

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 001 434**

51 Int. Cl.:

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.10.2010** **E 21198132 (9)**

97 Fecha y número de publicación de la concesión europea: **27.11.2024** **EP 3996089**

54 Título: **Aparato, método y programa informático para proporcionar parámetros ajustados**

30 Prioridad:

16.10.2009 US 25229809 P

30.07.2010 EP 10171459

30.07.2010 US 36925610 P

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

05.03.2025

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.00%)**

**Hansastr. 27c
80686 München, DE**

72 Inventor/es:

**HERRE, JÜRGEN;
FALCH, CORNELIA y
TERENTIV, LEON**

74 Agente/Representante:

RUÍZ GALLEGOS, Natalia

ES 3 001 434 T3

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

DESCRIPCIÓN

Aparato, método y programa informático para proporcionar parámetros ajustados

5 Campo técnico

Una realización según la invención está relacionada con un aparato para proporcionar uno o más parámetros ajustados para una provisión de una representación de señal de mezcla ascendente basándose en una representación de señal de mezcla descendente y una información lateral paramétrica asociada con la representación de señal de mezcla descendente.

Otra realización según la invención está relacionada con un aparato para proporcionar una representación de señal de mezcla ascendente basándose en la representación de señal de mezcla descendente y la información lateral paramétrica.

Otra realización según la invención está relacionada con un método para proporcionar uno o más parámetros ajustados para una provisión de una representación de señal de mezcla ascendente basándose en una representación de señal de mezcla descendente y una información lateral paramétrica asociada con una representación de señal de mezcla descendente.

Otra realización según la invención está relacionada con un programa informático para realizar dicho método.

Algunas realizaciones según la invención están relacionadas con un esquema de limitación de parámetro para control de distorsión en MPEG SAOC.

25 Antecedentes de la invención

En la técnica anterior del procesamiento de audio, la transmisión de audio y el almacenamiento de audio, hay un creciente deseo de manejar contenidos multicanal para mejorar la impresión auditiva. El uso de contenido de audio multicanal produce mejoras significativas para el usuario. Por ejemplo, se puede obtener una impresión auditiva tridimensional, lo cual produce una mejor satisfacción del usuario en aplicaciones de entretenimiento. Sin embargo, los contenidos de audio multicanal también son útiles para entornos profesionales, por ejemplo, en aplicaciones de conferencias telefónicas, porque se puede mejorar la inteligibilidad del locutor usando una reproducción de audio multicanal.

Sin embargo, también es deseable tener una buena solución de compromiso entre calidad de audio y requisitos de tasa de bits transmitidos para evitar una excesiva carga de recursos causada por aplicaciones multicanal.

Recientemente se han propuesto técnicas paramétricas para el almacenamiento y/o la transmisión eficiente en cuanto a tasa de bits de escenas de audio que contienen múltiples objetos de audio, por ejemplo, codificación de indicación binaural (tipo I) (véase, por ejemplo, la referencia [1]), codificación de fuente conjunta (véase, por ejemplo, la referencia [2]), y codificación de objeto de audio espacial MPEG (SAOC) (véanse, por ejemplo, las referencias [3], [4] y [5]).

En combinación con la interactividad de usuario en el lado receptor, tales técnicas pueden conducir a una baja calidad de audio de las señales de salida si se realiza la renderización de objeto extremo (véase, por ejemplo, la referencia [6]).

Estas técnicas pretenden reconstruir perceptualmente la escena de audio de salida deseada más que mediante una coincidencia de forma de onda.

La figura 8 muestra una vista general de un sistema de este tipo (en el presente documento: MPEG SAOC). El sistema MPEG SAOC 800 mostrado en la figura 8 comprende un codificador de SAOC 810 y un decodificador de SAOC 820. El codificador de SAOC 810 recibe una pluralidad de señales de objeto x_1 a x_N , las cuales pueden representarse, por ejemplo, como señales en el dominio del tiempo o como señales en el dominio de tiempo-frecuencia (por ejemplo, en la forma de un conjunto de coeficientes de transformada de una transformada de tipo Fourier, o en la forma de señales de subbanda QMF). El codificador de SAOC 810 normalmente también recibe coeficientes de mezcla descendente d_1 a d_N , los cuales se asocian con las señales de objeto x_1 a x_N . Conjuntos separados de coeficientes de mezcla descendente pueden estar disponibles para cada canal de la señal de mezcla descendente. El codificador de SAOC 810 normalmente está configurado para obtener un canal de la señal de mezcla descendente combinando las señales de objeto x_1 a x_N según los coeficientes de mezcla descendente asociados d_1 a d_N . Normalmente, hay menos canales de mezcla descendente que señales de objeto x_1 a x_N . Para permitir (al menos aproximadamente) una separación (o tratamiento separado) de las señales de objeto en el lado del decodificador de SAOC 820, el codificador de SAOC 810 proporciona tanto la una o más señales de mezcla descendente (designadas como canales de mezcla descendente) 812 como una información lateral 814. La información lateral 814 describe características de las señales

de objeto x_1 a x_N , para permitir un procesamiento específico de objeto del lado del decodificador.

El decodificador de SAOC 820 está configurado para recibir tanto la una o más señales de mezcla descendente 812 como la información lateral 814. También el decodificador de SAOC 820 normalmente está configurado para recibir una información de interacción de usuario y/o una información de control de usuario 822, la cual describe una configuración de renderización deseada. Por ejemplo, la información de interacción de usuario/información de control de usuario 822 puede describir una configuración de altavoz y la ubicación espacial deseada de los objetos, los cuales proporcionan las señales de objeto x_1 a x_N .

El decodificador de SAOC 820 está configurado para proporcionar, por ejemplo, una pluralidad de señales de canal de mezcla ascendente decodificadas $\hat{y}_1$ a $\hat{y}_M$. Las señales de canal de mezcla ascendente pueden estar asociadas, por ejemplo, con altavoces individuales de una disposición de renderización de múltiples altavoces. El decodificador de SAOC 820 puede comprender, por ejemplo, un separador de objeto 820a, el cual está configurado para reconstruir, al menos aproximadamente, las señales de objeto x_1 a x_N basándose en la una o más señales de mezcla descendente 812 y la información lateral 814, obteniendo de ese modo señales de objeto reconstruidas 820b. Sin embargo, las señales de objeto reconstruidas 820b pueden desviarse un poco de las señales de objeto originales x_1 a x_N , por ejemplo, porque la información lateral 814 no es suficiente para una reconstrucción perfecta debido a las restricciones de tasa de bits. El decodificador de SAOC 820 puede comprender además un mezclador 820c, el cual puede estar configurado para recibir las señales de objeto reconstruidas 820b y la información de interacción de usuario/información de control de usuario 822, y para proporcionar, basándose en ello, las señales de canal de mezcla ascendente $\hat{y}_1$ a $\hat{y}_M$. El mezclador 820c puede estar configurado para usar la información de interacción de usuario/información de control de usuario 822 para determinar la contribución de las señales de objeto reconstruidas individuales 820b a las señales de canal de mezcla ascendente $\hat{y}_1$ a $\hat{y}_M$. La información de interacción de usuario/información de control de usuario 822 puede comprender, por ejemplo, parámetros de renderización (también designados como coeficientes de renderización), los cuales determinan la contribución de las señales de objeto reconstruidas individuales 822 a las señales de canal de mezcla ascendente $\hat{y}_1$ a $\hat{y}_M$.

Sin embargo, debe observarse que, en muchas realizaciones, la separación de objetos, la cual está indicada mediante el separador de objetos 820a en la figura 8, y el mezclado, el cual está indicado mediante el mezclador 820c en la figura 8, se realizan en una única etapa. Con este propósito, se pueden calcular parámetros globales los cuales describen un mapeo directo de la una o más señales de mezcla descendente 812 sobre las señales de canal de mezcla ascendente $\hat{y}_1$ a $\hat{y}_M$. Estos parámetros pueden calcularse basándose en la información lateral y la información de interacción de usuario/información de control de usuario 820.

Haciendo referencia ahora a las figuras 9a, 9b y 9c, se describirán diferentes aparatos para obtener una representación de señal de mezcla ascendente basándose en una representación de señal de mezcla descendente e información lateral relacionada con el objeto. Debe observarse que la información lateral relacionada con el objeto es un ejemplo de una información lateral asociada con la señal de mezcla descendente. La figura 9a muestra un diagrama de bloques esquemático de un sistema de SAOC MPEG 900 que comprende un decodificador de SAOC 920. El decodificador de SAOC 920 comprende, como bloques funcionales separados, un decodificador de objeto 922 y un mezclador/renderizador 926. El decodificador de objeto 922 proporciona una pluralidad de señales de objeto reconstruidas 924 dependiendo de la representación de señal de mezcla descendente (por ejemplo, en forma de una o más señales de mezcla descendente representadas en el dominio del tiempo o en el dominio de tiempo-frecuencia) e información lateral relacionada con el objeto (por ejemplo, en forma de metadatos de objeto). El mezclador/renderizador 926 recibe las señales de objeto reconstruidas 924 asociadas con una pluralidad de N objetos y proporciona, basándose en ello y en la información de renderización, una o más señales de canal de mezcla ascendente 928. En el decodificador de SAOC 920, la extracción de las señales de objeto 924 se realiza de forma separada del mezclado/renderización, lo cual permite una separación de la funcionalidad de decodificación de objeto, de la funcionalidad de mezclado/renderización, pero produce una complejidad de cálculo relativamente alta.

Haciendo referencia ahora a la figura 9b, se comentará brevemente otro sistema de SAOC MPEG 930, el cual comprende un decodificador de SAOC 950. El decodificador de SAOC 950 proporciona una pluralidad de señales de canal de mezcla ascendente 958 dependiendo de una representación de señal de mezcla descendente (por ejemplo, en forma de una o más señales de mezcla descendente) y una información lateral relacionada con el objeto (por ejemplo, en forma de metadatos de objeto). El decodificador de SAOC 950 comprende un decodificador de objeto y mezclador/renderizador combinados, lo cual está configurado para obtener las señales de canal de mezcla ascendente 958 en un proceso de mezclado colectivo sin una separación de la decodificación de objeto y el mezclado/renderización, en el que los parámetros para dicho proceso de mezcla ascendente colectivo dependen tanto de la información lateral relacionada con el objeto como de la información de renderización. El proceso de mezcla ascendente colectivo depende también de la información de mezcla descendente, la cual se considera parte de la información lateral relacionada con el objeto.

Para resumir lo anterior, la provisión de las señales de canal de mezcla ascendente 928, 958 puede realizarse en un proceso de una única etapa o en un proceso de dos etapas.

Haciendo referencia ahora a la figura 9c, se describirá un sistema de SAOC MPEG 960. El sistema de SAOC 960 comprende un transcodificador de SAOC a MPEG Surround 980, en lugar de un decodificador de SAOC.

- 5 El transcodificador de SAOC a MPEG Surround comprende un transcodificador de información lateral 982, el cual está configurado para recibir la información lateral relacionada con el objeto (por ejemplo, en forma de metadatos de objeto) y, opcionalmente, información sobre la una o más señales de mezcla descendente y la información de renderización. El transcodificador de información lateral también está configurado para proporcionar una información lateral de MPEG Surround (por ejemplo, en forma de un flujo de bits de MPEG Surround) basándose en un dato recibido.
- 10 Consiguientemente, el transcodificador de información lateral 982 está configurado para transformar una información lateral (paramétrica) relacionada con el objeto, la cual se recibe del codificador de objeto, en una información lateral (paramétrica) relacionada con el canal, teniendo en cuenta la información de renderización y, opcionalmente, la información acerca del contenido de la una o más señales de mezcla descendente.
- 15 Opcionalmente, el transcodificador de SAOC a MPEG Surround 980 puede estar configurado para manipular la una o más señales de mezcla descendente, descritas, por ejemplo, por la representación de señal de mezcla descendente, para obtener una representación de señal de mezcla descendente manipulada 988. Sin embargo, el manipulador de señal de mezcla descendente 986 puede omitirse, de manera que la representación de señal de mezcla descendente de salida 988 del transcodificador de SAOC a MPEG Surround 980 sea idéntica a la representación de señal de mezcla descendente de entrada del transcodificador de SAOC a MPEG Surround. El manipulador de señal de mezcla descendente 986 puede usarse, por ejemplo, si la información lateral de MPEG Surround relacionada con el canal 984 no permitiera proporcionar una impresión auditiva deseada basándose en la representación de señal de mezcla descendente de entrada del transcodificador de SAOC a MPEG Surround 980, lo cual podría ser el caso en algunas constelaciones de renderización.
- 20 Consiguientemente, el transcodificador de SAOC a MPEG Surround 980 proporciona la representación de señal de mezcla descendente 988 y el flujo de bits MPEG Surround 984 de manera que una pluralidad de señales de canal de mezcla ascendente, las cuales representan los objetos de audio según la información de renderización introducida al transcodificador de SAOC a MPEG Surround 980 puede generarse usando un decodificador MPEG Surround que recibe el flujo de bits MPEG Surround 984 y la representación de señal de mezcla descendente 988.
- 25

- Para resumir lo anterior, se pueden usar diferentes conceptos para decodificar señales de audio codificadas por SAOC. En algunos casos, se usa un decodificador de SAOC, el cual proporciona señales de canal de mezcla ascendente (por ejemplo, señales de canal de mezcla ascendente 928, 958) dependiendo de la representación de señal de mezcla descendente y la información lateral paramétrica relacionada con el objeto. Ejemplos para este concepto se pueden ver en las figuras 9a y 9b. Alternativamente, la información de audio codificada por SAOC puede transcodificarse para obtener una representación de señal de mezcla descendente (por ejemplo, una representación de señal de mezcla descendente 988) y una información lateral relacionada con el canal (por ejemplo, el flujo de bits MPEG Surround relacionado con el canal 984), que puede usarse por un decodificador MPEG Surround para proporcionar las señales de canal de mezcla ascendente deseadas.
- 30

- En el sistema de SAOC MPEG 800, del cual se da una vista general del sistema en la figura 8, el procesamiento general se lleva a cabo de una manera selectiva en frecuencia y puede describirse como sigue dentro de cada banda de frecuencia:
- 45

- N señales de objeto de audio de entrada x_1 a x_N se someten a mezcla descendente como parte del procesamiento de codificador de SAOC. Para una mezcla descendente mono, los coeficientes de mezcla descendente se indican d_1 a d_N . Además, el codificador de SAOC 810, extrae información lateral 814 que describe las características de los objetos de audio de entrada. Para SAOC MPEG las relaciones de las potencias de objetos entre sí son la forma más básica de una información lateral de este tipo.
- 50

- La(s) señal(es) de mezcla descendente 812 y la información lateral 814 se transmiten y/o almacenan. Con este fin, la señal de audio de mezcla descendente puede comprimirse usando codificadores de audio perceptuales conocidos tales como MPEG-1 capa II o III (también conocido como ".mp3"), codificación de audio avanzada MPEG (AAC), o cualquier otro codificador de audio.
- 55

- En el extremo de recepción, el decodificador de SAOC 820 trata conceptualmente de recuperar la señal de objeto original ("separación de objeto") usando la información lateral transmitida 814 (y, naturalmente, la una o más señales de mezcla descendente 812). Estas señales de objeto aproximadas (también designadas como señales de objeto reconstruidas 820b) se mezclan entonces formando una escena objetivo representada por M canales de salida de audio (los cuales pueden estar representados, por ejemplo, por las señales de canal de mezcla ascendente $\hat{y}_1$ a $\hat{y}_M$) usando una matriz de renderización. Para una salida mono, los coeficientes de la matriz de renderización vienen dados por r_1 a r_N .
- 60

• Efectivamente, la separación de las señales de objetos se realiza raramente (o incluso nunca se realiza), ya que tanto la etapa de separación (indicada por el separador de objetos 820^a) como la etapa de mezclado (indicada por el mezclador 820c) se combinan formando una única etapa de transcodificación, la cual con frecuencia da como resultado una enorme reducción de complejidad de cálculo.

5

Se ha encontrado que un esquema de este tipo es muy eficiente, tanto en términos de tasa de bits de transmisión (solo es necesario transmitir unos pocos canales de mezcla descendente más alguna información lateral en lugar de N señales de audio de objetos discretas o un sistema discreto) como en complejidad de cálculo (la complejidad de procesamiento se refiere principalmente al número de canales de salida más que al número de objetos de audio).

10

Ventajas adicionales para el usuario en el extremo de recepción incluyen la libertad de elegir una configuración de renderización de su elección (mono, estéreo, envolvente, reproducción de auricular virtualizada, y así sucesivamente) y el rasgo de interactividad del usuario: la matriz de renderización, y por ende la escena de salida, puede fijarse y cambiarse de forma interactiva por el usuario según el deseo, la preferencia personal u otros criterios. Por ejemplo, es posible ubicar los locutores de un grupo juntos en un área espacial para maximizar la discriminación de otros locutores restantes. Esta interactividad se logra proporcionando una interfaz de usuario de decodificador.

15

Para cada objeto de sonido transmitido, se puede ajustar su nivel relativo y (para renderización no mono) posición espacial de renderización. Esto puede producirse en tiempo real conforme el usuario cambia la posición de los controles deslizantes de la interfaz gráfica de usuario (GUI) asociada (por ejemplo, nivel de objeto = +5 dB, posición de objeto = -30 grados).

20

Sin embargo, se ha encontrado que la elección de parámetros del lado del decodificador para la provisión de la representación de señal de mezcla ascendente (por ejemplo, las señales de canal de mezcla ascendente $\hat{y}_1$ a $\hat{y}_M$) produce degradaciones audibles en algunos casos. El documento "Technical provisions for limiting perceptible distortions in SAOC" de Jürgen Herre *et. al.*, 90. MPEG MEETING, octubre de 2009, Xi'an, China, divulga medios para limitar la cantidad de distorsión perceptible del sistema de SAOC para escenarios de renderización exigentes para garantizar una experiencia de usuario satisfactoria.

25

En vista de esta situación, el objetivo de la presente invención es crear un concepto que permita reducir o incluso evitar distorsión audible cuando se proporciona una representación de señal de mezcla ascendente (por ejemplo, en forma de señales de canal de mezcla ascendente $\hat{y}_1$ a $\hat{y}_M$).

30

Sumario de la invención

35

La invención se define en las reivindicaciones independientes. Las realizaciones preferidas se presentan en las reivindicaciones dependientes.

Breve descripción de las figuras

40

La figura 1 muestra un diagrama de bloques esquemático de un aparato para proporcionar uno o más parámetros ajustados, según una realización de la invención;

la figura 2 muestra un diagrama de bloques esquemático de un aparato para proporcionar una representación de señal de mezcla ascendente, según una realización de la invención;

45

la figura 3 muestra un diagrama de bloques esquemático de un aparato para proporcionar una representación de señal de mezcla ascendente, según otra realización de la invención;

la figura 4 muestra una representación esquemática de esquemas de limitación de parámetros usando un control indirecto y un control directo;

50

la figura 5a muestra una tabla que representa condiciones de prueba de audición;

la figura 5b muestra una tabla que representa objetos de audio de prueba de audición;

55

la figura 6 muestra una tabla que representa condiciones de renderización extremas sometidas a prueba;

la figura 7 muestra una representación gráfica de resultados de prueba de audición MUSHRA para diferentes esquemas de limitación de parámetros (PLS);

60

la figura 8 muestra un diagrama de bloques esquemático de un sistema de SAOC MPEG de referencia;

la figura 9a muestra un diagrama de bloques esquemático de un sistema de SAOC de referencia usando un decodificador y mezclador separados;

la figura 9b muestra un diagrama de bloques esquemático de un sistema de SAOC de referencia usando un decodificador y mezclador integrados;

- 5 la figura 9c muestra un diagrama de bloques esquemático de un sistema de SAOC de referencia usando un transcodificador de SAOC a MPEG; y

la figura 10 muestra una tabla que describe qué coeficientes de transcodificación pueden modificarse por el esquema de limitación de parámetros propuesto.

10

Descripción detallada de las realizaciones

1. Aparato para proporcionar uno o más parámetros ajustados, según la figura 1

- 15 A continuación, se describirá un aparato para proporcionar uno o más parámetros ajustados para una provisión de una representación de señal de mezcla ascendente basándose en una representación de señal de mezcla descendente y una información lateral paramétrica asociada con la representación de señal de mezcla descendente. La figura 1 muestra un diagrama de bloques esquemático de un aparato 100 de este tipo.

- 20 El aparato 100 está configurado para recibir uno o más parámetros de entrada 110 y para proporcionar, basándose en ello, uno o más parámetros ajustados 120. El aparato 100 comprende un ajustador de parámetros 130 el cual está configurado para recibir el uno o más parámetros de entrada 110 y para proporcionar, basándose en ello, el uno o más parámetros ajustados 120. El ajustador de parámetros 130 está configurado para proporcionar el uno o más parámetros ajustados 120 dependiendo de un valor promedio 132 de una pluralidad de valores de parámetros de entrada, de manera que una distorsión de una representación de señal de mezcla ascendente causada por el uso de parámetros no óptimos (por ejemplo, el uno o más parámetros de entrada 110) se reduce al menos para parámetros de entrada (por ejemplo, parámetros de entrada 110) que se desvían de los parámetros óptimos en más que una desviación predeterminada. Por ejemplo, el ajustador de parámetros 130 puede tener el efecto de que el uno o más parámetros ajustados 120 están “más cerca” (en el sentido de causar distorsiones más pequeñas) a parámetros óptimos (los cuales darían como resultado una representación de señal de mezcla ascendente sin distorsión) que el uno o más parámetros de entrada 110.

- Con este propósito, el ajustador de parámetros 130 implementa un cálculo de valor promedio, para obtener el valor promedio 132 (por ejemplo, como un promedio temporal o un promedio interobjeto) de un conjunto de parámetros de entrada relacionados 110 (por ejemplo, parámetros de entrada asociados con un intervalo de tiempo común, o parámetros de entrada del mismo tipo de parámetro asociado con diferentes instancias de tiempo). Con respecto a la operación del aparato 100, debe observarse que la provisión del uno o más parámetros ajustados 120 basándose en el uno o más parámetros de entrada 110 se hace dependiendo del valor promedio 132, porque se ha encontrado que el valor promedio 132 es una cantidad significativa para ajustar los parámetros. En particular, se ha encontrado que parámetros moderados (con respecto al valor promedio) normalmente producen distorsiones moderadas.

A continuación, se describirán detalles adicionales.

2. Aparato para proporcionar una representación de señal de mezcla ascendente, según la figura 2

- 45 A continuación, se describirá un aparato para proporcionar una representación de señal de mezcla ascendente según la figura 2. La figura 2 muestra un diagrama de bloques esquemático de un aparato 200 de este tipo, el cual puede considerarse un decodificador de señal de audio. Por ejemplo, el aparato 200 puede comprender la funcionalidad de un decodificador de SAOC o un transcodificador de SAOC.

- 50 El aparato 200 está configurado para recibir una representación de señal de mezcla descendente 210 y una información lateral paramétrica 212. Además, el aparato 200 está configurado para recibir parámetros de renderización especificados por el usuario 214. El aparato está configurado para proporcionar una representación de señal de mezcla ascendente 220.

55

- La representación de señal de mezcla descendente 210 puede ser, por ejemplo, una representación de una señal de audio monocanal o de una señal de audio de dos canales. La representación de señal de mezcla descendente 210 puede ser, por ejemplo, una representación del dominio del tiempo o una representación codificada. En algunas realizaciones, la representación de señal de mezcla descendente 210 puede ser una representación del dominio de tiempo-frecuencia, en la cual el uno o más canales de la representación de señal de mezcla descendente 210 están representados por conjuntos posteriores de valores espectrales.

60

La representación de señal de mezcla ascendente 220 puede ser, por ejemplo, una representación de canales de audio individuales, por ejemplo, en la forma de una representación del dominio del tiempo o una representación del

dominio de tiempo–frecuencia. Alternativamente, la representación de señal de mezcla ascendente 220 puede ser una representación codificada, que comprende tanto una representación de señal de mezcla descendente como una información lateral relacionada con el canal, por ejemplo, una información lateral de MPEG Surround.

- 5 Los parámetros de renderización especificados por el usuario 214 pueden proporcionarse en la forma de entradas de matriz de renderización que describen contribuciones deseadas de una pluralidad de objetos de audio al uno o más canales de la representación de señal de mezcla ascendente 220. Alternativamente, los parámetros de renderización especificados por el usuario 214 pueden proporcionarse de cualquier otra forma apropiada, por ejemplo, especificando una posición de renderización deseada y un volumen de renderización de los objetos de audio.

10 El aparato 200 comprende un procesador de señal 230, el cual está configurado para proporcionar la representación de señal de mezcla ascendente 220 basándose en la representación de señal de mezcla descendente 210 y la información lateral paramétrica 212. El procesador de señal 230 comprende una funcionalidad de remezclado 232 para proporcionar la representación de señal de mezcla ascendente 220 basándose en la representación de señal de mezcla descendente 210. Por ejemplo, la funcionalidad de remezclado 232 puede estar configurada para combinar linealmente una pluralidad de canales de la representación de señal de mezcla descendente 212 para obtener el uno o más canales de la representación de señal de mezcla ascendente 220. En este remezclado, se pueden determinar las contribuciones de los canales de la representación de señal de mezcla descendente 210 a los canales de la representación de señal de mezcla ascendente 220 mediante los elementos de matriz de mezcla de una matriz de mezcla **G**, en el que una primera dimensión (por ejemplo, un número de filas) de la matriz de mezcla **G** puede determinarse por el número de canales de la representación de señal de mezcla ascendente 220, y en el que una segunda dimensión (por ejemplo, un número de columnas) de la matriz de mezcla **G** puede determinarse por un número de canales de la representación de señal de mezcla descendente 210.

- 25 Por ejemplo, el proceso de remezclado 232 puede usarse para proporcionar uno o más vectores que comprenden valores espectrales asociados con uno o más canales de la representación de señal de mezcla ascendente 220 multiplicando uno o más vectores que comprenden valores espectrales de uno o más canales de la representación de señal de mezcla descendente 210 con la matriz de mezcla **G**.

30 El procesador de señal 230 también puede comprender un cálculo de parámetro de mezclado 236 el cual proporciona la matriz de mezcla **G** (o de forma equivalente, los elementos de la misma). Los elementos de la matriz de mezcla se determinan dependiendo de la información lateral paramétrica 212 y los parámetros de renderización modificados 252 mediante el cálculo de parámetro de mezclado 236. Los elementos de la matriz de mezcla de la matriz de mezcla **G** se proporcionan, por ejemplo, de manera que el uno o más canales de la representación de señal de mezcla ascendente 220 describen objetos de audio, los cuales están representados por el uno o más canales de la representación de señal de mezcla descendente 210, según los parámetros de renderización modificados 252. Con este propósito, la información lateral paramétrica 212 se evalúa por el cálculo de parámetro de mezclado 236, en el que la información lateral paramétrica 212 comprende, por ejemplo, una información de diferencia de nivel de objeto OLD, una información de correlación interobjeto IOC, una información de ganancia de mezcla descendente DMG y (opcionalmente) una información de diferencia de nivel de canal de mezcla descendente DCLD. La información de diferencia de nivel de objeto puede describir, por ejemplo, de una manera a modo de banda de frecuencia, diferencias de nivel entre una pluralidad de objetos de audio. De manera similar, la información de correlación interobjeto puede describir, por ejemplo, de una manera a modo de banda de frecuencia, correlaciones entre una pluralidad de objetos de audio. La información de ganancia de mezcla descendente y la información (opcional) de diferencia de nivel de canal de mezcla descendente puede describir la mezcla descendente, la cual se realiza para combinar señales de objeto de audio de una pluralidad de objetos de audio en el uno o más canales de la representación de señal de mezcla descendente, en el que normalmente hay más objetos de audio que canales de la representación de señal de mezcla descendente 210.

- 50 Consiguientemente, el cálculo de parámetro de mezclado 236 puede evaluar cómo deben elegirse los elementos de matriz de mezcla para obtener una representación de señal de mezcla ascendente 220 que comprende propiedades estadísticas esperadas basándose en la información lateral paramétrica 212 y los parámetros de renderización modificados 252.

55 El procesador de señal 230 opcionalmente puede comprender una modificación de información lateral o transformación de información lateral 240, la cual está configurada para recibir la información lateral paramétrica 212 y para proporcionar una información lateral modificada (por ejemplo, una información lateral de MPEG Surround), de manera que la información lateral modificada y la representación de señal de mezcla descendente remezclada asociada proporcionada por el proceso de remezclado 232 describen una escena de audio deseada.

60 Para resumir, el procesador de señal 230 puede cumplir, por ejemplo, la funcionalidad del decodificador de SAOC 820, en el que la representación de señal de mezcla descendente 210 desempeña el papel de la una o más señales de mezcla descendente 812, en el que la información lateral paramétrica 212 desempeña el papel de la información lateral 814, y en el que la representación de señal de mezcla ascendente 220 es equivalente a las señales de canal

de salida $\hat{y}_1$ a $\hat{y}_M$.

Alternativamente, el procesador de señal 230 puede comprender la funcionalidad del decodificador y mezclador separados 920, en el que la representación de señal de mezcla descendente 210 puede desempeñar el papel de la una o más señales de mezcla descendente, en el que la información lateral paramétrica 212 puede desempeñar el papel de los metadatos de objeto, y en el que la representación de señal de mezcla ascendente 220 puede desempeñar el papel de la una o más señales de canal de salida 928.

Alternativamente, el procesador de señal 230 puede comprender la funcionalidad del decodificador y mezclador integrados 950, en el que la representación de señal de mezcla descendente 210 puede desempeñar el papel de la una o más señales de mezcla descendente, en el que la información lateral paramétrica 212 puede desempeñar el papel de metadatos de objeto, y en el que la representación de señal de mezcla ascendente 220 puede desempeñar el papel de la una o más señales de canal de salida 958.

Alternativamente, el procesador de señal 230 puede comprender la funcionalidad del transcodificador de SAOC a MPEG Surround 980, en el que la representación de señal de mezcla descendente 210 puede desempeñar el papel de la una o más señales de mezcla descendente, en el que la información lateral paramétrica 212 puede desempeñar el papel de los metadatos de objeto, y en el que la representación de señal de mezcla ascendente puede ser equivalente a la una o más señales de mezcla descendente 988 cuando se toma en combinación con el flujo de bits MPEG Surround 984.

En cualquier caso, los parámetros de renderización modificados 252 pueden desempeñar el papel de la información de interacción/control de usuario 822 o de la información de renderización.

El aparato 200 también comprende un aparato 250 para proporcionar parámetros de renderización ajustados. El aparato 250 para proporcionar los parámetros de renderización ajustados recibe los parámetros de renderización especificados por el usuario 214 y proporciona, basándose en ellos, los parámetros de renderización modificados 252. El aparato 250 normalmente está configurado para calcular un valor promedio sobre una pluralidad de parámetros de renderización especificados por el usuario asociados con diferentes objetos de audio, para obtener un valor promedio. También, el aparato 250 está configurado para realizar una limitación de parámetro de renderización dependiendo de un valor promedio, para obtener los parámetros de renderización modificados 252 limitando los parámetros de renderización especificados por el usuario 214. Normalmente se determina un intervalo de tolerancia, al cual se limitan los parámetros de renderización modificados 252 dependiendo del valor promedio, de manera que se evitan fuertes desviaciones de los parámetros de renderización modificados 252, del valor promedio, incluso si uno o más de los parámetros de renderización especificados por el usuario 214 comprende una desviación fuerte de este tipo del valor promedio. De esta manera, normalmente se evitan distorsiones excesivas dentro de la representación de señal de mezcla ascendente 220, porque los parámetros de renderización modificados 252, los cuales comprenden limitada desviación interobjeto, darán como resultado una representación de señal de mezcla ascendente con bajas distorsiones, mientras una gran diferencia entre parámetros de renderización asociados con diferentes objetos de audio normalmente daría como resultado artefactos audibles.

Debe observarse en este caso que el aparato 250 para proporcionar coeficientes de renderización ajustados puede comprender la misma funcionalidad global que el aparato 100 para proporcionar uno o más parámetros ajustados, en el que los parámetros de renderización especificados por el usuario 214 pueden desempeñar el papel de uno o más parámetros de entrada 110, y en el que los parámetros de renderización ajustados 252 pueden desempeñar el papel del uno o más parámetros ajustados 120.

A continuación, se comentarán detalles con respecto a la provisión de los parámetros de renderización modificados 252, haciendo referencia a la figura 4.

3. Aparato para proporcionar una representación de señal de mezcla ascendente, según la figura 3

A continuación, se describirá un aparato para proporcionar una representación de señal de mezcla ascendente según otra realización de la invención, haciendo referencia a la figura 3, la cual muestra un diagrama de bloques esquemático de un aparato 300 de este tipo.

El aparato 300 normalmente recibe el mismo tipo de señales de entrada y proporciona el mismo tipo de señales de salida que el aparato 200, de manera que en el presente documento se usan números de referencia idénticos para describir señales idénticas o equivalentes. Para resumir, el aparato 300 recibe una representación de señal de mezcla descendente 210, información lateral paramétrica 212 y parámetros de renderización especificados por el usuario 214, y el aparato 300 proporciona, basándose en ellos, una representación de señal de mezcla ascendente 220.

El aparato 300 comprende un procesador de señal 330, el cual puede ser substancialmente equivalente en la funcionalidad, al procesador de señal 230. El procesador de señal 330 comprende una funcionalidad de remezclado

332, la cual es idéntica a la funcionalidad de remezclado 232 del procesador de señal 230 en que proporciona señales de canal de audio remezcladas basándose en la representación de señal de mezcla descendente. Sin embargo, el remezclado 332 usa una matriz de mezcla ajustada, en lugar de una matriz de mezcla obtenida directamente de un cálculo de parámetro de mezclado.

5

El procesador de señal 330 también comprende un cálculo de parámetro de mezclado 336, el cual puede ser idéntico en funcionamiento al cálculo de parámetro de mezclado 236 del procesador de señal 230. Consiguientemente, el cálculo de parámetro de mezclado 336 recibe la información lateral paramétrica 212 y los parámetros de renderización especificados por el usuario 214, y proporciona, basándose en ellos, una matriz de mezcla **G** (o de forma equivalente, elementos de matriz de mezcla de la matriz de mezcla **G**, los cuales también se designan con 337).

10

El procesador de señal 330 opcionalmente también comprende una modificación de información lateral 338, cuya funcionalidad es idéntica a la modificación de información lateral 240.

15

Además, el aparato 300 comprende un aparato 350 para proporcionar elementos de matriz de mezcla ajustados. El aparato 350 puede ser parte o no del procesador de señal 330. El aparato 350 está configurado para recibir la matriz de mezcla 337, **G** (o, de forma equivalente, los elementos de matriz de mezcla de la misma), los cuales se proporcionan mediante el cálculo de parámetro de mezclado 336, y para proporcionar, basándose en ello, una matriz de mezcla ajustada 352 **G'** (o de forma equivalente, elementos de matriz de mezcla ajustados de ella). Por ejemplo, se puede proporcionar un conjunto de elementos de matriz de mezcla y un conjunto de elementos de matriz de mezcla ajustada por banda de frecuencia y por trama de audio. En otras palabras, la matriz de mezcla **G** y la matriz de mezcla modificada **G'** pueden actualizarse una vez por trama de audio de la representación de señal de mezcla descendente 210, si se elige un procesamiento a modo de trama. Sin embargo, el intervalo de actualización puede ser diferente en algunos casos. Asimismo, no es necesario que haya múltiples matrices de mezcla y matrices de mezcla ajustadas **G**, **G'** para diferentes bandas de frecuencia.

20

25

Sin embargo, el aparato 350 está configurado para proporcionar elementos de matriz de mezcla ajustados de la matriz de mezcla ajustada 352 basándose en los elementos de matriz de mezcla de la matriz de mezcla 337 proporcionados por el cálculo de parámetro de mezclado 336. Por ejemplo, el procesamiento puede realizarse individualmente por posición de la matriz de mezcla (o matriz de mezcla ajustada), de manera que una secuencia de elementos de matriz de mezcla ajustada de una posición de matriz de mezcla dada puede depender de una secuencia de elementos de matriz de mezcla de la matriz de mezcla 337 en la misma posición de matriz de mezcla, pero independiente de elementos de matriz de mezcla en diferentes posiciones de matriz de mezcla.

30

35

El aparato 350 para proporcionar un elemento de matriz de mezcla ajustado está configurado para proporcionar el uno o más elementos de matriz de mezcla ajustados de la matriz de mezcla ajustada 352 dependiendo de uno o más valores promedio (por ejemplo, uno o más valores promedio individuales de posición de matriz) calculados basándose en la matriz de mezcla 337. El aparato 350 para proporcionar los elementos de matriz de mezcla ajustados de la matriz de mezcla ajustada 352 preferiblemente está configurado para calcular un valor promedio de elementos de matriz de mezcla en una posición de matriz de mezcla dada a lo largo del tiempo. Así, para una posición de matriz de mezcla dada, se puede calcular un valor promedio (preferiblemente, pero no necesariamente, un valor promedio temporal como, por ejemplo, un promedio flotante o un valor promedio de respuesta de impulsos casi infinita o un valor promedio obtenido mediante un filtrado de paso bajo recursivo u operaciones matemáticas similares muy conocidas para promediado en el tiempo), basándose en una secuencia de elementos de matriz de mezcla de la posición de matriz de mezcla dada. Por ejemplo, se puede usar una secuencia de elementos de matriz de mezcla que describe una contribución de un canal dado de la representación de señal de mezcla descendente 210 sobre un canal dado de la representación de señal de mezcla ascendente 220, cuyos elementos de matriz de mezcla están asociados con una pluralidad de tramas de audio, para obtener un valor promedio de este tipo (también designado como valor medio), cuyo valor promedio puede ser un valor promedio de respuesta de impulsos infinita o un valor promedio de respuesta de impulsos (casi) infinita (obtenido, por ejemplo, usando un filtrado de paso bajo recursivo u operaciones matemáticas similares muy conocidas para promediado en el tiempo). Un elemento de matriz de mezcla ajustado actual de la posición de matriz de mezcla dada (que describe la contribución del canal dado de la representación de señal de mezcla descendente 210 sobre el canal dado de la representación de señal de mezcla ascendente 220) puede estar limitado por el aparato 350 a un intervalo de tolerancia que se define dependiendo del valor promedio asociado con la posición de matriz de mezcla dada.

40

45

50

55

Consiguientemente, se evitan excesivas fluctuaciones temporales de elementos de matriz de mezcla, porque los elementos de matriz de mezcla ajustados se limitan a un intervalo de tolerancia el cual está determinado, por ejemplo, por un promedio (promedio de respuesta de impulsos finita o promedio de respuesta de impulsos infinita) de elementos de matriz de mezcla previos en la misma posición de matriz de mezcla. Se ha encontrado que una limitación de este tipo de los elementos de matriz de mezcla ajustados de la matriz de mezcla ajustada 352 normalmente produce una limitación de las distorsiones de la señal de mezcla ascendente 220 causadas por el uso de parámetros no óptimos (por ejemplo, parámetros de renderización especificados por el usuario no óptimos) al menos si los parámetros de renderización especificados por el usuario no óptimos se desvían de los parámetros de renderización especificados

60

por el usuario óptimos en más de una desviación predeterminada.

Debe observarse en este caso que el aparato 350 para proporcionar elementos de matriz de mezcla ajustados puede comprender la misma funcionalidad general que el aparato 100 para proporcionar uno o más parámetros ajustados, en el que los elementos de matriz de mezcla de la matriz de mezcla 337 pueden desempeñar el papel de uno o más parámetros de entrada 110, y en el que los elementos de matriz de mezcla ajustados de la matriz de mezcla ajustada 352 pueden desempeñar el papel del uno o más parámetros ajustados 120.

4. Esquemas de limitación de parámetros según la figura 4

A continuación, se describirán esquemas de limitación de parámetros según la invención haciendo referencia a la figura 4, la cual muestra una representación esquemática de tales esquemas de limitación de parámetros.

La figura 4 muestra la aplicación de esquemas de limitación de parámetros en combinación con un decodificador de SAOC 410. Sin embargo, los esquemas de limitación de parámetros pueden aplicarse en combinación con diferentes tipos de decodificadores o transcodificadores de audio, como, por ejemplo, un transcodificador de SAOC.

El decodificador de SAOC 410 recibe una mezcla descendente 420 y un flujo de bits de SAOC 422. También el decodificador de SAOC proporciona uno o más canales de salida 430a a 420M.

En una primera implementación, designada con (a), el esquema de limitación de parámetros 440 implementa un control indirecto. El esquema de limitación de parámetros 440 recibe una matriz de renderización de entrada R , por ejemplo, una matriz de renderización especificada por el usuario, y proporciona, basándose en ella, una matriz de renderización ajustada $\tilde{R}$ al decodificador de SAOC. En este caso, el decodificador de SAOC usa la matriz de renderización ajustada $\tilde{R}$ para una derivación de la matriz de mezcla G , tal como se describió anteriormente. El esquema de limitación de parámetros 440 también puede recibir parámetros Λ_{R-} , Λ_{R+} , los cuales pueden determinar límites de un intervalo de tolerancia.

Además, se puede aplicar un segundo esquema de limitación de parámetros 450. El segundo esquema de limitación de parámetros recibe parámetros de transcodificación T y proporciona, basándose en ello, parámetros de transcodificación ajustados $\tilde{T}$. Los parámetros de transcodificación T pueden calcularse en el decodificador de SAOC 410, y los parámetros de transcodificación ajustados $\tilde{T}$ pueden aplicarse por el decodificador de SAOC 410. Por ejemplo, los parámetros de transcodificación T pueden ser equivalentes a los elementos de matriz de mezcla de la matriz de mezcla G , tal como se comentó anteriormente, y los parámetros de transcodificación ajustados $\tilde{T}$ pueden ser equivalentes a los elementos de matriz de mezcla ajustada de la matriz de mezcla ajustada G^* .

El esquema de limitación de parámetros 450 puede recibir uno o más parámetros Λ_{R-} , Λ_{R+} , cuyos parámetros pueden determinar límites de intervalos de tolerancia.

4.1 Visión general

A continuación, se dará una visión general del esquema de limitación de parámetros para control de distorsión.

El procesamiento de SAOC general se lleva a cabo de una manera selectiva de tiempo/frecuencia y se describirá a continuación.

El codificador de SAOC extrae las características psicoacústicas (por ejemplo, relaciones y correlaciones de potencia de objeto) de varias señales de objeto de audio de entrada y luego las somete a mezcla descendente formando un canal mono o estéreo combinado (el cual puede designarse, por ejemplo, como una representación de señal de mezcla descendente). Esta señal de mezcla descendente e información lateral extraída se transmiten (o almacenan) en formato comprimido usando codificadores de audio perceptuales muy conocidos. En el extremo de recepción, el decodificador de SAOC trata conceptualmente de recuperar la señal de objeto original (esto es, objetos sometidos a mezcla descendente separados) usando la información lateral transmitida (por ejemplo, información de diferencia de nivel de objeto OLD, información de correlación interobjeto IOC, información de ganancia de mezcla descendente DMG e información de diferencia de nivel de canal de mezcla descendente DCLD). Estas señales de objeto aproximadas se mezclan después formando una escena objetivo usando una matriz de renderización (en el que la matriz de renderización normalmente describe contribuciones de diferentes objetos de audio a diferentes canales de la representación de señal de mezcla ascendente). La matriz de renderización está compuesta por los coeficientes de renderización relativos RC (o ganancias de objeto) especificados para cada objeto de audio transmitido y altavoz de configuración de mezcla ascendente. Estas ganancias de objeto determinan la posición espacial de todos los objetos separados/renderizados. Efectivamente, la separación de las señales de objeto raramente se realiza (o incluso nunca

se realiza) ya que la separación y el mezclado se realizan en una única etapa de procesamiento combinado, lo cual da como resultado una enorme reducción de la complejidad de cálculo. La única etapa de procesamiento combinado puede realizarse, por ejemplo, usando coeficientes de transcodificación, los cuales describen la combinación de la separación de objetos y el mezclado de los objetos separados.

5 Se ha encontrado que este esquema es enormemente efectivo, tanto en términos de tasa de bits de transmisión (solo se requiere transmitir uno o dos canales de mezcla descendente más alguna información lateral en lugar de un número de señales de audio de objetos individuales), como en la complejidad de cálculo (la complejidad de procesamiento se relaciona principalmente con el número de canales de salida más que con el número de objetos de audio).

10 El decodificador de SAOC transforma (en un nivel paramétrico) las ganancias de objeto y otra información lateral directamente en coeficientes de transcodificación (TC) los cuales se aplican a la señal de mezcla descendente para crear las correspondientes señales para la escena de audio de salida renderizada (o una señal de mezcla descendente preprocesada para una operación de decodificación adicional, esto es, normalmente renderización MPEG Surround multicanal).

15 Se ha encontrado que la calidad de audio subjetivamente percibida de la escena de salida renderizada puede mejorarse mediante aplicación de medidas de control de distorsión o DCM, tal como se describe en el documento no publicado previamente US 61/173.456. Esta mejora puede alcanzarse a cambio del precio de aceptar una moderada modificación dinámica de los ajustes de renderización objetivo. La modificación de la información de renderización tiene naturaleza variante con el tiempo y la frecuencia que bajo circunstancias específicas puede resultar en coloraciones de sonido no naturales y artefactos de fluctuación temporal.

20 En una alternativa a las medidas de control de distorsión (DCM) descritas en la referencia [6], realizaciones según la presente invención usan un número de esquemas de limitación de parámetros que se centra en la reducción de artefactos de audio (correlaciones de sonido, fluctuaciones temporales, etc.) y al mismo tiempo conservan una calidad de sonido natural.

25 Los conceptos de esquema de limitación de parámetros propuestos en el presente documento no ajustan coeficientes de renderización (RC) en base a una medida de distorsión calculada usando sofisticados algoritmos basados en modelos psicoacústicos. En cambio, los conceptos de esquema de limitación de parámetros propuestos muestran una baja complejidad de cálculo y estructural y por lo tanto son atractivos para su integración en la tecnología de SAOC. De todos modos, también pueden combinarse ventajosamente con los esquemas descritos en la referencia [6] para lograr mejor calidad de salida global complementándose entre sí.

30 Dentro del sistema de SAOC global, los esquemas de limitación de parámetros pueden incorporarse a la cadena de procesamiento del decodificador de SAOC de dos maneras. Por ejemplo, el esquema de limitación de parámetros puede estar ubicado en el extremo frontal para modificación indirecta (externa) de la salida de SAOC controlando los coeficientes de renderización (RC) R , lo cual se muestra como alternativa (a) en la figura 4. De acuerdo con una alternativa que no está cubierta por la invención reivindicada, los coeficientes de transcodificación inherentes (TC) T son modificados directamente (internamente) en el extremo posterior del decodificador de SAOC, antes de que los coeficientes se apliquen a la señal de mezcla descendente para producir las señales de canal de mezcla ascendente de salida, lo cual se muestra como la alternativa (b) de la figura 4.

45 4.2 Control indirecto

A continuación, se comentará en más detalle el concepto de control indirecto.

50 La hipótesis subyacente del método de control indirecto considera una relación entre nivel de distorsión y desviaciones de los RC de su valor promediado en objetos. Esto se basa en la observación de que cuanto más específica es la atenuación/intensificación aplicada por los RC a un objeto particular con respecto a los otros objetos, más agresiva es la modificación de la señal sometida a mezcla descendente transmitida que ha de realizarse por el codificador/transcodificador de SAOC. En otras palabras: cuanto mayor es la desviación de los valores de "ganancia de objeto" entre sí, mayor es la posibilidad de que se produzca una distorsión inaceptable (suponiendo idénticos coeficientes de mezcla descendente). Se ha encontrado que esto puede someterse a prueba examinando la desviación de los RC del promedio de RC sobre los objetos (por ejemplo, valor medio de renderización).

60 Sin pérdida de generalidad, la siguiente descripción está basada en la configuración que considera una mezcla descendente mono con ganancias de mezcla descendente de unidad para todos los objetos. Para el caso de mezclas descendentes no triviales (con ganancias de objeto diferentes y/o dinámicas) el algoritmo puede modificarse apropiadamente. Además, se supone que los RC son invariantes con la frecuencia para simplificar la anotación.

Basándose en el escenario de renderización especificado por el usuario representado por los coeficientes $R(i)$ con índice de objeto i , el PLS evita valores de renderización extremos produciendo valores de RC modificados $\tilde{R}(i)$ que

realmente se usan por el motor de renderización de SAOC. Pueden derivarse según la siguiente función

$$\tilde{R}(i) = F_{\tilde{R}}(R(i), \Lambda),$$

- 5 donde Λ es un parámetro de control del PLS (esto es, valor umbral). El parámetro de control del PLS puede considerarse como un parámetro de tolerancia.

La desviación $R_d(i)$ del coeficiente de renderización $R(i)$ de un valor de renderización promediado $\bar{R}$ (por ejemplo, la media aritmética) puede obtenerse como

10

$$R_d(i) = \frac{R(i)}{\bar{R}},$$

donde

$$\bar{R} = \frac{1}{N_{ob}} \sum_{i=1}^{N_{ob}} R(i).$$

15

Por consiguiente, $R_d(i)$ es una relación entre un coeficiente de renderización $R(i)$ y un valor de renderización promediado $\bar{R}$. El valor de renderización promediado $\bar{R}$ es un valor promedio, promediado sobre los objetos de audio que tienen índices de objeto de audio i , de los coeficientes de renderización $R(i)$.

20

La desviación limitada $\tilde{R}_d(i)$ está restringida a cierto intervalo de tolerancia Λ según

$$\begin{aligned} \tilde{R}_d(i) &= \Lambda & \text{para } R_d(i) > \Lambda, \\ \tilde{R}_d(i) &= \frac{1}{\Lambda} & \text{para } R_d(i) < \frac{1}{\Lambda}. \end{aligned}$$

25

Cabe observar que esto corresponde a una operación de limitación de RC la cual se lleva a cabo con respecto a un valor de referencia, por ejemplo, $\bar{R}$ el cual se calcula dinámicamente a partir de RC de entrada más que de un valor específico predefinido.

- 30 Para el enfoque de PLS descrito, la solución óptima puede formularse como un problema de minimización para el cual se minimiza la diferencia entre $\tilde{R}(i)$ de un RC dado y un valor $\tilde{R}(i)$ modificado (limitado)

$$\|\tilde{R}(i) - R(i)\| \rightarrow \min.$$

- 35 A continuación, se describirán algunas soluciones algorítmicas para proporcionar los coeficientes de renderización ajustados $\tilde{R}(i)$, en el que los coeficientes de renderización ajustados $\tilde{R}(i)$ pueden considerarse parámetros ajustados.

- 40 Las siguientes dos soluciones algorítmicas se basan en la desviación de esos valores de renderización que se encuentran fuera del intervalo de tolerancia, esto es,

$$R_{d, \text{lim}}(i) = R_d(i) \quad \text{para } R_d(i) > \Lambda, \text{ o } R_d(i) < \frac{1}{\Lambda}.$$

4.2.1 Solución de una etapa

45

Se puede emplear una solución sencilla y rápida de una única etapa para limitar todos los valores de renderización que están fuera del intervalo de tolerancia mediante

$$\begin{aligned}\tilde{R}(i) &= \Lambda \bar{R} & \text{para } R_{d, fuera}(i) > \Lambda, \\ \tilde{R}(i) &= \frac{\bar{R}}{\Lambda} & \text{para } R_{d, fuera}(i) < \frac{1}{\Lambda}.\end{aligned}$$

- 5 En contraste, los valores de renderización que están dentro del intervalo de tolerancia pueden no verse afectados, de manera que

$$\tilde{R}(i) = R(i)$$

- 10 para tales valores de renderización $\tilde{R}(i)$.

4.2.2 Solución iterativa

- 15 Se puede emplear otro método directo en el cual los valores de renderización que están fuera del intervalo con desviaciones asociadas $R_{d, fuera}(i)$ se limitan gradualmente. En cada iteración de este algoritmo, la desviación de renderización máxima $R_{d, máx}$ se define como

$$R_{d, máx} = \max \{ R_{d, fuera}(i) \} \text{ para } R_d > \Lambda,$$

$$R_{d, máx} = \min \{ R_{d, fuera}(i) \} \text{ para } R_d < \frac{1}{\Lambda}.$$

El correspondiente coeficiente de renderización está restringido de manera que

$$\tilde{R}(i) = (1 - \lambda) R(i) + \lambda \bar{R}, \quad \lambda \in (0, 1).$$

- 25 Este procesamiento puede realizarse hasta que todos los valores estén dentro de la región de tolerancia o con un número predeterminado de iteraciones.

- Consignientemente, en cada iteración, se selecciona un coeficiente de renderización $R(i_{máx})$ para el cual la desviación $R_{d, fuera}(i_{máx})$ (por ejemplo, a partir de un valor promedio $\bar{R}$) toma el valor máximo $R_{d, máx}$. En otras palabras, se selecciona el coeficiente de renderización $R(i_{máx})$ que comprende una desviación máxima (en términos del valor de desviación $R_{d, fuera}$) desde el promedio $\bar{R}$ de entre los coeficientes de renderización en la respectiva iteración. Además, el coeficiente de renderización seleccionado $R(i_{máx})$ se lleva más cerca del promedio sobre los coeficientes de renderización usando la combinación lineal mencionada anteriormente de $R(i)$ y $\bar{R}$ (lo cual puede aplicarse selectivamente para $i = i_{máx}$). En cada etapa del procedimiento iterativo, se puede realizar una nueva selección del coeficiente de renderización que tiene la máxima desviación desde el valor promedio, de manera que se pueden modificar diferentes coeficientes de renderización en diferentes etapas del algoritmo iterativo. En otras palabras, $i_{máx}$ normalmente se actualiza en cada iteración. También, el valor promedio, opcionalmente, puede volver a calcularse para cada etapa del algoritmo iterativo, considerando un coeficiente de renderización modificado previamente.

4.3 Control directo

- La hipótesis subyacente del método de control directo considera una relación entre nivel de distorsión y desviaciones de los TC de su valor promediado en el tiempo. Esto se basa en la observación de que cuanto más específica es la atenuación/intensificación aplicada a un objeto particular con respecto a los otros objetos, más agresiva es la modificación de la señal de mezcla descendente transmitida por los TC que va a realizarse por el decodificador/transcodificador de SAOC. En otras palabras: si el valor de un TC es inusualmente grande, se puede concluir que el algoritmo de SAOC intenta modificar una señal de objeto con pequeña potencia a una salida dominada por otras(s) señal(es) de objeto(s) con una potencia grande aplicando una fuerte intensificación. Inversamente, si un TC es inusualmente pequeño, se puede concluir que el algoritmo de la SAOC intenta modificar una señal de objeto con una potencia grande a una salida dominada por otra(s) señal(es) de objeto(s) con una potencia pequeña aplicando una fuerte atenuación. En ambos casos, hay un algo riesgo de producir una calidad de señal inaceptablemente baja en la salida de la SAOC. Así, la idea central es evitar grandes desviaciones de los TC de un valor promedio.

Este PLS puede considerarse como variante en tiempo y frecuencia, ya que incluye todas las dependencias de los parámetros de señal de SAOC (por ejemplo, OLD, IOC) y elementos heurísticos del proceso de transcodificación/decodificación.

5 Sin pérdida de generalidad, la siguiente descripción se basa en la configuración considerando una mezcla ascendente mono.

10 Basándose en la salida de la SAOC TC $T(k)$ con índice de frecuencia k , la PLS evita valores extremos de los TC reemplazándolos (por ejemplo, coeficientes de transcodificación que están fuera de un intervalo de tolerancia) con valores de TC modificados los cuales luego se usan por el proceso de renderización de la SAOC actual. Los valores de TC modificados $\tilde{T}(k)$ pueden derivarse con la siguiente función

$$\tilde{T}(k) = F_T(T(k), \Lambda),$$

15 donde Λ es un parámetro de control del PLS (esto es, valor umbral). El parámetro de control del PLS puede considerarse como un parámetro de tolerancia.

20 Como los TC son variables con el tiempo, se aplica un filtro de paso bajo recursivo para calcular la media

$$\bar{T}_n(k) = \mu T_n(k) + (1 - \mu) \bar{T}_{n-1}(k).$$

25 Se considera que la media $\bar{T}$ es un valor promedio, en el que se introduce una ponderación de los valores de transcodificación individuales mediante la aplicación del filtrado de paso bajo recursivo.

En este caso, n representa el índice de tiempo de los TC y $\mu \in (0, 1]$ es el parámetro de promediado. El intervalo de tolerancia para el valor de TC modificado $\tilde{T}(k)$ se define como

$$\frac{\bar{T}(k)}{\Lambda} \leq \tilde{T}(k) \leq \Lambda \bar{T}(k).$$

30 Cabe observar que esto corresponde a una operación de limitación de TC la cual se lleva a cabo con respecto a un valor de referencia el cual se calcula dinámicamente a partir de los TC más que un valor predefinido específico.

35 Para el enfoque de PLS descrito la solución óptima puede formularse como problema de minimización para el cual se minimiza la diferencia entre el valor de TC dado $T(k)$ y de TC modificado (limitado) $\tilde{T}(k)$.

$$\|\tilde{T}(k) - T(k)\| \rightarrow \min.$$

40 A continuación, se describirá un posible algoritmo de solución para este problema.

4.3.1 Algoritmo de solución

El valor de TC modificado $\tilde{T}(k)$ se puede obtener como

$$\tilde{T}(k) = \Lambda \bar{T}(k)$$

45 para $T(k) > \Lambda$,

$$\tilde{T}(k) = \frac{\bar{T}(k)}{\Lambda}$$

50 para

$$T(k) < \frac{1}{\Lambda}.$$

4.3.2 Ejemplos de coeficientes de transcodificación

- 5 El esquema de limitación de parámetro comentado anteriormente para coeficientes de transcodificación puede aplicarse a diferentes coeficientes de transcodificación los cuales se usan, por ejemplo, en decodificadores y transcodificadores de SAOC comentados anteriormente.

- 10 Por ejemplo, el esquema de limitación de parámetros para coeficientes de transcodificación puede aplicarse para limitar parámetros de la matriz de mezcla $\mathbf{G}$, la cual se usa en el procesador de señal 330 del aparato 300. En este caso, un elemento de matriz de mezcla en una posición de matriz dada de la matriz $\mathbf{G}$ puede ocupar el lugar de un coeficiente de transcodificación $T(k)$, en el que k es un índice de frecuencia. Un correspondiente elemento de matriz de mezcla de la matriz de mezcla $\mathbf{G}$ puede corresponder a un coeficiente de transcodificación ajustado $\tilde{T}(k)$. El esquema de limitación de parámetro de transcodificación puede aplicarse, por ejemplo, individualmente a las
- 15 diferentes posiciones de matriz de la matriz de mezcla. Por ejemplo, si la matriz de mezcla $\mathbf{G}$ comprende elementos de matriz de mezcla g_{11} , g_{12} , g_{21} y g_{22} , y la matriz de mezcla ajustada $\mathbf{G}'$ comprende los correspondientes elementos de matriz g_{11}' , g_{12}' , g_{21}' y g_{22}' , el elemento de matriz de mezcla ajustado $g_{11}'(n_0)$, puede derivarse de una secuencia de $g_{11}(1)$ hasta $g_{11}(n_0)$. Se pueden usar derivaciones equivalentes para los otros elementos de matriz de mezcla g_{12}' , g_{21}' y g_{22}' de la matriz de mezcla ajustada $\mathbf{G}'$.

- 20 La tabla de la figura 10 proporciona una lista de coeficientes de transcodificación los cuales pueden modificarse, por ejemplo, limitarse, mediante los esquemas de limitación de parámetros propuestos para todos los modos de operación de SAOC. La tabla de la figura 10 muestra, en una primera columna 1010, diferentes modos de SAOC. La tabla de la figura 10 muestra, además, en una segunda columna 1020, qué parámetros pueden modificarse (por ejemplo, limitarse) mediante el esquema de limitación de parámetros propuesto. Una tercera columna 1030 muestra una referencia a las correspondientes subcláusulas del documento FCD de SAOC MPEG de la referencia [8]. Para resumir, la tabla de la figura 10 muestra una lista de coeficientes de transcodificación los cuales pueden modificarse (por ejemplo, limitarse) mediante los esquemas de limitación de parámetros propuestos para todos los modos de operación de SAOC con referencias a las correspondientes subcláusulas del documento [8] de FCD de SAOC MPEG.

30 4.4 Formulación generalizada del esquema de limitación de parámetros para desviación relativa limitada

- Existe una formulación generalizada para los PLS comentados anteriormente. Esta formulación puede expresarse en forma del siguiente problema de minimización para el parámetro general variable $\tilde{X}_i$, según

$$\begin{cases} \frac{\bar{X}_i}{\Lambda} \leq \tilde{X}_i \leq \Lambda \bar{X}_i, \\ \|\tilde{X}_i - X_i\| \rightarrow \min. \end{cases}$$

- En este caso, el valor de X_i viene dado inicialmente y el valor de "referencia" $\bar{X}_i$, puede estimarse como una función de la variable $\tilde{X}_i$ modificada según $\bar{X}_i = F(\tilde{X}_i)$.

- 40 En lo anterior, la variable de parámetro X_i puede, por ejemplo, ser idéntica a $R(i)$ o $T(i)$. De manera similar, la variable de parámetro ajustado $\tilde{X}_i$ puede ser idéntica al coeficiente de renderización ajustado $\tilde{R}(i)$ o al coeficiente de transcodificación ajustado $\tilde{T}(i)$. Las variables X_i , $\tilde{X}_i$ también pueden, por ejemplo, ser equivalentes a elementos de matriz de mezcla $g_{mn}(i)$ y $g_{mn}'(i)$.

- 45 A continuación, se comentarán dos algoritmos de solución.

- En general, los enfoques analíticos para obtener la solución exacta de tales problemas de minimización son exigentes a nivel de cálculo. De todos modos, existen maneras alternativas sencillas y rápidas de proporcionar resultados subóptimos los cuales todavía son adecuados para propósitos de PLS. En el presente documento, se describen dos de tales enfoques sencillos.

4.4.1 Solución de una etapa

La solución de una etapa se basa en la suposición de que $\tilde{X}_i \approx F(\tilde{X}_i)$ limita todos los valores fuera del intervalo de tolerancia que van a encontrarse dentro del mismo mediante

$$\tilde{X}_i = \Lambda \bar{X}_i$$

$$\tilde{X}_i = \frac{\bar{X}_i}{\Lambda}$$

para

10 $X_i < \frac{1}{\Lambda}$.

Los valores que se encuentran dentro del intervalo de tolerancia (los cuales pueden considerarse como intervalo de tolerancia), pueden, por ejemplo, dejarse sin cambiar.

15 4.4.2 Solución iterativa

La solución iterativa modifica en cada etapa un valor seleccionado fuera del intervalo X_i a X_{i+1} .

$$\tilde{X}_{j^*} = (1-\lambda)X_{j^*} + \lambda\bar{X} \quad \text{con } \lambda \in (0,1).$$

Por ejemplo, el índice de procesamiento i^* puede elegirse usando la condición:

$$X_{i^*} = \max \left(\frac{X_i}{\bar{X}} \right)$$

25 y

$$\frac{X_i}{\bar{X}} > \Lambda,$$

0

$$X_{j^*} = \min \left(\frac{X_j}{\bar{X}} \right)$$

y

$$\frac{X_i}{\bar{X}} < \frac{1}{\Lambda}.$$

35 El número de iteraciones puede fijarse a un determinado valor o derivarse implícitamente del algoritmo.

Cabe observar que todos estos métodos pueden aplicarse para limitar RC y TC tal como se describió anteriormente.

40 4.5 Formulación lineal generalizada

Existe una formulación lineal generalizada para los PLS comentados anteriormente. En la sección previa, la desviación del parámetro general X_i se describe como una relación $\frac{X_i}{\bar{X}_i}$. En contraste, también puede definirse como $\|X_i - \bar{X}_i\|$, que conduce al siguiente problema de minimización para la variable de parámetro general $\bar{X}_i$, según

$$\begin{cases} (\bar{X}_i - \Lambda_{X_-}) \leq \tilde{X}_i \leq (\bar{X}_i + \Lambda_{X_+}), \\ \|\tilde{X}_i - X_i\| \rightarrow \min. \end{cases}$$

En este caso, el valor de X_i viene dado inicialmente y el valor de "referencia" $\tilde{X}_i$ puede estimarse como función de la variable $\tilde{X}_i$ modificada como $\bar{X}_i = F(\tilde{X}_i)$.

5

A continuación, se describirán dos algoritmos de solución para este problema.

En general, los enfoques analíticos para obtener la solución exacta de tales problemas de minimización son exigentes a nivel de cálculo. De todos modos, existen maneras alternativas sencillas y rápidas para proporcionar resultados subóptimos los cuales todavía son adecuados para propósitos de PLS. En este caso, se describen dos de tales enfoques sencillos:

10

4.5.1 Solución de una etapa

15

La solución de una etapa basada en la suposición de que $\tilde{X}_i \approx F(X_i)$ limita todos los valores fuera del intervalo de tolerancia que van a encontrarse dentro del mismo mediante

$$\tilde{X}_i = \min\left(\max\left(X_i, \bar{X}_i - \Lambda_{X_-}\right), \bar{X}_i + \Lambda_{X_+}\right).$$

20

4.5.2 Solución iterativa

La solución iterativa modifica en cada etapa un valor seleccionado X_{i^*} hasta $\tilde{X}_{i^*}$ si X_{i^*} está fuera del intervalo de tolerancia:

25

$$X_{i^*} > \bar{X}_{i^*}$$

y

30

$$\|X_{i^*} - \bar{X}_{i^*}\| > \|X_{i^*} - \Lambda_{X_+}\| \Rightarrow \tilde{X}_{i^*} = X_{i^*} - S,$$

$$X_{i^*} < \bar{X}_{i^*}$$

y

35

$$\|X_{i^*} - \bar{X}_{i^*}\| < \|X_{i^*} - \Lambda_{X_-}\| \Rightarrow \tilde{X}_{i^*} = X_{i^*} + S.$$

Por ejemplo, el índice de procesamiento i^* puede elegirse usando la condición: $\|X_{i^*} - \bar{X}_{i^*}\| \geq \|X_i - \bar{X}_i\|$ y el valor de tamaño de etapa de modificación según $S = \lambda \|X_{i^*} - \bar{X}_{i^*}\|$ con $\lambda \in (0,1)$. El número de iteraciones puede fijarse a un cierto valor o derivarse implícitamente del algoritmo.

40

Este algoritmo proporciona una manera flexible de usar el intervalo de tolerancia, es decir, es dinámicamente cambiante (dependiendo de X_{i^*}).

Cabe observar que todos estos métodos pueden aplicarse para limitar RC y TC tal como se describió anteriormente.

45

Alternativamente, se puede usar el siguiente algoritmo:

Si $X_{i^*} > \bar{X}_{i^*}$ y $\|X_{i^*} - \bar{X}_{i^*}\| > \Lambda_{X_+}$ entonces

50

$$\tilde{X}_{i^*} = X_{i^*} - S$$

y

si $X_{i^*} < X_i$ y $\|X_{i^*} - X_i\| > \Lambda_X$, entonces

$$\tilde{X}_{i^*} = X_{i^*} + S_i$$

- 5 Esta versión del algoritmo usa un intervalo de tolerancia fijo (estático) Λ_{x-} , Λ_{x+} .

4.6 Observaciones adicionales

- 10 Cabe observar que todos estos métodos pueden aplicarse para limitar coeficientes de renderización y coeficientes de transcodificación, tal como se describió anteriormente.

5. Aplicación de esquemas de limitación de parámetros a escenarios de mezcla descendente/ascendente multicanal

- 15 El PLS de TC (por ejemplo, control directo) de un escenario de mezcla descendente/ascendente mono se extiende a una matriz de TC considerando cualquier combinación de canales de mezcla descendente/ascendente. Consecuentemente, se puede aplicar el control directo a cada TC individualmente. El escenario de mezcla ascendente multicanal para el PLS de TC (por ejemplo, control indirecto) puede realizarse, por ejemplo, en un sencillo enfoque múltiple-mono donde todos los coeficientes de renderización individuales se manejan independientemente.

20 6 Resultados de pruebas de audición

6.1 Diseño de prueba y objetos

- 25 Se ha realizado la prueba de audición subjetiva para evaluar el rendimiento perceptual de los conceptos de medida de control de distorsión (DCM) propuestos y compararlos con el procesamiento de decodificación (SAOC RM) del modelo de referencia de SAOC.

- 30 El diseño de la prueba incluye los casos de aplicación individual de los enfoques de control directo e indirecto del esquema de limitación de parámetros propuesto, así como también su combinación. Se incluye en la prueba la señal de salida del decodificador de SAOC regular (sin procesar por el esquema de limitación de parámetros PLS) para demostrar el rendimiento básico de la SAOC. Además, el caso de renderización trivial, que corresponde a la señal de mezcla descendente, se usa en la prueba de audición con propósitos de comparación.

- 35 La tabla de la figura 5a describe condiciones de prueba de audición.

Para la prueba de audición actual se han elegido los cuatro objetos que representan los tipos de artefactos más normales y críticos para las condiciones de renderización extremas a partir del material de prueba de audición de la convocatoria.

- 40 La tabla de la figura 5b describe objetos de audio de la prueba de audición.

Se han aplicado las ganancias de objeto de renderización según la tabla de la figura 6 para los escenarios de mezcla ascendente considerados.

- 45 Como el PLS propuesto funciona usando los flujos de bits de SAOC regulares y somete a mezcla descendente (no se necesita ninguna actividad relacionada con el PLS sobre el lado del codificador de SAOC) y no se apoya en información residual, no se ha aplicado ningún codificador de núcleo a las correspondientes señales de SAOC de mezcla descendente.

- 50 Para todos los objetos de prueba y las condiciones de renderización consideradas, los ajustes globales para el PLS se toman según

$$\Lambda_{\{R-, S+\}} = \Lambda_{\{R-, X+\}} = 6$$

55 6.2 Metodología de prueba

Las pruebas de audición subjetiva se realizaron en una sala de audición aislada acústicamente que está diseñada para permitir audición de alta calidad. La reproducción se hizo usando auriculares (STAX SR Lambda Pro con un convertidor digital a analógico de Lake People y un monitor SRM STAX).

- 60 El método de prueba siguió el procedimiento usado en las pruebas de verificación de audio espacial, basadas en el método "estímulo múltiple con referencia y anclajes ocultos" (MUSHRA) para la evaluación subjetiva de audio de calidad intermedia [7]. El método de prueba se modificó en consecuencia para evaluar el rendimiento perceptual de

los conceptos de DCM propuestos. Según la metodología de prueba adoptada, los oyentes recibieron instrucciones para comparar todas las condiciones de prueba entre sí según las siguientes instrucciones de prueba de audición:

Para cada objeto de audio, se debe:

5

- leer primero la descripción de las mezclas de sonido deseadas que usted como usuario del sistema quisiera lograr:

Objeto "BlackCoffee":	Sonido de sección de trompeta suave dentro de la mezcla de sonido
Objeto "Fanta4":	Sonido de tambor fuerte dentro de la mezcla de sonido
Objeto "LovePop":	Sonido de sección de cuerda suave dentro de la mezcla de sonido
Objeto "Audition":	Música suave y sonido vocal fuerte

- calificar después las señales usando una calificación común para describir ambos

10

– logro del objetivo de la mezcla de sonido deseada

– calidad global de sonido de la escena (considere distorsiones, artefactos, falta de naturalidad...)

15

Participaron 9 oyentes en total en cada una de las pruebas realizadas. Todos los sujetos pueden considerarse oyentes experimentados. Las condiciones de la prueba se aleatorizaron automáticamente para cada objeto de prueba y para cada oyente. Se registraron las respuestas subjetivas mediante un programa informático de MUSHRA sobre una escala que oscila entre 0 y 100. Se permitió una conmutación instantánea entre objetos bajo prueba.

20

6.3 Resultados de prueba de audición

En el apéndice puede encontrarse un breve resumen en términos de diagramas que muestran los resultados obtenidos de la prueba de audición. Estos gráficos muestran la calificación MUSHRA promedio por objeto sobre todos los oyentes y el valor medio estadístico sobre los objetos evaluados junto con los intervalos de confianza de 95% asociados.

25

Se pueden hacer las siguientes observaciones basándose en los resultados de las pruebas de audición realizadas: para todas las pruebas de audición realizadas, la puntuación MUSHRA obtenida demuestra que la funcionalidad del PLS propuesto proporciona mejor rendimiento en comparación con el sistema de SAOC RM regular en el sentido de valores medios estadísticos globales. Cabe observar que la calidad de todos los objetos producida por el decodificador de SAOC regular (que muestra fuertes artefactos de audio para las condiciones de renderización extremas consideradas) se califica solo levemente superior en comparación con la calidad de los ajustes de renderización idénticos a la mezcla descendente lo cual no satisface en absoluto el escenario de renderización deseado. Por ende, se puede concluir que el PLS propuesto conduce a una considerable mejora de la calidad subjetiva de señal para todos los escenarios de prueba de audición considerados. También se puede concluir que el sistema de limitación más prometedor consiste en una combinación de PLS de RC y TC.

35

En la representación gráfica de la figura 7 se pueden ver detalles con respecto a los resultados de la prueba de audición.

40

7. Alternativas de implementación

A pesar de que algunos aspectos se han descrito en el contexto de un aparato, queda claro que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a una etapa del método o a una característica de una etapa del método. Análogamente, aspectos descritos en el contexto de una etapa del método también representan una descripción de un bloque u objeto o característica correspondiente de un aparato correspondiente. Algunos o todas las etapas del método puede realizarse mediante (o usando) un aparato de hardware, como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, uno o más de las etapas del método más importantes pueden realizarse por un aparato de este tipo.

50

La señal de audio codificada puede almacenarse en un medio de almacenamiento digital o puede transmitirse a través de un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión alámbrico tal como Internet.

55

Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EPROM o una memoria FLASH, que tienen señales de control legibles electrónicamente almacenadas en ellos, las cuales actúan conjuntamente (o son capaces de actuar conjuntamente) con un sistema informático programable de manera que se

realice el respectivo método. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

En algunas realizaciones, se puede usar un dispositivo lógico programable (por ejemplo, una matriz de puertas de campo programable) para realizar algunas o todas las funcionalidades de los métodos descritos en el presente documento. En algunas realizaciones, una matriz de puertas de campo programable puede actuar conjuntamente con un microprocesador para realizar uno de los métodos descritos en el presente documento. Generalmente, los métodos se realizan preferiblemente mediante cualquier aparato de hardware.

Las realizaciones descritas anteriormente son meramente ilustrativas para los principios de la presente invención. Se entiende que modificaciones y variaciones de las disposiciones y los detalles descritos en el presente documento serán evidentes para otros expertos en la técnica. Por lo tanto, pretende solo estar limitado por el alcance de las reivindicaciones de patente inminentes y no por detalles específicos presentados a modo de descripción y explicación de las realizaciones en el presente documento.

8. Conclusiones

Las realizaciones según la invención crean esquemas de limitación de parámetros para control de distorsión en decodificadores de audio. Algunas realizaciones según la invención se centran en codificación de objeto de audio espacial (SAOC), la cual proporciona un medio para una interfaz de usuario para una selección de la configuración de reproducción deseada (por ejemplo, mono, estéreo, 5.1, etc.) y una modificación en tiempo real interactiva de la escena de renderización de salida deseada controlando la matriz de renderización según una preferencia personal u otros criterios. Sin embargo, es una tarea directa adaptar el método propuesto para técnicas paramétricas en general.

Debido al enfoque paramétrico basado en mezcla descendente/separación/mezcla, la calidad subjetiva de la salida de audio renderizada depende de los ajustes de parámetro de renderización. La libertad de seleccionar ajustes de renderización de la elección de los usuarios ocasiona el riesgo de que el usuario seleccione opciones de renderización de objeto inapropiadas, tal como manipulaciones de ganancia extrema de un objeto dentro de la escena de sonido global.

Para un producto comercial es, sin duda, inaceptable producir mala calidad de sonido y/o artefactos de audio para cualquier ajuste en la interfaz de usuario. Para controlar el deterioro excesivo de la salida de audio de SAOC producida, se han descrito varias medidas de cálculo las cuales se basan en la idea de calcular una medida de calidad perceptual de la escena renderizada, y dependiendo de esta medida (y otra información), modificar los coeficientes de renderización realmente aplicados (véase, por ejemplo, la referencia [6]).

La presente invención crea ideas alternativas para salvaguardar la calidad subjetiva de sonido de la escena de SAOC renderizada

- para lo cual todo el procesamiento se lleva a cabo dentro del decodificador/transcodificador de SAOC, y
- las cuales no involucran el cálculo explícito de medidas sofisticadas de calidad de audio percibido de la escena de sonido renderizada.

Estas ideas, por ende, pueden implementarse de una manera estructuralmente sencilla y muy eficiente dentro del marco de trabajo de decodificador/transcodificador de SAOC. Dado que los mecanismos de control de distorsión (DCM) propuestos pretenden limitar parámetros inherentes al decodificador de SAOC, a saber, los coeficientes de renderización (RC) y los coeficientes de transcodificación (TC), se denominan esquemas de limitación de parámetros (PLS) en toda la presente descripción.

Sin embargo, los esquemas de limitación de parámetros pueden aplicarse a cualquier decodificador de audio diferente también.

9. Referencias

- [1] C. Faller y F. Baumgarte, "Binaural Cue Coding – Part II: Schemes and applications", IEEE Trans. sobre procesamiento de voz y audio, vol. 11, n.º 6, noviembre de 2003.
- [2] C. Faller, "Parametric Joint-Coding of Audio Sources", Convención de AES, París, 2006, Impresión previa 6752.
- [3] J. Herrer, S. Disch, J. Hilpert, O. Hellmuth: "From SAC to SAOC – Recent Developments in Parametric Coding of Spatial Audio", 22^{do} Congreso Regional de AES de Reino Unido, Cambridge, Reino Unido, abril de 2007.
- [4] J. Engdegård, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Hölzer, L. Terentiev, J. Breebart, J. Koppens, E. Schuijers y W. Oomen "Spatial Audio Object Coding (SAOC) – The Upcoming MPEG Standard for Parametric Object

ES 3 001 434 T3

Based Audio Coding”, 124^{da} Convención de AES, Ámsterdam 2008, Impresión previa 7377.

[5] ISO/IEC “MPEG Audio Technology – Part 2: Spatial Audio Object Coding (SAOC)”, ISO/IEC JTC1/SC29/WG11 (MPEG) FCD 23003–2.

5

[6] Solicitud de patente de EE.UU. n.º 61/173.456 “METHODS; APPARATUS; AND COMPUTER PROGRAMS FOR DISTORTION AVOIDING AUDIO SIGNAL PROCESSING”.

10

[7] Recomendación técnica de EBU: “MUSHRA-EBU Method for Subjective Listening Tests of Intermediate Audio Quality”, documento B/AIM022, octubre de 1999.

[8] ISO/IEC JTC1/SC29/WG11 MPEG, documento N10843, “Study on ISO/IEC 23003–2:200x Spatial Audio Object Coding (SAOC)”, 89º Congreso de MPEG, Londres, R.U. julio de 2009.

REIVINDICACIONES

1. Aparato (100; 250; 350; 440; 450) para proporcionar uno o más parámetros ajustados (120; 252; 352; $\tilde{R}$; $\tilde{T}$) para una provisión de una representación de señal de mezcla ascendente (220; 430a-430M) basándose en una representación de señal de mezcla descendente (210; 420) y una información lateral paramétrica (212; 422) asociada con la representación de señal de mezcla descendente, comprendiendo el aparato:

un ajustador de parámetros configurado para recibir uno o más parámetros (110; 214; 337) y para proporcionar, basándose en ello, uno o más parámetros ajustados (120; 252; 352), en el que el ajustador de parámetros está configurado para proporcionar el uno o más parámetros ajustados dependiendo de un valor promedio ($\overline{132}$; $\overline{R}$; $\overline{T}$) de una pluralidad de valores de parámetros (110; 214; 337; R; T), de manera que una distorsión de la representación de señal de mezcla ascendente causada por el uso de parámetros no óptimos para la provisión de la representación de señal de mezcla ascendente se reduce al menos para uno o más parámetros que se desvían de los parámetros óptimos más que una desviación predeterminada;

caracterizado por que el aparato está configurado para recibir uno o más coeficientes de renderización (214; R) que describen contribuciones deseadas de objetos de audio a uno a uno o más canales de la representación de la señal de mezcla ascendente (220; 430a-430M), y en el que el aparato está configurado para proporcionar uno o más coeficientes de renderizado ajustados (252; $\tilde{R}$) como los parámetros ajustados.
2. Aparato (100; 250; 350; 440; 450) según la reivindicación 1, en el que el ajustador de parámetros está configurado para proporcionar el uno o más parámetros ajustados dependiendo de un valor promedio el cual es un promedio ponderado de una pluralidad de valores de parámetros.
3. Aparato (100; 250; 350; 440; 450) según la reivindicación 1 o 2, en el que el ajustador de parámetros está configurado para proporcionar el uno o más parámetros ajustados de manera que el uno o más parámetros ajustados se desvían del valor promedio menos que los correspondientes parámetros recibidos.
4. Aparato (100; 250; 440) según una de las reivindicaciones 1 a 3, en el que el ajustador de parámetros está configurado para recibir, como parámetros de entrada, una pluralidad de coeficientes de renderización (214; R); y

en el que el ajustador de parámetros está configurado para calcular un promedio ($\overline{R}$) sobre los coeficientes de renderización asociados con una pluralidad de objetos de audio; y

en el que el ajustador de parámetros está configurado para proporcionar los coeficientes de renderización ajustados (252; $\tilde{R}$) de tal manera que una desviación de un coeficiente de renderización ajustado con respecto a la media de los coeficientes de renderización asociados con una pluralidad de objetos de audio esté restringida.
5. Aparato (100; 250; 440) según la reivindicación 4, en el que el ajustador de parámetros está configurado para dejar un coeficiente de renderización (214; R), que está dentro de un intervalo de tolerancia determinado en función del promedio ($\overline{R}$) sobre los coeficientes de renderización, sin cambios, y para establecer selectivamente un coeficiente de renderización (214; R), que sea mayor que un valor límite superior ($A \cdot \overline{R}$) del intervalo de tolerancia, a un valor que sea menor o igual que el valor límite superior, y

para establecer selectivamente un coeficiente de renderización (214; R), que es menor que un valor límite inferior ($\frac{1}{A} \cdot \overline{R}$) del intervalo de tolerancia a un valor que sea mayor o igual que el valor límite inferior.
6. Aparato (100; 250; 440) según la reivindicación 4, en el que el ajustador de parámetros está configurado para seleccionar iterativamente un respectivo ($R_{(i_{m\acute{a}x})}$) de los coeficientes de renderización, que comprende una desviación máxima ($R_{d,m\acute{a}x}$) del promedio ($\overline{R}$) sobre los coeficientes de renderización en la iteración respectiva, y llevar el seleccionado ($R_{(i_{m\acute{a}x})}$) de los coeficientes de renderización más cerca del promedio ($\overline{R}$) sobre los coeficientes de renderización, con el fin de llevar iterativamente los coeficientes de renderización, que están fuera de un intervalo de tolerancia determinado en función del promedio sobre los coeficientes de renderización, en el intervalo de tolerancia.
7. Aparato (100; 250; 440) según la reivindicación 6, en el que el ajustador de parámetros está configurado para repetir la selección iterativa de uno respectivo ($R_{(i_{m\acute{a}x})}$) de los coeficientes de renderización y la modificación iterativa del seleccionado de los coeficientes de renderización hasta que todos los coeficientes de renderización se ajusten para estar dentro de los intervalos de tolerancia aplicables.

8. Aparato (100; 250; 350; 440; 450) según una de las reivindicaciones 1 a 7, en el que el ajustador de parámetros está configurado para proporcionar uno dado de los uno o más parámetros ajustados, de manera que el uno dado de los parámetros ajustados está dentro de un intervalo de tolerancia, cuyos límites están definidos dependiendo del valor promedio (132; $\bar{R}; \bar{T}; \bar{X}$) de la pluralidad de valores de parámetros de entrada y uno o más parámetros de tolerancia ($\Delta R; \Delta R_{+}; \Delta T; \Delta T_{+}; \Delta X; \Delta X_{+}$) y de manera que una desviación entre un parámetro de entrada y un correspondiente parámetro ajustado se minimiza o se mantiene dentro de un intervalo admisible máximo predeterminado.
9. Aparato (100; 250; 350; 440; 450) según la reivindicación 8, en el que el ajustador de parámetros está configurado para fijar selectivamente un parámetro de entrada, que se encuentra fuera del intervalo de tolerancia, cuyos límites están definidos dependiendo del valor promedio (132; $\bar{R}; \bar{T}; \bar{X}$) de la pluralidad de valores de parámetros de entrada, a un valor de límite superior ($\Delta \bar{R}; \Delta \bar{T}; \Delta \bar{X}; \bar{X} + \Delta X_{+}$) o un valor de límite inferior ($\frac{1}{\Delta} \bar{R}; \frac{1}{\Delta} \bar{T}; \frac{1}{\Delta} \bar{X}; \bar{X} - \Delta X_{-}$) del intervalo de tolerancia, para obtener una versión ajustada ($\bar{R}; \bar{T}; \bar{X}$) del parámetro de entrada.
10. Aparato (100; 250; 350; 440; 450) según la reivindicación 8, en el que el ajustador de parámetros está configurado para seleccionar iterativamente uno respectivo ($R(i_{\max}); X_i$) de los parámetros de entrada, el cual comprende una desviación máxima del valor promedio (132; $\bar{R}; \bar{T}; \bar{X}$) en una respectiva iteración, y para llevar el seleccionado de los parámetros de entrada más cerca del promedio, para llevar iterativamente los parámetros de entrada, que se determinó que estaban fuera del intervalo de tolerancia, cuyos límites están definidos dependiendo del valor promedio, dentro del intervalo de tolerancia.
11. Aparato (100; 350; 450) según la reivindicación 10, en el que el ajustador de parámetros está configurado para elegir un tamaño de etapa de modificación usado para llevar el seleccionado ($R(i_{\max}); X_i$) de los parámetros de entrada, más cerca del valor promedio para ser una fracción predeterminada de una diferencia entre el seleccionado de los parámetros de entrada y el valor promedio.
12. Aparato (200; 300; 410) para proporcionar una representación de señal de mezcla ascendente (220; 430a-430M) basándose en una representación de señal de mezcla descendente (210; 420) y una información lateral paramétrica (212; 422), comprendiendo el aparato:
un aparato (100; 250; 350; 440; 450) para proporcionar uno o más parámetros ajustados (120; 252; 352; $\tilde{R}$, $\tilde{T}$) basándose en uno o más parámetros recibidos (110; 214; 337; R; T), según una de las reivindicaciones 1 a 11;
un procesador de señal (230; 330) configurado para obtener la representación de señal de mezcla ascendente basándose en la representación de señal de mezcla descendente y la información lateral paramétrica,
en el que el aparato para proporcionar uno o más parámetros ajustados está configurado para ajustar uno o más parámetros de procesamiento (252; 352; R; T) del procesador de señal.
13. Aparato (200; 300; 410) según la reivindicación 12, en el que el procesador de señal (230) está configurado para proporcionar la representación de señal de mezcla ascendente (220; 430a-430M) dependiendo de coeficientes de renderización ajustados (252; $\tilde{R}$) que describen contribuciones de objetos de audio a uno o más canales de la representación de señal de mezcla ascendente; y
en el que el aparato (100, 250; 440) para proporcionar uno o más parámetros ajustados está configurado para recibir una pluralidad de parámetros de renderización especificados por el usuario (214; R) como parámetros de entrada y para proporcionar, basándose en ello, uno o más parámetros de renderización ajustados (252; $\tilde{R}$) para que use el procesador de señal.
14. Método para proporcionar uno o más parámetros ajustados para la provisión de una representación de señal de mezcla ascendente basándose en una representación de señal de mezcla descendente y una información lateral paramétrica asociada con la representación de señal de mezcla descendente, comprendiendo el método:
recibir uno o más parámetros; y
proporcionar, basándose en ello, uno o más parámetros ajustados, en el que el uno o más parámetros

- ajustados se proporcionan dependiendo de un valor promedio de una pluralidad de valores de parámetros, de manera que una distorsión de la representación de señal de mezcla ascendente causada por el uso de parámetros no óptimos se reduce al menos para uno o más parámetros que se desvían de los parámetros óptimos más que una desviación predeterminada;
- 5
- caracterizado por que el método comprende recibir uno o más coeficientes de renderización (214; R) que describen contribuciones deseadas de objetos de audio a uno o más canales de la representación de señal de mezcla ascendente (220; 430a-430M), y
- 10
- en el que el método comprende proporcionar uno o más coeficientes de renderización ajustados (252; $\hat{R}$) como los parámetros ajustados.
- 15.
- Programa informático que hace que un ordenador realice el método según la reivindicación 14, cuando el programa informático se ejecuta en el ordenador.
- 15

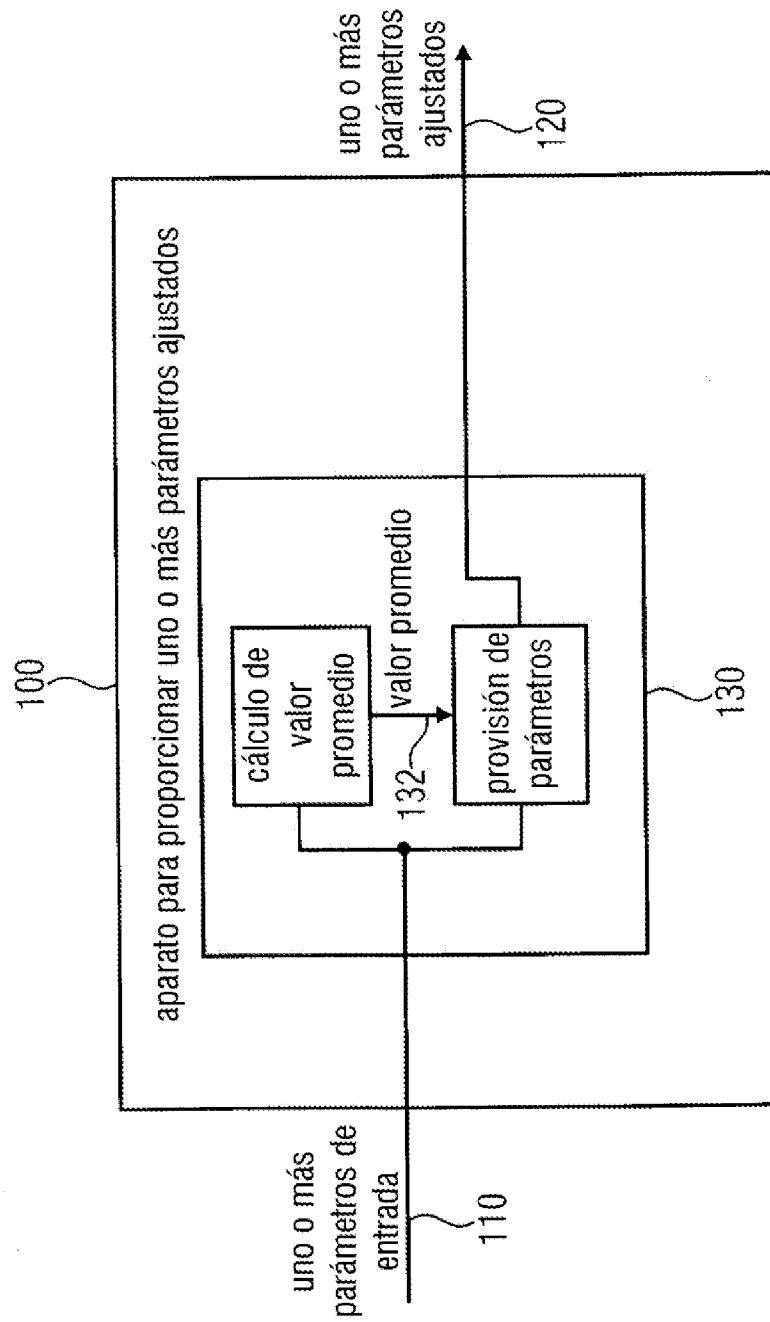


FIG 1

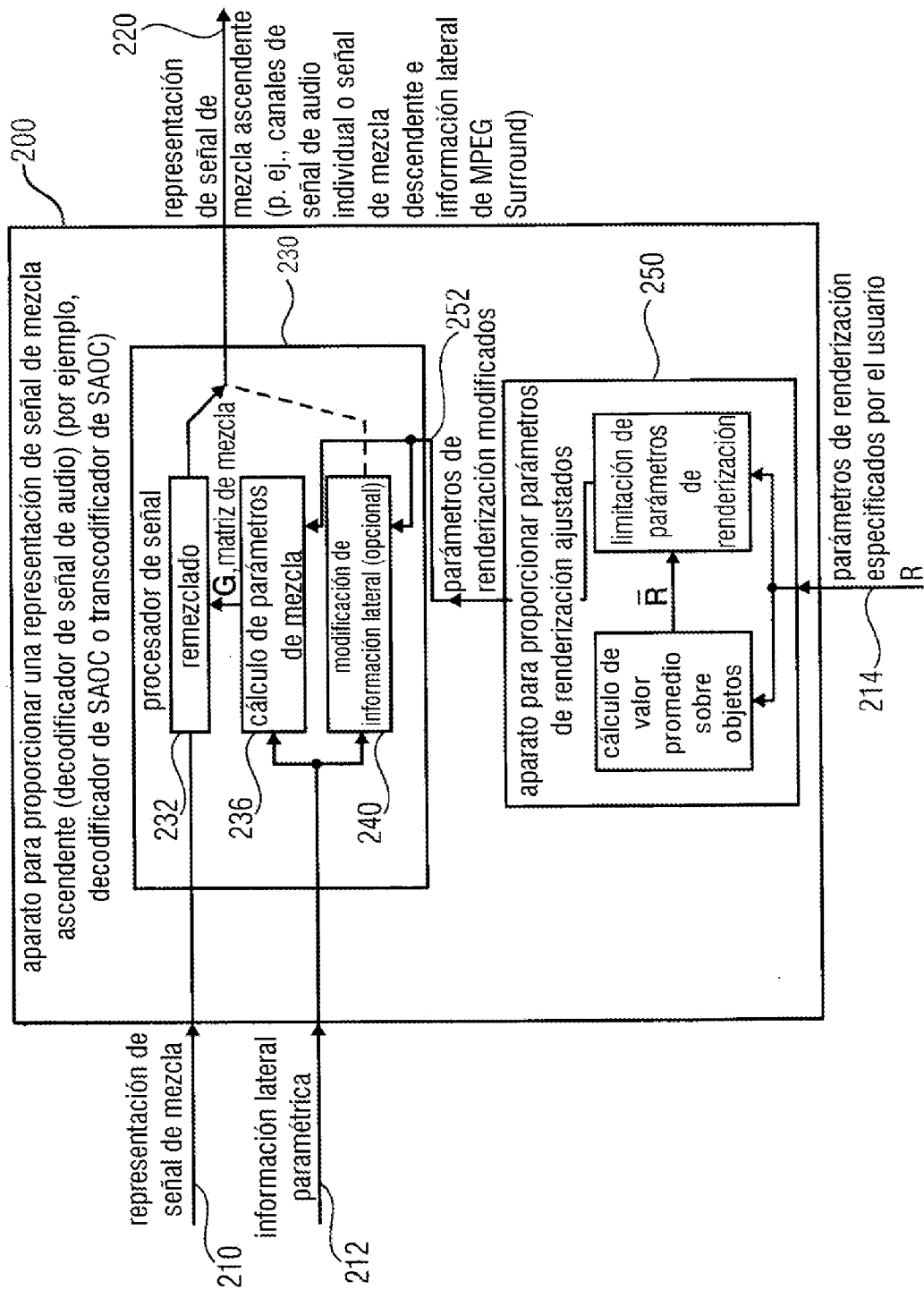


FIG 2

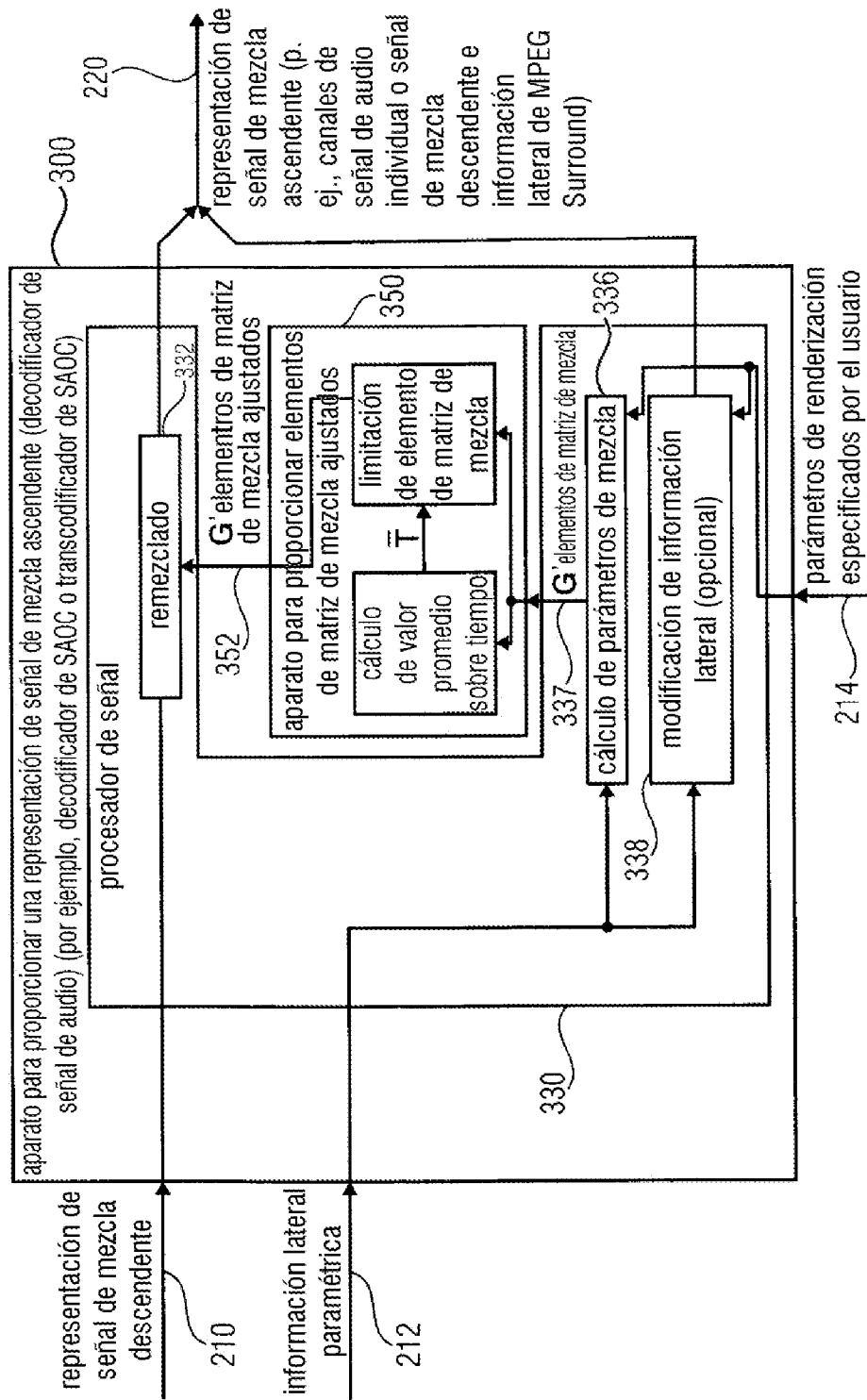
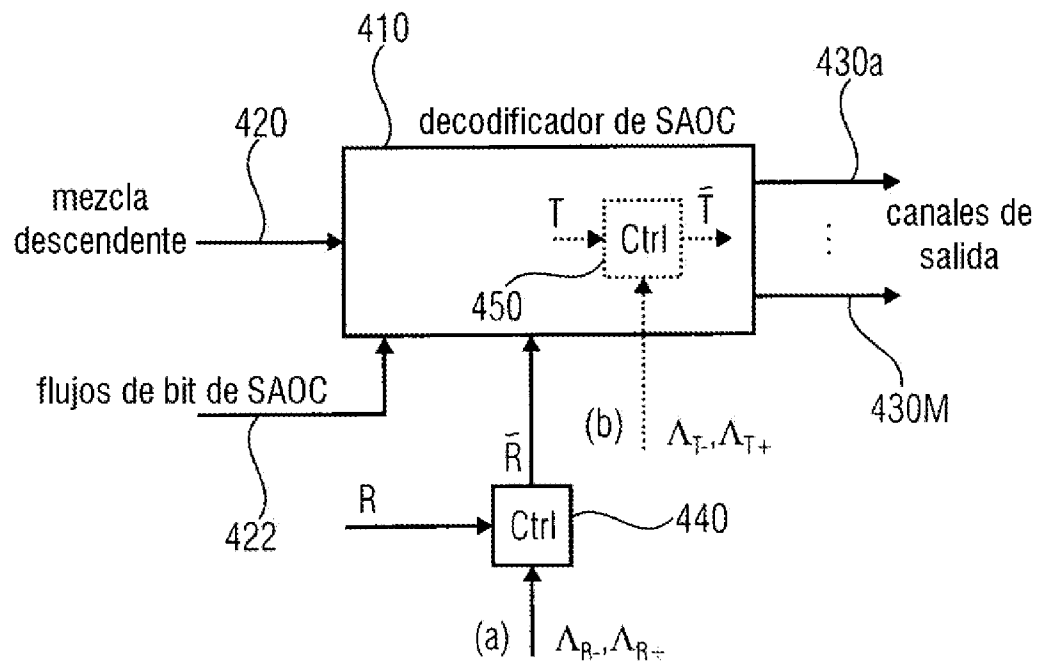


FIG 3



Esquemas de limitación de parámetro - control (a) indirecto y (b) directo

FIG 4

Tabla - Condiciones de prueba de audición

Nombre del codificador	Descripción
"DMX"	señal trivial de renderización similar a mezcla descendente del decodificador de SAOC regular
"noLim"	salida del decodificador de SAOC (no procesado por el PLS) regular
"TC"	salida del decodificador de SAOC con PLS de control directo de TC
"RO"	salida del decodificador de SAOC con PLS de control indirecto de RC
"RO TC 1"	salida del decodificador de SAOC con PLS de control TC/RC combinado

Fig. 5a

Tabla - Objetos de audio de las pruebas de audición

Objetos de audición	Matriz y tipo de renderización	Duración
"BlackCoffee"	sonido de sección de trompeta suave dentro de la mezcla de sonido	7 segundos
"Fanta4"	sonido de tambor fuerte dentro de la mezcla de sonido	7 segundos
"LovePop"	sonido de sección de cuerda suave dentro de la mezcla de sonido	7 segundos
"Audition"	música suave y sonido vocal fuerte	13 segundos

Fig. 5b

Tabla - Condiciones de renderización extrema sometidas a prueba

Objetos de audición	Obj. n.º	Descripción de objeto de audio	Renderización	Coefficiente
"BlackCoffee"	1	Metal 2	0.0001	~-80dB
	2	Metal 1	0.0001	~-80dB
	3	Organo	3.1623	~10dB
	4	Tambor y bajo	3.1623	~10dB
	5	Percusión	3.1623	~10dB
"Fanta4"	1	Voz	0.2512	~-12dB
	2	Bajo	0.2512	~-12dB
	3	Ritmo	2.0	~6dB
"LovePop"	1	Tambor	2.5119	~8dB
	2	Bajo	2.5119	~8dB
	3	Guitarra eléctrica	2.5119	~8dB
	4	Guitarra acústica	2.5119	~8dB
	5	Cuerdas	0.0001	~-80dB
"Audition"	1	Voz BG	0.1778	~-15dB
	2	Bajo	0.1778	~-15dB
	3	Tambor	0.1778	~-15dB
	4	Bombo	0.1778	~-15dB
	5	Guitarra principal	0.1778	~-15dB
	6	Segunda voz principal	3.1623	~10dB
	7	Voz principal	3.1623	~10dB
	8	Tambor superior	0.1778	~-15dB

FIG 6

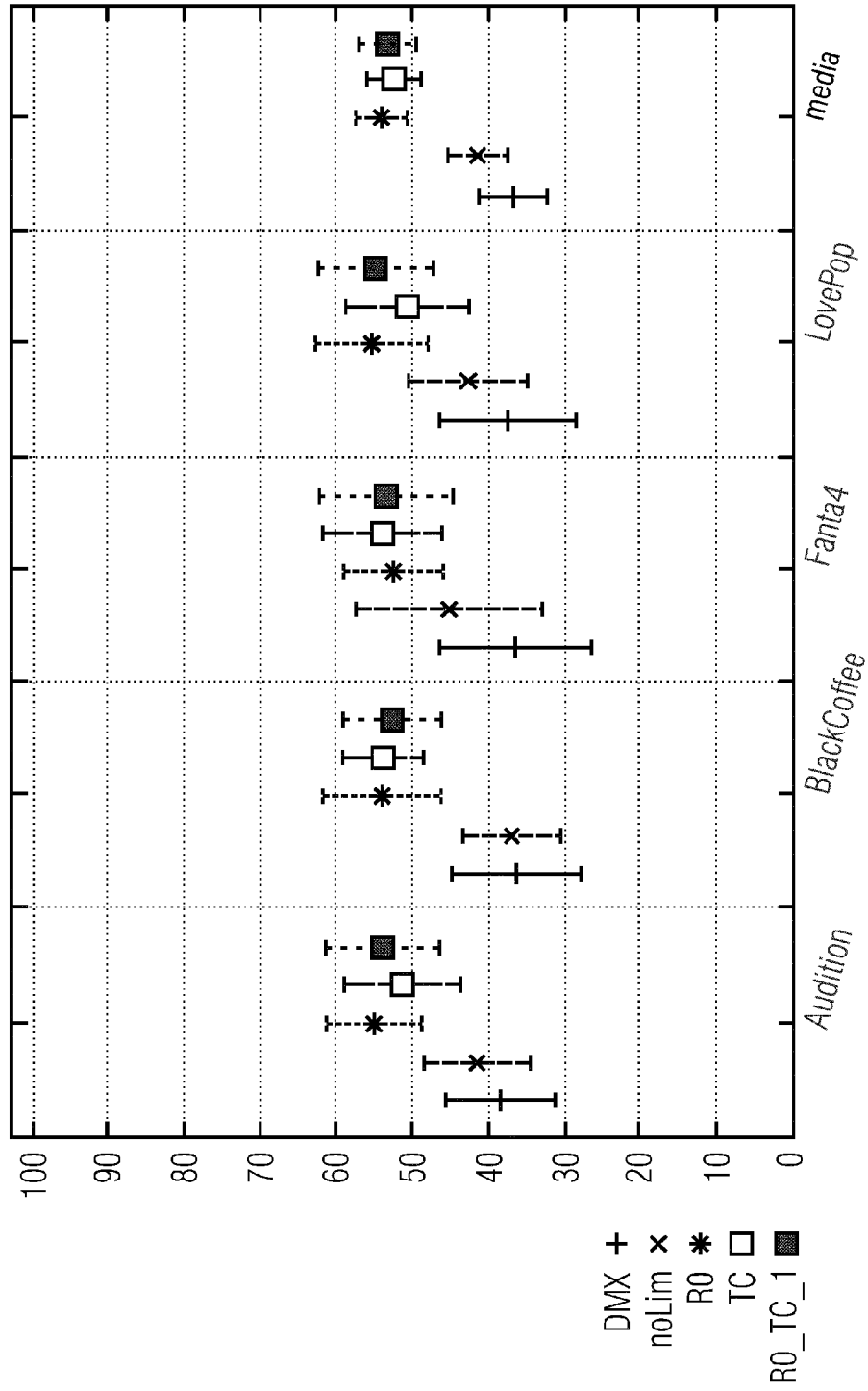
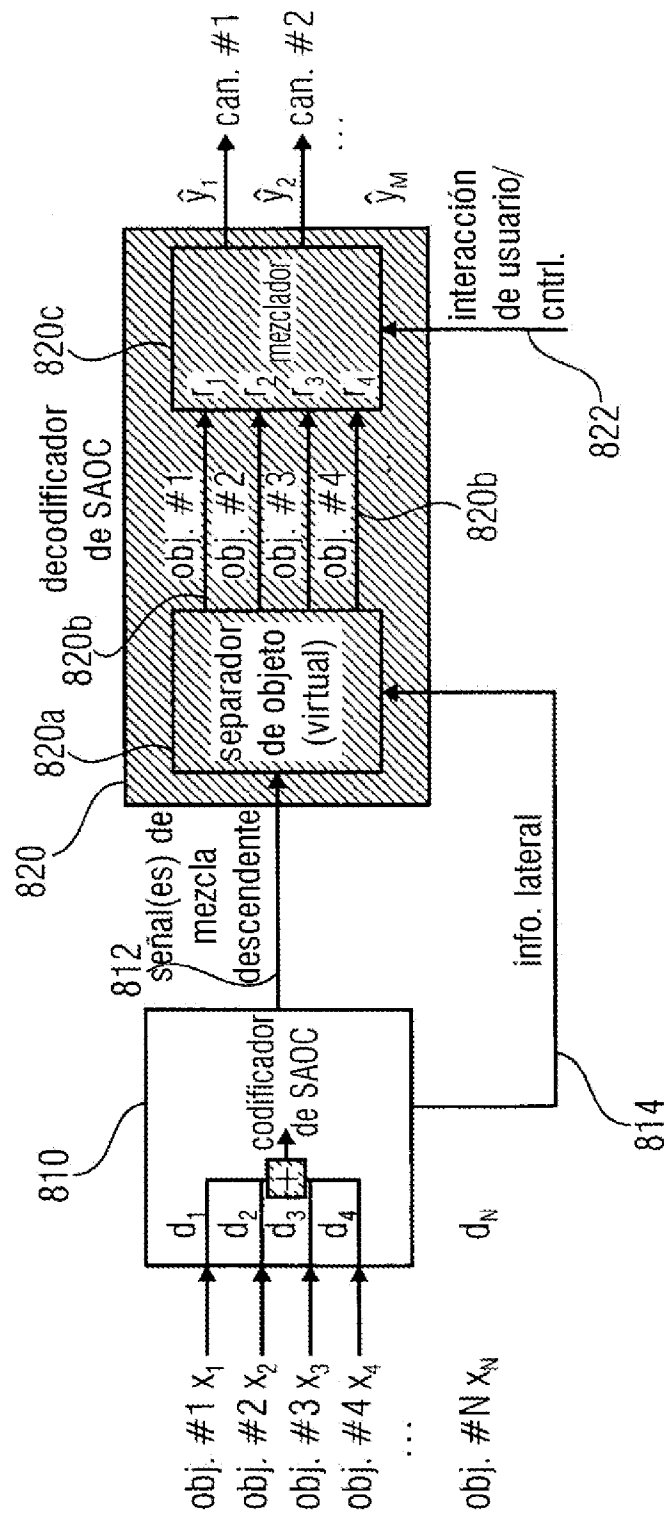
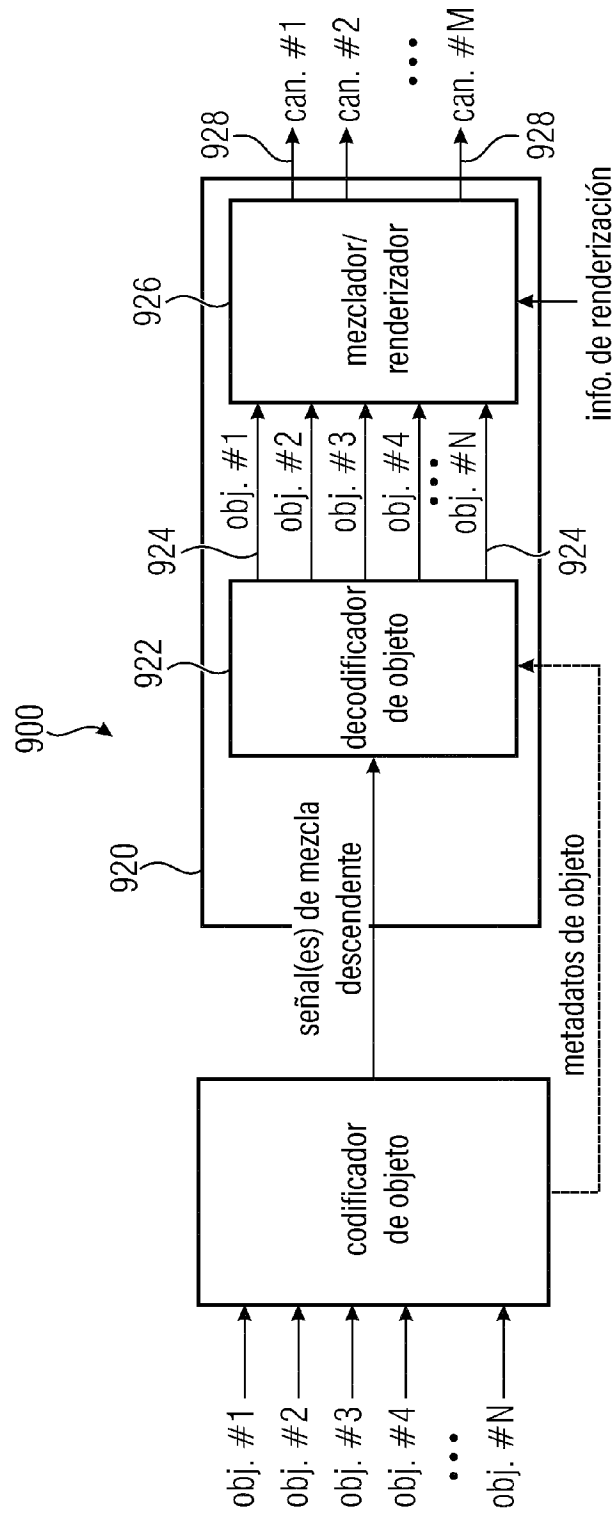


FIGURA - RESULTADOS DE LA PRUEBA DE AUDICIÓN MUSHRA PARA PLS
Fig. 7



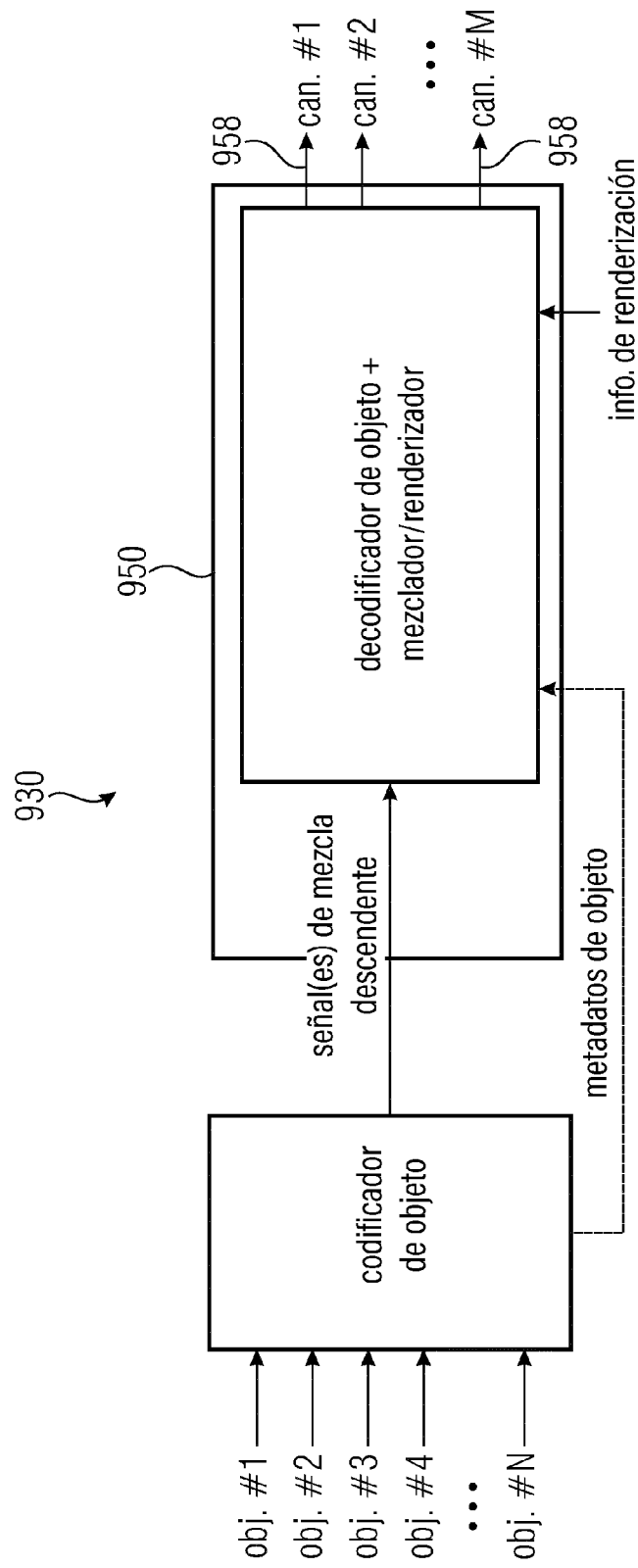
Visión general del sistema de SAOC MPG

FIG 8



MEZCLADOR Y DECODIFICADOR SEPARADOS

Fig. 9a



MEZCLADOR Y DECODIFICADOR INTEGRADOS

Fig. 9b

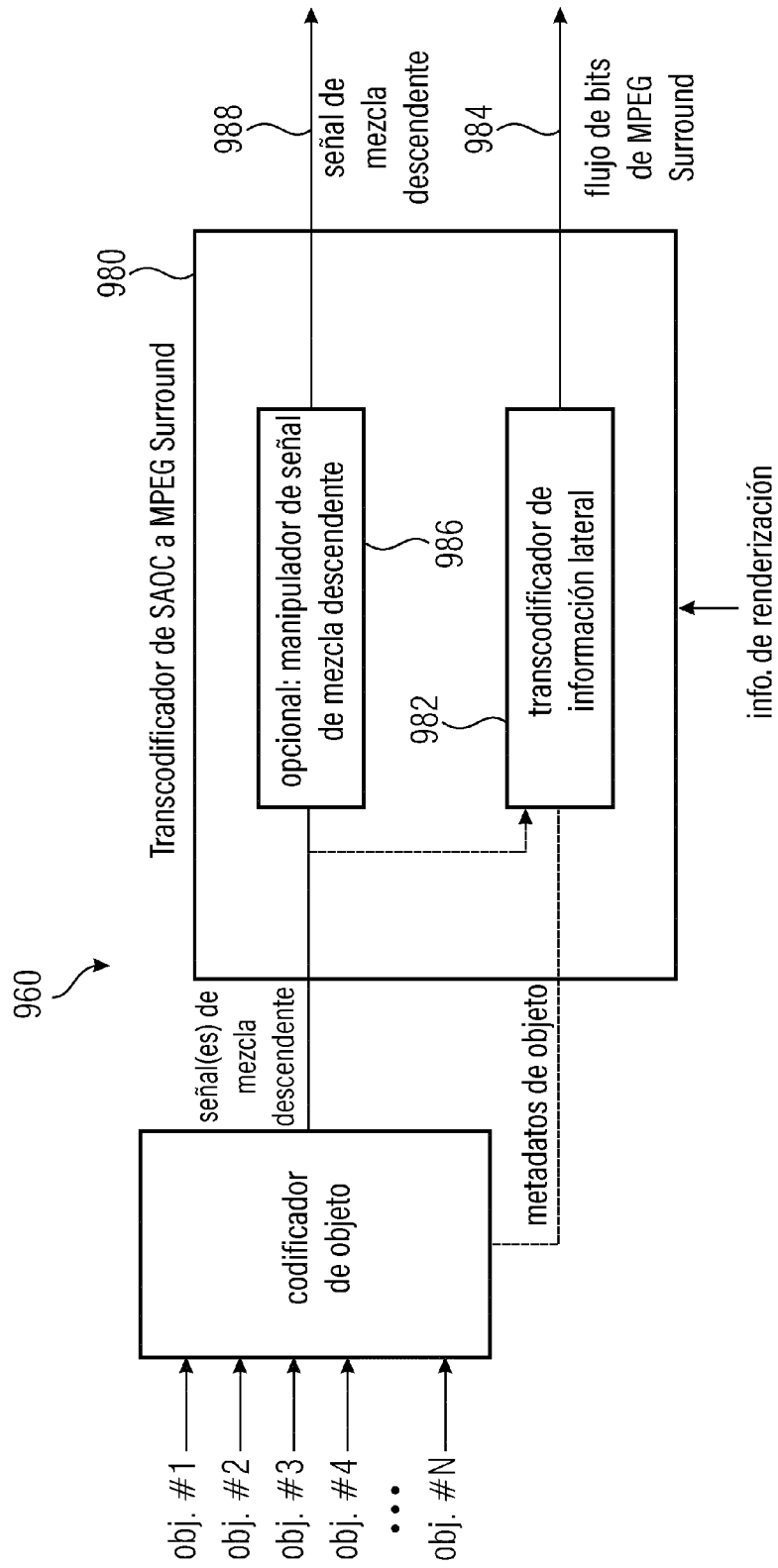


Fig. 9c

Modo de SAOC	Coeficientes modificados	Referencia
"X-1-1" modo de decodificación	coeficientes de matriz de mezcla ($G^{l,m}$)	6.6.1.3
"X-1-2" modo de decodificación	coeficientes de matriz de mezcla ($G^{l,m}$)	6.6.1.2
"X-1-b" modo de decodificación	coeficientes de matriz de mezcla ($G^{l,m}$)	6.6.1.1
"X-1-5" modo de transcodificación	coeficientes de ganancia de mezcla descendente arbitraria ($ADG^{l,m}$)	6.5.2.4
"X-2-1" modo de decodificación	coeficientes de matriz de mezcla ($G^{l,m}$)	6.6.1.6
"X-2-2" modo de decodificación	coeficientes de matriz de mezcla ($G^{l,m}$)	6.6.1.5
"X-2-b" modo de decodificación	coeficientes de matriz de mezcla ($G^{l,m}$)	6.6.1.4
"X-2-5" modo de transcodificación	coeficientes de matriz de predicción (C_3)	6.5.3.3.1

FIG 10