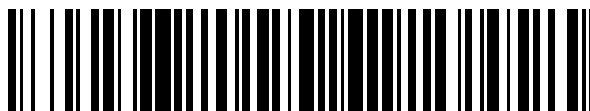


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 527 918**

51 Int. Cl.:

**G10L 19/00** (2013.01)

**H04S 7/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.09.2006** **E 10155535 (7)**

97 Fecha y número de publicación de la concesión europea: **12.11.2014** **EP 2216776**

54 Título: **Descodificador multi-canal binaural en el contexto de reglas de mezcla ascendente que no conservan la energía**

30 Prioridad:

**02.06.2006 US 803819 P**

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

**02.02.2015**

73 Titular/es:

**DOLBY INTERNATIONAL AB (100.0%)**  
**Apollo Building, 3E, Herikerbergweg 1-35**  
**1101 CN Amsterdam Zuid-Oost , NL**

72 Inventor/es:

**VILLEMOES, LARS**

74 Agente/Representante:

**PONTI SALES, Adelaida**

ES 2 527 918 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

Descodificador multi-canal binaural en el contexto de reglas de mezcla ascendente que no conservan la energía

5 Sector de la invención

**[0001]** La presente invención se refiere a decodificación binaural de señales de audio multi-canal a partir de señales de mezcla descendente disponibles y datos de control adicionales, mediante filtración HRTF.

10 Contexto técnico de la invención y Antecedentes

**[0002]** El reciente desarrollo en la codificación de audio ha hecho puesto procedimientos a disposición para recrear una representación multicanal de una señal de audio a partir de una señal estéreo (o mono) de la señal y los datos de control correspondientes. Estos procedimientos difieren sustancialmente de la solución más antigua basada en matriz tal como Dolby Prologic, puesto que los datos de control adicionales se transmiten para controlar la recreación, también denominada como upmix, de los canales envolventes basándose en los canales mono o estéreo transmitidos.

**[0003]** Por lo tanto, un decodificador de audio multicanal paramétrico como este, por ejemplo, MPEG Surround reconstruye  $N$  canales a partir de  $M$  canales transmitidos, donde  $N > M$ , y los datos de control adicionales. Los datos de control adicionales representan una velocidad de datos significativamente menor que la requerida para la transmisión de todos los  $N$  canales, haciendo que la codificación sea muy eficaz mientras que al mismo tiempo se garantiza la compatibilidad tanto con dispositivos de  $M$  canales y los dispositivos con  $N$  canales. [J. Breebaart et al. "MPEG codification of spatial audio / MPEG Surround: overview and current status", Proc. 119th AES convention, New York, USA, October 2005, Preprint 6447].

**[0004]** Estos procedimientos de codificación de sonido envolvente paramétricos generalmente comprenden una parametrización de la señal envolvente a partir de la Diferencia de nivel de Canales (CLD) y la Coherencia/correlación entre canales (ICC). Estos parámetros describen niveles de potencia y correlación entre pares de canales en el proceso de mezcla ascendente. También se usan otros Coeficientes de predicción de canal (CPC) en el estado de la técnica para predecir canales intermedios o de salida durante el proceso de mezcla ascendente.

**[0005]** Otros desarrollos en la codificación de audio han proporcionado medios para obtener una impresión de señal multicanal con auriculares estéreo. Esto se realiza comúnmente por mezcla descendente de una señal multicanal a estéreo utilizando la señal multicanal original y filtros HRTF (funciones de transferencia relacionadas con la cabeza).

**[0006]** Como alternativa, por supuesto, sería útil por razones de eficiencia computacional y también por razones de calidad de audio simplificar la generación de la señal binaural que tienen el canal binaural izquierdo y el canal binaural derecho.

**[0007]** Sin embargo, la cuestión es cómo se pueden combinar los filtros HRTF originales. Además surge un problema en un contexto de una regla de mezcla ascendente afectada de pérdida de energía, es decir, cuando la señal de entrada del decodificador multicanal incluye una señal de mezcla descendente que tiene, por ejemplo, un primer canal de mezclado descendente y un segundo canal de mezclado descendente, y que tiene además parámetros espaciales, que se utilizan para mezclar hacia arriba de una manera que no conserva la energía. Estos parámetros son también conocidos como parámetros de predicción o parámetros de CPC. Estos parámetros tienen, en contraste con la diferencia de nivel de parámetros de canal, la propiedad de que no se calculan para reflejar la distribución de energía entre dos canales, sino que se calculan para realizar una adaptación lo mejor-posible a la forma de onda lo que resulta automáticamente en un error de energía (por ejemplo, pérdida), ya que, cuando se generan los parámetros de predicción, a uno no le importa acerca de las propiedades de conservación de energía de una mezcla ascendente, sino que se preocupa de tener una mejor coincidencia de la forma de onda en el dominio del tiempo o de sub-banda con respecto a la señal reconstruida en comparación con la señal original.

**[0008]** Cuando simplemente se combinan linealmente filtros HRTF sobre la base de estos parámetros de predicción espaciales de transmisión, se recibirán artefactos que son especialmente graves, cuando la predicción de los canales se realiza mal. En esta situación, incluso las dependencias lineales sutiles conducen a la coloración espectral no deseada de la salida binaural. Se ha encontrado que este artefacto se produce con mayor frecuencia cuando los canales originales llevan señales a pares no correlacionadas y que tienen magnitudes comparables.

**[0009]** WO 2006/048203 A1 describe la utilización de una medida de energía para compensar pérdidas de energía debido a una mezcla ascendente predictiva para reconstrucción multi-canal de señales de audio a partir de al menos un canal de base. La medida de energía se puede aplicar en el codificador o el decodificador. Además, una señal descorrelacionada se añade a canales de salida generados por un proceso de mezcla ascendente que introduce pérdidas de energía. La energía de la señal descorrelacionada es menor o igual que un error de energía introducido por la mezcla ascendente predictiva.

[0010] WO 2004/028204 A2 describe un procedimiento y un sistema de medios de/para una generación de al menos una señal de salida a partir de al menos una señal de entrada a partir de un segundo conjunto de señales de sonido que tienen segundo conjunto relacionado de Funciones de Transferencia Relativas a la Cabeza. El sistema de medios puede ser una TV, un reproductor de CD, un reproductor de DVD, una radio, una pantalla visual, un amplificador, un auricular o un VCR. El procedimiento incluye las etapas de determinar, para cada señal en el segundo conjunto de señales de sonido, una relación ponderada que comprende al menos una señal a partir de una tercera etapa de señales de sonido intermedias y al menos un valor de ponderación; determinar un primer conjunto de Funciones de Transferencia Relativas a la Cabeza a partir del segundo conjunto de señales de sonido, el segundo conjunto de Funciones de Transferencia Relativas a la Cabeza y la relación ponderada; y transferir al menos una señal del tercer conjunto de señales de sonido intermedias mediante al menos un HRTF a partir de dicho primer conjunto de Funciones de Transferencia Relativas a la Cabeza para generar al menos una señal de salida, que pertenecen a dicho primer conjunto de señales de sonido.

#### Resumen de la invención

[0011] Es el objeto de la presente invención proporcionar un concepto eficiente y cualitativamente aceptable para la decodificación multicanal para obtener una señal binaural que se pueda utilizar, por ejemplo, para la reproducción de una señal multicanal en auriculares.

[0012] Este objeto se logra con un descodificador multi-canal según la reivindicación 1.

[0013] Además, este objeto se logra con un procedimiento de decodificación multi-canal según la reivindicación 2.

[0014] Otros aspectos de esta invención se refieren a un programa de ordenador que tiene un código legible por ordenador que implementa, cuando se ejecuta en un ordenador, el procedimiento de decodificación multi-canal.

[0015] La presente invención se basa en el hallazgo de que se puede incluso utilizar ventajosamente información de regla de mezcla ascendente en una mezcla ascendente que resulta en un error de energía para el filtrado de una señal de mezcla descendente para obtener una señal binaural sin tener que representar totalmente la señal multicanal y aplicar posteriormente un gran número de filtros HRTF. En lugar de ello, de acuerdo con la presente invención, la información de regla de mezcla ascendente relativa a una regla de mezcla ascendente afectada de error de energía puede ser usada ventajosamente para la representación binaural acortada de una señal de mezcla descendente, cuando, de acuerdo con la presente invención, se calcula un factor de ganancia y se utiliza en el filtrado de la señal de mezcla descendente, en el que este factor de ganancia se calcula de manera que el error de energía se reduce o se elimina por completo.

[0016] En particular, el factor de ganancia no sólo depende de la información sobre la regla de mezcla ascendente como los parámetros de predicción, pero, sobre todo, también depende de filtros a base de funciones de transferencia relativas a la cabeza correspondientes a canales de mezcla ascendente, para los que se proporciona la regla de mezcla ascendente. En particular, estos canales de mezcla ascendente nunca existen en la realización preferida de la presente invención, puesto que los canales binaurales se calculan en primer lugar sin la representación, por ejemplo, tres canales intermedios. Sin embargo, se pueden obtener o proporcionar filtros basados en HRTF correspondientes a los canales de mezcla ascendente aunque los propios canales de mezcla ascendente nunca existen en la realización preferida. Se ha encontrado que el error de energía introducido por esta regla de mezcla ascendente afectada de pérdida de energía no sólo corresponde a la información de regla de mezcla ascendente que se transmite desde el codificador al decodificador, sino que también depende de los filtros basados en HRTF de manera que, cuando se genera el factor de ganancia, los filtros basados en HRTF también influyen en el cálculo del factor de ganancia.

[0017] En vista de ello, la presente invención representa la interdependencia entre la información de regla de mezcla ascendente tal como parámetros de predicción y la aparición específica de los filtros basados en HRTF para los canales que serían el resultado de mezcla ascendente utilizando la regla de mezcla ascendente.

[0018] Por lo tanto, la presente invención proporciona una solución al problema de la coloración espectral derivada de la utilización de una mezcla ascendente predictiva en combinación con la decodificación de binaural paramétrica de audio multicanal.

[0019] Las realizaciones preferidas de la presente invención comprenden las siguientes características: un decodificador de audio para generar una señal de audio binaural a partir de  $M$  señales decodificadas y parámetros espaciales pertinentes a la creación de  $N > M$  canales, comprendiendo el decodificador un calculador de ganancia para estimar, en varias sub-bandas, dos ganancias de compensación a partir de  $P$  pares de filtros sub-banda binaurales y un subconjunto de los parámetros espaciales pertinentes para la creación de  $P$  canales intermedios, y un ajustador de ganancia para modificar, en varias sub-bandas,  $M$  pares de filtros sub-banda binaurales obtenidos por combinación lineal de los  $P$  pares de filtros sub-banda binaurales, consistiendo la modificación en multiplicar cada uno de los  $M$  pares con las dos ganancias calculadas mediante el calculador de ganancia.

Breve descripción de los dibujos

**[0020]** La presente invención se describirá ahora mediante ejemplos ilustrativos, no limitativos del alcance de la invención, con referencia a los dibujos adjuntos, en los que:

La figura 1 ilustra la síntesis binaural de señales multicanal paramétricas utilizando filtros relacionados con HRTF;  
 La figura 2 ilustra la síntesis binaural de señales multicanal paramétricas utilizando filtración combinada;  
 La figura 3 ilustra los componentes del combinador parámetro / filtro;  
 La figura 4 ilustra la estructura de la decodificación MPEG Surround espacial;  
 La figura 5 ilustra el espectro de una señal binaural decodificada sin la compensación de la ganancia de la invención;  
 La figura 6 ilustra el espectro de la decodificación de la invención de una señal binaural.  
 La figura 7 ilustra una síntesis binaural convencional usando HRTF;  
 La figura 8 ilustra un codificador MPEG envolvente;  
 La figura 9 ilustra una cascada de decodificador MPEG surround y sintetizador binaural;  
 La figura 10 ilustra un decodificador binaural 3D conceptual para ciertas configuraciones;  
 La figura 11 ilustra un codificador espacial para ciertas configuraciones;  
 La figura 12 ilustra un decodificador espacial (MPEG Surround);  
 La figura 13 ilustra el filtrado de dos canales de mezcla descendente utilizando cuatro filtros para obtener señales binaurales sin corrección del factor de ganancia;  
 La figura 14 ilustra una configuración espacial para explicar diferentes filtros HRTF 1-10 en una configuración de cinco canales;  
 La figura 15 ilustra una situación de la fig. 14, cuando se han combinado los canales L, Ls y R, Rs;  
 La figura 16a ilustra la configuración de la figura 14 o la figura 15, cuando se ha realizado una combinación máxima de filtros HRTF y sólo permanecen los cuatro filtros de la figura 13;  
 La figura 16b ilustra una regla de mezcla ascendente tal como se determina mediante el codificador de la figura 20 que tiene coeficientes de mezcla ascendente que resulta en una mezcla ascendente que no conserva la energía;  
 La figura 17 ilustra cómo se combinan filtros HRTF para obtener finalmente cuatro filtros basados en HRTF;  
 La figura 18 ilustra una forma de realización preferida de un decodificador multicanal de la invención;  
 La figura 19a ilustra una primera realización del decodificador multicanal de la invención que tiene una etapa de escala después de la filtración a base de HRTF sin corrección de la ganancia;  
 La figura 19b ilustra un dispositivo de la invención que tiene filtros basados en HRTF ajustados que resultan en una señal de salida del filtro de ganancia ajustada; y  
 La figura 20 muestra un ejemplo de un codificador para generar la información para una regla de mezcla ascendente que no conserva la energía.

Descripción detallada de realizaciones preferidas

**[0021]** Antes de discutir la ganancia de la invención ajustando aspecto en detalle, una combinación de filtros HRTF y el uso de filtros basados en HRTF se discutirá en relación con las Figs. 7 a 11.

**[0022]** Con el fin de delinear mejor las características y ventajas de la presente invención una descripción más elaborada se da primero. Un algoritmo de síntesis binaural se describe en la Fig. 7. Un conjunto de canales de entrada se filtra mediante un conjunto de HRTFs. Cada señal de entrada se divide en dos señales (un componente izquierdo 'L', y un componente derecho 'R'); cada una de estas señales se filtra posteriormente mediante un HRTF correspondiente a la posición de fuente de sonido deseada. Todas las señales del oído izquierdo se suman posteriormente para generar la señal de salida binaural izquierda, y las señales del oído derecho se suman para generar la señal de salida binaural derecha.

**[0023]** La convolución de HRTF se puede realizar en el dominio del tiempo, pero a menudo se prefiere realizar el filtrado en el dominio de la frecuencia debido a la eficiencia computacional. En ese caso, la suma tal como se muestra en la figura 7 también se realiza en el dominio de la frecuencia.

**[0024]** En principio, el procedimiento de síntesis binaural tal como se indica en la figura 7 se podría utilizar directamente en combinación con un codificador / decodificador MPEG surround. El codificador MPEG surround se muestra esquemáticamente en la figura 8. Se analiza una señal de entrada multi-canal por un codificador espacial, que da como resultado una señal de mezcla descendente mono o estéreo, combinada con parámetros espaciales. La mezcla descendente se puede codificar con cualquier códec de audio convencional mono o estéreo. El flujo de bits de mezcla descendente resultante se combina con los parámetros espaciales por un multiplexor, lo que resulta en el flujo de bits total de salida.

**[0025]** Un esquema de síntesis binaural en combinación con un decodificador de sonido MPEG surround se muestra en la figura 9. El flujo de bits de entrada se multiplexa dando como resultado unos parámetros espaciales y un flujo de bits de mezcla descendente. Este último flujo de bits se decodifica utilizando un decodificador mono o estéreo convencional. La mezcla descendente decodificada se decodifica mediante un decodificador espacial, que genera una salida multicanal a partir de los parámetros espaciales transmitidos. Finalmente, la salida multi-canal se procesa

mediante una etapa de síntesis binaural tal como se representa en la figura 7, lo que resulta en una señal de salida binaural.

**[0026]** Sin embargo, hay al menos tres desventajas de esta cascada de un decodificador MPEG surround y un módulo de síntesis binaural:

- Una representación de la señal multi-canal se calcula como un paso intermedio, seguido de la convolución de HRTF y mezcla descendente en la etapa de síntesis binaural. Aunque la convolución de HRTF debe realizarse por canal, debido a que cada canal de audio puede tener una posición espacial diferente, esta es una situación indeseable desde un punto de vista de la complejidad.
- El decodificador espacial opera en un banco de filtros (MGC) de dominio. La convolución de HRTF, por otro lado, se aplica típicamente en el dominio de FFT. Por lo tanto, son necesarias una cascada de un banco de filtros de síntesis QMF multi-canal, una transformada multi-canal DFT, y una transformada DFT inversa estéreo, lo que da como resultado un sistema con altas exigencias computacionales.
- La codificación de artefactos creados por el decodificador espacial para crear una reconstrucción multicanal será audible, y posiblemente mejorada en la salida binaural (estéreo).

**[0027]** El codificador espacial se muestra en la figura 11. Una señal de entrada multi-canal que consiste en señales  $L_f$ ,  $L_s$ ,  $C$ ,  $R_f$  y  $R_s$ , para los canales delantero izquierdo, surround izquierdo, central, delantero derecho y derecho surround se procesa mediante dos unidades "OTT", que tanto generan una mezcla descendente mono y parámetros para dos señales de entrada. Las señales de mezcla descendente resultantes, junto con el canal central y parámetros espaciales adicionales.

**[0028]** Los parámetros resultantes del codificador 'TTT' consisten típicamente en un par de coeficientes de predicción para cada banda de parámetro, o un par de diferencias de nivel para describir las relaciones de energía de las tres señales de entrada. Los parámetros de los codificadores 'OTT' consisten en diferencias de nivel y valores de coherencia o de correlación cruzada entre las señales de entrada para cada banda de frecuencia.

**[0029]** En la figura 12 se representa un decodificador MPEG Surround. Las señales de mezcla descendente  $10$  y  $r0$  se introducen en el módulo Dos a Tres, que reproduce un canal central, un canal lateral derecho y un canal lateral izquierdo. Estos tres canales se procesan adicionalmente mediante varios módulos OTT (de uno a dos) que producen los seis canales de salida.

**[0030]** El decodificador binaural correspondiente como se ve desde un punto de vista conceptual se muestra en la figura 10. En el dominio del banco de filtros, la señal de entrada estéreo ( $L_0$ ,  $R_0$ ) se procesa mediante un decodificador TTT, lo que da como resultado tres señales  $L$ ,  $R$  y  $C$ . Estas tres señales se someten a un procesamiento de parámetros HRTF. Los 6 canales resultantes se suman para generar el par de salida estéreo binaural ( $L_0$ ,  $R_0$ ).

**[0031]** El decodificador TTT puede describirse como la siguiente operación matricial:

$$\begin{bmatrix} L \\ R \\ C \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} \begin{bmatrix} L_0 \\ R_0 \end{bmatrix},$$

con entradas de matriz  $m_{xy}$  dependientes de los parámetros espaciales. La relación de parámetros espaciales y entradas de matriz es idéntica a aquellas relaciones como en el decodificador MPEG surround 5.1-multicanal. Cada uno de las tres señales resultantes  $L$ ,  $R$ , y  $C$  se dividen en dos y se procesan con parámetros HRTF correspondientes a la posición deseada (percibida) de estas fuentes de sonido. Para el canal central ( $C$ ), los parámetros espaciales de la posición de la fuente de sonido se pueden aplicar directamente, lo que da como resultado dos señales de salida para el centro,  $L_B(C)$  y  $R_B(C)$ :

$$\begin{bmatrix} L_B(C) \\ R_B(C) \end{bmatrix} = \begin{bmatrix} H_L(C) \\ H_R(C) \end{bmatrix} C.$$

**[0032]** Para el canal izquierdo ( $L$ ), se combinan los parámetros HRTF de los canales izquierdo frontal e izquierdo surround en un único conjunto de parámetros HRTF, empleando las ponderaciones  $w_{lf}$  y  $w_{lr}$ . Los parámetros HRTF resultantes 'compuestos' simulan el efecto de los canales frontal y surround en un sentido estadístico. Se utilizan las siguientes ecuaciones para generar el par de salida binaural ( $L_B$ ,  $R_B$ ) para el canal izquierdo:

$$\begin{bmatrix} L_B(L) \\ R_B(L) \end{bmatrix} = \begin{bmatrix} H_L(L) \\ H_R(L) \end{bmatrix} L ,$$

**[0033]** De manera similar, la salida binaural para el canal derecho se obtiene según:

$$\begin{bmatrix} L_B(R) \\ R_B(R) \end{bmatrix} = \begin{bmatrix} H_L(R) \\ H_R(R) \end{bmatrix} R ,$$

5 **[0034]** Teniendo en cuenta las definiciones anteriores de  $L_B(C)$ ,  $R_B(C)$ ,  $L_B(L)$ ,  $R_B(L)$ ,  $L_B(R)$  y  $R_B(R)$ , las señales completas  $L_B$  y  $R_B$  se pueden derivar de una sola matriz 2 por 2 dada la señal de entrada estéreo:

$$\begin{bmatrix} L_B \\ R_B \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} L_0 \\ R_0 \end{bmatrix} ,$$

Con

$$h_{11} = m_{11}H_L(L) + m_{21}H_L(R) + m_{31}H_L(C) ,$$

$$h_{12} = m_{12}H_L(L) + m_{22}H_L(R) + m_{32}H_L(C) ,$$

$$h_{21} = m_{11}H_R(L) + m_{21}H_R(R) + m_{31}H_R(C) ,$$

$$h_{22} = m_{12}H_R(L) + m_{22}H_R(R) + m_{32}H_R(C) .$$

10 **[0035]** Los filtros  $H_X(Y)$  se pueden expresar como combinaciones ponderadas paramétricas de versiones paramétricas de los filtros HRTF originales. Para que esto funcione, los filtros HRTF originales se expresan como

- Un nivel (promedio) por banda de frecuencia para la respuesta al impulso del oído izquierdo;
- Un nivel (promedio) por banda de frecuencia para la respuesta al impulso del oído derecho;
- 15 • Un tiempo de llegada o diferencia de fase (promedio) entre las respuestas al impulso oído izquierdo y oído derecho.

**[0036]** Por lo tanto, los filtros HRTF para el oído izquierdo y derecho dada la señal de entrada del canal central se expresan como:

$$\begin{bmatrix} H_L(C) \\ H_R(C) \end{bmatrix} = \begin{bmatrix} P_l(C)e^{+j\phi(C)/2} \\ P_r(C)e^{-j\phi(C)/2} \end{bmatrix} ,$$

20 donde  $P_l(C)$  es el nivel promedio para una banda de frecuencia dada para el oído izquierdo, y  $\phi(C)$  es la diferencia de fase.

25 **[0037]** Por lo tanto, el procesamiento de parámetros HRTF consiste simplemente en una multiplicación de la señal con  $P_l$  y  $P_r$  correspondientes a la posición de la fuente de sonido del canal central, mientras que la diferencia de fase se distribuye simétricamente. Este proceso se lleva a cabo independientemente para cada banda QMF, utilizando el cartografiado de los parámetros HRTF al banco de filtros QMF, por un lado, y el cartografiado de parámetros espaciales a banda QMF por otro lado.

**[0038]** Del mismo modo los filtros HRTF para los oídos izquierdo y derecho dado el canal izquierdo y el canal derecho están dados por:

$$H_L(L) = \sqrt{w_{lf}^2 P_l^2(Lf) + w_{ls}^2 P_l^2(Ls)} ,$$

$$H_R(L) = e^{-j(w_{lf}^2 \phi(Lf) + w_{ls}^2 \phi(Ls))} \sqrt{w_{lf}^2 P_r^2(Lf) + w_{ls}^2 P_r^2(Ls)} .$$

$$H_L(R) = e^{+j(w_{rf}^2 \phi(Rf) + w_{rs}^2 \phi(Rs))} \sqrt{w_{rf}^2 P_l^2(Rf) + w_{rs}^2 P_l^2(Rs)} ,$$

$$H_R(R) = \sqrt{w_{rf}^2 P_r^2(Rf) + w_{rs}^2 P_r^2(Rs)}$$

5 **[0039]** Claramente, los HRTFs son combinaciones ponderadas de los niveles y diferencias de fase para los filtros HRTF parametrizados para los seis canales originales.

**[0040]** Las ponderaciones  $w_{lf}$  y  $w_{ls}$  dependen del parámetro CLD de la caja "OTT" para Lf y Ls:

$$w_{lf}^2 = \frac{10^{CLD_l/10}}{1 + 10^{CLD_l/10}} , \quad w_{ls}^2 = \frac{1}{1 + 10^{CLD_l/10}} .$$

10 **[0041]** y las ponderaciones  $w_{rf}$  y  $w_{rs}$  dependen del parámetro CLD de la caja "OTT" para Rf y Rs:

$$w_{rf}^2 = \frac{10^{CLD_r/10}}{1 + 10^{CLD_r/10}} , \quad w_{rs}^2 = \frac{1}{1 + 10^{CLD_r/10}} .$$

**[0042]** El enfoque anterior funciona bien para los filtros HRTF cortos que de manera suficientemente exacta pueden expresarse como un nivel promedio por banda de frecuencia, y una diferencia de fase promedio por banda de frecuencia. Sin embargo, para HRTF ecoicas largas este no es el caso.

15 **[0043]** La presente invención enseña cómo extender el enfoque de un decodificador binaural de matriz 2 por 2 para manejar filtros HRTF de longitud arbitraria. Con el fin de lograr esto, la presente invención comprende las siguientes etapas:

- 20
- Transformar las respuestas de filtro HRTF a un dominio del banco de filtros;
  - Extracción de la Diferencia de retardo general o diferencia de fase de pares de filtros HRTF;
  - Dar forma a las respuestas del par de filtros HRTF como una función de los parámetros CL
  - Ajuste de ganancia

25 **[0044]** Esto se logra mediante la sustitución de las seis ganancias complejas  $H_Y(X)$  para  $Y=L_0, R_0$  y  $X=L, R, C$  con seis filtros. Estos filtros se derivan de los diez filtros  $H_Y(X)$  para  $Y = L_0, R_0$  y  $X=Lf, Ls, Rf, Rs, C$ , que describen las respuestas de filtro HRTF dadas en el dominio de QMF. Estas representaciones QMF se pueden lograr de acuerdo con el procedimiento descrito a continuación.

30 **[0045]** El conformado de los filtros de canal frontales y surround se realiza con una combinación lineal compleja según

$$H_Y(X) = g w_f \exp(-j \phi_{Xf} w_s^2) H_Y(Xf) + g w_s \exp(j \phi_{Xf} w_f^2) H_Y(Xs) .$$

**[0046]** El parámetro de fase  $\phi_{xy}$  se puede definir a partir de la principal diferencia entre el tiempo de retardo  $\tau_{xy}$  entre los filtros HRTF frontal y posterior y el índice de subbanda del banco QMF a través de

$$\phi_{xy} = \frac{\pi(n + \frac{1}{2})}{64} \tau_{xy} ,$$

**[0047]** La función de este parámetro de fase en la conformación de filtros es doble. En primer lugar, se realiza una compensación de retardo de los dos filtros antes de la superposición que lleva a una respuesta combinada que modela un tiempo de retardo principal correspondiente a una posición de la fuente entre los altavoces delanteros y traseros. En segundo lugar, hace que el factor de compensación de ganancia  $g$  necesario sea mucho más estable y que varíe más lentamente sobre la frecuencia que en el caso de superposición simple, con  $\phi_{xy} = 0$ .

**[0048]** El factor de ganancia  $g$  se determina por la misma regla de potencia de adición incoherente como para el caso de HRTF paramétrico,

$$P_Y(X)^2 = w_f^2 P_Y(Xf)^2 + w_s^2 P_Y(Xs)^2 ,$$

Donde

$$P_Y(X)^2 = g^2 \left( w_f^2 P_Y(Xf)^2 + w_s^2 P_Y(Xs)^2 + 2w_f w_s P_Y(Xf) P_Y(Xs) \rho_{xy} \right)$$

y  $\rho_{xy}$  es el valor real de la correlación cruzada compleja normalizada entre los filtros

$$\exp(-j\phi_{xy}) H_Y(Xf) \text{ and } H_Y(Xs) .$$

**[0049]** En el caso de una simple superposición con  $\phi_{xy} = 0$ , el valor de  $\rho_{xy}$  varía de un modo irregular y oscilatorio como una función de la frecuencia, lo que conduce a la necesidad de una amplio ajuste de ganancia. En la aplicación práctica, es necesario limitar el valor de la ganancia  $g$  y no se puede evitar una coloración espectral restante de la señal.

**[0050]** En contraste, el uso de conformación con una compensación de fase basada en retardo como se enseña por la presente invención conduce a un comportamiento suave de  $\rho_{xy}$  como una función de la frecuencia. Este valor es a menudo próximo de uno para pares de filtros derivada de HRTF natural de ya que difieren principalmente en un retraso y una amplitud, y el propósito del parámetro de fase es considerar la diferencia de retardo en el dominio del banco de filtros QMF.

**[0051]** Una opción beneficiosa alternativa de parámetro de fase  $\phi_{xy}$  está dada por el cálculo del ángulo de fase de la correlación cruzada compleja normalizada entre los filtros

$$H_Y(Xf) \text{ y } H_Y(Xs) ,$$

y despliegue de los valores de fase con técnicas de despliegue estándar como una función del índice de sub-banda  $n$  del banco QMF. Esta selección tiene la consecuencia de que  $\rho_{xy}$  nunca es negativa y por lo tanto la ganancia de compensación  $g$  satisface

$$1/\sqrt{2} \leq g \leq 1$$

para todas las sub-bandas. Por otra parte la elección de este parámetro de fase permite el conformado de los filtros de canales delanteros y de sonido envolvente en situaciones en las que la diferencia principal de tiempo de retardo  $\tau_{xy}$  no está disponible.

**[0052]** Todas las señales consideradas a continuación son muestras de subbanda de un banco de filtros modulado o análisis FFT por ventanas de las señales de tiempo discreto o señales de tiempo discreto. Se entiende que estas sub-bandas tienen que transformarse de nuevo al dominio de tiempo discreto mediante las operaciones del banco de filtros de síntesis correspondiente.

**[0053]** La figura 1 ilustra un proceso para síntesis binaural de señales multi-canal paramétricas empleando filtros relacionados con HRTF. Una señal multi-canal que comprende  $N$  canales se produce por descodificación espacial 101 a partir de  $M < N$  canales transmitidos y parámetros espaciales transmitidos. Estos  $N$  canales se convierten a su vez en dos canales de salida destinados a escucha binaural mediante filtrado HRTF. Este filtrado HRTF 102 superpone los resultados de filtrar cada canal de entrada con un filtro HRTF para el oído izquierdo y un filtro HRTF para el oído derecho. Con todo, esto requiere  $2N$  filtros. Mientras que la señal multicanal paramétrica logra una experiencia de escucha de alta calidad cuando se escucha a través de los  $N$  altavoces, unas interdependencias sutiles de las señales  $N$  conducirán a artefactos para la escucha binaural. Estos artefactos están dominados por la desviación en el contenido espectral a partir de la señal binaural de referencia definida por el filtrado HRTF de los  $N$  canales originales antes de la codificación. Una desventaja adicional de esta concatenación es que el coste computacional total para la síntesis binaural es la adición del coste requerido para cada uno de los componentes 101 y 102.

**[0054]** La figura 2 ilustra la síntesis binaural de señales multi-canal paramétricas empleando el filtrado combinado enseñado por la presente invención. Los parámetros espaciales transmitidos se dividen en 201 en dos conjuntos, Conjunto 1 y Conjunto 2. Aquí, el Conjunto 2 comprende parámetros pertinentes para la creación de  $P$  canales intermedios a partir de  $M$  canales transmitidos y el Conjunto 1 comprende parámetros pertinentes para la creación de  $N$  canales a partir de  $P$  canales intermedios. El precombinador del estado de la técnica 202 combina pares seleccionados de los  $2N$  filtros de sub-banda relacionados con HRTF  $r$  con ponderaciones que dependen del Conjunto 1 de parámetros y los pares de filtros seleccionados. El resultado de esta precombinación con  $2P$  filtros sub-banda binaurales que representan un par de filtros binaurales para cada uno de los  $P$  canales intermedios. El combinador 203 combina los  $2P$  filtros sub-banda binaurales en un conjunto de  $2M$  filtros sub-banda binaurales mediante la aplicación de ponderaciones que dependen tanto del Conjunto de parámetros 2 como de los  $2P$  filtros sub-banda binaurales. En comparación, un combinador lineal del estado de la técnica aplicaría ponderaciones que dependen solamente del Conjunto de parámetros 2. El conjunto resultante de  $2M$  filtros consiste en un par de filtros binaurales para cada uno de los  $M$  canales transmitidos. La unidad de filtrado combinado 204 obtiene un par de contribuciones a la salida de dos canales para cada uno de los  $M$  canales transmitidos mediante el filtrado con el par de filtros correspondiente. Posteriormente, todas las  $M$  contribuciones se añaden para formar una salida de dos canales en el dominio sub-banda.

**[0055]** La figura 3 ilustra los componentes del combinador 203 para combinación de parámetros espaciales y filtros binaurales. El combinador lineal 301 combina los  $2P$  filtros sub-banda binaurales en  $2M$  filtros binaurales mediante la aplicación de ponderaciones que se derivan de los parámetros espaciales dados, donde estos parámetros espaciales son pertinentes para la creación de  $P$  canales intermedios a partir de los  $M$  canales transmitidos. Específicamente, esta combinación lineal simula la concatenación de una mezcla ascendente a partir de  $M$  canales transmitidos a  $P$  canales intermedios seguida de un filtrado binaural a partir de  $P$  fuentes. El ajustador de ganancia 303 modifica los  $2M$  filtros binaurales que salen del combinador lineal 301 mediante la aplicación de una ganancia izquierda común a cada uno de los filtros que corresponden a la salida de oído izquierdo y mediante la aplicación de una ganancia derecha común a cada uno de los filtros que corresponden a la salida de oído derecha. Estas ganancias se obtienen a partir del calculador de ganancia 302 que deriva las ganancias a partir de los parámetros espaciales y los  $2P$  filtros binaurales. El propósito del ajuste de ganancia de los componentes 302 y 303 es compensar la situación en que los  $P$  canales intermedios de la descodificación espacial llevan dependencias lineales que llevan a una coloración espectral no deseada debido al combinador lineal 301. El calculador de ganancia 302 revelado por la presente invención incluye medios para estimar una distribución de energía de los  $P$  canales intermedios como una función de los parámetros espaciales.

**[0056]** La figura 4 ilustra la estructura de la descodificación espacial MPEG surround en el caso de una señal transmitida estéreo. Las sub-bandas de análisis de las  $M=2$  señales transmitidas se alimentan a una caja 2  $\rightarrow$  3 401 que da como salida  $P=3$  señales intermedias, una izquierda combinada, una combinada derecha, y una combinada central. Esta mezcla ascendente depende de un subconjunto de los parámetros espaciales transmitidos que corresponde al conjunto 2 en la figura 2. Las tres señales intermedias se alimentan a continuación a tres cajas 1  $\rightarrow$  2 402-404 que generan un total de  $N=6$  señales 405:  $lr$  (frontal izquierda),  $ls$  (surround izquierda),  $rr$  (frontal derecha),  $rs$  (surround derecha),  $c$  (central), y  $lfe$  (extensión de baja frecuencia). Esta mezcla ascendente depende de un subconjunto de los parámetros espaciales transmitidos que corresponde al conjunto 1 en la figura 2. La salida de audio digital multicanal final se crea pasando las seis señales de subbanda a seis bancos de filtros de síntesis.

**[0057]** La figura 5 ilustra el problema a resolver por la compensación de ganancia según la invención. El espectro de una salida binaural HRTF filtrada de referencia para el oído izquierdo se representa como un gráfico sólido. El gráfico de trazos representa el espectro de la señal decodificada correspondiente tal como se genera mediante el procedimiento de la figura 2, en el caso en que el combinador 203 consiste en el combinador lineal 301 solamente. Como se puede ver, hay una pérdida de energía espectral sustancial en relación con el espectro de referencia deseado en los intervalos de frecuencia de 3-4 kHz y 11-13 kHz. Hay también un impulso más pequeño espectral alrededor de 1 kHz y 10 kHz.

**[0058]** La figura 6 ilustra la ventaja de utilizar la compensación de ganancia según la invención. El gráfico sólido es el mismo espectro de referencia que en la figura 5, pero ahora el gráfico de trazos representa el espectro de la señal

decodificada tal como se genera mediante el procedimiento de la figura 2, en el caso en que el combinador 203 consiste en todos los componentes de la figura 3. Como se puede ver, hay una coincidencia espectral significativamente mejorada entre las dos curvas en comparación con la de las dos curvas de la figura 5.

- 5 **[0059]** En el texto que sigue, se pondrá de relieve la descripción matemática de la compensación de ganancia según la invención. Para señales discretas complejas  $x, y$  Para señales discretas complejas  $x, y$ , el producto interno complejo y la norma al cuadrado (energía) se definen mediante

$$\left\{ \begin{array}{l} \langle x, y \rangle = \sum_k x(k) \bar{y}(k), \\ X = \|x\|^2 = \langle x, x \rangle = \sum_k |x(k)|^2, \\ Y = \|y\|^2 = \langle y, y \rangle = \sum_k |y(k)|^2, \end{array} \right. \quad (1)$$

donde  $y(k)$  indica la señal conjugada compleja de  $y(k)$ .

- 10 **[0060]** La señal multicanal original consta de  $N$  canales, y cada canal tiene un par de filtros binaurales relacionados con HRTF asociado este. No obstante, se supone aquí que la señal multicanal paramétrica se crea con un paso intermedio de mezcla ascendente predictivo de los  $M$  canales transmitidos a los  $P$  canales predichos. Esta estructura se utiliza en MPEG Surround tal como se describe en la figura 4. Se supondrá que el conjunto original de  $2N$  filtros  
15 relacionados con HRTF se han reducido mediante el precombinador del estado de la técnica 202 a un par de filtros para cada uno de los  $P$  canales predichos en los que  $M \leq P \leq N$ . Las  $P$  señales de canales predichos  $x_p$ ,  $p=1,2,\dots,P$ , tienen por objeto aproximar las  $P$  señales  $x_p$   $p=1,2,\dots,P$ , que se derivan de los  $N$  canales originales a través de la mezcla descendente original. En MPEG Surround, estas señales son una izquierda combinada, una combinada derecha y un canal central combinado y escalado. Se supone que el par de filtros HRTF correspondiente a la señal  
20  $x_p$  se describe mediante un filtro de subbanda  $b_{1,p}$  para el oído izquierdo y un filtro de sub-banda  $b_{2,p}$  para el oído derecho. La señal de salida binaural de referencia está por lo tanto dada por la superposición lineal de señales filtradas para  $n=1,2$ ,

$$y_n(k) = \sum_{p=1}^P (b_{n,p} * x_p)(k) , \quad (2)$$

- 25 donde el asterisco indica convolución en la dirección temporal. Los filtros de sub-banda se pueden dar en forma de filtros de respuesta de impulsos finitos (FIR), filtros de respuesta de impulso infinito (IIR) o que se derivan de una familia parametrizada de filtros.

- [0061]** En el codificador, la mezcla descendente se forma mediante la aplicación de una matriz  $M \times P$  de mezcla descendente  $\mathbf{D}$  a un vector columna de señales formado por  $x_p$   $p=1,2,\dots,P$  y la predicción en el decodificador se  
30 realiza mediante la aplicación de una  $P \times M$  matriz de predicción  $\mathbf{C}$  al vector columna de señales formado por los  $M$  canales mezclados hacia abajo transmitidos  $z_m$   $m=1,\dots,M$ .

$$\hat{x}_p(k) = \sum_{m=1}^M c_{p,m} z_m(k) , \quad (3)$$

Ambas matrices son conocidas en el decodificador, e ignorando los efectos de la codificación de los canales mezclados hacia abajo, el efecto combinado de predicción se puede modelizar mediante

$$\hat{x}_p(k) = \sum_{q=1}^P a_{p,q} x_q(k) , \quad (4)$$

- 35 donde  $a_{p,q}$  con las entradas del producto de matrices  $\mathbf{A}=\mathbf{CD}$ .

**[0062]** Un procedimiento sencillo para producir una salida binaural en el decodificador es simplemente insertar las señales predichas  $x_p$  en (2) lo que da como resultado

$$\hat{y}_n(k) = \sum_{p=1}^P (b_{n,p} * \hat{x}_p)(k) . \quad (5)$$

**[0063]** En términos de cálculos, el filtrado binaural se combina con el mezcla ascendente predictiva de antemano de tal manera que (5) puede escribirse como

$$\hat{y}_n(k) = \sum_{m=1}^M (h_{n,m} * z_m)(k) , \quad (6)$$

5 con los filtros combinados definidos por

$$h_{n,m}(k) = \sum_{p=1}^P c_{p,m} b_{n,p}(k) . \quad (7)$$

**[0064]** Esta fórmula describe la acción del combinador lineal 301 que combina los coeficientes  $c_{p,m}$  derivados de parámetros espaciales con los filtros de dominio de sub-banda binaurales  $b_{n,p}$ . Cuando las  $P$  señales originales  $x_p$  tienen un rango numérico esencialmente delimitado por  $M$ , la predicción puede ser diseñada para funcionar muy

10 bien y la aproximación  $\hat{x}_p \approx \hat{x}_p$  es válida. Esto sucede por ejemplo si sólo  $M$  de los  $P$  canales están activos, o si los componentes de señal importantes se originan a partir de expansión de amplitud. En este caso, la señal binaural descodificada (5) es una muy buena coincidencia con la referencia (2). Por otro lado, en el caso general y especialmente en el caso en que las  $P$  señales originales  $x_p$  no están correlacionadas, habrá una pérdida sustancial de predicción y la salida de (5) puede tener una energía que se desvía considerablemente de la energía de (2). Como la desviación será diferente en diferentes bandas de frecuencia, la salida de audio final sufre de artefactos colorantes espectrales tal como se describe en la figura 5. La presente invención enseña cómo eludir este problema mediante la compensación de ganancia de la salida de acuerdo con

$$\tilde{y}_n = g_n \cdot \hat{y}_n . \quad (8)$$

20 **[0065]** En términos de cálculos, la compensación de ganancia se realiza ventajosamente mediante la alteración de los filtros combinados de acuerdo con el ajustador de ganancia 303,  $h_{n,m}(k) = g_n h_{n,m}(k)$ . El filtrado combinado modificado se convierte entonces en

$$\tilde{y}_n(k) = \sum_{m=1}^M (\tilde{h}_{n,m} * z_m)(k) . \quad (9)$$

**[0066]** Los valores óptimos de las ganancias de compensación en (8) son

$$g_n = \frac{\|y_n\|}{\|\hat{y}_n\|} . \quad (10)$$

25 **[0067]** El propósito del calculador de ganancia 302 es estimar estas ganancias a partir de la información disponible en el descodificador. A continuación se describen varias herramientas para este fin. La información disponible está representado aquí por las entradas de matriz  $a_{p,q}$  y los filtros relacionados con HRTF de sub-banda  $b_{n,p}$ . En primer lugar, se supondrá la siguiente aproximación para el producto interno entre las señales  $x,y$  que se han filtrado mediante filtros relacionados con HRTF de sub-banda  $b,d$ ,

$$\langle b * x, d * y \rangle \approx \langle b, d \rangle \langle x, y \rangle . \quad (11)$$

**[0068]** Esta aproximación se basa en el hecho de que a menudo la mayor parte de la energía de los filtros se concentra en un solo intervalo dominante, que a su vez supone que el paso de tiempo de la transformada de tiempo frecuencia aplicada es suficientemente grande en comparación con las diferencias de retardo principales de filtros HRTF. Aplicando la aproximación (11) en combinación con (2) conduce a

$$\|y_n\|^2 \approx \sum_{p,q=1}^P \langle b_{n,p}, b_{n,q} \rangle \langle x_p, x_q \rangle. \quad (12)$$

5 **[0069]** La siguiente aproximación consiste en asumir que las señales originales no están correlacionadas, es decir  $\langle x_p, x_q \rangle = 0$  para  $p \neq q$ . Entonces (12) se reduce a

$$\|y_n\|^2 \approx \sum_{p=1}^P \|b_{n,p}\|^2 \|x_p\|^2. \quad (13)$$

**[0070]** Para la energía decodificada el resultado correspondiente a (12) es

$$\|\hat{y}_n\|^2 \approx \sum_{p,q=1}^P \langle b_{n,p}, b_{n,q} \rangle \langle \hat{x}_p, \hat{x}_q \rangle. \quad (14)$$

10 **[0071]** La inserción de las señales predichas (4) en (14) y la aplicación de la presunción de que las señales originales no están correlacionadas da

$$\|\hat{y}_n\|^2 \approx \sum_{p=1}^P \left( \sum_{q,r=1}^P a_{q,p} a_{r,p} \langle b_{n,q}, b_{n,r} \rangle \right) \|x_p\|^2. \quad (15)$$

15 **[0072]** Lo que queda con el fin de ser capaz de calcular la ganancia de compensación dada por el cociente (10) es estimar la distribución de energía  $\|x_p\|^2$ ,  $p = 1, 2, \dots, P$  de los canales originales hasta un factor arbitrario. La presente invención enseña a hacer esto mediante la computación, como una función de la distribución de energía, correspondiendo la matriz de predicción  $\mathbf{C}_{\text{model}}$  a la suposición de que estos canales no están correlacionados y que el codificador tiene por objeto reducir al mínimo el error de predicción. La distribución de energía se estima entonces resolviendo el sistema de ecuaciones no lineal  $\mathbf{C}_{\text{model}} = \mathbf{C}$  si es posible. Para parámetros de predicción que llevarían a un sistema de ecuaciones sin soluciones, los factores de compensación de ganancia se establecen como  $g_n = 1$ . Este procedimiento se detallará en el siguiente apartado en el caso especial más importante.

**[0073]** La carga de cálculo impuesta por (15) se puede reducir en el caso en que  $P = M + 1$  mediante la aplicación de la expansión (véase, por ejemplo PCT/EP2005/011586),

$$\langle x_p, x_q \rangle = \langle \hat{x}_p, \hat{x}_q \rangle + \Delta E \cdot v_p \cdot v_q, \quad (16)$$

25 donde  $v$  es un vector unitario con componentes  $v_p$  de modo que  $\mathbf{D}v = 0$ , y  $\Delta E$  es la energía de la pérdida de predicción,

$$\Delta E = E - \hat{E} = \sum_{p=1}^P \|x_p\|^2 - \sum_{p=1}^P \|\hat{x}_p\|^2. \quad (17)$$

**[0074]** El cálculo de (15) se sustituye ventajosamente por la aplicación de (16) en (14), dando lugar a

$$\|\hat{y}_n\|^2 \approx \|y_n\|^2 - \Delta E \cdot \left\| \sum_{p=1}^P v_p b_{n,p} \right\|^2. \quad (18)$$

30 **[0075]** Posteriormente, se discutirá una especialización preferente a la predicción de los tres canales a partir de dos canales. Se utiliza el caso en el que  $M = 2$  y  $P = 3$  en MPEG Surround. Las señales son un canal izquierdo

combinada  $x_1=l$ , un canal combinado derecho  $x_2=r$  y un canal combinado central izquierdo (escalado)  $x_3=c$ . La matriz de mezcla descendente es

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad (19)$$

y la matriz de predicción se construye a partir de dos parámetros  $c_1, c_2$  reales transmitidos, según

$$\mathbf{C} = \frac{1}{3} \begin{bmatrix} 2+c_1 & c_2-1 \\ c_1-1 & 2+c_2 \\ 1-c_1 & 1-c_2 \end{bmatrix}. \quad (20)$$

- 5 **[0076]** Suponiendo que los canales originales no están correlacionados la matriz de predicción que consigue el error de predicción mínimo se da mediante:

$$\mathbf{C}_{\text{model}} = \frac{1}{LC+RC+LR} \begin{bmatrix} LC+LR & -LC \\ -RC & RC+LR \\ RC & LC \end{bmatrix}. \quad (21)$$

- 10 **[0077]** Imponiendo la igualdad  $\mathbf{C}_{\text{model}} = \mathbf{C}$  se llega a la distribución de energía (no normalizada) enseñada por la presente invención

$$\begin{bmatrix} L \\ R \\ C \end{bmatrix} = \begin{bmatrix} \beta(1-\sigma) \\ \alpha(1-\sigma) \\ p \end{bmatrix}, \quad (22)$$

donde  $\alpha = (1-c_1) / 3$ ,  $\beta = (1-c_2) / 3$ ,  $\sigma = \alpha + \beta$ , y  $p = \alpha\beta$ . Esto se mantiene en el intervalo viable definido por

$$\alpha > 0, \beta > 0, \sigma < 1, \quad (23)$$

en cuyo caso el error de predicción se puede encontrar en la misma escala a partir de

$$\Delta E = 3p(1-\sigma). \quad (24)$$

- 15 **[0078]** Puesto que  $P=3=2+1=M+1$ , el procedimiento descrito por (16) - (18) es aplicable. El vector unitario es

$$[v_1, v_2, v_3] = [1, 1, -1] / \sqrt{3}$$

y con las definiciones

$$\Delta E_n^B = p(1-\sigma) \|b_{n,1} + b_{n,2} - b_{n,3}\|^2, \quad (25)$$

- 20 Y

$$E_n^B = \beta(1-\sigma)\|b_{n,1}\|^2 + \alpha(1-\sigma)\|b_{n,2}\|^2 + p\|b_{n,3}\|^2, \quad (26)$$

la ganancia compensación para cada oído  $n=1,2$  tal como se ha calculado en una realización preferida del calculador de ganancia 302 se puede expresar mediante:

$$g_n = \begin{cases} \min \left\{ g_{\max}, \sqrt{\frac{E_n^B + \varepsilon}{E_n^B - \Delta E_n^B + \varepsilon}} \right\}, & \text{if } \alpha > 0, \beta > 0, \sigma < 1; \\ 1, & \text{En los demás casos.} \end{cases} \quad (27)$$

- 5 **[0079]** Aquí  $\varepsilon > 0$  es un número pequeño, cuyo propósito es estabilizar la fórmula cerca del borde de la gama de parámetros viables y  $g_{\max}$  es un límite superior a la ganancia de compensación aplicado. Las ganancias de (27) son diferentes para los oídos izquierdo y derecho,  $n = 1, 2$ . Una variante del procedimiento es utilizar una ganancia común  $g_0 = g_1 = g$ , donde

$$g = \begin{cases} \min \left\{ g_{\max}, \sqrt{\frac{E_0^B + E_1^B + \varepsilon}{E_0^B + E_1^B - \Delta E_0^B - \Delta E_1^B + \varepsilon}} \right\}, & \text{if } \alpha > 0, \beta > 0, \sigma < 1; \\ 1, & \text{En los demás casos.} \end{cases} \quad (28)$$

- 10 **[0080]** El factor de ganancia de corrección de la invención puede ser llevado a la coexistencia con una compensación ganancia multicanal directa hacia adelante disponible sin ningún problema relacionados con HRTF.

- 15 **[0081]** En MPEG Surround, la compensación para la pérdida de predicción ya se aplica en el decodificador multiplicando la matriz de mezcla ascendente  $\mathbf{C}$  por un factor  $1/p$  donde  $0 < p \leq 1$  es una parte de los parámetros espaciales transmitidos. En ese caso, las ganancias de (27) y (28) tienen que ser sustituidos por los productos  $pg_n$  y  $pg$  respectivamente. Dicha compensación se aplica para la decodificación binaural estudiada en las figuras 5 y 6. Esa es la razón por la cual la decodificación de la técnica anterior de la figura 5 aumenta partes del espectro en comparación con la referencia. Para las sub-bandas correspondientes a dichas regiones de frecuencia, la
- 20 compensación de ganancia de la invención sustituye eficazmente el factor de ganancia de parámetro transmitido  $1/p$  por un valor más pequeño derivado de fórmula (28).

- [0082]** Además, puesto que en el caso en que  $p = 1$  corresponde a una predicción exitosa, una variante más conservadora de la compensación de ganancia enseñada por la presente invención desactivará la compensación de
- 25 ganancia binaural para  $p = 1$ .

- [0083]** Además, la presente invención se puede utilizar junto con una señal residual. En MPEG Surround, se puede transmitir una señal residual de predicción adicional  $z_3$  que permite reproducir las  $P=3$  señales originales  $x_p$  más fielmente. En este caso, la compensación de ganancia debe ser reemplazada por una adición de la señal residual
- 30 binaural que se describe a continuación. La mezcla ascendente predictiva aumentada por un residuo se forma mediante

$$\tilde{x}_p(k) = \sum_{m=1}^2 c_{p,m} z_m(k) + w_p \cdot z_3(k), \quad (29)$$

Donde  $[w_1, w_2, w_3] = [1, 1, -1]/3$ . Sustituyendo  $x_p$  por  $\tilde{x}_p$  en (5) se obtiene el filtrado combinado correspondiente,

$$\tilde{y}_n(k) = \sum_{m=1}^3 (h_{n,m} * z_m)(k), \quad (30)$$

donde los filtros combinados  $h_{n,m}$  se definen mediante (7) para  $m = 1,2$ , y los filtros combinados para la adición residual se definen mediante

$$h_{n,3} = \frac{1}{3}(b_{n,1} + b_{n,2} - b_{n,3}) . \quad (31)$$

**[0084]** Por lo tanto, la estructura global de este modo de decodificación se describe también en la figura 2 estableciendo  $P = M = 3$ , y modificando el combinador 203 para realizar sólo la combinación lineal definida por (7) y (31).

**[0085]** La figura 13 ilustra en una representación modificada el resultado del combinador lineal 301 en la figura 3. El resultado del combinador son cuatro filtros basados en HRTF  $h_{11}$ ,  $h_{12}$ ,  $h_{21}$  y  $h_{22}$ . Como será más claro a partir de la descripción de la figura 16a y la figura 17, estos filtros corresponden a los filtros indicados por 15, 16, 17 18 en la figura 16a.

**[0086]** La figura 16a muestra una cabeza de un oyente que tiene un oído izquierdo o un punto binaural izquierdo y que tiene un oído derecho o un punto binaural derecho. Cuando la figura 16a corresponda solamente a un escenario estéreo, entonces los filtros 15, 16, 17, 18 serían típicas Funciones de Transferencia Relativas a la Cabeza que pueden medirse individualmente u obtenerse a través de Internet o en libros de texto correspondientes para diferentes posiciones entre el oyente y el altavoz de canal izquierdo y el altavoz del canal derecho.

**[0087]** Sin embargo, puesto que la presente invención está dirigida a un decodificador binaural multi-canal, los filtros ilustrados por 15, 16, 17, 18 no son filtros HRTF puros, sino que son filtros basados en HRTF, que no sólo reflejan propiedades de HRTF sino que también dependen de los parámetros espaciales y, en particular, como se discute en relación con la figura 2, dependen del conjunto de parámetros espaciales 1 y el conjunto de parámetros espaciales 2.

**[0088]** La figura 14 muestra la base de los filtros basados en HRTF utilizados en la figura 16a. En particular, se ilustra una situación en la que un oyente se coloca en un punto dulce entre cinco altavoces en una configuración de altavoces de cinco canales que se pueden encontrar, por ejemplo, en típicos sistemas envolventes de entretenimiento de cine o domésticos. Para cada canal, existen dos HRTF que se pueden convertir para canalizar respuestas de impulso de un filtro que tiene la HRTF como la función de transferencia. Particularmente, como se conoce en la técnica, un filtro basado en HRTF da cuenta de la propagación del sonido dentro de la cabeza de una persona de manera que, por ejemplo, el HRTF1 en la figura 14 da cuenta de la situación en que un sonido emitido por el altavoz Ls se encuentra con el oído derecho tras haber pasado alrededor de la cabeza del oyente. Contrariamente este, el sonido emitido por el altavoz de sonido envolvente izquierdo Ls se encuentra con el oído izquierdo casi directamente y está sólo parcialmente afectado por la posición del oído en la cabeza y también la forma de la oreja etc. Por lo tanto, se hace evidente que las HRTF 1 y 2 son diferentes entre sí.

**[0089]** Lo mismo es cierto para los HRTFs 3 y 4 para el canal izquierdo, ya que las relaciones de ambos oídos con el canal izquierdo L son diferentes. Esto también se aplica para todos los otros HRTFs, aunque, como se aclara en la figura 14, los HRTF 5 y 6 para el canal central serán casi idénticos o incluso completamente idénticos entre sí, a menos que la asimetría de los oyentes individuales se acomode mediante los datos de HRTF.

**[0090]** Como se dijo anteriormente, estos HRTFs se han determinado para las cabezas modelo y pueden ser descargados por cualquier " cabeza promedio " específica, y la configuración de altavoces.

**[0091]** Ahora, como se aclara en 171 y 172 en la figura 17, se lleva a cabo una combinación para combinar el canal izquierdo y el canal envolvente izquierdo para obtener dos filtros basados en HRTF para el lado izquierdo indicado por L 'en la figura 15. Se lleva a cabo el mismo procedimiento para el lado derecho tal como se ilustra por R ' en la figura 15 que se traduce en HRTF 13 y HRTF 14. Para este fin, también se hace referencia al punto 173 y al punto 174 en la figura 17. Sin embargo, es de notar aquí que, para combinar los respectivos HRTF en los elementos 171, 172, 173 y 174, se contabilizan los parámetros de diferencia de nivel entre canales que reflejan la distribución de energía entre el canal L y el canal Ls de la configuración original o entre el canal R y el canal Rs de la configuración original de multicanal. En particular, estos parámetros definen un factor de ponderación cuando se combinan linealmente mediante HRTF.

**[0092]** Como se ha indicado antes, también se puede aplicar un factor de fase cuando se combinan HRTF, cuyo factor de fase se define por retardos de tiempo o diferencias de fase desplegadas entre los HRTF a combinar. Sin embargo, este factor de fase no depende de los parámetros de transmisión.

**[0093]** Por lo tanto, los HRTF 11, 12, 13 y 14 no son verdaderos filtros HRTF sino que son filtros basados en HRTF, puesto que estos filtros no sólo dependen de los HRTF, sino que son independientes de la señal transmitida. En lugar de ello, los HRTF 11, 12, 13 y 14 son también dependientes de la señal transmitida debido a que los Parámetros de Diferencia de nivel de Canales  $cl_{d1}$  y  $cl_{d2}$  se utilizan para calcular estos HRTFs 11, 12, 13 y 14.

**[0094]** Ahora se obtiene la situación de la figura 15, que todavía tiene tres canales en lugar de dos canales transmitidos tal como se incluyen en una señal de mezcla descendente preferida. Por lo tanto, se debe hacer una combinación de los seis HRTFs 11, 12, 5, 6, 13, 14 en cuatro HRTFs 15, 16, 17, 18 tal como se ilustra en la figura 16a.

**[0095]** Para este fin, los HRTFs 11, 5, 13 se combinan empleando una regla de mezcla ascendente izquierda, que se vuelve clara a partir de la matriz de mezcla ascendente de la figura 16b. En particular, la regla de mezcla ascendente izquierda tal como se muestra en la figura 16b y como se indica en el bloque 175 incluye parámetros  $m_{11}$ ,  $m_{21}$  y  $m_{31}$ . Esta regla de mezcla ascendente izquierda está en la ecuación de la matriz de la figura 16 sólo para que se multiplique por el carril izquierdo. Por lo tanto, estos tres parámetros se denominan la regla de mezcla ascendente izquierda.

**[0096]** Como se indica en el bloque 176, se combinan los mismos HRTFs 11, 5, 13, pero ahora usando la regla de mezcla ascendente derecha, es decir, en la realización de la figura 16b, los parámetros  $m_{12}$ ,  $m_{22}$  y  $m_{32}$ , que todos se utilizan para ser multiplicados por el canal derecho  $R_0$  en la figura 16b.

**[0097]** De este modo, se generan HRTF 15 y HRTF 17. Análogamente HRTF 12, HRTF 6 y HRTF 14 de la figura 15 se combinan utilizando los parámetros de mezcla ascendente izquierda  $m_{11}$ ,  $m_{21}$  y  $m_{31}$  para obtener HRTF 16. Una combinación correspondiente se realiza empleando HRTF 12, HRTF 6 HRTF 14, pero ahora con los parámetros de mezcla ascendente derecha o regla de mezcla ascendente derecha indicados por  $m_{12}$ ,  $m_{22}$  y  $m_{32}$  para obtener HRTF 18 de la figura 16a.

**[0098]** De nuevo, se enfatiza que, mientras los HRTF originales de la figura 14 no dependían todos de la señal transmitida, los nuevos filtros basados en HRTF 15, 16, 17, 18 dependen ahora de la señal transmitida, ya que los parámetros espaciales incluidos en la señal multi-canal se utilizaron para el cálculo de estos filtros 15, 16, 17 y 18.

**[0099]** Para obtener finalmente un canal binaural izquierdo  $L_b$  y un canal binaural derecho  $R_b$  las salidas de los filtros 15 y 17 se deben combinar en un sumador 130a. de manera análoga, la salida de los filtros 16 y 18 se deben combinar en un sumador 130b. Estos sumadores 130a, 130b reflejan la superposición de dos señales en el oído humano.

**[0100]** A continuación se discutirá la figura 18. La figura 18 muestra una forma de realización preferida de un decodificador multicanal de la invención para generar una señal binaural empleando una señal de mezcla descendente derivada de una señal multi-canal original. La señal de mezcla descendente se ilustra en  $z_1$  y  $z_2$  o también se indica con "L" y "R". Además, la señal de mezcla descendente tiene parámetros asociados con ella, parámetros que incluyen al menos una Diferencia de nivel de Canales para canales izquierdo y surround izquierdo o una Diferencia de nivel de Canales para canales derecho y derecho surround e información acerca de la regla de mezcla ascendente.

**[0101]** Naturalmente, cuando la señal multicanal original era sólo señal de tres canales,  $cld_1$  o  $cld_r$  no se transmiten y la única información lateral paramétrica será información sobre la regla de mezcla ascendente que, como se indicó antes, es una regla de mezcla ascendente tal que resulta en un error de energía en la señal de mezcla ascendente. Así, aunque las formas de onda de las señales de mezcla ascendente cuando se realiza una representación no binaural, coincide lo mejor posible con las formas de onda originales, la energía de los canales de mezcla ascendente es diferente de la energía de los canales originales correspondientes.

**[0102]** En la realización preferida de la figura 18, la información de regla de mezcla ascendente se refleja por dos parámetros de mezcla ascendente  $cpc_1$ ,  $cpc_2$ . Sin embargo, se podría aplicar cualquier otra información de regla de mezcla ascendente y se señaliza a través de un cierto número de bits. En particular, se podrían señalar ciertos escenarios de mezcla ascendente y parámetros de mezcla ascendente mezcla ascendente utilizando una tabla predeterminada en el decodificador de manera que sólo los índices de la tabla tengan que transmitirse desde un codificador al decodificador. Como alternativa, también se podría utilizar diferentes escenarios de mezcla ascendente tal como un mezcla ascendente de dos a más de tres canales. Como alternativa, también se podrían transmitir más de dos parámetros de mezcla ascendente predictiva que requeriría entonces una regla de mezcla descendente diferente correspondiente que tendría que encajar con la regla de mezcla ascendente tal como se discutirá con más detalle con respecto a la figura 20.

**[0103]** Independientemente de dicha realización preferida para la información de regla de mezcla ascendente, cualquier información de regla de mezcla ascendente es suficiente, siempre y cuando sea posible una mezcla ascendente para generar un conjunto afectado de pérdida de energía de los canales de mezcla ascendente, que es la forma de onda a juego con el conjunto correspondiente de señales originales.

**[0104]** El decodificador multicanal de la invención incluye un calculador de factor de ganancia 180 para calcular al menos un factor de ganancia  $g_1$ ,  $g_r$  o  $g$ , para reducir o eliminar el error de energía. El calculador de factor de ganancia calcula el factor de ganancia a partir de la información de regla de mezcla ascendente y características de filtros de filtros basados en HRTF correspondientes a canales de mezcla ascendente que se podrían obtener,

cuando se aplicase la regla de mezcla ascendente. Sin embargo, como se indicó antes, en la representación binaural, esta mezcla ascendente no tiene lugar. Sin embargo, como se analiza en relación con la figura 15 y los bloques 175, 176, 177, 178 de la figura 17, se utilizaron no obstante filtros basados en HRTF correspondientes a estos canales de mezcla ascendente.

**[0105]** Como se discutió antes, la calculadora de factor de ganancia 180 puede calcular diferentes factores de ganancia  $g_1$  y  $g_r$  como se indica en la ecuación (27), cuando, en lugar de  $n$ , se inserta 1 o  $r$ . Como alternativa, el calculador de factor de ganancia podría generar un único factor de ganancia para ambos canales tal como se indica mediante la ecuación (28).

**[0106]** Es importante destacar que el calculador de factor de ganancia 180 de la invención calcula el factor de ganancia no solo a partir de la regla de mezcla ascendente, sino también a partir de las características de filtro de los filtros basados en HRTF correspondientes a canales de mezcla ascendente. Esto refleja la situación que los propios filtros también dependen de las señales transmitidas y también se ven afectados por un error de energía. Por lo tanto, el error de energía no sólo está causado por la información de regla de mezcla ascendente tal como los parámetros de predicción  $CPC_1$ ,  $CPC_2$ , sino que también se ve influido por los propios filtros.

**[0107]** Por lo tanto, para obtener una corrección de ganancia bien adaptada, el factor de ganancia de la invención no sólo depende del parámetro de predicción, sino que también depende de los filtros correspondientes a los canales de mezcla ascendente.

**[0108]** El factor de ganancia y los parámetros de mezcla descendente así como los filtros basados en HRTF se usan en el procesador de filtro 182 para filtrar la señal de mezcla descendente para obtener una señal binaural con energía corregida que tiene un canal binaural izquierdo  $L_b$  y que tiene un canal binaural derecho  $R_b$ .

**[0109]** En una realización preferida, el factor de ganancia depende de una relación entre la energía total incluida en las respuestas de impulso de canal de los filtros correspondientes a los canales de mezcla ascendente y una diferencia entre esta energía total y un error de energía de mezcla ascendente estimado  $\Delta E$ .  $\Delta E$  puede calcularse preferentemente mediante la combinación de las respuestas de impulsos de canal de los filtros correspondientes a canales de mezcla ascendente y para calcular entonces la energía de la respuesta de impulso de canal combinada. Dado que todos los números en las relaciones de  $G_L$  y  $G_R$  en la figura 18 son números positivos, que se hace evidente a partir de las definiciones de  $\Delta E$  y  $E$  es evidente que ambos factores de ganancia son más grandes que 1. Esto refleja la experiencia que se ilustra en la figura 5 que, en la mayoría de las veces, la energía de la señal binaural es menor que la energía de la señal multicanal original. También es de notar, que incluso cuando se aplica la compensación de ganancia multi-canal, es decir, cuando se utiliza el factor  $p$  en la mayoría de las señales, se provoca sin embargo una pérdida de energía.

**[0110]** La figura 19a ilustra una realización preferida del procesador de filtro 182 de la figura 18. En particular, la figura 19a ilustra la situación, cuando en el bloque 182a se utilizan los filtros combinados 15, 16, 17, y 18 de la figura 16a sin compensación de ganancia y las señales de salida de filtro se añaden tal como se indica en la figura 13. Entonces, la salida de la caja 182a se introduce en una caja de escalado 182b para escalar la salida utilizando el factor de ganancia calculado por la caja 180.

**[0111]** Como alternativa, el procesador de filtro puede ser construido tal como se muestra en la figura 19b. Aquí, los filtros HRTFs 15 a 18 se calculan tal como se ilustra en la caja 182c. Por lo tanto, el calculador 182c realiza la combinación HRTF sin ajuste de ganancia alguno. A continuación, se proporciona un ajustador de filtro 182d, que utiliza el factor de ganancia calculado de la invención. Los resultados del ajustador de filtro en filtros ajustados como se muestra en el bloque 180e, donde el bloque 180e realiza el filtrado mediante el filtro ajustado y lleva a cabo la posterior adición de la salida del filtro correspondiente tal como se muestra en la figura 13. Por lo tanto, no es necesario post-escalado como en la figura 19a para obtener canales binaurales de ganancia corregida  $L_b$  y  $R_b$ .

**[0112]** En general, como se ha descrito en relación con la ecuación 16, la ecuación 17 y la ecuación 18, el cálculo de la ganancia se lleva a cabo usando el error de mezcla ascendente estimado  $\Delta E$ . Esta aproximación es especialmente útil para el caso en que el número de canales de mezcla ascendente es igual al número de canales de mezcla descendente +1. Así, en caso de dos canales de mezclado descendente, esta aproximación funciona bien para tres canales de mezcla ascendente. Como alternativa, cuando se tuvieran tres canales de mezclado descendente, esta aproximación también funcionaría bien en un escenario en el que hay cuatro canales de mezcla ascendente.

**[0113]** Sin embargo, es de notar que el cálculo del factor de ganancia en base a una estimación del error de mezcla ascendente también se puede realizar para escenarios en los que por ejemplo, cinco canales se predicen a partir de tres canales de mezclado descendente. Como alternativa, también se podría utilizar una mezcla ascendente basada en predicción a partir de dos canales de mezclado descendente a cuatro canales de mezcla ascendente. En cuanto al error de energía de mezcla ascendente estimado  $\Delta E$ , no solo se puede calcular directamente este error estimado tal como se indica en la ecuación (25) para el caso preferido, sino que también se podría transmitir alguna información sobre el error de mezcla ascendente real en un flujo de bits. Sin embargo, incluso en otros casos que en

el caso especial como el ilustrado en relación con las ecuaciones (25) a (28), se podría entonces calcular el valor  $E_b$  a partir de los filtros basados en HRTF para los canales de mezcla ascendente empleando parámetros de predicción. Cuando se considera la ecuación (26), se hace claro que esta ecuación también puede ser fácilmente aplicada a un esquema de mezcla ascendente de predicción 2/4, cuando los factores de ponderación para las energías de las respuestas de impulso del filtro basado en HRTF se adaptan correspondientemente.

[0114] En vista de ello, se hace evidente que la estructura general de la ecuación (27), es decir, calcular el factor de ganancia a partir de la relación de  $E_b/(E_b - \Delta E_b)$  también se aplica para otros escenarios.

[0115] Posteriormente, se discutirá la figura 20 para mostrar una implementación esquemática de un codificador basado en predicciones que podría ser utilizado para generar la señal de mezcla descendente L, R y la información de regla de mezcla ascendente transmitida a un descodificador de modo que el descodificador pueda realizar la compensación de ganancia en el contexto del procesador de filtro binaural.

[0116] Un mezclador descendente 191 recibe cinco canales originales o, como alternativa, tres canales originales tal como se ilustra mediante ( $L_s$  y  $R_s$ ). El mezclador descendente 191 puede funcionar a partir de una regla de mezcla descendente predeterminada. En este caso, la indicación de regla de mezcla descendente tal como se ilustra mediante la línea 192 no es necesaria. Naturalmente, el minimizador de error 193 podría variar la regla de mezcla descendente también para minimizar el error entre canales reconstruidos en la salida de un mezclador descendente 194 con respecto a los canales de entrada originales correspondientes.

[0117] Por lo tanto, el minimizador de error 193 puede variar la regla de mezcla descendente 192 o la regla de mezclador ascendente 196 de modo que los canales reconstruidos tengan una pérdida de predicción mínima  $\otimes E$ . Este problema de optimización se resuelve mediante cualquiera de los algoritmos bien conocidos dentro del minimizador de error 193, que opera preferentemente en un modo a nivel de sub-banda para minimizar la diferencia entre los canales de reconstrucción y los canales de entrada.

[0118] Como se dijo antes, los canales de entrada pueden ser canales originales L,  $L_s$ , R,  $R_s$ , C. Como alternativa, los canales de entrada puede ser solamente tres canales L, R, C, en los que, en este contexto, los canales de entrada L, R, pueden derivarse por las correspondientes cajas OTT ilustradas en la figura 11. Como alternativa, cuando la señal original solo tiene canales L, R, C, entonces también se puede llamar a estos canales "canales originales".

[0119] La figura 20 ilustra además que se puede utilizar cualquier información de regla de mezcla ascendente aparte de la transmisión de dos parámetros de predicción siempre que un descodificador esté en la posición para realizar una mezcla ascendente empleando esta información de regla de mezcla ascendente. Por lo tanto, la información de regla de mezcla ascendente también puede ser una entrada en una tabla de consulta o cualquier otra información relacionada con la mezcla ascendente.

[0120] Por consiguiente, la presente invención, proporciona una manera eficaz de realizar la decodificación binaural de señales de audio multi-canal a partir de señales de mezcla descendente disponibles y datos de control adicionales mediante HRTF filtrar. La presente invención proporciona una solución al problema de la coloración espectral derivada de la combinación de mezcla ascendente predictiva con la decodificación binaural.

[0121] Ejemplos relacionados con un descodificador multi-canal para generar una señal binaural a partir de una señal de mezcla descendente derivada de una señal multi-canal original empleando parámetros que incluyen una información de regla de mezcla ascendente utilizable para mezclar hacia arriba la señal de mezcla descendente con una regla de mezcla ascendente, dando como resultado la regla de mezcla ascendente un error de energía, comprenden:

un calculador de factor de ganancia (180, 302) para calcular al menos un factor de ganancia ( $g_1$ ,  $g_2$ ) para reducir o eliminar el error de energía, a partir de la información de regla de mezcla ascendente y características de filtros de filtros a base de función de transferencia relativa a la cabeza correspondiente a canales de mezcla ascendente; y un procesador de filtro (182) para filtrar la señal de mezcla descendente empleando el al menos un factor de ganancia ( $g_1$ ,  $g_2$ ), las características de filtro y la información de regla de mezcla ascendente para obtener una señal binaural con energía corregida.

[0122] Según otro ejemplo, el procesador de filtro (182) es capaz de calcular coeficientes de filtro para dos filtros ajustados de ganancia para cada canal de la señal de mezcla descendente y para filtrar el canal de mezcla descendente empleando cada uno de los dos filtros ajustados de ganancia.

[0123] Según otro ejemplo, el procesador de filtro (182) es capaz de calcular coeficientes de filtro para dos filtros para cada canal de el canal de mezcla descendente sin emplear el factor de ganancia y para filtrar los canales de mezcla descendente y ajustar la ganancia tras filtrar el canal de mezcla descendente.

[0124] Según otro ejemplo, el calculador de factor de ganancia (180, 302) es capaz de calcular el factor de ganancia a partir de una combinación de potencias de respuestas de impulso de filtro individuales.

**[0125]** Según otro ejemplo, el calculador de factor de ganancia (180, 302) es capaz de calcular el factor de ganancia a partir de una adición ponderada de potencias de respuestas de impulso de filtro individuales, en el que coeficientes de ponderación utilizados en la adición ponderada dependen de la información de regla de mezcla ascendente.

**[0126]** Según otro ejemplo, el calculador de factor de ganancia (180, 302) es capaz de calcular el factor de ganancia a partir de una expresión que tiene un numerador y un denominador, teniendo el numerador una combinación de potencias de respuestas de impulso de filtro individuales, y teniendo el denominador una adición ponderada de potencias de respuestas de impulso de filtro individuales, en el que los coeficientes de ponderación utilizados en la adición ponderada dependen de la información de regla de mezcla ascendente.

**[0127]** Según otro ejemplo, el calculador de factor de ganancia (180, 302) es capaz de calcular el factor de ganancia a partir de la ecuación siguiente:

$$g_n = \begin{cases} \min \left\{ g_{\max}, \sqrt{\frac{E_n^B + \varepsilon}{E_n^B - \Delta E_n^B + \varepsilon}} \right\}, & \text{if } \alpha > 0, \beta > 0, \sigma < 1; \\ 1, & \text{En los demás casos.} \end{cases}$$

en la que  $g_n$  es el factor de ganancia para el primer canal, cuando  $n$  se establece a 1, en el que  $g_2$  es el factor de ganancia de un segundo canal, cuando  $n$  se establece a 2, en la que  $E_n^B$  es una energía de adición ponderada calculada mediante la ponderación energías de respuestas de impulso de canal empleando parámetros de ponderación, y en la que  $\Delta E_n^B$  es una estimación para el error de energía introducido por la regla de mezcla ascendente, en la que  $\alpha$ ,  $\beta$  y  $\sigma$  son parámetros dependientes de regla de mezcla ascendente, y en la que  $\varepsilon$  es un número mayor o igual a cero.

**[0128]** Según otro ejemplo, el calculador de factor de ganancia (180, 302) es capaz de calcular  $E_n^B$  y  $\Delta E_n^B$  a partir de la siguiente ecuación:

$$\Delta E_n^B = p(1-\sigma) \|b_{n,1} + b_{n,2} - b_{n,3}\|^2,$$

$$E_n^B = \beta(1-\sigma) \|b_{n,1}\|^2 + \alpha(1-\sigma) \|b_{n,2}\|^2 + p \|b_{n,3}\|^2,$$

Donde  $b_{n,1}$  es un filtro basado en HRTF correspondiente al primer canal de mezcla ascendente y al canal binaural enésimo, en el que  $b_{n,2}$  es un filtro basado en respuesta de impulso HRTF correspondiente a un segundo canal de mezcla ascendente y al canal binaural enésimo, en el que  $b_{n,3}$  es un filtro basado en respuesta de impulso HRTF correspondiente a un tercer canal de mezcla ascendente para un canal binaural enésimo, en el que las siguientes definiciones son válidas

$$\alpha = (1-c_1)/3, \quad \beta = (1-c_2)/3, \quad \sigma = \alpha + \beta, \quad \text{y} \quad p = \alpha\beta,$$

en las que  $c_1$  es un primer parámetro de predicción,  $c_2$  es un segundo parámetro de predicción, y en el que el primer parámetro de predicción y el segundo parámetro de predicción constituyen la información de regla de mezcla ascendente.

**[0129]** Según otro ejemplo, el calculador de factor de ganancia (180, 302) es capaz de calcular un factor de ganancia común para un canal binaural izquierdo y un canal binaural derecho.

**[0130]** Según otro ejemplo, el procesador de filtro (182) es capaz de usar, como las características de filtro, los filtros a base de función de transferencia relativa a la cabeza para el canal binaural izquierdo y el canal binaural derecho para posiciones virtuales central, izquierda y derecha o para utilizar características de filtros derivadas mediante la combinación de filtros HRTF para una posición virtual frontal izquierda y una posición virtual surround izquierda o mediante la combinación de filtros HRTF para una posición virtual frontal derecha y una posición virtual surround derecha.

[0131] Según otro ejemplo, unos parámetros relativos a canales izquierdo y surround izquierdo originales o canales derecho y surround derecho originales se incluyen en una señal de entrada de descodificador, y en el que el procesador de filtro (182) es capaz de usar los parámetros para combinar los filtros de función de transferencia relativos a la cabeza.

[0132] Según otro ejemplo, el calculador de factor de ganancia (180, 302) es capaz de calcular el factor de ganancia para un canal binaural a partir de una relación de una combinación lineal ponderada de energías de respuestas de impulso de canal de los filtros basados en HRTF para el canal binaural y un valor obtenido mediante la resta de un error de energía estimado a partir de la combinación lineal ponderada.

[0133] Según otro ejemplo, el calculador de factor de ganancia (180, 302) es capaz de usar información en la regla de mezcla ascendente para determinar el factor de ponderación.

[0134] Según otro ejemplo, la información de regla de mezcla ascendente incluye al menos dos parámetros de predicción utilizables para construir una matriz de mezcla ascendente lo que da como resultado canales de salida que tienen el error de energía con respecto a tres canales de entrada correspondientes.

[0135] Según otro ejemplo, el procesador de filtro (182) es operativo para tener, como características de filtros, un primer filtro (15) para filtrar un canal de mezcla descendente izquierdo para obtener una primera salida binaural izquierda, un segundo filtro (17) para filtrar un canal de mezcla descendente derecho para obtener una segunda salida binaural izquierda, un tercer filtro (16) para filtrar un canal de mezcla descendente izquierdo para obtener una primera salida binaural derecha, un cuarto filtro (18) para filtrar un canal de mezcla descendente derecho para obtener una segunda salida binaural derecha, un sumador (182a) para sumar la primera salida binaural izquierda y la segunda salida binaural izquierda para obtener un canal binaural izquierdo y para sumar la primera salida binaural derecha y la segunda salida binaural derecha para obtener un canal binaural derecho, en el que el procesador de filtro (182) es capaz de aplicar un factor de ganancia para el canal binaural izquierdo a los filtros primero y segundo o a la salida binaural izquierda antes o después de sumar y para aplicar el factor de ganancia para el canal binaural derecho al tercer filtro y al cuarto filtro o a la salida binaural derecha antes o después de sumar.

[0136] Según otro ejemplo, la información de regla de mezcla ascendente incluye parámetros de mezcla ascendente utilizables para construir una matriz de mezcla ascendente lo que da como resultado una mezcla ascendente a partir de dos a tres canales.

[0137] Según otro ejemplo, la regla de mezcla ascendente se define tal como sigue:

$$\begin{bmatrix} L \\ R \\ C \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} \begin{bmatrix} L_0 \\ R_0 \end{bmatrix},$$

Donde L es un primer canal de mezcla ascendente, R es un segundo canal de mezcla ascendente, y C es un tercer canal de mezcla ascendente,  $L_0$  es un primer canal de mezcla descendente,  $R_0$  es un segundo canal de mezcla descendente, y  $m_{ij}$  son parámetros de información de regla de mezcla ascendente.

[0138] Según otro ejemplo, se incluye un parámetro de pérdida de predicción en una señal de entrada de descodificador multi-canal, y en el que el procesador de filtro (182) puede escalar el factor de ganancia empleando el parámetro de pérdida de predicción.

[0139] Según otro ejemplo, el calculador de factor de ganancia (180, 302) es capaz de calcular el factor de ganancia a nivel de sub-banda, y en el que el procesador de filtro (182) es capaz de aplicar el factor de ganancia a nivel de sub-banda.

[0140] Según otro ejemplo, el procesador de filtro (182) es capaz de combinar filtros HRTF asociados con dos canales mediante la suma de versiones ponderadas o desplazadas en fase de respuestas de impulso de canal de los filtros HRTF, en el que los factores de ponderación para ponderar las respuestas de impulso de canal de los filtros HRTF depende de una diferencia de nivel entre los canales, y un desplazamiento de fase aplicado depende de un retardo de tiempo entre respuestas de impulso de canal de los filtros HRTF.

[0141] Según otro ejemplo, las características de filtros basados en filtros HRTF o filtros HRTF son filtros de sub-banda complejos obtenidos filtrando una respuesta de impulso de filtro de valor real de un filtro HRTF empleando un banco de filtros modelado complejo-exponencial.

[0142] Dependiendo de ciertos requisitos de implementación de los procedimientos de la invención, los procedimientos de la invención pueden implementarse en hardware o en software. La implementación puede

realizarse utilizando un medio de almacenamiento digital, en particular un disco, DVD o un CD que tiene señales de control legibles electrónicamente almacenadas en el mismo, que cooperan con un sistema informático programable de tal manera que se llevan a cabo los procedimientos de la invención. En general, la presente invención es, por lo tanto, un producto de programa informático con un código de programa almacenado en un soporte legible por máquina, siendo el código de programa operativo para realizar los procedimientos de la invención cuando el producto de programa informático se ejecuta en un ordenador. En otras palabras, los procedimientos de la invención son, por lo tanto, un programa informático que tiene un código de programa para realizar al menos uno de los procedimientos de la invención cuando el programa informático se ejecuta en un ordenador.

- 5
  - 10
- [0143]** Aunque lo anterior se ha mostrado y descrito particularmente con referencia a realizaciones particulares de la misma, se entenderá por los expertos en la materia que otros diversos cambios en la forma y detalles pueden hacerse quedando dentro del alcance de las reivindicaciones.

# REIVINDICACIONES

- 5     **1.** Descodificador multi-canal para generar una señal binaural a partir de una señal de mezcla descendente derivada de una señal multi-canal original empleando parámetros que incluyen una información de regla de mezcla ascendente utilizable para mezclar hacia arriba la señal de mezcla descendente con una regla de mezcla ascendente, dando como resultado la regla de mezcla ascendente un error de energía, que comprende:  
  
10     un calculador de factor de ganancia (180, 302) para calcular al menos un factor de ganancia ( $g_1$ ,  $g_2$ ) para reducir o eliminar el error de energía, a partir de la información de regla de mezcla ascendente y características de filtros de  
15     filtros a base de función de transferencia relativa a la cabeza correspondientes a canales de mezcla ascendente, en el que el calculador de factor de ganancia (180, 302) es capaz de calcular el factor de ganancia a partir de una energía de una respuesta de impulso combinada de las características del filtro, calculándose la respuesta de impulso combinada añadiendo o restando respuestas de impulso de filtro individuales; y un procesador de filtro (182) para filtrar la señal de mezcla descendente empleando el al menos un factor de ganancia ( $g_1$ ,  $g_2$ ), las características de filtro y la información de regla de mezcla ascendente para obtener una señal binaural con energía corregida.  
  
20     **2.** Procedimiento de decodificación multi-canal para generar una señal binaural a partir de una señal de mezcla descendente derivada de una señal multi-canal original empleando parámetros que incluyen una información de regla de mezcla ascendente utilizable para mezclar hacia arriba la señal de mezcla descendente con una regla de mezcla ascendente, dando como resultado la regla de mezcla ascendente un error de energía, que comprende:  
  
25     calcular (180, 302) al menos un factor de ganancia ( $g_1$ ,  $g_2$ ) para reducir o eliminar el error de energía, a partir de la información de regla de mezcla ascendente y características de filtros de filtros a base de función de transferencia relativa a la cabeza correspondiente a canales de mezcla ascendente, en el que el factor de ganancia se calcula a partir de una energía de una respuesta de impulso combinada de las características del filtro, calculándose la respuesta de impulso combinada añadiendo o restando respuestas de impulso de filtro individuales; y  
  
30     filtrar (182) la señal de mezcla descendente empleando el al menos un factor de ganancia ( $g_1$ ,  $g_2$ ), las características de filtro y la información de regla de mezcla ascendente para obtener una señal binaural con energía corregida.  
  
3. Programa de ordenador que tiene un código de programa para realizar el procedimiento según la reivindicación 2, cuando el programa de ordenador se ejecuta en un ordenador.

FIG 1

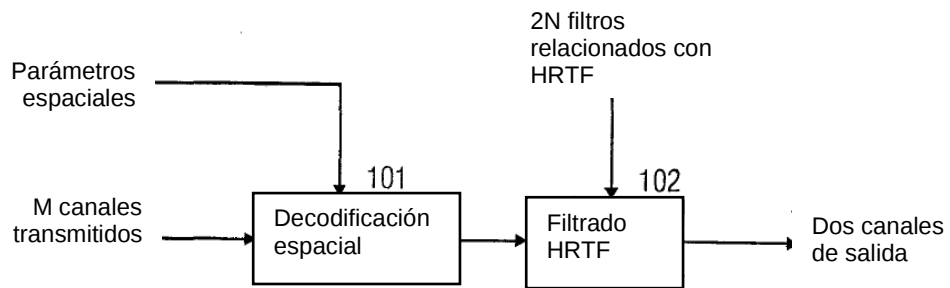


FIG 2

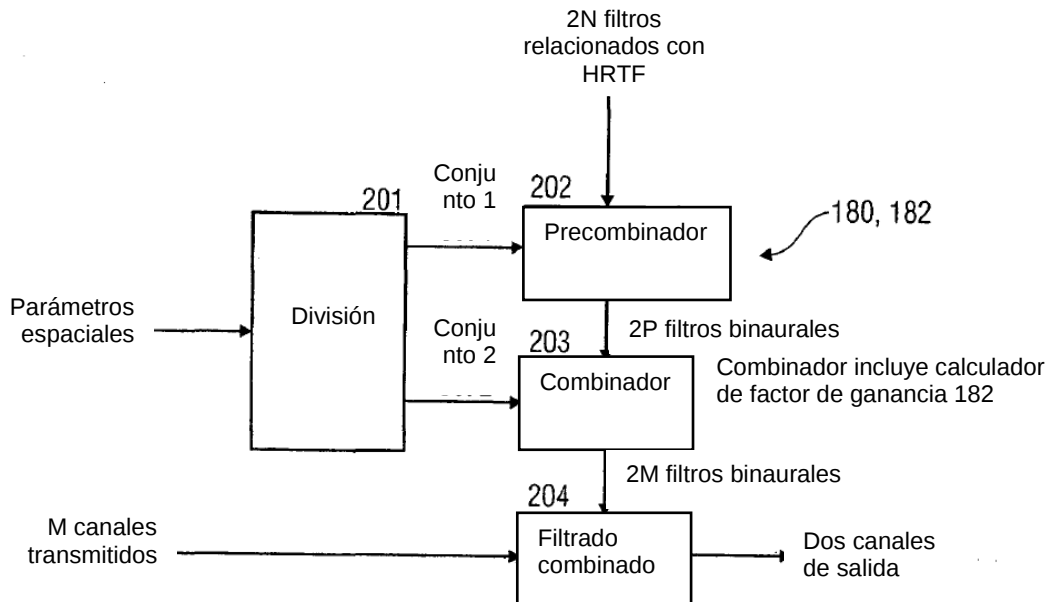


FIG 3

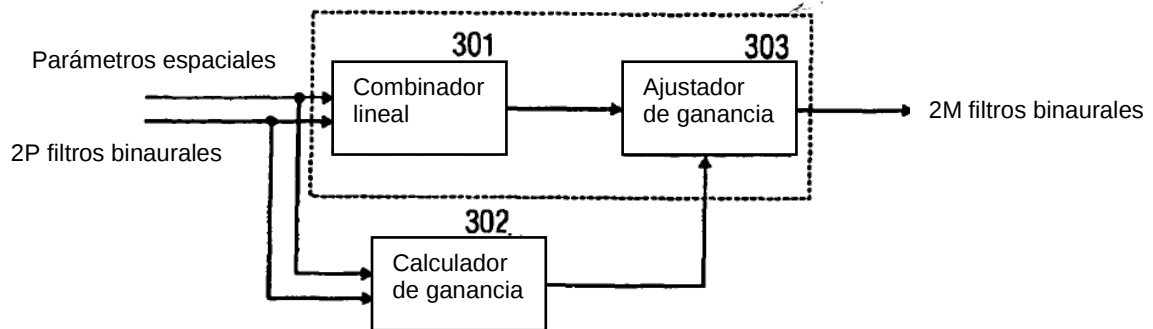


FIG 4

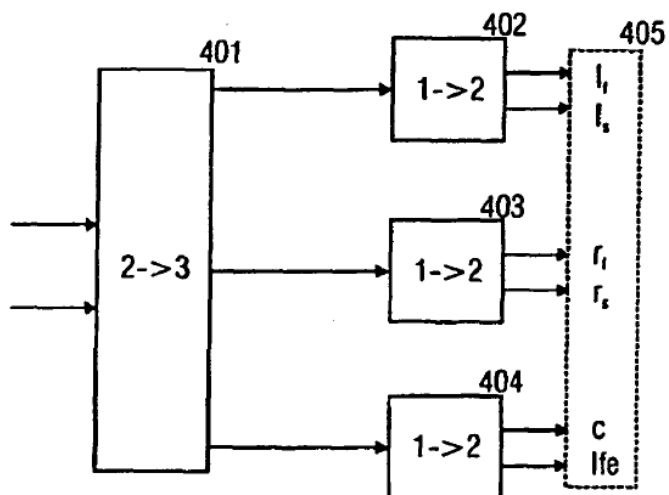


FIG 5

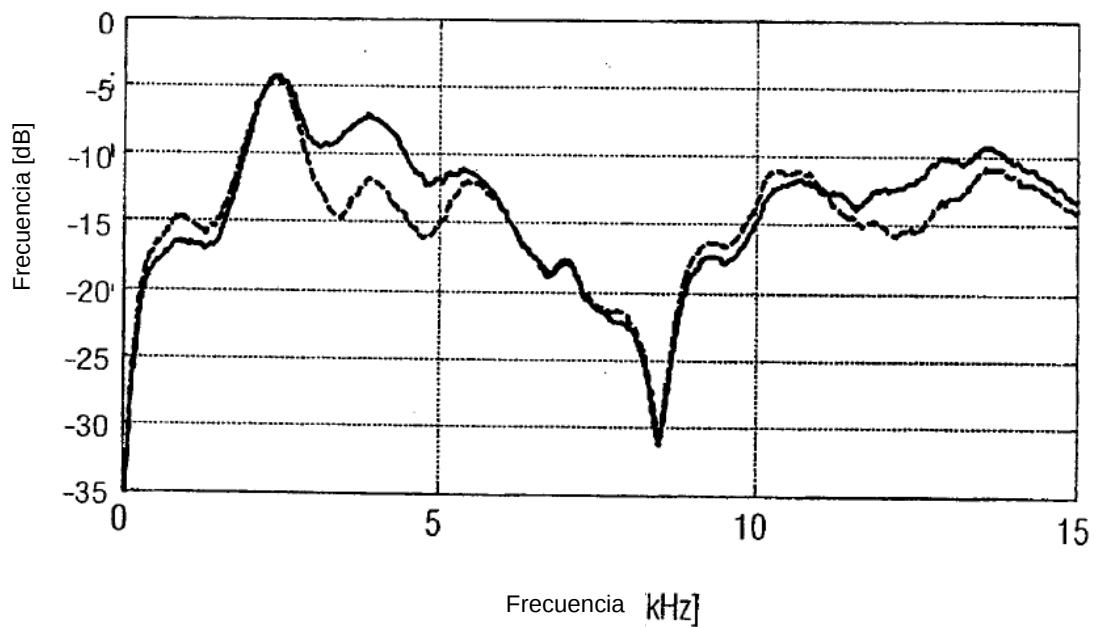


FIG 6

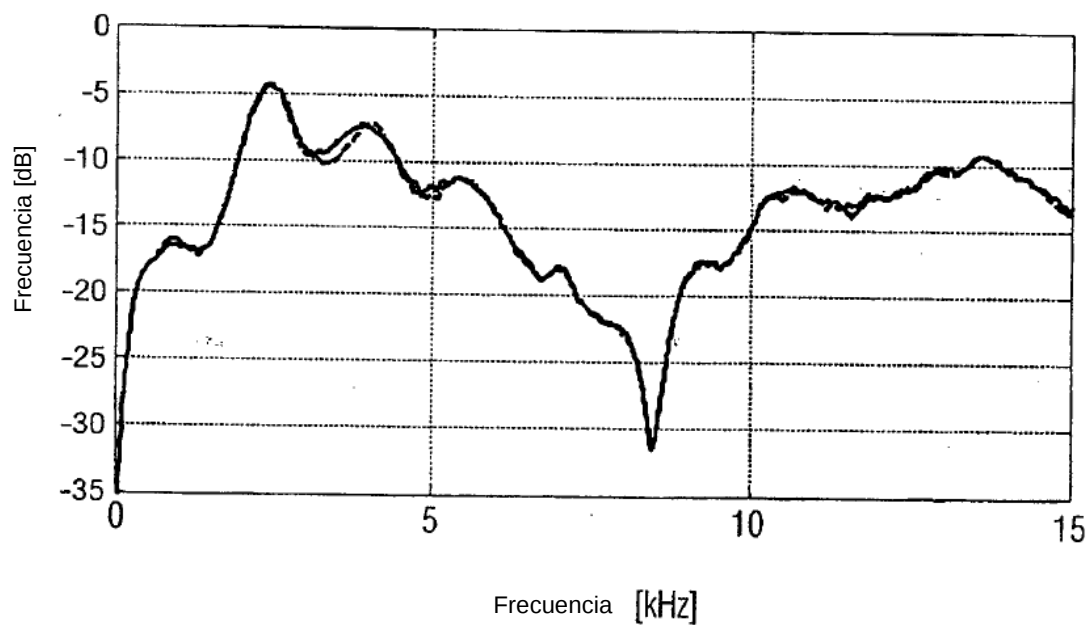


FIG 7

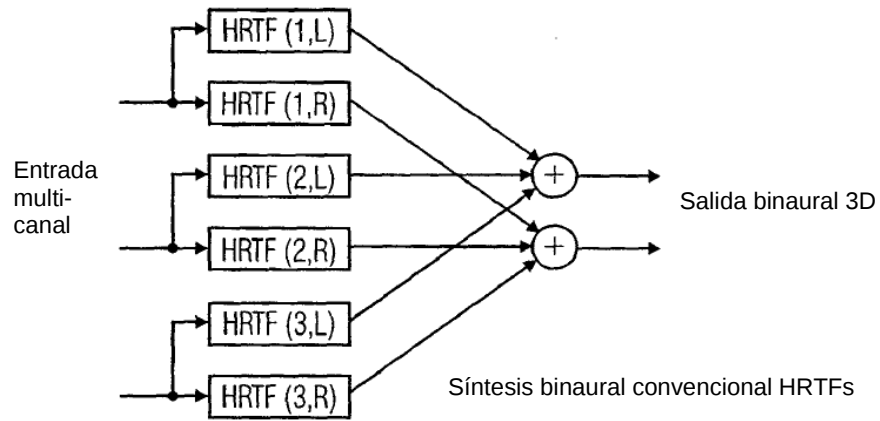


FIG 8

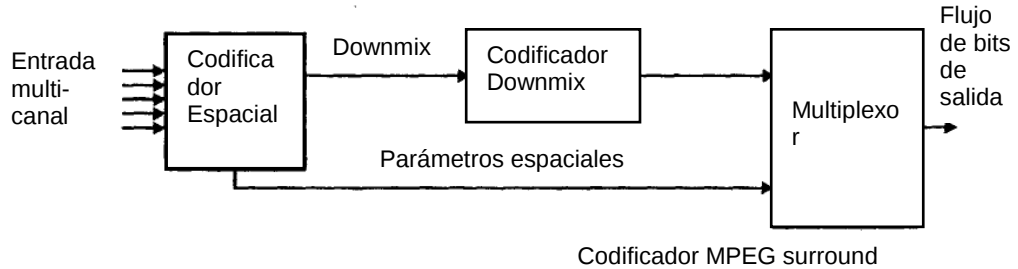
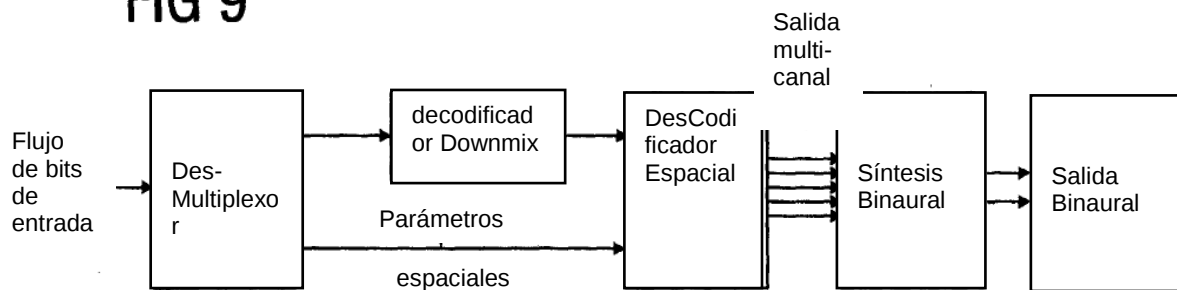


FIG 9



Cascada de decodificador MPEG surround y sintetizador binaural

FIG 10

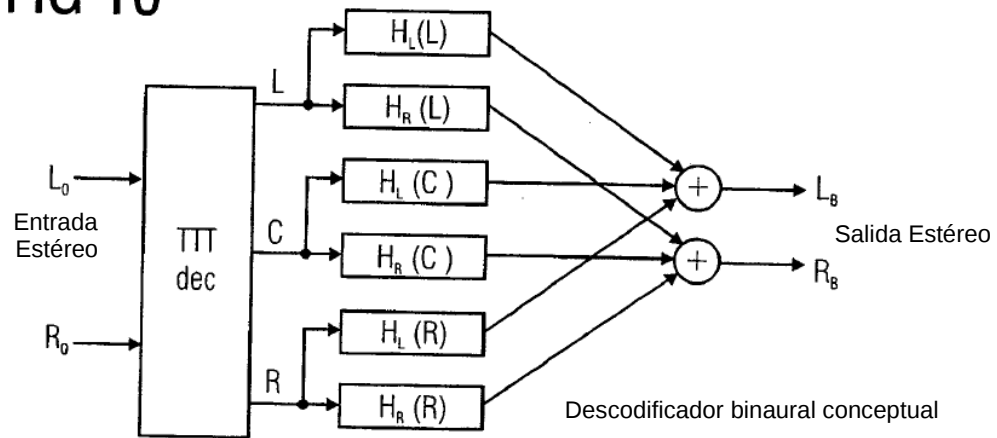


FIG 11

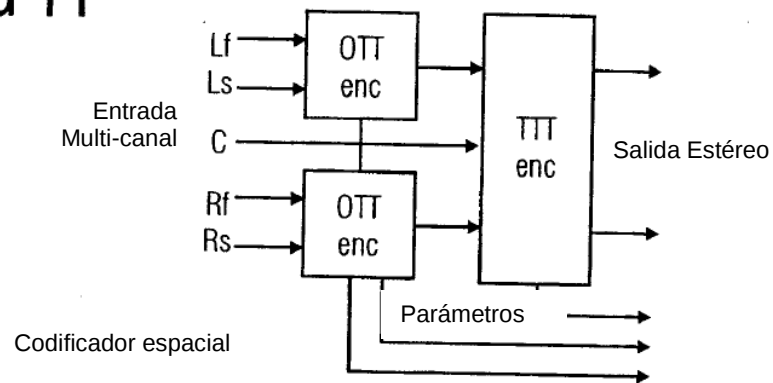


FIG 12

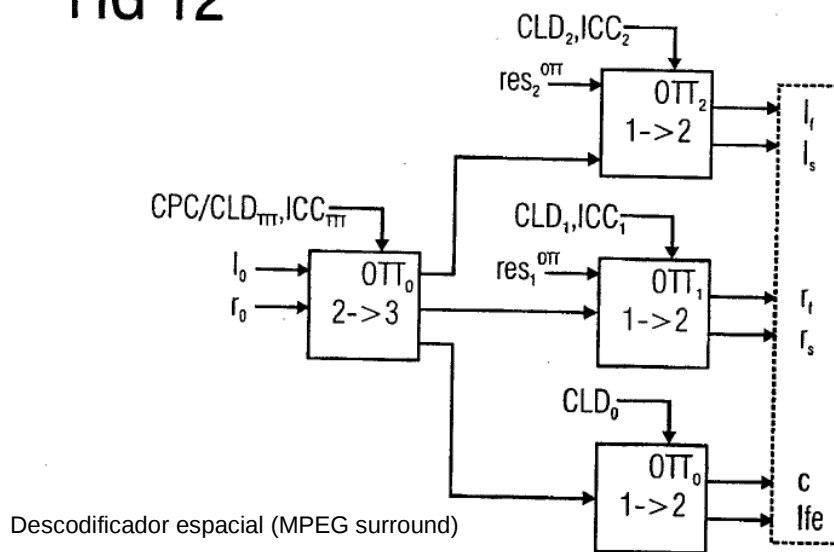


FIG 13

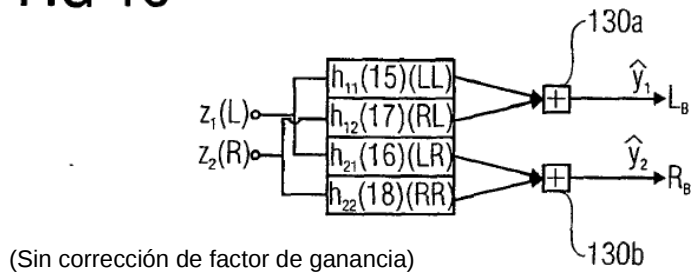


FIG 14

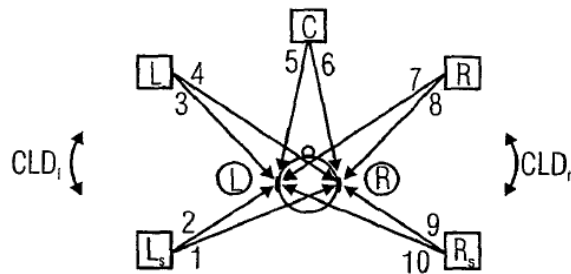


FIG 15

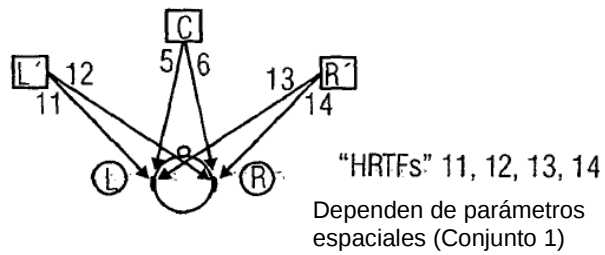


FIG 16a

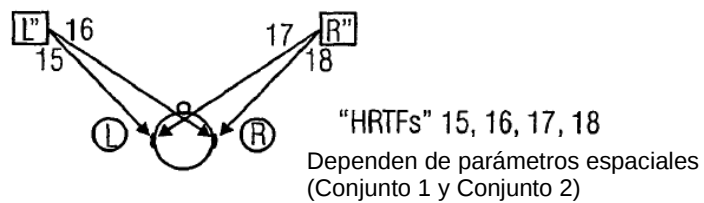


FIG 16b

$$\begin{bmatrix} L \\ R \\ C \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} \begin{bmatrix} L_0 \\ R_0 \end{bmatrix}$$

Regla de upmix (determinada p.e. por codificador de fig. 20)

$m_{11}, m_{21}, m_{31}$ : Upmix – izquierdo  
 $m_{12}, m_{22}, m_{32}$ : Upmix - derecho

FIG 17

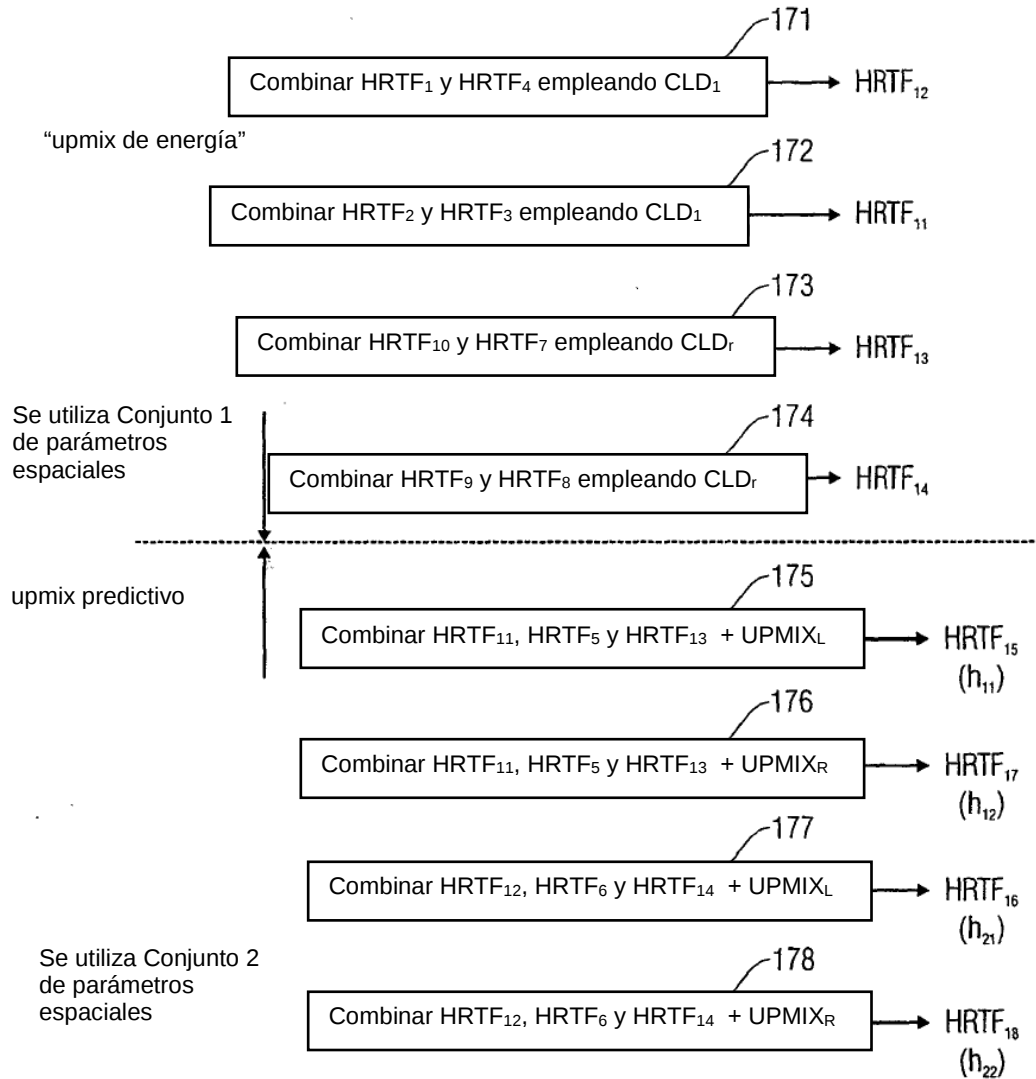
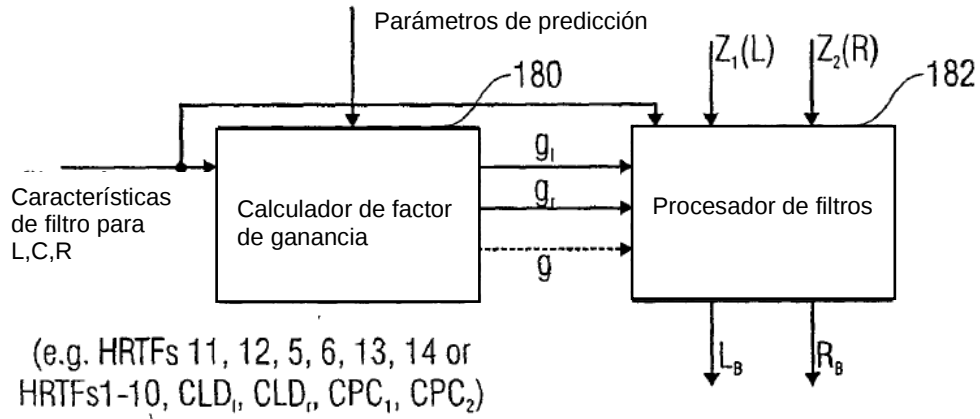


FIG 18



$$g_L \approx \sqrt{E_L^B / E_L^B - \Delta E_L^B};$$

$$g_R \approx \sqrt{E_R^B / E_R^B - \Delta E_R^B};$$

$$\Delta E = f(\|\sum_{i=11, 5, 13, 12, 6, 14} \text{HRTF}_i\|)$$

$$E = f(\|\sum \text{HRTF}_i\|)$$

FIG 19a

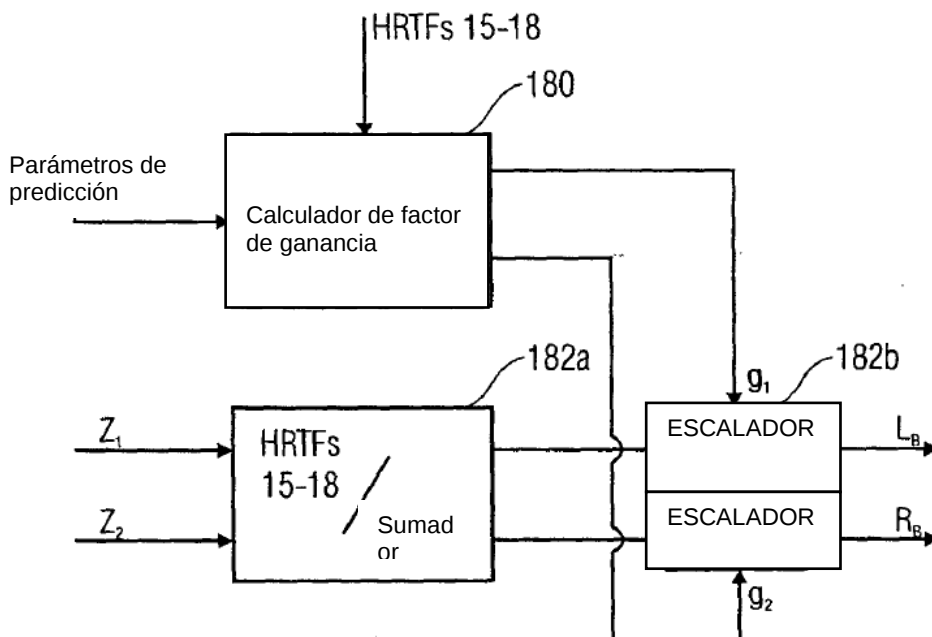


FIG 19b

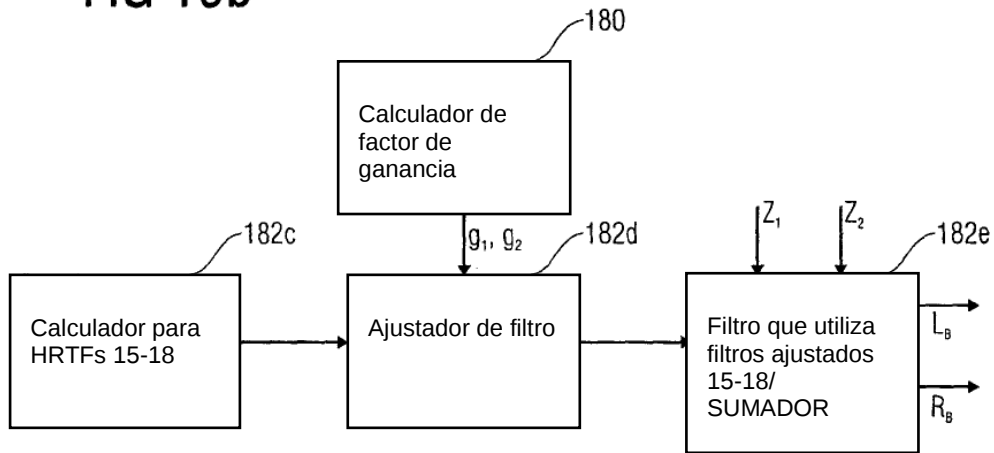


FIG 20

