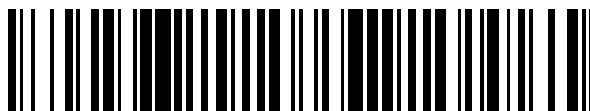


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 452 348**

51 Int. Cl.:

G10L 19/008 (2013.01)

H04S 1/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.04.2008** **E 08749081 (9)**

97 Fecha y número de publicación de la concesión europea: **08.01.2014** **EP 2137725**

54 Título: **Aparato y procedimiento para sintetizar una señal de salida**

30 Prioridad:

26.04.2007 US 914267 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
01.04.2014

73 Titular/es:

**DOLBY INTERNATIONAL AB (50.0%)
Apollo Building, 3E, Herikerbergweg 1-35
1101 CN Amsterdam Zuid-Oost, NL y
FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (50.0%)**

72 Inventor/es:

**ENGDEGARD, JONAS;
VILLEMORS, LARS;
PURNHAGEN, HEIKO;
RESCH, BARBARA;
FALCH, CORNELIA;
HERRE, JUERGEN;
HILPERT, JOHANNES;
HOELZER, ANDREAS y
TERENTIEV, LEONID**

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 452 348 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Aparato y procedimiento para sintetizar una señal de salida

5 **[0001]** La presente invención se relaciona con la síntesis de una señal de salida reproducida, como una señal de salida estéreo o una señal de salida que tiene más señales de canal de audio basadas en una conversión de multicanal a estéreo de varios canales disponibles y datos de control adicionales. Específicamente, la conversión multicanal a estéreo de varios canales es una conversión de multicanal a estéreo de una pluralidad de señales de objeto de audio.

10 **[0002]** El desarrollo reciente en el audio facilita la reconstrucción de una representación de varios canales de una señal de audio basada en un estéreo (o mono) señal y unos datos de control correspondientes. Estos procedimientos de codificación de sonido envolvente paramétricos por lo general comprenden una parametrización. Un decodificador de audio de varios canales paramétrico, (p.ej el Decodificador de sonido envolvente MPEG definido en ISO/IEC 23003-1 [1], [2]), reconstruye canales M basados en los canales transmitidos de K, con $M > K$, por el uso de los datos de control adicionales. Los datos de control comprenden una parametrización de la señal de varios canales basada en IID (Diferencia de Intensidad de Intercanal) e ICC (Coherencia de Intercanal). Estos parámetros son normalmente extraídos en la etapa de codificación y describen la relación de potencias y la correlación entre pares de canal usados en el proceso de conversión de estéreo a multicanal. Usar tal esquema de codificación permite la codificación a de una velocidad de transferencia de datos inferior considerablemente significativa que la transmisión de todos los canales M, haciendo la codificación muy eficiente a un al mismo tiempo contrato de un seguro de la compatibilidad tanto por dispositivos de canal K como por dispositivos de canal M.

25 **[0003]** Mucho el sistema de codificación relacionado es el codificador de objeto de audio correspondiente [3], [4], donde varios objetos de audio se convierten de multicanal a estéreo al codificador y más tarde convertido de estéreo a multicanal, dirigido por datos de control. El proceso de conversión de estéreo a multicanal también puede observarse como una separación de los objetos que se mezclan en la conversión de multicanal a estéreo. La señal convertida de estéreo a multicanal resultante puede volverse en uno o varios canales de repetición. Más exactamente, [3, 4], presentan un procedimiento para sintetizar canales de audio de una conversión de multicanal a estéreo (referido como señal suma), información estadística sobre los objetos de la fuente, y datos que describen el formato de salida deseado. Por si varias señales de conversión de multicanal a estéreo se usen, estas señales de conversión de multicanal a estéreo comprenden diferentes subconjuntos de los objetos, y la conversión de estéreo a multicanal se lleva a cabo para cada canal de conversión de multicanal a estéreo individualmente.

35 **[0004]** En caso de una conversión de multicanal a estéreo de objeto y renderización de objeto al estéreo, o generación de una señal estéreo adecuada para procesar adicionalmente por por ejemplo un decodificador de sonido envolvente MPEG, esto se conoce de la técnica anterior que una ventaja de desempeño significativa se logra por el procesamiento conjunto de los dos canales con esquema de matrización dependiente de tiempo y frecuencia. Fuera del alcance de la codificación de objeto de audio, un procedimiento relacionado se aplica para transformar parcialmente una señal de audio estéreo en otra señal de audio estéreo en el documento WO2006/103584. También es conocido que para un sistema de codificación de objeto de audio general es necesario introducir la adición de un proceso de decorrelación a la renderización a fin de reproducir perceptivamente la escena de referencia deseada. Sin embargo, no hay ninguna técnica anterior que describe una combinación conjuntamente optimizada de matrización y decorrelación. Una combinación simple de los procedimientos de técnica anterior conduce al uso ineficaz e inflexible de las capacidades ofrecidas por una conversión de multicanal a estéreo de objeto de varios canales o a una baja calidad de imagen estéreo en las renderizaciones de decodificador de objeto resultantes.

Referencias:

50 **[0005]**

[1] L. Villemoes, J. Herre, J. Breebaart, G. Hotho, S. Disch, H. Purnhagen, y K. Kjörling, "MPEG Surround: The Forthcoming ISO Standard for Spatial Audio Coding," en la 28va International AES Conference, The Future of Audio Technology Surround and Beyond, Piteå, Suecia, 30 de junio a 2 de julio, 2006.

55 [2] J. Breebaart, J. Herre, L. Villemoes, C. Jin, , K. Kjörling, J. Plogsties, y J. Koppens, "Multi-Channels goes Mobile: MPEG Surround Binaural Rendering," en la 29va International AES Conference, Audio for Mobile and Handheld Devices, Seul, 2-4 de septiembre, 2006.

60 [3] C. Faller, "Parametric Joint-Coding of Audio Sources," Convention Paper 6752 presentado en la 120va AES Convention, Paris, Francia, 20-23 de mayo, 2006.

[4] C. Faller, "Parametric Joint-Coding of Audio Sources," Solicitud de patente PCT/EP2006/050904, 2006.

65 **[0006]** The "Call for Proposals on Spatial Audio Object Coding", January 2007, Marrakech, Morocco, MPEG 2007/N8853, XP090015347, se refiere a la codificación de audio espacial (SAC) , donde los canales originales se

codifican por un codificador SAC para producir la(s) señal(es) de mezcla descendente (s) dentro de la información , y donde un decodificador SAC decodifica la información transmitida para reproducir canales de salida . Un enfoque alternativo de codificación de objeto (" orientado a objetos ") de audio espacial (SAOC) incluye un codificador SAOC para generar señal(es) de mezcla descendente (s) dentro de la información de los objetos de audio y un decodificador SAOC para decodificar la información transmitida para generar objetos decodificados. Los objetos decodificados se introducen en un procesador que proporciona , como una entrada adicional, información sobre interacción / control con el fin último de dar como salida dos o más canales. En lugar de utilizar un decodificador de envolvente MPEG con N canales de salida para reproducir N objetos y utilizar una etapa de representación posterior de representación de N objetos en M canales de salida (típicamente con $N > M$) , es económico utilizar directamente una representación de envolvente MPEG de M canales para el número deseado de canales de salida que son accionados por los parámetros espaciales apropiados . Particularmente , se utilizan las posiciones de los objetos y la reproducción de configuración para generar una matriz de representación . La matriz de representación se utiliza en un transcodificador SAOC para transcodificar un flujo de bits SAOC en un flujo de bits MPS , y se entra el flujo de bits de entrada MPS, junto con una mezcla descendente preprocesada , en el que el procesamiento previo es controlado por parámetros SAOC , en un decodificador MPS para generar una escena renderizada de salida .

[0007] Es el objetivo de la presente invención el proporcionar un concepto mejorado a sintetizar una señal de salida reproducida.

[0008] Este objetivo se logra por un aparato para sintetizar una señal de salida reproducida de acuerdo con la reivindicación 1, un procedimiento para sintetizar una señal de salida reproducida de acuerdo con la reivindicación 13 o un programa de computadora de acuerdo con la reivindicación 14.

[0009] La presente invención proporciona una síntesis de una señal de salida reproducida que tiene dos señales de canal de audio (estéreo) o más de dos señales de canal de audio. En caso de muchos objetos de audio, vario canal de audio sintetizado hace señas es, sin embargo, más pequeño que el número de objetos de audio originales. Sin embargo, cuando el número de objetos de audio es pequeño (p.ej 2) o el número de canales de salida es 2, 3 o incluso más grande, el número de canales de salida de audio puede ser mayor que el número de objetos. La síntesis de la señal de salida reproducida se lleva a cabo sin una completa operación de decodificación de objeto de audio en objetos de audio decodificados y una renderización con especificidad de objetivo subsecuente de los objetos de audio sintetizados. En cambio, un cálculo de las señales de salida reproducidas se lleva a cabo en el dominio de parámetro basado en la información de conversión de multicanal a estéreo, en la información de renderización de objetivo y en la información de objeto de audio que describe los objetos de audio, como información de energía e información de correlación. Así, el número de decorrelacionadores que pesadamente contribuyen a la complejidad de realización de un aparato que sintetiza puede reducirse para ser más pequeño que el número de canales de salida e incluso sustancialmente más pequeños que el número de objetos de audio. Específicamente, los sintetizadores con sólo un decorrelacionador individual o dos decorrelacionadores pueden ponerse en práctica para la síntesis de audio de alta calidad. Más aún, debido a que una completa decodificación de objeto de audio y la renderización con especificidad de objetivo subsecuente no deben llevarse a cabo, la memoria y los recursos computacionales pueden ser ahorrados. Más aún, cada operación introduce artefactos potenciales. Por lo tanto, el cálculo de acuerdo con la presente invención se hace preferiblemente en el dominio de parámetro sólo de modo que las únicas señales de audio que no proporcionan en parámetros, pero que se les proporcionan como, por ejemplo, dominio temporal o señales de dominio de subbanda sean las al menos dos señales de conversión de multicanal a estéreo de objeto. Durante la síntesis de audio, ellos son introducidos en el decorrelacionador en una forma convertida de multicanal a estéreo cuando un decorrelacionador individual se usa o en una forma mixta, cuando un decorrelacionador para cada canal se usa. Otras operaciones hechas en el dominio temporal o dominio de banco de filtros o señales de canal mezcladas son combinaciones sólo ponderadas, como adiciones ponderadas o subtracciones ponderadas, es decir, operaciones lineales. Así, la introducción de artefactos debido a una completa operación de decodificación de objeto de audio y una operación de renderización de objetivo subsecuente es evitada.

[0010] Preferentemente, la información de objeto de audio se proporciona como una información de energía e información de correlación, por ejemplo en la forma de una matriz de covariancia de objeto. Más aún, se prefiere que esta matriz esté disponible para cada subbanda y cada bloque temporal de modo que un mapa de frecuencias y tiempo exista, donde cada entrada de mapa incluye una matriz de covariancia de objeto de audio descripción de la energía de los objetos de audio respectivos en esta subbanda y la correlación entre pares respectivos de objetos de audio en la subbanda correspondiente. Naturalmente, esta información se relaciona con un cierto bloque temporal o la porción de tiempo o marca de tiempo de una señal de subbanda o una señal de audio.

[0011] Preferentemente, la síntesis de audio se lleva a cabo en una señal de salida estéreo reproducida que tiene una primera o izquierda señal de canal de audio y una segunda o derecha señal de canal de audio. Así, uno puede acercarse a una solicitud de la codificación de objeto de audio, en la cual la renderización de los objetos al estéreo está tan cerca como sea posible a la renderización de estéreo de referencia.

[0012] En muchas aplicaciones de codificación de objeto de audio es de gran importancia que la renderización de los objetos al estéreo esté tan cerca como sea posible a la renderización de estéreo de referencia. El logro de una renderización estéreo de alta calidad, cuando una aproximación a la renderización de estéreo de referencia es importante ambos en términos de calidad de audio para el caso donde la renderización estéreo es la salida final del decodificador de objeto, y en el caso donde la señal estéreo se alimenta a un dispositivo subsecuente, como un Decodificador de sonido envolvente MPEG que funciona en el modo de conversión de multicanal a estéreo estéreo.

[0013] La presente invención proporciona una combinación conjuntamente optimizada de una matrización y procedimiento de decorrelación que permite a un decodificador de objeto de audio explotar el potencial lleno de esquema de codificación de objeto de audio usando una conversión de multicanal a estéreo de objeto con más de un canal.

[0014] Las modalidades de la presente invención comprenden las características que siguen:

- un decodificador de objeto de audio para renderizar una pluralidad de objetos de audio individuales usando una conversión de multicanal a estéreo de varios canales, datos de control que describen los objetos, datos de control que describen la conversión de multicanal a estéreo, y renderizan información, comprendiendo

- un procesador estéreo que comprende una unidad de matrización potenciada, operacional en combinar en línea recta los canales de conversión de multicanal a estéreo de varios canales en una mezcla seca hace señas y una señal de entrada de decorrelacionador y posteriormente alimentar la señal de entrada de decorrelacionador en una unidad de decorrelacionador, la señal de salida de que es en línea recta combinada en una señal la adición que al en lo referente a canales con la señal de mezcla seca constituye la salida estéreo de la unidad de matrización potenciada; o

- una calculadora de la matriz para calcular los pesos para la combinación lineal usada por la unidad de matrización potenciada, basada en los datos de control que describen los objetos, los datos de control describen la conversión de multicanal a estéreo e información de renderización de estéreo.

[0015] La presente invención se describirá ahora por vía de ejemplos ilustrativos, no limitando el alcance o el espíritu de la invención, en cuanto a las figuras acompañantes, donde:

La Figura 1 ilustra la operación de la codificación de objeto de audio codificación que comprende y decodificar;

La Figura 2a ilustra la operación de decodificación de objeto de audio al estéreo;

La Figura 2b ilustra la operación de decodificación de objeto de audio;

La Figura 3a ilustra la estructura de un procesador estéreo;

La Figura 3b ilustra un aparato para sintetizar una señal de salida reproducida;

La Figura 4a ilustra el primer aspecto de la invención que incluye una matriz de mezclado de señal seca C_0 , una prematriz de mezcla del predecorrelacionador Q y una matriz de conversión de estéreo a multicanal de decorrelacionador P ;

La Figura 4b ilustra otro aspecto de la presente invención que se pone en práctica sin una prematriz de mezcla del predecorrelacionador;

La Figura 4c ilustra otro aspecto de la presente invención que se pone en práctica sin la matriz de conversión de estéreo a multicanal de decorrelacionador;

La Figura 4d ilustra otro aspecto de la presente de la presente invención que se pone en práctica con una matriz de compensación de ganancia adicional G ;

La Figura 4e ilustra una realización de la matriz de conversión de multicanal a estéreo de decorrelacionador Q y la matriz de conversión de estéreo a multicanal de decorrelacionador P cuando un decorrelacionador individual se usa;

La Figura 4f ilustra una realización de la matriz de mezclado seco C_0 ;

La Figura 4g ilustra una vista detallada de la combinación actual del resultado de la mezcla de señal seca y el resultado del decorrelacionador u operación de conversión de estéreo a multicanal de decorrelacionador;

La Figura 5 ilustra una operación de una etapa de decorrelacionador de varios canales que tiene muchos decorrelacionadores;

La Figura 6 ilustra un mapa que indica que varios objetos de audio identificados por cierto ID, que tienen un archivo de audio de objeto, y una matriz de información de objeto de audio conjunta **E**;

La Figura 7 ilustra una explicación de una matriz de covarianza de objeto **E** de la Figura 6:

La Figura 8 ilustra una matriz de conversión de multicanal a estéreo y un codificador de objeto de audio controlado por la matriz de conversión de multicanal a estéreo **D**;

La Figura 9 ilustra una matriz de renderización objetivo **A** que es normalmente proporcionada por un usuario y un ejemplo para un caso de renderización de objetivo específico;

La Figura 10 ilustra una recolección de etapas de precálculo llevadas a cabo para determinar los elementos de la matriz de las matrices en Figuras 4a a 4d de acuerdo con cuatro diferentes modalidades;

La Figura 11 ilustra una recolección de etapas de cálculo de acuerdo con la primera modalidad;

La Figura 12 ilustra una recolección de etapas de cálculo de acuerdo con la segunda modalidad;

La Figura 13 ilustra una recolección de etapas de cálculo de acuerdo con la tercera modalidad; y

La Figura 14 ilustra una recolección de etapas de cálculo de acuerdo con la cuarta modalidad.

[0016] Las modalidades descritas a continuación son simplemente ilustrativas para los principios de la presente invención para el APARATO Y PROCEDIMIENTO PARA SINTETIZAR UNA SEÑAL DE SALIDA. Se entiende que las modificaciones y las variaciones de las configuraciones y los detalles descritos aquí serán evidentes a otros expertos en la técnica. Es la intención, por lo tanto, sólo para ser limitado por el alcance de las reivindicaciones de patente inminentes y no por los detalles específicos presentados por vía de descripción y explicación de las modalidades aquí.

[0017] La Figura 1 ilustra la operación de la codificación de objeto de audio, que comprende un codificador de objeto 101 y un decodificador de objeto 102. El codificador de objeto de audio espacial 101 codifica N objetos en un objeto de mezcla descendente que comprende $K > 1$ canales de audio, según parámetros de codificador. La información sobre la matriz ponderada **D** de conversión de multicanal a estéreo aplicada es enviada por el codificador de objeto conjuntamente con datos opcionales acerca de la energía y la correlación de la conversión de multicanal a estéreo. La matriz **D** es a menudo, pero no necesariamente siempre, constante sobre el tiempo y frecuencia, y por lo tanto representa una relativamente pequeña cantidad de la información. Finalmente, los extractos de codificador de objeto objetan parámetros para cada objeto como una función de ambo tiempo y frecuencia a una resolución definida por consideraciones perceptuales. El decodificador de objeto espacial de audio 102 considera los canales de conversión de multicanal a estéreo de objeto, la información de conversión de multicanal a estéreo, y los parámetros de objeto (tan generado por el codificador) como introducido y genera una salida con M canales de audio para la presentación al usuario. La renderización de N objetos en M canales de audio usa una matriz de renderización proporcionada como la introducción de datos por el usuario al decodificador de objeto.

[0018] La Figura 2a ilustra los componentes de un decodificador de objeto de audio 102, en el caso donde la salida deseada es el audio estéreo. La conversión de multicanal a estéreo de objeto de audio se alimenta en un procesador estéreo 201, que lleva a cabo el procesamiento de señal que da como resultado a una salida de audio estéreo. Este procesamiento depende de la información de la matriz proporcionada por la calculadora de la matriz 202. La información de la matriz se deriva de los parámetros de objeto, la información de conversión de multicanal a estéreo y la información de renderización de objeto suministrada, que describe la renderización del objetivo deseada de los N objetos en el estéreo por medio de una matriz de renderización.

[0019] La Figura 2b ilustra los componentes de un decodificador de objeto de audio 102, en el caso donde la salida deseada es una señal de audio de varios canales general. La conversión de multicanal a estéreo de objeto de audio se alimenta en un procesador estéreo 201, que lleva a cabo el procesamiento de señal que da como resultado a una salida de señal estéreo. Este procesamiento depende de la información de la matriz proporcionada por la calculadora de la matriz 202. La información de la matriz se deriva de los parámetros de objeto, la información de conversión de multicanal a estéreo y una información de renderización de objeto reducida, que es enviada por el reductor de renderización 204. La información de renderización de objeto reducida describe la renderización deseada de los N objetos en el estéreo por medio de una matriz de renderización, y esto se deriva de la información de renderización que describe la renderización de N objetos en M canales de audio suministrados al decodificador de objeto de audio 102, los parámetros de objeto, y la información de conversión de multicanal a estéreo de objeto. El procesador adicional 203 conversos la señal estéreo proporcionada por el procesador estéreo 201 en la salida de audio de varios canales final, basada en la información de renderización, la información de conversión de multicanal a estéreo y los parámetros de objeto. Un Decodificador de sonido envolvente MPEG que funciona en el modo de conversión de multicanal a estéreo estéreo es un componente principal común del procesador adicional 203.

5 **[0020]** La Figura 3a ilustra la estructura del procesador estéreo 201. Considerando la conversión de multicanal a estéreo de objeto transmitida en el formato de una salida de tren de bits de un codificador de audio de K canales , este tren de bits es primero decodificado por el decodificador de audio 301 en K señales de audio de dominio temporal . Estas señales son todas transformadas luego al dominio de frecuencia por la unidad T/F 302. El tiempo y frecuencia que varía la matrización potenciada inventiva definida por la información de la matriz suministró al procesador estéreo 201 se lleva a cabo en de señales de dominio de frecuencia resultante **X** por la unidad de matrización potenciada 303. Esta unidad envía una señal estéreo en el dominio de frecuencia que se convierte en la señal de dominio temporal **Y'** por la unidad F/T 304.

10 **[0021]** La Figura 3b ilustra un aparato para sintetizar una señal de salida reproducida 350 que tiene una primera señal de canal de audio y una segunda señal de canal de audio en caso de una operación de renderización de estéreo, o que tiene más de dos señales de canal de salida en caso de una renderización de canal más elevada. Sin embargo, para un número más elevado de objetos de audio tal como tres o más el número de canales de salida es preferentemente más pequeño que el número de objetos de audio originales, que tener contribuido a la señal de conversión de multicanal a estéreo 352. Específicamente, la señal de conversión de multicanal a estéreo 352 tiene al menos una primera señal de conversión de multicanal a estéreo de objeto y una segunda señal de conversión de multicanal a estéreo de objeto, donde la señal de conversión de multicanal a estéreo representa una conversión de multicanal a estéreo de una pluralidad de señales de objeto de audio de acuerdo con la información de conversión de multicanal a estéreo 354. Específicamente, el sintetizador de audio inventivo como ilustrado en Figura 3b incluye una etapa 356 de decorrelacionador mientras la generación de una señal no correlacionada que tiene una señal de canal individual no correlacionada o una primera señal de canal no correlacionada y una segunda señal de canal no correlacionada en caso de dos decorrelacionadores o que tiene más de dos señales de canal de decorrelacionador en caso de una realización que tiene tres o más decorrelacionadores. Sin embargo, un número más pequeño de decorrelacionadores y, por lo tanto, un número más pequeño de señales de canal no correlacionadas se prefiere sobre un número más elevado debido a la complejidad de realización incurrida por un decorrelacionador. Preferentemente, el número de decorrelacionadores es más pequeño que el número de objetos de audio incluidos en la señal de conversión de multicanal a estéreo 352 y será preferentemente igual al número de señales de canal en la señal de salida 352 o más pequeño que el número de señales de canal de audio en la señal de salida reproducida 350. Para un pequeño número de objetos de audio (p.ej. 2 o 3), sin embargo, el número de decorrelacionadores puede ser igual o incluso mayor que el número de objetos de audio.

35 **[0022]** Como indicado en la Figura 3b, la etapa de decorrelacionador recibe, como una entrada, la señal de conversión de multicanal a estéreo 352 y genera, como una señal de salida, la señal 358 no correlacionada. Además de la información de conversión de multicanal a estéreo 354, la información de renderización de objetivo 360 y la información de parámetro de objeto de audio 362 se proporcionan. Específicamente, la información de parámetro de objeto de audio es al menos usada en un combinador 364 y puede usarse opcionalmente en la etapa 356 de decorrelacionador como se describirá más adelante. La información de parámetro de objeto de audio 362 preferentemente comprende la energía y la información de correlación que describe el objeto de audio en una forma dada parámetros, como un número entre 0 y 1 o un cierto número que se define en un cierto intervalo de valor, y que indica una energía, una energía o una medida de correlación entre dos objetos de audio como descritos más adelante.

45 **[0023]** El combinador 364 se configura para llevar a cabo una combinación ponderada de la señal de conversión de multicanal a estéreo 352 y la señal 358 no correlacionada. Más aún, el combinador 364 es operativo para calcular factores de ponderación para la combinación ponderada de la información de conversión de multicanal a estéreo 354 y la información de renderización de objetivo 360. La información de renderización de objetivo indica posiciones virtuales de los objetos de audio en un sistema de repetición virtual e indica la colocación específica de los objetos de audio a fin de determinar, si un cierto objeto debe volverse en el primer canal de salida o el segundo canal de salida, es decir, en un canal de salida de izquierda o un canal de salida derecho para una renderización estéreo. Cuando, sin embargo, una renderización de varios canales se lleva a cabo, luego la información de renderización de objetivo además indica si un cierto canal debe colocarse más o menos en un sonido envolvente de izquierda o un sonido envolvente derecho o canal de centro etc. Cualquier caso de renderización puede ponerse en práctica, pero será diferente de entre sí debido a la información de renderización de objetivo preferentemente en la forma de la matriz de renderización de objetivo, que es normalmente proporcionada por el usuario y de que mencionarán más adelante.

60 **[0024]** Finalmente, el combinador 364 usa la información de parámetro de objeto de audio 362 indicación preferentemente información de energía e información de correlación que describe los objetos de audio. En una modalidad, la información de parámetro de objeto de audio se les proporciona como una matriz de covariancia de objeto de audio para cada "loseta" en el plano de tiempo/frecuencia. Declarado diferentemente, para cada subbanda y para cada bloque temporal, en el cual esta subbanda se define, una completa matriz de covariancia de objeto, es decir, una matriz que tiene la información de fuerza/energía y la información de correlación se proporciona como la información de parámetro de objeto de audio 362.

65 **[0025]** Cuando la Figura 3b y la Figura 2a o 2b se comparan, se convierte claro que el decodificador de objeto de audio 102 en la Figura 1 corresponde al aparato para sintetizar una señal de salida reproducida.

5 **[0026]** Más aún, el procesador estéreo 201 incluye la etapa 356 de decorrelacionador de la Figura 3b. Por otra parte, el combinador 364 incluye la calculadora de la matriz 202 en la Figura 2a. Más aún, cuando la etapa 356 de decorrelacionador incluye una operación de conversión de multicanal a estéreo de decorrelacionador, esta porción de la calculadora de la matriz 202 se incluye en la etapa 356 de decorrelacionador en vez de en el combinador 364.

10 **[0027]** Sin embargo, cualquier ubicación específica de una cierta función no es decisiva aquí, ya que una realización de la presente invención en el software o dentro de un procesador de señal digital dedicado o incluso dentro de una computadora personal de objetivo general es en alcance de la presente invención. Por lo tanto, la atribución de una cierta función a un cierto bloque es un modo de poner en práctica la presente invención en el hardware. Cuando, sin embargo, todos los diagramas de circuito de bloque se consideran como diagramas de flujo para ilustrar un cierto flujo de etapas operacionales, se convierte claro que la contribución de ciertas funciones a un cierto bloque es libremente posible y puede llevarse a cabo según requerimientos de programación o realización.

15 **[0028]** Más aún, cuando la Figura 3b se compara a la Figura 3a, se convierte claro que la funcionalidad del combinador 364 para calcular factores de ponderación para la combinación ponderada se incluye en la calculadora de la matriz 202. Declarado diferentemente, la información de la matriz constituye una recolección de factores de ponderación que se aplican a la unidad de la matriz potenciada 303, que se pone en práctica en el combinador 364, pero que también puede incluir la porción de la etapa 356 de decorrelacionador (con respecto a la matriz **Q** como será mencionado más adelante). Así, la unidad de matriz potenciada 303 lleva a cabo la operación de combinación de preferentemente subbandas del al menos dos objeto abajo mezcla señales, donde la información de la matriz incluye factores de ponderación para ponderar a éstos al menos dos abajo señales de mezcla o la señal no correlacionada antes de llevar a cabo la operación de combinación.

25 **[0029]** Posteriormente, mencionan de la estructura detallada de una modalidad preferida del combinador 364 y la etapa 356 de decorrelacionador. Específicamente, mencionan de varias diferentes realizaciones de la funcionalidad de la etapa 356 de decorrelacionador y el combinador 364 con respecto a Figuras 4a a 4d. Las Figuras 4e a la Figura 4g ilustran realizaciones específicas de artículos en la Figura 4a a la Figura 4d. Antes de mencionar la Figura 4a a la Figura 4d detalladamente, mencionan de la estructura general de estas figuras. Cada figura incluye una rama superior relacionada con la señal no correlacionada y una rama inferior relacionada con la señal seca. Más aún, la señal de salida de cada rama, es decir, una señal a línea 450 y una señal a línea 452 se combina en un combinador 454 a fin de obtener finalmente la señal de salida reproducida 350. En términos generales, el sistema en la Figura 4a ilustra tres unidades de procesamiento de matriz 401, 402, 404. 401 es la unidad de mezcla de señal seca. Las al menos dos señales de conversión de multicanal a estéreo de objeto 352 son ponderadas y/o mezcladas entre sí para obtener dos mezclan en seco señales de objeto que corresponden las señales de la rama de señal seca que es introducida en el circuito sumador 454. Sin embargo, la rama de señal seca puede tener otra unidad de procesamiento de matriz, es decir, la unidad de compensación de ganancia 409, en la Figura 4d que se conecta río abajo de la unidad de mezcla de señal seca 401.

40 **[0030]** Más aún, la unidad de combinador 364 puede o no puede incluir la unidad de conversión de estéreo a multicanal de decorrelacionador 404 que tiene la matriz de conversión de estéreo a multicanal de decorrelacionador **P**.

45 **[0031]** Naturalmente, la separación de las unidades de matrización 404, 401 y 409 (Figura 4d) y la unidad de combinador 454 sólo es verdad artificialmente, aunque una realización correspondiente sea, por supuesto, posible. Alternativamente, sin embargo, las funcionalidades de estas matrices pueden ponerse en práctica vía una matriz "grande" individual que recibe, como una entrada, la señal 358 no correlacionada y la señal de conversión de multicanal a estéreo 352, y que envía los dos o tres o más canales de salida reproducidos 350. En una tan realización de la matriz grande, las señales a las líneas 450 y 452 no necesariamente pueden ocurrir, pero la funcionalidad de una tan "matriz grande" puede describirse en cierto modo que un resultado de una solicitud de esta matriz es representado por las diferentes suboperaciones llevadas a cabo por las unidades de matrización 404, 401 o 409 y una unidad de combinador 454, aunque los resultados intermedios 450 y 452 nunca puedan ocurrir de un modo explícito.

55 **[0032]** Más aún, la etapa 356 de decorrelacionador puede incluir la unidad de mezcla del predecorrelacionador 402 o no. La Figura 4b ilustra una situación, en la cual esta unidad no es proporcionada. Esto es específicamente útil cuando dos decorrelacionadores para las dos señales de canal de conversión de multicanal a estéreo se proporcionan y una conversión de multicanal a estéreo específica no es necesaria. Naturalmente, uno podría aplicar ciertos factores de ganancia a ambos canales de conversión de multicanal a estéreo o uno podría mezclar los dos canales de conversión de multicanal a estéreo antes de que ellos sean introducidos en una etapa de decorrelacionador según un requerimiento de realización específico. Por otra parte, sin embargo, la funcionalidad de matriz **Q** también puede incluirse en una matriz específica **P**. Esto significa que la matriz **P** en la Figura 4b es diferente de la matriz **P** en la Figura 4a, aunque el mismo resultado se obtenga. En vista de esto, la etapa 356 de decorrelacionador no puede incluir ninguna matriz en absoluto, y el completo cálculo de información de la matriz se lleva a cabo en el combinador y la completa solicitud de las matrices se lleva a cabo en el combinador también. Sin embargo, para la mejor ilustración de las funcionalidades técnicas detrás de éstos matemáticas, la descripción

subsecuente de la presente invención se llevará a cabo con respecto a esquema de procesamiento de matriz específico y técnicamente transparente ilustrado en Figuras 4a a 4d.

5 [0033] La Figura 4a ilustra la estructura de la unidad de matrización potenciada inventiva 303. La entrada **X** que comprende al menos dos canales se alimenta en la unidad de mezcla de señal seca 401 que lleva a cabo una operación de la matriz según la matriz de mezclado seco **C** y envía la señal de conversión de estéreo a multicanal seca **Y**. La entrada **X** también es alimentado en la unidad de mezcla del predecorrelacionador 402 que lleva a cabo una operación de la matriz según la prematriz de mezcla **Q** del predecorrelacionador y envía una señal de canal N_d a alimentar a la unidad de decorrelacionador 403. La señal decorrelacionada **Z** del canal resultante N_d es alimentada posterior- mente a la unidad de conversión de estéreo a multicanal de decorrelacionador 404 que lleva a cabo una operación de la matriz según la matriz de conversión de estéreo a multicanal de decorrelacionador **P** y envía una señal estéreo no correlacionada. Finalmente, la señal estéreo no correlacionada se mezcla por la adición en lo referente a canales simple con la señal de conversión de estéreo a multicanal seca estéreo **Y** a fin de formar la señal de salida **Y** de la unidad de matrización potenciada. Las tres matrices de mezcla (**C**, **Q**, **P**) están todas descritas por la información de la matriz suministrada al procesador estéreo 201 por la calculadora de la matriz 202. Un sistema de la técnica anterior sólo contendría la rama de señal seca inferior. Tal sistema funcionaría mal en el caso simple donde un objeto de música estéreo está contenido en un canal de conversión de multicanal a estéreo de objeto y un objeto de voz mono está contenido en el otro canal de conversión de multicanal a estéreo de objeto. Esto es así porque la renderización de la música al estéreo se basaría completamente en el paneo de frecuencia selectiva aunque un enfoque estéreo paramétrico que incluye la decorrelación se conozca para lograr la calidad de audio percibida mucho más elevada. Un completamente diferente sistema de técnica anterior que incluye la decorrelación pero basado en dos conversiones de multicanal a estéreo de objeto mono separadas llevaría a cabo mejor para este ejemplo particular, pero alcanzaría por otra parte la misma calidad que el primer sistema estéreo seco mencionado para un caso de conversión de multicanal a estéreo hacia atrás compatible donde la música es escrita en el estéreo verdadero y la voz se mezcla con pesos iguales a los dos canales de conversión de multicanal a estéreo de objeto. Como un ejemplo se toma en cuenta el caso de un objetivo de tipo del Karaoke que renderiza comprender el objeto de música estéreo por sí solo. Un tratamiento separado de cada uno de los canales de conversión de multicanal a estéreo luego permite una extinción menos óptima del objeto de voz que un tratamiento conjunto que tiene en cuenta la información de objeto de audio estéreo transmitida, como la correlación de intercanal. La característica crucial de la presente invención debe permitir la calidad de audio más elevada posible, no sólo en ambos de estas situaciones simples, pero también para mucho más combinaciones complejas de conversión de multicanal a estéreo de objeto y renderización.

35 [0034] La Figura 4b ilustra, como se declara anteriormente una situación donde, en contraste con la Figura 4a, la prematriz de mezcla del predecorrelacionador **Q** no se requiere o es "absorbida" en la matriz de conversión de estéreo a multicanal de decorrelacionador **P**. La Figura 4c ilustra una situación, en la cual la matriz de predecorrelacionador **Q** se proporciona y puesto en práctica en la etapa 356 de decorrelacionador, y en que la matriz de conversión de estéreo a multicanal de decorrelacionador **P** no se requiere o es "absorbida" en la matriz **Q**.

40 [0035] Más aún, la Figura 4d ilustra una situación, en la cual las mismas matrices que en la Figura 4a se encuentran, pero en que una matriz de compensación de ganancia adicional **G** se proporciona que es específicamente útil en la tercera modalidad para ser mencionada en relación con Figura 13 y la cuarta modalidad para ser mencionada en la Figura 14.

45 [0036] La etapa 356 de decorrelacionador puede incluir un decorrelacionador individual o dos decorrelacionadores. La Figura 4e ilustra una situación, en la cual un decorrelacionador individual 403 se proporciona y en que la señal de conversión de multicanal a estéreo es una señal de conversión de multicanal a estéreo de objeto de dos canales, y la señal de salida es una señal de salida de audio de dos canales. En este caso, la matriz de conversión de multicanal a estéreo de decorrelacionador **Q** tiene una línea y dos columnas, y la matriz de conversión de estéreo a multicanal de decorrelacionador tiene una columna y dos líneas. Cuando, sin embargo, la señal de conversión de multicanal a estéreo tendría más de dos canales, luego el número de columnas de **Q** igualaría al número de canales de la señal de conversión de multicanal a estéreo, y cuando la señal de salida reproducida sintetizada tendría más de dos canales, luego la matriz de conversión de estéreo a multicanal de decorrelacionador **P** tendría varias líneas iguales al número de canales de la señal de salida reproducida.

55 [0037] La Figura 4f ilustra una realización parecida a un circuito de la unidad de mezcla de señal seca 401, que se indica como **C₀** y que tiene, en los dos por dos modalidad, dos líneas en dos columnas. Los elementos de la matriz se ilustran en la estructura parecida a un circuito como los factores de ponderación c_{ij} . Más aún, los canales ponderados se combinan usando circuitos sumadores como es visible de la Figura 4f. Cuando, sin embargo, el número de canales de conversión de multicanal a estéreo es diferente del número de canales de señal de salida reproducidos, luego la matriz de mezclado seco **C₀** no será una matriz cuadrática, pero tendrá varias líneas que es diferente del número de columnas.

65 [0038] La Figura 4g ilustra detalladamente la funcionalidad de adicionar la etapa 454 en la Figura 4a. Específicamente, para el caso de dos canales de salida, como la señal de canal de estéreo de izquierda y la señal de canal estéreo derecha, dos diferentes etapas de circuito sumador 454 se proporcionan, que combinan señales de

salida de la rama superior relacionada con la señal de decorrelacionador y la rama inferior relacionada con la señal seca como ilustrado en la Figura 4g.

5 **[0039]** En cuanto a la matriz de compensación de ganancia **G** 409, los elementos de la matriz de compensación de ganancia sólo están en la diagonal de la matriz **G**. En los dos por dos caso, que se ilustra en la Figura 4f para la matriz de mezclado de señal seca **C₀**, un factor de ganancia para la compensación de la ganancia la izquierda la señal seca sería a la posición de **c₁₁**, y un factor de ganancia para la compensación de la ganancia la señal seca derecha sería a la posición de **c₂₂** de **C₀** de la matriz en la Figura 4f. Los valores para **c₁₂** y **c₂₁** serían iguales a 0 en los dos por dos matriz de ganancia **G** como ilustrado a 409 en la Figura 4d.

10 **[0040]** La Figura 5 ilustra la operación de técnica anterior de un decorrelacionador de varios canales 403. Tal herramienta se usa por ejemplo en el Sonido envolvente de MPEG. Las señales **N_d** de la señal 1, la señal 2..., la señal son por separado alimentadas en, decorrelacionador 1, decorrelacionador 2, ... decorrelacionador **N_d**. Cada decorrelacionador por lo común comprende un filtro que apunta a producir una salida que es tan no correlacionada como sea posible con la entrada, a un mantenimiento de la energía de señal de entrada. Más aún, los diferentes filtros de decorrelacionador se seleccionan tal que la señal 1 de decorrelacionador de salidas, la señal 2 de decorrelacionador,... la señal de decorrelacionador **N_d** también sea tan no correlacionadas como sea posible en un sentido par. Ya que los decorrelacionadores son por lo común de la elevada complejidad computacional comparado con otras partes de un decodificador de objeto de audio, es de interés para mantener el número **N_d** tan bajo como sea posible.

15 **[0001]** La presente invención ofrece soluciones para **N_d** igual a 1, 2 o más, pero preferentemente menor que el número de objetos de audio. Específicamente, el número de decorrelacionadores es, en una modalidad preferida, igual al número de señales de canal de audio de la señal de salida reproducida o incluso más pequeño que el número de señales de canal de audio de la señal de salida reproducida 350.

20 **[0041]** A continuación texto, una descripción matemática de la presente invención será detallada. Todas las señales consideradas aquí son muestras de subbanda de un banco de filtros modulado o el análisis FFT con ventanaje de señales de tiempo discreto. Se entiende que estas subbandas tienen que ser transformadas nuevamente al dominio temporal individual por operaciones de banco de filtros de síntesis correspondientes. Una bloque de señal de **L** muestras representa la señal en un intervalo de tiempo y frecuencia que es una parte de tejar perceptivamente motivado del plano de frecuencia del tiempo que se aplica para la descripción de propiedades de señal. En este fraguado, los objetos de audio proporcionados pueden ser representados como **N** hileras de longitud **L** en una matriz,

$$\mathbf{S} = \begin{bmatrix} s_1(0) & s_1(1) & \dots & s_1(L-1) \\ s_2(0) & s_2(1) & \dots & s_2(L-1) \\ \vdots & \vdots & & \vdots \\ s_N(0) & s_N(1) & \dots & s_N(L-1) \end{bmatrix}. \quad (1)$$

35 **[0043]** La Figura 6 ilustra una modalidad de un mapa de objeto de audio que ilustra varios **N** objetos. En la explicación ejemplificante de la Figura 6, cada objeto tiene un objeto ID, un archivo de audio de objeto correspondiente y, importantemente, información de parámetro de objeto de audio que es, preferentemente, la información que se relaciona con la energía del objeto de audio y a la correlación de interobjeto del objeto de audio. Específicamente, la información de parámetro de objeto de audio incluye una matriz de covariancia de objeto **E** para cada subbanda y para cada bloque temporal.

40 **[0044]** Un ejemplo para tal matriz de información de parámetro de audio de objeto **E** se ilustra en la Figura 7. Los elementos diagonales **e_{ii}** incluyen la energía o la información de energía del objeto de audio **i** en la subbanda correspondiente y el bloque temporal correspondiente. Con este fin, la señal de subbanda que representa un cierto objeto de audio soy introducido en una energía o calculadora de energía que puede llevar a cabo, por ejemplo, una función de correlación automática (acf) para obtener el valor **e₁₁** con o sin un poco de normalización. Alternativamente, la energía puede calcularse como la suma de los cuadrados de la señal sobre una cierta longitud (es decir el producto de vector: **ss***). El acf puede en algún sentido describir la distribución espectral de la energía, pero debido a que una transformada T/F para la selección de frecuencia se utiliza preferiblemente de todos modos, el cálculo de energía puede llevarse a cabo sin un acf para cada subbanda por separado. Así, los elementos diagonales principales de la matriz de parámetro de audio de objeto **E** indican una medida para la energía de energía de un objeto de audio en una cierta subbanda en un cierto bloque temporal.

55 **[0045]** Por otra parte, el elemento no diagonal **e_{ij}** indica una medida de correlación respectiva entre objetos de audio **i**, **j** en la subbanda correspondiente y bloque temporal. Es claro de la Figura 7 que la matriz **E** es, para entradas

valoradas reales, simétricas con respecto a la diagonal principal. En términos generales, esta matriz es una matriz hermítica. El e_{ij} de elemento de medida de correlación puede calcularse, por ejemplo, por una correlación cruzada de las dos señales de subbanda de los objetos de audio respectivos de modo que una medida de correlación cruzada se obtenga que puede o no puede ser normalizado. Otras medidas de correlación pueden usarse que no son calculados usando una operación de correlación cruzada, pero que son cuentan por otros modos de determinar la correlación entre dos señales. Por motivos prácticos, todos los elementos de la matriz **E** son normalizados de modo que ellos tengan magnitudes entre 0 y 1, donde 1 indica una energía máxima o una correlación máxima y 0 indica que una energía mínima (energía cero) y -1 indica una correlación mínima (desfasada).

5
10 **[0046]** La matriz de conversión de multicanal a estéreo **D**, de tamaño $K \times N$ donde $K > 1$ determina la señal de conversión de multicanal a estéreo K de canal en la forma de una matriz con K hileras a través de la multiplicación de la matriz

$$\mathbf{X} = \mathbf{DS} . \quad (2)$$

15 **[0047]** La Figura 8 ilustra un ejemplo de una matriz de conversión de multicanal a estéreo **D** que tiene elementos de matriz de conversión de multicanal a estéreo d_{ij} . Estos elementos d_{ij} indican si una porción o el objeto j completo se incluye en la señal de conversión de multicanal a estéreo de objeto i o no. Cuando, por ejemplo, d_{12} es igual al cero, esto significa que el objeto 2 no es incluido en la señal de conversión de multicanal a estéreo de objeto 1. Por otra parte un valor del d_{23} igual a 1 indica que se oponen 3 es completamente incluido en la señal de conversión de multicanal a estéreo de objeto 2.

20 **[0048]** Los valores de elementos de matriz de conversión de multicanal a estéreo entre 0 y 1 son posibles. Específicamente, el valor de 0.5 indica que un cierto objeto se incluye en una señal de conversión de multicanal a estéreo, pero sólo con la mitad de su energía. Así, cuando un objeto de audio tal objeto el número 4 es igualmente distribuido a ambos canales de señal de conversión de multicanal a estéreo, luego d_{24} y d_{14} sería igual a 0.5. Este modo de la conversión de multicanal a estéreo es una operación de conversión de multicanal a estéreo que conserva la energía que se prefiere para algunas situaciones. Alternativamente, sin embargo, una conversión de multicanal a estéreo de conservación de no energía puede usarse también, en que el objeto de audio completo es introducido en el canal de conversión de multicanal a estéreo de izquierda y el canal de conversión de multicanal a estéreo derecho de modo que la energía de este objeto de audio haya sido duplicada con respecto a los otros objetos de audio dentro de la señal de conversión de multicanal a estéreo.

25 **[0049]** A la porción inferior de la Figura 8, un diagrama esquemático del codificador de objeto 101 de la Figura 1 se les proporciona. Específicamente, el codificador de objeto 101 incluye dos diferentes porciones 101a y 101b. La porción 101a es un convertidor de multicanal a estéreo que preferentemente lleva a cabo una combinación lineal ponderada de objetos de audio 1, 2, ..., N , y la segunda porción del codificador de objeto 101 es una calculadora de parámetro de objeto de audio 101b, que calcula la información de parámetro de objeto de audio, como la matriz **E** para cada bloque temporal o subbanda a fin de proporcionar la energía de audio e información de correlación que es una información paramétrica y puede transmitirse, por lo tanto, con una baja velocidad binaria o puede almacenarse consumiendo una pequeña cantidad de recursos de memoria.

30 **[0050]** El usuario controló la matriz de renderización de objeto de tamaño determina la renderización de objetivo de canal de los objetos de audio en la forma de una matriz con hileras a través de la multiplicación de la matriz;

$$\mathbf{Y} = \mathbf{AS} . \quad (3)$$

35 **[0051]** Será asumido durante todo la derivación que $M=2$ sigue que ya que el foco está en la renderización estéreo. Considerando una matriz de renderización inicial a más de dos canales, y una conversión de multicanal a estéreo gobiernan de aquellos varios canales en dos canales es obvio para aquellos expertos en la técnica derivar la matriz de renderización **A** correspondiente de tamaño $2 \times N$ para la renderización estéreo. Esta reducción se lleva a cabo en el reductor de renderización **204**. También se supondrá para la simplicidad que $K=2$ tal que la conversión de multicanal a estéreo de objeto también sea una señal estéreo. El caso de una conversión de multicanal a estéreo de objeto estéreo es además el caso especial más importante en términos de casos de solicitud.

40 **[0052]** La Figura 9 ilustra una explicación detallada de la matriz de renderización de objetivo **A**. Según la solicitud, el objetivo que renderiza la matriz **A** se proporciona por el usuario. El usuario tiene la libertad total de indicar, donde un objeto de audio debería ubicarse en una manera virtual para un sistema de repetición. La fuerza del concepto de objeto de audio es que la información de conversión de multicanal a estéreo y la información de parámetro de objeto de audio son completamente independientes en una localización específica de los objetos de audio. Esta localización de objetos de audio se proporciona por un usuario en la forma de la información de renderización de objetivo. Preferentemente, la información de renderización objetivo puede ponerse en práctica como una matriz de renderización objetivo **A** que puede estar en forma de la matriz en la Figura 9. Específicamente, la matriz de

renderización **A** tiene M de líneas y columnas N, donde M es igual al número de canales en la señal de salida reproducida, y donde el N es igual al número de objetos de audio. M es igual a dos del caso de renderización de estéreo preferido, pero si una renderización de canal M se lleva a cabo, entonces la matriz **A** tiene M de líneas.

5 **[0053]** Específicamente, un elemento de la matriz a_{ij} , indica si una porción o el objeto j completo deben volverse en el canal de salida específico i o no. La porción inferior de la Figura 9 proporciona un ejemplo simple para la matriz de renderización de objetivo de un caso, en el cual hay seis objetos de audio AO1 a AO6 donde sólo los cinco primeros objetos de audio deberían volverse posiciones a específicas y que el sexto objeto de audio no debería volverse en absoluto.

10 **[0054]** En cuanto al objeto de audio AO1, el usuario quiere aquel este objeto de audio se vuelve a la izquierda de un caso de repetición. Por lo tanto, este objeto se coloca a la posición de un altavoz de izquierda en un cuarto de repetición (virtual), que da como resultado la primera columna de la matriz de renderización un para ser (10). En cuanto al segundo objeto de audio, a_{22} es un y a_{12} es 0 que significa que el segundo objeto de audio debe volverse en el lado derecho.

15 **[0055]** El objeto de audio 3 debe volverse en el medio entre el orador de izquierda y el orador derecho de modo que el 50 % del nivel o señal de este objeto de audio entre en el canal de izquierda y el 50 % del nivel o señal entra en el canal derecho de modo que la tercera columna correspondiente de la matriz de renderización de objetivo **A** sea (longitud 0.5).

20 **[0056]** De forma similar, cualquier colocación entre el orador de izquierda y el orador derecho puede indicarse por la matriz de renderización de objetivo. En cuanto al objeto de audio 4, la colocación es más al lado derecho, ya que el elemento de la matriz a_{24} es más grande que a_{14} . De forma similar, el quinto objeto de audio que A_{05} se vuelve para ser más al altavoz de izquierda como se indica por el objetivo que renderiza elementos de la matriz a_{15} y a_{25} . La matriz de renderización de objetivo **A** además permite no renderizar un cierto objeto de audio en absoluto. Esto es ejemplarmente ilustrado por la sexta columna de la matriz de renderización de objetivo **A**, que tiene elementos cero.

25 **[0057]** Será asumido durante todo la derivación que $M = 2$ puesto que el foco está en la renderización estéreo. Considerando una matriz de renderización inicial a más de dos canales, y una conversión de multicanal a estéreo gobiernan de aquellos varios canales en dos canales es obvio para aquellos expertos en la técnica derivar la matriz de renderización **A** correspondiente de tamaño $2 \times N$ para la renderización estéreo. Esta reducción se lleva a cabo en el reductor de renderización 204. También se supondrá para la simplicidad que $K = 2$ tal que la conversión de multicanal a estéreo de objeto también sea una señal estéreo. El caso de una conversión de multicanal a estéreo de objeto estéreo es además el caso especial más importante en términos de casos de solicitud.

30 **[0058]** Omitiendo durante un momento los efectos de la codificación con pérdidas de la señal de audio de conversión de multicanal a estéreo de objeto, la tarea del decodificador de objeto de audio es generar una aproximación en el sentido perceptual de la renderización objetivo **Y** de los objetos de audio originales, considerando la matriz de renderización **A**, la conversión de multicanal a estéreo **X** la matriz de conversión de multicanal a estéreo **D**, y parámetros de objeto. La estructura de la unidad de matrización potenciada inventiva 303 se les proporciona en Figura 4. Considerando un número N_d de decorrelacionadores mutuamente ortogonales en 403, hay tres matrices mezcladas.

- 35
- 45 • **C** de tamaño 2×2 lleva a cabo la mezcla de señal seca
 - **D** de tamaño $N_d \times 2$ lleva a cabo la mezcla de predecorrelacionador
 - 50 • **P** de tamaño $2 \times N_d$ lleva a cabo la conversión de estéreo a multicanal de decorrelacionador

[0059] La asunción de los decorrelacionadores es la conservación de energía, la matriz de señal no correlacionada **Z** tiene una matriz de covariancia de diagonal $N_d \times N_d$ $\mathbf{RZ} = \mathbf{ZZ}^*$ cuyos valores diagonales son iguales a aquellos de la matriz de covariancia;

$$\mathbf{QXX}^*\mathbf{Q}^* \quad (4)$$

55 de la conversión de multicanal a estéreo de objeto procesada de mezcla de predecorrelacionador. (Aquí y en el seguir, el asterisco denota que el conjugado complejo transporta la operación de la matriz. También se entiende que las matrices de covariancia deterministas de la forma \mathbf{UV}^* que se usan durante todo para conveniencia computacional pueden ser sustituidas por expectativas $\mathbf{E}\{\mathbf{UV}^*\}$. Más aún, puede suponerse que todas las señales no correlacionadas sean no correlacionadas de las señales de conversión de multicanal a estéreo de objeto. Por lo tanto, la covariancia **R'** de la salida combinada de la unidad de matrización potenciada inventiva 303,

60

$$\mathbf{Y}' = \hat{\mathbf{Y}} + \mathbf{PZ} = \mathbf{CX} + \mathbf{PZ}, \quad (5)$$

puede ser escrito como una suma de la covariancia $\hat{\mathbf{R}} = \hat{\mathbf{Y}}\hat{\mathbf{Y}}^*$ de la señal seca $\hat{\mathbf{Y}} = \mathbf{CX}$ mezclan y la covariancia de salida de decorrelacionador resultante;

$$\mathbf{R}' = \hat{\mathbf{R}} + \mathbf{PR}_z\mathbf{P}' . \quad (6)$$

5 **[0060]** Los parámetros de información de objeto por lo común transportable sobre energías de objeto y correlaciones seleccionadas del interobjeto. De estos parámetros, se logra un modelo \mathbf{E} de la covariancia $N \times N$ del objeto \mathbf{SS}' .

$$\mathbf{SS}' = \mathbf{E} . \quad (7)$$

10 **[0061]** Los datos disponibles para el decodificador de objeto de audio son en este caso descrito por el trillizo de matrices $(\mathbf{D}, \mathbf{E}, \mathbf{A})$ y el procedimiento mostrado por la presente invención comprende usar estos datos para optimizar conjuntamente la correspondencia de forma de onda de la salida combinada (5) y su covariancia (6) a la señal (4) de renderización de objetivo. Para una matriz de mezclado de señal seca proporcionada, el problema a mano es apuntar a la covariancia con especificidad de objetivo correcta $\mathbf{R} = \mathbf{R}'$ que puede ser estimada por;

$$\mathbf{R} = \mathbf{Y}\mathbf{Y}' = \mathbf{A}\mathbf{S}\mathbf{S}'\mathbf{A}' = \mathbf{A}\mathbf{E}\mathbf{A}' . \quad (8)$$

[0062] Con la definición de la matriz de error;

$$\Delta\mathbf{R} = \mathbf{R} - \hat{\mathbf{R}} , \quad (9)$$

15 una comparación con (6) da como resultado al requerimiento de diseño

$$\mathbf{PR}_z\mathbf{P}' = \Delta\mathbf{R} . \quad (10)$$

20 **[0063]** Ya que el lado de mano izquierda (de 10) es una matriz semidefinida positiva para cualquier opción de la matriz de mezcla del predecorrelacionador \mathbf{P} , es necesario que la matriz de error (de 9) sea una matriz semidefinida positiva también. A fin de esclarecer los detalles de las fórmulas subsecuentes, deje a las covariancias de la mezcla de señal seca y la renderización del objetivo de parámetros como sigue;

$$\mathbf{R} = \begin{bmatrix} L & p \\ p & R \end{bmatrix}, \quad \hat{\mathbf{R}} = \begin{bmatrix} \hat{L} & \hat{p} \\ \hat{p} & \hat{R} \end{bmatrix} . \quad (11)$$

[0064] Para la matriz de error

$$\Delta\mathbf{R} = \begin{bmatrix} \Delta L & \Delta p \\ \Delta p & \Delta R \end{bmatrix} = \begin{bmatrix} L - \hat{L} & p - \hat{p} \\ p - \hat{p} & R - \hat{R} \end{bmatrix}, \quad (12)$$

25 el requerimiento necesario para ser positivo semidefinido puede expresarse como las tres condiciones

$$\Delta L \geq 0, \Delta R \geq 0, \Delta L \Delta R - (\Delta p)^2 \geq 0. \quad (13)$$

5 [0065] Posteriormente, mencionan de la Figura 10. La Figura 10 ilustra una recolección de algunas etapas que precuentan que son preferentemente preformadas para cuatro modalidades para ser mencionadas en relación con Figuras 11 a 14. Una tal etapa de precálculo es el cálculo de R de matriz de covariancia de la señal de renderización de objetivo como indicado a 1000 en la Figura 10. Obstrúyase 1000 corresponde a la ecuación (8).

10 [0066] Como es indicado en el bloque 1002, la matriz de mezclado seco puede calcularse usando la ecuación (15). Particularmente, la matriz de mezclado seco C_0 se calcula tal que el una mejor correspondencia de la señal de renderización de objetivo se obtenga al usar las señales de conversión de multicanal a estéreo, al suponer que la señal no correlacionada no fuera agregarse en absoluto. Así, la matriz de mezclado seco se asegura que una forma de onda de señal de salida de matriz de mezclado iguala la señal de renderización de objetivo como cerca como sea posible sin cualquier señal no correlacionada adicional. Este requisito previo para la matriz de mezclado seco es particularmente útil para mantener la porción de la señal no correlacionada en el canal de salida tan bajo como sea posible. En términos generales, la señal no correlacionada es una señal que tiene sido modificado por el decorrelacionador en gran medida. Así, esta señal por lo general tiene artefactos tal colorización, tiempo corriéndose y respuesta mal transitoria. Por lo tanto, esta modalidad proporciona la ventaja lo que menos hace señas del proceso de decorrelación por lo general da como resultado una mejor calidad de salida de audio. A una realización una correspondencia de forma de onda, es decir, ponderación y combinación de los dos canales o más canales en la señal de conversión de multicanal a estéreo de modo que estos canales después de la operación de mezcla seca se acerquen a la señal de renderización de objetivo como cerca como sea posible, sólo una cantidad mínima de la señal no correlacionada es necesaria.

25 [0067] El combinador 364 es operativo para calcular los factores de ponderación tan el resultado 452 de una operación de mezclado de la primera señal de conversión de multicanal a estéreo de objeto y la segunda señal de conversión de multicanal a estéreo de objeto se igualan por la forma de onda a un resultado de renderización de objetivo, que iba por lo que posible corresponden a una situación que se obtendría, a una renderización de los objetos de audio originales usando la información de renderización de objetivo 360 a condición de que la información de objeto de audio paramétrica 362 fuera una pérdida menos representación de los objetos de audio. Por lo tanto, la reconstrucción exacta de la señal nunca puede ser garantizada, incluso con una matriz E no cuantificada. Uno minimiza el error en una media sentido cuadrado. Por lo tanto, uno apunta a la adquisición de una correspondencia de forma de onda, y las energías y las correlaciones cruzadas son reconstruidas.

30 [0068] Tan pronto como la matriz de mezclado seco C_0 se calcula p.ej del modo superior, luego la matriz de covariancia \hat{R}_0 de la señal de mezcla seca puede calcularse. Específicamente, se prefiere usar la ecuación escrita a la derecha de la Figura 10, es decir, $C_0 D E D' C_0'$. Esta fórmula de cálculo se asegura que, para el cálculo de

40 la matriz de covariancia \hat{R}_0 del resultado de la mezcla de señal seca, sólo los parámetros son necesarios, y las muestras de subbanda no se requieren. Alternativamente, sin embargo, uno podría calcular la matriz de covariancia del resultado de la mezcla de señal seca usando la matriz de mezclado seco C_0 y las señales de conversión de multicanal a estéreo también, pero el primer cálculo que ocurre en el dominio de parámetro sólo es de la complejidad inferior.

45 [0069] Subsecuente a las etapas de cálculo 1000, 1002, 1004 la matriz de mezclado de señal seca C_0 , R de matriz de covariancia de la señal de renderización de objetivo y la matriz de covariancia de la señal de mezcla seca están disponibles.

50 [0070] Para la determinación específica de matrices Q , P cuatro diferentes modalidades son descritos posteriormente. Además, una situación de Figura 4d (por ejemplo para la tercera modalidad y la cuarta modalidad) se describe, en que los valores de la matriz de compensación de ganancia G se determinan también Aquellos expertos en la técnica observarán que allí existen otras modalidades para calcular los valores de estas matrices, ya que allí existe un poco de grado de la libertad para determinar los factores de ponderación de la matriz requeridos.

[0071] En una primera modalidad de la presente invención, la operación de la calculadora de la matriz 202 es diseñada como sigue. La matriz de conversión de estéreo a multicanal seca es primero derivada para lograr la menor parte de solución de cuadrados con la correspondencia de forma de onda de señal

$$\hat{Y} = CX \approx Y = AS, \quad (14)$$

$$\hat{Y}_0 = C_0 \cdot X = C_0 \cdot D \cdot S$$

[0072] En este contexto, esto se observa que es válido. Más aún, la ecuación que sigue es verdadera:

$$\begin{aligned} \hat{R}_0 &= \hat{Y}_0 \hat{Y}_0^* = C_0 \cdot D \cdot S \cdot (C_0 \cdot D \cdot S)^* = \\ &= C_0 \cdot D \cdot (S \cdot S^*) \cdot D^* \cdot C_0^* = C_0 \cdot D \cdot E \cdot D^* \cdot C_0^* \end{aligned}$$

5 [0073] La solución a este problema se les proporciona por

$$C \approx C_0 = AED^*(DED^*)^{-1} \quad (15)$$

y esto tiene la propiedad conocida adicional de la menor parte de soluciones de cuadrados, que también pueden ser fácilmente verificadas (de 13) que el error $\Delta Y = Y - \hat{Y}_0 = AS - C_0 X$ es ortogonal a la aproximación $\hat{Y} = C_0 X$. Por lo tanto, los términos cruzados desaparecen en el cálculo que sigue,

$$\begin{aligned} R &= YY^* = (\hat{Y}_0 + \Delta Y)(\hat{Y}_0 + \Delta Y)^* \\ &= \hat{Y}_0 \hat{Y}_0^* + (\Delta Y)(\Delta Y)^* \\ &= \hat{R}_0 + (\Delta Y)(\Delta Y)^* \end{aligned} \quad (16)$$

10 [0074] se deduce que;

$$\Delta R = (\Delta Y)(\Delta Y)^*, \quad (17)$$

que trivialmente es positivamente semidefinido tal que (10) pueda solucionarse. De un modo simbólico la solución es;

$$P = TR_z^{-1/2}, \quad (18)$$

15 [0075] Aquí el segundo factor $R_z^{-1/2}$ es simplemente definido por la operación sabia por el elemento en la diagonal, y de la matriz T soluciona la ecuación de la matriz $TT^* = \Delta R$. Hay una libertad grande en la opción de solución con esta ecuación de la matriz. El procedimiento mostrado por la presente invención es comenzar desde la descomposición de valor singular de ΔR . Para esta matriz simétrica esto reduce la descomposición de vector propio habitual,

20

$$\Delta\mathbf{R} = \mathbf{U} \begin{bmatrix} \lambda_{\max} & 0 \\ 0 & \lambda_{\min} \end{bmatrix} \mathbf{U}^*; \quad \mathbf{U} = \begin{bmatrix} u_1 & u_2 \\ u_2 & -u_1 \end{bmatrix}, \quad (19)$$

donde la matriz de vector propio \mathbf{U} es unitaria y sus columnas contienen los vectores propios correspondientes a los valores propios clasificados en el tamaño decreciente $\lambda_{\max} \geq \lambda_{\min} \geq 0$. La primera solución con un decorrelacionador ($N_d = 1$) mostrado por la presente invención se obtiene estableciendo $\lambda_{\min} = 0$ en (19), y al insertar la aproximación natural correspondiente;

$$\mathbf{T} \approx \begin{bmatrix} u_1 \sqrt{\lambda_{\max}} \\ u_2 \sqrt{\lambda_{\max}} \end{bmatrix} \quad (20)$$

en (18). La solución completa con $N_d = 2$ decorrelacionadores se obtiene agregando la contribución ausente menos significativa del valor propio λ_{\min} más pequeño de $\Delta\mathbf{R}$ y adición de una segunda columna (a 20) correspondiente a un producto del primer factor \mathbf{U} (de 19) y la raíz cuadrada de todo elemento de la matriz de valores propios diagonal. Escrito detalladamente esto asciende a;

$$\mathbf{T} = \begin{bmatrix} u_1 \sqrt{\lambda_{\max}} & u_2 \sqrt{\lambda_{\min}} \\ u_2 \sqrt{\lambda_{\max}} & -u_1 \sqrt{\lambda_{\min}} \end{bmatrix}. \quad (21)$$

[0076] Posteriormente, el cálculo de matriz \mathbf{P} de acuerdo con la primera modalidad es resumido en relación con Figura 11. En la etapa 1101, la matriz de covariancia $\Delta\mathbf{R}$ de la señal de error \mathbf{o} , cuando la Figura 4a se considera, que la señal correlacionada a la rama superior se calcula al usar los resultados de etapa 1000 y etapa 1004 de la Figura 10. Luego, una descomposición de valor propio de esta matriz se lleva a cabo que tiene sido mencionado en relación con ecuación (19). Luego, la matriz \mathbf{Q} se selecciona de acuerdo con una de una pluralidad de estrategias disponibles de que mencionarán más adelante. Basado en la matriz seleccionada \mathbf{Q} , la matriz de covariancia \mathbf{R}_z de la señal no correlacionada matrizada se calcula usando la ecuación escrita a la derecha de la caja 1103 en la Figura 11, es decir, la multiplicación de matrices $\mathbf{Q}\mathbf{E}\mathbf{D}^*\mathbf{Q}^*$. Luego, basado en \mathbf{R}_z como obtenido en la etapa 1103, la matriz de conversión de estéreo a multicanal de decorrelacionador \mathbf{P} se calcula. Es claro que esta matriz no necesariamente tiene que llevar a cabo una conversión de estéreo a multicanal actual mencionando que a la salida de bloque \mathbf{P} 404 en la Figura 4a es más señales de canal que a la entrada. Esto puede llevarse a cabo en caso de un correlador individual, pero en caso de dos decorrelacionadores, la matriz de conversión de estéreo a multicanal de decorrelacionador \mathbf{P} recibe dos canales de entrada y envía dos canales de salida y puede ponerse en práctica como la matriz de conversión de estéreo a multicanal seca ilustrada en la Figura 4f.

[0077] Así, la primera modalidad es exclusiva en aquella \mathbf{C}_0 y \mathbf{P} se calculan. Es referido que, a fin de garantizar la estructura de correlación resultante correcta de la salida, uno necesita dos decorrelacionadores. Por otra parte, es una ventaja para ser capaz de usar sólo un decorrelacionador. Esta solución se indica por la ecuación (20). Específicamente, el decorrelacionador que tiene el valor propio más pequeño se pone en práctica.

[0078] En una segunda modalidad de la presente invención la operación de la calculadora de la matriz 202 es diseñada como sigue. La matriz de mezcla del predecorrelacionador es restringida para ser de la forma;

$$\mathbf{P} = c \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \quad (22)$$

[0079] Con esta restricción la matriz de covariancia de señal no correlacionada individual es un escalar $R_z = r_z$ y la covariancia de la salida combinada (6) se convierte;

$$\mathbf{R}' = \hat{\mathbf{R}} + \mathbf{P}\mathbf{R}_z\mathbf{P}' = \begin{bmatrix} \hat{L} & \hat{p} \\ \hat{p} & \hat{R} \end{bmatrix} + \alpha \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad (23)$$

5 donde $\alpha = c^2 r_z$. Una correspondencia total de la covarianza objetivo $\mathbf{R}' = \mathbf{R}$ es imposible en general, pero la correlación normalizada perceptivamente importante entre los canales de salida puede ser ajustada a aquel del objetivo en un amplio intervalo de situaciones. Aquí, la correlación con especificidad de objetivo se define por;

$$\rho = \frac{p}{\sqrt{LR}}, \quad (24)$$

correlación lograda por la salida combinada (23) se les proporciona por y la

$$\rho' = \frac{\hat{p} - \alpha}{\sqrt{(\hat{L} + \alpha)(\hat{R} + \alpha)}}. \quad (25)$$

[0080] Comparando (24) (y 25) da como resultado a una ecuación cuadrática en α ,

$$\rho^2 (\hat{L} + \alpha)(\hat{R} + \alpha) = (\hat{p} - \alpha)^2. \quad (26)$$

10 [0081] Para los casos donde (26) tiene una solución positiva $\alpha = \alpha_0 > 0$, la segunda modalidad de la presente

$$c = \sqrt{\alpha_0 / r_z}$$

invencción enseña usar constante en la definición (22) de matriz de mezclado. Si ambas soluciones (de 26) son positivas, el que produce una norma más pequeña debe usarse. En el caso donde ninguna tal solución existe, la contribución de decorrelacionador se configura al cero mediante una elección $c = 0$, puesto que las soluciones complejas de c dan como resultado distorsiones de fase perceptibles en las señales no

15 correlacionadas. El cálculo de \hat{p} puede ponerse en práctica de dos diferentes modos, directamente de la señal \mathbf{Y} o incorporación de la matriz de covarianza de objeto combinada con la conversión de multicanal a estéreo y renderización de la información, como $\hat{\mathbf{R}} = \mathbf{C}\mathbf{D}\mathbf{E}\mathbf{D}^*\mathbf{C}^*$. Aquí el primer procedimiento dará como resultado de

valores complejos \hat{p} y por lo tanto, en el lado derecho de (26) el cuadrado debe obtenerse de la parte real o magnitud de $(\hat{p} - \alpha)$, respectivamente. Alternativamente, sin embargo, incluso un valor complejo \hat{p} puede usarse. Un valor tan complejo indica una correlación con un término de fase específico que también es útil para modalidades específicas.

[0082] Una característica de esta modalidad, cuando esto puede observarse (de 25), es que esto sólo puede

25 disminuir la correlación comparada con aquella de la mezcla seca. Es decir $\rho' \leq \hat{\rho} = \hat{p} / \sqrt{\hat{L}\hat{R}}$.

[0083] Para resumir, la segunda modalidad se ilustra como se muestra en la Figura 12. Esto comienza con el cálculo de la matriz de covarianza $\Delta\mathbf{R}$ en la etapa 1101, que es idéntico a la etapa 1101 en la Figura 11. Luego, la ecuación (22) se pone en práctica. Específicamente, el aspecto de matriz \mathbf{P} es predeterminado y sólo el factor de ponderación

c que es idéntico para ambos elementos de P está abierto para calcularse. Específicamente, una matriz P que tiene una columna individual indica que sólo un decorrelacionador individual se usa en esta segunda modalidad. Más aún, las señales de los elementos de p aclaran que la señal no correlacionada se agrega a un canal, como el canal de izquierda de la señal de mezcla seca y es restada del canal derecho de la señal de mezcla seca. Así, una decorrelación máxima se obtiene agregando la señal no correlacionada a un canal y restar la señal no correlacionada del otro canal. A fin de determinar el valor c, etapas 1203, 1206, 1103, y 1208 se lleva a cabo. Específicamente, la hilera de correlación con especificidad de objetivo como indicado en la ecuación (24) se calcula en la etapa 1203. Este valor es el valor de correlación cruzado de intercanal entre las dos señales de canal de audio cuando una renderización estéreo se lleva a cabo. Basado en el resultado de etapa 1203, el factor de ponderación Δ se determina como indicado en la etapa 1206 basado en la ecuación (26). Más aún, los valores para los elementos de la matriz de la matriz Q se seleccionan y la matriz de covariancia, que es en este caso sólo un valor escalar Rz se calcula como indicado en la etapa 1103, y como ilustrado por la ecuación a la derecha de la caja 1103, en la Figura 12. Finalmente, el factor c se calcula como indicado en la etapa 1208. La ecuación (26) es una ecuación cuadrática que puede proporcionar dos soluciones positivas Δ . En este caso, como declarado antes, la solución ceder es la norma más pequeña de c debe usarse. Cuando, sin embargo, ninguna tal solución positiva se obtiene, el c se configura a 0.

[0084] Así, en la segunda modalidad, uno calcula P usar un caso especial de una distribución de decorrelacionador para los dos canales indicados por la matriz P en la caja 1201. Para algunos casos, la solución no existe y uno simplemente cierra el decorrelacionador. Una ventaja de esta modalidad consiste en que nunca adiciona una señal sintética con la correlación positiva. Esto es beneficioso, ya que tal señal podría ser percibida como una fuente de fantasma localizada que es un artefacto que disminuye la calidad de audio de la señal de salida reproducida. En vista del hecho que las cuestiones de energía no se consideran en la derivación, uno podría conseguir un apareamiento erróneo en la señal de salida que significa que la señal de salida tiene más o menos la energía que la señal de conversión de multicanal a estéreo. En este caso, uno podría poner en práctica una compensación de ganancia adicional en una modalidad preferida a fin de potenciar aún más la calidad de audio.

[0085] En una tercera modalidad de la presente invención la operación de la calculadora de la matriz 202 es diseñada como sigue. El punto de partida es una mezcla seca de ganancia compensada;

$$\hat{\mathbf{Y}} = \begin{bmatrix} g_1 & 0 \\ 0 & g_2 \end{bmatrix} \mathbf{Y}_0, \quad (27)$$

donde, por ejemplo, la mezcla en seco no compensada \mathbf{Y}_0 es el resultado de la aproximación de mínimos cuadrados

$\hat{\mathbf{Y}}_0 = \mathbf{C}_0 \mathbf{X}$ con la matriz de mezclado proporcionada por (15). Más aún, $\mathbf{C} = \mathbf{G} \mathbf{C}_0$, donde G es una matriz diagonal con entradas g_1 y g_2 . En este caso;

$$\hat{\mathbf{R}} = \begin{bmatrix} \hat{L} & \hat{p} \\ \hat{p} & \hat{R} \end{bmatrix} = \begin{bmatrix} g_1 & 0 \\ 0 & g_2 \end{bmatrix} \begin{bmatrix} \hat{L}_0 & \hat{p}_0 \\ \hat{p}_0 & \hat{R}_0 \end{bmatrix} \begin{bmatrix} g_1 & 0 \\ 0 & g_2 \end{bmatrix}, \quad (28)$$

$$= \begin{bmatrix} g_1^2 \hat{L}_0 & g_1 g_2 \hat{p}_0 \\ g_1 g_2 \hat{p}_0 & g_2^2 \hat{R}_0 \end{bmatrix}$$

y la matriz de error es;

$$\Delta \mathbf{R} = \begin{bmatrix} \Delta L & \Delta p \\ \Delta p & \Delta R \end{bmatrix} = \begin{bmatrix} L - g_1^2 \hat{L}_0 & p - g_1 g_2 \hat{p}_0 \\ p - g_1 g_2 \hat{p}_0 & R - g_2^2 \hat{R}_0 \end{bmatrix}, \quad (29)$$

[0086] Es mostrado luego por la tercera modalidad de la presente invención elegir las ganancias de compensación (g_1, g_2) para minimizar una suma ponderada de las potencias de error;

$$w_1 \Delta L + w_2 \Delta R = w_1 (L - g_1^2 \hat{L}_0) + w_2 (R - g_2^2 \hat{R}_0), \quad (30)$$

proporcionado por (13). Las opciones de ejemplo de pesos en (30) son $(w_1, w_2) = (1, 1)$ o

$$(w_1, w_2) = (R, L).$$

La matriz de error ΔR resultante es usado luego como introducido para el cálculo de la matriz P de mezcla del predecorrelacionador según las etapas de ecuaciones (18) - (21) una característica atractiva

5 de esta modalidad es que en casos donde señal de error $Y - \hat{Y}_0$ es similar a la conversión de estéreo a multicanal seca, la cantidad de la señal no correlacionada adicionada a la salida final es más pequeña que esto adicionado al final enviado por la primera modalidad de la presente invención.

10 **[0087]** En la tercera modalidad, que es resumida en relación con Figura 13, una matriz de ganancia adicional G es asumida como indicado en la Figura 4d. De acuerdo con lo que es escrito en la ecuación (29) (y 30), los factores de ganancia g_1 y g_2 se calcula usando w_1, w_2 seleccionado como se indican en el texto a continuación la ecuación (30) y basado en las coacciones en la matriz de error como indicado en la ecuación (13). Después de llevar a cabo estas dos etapas 1301, 1302, uno puede calcular una matriz de covariancia de señal de error ΔR usando g_1, g_2 como indicado en la etapa 1303. Esto se observa que esta matriz de covariancia de señal de error calculada en la etapa

15 1303, es diferente de R de matriz de covariancia como calculado en etapas 1101, en Figura 11 y Figura 12. Luego, las mismas etapas 1102, 1103, 1104 se llevan a cabo como han sido mencionados ya en relación con la primera modalidad de Figura 11.

20 **[0088]** La tercera modalidad es ventajosa en aquella la mezcla seca no es sólo la onda igualada por la forma, pero, además, de ganancia compensada. Esto ayuda a reducir adicionalmente la cantidad de la señal no correlacionada de modo que cualquier artefacto incurrido agregando la señal no correlacionada se reduzca también. Así, la tercera modalidad intenta conseguir el mejor posible de una combinación de compensación de ganancia y adición de decorrelacionador. Nuevamente, el objetivo es reproducir completamente la estructura de covariancia que incluye energías de canal y usar tan poco como sea posible de la señal sintética tal como a una reducción al mínimo de la

25 ecuación (30).

[0089] Posteriormente, mencionan de una cuarta modalidad. En la etapa 1401, el decorrelacionador individual se pone en práctica. Así, una baja modalidad de complejidad se forma ya que un decorrelacionador individual es, para una realización práctica, el más ventajosa. En la etapa subsecuente 1101, R de datos de matriz de covariancia se calcula como detallado y mencionado en relación con la etapa 1101 de la primera modalidad. Alternativamente, sin embargo, R de datos de matriz de covariancia también puede calcularse como indicado en la etapa 1303 de la

30 Figura 13, donde hay compensación de ganancia además de la correspondencia de forma de onda. Posteriormente, la señal Δp que es el elemento no diagonal de la matriz de covariancia ΔR es comprobada. Cuando la etapa 1402 determina que este signo es negativo, luego etapas 1102, 1103, 1104 de la primera modalidad se procesa, donde la

35 etapa 1103 es particularmente no complejo debido a que r_z es un valor escalar, ya que hay sólo un decorrelacionador individual.

[0090] Cuando, sin embargo, esto se determina que la señal Δp es positiva, una adición de la señal no correlacionada es completamente eliminada tal como a un fraguado al cero, los elementos de matriz P . Alternativamente, la adición de una señal no correlacionada puede reducirse a un valor el cero superior, pero a un valor más pequeño que un valor que sería allí debería el signo ser negativo. Preferentemente, sin embargo, los

40 elementos de la matriz de la matriz P no sólo son juego a valores más pequeños pero se configuran al cero como indicado en el bloque 1404, en la Figura 14. De acuerdo con la Figura 4d, sin embargo, los factores de ganancia g_1, g_2 se determinan a fin de llevar a cabo una compensación de ganancia como indicado en el bloque 1406. Específicamente, los factores de ganancia se calculan tal que los elementos diagonales principales de la matriz a la

45 derecha de ecuación (29) se conviertan el cero. Esto significa que la matriz de covariancia de la señal de error tiene elementos cero a su diagonal principal. Así, una compensación de ganancia se logra en el caso, cuando la señal de decorrelacionador se reduce o completamente apagado debido a la estrategia para evitar artefactos de la fuente de fantasma que podrían ocurrir cuando una señal no correlacionada que tiene propiedades de correlación específicas se agrega.

50

[0091] Así, la cuarta modalidad combina algunas características de la primera modalidad y se basa en una solución de decorrelacionador individual, pero incluye una prueba de determinar la calidad de la señal no correlacionada de modo que la señal no correlacionada pueda reducirse o completamente eliminado, cuando un indicador de calidad, como el valor Δp en la matriz de covariancia ΔR de la señal de error (señal adicionada) se convierte positivo.

55

[0092] La opción de la matriz de predecorrelacionador \mathbf{Q} debería basarse en consideraciones perceptuales, ya que la segunda teoría de orden encima de es insensible a la matriz específica usada. Esto también implica que las consideraciones que dan como resultado a una opción de son independientes de la selección entre cada una de las modalidades ya mencionadas.

5 [0093] Una primera solución preferida mostrada por la presente invención comprende usar la conversión de multicanal a estéreo mono de la mezcla estéreo seca como introducido a todos los decorrelacionadores. En términos de elementos de la matriz esto significa que;

$$q_{n,k} = c_{1,k} + c_{2,k}, \quad k = 1, 2; \quad n = 1, 2, \dots, N_d, \quad (31)$$

10 donde $\{q_{n,k}\}$ son los elementos de matriz de \mathbf{Q} y $\{c_{n,k}\}$ son los elementos de la matriz de \mathbf{C}_0 .

[0094] Una segunda solución mostrada por la presente invención da como resultado a una matriz de predecorrelacionador \mathbf{Q} derivada de la matriz de conversión de multicanal a estéreo por sí solo. La derivación se basa asumiendo que todos los objetos tienen la energía de unidad y son no correlacionados. Una matriz de conversión de estéreo a multicanal de los objetos a sus errores de predicción individuales se forma proporcionada aquella suposición. Luego el cuadrado de los pesos de predecorrelacionador se selecciona en la proporción a la energía de error de objeto pronosticada total a través de canales de conversión de multicanal a estéreo. Los mismos pesos son finalmente usados para todos los decorrelacionadores. Detalladamente, estos pesos se obtienen por la primera formación de la matriz $N \times N$,

$$\mathbf{W} = \mathbf{I} - \mathbf{D}'(\mathbf{D}\mathbf{D}')^{-1}\mathbf{D}, \quad (32)$$

20 y luego derivar una matriz de energía de error de predicción \mathbf{W}_0 de objeto estimada definido a un fraguado de todos los valores no diagonales (de 32) a cero. A una denotación de los valores diagonales de $\mathbf{P}\mathbf{W}_0\mathbf{D}'$ por t_1, t_2 , que representan las contribuciones de energía de error de objeto totales a cada canal de conversión de multicanal a estéreo, la opción final de elementos de matriz de predecorrelacionador se les proporciona por;

$$q_{n,k} = \sqrt{\frac{t_k}{t_1 + t_2}}, \quad k = 1, 2; \quad n = 1, 2, \dots, N_d, \quad (33)$$

25 [0095] En cuanto a una realización específica de los decorrelacionadores, todos los decorrelacionadores, como reverberadores o cualquier otro decorrelacionador pueden usarse. En una modalidad preferida, sin embargo, los decorrelacionadores deberían ser la conservación de la energía. Esto significa que la energía de la señal de salida de decorrelacionador debería ser igual que la energía de la señal de entrada de decorrelacionador. Sin embargo, desviaciones incurridas por un decorrelacionador de no conservación de energía" también puede ser absorbida, por ejemplo a una toma de esto en cuenta cuando la matriz \mathbf{P} se calcula.

30 [0096] Como declarado antes, las modalidades preferidas tratan de evitar adicionar una señal sintética con la correlación positiva, ya que tal señal podría ser percibida como una fuente de fantasma sintética localizada. En la segunda modalidad, esto es explícitamente evitado debido a la estructura específica de la matriz \mathbf{P} como indicado en el bloque 1201. Más aún, este problema es explícitamente burlado en la cuarta modalidad debido a la operación que comprueba en la etapa 1402. Otros modos de determinar la calidad de la señal no correlacionada y, específicamente, las características de correlación de modo que tales artefactos de la fuente de fantasma puedan ser evitados están disponibles para aquellos expertos en la técnica y pueden usarse para apagar la adición de la señal no correlacionada como en la forma de algunas modalidades o pueden usarse para reducir la energía de la señal no correlacionada y aumentar la energía de la señal seca, a fin de tener una señal de salida de ganancia compensada.

35 [0097] Aunque todas las matrices $\mathbf{E}, \mathbf{D}, \mathbf{A}$ se hayan descrito como matrices complejas, estas matrices también pueden ser valoradas del modo real. Sin embargo, la presente invención también es útil en relación con se forman en complejos matrices $\mathbf{D}, \mathbf{A}, \mathbf{E}$ de hecho teniendo coeficientes complejos con una parte imaginaria diferente del cero.

40 [0098] Más aún, a menudo será el caso que la matriz \mathbf{D} y la matriz \mathbf{A} tienen mucho inferior espectral y resolución de tiempo comparado con la matriz el \mathbf{E} que tiene la resolución de tiempo y frecuencia más elevada de todas las

matrices. Específicamente, la matriz de renderización de objetivo y la matriz de conversión de multicanal a estéreo no dependerán de la frecuencia, pero pueden depender a tiempo. Con respecto a la matriz de conversión de multicanal a estéreo, esto podría ocurrir en una operación de conversión de multicanal a estéreo optimizada específica. En cuanto a la matriz de renderización de objetivo, esto podría ser el caso en relación con mueven objetos de audio que pueden cambiar su posición entre derecho e izquierdo de vez en cuando.

[0099] Las modalidades descritas a continuación son simplemente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y las variaciones de las configuraciones y los detalles descritos aquí serán evidentes a otros expertos en la técnica. Es la intención, por lo tanto, sólo para ser limitado por el alcance de las reivindicaciones de patente inminentes y no por los detalles específicos presentados por vía de descripción y explicación de las modalidades aquí.

[0100] Según ciertos requerimientos de realización de los procedimientos inventivos, los procedimientos inventivos pueden ponerse en práctica en el hardware o en el software. La realización puede llevarse a cabo usando un medio de almacenamiento digital, en particular, un disco, un DVD o un CD que tienen señales de control legibles por medios electrónicos almacenadas sobre eso, que cooperan con sistemas de computadora programables tal que los procedimientos inventivos se lleven a cabo. En términos generales, la presente invención es por lo tanto un producto de programa de computadora con un código de programación almacenado en una portadora de lectura por máquina, el código de programación que se hace funcionar para llevar a cabo los procedimientos inventivos cuando el producto de programa de computadora corre en una computadora. En otras palabras, los procedimientos inventivos son, por lo tanto, un programa de computadora que tiene un código de programación para llevar a cabo al menos uno de los procedimientos inventivos cuando el programa de computadora corre en una computadora.

[0101] Un ejemplo de la invención comprende un aparato para sintetizar una señal de salida (350) que presenta una primera señal de canal audio y una segunda señal de canal audio, el combinador (364) es operativo para calcular los factores de ponderación para la combinación ponderada de modo que un resultado 452 de una operación de mezclado de la primera señal de mezcla descendente de objeto audio y la segunda señal de mezcla descendente de objeto audio se hacen coincidir en forma de onda con un resultado de representación objetivo.

[0102] Según otro ejemplo del aparato para sintetizar una señal de salida (350) que presenta una primera señal de canal audio y una segunda señal de canal audio, el combinador (364) es operativo para calcular una matriz de mezcla C_0 destinada a mezclar la primera señal de mezcla descendente de objeto audio y la segunda señal de mezcla descendente de objeto audio a partir de la ecuación siguiente:

$$C_0 = A E D^* (D E D^*)^{-1},$$

donde C_0 es la matriz de mezcla, donde A es una matriz de renderización de objetivo que representa las informaciones de renderización de objetivo (360), donde D es una matriz de mezcla descendente que representa las informaciones de mezcla descendente (354), donde $*$ representa una operación de transposición conjugada compleja, y donde E es una matriz de covarianza de objeto que representa las informaciones de objeto audio paramétricas (362).

[0103] Según otro ejemplo del aparato para sintetizar una señal de salida (350) que presenta una primera señal de canal audio y una segunda señal de canal audio, el combinador (364) es operativo para calcular los factores de ponderación a partir de la siguiente ecuación:

$$R = A E A^*,$$

Donde R es una matriz de covarianza de la señal de salida representada 350 obtenida aplicando la información de representación objetivo a los objetos de audio, donde A es la matriz de renderización de objetivo que representa las informaciones de renderización de objetivo (360) y donde E es una matriz de covarianza de objeto que representa las informaciones de objeto audio paramétricas (362).

[0104] Según otro ejemplo del aparato para sintetizar una señal de salida (350) que presenta una primera señal de canal audio y una segunda señal de canal audio, el combinador (364) es operativo para calcular los factores de ponderación a partir de la siguiente ecuación:

$$R_0 = C_0 D E D^* C_0^*,$$

Donde R_0 es la matriz de covarianza del resultado de la operación de mezclado 401 de la señal de mezcla descendente.

5 [0105] Según otro ejemplo del aparato para sintetizar una señal de salida (350) que presenta una primera señal de canal audio y una segunda señal de canal audio, la operación de pre-desacorrelacionador incluye una operación de me mezclado para mezclar el primer canal de mezcla descendente de objeto de audio y el segundo canal de mezcla descendente de objeto de audio a partir de la información de mezcla descendente 354 que indica una distribución del objeto de audio en la señal de mezcla descendente.

10 [0106] Según otro ejemplo del aparato para sintetizar una señal de salida (350) que presenta una primera señal de canal audio y una segunda señal de canal audio, el combinador (364) es operativo para realizar una operación de mezclado seco 401 de las señales de mezcla descendente de objeto de audio primera y segunda, en el que la operación de pre-desacorrelacionador402 es similar a la operación de mezclado seco 401.

15 [0107] Según otro ejemplo del aparato para sintetizar una señal de salida 350, el combinador 364 es operativo para utilizar la matriz de mezclado seco C_0 en el que la manipulación de pre-desacorrelacionador 402 se implementa utilizando una matriz de pre-desacorrelacionador Q que es idéntica a la matriz de mezclado seco C_0 .

20 [0108] Según otro ejemplo del aparato para sintetizar una señal de salida 350 que tiene una primera señal de canal audio y una segunda señal de canal audio, el combinador 364 es operativo para calcular los factores de ponderación a partir de la multiplicación 1104 de una matriz (T) derivada de valores propios obtenidos mediante la descomposición de valores propios 1102 y una matriz de covarianza de la señal de desacorrelacionador 358.

25 [0109] Según otro ejemplo del aparato para sintetizar una señal de salida 350 que tiene una primera señal de canal audio y una segunda señal de canal audio, el combinador 364 es operativo para calcular los factores de ponderación de modo que se utiliza un único desacorrelacionador 403 y la matriz de post-procesamiento de desacorrelacionador P es una matriz que tiene una única columna y un número de líneas igual a el número de señales de canal en la señal de salida representada, o en el se emplean que dos desacorrelacionadores 403, y la matriz de post-procesamiento de desacorrelacionador P tiene dos columnas y un número de líneas igual a el número de señales de canal de la señal de salida representada.

30 [0110] Según otro ejemplo del aparato para sintetizar una señal de salida 350 que tiene una primera señal de canal audio y una segunda señal de canal audio, el combinador es operativo para calcular los factores de ponderación a partir de una matriz de covarianza de la señal desacorrelacionada, que se calcula a partir de la siguiente ecuación:

$$R_z = Q D E D^* Q^*,$$

35 donde R_z es la matriz de covarianza de la señal desacorrelacionada 358, Q es una matriz de mezclado de pre-desacorrelacionador, D es una matriz de mezcla descendente que representa la información de mezcla descendente 354, E es matriz de covarianza de objeto de audio que representa la información de objeto de audio paramétrica 362.

40 [0111] Según otro ejemplo del aparato para sintetizar una señal de salida 350 que tiene una primera señal de canal audio y una segunda señal de canal audio, se resuelve una ecuación cuadrática 26 para determinar el factor de ponderación (c) y en el que, si no hay solución real para esta ecuación cuadrática, se reduce o desactiva la suma un señal desacorrelacionada 1208.

45 [0112] Según otro ejemplo del aparato para sintetizar una señal de salida 350 que tiene una primera señal de canal audio y una segunda señal de canal audio, también comprende: un convertidor de tiempo/ frecuencia 302 para convertir la señal de mezcla descendente en una representación espectral que comprende a pluralidad de señales de mezcla descendente de sub-banda: en el que, para cada señal de sub-banda, se utiliza una operación de desacorrelacionador 403 y una operación de combinador 364 de modo que se genera la pluralidad señal de sub-bandas de representación de salida, y un convertidor de frecuencia / tiempo 304 para convertir la pluralidad de señal de sub-bandas de la señal de salida representada en una representación en el dominio del tiempo.

55 [0113] Según otro ejemplo del aparato para sintetizar una señal de salida 350 que tiene una primera señal de canal audio y una segunda señal de canal audio, también comprende un controlador de procesamiento de bloques para generar bloques de valores de muestreo de la señal de mezcla descendente y para controlar el desacorrelacionador 356 y el combinador 364 para procesar bloques individuales de valores de muestra.

60 [0114] Según otro ejemplo del aparato para sintetizar una señal de salida 350 que tiene una primera señal de canal audio y una segunda señal de canal audio, se proporciona la información de objeto de audio para cada bloque y

para cada señal de sub-banda, y la información de representación objetivo y la información de mezcla descendente de objeto de audio son constantes en frecuencia para un bloque de tiempo.

REIVINDICACIONES

1. Aparato para sintetizar una señal de salida (350) que presenta una primera señal de canal audio y una segunda señal de canal audio, comprendiendo el aparato; una etapa de descorrelación (356) destinada a generar una señal descorrelacionada (358) que presenta una señal de canal única descorrelacionada o una señal de premier canal descorrelacionada y una señal de segundo canal descorrelacionada a partir de una señal de mezcla descendente, presentando la señal de mezcla descendente una primera señal de mezcla descendente de objeto audio y una segunda señal de mezcla descendente de objeto audio, representando la señal de mezcla descendente una mezcla descendente de una pluralidad de señales de objeto audio según las informaciones de mezcla descendente (354); y un combinador (364) destinado a realizar una combinación ponderada de la señal de mezcla descendente (352) y de la señal descorrelacionada (358) con ayuda de factores de ponderación, donde el combinador (364) es operativo para calcular los factores de ponderación para la combinación ponderada a partir de las informaciones de mezcla descendente (354), a partir de las informaciones de renderización de objetivo (360) que indican las posiciones virtuales de los objetos audio en una disposición de reproducción virtual, y de informaciones de objeto audio paramétricas (362) que describen los objetos audio, en el cual el combinador (364) es operativo para calcular una matriz de mezcla C_0 destinada a mezclar la primera señal de mezcla descendente de objeto audio y la segunda señal de mezcla descendente de objeto audio a partir de la ecuación siguiente:

$$C_0 = A E D^* (D E D^*)^{-1},$$

donde C_0 es la matriz de mezcla, donde A es una matriz de renderización de objetivo que representa las informaciones de renderización de objetivo (360), donde D es una matriz de mezcla descendente que representa las informaciones de mezcla descendente (354), donde $*$ representa una operación de transposición conjugada compleja, y donde E es una matriz de covarianza de objeto que representa las informaciones de objeto audio paramétricas (362).

2. Aparato según la reivindicación precedente, en el cual el combinador (364) es operativo para calcular los factores de ponderación para la combinación ponderada de manera que la combinación ponderada pueda ser obtenida calculando una matriz de mezcla ascendente de descorrelacionador (P) y aplicando (404) la matriz de mezcla ascendente de descorrelacionador (P) a la señal descorrelacionada (358), y combinando los resultados (454) de las operaciones de aplicación (404, 401), para obtener la señal de salida renderizada (550).

3. Aparato según cualquiera de las reivindicaciones anteriores, en el cual la etapa de descorrelacionador (356) es operativa para realizar una operación (402) para manipular la señal de mezcla descendente (352) donde la señal de mezcla descendente manipulada se alimenta hacia un descorrelacionador (403).

4. Aparato según la reivindicación 2, en el cual la matriz de mezcla ascendente de descorrelacionador (P) está basada en la realización (1102) de una descomposición de valor propio de una matriz de covarianza de la señal descorrelacionada añadida a un resultado de mezcla de señal en seco (452).

5. Aparato según la reivindicación 2, en el cual el combinador (364) es operativo para calcular los factores de ponderación para la combinación ponderada de manera que la matriz de mezcla ascendente de descorrelacionador (P) se calcule de manera que la señal descorrelacionada se añada a dos canales resultantes (452) de una operación de mezcla en seco con signos opuestos (1201).

6. Aparato según la reivindicación 5, en el cual el combinador (364) es operativo para calcular los factores de ponderación de manera que la señal descorrelacionada (358) sea ponderada por un factor de ponderación (c) determinado por una referencia de correlación entre dos canales de la señal de salida renderizada, siendo la referencia de correlación un valor de correlación determinado por una operación de renderización de objetivo virtual a una matriz de renderización de objetivo (A) (1203).

7. Aparato según la reivindicación 2, en el cual el combinador (364) es operativo para calcular los factores de ponderación de manera que la combinación ponderada efectuada por el combinador (364) resulte en una compensación de ganancia (409) en la cual un resultado de mezcla de señal en seco sea ponderado de manera que un error de energía en el resultado de mezcla de señal en seco comparado con la energía de la señal de mezcla descendente sea reducida (1302).

8. Aparato según cualquiera de las reivindicaciones 1 à 2, en el cual el combinador (364) es operativo para determinar si una adición de una señal descorrelacionada resultará en un artefacto (1402), y en el cual el combinador (364) es operativo para desactivar o reducir una adición de la señal descorrelacionada (1404) cuando se determina una situación de creación de artefacto, y

para reducir (1406) un error de potencia incurrida por la reducción o desactivación (1404) de la señal descorrelacionada.

9. Aparato según la reivindicación 8,

5 en el cual el combinador (364) es operativo para calcular los factores de ponderación de manera que se aumente la potencia de un resultado de la operación de mezcla en seco (401).

10. Aparato según la reivindicación 8, en el cual el combinador (364) es operativo para calcular un dato de matriz de covarianza de error (**R**) (1104) que representa una estructura de correlación de la señal de error entre la señal de mezcla ascendente en seco y una señal de salida determinada por un esquema de renderización de objetivo virtual con ayuda de las informaciones de renderización de objetivo (360), y

en el cual el combinador (364) es operativo para determinar un signo (1402) de un elemento fuera de la diagonal de los datos de matriz de covarianza de error (**R**) y para desactivar (1104) o reducir la adición si el signo es positivo.

15 11. Aparato según cualquiera de las reivindicaciones anteriores, en el cual el combinador (364) comprende una unidad de matrizado mejorada (303) operativa para combinar linealmente la primera señal de mezcla descendente de objeto audio y la segunda señal de mezcla descendente de objeto audio para obtener una señal de mezcla en seco (452), y en el cual el combinador (364) es operativo para combinar linealmente la señal descorrelacionada (358) para obtener una señal que constituye, tras la adición por canal con la señal de mezcla en seco, una salida etéreo de la unidad de matrizado mejorada (303), y

20 en el cual el combinador (364) comprende un calculador de matriz (202) destinado a calcular los factores de ponderación para la combinación lineal utilizada por la unidad de matrizado mejorada (303) a partir de las informaciones de objeto audio paramétricas (362) de las informaciones de mezcla hacia abajo (354) y de las informaciones de renderización de objetivo (360).

25 12. Aparato según cualquiera de las reivindicaciones anteriores, en el cual el combinador (364) es operativo para calcular los factores de ponderación de manera que una parte de energía de la señal descorrelacionada (358) en la señal de salida renderizada sea mínima y una parte de energía de la señal de mezcla en seco (452) obtenida combinando linealmente la primera señal de mezcla descendente de objeto audio y la segunda señal de mezcla descendente de objeto audio sea máxima.

30 13. Procedimiento para sintetizar una señal de salida (350) que presenta una primera señal de canal audio y una segunda señal de canal audio, que comprende:

35 generar (356) una señal descorrelacionada (358) que presenta una señal de canal única descorrelacionada o una primera señal de canal descorrelacionada y una segunda señal de canal descorrelacionada a partir de una señal de mezcla descendente, presentando la señal de mezcla descendente una primera señal de mezcla descendente de objeto audio y una segunda señal de mezcla descendente de objeto audio, representando la señal de mezcla descendente una mezcla descendente de una pluralidad de señales de objeto audio según las informaciones de mezcla descendente (354); y

40 realizar (364) una combinación ponderada de la señal de mezcla descendente (352) y de la señal descorrelacionada (358) con ayuda de factores de ponderación a partir de un cálculo de los factores de ponderación para la combinación ponderada a partir de las informaciones de mezcla descendente (354), a partir de las informaciones de renderización de objetivo (360) que indican partes virtuales de los objetos audio en una disposición de reproducción virtual, y de las informaciones de objeto audio paramétricas (362) que describen los objetos audio, en el cual el combinador (364) es operativo para calcular una matriz de mezcla C_0 destinada a mezclar la primera señal de mezcla descendente de objeto audio y la segunda señal de mezcla descendente de objeto audio a partir de la ecuación siguiente:

$$C_0 = A E D^* (D E D^*)^{-1},$$

50 donde C_0 es la matriz de mezcla, donde **A** es la matriz de renderización de objetivo que representa las informaciones de renderización de objetivo (360), donde **D** es una matriz de mezcla descendente que representa las informaciones de mezcla descendente (354), donde * representa una operación de transposición conjugada compleja, y donde **E** es una matriz de covarianza de objeto que representa las informaciones de objeto audio paramétricas (362).

55 14. Programa de ordenador que tiene un código de programa adaptado para realizar el procedimiento según la reivindicación 13 cuando se ejecuta en un procesador.

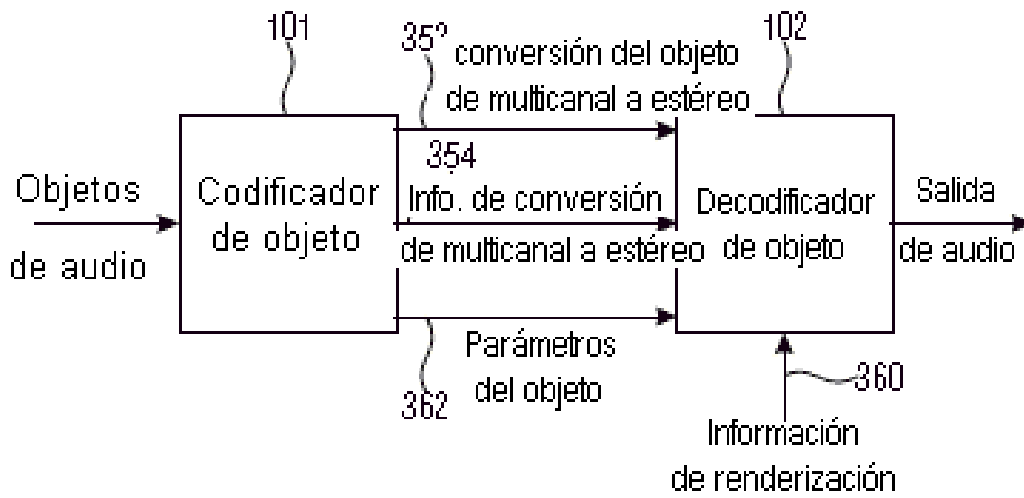


FIGURA 1

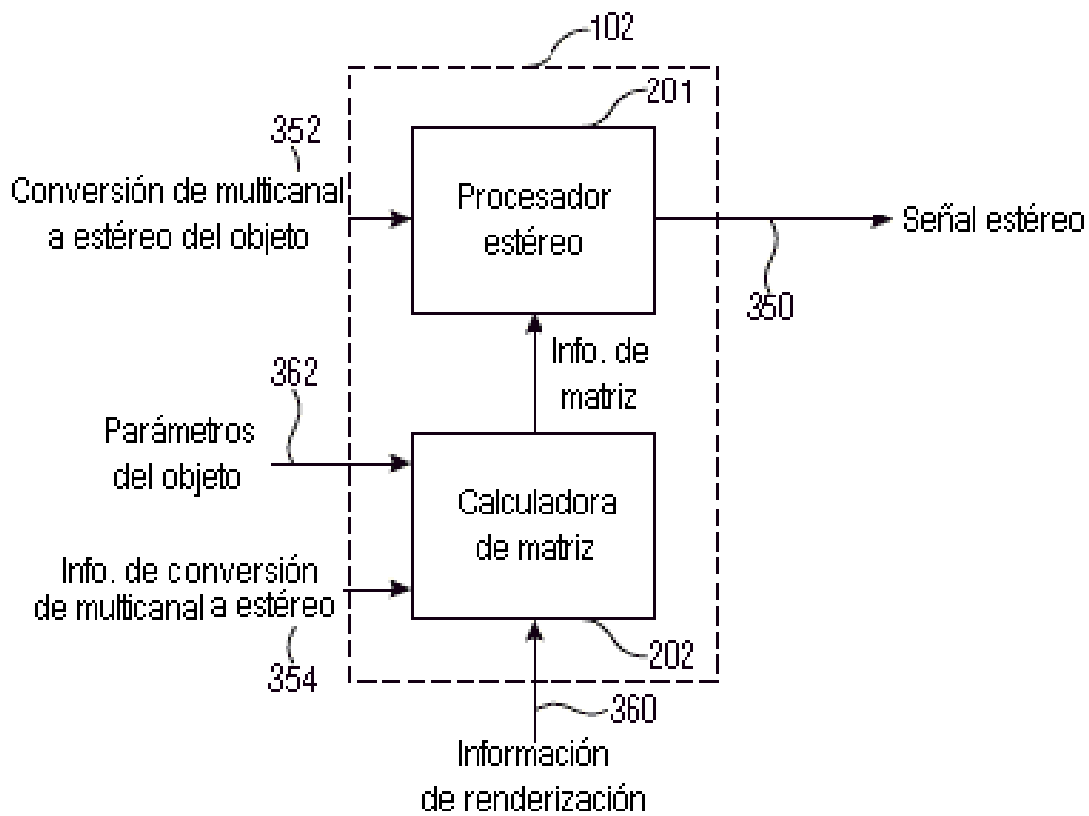


FIGURA 2A

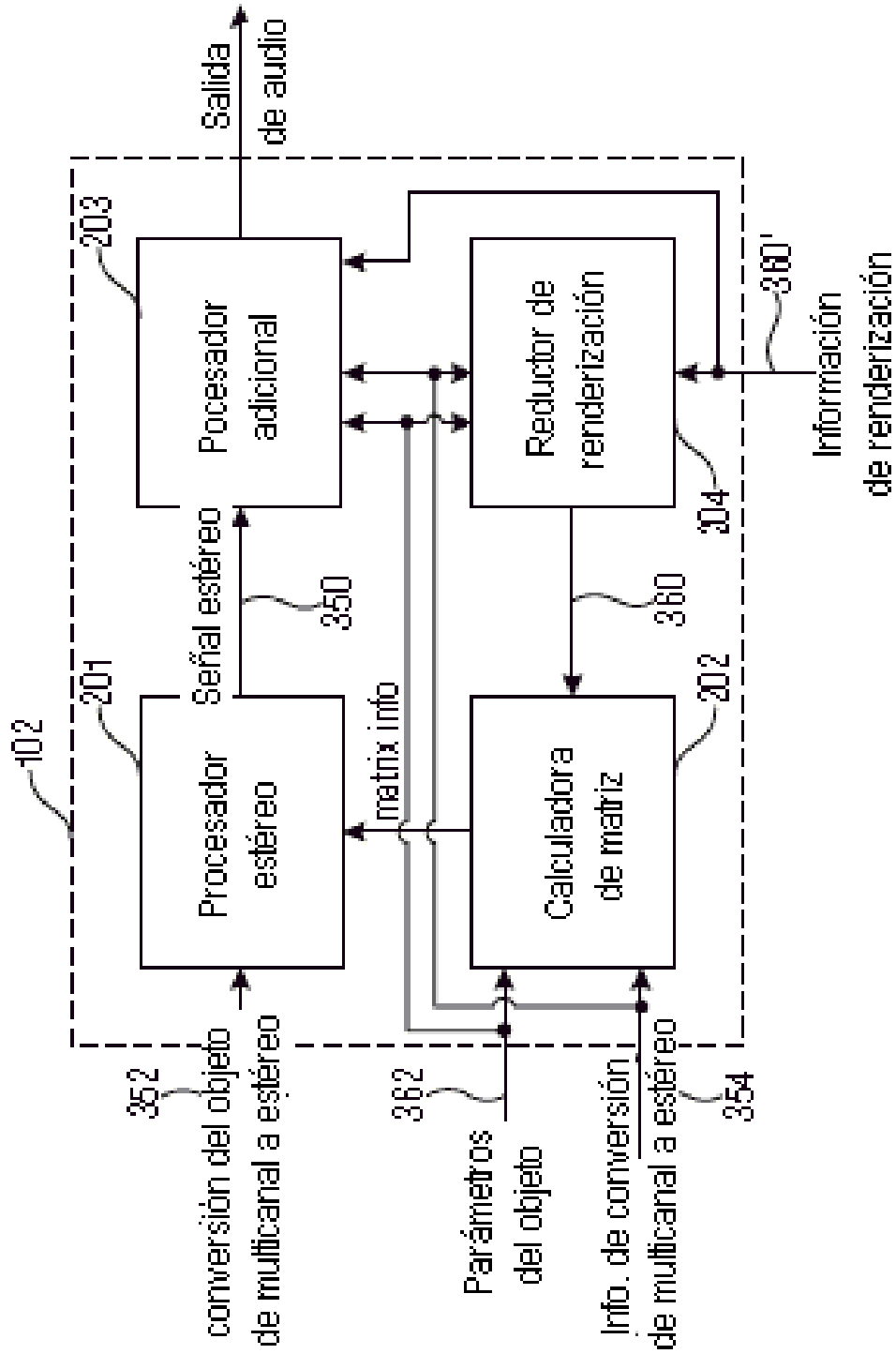


FIGURA 2B

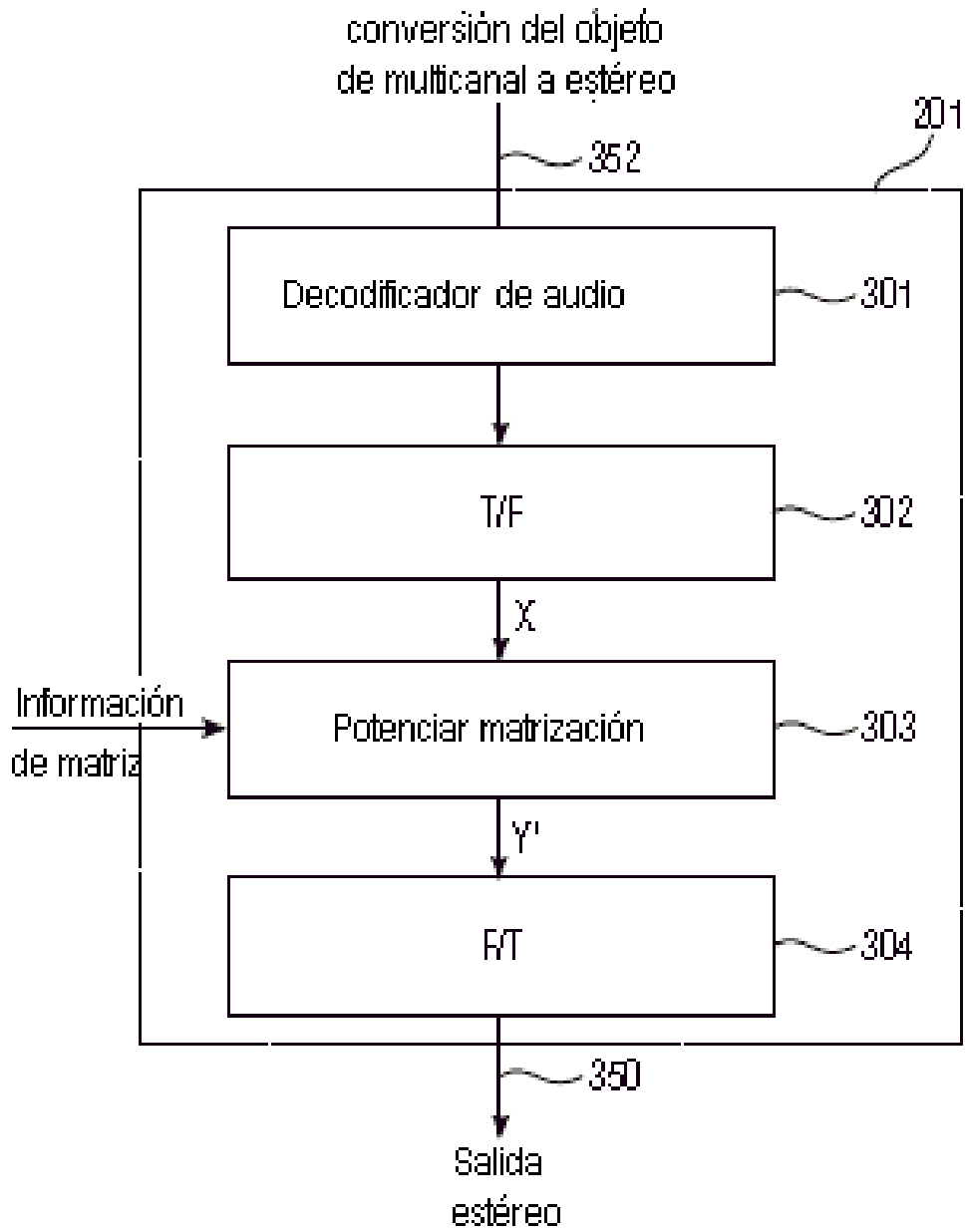


FIGURA 3A

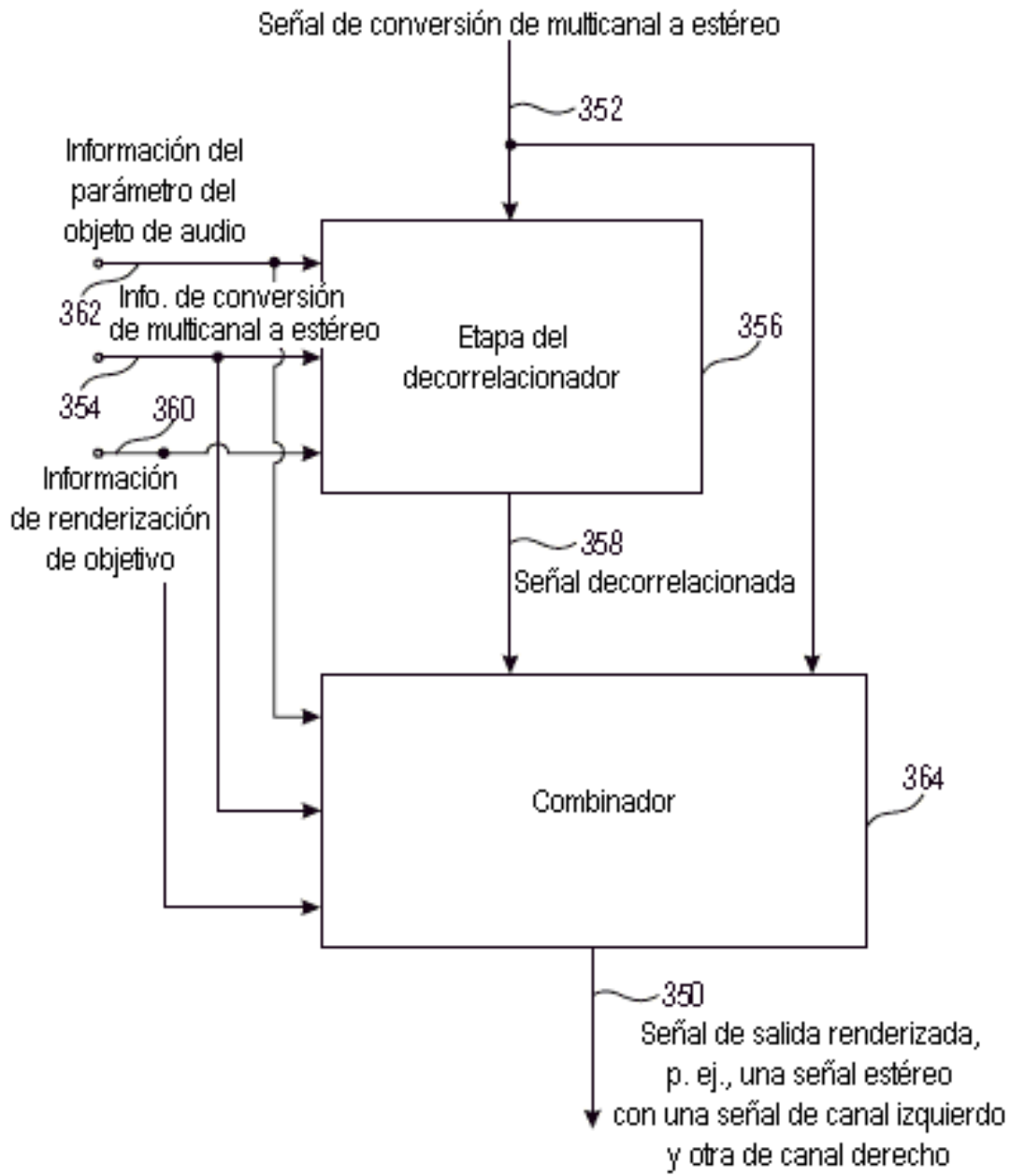


FIGURA 3B

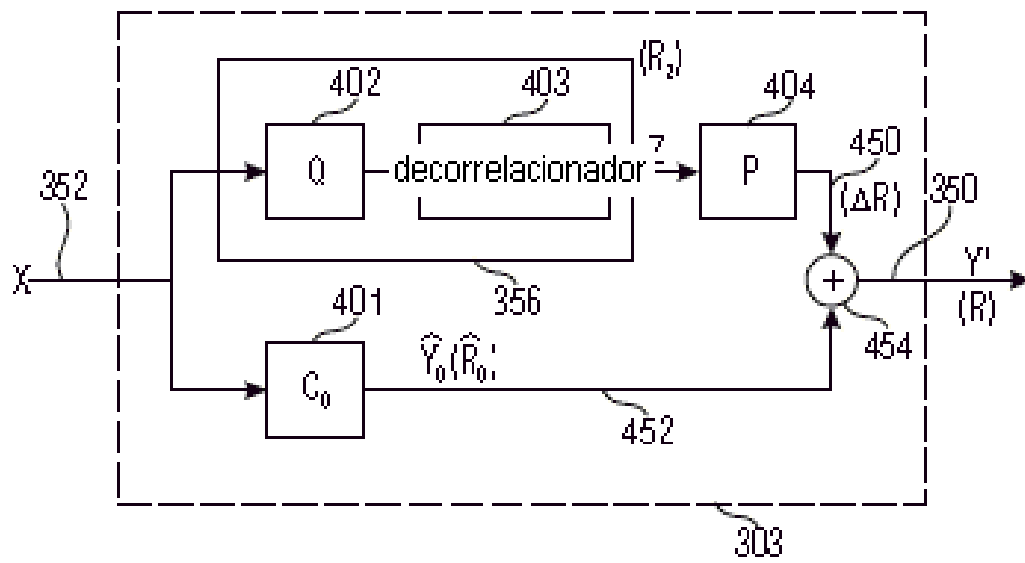


FIGURA 4A

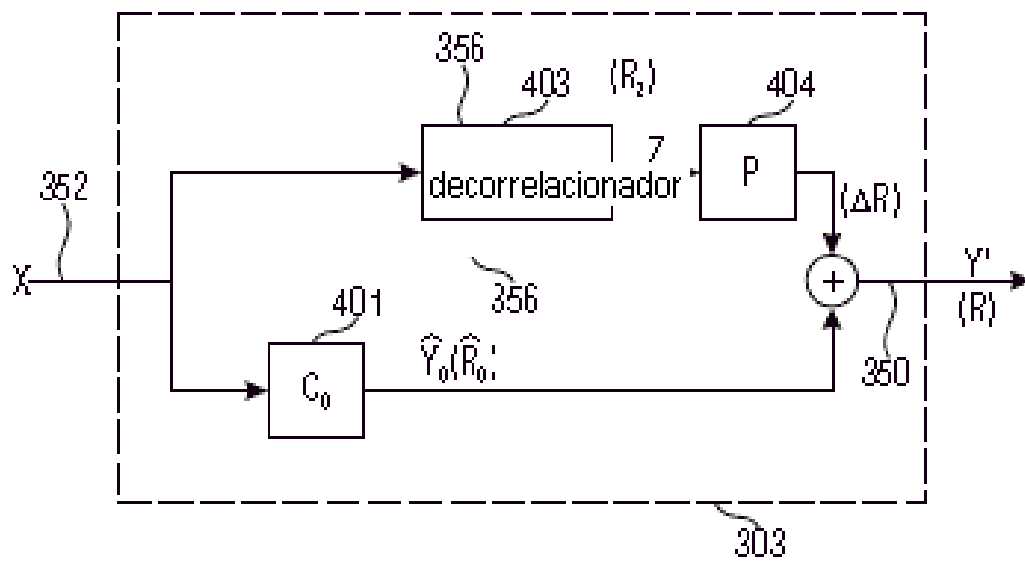
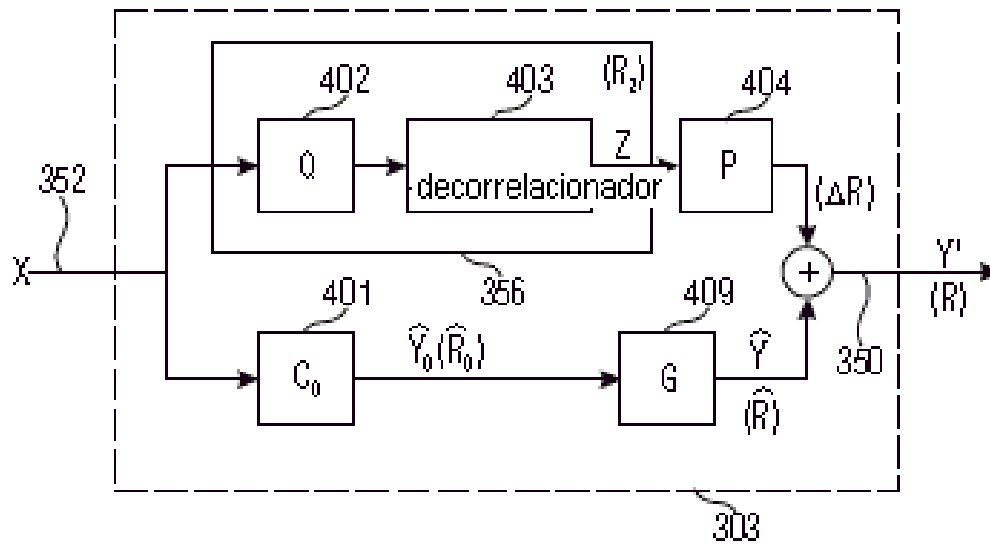
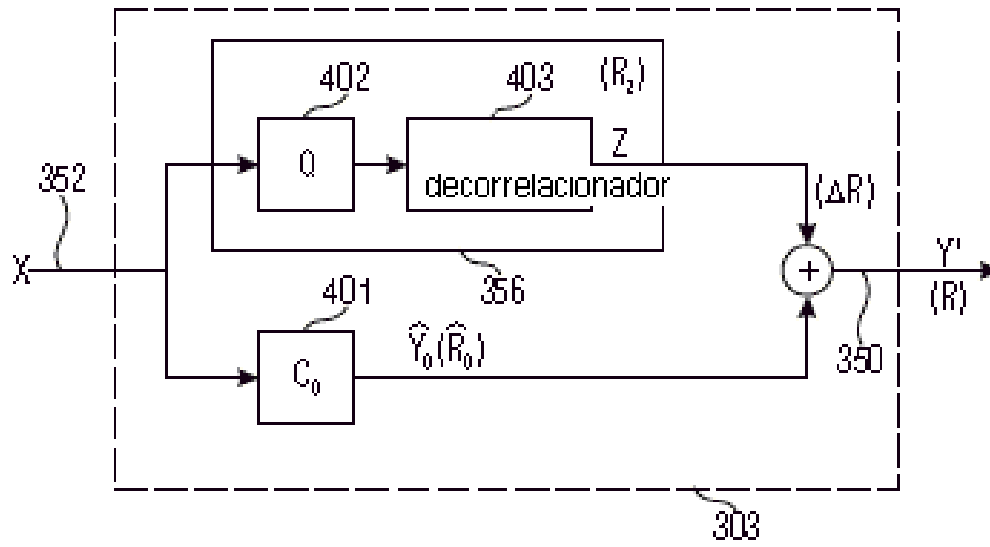


FIGURA 4B



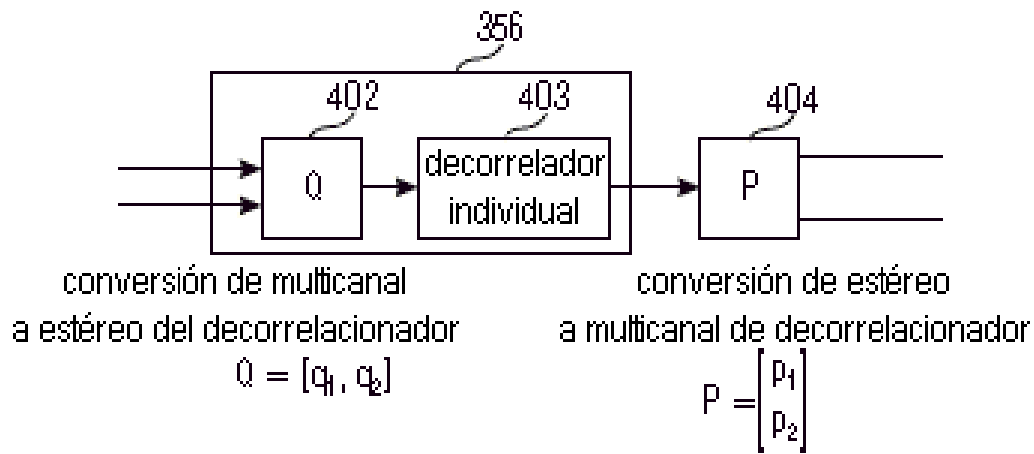


FIGURA 4E

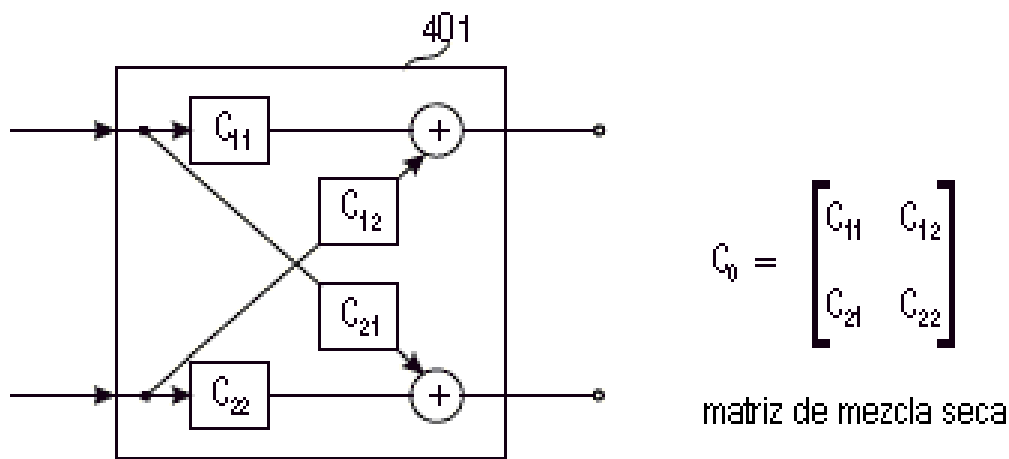


FIGURA 4F

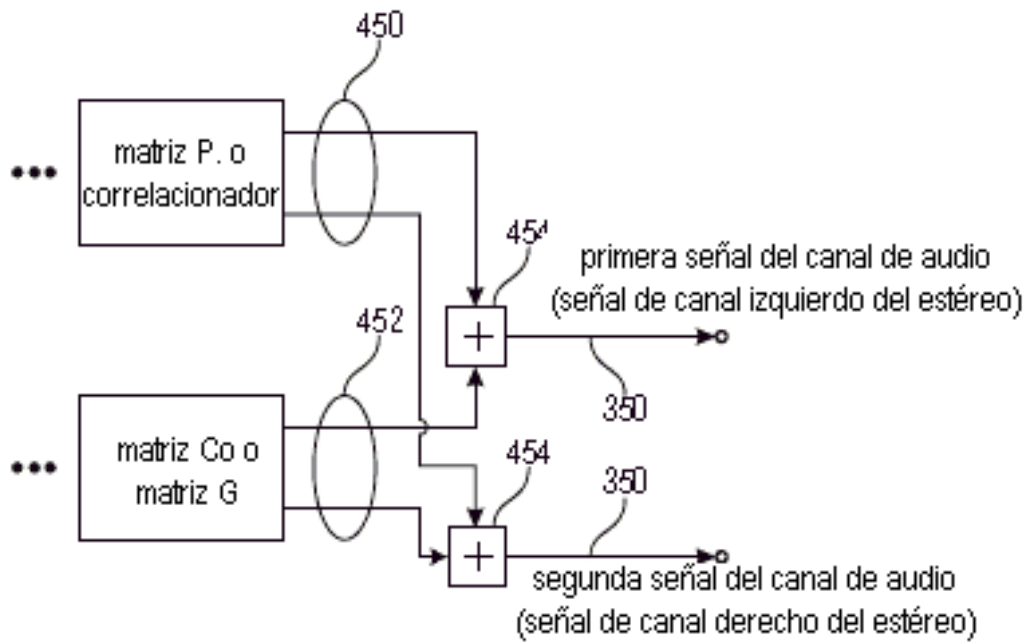


FIGURA 4G

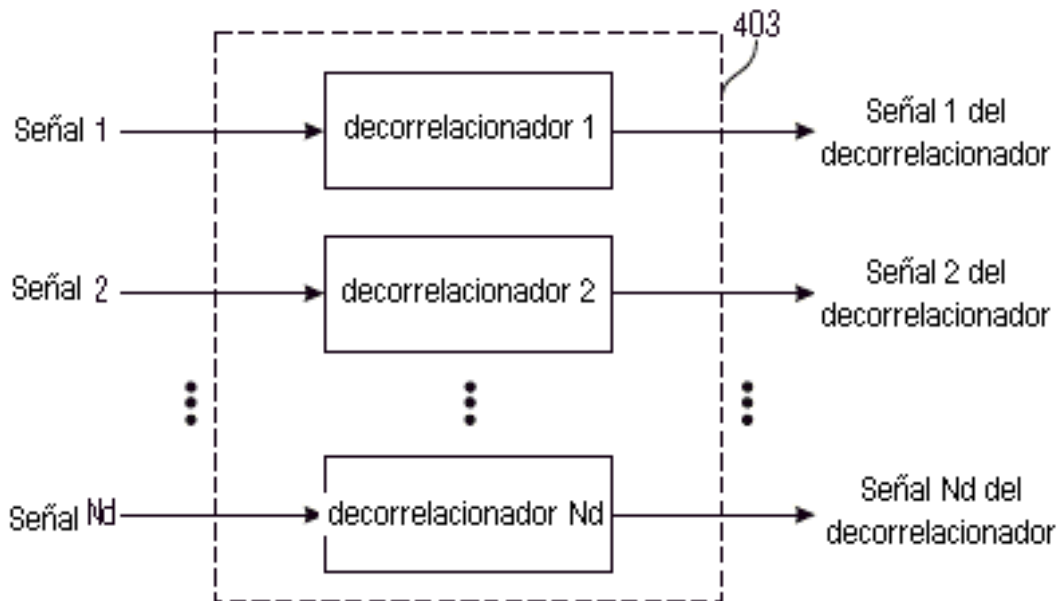


FIGURA 5

ID del objeto	archivo de audio del objeto	correlación y energía del parámetro del objeto de audio
1		matriz E para cada subbanda y bloque temporal
2		
3		
⋮	⋮	
N		

FIGURA 6

matriz de covarianza de objeto para una subbanda y bloque temporal

$$E = \begin{bmatrix} e_{11} & & & & & \\ e_{21} & e_{22} & & & & \\ e_{31} & e_{32} & e_{33} & & & \\ e_{41} & e_{42} & e_{43} & e_{44} & & \\ \vdots & & & & e_{55} & \\ \vdots & & & & & e_{66} \\ & & & & & \vdots \\ & & & & & & e_{nn} \end{bmatrix}$$

e_{ii} : energía de objeto de audio en la subbanda y bloque temporal correspondientes
 e_{ij} : medida de correlación entre objetos de audio i,j en el bloque temporal y subbanda correspondientes

FIGURA 7

**matriz de conversión
de multicanal a estéreo**

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & \dots & d_{1N} \\ d_{21} & d_{22} & d_{23} & d_{24} & \dots & d_{2N} \\ \vdots & & & & & \\ d_{K1} & & & & \dots & d_{KN} \end{bmatrix}$$

d_{ij} indica, si una porción de todo el objeto j se incluye o no en la señal i de conversión de multicanal a estéreo

por ejemplo: $d_{12} = 0 \Rightarrow$ objeto 2 NO se incluye en la señal 1 de conversión de multicanal a estéreo

$d_{23} = 1 \Rightarrow$ objeto 3 se incluye COMPLETAMENTE en la señal 1 de conversión de multicanal a estéreo

$d_{24} = d_{14} = 0.5 \Rightarrow$ objeto 4 está en ambas señales de conversión de multicanal a estéreo del objeto, pero con la mitad de energía en cada señal de conversión de multicanal a estéreo

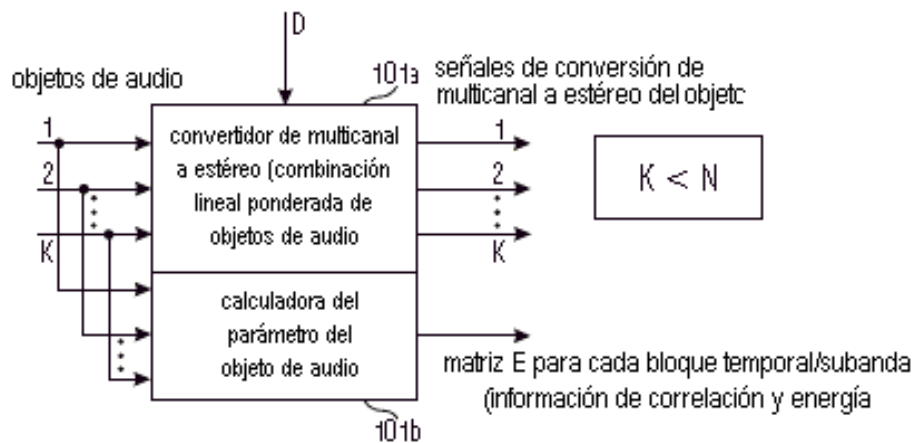


FIGURA 8

matriz A de renderización del objetivo (normalmente proporcionada por usuario)

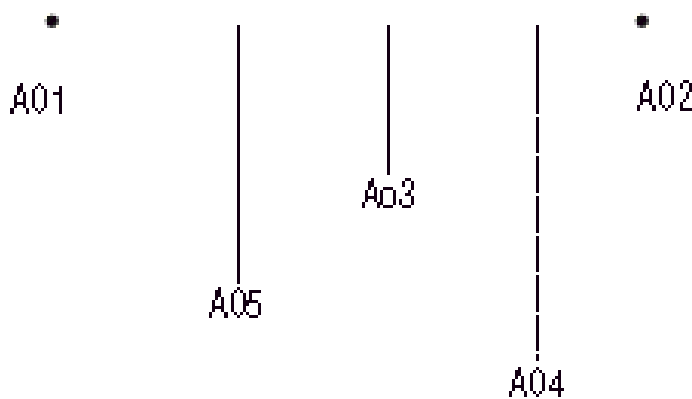
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \dots & a_{1N} \\ a_{21} & a_{22} & a_{23} & a_{24} & \dots & a_{2N} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{M1} & \dots & \dots & \dots & \dots & a_{MN} \end{bmatrix}$$

$M=2$ para renderización estéreo
 $M=M$ para renderización de canal M

a_{ij} indica, si todo el objeto o una porción se ha renderizar o no en el canal de salida

altavoz izquierdo (canal 1)

altavoz derecho (canal 2)



matriz A ejemplificante = $\begin{bmatrix} 1 & 0 & 0.5 & 0.25 & 0.75 & 0 \\ 0 & 1 & 0.5 & 0.75 & 0.25 & 0 \end{bmatrix}$

(el objeto 6 NO se va a renderizar en absoluto)

FIGURA 9

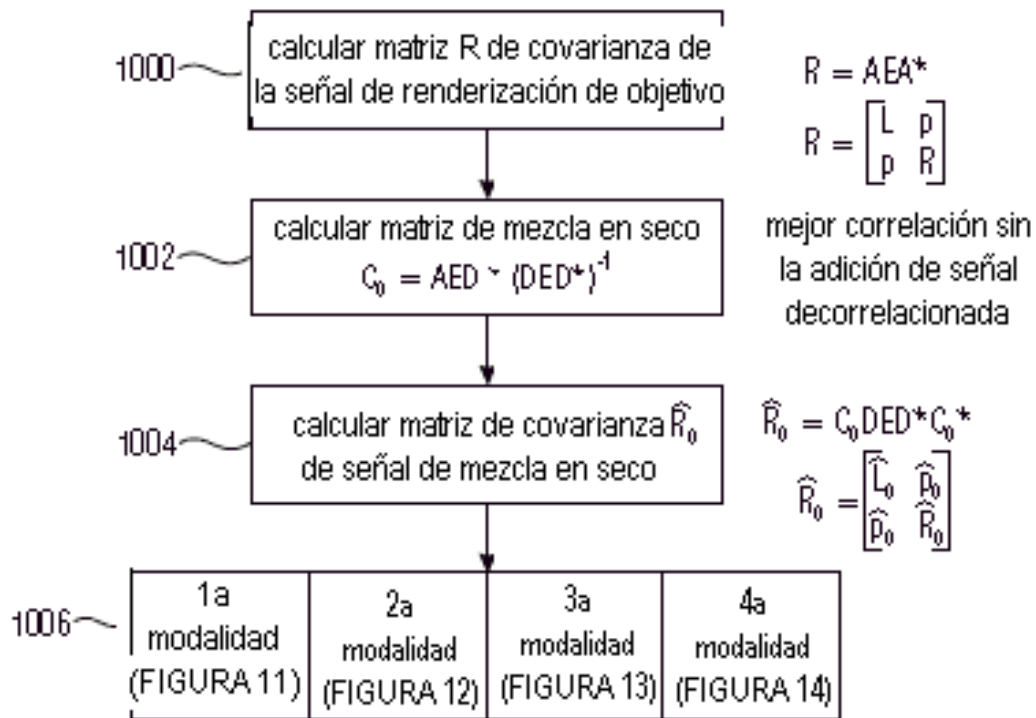


FIGURA 10

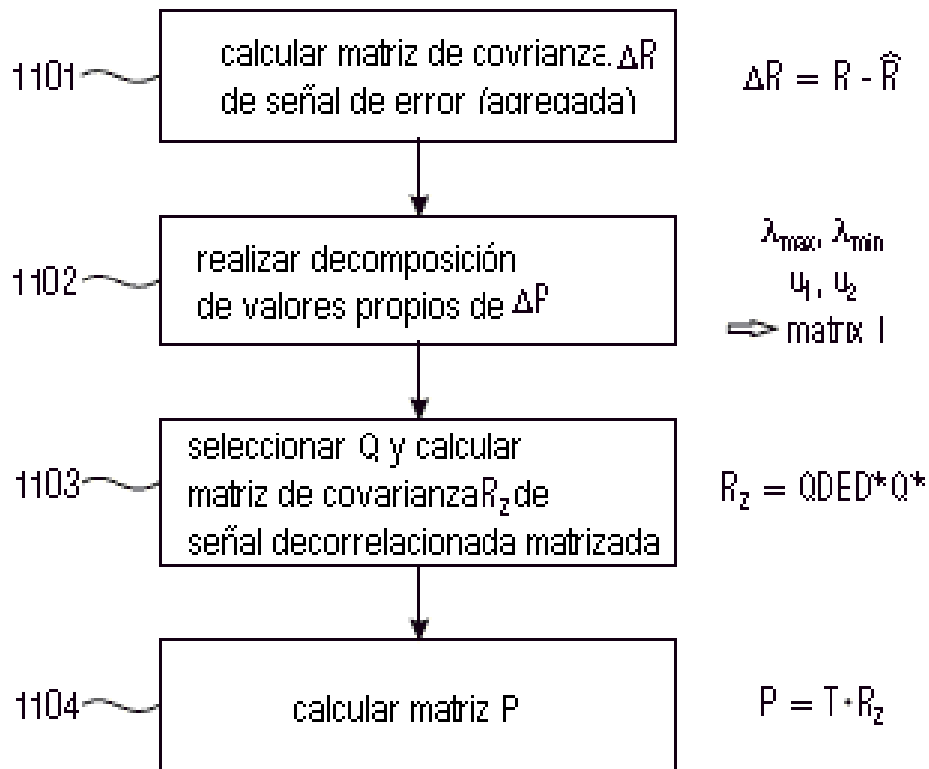


FIGURA 11

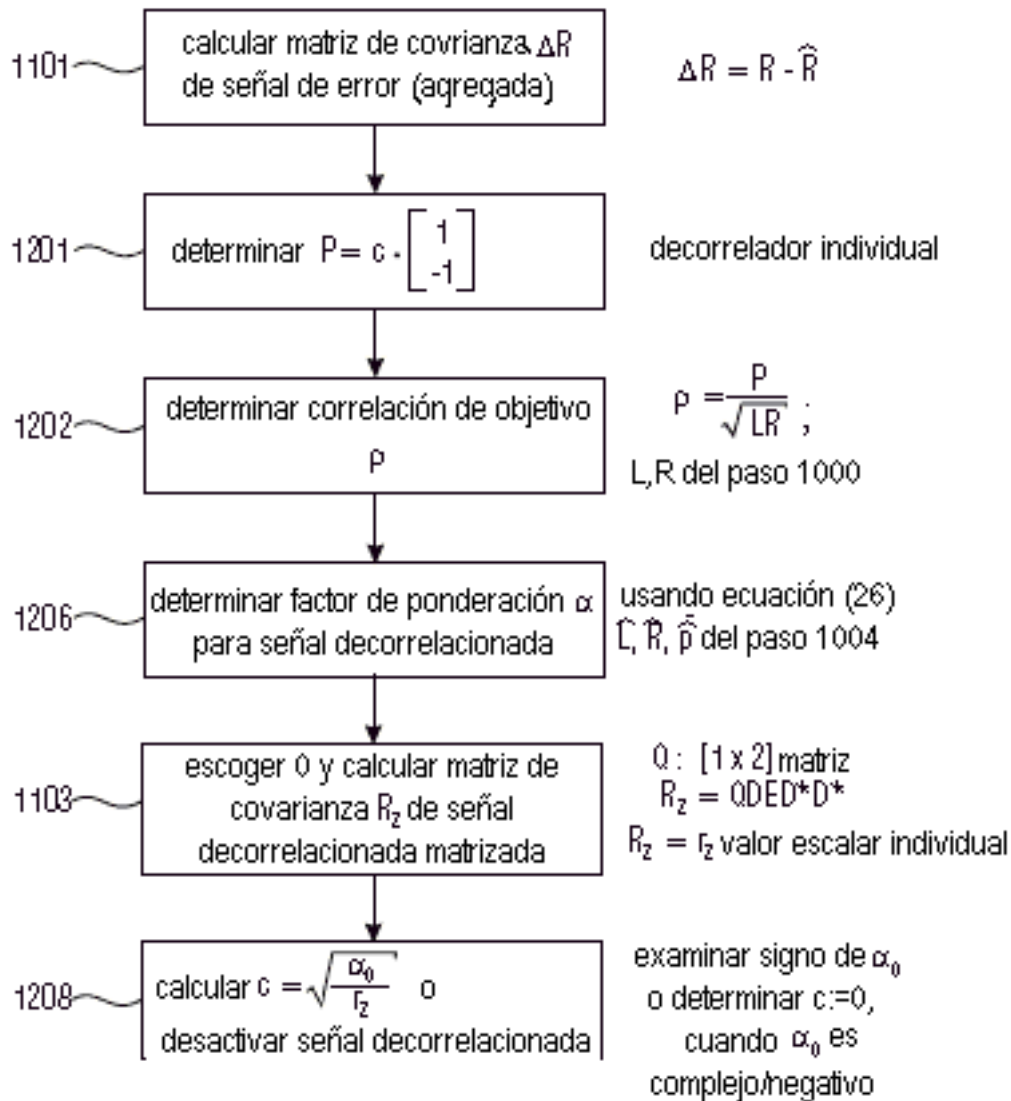


FIGURA 12

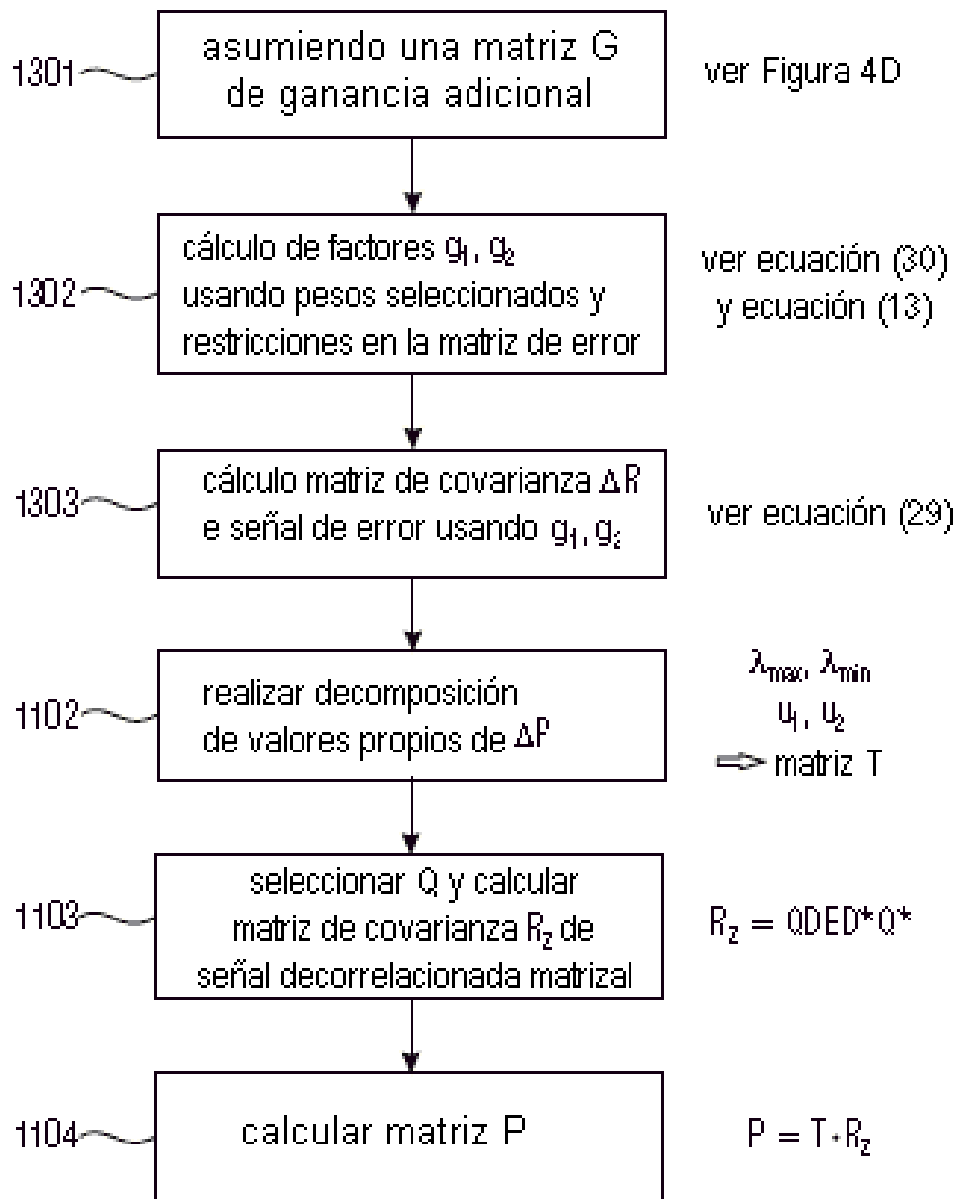


FIGURA 13

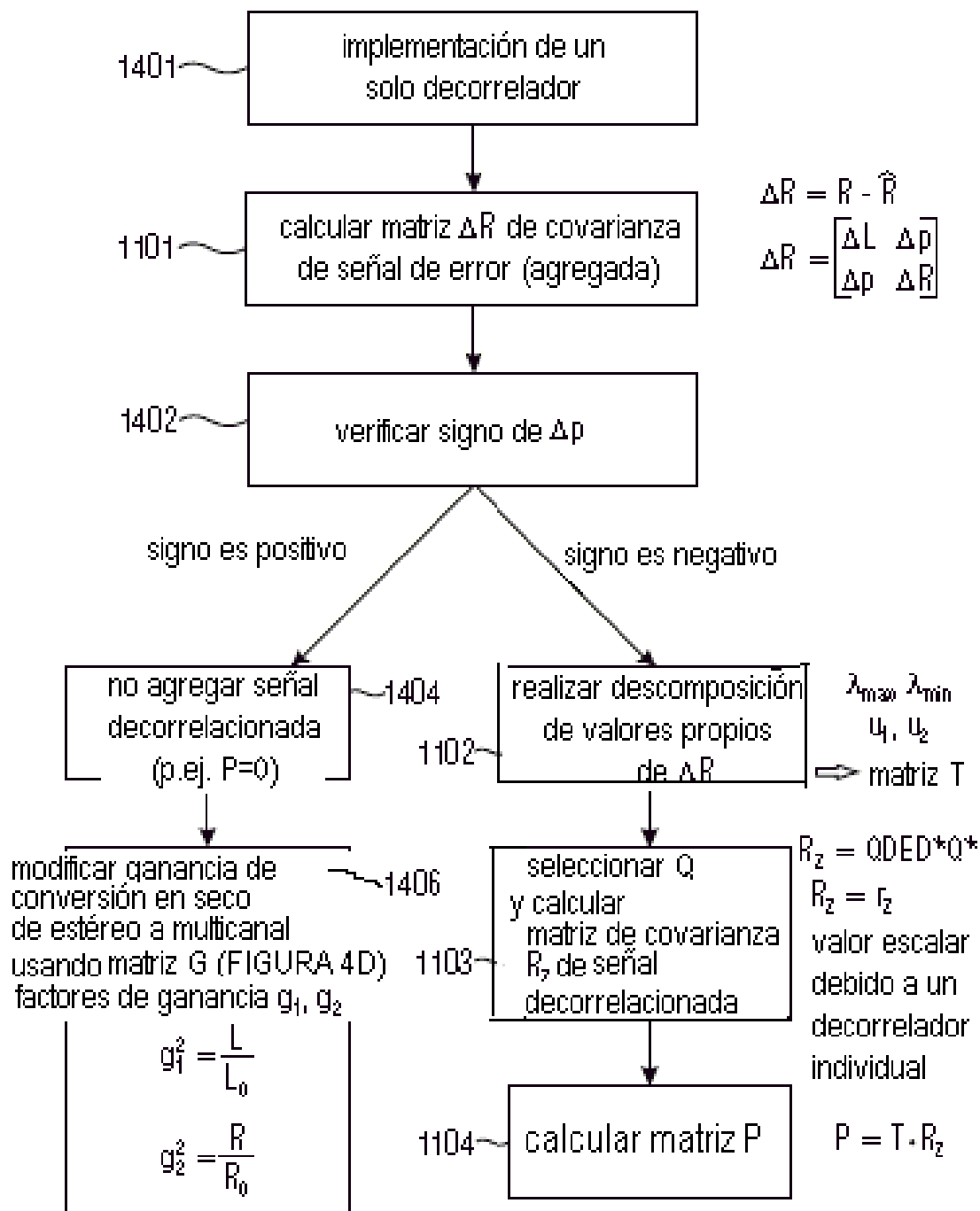


FIGURA 14