

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 524 391**

51 Int. Cl.:

**H04S 3/00** (2006.01)

**H04S 7/00** (2006.01)

**G10K 15/12** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **30.07.2009 E 09777567 (0)**

97 Fecha y número de publicación de la concesión europea: **27.08.2014 EP 2304975**

54 Título: **Generación de señal para señales binaurales**

30 Prioridad:

**31.07.2008 US 85286**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**09.12.2014**

73 Titular/es:

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

72 Inventor/es:

**MUNDT, HARALD;  
NEUGEBAUER, BERNHARD;  
HILPERT, JOHANNES;  
SILZLE, ANDREAS y  
PLOGSTIES, JAN**

74 Agente/Representante:

**PONTI SALES, Adelaida**

**ES 2 524 391 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

Generación de señal para señales binaurales

5 **[0001]** La presente invención se refiere a la generación de una contribución relacionada con la reflexión y/o reverberación de sala de una señal binaural, la propia generación de una señal binaural, y la formación de un conjunto decreciente de similitud mutua de funciones de transferencia relativas a la cabeza.

10 **[0002]** El sistema auditivo humano es capaz de determinar la dirección o las direcciones de donde proceden los sonidos percibidos. Para este fin, el sistema auditivo humano evalúa ciertas diferencias entre el sonido recibido en el oído derecho y el sonido recibido en el oído izquierdo. Esta última información comprende, por ejemplo, las llamadas señales inter-aurales que pueden, a su vez, referirse a la diferencia de la señal de sonido entre oídos. Las señales Inter-aurales son el medio más importante para la localización. La diferencia de nivel de presión entre los oídos, es decir, la diferencia de nivel inter-aural (ILD), es la señal más importante para la localización. Cuando el sonido llega desde el plano horizontal con un acimut distinto de cero, tiene un nivel diferente en cada oído. El oído sombreado tiene una imagen de sonido suprimida naturalmente, en comparación con el oído sin sombra. Otra propiedad muy importante al tratar con la localización es la diferencia horaria inter-aural (ITD). El oído sombreado tiene una distancia más larga a la fuente de sonido, y por lo tanto el frente de onda de sonido llega más tarde a este que al oído sin sombra. El significado de la ITD se enfatiza en las frecuencias bajas que no se atenúan mucho cuando llega al oído sombreado en comparación con el oído sin sombra. La ITD es menos importante en las frecuencias más altas debido a que la longitud de onda del sonido es más cercana a la distancia entre oídos. Por lo tanto, en otras palabras, la localización explota el hecho de que el sonido está sujeto a diferentes interacciones con la cabeza, los oídos y los hombros del oyente cuando viaja desde la fuente de sonido hasta el oído izquierdo y el derecho, respectivamente.

25 **[0003]** Los problemas se producen cuando una persona escucha una señal estéreo que está destinada a ser reproducida por una configuración de altavoces a través de auriculares. Es muy probable que el oyente considere el sonido como poco natural, incómodo, y perturbador ya que el oyente siente que la fuente de sonido se encuentra en la cabeza. Este fenómeno se denomina a menudo en la literatura como la localización "en-la-cabeza" ("in the head"). Escuchar el sonido "en-la-cabeza" a largo plazo puede conducir a la fatiga auditiva. Esto ocurre porque la información en la que se basa el sistema auditivo humano, al colocar las fuentes de sonido, es decir, las señales inter-aurales, falta o es ambigua.

35 **[0004]** A fin de representar las señales estéreo, o incluso señales de múltiples canales con más de dos canales para la reproducción de auriculares, se pueden utilizar filtros direccionales con el fin de modelar estas interacciones. Por ejemplo, la generación de una salida de auriculares de una señal multicanal descodificada puede comprender filtrar cada señal después de la decodificación por medio mediante un par de filtros direccionales. Estos filtros suelen modelar la transmisión acústica a partir de una fuente de sonido virtual en una sala para el canal del oído de un oyente, que es la así llamada función de transferencia de habitación binaural (BRTF). La BRTF realiza modificaciones de tiempo, de nivel y espectrales y modeliza las reflexiones y la reverberación de la sala. Los filtros direccionales pueden implementarse en el dominio del tiempo o de la frecuencia.

45 **[0005]** Sin embargo, ya que son necesarios muchos filtros, a saber,  $N \times 2$  siendo  $N$  el número de canales descodificados, estos filtros direccionales son bastante largos, tales como 20.000 tomas de filtro a 44,1 kHz, y el proceso de filtrado es computacionalmente exigente. Por lo tanto, a veces se reducen los filtros direccionales a un mínimo. Las así llamadas funciones de transferencia relacionadas con la cabeza (HRTF) contienen la información direccional incluyendo las señales interaurales. Se utiliza un bloque de procesamiento común para modelizar las reflexiones y la reverberación de la sala. El módulo de procesamiento de sala puede ser un algoritmo de reverberación en el dominio del tiempo o de la frecuencia, y puede funcionar en una señal de entrada de uno o dos canales obtenida a partir de la señal de entrada multicanal mediante una suma de los canales de la señal de entrada multicanal. Esta estructura se describe, por ejemplo, en el documento WO 99/14983 A1. Tal como se acaba de describir, el bloque de procesamiento de sala implementa reflexiones y / o la reverberación de la sala. Las reflexiones y la reverberación de la sala son esenciales para sonidos localizados, especialmente en relación con la distancia y la externalización – lo que significa que los sonidos se perciben fuera de la cabeza del oyente. El citado documento también sugiere la aplicación de los filtros direccionales como un conjunto de filtros FIR que operan en versiones retardadas de manera diferente del canal respectivo, a fin de modelar el camino directo desde la fuente de sonido al oído respectivo y las reflexiones distintas. Por otra parte, en la descripción de una serie de medidas para proporcionar una experiencia de sonido más agradable con un par de auriculares, este documento también sugiere retrasar una mezcla del canal central y el canal delantero izquierdo y el canal central y el canal delantero derecho, respectivamente, en relación con una suma y una diferencia de los canales izquierdo y derecho traseros, respectivamente.

60 **[0006]** Sin embargo, los resultados de escucha alcanzados hasta ahora todavía carecen en gran medida de una anchura espacial reducida de la señal de salida binaural y de la falta de externalización. Además, se ha comprobado que a pesar de las medidas antes mencionadas para la representación de señales multicanal para la reproducción

por auriculares, partes de voz en los diálogos y la música de películas se perciben a menudo como reverberantes de manera poco natural y espectralmente desiguales.

5 **[0007]** Por lo tanto, es el objeto de la presente invención proporcionar un esquema para la generación de la señal binaural, que produzca una reproducción por auriculares más estable y agradable.

**[0008]** Este objeto se consigue mediante el contenido de las reivindicaciones independientes.

10 **[0009]** Se puede conseguir una señal binaural más estable y agradable para la reproducción con auriculares mediante el procesamiento diferente, y reduciendo así la similitud entre, al menos uno de entre un canal izquierdo y un canal derecho de la pluralidad de canales de entrada, un canal frontal y posterior de la pluralidad de canales de entrada, y un canal central y no central de la pluralidad de canales, obteniendo de este modo un conjunto de canales de similitud mutua reducida. Este conjunto de canales de similitud mutua reducida se suministra entonces a una pluralidad de filtros direccionales seguidos por respectivos mezcladores para los oídos izquierdo y derecho, respectivamente. Al reducir la inter-similitud de canales de la señal de entrada multi-canal, se puede aumentar la anchura espacial de la señal de salida binaural y se puede aumentar la externalización.

15 **[0010]** Además, se puede lograr una señal binaural más estable y agradable para reproducción con auriculares realizando - en un sentido variable espectralmente - una modificación de manera diferente de fase y/o magnitud entre al menos dos canales de la pluralidad de canales, obteniendo de este modo el conjunto de canales de similitud mutua reducida el cual, a su vez, puede ser entonces suministrado a una pluralidad de filtros direccionales seguidos de mezcladores respectivos para los oídos izquierdo y derecho, respectivamente. De nuevo, mediante la reducción de la similitud mutua de canales de la señal de entrada multi-canal, se puede aumentar la anchura espacial de la señal de salida binaural y se puede aumentar la externalización.

20 **[0011]** Las ventajas antes mencionadas también son alcanzables cuando se forma un conjunto decreciente de similitud mutua de funciones de transferencia relativas a la cabeza provocando que las respuestas de impulso de una pluralidad original de funciones de transferencia relacionadas con la cabeza se retrasen entre sí, o - en un sentido variable espectralmente - respuestas de fase y/o magnitud de la pluralidad original de funciones de transferencia relacionadas con la cabeza de manera diferente en relación unas con otras. La formación se puede realizar fuera de línea como etapa de diseño, o en línea durante la generación de la señal binaural, mediante el uso de las funciones de transferencia relacionadas con la cabeza como filtros direccionales tales como, por ejemplo, en respuesta a una indicación de las ubicaciones de fuente de sonido virtual a utilizar.

25 **[0012]** Algunas partes de las películas y la música dan como resultado una reproducción por auriculares percibida de forma más natural, cuando la mezcla descendente mono o estéreo de los canales de la señal multicanal que tiene que someterse al procesador de sala para la generación de la contribución relacionada con las reflexiones / reverberación de sala de la señal binaural, se forma de tal manera que la pluralidad de canales contribuyen a la mezcla descendente mono o estéreo en un nivel diferente entre los al menos dos canales de la señal multicanal. Por ejemplo, los inventores se dieron cuenta de que las voces en los diálogos y la música de las películas se mezclan típicamente principalmente en el canal central de una señal multicanal, y que la señal de canal central, cuando se alimenta al módulo de procesamiento de sala, da como resultado una salida reverberante de manera no natural y percibida de manera espectralmente desigual. Los inventores descubrieron, sin embargo, que estas deficiencias pueden ser superadas mediante el suministro del canal central al módulo de procesamiento de sala con una reducción del nivel de, por ejemplo, de una atenuación de 3-12 dB, o específicamente, 6 dB.

**[0013]** En lo que sigue, se describen realizaciones preferidas en más detalle con respecto a las figuras, entre las cuales:

50 La figura 1 muestra un diagrama de bloques de un dispositivo para generar una señal binaural acuerdo con una realización de comparación;  
La figura 2 muestra un diagrama de bloque de un dispositivo para formar un conjunto decreciente de similitud mutua de funciones de transferencia relativas a la cabeza según una realización;  
La figura 3 muestra un dispositivo para generar una contribución relacionada con la reflexión y/o reverberación de sala de una señal binaural según un ejemplo;  
55 Las figuras 4a y 4b muestra diagramas de bloque del procesador de sala de la figura 3 según diferentes ejemplos;  
La figura 5 muestra un diagrama de bloque del generador de mezcla descendente de la figura 3 según un ejemplo;  
La figura 6 muestra un diagrama esquemático que ilustra una representación de una señal multi-canal que utiliza codificación de audio espacial según un ejemplo;  
60 La figura 7 muestra un generador de señal de salida binaural según un ejemplo;  
La figura 8 muestra un diagrama de bloque de un generador de señal de salida binaural según un ejemplo adicional;  
La figura 9 muestra un diagrama de bloque de un generador de señal de salida binaural según otro ejemplo adicional;  
La figura 10 muestra un diagrama de bloque de un generador de señal de salida binaural según un ejemplo adicional;

La figura 11 muestra un diagrama de bloque de un generador de señal de salida binaural según un ejemplo adicional;

La figura 12 muestra un diagrama de bloque del descodificador de audio espacial binaural de la figura 11 según un ejemplo; y

5 La figura 13 muestra un diagrama de bloque del descodificador de audio espacial modificado de la figura 11 según un ejemplo.

10 **[0014]** La figura 1 muestra un dispositivo para generar una señal binaural deseada, por ejemplo, para reproducción con auriculares a partir de una señal multi-canal que representa una pluralidad de canales y prevista para reproducción por una configuración de altavoz que tiene una posición de fuente sonora virtual asociada a cada canal. El dispositivo que se indica generalmente mediante el signo de referencia 10, comprende un reductor de similitud 12, una pluralidad 14 de filtros direccionales 14a-14h, un primer mezclador 16a y un segundo mezclador 16b.

15 **[0015]** El reductor de similitud 12 está configurado para convertir la señal multi-canal 18 que representa la pluralidad de canales 18a-18d, en un conjunto reducido de similitud mutua 20 de canales 20a-20d. El número de canales 18a-18d representado por la señal multi-canal 18 puede ser dos o más. Solamente con fines ilustrativos, se muestran cuatro canales 18a-18d en la figura 1. La pluralidad 18 de canales puede, por ejemplo, comprender un canal central, un canal frontal izquierdo, un canal frontal derecho, un canal posterior izquierdo, y un canal posterior derecho. Los canales 18a-18d, por ejemplo, se han mezclado hacia arriba por un diseñador de sonido a partir de una pluralidad de señales de audio individuales que representan, por ejemplo, instrumentos, vocales individuales, u otras fuentes de sonido individuales, suponiendo que o con la intención de que los canales 18a-18d se reproduzcan por una disposición de altavoces (no mostrada en la figura 1), que tiene los altavoces posicionados en posiciones predefinidas de fuentes sonoras virtuales asociadas a cada canal 18a-18d.

20 **[0016]** Según la figura 1, la pluralidad de canales 18a-18d comprende, al menos, un par de un canal izquierdo y un canal derecho, un par de un canal frontal y posterior, o un par de un canal central y no central. Obviamente, puede haber más de uno de los pares que se acaban de mencionar en la pluralidad 18 de canales 18a-18d. El reductor de similitud 12 está configurado para procesar de manera diferente, y de este modo reducir una similitud entre canales de la pluralidad de canales, con la finalidad de obtener el conjunto reducido de similitud mutua 20 de canales 20a-20d. Según un primer aspecto, la similitud entre al menos uno de, un canal izquierdo y un canal derecho de la pluralidad 18 de canales, un canal frontal y posterior de una pluralidad 18 de canales, y un canal central y no central de la pluralidad 18 de canales puede reducirse mediante el reductor de similitud 12, con la finalidad de obtener el conjunto reducido de similitud mutua 20 de canales 20a-20d. Según un segundo aspecto, el reductor de similitud (12) puede - de manera adicional o de manera alternativa - realizar - en un sentido variable espectralmente - una modificación de manera diferente de fase y/o magnitud entre al menos dos canales de la pluralidad de canales, con la finalidad de obtener el conjunto reducido de similitud mutua 20 de canales.

30 **[0017]** Como se detallará con mayor detalle más abajo, el reductor de similitud 12 puede, por ejemplo, lograr el procesamiento diferente haciendo que los respectivos pares se retrasen entre sí, o sometiendo los pares de canales respectivos a retardos de diferentes cantidades en, por ejemplo, cada uno de una pluralidad de bandas de frecuencia, obteniendo de este modo un conjunto reducido de inter-correlación 20 de canales. Hay, por supuesto, otras posibilidades para disminuir la correlación entre los canales. Dicho de otro modo, el reductor de correlación 12 puede tener una función de transferencia según la cual la distribución de energía espectral de cada canal sigue siendo la misma, es decir la función de transferencia una magnitud de una sobre el rango de espectro de audio correspondiente en el que, sin embargo, el reductor de similitud 12 de manera diferente modifica fases de sub-bandas o componentes de frecuencia de estas. Por ejemplo, el reductor de correlación 12 se podría configurar de modo que este provoque una modificación de fase en uno de, o uno o varios de, los canales 18 de modo que una señal de un primer canal para una determinada banda de frecuencia se retrase con respecto a otro de los canales en al menos una muestra. Además, el reductor de correlación 12 se podría configurar de modo que este provoque la modificación de fase de modo que los retardos de grupo de un primer canal con respecto a otro de los canales para una pluralidad de bandas de frecuencia, muestre una desviación estándar de al menos una octava parte de una muestra. Las bandas de frecuencia consideradas podrían ser las bandas de Bark o un subconjunto de estas o cualquier otra sub-división de banda de frecuencia.

40 **[0018]** La reducción de la correlación no es la única manera de evitar la localización en la cabeza. Más bien, la correlación es una de varias medidas posibles mediante el uso de la cual el sistema auditivo humano mide la similitud del sonido que llega a ambos oídos, y por lo tanto, la dirección de origen del sonido. En consecuencia, el reductor de similitud 12 también puede lograr el procesamiento diferente sometiendo los respectivos pares de canales a reducción de nivel en diferentes cantidades en, por ejemplo, cada una de una pluralidad de bandas de frecuencias, obteniendo de este modo un conjunto reducido de similitud mutua 20 de canales de una manera formada espectralmente. La formación espectral puede, por ejemplo, exagerar la reducción espectralmente formada relativa que ocurre, por ejemplo, para el sonido del canal posterior con respecto al sonido del canal frontal debido a la sombra provocada por el lóbulo de la oreja. Por consiguiente, el reductor de similitud 12 puede someter el o los canales traseros a reducciones de nivel variables espectralmente con respecto a los otros canales. En esta formación espectral, el reductor de similitud 12 puede tener una respuesta de fase que sea constante sobre el rango

de espectro de audio relevante en el que, sin embargo, el reductor de similitud 12 de manera diferente modifica magnitudes de sub-bandas o componentes de frecuencia de este.

5 **[0019]** La manera en la que la señal multicanal 18 representa una pluralidad de canales 18a-18d, en principio, no se limita a ninguna representación específica. Por ejemplo, la señal multi-canal 18 podría representar la pluralidad de canales 18a-18d de una manera comprimida, empleando codificación de audio espacial. Según la codificación de audio espacial, la pluralidad de canales 18a-18d podría ser representada por mediante una señal de mezcla descendente en la cual los canales están mezclados hacia abajo, acompañada de información de mezcla descendente que informa de la relación de mezclado según la cual los canales individuales 18a-18d se han  
10 mezclado en uno o más canales de mezcla descendente, y parámetros espaciales que describen la imagen espacial de la señal multicanal mediante, por ejemplo, diferencias de nivel / intensidad, diferencias de fase, diferencias de tiempo y/o medidas de correlación / coherencia entre canales individuales 18a-18d. La salida del reductor de correlación 12 se divide en los canales individuales 20a-20d. Los últimos canales pueden, por ejemplo, ser emitidos como señales de tiempo o como espectrogramas tal como, por ejemplo, espectralmente descompuestos en sub-  
15 bandas.

**[0020]** Los filtros direccionales 14a-14h están configurados para modelizar una transmisión acústica de uno respectivo de los canales 20a-20d a partir de una posición de fuente de sonido virtual asociada con el canal correspondiente a un canal del oído respectivo del oyente. En la figura 1, unos filtros direccionales 14a- 14d modelan la transmisión acústica para, por ejemplo, el canal del oído izquierdo, mientras los filtros direccionales 14e-14h modelan la transmisión acústica para el canal del oído derecho. Los filtros direccionales pueden modelizar la transmisión acústica a partir de una posición de fuente de sonido virtual en una habitación a un canal del oído del oyente y pueden realizar este modelado mediante la realización de modificaciones temporales, de nivel y espectrales, y opcionalmente, modelizando reflexiones de la sala y reverberaciones. Los filtros direccionales 18a-18h pueden ser implementados en el dominio del tiempo o de la frecuencia. Es decir, los filtros direccionales pueden ser filtros en el dominio del tiempo tales como filtros FIR, o pueden operar en el dominio de la frecuencia multiplicando los valores de muestra de función de transferencia respectivos con los respectivos valores espectrales de los canales 20a-20d. En particular, los filtros direccionales 14a-14h se pueden seleccionar para modelizar la respectiva función de transferencia relacionada con la cabeza que describe la interacción de la señal de canal respectiva 20a-20d a partir de la posición respectiva de la fuente de sonido virtual en el canal auditivo respectivo, incluyendo, por ejemplo, las interacciones con la cabeza, las orejas y los hombros de un humano. El primer mezclador 16a está configurado para mezclar las salidas de los filtros direccionales 14a-14d modelizando la transmisión acústica al canal auditivo izquierdo del oyente para obtener una señal 22a destinada a contribuir a, o incluso ser el canal izquierdo de la señal de salida binaural, mientras que el segundo mezclador 16b está  
20 configurado para mezclar las salidas de los filtros direccionales 14e-14h modelizando la transmisión acústica al canal del oído derecho del oyente para obtener una señal 22b, y destinada a contribuir a o incluso ser el canal derecho de la señal de salida binaural.  
25

**[0021]** Tal como se describirá con más detalle más adelante, se pueden añadir más contribuciones a las señales 22a y 22b, con la finalidad de tener en cuenta las reflexiones y/o la reverberación de la sala. Mediante esta medida, se puede reducir la complejidad de los filtros direccionales 14a-14h.  
30

**[0022]** En el dispositivo de la figura 1, el reductor de similitud 12 contrarresta los efectos laterales negativos de la suma de las señales correlacionadas entradas en los mezcladores 16a y 16b, respectivamente, según la cual resulta una anchura espacial muy reducida de la señal de salida binaural 22a y 22b y una falta de externalización. La descorrelación alcanzada con el reductor de similitud 12 reduce estos efectos secundarios negativos.  
35

**[0023]** Antes de volver a la realización de la figura 2, la figura 1 muestra, en otras palabras, un flujo de señal para la generación de una salida de auricular a partir de, por ejemplo, una señal multi- canal descodificada. Cada señal se filtra mediante un par de pares de filtros direccionales. Por ejemplo, el canal 18a se filtra mediante el par de filtros direccionales 14a-14e. Desafortunadamente, existe una cantidad significativa de similitud tal como la correlación entre los canales 18a-18d en producciones típicas de sonido multicanal. Esto afectaría negativamente a la señal de salida binaural. Es decir, después de procesar las señales multicanal con un filtro direccional 14a-14h, la salida de señales intermedias por los filtros direccionales 14a-14h se añaden en el mezclador 16a y 16b para formar la señal de salida para auriculares 20a y 20b. La suma de las señales de salida similares / correlacionadas daría lugar a una anchura espacial muy reducida de la señal de salida 20a y 20b, y una falta de externalización. Esto es particularmente problemático para la similitud / correlación de las señales izquierda y derecha y el canal central. De acuerdo con ello, el reductor de similitud 12 debe reducir la similitud entre estas señales en la medida de lo posible.  
40

**[0024]** Cabe señalar que la mayoría de las medidas realizadas por el reductor de similitud 12 para reducir la similitud entre los canales de la pluralidad 18 de canales 18a-18d también puede alcanzarse mediante la eliminación del reductor de similitud 12 concurrentemente con la modificación de los filtros direccionales para llevar a cabo no sólo la modelización de la transmisión acústica antes mencionada, sino también lograr la disimilitud tal como la descorrelación que se acaba de mencionar. En consecuencia, los filtros direccionales, por ejemplo, no modelarían los HRTF, sino funciones de transferencia modificadas relacionadas con la cabeza.  
45  
50  
55  
60  
65

5 **[0025]** La figura 2, por ejemplo, muestra un dispositivo para formar un conjunto decreciente de similitud mutua de funciones de transferencia relativas a la cabeza para modelizar una transmisión acústica de un conjunto de canales a partir de una posición de fuente sonora virtual asociada con el canal respectivo en los canales de oído de un oyente. El dispositivo que se indica generalmente por 30 comprende un proveedor de HRTF 32, así como un procesador HRTF 34.

10 **[0026]** El proveedor de HRTF 32 está configurado para proporcionar una pluralidad original de HRTFs. La etapa 32 puede comprender medidas empleando una cabeza de muñeco estándar, con la finalidad de medir las funciones de transferencia relacionadas con la cabeza desde determinadas posiciones de sonido a los canales de oído de un muñeco oyente estándar. De manera similar, el proveedor de HRTF 32 puede ser configurado para simplemente consultar o descargar las HRTFs originales de una memoria. Incluso de manera alternativa, el proveedor de HRTF 32 puede ser configurado para calcular los HRTFs según una fórmula predeterminada, que depende de, por ejemplo, posiciones de fuentes sonoras virtuales de interés. Por consiguiente, el proveedor de HRTF 32 puede ser configurado para funcionar en un entorno de diseño para diseñar un generador de señal de salida binaural, o puede ser parte de este mismo generador de señal de salida binaural, con la finalidad de proporcionar las HRTFs originales online tal como, por ejemplo, en respuesta a una selección o cambio de las posiciones de fuentes sonoras virtuales. Por ejemplo, el dispositivo 30 puede ser parte de un generador de señal de salida binaural que es capaz de alojar señales multi-canal previstas para diferentes configuraciones de altavoz que tienen diferentes posiciones de fuentes sonoras virtuales asociadas con sus canales. En este caso, el proveedor de HRTF 32 puede ser configurado para proporcionar las HRTFs originales de manera adaptada a las posiciones previstas actualmente de fuentes sonoras virtuales.

20 **[0027]** El procesador HRTF 34, a su vez, está configurado para provocar que las respuestas de impulso de al menos un par de los HRTFs a desplazar en relación unas con otras o modificar en un sentido variable espectralmente - sus respuestas de fase y/o amplitud de manera diferente en relación unas con otras. El par de HRTFs puede modelizar la transmisión acústica de uno de los canales izquierdo y derecho, canales frontal y posterior, y canales central y no central. En efecto, esto se puede lograr mediante una o una combinación de las siguientes técnicas aplicadas a uno o varios canales de la señal multicanal, es decir, el retraso de la HRTF de un canal respectivo, la modificación de la respuesta de fase de una HRTF respectiva y / o la aplicación de un filtro de descorrelación tal como un filtro de pasatodo a la respectiva HRTF, obteniéndose de este modo un conjunto reducido de inter-correlación de HRTFs, y / o modificar - en un sentido de modificación espectral - la respuesta de magnitud de una HRTF respectiva, obteniendo de este modo, al menos, un conjunto reducido de similitud mutua de HRTFs. En cualquier caso, la descorrelación / disimilitud resultante entre los respectivos canales puede ayudar al sistema auditivo humano a localizar externamente la fuente de sonido y de este modo impedir que se produzca la localización en la cabeza. Por ejemplo, el procesador HRTF 34 se podría configurar de modo que este provoque una modificación de la respuesta de fase de todos, o de uno o varios de, los canales HRTFs de modo que se introduzca un retraso de grupo de una primera HRTF para una determinada banda de frecuencia - o se retrase una determinada banda de frecuencia de una primera HRTF - con respecto a otra de las HRTFs en al menos una muestra. Además, el procesador HRTF 34 se podría configurar de modo que este provoque la modificación de la respuesta de fase de modo que los retrasos de grupo de una primera HRTF con respecto a otra de los HRTFs para una pluralidad de bandas de frecuencia, muestra una desviación estándar de al menos un octavo de una muestra. Las bandas de frecuencia consideradas podrían ser las bandas de Bark o un subconjunto de estas o cualquier otra sub-división de banda de frecuencia.

35 **[0028]** El conjunto de HRTFs decreciente de similitud mutua resultante del procesador HRTF 34 puede utilizarse para ajustar las HRTFs de los filtros direccionales 14a-14h del dispositivo de la figura 1, donde el reductor de similitud 12 puede estar presente o ausente. Debido a la propiedad de disimilitud de las HRTF modificadas, las ventajas antes mencionadas con respecto a la anchura espacial de la señal de salida binaural y la mejora de la externalización se consiguen de manera similar incluso cuando no hay reductor de similitud 12.

45 **[0029]** Como ya se ha descrito arriba, el dispositivo de la figura 1 puede estar acompañado por otra pasada configurada para obtener contribuciones de la señal de salida binaural relacionadas con la reflexión y/o reverberación de sala a partir de una mezcla descendente de al menos algunos de los canales de entrada 18a-18d. Esto alivia la complejidad que plantean los filtros de dirección 14a-14h. Un dispositivo para generar esta contribución relacionada con la reflexión y/o reverberación de sala de una señal de salida binaural se muestra en la figura 3. El dispositivo 40 comprende un generador de mezcla descendente 42 y un procesador de sala 44 conectados en serie entre sí con el procesador de sala 44 tras el generador de mezcla descendente 42. El dispositivo 40 puede conectarse entre la entrada del dispositivo de la figura 1 en el que la señal multi-canal 18 se entra, y la salida de la señal de salida binaural donde la contribución de canal izquierdo 46a del procesador de sala 44 se añade a la salida 22a, y la salida de canal derecho 46b del procesador de sala 44 se añade a la salida 22b. El generador de mezcla descendente 42 forma una mezcla descendente mono o estéreo 48 a partir de los canales de la señal multi-canal 18, y el procesador 44 está configurado para generar el canal izquierdo 46a y el canal derecho 46b de las contribuciones relacionadas con la reflexión y/o la reverberación de sala de la señal binaural modelizando la reflexión y/o reverberación de sala a partir de la señal mono o estéreo 48.

60 **[0030]** La idea que subyace al procesador de sala 44 es que la reflexión / reverberación de sala que se produce, por ejemplo, en una sala, puede ser modelizada de una manera transparente para el oyente, a partir de una mezcla

descendente como una simple suma de los canales de la señal multi-canal 18. Puesto que las reflexiones/reverberaciones de sala ocurren más tarde que los sonidos que viajan a lo largo del camino directo o línea de visión desde la fuente de sonido a los canales auditivos, la respuesta de impulso del procesador de sala es representativa de, y los sustituye a, la cola de las respuestas de impulso de los filtros direccionales que se muestran en la figura 1. Las respuestas de impulso de los filtros direccionales pueden, a su vez, limitarse a modelar la trayectoria directa y las reflexiones y atenuaciones que ocurren en la cabeza, las orejas y los hombros del oyente, lo que permite acortar las respuestas de impulso de los filtros direccionales. Por supuesto, la frontera entre lo que se modeliza mediante el filtro direccional y lo que se modeliza mediante el procesador de sala 44 se puede variar libremente para que el filtro direccional pueda, por ejemplo, también modelizar las primeras reflexiones/reverberaciones de la sala.

**[0031]** Las figuras 4a y 4b muestran posibles implementaciones de la estructura interna del procesador de sala. Según la figura 1a, al procesador de sala 44 se le suministra una señal mono de mezcla descendente 48 y comprende dos filtros de reverberación 50a y 50b. De manera análoga a los filtros direccionales, los filtros de reverberación 50a y 50b pueden ser implementados para funcionar en el dominio del tiempo o en el dominio de la frecuencia. Las entradas de ambos reciben la señal mono de mezcla descendente 48. La salida del filtro de reverberación 50a proporciona la salida de contribución de canal izquierdo 46a, mientras que el filtro de reverberación 50b da como salida la señal de contribución de canal derecho 46b. La figura 4b muestra un ejemplo de la estructura interna del procesador de sala 44, en el caso del procesador de sala 44 al que se le suministra una señal estéreo de mezcla descendente 48. En este caso, el procesador de sala comprende cuatro filtros de reverberación 50a-50d. Las entradas de los filtros de reverberación 50a y 50b están conectadas a un primer canal 48a de la mezcla descendente estéreo 48, mientras que la entrada de los filtros de reverberación 50c y 50d está conectada al otro canal 48b de la mezcla descendente estéreo 48. Las salidas de los filtros de reverberación 50a y 50c están conectadas a la entrada de un sumador 52a, cuya salida proporciona la contribución de canal izquierdo 46a. La salida de los filtros de reverberación 50b y 50d está conectada a las entradas de otro sumador 52b, cuya salida proporciona la contribución de canal derecho 46b.

**[0032]** Aunque se ha descrito que el generador de mezcla descendente 42 puede simplemente sumar los canales de la señal multi-canal 18 - ponderando cada canal de la misma manera-, este no es exactamente el caso de la realización de la figura 3. Más bien, el generador de mezcla descendente 42 de la figura 3 está configurado para formar la mezcla descendente mono o estéreo 48, de modo que la pluralidad de canales contribuye a la mezcla descendente mono o estéreo en un nivel que difiere al menos entre dos canales de la señal multicanal 18. Mediante esta medida, se puede impedir o bien favorecer a determinados contenidos de señales multicanal, como el habla o la música de fondo que se mezclan en un canal específico o canales específicos o la señal multicanal, que se sometan a un procesado de sala, evitando así un sonido poco natural.

**[0033]** Por ejemplo, el generador de mezcla descendente 42 de la figura 3 puede ser configurado para formar la mezcla descendente mono o estéreo 48 de modo que un canal central de la pluralidad de canales de la señal multi-canal 18 contribuya a la señal mono o estéreo de mezcla descendente 48 de manera reducida con respecto a otros canales de la señal multi-canal 18. Por ejemplo, la cantidad de reducción de nivel puede ser de entre 3 dB y 12 dB. La reducción de nivel puede ser repartida de manera uniforme en el rango espectral efectivo de los canales de la señal multi-canal 18, o puede ser dependiente de la frecuencia como por ejemplo concentrado en una parte espectral específica, tal como la parte espectral normalmente ocupada por las señales de voz. La cantidad de reducción de nivel con respecto a los otros canales puede ser la misma para todos los demás canales. Es decir, los otros canales se pueden mezclar en la señal de mezcla descendente 48 al mismo nivel. De manera alternativa, los otros canales se pueden mezclar en la señal de mezcla descendente 48 con un nivel distinto. Por lo tanto, la cantidad de reducción de nivel con respecto a los otros canales se puede medir con respecto al valor medio de los otros canales o el valor medio de todos los canales incluyendo el reducido. Si es así, la desviación estándar de las ponderaciones de mezclado de los otros canales o la desviación estándar de las ponderaciones de mezclado de todos los canales puede ser menor que 66% de la reducción de nivel de la ponderación de mezclado del canal de nivel reducido con respecto al valor medio que se acaba de mencionar.

**[0034]** El efecto de la reducción de nivel con respecto al canal central es que la señal de salida binaural obtenida a través de las contribuciones 56a y 56b se percibe - al menos en algunas circunstancias que se discuten con más detalle a continuación - con más naturalidad por los oyentes que sin la reducción del nivel. En otras palabras, el generador de mezcla descendente 42 forma una suma ponderada de los canales de los canales de la señal multi-canal 18, con el valor de ponderación asociado con el canal central reducido con respecto a los valores de ponderación de los otros canales.

**[0035]** La reducción del nivel del canal central es especialmente ventajosa durante partes con voz de diálogos o música de películas. La mejora de impresión de audio obtenida durante estas partes de voz sobre compensa las sanciones menores debidas a la reducción en el nivel de fases que no son de voz. Sin embargo, de acuerdo con una alternativa de realización, la reducción del nivel no es constante. Más bien, el generador de mezcla descendente 42 puede ser configurado para conmutar entre un modo en el que la reducción de nivel se desactiva, y un modo en que la reducción de nivel se activa. En otras palabras, el generador de mezcla descendente 42 puede ser configurado para variar la cantidad de reducción de nivel de una manera variable en el tiempo. La variación puede ser de naturaleza binaria o analógica, entre cero y un valor máximo. El generador de mezcla descendente 42 puede ser

configurado para realizar el modo que activa o reduce el nivel de cantidad de variación en función de la información contenida en la señal multi-canal 18. Por ejemplo, el generador de mezcla descendente 42 puede ser configurado para detectar las fases de voz o distinguir estas fases de voz de fases que no son de voz, o puede asignar una medida de contenido de voz midiendo el contenido de voz, siendo al menos de escala ordinal, en tramas consecutivas del canal central. Por ejemplo, el generador de mezcla descendente 42 detecta la presencia de la voz en el canal central por medio de un filtro de voz y determina si el nivel de salida de este filtro supera el umbral de suma. Sin embargo, la detección de fases de voz dentro del canal central por el generador de mezcla descendente 42 no es la única manera de hacer el cambio de modo antes mencionado de reducción de variación de cantidad de nivel dependiente del tiempo. Por ejemplo, la señal multi-canal 18 podría tener información lateral asociada con ella, especialmente prevista para distinguir entre las fases de voz y las fases que no son de voz, o medir cuantitativamente el contenido de voz. En este caso, el generador de mezcla descendente 42 funcionaría en respuesta a esta información lateral. Otra probabilidad sería que el generador de mezcla descendente 42 realizara la conmutación de modo o variaciones de reducción de cantidad de nivel de las antes mencionadas dependiendo de una comparación entre, por ejemplo, los niveles actuales del canal central, el canal izquierdo y el canal derecho. En el caso de que el canal central sea mayor que los canales izquierdo y derecho, ya sea individualmente o en relación con la suma de estos, en más de una cierta relación de umbral, entonces el generador de mezcla descendente 42 podría asumir que una fase de voz está actualmente presente y actuaría en consecuencia, es decir, mediante la realización de la reducción del nivel. De manera similar, el generador de mezcla descendente 42 podría utilizar las diferencias de nivel entre los canales central, izquierdo y derecho con la finalidad de realizar las dependencias antes mencionadas.

**[0036]** Además de esto, el generador de mezcla descendente 42 puede ser sensible a parámetros espaciales utilizados para describir la imagen espacial de los múltiples canales de la señal multi-canal 18. Esto se muestra en la figura 5. La figura 5 muestra un ejemplo del generador de mezcla descendente 42 en el caso en que la señal multicanal 18 representa una pluralidad de canales empleando una codificación de audio especial, es decir empleando una señal de mezcla descendente 62 en la que la pluralidad de canales se han mezclado hacia abajo y parámetros espaciales 64 que describen la imagen espacial de la pluralidad de canales. De manera opcional, la señal multi-canal 18 también puede comprender información de mezcla descendente que describe las relaciones con las que se han mezclado los canales individuales en la señal de mezcla descendente 62, o los canales individuales de la señal de mezcla descendente 62, puesto que el canal de mezcla descendente 62 puede ser por ejemplo una señal de mezcla descendente normal 62 o una señal estéreo de mezcla descendente 62. El generador de mezcla descendente 42 de la figura 5 comprende un descodificador 64 y un mezclador 66. El descodificador 64 descodifica, según una descodificación de audio espacial, la señal multi-canal 18 con la finalidad de obtener la pluralidad de canales incluyendo, entre otros, el canal central 66, y otros canales 68. El mezclador 66 está configurado para mezclar el canal central 66 y los otros canales no centrales 68 para derivar la señal mono o estéreo 48 realizando la reducción de nivel antes mencionada. Tal como se indica mediante la línea a trazos 70, el mezclador 66 puede estar configurado para utilizar el parámetro espacial 64 con la finalidad de conmutar entre el modo de reducción de nivel y el modo de no-reducción de nivel de la cantidad de reducción de nivel variada, tal como se ha mencionado antes. El parámetro espacial 64 utilizado por el mezclador 66 puede, por ejemplo, ser unos coeficientes de predicción de canal que describen como el canal central 66, el canal izquierdo o el canal derecho pueden derivarse de la señal de mezcla descendente 62, en el que mezclador 66 puede de manera adicional utilizar parámetros de coherencia/correlación cruzada entre canales que representan la coherencia o correlación cruzada entre los canales izquierdo y derecho que se acaban de mencionar los cuales, a su vez, pueden ser mezclas descendentes de canales izquierdo frontal y posterior, y canales derecho frontal y posterior, respectivamente. Por ejemplo, el canal central se puede mezclar con una relación fija en los canales izquierdo y derecho antes mencionados de la señal estéreo de mezcla descendente 62. En este caso, dos coeficientes de predicción de canal son suficientes con la finalidad de determinar cómo pueden derivarse los canales central, izquierdo y derecho de una combinación lineal respectiva de los dos canales de la señal estéreo de mezcla descendente 62. Por ejemplo, el mezclador 66 puede utilizar una relación entre una suma y una diferencia de los coeficientes de predicción de canal con la finalidad de diferenciar entre las fases de voz y fases que no son de voz.

**[0037]** Aunque la reducción de nivel con respecto al canal central se ha descrito con el fin de ejemplificar la suma ponderada de la pluralidad de canales de tal manera que contribuyen a la mezcla descendente mono o estéreo en un nivel que difiere entre al menos dos canales de la señal multicanal 18, hay también otros ejemplos en los que otros canales ventajosamente reducen o amplifican su nivel en relación con otro canal u otros canales debido a que cierto contenido de fuente de sonido presente en este o estos canales está o no sujeto al procesamiento de sala en el mismo nivel que otros contenidos en la señal multi-canal, pero a un menor / mayor nivel.

**[0038]** La figura 5 se explicó de manera general, con respecto a una posibilidad para la representación de la pluralidad de canales de entrada por medio de una señal de mezcla descendente 62 y parámetros espaciales 64. Con respecto a la figura 6, esta descripción se intensifica. La descripción con respecto a la figura 6 también se utiliza para la comprensión de los siguientes ejemplos de realización descritos con respecto a las figuras 10 a 13. La figura 6 muestra la señal de mezcla descendente 62 espectralmente descompuesta en una pluralidad de sub-bandas 82. En la figura 6, las sub-bandas 82 se muestran a modo de ejemplo como extendiéndose horizontalmente con las sub-bandas 82 dispuestas con el aumento de la frecuencia de sub-banda de abajo a arriba tal como se indica mediante la flecha 84 de dominio de la frecuencia. La extensión a lo largo de la dirección horizontal indica el eje del tiempo 86.



Por ejemplo, la señal de mezcla descendente 62 comprende una secuencia de valores espectrales 88 por sub-banda 82. La resolución de tiempo con la que se muestrean las sub-bandas 82 por los valores de muestra 88 puede definirse mediante ranuras de banco de filtros 90. Por lo tanto, las ranuras de tiempo 90 y las sub-bandas 82 definen alguna resolución o rejilla de tiempo / frecuencia. Una rejilla de tiempo / frecuencia más basta se define mediante la unión de valores de muestra vecinos 88 o baldosas de tiempo / frecuencia 92 tal como se indica mediante las líneas de trazos de la figura 6, definiendo estas baldosas la resolución o rejilla de tiempo / frecuencia. Los parámetros espaciales mencionados 62 se definen en esta resolución de parámetro de tiempo / frecuencia 92. La resolución de parámetro de tiempo / frecuencia 92 puede cambiar en el tiempo. Con este fin, la señal de múltiples canales 62 puede ser dividida hasta en 94 tramas consecutivas. Para cada trama, la rejilla de resolución de tiempo / frecuencia 92 se puede ajustar de forma individual. En caso de que el decodificador 64 reciba la señal de mezcla descendente 62 en el dominio del tiempo, el decodificador 64 puede comprender un banco de filtros de análisis interno a fin de obtener la representación de la señal de mezcla descendente 62, tal como se muestra en la figura 6. Como alternativa, la señal de mezcla descendente 62 entra en el decodificador 64 en la forma mostrada en la figura 6, en cuyo caso no es necesario el banco de filtros de análisis en el decodificador 64. Como ya se ha mencionado en la figura 5, para cada baldosa 92 pueden estar presentes dos coeficientes de predicción de canal que informan de cómo, con respecto a la baldosa de tiempo / frecuencia respectiva 92, los canales derecho e izquierdo se pueden derivar de los canales izquierdo y derecho de la señal de mezcla descendente estéreo 62. Además, un parámetro de coherencia/ correlación cruzada entre canales (ICC) puede estar presente para la baldosa 92 el cual indica las similitudes ICC entre los canales izquierdo y derecho a derivar de la señal estéreo de mezcla descendente 62, en el que un canal se ha mezclado completamente en un canal de la señal estéreo de mezcla descendente 62, mientras que los otros se han mezclado completamente en el otro canal de la señal estéreo de mezcla descendente 62. Sin embargo, un parámetro de diferencia de nivel de canal (CLD) puede estar además presente en cada baldosa 92 el cual indica la diferencia de nivel entre los canales izquierdo y derecho que se acaban de mencionar. Se puede aplicar una cuantificación no uniforme en una escala logarítmica a los parámetros CLD, donde la cuantificación tiene una alta precisión cercana a cero dB y una resolución más gruesa cuando hay una gran diferencia de nivel entre los canales. Además, otros parámetros pueden estar presentes dentro del parámetro espacial 64. Estos parámetros pueden, entre otras cosas, definir la CLD y la ICC relativas a los canales que servían para la formación, mediante la mezcla de los canales izquierdo y derecho que se acaban de mencionar, tal como los canales posterior izquierdo, delantero izquierdo, posterior derecho, y frontal derecho.

**[0039]** Cabe señalar que las realizaciones mencionadas anteriormente pueden combinarse entre sí. Algunas posibilidades de combinación ya se han mencionado anteriormente. Se mencionan otras posibilidades a continuación con respecto a las realizaciones de las figuras 7 a 13. Además, en las realizaciones de las figuras 1 y 5 mencionadas anteriormente se supone que los canales intermedios 20, 66, y 68, respectivamente, están realmente presentes en el dispositivo. Sin embargo, este no es necesariamente el caso. Por ejemplo, el HRTF modificado tal como se deriva por el dispositivo de la figura 2 puede ser utilizado para definir los filtros direccionales de la figura 1 dejando fuera el reductor de similitud 12, y en este caso, el dispositivo de la figura 1 puede funcionar sobre una señal de mezcla descendente tal como la señal de mezcla descendente 62 que se muestra en la figura 5, que representa la pluralidad de canales 18a-18d, combinando adecuadamente los parámetros espaciales y los HRTF modificados en la resolución de parámetro tiempo / frecuencia 92, y en consecuencia mediante la aplicación de coeficientes de combinación lineal obtenidos con el fin de formar señales binaurales 22a y 22b.

**[0040]** Del mismo modo, el generador de mezcla descendente 42 puede estar configurado para combinar adecuadamente los parámetros espaciales 64 y la cantidad de reducción de nivel a alcanzar para el canal central con el fin de derivar la mezcla descendente mono o estéreo 48 prevista para el procesador de sala 44. La figura 7 muestra un generador de señal de salida binaural. Un generador que se indica generalmente mediante el signo de referencia 100 comprende un decodificador multicanal 102, una salida binaural 104, y dos caminos que se extienden entre la salida del decodificador multicanal 102 y la salida binaural 104, respectivamente, es decir un camino directo 106 y un camino de reverberación 108. En el camino directo, los filtros direccionales 110 están conectados a la salida del decodificador multicanal 102. El camino directo también comprende un primer grupo de sumadores 112 y un segundo grupo de sumadores 114. Los sumadores 112 suman la señal de salida de una primera mitad de los filtros direccionales 110 y los segundos sumadores 114 suman la señal de salida de una segunda mitad de los filtros direccionales 110. La suma de como salida de los sumadores primero y segundo 112 y 114 representan antes mencionada contribución de la señal de salida binaural al camino directo 22a y 22b. Se proporcionan los sumadores 116 y 118 con el fin de combinar las señales de contribución 22a y 22b con las señales de contribución binaurales proporcionadas por el camino de reverberación 108 es decir las señales 46a y 46b. En el camino de reverberación 108, un mezclador 120 y un procesador de sala 122 están conectados en serie entre la salida del decodificador multicanal 102 y las entradas respectivas de los sumadores 16 y 118, salidas que definen la señal de salida binaural en la salida 104.

**[0041]** Con el fin de facilitar la comprensión de la siguiente descripción del dispositivo de la figura 7, se han utilizado parcialmente los signos de referencia utilizados en las figuras 1 a 6 con el fin de indicar los elementos de la figura 7, que corresponden a aquellos, o asume la responsabilidad de la funcionalidad, elementos que salen en las figuras 1 a 6. La descripción correspondiente se hará más patente en la siguiente descripción. Sin embargo, se observa que, a fin de facilitar la siguiente descripción, las siguientes realizaciones se han descrito con la suposición de que el reductor de similitud realiza una reducción de correlación. En consecuencia, este último se indica a continuación

como reductor de correlación. Sin embargo, como se hizo evidente de lo anterior, las realizaciones descritas a continuación son fácilmente transferibles a los casos en los que el reductor de similitud realiza una reducción en la similitud distinta a la de correlación. Además, la siguiente descripción se ha elaborado suponiendo que el mezclador para generar la mezcla descendente para el procesado de sala genera una reducción de nivel del canal central aunque, como se ha descrito, una transferencia a formas alternativas de realización sería fácilmente realizable.

**[0042]** El dispositivo de la figura 7 utiliza un flujo de señal para la generación de una salida de auriculares en la salida 104 a partir de una señal multi-canal descodificada 124. El multicanal descodificado 124 se deriva mediante el decodificador multicanal 102 a partir de una corriente de bits de entrada en una corriente de bits de entrada 126, tal como, por ejemplo, mediante descodificación de audio espacial. Después de la decodificación, cada señal o canal de la señal multi-canal descodificada 124 se filtra mediante un par de filtros direccionales 110. Por ejemplo, el primer (superior) canal de la señal multi-canal descodificada 124 se filtra mediante filtros direccionales 20 Dir- Filter(1,L) y DirFilter(1,R), y una segunda (segunda desde arriba) señal o canal se filtra mediante los filtros direccionales DirFilter(2,L) y DirFilter(2,R), y así sucesivamente. Estos filtros 110 pueden modelizar la transmisión acústica a partir de una fuente de sonido virtual en una sala al canal del oído de un oyente, una así llamada función de transferencia de sala binaural (BRTF). Pueden realizar modificaciones de tiempo, de nivel y espectrales, y puede en parte también modelizar reflexión y reverberación de la sala. Los filtros direccionales 110 pueden ser implementados en los dominios de tiempo o de la frecuencia. Puesto que se requieren muchos filtros 110 ( $N \times 2$ , siendo  $N$  el número de canales descodificados), estos filtros direccionales podrían, si deben modelizar completamente la reflexión y la reverberación de sala, ser bastante largos, es decir, 20000 tomas de filtro a 44.1 kHz, en cuyo caso el proceso de filtrado sería computacionalmente exigente. Los filtros direccionales 110 se reducen ventajosamente al mínimo, las así llamadas funciones de transferencia de cabeza (HRTFs) y se utiliza el bloque de procesamiento común 122 para modelizar las reflexiones y reverberaciones de sala. El módulo de procesado de sala 122 puede implementar un algoritmo de reverberación en un dominio del tiempo o de la frecuencia y puede operar desde un una o dos señales de canal de entrada 48, que se calcula a partir de la señal de entrada multicanal decodificada 124 por una matriz de mezcla dentro del mezclador 120. El bloque de procesado de sala implementa reflexiones y/o reverberación de sala. Las reflexiones y la reverberación de sala son esenciales para localizar sonidos, especialmente con respecto a la distancia y la externalización - lo que significa que son sonidos que se perciben fuera de la cabeza del oyente.

**[0043]** Por lo general, se produce el sonido multicanal de tal manera que la energía sonora que domina está contenida en los canales frontales, es decir, frontal izquierdo, frontal derecho, central. Las voces en diálogos de películas y música se mezclan típicamente principalmente en el canal central. Si las señales de canal central se alimentan al módulo de procesamiento de sala 122, la salida resultante se percibe a menudo como poco natural reverberante y espectralmente desigual. Por lo tanto, según la figura 7, el canal central se alimenta al módulo de procesamiento de sala 122 con una reducción de nivel significativa, tal como atenuada en 6 dB, reducción de nivel que se lleva a cabo, como ya se ha indicado anteriormente, dentro del mezclador 120. En tanto, la realización de la figura 7 comprende una configuración según las figuras 3 y 5, donde los signos de referencia 102, 124, 120, y 122 de la figura 7 corresponden a los signos de referencia 18, 64, la combinación de los signos de referencia 66 y 68, el signo de referencia 66 y el signo de referencia 44 de las figuras 3 y 5, respectivamente.

**[0044]** La figura 8 muestra otro generador de señal de salida binaural. El generador se indica en general con el signo de referencia 140. Con el fin de facilitar la descripción de la figura 8, se han utilizado los mismos signos de referencia que en la figura 7. Con el fin de indicar que el mezclador 120 no tiene necesariamente la funcionalidad tal como se indica con respecto a las figuras 3, 5 y 7, a saber, la realización de la reducción del nivel con respecto al canal central, se ha utilizado el signo de referencia 40 con el fin de indicar la disposición de los bloques 102, 120, y 122, respectivamente. En otras palabras, la reducción de nivel dentro del mezclador 122 es opcional en el caso de la figura 8. A diferencia de la figura 7, sin embargo, unos decorrelacionadores están conectados entre cada par de filtros direccionales 110 y la salida del decodificador 102 para el canal asociado de la señal multicanal descodificada 124, respectivamente. Los decorrelacionadores se indican con los signos de referencia 1421, 1422, y así sucesivamente. Los decorrelacionadores 1421-1424 actúan como el reductor de correlación 12 indicado en la figura 1. Aunque se muestra en la figura 8, no es necesario que se proporcione un decorrelacionador 1421-1424 para cada una de los canales de la señal multi-canal descodificada 124. Más bien, un decorrelacionador sería suficiente. Los decorrelacionadores 142 podrían ser simplemente un retraso. Preferentemente, la cantidad de retraso causado por cada uno de los retrasos 1421-1424 será diferente. Otra posibilidad sería que los decorrelacionadores 1421- 1424 fuesen filtros pasa-todo, es decir filtros que tienen una función de transferencia de magnitud constantemente igual a 1 pero, sin embargo, cambiando las fases de las componentes espectrales del canal respectivo. Las modificaciones de fase causadas por los decorrelacionadores 1421-1424 serían preferentemente diferentes para cada uno de los canales. También existen otras posibilidades, por supuesto. Por ejemplo, los decorrelacionadores 1421- 1424 podrían ser implementados como filtros FIR, o similar.

**[0045]** Por lo tanto, según la figura 8, los elementos 1421- 1424, 110, 112, y 114 actúan de acuerdo con el dispositivo 10 de la figura 1.

**[0046]** De manera similar a la figura 8, la figura 9 muestra una variación del generador de señal de salida binaural de la figura 7. Por lo tanto, la figura 9 también se explica más abajo empleando los mismos signos de referencia que los que se utilizan en la figura 7. De manera similar a la figura 8, la reducción de nivel de mezclador 122 es meramente

opcional en el caso de la figura 9, y por lo tanto, se ha utilizado el signo de referencia 40' en la figura 9 antes que '40, como fue el caso en la figura 7. La variación de la figura 9 aborda el problema de que existe una correlación significativa entre todos los canales en producciones de sonido multicanal. Después del procesamiento de las señales multicanal con los filtros direccionales 110, las señales intermedias de dos canales de cada par de filtros se añaden por los sumadores 112 y 114, para formar la señal de salida de auriculares en la salida 104. La suma de las señales de salida correlacionadas por los sumadores 112 y 114 da como resultado una anchura espacial muy reducida de la señal de salida en la salida 104, y una falta de externalización. Esto es particularmente problemático para la correlación de las señales izquierda y derecha y el canal central dentro de la señal multicanal decodificada 124. Según la realización de la figura 9, los filtros direccionales están configurados para tener una salida descorrelacionada en la medida de lo posible. Para este fin, el dispositivo de la figura 9 comprende el dispositivo 30 para formar un conjunto decreciente de inter-correlación de HRTF a utilizar por los filtros direccionales 110 a partir de un conjunto original de HRTF. Como se describió anteriormente, el dispositivo 30 puede usar una, o una combinación de, las siguientes técnicas con respecto a las HRTF del par de filtros direccionales asociados con uno o varios canales de la señal multicanal decodificada 124: retrasar el filtro direccional o el par de filtros direccionales respectivo tal como por ejemplo mediante el desplazamiento de su respuesta de impulso que se podría hacer, por ejemplo, mediante el desplazamiento de las tomas de filtro; la modificación de la respuesta de fase de los filtros direccionales respectivos; y aplicar un filtro de descorrelación, tal como un filtro pasa-todo a los respectivos filtros direccionales del canal respectivo. Este filtro de pasa-todo podría ser implementado como un filtro FIR.

**[0047]** Como se describió anteriormente, el dispositivo 30 podría funcionar en respuesta al cambio en la configuración del altavoz al que está destinada la corriente de bits en la entrada de corriente de bits 126.

**[0048]** Las figuras 7 a 9 se refieren a una señal multi-canal descodificada. Lo que sigue se refiere a la descodificación multi-canal paramétrica para auriculares.

**[0049]** En términos generales, la codificación de audio espacial es una técnica de compresión de múltiples canales que explota la irrelevancia entre canales perceptual en señales de audio multicanal para lograr mayores tasas de compresión. Esto puede ser capturado en términos de señales espaciales o parámetros espaciales, es decir, parámetros que describen la imagen espacial de una señal de audio multicanal. Las señales espaciales suelen incluir diferencias de nivel / intensidad, diferencias de fase y medidas de correlación / coherencia entre canales, y pueden ser representadas de una manera extremadamente compacta. El concepto de codificación de audio espacial ha sido adoptado por MPEG lo que ha dado como resultado el estándar MPEG Surround, es decir, ISO / IEC23003-1. Los parámetros espaciales tales como los empleados en la codificación de audio espacial también se pueden emplear para describir filtros direccionales. Procediendo de este modo, la etapa de descodificación de datos de audio espaciales y la aplicación de filtros direccionales se puede combinar para decodificar de manera eficiente y representar audio de múltiples -canales para la reproducción con auriculares.

**[0050]** La estructura general de un descodificador de audio espacial para la salida de auriculares se da en la figura 10. El descodificador de la figura 10 se indica en general con el signo de referencia 200, y comprende un modificador de sub-banda espacial binaural 202 que comprende una entrada para una señal de mezcla descendente estéreo o mono 204, otra entrada para parámetros espaciales 206, y una salida para la señal de salida binaural 208. La señal de mezcla descendente junto con los parámetros espaciales 206 forman la antes mencionada señal multicanal 18 y representa la pluralidad de sus canales.

**[0051]** Internamente, el modificador de sub-banda 202 comprende un banco de filtros de análisis 208, una unidad de matrizado o combinador lineal 210 y un banco de filtros de síntesis 212 conectados en el orden mencionado entre la entrada de señal de mezcla descendente y la salida del modificador de sub-banda 202. Además, el modificador de sub-banda 202 comprende un convertidor de parámetros 214 que es alimentado por los parámetros espaciales 206 y un conjunto modificado de HRTF tal como se obtiene mediante el dispositivo 30.

**[0052]** En la figura 10, se supone que la señal de mezcla descendente ya se ha descodificado de antemano, incluyendo por ejemplo, la codificación de entropía. El descodificador de audio espacial binaural se alimenta con la descripción paramétrica de los filtros direccionales en la forma del parámetro de HRTF modificado 216 para formar parámetros binaurales 218. Estos parámetros 218 son aplicados por la unidad de matrizado 210 desde una matriz de dos por dos (en caso de una señal de mezcla descendente estéreo) y en forma de una matriz de uno por (en el caso de una señal de mezcla descendente mono 204), en el dominio de frecuencia, a los valores espectrales 88 que salen del banco de filtros de análisis 208 (véase la figura 6). En otras palabras, los parámetros binaurales 218 varían en la resolución del parámetro de tiempo / frecuencia 92 que se muestra en la figura 6 y se aplican a cada valor de la muestra 88. Se puede utilizar la interpolación para suavizar los coeficientes de la matriz y los parámetros binaurales 218, respectivamente, a partir del dominio de parámetro tiempo / frecuencia 92 más aproximado a la resolución de tiempo / frecuencia del banco de filtros de análisis 208. Es decir, en el caso de una mezcla descendente estéreo 204, el matrizado realizado por unidad 210 da como resultado dos valores de muestra por cada par de valores de muestra del canal izquierdo de la señal de mezcla descendente 204 y el valor de muestra correspondiente del canal derecho de la señal de mezcla descendente 204. Los dos valores de muestra resultantes son parte de los canales izquierdo y derecho de la señal de salida binaural 208, respectivamente. En caso de una señal de mezcla descendente mono204, el matrizado por la unidad 210 da como resultado dos valores de muestra por valor de

muestra de la señal de mezcla descendente mono 204, a saber, uno para el canal izquierdo y uno para el canal derecho de la señal de salida binaural 208. Los parámetros binaural es218 definen la operación de matriz que conduce de uno o dos valores de muestra de la señal de mezcla descendente 204 a los respectivos valores de muestra de canal izquierdo y derecho de la señal de salida binaural 208. Los parámetros binaurales 218 ya reflejan los parámetros HRTF modificados. Por lo tanto, eliminan la correlación de los canales de entrada de la señal multi-canal 18 tal como se indicó anteriormente.

**[0053]** Por lo tanto, la salida de la unidad de matrizado 210 es un espectrograma modificado tal como se muestra en la figura 6. El banco de filtros de síntesis 212 reconstruye a partir de esta la señal de salida binaural 208. En otras palabras, el banco de filtros de síntesis 212 convierte la salida de señal de dos canales resultante por la unidad de matrizado 210 en el dominio del tiempo. Esto es, por supuesto, opcional.

**[0054]** En el caso de la figura 10, los efectos de reflexión y reverberación de sala no se abordan por separado. Si alguna vez se hace, estos efectos tienen que ser considerados en la HRTF 216. La figura 11 muestra un generador de señal de salida binaural que combina un descodificador de audio espacial binaural 200' con un procesamiento de reflexión/reverberación de sala separado. La ' del signo de referencia 200' en la figura 11 indica que el descodificador de audio espacial binaural 200' de la figura 11 puede utilizar HRTFs no modificadas, es decir las HRTFs originales tal como se indican en la figura 2. De manera opcional, sin embargo, el descodificador de audio espacial binaural 200' de la figura 11 puede ser el mostrado en la figura 10. En cualquier caso, el generador de señal de salida binaural de la figura 11 que se indica generalmente mediante el signo de referencia 230, comprende además el descodificador espacial binaural 200', un descodificador de audio de mezcla descendente 232, un modificador de sub-banda de audio espacial modificado 234, un procesador de sala 122, y dos sumadores 116 y 118. El descodificador de audio de mezcla descendente 232 está conectado entre una corriente de bits de entrada 126 y un modificador de sub-banda de audio espacial binaural 202 del descodificador de audio espacial binaural 200'. El descodificador de audio de mezcla descendente 232 está configurado para descodificar la entrada de corriente de datos en la entrada 126 para derivar la señal de mezcla descendente 214 y los parámetros espaciales 206. A ambos, el modificador de sub-banda de audio espacial binaural 202, así como el modificador de sub-banda de audio espacial modificado 234 se les proporciona una señal de mezcla descendente 204 además de los parámetros espaciales 206. El modificador de sub-banda de audio espacial modificado 234 calcula a partir de la señal de mezcla descendente 204 – empleando los parámetros espaciales 206 así como los parámetros modificados 236 que reflejan la cantidad de reducción de nivel del canal central antes mencionada - la mezcla descendente mono o estéreo 48 que sirve como entrada para el procesador de sala 122. Las salidas de contribución para el modificador de sub-banda de audio espacial binaural 202 y el procesador de sala 122, respectivamente, se suman al nivel del canal en los sumadores 116 y 118 para dar como resultado la señal de salida binaural en la salida 238.

**[0055]** La figura 12 muestra un diagrama de bloque que ilustra la funcionalidad del descodificador de audio binaural 200' de la figura 11. Hay que destacar que la figura 12 no muestra la estructura interna real del descodificador de audio espacial binaural 200' de la figura 11, sino que ilustra las modificaciones de señal obtenidas mediante el descodificador de audio espacial binaural 200'. Se recuerda que la estructura interna del descodificador de audio espacial binaural 200' satisface generalmente la estructura mostrada en la figura 10, con la excepción de que el dispositivo 30 puede dejarse en el caso de que este funcione con las HRTFs originales. De manera adicional, la figura 12 muestra la funcionalidad del descodificador de audio espacial binaural 200' a modo de ejemplo para el caso de que solamente se utilicen tres canales representados por la señal multi-canal 18 por el descodificador de audio espacial binaural 200' con la finalidad de formar la señal de salida binaural 208. En particular, se utiliza una caja "2 a 3", es decir TTT, para derivar un canal central 242, un canal derecho 244, y un canal izquierdo 246 desde los dos canales de la mezcla descendente estéreo 204. En otras palabras, la figura 12 supone a modo de ejemplo que la mezcla descendente 204 es una mezcla descendente estéreo. Los parámetros espaciales 206 utilizados por la caja TTT 248 comprenden los coeficientes de predicción de canal mencionados más arriba. La reducción de correlación se logra con tres descorrelacinadores, llamados DelayL, DelayR, y DelayC en la figura 12. Corresponden a la descorrelación introducida en el caso de, por ejemplo, las figuras 1 y 7. Sin embargo, se vuelve a recordar que la figura 12 solamente muestra las modificaciones de señal logradas por el descodificador de audio espacial binaural 200', aunque la estructura real corresponde a la mostrada en la figura 10. Por lo tanto, aunque los retrasos que forman el reductor de correlación 12 se muestran como características separadas con respecto a las los HRTFs que forman los filtros direccionales 14, la existencia de los retrasos en el reductor de correlación 12 puede verse como una modificación de los parámetros HRTF que forman las HRTFs originales de los filtros direccionales 14 de la figura 12. Primero, la figura 12 solamente muestra que el descodificador de audio espacial binaural 200' descorrelaciona los canales para la reproducción con auriculares. La descorrelación se logra mediante unos medios simples, a saber, añadiendo un bloque de retraso en el procesado paramétrico para la matriz M y el descodificador de audio espacial binaural 200'. Por lo tanto, el descodificador de audio espacial binaural 200' puede aplicar las siguientes modificaciones a los canales individuales, a saber retrasar el canal central preferentemente al menos una muestra, retrasar el canal central en diferentes intervalos en cada banda de frecuencia, retrasar los canales izquierdo y derecho preferentemente al menos una muestra y/o retrasar canales izquierdo y derecho en diferentes intervalos en cada banda de frecuencia.

**[0056]** La figura 13 muestra un ejemplo de una estructura del modificador de sub-banda de audio espacial modificado de la figura 11. El modificador de sub-banda 234 de la figura 13 comprende una caja dos a tres o caja

TTT 262, unas etapas de ponderación 264a-264e, unos primeros sumadores 266a y 266b, unos segundos sumadores 268a y 268b, una entrada para la mezcla descendente estéreo 204, una entrada para los parámetros espaciales 206, una entrada adicional para una señal residual 270 y una salida para la mezcla descendente 48 prevista para ser procesada por el procesador de sala, y que es, de acuerdo con la figura 13, una señal estéreo.

**[0057]** Mientras que la figura 13 define en un sentido estructural un ejemplo de modificador de sub-banda de audio espacial modificado 234, la caja TTT 262 de la figura 13 solamente reconstruye el canal central, el canal derecho.244, y el canal izquierdo 246 a partir de la mezcla descendente estéreo 204 empleando los parámetros espaciales 206. Se vuelve a recordar en el caso de la figura 12, los canales 242-246 no se calculan realmente. Más bien, el modificador de sub-banda de audio espacial binaural modifica la matriz M de tal manera que la señal estéreo de mezcla descendente 204 se convierte directamente en la contribución binaural que refleja las HRTFs. La caja TTT 262 de la figura 13, sin embargo, realiza realmente la reconstrucción. De manera opcional, tal como se muestra en la figura 13, la caja TTT 262 puede utilizar a señal residual 270 que refleja la predicción residual cuando reconstruye los canales 242-246 a partir de la mezcla descendente estéreo 204 y los parámetros espaciales 206, indicados más arriba, que comprenden los coeficientes de predicción de canal y, de manera opcional, los valores ICC. Los primeros sumadores 266a están configurados para añadir canales 242-246 para formar el canal izquierdo de la mezcla descendente estéreo 48. En particular, una suma ponderada está formada por los sumadores 266a y 266b, en el que los valores de ponderación se definen mediante las etapas de ponderación 264a, 264b, 264c, y 264E que se podrían aplicar al canal respectivo 246 a 242, un valor de ponderación respectivo EQLL, EQRL y EQCL. De manera similar, los sumadores 268a y 268b forman una suma ponderada de canales 246 a 242 mediante las etapas de ponderación 264b, 264d, y 264e formando los valores de ponderación, formando la suma ponderada el canal derecho de la mezcla descendente estéreo 48.

**[0058]** Los parámetros 270 para las etapas de ponderación de 264a-264E se seleccionan, tal como se describió anteriormente, de tal manera que se consigue la reducción del nivel de canal central descrita anteriormente en la mezcla descendente estéreo 48, lo que resulta, tal como se ha descrito anteriormente, en las ventajas con respecto a la percepción de sonido natural.

**[0059]** Por lo tanto, en otras palabras, la figura 13 muestra un módulo de procesamiento de sala que puede ser aplicado en combinación con el decodificador paramétrico binaural 200' de la figura 12. En la figura 13, se utiliza la señal de mezcla descendente 204 para alimentar el módulo. La señal de mezcla descendente 204 contiene todas las señales de la señal multi-canal para que pueda proporcionar compatibilidad estéreo. Como se mencionó anteriormente, es deseable alimentar el módulo de procesamiento de sala con una señal que contiene sólo una señal central reducida. El modificador de sub-banda de audio espacial modificado de la figura 13 sirve para realizar esta reducción de nivel. En particular, según la figura 13, se puede utilizar una señal residual 270 con el fin de reconstruir los canales central, derecho e izquierdo 242-246. La señal residual de los canales central, derecho e izquierdo 242-246 pueden ser decodificados por el decodificador de audio de mezcla descendente 232, aunque no se muestra en la figura 11. Los parámetros EQ o los valores de ponderación aplicados por las etapas de ponderación 264a-264E pueden ser de valor real para los canales central, derecho e izquierdo 242-246. Un único conjunto de parámetros para el canal central 242 puede ser almacenado y aplicado, y el canal central está, de acuerdo con la figura 13, y a modo de ejemplo, mezclado en igual cantidad en las salidas izquierda y derecha de la mezcla descendente estéreo 48.

**[0060]** Los parámetros EQ 270 alimentados al modificador de sub-banda de audio espacial modificado 234 puede tener las siguientes propiedades. En primer lugar, la señal del canal central se puede atenuar preferiblemente en al menos 6 dB. Además, la señal del canal central puede tener una característica de pasa bajos. Aún más, la señal de diferencia de los canales restantes puede ser impulsada a bajas frecuencias. A fin de compensar el nivel inferior del canal central 242 en relación con los otros canales 244 y 246, la ganancia de los parámetros de HRTF para el canal central utilizada en el modificador de sub-banda de audio espacial binaural 202 debe aumentarse en consecuencia.

**[0061]** El objetivo principal de la configuración de los parámetros EQ es la reducción de la señal del canal central en la salida para el módulo de procesamiento de sala. Sin embargo, el canal central sólo debe ser suprimido de forma limitada: la señal del canal central se resta de la de los canales de mezclado descendente derecho e izquierdo dentro de la caja TTT. Si se reduce el nivel central, los artefactos en los canales izquierdo y derecho pueden ser audibles. Por lo tanto, la reducción del nivel central en la etapa EQ es un compromiso entre la supresión y los artefactos. Es posible encontrar un ajuste fijo de parámetros EQ, pero puede no ser óptimo para todas las señales. En consecuencia, según un ejemplo, se puede utilizar un algoritmo o módulo de adaptación 274 para controlar la cantidad de reducción del nivel central mediante uno, o una combinación de los siguientes parámetros:

**[0062]** Se pueden utilizar los parámetros espaciales 206 utilizados para decodificar el canal central 242 desde los canales de mezclado descendente izquierdo y derecho 204 dentro de la caja TTT 262 tal como se indica mediante la línea discontinua 276.

**[0063]** Se pueden utilizar los niveles de los canales central, izquierdo y derecho tal como se indica mediante la línea discontinua 278.

**[0064]** Se pueden utilizar las diferencias de nivel entre los canales central, izquierdo y derecho 242-246 tal como se indica también mediante la línea discontinua 278.

5 **[0065]** Se pueden utilizar la salida de un algoritmo de detección de tipo único, tal como un detector de actividad de voz, tal como se indica también mediante la línea discontinua 278.

**[0066]** Por último, se puede utilizar estática de metadatos dinámica que describe el contenido de audio con la finalidad de determinar la cantidad de reducción de nivel central, tal como se indica mediante la línea de trazos 280.

10 **[0067]** Aunque algunos aspectos se han descrito en el contexto de un aparato, está claro que estos aspectos también representan una descripción del procedimiento correspondiente, en el que un bloque o dispositivo corresponde a una etapa de procedimiento o una función de una etapa de procedimiento. Análogamente, los aspectos descritos en el contexto de una etapa de procedimiento también representan una descripción de un bloque correspondiente o un elemento o característica de un aparato correspondiente tal como una parte de una ASIC, una sub-rutina de un código de programa o una parte de una lógica programable programada.

15 **[0068]** La señal de audio codificada de la invención puede ser almacenada en un medio de almacenamiento digital o se puede transmitir por un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable, tal como Internet.

20 **[0069]** Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden ser implementadas en hardware o en software. La implementación puede realizarse utilizando un medio de almacenamiento digital, como por ejemplo un disquete, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control legibles electrónicamente almacenadas en esta, que cooperan (o son capaces de cooperar) con un sistema informático programable de tal manera que se lleva a cabo el procedimiento respectivo.

25 **[0070]** Algunas formas de realización de acuerdo con la invención comprenden un portador 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 este documento.

30 **[0071]** Generalmente, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, siendo el código de programa capaz de llevar a cabo uno de los procedimientos, cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede, por ejemplo ser almacenado en un soporte legible por máquina.

35 **[0072]** Otras realizaciones comprenden el programa de ordenador para realizar uno de los procedimientos descritos en este documento, almacenado en un soporte legible por máquina.

40 **[0073]** En otras palabras, una realización del procedimiento de la invención es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los procedimientos descritos en este documento, cuando el programa informático se ejecuta en un ordenador.

45 **[0074]** Una realización adicional de los procedimientos de la invención es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, almacenado este, el programa de ordenador para realizar uno de los procedimientos descritos en este documento.

50 **[0075]** Una realización adicional del procedimiento de la invención es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa de ordenador para realizar uno de los procedimientos descritos en este documento. El flujo de datos o la secuencia de señales pueden, por ejemplo ser configurados para ser transferidos a través de una conexión de comunicación de datos, por ejemplo a través de Internet.

55 **[0076]** Otra forma de realización comprende un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado o adaptado para realizar uno de los procedimientos descritos en este documento.

**[0077]** Una forma de realización comprende además un ordenador que tiene instalado él mismo el programa informático para realizar uno de los procedimientos descritos en este documento.

60 **[0078]** En algunas realizaciones, se puede usar un dispositivo lógico programable (por ejemplo, una matriz de puertas programable en campo) para realizar algunas o todas las funcionalidades de los procedimientos descritos aquí. En algunas realizaciones, una matriz de puertas programable de campo podrá cooperar con un microprocesador para llevar a cabo uno de los procedimientos descritos en este documento. Generalmente, los procedimientos se realizan preferiblemente mediante cualquier aparato de hardware.

65 **[0079]** Las realizaciones anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en el

presente documento serán evidentes para otros expertos en la materia. Es la intención, por lo tanto, estar limitado sólo por el alcance de las reivindicaciones de patente adjuntas y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones del presente documento.

## REIVINDICACIONES

- 5 **1.** Dispositivo para formar un conjunto de HRTFs decreciente de similitud mutua para modelizar una transmisión acústica de una pluralidad de canales a partir de una posición de fuente sonora virtual asociada con el canal respectivo a canales de oído de un oyente, comprendiendo el dispositivo: un proveedor de HRTF (32) adaptado para proporcionar una pluralidad original de HRTFs implementados como filtros FIR, mediante la consulta o el cálculo de tomas de filtro para cada una de la pluralidad original de HRTFs en respuesta a una selección o cambio de las posiciones de fuentes sonoras virtuales; y un procesador HRTF (34) adaptado para provocar que las respuestas de impulso de las HRTFs que modelizan las transmisiones acústicas de un par predeterminado de canales se retrasen entre sí, o para modificar de manera diferente - en un sentido variable espectralmente - sus respuestas de fase y/o amplitud, siendo el par de canales uno de entre un canal izquierdo y un canal derecho de la pluralidad de canales, un canal frontal y posterior de la pluralidad de canales, y un canal central y no central de la pluralidad de canales.
- 10
- 15 **2.** Dispositivo según la reivindicación 1, en el que el procesador HRTF (34) está configurado para provocar que las respuestas de impulso de las HRTFs que modelizan las transmisiones acústicas de un par predeterminado de canales se retrasen entre sí mediante el desplazamiento de las tomas de filtro.
- 20 **3.** Dispositivo según la reivindicación 1 o la 2, en el que el procesador HRTF (34) está configurado para provocar que las respuestas de impulso de las HRTFs que modelizan las transmisiones acústicas de un par predeterminado de canales se retrasen entre sí, o modifiquen de manera diferente - en un sentido variable espectralmente - sus respuestas de fase y/o amplitud de modo que los retardos de grupo de un primero de entre los HRTFs con respecto a otro de los HRTFs, muestren, para bandas de Bark, una desviación estándar de al menos un octavo de una muestra.
- 25 **4.** Dispositivo según cualquiera de las reivindicaciones 1 a 3, en el que el proveedor de HRTF (32) está configurado para proporcionar la pluralidad original de HRTFs a partir de las posiciones de fuentes sonoras virtuales y parámetros HRTF.
- 30 **5.** Dispositivo según cualquiera de las reivindicaciones 1 a 4, en el que el procesador HRTF (34) está configurado para filtrar pasa-todo de manera diferente las respuestas de impulso del par predeterminado de canales.
- 35 **6.** Procedimiento para formar un conjunto decreciente de similitud mutua de funciones de transferencia relativas a la cabeza para modelizar una transmisión acústica de una pluralidad de canales a partir de una posición de fuente sonora virtual asociada con el canal respectivo a canales de oído de un oyente, comprendiendo el procedimiento:
- 40 proporcionar una pluralidad original de HRTFs implementados como filtros FIR, mediante la consulta o el cálculo de tomas de filtro para cada una de la pluralidad original de HRTFs en respuesta a una selección o cambio de las posiciones de fuentes sonoras virtuales; y modificar de manera diferente - en un sentido variable espectralmente - respuestas de fase y/o magnitud de respuestas de impulso de las HRTFs que modelizan las transmisiones acústicas de un par predeterminado de canales de modo que los retardos de grupo de un primero de entre los HRTFs con respecto a otro de los HRTFs muestran, para bandas de Bark, una desviación estándar de al menos un octavo de una muestra, siendo el par de canales uno de entre un canal izquierdo y un canal derecho de la pluralidad de canales, un canal frontal y posterior de la pluralidad de canales, y un canal central y no central de la pluralidad de canales.
- 45 **7.** Programa de ordenador que tiene instrucciones para realizar, cuando se ejecuta en un ordenador, un procedimiento según la reivindicación 6.



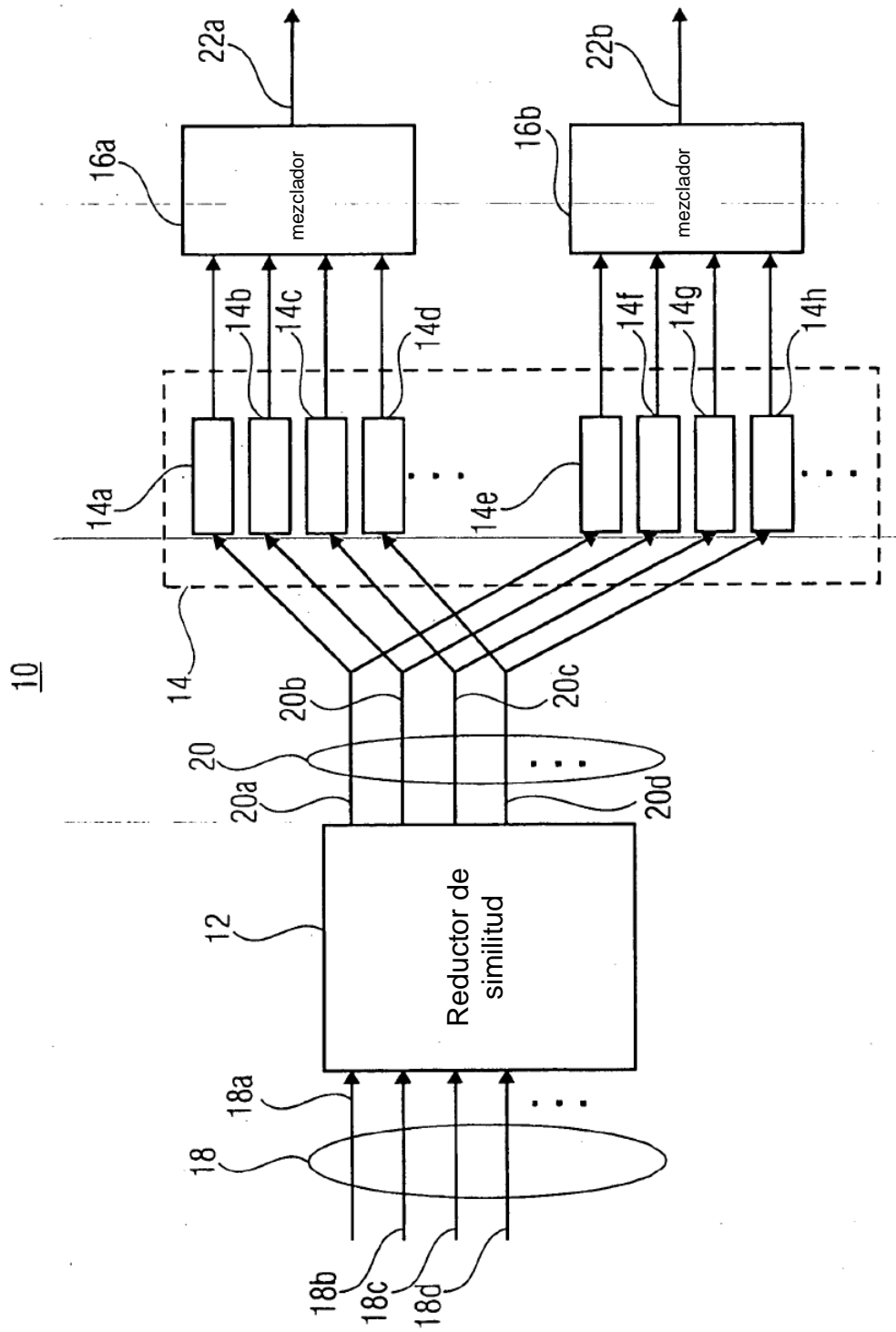


FIG 1

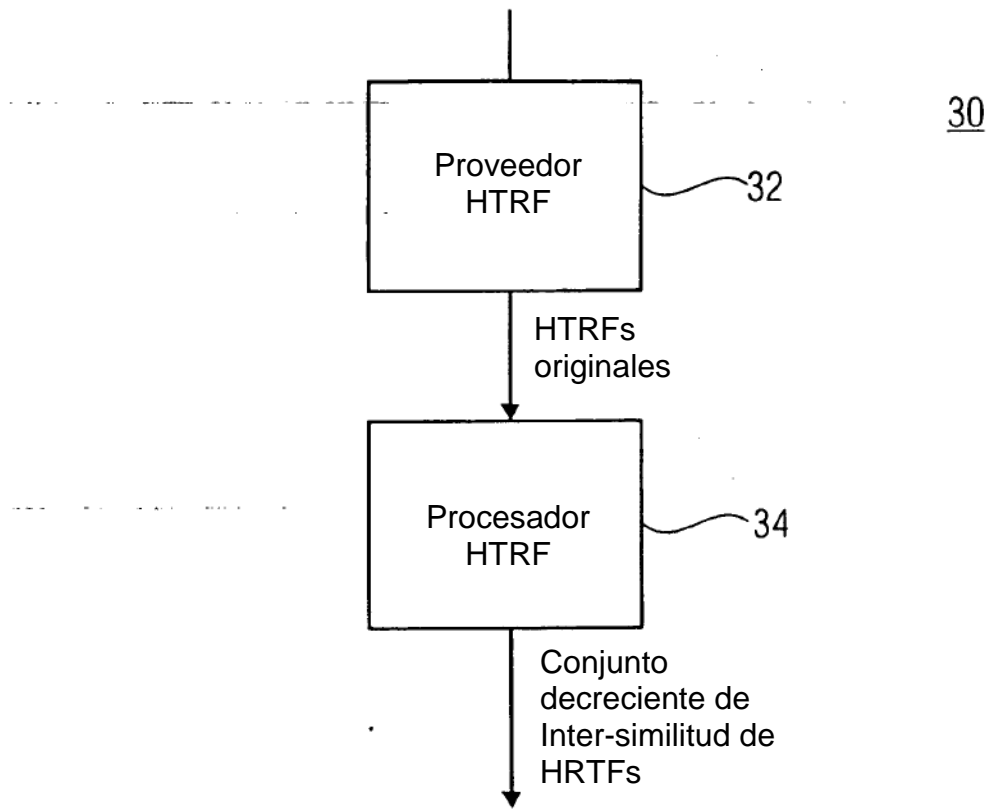


FIG 2

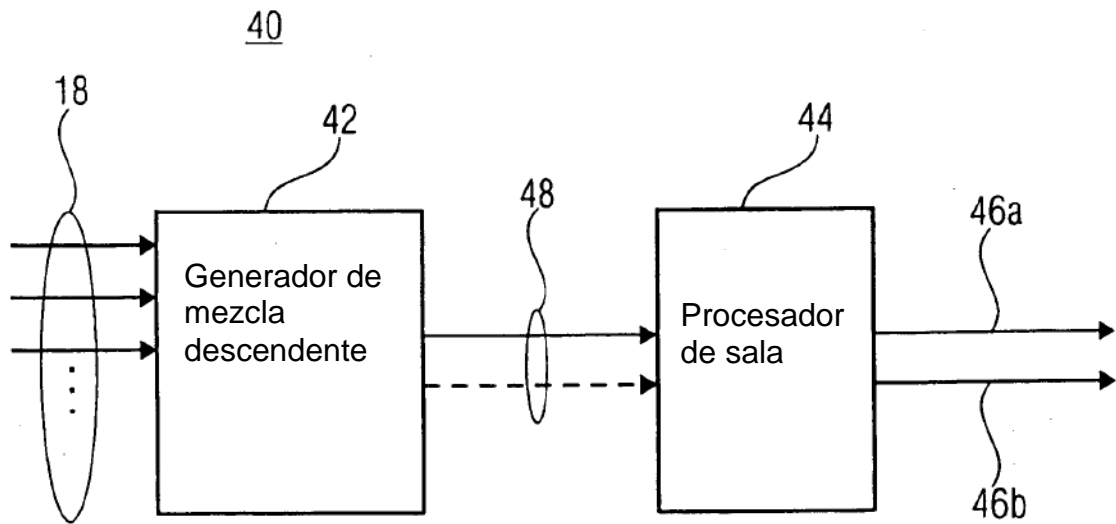


FIG 3

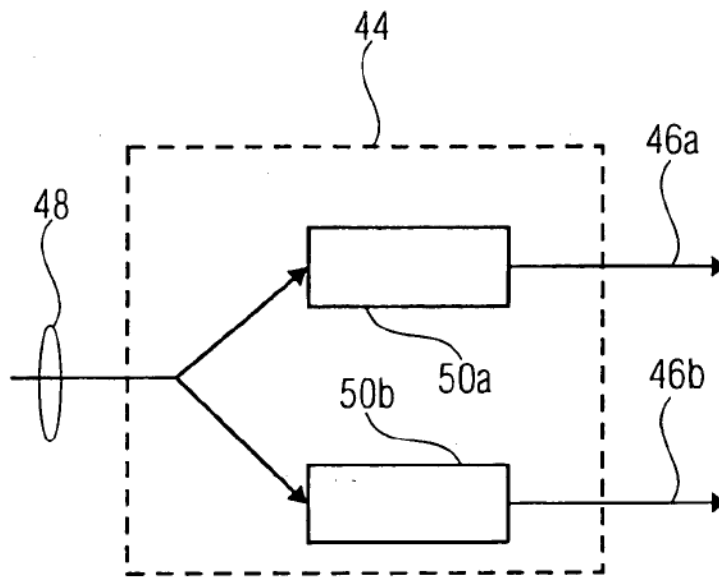


FIG 4A

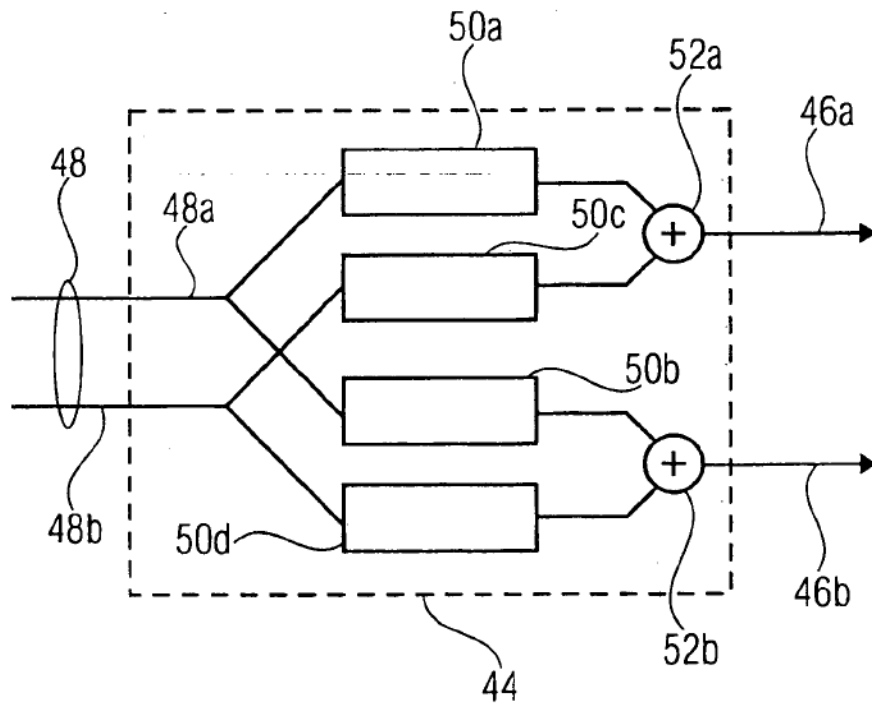


FIG 4B

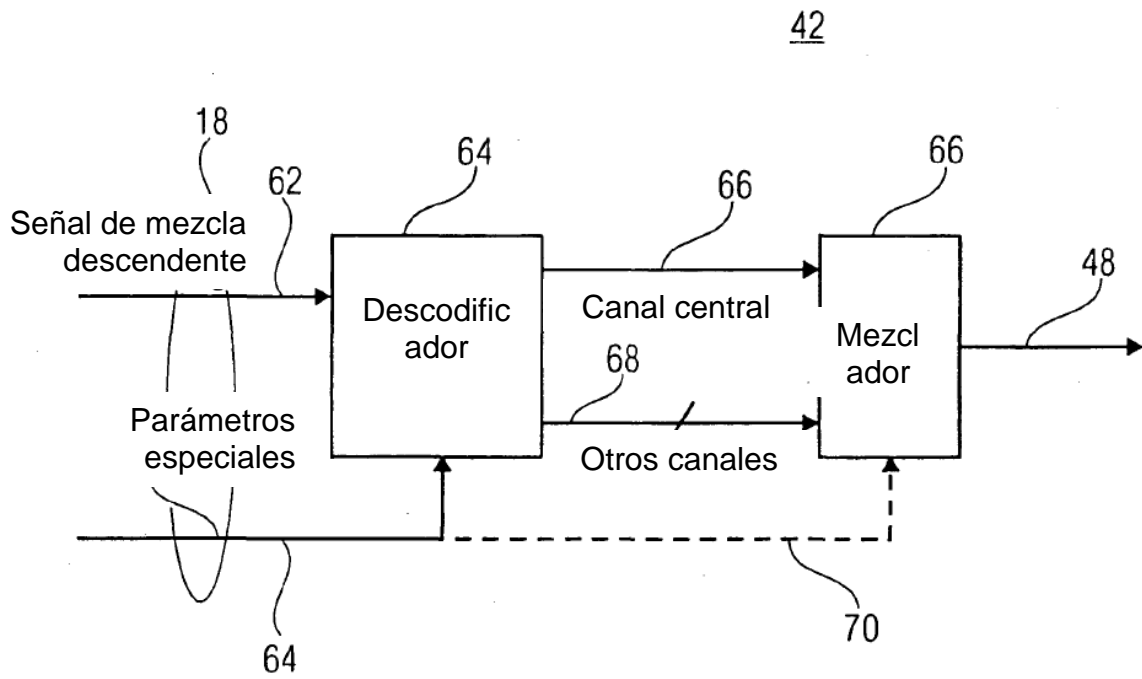


FIG 5

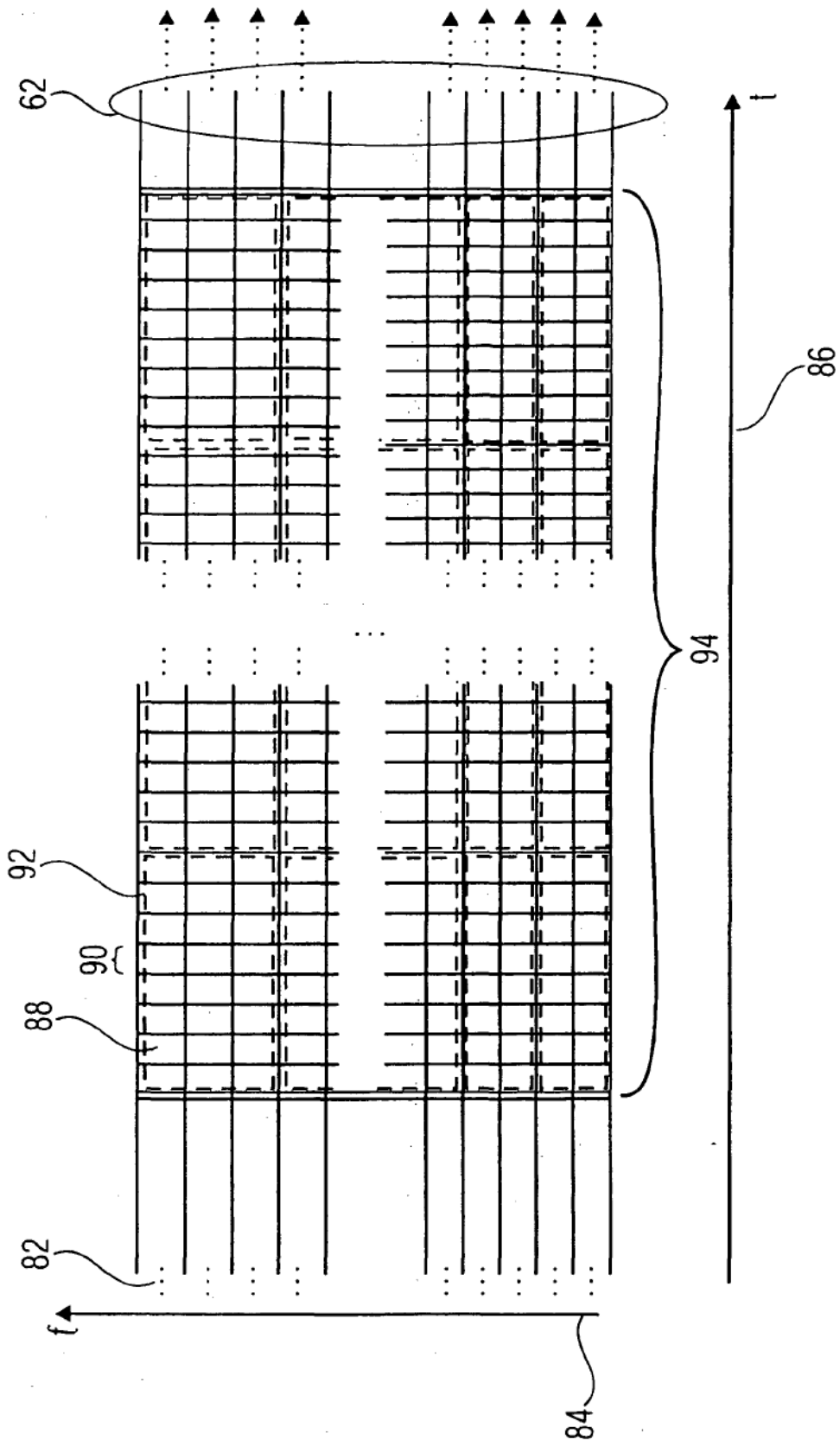
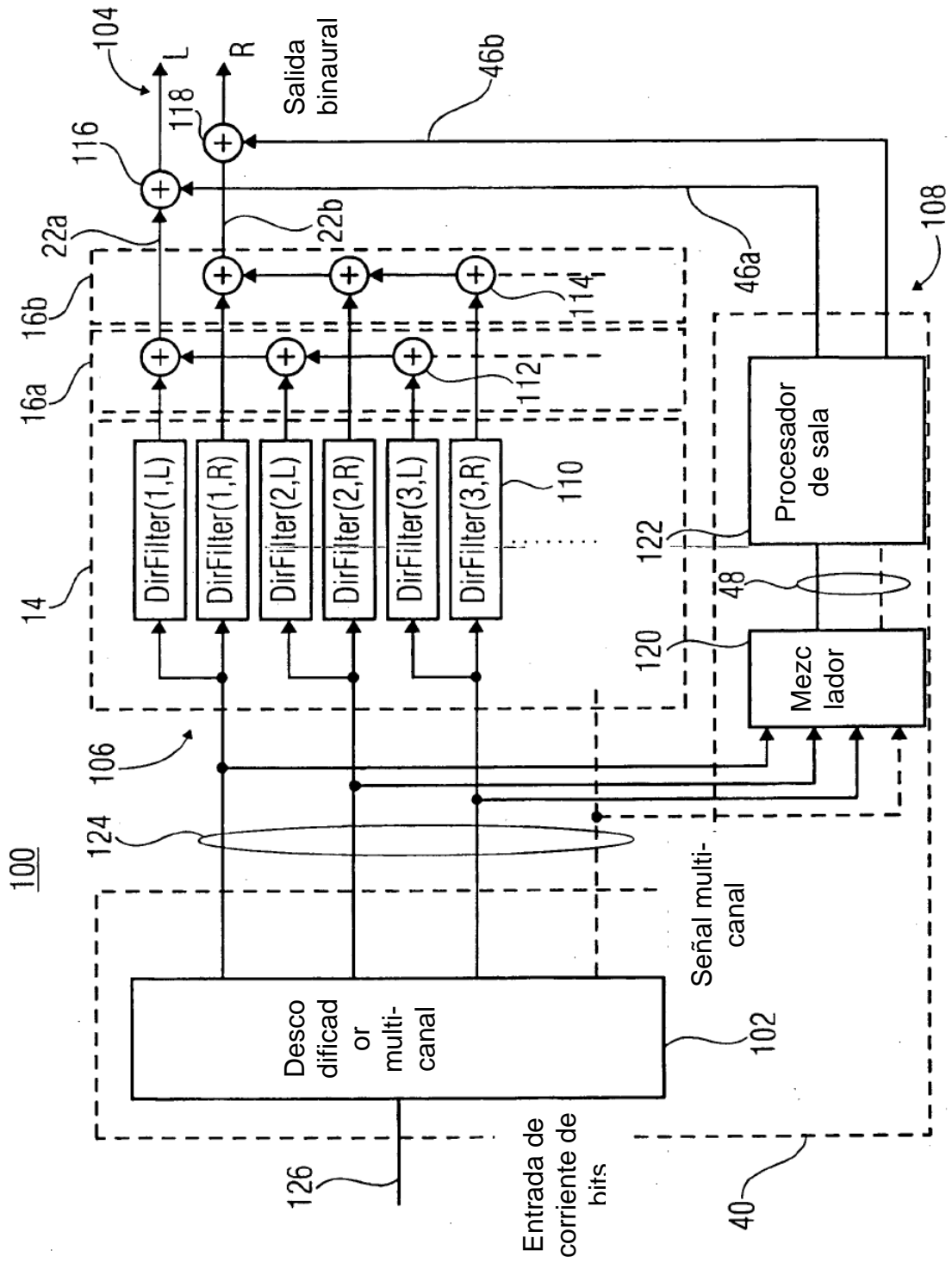
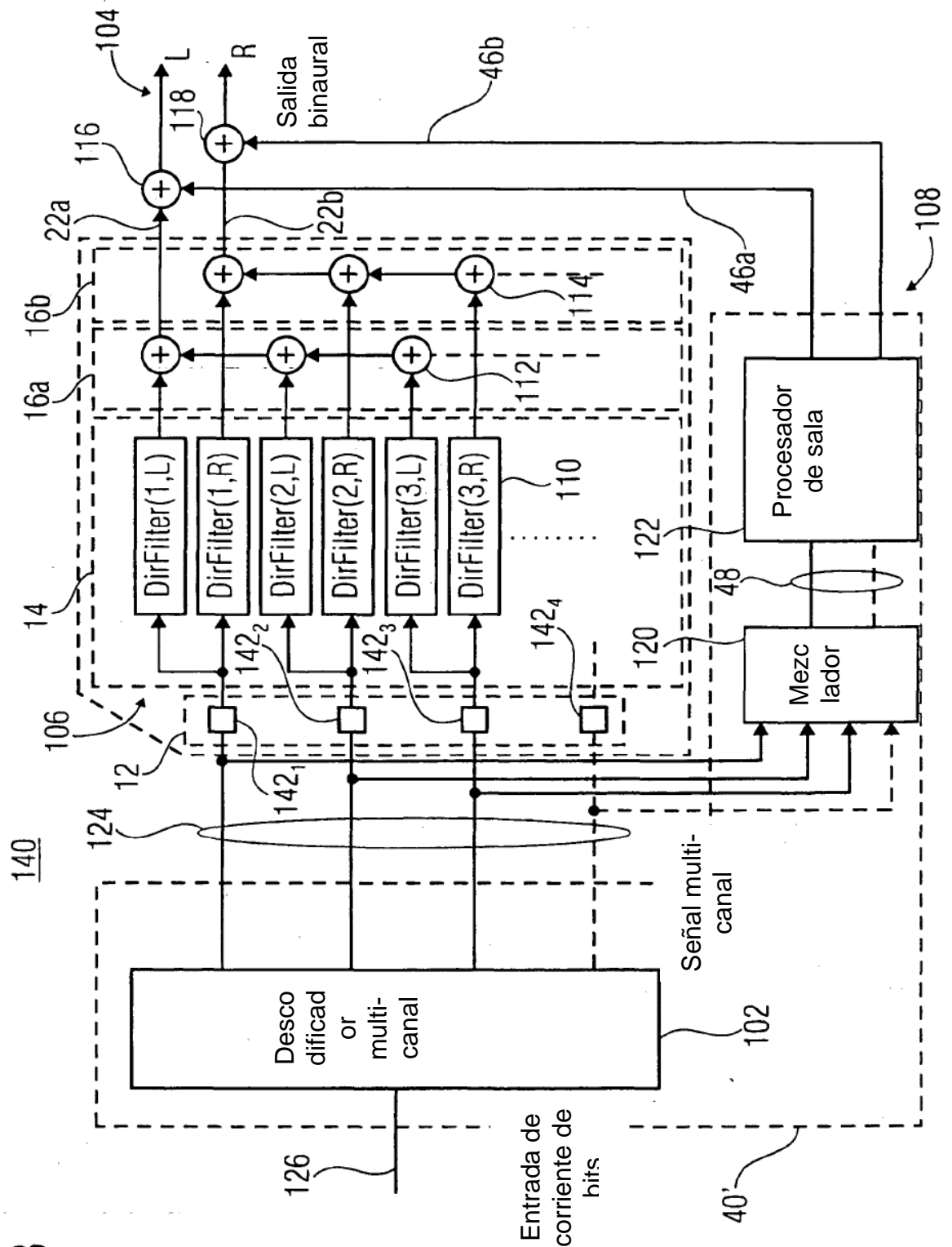


FIG 6

FIG 7





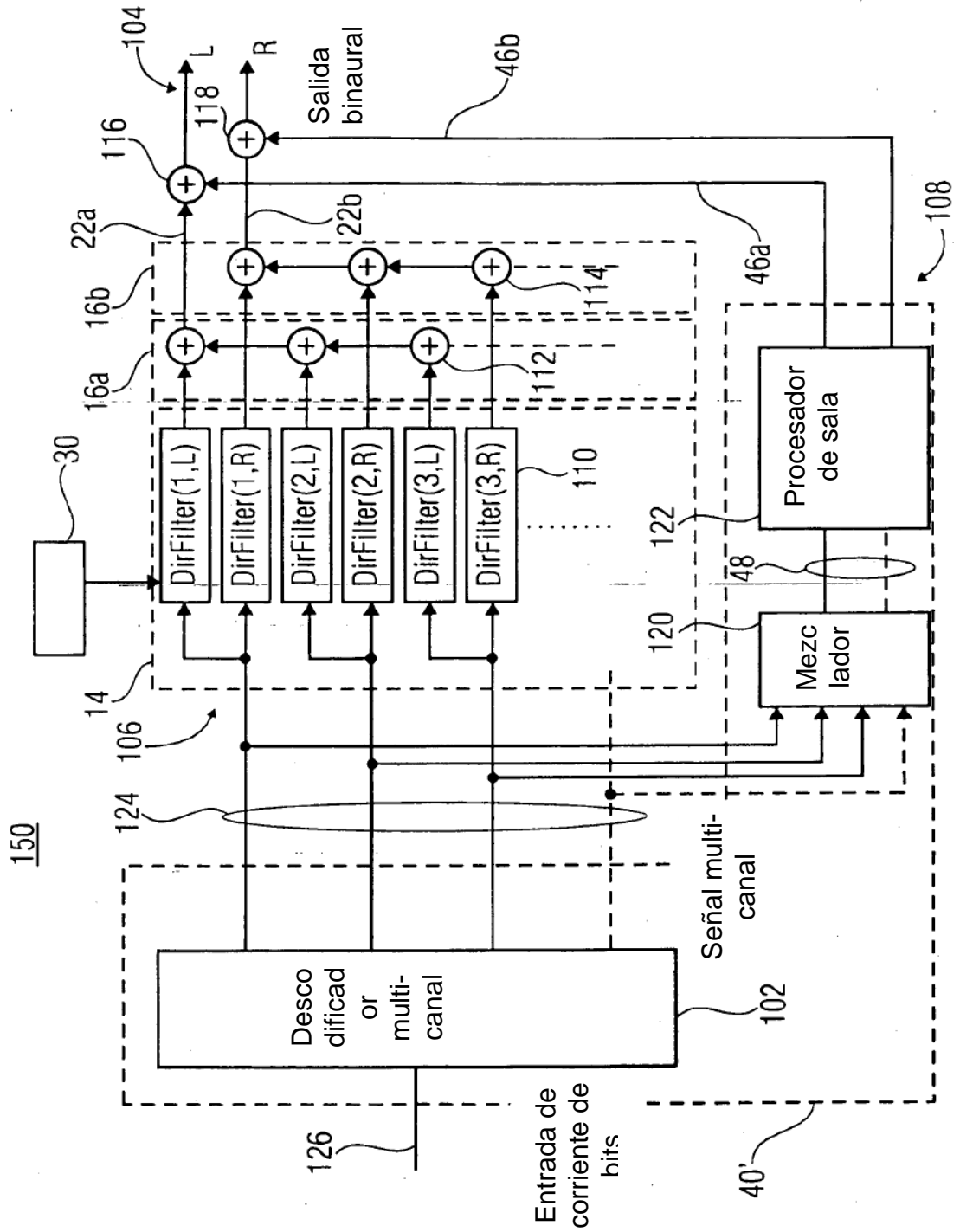


FIG 9



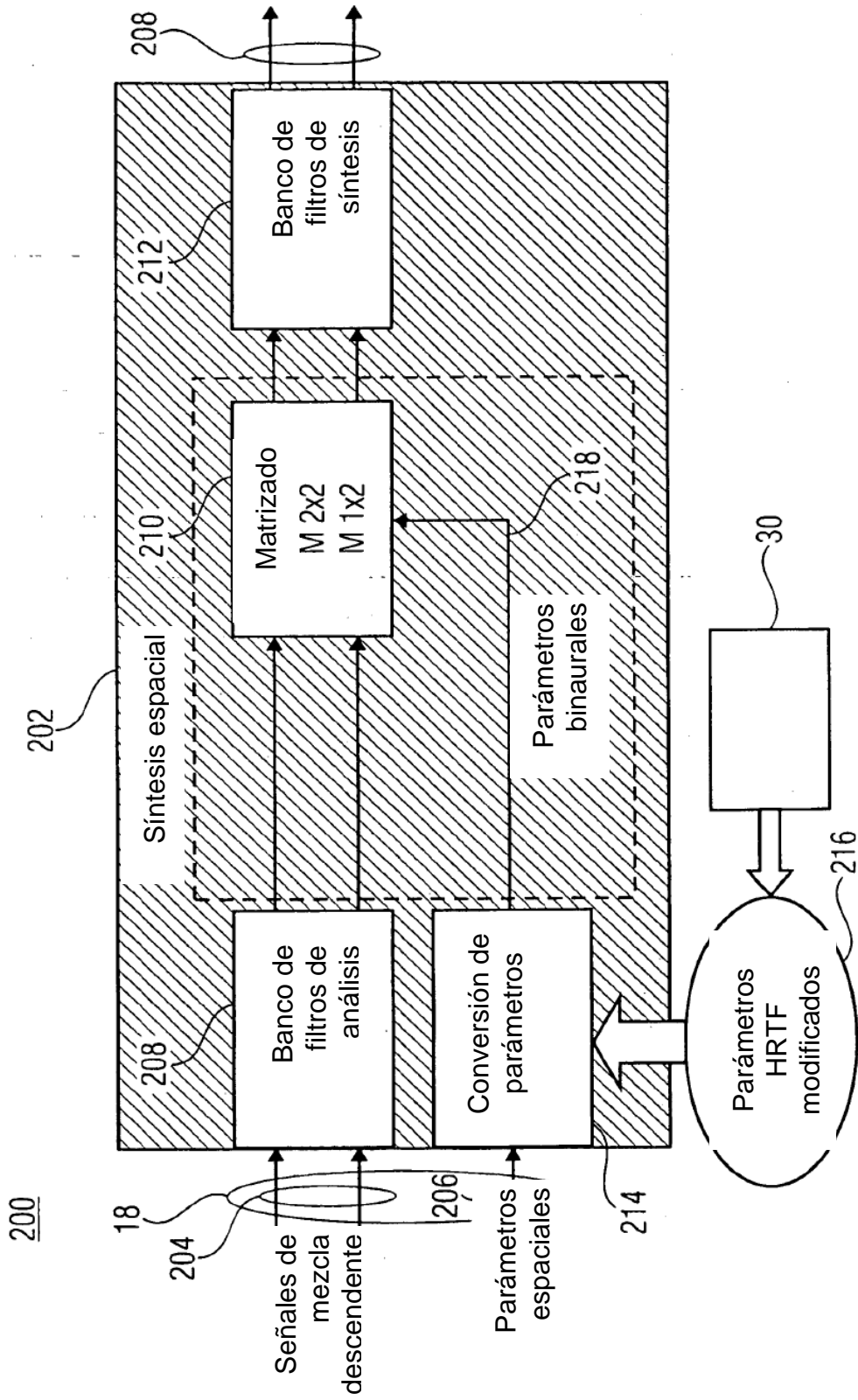


FIG 10

FIG 11

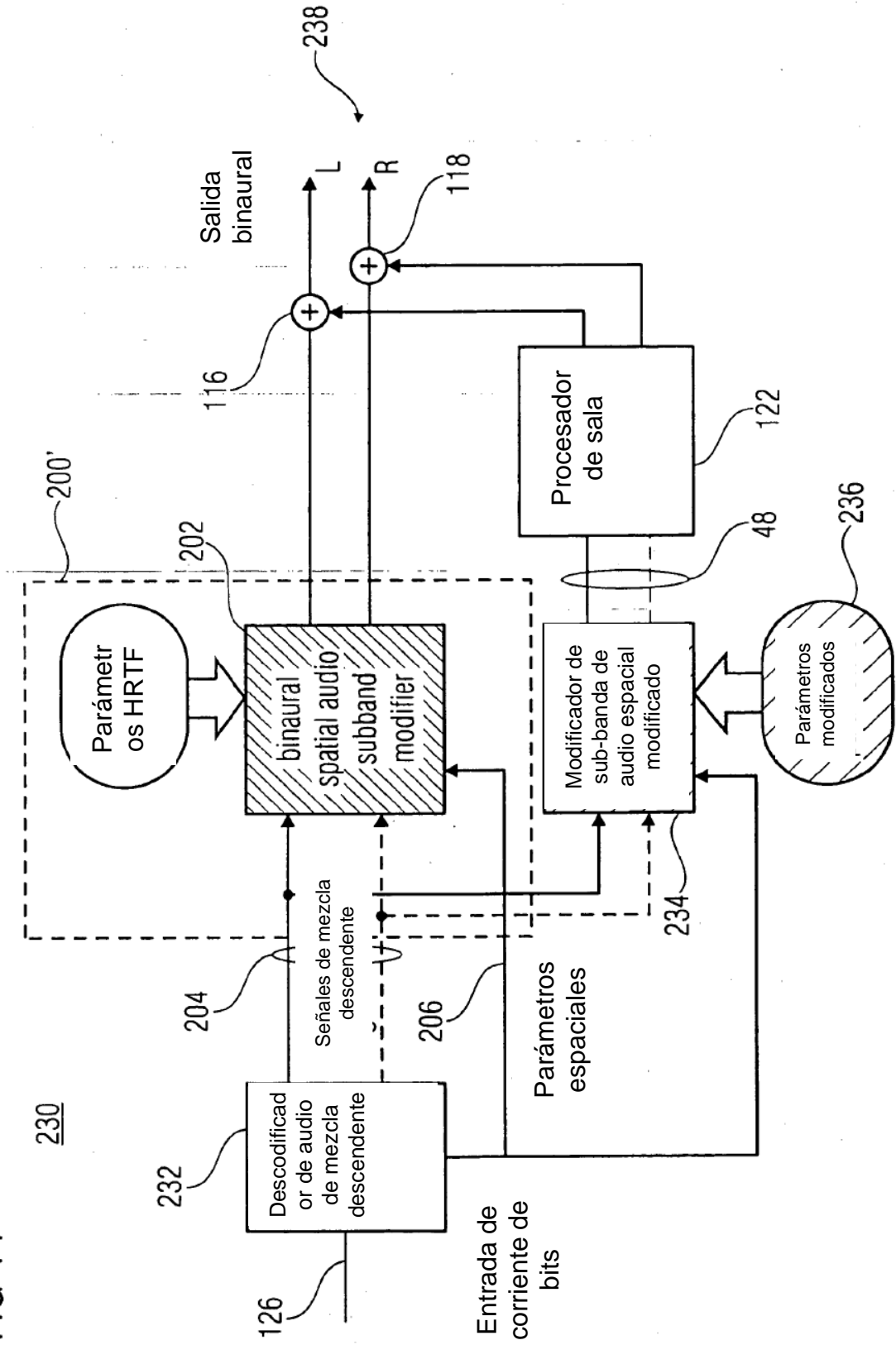


FIG 12

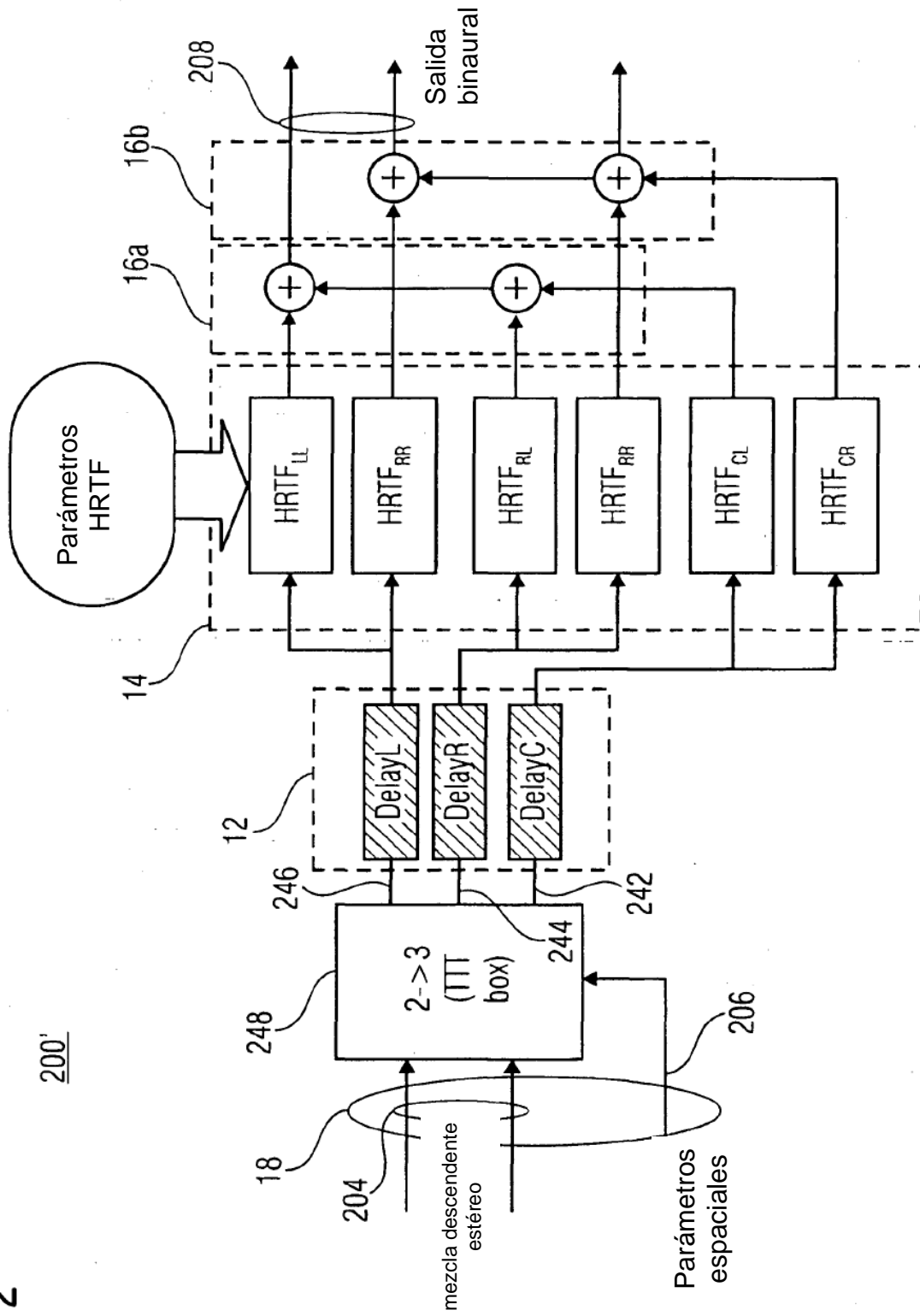
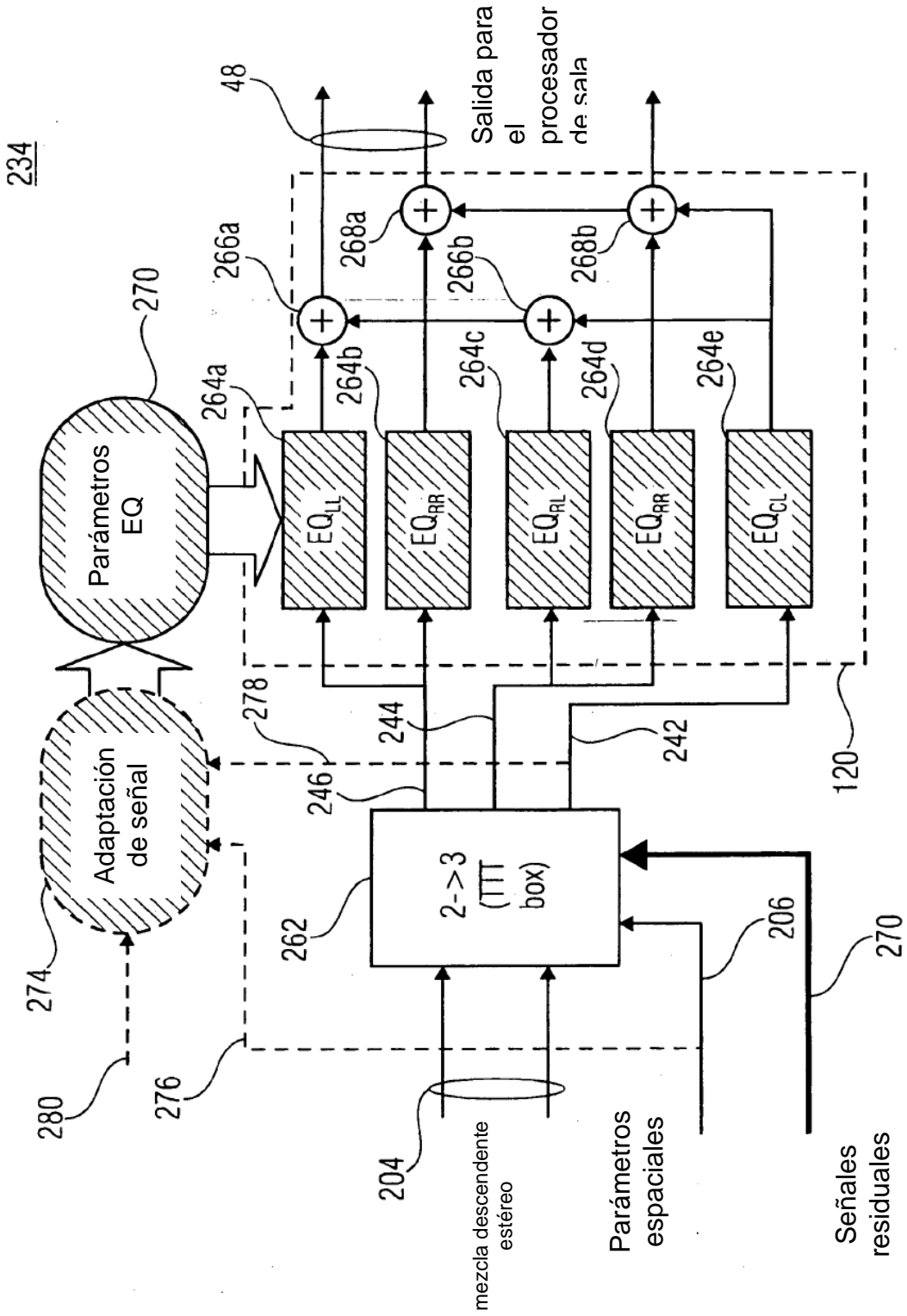


FIG 13



234