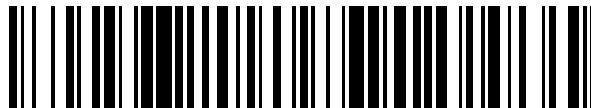


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 638 391**

51 Int. Cl.:

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **16.04.2013 PCT/EP2013/057932**

87 Fecha y número de publicación internacional: **13.02.2014 WO14023443**

96 Fecha de presentación y número de la solicitud europea: **16.04.2013 E 13716016 (4)**

97 Fecha y número de publicación de la concesión europea: **07.06.2017 EP 2883225**

54 Título: **Codificador, decodificador, sistema y procedimiento que emplea un concepto residual para una codificación paramétrica de un objeto de audio**

30 Prioridad:

10.08.2012 US 201261681730 P

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

20.10.2017

73 Titular/es:

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

72 Inventor/es:

**KASTNER, THORSTEN;
HERRE, JÜRGEN;
PAULUS, JOUNI;
TERENTIV, LEON;
HELLMUTH, OLIVER y
FUCHS, HARALD**

74 Agente/Representante:

SALVA FERRER, Joan

ES 2 638 391 T3

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

DESCRIPCIÓN

Codificador, decodificador, sistema y procedimiento que emplea un concepto residual para una codificación paramétrica de un objeto de audio

5

[0001] La presente invención se refiere a la codificación, decodificación y procesamiento de señal de audio, y, en particular, a un codificador, un decodificador y un procedimiento que emplean conceptos residuales para la codificación paramétrica de un objeto de audio.

10 **[0002]** Recientemente se han propuesto técnicas paramétricas para la transmisión eficiente de la velocidad binaria/almacenamiento de escenas de audio que comprenden múltiples objetos de audio en el campo de la codificación de audio (véase, por ejemplo, [BCC], [JSC], [SAOC], [SAOC1] y [SAOC2]) y la separación de fuentes informada (véase, por ejemplo, [ISS1], [ISS2], [ISS3], [ISS4], [ISS5] e [ISS6]). Estas técnicas tienen como objetivo la reconstrucción de una escena de audio de salida deseada o un objeto de fuente de audio deseado sobre la base de
15 informaciones secundarias adicionales que describen la escena de audio transmitida y/o almacenada y/o los objetos de fuente de audio en la escena de audio.

[0003] La Fig. 5 representa una visión general del sistema SAOC (SAOC = Codificación de objetos de audio espaciales) que ilustra el principio de tales sistemas paramétricos utilizando el ejemplo de MPEG SAOC (MPEG =
20 Grupo de Expertos en Imágenes en Movimiento) (véase, por ejemplo, [SAOC], [SAOC1] y [SAOC2]).

[0004] El procesamiento general se lleva a cabo de una manera selectiva en tiempo/frecuencia y puede describirse según se indica:

25 El codificador SAOC 510, en particular, un estimador de informaciones secundarias 530 del codificador SAOC 510, extrae las informaciones secundarias que describen las características de las señales de objeto de audio de entrada 32 máximas $s_1 \dots s_{32}$ (en su forma más simple, las relaciones de las potencias del objeto de las señales de objeto de audio). Un mezclador 520 del codificador SAOC 510 mezcla de forma descendente las señales de objeto de audio $s_1 \dots s_{32}$ para obtener una mezcla de señal mono o de 2 canales (es decir, una o dos señales de mezcla descendente)
30 utilizando los factores de ganancia de mezcla descendente $d_{1,1} \dots d_{32,2}$.

[0005] La señal o señales de mezcla descendente y las informaciones secundarias se transmiten o almacenan. Con este fin, la señal o señales de audio de mezcla descendente pueden codificarse utilizando un codificador de audio 540. El codificador de audio 540 puede ser un codificador de audio perceptual bien conocido,
35 por ejemplo, un codificador de audio MPEG-1 Capa II o III (también conocido como .mp3), un codificador de audio de codificación de audio avanzada (AAC) MPEG, etc.

[0006] Del lado del receptor, un decodificador de audio correspondiente 550, por ejemplo, un decodificador de audio perceptual, tal como un decodificador de audio MPEG-1 Capa II o III (también conocido como .mp3),
40 decodificador de audio de codificación de audio avanzada (AAC) MPEG, etc. decodifica la señal o señales de audio de mezcla descendente codificadas.

[0007] Un decodificador SAOC 560 intenta conceptualmente restablecer las señales de objeto (de audio) originales ("separación de objetos") de una o dos señales de mezcla descendente utilizando las informaciones
45 secundarias transmitidas y/o almacenadas, por ejemplo, mediante el empleo de un separador de objeto virtual 570. Estas señales de objeto (de audio) aproximadas $s_{1,est} \dots s_{32,est}$ se mezclan a continuación mediante un renderizador 580 del decodificador SAOC 560 en una escena diana representada por un máximo de 6 canales de salida de audio $y_{1,est} \dots y_{6,est}$ utilizando una matriz de renderización (descrita por los coeficientes $r_{1,1} \dots r_{32,6}$). La salida puede ser una escena diana de un solo canal, un estéreo de 2 canales o de múltiples canales 5.1 (por ejemplo, una, dos o seis
50 señales de salida de audio).

[0008] Debido a las limitaciones subyacentes de la estimación paramétrica de los objetos de audio en el lado de decodificación; en la mayoría de los casos, la escena de salida diana deseada no puede generarse perfectamente. En los puntos de funcionamiento extremos (por ejemplo, la reproducción en solitario de un objeto de audio), a menudo, el procesamiento ya no puede lograr un sonido subjetivo adecuado. Con este fin, el esquema SAOC se ha extendido mediante la introducción de objetos de audio potenciados (OAP) (véase, por ejemplo, [Dfx], véase, por ejemplo, además, [SAOC]). Los objetos de audio que se codifican como OAPs exhiben una capacidad de separación aumentada de los otros objetos de audio no potenciados (regulares) (OAnPs) codificados en la misma señal de mezcla descendente a expensas de un aumento del índice de informaciones secundarias. El concepto OAP

considera para cada OAP el error de predicción (señal residual) del modelo paramétrico.

[0009] La Fig. 6 representa la estimación residual en el lado del codificador, que ilustra esquemáticamente el cálculo de las señales residuales para cada OAP. En el codificador SAOC, las señales residuales (hasta 4 OAPs) se estiman utilizando las informaciones secundarias paramétricas (ISP) extraídas y las señales de fuente originales, la forma de onda codificada e incluida en el flujo de bits SAOC como informaciones secundarias residuales no paramétricas (ISR). Con más detalle, un decodificador SAOC ISP para OAPs 610 genera señales de objeto de audio estimadas $s_{est,OAP}$ a partir de una mezcla descendente X . Una unidad de generación ISR 620 genera entonces hasta cuatro señales residuales $s_{res,ISR, \{1, \dots, 4\}}$ basándose en las señales de objeto de audio estimadas generadas $s_{est,OAP}$ y basándose en las señales de objeto de audio OAP originales s_1, \dots, s_4 .

[0010] La Fig. 7 representa una estructura básica del decodificador SAOC con el soporte OAP, que ilustra una visión general conceptual del esquema de procesamiento de OAP integrado en la cadena de decodificación/transcodificación SAOC (transcodificación = conversión de datos de una codificación a otra codificación).

[0011] Los parámetros orientados a la señal de mezcla descendente, a saber, los coeficientes de predicción de canal (CPC) se derivan de las informaciones secundarias paramétricas (ISP) por una unidad de estimación de CPC 710.

[0012] Los CPC junto con la señal de mezcla descendente se introducen en dos a N cajas (caja TTN) 720. La caja TTN 720 trata conceptualmente de estimar los OAP ($s_{est,OAP}$) a partir de la señal de mezcla descendente transmitida (X) y de proporcionar una mezcla descendente OAnP estimado ($X_{est, OAnP}$) que consiste en sólo OAnP.

[0013] Las señales residuales transmitidas/almacenados (y decodificas) ($s_{res,ISR}$) se utilizan por una unidad de procesamiento ISR 730 para potenciar las estimaciones de OAPs ($s_{est,OAP}$) y la correspondiente mezcla descendente de sólo objetos OAnP (X_{OAnP}).

[0014] Según el estado de la técnica, en la siguiente etapa, la unidad de procesamiento ISR 730 introduce la señal de mezcla descendente OAnP (X_{OAnP}) en un procesador de mezcla descendente SAOC (una unidad de decodificación ISP) 740 para estimar los objetos OAnP $s_{est, OAnP}$. La unidad de decodificación ISP 740 pasa los objetos de audio OAnP estimados $s_{est,OAnP}$ a la unidad de renderización 750. Además, la unidad de procesamiento ISR introduce directamente los OAP potenciados $s_{est,OAP}$ en la unidad de renderización 750. La unidad de renderización 750 genera entonces señales de salida mono o estéreo basándose en los objetos de audio OAnP estimados $s_{est,OAnP}$ y basándose en los OAP potenciados $s_{est, OAP}$.

[0015] El estado del sistema de la técnica tiene los siguientes inconvenientes:

Antes de que las señales residuales se apliquen para calcular los OAP en el decodificador SAOC, los CPC orientados a la mezcla descendente tienen que calcularse a partir de las informaciones secundarias paramétricas transmitidas/almacenadas.

[0016] Todas las señales de mezcla descendente tienen que ser procesadas en el concepto residual SAOC independientemente de su utilidad para el procesamiento de OAP.

[0017] El concepto residual SAOC sólo se puede utilizar con mezclas de señal de uno o dos canales debido a las limitaciones de la caja TTN. El concepto residual OAP no se puede utilizar en combinación con mezclas de múltiples canales (por ejemplo, mezclas de múltiples canales 5,1).

[0018] Además, debido a la correspondiente complejidad computacional de su estimación, el procesamiento OAP SAOC establece limitaciones sobre el número de OAPs (es decir, hasta 4).

[0019] Debido a estas limitaciones, el concepto de manipulación residual de OAP SAOC no se puede aplicar a señales de mezcla descendente de múltiples canales (por ejemplo, 5.1) o utilizarse durante más de 4 OAPs.

[0020] Por lo tanto, sería muy apreciado, si se proporcionasen conceptos mejorados para la codificación de señales de audio, decodificación de señales de audio y procesamiento de señales de audio.

[0021] Un objeto de la presente invención es proporcionar conceptos mejorados para la codificación de

señales de audio, decodificación de señales de audio y procesamiento de señales de audio. El objeto de la presente invención se resuelve por un decodificador según la reivindicación 1, por un generador de señales residuales según la reivindicación 11, por un codificador según la reivindicación 19, por un sistema según la reivindicación 21, por una señal codificada según la reivindicación 22, por un procedimiento según la reivindicación 23, por un procedimiento según la reivindicación 24 y por un programa informático según la reivindicación 25.

[0022] Se proporciona un decodificador. El decodificador comprende una unidad de decodificación paramétrica para generar una pluralidad de primeras señales de objeto de audio estimadas por la mezcla ascendente de tres o más señales de mezcla descendente, en el que las tres o más señales de mezcla descendente codifican una pluralidad de señales de objeto de audio originales, en el que la unidad de decodificación paramétrica se configura para mezclar de forma ascendente tres o más señales de mezcla descendente en función de las informaciones secundarias paramétricas que indican informaciones sobre la pluralidad de señales de objeto de audio originales. Además, el decodificador comprende una unidad de procesamiento residual para generar una pluralidad de segundas señales de objeto de audio estimadas mediante la modificación de una o más de las primeras señales de objeto de audio estimadas, en el que la unidad de procesamiento residual se configura para modificar dichas una o más de las primeras señales de objeto de audio estimadas en función de una o más señales residuales.

[0023] La realización presenta un concepto residual orientado al objeto que mejora la calidad percibida de los OAPs. A diferencia del estado del sistema de la materia, el concepto presentado no queda restringido al número de señales de mezcla descendente ni al número de OAPs. Se presentan dos procedimientos para derivar señales residuales relacionadas con el objeto. Un concepto en cascada con el cual la energía de la señal residual se reduce iterativamente al aumentar el número de OAPs a costa de una mayor complejidad computacional, y un segundo concepto con menos complejidad computacional en el cual todos los residuos se calculan simultáneamente.

[0024] Además, las realizaciones proporcionan un concepto mejorado de la aplicación de señales residuales orientadas a un objeto en el lado del decodificador, y conceptos con complejidad reducida diseñada para escenarios de aplicación en los que sólo los OAP se manipulan en el lado del decodificador, o la modificación de los OAP se restringe a una escala de ganancia.

[0025] Según una realización, la unidad de procesamiento residual puede configurarse para modificar dichas una o más de las primeras señales de objeto de audio estimadas en función de al menos tres señales residuales. El decodificador se adapta para generar al menos tres canales de salida de audio basándose en la pluralidad de segundas señales de objeto de audio estimadas.

[0026] Según una realización, el decodificador puede comprender además una unidad de modificación de mezcla descendente. La unidad de procesamiento residual puede determinar una o más señales de objeto de audio de la pluralidad de segundas señales de objeto de audio estimadas. La unidad de modificación de mezcla descendente puede adaptarse para eliminar una o más segundas señales de objeto de audio estimadas determinadas de tres o más señales de mezcla descendente para obtener tres o más señales de mezcla descendente modificadas. La unidad de decodificación paramétrica puede configurarse para determinar una o más señales de objeto de audio de las primeras señales de objeto de audio estimadas basándose en las tres o más señales de mezcla descendente modificadas.

[0027] En una realización particular, la unidad de modificación de mezcla descendente puede, por ejemplo, adaptarse para aplicar la fórmula

$$\tilde{X}_{OAP} = X - DZ_{obp}^* S_{osp} .$$

[0028] Además, el decodificador puede adaptarse para llevar a cabo dos o más etapas de iteración. Para cada etapa de iteración, la unidad de decodificación paramétrica puede adaptarse para determinar exactamente una señal de objeto de audio de la pluralidad de primeras señales de objeto de audio estimadas. Además, para dicha etapa de iteración, la unidad de procesamiento residual puede adaptarse para determinar exactamente una señal de objeto de audio de la pluralidad de segundas señales de objeto de audio estimadas modificando dicha señal de objeto de audio de la pluralidad de primeras señales de objeto de audio estimadas. Es más, para dicha etapa de iteración, la unidad de modificación de mezcla descendente puede adaptarse para eliminar dicha señal de objeto de audio de la pluralidad de segundas señales de objeto de audio estimadas de las tres o más señales de mezcla descendente para modificar las tres o más señales de mezcla descendente. En la siguiente etapa de iteración posterior a dicha etapa de iteración, la unidad de decodificación paramétrica puede adaptarse para determinar

exactamente una señal de objeto de audio de la pluralidad de primeras señales de objeto de audio estimadas basándose en las tres o más señales de mezcla descendente que se han modificado.

5 **[0029]** En una realización, cada una de una o más señales residuales puede indicar una diferencia entre una de la pluralidad de señales de objeto de audio originales y una de una o más primeras señales de objeto de audio estimadas.

10 **[0030]** Según una realización, en la que la unidad de procesamiento residual puede adaptarse para generar la pluralidad de segundas señales de objeto de audio estimadas mediante la modificación de cinco o más de las primeras señales de objeto de audio estimadas, en la que la unidad de procesamiento residual puede configurarse para modificar dichas cinco o más de las primeras señales de objeto de audio estimadas en función de cinco o más señales residuales.

15 **[0031]** En otra realización, el decodificador puede configurarse para generar siete o más canales de salida de audio basándose en la pluralidad de segundas señales de objeto de audio estimadas.

20 **[0032]** Según una realización adicional, el decodificador puede adaptarse para no determinar coeficientes de predicción de canal para determinar la pluralidad de segundas señales de objeto de audio estimadas. Las realizaciones proporcionan conceptos de modo que el cálculo de los coeficientes de predicción de canal que han sido hasta ahora necesarios para decodificar SAOC del estado de la técnica, ya no son necesarios para la decodificación.

[0033] En una realización adicional, el decodificador puede ser un decodificador SAOC.

25 **[0034]** Además, se proporciona un generador de señal residual. El generador de señal residual comprende una unidad de decodificación paramétrica para generar una pluralidad de señales de objeto de audio estimadas por mezcla ascendente de tres o más señales de mezcla descendente, en el que las tres o más señales de mezcla descendente codifican una pluralidad de señales de objeto de audio originales, en la que la unidad de decodificación paramétrica se configura para mezclar de forma ascendente las tres o más señales de mezcla descendente en función de la informaciones secundarias paramétricas que indican informaciones sobre la pluralidad de señales de objeto de audio originales. Además, el generador de señal residual comprende una unidad de estimación residual para generar una pluralidad de señales residuales basándose en la pluralidad de señales de objeto de audio originales y basándose en la pluralidad de señales de objeto de audio estimadas, de manera tal que cada una de la pluralidad de señales residuales es una señal de diferencia que indica una diferencia entre una de la pluralidad de
35 señales de objeto de audio originales y una de la pluralidad de señales de objeto de audio estimadas.

40 **[0035]** En una realización, la unidad de estimación residual puede adaptarse para generar al menos cinco señales residuales basándose en al menos cinco señales de objeto de audio originales de la pluralidad de señales de objeto de audio originales y basándose en al menos cinco señales de objeto de audio estimadas de la pluralidad de señales de objeto de audio estimadas.

45 **[0036]** En una realización, el generador de señal residual puede comprender además una unidad de modificación de mezcla descendente que se adapta para modificar las tres o más señales de mezcla descendente para obtener tres o más señales de mezcla descendente modificadas. La unidad de decodificación paramétrica puede configurarse para determinar una o más señales de objeto de audio de las primeras señales de objeto de audio estimadas basándose en las tres o más señales de mezcla descendente modificadas.

50 **[0037]** En una realización, la unidad de modificación de mezcla descendente puede, por ejemplo, configurarse para modificar las tres o más señales de mezcla descendente originales para obtener las tres o más señales de mezcla descendente modificadas, mediante la eliminación de una o más de la pluralidad de señales de objeto de audio originales de las tres o más señales de mezcla descendente originales.

55 **[0038]** En otra realización, la unidad de modificación de mezcla descendente puede, por ejemplo, configurarse para modificar las tres o más señales de mezcla descendente originales para obtener las tres o más señales de mezcla descendente modificadas mediante la generación de una o más señales de objeto de audio modificadas basándose en una o más de las señales de objeto de audio estimadas y basándose en una o más de las señales residuales, y mediante la eliminación de una o más señales de objeto de audio modificadas de las tres o más señales de mezcla descendente originales. Por ejemplo, cada una de una o más señales de objeto de audio modificadas pueden generarse por la unidad de modificación de mezcla descendente mediante la modificación de

una de las señales de objeto de audio estimadas, en la que la unidad de modificación de mezcla descendente puede adaptarse para modificar dicha señal de objeto de audio estimada en función de una de una o más señales residuales.

5 **[0039]** En las dos realizaciones descritas anteriormente, la unidad de modificación de mezcla descendente

puede, por ejemplo, adaptarse para aplicar la fórmula
$$\tilde{\mathbf{X}} = \mathbf{X} - \mathbf{DZ}_{oap}^* \mathbf{S}_{oap}$$
, en la que \mathbf{X} es la mezcla descendente que se va a modificar, en la que \mathbf{D} indica informaciones de mezcla descendente, en la que \mathbf{S}_{oap} comprende las

señales de objeto de audio originales a eliminar o las señales de objeto de audio modificadas, en la que con \mathbf{Z}_{oap}^* indica las ubicaciones de las señales a eliminar, y en la que $\tilde{\mathbf{X}}$ es la señal de mezcla descendente modificada. Por

10 ejemplo, una ubicación (posición) de una señal de objeto de audio corresponde a la ubicación (posición) de su objeto de audio en el listado de todos los objetos.

[0040] Según una realización, el generador de señal residual puede adaptarse para llevar a cabo dos o más

15 etapas de iteración. Para cada etapa de iteración, la unidad de decodificación paramétrica puede adaptarse para determinar exactamente una señal de objeto de audio de la pluralidad de señales de objeto de audio estimadas.

Además, para dicha etapa de iteración, la unidad de estimación residual puede adaptarse para determinar exactamente una señal residual de la pluralidad de señales residuales mediante la modificación de dicha señal de objeto de audio de la pluralidad de señales de objeto de audio estimadas. Además, para dicha etapa de iteración, la

20 descendente. En la siguiente etapa de iteración posterior a dicha etapa de iteración, la unidad de decodificación paramétrica puede adaptarse para determinar exactamente una señal de objeto de audio de la pluralidad de señales de objeto de audio estimadas basándose en las tres o más señales de mezcla descendente que se han modificado.

[0041] En una realización, se proporciona un codificador para codificar una pluralidad de señales de objeto de

25 audio originales mediante la generación de tres o más señales de mezcla descendente, mediante la generación de informaciones secundarias paramétricas y mediante la generación de una pluralidad de señales residuales. El codificador comprende un generador de mezcla descendente para proporcionar las tres o más señales de mezcla descendente que indican una mezcla descendente de la pluralidad de señales de objeto de audio originales.

30 Además, el codificador comprende un estimador de informaciones secundarias paramétricas para generar las informaciones secundarias paramétricas que indican las informaciones sobre la pluralidad de señales de objeto de audio originales, para obtener las informaciones secundarias paramétricas. Además, el codificador comprende un generador de señal residual según una de las realizaciones descritas anteriormente. La unidad de decodificación

35 paramétrica del generador de señal residual se adapta para generar una pluralidad de señales de objeto de audio estimadas por la mezcla de forma ascendente de las tres o más señales de mezcla descendente proporcionadas por el generador de mezcla descendente, en el que las señales de mezcla descendente codifican la pluralidad de señales de objeto de audio originales. La unidad de decodificación paramétrica se configura para mezclar de forma

40 ascendente las tres o más señales de mezcla descendente en función de las informaciones secundarias paramétricas generadas por el estimador de informaciones secundarias paramétricas. La unidad de estimación residual del generador de señal residual se adapta para generar la pluralidad de señales residuales basándose en la pluralidad de señales de objeto de audio originales y basándose en la pluralidad de señales de objeto de audio

estimadas, de manera tal que cada una de la pluralidad de señales residuales indica una diferencia entre una de la pluralidad de señales de objeto de audio originales y una de la pluralidad de señales de objeto de audio estimadas.

[0042] En una realización, el codificador puede ser un codificador SAOC.

45

[0043] Además, se proporciona un sistema. El sistema comprende un codificador según una de las realizaciones descritas anteriormente para codificar una pluralidad de señales de objeto de audio originales mediante la generación de tres o más señales de mezcla descendente, mediante la generación de informaciones secundarias paramétricas y mediante la generación de una pluralidad de señales residuales. Además, el sistema

50 comprende un decodificador según una de las realizaciones descritas anteriormente, en el que el decodificador se configura para generar una pluralidad de canales de salida de audio basándose en las tres o más señales de mezcla descendente que se generan por el codificador, basándose en las informaciones secundarias paramétricas que se generan por el codificador y basándose en la pluralidad de señales residuales que se generan por el codificador.

55 **[0044]** Además, se proporciona una señal de audio codificada. La señal de audio codificada comprende tres o más señales de mezcla descendente, informaciones secundarias paramétricas y una pluralidad de señales residuales. Las tres o más señales de mezcla descendente son una mezcla descendente de una pluralidad de

señales de objeto de audio originales. Las informaciones secundarias paramétricas comprenden parámetros que indican informaciones secundarias sobre la pluralidad de señales de objeto de audio originales. Cada una de la pluralidad de señales residuales es una señal de diferencia que indica una diferencia entre una de la pluralidad de señales de audio originales y una de una pluralidad de señales de objeto de audio estimadas.

5

[0045] Además, se proporciona un procedimiento. El procedimiento comprende;

- generar una pluralidad de primeras señales de objeto de audio estimadas por mezcla ascendente de tres o más señales de mezcla descendente, en el que las tres o más señales de mezcla descendente codifican una pluralidad de señales de objeto de audio originales, en el que generar la pluralidad de primeras señales de objeto de audio estimadas comprende la mezcla ascendente de tres o más señales de mezcla descendente en función de las informaciones secundarias paramétricas que indican informaciones sobre la pluralidad de señales de objeto de audio originales. Y:

15 - generar una pluralidad de segundas señales de objeto de audio estimadas mediante la modificación de una o más de las primeras señales de objeto de audio estimadas, en el que la generación de una pluralidad de segundas señales de objeto de audio estimadas comprende la modificación de dicha una o más de las primeras señales de objeto de audio estimadas en función de una o más señales residuales.

20 **[0046]** Además, se proporciona otro procedimiento. Dicho procedimiento comprende:

- generar una pluralidad de señales de objeto de audio estimadas por mezcla ascendente de tres o más señales de mezcla descendente, en el que las tres o más señales de mezcla descendente codifican una pluralidad de señales de objeto de audio originales, en el que generar la pluralidad de señales de objeto de audio estimadas comprende la mezcla ascendente de las tres o más de señales de mezcla descendente en función de las informaciones secundarias paramétricas que indican informaciones sobre la pluralidad de señales de objeto de audio originales. Y:

30 - generar una pluralidad de señales residuales basándose en la pluralidad de señales de objeto de audio originales y basándose en la pluralidad de señales de objeto de audio estimadas, de manera tal que cada una de la pluralidad de señales residuales es una señal de diferencia que indica una diferencia entre una de la pluralidad de señales de objeto de audio originales y una de la pluralidad de señales de objeto de audio estimadas.

[0047] Además, se proporciona un programa informático para implementar uno de los procedimientos descritos anteriormente cuando se ejecuta en un ordenador o en un procesador de señal.

35

[0048] En lo que sigue, las realizaciones de la presente invención se describen con más detalle con referencia a las figuras, en las que:

40 La Fig. 1a ilustra un decodificador según realización,

La Fig. 1b ilustra un decodificador según otra realización, en la que el decodificador comprende además un renderizador,

45 La Fig. 2a ilustra un generador de señal residual según una realización,

La Fig. 2b ilustra un codificador según una realización,

La Fig. 3 ilustra un sistema según una realización,

50 La Fig. 4 ilustra una señal de audio codificada según una realización,

La Fig. 5 representa una vista general del sistema SAOC que ilustra el principio de tales sistemas paramétricos utilizando el ejemplo de MPEG SAOC,

55 La Fig. 6 representa la estimación residual en el lado del codificador, que ilustra esquemáticamente el cálculo de las señales residuales para cada OAP,

La Fig. 7 representa una estructura básica del decodificador SAOC con soporte OAP, que ilustra una visión general conceptual del esquema de procesamiento de OAP integrado en la cadena de decodificación/transcodificación

SAOC,

La Fig. 8 representa una visión general conceptual del esquema de codificación de objeto de audio basado en paramétrico y residual presentado según una realización,

5

La Fig. 9 representa un concepto para estimar conjuntamente la señal residual para cada señal de OAP en el lado del codificador según una realización,

La Fig. 10 ilustra un concepto de decodificación residual conjunta en el lado del decodificador según una realización,

10

La Fig. 11 ilustra un generador de señal residual según una realización, en el que el generador de señal residual comprende además una unidad de modificación de mezcla descendente,

La Fig. 12 ilustra un decodificador según realización, en el que el decodificador comprende además una unidad de modificación de mezcla descendente,

15

La Fig. 13 ilustra un concepto de cálculo de los componentes residuales de una manera en cascada en un lado del codificador según una realización,

20 La Fig. 14 ilustra la unidad de "decodificación de ISR" en cascada empleada en combinación con el cálculo residual en cascada en el lado del decodificador según una realización,

La Fig. 15 ilustra un generador de señal residual según una realización que emplea un concepto en cascada, y

25 La Fig. 16 ilustra un decodificador según una realización, que emplea un concepto en cascada.

[0049] La Fig. 2a ilustra un generador de señal residual 200 según una realización.

[0050] El generador de señal residual 200 comprende una unidad de decodificación paramétrica 230 para generar una pluralidad de señales de objeto de audio estimadas (señal de objeto de audio estimada n.^o 1, ... señal de objeto de audio estimada n.^o M) por mezcla ascendente de tres o más señales de mezcla descendente (señal de mezcla descendente n.^o 1, señal de mezcla descendente n.^o 2, señal de mezcla descendente n.^o 3, ..., señal de mezcla descendente n.^o N). Las tres o más señales de mezcla descendente (señal de mezcla descendente n.^o 1, señal de mezcla descendente n.^o 2, señal de mezcla descendente n.^o 3, ..., señal de mezcla descendente n.^o N) codifican una pluralidad de señales de objeto de audio originales (señal de objeto de audio original n.^o 1, ..., señal de objeto de audio original n.^o M). La unidad de decodificación paramétrica 230 se configura para la mezcla ascendente de las tres o más señales de mezcla descendente (señal de mezcla descendente n.^o 1, señal de mezcla descendente n.^o 2, señal de mezcla descendente n.^o 3, ..., señal de mezcla descendente n.^o N) en función de las informaciones secundarias paramétricas que indican informaciones sobre la pluralidad de señales de objeto de audio originales (señal de objeto de audio original n.^o 1, ..., señal de objeto de audio original n.^o M).

[0051] Además, el generador de señal residual 200 comprende una unidad de estimación residual 240 para generar una pluralidad de señales residuales (señal residual n.^o 1, ..., señal residual n.^o M) basándose en la pluralidad de señales de objeto de audio originales (señal de objeto de audio original n.^o 1, ..., señal de objeto de audio original n.^o M) y basándose en la pluralidad de señales de objeto de audio estimadas (señal de objeto de audio estimada n.^o 1, ..., señal de objeto de audio estimada n.^o M), de manera tal que cada una de la pluralidad de señales residuales (señal residual n.^o 1, ..., señal residual n.^o M) es una señal de diferencia que indica una diferencia entre una de la pluralidad de señales de objeto de audio originales (señal de objeto de audio original n.^o 1, ..., señal de objeto de audio original n.^o M) y una de la pluralidad de señales de objeto de audio estimadas (señal de objeto de audio estimada n.^o 1, ..., señal de objeto de audio estimada n.^o M).

[0052] El codificador según la realización anteriormente descrita supera las restricciones SAOC (véase [SAOC]) del estado de la técnica.

[0053] Los sistemas actuales SAOC llevan a cabo la mezcla descendente empleando una o más dos a una cajas o una o más tres a dos cajas. Entre otras cosas, debido a estas restricciones subyacentes, los sistemas SAOC actuales pueden mezclar de forma descendente señales de objeto de audio como máximo en dos canales de mezcla descendente/dos señales de mezcla descendente.

[0054] Se proporcionan conceptos para generadores de señal residual y para codificadores, que permiten superar las restricciones de SAOC de modo que la codificación de objetos de audio resulta ahora ventajosa para los sistemas de transmisión que emplean más de dos canales de transmisión.

5 **[0055]** En una realización, la unidad de estimación residual 240 se adapta para generar al menos cinco señales residuales basándose en al menos cinco señales de objeto de audio originales de la pluralidad de señales de objeto de audio originales y basándose en al menos cinco señales de objeto de audio estimadas de la pluralidad de señales de objeto de audio estimadas.

10 **[0056]** La Fig. 2b ilustra un codificador según una realización. El codificador de la Fig. 2b comprende un generador de señal residual 200.

[0057] Además, el codificador comprende un generador de mezcla descendente 210 para proporcionar las tres o más señales de mezcla descendente (señal de mezcla descendente n.º 1, señal de mezcla descendente n.º 2, 15 señal de mezcla descendente n.º 3, ..., señal de mezcla descendente n.º N) que indican una mezcla descendente de la pluralidad de señales de objeto de audio originales (señal de objeto de audio original n.º 1, ..., señal de objeto de audio original n.º M, señal o señales de objeto de audio originales adicionales).

[0058] Con respecto a la señal de objeto de audio original n.º 1, ..., señal de objeto de audio original n.º M, la 20 unidad de estimación residual 240 genera una señal residual (señal residual n.º 1, ..., señal residual n.º M). De este modo, señal de objeto de audio original n.º 1, ..., señal de objeto de audio original n.º M se refieren a objetos de audio potenciados (OAPs).

[0059] No obstante, como se puede apreciar en la Fig. 2b, pueden existir opcionalmente señal o señales de 25 objeto de audio originales, que se mezclan de forma descendente, pero para las cuales no se generarán señales residuales. Esta señal o señales de objeto de audio originales adicionales se refieren de este modo a los objetos de audio no potenciados (OAnP).

[0060] El codificador de la Fig. 2b comprende además un estimador de informaciones secundarias 30 paramétricas 220 para generar las informaciones secundarias paramétricas que indican informaciones sobre la pluralidad de señales de objeto de audio originales (señal de objeto de audio original n.º 1, ..., señal de objeto de audio original n.º M, señal o señales de objeto de audio originales adicionales, para obtener las informaciones secundarias paramétricas. En la realización de la Fig. 2b, el estimador de informaciones secundarias paramétricas también toma señales de objeto de audio originales (señal o señales de objeto de audio originales adicionales que 35 tienen en cuenta los OAnP).

[0061] En una realización, el número de señales de objeto de audio originales puede ser igual al número de señales residuales, por ejemplo, cuando todas las señales de objeto de audio originales se refieren a OAPs.

40 **[0062]** En otras realizaciones, sin embargo, el número de señales residuales puede diferir del número de señales de objeto de audio originales y/o puede diferir del número de señales de objeto de audio estimadas, por ejemplo, cuando las señales de objeto de audio originales se refieren a OAnPs.

[0063] En algunas realizaciones, el codificador es un codificador SAOC.

45

[0064] La Fig. 1 ilustra un decodificador según una realización.

[0065] El decodificador comprende una unidad de decodificación paramétrica 110 para generar una 50 pluralidad de primeras señales de objeto de audio estimadas (1ª señal de objeto de audio estimada n.º 1, ... 1ª señal de objeto de audio estimada n.º M) por mezcla ascendente de tres o más señales de mezcla descendente (señal de mezcla descendente n.º 1, señal de mezcla descendente n.º 2, señal de mezcla descendente n.º 3, ..., señal de mezcla descendente n.º N), en el que las tres o más señales de mezcla descendente (señal de mezcla descendente n.º 1, señal de mezcla descendente n.º 2, señal de mezcla descendente n.º 3, ..., señal de mezcla descendente n.º N) codifican una pluralidad de señales de objeto de audio originales, en el que la unidad de decodificación 55 paramétrica 110 se configura para la mezcla ascendente de las tres o más señales de mezcla descendente (señal de mezcla descendente n.º 1, señal de mezcla descendente n.º 2, señal de mezcla descendente n.º 3, ..., señal de mezcla descendente n.º N) en función de las informaciones secundarias paramétricas que indican informaciones sobre la pluralidad de señales de objeto de audio originales.

[0066] Además, el decodificador comprende una unidad de procesamiento residual 120 para generar una pluralidad de segundas señales de objeto de audio estimadas (2ª señal de objeto de audio estimada n.º 1, ... 2ª señal de objeto de audio estimada n.º M) modificando una o más de las primeras señales de objeto de audio estimadas (1ª señal de objeto de audio estimada n.º 1, ... 1ª señal de objeto de audio estimada n.º M), en el que la
5 unidad de procesamiento residual 120 se configura para modificar dichas una o más de las primeras señales de objeto de audio estimadas (1ª señal de objeto de audio estimada n.º 1, ... 1ª señal de objeto de audio estimada n.º M) en función de una o más señales residuales (señal residual n.º 1, ..., señal residual n.º M).

[0067] El decodificador según la realización anteriormente descrita supera las restricciones de SAOC (véase
10 [SAOC]) del estado de la técnica.

[0068] Además, los actuales sistemas de SAOC llevan a cabo la mezcla ascendente mediante el empleo de una a dos cajas (cajas OTT) o una o más de dos a tres cajas (cajas TTT). Entre otras cosas, debido a estas restricciones, las señales de objeto de audio codificadas con más de dos señales de mezcla descendente/canales
15 de mezcla descendente no pueden mezclarse de forma ascendente por los decodificadores SAOC del estado de la técnica.

[0069] Se proporcionan conceptos para decodificadores, que permiten superar las restricciones de SAOC de modo que la codificación de objetos de audio resulta ahora ventajosa para los sistemas de transmisión que emplean
20 más de dos canales de transmisión.

[0070] La Fig. 1b ilustra un decodificador según otra realización, en el que el decodificador comprende además una unidad de renderización 130 para generar la pluralidad de canales de salida de audio (canal de salida de audio n.º 1, ..., canal de salida de audio n.º R) a partir de las segundas señales de objeto de audio estimadas (2ª
25 señal de objeto de audio estimada n.º 1, ... 2ª señal de objeto de audio estimada n.º M) en función de las informaciones de renderización. Por ejemplo, las informaciones de renderización pueden ser una matriz de renderización y/o los coeficientes de una matriz de renderización y la unidad de renderización 130 puede configurarse para aplicar la matriz de renderización en las segundas señales de objeto de audio estimadas (2ª señal de objeto de audio estimada n.º 1, ... 2ª señal de objeto de audio estimada n.º M) para obtener la pluralidad de
30 canales de salida de audio (canal de salida de audio n.º 1, ..., canal de salida de audio n.º R).

[0071] Según una realización, la unidad de procesamiento residual 120 se configura para modificar dichas una o más de las primeras señales de objeto de audio estimadas en función de al menos tres señales residuales. El decodificador se adapta para generar al menos tres canales de salida de audio basándose en la pluralidad de
35 segundas señales de objeto de audio estimadas.

[0072] En otra realización, cada una de una o más señales residuales indica una diferencia entre una de la pluralidad de señales de objeto de audio originales y una de una o más primeras señales de objeto de audio
40 estimadas.

[0073] Según una realización, la unidad de procesamiento residual 120 se adapta para generar la pluralidad de segundas señales de objeto de audio estimadas mediante la modificación de cinco o más de las primeras señales de objeto de audio estimadas. La unidad de procesamiento residual 120 se adapta para modificar dichas cinco o
45 más de las primeras señales de objeto de audio estimadas en función de cinco o más señales residuales.

[0074] En otra realización, el decodificador se configura para generar siete o más canales de salida de audio basándose en la pluralidad de segundas señales de objeto de audio estimadas.

[0075] Según una realización adicional, el decodificador se adapta para no determinar coeficientes de predicción de canal para determinar la pluralidad de segundas señales de objeto de audio estimadas.
50

[0076] En una realización adicional, el decodificador es un decodificador SAOC.

[0077] La Fig. 3 ilustra un sistema según una realización. El sistema comprende un codificador 310 según una de las realizaciones descritas anteriormente para codificar una pluralidad de señales de objeto de audio originales (señal de objeto de audio original n.º 1, ..., señal de objeto de audio original n.º M) mediante la generación de tres o más señales de mezcla descendente, mediante la generación de informaciones secundarias paramétricas y mediante la generación de una pluralidad de señales residuales. Además, el sistema comprende un decodificador
55 320 según una de las realizaciones descritas anteriormente, en el que el decodificador 320 se configura para

generar una pluralidad de segundas señales de objeto de audio estimadas basándose en las tres o más señales de mezcla descendente que se generan por el codificador 310, basándose en las informaciones secundarias paramétricas que se generan por el codificador 310 y basándose en la pluralidad de señales residuales que se generan por el codificador 310.

5

[0078] La Fig. 4 ilustra una señal de audio codificada según una realización. La señal de audio codificada comprende tres o más señales de mezcla descendente 410, informaciones secundarias paramétricas 420 y una pluralidad de señales residuales 430. Las tres o más señales de mezcla descendente 410 son una mezcla descendente de una pluralidad de señales de objeto de audio originales. Las informaciones secundarias paramétricas 420 comprenden parámetros que indican informaciones secundarias sobre la pluralidad de señales de objeto de audio originales. Cada una de la pluralidad de señales residuales 430 es una señal de diferencia que indica una diferencia entre una de la pluralidad de señales de audio originales y una de una pluralidad de señales de objeto de audio estimadas.

10

15 **[0079]** En lo que sigue, se proporciona una visión general del concepto según una realización.

[0080] La Fig. 8 representa una visión general conceptual del esquema de codificación de objeto de audio basado en paramétrica y residual presentado según una realización, en el que el esquema de codificación exhibe una señal de mezcla descendente avanzada y soporte de OAP avanzado.

20

[0081] En el lado del codificador, un estimador de informaciones secundarias paramétricas ("unidad de generación de ISP") 220 calcula la ISP para la estimación de las señales de objetos en la fuente de explotación del decodificador y las características relacionadas con la mezcla descendente. Una unidad de generación de ISR 245 calcula para cada señal de objetos a potenciar las informaciones residuales mediante el análisis de las diferencias entre las señales de objeto estimadas y originales. La unidad de generación de ISR 245 puede, por ejemplo, comprender una unidad de decodificación paramétrica 230 y una unidad de estimación residual 240.

25

[0082] En el lado del decodificador, una unidad de decodificación paramétrica (unidad de "decodificación de ISP") 110 estima las señales de objeto a partir de las señales de mezcla descendente con ISP dado. En una segunda etapa, una unidad de procesamiento residual (unidad de "decodificación de ISR") 120 utiliza ISR para mejorar la calidad de las señales de objeto estimadas a potenciar. Todas las señales de objeto (objetos de audio potenciados y no potenciados) pueden, por ejemplo, pasarse a una unidad de renderización 130 para generar la escena de salida diana.

30

35 **[0083]** Cabe señalar que no es necesario tomar todas las señales de mezcla descendente en consideración. Las señales de mezcla descendente pueden omitirse en el cálculo si su contribución en la estimación y/o la estimación y potenciamiento de las señales de objeto puede ignorarse.

[0084] Para facilitar la comprensión, las etapas de procesamiento de la Fig. 8 y las figuras siguientes se visualizan como unidades de procesamiento separadas. En la práctica, se pueden combinar de manera eficiente para reducir la complejidad computacional.

40

[0085] En lo que sigue, se proporciona un concepto conjunto de codificación/decodificación residual.

45 **[0086]** La Fig. 9 representa un concepto para estimar conjuntamente la señal residual para cada señal de OAP en el lado del codificador según una realización.

[0087] La unidad de decodificación paramétrica (unidad de "decodificación de ISP") 230 produce una estimación de las señales de objeto de audio (señales de objeto de audio estimadas $S_{est,ISP, \{1, \dots, M\}}$) dando ISP estimado y la señal o señales de mezcla descendente como entrada. Las señales de objeto de audio estimadas $S_{est,ISP \{1, \dots, M\}}$ se comparan con las señales de fuente originales inalteradas s_1, \dots, s_M en la unidad de estimación residual ("unidad de estimación de ISR") 240. La unidad de estimación residual 240 proporciona un término de señal residual/de error $S_{res,ISR, \{1, \dots, M\}}$ para cada objeto de audio a potenciar.

50

55 **[0088]** La Fig. 10 muestra la unidad de "decodificación de ISR" que se utiliza en combinación con el cálculo conjunto residual en el decodificador. En particular, la Fig. 10 ilustra un concepto de decodificación conjunta residual en el lado del decodificador según una realización.

[0089] Las (primeras) señales de objeto de audio estimadas $S_{est,ISP, \{1, \dots, M\}}$ de la unidad de decodificación

paramétrica (unidad de "decodificación de ISP") 110 se introducen junto con las informaciones residuales ("informaciones secundarias residuales") en la unidad de procesamiento residual 120 ("decodificación de ISR"). La unidad de procesamiento residual 120 calcula a partir de las informaciones (secundarias) residuales y las señales de objeto de audio estimadas $S_{est,ISP, \{1, \dots, M\}}$, las segundas señales de objeto de audio estimadas $S_{est,ISR, \{1, \dots, M\}}$, por ejemplo, las señales de objeto de audio potenciadas y no potenciadas, y proporciona las segundas señales de objeto de audio estimadas $S_{est,ISR, \{1, \dots, M\}}$, por ejemplo, las señales de objeto de audio potenciadas y no potenciadas, como salida de la unidad de procesamiento residual 120.

[0090] Adicionalmente, una reestimación de los OAnP puede llevarse a cabo (no ilustrados en la Fig. 10). Los OAP se eliminan de la mezcla de señales y los OAnP restantes se vuelven a estimar a partir de esta mezcla. Esto produce una estimación mejorada de estos objetos en comparación con la estimación de la mezcla de señales que comprende todas las señales de objetos. Esta reestimación se puede omitir, si el objetivo es manipular solamente las señales de objeto potenciadas en la mezcla.

[0091] La Fig. 11 ilustra un generador de señal residual según una realización, en la que.

[0092] En la Fig. 11, el generador de señal residual 200 comprende además una unidad de modificación de mezcla descendente 250 que se adapta para modificar las tres o más señales de mezcla descendente para obtener tres o más señales de mezcla descendente modificadas.

[0093] La unidad de decodificación paramétrica 230 se configura para determinar una o más señales de objeto de audio de las primeras señales de objeto de audio estimadas basándose en las tres o más señales de mezcla descendente modificadas.

[0094] Entonces, la unidad de estimación residual 240 puede, por ejemplo, determinar una o más señales residuales basándose en dichas una o más señales de objeto de audio de las primeras señales de objeto de audio estimadas.

[0095] En una realización, la unidad de modificación de mezcla descendente 250 puede, por ejemplo, configurarse para modificar las tres o más señales de mezcla descendente originales para obtener las tres o más señales de mezcla descendente modificadas, mediante la eliminación de una o más de la pluralidad de señales de objeto de audio originales de la tres o más señales de mezcla descendente originales.

[0096] En otra realización, la unidad de modificación de mezcla descendente 250 puede, por ejemplo, configurarse para modificar las tres o más señales de mezcla descendente originales para obtener las tres o más señales de mezcla descendente modificadas mediante la generación de una o más señales de objeto de audio modificadas basándose en una o más de las señales de objeto de audio estimadas y basándose en una o más de las señales residuales, y mediante la eliminación de una o más señales de objeto de audio modificadas a partir de las tres o más señales de mezcla descendente originales. Por ejemplo cada una de las una o más señales de objeto de audio modificadas pueden generarse por la unidad de modificación de mezcla descendente mediante la modificación de una de las señales de objeto de audio estimadas, en la que la unidad de modificación de mezcla descendente puede adaptarse para modificar dicha señal de objeto de audio estimada en función de una de las una o más señales residuales.

[0097] En ambas realizaciones descritas anteriormente, la unidad de modificación de mezcla descendente puede, por ejemplo, adaptarse para aplicar la fórmula

$$\tilde{\mathbf{X}} = \mathbf{X} - \mathbf{D}\mathbf{Z}_{oap}^* \mathbf{S}_{oap} ,$$

en la que \mathbf{X} es el mezcla descendente a modificar,

en la que \mathbf{D} indica las informaciones de mezcla descendente relacionada,

en la que \mathbf{S}_{oap} comprende las señales de objeto de audio originales a eliminar o las señales de objeto de audio modificadas a eliminar,

en la que \mathbf{Z}_{oap}^* indica las ubicaciones de las señales a eliminar, y

en la que \tilde{X} es la señal de mezcla descendente modificada.

[0098] Por ejemplo, una ubicación (posición) de una señal de objeto de audio corresponde a la ubicación (posición) de su objeto de audio en el listado de todos los objetos.

[0099] La Fig. 12 ilustra un decodificador según una realización.

[0100] En la realización de la Fig. 12, el decodificador comprende además una unidad de modificación de mezcla descendente 140.

[0101] La unidad de procesamiento residual 120 determina una o más señales de objeto de audio de la pluralidad de segundas señales de objeto de audio estimadas.

[0102] La unidad de modificación de mezcla descendente 140 se adapta para eliminar una o más segundas señales de objeto de audio estimadas determinadas a partir de las tres o más señales de mezcla descendente para obtener tres o más señales de mezcla descendente modificadas.

[0103] La unidad de decodificación paramétrica 110 se configura para determinar una o más señales de objeto de audio de las primeras señales de objeto de audio estimadas basándose en las tres o más señales de mezcla descendente modificadas.

[0104] La unidad de procesamiento residual 120 puede entonces por ejemplo determinar una o más segundas señales de objeto de audio estimadas adicionales basándose en una o más señales de objeto de audio determinadas de las primeras señales de objeto de audio estimadas.

[0105] En una realización particular, la unidad de modificación de mezcla descendente 130 puede, por ejemplo, adaptarse para aplicar la fórmula:

$$\tilde{X}_{OANP} = X - DZ_{oap}^* S_{oap} .$$

para eliminar una o más señales de objeto de audio de la pluralidad de segundas señales de objeto de audio estimadas determinadas por la unidad de procesamiento residual 120 a partir de las tres o más señales de mezcla descendente para obtener tres o más señales de mezcla descendente modificadas, en las que

X indica las tres o más señales de mezcla descendente antes de ser modificadas

X_{OANP} indica las tres o más señales de mezcla descendente modificadas

D indica una matriz de mezcla descendente

Z_{oap} indica una sub-matriz de mapeo que denota las posiciones (ubicaciones) de OAPs.

[0106] (Para más detalles sobre variantes particulares de esta realización, véase la descripción a continuación).

[0107] En lo siguiente, se presenta un concepto de codificación/decodificación residual en cascada.

[0108] La Fig. 13 ilustra un concepto de cálculo de los componentes residuales de una manera en cascada en un lado del codificador según una realización. En comparación con el concepto de cálculo conjunto residual, la estrategia en cascada reduce en cada etapa de iteración la energía de la energía residual a costa de una mayor complejidad computacional. En cada etapa, una de las señales de objeto de audio originales (s_M) (o, en una realización alternativa, una señal de objeto de audio estimada; véase las flechas de líneas discontinuas 2461, 2462) de un objeto de audio potenciado se elimina de la mezcla de señal (mezcla descendente) antes de que la mezcla de señal (mezcla descendente) pase a la siguiente unidad de procesamiento 2452. De esta manera, el número de señales de objeto en la mezcla de señal (mezcla descendente) disminuye con cada etapa de procesamiento. La estimación de la señal de objeto de audio potenciado (la segunda señal de objeto de audio estimada) en la etapa siguiente mejora por ende, lo que reduce sucesivamente la energía de las señales residuales.

[0109] (Cabe señalar, que en la realización alternativa, cuando en cada etapa de iteración, se elimina una señal de objeto de audio estimada a partir de la mezcla de señal, las subunidades de modificación de mezcla descendente 2501, 2502 no necesitan recibir las señales de objeto de audio originales s_M).

5 **[0110]** Por el contrario, en la realización, cuando en cada etapa de iteración, se elimina una señal de objeto de audio original a partir de la mezcla de señal, las subunidades de modificación de mezcla descendente 2501, 2502 no necesitan recibir las señales de objeto de audio estimadas).

[0111] Con más detalle, la Fig. 13 ilustra una pluralidad de subunidades de generación de ISR 2451, 2452. La pluralidad de subunidades de generación de ISR 2451, 2452 forman en conjunto una unidad de generación de ISR.

[0112] Cada una de la pluralidad de subunidades de generación de ISR 2451, 2452 comprende una subunidad de decodificación paramétrica 2301. La pluralidad de subunidades de decodificación paramétrica 2301 forma en conjunto una unidad de decodificación paramétrica. Las subunidades de decodificación paramétrica 2301 generan las primeras señales de objeto de audio estimadas $s_{est,ISP, \{1, \dots, M\}}$.

[0113] Cada una de la pluralidad de subunidades de generación de ISR 2451, 2452 comprende una subunidad de estimación residual 2401. La pluralidad de subunidades de estimación residuales 2401 forma en conjunto una unidad de estimación residual. Las subunidades de estimación residuales 2401 generan las segundas señales de objeto de audio estimadas $s_{est,ISR,M}, s_{est,ISR,M-1}$.

[0114] Además, la Fig. 13 ilustra una pluralidad de subunidades de modificación de mezcla descendente 2501, 2502. Cada una de las subunidades de modificación de mezcla descendente 2501, 2502 forma en conjunto una unidad de modificación de mezcla descendente.

[0115] La Fig. 14 muestra la unidad de "decodificación de ISR" en cascada empleada en combinación con el cálculo residual en cascada en el lado del decodificador según una realización.

[0116] En cada etapa, una de las señales de objeto a potenciar se estima por una subunidad de decodificación paramétrica 1101 ("decodificación de ISP") (para obtener una de las primeras señales de objeto de audio estimadas $s_{est,ISP,M}$), y una de las primeras señales de objeto de audio estimadas $s_{est,ISP,M}$ se procesa junto con las correspondientes señales residuales $s_{res,ISR,M}$ por una subunidad de procesamiento residual 1201 ("procesamiento de ISR"), para producir la versión de la señal de objeto potenciado (una de las segundas señales de objeto de audio estimadas) $s_{est,ISR,M}$. La señal de objeto potenciado $s_{est,ISR,M}$ se cancela desde la señal de mezcla descendente por una subunidad de modificación de mezcla descendente 1401 ("modificación de mezcla descendente") antes de que las señales de mezcla descendente modificadas se introduzcan en la siguiente subunidad de decodificación residual 1252 ("decodificación residual").

[0117] Igual al concepto conjunto de codificación/decodificación residual, los OAnP pueden, además, volverse a estimar.

[0118] Con más detalle, la Fig. 14 ilustra una pluralidad de subunidades de decodificación residuales 1251, 1252. La pluralidad de subunidades de decodificación residuales 1251, 1252 forman en conjunto una unidad de decodificación residual.

[0119] Cada una de la pluralidad de subunidades de decodificación residuales 1251, 1252 comprende una subunidad de decodificación paramétrica 1101. La pluralidad de subunidades de decodificación paramétrica 1101 forma en conjunto una unidad de decodificación paramétrica. Las subunidades de decodificación paramétrica 1101 generan las primeras señales de objeto de audio estimadas $s_{est,ISP, \{1, \dots, M\}}$.

[0120] Cada una de la pluralidad de subunidades de decodificación residuales 1251, 1252 comprende una subunidad de procesamiento residual 1201. La pluralidad de subunidades de procesamiento residual 1201 forma en conjunto una unidad de procesamiento residual. Las subunidades de procesamiento residual 1201 generan las segundas señales de objeto de audio estimadas $s_{est,ISR,M}, s_{est,ISR,M-1}$.

[0121] Además, la Fig. 14 ilustra una pluralidad de subunidades de modificación de mezcla descendente 1401, 1402. Cada una de las subunidades de modificación de mezcla descendente 1401, 1402 forma en conjunto una unidad de modificación de mezcla descendente.

[0122] La Fig. 15 ilustra un generador de señal residual según una realización que emplea un concepto en cascada.

[0123] En la Fig. 15, el generador de señal residual comprende una unidad de modificación de mezcla descendente 250.

[0124] El generador de señal residual 200 se adapta para llevar a cabo dos o más etapas de iteración:

Para cada etapa de iteración, la unidad de decodificación paramétrica 230 se adapta para determinar exactamente una señal de objeto de audio de la pluralidad de señales de objeto de audio estimadas.

[0125] Además, para dicha etapa de iteración, la unidad de estimación residual 240 se adapta para determinar exactamente una señal residual de la pluralidad de señales residuales mediante la modificación de dicha señal de objeto de audio de la pluralidad de señales de objeto de audio estimadas.

[0126] Además, para dicha etapa de iteración, la unidad de modificación de mezcla descendente 250 se adapta para modificar las tres o más señales de mezcla descendente.

[0127] En la siguiente etapa de iteración posterior a dicha etapa de iteración, la unidad de decodificación paramétrica 230 se adapta para determinar exactamente una señal de objeto de audio de la pluralidad de señales de objeto de audio estimadas basándose en las tres o más señales de mezcla descendente que se han modificado.

[0128] La Fig. 16 ilustra un decodificador según una realización, que emplea un concepto en cascada. En la Fig. 16, el decodificador comprende de nuevo una unidad de modificación de mezcla descendente 140.

[0129] El decodificador de la Fig. 16 se adapta para llevar a cabo dos o más etapas de iteración:

Para cada etapa de iteración, la unidad de decodificación paramétrica 110 se adapta para determinar exactamente una señal de objeto de audio de la pluralidad de primeras señales de objeto de audio estimadas.

[0130] Además, para dicha etapa de iteración, la unidad de procesamiento residual 120 se adapta para determinar exactamente una señal de objeto de audio de la pluralidad de segundas señales de objeto de audio estimadas modificando dicha señal de objeto de audio de la pluralidad de primeras señales de objeto de audio estimadas.

[0131] Además, para dicha etapa de iteración, la unidad de modificación de mezcla descendente 140 se adapta para eliminar dicha señal de objeto de audio de la pluralidad de segundas señales de objeto de audio estimadas a partir de las tres o más señales de mezcla descendente para modificar las tres o más señales de mezcla descendente.

[0132] En la siguiente etapa de iteración posterior a dicha etapa de iteración, la unidad de decodificación paramétrica 110 se adapta para determinar exactamente una señal de objeto de audio de la pluralidad de primeras señales de objeto de audio estimadas basándose en las tres o más señales de mezcla descendente que se han modificado.

[0133] En lo que sigue, se describe una derivación matemática en el ejemplo del concepto de codificación/decodificación residual conjunto:

La siguiente notación se utiliza en lo siguiente:

Dimensiones:

$N_{objetos}$ - número de señales de objeto de audio

N_{cmd} - número de señales de mezcla descendente

$N_{mezclaascendente}$ - número de canales de mezcla ascendente

$N_{muestras}$ - número de datos procesados

N_{OAP} - número de OAPs

Términos:

\mathbf{Z}^* - la estrella-operador (*) indica la transpuesta conjugada de la matriz dada

\mathbf{S} - señal de objeto de audio original proporcionada al codificador (tamaño $N_{Objetos} \times N_{Muestras}$)

5 \mathbf{D} - matriz de mezcla descendente (tamaño $N_{cmd} \times N_{Objetos}$)

\mathbf{R} - matriz de renderización (tamaño $N_{cmezclaascendente} \times N_{Objetos}$)

10 \mathbf{X} - señal de audio de mezcla descendente $\mathbf{X} = \mathbf{DS}$ (tamaño $N_{cmd} \times N_{Muestras}$)

\mathbf{Y} - señal de salida de audio ideal $\mathbf{Y} = \mathbf{RS}$ (tamaño $N_{cmezclaascendente} \times N_{Muestras}$)

15 \mathbf{S}_{est} - señal de objeto reconstruida paraméricamente de manera aproximada $\mathbf{S}_{est} \equiv \mathbf{S}$ definida como $\mathbf{S}_{est} = \mathbf{GX}$ (tamaño $N_{Objetos} \times N_{Muestras}$)

$\hat{\mathbf{S}}_{est}$ - la salida del decodificador que comprende todas las señales de OAnPs (paraméricamente estimados) y OAP (paraméricamente más residuales) estima el tamaño $N_{Objetos} \times N_{Muestras}$

20 $\hat{\mathbf{Y}}_{est}$ - señal de salida de audio de mezcla ascendente de manera aproximada $\hat{\mathbf{Y}}_{est} \equiv \mathbf{Y}$ definida como $\hat{\mathbf{Y}}_{est} = \mathbf{R}\hat{\mathbf{S}}_{est}$ (tamaño $N_{cmezclaascendente} \times N_{Muestras}$)

\mathbf{Z}_{Oanp} ; \mathbf{Z}_{oap} - sub-matriz de mapeo que indica las ubicaciones de OAnPs y OAPs en el listado de todos los objetos.

25 Nota $\mathbf{Z}_{Oanp} \mathbf{Z}_{oap}^* = \begin{bmatrix} 0 \end{bmatrix}$ (tamaño $(N_{Objetos}-N_{OAP}) \times N_{Objetos}$, $N_{OAP} \times N_{Objetos}$). El OAnP \mathbf{Z}_{Oanp} y correspondientes matrices de mapeo \mathbf{Z}_{Oap} se definen como

$$\mathbf{Z}_{Oanp}(i, j) = \begin{cases} 1, & \text{Si el objeto } j \text{ es } i - \text{ en caso contrario,} \\ & \text{es OAnP,} \\ 0, & \end{cases}$$

$$\mathbf{Z}_{Oap}(i, j) = \begin{cases} 1, & \text{Si el objeto } j \text{ es } i - \text{ en caso} \\ & \text{contrario, es OAP.} \\ 0, & \end{cases}$$

30

[0134] Por ejemplo, para $N_{Objetos} = 5$ y los objeto número 2 y 4 son OAPs, estas matrices son

$$\mathbf{Z}_{Oanp} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{Z}_{Oap} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

35

\mathbf{D}_{Oanp} - sub-matriz de mezcla descendente correspondiente a OAnPs, definida como $\mathbf{D}_{Oanp} = \mathbf{DZ}_{Oanp}^*$ (tamaño $N_{cmd} \times (N_{Objetos}-N_{OAP})$)

\mathbf{D}_{Oap} - sub-matriz de mezcla descendente correspondiente a OAPs, definida como $\mathbf{D}_{Oap} = \mathbf{DZ}_{Oap}^*$ (tamaño $N_{cmd} \times N_{OAP}$)

40

\mathbf{G} - matriz de estimación de fuente paramétrica (tamaño $N_{Objetos} \times N_{cmd}$)

E - matriz de covarianza del objeto (tamaño $N_{Objetos} \times N_{Objetos}$)

\mathbf{E}_{Oanp} - sub-matriz de covarianza correspondiente a OAnPs, definida como $\mathbf{E}_{Oanp} = \mathbf{Z}_{Oanp} \mathbf{E} \mathbf{Z}_{Oanp}^*$ (tamaño $(N_{Objetos}-N_{OAP}) \times (N_{Objetos}-N_{OAP})$)

5 **S_{oap}** - señal de OAP que comprende las reconstrucciones de OAPs (tamaño $N_{OAP} \times N_{muestras}$)

S_{oanp} - señal OAnP que comprende las reconstrucciones de OAnPs (tamaño $(N_{Objetos}-N_{OAP}) \times N_{muestras}$)

S_{res} - señales residuales para OAPs (tamaño $N_{OAP} \times N_{muestras}$)

10

X_{oanp} - señal de mezcla descendente modificada que comprende sólo señales de OAnPs; calculada como la diferencia entre la mezcla descendente de SAOC y la mezcla descendente de OAPs reconstruidos (tamaño $N_{cmd} \times N_{muestras}$)

15 **[0135]** Todas las matrices introducidas son (en general) una variante de tiempo y de frecuencia.

[0136] Ahora, se considera un procedimiento general con una reestimación de la señal de OAnP en el lado del decodificador:

20 El procedimiento general se puede describir como una estrategia de dos etapas con una primera extracción de todas las señales de OAP de la señal de mezcla descendente correspondiente, y a continuación una reconstrucción de todas las señales de OAnP considerando los OAPs. Las señales de objeto se recuperan de la señal de mezcla descendente (**X**) utilizando ISP (**E**, **D**) y la señal residual incorporada (**S_{res}**).

25 **[0137]** Se considera que la señal de salida renderizada final $\hat{\mathbf{Y}}_{est}$ se da como:

$$\hat{\mathbf{Y}}_{est} = \mathbf{R} \hat{\mathbf{S}}_{est} .$$

30 **[0138]** La señal de objeto de salida del decodificador $\hat{\mathbf{S}}_{est}$ se puede representar como la siguiente suma:

$$\hat{\mathbf{S}}_{est} = \mathbf{Z}_{oap}^* \mathbf{S}_{oap} + \mathbf{Z}_{Oanp}^* \mathbf{S}_{Oanp} .$$

[0139] La señal de OAP **S_{oap}** se calcula a partir de la mezcla descendente **X** con la ayuda de la matriz de reconstrucción paramétrica de OAP **G_{oap}** y los residuos correspondientes de OAP **S_{res}** según se indica:

35

$$\mathbf{S}_{oap} = \mathbf{G}_{oap} \mathbf{X} + \mathbf{S}_{res} .$$

[0140] La señal de OAnP **S_{Oanp}** se calcula a partir de la mezcla descendente modificada **X_{Oanp}** con la ayuda de la matriz de reconstrucción paramétrica de OAnP **G_{Oanp}** según se indica:

40

$$\mathbf{S}_{Oanp} = \tilde{\mathbf{G}}_{Oanp} \tilde{\mathbf{X}}_{Oanp} .$$

[0141] La señal de mezcla descendente modificada **X_{Oanp}** se determina como la diferencia entre la mezcla descendente **X** y la correspondiente mezcla descendente de OAPs reconstruidos según se indica, cancelando de este modo los OAP de la señal de mezcla descendente **X**:

45

$$\tilde{\mathbf{X}}_{OAnP} = \mathbf{X} - \mathbf{D} \mathbf{Z}_{oap}^* \mathbf{S}_{oap} .$$

[0142] En este caso, las matrices de reconstrucción de objeto paramétricas para OAPs **G_{oap}** y OAnPs **G_{Oanp}** se determinan utilizando ISP (**E**, **D**) según se indica:

50

$$\mathbf{G}_{oap} = \mathbf{Z}_{oap} \mathbf{E} \mathbf{D}^* \mathbf{J} , \quad \mathbf{J} \approx (\mathbf{D} \mathbf{E} \mathbf{D}^*)^{-1} ,$$

$$\tilde{\mathbf{G}}_{Oanp} = \mathbf{E}_{Oanp} \mathbf{D}_{Oanp}^* \mathbf{J}_{Oanp} \dots, \mathbf{J}_{Oanp} \approx \left(\mathbf{D}_{Oanp} \mathbf{E}_{Oanp} \mathbf{D}_{Oanp}^* \right)^{-1}.$$

[0143] En lo que sigue, se describe un procedimiento simplificado "A" sin reestimación de la señal de OAnP en el lado del decodificador:

5

Si sólo se manipulan OAPs en la mezcla de la señal, la escena diana puede interpretarse como una combinación lineal de las señales de mezcla descendente y las señales de OAP. Por ello, la reestimación adicional de las señales OAnP puede omitirse. El procedimiento general con la reestimación de la señal de OAnP se puede simplificar en un procedimiento de una única etapa:

10

$$\hat{\mathbf{S}}_{est} = \mathbf{S}_{est} + \mathbf{X}_{dif}.$$

[0144] La señal $\mathbf{X}_{dif} = f(\mathbf{S}_{res}, \mathbf{D})$ comprende las señales residuales transmitidas de OAPs y los términos de compensación residual de modo que se aplica la siguiente definición:

15

$$\mathbf{D}\hat{\mathbf{S}}_{est} = \mathbf{X}.$$

[0145] Esta condición es suficiente para hacer cualquier escena acústica, que se restringe a manipular solamente OAPs.

20

[0146] Con $\mathbf{D}\hat{\mathbf{S}}_{est} = \mathbf{D}(\mathbf{S}_{est} + \mathbf{X}_{dif}) = \mathbf{X}$ y $\mathbf{D}\mathbf{S}_{est} = \mathbf{X}$, la siguiente restricción para el término \mathbf{X}_{dif} se tiene que cumplir:

$$\mathbf{D}\mathbf{X}_{dif} = 0.$$

25

[0147] El término \mathbf{X}_{dif} consiste en componentes que se determinan por el codificador (y transmitidos o almacenados) \mathbf{S}_{res} y componentes \mathbf{X}_{Oanp} a determinar utilizando esta ecuación.

[0148] Utilizando las definiciones de la matriz de mezcla descendente ($\mathbf{D} = \mathbf{D}_{oap}\mathbf{Z}_{oap} + \mathbf{D}_{Oanp}\mathbf{Z}_{Oanp}$) y el término

30

de compensación ($\mathbf{X}_{dif} = \mathbf{Z}_{oap}^* \mathbf{S}_{res} + \mathbf{Z}_{Oanp}^* \mathbf{X}_{Oanp}$) se puede derivar la siguiente ecuación:

$$\mathbf{D}\mathbf{X}_{dif} = \mathbf{D}_{oap} \mathbf{Z}_{oap} \mathbf{Z}_{oap}^* \mathbf{S}_{res} + \mathbf{D}_{Oanp} \mathbf{Z}_{Oanp} \mathbf{Z}_{Oanp}^* \mathbf{X}_{Oanp} + \mathbf{D}_{oap} \mathbf{Z}_{oap} \mathbf{Z}_{Oanp}^* \mathbf{X}_{Oanp} + \mathbf{D}_{Oanp} \mathbf{Z}_{Oanp} \mathbf{Z}_{oap}^* \mathbf{S}_{res} = 0$$

[0149] Con $\mathbf{Z}_{oap} \mathbf{Z}_{oap}^* = \mathbf{I}$, $\mathbf{Z}_{Oanp} \mathbf{Z}_{Oanp}^* = \mathbf{I}$ y $\mathbf{Z}_{Oanp} \mathbf{Z}_{oap}^* = [0]$; $\mathbf{Z}_{oap} \mathbf{Z}_{Oanp}^* = [0]$, la ecuación se puede

35

simplificar a:

$$\mathbf{D}_{oap} \mathbf{S}_{res} + \mathbf{D}_{Oanp} \mathbf{X}_{Oanp} = 0.$$

[0150] Resolviendo la ecuación lineal para \mathbf{X}_{Oanp} da:

40

$$\mathbf{X}_{Oanp} = -\left(\mathbf{D}_{Oanp} \mathbf{D}_{Oanp} \right)^{-1} \mathbf{D}_{Oanp}^* \mathbf{D}_{oap} \mathbf{S}_{res}.$$

[0151] Después de resolver este sistema de ecuaciones lineales, la escena diana deseada se puede calcular como la siguiente suma del término de predicción paramétrica y el término de mejora residual como:

45

$$\hat{\mathbf{Y}}_{est} = \mathbf{R}\hat{\mathbf{S}}_{est}, \quad \hat{\mathbf{S}}_{est} = \mathbf{S}_{est} + \mathbf{X}_{dif}, \quad \mathbf{X}_{dif} = \mathbf{Z}_{oap}^* \mathbf{S}_{res} - \mathbf{Z}_{Oanp}^* \left(\mathbf{D}_{Oanp}^* \mathbf{D}_{Oanp} \right)^{-1} \mathbf{D}_{Oanp}^* \mathbf{D}_{oap} \mathbf{S}_{res}.$$

[0152] En lo que sigue, se proporciona un procedimiento simplificado "B", sin reestimación de señal de OANP en el lado del decodificador:

Considere el término de compensación \mathbf{X}_{dif} como anteriormente ($\hat{\mathbf{S}}_{est} = \mathbf{S}_{est} + \mathbf{X}_{dif}$) para la predicción de la señal

5 paramétrica \mathbf{S}_{est} y represéntela como la siguiente función $\mathbf{X}_{dif} = \mathbf{H}_{pot} \mathbf{Z}_{oap}^* \mathbf{S}_{res}$ de las señales residuales \mathbf{S}_{res} que dan lugar a:

$$\hat{\mathbf{S}}_{est} = \mathbf{S}_{est} + \mathbf{H}_{pot} \mathbf{Z}_{oap}^* \mathbf{S}_{res}$$

10 **[0153]** Una formulación alternativa comprende las siguientes tres partes incluyendo la combinación apropiada lineal de las señales de mezcla descendente ($\mathbf{H}_{md}\mathbf{X}$), los objetos potenciados ($\mathbf{H}_{pot} \mathbf{Z}_{oap}^* \mathbf{Z}_{oap} \mathbf{S}_{pot}$), y los objetos no potenciados ($\mathbf{H}_{est}\mathbf{S}_{est}$) tal como se indica:

$$\hat{\mathbf{S}}_{est} = \mathbf{H}_{md} \mathbf{X} + \mathbf{H}_{pot} \mathbf{Z}_{oap}^* \mathbf{Z}_{oap} \mathbf{S}_{pot} + \mathbf{H}_{est} \mathbf{S}_{est}.$$

15 **[0154]** Las matrices son de los tamaños \mathbf{H}_{md} : $N_{objetos} \times N_{cmd}$, \mathbf{H}_{pot} : $N_{objetos} \times N_{objetos}$, \mathbf{S}_{pot} : $N_{objetos} \times N_{muestras}$, y \mathbf{H}_{est} : $N_{objetos} \times N_{objetos}$

[0155] Suponiendo que $\mathbf{D}\mathbf{S}_{est} = \mathbf{X}$ y la definición de $\mathbf{S}_{pot} = \mathbf{S}_{est} + \mathbf{Z}_{oap}^* \mathbf{S}_{res}$, esto se puede escribir como:

$$\hat{\mathbf{S}}_{est} = (\mathbf{H}_{md} \mathbf{D} + \mathbf{H}_{pot} \mathbf{Z}_{oap}^* \mathbf{Z}_{oap} + \mathbf{H}_{est}) \mathbf{S}_{est} + \mathbf{H}_{pot} \mathbf{Z}_{oap}^* \mathbf{S}_{res}.$$

[0156] Comparando esto, y la definición anterior de las señales reconstruidas $\hat{\mathbf{S}}_{est} = \mathbf{S}_{est} + \mathbf{H}_{pot} \mathbf{Z}_{oap}^* \mathbf{S}_{res}$, se desprende que:

$$\mathbf{H}_{md} \mathbf{D} + \mathbf{H}_{pot} \mathbf{Z}_{oap}^* \mathbf{Z}_{oap} + \mathbf{H}_{est} = \mathbf{I}.$$

[0157] Se puede derivar el término \mathbf{H}_{est} como:

$$\mathbf{H}_{est} = \mathbf{I} - \mathbf{H}_{ext} \mathbf{D}_{ext}.$$

[0158] Se minimizará el error en la reconstrucción final, cuando se minimice la contribución de las señales no potenciadas. De este modo, la orientación de $\mathbf{H}_{est} = \mathbf{0}$ permite resolver el término \mathbf{H}_{ext} de un sistema de ecuaciones lineales:

$$\mathbf{H}_{ext} = \mathbf{D}_{ext}^* (\mathbf{D}_{ext} \mathbf{D}_{ext}^*)^{-1},$$

cuando la matriz de mezcla descendente extendida \mathbf{D}_{ext} y la matriz de mezcla ascendente \mathbf{H}_{ext} se definan como matrices concatenadas:

$$\mathbf{D}_{ext} = \begin{bmatrix} \mathbf{D} \\ \mathbf{Z}_{oap}^* \mathbf{Z}_{oap} \end{bmatrix} \text{ y } \mathbf{H}_{ext} = [\mathbf{H}_{md} \quad \mathbf{H}_{pot}] \text{ y, por tanto } \mathbf{H}_{pot} = \mathbf{H}_{ext} \begin{bmatrix} \mathbf{0}^{N_{cmd} \times N_{objetos}} \\ \mathbf{I}^{N_{objetos} \times N_{objetos}} \end{bmatrix}$$

[0159] Después de resolver este sistema de ecuaciones lineales, el término de corrección deseado \mathbf{X}_{dif} puede obtenerse como:

$$\mathbf{X}_{dif} = \mathbf{D}_{ext}^* \left(\mathbf{D}_{ext} \mathbf{D}_{ext}^* \right)^{-1} \begin{bmatrix} \mathbf{0}^{N_{cmd} \times N_{Objetos}} \\ \mathbf{I}^{N_{Objetos} \times N_{Objetos}} \end{bmatrix} \mathbf{Z}_{oap}^* \mathbf{S}_{res}$$

[0160] Que conduce a las salidas finales de $\hat{\mathbf{Y}}_{est} = \mathbf{R}\hat{\mathbf{S}}_{est}$, $\hat{\mathbf{S}}_{est} = \mathbf{S}_{est} + \mathbf{X}_{dif}$.

5 [0161] En lo que sigue, se considera un procedimiento simplificado "C":

Si sólo los OAP se manipulan de una manera arbitraria, cualquier escena diana puede generarse por una combinación lineal de las señales de mezcla descendente y los OAP. Tenga en cuenta que en lugar de la mezcla descendente, también se puede utilizar la mezcla descendente con los OAP cancelados. La escena diana puede generarse perfectamente si el procesamiento residual restablece perfectamente los OAP. La renderización de cualquier escena diana se puede efectuar utilizando la búsqueda de las dos matrices de renderización de los componentes \mathbf{R}_D y \mathbf{R}_{oap} para la mezcla descendente y las reconstrucciones de OAP. Las matrices tienen los tamaños \mathbf{R}_D : $N_{mezclaascendente} \times N_{cmd}$ y \mathbf{R}_{oap} : $N_{mezclaascendente} \times N_{OAP}$. La matriz de renderización diana \mathbf{R} puede representarse como un producto de las matrices de renderización combinadas y la matriz de mezcla descendente como

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_D & \mathbf{R}_{oap} \end{bmatrix} \begin{bmatrix} \mathbf{D} \\ \mathbf{Z}_{oap}^* \mathbf{Z}_{oap} \end{bmatrix} = \mathbf{R}_{ext} \mathbf{D}_{ext}$$

[0162] A partir de ello, \mathbf{R}_{ext} se puede resolver con

20

$$\mathbf{R}_{ext} = \mathbf{R} \mathbf{D}_{ext}^* \left(\mathbf{D}_{ext} \mathbf{D}_{ext}^* \right)^{-1}$$

y las sub-matrices \mathbf{R}_D y \mathbf{R}_{oap} pueden extraerse de la solución con

$$\mathbf{R}_D = \mathbf{R}_{ext} \begin{bmatrix} \mathbf{I}^{N_{cmd} \times N_{cmd}} \\ \mathbf{0}^{N_{Objetos} \times N_{Objetos}} \end{bmatrix} \quad \text{y} \quad \mathbf{R}_{oap} = \mathbf{R}_{ext} \begin{bmatrix} \mathbf{0}^{(N_{Objetos} + N_{cmd}) \times N_{OAP}} \\ \mathbf{I}^{N_{OAP} \times N_{OAP}} \end{bmatrix}$$

25

[0163] La escena diana se puede calcular ahora como:

$$\hat{\mathbf{Y}}_{est} = \mathbf{R}_D \mathbf{X} + \mathbf{R}_{oap} \mathbf{S}_{oap}$$

30

en el que \mathbf{S}_{oap} comprende las reconstrucciones completas de OAPs y se define (como antes)

$$\mathbf{S}_{oap} = \mathbf{G}_{oap} \mathbf{X} + \mathbf{S}_{res}$$

35 [0164] Una ecuación similar se puede formular para la renderización de la diana mediante la mezcla descendente con OAPs cancelados de la mezcla restando $\mathbf{D}_{oap} \mathbf{S}_{oap}$ de la mezcla descendente.

[0165] En lo siguiente, se describen otra derivación matemática y más detalles sobre el concepto de codificación/decodificación residual conjunto, y se proporciona una unificación entre el procedimiento general y la simplificación "A".

40

[0166] De ahora en adelante en la descripción, se aplica la siguiente notación. Si para algunos elementos, la siguiente notación es incompatible con la notación proporcionada anteriormente, a partir de ahora en la descripción,

sólo la siguiente notación se aplica a estos elementos.

Definiciones:

5 [0167]

S es la señal de objeto de tamaño $N_{Objetos} \times N_{muestras}$

10 $E = SS^*$ es la matriz de covarianza del objeto de tamaño $N_{Objetos} \times N_{Objetos}$

D es la matriz de mezcla descendente de tamaño $N_{cmd} \times N_{Objetos}$

$X = DS$ es la señal de mezcla descendente de tamaño $N_{cmd} \times N_{muestras}$

15 $G = ED^*J$ es la matriz de mezcla ascendente de tamaño $N_{Objetos} \times N_{cmd}$

M_{ren} es la matriz de renderización de tamaño $N_{mezclaascendente} \times N_{Objetos}$

20 X_{res} es la señal residual de tamaño $N_{OAP} \times N_{muestras}$

R_{oap} es una matriz de tamaño $N_{OAP} \times N_{Objetos}$ que denota las posiciones (ubicaciones) de OAPs definidos como

$$R_{oap}(i, j) = \begin{cases} 1, & \text{si el objeto } j \text{ es } i, \text{ en caso} \\ 0, & \text{contrario, es OAP} \end{cases}$$

25 R_{Oanp} es una matriz de tamaño $(N_{Objetos} - N_{OAP}) \times N_{Objetos}$ que denota las posiciones (ubicaciones) de OAnPs definidos como

$$R_{Oanp}(i, j) = \begin{cases} 1, & \text{si el objeto } j \text{ es } i, \text{ en} \\ 0, & \text{caso contrario, es OAnP} \end{cases}$$

30 [0168] Las sub-matrices de algunos de los anteriores correspondientes a OAnPs se pueden especificar con la ayuda de las matrices de selección R_{Oanp} como:

$$E_{Oanp} = R_{Oanp} E R_{Oanp}^*$$

$$D_{Oanp} = D R_{Oanp}^*$$

$$\begin{aligned} G_{Oanp} &= E_{Oanp} D_{Oanp}^* J_{Oanp} = E_{Oanp} D_{Oanp}^* (D_{Oanp} E_{Oanp} D_{Oanp}^*)^{-1} \\ &= R_{Oanp} E R_{Oanp}^* R_{Oanp} D^* (D R_{Oanp}^* R_{Oanp} E R_{Oanp}^* R_{Oanp} D^*)^{-1} \end{aligned}$$

35 [0169] En lo que sigue, se proporciona otra descripción matemática detallada sobre el procedimiento general (con reestimación de la señal de OAnP en el decodificador):

Las señales de objeto se recuperan de la mezcla descendente utilizando las informaciones secundarias y

señales residuales incorporadas. La salida del decodificador $\hat{\mathbf{X}}$ se produce según se indica

$$\hat{\mathbf{X}} = \mathbf{M}_{ren} \mathbf{R}_{oap}^* \mathbf{X}_{oap} + \mathbf{M}_{ren} \mathbf{R}_{Oanp}^* \mathbf{X}_{Oanp} .$$

El término de OAP \mathbf{X}_{oap} de tamaño N_{OAP} con los OAP se calcula según se indica

5

$$\mathbf{X}_{oap} = \mathbf{R}_{oap} \mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{res} ,$$

en el que el término de señal residual \mathbf{X}_{res} de tamaño N_{OAP} comprende las señales residuales para OAPs.

10 [0171] El término OAnP \mathbf{X}_{Oanp} de tamaño $N_{Objetos} - N_{OAP}$ que comprende OAnPs se calcula como

$$\mathbf{X}_{Oanp} = \mathbf{E}_{Oanp} \mathbf{D}_{Oanp}^* \mathbf{J}_{Oanp} \tilde{\mathbf{X}}_{Oanp} , \quad \mathbf{J}_{Oanp} \approx \left(\mathbf{D}_{Oanp} \mathbf{E}_{Oanp} \mathbf{D}_{Oanp}^* \right)^{-1}$$

cuando la señal de mezcla descendente modificada \mathbf{X}_{Oanp} que comprende sólo señales de OAnPs se calcula como la

15 diferencia entre la mezcla descendente de SAOC y la mezcla descendente de OAPs reconstruidos

$$\tilde{\mathbf{X}}_{Oanp} = \mathbf{X} - \mathbf{D} \mathbf{R}_{oap}^* \mathbf{X}_{oap} .$$

[0172] La sub-matriz de covarianza \mathbf{E}_{Oanp} de tamaño $(N_{Objetos} - N_{OAP}) \times (N_{Objetos} - N_{OAP})$ que corresponde a

20 OAnPs se calcula como

$$\mathbf{E}_{Oanp} = \mathbf{R}_{Oanp} \mathbf{E} \mathbf{R}_{Oanp}^* .$$

[0173] La sub-matriz de mezcla descendente \mathbf{D}_{Oanp} de tamaño $N_{cmd} \times (N_{Objetos} - N_{OAP})$ que corresponde a

25 OAnPs se calcula como

$$\mathbf{D}_{Oanp} = \mathbf{D} \mathbf{R}_{Oanp}^* .$$

[0174] En lo que sigue, se proporciona otra descripción matemática detallada en el procedimiento "A"

30 simplificado (sin reestimación de la señal de OAnP en el decodificador):

Las señales de objeto se recuperan de la mezcla descendente utilizando las informaciones secundarias y las señales residuales incorporadas. La salida final del decodificador \mathbf{X} se produce según se indica

$$\hat{\mathbf{X}} = \mathbf{M}_{ren} \left(\mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{dif} \right) .$$

35

[0175] El término \mathbf{X}_{dif} de tamaño $N_{Objetos}$ incorpora señales residuales N_{OAP} \mathbf{X}_{res} para OAPs y el término previsto \mathbf{X}_{Oanp} para OAnPs según se indica

40

$$\mathbf{X}_{dif} = \mathbf{R}_{oap}^* \mathbf{X}_{res} + \mathbf{R}_{Oanp}^* \mathbf{X}_{Oanp} .$$

[0176] El término previsto \mathbf{X}_{Oanp} se estima según se indica

$$\mathbf{X}_{Oanp} = - \left(\mathbf{D}_{Oanp}^* \mathbf{D}_{Oanp} \right)^{-1} \mathbf{D}_{Oanp}^* \mathbf{D}_{Oanp} \mathbf{X}_{res}$$

45

[0177] La sub-matriz de mezcla descendente \mathbf{D}_{OAP} correspondiente a OAPs y \mathbf{D}_{Oanp} correspondiente a los

objetos regulares se definen como

$$\mathbf{D} = \mathbf{D}_{oap} \mathbf{R}_{oap} + \mathbf{R}_{Oanp} \mathbf{D}_{Oanp} .$$

5 [0178] En lo siguiente, se considera un caso especial de la matriz de renderización 1:

Considere el siguiente caso especial de la matriz de renderización similar a la mezcla descendente \mathbf{M}_D del tamaño $N_{cmd} \times N_{objetos}$ con modificación arbitraria de OAPs y sólo una escala uniforme (en comparación con la mezcla descendente) de OAnPs

10

$$\mathbf{M}_D = \mathbf{M} \mathbf{R}_{oap}^* \mathbf{R}_{oap} + a \mathbf{D} \mathbf{R}_{Oanp}^* \mathbf{R}_{Oanp} .$$

[0179] Ahora, se proporciona una descripción matemática detallada del procedimiento general:

$$\begin{aligned} \hat{\mathbf{X}} &= \mathbf{M}_D \left(\mathbf{R}_{oap}^* \mathbf{X}_{oap} + \mathbf{R}_{Oanp}^* \mathbf{X}_{Oanp} \right) \\ &= \mathbf{M}_D \mathbf{R}_{oap}^* \left(\mathbf{R}_{oap} \mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{res} \right) + \mathbf{M}_D \mathbf{R}_{Oanp}^* \mathbf{G}_{Oanp} \left(\mathbf{X} - \mathbf{D} \mathbf{R}_{oap}^* \mathbf{X}_{oap} \right) \\ &= \mathbf{M}_D \mathbf{R}_{oap}^* \left(\mathbf{R}_{oap} \mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{res} \right) + \mathbf{M}_D \mathbf{R}_{Oanp}^* \mathbf{G}_{Oanp} \left(\mathbf{X} - \mathbf{D} \mathbf{R}_{oap}^* \left(\mathbf{R}_{oap} \mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{res} \right) \right) \\ &= \mathbf{M} \mathbf{R}_{oap}^* \left(\mathbf{R}_{oap} \mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{res} \right) + a \mathbf{D} \mathbf{R}_{Oanp}^* \mathbf{G}_{Oanp} \left(\mathbf{X} - \mathbf{D} \mathbf{R}_{oap}^* \left(\mathbf{R}_{oap} \mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{res} \right) \right) \\ &= \mathbf{M} \mathbf{R}_{oap}^* \left(\mathbf{R}_{oap} \mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{res} \right) + \\ &\quad a \mathbf{D} \mathbf{R}_{Oanp}^* \mathbf{R}_{Oanp} \mathbf{E} \mathbf{R}_{Oanp}^* \mathbf{R}_{Oanp} \mathbf{D}^* \left(\mathbf{D} \mathbf{R}_{Oanp}^* \mathbf{R}_{Oanp} \mathbf{E} \mathbf{R}_{Oanp}^* \mathbf{R}_{Oanp} \mathbf{D}^* \right)^{-1} \left(\mathbf{X} - \mathbf{D} \mathbf{R}_{oap}^* \left(\mathbf{R}_{oap} \mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{res} \right) \right) \\ &= \mathbf{M} \mathbf{R}_{oap}^* \left(\mathbf{R}_{oap} \mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{res} \right) + a \left(\mathbf{X} - \mathbf{D} \mathbf{R}_{oap}^* \left(\mathbf{R}_{oap} \mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{res} \right) \right) \\ &= \mathbf{M} \mathbf{R}_{oap}^* \mathbf{X}_{oap} + a \left(\mathbf{X} - \mathbf{D} \mathbf{R}_{oap}^* \mathbf{X}_{oap} \right) \end{aligned}$$

15

[0180] Ahora, se proporciona una descripción matemática detallada del procedimiento simplificado "A":

$$\begin{aligned}
 \hat{X} &= M_D (GX + X_{dif}) \\
 &= M_D (GX + R_{oap}^* X_{res} + R_{Oanp}^* X_{Oanp}) \\
 &= M_D (GX + R_{oap}^* X_{res} - R_{Oanp}^* (D_{Oanp}^* D_{Oanp})^{-1} D_{Oanp}^* D_{oap} X_{res}) \\
 &= M_D (GX + R_{oap}^* X_{res} - R_{Oanp}^* D_{Oanp}^* (D_{Oanp} D_{Oanp}^*)^{-1} D_{oap} X_{res}) \\
 &= M_D (R_{oap}^* R_{oap} GX + R_{oap}^* X_{res} + R_{Oanp}^* R_{Oanp} GX - R_{Oanp}^* D_{Oanp}^* (D_{Oanp} D_{Oanp}^*)^{-1} D_{oap} X_{res}) \\
 &= M_D (R_{oap}^* X_{oap} + R_{Oanp}^* (R_{Oanp} GX - D_{Oanp}^* (D_{Oanp} D_{Oanp}^*)^{-1} D_{oap} X_{res})) \\
 &= MR_{oap}^* X_{oap} + aDR_{Oanp}^* R_{Oanp} R_{Oanp}^* (R_{Oanp} GX - D_{Oanp}^* (D_{Oanp} D_{Oanp}^*)^{-1} D_{oap} X_{res}) \\
 &= MR_{oap}^* X_{oap} + aDR_{Oanp}^* R_{Oanp} GX - aD_{Oanp} D_{Oanp}^* (D_{Oanp} D_{Oanp}^*)^{-1} D_{oap} X_{res} \\
 &= MR_{oap}^* X_{oap} + aDR_{Oanp}^* R_{Oanp} GX - aD_{oap} X_{res} \\
 &= MR_{oap}^* X_{oap} + a(X - DR_{oap}^* R_{oap} GX) - aD_{oap} X_{res} \\
 &= MR_{oap}^* X_{oap} + a(X - DR_{oap}^* X_{oap})
 \end{aligned}$$

[0181] Se puede apreciar que los dos resultados son idénticos cuando la asunción de la matriz de reproducción se aplica.

5

[0182] Ahora se considera un caso especial de la matriz de renderización 2:

Incluyendo una restricción adicional en la estructura de la matriz de renderización M_s del tamaño $N_{cmd} \times N_{objetos}$: todos los OANPs se modifican solamente por un factor de escala común en comparación con la mezcla descendente, y también todos los OAPs se modifican solamente por un factor de escala común b en comparación con la mezcla descendente:

10

$$M_D = bDR_{oap}^* R_{oap} + aDR_{Oanp}^* R_{Oanp} = D(bR_{oap}^* R_{oap} + aR_{Oanp}^* R_{Oanp}).$$

15 [0183]

Continuando con los resultados anteriores, la salida del sistema será

$$\begin{aligned}
 \hat{X} &= bDR_{oap}^* X_{oap} + a(X - DR_{oap}^* X_{oap}) \\
 &= aX + (b - a)DR_{oap}^* X_{oap} \\
 &= aX + (b - a)DR_{oap}^* (R_{oap} ED^* JX + X_{res})
 \end{aligned}$$

[0184] Aunque algunos aspectos se han descrito en el contexto de un aparato, resulta evidente que estos aspectos también representan una descripción del procedimiento correspondiente, en el que un bloque o dispositivo corresponde a una etapa de procedimiento o a una característica de una etapa del procedimiento. Análogamente, los aspectos descritos en el contexto de una etapa del procedimiento también representan una descripción de un bloque correspondiente o un elemento o característica de un aparato correspondiente.

25 [0185]

La señal descompuesta de la invención puede almacenarse en un medio de almacenamiento digital o puede transmitirse sobre un medio de transmisión, tal como un medio de transmisión inalámbrico o un medio de transmisión por cable, tal como internet.

[0186]

En función de ciertos requisitos de implementación, las realizaciones de la invención pueden

implementarse en hardware o en software. La implementación puede realizarse utilizando un medio de almacenamiento digital, por ejemplo un disquete, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tengan señales de control legibles electrónicamente almacenadas en el mismo, que cooperen (o sean capaces de cooperar) con un sistema informático programable de manera tal que se lleve a cabo el procedimiento respectivo.

[0187] Algunas realizaciones según la invención comprenden un portador de datos no transitorio que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema informático programable, de modo que se lleva a cabo uno de los procedimientos descritos en esta invención.

[0188] Generalmente, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, el código de programa es operativo para realizar uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede almacenarse en un portador legible en máquina, por ejemplo.

[0189] Otras realizaciones comprenden el programa informático para realizar uno de los procedimientos descritos en esta invención, almacenado en un portador legible en máquina.

[0190] En otras palabras, una realización del procedimiento según la invención es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los procedimientos descritos en esta invención, cuando el programa informático se ejecuta en un ordenador.

[0191] Una realización adicional de los procedimientos de la invención es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, registrado en el mismo, el programa informático para realizar uno de los procedimientos descritos en esta invención.

[0192] Una realización adicional del procedimiento de la invención es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para realizar uno de los procedimientos descritos en esta invención. El flujo de datos o la secuencia de señales puede por ejemplo configurarse para transferirse por medio de una conexión de comunicación de datos, por ejemplo a través de internet.

[0193] Una realización adicional comprende un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado o adaptado para realizar uno de los procedimientos descritos en esta invención.

[0194] Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para realizar uno de los procedimientos descritos en esta invención.

[0195] En algunas realizaciones, se puede utilizar un dispositivo lógico programable (por ejemplo una matriz de puertas programables en campo) para realizar algunas o todas las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, una matriz de puertas programables en campo podrá cooperar con un microprocesador con el fin de realizar uno de los procedimientos descritos en esta invención. Generalmente, los procedimientos se realizan preferentemente por cualquier aparato de hardware.

[0196] Las realizaciones anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en esta invención resultarán evidentes para otros expertos en la técnica. El propósito, por lo tanto, es que se limiten sólo por el alcance de las reivindicaciones de patente inminentes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones de esta invención.

REFERENCIAS

[0197]

[BCC] C. Faller and F. Baumgarte, "Binaural Cue Coding - Part II: Schemes and applications," IEEE Trans. on Speech and Audio Proc., vol. 11, no. 6, Nov. 2003

[JSC] C. Faller, "Parametric Joint-Coding of Audio Sources", 120th AES Convention, Paris, 2006

- [SAOC1] J. Herre, S. Disch, J. Hilpert, O. Hellmuth: "From SAC To SAOC - Recent Developments in Parametric Coding of Spatial Audio", 22nd Regional UK AES Conference, Cambridge, UK, April 2007
- 5 [SAOC2] J. Engdegård, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Hölzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers and W. Oomen: " Spatial Audio Object Coding (SAOC) - The Upcoming MPEG Standard on Parametric Object Based Audio Coding", 124th AES Convention, Amsterdam 2008
- 10 [SAOC] ISO/IEC, "MPEG audio technologies - Part 2: Spatial Audio Object Coding (SAOC)," ISO/IEC JTC1/SC29/WG11 (MPEG) International Standard 23003-2:2010.
- [ISS1] M. Parvaix and L. Girin: "Informed Source Separation of underdetermined instantaneous Stereo Mixtures using Source Index Embedding", IEEE ICASSP, 2010
- 15 [ISS2] M. Parvaix, L. Girin, J.-M. Brossier: "A watermarking-based method for informed source separation of audio signals with a single sensor", IEEE Transactions on Audio, Speech and Language Processing, 2010
- [ISS3] A. Liutkus and J. Pinel and R. Badeau and L. Girin and G. Richard: "Informed source separation through spectrogram coding and data embedding", Signal Processing Journal, 2011
- 20 [ISS4] A. Ozerov, A. Liutkus, R. Badeau, G. Richard: "Informed source separation: source coding meets source separation", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2011
- [ISS5] Shuhua Zhang and Laurent Girin: "An Informed Source Separation System for Speech Signals", INTERSPEECH, 2011
- 25 [ISS6] L. Girin and J. Pinel: "Informed Audio Source Separation from Compressed Linear Stereo Mixtures", AES 42nd International Conference: Semantic Audio, 2011
- 30 [Dfx] C. Falch and L. Terentiev and J. Herre: "Spatial Audio Object Coding with Enhanced Audio Object Separation", 10th International Conference on Digital Audio Effects, 2010

REIVINDICACIONES

1. Un decodificador, que comprende

5 una unidad de decodificación paramétrica (110) para generar una pluralidad de primeras señales de objeto de audio estimadas mezclando de forma ascendente tres o más señales de mezcla descendente, en la que las tres o más señales de mezcla descendente codifican una pluralidad de señales de objeto de audio originales, en la que la unidad de decodificación paramétrica (110) se configura para mezclar de forma ascendente las tres o más
 10 señales de mezcla descendente en función de las informaciones secundarias paramétricas que indican informaciones sobre la pluralidad de señales de objeto de audio originales, y una unidad de procesamiento residual (120) para generar una pluralidad de segundas señales de objeto de audio estimadas mediante la modificación de una o más de las primeras señales de objeto de audio estimadas, en la que la unidad de procesamiento residual (120) se configura para modificar dichas una o más de las primeras
 15 señales de objeto de audio estimadas en función de una o más señales residuales,

2. Un decodificador según la reivindicación 1,

en el que el decodificador se adapta para generar al menos tres canales de salida de audio basándose en la pluralidad de segundas señales de objeto de audio estimadas.

3. Un decodificador según una de las reivindicaciones anteriores,

en el que el decodificador comprende además una unidad de modificación de mezcla descendente (140) que se adapta para eliminar una o más señales de objeto de audio de la pluralidad de segundas señales de objeto de audio
 25 estimadas determinadas por la unidad de procesamiento residual (120) de las tres o más señales de mezcla descendente para obtener tres o más señales de mezcla descendente modificadas, y en el que la unidad de decodificación paramétrica (110) se configura para determinar una o más señales de objeto de audio de las primeras señales de objeto de audio estimadas basándose en las tres o más señales de mezcla descendente modificadas.

4. Un decodificador según la reivindicación 3,

en el que la unidad de modificación de mezcla descendente (140) se adapta para aplicar la fórmula:

35
$$\tilde{X}_{OAnP} = X - DZ_{oap}^* S_{oap} .$$

para eliminar una o más señales de objeto de audio de la pluralidad de segundas señales de objeto de audio estimadas determinadas por la unidad de procesamiento residual (120) de las tres o más señales de mezcla descendente para obtener tres o más señales de mezcla descendente modificadas,
 40 en las que

- X** indica las tres o más señales de mezcla descendente antes de ser modificadas
- \tilde{X}_{OAnP}** indica las tres o más señales de mezcla descendente modificadas
- D** indica las informaciones de mezcla descendente
- 45 **S_{oap}** comprende dichas una o más señales de objeto de audio de la pluralidad de segundas señales de objeto de audio estimadas, y
- Z_{oap}^*** indica las posiciones de dichas una o más señales de objeto de audio de la pluralidad de segundas señales de objeto de audio estimadas.

50 5. Un decodificador según la reivindicación 4,

en el que **S_{oap}** se define según:

55
$$S_{oap} = G_{oap} X + S_{res},$$

en el que **G_{oap}** es una matriz de reconstrucción de objetos de audio potenciados, y en el que **S_{res}** son una o más señales residuales que son una o más señales residuales de objetos de audio

potenciados.

6. Un decodificador según la reivindicación 3 o 4,
- 5 en el que el decodificador se adapta para llevar a cabo dos o más etapas de iteración, en el que, para cada etapa de iteración, la unidad de decodificación paramétrica (110) se adapta para determinar exactamente una señal de objeto de audio de la pluralidad de primeras señales de objeto de audio estimadas, en el que, para dicha etapa de iteración, la unidad de procesamiento residual (120) se adapta para determinar exactamente una señal de objeto de audio de la pluralidad de segundas señales de objeto de audio estimadas
- 10 modificando dicha señal de objeto de audio de la pluralidad de primeras señales de objeto de audio estimadas, en el que, para dicha etapa de iteración, la unidad de modificación de mezcla descendente (140) se adapta para eliminar dicha señal de objeto de audio de la pluralidad de segundas señales de objeto de audio estimadas de tres o más señales de mezcla descendente para modificar las tres o más señales de mezcla descendente, y en el que, para la siguiente etapa de iteración posterior a dicha etapa de iteración, la unidad de decodificación
- 15 paramétrica (110) se adapta para determinar exactamente una señal de objeto de audio de la pluralidad de primeras señales de objeto de audio estimadas basándose en tres o más señales de mezcla descendente que se han modificado.
7. Un decodificador según una de las reivindicaciones 1 a 4 o según la reivindicación 6, en el que cada
- 20 una de una o más señales residuales indica una diferencia entre una de la pluralidad de señales de objeto de audio originales y una de una o más primeras señales de objeto de audio estimadas.
8. Un decodificador según la reivindicación 1 o 2,
- 25 en el que la unidad de procesamiento residual (120) se adapta para generar la pluralidad de segundas señales de objeto de audio estimadas mediante la modificación de cinco o más de las primeras señales de objeto de audio estimadas, en el que la unidad de procesamiento residual (120) se configura para modificar dichas cinco o más de las primeras señales de objeto de audio estimadas en función de cinco o más señales residuales.
- 30 9. Un decodificador según la reivindicación 1 o 2, en el que el decodificador se configura para generar siete o más canales de salida de audio basándose en la pluralidad de segundas señales de objeto de audio estimadas.
- 35 10. Un decodificador según una de las reivindicaciones 1 a 4 o según una de las reivindicaciones 6 a 9, en el que el decodificador se adapta para no determinar coeficientes de predicción de canal para determinar la pluralidad de segundas señales de objeto de audio estimadas.
11. Un decodificador según una de las reivindicaciones 1 a 4 o según una de las reivindicaciones 6 a 10,
- 40 en el que el decodificador es un decodificador SAOC de codificación de objetos de audio espaciales.
12. Un generador de señal residual (200), que comprende:
- 45 una unidad de decodificación paramétrica (230) para generar una pluralidad de señales de objeto de audio estimadas por mezcla ascendente de tres o más señales de mezcla descendente, en la que las tres o más señales de mezcla descendente codifican una pluralidad de señales de objeto de audio originales, en la que la unidad de decodificación paramétrica (230) se configura para mezclar de forma ascendente las tres o más señales de mezcla descendente en función de informaciones secundarias paramétricas que indican informaciones sobre la pluralidad de señales de objeto de audio originales, y
- 50 una unidad de estimación residual (240) para generar una pluralidad de señales residuales basándose en la pluralidad de señales de objeto de audio originales y basándose en la pluralidad de señales de objeto de audio estimadas, de manera tal que cada una de la pluralidad de señales residuales es una señal de diferencia que indica una diferencia entre una de la pluralidad de señales de objeto de audio originales y una de la pluralidad de señales de objeto de audio estimadas.
- 55 13. Un generador de señal residual (200) según la reivindicación 12,
- en el que el generador de señal residual (200) comprende además una unidad de modificación de mezcla descendente (250) que se adapta para modificar las tres o más señales de mezcla descendente para obtener tres o

más señales de mezcla descendente modificadas, y en el que la unidad de decodificación paramétrica (230) se configura para determinar una o más señales de objeto de audio de las primeras señales de objeto de audio estimadas basándose en las tres o más señales de mezcla descendente modificadas.

5

14. Un generador de señal residual (200) según la reivindicación 13, en el que la unidad de modificación de mezcla descendente (250) se configura para modificar las tres o más señales de mezcla descendente originales para obtener las tres o más señales de mezcla descendente modificadas, mediante la eliminación de una o más de la pluralidad de señales de objeto de audio originales de las tres o más señales de mezcla descendente originales.

10

15. Un generador de señal residual según la reivindicación 14, en el que la unidad de modificación de mezcla descendente (250) se adapta para aplicar la fórmula:

$$\tilde{\mathbf{X}}_{OAnP} = \mathbf{X} - \mathbf{D}\mathbf{Z}_{oap}^* \mathbf{S}_{oap} .$$

15

para eliminar una o más de la pluralidad de señales de objeto de audio originales de las tres o más señales de mezcla descendente para obtener tres o más señales de mezcla descendente modificadas, en las que

20

\mathbf{X} indica las tres o más señales de mezcla descendente antes de ser modificadas

$\tilde{\mathbf{X}}_{OAnP}$ indica las tres o más señales de mezcla descendente modificadas

\mathbf{D} indica informaciones de mezcla descendente

\mathbf{S}_{oap} comprende dichas una o más de la pluralidad de señales de objeto de audio originales, y

\mathbf{Z}_{oap}^* indica las ubicaciones de dichas una o más de la pluralidad de señales de objeto de audio originales.

25

16. Un generador de señal residual (200) según la reivindicación 13, en el que la unidad de modificación de mezcla descendente (250) se configura para modificar las tres o más señales de mezcla descendente originales para obtener las tres o más señales de mezcla descendente modificadas mediante la generación de una o más señales de objeto de audio modificadas basándose en una o más de las señales de objeto de audio estimadas y basándose en una o más de las señales residuales, y mediante la eliminación de una o más señales de objeto de audio modificadas de las tres o más señales de mezcla descendente originales.

30

17. Un generador de señal residual según la reivindicación 16, en el que la unidad de modificación de mezcla descendente (250) se adapta para aplicar la fórmula:

35

$$\tilde{\mathbf{X}}_{OAnP} = \mathbf{X} - \mathbf{D}\mathbf{Z}_{oap}^* \mathbf{S}_{oap} .$$

para eliminar una o más de la pluralidad de señales de objeto de audio modificadas de las tres o más señales de mezcla descendente para obtener tres o más señales de mezcla descendente modificadas,

40

en las que

\mathbf{X} indica las tres o más señales de mezcla descendente antes de ser modificadas

$\tilde{\mathbf{X}}_{OAnP}$ indica las tres o más señales de mezcla descendente modificadas

\mathbf{D} indica informaciones de mezcla descendente

45

\mathbf{S}_{oap} comprende dichas una o más señales de objeto de audio modificadas, y

\mathbf{Z}_{oap}^* indica las ubicaciones de dichas una o más de señales de objeto de audio modificadas.

18. Un generador de señal residual según la reivindicación 15 o 17, en el que \mathbf{S}_{oap} se define según:

50

$$\mathbf{S}_{oap} = \mathbf{G}_{oap} \mathbf{X} + \mathbf{S}_{res},$$

en el que \mathbf{G}_{oap} es una matriz de reconstrucción de objetos de audio potenciados, y

en el que \mathbf{S}_{res} son una o más señales residuales que son una o más señales residuales de objetos de audio potenciados.

55

19. Un generador de señal residual (200) según una de las reivindicaciones 13 a 17,

en el que el generador de señal residual (200) se adapta para llevar a cabo dos o más etapas de iteración,
 5 en el que, para cada etapa de iteración, la unidad de decodificación paramétrica (230) se adapta para determinar exactamente una señal de objeto de audio de la pluralidad de señales de objeto de audio estimadas,
 en el que, para dicha etapa de iteración, la unidad de estimación residual (240) se adapta para determinar exactamente una señal residual de la pluralidad de señales residuales modificando dicha señal de objeto de audio de la pluralidad de señales de objeto de audio estimadas,
 10 en el que, para dicha etapa de iteración, la unidad de modificación de mezcla descendente (250) se adapta para modificar las tres o más señales de mezcla descendente, y
 en el que, para la siguiente etapa de iteración posterior a dicha etapa de iteración, la unidad de decodificación paramétrica (230) se adapta para determinar exactamente una señal de objeto de audio de la pluralidad de señales de objeto de audio estimadas basándose en tres o más señales de mezcla descendente que se han modificado.

15 20. Un generador de señal residual (200) según una de las reivindicaciones 12 a 16 o según la reivindicación 18, en el que la unidad de estimación residual (240) se adapta para generar al menos cinco señales residuales basándose en al menos cinco señales de objeto de audio originales de la pluralidad de señales de objeto de audio originales y basándose en al menos cinco señales de objeto de audio estimadas de la pluralidad de señales de objeto de audio estimadas.

21. Un codificador para codificar una pluralidad de señales de objeto de audio originales mediante la generación de tres o más señales de mezcla descendente, mediante la generación de informaciones secundarias paramétricas y mediante la generación de una pluralidad de señales residuales, en el que el codificador comprende:

25 un generador de mezcla descendente (210) para proporcionar las tres o más señales de mezcla descendente que indican una mezcla descendente de la pluralidad de señales de objeto de audio originales,
 un estimador de informaciones secundarias paramétricas (220) para generar las informaciones secundarias paramétricas que indican informaciones sobre la pluralidad de señales de objeto de audio originales, para
 30 obtener las informaciones secundarias paramétricas, y
 un generador de señal residual (200) según una de las realizaciones 12 a 20,
 en el que la unidad de decodificación paramétrica (230) del generador de señal residual (200) se adapta para generar una pluralidad de señales de objeto de audio estimadas por mezcla descendente de las tres o más
 35 señales de mezcla descendente proporcionadas por el generador de mezcla descendente (210), en el que las señales de mezcla descendente codifican la pluralidad de señales de objeto de audio originales, en el que la unidad de decodificación paramétrica (230) se configura para mezclar de forma ascendente las tres o más señales de mezcla descendente en función de las informaciones secundarias paramétricas generadas por el estimador de informaciones secundarias paramétricas (220), y
 40 en el que la unidad de estimación residual (240) del generador de señal residual (200) se adapta para generar la pluralidad de señales residuales basándose en la pluralidad de señales de objeto de audio originales y basándose en la pluralidad de señales de objeto de audio estimadas, de manera tal que cada una de la pluralidad de señales residuales indica una diferencia entre una de la pluralidad de señales de objeto de audio originales y una de la pluralidad de señales de objeto de audio estimadas.

45 22. Un codificador según la reivindicación 21, en el que el codificador es un codificador SAOC.

23. Un sistema, que comprende:

50 un codificador (310) según la reivindicación 21 o 22 para codificar una pluralidad de señales de objeto de audio originales mediante la generación de tres o más señales de mezcla descendente, mediante la generación de informaciones secundarias paramétricas y mediante la generación de una pluralidad de señales residuales, y
 un decodificador (320) según una de las reivindicaciones 1 a 11, en el que el decodificador (320) se configura para generar una pluralidad de segundas señales de objeto de audio estimadas basándose en las tres o más
 55 señales de mezcla descendente que se generan por el codificador (310), basándose en las informaciones secundarias paramétricas que se generan por el codificador (310) y basándose en la pluralidad de señales residuales que se generan por el codificador (310).

24. Una señal de audio codificada, que comprende tres o más señales de mezcla descendente (410), informaciones secundarias paramétricas (420) y una pluralidad de señales residuales (430),

en la que las tres o más señales de mezcla descendente (410) son una mezcla descendente de una pluralidad de señales de objeto de audio originales,

en la que las informaciones secundarias paramétricas (420) comprenden parámetros que indican informaciones secundarias en la pluralidad de señales de objeto de audio originales,

5 en la que cada una de la pluralidad de señales residuales (430) es una señal de diferencia que indica una diferencia entre una de la pluralidad de señales de audio originales y una de una pluralidad de señales de objeto de audio estimadas.

25. Un procedimiento, que comprende:

10

generar una pluralidad de primeras señales de objeto de audio estimadas por mezcla ascendente de tres o más señales de mezcla descendente, en el que las tres o más señales de mezcla descendente codifican una pluralidad de señales de objeto de audio originales, en el que generar la pluralidad de primeras señales de objeto de audio estimadas comprende la mezcla ascendente de las tres o más de señales de mezcla descendente en función de las informaciones secundarias paramétricas que indican informaciones sobre la pluralidad de señales de objeto de audio originales, y:

15

generar una pluralidad de segundas señales de objeto de audio estimadas modificando una o más de las primeras señales de objeto de audio estimadas, en el que generar una pluralidad de segundas señales de objeto de audio estimadas comprende modificar dicha una o más de las primeras señales de objeto de audio estimadas en función de una o más señales residuales.

20

26. Un procedimiento, que comprende:

25

generar una pluralidad de señales de objeto de audio estimadas por mezcla ascendente de tres o más señales de mezcla descendente, en el que las tres o más señales de mezcla descendente codifican una pluralidad de señales de objeto de audio originales, en el que generar la pluralidad de señales de objeto de audio estimadas comprende la mezcla ascendente de las tres o más de señales de mezcla descendente en función de las informaciones secundarias paramétricas que indican informaciones sobre la pluralidad de señales de objeto de audio originales. y

30

generar una pluralidad de señales residuales basándose en la pluralidad de señales de objeto de audio originales y basándose en la pluralidad de señales de objeto de audio estimadas, de manera tal que cada una de la pluralidad de señales residuales es una señal de diferencia que indica una diferencia entre una de la pluralidad de señales de objeto de audio originales y una de la pluralidad de señales de objeto de audio estimadas.

35 27.

Un programa informático adaptado para implementar el procedimiento de la reivindicación 25 o 26 cuando se ejecuta en un ordenador o en un procesador de señal.

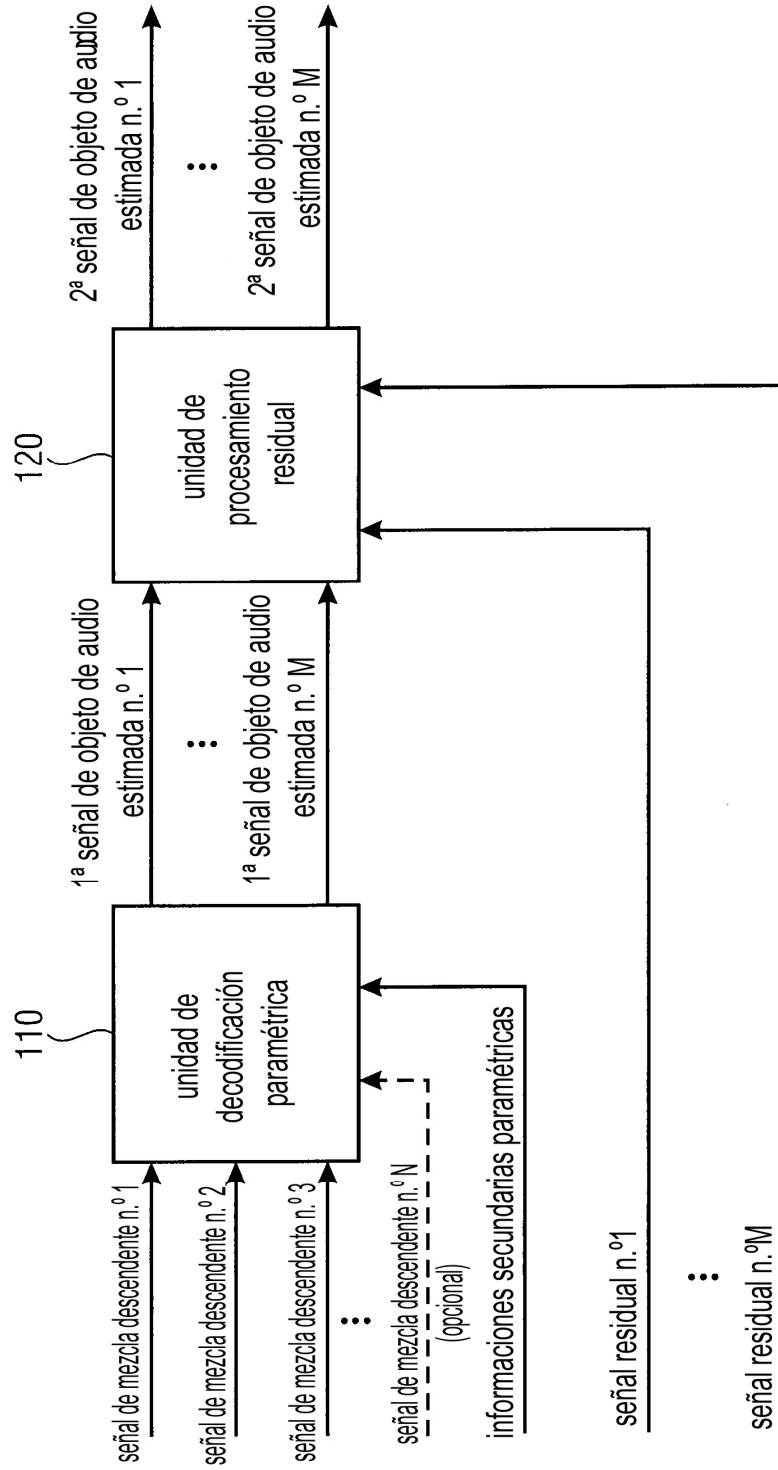


FIG 1A

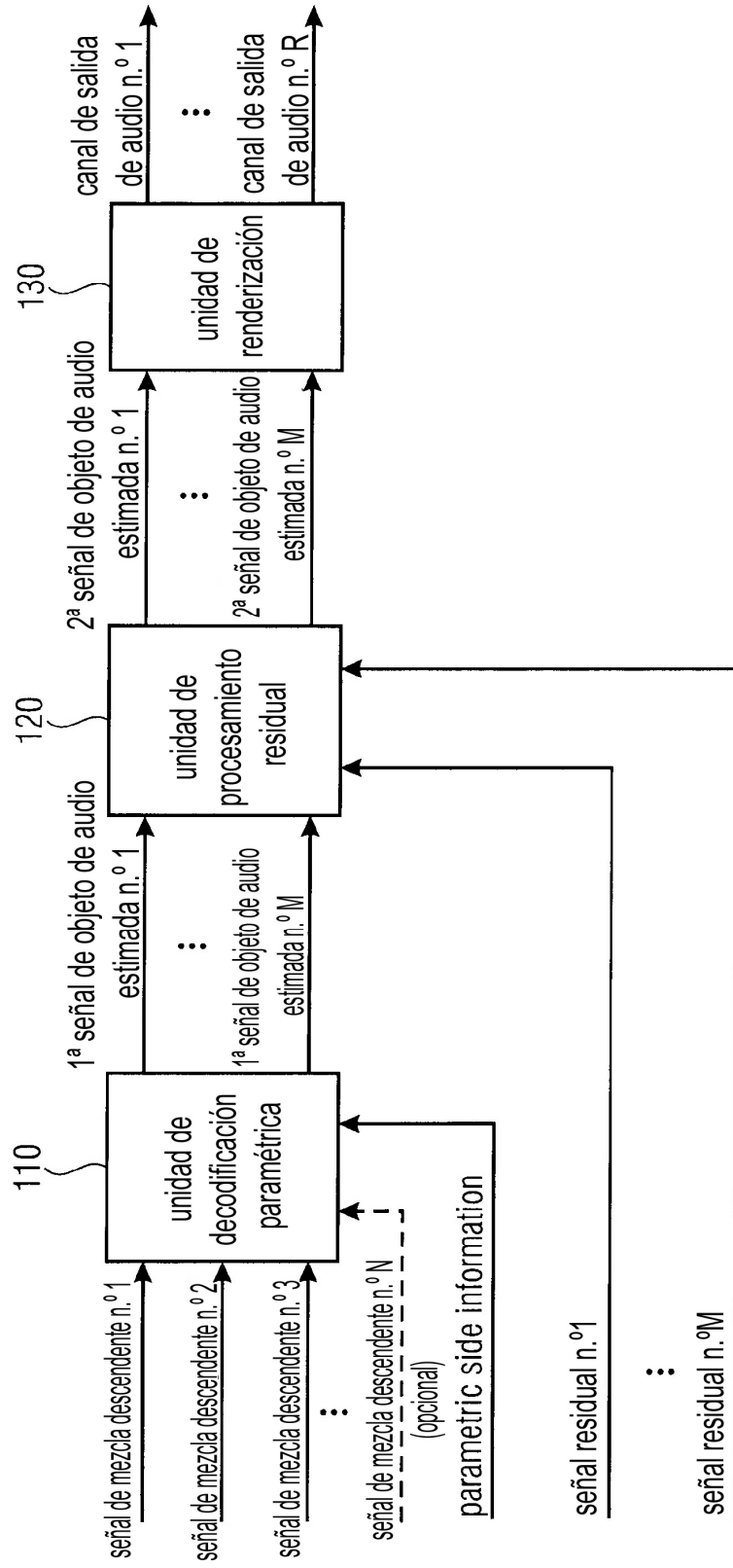


FIG 1B

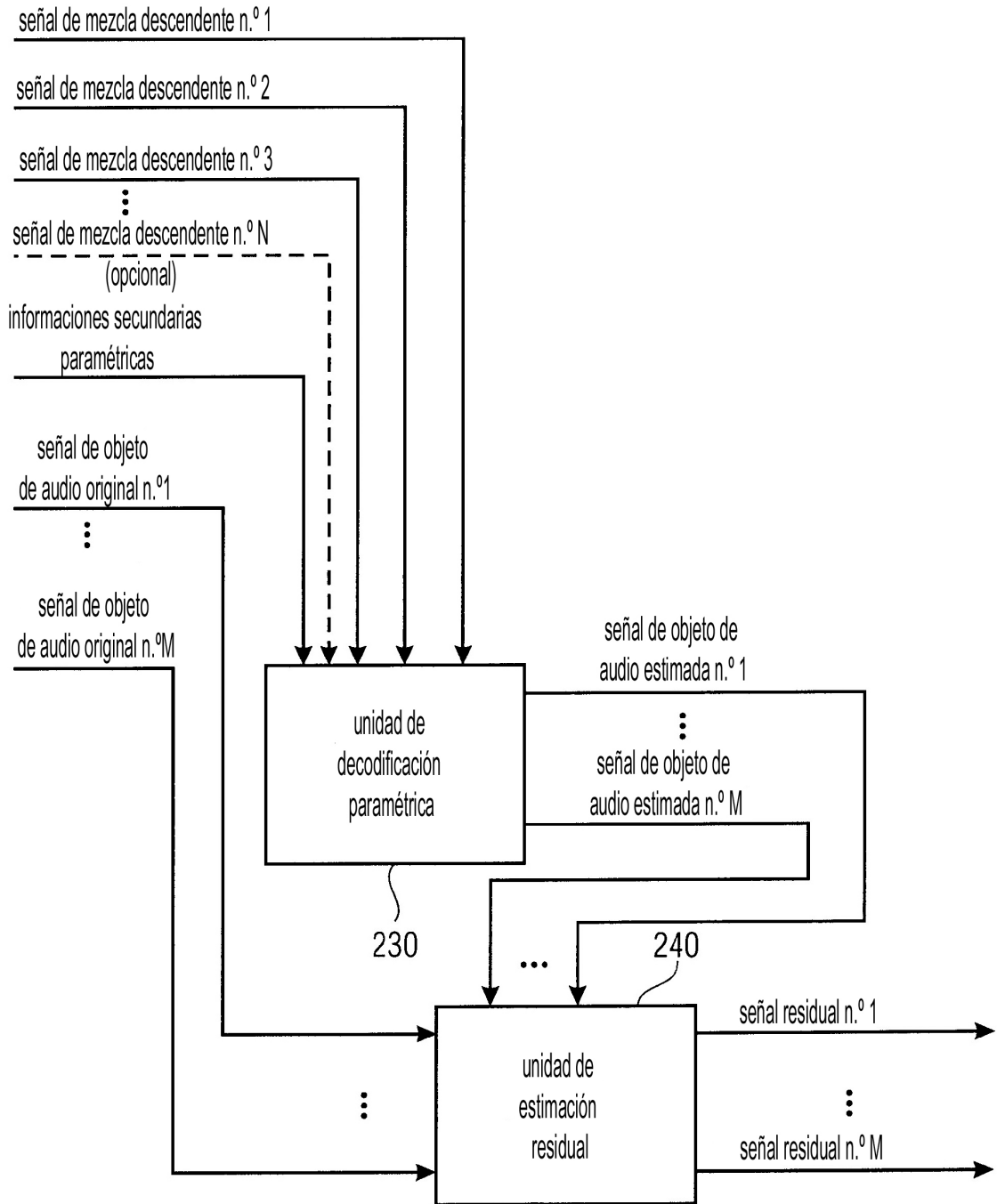


FIG 2A

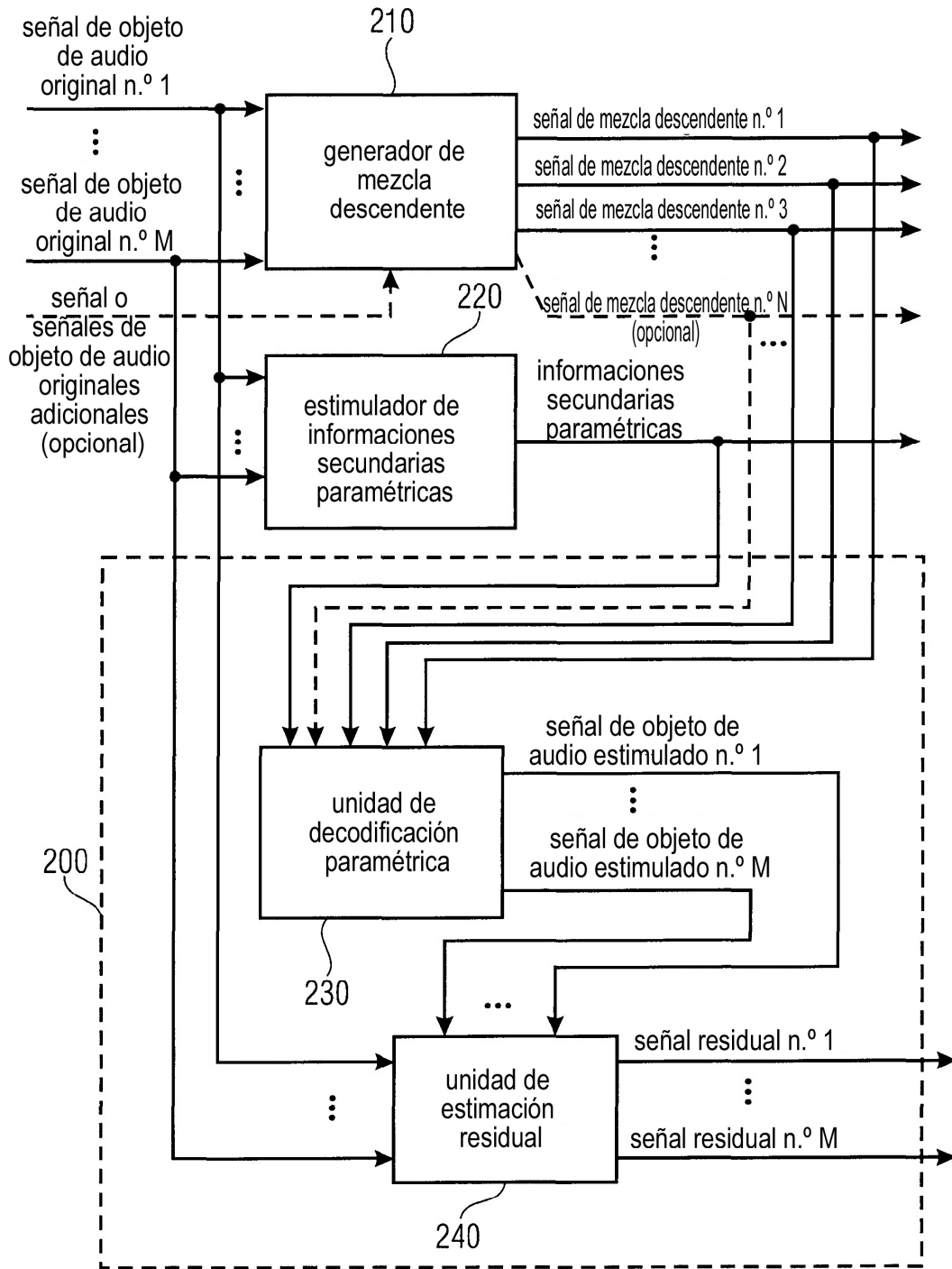


FIG 2B

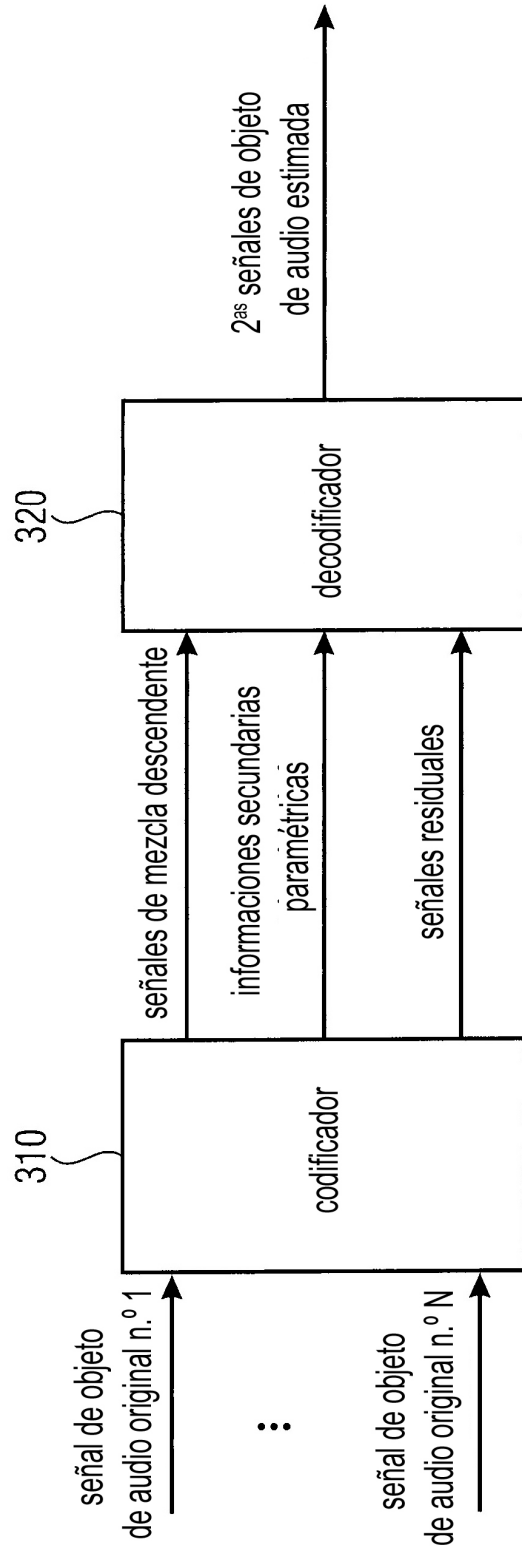


FIG 3

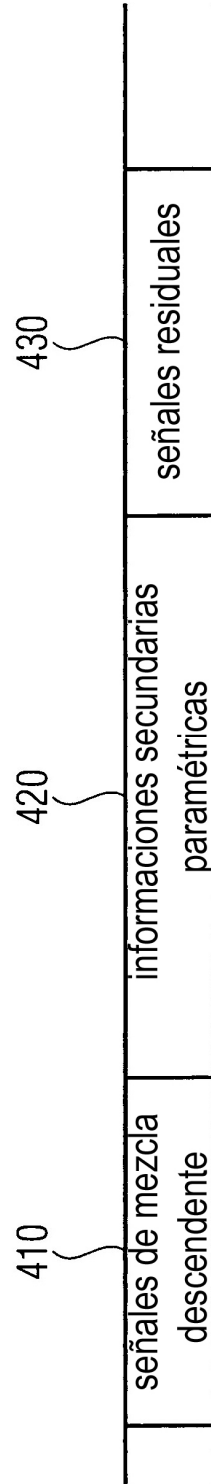


FIG 4

Salida de canal mono, estéreo o 5.1

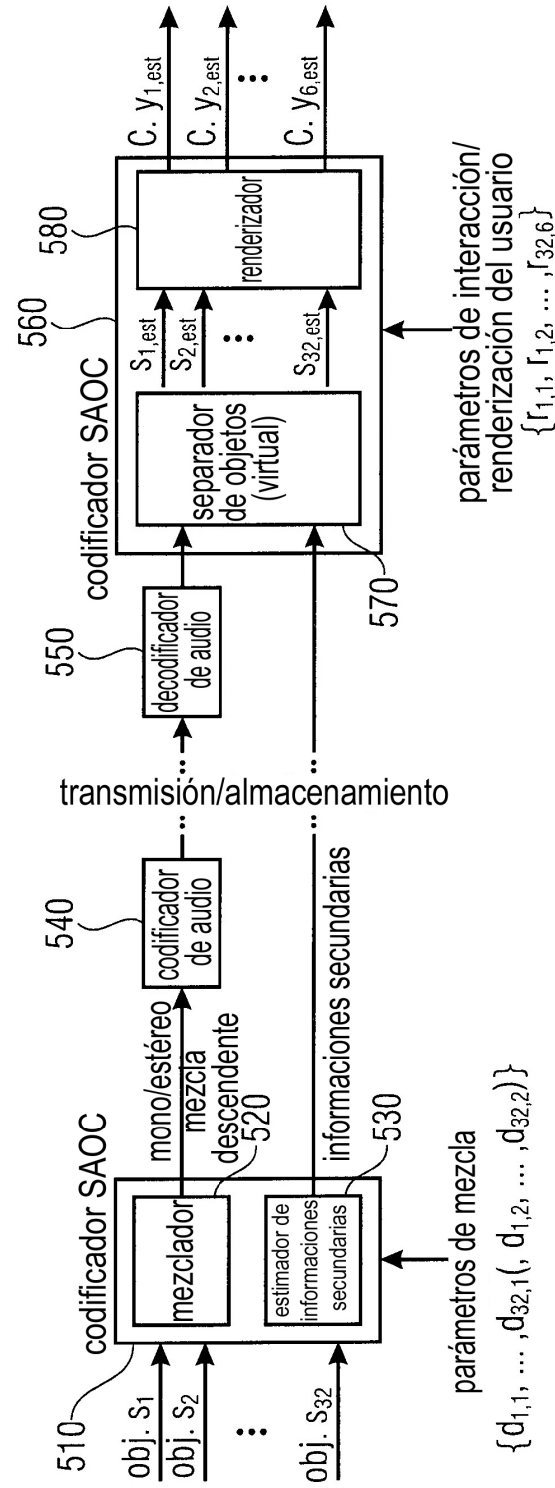


FIG 5

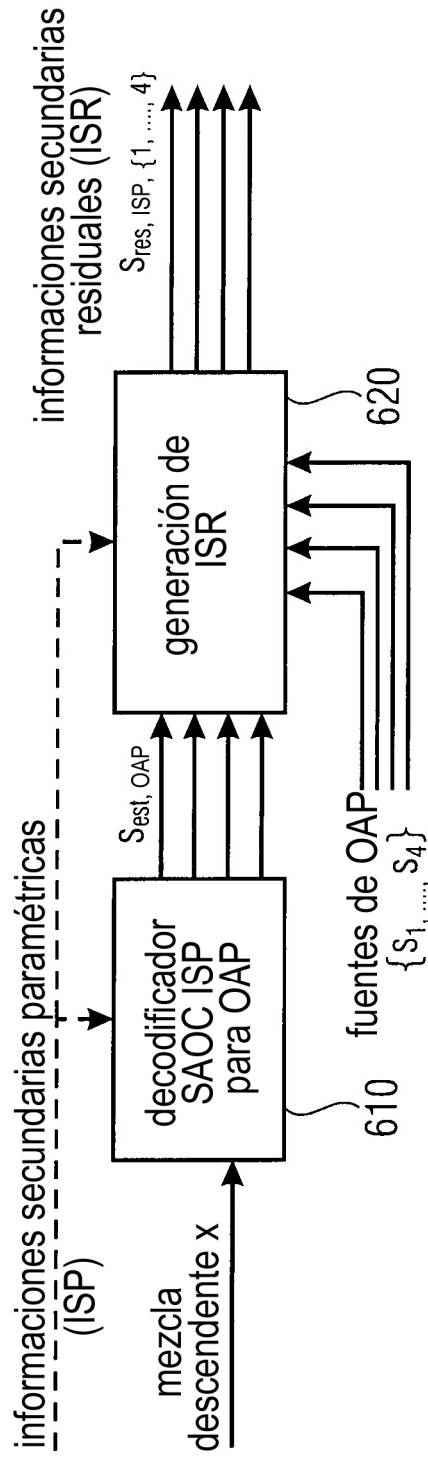


FIG 6

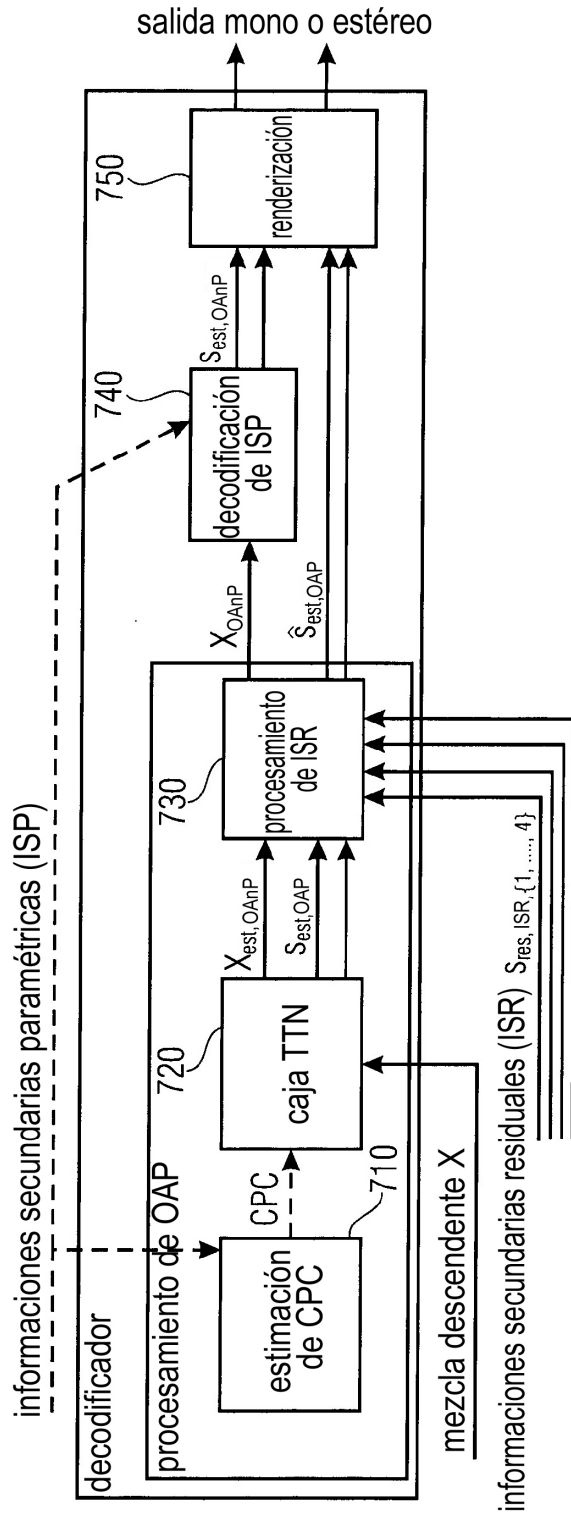
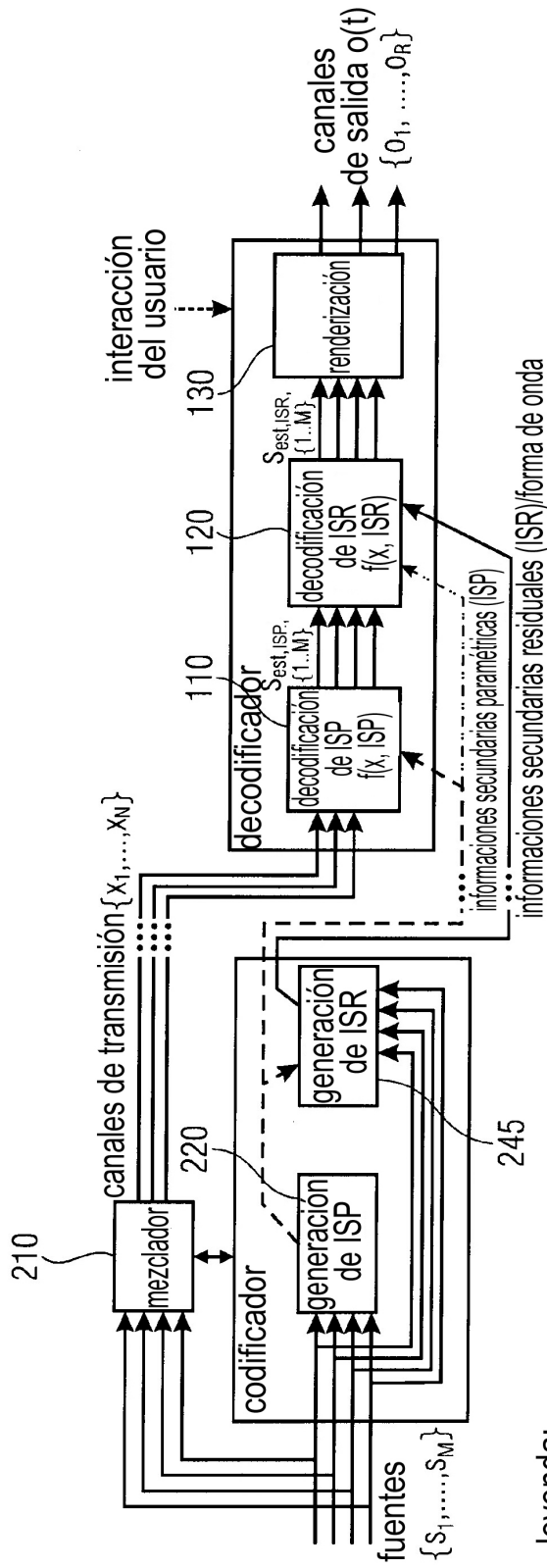


FIG 7



leyenda:

ISP: informaciones secundarias paramétricas

ISR: informaciones secundarias residuales/forma de onda

$S_{est, ISP}$: señales de fuente estimadas tras la decodificación de ISP

$S_{est, ISR}$: señales de fuente estimadas tras la decodificación de ISR

FIG 8

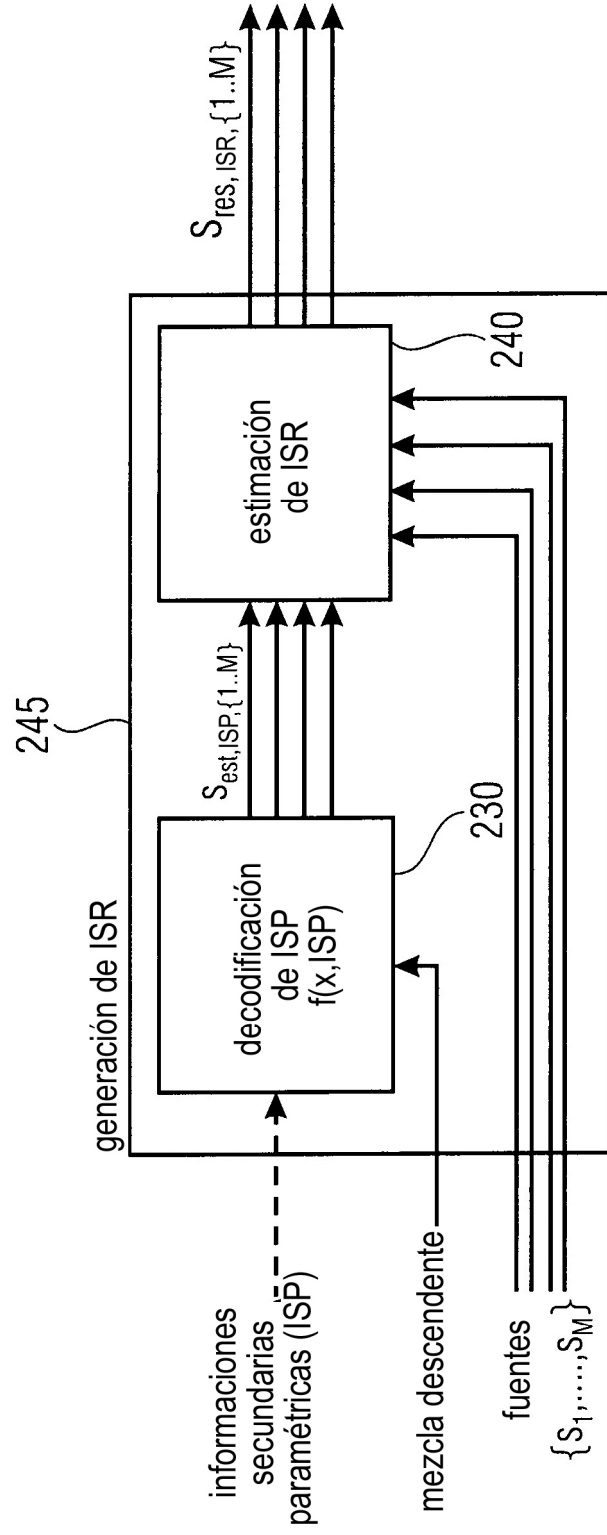


FIG 9

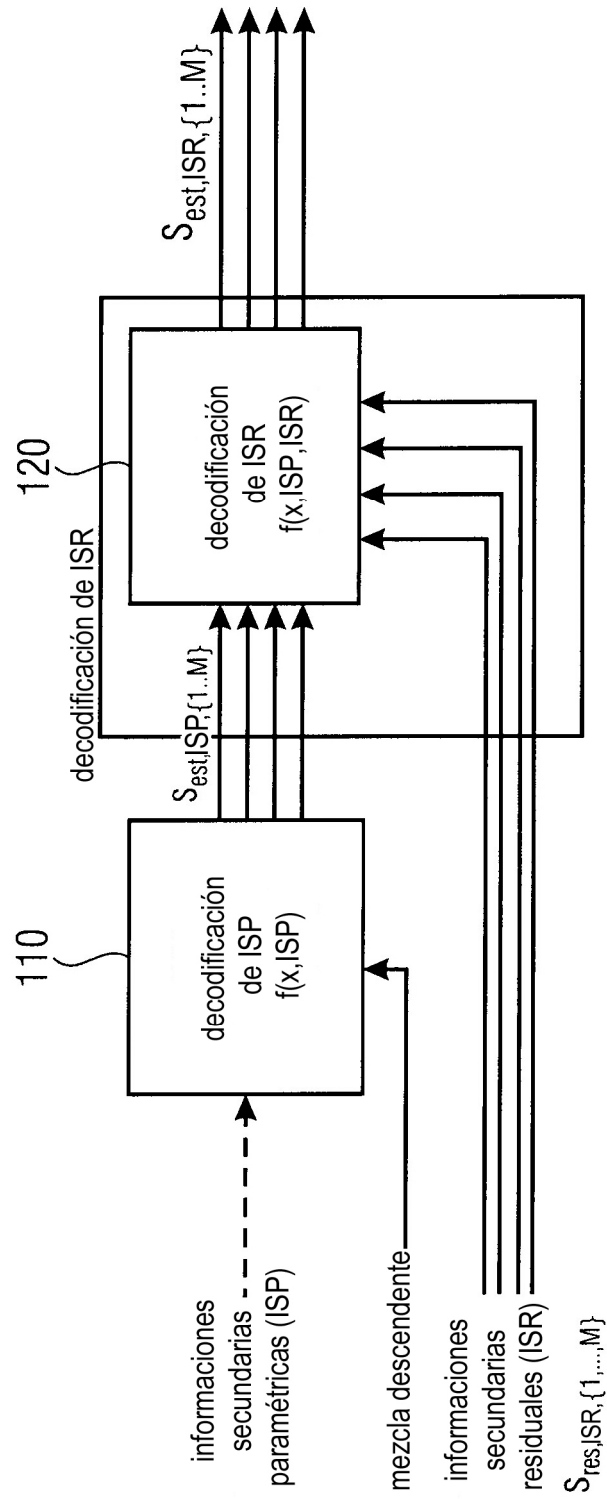


FIG 10

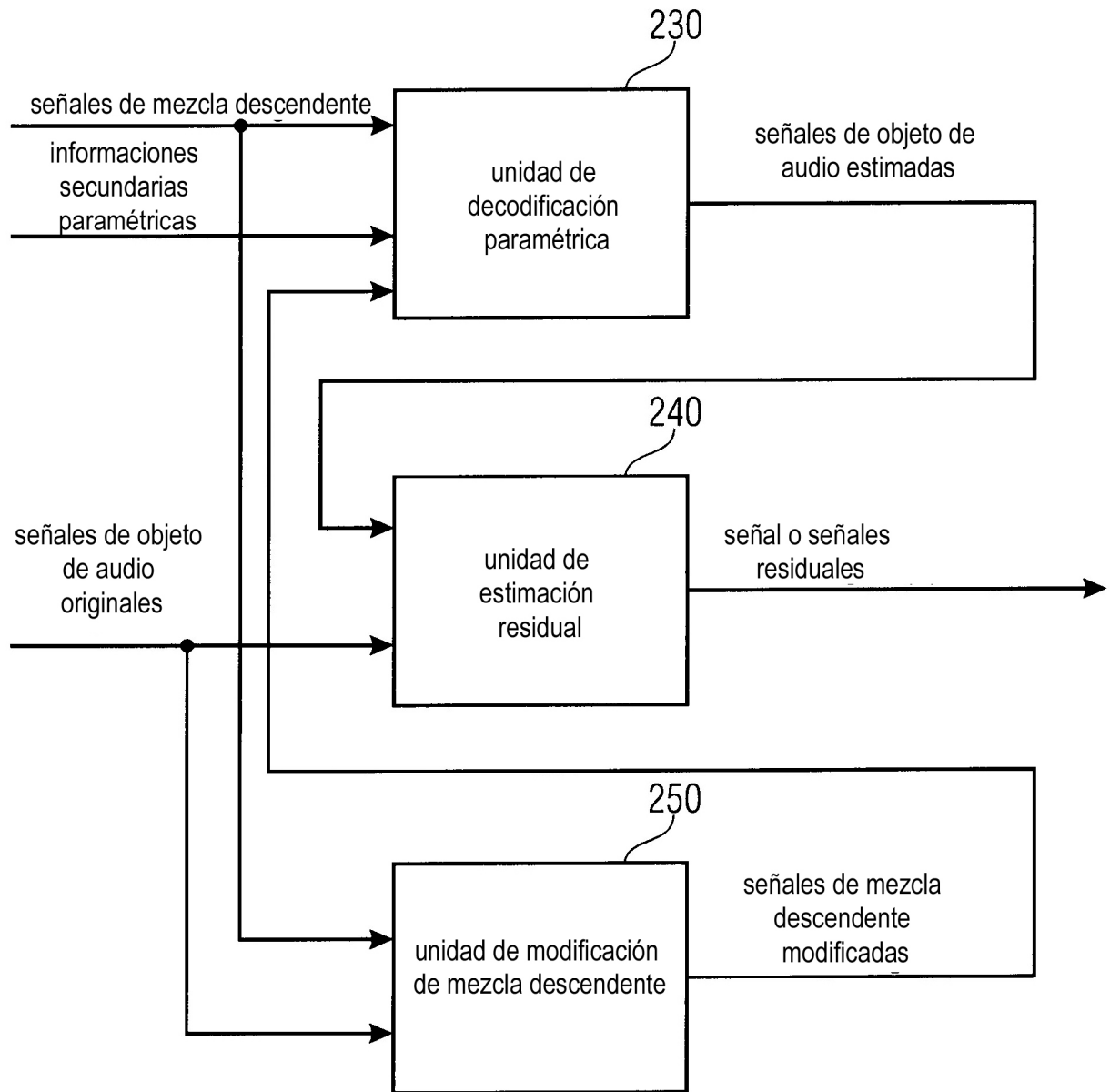


FIG 11

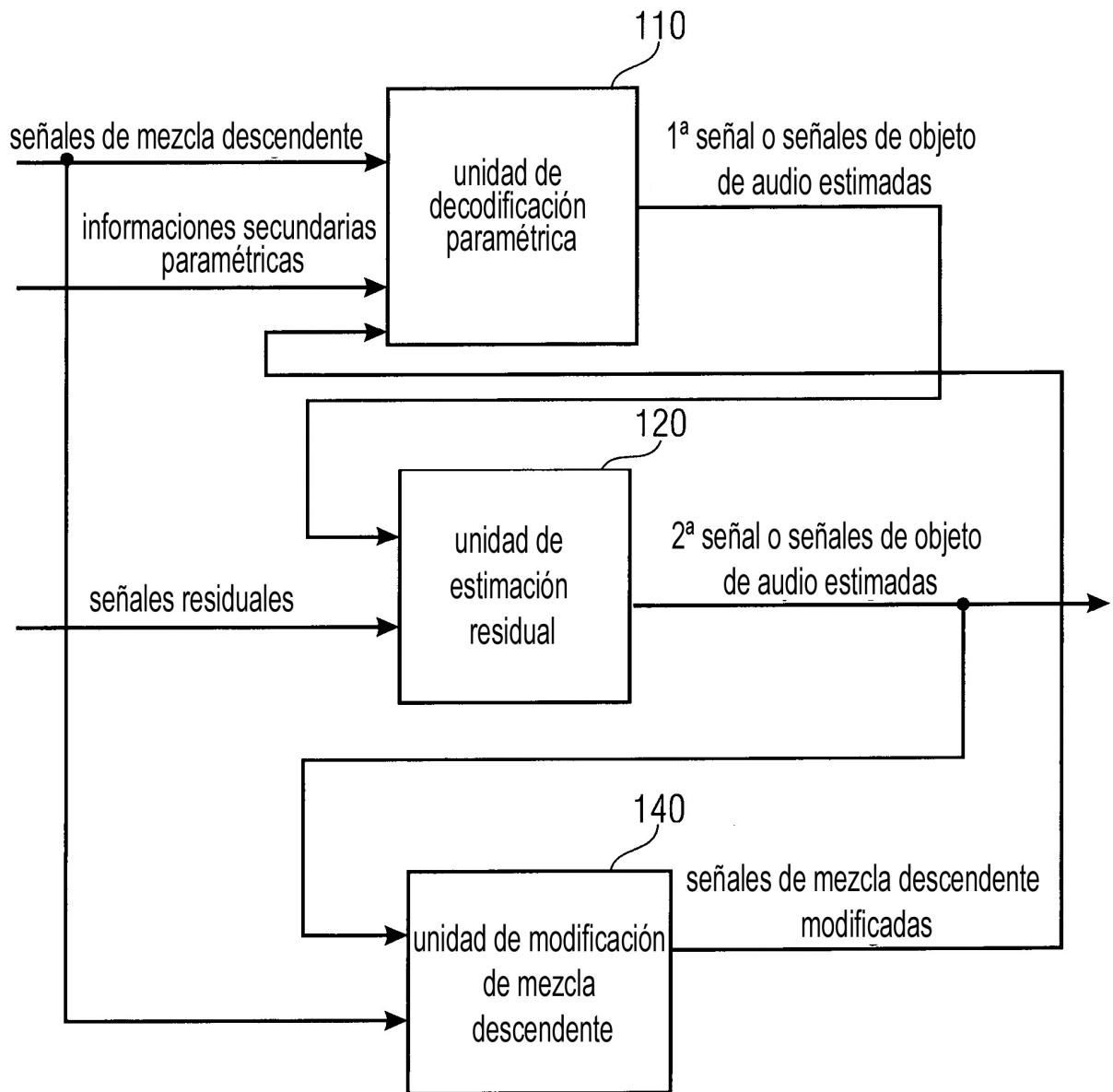


FIG 12

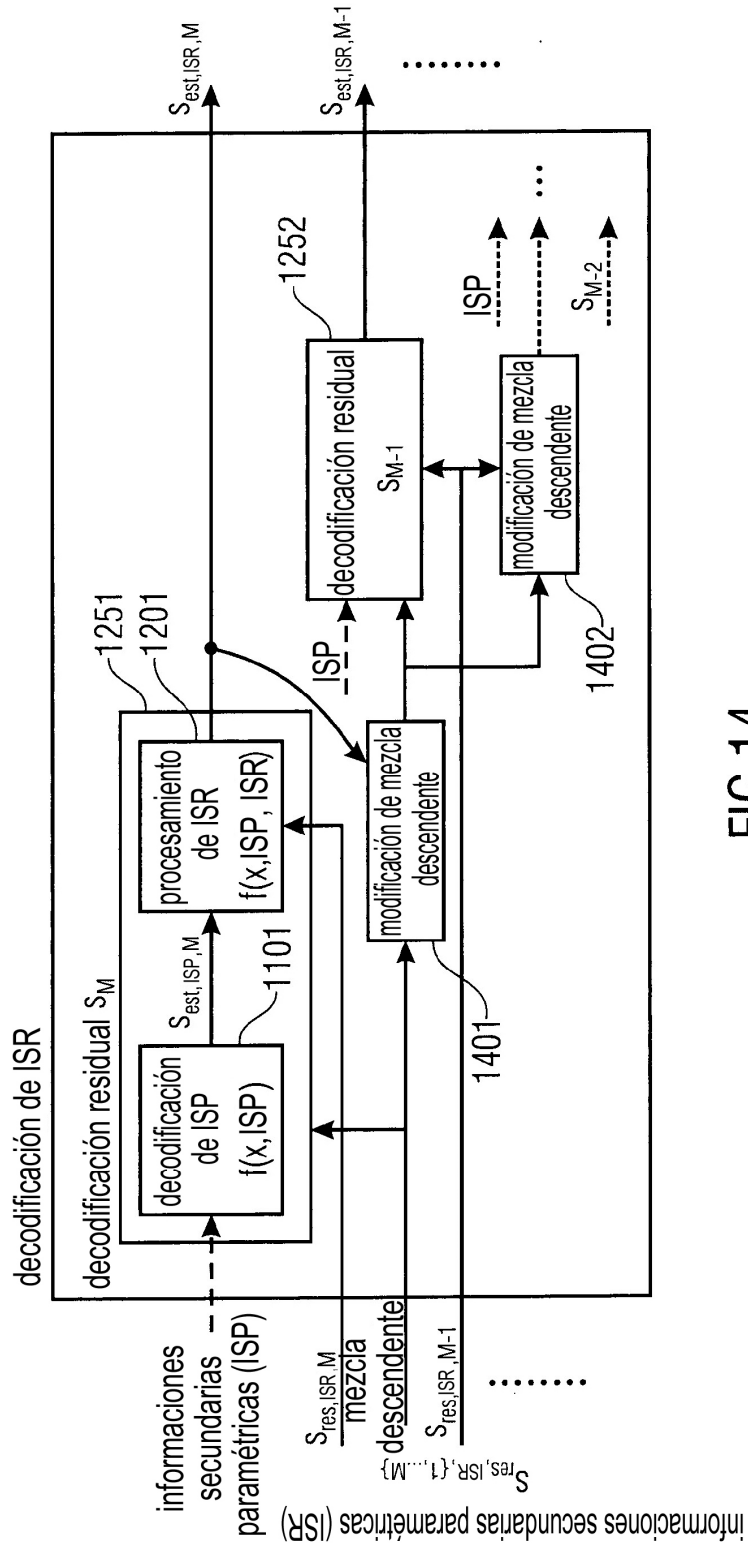


FIG 14

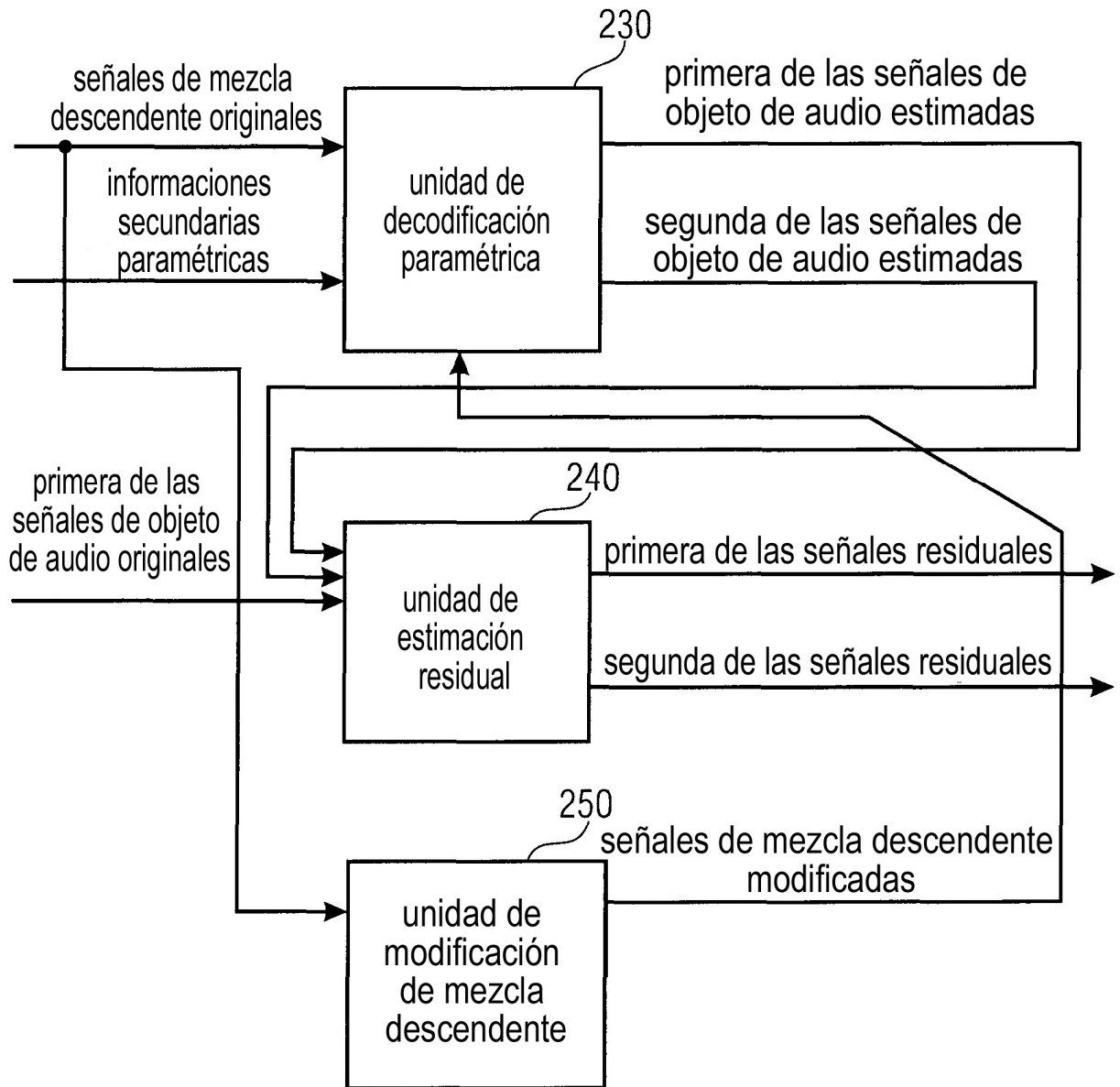


FIG 15

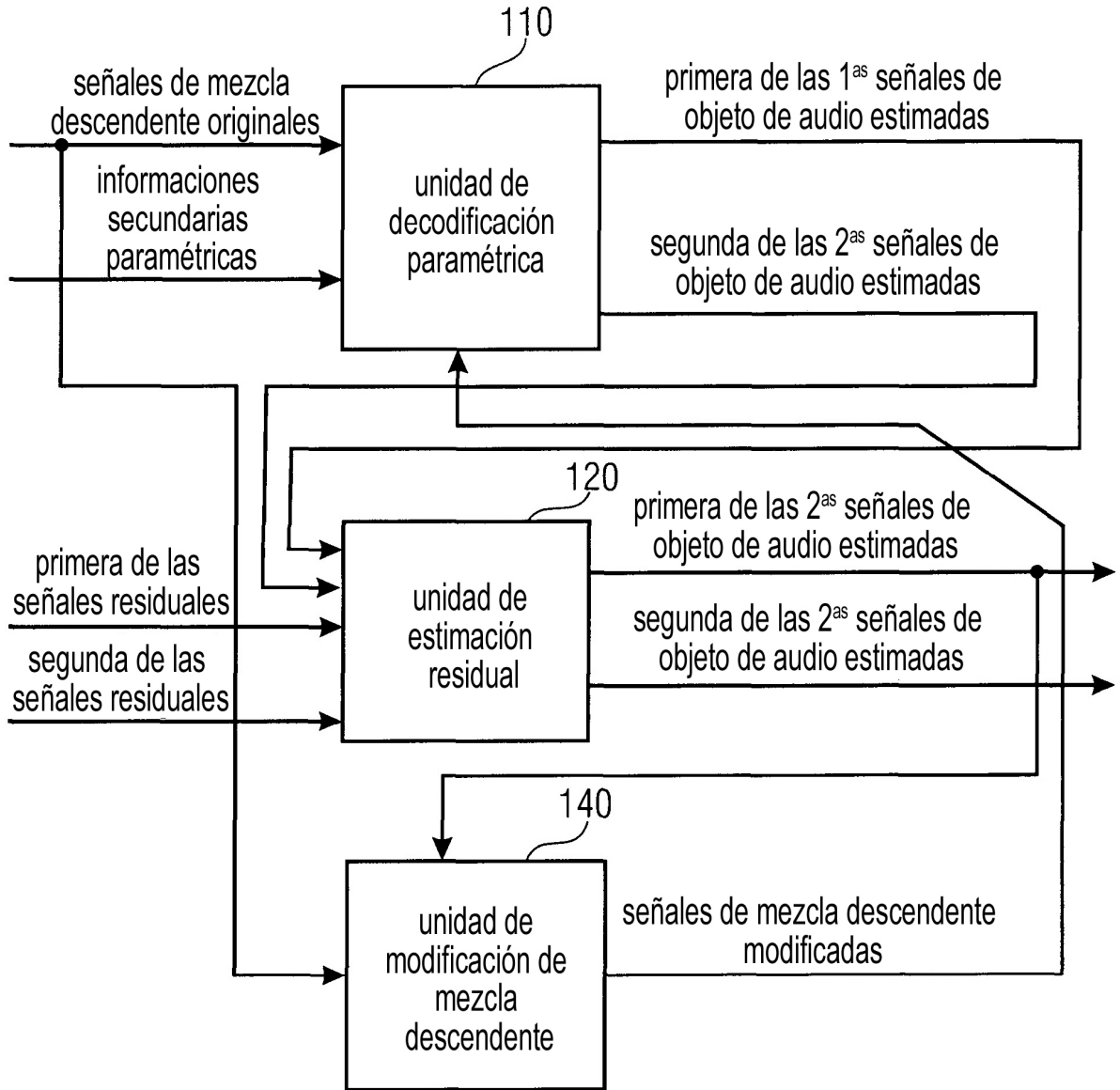


FIG 16