un filtro 32, que representa un filtro de ecualización. El aparato 30 recibe una serie de canales de entrada 34 y emite un número de canales de salida 36. Los canales de entrada 34 representan una configuración de canales de entrada y los canales de salida 36 representan una configuración de canales de salida. Tal como se muestra en la Fig. 8, un tercer canal de entrada 38 se pone en correspondencia directamente con un segundo canal de salida 42 y un cuarto canal de entrada 40 se pone  
10 en correspondencia directamente con un tercer canal de salida 44. El tercer canal de entrada 38 puede ser un canal izquierdo asociado al altavoz izquierdo LC. El cuarto canal de entrada 40 puede ser un canal de entrada derecho asociado al altavoz derecho RC. El segundo canal de salida 42 puede ser un canal izquierdo asociado al altavoz izquierdo LC y el tercer canal de salida 44 puede ser un canal derecho asociado al altavoz derecho RC. El primer canal de entrada 12 puede ser el canal horizontal central asociado al altavoz central CC y el segundo canal de entrada  
15 14 puede ser el canal central de altura asociado al altavoz central elevado ECC. El filtro 32 se aplica al segundo canal de entrada 14, es decir el canal central de altura. Después del filtrado, se encamina el segundo canal de entrada hacia el altavoz horizontal central, es decir, el primer canal de salida 16 asociado al altavoz CC en la posición  $x_2$ . De esta manera, los dos canales de entrada 12 y 14 se ponen en correspondencia con el primer canal de salida 16, tal como se indica en el bloque 46 de la Fig. 8. En las realizaciones, el primer canal de entrada 12 y la versión procesada del  
20 segundo canal de entrada 14 pueden añadirse en el bloque 46 y pueden ser enviados al altavoz asociado al canal de salida 16, es decir, el altavoz horizontal central CC.

**[0054]** La Fig. 7a muestra una vista esquemática de un aparato 50 según un ejemplo que no forma parte de la invención. El aparato 50 está configurado para recibir el primer canal de entrada 12 y el segundo canal de entrada 14.  
25 El aparato 50 está configurado para la correspondencia del primer canal de entrada 12 directamente con el primer canal de salida 16. El aparato 50 está configurado asimismo para generar una fuente fantasma mediante el ajuste panorámico entre el segundo y el tercer canal de salida, que pueden ser el segundo canal de salida 42 y el tercer canal de salida 44. Esto está indicado en la Fig. 7a por el bloque 52. De esa manera se genera una fuente fantasma que tiene un ángulo azimutal correspondiente al ángulo azimutal del segundo canal de entrada.

30 **[0055]** Cuando se considera el escenario de la Fig. 5, el primer canal de entrada 12 puede estar asociado con el altavoz horizontal central CC, el segundo canal de entrada 14 puede estar asociado con el altavoz central elevado ECC, el primer canal de salida 16 puede estar asociado con el altavoz central CC, el segundo canal de salida 42 puede estar asociado con el altavoz izquierdo LC y el tercer canal de salida 44 puede estar asociado con el altavoz derecho  
35 RC. Por consiguiente, en el ejemplo mostrado en la Fig. 7a, se coloca una fuente fantasma en la posición  $x_2$  mediante el ajuste panorámico de los altavoces en las posiciones  $x_1$  y  $x_3$  en lugar de aplicar directamente la correspondiente señal al altavoz en la posición  $x_2$ . Por consiguiente, se realiza el ajuste panorámico entre altavoces en las posiciones  $x_1$  y  $x_3$  aunque exista otro altavoz en la posición  $x_2$ , que está más próxima a la posición  $x_4$  que las posiciones  $x_1$  y  $x_3$ . Dicho de otro modo, el ajuste panorámico entre altavoces en las posiciones  $x_1$  y  $x_3$  se lleva a cabo aunque las  
40 desviaciones del ángulo azimutal  $\Delta\alpha$  entre los respectivos canales 42, 44 y el canal 14 sean mayores que la desviación del ángulo azimutal entre los canales 14 y 16, que es de  $0^\circ$ ; véase la Fig. 7b. Al actuar así se conserva la diversidad espacial introducida por los altavoces en las posiciones  $x_2$  y  $x_4$  mediante el uso de un altavoz separado en la posición  $x_2$  correspondiente a la señal asignada originariamente al canal de entrada correspondiente y una fuente fantasma en la misma posición. La señal de la fuente fantasma corresponde a la señal del altavoz en la posición  $x_4$  de la  
45 configuración de canales de entrada original.

**[0056]** La Fig. 7b muestra esquemáticamente la correspondencia del canal de entrada asociado al altavoz en la posición  $x_4$  mediante el ajuste panorámico 52 entre los altavoces en las posiciones  $x_1$  y  $x_3$ .

50 **[0057]** En los ejemplos descritos con respecto a las Figs. 7a y 7b, se supone que una configuración de canales de entrada presenta una capa de altura y una horizontal que incluye un altavoz central de altura y un altavoz horizontal central. Más aún, se supone que la configuración de los canales de salida sólo presenta una capa horizontal que incluye un altavoz horizontal central y altavoces horizontales izquierdo y derecho, que pueden generar una fuente fantasma en la posición del altavoz horizontal central. Tal como se explica, en una técnica simple común, el canal de  
55 entrada central de altura se reproduciría con el altavoz de salida horizontal central. Por el contrario, según el ejemplo descrito, se realiza un ajuste panorámico intencionado en el canal de entrada central de altura entre los altavoces horizontales izquierdo y derecho de salida. De esa manera se conserva la diversidad espacial del altavoz central de altura y el altavoz horizontal central de la configuración de canales de entrada mediante el uso del altavoz horizontal central y una fuente fantasma alimentada por el canal de entrada central de altura.

60 **[0058]** En la invención, además del ajuste panorámico, se puede aplicar un filtro de ecualización para compensar los posibles cambios de timbre debido a diferentes BRIR.

**[0059]** En la Fig. 9 se muestra una realización de un aparato 60 que implementa la estrategia de ajuste  
65 panorámico y que no forma parte de la invención. En la Fig. 9, los canales de entrada y los canales de salida

corresponden a los canales de entrada y el canal de salida mostrados en la Fig. 8 y se omite una descripción repetida de los mismos. El aparato 60 está configurado para generar una fuente fantasma mediante el ajuste panorámico entre los canales de salida segundo y tercero 42 y 44, tal como se muestra en la Fig. 9 por los bloques 62.

5 **[0060]** En las realizaciones de la invención, el ajuste panorámico se puede efectuar usando algoritmos de ajuste panorámico comunes, tales como los algoritmos de ajuste panorámico genérico como el ajuste panorámico por ley de tangentes en 2D o el ajuste panorámico vectorial basado en la amplitud en 3D; véase V. Pulkki: "Virtual Sound Source Positioning Using Vector Base Amplitude Panning", Journal of the Audio Engineering Society, vol. 45, págs. 456-466, 1997, y no es necesario describirlo en más detalle en esta invención. Las ganancias de ajuste panorámico de la ley de ajuste panorámico aplicada determinan las ganancias que se aplican a la correspondencia de los canales de entrada con los canales de salida. Las señales respectivas obtenidas se suman a los canales de salida segundo y tercero 42 y 44, véanse los bloques sumadores 64 de la Fig. 9. Por consiguiente, se pone en correspondencia el segundo canal de entrada 14 con los canales de salida segundo y tercero 42 y 44 realizando un ajuste panorámico con el fin de generar una fuente fantasma en la posición  $x_2$ , el primer canal de entrada 12 se pone en correspondencia directamente con el primer canal de salida 16, y también se ponen en correspondencia directamente los canales de entrada tercero y cuarto 38 y 40 con los canales de salida segundo y tercero 42 y 44.

**[0061]** El bloque 62 se puede modificar para dar lugar también a la funcionalidad de un filtro de ecualización además de la funcionalidad de ajuste panorámico. Por consiguiente, mediante la estrategia de ajuste panorámico se pueden compensar los posibles cambios de timbre debido a diferentes BRIR además de conservar la diversidad espacial.

**[0062]** La Fig. 10 muestra un sistema para generar una matriz de DMX, donde se puede incorporar la presente invención. El sistema comprende series de reglas que describen correspondencias potenciales de los canales de entrada y salida, bloque 400, y un selector 402 que selecciona las reglas más apropiadas para una determinada combinación de una configuración de canales de entrada 404 y una combinación de configuraciones de los canales de salida 406 basándose en las series de reglas 400. El sistema puede comprender una interfaz apropiada para recibir información sobre la configuración de los canales de entrada 404 y la configuración de los canales de salida 406. La configuración de los canales de entrada define los canales presentes en una disposición de entrada, donde cada canal de entrada tiene una dirección o posición asociada al mismo. La configuración de los canales de salida define los canales presentes en la configuración de salida, donde cada canal de salida tiene una dirección o posición asociada. El selector 402 suministra las reglas seleccionadas 408 a un evaluador 410. El evaluador 410 recibe las reglas seleccionadas 408 y evalúa las reglas seleccionadas 408 para deducir coeficientes de DMX 412 sobre la base de las reglas seleccionadas 408. Se puede generar una matriz de DMX 414 a partir de los coeficientes de mezcla descendente derivados. El evaluador 410 puede estar configurado para deducir la matriz de mezcla descendente de los coeficientes de mezcla descendente. El evaluador 410 puede recibir información sobre la configuración de los canales de entrada y la configuración de los canales de salida, como por ejemplo, información sobre la geometría de la configuración de salida (por ejemplo, las posiciones de los canales) e información sobre la geometría de la configuración de entrada (por ejemplo, las posiciones de los canales) y tener en cuenta la información al deducir los coeficientes de DMX. Tal como se muestra en la Fig. 11, el sistema se puede implementar en una unidad de procesamiento de señales 420 que comprende un procesador 422 programado o configurado para actuar como selector 402 y evaluador 410 y una memoria 424 configurada para almacenar al menos parte de las series 400 de reglas de correspondencia. Otra parte de las reglas de correspondencia puede ser verificada por el procesador sin acceder a las reglas almacenadas en la memoria 422. En cualquier caso, las reglas se envían al procesador para ejecutar los procedimientos descritos. La unidad de procesamiento de señales puede incluir una interfaz de entrada 426 para recibir las señales de entrada 228 asociadas a los canales de entrada y una interfaz de salida 428 para emitir las señales de salida 234 asociadas a los canales de salida.

**[0063]** Algunas de las reglas 400 se pueden concebir de tal manera que la unidad de procesamiento de señales 420 implemente una realización de la invención. En la Tabla 1 se enumeran ejemplos de reglas para la correspondencia de un canal de entrada con uno o más canales de salida.

**Tabla 1: Reglas de correspondencia**

Entrada (Origen)	Salida (Destino)	Ganancia	Índice EQ
CH_M_000	CH_M_L030, CH_M_R030	1,0	0 (inactivo)
CH_M_L060	CH_M_L030, CH_M_L110	1,0	0 (inactivo)
CH_M_L060	CH_M_L030	0,8	0 (inactivo)
CH_M_R060	CH_M_R030, CH_M_R110,	1,0	0 (inactivo)
CH_M_R060	CH_M_R030,	0,8	0 (inactivo)
CH_M_L090	CH_M_L030, CH_M_L110	1,0	0 (inactivo)
CH_M_L090	CH_M_L030	0,8	0 (inactivo)
CH_M_R090	CH_M_R030, CH_M_R110	1,0	0 (inactivo)
CH_M_R090	CH_M_R030	0,8	0 (inactivo)
CH_M_L110	CH_M_L135	1,0	0 (inactivo)

ES 3 010 378 T3

Entrada (Origen)	Salida (Destino)	Ganancia	Índice EQ
CH_M_L110	CH_M_L030	0,8	0 (inactivo)
CH_M_R110	CH_M_R135	1,0	0 (inactivo)
CH_M_R110	CH_M_R030	0,8	0 (inactivo)
CH_M_L135	CH_M_L110	1,0	0 (inactivo)
CH_M_L135	CH_M_L030	0,8	0 (inactivo)
CH_M_R135	CH_M_R110	1,0	0 (inactivo)
CH_M_R135	CH_M_R030	0,8	0 (inactivo)
CH_M_180	CH_M_R135, CH_M_L135	1,0	0 (inactivo)
CH_M_180	CH_M_R110, CH_M_L110	1,0	0 (inactivo)
CH_M_180	CH_M_R030, CH_M_L030	0,6	0 (inactivo)
CH_U_000	CH_U_L030, CH_U_R030	1,0	0 (inactivo)
CH_U_000,	CH_M_L030, CH_M_R030	0,85	0 (inactivo)
CH_U_L045	CH_U_L030	1,0	0 (inactivo)
CH_U_L045	CH_M_L030	0,85	1
CH_U_R045	CH_U_R030	1,0	0 (inactivo)
CH_U_R045	CH_M_R030	0,85	1
CH_U_L030	CH_U_L045	1,0	0 (inactivo)
CH_U_L030	CH_M_L030	0,85	1
CH_U_R030	CH_U_R045	1,0	0 (inactivo)
CH_U_R030	CH_M_R030	0,85	1
CH_U_L090	CH_U_L030, CH_U_L110	1,0	0 (inactivo)
CH_U_L090	CH_U_L030, CH_U_L135	1,0	0 (inactivo)
CH_U_L090	CH_U_L045	0,8	0 (inactivo)
CH_U_L090	CH_U_L030	0,8	0 (inactivo)
CH_U_L090	CH_M_L030, CH_M_L110	0,85	2
CH_U_L090	CH_M_L030	0,85	2
CH_U_R090	CH_U_R030, CH_U_R110	1,0	0 (inactivo)
CH_U_R090	CH_U_R030, CH_U_R135	1,0	0 (inactivo)
CH_U_R090	CH_U_R045	0,8	0 (inactivo)
CH_U_R090	CH_U_R030	0,8	0 (inactivo)
CH_U_R090	CH_M_R030, CH_M_R110	0,85	2
CH_U_R090	CH_M_R030	0,85	2
CH_U_L110	CH_U_L135	1,0	0 (inactivo)
CH_U_L110	CH_U_L030	0,8	0 (inactivo)
CH_U_L110	CH_M_L110	0,85	2
CH_U_L110	CH_M_L030	0,85	2
CH_U_R110	CH_U_R135	1,0	0 (inactivo)
CH_U_R110	CH_U_R030	0,8	0 (inactivo)
CH_U_R110	CH_M_R110	0,85	2
CH_U_R110	CH_M_R030	0,85	2
CH_U_L135	CH_U_L110	1,0	0 (inactivo)
CH_U_L135	CH_U_L030	0,8	0 (inactivo)
CH_U_L135	CH_M_L110	0,85	2
CH_U_L135	CH_M_L030	0,85	2
CH_U_R135	CH_U_R110	1,0	0 (inactivo)
CH_U_R135	CH_U_R030	0,8	0 (inactivo)
CH_U_R135	CH_M_R110	0,85	2
CH_U_R135	CH_M_R030	0,85	2
CH_U_180	CH_U_R135, CH_U_L135	1,0	0 (inactivo)
CH_U_180	CH_U_R110, CH_U_L110	1,0	0 (inactivo)
CH_U_180	CH_M_180	0,85	2
CH_U_180	CH_M_R110, CH_M_L110	0,85	2
CH_U_180	CH_U_R030, CH_U_L030	0,8	0 (inactivo)
CH_U_180	CH_M_R030, CH_M_L030	0,85	2
CH_T_000	ALL_U	1,0	3
CH_T_000	ALL_M	1,0	4
CH_L_000	CH_M_000	1,0	0 (inactivo)
CH_L_000	CH_M_L030, CH_M_R030	1,0	0 (inactivo)
CH_L_000	CH_M_L030, CH_M_R060	1,0	0 (inactivo)
CH_L_000	CH_M_L060, CH_M_R030	1,0	0 (inactivo)
CH_L_L045	CH_M_L030	1,0	0 (inactivo)

{continuación}

Entrada (Origen)	Salida (Destino)	Ganancia	Índice EQ
CH_L_R045	CH_M_R030	1,0	0 (inactivo)
CH_LFE1	CH_LFE2	1,0	0 (inactivo)
CH_LFE1	CH_M_L030, CH_M_R030	1,0	0 (inactivo)
CH_LFE2	CH_LFE1	1,0	0 (inactivo)
CH_LFE2	CH_M_L030, CH_M_R030	1,0	0 (inactivo)

**[0064]** Las etiquetas utilizadas en la tabla 1 correspondientes a los respectivos canales se deben interpretar de la siguiente manera. Las letras "CH" significan "Canal". La letra "M" representa "plano horizontal del oyente", es decir, un ángulo de elevación de 0°. Este es el plano en el cual están situados los altavoces en la configuración normal 2D  
 5 tales como estéreo o 5.1. La letra "L" representa un plano inferior, es decir, un ángulo de elevación < 0°. La letra "U" representa un plano superior, es decir, un ángulo de elevación > 0°, como por ejemplo, 30° como altavoz superior en una configuración 3D. La letra "T" representa el canal superior, es decir, un ángulo de elevación de 90°, que también se conoce como canal "la voz de Dios". Situado después de una de las etiquetas M/L/U/T es un rótulo correspondiente a izquierdo (L) o derecho (R) seguido por el ángulo azimutal.

10 **[0065]** Por ejemplo, CH\_M\_L030 y CH\_M\_R030 representan el canal izquierdo y derecho de una configuración estéreo convencional. El ángulo azimutal y el ángulo de elevación correspondientes a cada canal están indicados en la Tabla 1, excepto por los canales LFE y el último canal vacío.

15 **[0066]** La Tabla 1 muestra una matriz de reglas en la cual una o más reglas están asociadas a cada canal de entrada (canal de origen). Como se puede ver en la Tabla 1, cada regla define uno o más canales de salida (canales de destino), con los cuales se ha de poner en correspondencia el canal de entrada. Además, cada regla define el valor de ganancia G en su tercera columna. Cada regla define asimismo un índice EQ que indica si se debe aplicar un filtro de ecualización o no, y en caso afirmativo, el filtro de ecualización específico (índice EQ 1 a 4) que se debe aplicar.  
 20 La correspondencia del canal de entrada con un canal de salida se realiza con la ganancia G dada en la columna 3 de la Tabla 1. La correspondencia del canal de entrada con dos canales de salida (indicado en la segunda columna) se realiza aplicando ajuste panorámico entre los dos canales de salida, donde las ganancias de ajuste panorámico  $g_1$  y  $g_2$  que surgen como resultado de la aplicación de la ley de ajuste panorámico se multiplican además por la regla respectiva (columna tres de la Tabla 1). Se aplican reglas especiales para el canal superior. Según una primera regla,  
 25 se pone en correspondencia el canal superior con todos los canales de salida del plano superior, lo que está indicado por ALL\_U, y según una segunda regla (con menos prioridad), el canal superior se pone en correspondencia con todos los canales de salida del plano horizontal del oyente, indicado por ALL\_M.

30 **[0067]** Al considerar las reglas indicadas en la Tabla 1, las reglas que definen la correspondencia de un canal de entrada con un canal de salida y que tienen un índice EQ de 1 o 2 forman parte de la invención.

**[0068]** Como se puede apreciar en la Tabla 1, se aplica uno de los filtros ecualizadores 1 a 4 si se pone en correspondencia un canal de entrada elevado con uno o más canales inferiores. Los valores de ganancia del ecualizador  $G_{EQ}$  se pueden determinar de la siguiente manera sobre la base de las frecuencias centrales normalizadas  
 35 dadas en la Tabla 2 y basándose en los parámetros consignados en la Tabla 3.

**Tabla 2: Frecuencias centrales normalizadas de las bandas de 77 bancos de filtros**

Frecuencia normalizada [0, 1]
0,00208330
0,00587500
0,00979170
0,01354200
0,01691700
0,02008300
0,00458330
0,00083333
0,03279200
0,01400000
0,01970800
0,02720800
0,03533300
0,04283300
0,04841700
0,02962500
0,05675000

# ES 3 010 378 T3

Frecuencia normalizada [0, 1]
0,07237500
0,08800000
0,10362000
0,11925000
0,13487000
0,15050000
0,16612000
0,18175000
0,19737000
0,21300000
0,22862000
0,24425000
0,25988000
0,27550000
0,29113000
0,30675000
0,32238000
0,33800000
0,35363000
0,36925000
0,38488000
0,40050000
0,41613000
0,43175000
0,44738000
0,46300000
0,47863000
0,49425000
0,50987000
0,52550000
0,54112000
0,55675000
0,57237000
0,58800000
0,60362000
0,61925000
0,63487000
0,65050000
0,66612000
0,68175000
0,69737000
0,71300000
0,72862000
0,74425000
0,75987000
0,77550000
0,79112000
0,80675000
0,82237000
0,83800000
0,85362000
0,86925000
0,88487000
0,90050000
0,91612000
0,93175000
0,94737000
0,96300000
0,97454000
0,99904000

Tabla 3: Parámetros del ecualizador

Ecualizador	$P_f$ [Hz]	$P_Q$	$P_g$ [dB]	$g$ [dB]
$G_{EQ,1}$	12000	0,3	-2	1,0
$G_{EQ,2}$	12000	0,3	-3,5	1,0
$G_{EQ,3}$	200, 1300, 600	0,3, 0,5, 1,0	-6,5, 1,8, 2,0	0,7
$G_{EQ,4}$	5000, 1100	1,0, 0,8	4,5, 1,8	-3,1
$G_{EQ,5}$	35	0,25	-1,3	1,0

[0069]  $G_{EQ}$  consiste en valores de ganancia por banda  $k$  de frecuencia y el índice de ecualizador  $e$ . Cinco ecualizadores predefinidos son combinaciones de diferentes filtros de pico (o filtros de banda). Como se puede ver en la Tabla 3, los ecualizadores  $G_{EQ,1}$ ,  $G_{EQ,2}$  y  $G_{EQ,5}$  incluyen un solo filtro de pico, el ecualizador  $G_{EQ,3}$  incluye tres filtros de pico y el ecualizador  $G_{EQ,4}$  incluye dos filtros de pico. Cada ecualizador es una cascada en serie de uno o más filtros de pico y una ganancia:

$$G_{EQ,e}^k = 10^{\frac{g}{20}} \prod_{n=1}^N \text{pico}(\text{banda}(k) \cdot f_s/2, P_{f,n}, P_{Q,n}, P_{g,n})$$

10

donde  $\text{banda}(k)$  es la frecuencia central normalizada de la banda de frecuencia  $j$ , especificada en la Tabla 2,  $f_s$  es la frecuencia de muestreo y la función  $\text{pico}()$  corresponde a  $G$  negativo

$$\text{pico}(b, f, Q, G) = \frac{b^4 + \left(\frac{1}{Q^2} - 2\right) f^2 b^2 + f^4}{\sqrt{b^4 + \left(\frac{10^{\frac{-G}{20}}}{Q^2} - 2\right) f^2 b^2 + f^4}}$$

Ecuación 1

15 y por otra parte

$$\text{pico}(b, f, Q, G) = \frac{b^4 + \left(\frac{10^{\frac{G}{20}}}{Q^2} - 2\right) f^2 b^2 + f^4}{\sqrt{b^4 + \left(\frac{1}{Q^2} - 2\right) f^2 b^2 + f^4}}$$

Ecuación 2

[0070] Los parámetros correspondientes a los ecualizadores están especificados en la Tabla 3 En las Ecuaciones anteriores 1 y 2,  $b$  viene dado por  $\text{banda}(k) \cdot f_s/2$ ,  $Q$  viene dado por  $P_Q$  para el filtro de pico respectivo (1 a  $n$ ),  $G$  viene dado por  $P_g$  para el filtro de pico respectivo y  $f$  viene dado por  $P_f$  para el filtro de pico respectivo.

[0071] A modo de ejemplo, los valores de ganancia del ecualizador  $G_{EQ,4}$  correspondientes al ecualizador que tiene el índice 4 se calculan con los parámetros de filtro tomados de la fila pertinente de la Tabla 3. La Tabla 3 enumera dos series de parámetros correspondientes a los filtros de pico para  $G_{EQ,4}$ , es decir, series de parámetros correspondientes a  $n=1$  y  $n=2$ . Los parámetros son la frecuencia de los picos  $P_f$  en Hz, el factor de calidad del filtro de pico  $P_Q$ , la ganancia  $P_g$  (en dB) que se aplica a la frecuencia de los picos y una ganancia general  $g$  en dB que se aplica a la cascada de los dos filtros de pico (cascada de filtros para los parámetros  $n=1$  y  $n=2$ ).

[0072] Por consiguiente

30

$$\begin{aligned} G_{EQ,4} &\approx 10^{\frac{-3,1}{20}} \cdot \text{pico}(\text{banda}(k) \cdot f_s/2, P_{f,1}, P_{Q,1}, P_{g,1}) \cdot \text{pico}(\text{banda}(k) \cdot f_s/2, P_{f,2}, P_{Q,2}, P_{g,2}) \\ &= 10^{\frac{-3,1}{20}} \cdot \text{pico}(\text{banda}(k) \cdot f_s/2, 5000, 1,0, 4,5) \cdot \text{pico}(\text{banda}(k) \cdot f_s/2, 1100, 0,8, 1,8) \\ &= 10^{\frac{-3,1}{20}} \cdot \frac{b^4 + \left(\frac{10^{\frac{4,5}{20}}}{1^2} - 2\right) 5000^2 b^2 + 5000^4}{\sqrt{b^4 + \left(\frac{1}{1^2} - 2\right) 5000^2 b^2 + 5000^4}} \cdot \frac{b^4 + \left(\frac{10^{\frac{1,8}{20}}}{0,8^2} - 2\right) 1100^2 b^2 + 1100^4}{\sqrt{b^4 + \left(\frac{1}{0,8^2} - 2\right) 1100^2 b^2 + 1100^4}} \end{aligned}$$

[0073] La definición del ecualizador antes citada define ganancias  $G_{EQ,4}$  de fase cero de modo independiente por cada banda  $k$  de frecuencia. Cada banda  $k$  está especificada por su  $\text{banda}(k)$  de frecuencia central normalizada donde  $0 \leq \text{banda}(k) \leq 1$ . Obsérvese que la banda de frecuencia normalizada=1 corresponde a la frecuencia no normalizada  $f_s/2$ , donde  $f_s$  denota la frecuencia de muestreo. Por lo tanto,  $\text{banda}(k) \cdot f_s/2$  denota la frecuencia central no normalizada de la banda  $k$  en Hz.

[0074] Por consiguiente, se pueden usar filtros ecualizadores diferentes en los ejemplos que se han descrito.

[0075] La Tabla 4 muestra canales de ejemplo con su ángulo azimutal y su ángulo de elevación respectivos asociados.

5

**Tabla 4: Canales con los correspondientes ángulos azimutal y de elevación**

Canal	Azimut [grados]	Elevación [grados]
CH_M_000	0	0
CH_M_L030	+30	0
CH_M_R030	-30	0
CH_M_L060	+60	0
CH_M_R060	-60	0
CH_M_L090	+90	0
CH_M_R090	-90	0
CH_M_L110	+110	0
CH_M_R110	-110	0
CH_M_L135	+135	0
CH_M_R135	-135	0
CH_M_180	180	0
CH_U_000	0	+35
CH_U_L045	+45	+35
CH_U_R045	-45	+35
CH_U_L030	+30	+35
CH_U_R030	-30	+35
CH_U_L090	+90	+35
CH_U_R090	-90	+35
CH_U_L110	+110	+35
CH_U_R110	-110	+35
CH_U_L135	+135	+35
CH_U_R135	-135	+35
CH_U_180	180	+35
CH_T_000	0	+90
CH_L_000	0	-15
CH_L_L045	+45	-15
CH_L_R045	-45	-15
CH_LFE1	n/d	n/d
CH_LFE2	n/d	n/d
CH_EMPTY	n/d	n/d

[0076] En las realizaciones de la invención, el ajuste panorámico entre dos canales de destino se puede realizar mediante la aplicación de ajuste panorámico basado en la amplitud por la ley de tangentes. En el ajuste panorámico de un canal de origen con canales de destino primero y segundo se calcula un coeficiente de ganancia  $G_1$  correspondiente al primer canal de destino y se calcula un coeficiente de ganancia  $G_2$  correspondiente al segundo canal de destino:

15  $G_1 = (\text{columna de valores de Ganancia de la Tabla 1}) * g_1$   
 $G_2 = (\text{columna de valores de Ganancia de la Tabla 1}) * g_2$

[0077] Las ganancias  $g_1$  y  $g_2$  se calculan aplicando el ajuste panorámico basado en amplitud por la ley de tangentes de la siguiente manera:

- 20 • se repliegan los ángulos azimutales del canal de destino para que sean positivos  
 • los ángulos azimutales de los canales de destino son  $\alpha_1$  y  $\alpha_2$  (véase la Tabla 4).  
 • el ángulo azimutal del canal de origen (objetivo de ajuste panorámico) es  $\alpha_{src}$ .

25 
$$\alpha_0 = \frac{|\alpha_1 - \alpha_2|}{2}$$

$$\alpha_{centro} = \frac{\alpha_1 + \alpha_2}{2}$$

$$\alpha = (\alpha_{centro} - \alpha_{src}) \cdot \text{sgn}(\alpha_2 - \alpha_1)$$

con

$$g = \frac{\tan \alpha_0 - \tan \alpha + 10^{-10}}{\tan \alpha_0 + \tan \alpha + 10^{-10}}$$

**[0078]** En otras realizaciones, se pueden aplicar diferentes leyes de ajuste panorámico.

5 **[0079]** En principio, los ejemplos pretenden modelizar un mayor número de canales acústicos en la configuración de canales de entrada por medio de correspondencias de canal cambiadas y modificaciones de las señales en la configuración de los canales de salida. En comparación con las técnicas simples que con frecuencia se consideran más exigentes, menos diversas y menos envolventes que la configuración de canales de entrada, mediante el empleo de las realizaciones de la invención se puede mejorar la diversidad espacial y la experiencia auditiva en general y hacerlas más agradables.

**[0080]** Dicho de otro modo, en las realizaciones se mezclan dos o más canales de entrada entre sí en una aplicación de mezcla descendente, en la cual se aplica un módulo de procesamiento en forma de un filtro de ecualización a una de las señales de entrada para conservar las diferentes características de las diferentes trayectorias de transmisión desde el canal original de entrada a los oídos del oyente. En particular, los filtros ecualizadores pueden compensar la pérdida de diferentes timbres de los canales de entrada con diferentes elevaciones asignadas a los mismos. En los ejemplos, el módulo de procesamiento puede encaminar al menos una de las señales de entrada hacia varios altavoces de salida para generar una trayectoria de transmisión diferente hacia el oyente, conservando de esa manera la diversidad espacial de los canales de entrada. En los ejemplos, se pueden aplicar modificaciones de filtro y encaminamiento por separado o en combinación. En los ejemplos de la invención, el módulo de procesamiento se puede reproducir a través de uno o varios altavoces.

**[0081]** Aunque se han descrito algunos aspectos en el contexto de un aparato, es obvio que estos aspectos también representan una descripción del procedimiento correspondiente, en el cual un bloque o dispositivo corresponde a una etapa del procedimiento o a una característica de una etapa del procedimiento. De manera análoga, los aspectos descritos en el contexto de una etapa del procedimiento también representan una descripción de un bloque o elemento correspondiente o de una característica de un aparato correspondiente. Algunas o todas las etapas del procedimiento pueden ser ejecutadas por medio de (o usando) un aparato de hardware, como por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. Algunas o más de las etapas más importantes del procedimiento pueden ser ejecutadas por dicho aparato. En las realizaciones de la invención, los procedimientos descritos en esta invención son implementados por un procesador o implementados por un ordenador.

**[0082]** Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención se pueden implementar en hardware o en software. La implementación se puede realizar empleando un medio de almacenamiento no transitorio tal como un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un Blue-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene almacenadas en la misma, señales de control legibles electrónicamente, que cooperan (o tienen capacidad para cooperar) con un sistema informático programable de tal manera que se ejecute el procedimiento respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por un ordenador.

**[0083]** Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable, de modo que se lleva a cabo uno de los procedimientos descritos en esta invención.

45 **[0084]** En general, las realizaciones de la presente invención se pueden implementar en forma de producto de programa informático con un código de programa, donde el código de programa cumple la función de ejecutar uno de los procedimientos al ejecutarse el producto de programa informático en un ordenador. El código de programa se puede almacenar, por ejemplo, en un soporte legible por una máquina.

50 **[0085]** Otras realizaciones comprenden el programa informático para ejecutar uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por una máquina.

**[0086]** Dicho de otro modo, una realización del procedimiento de la invención consiste, por lo tanto, en un programa informático que tiene un código de programa para realizar uno de los procedimientos descritos en esta invención al ejecutarse el programa informático en un ordenador.

**[0087]** Una realización adicional del procedimiento inventivo es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, registrado en él, el programa informático para realizar uno de los procedimientos descritos en esta invención. El soporte de datos, el medio de almacenamiento digital o el medio grabado suelen ser tangibles y/o no transitorios.

**[0088]** Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, un dispositivo lógico programable, programado, configurado o adaptado para ejecutar uno de los procedimientos descritos en esta invención.

5 **[0089]** Una realización adicional comprende un ordenador donde se ha instalado el programa informático para ejecutar uno de los procedimientos descritos en esta invención.

10 **[0090]** En algunas realizaciones, se puede usar un dispositivo lógico programable (por ejemplo, una matriz de puertas programables de campo) para ejecutar algunas o todas las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, una matriz de puertas programables de campo puede cooperar con un microprocesador con el fin de ejecutar uno de los procedimientos descritos en esta invención. Por lo general, los procedimientos son ejecutados preferentemente por cualquier aparato de hardware.

15 **[0091]** Las realizaciones descritas anteriormente son meramente ilustrativas de los principios de la presente invención. Se entiende que para los expertos en la materia serán evidentes las modificaciones y variaciones de las disposiciones y detalles descritos en esta invención. Por lo tanto, se pretende estar limitado sólo por el alcance de las siguientes reivindicaciones de patente y no por los detalles específicos presentados a modo de descripción y explicación de los ejemplos mostrados en esta invención.

REIVINDICACIONES

1. Un aparato (60) para la correspondencia de un primer canal de entrada de altavoz (38) y un segundo canal de entrada de altavoz (14) de una configuración de canales de entrada de altavoz con un primer y segundo canales de altavoz de salida de una configuración de canal de altavoz de salida, donde cada canal de entrada de altavoz y cada canal de salida de altavoz tiene una dirección con respecto a una posición central del oyente (P), donde el primer y segundo canales de altavoz de entrada (38, 14) tienen diferentes ángulos de elevación con respecto a un plano horizontal del oyente (300), donde el aparato está configurado para:
- 10 la puesta en correspondencia del primer canal de entrada de altavoz (38) con un primer canal de salida de altavoz (42) de la configuración de canales de salida de altavoz; y  
la puesta en correspondencia del segundo canal de entrada de altavoz (14) con el primer canal de salida de altavoz (42) y el segundo canal de salida de altavoz (44) de la configuración de canal de salida de altavoz, que comprende el procesamiento del segundo canal de entrada de altavoz (14) mediante la aplicación de al menos un filtro de
- 15 **ecualización (62) al segundo canal de entrada de altavoz (14),**  
**caracterizado porque** el aparato está configurado para realizar la puesta en correspondencia del segundo canal de altavoz de entrada (14) al primer canal de altavoz de salida (42) y al segundo canal de altavoz de salida (44) según las siguientes reglas de correspondencia:

Entrada (origen)	Salida (destino)	Ganancia	Índice de EQ
CH_U_L090	CH_M_L030, CH_M_L110	0,85	2
CH_U_R090	CH_M_R030, CH_M_R110	0,85	2
CH_U_180	CH_M_R110, CH_M_L110	0,85	2
CH_U_180	CH_M_R030, CH_M_L030	0,85	2

- 20 donde  
la Entrada (Fuente) representa el segundo canal de altavoz de entrada (14),  
la Salida (Destino) representa el primer y segundo canal de altavoz de salida (42, 44),  
los caracteres "CH" representan "canal",  
25 el carácter "M" representa un ángulo de elevación de 0°,  
el carácter "U" representa un ángulo de elevación > 0°,  
después de una de las etiquetas M/U hay una etiqueta para izquierda (L) o derecha (R) seguida del ángulo de azimut,  
la Ganancia es una ganancia utilizada en la puesta en correspondencia del canal de altavoz de entrada respectivo  
30 al canal de altavoz de salida respectivo, y  
el índice EQ indica qué ecualizador se va a aplicar,  
donde la puesta en correspondencia del segundo canal de altavoz de entrada (14) al primer y segundo canal de altavoz de salida (42, 44) se realiza aplicando un ajuste panorámico entre el primer y segundo canal de altavoz de salida (42, 44),  
35 donde las ganancias de ajuste panorámico se multiplican adicionalmente por la Ganancia,  
con el siguiente parámetro de ecualizador para el ecualizador 2:

Ecualizador	P <sub>f</sub> [Hz]	P <sub>Q</sub>	P <sub>g</sub> [dB]	g [dB]
G <sub>EQ,2</sub>	12000	0,3	-3,5	1,0

- 40 donde G<sub>EQ,2</sub> consta de valores de ganancia por banda k de frecuencia y el índice del ecualizador P<sub>f</sub> es una frecuencia pico en Hz, P<sub>Q</sub> es un factor de calidad de filtro pico, P<sub>g</sub> es una ganancia en dB aplicada a la frecuencia pico y g en dB es una ganancia general aplicada a un filtro de pico,  
donde el ecualizador G<sub>EQ,2</sub> incluye un filtro de pico único, donde el ecualizador G<sub>EQ,2</sub> es una cascada en serie del filtro de pico único y una ganancia:

$$G_{EQ,k}^k = 10^{\frac{g}{20}} \text{pico}(banda(k) \cdot f_s/2, P_f, P_Q, P_g)$$

- donde (banda(k)) es una frecuencia central normalizada de cada banda de frecuencia, f<sub>s</sub> es una frecuencia de muestreo y función (pico())

$$pico(b, f, Q, G) = \frac{b^4 + \left(\frac{1}{Q^2} - 2\right) f^2 b^2 + f^4}{\sqrt{b^4 + \left(\frac{10^{16}}{Q^2} - 2\right) f^2 b^2 + f^4}}$$

donde b viene dada por banda(k) fs/2, Q viene dada por P<sub>Q</sub> para el respectivo filtro de pico, G viene dada por P<sub>g</sub> para el respectivo filtro de pico y f viene dada por P<sub>f</sub> para el filtro de pico,  
 5 donde cada banda k está especificada por su frecuencia central normalizada *banda(k)* donde 0<=banda<=1, donde la banda de frecuencia normalizada=1 corresponde a la frecuencia no normalizada f<sub>s</sub>/2, donde f<sub>s</sub> denota la frecuencia de muestreo, y *banda(k) • fs/2* denota la frecuencia central no normalizada de la banda k en Hz.

2. Un procedimiento para la correspondencia de un primer canal de entrada de altavoz (12) y un segundo canal de entrada de altavoz (14) de una configuración de canales de entrada de altavoz a un primer y segundo canal de salida de altavoz de una configuración de canales de salida de altavoz, donde cada canal de entrada de altavoz y cada canal de salida de altavoz tienen una dirección con respecto a una posición central del oyente (P), donde el primer y segundo canal de entrada de altavoz (38, 14) tienen diferentes ángulos de elevación con respecto a un plano horizontal del oyente (300), que comprende:

15 la puesta en correspondencia del primer canal de entrada de altavoz (38) con el primer canal de salida de altavoz (42) de la configuración de canales de salida de altavoz; y  
 la puesta en correspondencia del segundo canal de entrada de altavoz (14) al primer canal de salida de altavoz (42) y al segundo canal de salida de altavoz (44) de la configuración de canal de salida de altavoz, que comprende  
 20 el procesamiento del segundo canal de entrada de altavoz (14) mediante la aplicación de un filtro de ecualización (62) al segundo canal de entrada de altavoz (14),  
**caracterizado porque** la puesta en correspondencia del segundo canal de altavoz de entrada (14) al primer canal de altavoz de salida (42) y al segundo canal de altavoz de salida (44) se realiza según las siguientes reglas de puesta en correspondencia:

Entrada (origen)	Salida (destino)	Ganancia	Índice de EQ
CH_U_L090	CH_M_L030, CH_M_L110	0,85	2
CH_U_R090	CH_M_R030, CH_M_R110	0,85	2
CH_U_180	CH_M_R110, CH_M_L110	0,85	2
CH_U_180	CH_M_R110, CH_M_L110	0,85	2
CH_U_180	CH_M_R030, CH_M_L030	0,85	2

donde  
 la Entrada (Fuente) representa el segundo canal de altavoz de entrada (14),  
 la Salida (Destino) representa el primer y segundo canal de altavoz de salida (42, 44),  
 30 los caracteres "CH" representan "canal",  
 el carácter "M" representa un ángulo de elevación de 0°,  
 el carácter "U" representa un ángulo de elevación > 0°,  
 después de una de las etiquetas M/U hay una etiqueta para izquierda (L) o derecha (R) seguida del ángulo de azimut,  
 35 la Ganancia es una ganancia utilizada en la puesta en correspondencia del canal de altavoz de entrada respectivo al canal de altavoz de salida respectivo, y  
 el índice EQ indica qué ecualizador se va a aplicar,  
 donde la puesta en correspondencia del segundo canal de altavoz de entrada (14) al primer y segundo canal de altavoz de salida (42, 44) se realiza aplicando un ajuste panorámico entre el primer y segundo canal de altavoz de salida (42, 44),  
 40 donde las ganancias de ajuste panorámico se multiplican adicionalmente por la Ganancia,  
 con el siguiente parámetro de ecualizador para el ecualizador 2:

Ecualizador	P <sub>f</sub> [Hz]	P <sub>Q</sub>	P <sub>g</sub> [dB]	g [dB]
G <sub>EQ,2</sub>	12000	0,3	-3,5	1,0

45 donde G<sub>EQ,2</sub> consta de valores de ganancia por banda k de frecuencia y el índice del ecualizador P<sub>f</sub> es una

frecuencia pico en Hz,  $P_Q$  es un factor de calidad de filtro pico,  $P_g$  es una ganancia en dB aplicada a la frecuencia pico y  $g$  en dB es una ganancia general aplicada a un filtro de pico, donde el ecualizador  $G_{EQ,2}$  incluye un filtro de pico único, donde el ecualizador  $G_{EQ,2}$  es una cascada en serie del filtro de pico único y una ganancia:

5

$$G_{EQ,e}^k = 10^{\frac{g}{20}} \text{pico}(\text{banda}(k) \cdot f_s/2, P_f, P_Q, P_g)$$

donde  $\text{banda}(k)$  es una frecuencia central normalizada de cada banda de frecuencia,  $f_s$  es una frecuencia de muestreo y la función  $\text{pico}()$  es

10

$$\text{pico}(b, f, Q, G) = \frac{\sqrt{b^4 + \left(\frac{1}{Q^2} - 2\right) f^2 b^2 + f^4}}{\sqrt{b^4 + \left(\frac{10^{\frac{-g}{20}}}{Q^2} - 2\right) f^2 b^2 + f^4}}$$

donde  $b$  viene dada por  $\text{banda}(k) f_s/2$ ,  $Q$  viene dada por  $P_Q$  para el filtro de pico,  $G$  viene dada por  $P_g$  para el filtro de pico y  $f$  viene dada por  $P_f$  para el filtro de pico,

15

donde cada banda  $k$  está especificada por su  $\text{banda}(k)$  de frecuencia central normalizada donde  $0 \leq \text{band} \leq 1$ , donde la banda de frecuencia normalizada=1 corresponde a la frecuencia no normalizada  $f_s/2$ , donde  $f_s$  denota la frecuencia de muestreo, y  $\text{banda}(k) f_s/2$  denota la frecuencia central no normalizada de la banda  $k$  en Hz.

3. Un programa informático para realizar, cuando se ejecuta en un ordenador o un procesador, el  
20 procedimiento según la reivindicación 2.

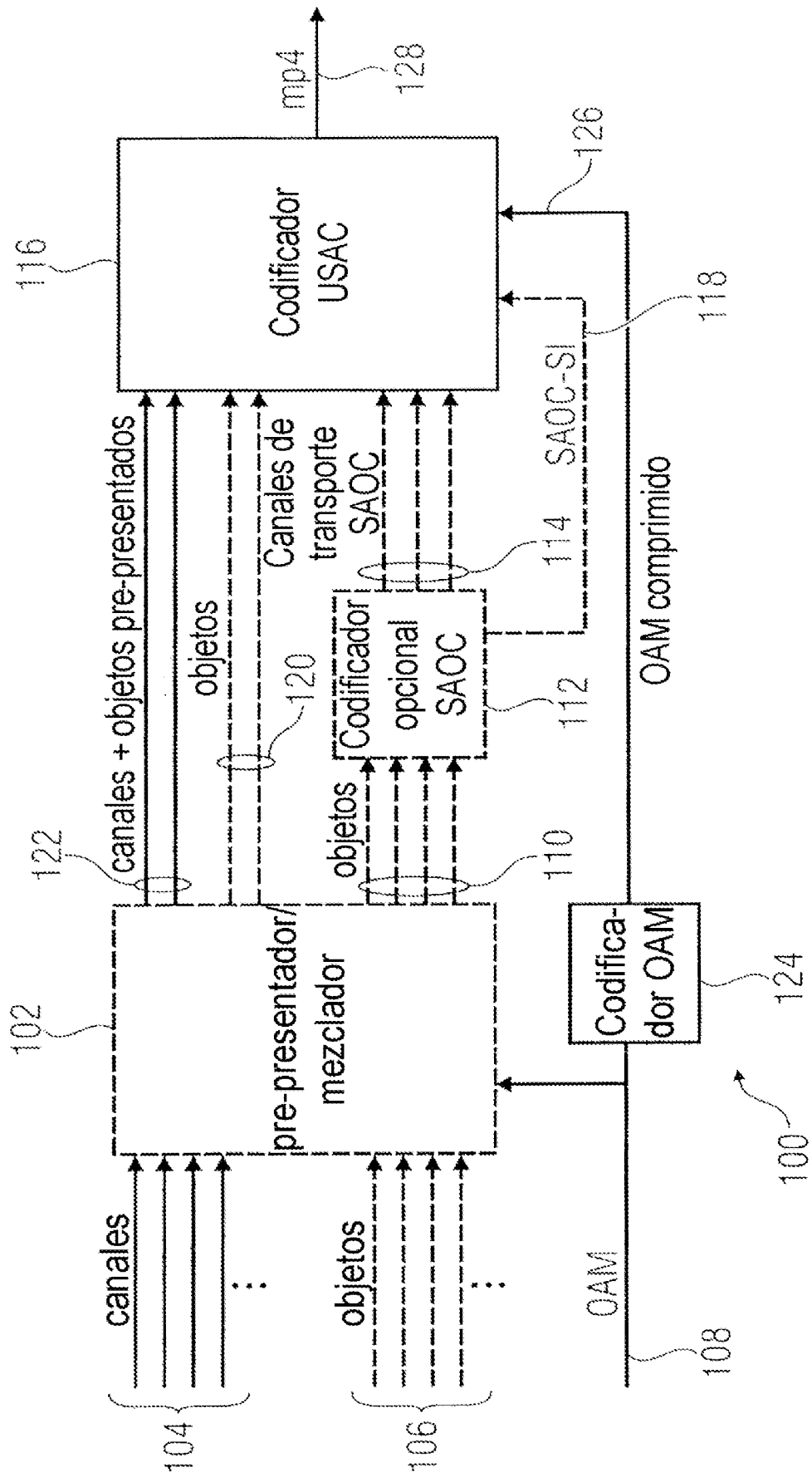


FIG 1

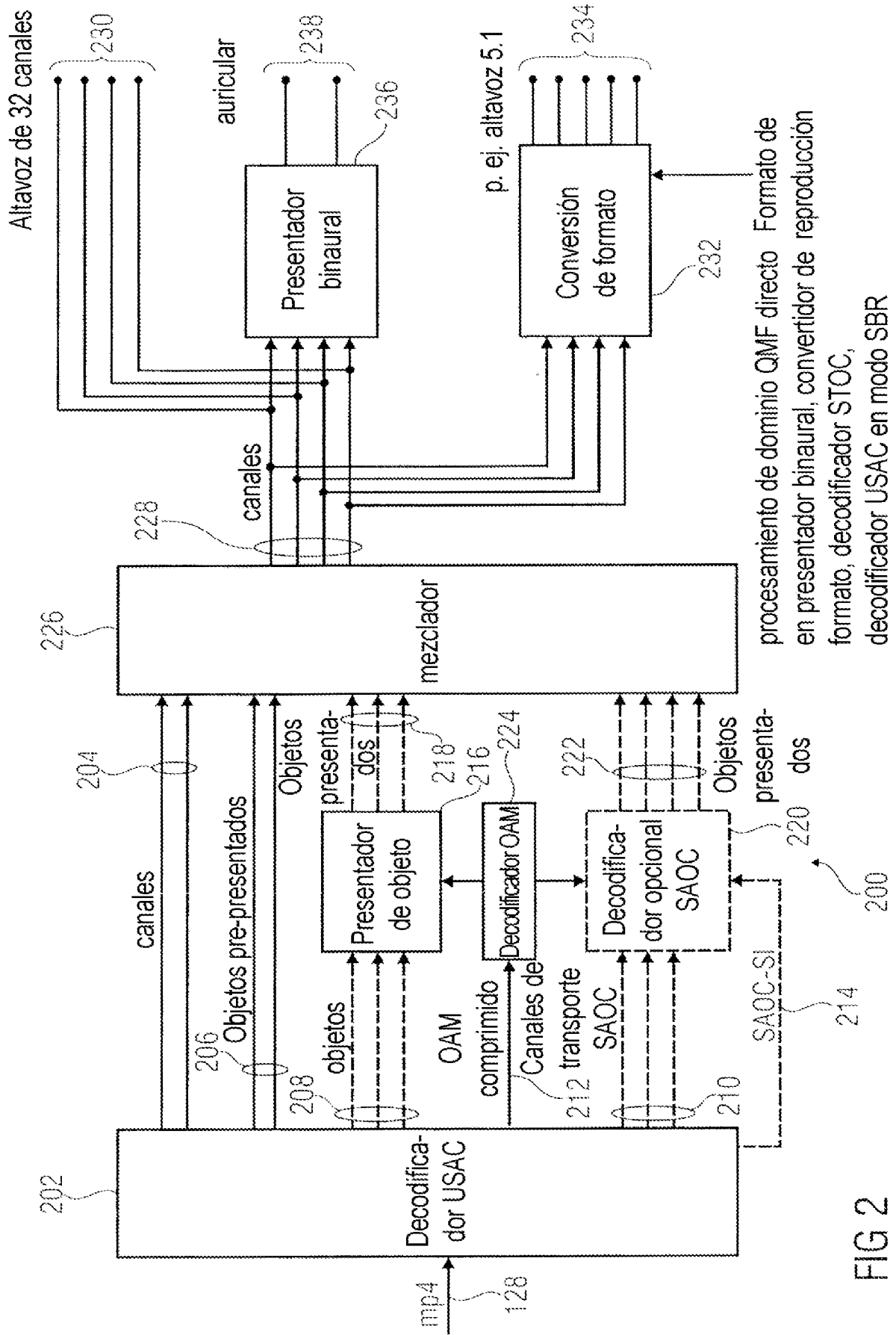


FIG 2

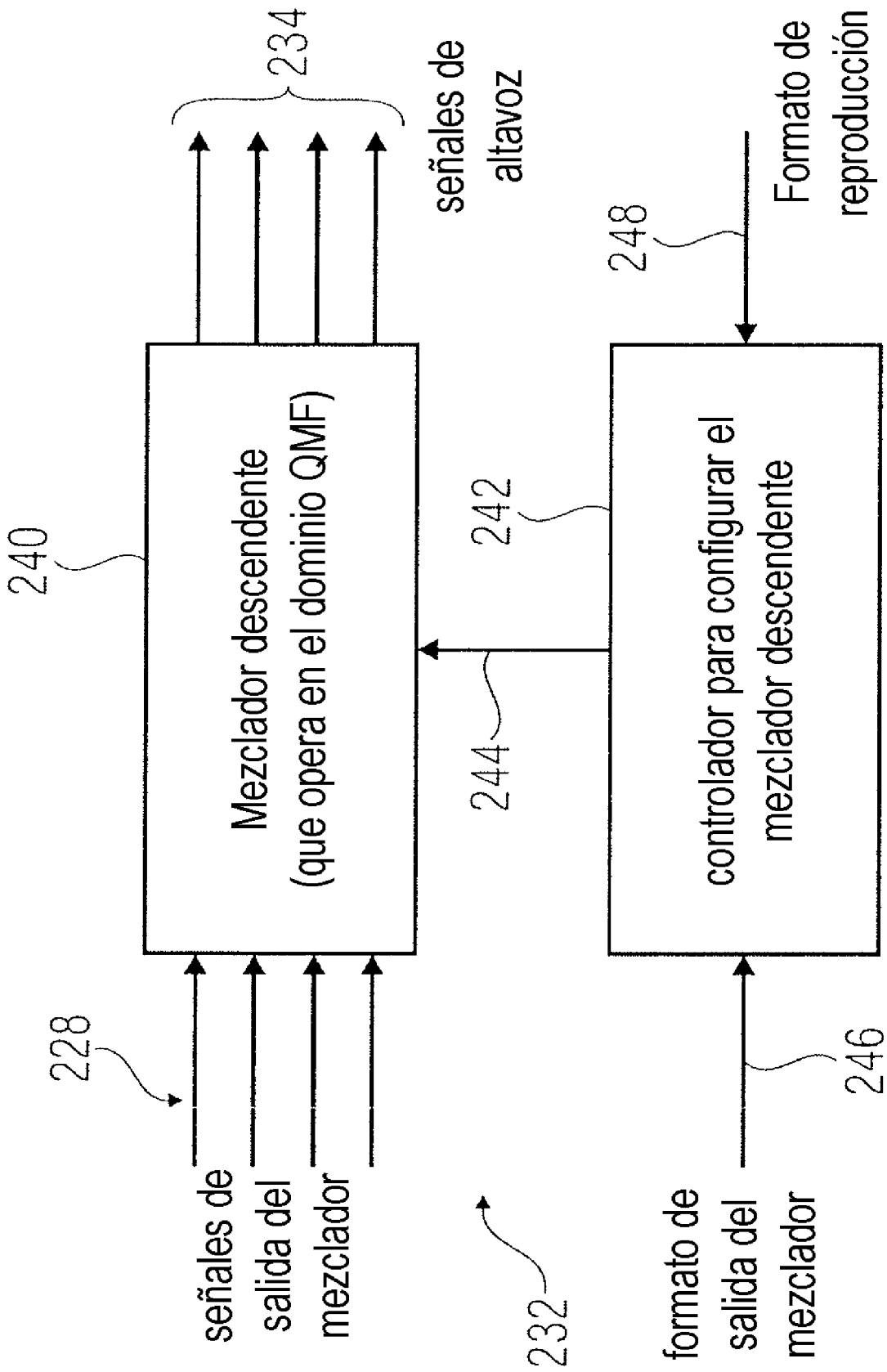


FIG 3

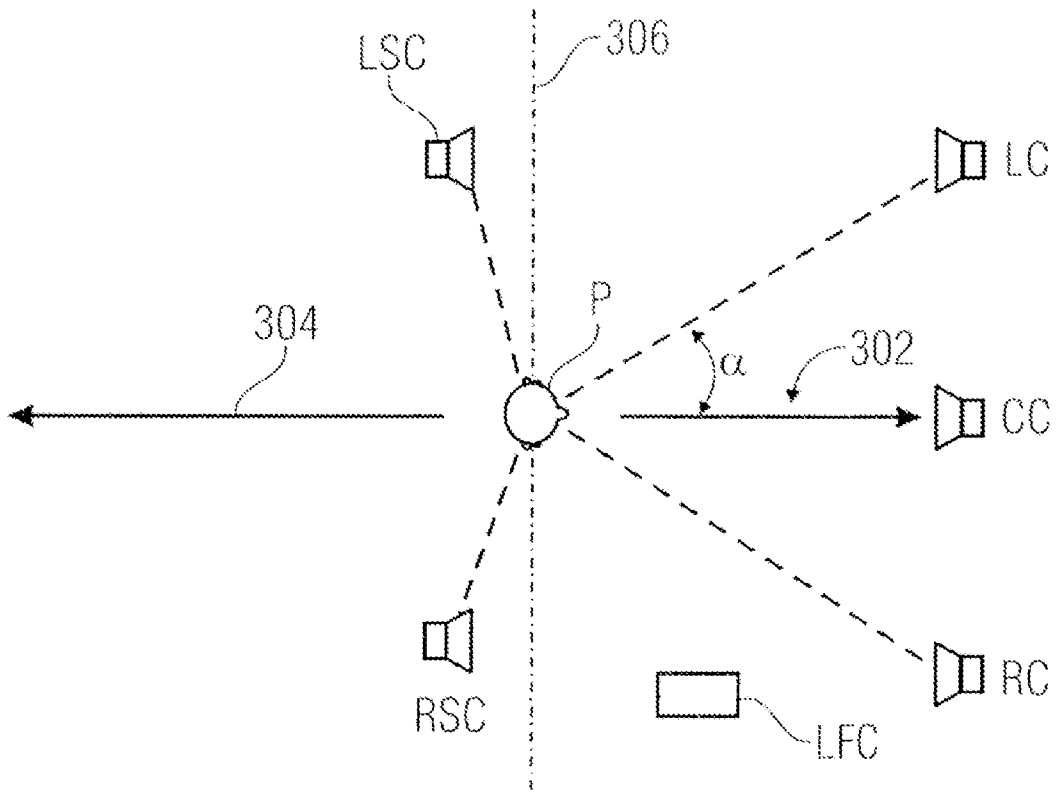


FIG 4

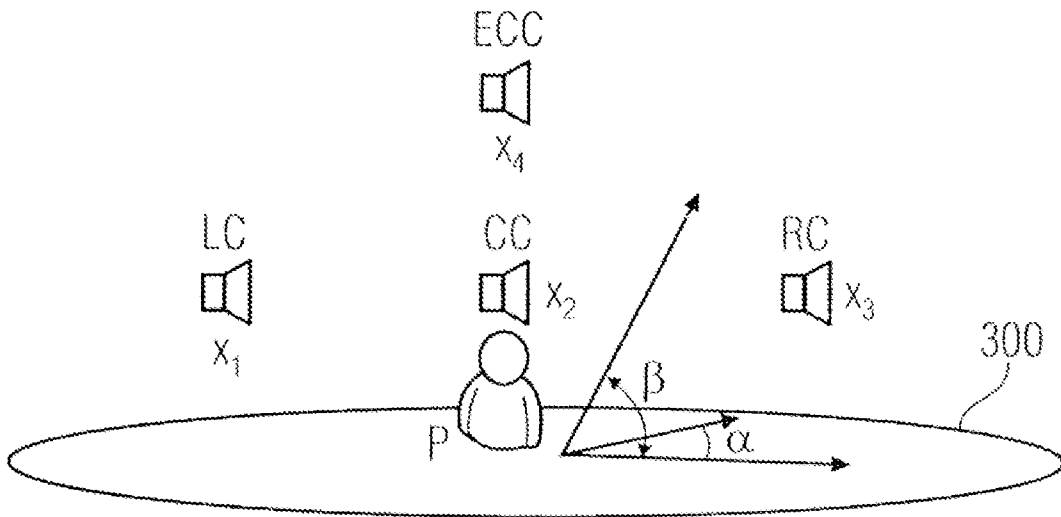


FIG 5

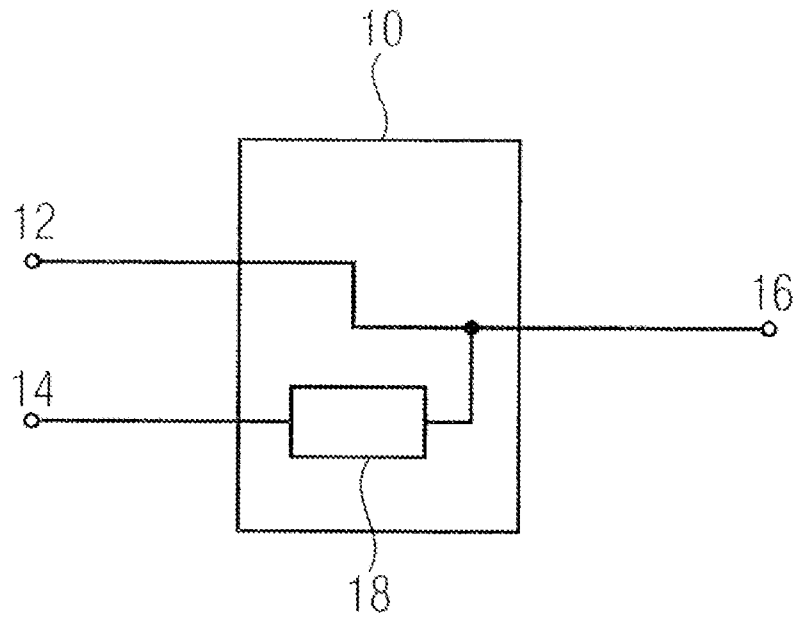


FIG 6A

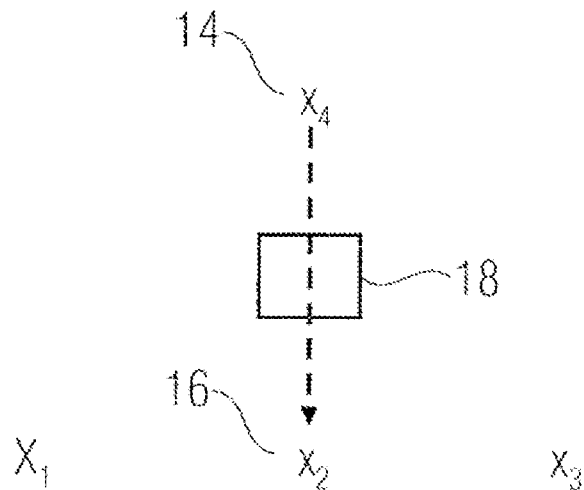


FIG 6B

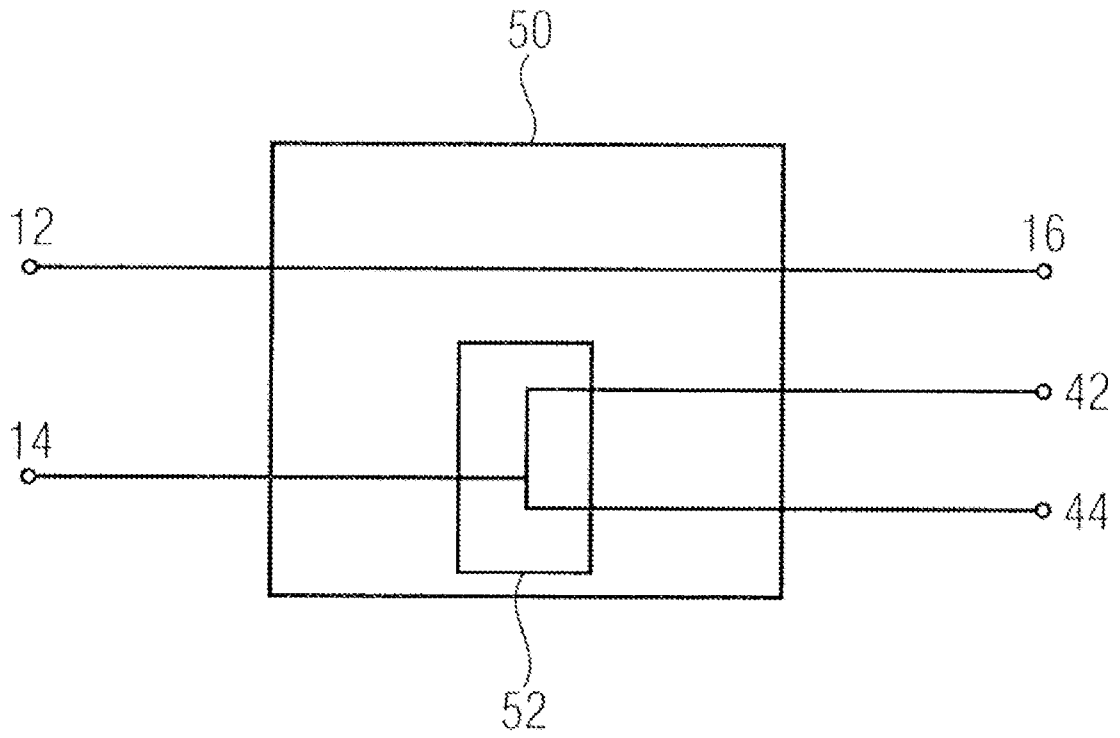


FIG 7A

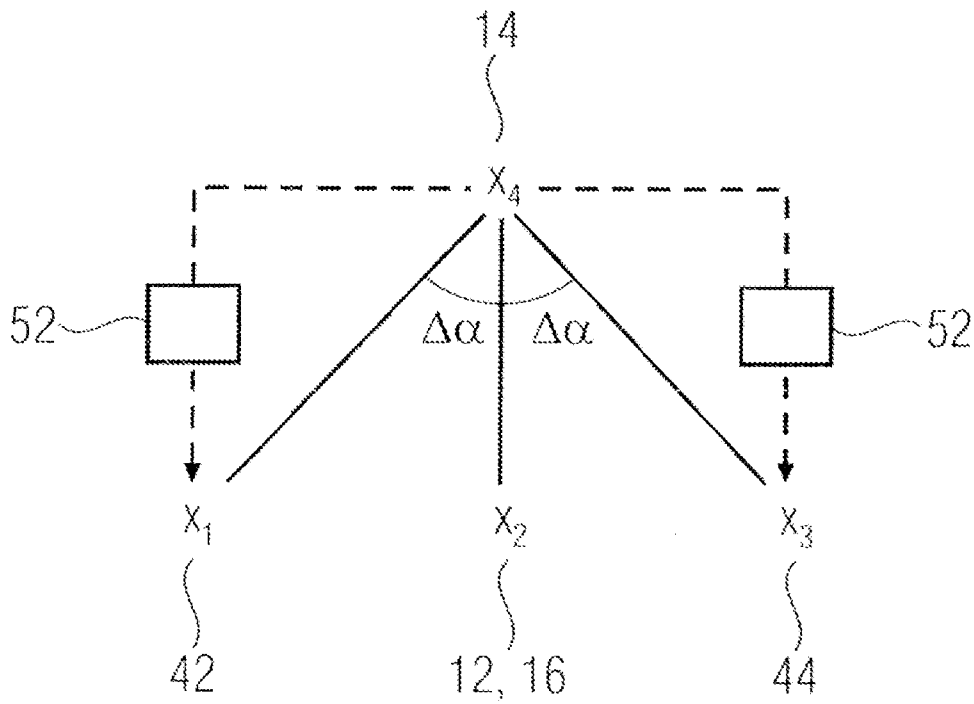


FIG 7B

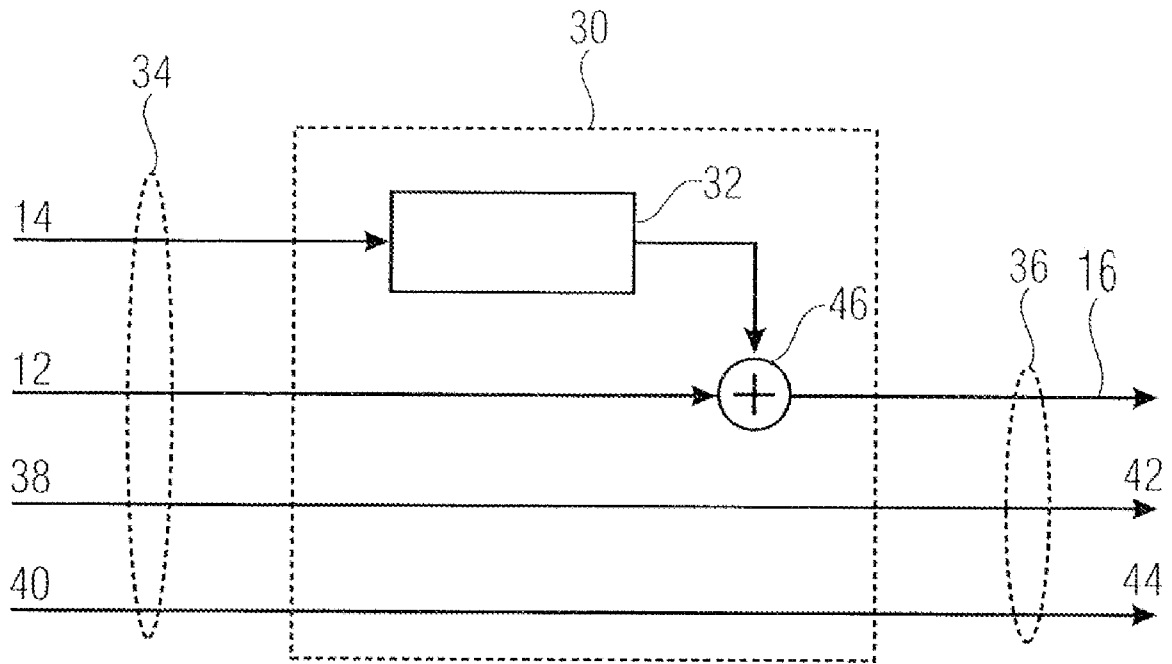


FIG 8

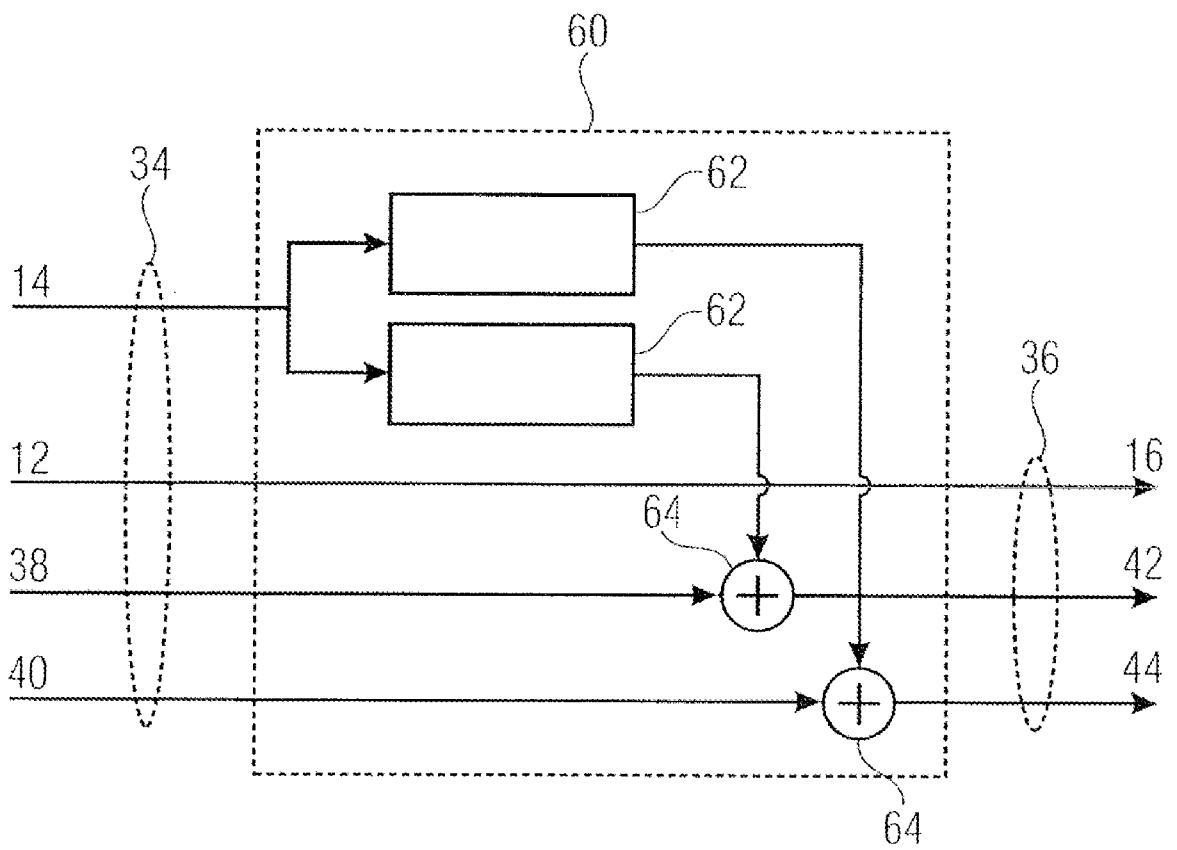


FIG 9

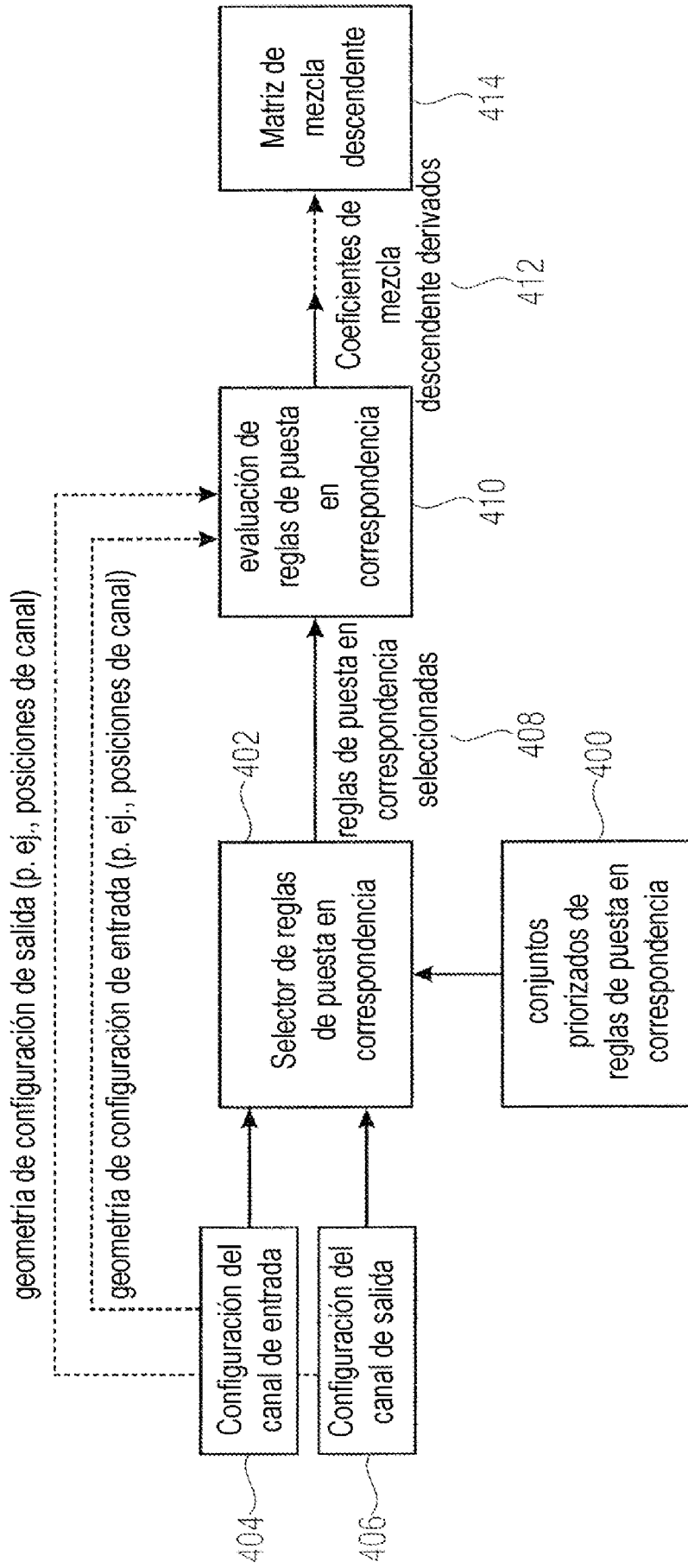


FIG 10

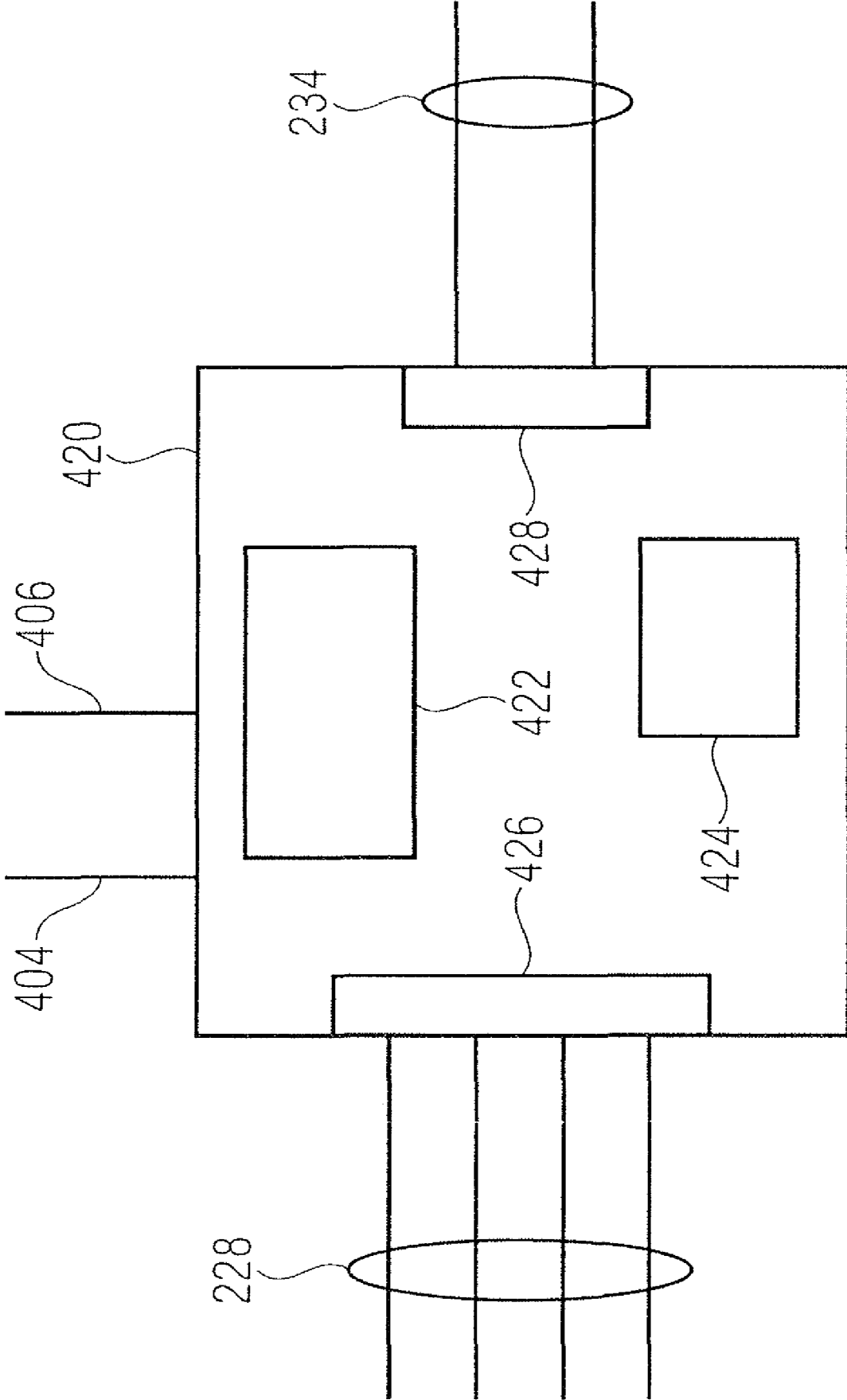


FIG 11

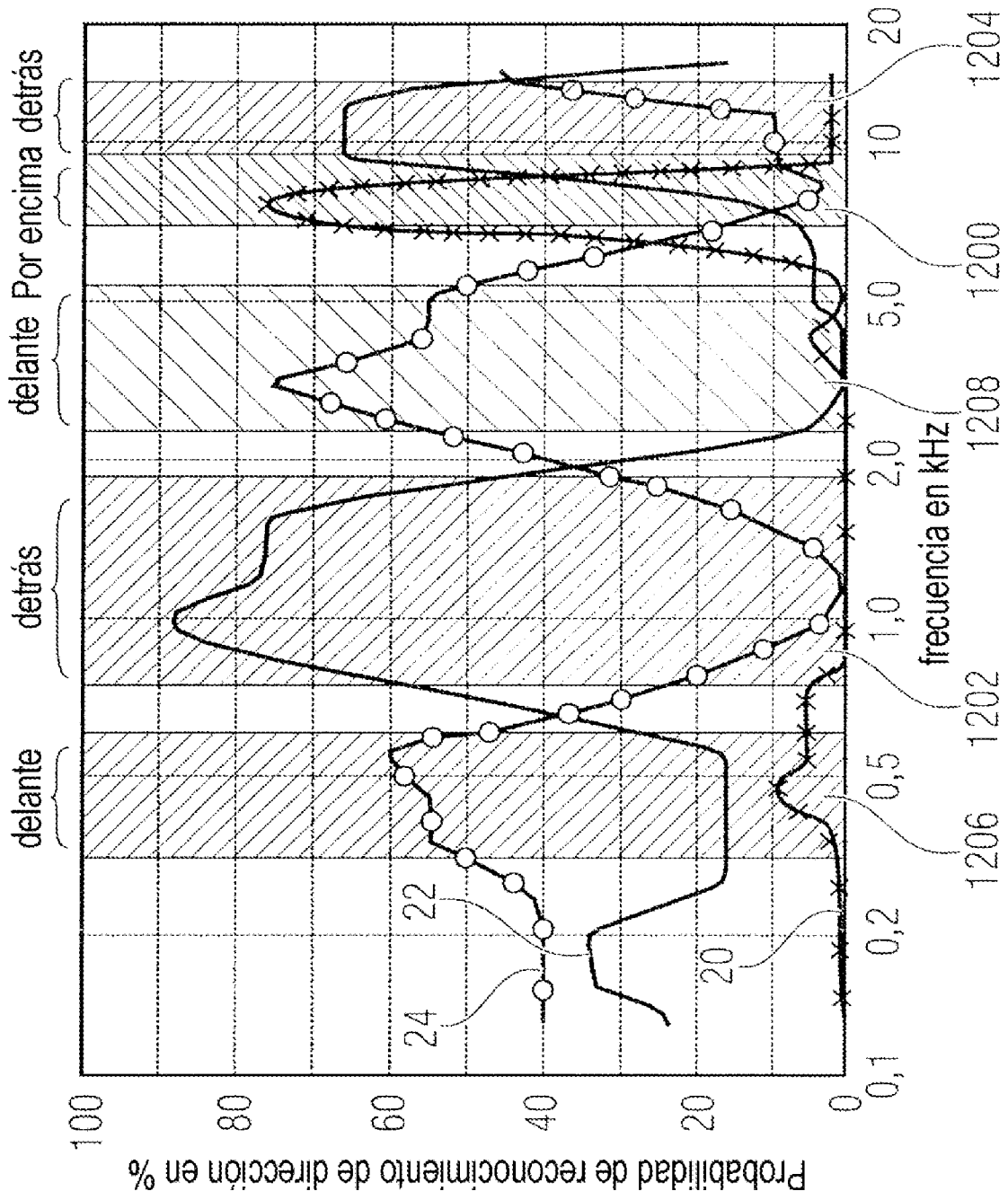


FIG 12