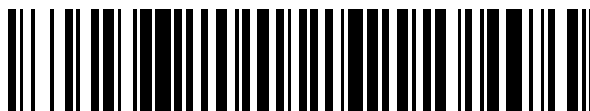


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 528 006**

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 11168513 (7)**

97 Fecha y número de publicación de la concesión europea: **05.11.2014 EP 2384028**

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

30 Prioridad:

31.07.2008 US 85286

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

03.02.2015

73 Titular/es:

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

72 Inventor/es:

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

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 528 006 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 señales para señales binaurales

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

10 **[0002]** El sistema auditivo humano es capaz de determinar la dirección o direcciones desde dónde 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 interaurales 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 con sombras tiene una imagen de sonido suprimida naturalmente, en comparación con el oído sin sombras. Otra propiedad muy importante al tratar con la localización es la diferencia temporal inter-aural (ITD). El oído con sombras tiene una distancia más larga a la fuente de sonido, y por lo tanto consigue el frente de onda de sonido más tarde que el oído sin sombras. El significado del ITD se enfatiza en las frecuencias bajas que no se atenúan mucho cuando llegan al oído con sombras en comparación con el oído sin sombras. La ITD es menos importante en las frecuencias más altas debido a que la longitud de onda del sonido se aproxima más 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 hacia los oídos izquierdo y derecho, respectivamente.

25 **[0003]** Los problemas se producen cuando una persona escucha a través de auriculares una señal estéreo que está destinada a ser reproducida por una configuración de altavoces. Es muy probable que el oyente considere el sonido como poco natural, incómodo, y perturbador puesto 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". A largo plazo, escuchar el sonido "en la cabeza" puede conducir a fatiga auditiva. Esto se produce porque la información en la que el sistema auditivo humano se basa, al ubicar las fuentes de sonido, es decir, las señales interaurales, falta o es ambigua.

35 **[0004]** A fin de representar señales estéreo, o incluso señales multicanal con más de dos canales para la reproducción de auriculares, se pueden usar filtros direccionales para modelar estas interacciones. Por ejemplo, la generación de una salida de auriculares a partir de una señal multicanal descodificada puede comprender filtrar cada señal después de la decodificación por medio de un par de filtros direccionales. Estos filtros suelen modelizar la transmisión acústica desde una fuente de sonido virtual en una sala hasta el canal auditivo de un oyente, la llamada función de transferencia de habitación binaural (BRTF). La BRTF realiza modificaciones de tiempo, 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 más bien largos, tales como 20.000 intervalos de filtro a 44,1 kHz, y el proceso de filtrado es computacionalmente exigente. Por lo tanto, los filtros direccionales a veces se reducen 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 por medio de 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 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 implica que 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 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 frontal izquierdo y el canal central y el canal delantero derecho, respectivamente, en relación con una suma y una diferencia de los canales posterior izquierdo y posterior derecho 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 falta de externalización. Además, se ha descubierto que a pesar de las medidas antes mencionadas para la representación de señales multicanal para la reproducción por auriculares, las partes de voz en los diálogos de películas y música a menudo se perciben como reverberantes de manera poco natural y espectralmente desiguales.

5 **[0007]** US5371799 describe un sistema para procesar una señal de audio para reproducción en auriculares en la que una fuente de sonido aparente que está localizada fuera de la cabeza del oyente que procesa la señal de entrada como si estuviera hecha de una parte de onda directa, una parte de reflexión temprana, y una parte de reverberación. La parte de onda directa de la señal se procesa en filtros cuyos coeficientes de filtro se escogen a partir del acimut deseado de la posición de fuente de sonido virtual. La parte de reflexión temprana se hace pasar a través de un banco de filtros conectados en paralelo cuyos coeficientes se escogen a partir de cada acimut de reflexión. Las salidas de estos filtros se hacen pasar a través de escalares para ajustar la amplitud para simular un intervalo deseado de la fuente de sonido virtual. La parte de reverberación se procesa sin información alguna de ubicación de fuente de sonido, empleando un generador de números aleatorios, por ejemplo, y la salida se atenúa en un atenuador exponencial para desvanecerse. Entonces las salidas de los escalares y atenuadores se suman para producir señales de auriculares izquierda y derecha para su reproducción en transductores de auriculares respectivos.

15 **[0008]** Por lo tanto, es el objeto de la presente invención proporcionar un esquema de generación de señales binaurales, que dé como resultado una reproducción por auriculares más estable y agradable.

20 **[0009]** Este objeto se logra mediante dispositivos según la reivindicación 1 y un procedimiento según la reivindicación 9.

25 **[0010]** Otra idea que subyace en la presente solicitud es que algunas partes en las películas y la música dan como resultado de una reproducción en auriculares percibida de forma más natural, cuando la mezcla descendente mono o estéreo de los canales de la señal multi-canal que se deben someter al procesador para generar la contribución relativa a la reflexión/ reverberación de local de la señal binaural, se forma de modo que la pluralidad de canales contribuyen a la mezcla descendente mono o estéreo en un nivel que difiere entre al menos dos canales de la señal multi-canal. Por ejemplo, los inventores se dieron cuenta de que las voces en los diálogos de películas y música 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, suele dar 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 tal como, por ejemplo, una atenuación de 3-12 dB, o específicamente, 6 dB.

35 **[0011]** Además, una señal binaural más estable y agradable para la reproducción por auriculares se puede lograr mediante el procesamiento diferente de, y reduciendo así la similitud entre, al menos uno de un canal izquierdo y un canal derecho de la pluralidad de canales de entrada, un canal frontal y un canal posterior de la pluralidad de canales de entrada, y un canal central y un canal no central de la pluralidad de canales, obteniendo de este modo un conjunto de canales de inter-similitud reducida. Este conjunto de canales de inter-similitud reducida se alimenta 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 multicanal, la anchura espacial de la señal de salida binaural puede ser aumentada y la externalización puede ser mejorada.

45 **[0012]** Además, una señal binaural más estable y agradable para la reproducción por auriculares se puede lograr mediante la modificación de manera diferente de una fase y/o magnitud - en un sentido variable espectralmente - entre al menos dos canales de la pluralidad de canales, de modo que se obtiene un conjunto de canales de inter-similitud reducida que, a su vez, se puede suministrar a una pluralidad de filtros direccionales seguidos por respectivos mezcladores para los oídos izquierdo y derecho, respectivamente. Una vez más, mediante la reducción de la inter-similitud entre canales de la señal de entrada de múltiples canales, la anchura espacial de la señal de salida binaural puede ser aumentada y la externalización puede ser mejorada.

50 **[0013]** Además se puede formar un conjunto de funciones de transferencia relacionadas con la cabeza de inter-similitud decreciente 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 - las respuestas de fase y/o magnitud de la pluralidad original de funciones de transferencia relacionadas con la cabeza sean diferentes entre sí. La formación se puede realizar fuera de línea como una etapa de fabricación, o en línea durante la generación de señales binaurales empleando las funciones de transferencia relacionadas con la cabeza como filtros direccionales tales como, por ejemplo, que respondan a una indicación de posiciones de fuente de sonido virtual a utilizar.

60 **[0014]** A continuación, se describen realizaciones preferidas con más detalle haciendo referencia a las figuras, entre las cuales:

La figura 1 muestra un diagrama de bloque de un dispositivo para generar una señal binaural;

La figura 2 muestra un diagrama de bloque de un dispositivo para formar un conjunto de funciones de transferencia relacionadas con la cabeza de inter-similitud decreciente;

65 La figura 3 muestra un dispositivo para generar una contribución relacionada con la reflexión y/o la reverberación de una señal binaural según una realización adicional;

Las figuras 4a y 4b muestran diagramas de bloque del procesador de local de la figura 3 según realizaciones distintas;

La figura 5 muestra un diagrama de bloque del generador de mezcla descendente de la figura 3 según una realización;

5 La figura 6 muestra un diagrama esquemático que ilustra una representación de una señal multi-canal empleando codificación de audio espacial según una realización;

La figura 7 muestra un generador de señal de salida binaural según una realización;

La figura 8 muestra un diagrama de bloque de un generador de señal de salida binaural según una realización adicional;

10 La figura 9 muestra un diagrama de bloque de un generador de señal de salida binaural según otra realización adicional;

La figura 10 muestra un diagrama de bloque de un generador de señal de salida binaural según una realización adicional;

15 La figura 11 muestra un diagrama de bloque de un generador de señal de salida binaural según una realización adicional;

La figura 12 muestra un diagrama de bloque del decodificador de audio espacial binaural de la figura 11 según una realización; y

La figura 13 muestra un diagrama de bloque del decodificador de audio espacial modificado de la figura 11 según una realización.

20 **[0015]** La figura 1 muestra un dispositivo para generar una señal binaural deseada, por ejemplo, para reproducción por auriculares a partir de una señal multi-canal que representa una pluralidad de canales y destinado a la reproducción por una configuración de altavoces que tienen una posición de fuente de sonido virtual asociada a cada canal. El dispositivo que se indica de manera general 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.

30 **[0016]** 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 20 de canales 20a-20d de inter-similitud reducida. El número de canales 18a-18d representados por la señal multi-canal 18 puede ser de dos o más. Solamente con fines ilustrativos, se han mostrado 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 reproducen por una configuración de altavoces (que no se muestra la figura 1), que tiene los altavoces colocados en posiciones de fuentes de sonido virtuales asociadas a cada canal 18a-18d.

40 **[0017]** Según la figura 1, la pluralidad de canales 18a-18d comprende, al menos, un par canales izquierdo y derecho, un par de canales frontal y posterior, o un par canales central y no central. Por supuesto, más de una de las parejas que se acaban de mencionar, pueden estar presentes dentro de la pluralidad 18 de canales 18a-18d. El reductor de similitud 12 está configurado para procesar de forma diferente, y por lo tanto reducir una similitud entre los canales de la pluralidad de canales, con la finalidad de obtener el conjunto de inter-similitud reducida 20 de canales 20a-20d. De acuerdo con un primer aspecto, la similitud entre al menos uno de, un canal izquierdo y un canal derecho de la pluralidad de canales 18, un canal frontal y un canal posterior de la pluralidad de canales 18, y un canal central y un canal no central de la pluralidad de canales 18 puede ser reducida por el reductor de similitud 12, con la finalidad de obtener el conjunto de inter-similitud reducida 20 de canales 20a-20d. De acuerdo con un segundo aspecto, el reductor de similitud (12) puede - adicionalmente o como alternativa - realizar - en un sentido espectralmente variable - una modificación de fase y/o magnitud de manera diferente entre al menos dos canales de la pluralidad de canales, con la finalidad de obtener el conjunto de inter-similitud reducida 20 de canales.

50 **[0018]** Como se describe con más detalle a continuación, el reductor de similitud 12 puede, por ejemplo, lograr el procesamiento diferente haciendo que los respectivos pares se retrasen entre sí, o sometiendo los respectivos pares de canales a retrasos de diferentes cantidades en, por ejemplo, cada una de una pluralidad de bandas de frecuencias, obteniendo con ello un conjunto de canales de inter-correlación reducida 20. Hay, por supuesto, otras posibilidades para disminuir la correlación entre los canales. En incluso otras palabras, 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 como una magnitud de uno sobre el rango de espectro de audio correspondiente en el que, sin embargo, el reductor de similitud 12 modifica de manera diferente las fases de sub-bandas o componentes de frecuencia de los mismos. Por ejemplo, el reductor de correlación 12 se podría configurar de modo que este provoque una modificación de fase en todos, 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 el grupo se retrase 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 un octavo de

una muestra. Las bandas de frecuencia consideradas podrían ser las bandas de Bark o un subconjunto de las mismas o cualquier otra sub-división de bandas de frecuencia.

5 **[0019]** La reducción de la correlación no es la única manera de evitar al sistema auditivo humano la localización en la cabeza. Más bien, la correlación es una de varias medidas posibles mediante el uso de las cuales el sistema auditivo humano mide la similitud del sonido que llega a ambos oídos, y por lo tanto, la dirección de proveniencia del sonido. En consecuencia, el reductor de similitud 12 también puede lograr el procesamiento diferente sometiendo los pares de canales respectivos a reducciones de nivel de diferentes cantidades en, por ejemplo, cada de una pluralidad de bandas de frecuencia, obteniéndose de este modo un conjunto de inter-similitud reducida 20 de canales en de manera espectralmente conformada. La formación espectral puede, por ejemplo, exagerar la reducción formada espectralmente relativa que ocurre, por ejemplo, para sonido de canal posterior en relación con un canal de sonido frontal debido al sombreado por solapamiento de oídos. En consecuencia, el reductor de similitud 12 puede someter el o los canales posteriores a reducciones de nivel variables espectralmente en relación con otros canales. En esta formación espectral, el reductor de similitud 12 puede tener respuesta constante sobre el intervalo de audio relevante en el que, sin embargo, el reductor de similitud 12 de manera diferente modifica magnitudes de subbandas o componentes de frecuencia de estas.

20 **[0020]** La manera en que la señal multi-canal 18 representa una pluralidad de canales 18a-18d no está, en principio, restringida a ninguna representación específica. Por ejemplo, la señal multi-canal 18 podría representar la pluralidad de canales 18a-18d de manera comprimida, empleando codificación de audio espacial. Según la codificación de audio espacial, la pluralidad de canales 18a-18d podría representarse mediante una señal de mezcla descendente a partir de la cual los canales se mezclan hacia abajo, con la información de mezcla descendente que revela la relación de mezclado según la cual los canales individuales 18a-18d se han mezclado en el canal de mezcla descendente o canales de mezcla descendente, y parámetros espaciales que describen la imagen espacial de la señal multi-canal 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 entre los canales individuales 20a-20d. Estos últimos canales pueden, por ejemplo, ser emitidos como señales temporales o como espectrogramas que estén, por ejemplo, descompuestos espectralmente en subbandas.

30 **[0021]** Los filtros direccionales 14a-14h están configurados para modelizar una transmisión acústica de uno respectivo de los canales 20a-20d desde una posición de fuente de sonido virtual asociada con el canal respectivo a un canal auditivo respectivo del oyente. En la figura 1, los filtros direccionales 14a- 14d modelizan la transmisión acústica a, por ejemplo, el canal auditivo izquierdo, mientras que los filtros direccionales 14e-14h modelizan la transmisión acústica al canal auditivo derecho. Los filtros direccionales pueden modelizar la transmisión acústica desde una posición de fuente de sonido virtual en una sala a un canal auditivo del oyente y pueden modelizarla mediante modificaciones de tiempo, nivel y espectral, y opcionalmente, modelizando reflexiones y reverberación. 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, filtros FIR, o pueden operar en el dominio de la frecuencia multiplicando los valores de muestra de función de transferencia respectivos por 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 respectiva señal del canal 20a-20d desde la posición de la fuente de sonido virtual respectiva con el canal auditivo respectivo, incluyendo, por ejemplo, las interacciones con la cabeza, los oídos y los hombros de un humano. El primer mezclador 16a está configurado para mezclar las salidas de los filtros direccionales 14a-14d que modelizan 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 configurado para mezclar las salidas de los filtros direccionales 14e-14h que modelizan la transmisión acústica al canal auditivo 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.

55 **[0022]** Como se describirá con más detalle a continuación con las respectivas formas de realización, se pueden añadir contribuciones adicionales a las señales 22a y 22b, con el fin de tener en cuenta las reflexiones y / o la reverberación de la sala. Con esta medida, se puede reducir la complejidad de los filtros direccionales 14a-14h.

60 **[0023]** En el dispositivo de la figura 1, el reductor de similitud 12 contrarresta los efectos secundarios negativos de la suma de las señales correlacionadas de entrada de los mezcladores 16a y 16b, respectivamente, según la cual resulta una anchura espacial reducida de la señal de salida binaural 22a y 22b y una falta de externalización. La decorrelación alcanzada por el reductor de similitud 12 reduce estos efectos secundarios negativos.

65 **[0024]** La figura 1 muestra, en otras palabras, en otras palabras, un flujo de señal para la generación de una salida de auriculares de, por ejemplo, una señal multicanal decodificada. Cada señal se filtra por un par de pares de filtros direccionales. Por ejemplo, el canal 18a se filtra por el par de filtros direccionales 14a-14e. Desafortunadamente, existe una cantidad significativa de similitud como existe 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 de múltiples canales con un filtro direccional 14a-14h, la salida de señales intermedia por los

5 filtros direccionales 14a-14h se añaden en el mezclador 16a y 16b del para formar la señal de salida de auriculares 20a y 20b. La suma de las señales de salida similares / correlacionadas daría lugar a una anchura espacial reducida de la señal de salida 20a y 20b, y una falta de externalización. Esto es particularmente problemático para la similitud / correlación de la señal izquierda y derecha y el canal central. De acuerdo con ello, el reductor de similitud 12 es para reducir la similitud entre estas señales en la medida de lo posible.

10 **[0025]** 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 de reductor de similitud 12 concurrentemente con la modificación de los filtros direccionales para llevar a cabo no sólo el modelado antes mencionado de la transmisión acústica, sino también lograr la dis-similitud tal como la descorrelación que se acaba de mencionar. En consecuencia, los filtros direccionales, por ejemplo, no modelarían las HRTFs, sino funciones de transferencia relacionadas con la cabeza modificadas.

15 **[0026]** La figura 2, por ejemplo, muestra un dispositivo para formar un conjunto de inter-similitud decreciente de funciones de transferencia relacionadas con la cabeza para modelizar una transmisión acústica de un conjunto de canales desde una posición de fuente de sonido virtual asociada con el canal respectivo a los canales auditivos de un oyente. El dispositivo que se indica en general con 30 comprende un proveedor de HRTF 32, así como un procesador de HRTF 34.

20 **[0027]** El proveedor de HRTF 32 está configurado para proporcionar una pluralidad original de HRTFs. La etapa 32 puede comprender mediciones utilizando una cabeza de maniquí estándar, con el fin de medir las funciones de transferencia relacionadas con la cabeza de ciertas posiciones de sonido a los canales del oído de un oyente maniquí estándar. Del mismo modo, el proveedor de HRTF 32 puede estar configurado para simplemente consultar o cargar las HRTF originales de una memoria. Incluso como alternativa, el proveedor de HRTF 32 puede estar configurado para calcular los HRTF de acuerdo con una fórmula predeterminada, dependiendo de, por ejemplo, posiciones de las fuentes de sonido virtuales de interés. En consecuencia, el proveedor de HRTF 32 puede estar configurado para operar en un entorno de diseño para el diseño de un generador de señal de salida binaural, o puede ser parte del propio generador de señal binaural de salida, a fin de proporcionar las HRTF originales en línea, tales como, por ejemplo, que respondan a una selección o cambio de las posiciones de las fuentes de sonido virtuales. Por ejemplo, el dispositivo 30 puede ser parte de un generador de señal de salida binaural que es capaz de acomodar señales multicanal destinadas a diferentes configuraciones de altavoces que tienen diferentes posiciones de fuentes de sonido virtuales asociadas con sus canales. En este caso, el proveedor de HRTF 32 puede estar configurado para proporcionar los HRTF originales de una manera adaptada a las posiciones de las fuentes de sonido virtuales actualmente tratadas.

35 **[0028]** El procesador HRTF 34, a su vez, está configurado para hacer que las respuestas de impulso de al menos un par de las HRTF se desplacen entre sí o modificar- en un sentido espectralmente variable - las respuestas de fase y/o magnitud de estas de manera diferente entre sí. 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 retrasar la HRTF de un canal respectivo, modificar la respuesta de fase de una HRTF respectiva y/o aplicar un filtro de decorrelación tal como un filtro pasa todo a la HRTF respectiva, obteniendo de este modo un conjunto de HRTFs de inter-correlación reducida, y/o modificando - en un sentido de modificación espectral - la respuesta de magnitud de una HRTF respectiva, obteniendo de este modo, al menos, un conjunto de HRTFs inter-similitud reducida. En cualquier caso, la descorrelación / diferencia resultante entre los respectivos canales puede apoyar el sistema auditivo humano en la localización de la fuente de sonido externa y de ese modo prevenir la localización en la cabeza que se produzca. Por ejemplo, el procesador de 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 introduce un retraso de grupo de una primera HRTF para una determinada banda de frecuencia - o una determinada banda de frecuencia de una primera HRTF se retrasa - en relación con una de las HRTFs en al menos una muestra. Además, el procesador de 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 una de las HRTFs para una pluralidad de bandas de frecuencia, muestre 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 las mismas o cualquier otra sub-división de bandas de frecuencias.

60 **[0029]** El conjunto de HRTFs de inter-similitud decreciente resultante del procesador de HRTF 34 se puede utilizar para ajustar los HRTFs de los filtros direccionales 14a-14h del dispositivo de la figura 1, en el que el reductor de similitud 12 puede estar presente o ausente. Debido a la propiedad de dis-similitud 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 consigue de manera similar incluso cuando no hay reductor de similitud 12.

65 **[0030]** Como ya se ha descrito anteriormente, el dispositivo de la figura 1 podrá ir acompañado por una pasada adicional configurada para obtener contribuciones relacionadas con la reflexión y / o reverberación de sala de la señal de salida binaural 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 direccionales 14a-14h. Un dispositivo para generar dicha

contribución relacionada con la reflexión y / o reverberación de la sala de una señal de salida binaural se muestra en la figura 3. El dispositivo 40 comprende el generador de mezcla descendente 42 y un procesador de sala 44 conectados en serie entre sí con el procesador de sala 44 siguiendo el generador de mezcla descendente 42. El dispositivo 40 puede estar conectado entre la entrada del dispositivo de la figura 1 en la que la señal multi-canal 18 se entra, y la salida de la señal de salida binaural la contribución del canal izquierdo 46a del procesador de local 44 se añade a la salida 22a, y la salida del canal derecho 46b del procesador de local 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 están configurados para generar el canal izquierdo 46a y el canal derecho 46b de las contribuciones relacionadas con la reflexión y/o reverberación de sala de la señal binaural que modelizan la reflexión y/o reverberación de sala a partir de la señal mono o estéreo 48.

[0031] La idea que subyace en el procesador de local 44 es que la reflexión/reverberación de sala que ocurre en, por ejemplo, una sala, puede ser modelizada de una manera transparente para el oyente, a partir de una mezcla descendente tal como una simple suma de los canales de la señal multi-canal 18. Puesto que las reflexiones/reverberación de sala se producen más tarde que los sonidos que viajan a lo largo de la trayectoria directa 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 sustituye, 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, restringirse a modelizar el camino directo y la reflexión y atenuaciones que ocurren en la cabeza, oídos, y hombros del oyente, lo que permite así 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.

[0032] Las figuras 4a y 4b muestran posibles implementaciones para la estructura interna del procesador de sala. De acuerdo con la figura 1a, el procesador de sala 44 se alimenta con una señal de mezcla descendente mono 48 y comprende dos filtros de reverberación 50a y 50b. Análogamente a los filtros direccionales, los filtros de reverberación 50a y 50b pueden implementarse para operar en el dominio del tiempo o en el dominio de la frecuencia. Las entradas de ambos reciben la señal de mezcla descendente mono 48. La salida del filtro de reverberación 50a proporciona la contribución de salida del canal izquierdo 46a, mientras que el filtro de reverberación 50b emite la contribución de salida del canal izquierdo 46b. La figura 4b muestra un ejemplo de la estructura interna del procesador de sala 44, en el caso de que el procesador de sala 44 está provisto de una señal de mezcla descendente estéreo 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 las entradas de los filtros de reverberación 50c y 50d están conectadas al otro canal 48b de la mezcla descendente estéreo 48. Las salidas de filtros de reverberación 50a y 50c están conectadas a la entrada de un sumador 52a, cuya salida proporciona la contribución de canal izquierdo 46a. Las salidas de los filtros de reverberación 50b y 50d están conectadas a las entradas de un sumador adicional 52b, cuya salida proporciona la contribución de canal derecho 46b.

[0033] El generador de mezcla descendente 42 de la realización de la figura 3 está configurados para formar la mezcla descendente mono o estéreo 48, de modo que la pluralidad de canales contribuya a la mezcla descendente mono o estéreo en un nivel que difiere entre al menos dos canales de la señal multi-canal 18. Por esta medida, se puede impedir o promover que 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 estén sujetos a procesamiento de sala, evitando de ese modo un sonido poco natural.

[0034] Por ejemplo, el generador de mezcla descendente 42 de la figura 3 puede estar 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 de mezcla descendente mono o estéreo 48 según un nivel reducido con respecto a los otros canales de la señal multi-canal 18. Por ejemplo, la cantidad de reducción de nivel puede estar entre 3 dB y 12 dB. La reducción del nivel se puede distribuir de manera uniforme en el intervalo espectral efectivo de los canales de la señal multicanal 18, o puede ser dependiente de la frecuencia tal como concentrada en una parte espectral específica, tal como la parte espectral normalmente ocupada por las señales de voz. La cantidad de reducción de nivel relativo a los otros canales puede ser la misma para todos los otros canales. Es decir, los otros canales se pueden mezclar en la señal de mezcla descendente 48 en el mismo nivel. Como alternativa, los otros canales se pueden mezclar en la señal de mezcla descendente 48 en un nivel desigual. Entonces, la cantidad de la reducción de nivel con respecto a los otros canales se puede medir frente al valor medio de los otros canales o el valor medio de todos los canales, incluyendo la reducción de uno. Si es así, la desviación estándar de las ponderaciones de mezcla de los otros canales o la desviación estándar de las ponderaciones de todos los canales de mezcla puede ser menor que 66% de la reducción del nivel de la ponderación de mezcla del canal en relación con el valor medio mencionado.

[0035] El efecto de la reducción del nivel con respecto al canal central es que la señal de salida binaural obtenida a través de las contribuciones 56a y 56b es - al menos en algunas circunstancias que se discuten en más detalle a continuación - percibida 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 multicanal 18, con el valor de ponderación asociado con el canal central reducido con relación a los valores de ponderación de los otros canales.

5 **[0036]** La reducción del nivel del canal central es especialmente ventajosa en partes con voz de diálogos de películas o música. La mejora de impresión de audio obtenida en estas partes voz sobre compensa 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 el nivel de reducción está apagado, y un modo en el que el nivel de reducción está encendido. En otras palabras, el generador de mezcla descendente 42 puede estar configurado para variar la cantidad de reducción del nivel de una manera variable en el tiempo. La variación puede ser de una naturaleza binaria o analógica, entre cero y un valor máximo. El generador de mezcla descendente 42 puede estar configurado para realizar el cambio de modo o reducción del nivel de cantidad de variación en función de la información contenida dentro de la señal multi-canal 18. Por ejemplo, el generador de mezcla descendente 42 puede estar 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, que son de al menos escala ordinal, de tramas consecutivas del canal central. Por ejemplo, el generador de mezcla descendente 42 detecta la presencia de 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 las fases de voz dentro del canal central por el generador de mezcla descendente 42 no es la única manera de hacer la conmutación del modo antes mencionada de variación de la reducción de nivel dependiente del tiempo. Por ejemplo, la señal 18 multicanal podría tener información lateral asociada con la misma, que está especialmente diseñada para distinguir entre las fases de voz y las fases que no son de voz, o medir el contenido de voz cuantitativamente. En este caso, el generador de mezcla descendente 42 operaría en respuesta a esta información lateral. Otra posibilidad sería que el generador de mezcla descendente 42 realice la conmutación de modo o de las variaciones cantidad reducción de nivel antes mencionadas en función de una comparación entre, por ejemplo, los niveles actuales del canal central, el canal izquierdo y el canal derecho. En caso de que el canal central sea mayor que los canales izquierdo y derecho, ya sea individualmente o en relación con la suma de los mismos, por más de una cierta relación de umbral, entonces el generador de mezcla descendente 42 puede asumir que una fase de voz está presente actualmente y actuar en consecuencia, es decir, mediante la realización de la reducción del nivel. Del mismo modo, el generador de mezcla descendente 42 puede utilizar las diferencias de nivel entre canales central, izquierdo y derecho con el fin de darse cuenta de las dependencias antes mencionadas.

35 **[0037]** 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 multicanal 18. Esto se muestra en la figura 5. La figura 5 muestra un ejemplo del generador de mezcla descendente 42 en el caso de que la señal multicanal 18 representa una pluralidad de canales mediante el uso de codificación de audio espacial, es decir, mediante el uso de una señal de mezcla descendente 62 en la que se han mezclado hacia abajo la pluralidad de canales y los parámetros espaciales 64 que describen la imagen espacial de la pluralidad de canales. Opcionalmente, la señal multi-canal 18 también puede comprender mezclar hacia abajo información que describe las relaciones con las que los canales individuales han sido mezclados en la señal de mezcla descendente 62, o los canales individuales de la señal de mezcla descendente 62, ya que el canal de mezclado descendente 62 puede ser por ejemplo una señal de mezcla descendente normal 62 o una señal de mezcla descendente estéreo 62. El generador de mezcla descendente 42 de la figura 5 comprende un descodificador 64 y un mezclador 66. El descodificador 64 descodifica, según descodificación de audio espacial, la señal multi-canal 18 con la finalidad de obtener la pluralidad de canales que incluye, 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 variada de reducción de nivel, tal como se ha mencionado más arriba. El parámetro espacial 64 utilizado por el mezclador 66 puede, por ejemplo, ser coeficientes de predicción de canal que describen como el canal central 66, un canal izquierdo o el canal derecho se pueden derivar de la señal de mezcla descendente 62, en el que mezclador 66 también puede utilizar parámetros de coherencia/correlación cruzada entre canales que representa 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 los canales frontal izquierdo y posterior izquierdo, y canales frontal derecho y posterior derecho, respectivamente. Por ejemplo, el canal central se puede mezclar con una relación fija en el canal izquierdo y el canal derecho de la señal de mezcla descendente estéreo 62 antes mencionados. En este caso, dos coeficientes de predicción de canal son suficientes para determinar cómo los canales central, izquierdo y derecho se pueden derivar de una combinación lineal respectiva de los dos canales de la señal de mezcla descendente estéreo 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 fases de voz y fases que no son de voz.

65 **[0038]** Aunque la reducción de nivel con respecto al canal central ha sido descrita con el fin de poner un ejemplo de la suma ponderada de la pluralidad de canales de modo que estos contribuyen a la mezcla descendente mono o estéreo en un nivel que difiere entre al menos dos canales de la señal multi-canal 18, también hay otros ejemplos en

los que ventajosamente otros canales reduce o aumentan 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 debe/deben o no debe/ no deben ser sometidos al procesamiento de sala en el mismo nivel que otros contenidos en la señal multi-canal, pero en un menor / mayor nivel.

5
 [0039] La figura 5 se explicó de manera general, con respecto a una posibilidad para la representación de la pluralidad de canales de entrada mediante 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 las siguientes formas de realización descritas con respecto a las figuras 10 al 13. La figura 6 muestra la señal de mezcla descendente 62 espectralmente descompuesta en una pluralidad de subbandas 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 como se indica por la flecha de dominio de la frecuencia 84. La extensión a lo largo de la dirección horizontal debe indicar 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 en la que las sub-bandas 82 son muestreadas por los valores de muestra 88 puede ser definida mediante las ranuras del banco de filtros 90. Por lo tanto, las ranuras de tiempo 90 y las subbandas 82 definen alguna resolución o rejilla de tiempo / frecuencia. Una rejilla de tiempo / frecuencia más gruesa se define mediante la unión de valores de muestra vecinos 88 o baldosas de tiempo / frecuencia 92, tal como se indica por las líneas de trazos en la figura 6, definiendo estas baldosas la resolución/ rejilla de parámetro de tiempo/ frecuencia. Los parámetros espaciales mencionados 62 se definen en esa resolución de parámetro de tiempo/ frecuencia 92. La resolución de parámetro de tiempo/ frecuencia 92 puede cambiar en el tiempo. Para este fin, la señal multi-canal 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 recibe 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 tal como se muestra 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 puede haber dos coeficientes de predicción de canal que revelan cómo, con respecto a las respectivas baldosas de tiempo / frecuencia 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 el azulejo 92 que indica las similitudes ICC entre los canales izquierdo y derecho a derivar de la señal de mezcla descendente estéreo 62, en el que un canal ha sido completamente mezclado en un canal de la señal de mezcla descendente estéreo 62, mientras que el otro ha sido completamente mezclado en el otro canal de la señal de mezcla descendente estéreo 62. Sin embargo, puede haber además un parámetro de diferencia de nivel de canal (CLD) en cada baldosa 92 que indica la diferencia de nivel entre los canales izquierdo y derecho que se acaban de mencionar. Una cuantificación no uniforme en una escala logarítmica se puede aplicar 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, puede haber otros parámetros dentro del parámetro espacial 64. Estos parámetros pueden, entre otras cosas, definir CLD y ICC relativos a los canales que sirvieron para formar, por mezclado, los canales izquierdo y derecho que se acaban de mencionar, tales como el posterior izquierdo, frontal izquierdo, posterior derecho, y los canales frontales derechos.

45
 [0040] Cabe señalar que las realizaciones mencionadas anteriormente pueden combinarse entre sí. Algunas posibilidades de combinación ya se han mencionado anteriormente. Se mencionarán otras posibilidades a continuación con respecto a las realizaciones de las figuras 7 a 13. Además, las realizaciones mencionadas anteriormente de las figuras 1 y 5 suponen 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, las HRTFs modificadas tal como se han derivado mediante el dispositivo de la figura 2 se pueden usar 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 operar sobre una señal de mezcla descendente tal como la señal de mezcla descendente 62 mostrada en la figura 5, que representa la pluralidad de canales 18a-18d, combinando adecuadamente los parámetros espaciales y las HRTFs modificadas en la resolución de parámetro de tiempo/ frecuencia 92, y aplicando en consecuencia los coeficientes de combinación lineal obtenidos con el fin de formar señales binaurales 22a y 22b.

55
 [0041] De manera similar, el generador de mezcla descendente 42 puede estar configurado para combinar adecuadamente los parámetros espaciales 64 y la cantidad de reducción de nivel a conseguir para el canal central con la finalidad de derivar la mezcla descendente mono o estéreo 48 deseada para el procesador de local 44. La figura 7 muestra un generador de señal de salida binaural según una realización. Un generador que se indica de manera general mediante el signo de referencia 100 comprende un descodificador multi-canal 102, una salida binaural 104, y dos caminos que se extienden entre la salida del descodificador multi-canal 102 y la salida binaural 104, respectivamente, a saber un camino directo 106 y un camino de reverberación 108. En el camino directo, los filtros direccionales 110 están conectados a la salida del descodificador multi-canal 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. Las salidas sumadas de los sumadores

primero y segundo 112 y 114 representan la antes mencionada contribución de camino directo de la señal de salida binaural 22a y 22b. Los sumadores 116 y 118 se proporcionan para combinar las señales de contribución 22a y 22b con las señales de contribución binaural 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 local 122 están conectados en serie entre la salida del descodificador multi-canal 102 y la entrada respectiva de los sumadores 16 y 118, cuyas salidas definen la salida de la señal de salida binaural en la salida 104.

[0042] Con el fin de facilitar la comprensión de la siguiente descripción del dispositivo de la figura 7, los signos de referencia utilizados en las figuras 1 a 6 se han utilizado parcialmente con el fin de indicar elementos de la figura 7, que corresponden a aquellos, o asumen la funcionalidad de, elementos que aparecen 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 denota un reductor de correlación, en lo que sigue. Sin embargo, como quedó claro de lo anterior, las formas de realización descritas a continuación son fácilmente transferibles a los casos en que el reductor de similitud realiza una reducción en la similitud diferente que aquella en términos de correlación. Además, las formas de realización descritas a continuación se han redactado suponiendo que el mezclador para generar la mezcla descendente para el procesamiento de sala genera un nivel de reducción del canal central aunque, como se ha descrito anteriormente, una transferencia a formas de realización alternativas sería fácilmente realizable.

[0043] El dispositivo de la figura 7 utiliza un flujo de señal para la generación de una salida de auriculares en la salida 104 de una señal multicanal decodificada 124. El multicanal decodificada 124 se deriva por el decodificador multi-canal 102 desde una entrada de flujo de bits a una entrada de flujo de bits 126, tal como, por ejemplo, mediante la decodificación de audio espacial. Después de la decodificación, cada señal o canal de la señal multicanal decodificada 124 es filtrada por un par de filtros direccionales 110. Por ejemplo, el primer canal (superior) de la señal multicanal decodificada 124 es filtrada por filtros direccionales 20 DirFilter (1, L) y DirFilter (1, R), y un segundo (segundo desde la parte superior) de la señal o canal se filtra mediante el filtro direccional DirFilter (2, L) y DirFilter (2, R), y así sucesivamente. Estos filtros 110 pueden modelizar la transmisión acústica de una fuente de sonido virtual en un espacio al canal del oído de un oyente, una así llamada función de transferencia de sala binaural (BRTF). Pueden realizar modificaciones de tiempo, nivel y espectrales, y pueden en parte también modelizar reflexión y reverberación de sala. Los filtros direccionales 110 pueden ser implementados en los dominios de tiempo o de la frecuencia. Puesto que se necesitan muchos filtros 110 (Nx2, siendo N el número de canales decodificados), estos filtros direccionales podrían, si se debe modelizar la reflexión y la reverberación de sala completamente, ser bastante largos, es decir, 20000 intervalos de filtro a 44,1 kHz, en cuyo caso el proceso de filtrado sería computacionalmente exigente. El filtro direccional 110 se reduce ventajosamente al mínimo, las denominadas funciones de transferencia relacionadas con la cabeza (HRTF) y el bloque de procesamiento común 122 se utilizan para modelizar las reflexiones y reverberaciones de sala. El módulo de procesamiento de sala 122 puede implementar un algoritmo de reverberación en un dominio del tiempo o de la frecuencia y puede operar desde un uno o dos canales de señal de entrada 48, que se calculan a partir de la señal de entrada multicanal decodificada 124 por una matriz de mezcla dentro del mezclador 120. 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 localizar los sonidos, sobre todo con respecto a la distancia y la externalización – los sonidos significativos se perciben fuera de la cabeza del oyente.

[0044] Por lo general, se produce el sonido multicanal de tal manera que la energía sonora dominante 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 reverberante poco natural y espectralmente desigual. Por lo tanto, según la realización de la figura 7, el canal central se alimenta al módulo de procesamiento de sala 122 con una reducción de nivel significativo, tal como atenuada en 6 dB, nivel de reducción 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, en la que 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.

[0045] La figura 8 muestra otro generador de señal de salida binaural según una realización adicional. 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 las realizaciones de 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 del nivel dentro del mezclador 122 es opcional en el caso de la figura 8. A diferencia de la figura 7, sin embargo, los decorrelacionadores están conectados entre cada par de filtros direccionales 110 y la salida del decodificador 102 para el canal asociado de la señal multicanal decodificada 124, respectivamente. Los decorrelacionadores se indican con signos de referencia 1421, 1422, y así sucesivamente. Los decorrelacionadores 1421-1424 actúan como el reductor de correlación 12 que se indica en la figura 1. Aunque se muestra en la figura 8, no es necesario que se

proporcione un decorrelacionador 1421-1424 para cada uno de los canales de la señal multicanal decodificada 124. Más bien, un decorrelacionador sería suficiente. El decorrelacionador 142 podría ser simplemente un retraso. Preferentemente, la cantidad de retardo causado por cada uno de los retrasos 1421-1424 sería diferente el uno del otro. Otra posibilidad sería que los decorrelacionadores 1421-1424 fueran filtros pasa-todo, es decir, filtros que tienen una función de transferencia de una magnitud constantemente igual a uno, pero que sin embargo, cambien las fases de los 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, el decorrelacionador 1421-1424 podría ser implementado como filtros FIR, o similares.

[0046] Por lo tanto, según la realización de la figura 8, los elementos 1421-1424, 110, 112, y 114 actúan de acuerdo con el dispositivo 10 de la figura 1.

[0047] 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 a continuación utilizando los mismos signos de referencia tal como se utilizan en la figura 7. De manera similar a la realización de la figura 8, la reducción del nivel de mezclador 122 es meramente opcional en el caso de la figura 9, y por lo tanto, el signo de referencia es 40' en la figura 9 en lugar de '40, como fue el caso en la figura 7. La forma de realización 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 de múltiples canales con los filtros direccionales 110, las señales intermedias de dos canales de cada par de filtros se añaden mediante 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 la falta de una externalización. Esto es particularmente problemático para la correlación de los canales izquierdo, derecho y central dentro de la señal multicanal decodificada 124. Según la realización de la figura 9, los filtros direccionales están configurados para tener una salida decorrelacionada en la medida de lo posible. Para este fin, el dispositivo de la figura 9 comprende el dispositivo 30 para formar conjunto de HRTFs de inter-correlación decreciente a utilizar por los filtros direccionales 110 a partir de algún conjunto original de HRTFs. Tal como se ha descrito más arriba, el dispositivo 30 puede utilizar una, o una combinación de, las siguientes técnicas con respecto a las HRTFs del par de filtros direccionales asociados con uno o varios canales de la señal multicanal decodificada 124: retrasar el filtro direccional o el respectivo par de filtros direccionales tales como, por ejemplo, mediante el desplazamiento de su respuesta de impulso que se podría hacer, por ejemplo, mediante el desplazamiento de los intervalos de filtro; la modificación de la respuesta de fase de los respectivos filtros direccionales; y aplicar un filtro de decorrelación, tal como un filtro de pasa todo a los respectivos filtros direccionales del canal respectivo. Este filtro pasa-todo podría ser implementado como un filtro FIR.

[0048] Como se describió anteriormente, el dispositivo 30 podría operar en respuesta al cambio en la configuración del altavoz para el que está destinado el flujo de bits en la entrada de flujo de bits 126.

[0049] Las realizaciones de las figuras 7 a 9 se refieren a una señal multicanal decodificada. Las siguientes realizaciones se refieren a la decodificación multi-canal paramétrica para auriculares. En términos generales, la codificación de audio espacial es una técnica de compresión de múltiples canales que explota la irrelevancia entre canales de percepción de las señales de audio multicanal para lograr mayores tasas de compresión. Esto se puede capturar en términos de señales espaciales o parámetros espaciales, es decir, los 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 correlaciones/coherencia entre canales, y pueden ser representadas de una manera extremadamente compacta. El concepto de codificación de audio espacial ha sido adoptado por MPEG resultante en el estándar MPEG envolvente, es decir, ISO / IEC23003-1. También se pueden emplear parámetros espaciales tales como los empleados en la codificación de audio espacial para describir filtros direccionales. De esta manera, la etapa de decodificación de datos de audio espaciales y la aplicación de filtros direccionales se puede combinar para decodificar de manera eficiente y reproducir audio multi-canal en auriculares.

[0050] La estructura general de un decodificador de audio espacial para la salida de auriculares se da en la figura 10. El decodificador de la figura 10 se indica en general con el signo de referencia 200, y comprende un modificador de subbanda 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 señal multicanal antes mencionada 18 y representan la pluralidad de canales de la misma.

[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 de modificador de subbanda 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 por el dispositivo 30.

[0052] En la figura 10, se supone que la señal de mezcla descendente que ya se ha decodificado de antemano, incluyendo, por ejemplo, la codificación de entropía. El decodificador de audio espacial binaural se alimenta con la señal de mezcla descendente 204. El convertidor de parámetros 214 utiliza los parámetros espaciales 206 y 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 unidad de matrizado 210 en forma de una matriz dos por dos (en caso de una señal de mezcla descendente estéreo) y en forma de una matriz de uno por dos (en el caso de una señal de mezcla descendente mono 204), en el dominio de frecuencia, a los valores espectrales de salida 88 por banco de filtros de análisis 208 (véase la figura 6). En otras palabras, los parámetros binaurales 218 varían en la resolución de parámetro de tiempo/ frecuencia 92 mostrada en la figura 6 y se aplican a cada valor de la muestra 88. Se puede usar interpolación para suavizar los coeficientes de matriz y los parámetros binaurales 218, respectivamente, a partir del parámetro más grueso de dominio de tiempo / frecuencia 92 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 la unidad 210 da como resultado dos valores de muestra por cada par de valores de la muestra del canal izquierdo de la señal de mezcla descendente 204 y el valor de la muestra correspondiente del canal derecho de la señal de mezcla descendente 204. Los dos valores de las muestras 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 mono 204, 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 binaurales 218 definen la operación de matriz que va desde el 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 de HRTF modificados. Por lo tanto, se elimina la correlación de los canales de entrada de la señal de múltiples canales 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 la misma 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 mediante 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 trataron separadamente. Si acaso, estos efectos tienen que ser considerados en las HRTFs 216. La figura 11 muestra un generador de señal de salida binaural que combina un decodificador de audio espacial binaural 200' con el procesamiento de reflexión / reverberación de sala independiente. 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. Opcionalmente, sin embargo, el descodificador de audio espacial binaural 200' de la figura 11 puede ser el que se muestra en la figura 10. En cualquier caso, el generador de señal de salida binaural de la figura 11 que se indica de manera general mediante el signo de referencia 230, comprende además del decodificador binaural espacial 200', un decodificador de audio de mezcla descendente 232, un modificador de la subbanda de audio espacial modificado 234, un procesador de local 122, y dos sumadores 116 y 118. El decodificador de audio de mezcla descendente 232 está conectado entre una entrada de flujo de bits 126 y un modificador de subbanda de audio espacial binaural 202 del decodificador de audio espacial binaural 200'. El decodificador de audio de mezcla descendente 232 está configurado para descodificar el flujo de bits de entrada en la entrada 126 para derivar la señal de mezcla descendente 214 y los parámetros espaciales 206. Ambos, el modificador de sub-banda de audio espacial binaural 202, así como el modificador de subbanda de audio espacial modificado 234 se alimentan con una señal de mezcla descendente 204, además de los parámetros espaciales 206. El modificador de subbanda audio espacial modificada 234 calcula a partir de la señal de mezcla descendente 204 – mediante el uso de los parámetros espaciales 206, así como los parámetros modificados 236 que reflejan la cantidad antes mencionada de reducción de nivel del canal central - la mezcla descendente mono o estéreo 48 que sirve como entrada al procesador de local 122. La salida de contribuciones tanto por el modificador de subbanda de audio espacial binaural 202 y el procesador de sala 122, se suman a nivel de canal en sumadores 116 y 118 para dar como salida la señal de salida binaural en la salida 238.

[0055] La figura 12 muestra un diagrama de bloques que ilustra la funcionalidad del decodificador de audio binaural 200' de la figura 11. Cabe señalar que la figura 12 no muestra la estructura interna real del decodificador de audio espacial binaural 200' de la figura 11, sino que ilustra las modificaciones de la señal obtenida por el decodificador de audio binaural espacial 200'. Se recuerda que la estructura interna del decodificador de audio binaural espacial 200' generalmente se ajusta a la estructura que se muestra en la figura 10, con la excepción de que el dispositivo 30 se puede omitir, en el caso de que este mismo esté operando con las HRTFs originales. Además, la figura 12 muestra la funcionalidad del decodificador de audio binaural espacial 200' a modo de ejemplo para el caso de que sólo se utilizan tres canales representados por la señal multicanal 18 en el decodificador de audio binaural espacial 200' con el fin de formar la señal de salida binaural 208. En particular, se utiliza una "2 a 3", es decir caja TTT, para derivar un canal central 242, un canal derecho 244, y un canal izquierdo 246 a partir de los dos canales de la mezcla descendente estéreo 204. En otras palabras, la figura 12 asume 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 antes mencionados. La reducción de correlación se logra mediante tres descorrelacionadores, indicados DelayL, DelayR, y DelayC en la figura 12. Corresponden a la descorrelación

introducido en el caso de, por ejemplo, las figuras 1 y 7. Sin embargo, se recuerda de nuevo que la figura 12 simplemente muestra las modificaciones de la señal obtenidas por el decodificador de audio binaural espacial 200', aunque la estructura real corresponde a la mostrada en la figura 10. Así, aunque los retrasos que forman el reductor de correlación 12 se muestran como funciones separadas en relación con las HRTF que forman los filtros direccionales 14, la existencia de los retrasos en el reductor de correlación 12 puede ser vista como una modificación de los parámetros de HRTF que forman las HRTF originales de los filtros direccionales 14 de la figura 12. En primer lugar, la figura 12 simplemente muestra que el decodificador de audio espacial binaural 200' descorrelaciona los canales para la reproducción por auriculares. La decorrelación se consigue por medios sencillos, a saber, mediante la adición de un bloque de retardo en el procesamiento paramétrico para la matriz M y el decodificador de audio binaural espacial 200'. De este modo, el decodificador de audio binaural espacial 200' podrá aplicar las siguientes modificaciones a los canales individuales, a saber retrasar el canal central preferentemente al menos una muestra, retrasar el canal central por diferentes intervalos en cada banda de frecuencia, retrasar canales izquierdo y derecho preferentemente al menos una muestra y / o retrasar los canales izquierdo y derecho por diferentes intervalos en cada banda de frecuencia.

[0056] La figura 13 muestra un ejemplo de una estructura del modificador de audio espacial modificado subbanda de la figura 11. El modificador de sub-banda 234 de la figura 13 comprende una de dos a tres o caja TTT 262, etapas de ponderación 264a-264E, primeros sumadores 266a y 266b, segundos sumadores 268a y 268b, una entrada para la mezcla descendente estéreo 204, una entrada para los parámetros espaciales 206, una entrada adicional para para una señal residual 270 y una salida para la mezcla descendente 48 destinada a ser procesada por el procesador de local, y que es, de acuerdo con la figura 13, una señal estéreo.

[0057] Como la figura 13 define en un sentido estructural una realización para el modificador de subbanda de audio espacial modificado 234, la caja TTT 262 de la figura 13 simplemente 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 recuerda una vez más que en el caso de la figura 12, los canales 242 a 246 en realidad no se calculan. Más bien, el modificador de subbanda de audio espacial binaural modifica la matriz M de tal manera que la señal de mezcla descendente estéreo 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. Opcionalmente, como se muestra en la figura 13, la caja TTT 262 puede utilizar una señal residual 270 que refleja la predicción residual al reconstruir canales 242-246 a partir de la mezcla descendente estéreo 204 y los parámetros espaciales 206, que como se indicó anteriormente, comprenden los coeficientes de predicción de canal y, opcionalmente, valores ICC. Los primeros sumadores 266a están configurados para sumar canales 242-246 para formar el canal izquierdo de la mezcla descendente estéreo 48. En particular, se forma una suma ponderada mediante 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 podrían aplicarse al canal respectivo 246 a 242, un valor de ponderación respectivo EQLL, EQRL y EQCL. Del mismo modo, los sumadores 268a y 268b forman una suma ponderada de los canales 246 to 242 con etapas de ponderación 264b, 264d, y 264e que forman 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 264a-264E, como se describió anteriormente, se seleccionaron de tal manera que se consigue la reducción del nivel de canal central descrita anteriormente en la mezcla descendente estéreo 48 lo que da como resultado, tal como se describió anteriormente, 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 multicanal para poder 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 subbanda espacial de audio 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 la los canales central, izquierdo y derecho 242-246 pueden ser decodificados por el decodificador de audio 232 de mezcla descendente, aunque no se muestra en la figura 11. Los parámetros EQ o valores de ponderación aplicados por las etapas de ponderación 264a-264e puede ser de valor real para los canales central, izquierdo y derecho 242-246. Un único conjunto de parámetros para el canal central 242 puede ser almacenado y aplicado, y el canal central es, de acuerdo con la figura 13, a modo de ejemplo, mezclado de la misma manera tanto en la salida izquierda como la derecha en la mezcla descendente estéreo 48.

[0060] Los parámetros EQ 270 suministrados al modificador de subbanda de audio espacial modificada 234 pueden 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 audio espacial binaural de sub-banda 202 debe aumentarse en consecuencia.

- 5 [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 de descendente izquierdo y derecho dentro de la caja TTT. Si el nivel de centro se reduce, pueden ser audibles artefactos en los canales izquierdo y derecho. Por lo tanto, la reducción del nivel de centro en la etapa EQ es un compromiso entre la supresión y los artefactos. Encontrar un ajuste fijo de parámetros EQ es posible, pero puede no ser óptimo para todas las señales. En consecuencia, de acuerdo con una forma de realización, se puede utilizar un algoritmo o módulo de adaptación 274 para controlar la cantidad de reducción del nivel de centro por uno, o una combinación de los siguientes parámetros:
- 10 [0062] Los parámetros espaciales 206 utilizados para decodificar el canal central 242 desde el canal de mezclado descendente izquierdo y derecho 204 dentro de la caja TTT 262 se pueden utilizar como indicado por la línea discontinua 276.
- 15 [0063] El nivel de los canales central, izquierdo y derecho se puede utilizar tal como se indica mediante la línea discontinua 278.
- 20 [0064] Las diferencias de nivel entre canales central, izquierdo y derecho 242 a 246 pueden utilizarse como también indica la línea de trazos 278.
- [0065] Se puede utilizar la salida de un algoritmo de detección de tipo único, tal como un detector de actividad de voz, como también se indica por la línea de trazos 278.
- 25 [0066] Por último, se puede utilizar una estática de metadatos dinámica que describa el contenido de audio con el fin de determinar la cantidad de reducción de nivel central como indica la línea discontinua 280.
- 30 [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 un ASIC, una sub-rutina de un código de programa o una parte de una lógica programable programada.
- 35 [0068] La señal de audio codificada de la invención puede ser almacenada en un medio de almacenamiento digital o puede transmitirse en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable, tales como Internet.
- 40 [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, 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 el mismo, que cooperan (o son capaces de cooperar) con un sistema informático programable de tal manera que se lleve a cabo el procedimiento respectivo.
- 45 [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 tal manera que se lleve a cabo uno de los procedimientos descritos en este documento.
- 50 [0071] En general, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, el código de programa siendo operativo para 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.
- 55 [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.
- 60 [0073] En otras palabras, una forma de 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.
- [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, registrado en el mismo, el programa de ordenador para realizar uno de los procedimientos descritos en este documento.
- 65 [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.

5 **[0076]** Una forma de realización comprende además un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado para o adaptado para llevar a cabo uno de los procedimientos descritos en este documento.

10 **[0077]** Una forma de realización comprende además un ordenador que tiene instalado el programa informático para realizar uno de los procedimientos descritos en este documento.

15 **[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 de las funcionalidades de los procedimientos descritos aquí. En algunas realizaciones, una matriz de puertas programable en campo podrá cooperar con un microprocesador para llevar a cabo uno de los procedimientos descritos en este documento. Generalmente, los procedimientos se realizan preferiblemente por cualquier aparato de hardware.

20 **[0079]** Las realizaciones descritas anteriormente 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.

25 **[0080]** Por lo tanto, entre otras, las realizaciones de más arriba describían un dispositivo para generar una señal binaural a partir de una señal multi-canal que representa una pluralidad de canales y destinado a la reproducción por una configuración de altavoces que tienen una posición de fuente de sonido virtual asociada a cada canal, que comprende:

30 un reductor de similitud 12 para procesar de manera diferente, y por lo tanto reducir una similitud entre, al menos uno de un canal izquierdo y derecho de la pluralidad de canales, un canal frontal y un canal posterior de la pluralidad de canales, y un canal central y canal no central de la pluralidad de canales, con la finalidad de obtener un conjunto de inter-similitud reducida 20 de canales;

una pluralidad 14 de filtros direccionales para modelizar una transmisión acústica de uno respectivo del conjunto de inter-similitud reducida 20 de canales a partir de una posición de fuente de sonido virtual asociada con el canal respectivo de conjunto de canales de inter-similitud reducida a un canal auditivo respectivo de un oyente;

35 un primer mezclador 16a para mezclar salidas de los filtros direccionales que modelizan la transmisión acústica al primer canal auditivo del oyente para obtener un primer canal 22a de la señal binaural; y

40 un segundo mezclador 16b para mezclar salidas de los filtros direccionales que modelizan la transmisión acústica al segundo canal auditivo del oyente para obtener un segundo canal 22b de la señal binaural. El reductor de correlación 12 puede estar configurado para realizar el procesamiento diferente provocando un retardo relativo entre, o realizar - en un sentido variable espectralmente - modificación de fase de manera diferente entre, el al menos uno de los canales izquierdo y derecho de la pluralidad de canales, los canales frontal y posterior de la pluralidad de canales, y los canales central y no central de la pluralidad de canales, y/o realizar - en un sentido variable espectralmente - una modificación de magnitud de manera diferente entre, el al menos uno de los canales izquierdo y derecho de la pluralidad de canales, los canales frontal y posterior de la pluralidad de canales, y los canales central y no central de la pluralidad de canales.

45 **[0081]** Entre las realizaciones anteriores, también se describe un dispositivo para generar una señal binaural a partir de una señal multi-canal que representa una pluralidad de canales y destinado a la reproducción por una configuración de altavoces que tienen una posición de fuente de sonido virtual asociada a cada canal, que comprende:

50 un reductor de similitud 12 para realizar - en un sentido variable espectralmente - una modificación de fase y/o magnitud de manera diferente entre al menos dos canales de la pluralidad de canales, con la finalidad de obtener un conjunto de inter-similitud reducida 20 de canales;

55 una pluralidad 14 de filtros direccionales para modelizar una transmisión acústica de una respectiva del conjunto de inter-similitud reducida 20 de canales desde una posición de fuente de sonido virtual asociada con el canal respectivo del conjunto de canales de inter-similitud reducida a un canal auditivo respectivo de un oyente;

un primer mezclador 16a para mezclar salidas de los filtros direccionales que modelizan la transmisión acústica al primer canal auditivo del oyente para obtener un primer canal 22a de la señal binaural; y

60 un segundo mezclador 16b para mezclar salidas de los filtros direccionales que modelizan la transmisión acústica al segundo canal auditivo del oyente para obtener un segundo canal 22b de la señal binaural.

[0082] Además, las realizaciones de más arriba también describen un dispositivo para formar un conjunto de inter-similitud decreciente de funciones de transferencia relacionadas con la cabeza para modelizar una transmisión acústica de una pluralidad de canales a partir de una posición de fuente de sonido virtual asociada con el canal respectivo a canales auditivos de un oyente, comprendiendo el dispositivo:

65

5 un proveedor de HRTF 32 para proporcionar una pluralidad original de HRTFs; y un procesador de HRTF 34 para provocar que respuestas de impulso de las HRTFs que modelizan las transmisiones acústicas de un par de canales predeterminado se retrasen entre sí, o de manera diferente modifique - en un sentido variable espectralmente – sus respuestas de fase y/o magnitud, siendo el par de canales uno de un canal izquierdo o derecho de la pluralidad de canales, un canal frontal y un canal posterior de la pluralidad de canales, y un canal central y canal no central de la pluralidad de canales. El proveedor de HRTF 32 puede estar configurado para proporcionar la pluralidad original de HRTFs a partir de las posiciones de fuente de sonido virtual y parámetros HRTF. El procesador de HRTF 34 puede estar configurado para filtrar de manera diferente por pasa todo las respuestas de impulso del par de canales predeterminado.

10

REIVINDICACIONES

- 5 **1.** Dispositivo para generar una contribución relativa a la reflexión/ reverberación de local (46a, 46b) de una señal binaural a partir de una señal multi-canal (18) que representa una pluralidad de canales y que está destinado para la reproducción por una configuración de altavoces que tiene una posición de fuente de sonido virtual asociada a cada canal, que comprende:
- 10 un generador de mezcla descendente (42) que forma una mezcla descendente mono o estéreo de los canales de la señal multi-canal; y
- 15 un procesador de local (44) para generar la contribución relativa a la reflexión/ reverberación de local (46a, 46b) de la señal binaural modelizando las reflexiones/ reverberaciones de local a partir de la señal mono o estéreo, en el que el generador de mezcla descendente (42) está configurado para formar la mezcla descendente mono o estéreo de modo que la pluralidad de canales contribuyen a la mezcla descendente mono o estéreo en un nivel que difiere entre al menos dos canales de la señal multi-canal, en el que el generador de mezcla descendente (42) está configurado para formar la mezcla descendente mono o estéreo (48) de modo que un canal central de la pluralidad de canales contribuye a la mezcla descendente mono o estéreo según un nivel reducido con respecto a los otros canales de la señal multi-canal (18).
- 20 **2.** Dispositivo según la reivindicación 1, en el que el generador de mezcla descendente (42) está configurado para reconstruir, mediante codificación de audio espacial, la pluralidad de canales a partir de una señal de mezcla descendente y parámetros espaciales asociados que describen diferencias de nivel, diferencias de fase, diferencias de tiempo y/o medidas de correlación entre las pluralidades de canales.
- 25 **3.** Dispositivo según la reivindicación 2, en el que el generador de mezcla descendente (42) está configurado para realizar la formación de modo que una cantidad de reducción de nivel de un primero de los al menos dos canales relativos a un segundo de los al menos dos canales depende de los parámetros espaciales.
- 30 **4.** Dispositivo según la reivindicación 2, en el que el generador de mezcla descendente (42) está configurado para reconstruir, mediante codificación de audio espacial, la pluralidad de canales a partir de una señal de mezcla descendente estéreo, coeficientes de predicción de canal que describen como los canales de la señal de mezcla descendente estéreo se deben combinar linealmente para predecir un triplete de canales central, derecho e izquierdo, y una señal residual (270) que refleja una predicción residual cuando se predice el triplete.
- 35 **5.** Dispositivo según cualquiera de las reivindicaciones 1 a 4, en el que el generador de mezcla descendente (42) está configurado para realizar la formación de modo que una cantidad de reducción de nivel de un primero de los al menos dos canales relativos a un segundo de los al menos dos canales depende de una diferencia de nivel y/o una correlación entre canales individuales de la pluralidad de canales.
- 40 **6.** Dispositivo según la reivindicación 5, en el que el generador de mezcla descendente (42) está configurado para obtener la diferencia de nivel y/o la correlación entre canales individuales de la pluralidad de canales a partir de parámetros espaciales que acompañan una señal de mezcla descendente que representan conjuntamente la pluralidad de canales.
- 45 **7.** Dispositivo según cualquiera de las reivindicaciones 1 a 4, en el que el generador de mezcla descendente (42) está configurado para realizar la formación de modo que una cantidad de reducción de nivel de un primero de los al menos dos canales relativos a un segundo de los al menos dos canales varía en el tiempo tal como se indica mediante un indicador de variación temporal transmitido dentro de información lateral de la señal multi-canal.
- 50 **8.** Dispositivo según la reivindicación 1, comprendiendo el dispositivo además:
- un detector de tipo de señal para detectar fases de habla y de no - habla dentro de la señal multi-canal, en el que el generador de mezcla descendente está configurado para realizar la formación de modo que una cantidad de reducción de nivel es mayor durante fases de habla que durante fases de no habla.
- 55 **9.** Procedimiento para generar una contribución relativa a la reflexión/ reverberación de local de una señal binaural a partir de una señal multi-canal que representa una pluralidad de canales y que está destinado para la reproducción por una configuración de altavoces que tienen una posición de fuente de sonido virtual asociada a cada canal, que comprende:
- 60 formar una mezcla descendente mono o estéreo de los canales de la señal multi-canal; y
- generar la contribución relativa a la reflexión/ reverberación de local de la señal binaural modelizando reflexiones/ reverberaciones de local a partir de la señal mono o estéreo, en el que el generador de mezcla descendente está configurado para formar la mezcla descendente mono o estéreo de modo que la pluralidad de canales contribuyen a la mezcla descendente mono o estéreo en un nivel que difiere entre al menos dos canales de la señal multi-canal, en el que la formación de la mezcla descendente mono o estéreo se realiza de modo que un canal central de la
- 65

pluralidad de canales contribuye a la mezcla descendente mono o estéreo según un nivel reducido con respecto a los otros canales de la señal multi-canal.

- 5 **10.** Programa de ordenador que tiene instrucciones para realizar, cuando se ejecuta en un ordenador, un procedimiento según la reivindicación 9.

10

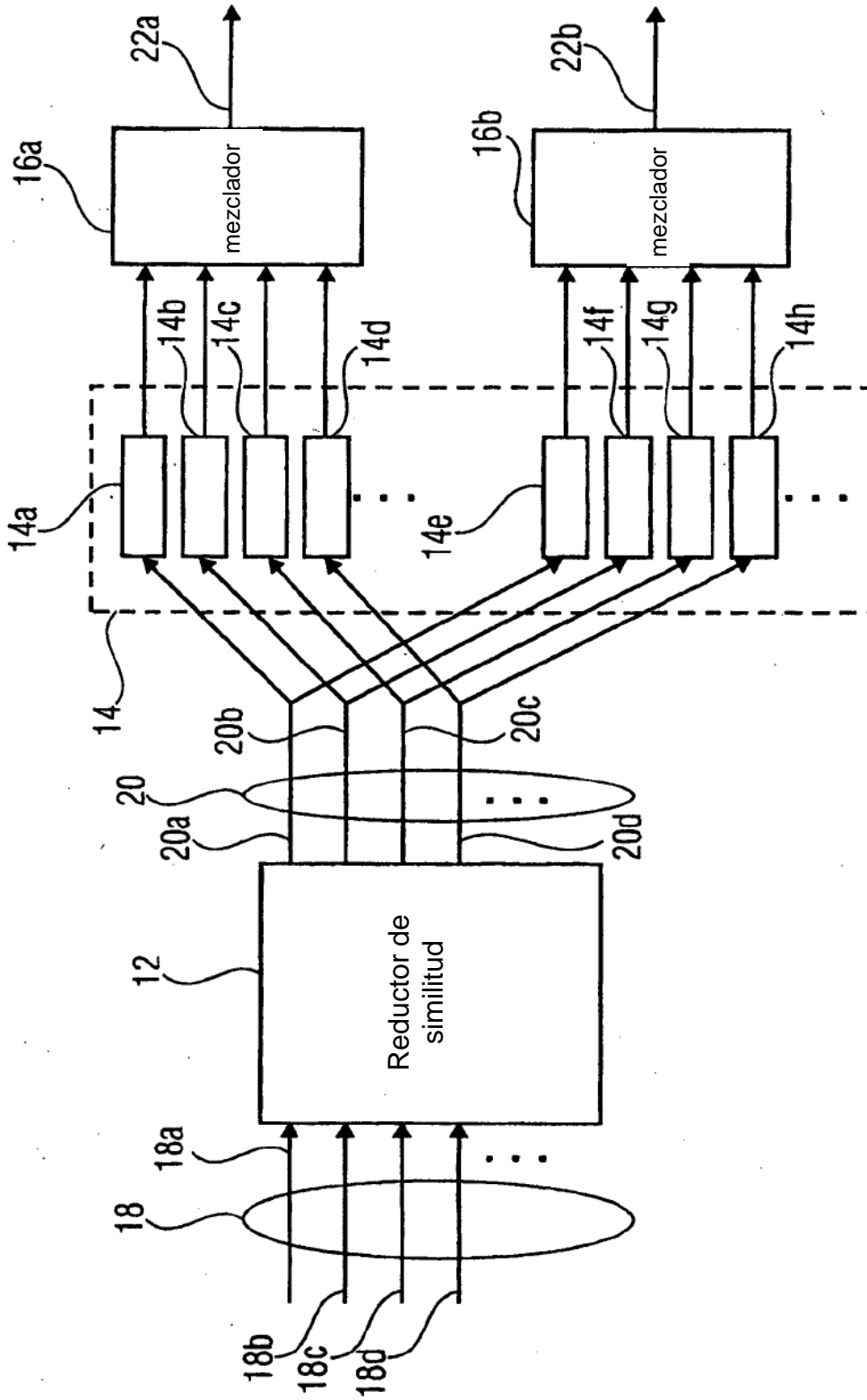


FIG 1

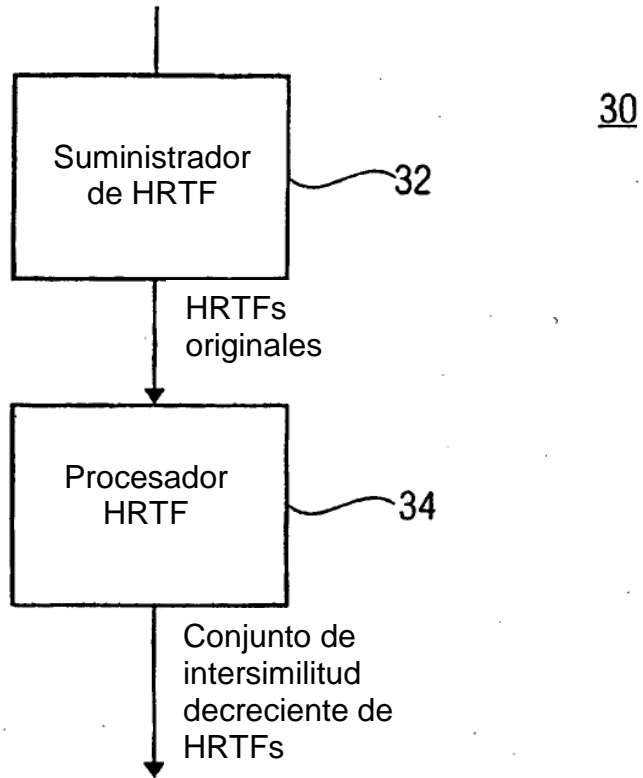


FIG 2

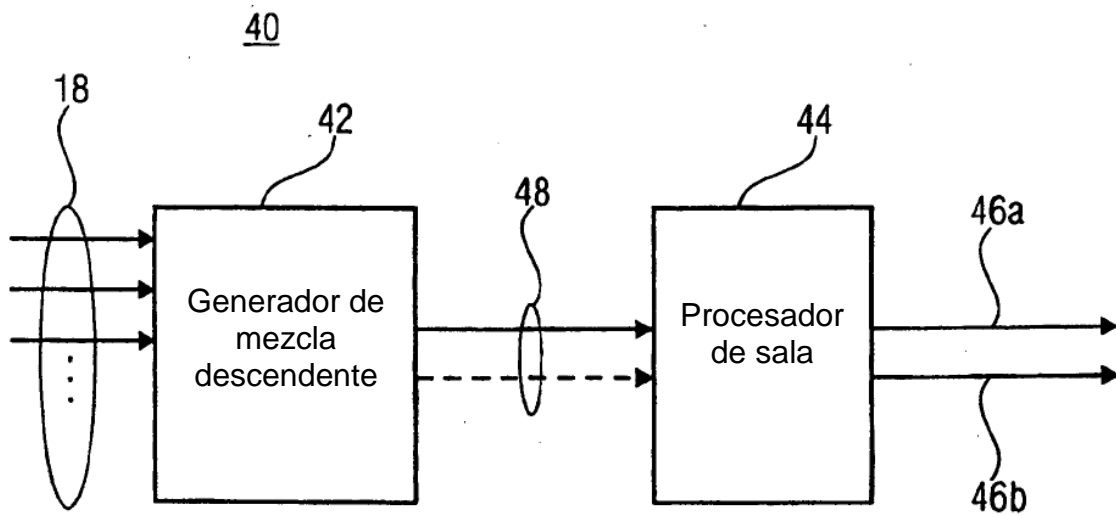


FIG 3

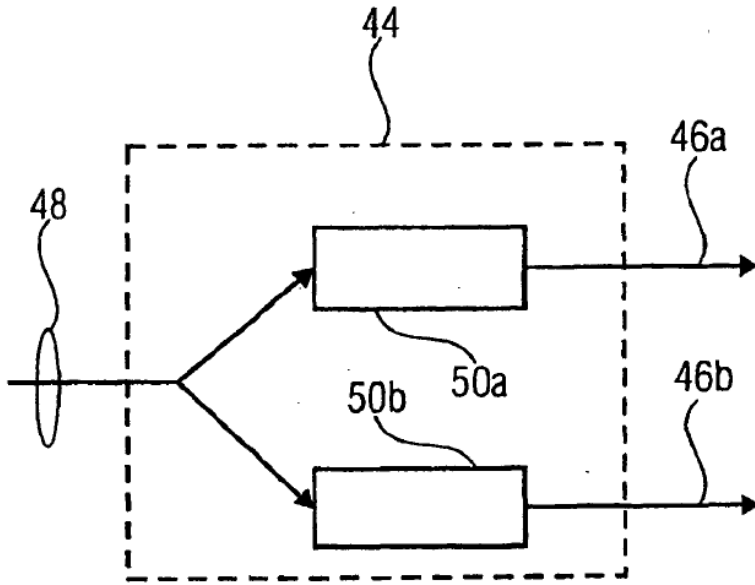


FIG 4A

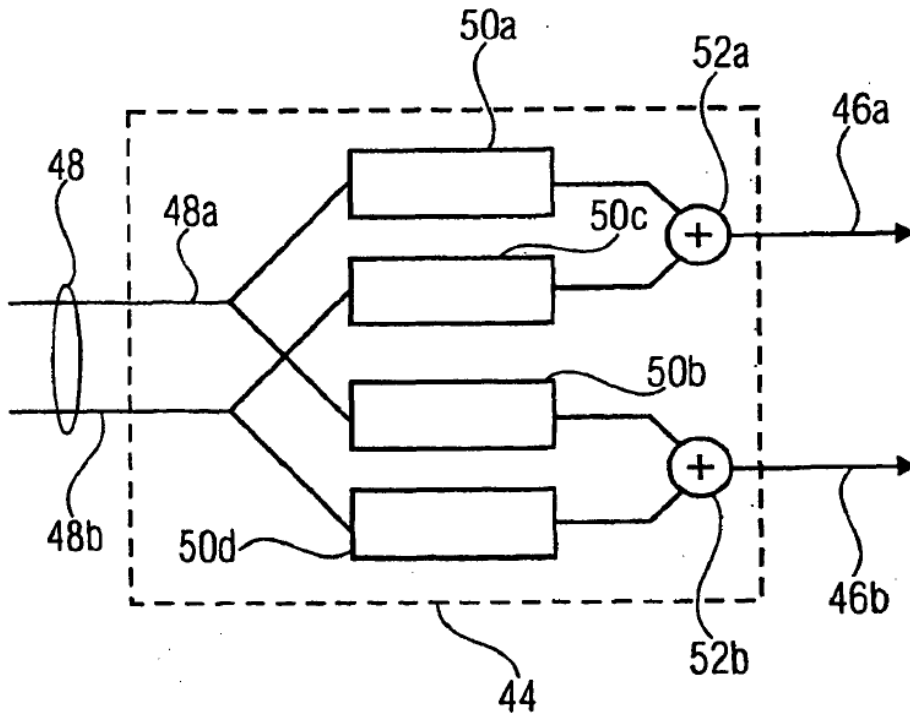


FIG 4B

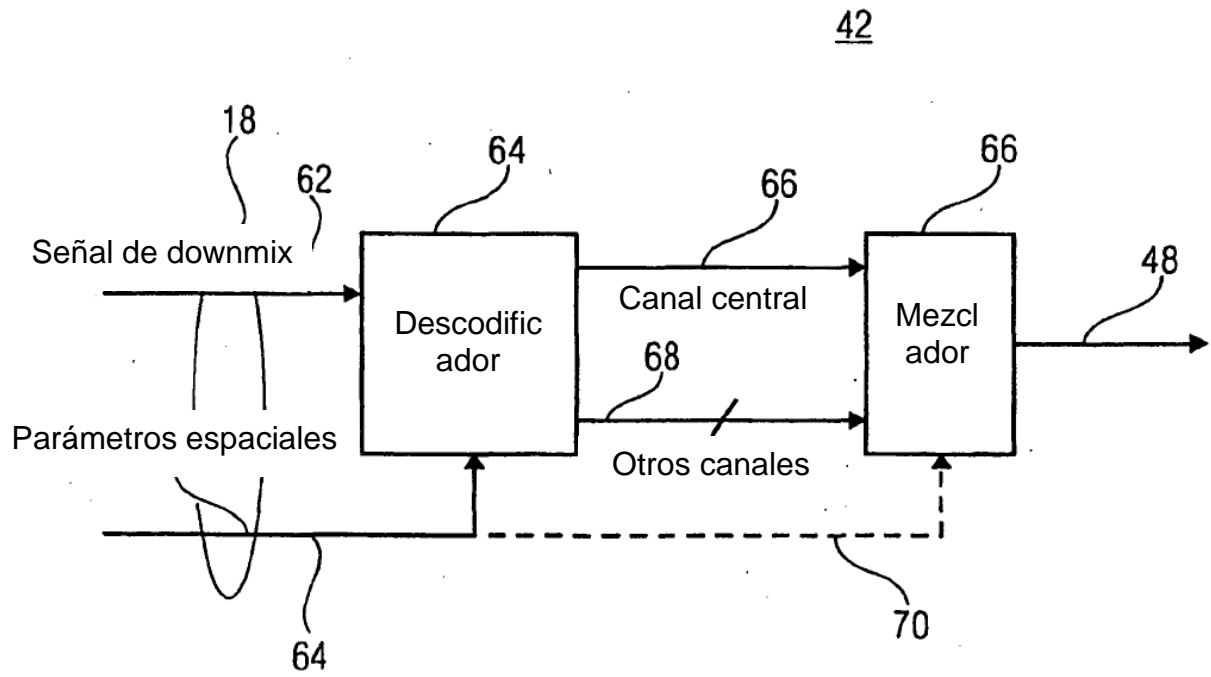


FIG 5

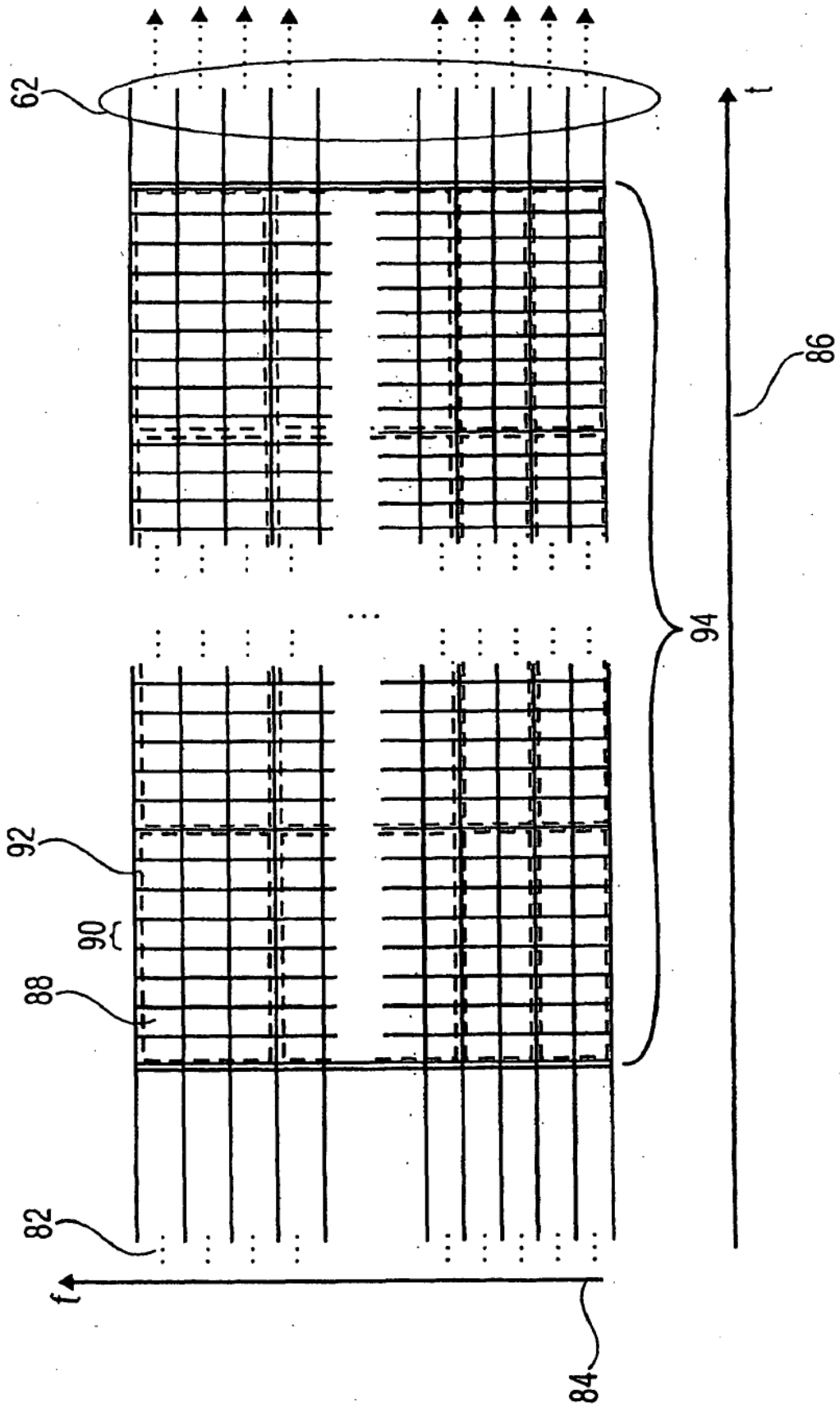


FIG 6

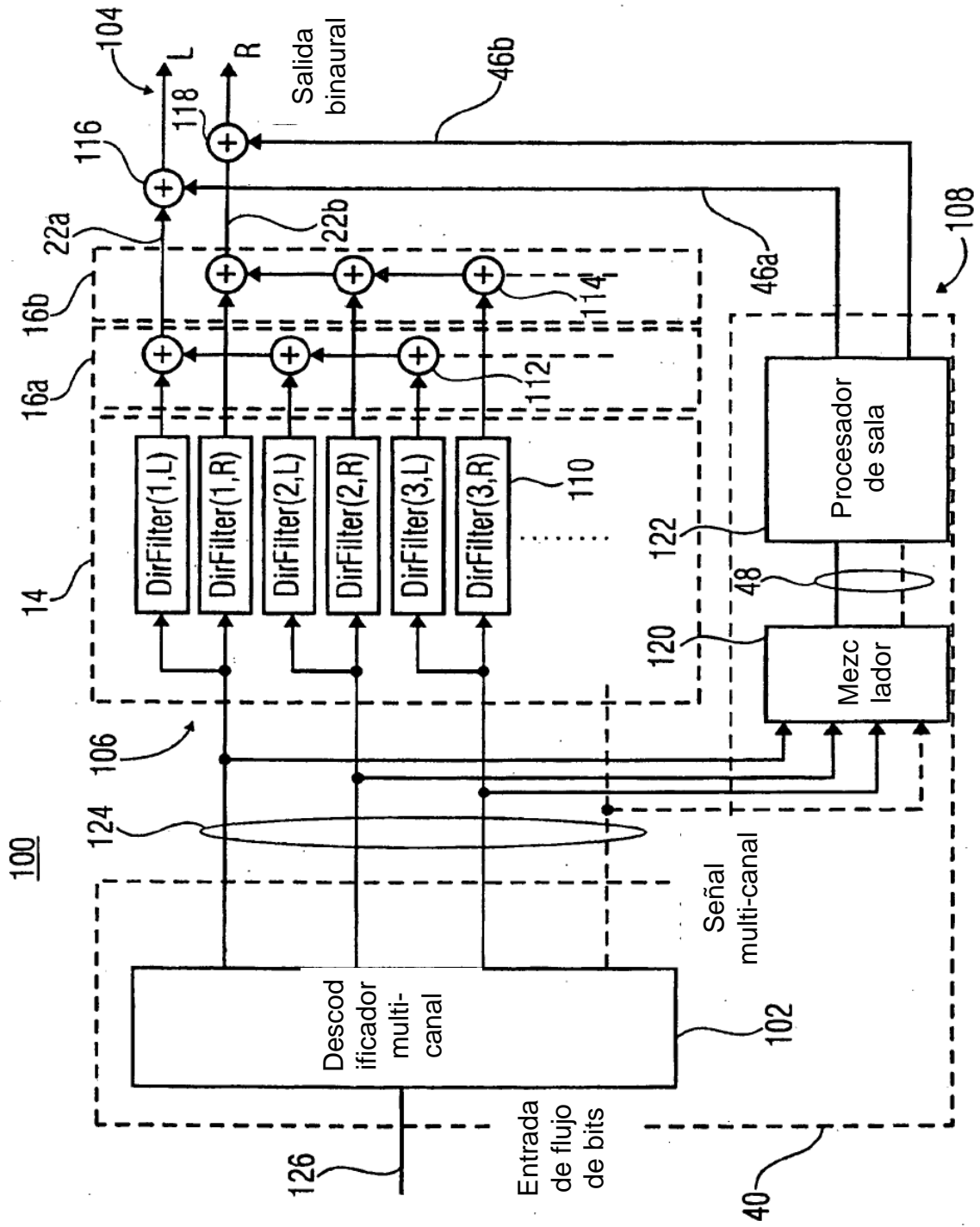


FIG 7

FIG 8

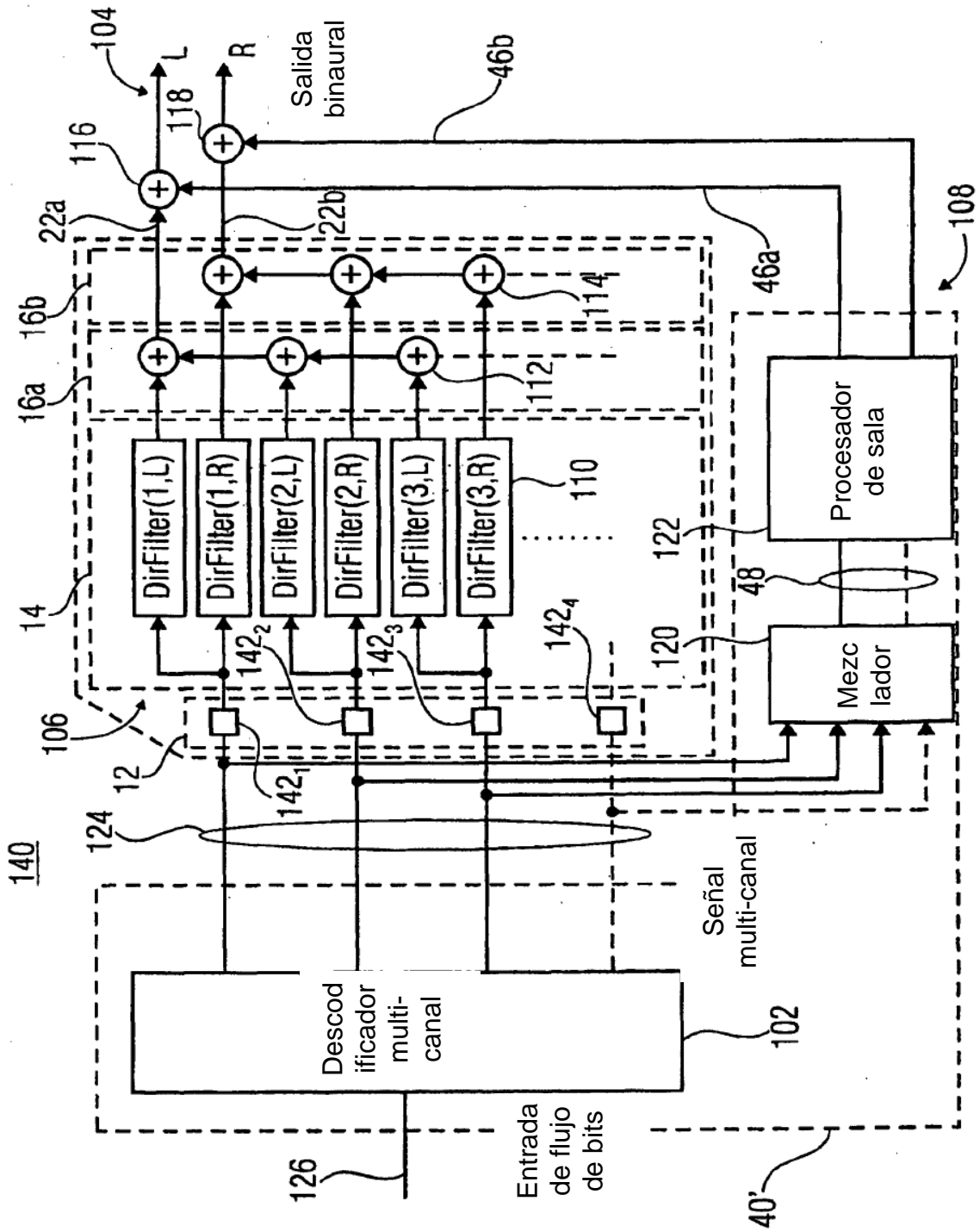
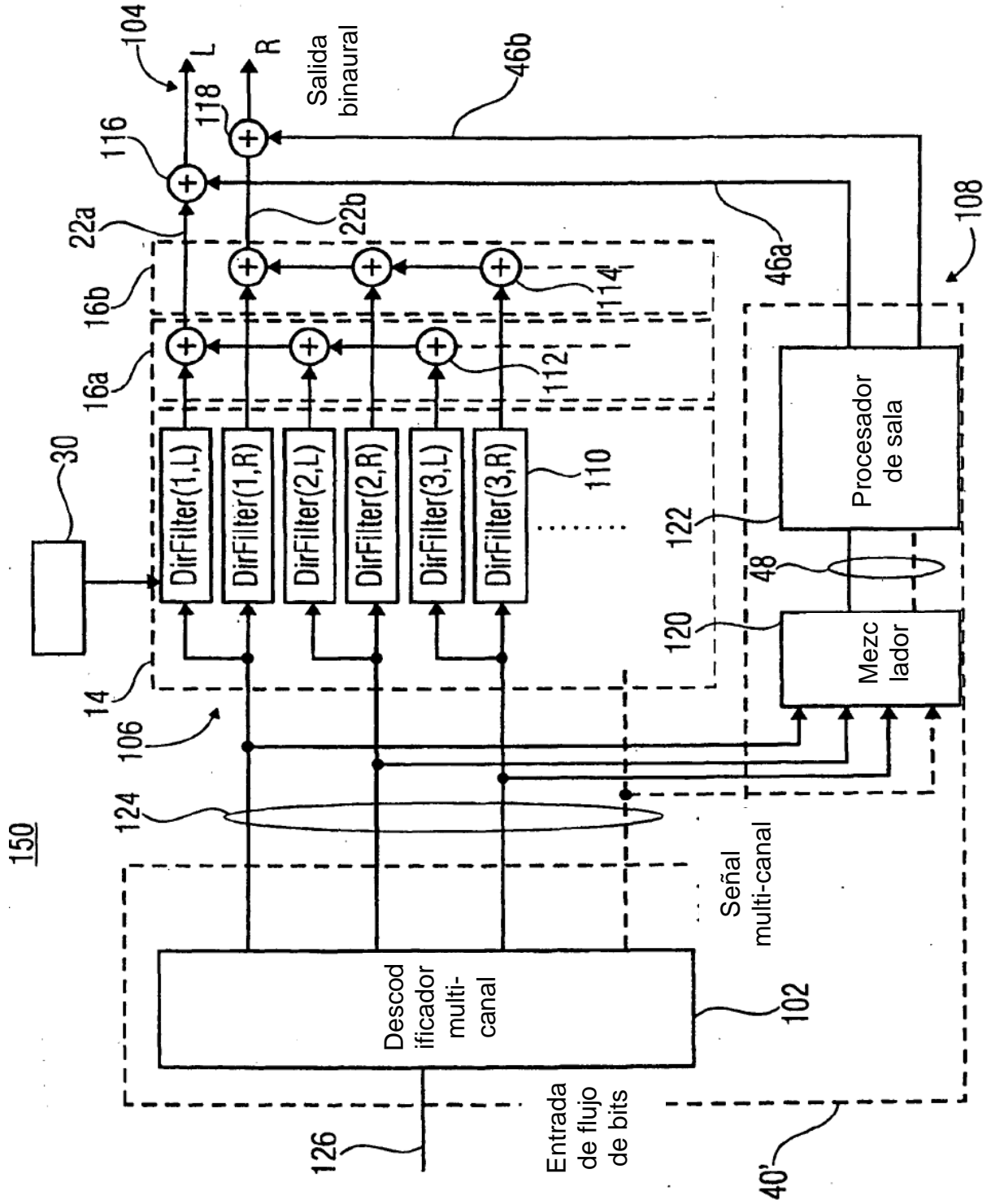


FIG 9



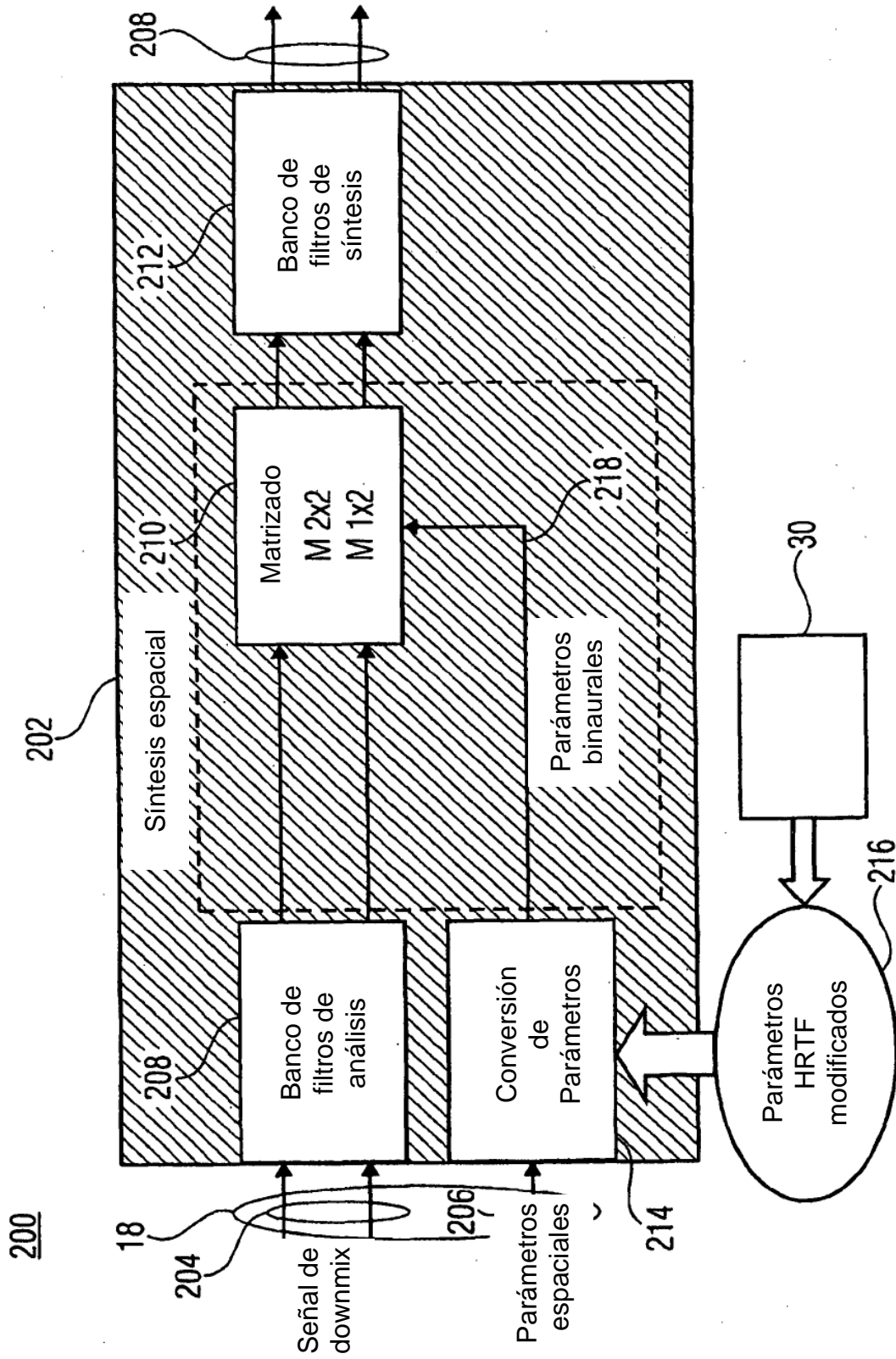


FIG 10

FIG 11

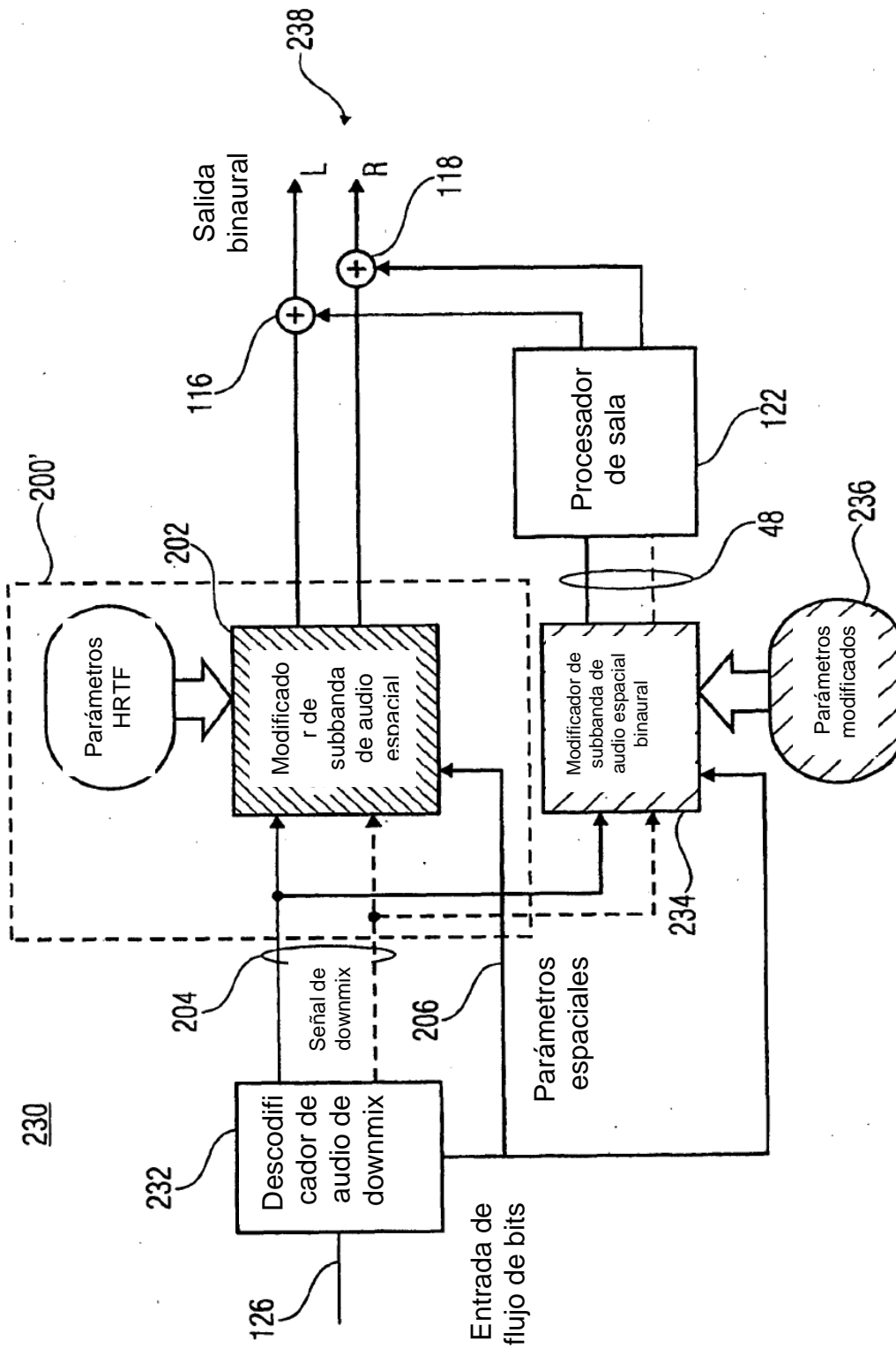


FIG 12

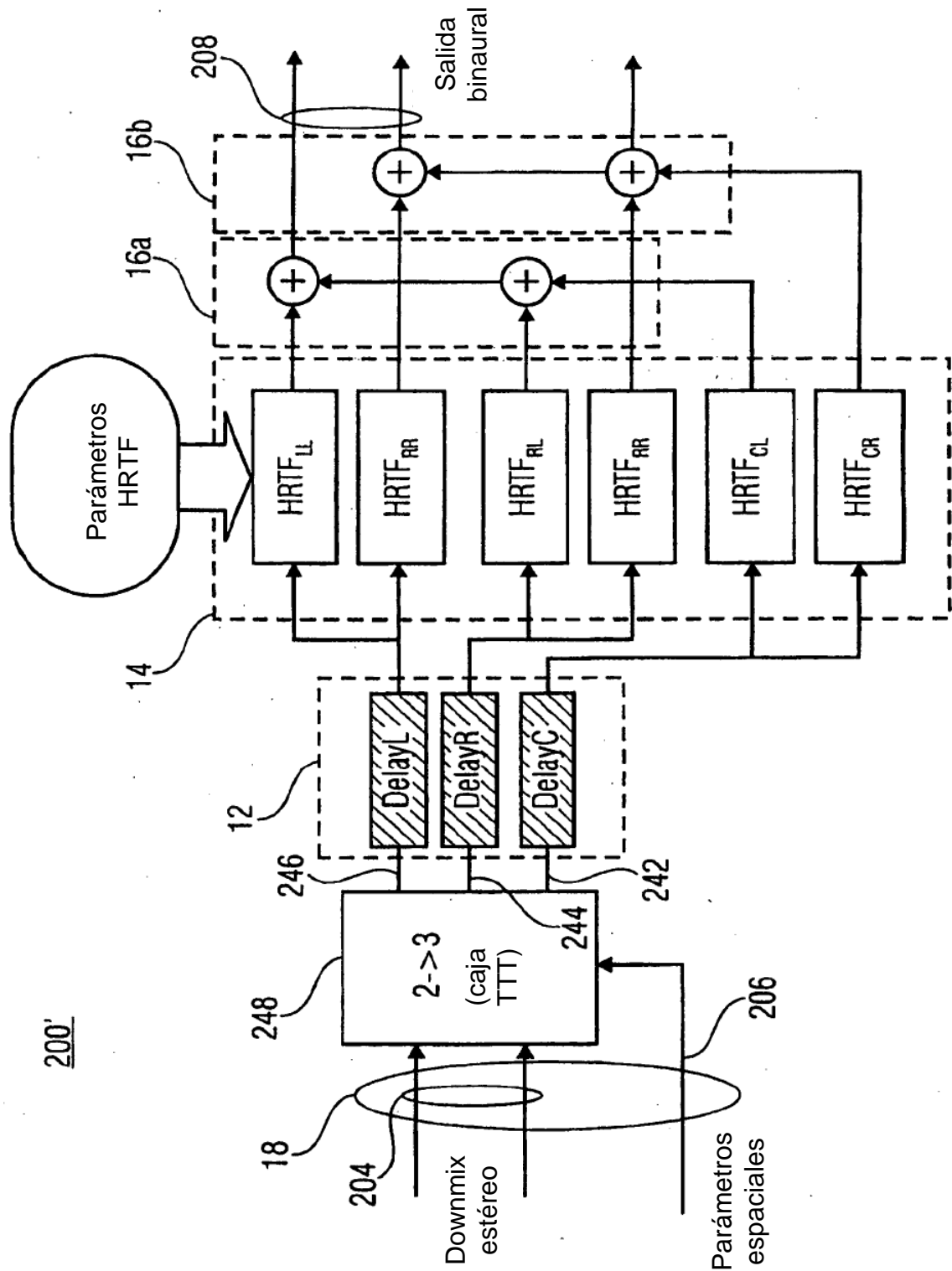
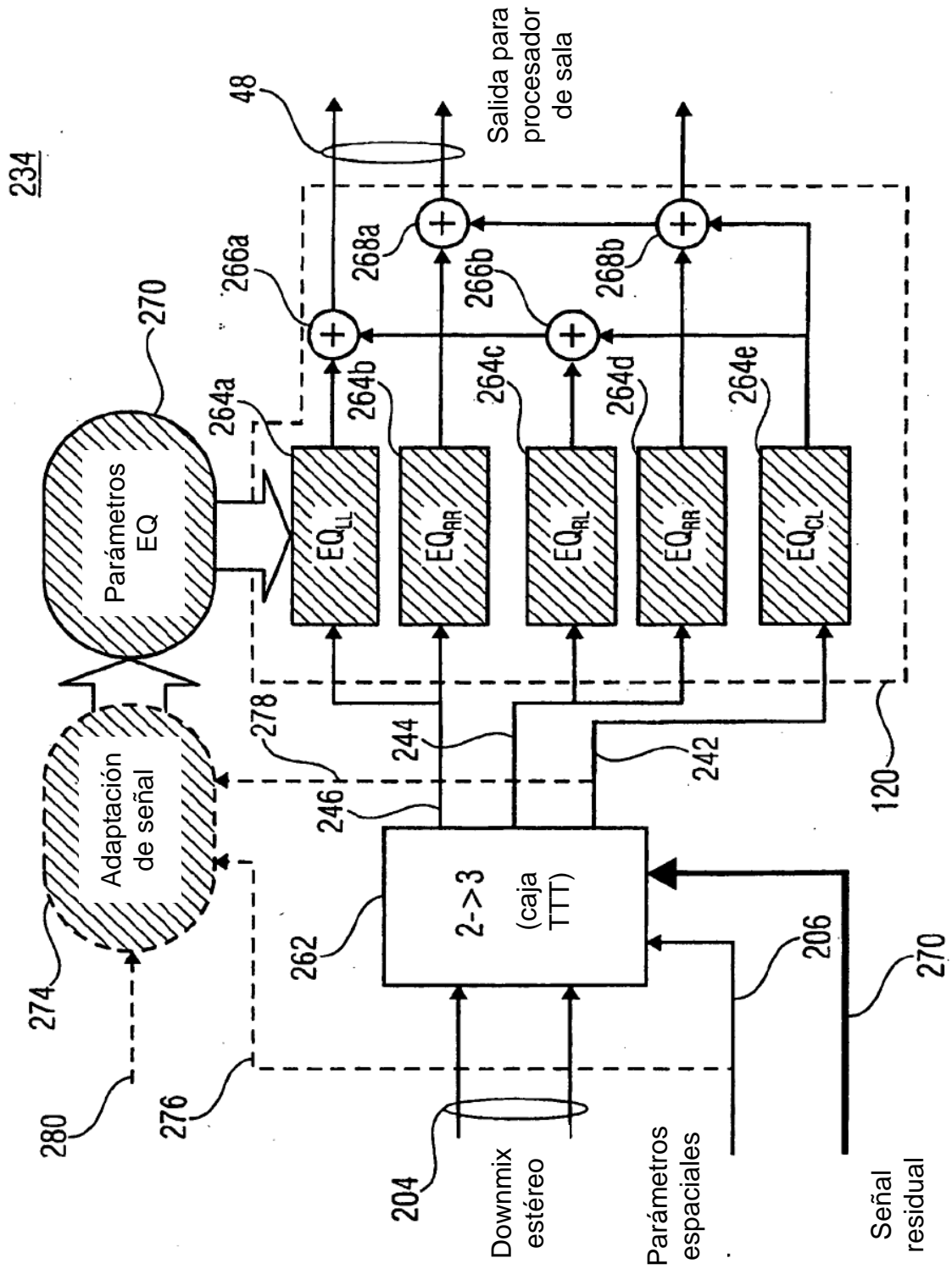


FIG 13



234