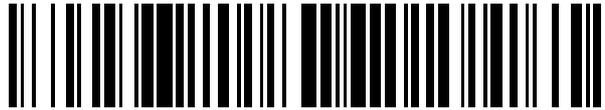


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 569 779**

51 Int. Cl.:

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.11.2010 E 10779542 (9)**

97 Fecha y número de publicación de la concesión europea: **13.01.2016 EP 2489038**

54 Título: **Aparato para proporcionar una representación de señal de mezcla ascendente con base en la representación de señal de mezcla descendente, aparato para proporcionar un flujo de bits que representa una señal de audio multicanal, métodos, programas informáticos y flujo de bits que representan una señal de audio multicanal usando un parámetro de combinación lineal**

30 Prioridad:

20.11.2009 US 263047 P
30.07.2010 EP 10171452
30.07.2010 US 369261 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
12.05.2016

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (50.0%)**
Hansastraße 27c
80686 München, DE y
DOLBY INTERNATIONAL AB (50.0%)

72 Inventor/es:

ENGDEGARD, JONAS;
PURNHAGEN, HEIKO;
HERRE, JUERGEN;
FALCH, CORNELIA;
HELLMUTH, OLIVER y
TERENTIV, LEON

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 569 779 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

5 Aparato para proporcionar una representación de señal de mezcla ascendente con base en la representación de señal de mezcla descendente, aparato para proporcionar un flujo de bits que representa una señal de audio multicanal, métodos, programas informáticos y flujo de bits que representan una señal de audio multicanal usando un parámetro de combinación lineal

CAMPO DE LA TÉCNICA

10 Las realizaciones de acuerdo con la invención se refieren a un aparato para proporcionar una representación de señal de mezcla ascendente con base en una representación de señal de mezcla descendente y una información paramétrica relacionada con el objeto, que se incluyen en una representación de flujo de bits de un contenido de audio y, que dependen de una matriz de reproducción especificada por el usuario.

15 Otras realizaciones de acuerdo con la invención se refieren a un aparato para proporcionar un flujo de bits que representa una señal de audio multicanal.

20 Otras realizaciones de acuerdo con la invención se refieren a un método para proporcionar una representación de señal de mezcla ascendente con base en una representación de señal de mezcla descendente y una información paramétrica relacionada con el objeto que se incluyen en una representación del flujo de bits del contenido de audio y, que dependen de una matriz de reproducción especificada por el usuario.

25 Otras realizaciones de acuerdo con la invención se refieren a un método para proporcionar un flujo de bits que representa una señal de audio multicanal.

Otras realizaciones de acuerdo con la invención se refieren a un programa informático que realiza uno de los métodos.

30 Otra realización de acuerdo con la invención se refiere a un flujo de bits que representa una señal de audio multicanal.

ANTECEDENTES DE LA INVENCION

35 En la técnica de procesamiento de audio, transmisión de audio, y almacenamiento de audio hay un deseo creciente de manejar contenidos multicanal con el fin de mejorar la impresión de la audiencia. El uso de un contenido de audio multicanal trae consigo importantes mejoramientos para el usuario. Por ejemplo, se puede obtener una impresión de audiencia tridimensional, que brinda una mejor satisfacción del usuario en aplicaciones de entretenimiento. Sin embargo, los contenidos de audio multicanal también son útiles en ambientes profesionales, por ejemplo, aplicaciones de teleconferencias, porque la inteligibilidad del altavoz puede ser mejorada mediante el uso de un playback de audio multicanal.

40 Sin embargo, también es deseable tener un buen equilibrio entre los requisitos de la calidad de audio y de velocidad de bits con el fin de evitar un consumo excesivo de recursos a bajo costo o aplicaciones profesionales multicanal.

45 Recientemente se han propuesto técnicas paramétricas para la transmisión eficiente en cuanto a la velocidad de bits y/o el almacenamiento de escenas de audio que contienen varios objetos de audio. Por ejemplo, también se ha propuesto una codificación de apunte binaural, que se describe, por ejemplo, en la referencia [1], y una codificación combinada paramétrica de fuentes de audio, que se describe, por ejemplo, en la referencia [2]. También, se ha propuesto una codificación del objeto de audio espacial (SAOC) de MPEG, que se describe, por ejemplo, en las referencias [3] y [4]. La codificación del objeto de audio espacial de MPEG se encuentra actualmente en la estandarización, y se describe en la referencia [5] no publicada previamente.

55 Estas técnicas tienen por objeto reconstruir perceptualmente la escena de salida deseada en lugar de por una coincidencia de forma de onda.

Sin embargo, en combinación con la interactividad del usuario en el lado receptor, tales técnicas pueden

conducir a una baja calidad de audio de las señales de salida de audio si se lleva a cabo la reproducción extrema del objeto. Esto se describe, por ejemplo, en la referencia [6]. En la referencia [9] se describe una generación de señal de mezcla ascendente a partir de una señal de mezcla descendente, mediante la aplicación de una información de rango de ganancia, vinculada a información de parámetro de objeto en el lado del mezclador ascendente. La referencia [10] describe la tecnología envolvente MPEG, una técnica de generación de sonido multicanal de alta calidad, utilizando técnicas de codificación paramétricas.

En lo siguiente, se describirán tales sistemas, y hay que señalar que los conceptos básicos también se aplican a las realizaciones de la invención.

La Figura 8 muestra una vista general de tal sistema (aquí: MPEG SAOC). El sistema MPEG SAOC 800 mostrado en la Figura 8 comprende un codificador SAOC 810 y un descodificador SAOC 820. El codificador SAOC 810 recibe una pluralidad de señales del objeto x_1 a x_N , que puede ser representado, por ejemplo, como señales de dominio de tiempo o como señales de dominio de tiempo-frecuencia (por ejemplo, en la forma de un conjunto de coeficientes de transformación de una transformación tipo Fourier, o en forma de señales de subbandas QMF). El codificador SAOC 810 típicamente también recibe coeficientes de mezcla descendente d_1 a d_N , que se asocian con las señales del objeto x_1 a x_N . Pueden estar disponibles dos grupos de coeficientes de mezcla descendente para cada canal de la señal de mezcla descendente. El codificador SAOC 810 se configura típicamente para obtener un canal de la señal de mezcla descendente mediante la combinación de las señales del objeto x_1 a x_N de acuerdo con coeficientes de mezcla descendente asociados d_1 a d_N . Típicamente, hay menos canales de mezcla descendente que señales del objeto x_1 a x_N . A fin de permitir (al menos aproximadamente) una separación (o un tratamiento separado) de las señales del objeto en el lado del descodificador SAOC 820, el codificador SAOC 810 proporciona tanto 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 del objeto x_1 a x_N , con el fin de permitir el procesamiento específico de objeto del lado del descodificador.

El descodificador SAOC 820 está configurado para recibir tanto una o más señales de mezcla descendente 812 como la información lateral 814. También, el descodificador SAOC 820 típicamente se configura para recibir una información de interacción del usuario y/o una información de control del usuario 822, que describe una configuración de reproducción deseada. Por ejemplo, la información de interacción del usuario/información de control del usuario 822 puede describir una configuración de altavoz y la colocación espacial deseada de los objetos que proporcionan las señales del objeto x_1 a x_N .

El descodificador SAOC 820 está configurado para proporcionar, por ejemplo, una pluralidad de señales descodificadas del canal de mezcla ascendente \hat{y}_1 a \hat{y}_M . Las señales del canal de mezcla ascendente por ejemplo se pueden asociar con altavoces individuales de una disposición de reproducción multi-altavoz. El descodificador SAOC 820, por ejemplo, puede comprender un separador de objeto 820a, que está configurado para reconstruir, al menos aproximadamente, las señales del objeto x_1 a x_N con base en una o más señales de mezcla descendente 812 y la información lateral 814, obteniendo de este modo las señales del objeto reconstruidas 820b. Sin embargo, las señales del objeto reconstruidas 820b pueden desviarse ligeramente de las señales del objeto originales x_1 a x_N , por ejemplo, porque la información lateral 814 no es suficiente para una reconstrucción perfecta debido a las limitaciones de velocidad de bits. El descodificador SAOC 820 puede comprender además un mezclador 820c, que puede ser configurado para recibir las señales del objeto reconstruidas 820b y la información de interacción del usuario/información de control del usuario 822, y proporcionar, con base en el mismo, las señales del canal de mezcla ascendente \hat{y}_1 a \hat{y}_M . El mezclador 820 puede ser configurado para utilizar la información de interacción del usuario/información de control del usuario 822 para determinar la contribución de las señales del objeto reconstruidas individuales 820b a las señales del canal de mezcla ascendente \hat{y}_1 a \hat{y}_M . La información de interacción del usuario/información de control del usuario 822, por ejemplo, puede comprender parámetros de reproducción (también designados como coeficientes de reproducción), que determinan la contribución de las señales del objeto reconstruidas individuales 822 a las señales del canal de mezcla ascendente \hat{y}_1 a \hat{y}_M .

Sin embargo, debe señalarse que en muchas realizaciones, la separación del objeto, que está indicada por el separador de objeto 820a en la Figura 8, y la mezcla, que se indica por el mezclador 820c en la Figura 8, se llevan a cabo en un solo paso. Para este propósito, los parámetros globales pueden computarse, los cuales describen un mapeo directo de una o más señales de mezcla descendente 812 en las señales del canal de mezcla ascendente \hat{y}_1 a \hat{y}_M . Estos parámetros pueden ser computados con base en la información

lateral y la información de interacción del usuario/información de control del usuario 820.

Haciendo referencia ahora a las Figuras 9a, 9b y 9c, se describirán diferentes aparatos para la obtención de una representación de señal de mezcla ascendente con base en una representación de señal de mezcla descendente e información lateral relacionada con el objeto. La Figura 9a muestra un diagrama esquemático de bloques de un sistema MPEG SAOC 900 que comprende un descodificador SAOC 920. El descodificador SAOC 920 comprende, como bloques funcionales separados, un descodificador del objeto 922 y un mezclador/reproductor 926. El descodificador del objeto 922 proporciona una pluralidad de señales del objeto reconstruidas 924 que dependen de la representación de señal de mezcla descendente (por ejemplo, en la forma de una o más señales de mezcla descendente representadas en el dominio de tiempo o en el dominio de tiempo-frecuencia) e información lateral relacionada con el objeto (por ejemplo, en la forma de metadatos del objeto). El mezclador/reproductor 924 recibe las señales del objeto reconstruidas 924 asociadas con una pluralidad de objetos N y proporciona, con base en las mismas, una o más señales del canal de mezcla ascendente 928. En el descodificador SAOC 920, la extracción de las señales del objeto 924 se realiza por separado de la mezcla/reproducción, lo que permite una separación de la funcionalidad de la descodificación del objeto de la funcionalidad de mezcla/reproducción pero trae consigo una complejidad computacional relativamente alta.

Haciendo referencia ahora a la Figura 9b, se expondrá brevemente otro sistema MPEG SAOC 930, que comprende un descodificador SAOC 950. El descodificador SAOC 950 proporciona una pluralidad de señales del canal de mezcla ascendente 958 que dependen de una representación de señal de mezcla descendente (por ejemplo, en la forma de una o más señales de mezcla descendente) y una información lateral relacionada con el objeto (por ejemplo, en la forma de metadatos del objeto). El descodificador SAOC 950 comprende un descodificador del objeto combinado y un mezclador/reproductor, que está configurado para obtener las señales del canal de mezcla ascendente 958 en un proceso de mezcla conjunta sin una separación de la descodificación y la mezcla/reproducción objeto, en donde los parámetros para el proceso de mezcla ascendente conjunta dependen tanto de la información lateral relacionada con el objeto como con la información de la reproducción. El proceso de mezcla ascendente conjunta depende también de la información de mezcla descendente, que se considera es parte de la información lateral relacionada con el objeto.

Para resumir lo anterior, la provisión de las señales del canal de mezcla ascendente 928, 958 se puede realizar en un proceso de un solo paso o un proceso de dos pasos.

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

El transcodificador Surround de SAOC a MPEG comprende un transcodificador de información lateral 982, que está configurado para recibir la información lateral relacionada con el objeto (por ejemplo, en la forma de metadatos del objeto) y, opcionalmente, información sobre una o más señales de mezcla descendente y la información de reproducción. El transcodificador de información lateral también está configurado para proporcionar una información lateral Surround MPEG (por ejemplo, en la forma de un flujo de bits Surround MPEG) con base en un dato recibido. Por consiguiente, el transcodificador de información lateral 982 está configurado para transformar una información lateral relacionada con el objeto (paramétrica), que se libera del codificador objeto, en una información lateral relacionada con el canal (paramétrica), teniendo en cuenta la información de reproducción y, opcionalmente, la información sobre el contenido de una o más señales de mezcla descendente.

Opcionalmente, el transcodificador Surround de SAOC a MPEG 980 puede ser configurado para manipular una o más señales de mezcla descendente, descritos, 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 ser omitido, de tal manera que la representación de señal de mezcla descendente de salida 988 del transcodificador Surround de SAOC a MPEG 980 es idéntica a la representación de señal de mezcla descendente de entrada del transcodificador Surround de SAOC a MPEG. El manipulador de señal de mezcla descendente 986, por ejemplo, puede ser usado si la información lateral Surround MPEG relacionada con el canal 984 no permitiría proporcionar una impresión de audiencia deseada con base en la representación de señal de mezcla descendente de entrada del transcodificador Surround de SAOC a MPEG 980, que puede ser el caso en algunas constelaciones de reproducción.

5 En consecuencia, el transcodificador Surround de SAOC a MPEG 980 proporciona la representación de señal de mezcla descendente 988 y el flujo de bits Surround MPEG 984 de tal manera que una pluralidad de señales del canal de mezcla ascendente, que representan los objetos de audio de acuerdo con la entrada de información de reproducción al transcodificador Surround de SAOC a MPEG se pueden generar utilizando un descodificador Surround MPEG, que recibe el flujo de bits Surround MPEG 984 y la representación de señal de mezcla descendente 988.

10 Para resumir lo anterior, se pueden utilizar diferentes conceptos para descodificar señales de audio codificadas por SAOC. En algunos casos, se utiliza un descodificador SAOC, que proporciona señales del canal de mezcla ascendente (por ejemplo, señales del canal de mezcla ascendente 928, 958) que dependen de la representación de señal de mezcla descendente y la información lateral paramétrica relacionada con el objeto. Ejemplos de este concepto se pueden ver en las Figuras 9a y 9b. Alternativamente, la información de audio codificada por SAOC puede ser transcodificada 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 Surround MPEG relacionado con el canal 984), que puede ser utilizada por un descodificador Surround MPEG para proporcionar las señales del canal de mezcla ascendente, deseadas.

20 En el sistema MPEG SAOC 800, en la Figura 8 se da una visión general del sistema, el procesamiento general se lleva a cabo en un modo selectivo de frecuencia y se puede describir como sigue dentro de cada banda de frecuencia:

25 • Las señales del objeto de audio de entrada N x_1 a x_N se mezclan en forma descendente como parte del procesamiento del codificador SAOC. Para una mezcla ascendente mono, los coeficientes de mezcla descendente se denotan por d_1 a d_N . Además, el codificador SAOC 810 extrae información lateral 814 que describe las características de los objetos de audio de entrada. Para MPEG SAOC, las relaciones de las potencias del objeto con respecto a otras son la forma más básica de tal información lateral.

30 • La señal (o señales) de mezcla descendente 812 y la información lateral 814 se transmiten y/o almacenan. Para este fin, la señal de audio de mezcla descendente puede comprimirse utilizando codificadores perceptuales de audio bien conocidos, tales como Capa II o III de MPEG-1 (también conocidos como “.mp3”), Codificación Avanzada de Audio MPEG (AAC), o cualquier otro codificador de audio.

35 • En el extremo receptor, el descodificador SAOC 820 conceptualmente intenta restaurar la señal del objeto original (“separación de objeto”) utilizando la información lateral transmitida 814 (y, naturalmente, una o más señales de mezcla descendente 812). Estas señales aproximadas del objeto (también designadas como señales reconstruidas del objeto 820b) después se mezclan en una escena objetivo representada por M canales de salida de audio (que, por ejemplo, pueden ser representadas por las señales del canal de mezcla ascendente \hat{y}_1 a \hat{y}_M) utilizando una matriz de reproducción. Para una salida mono, se dan los coeficientes de la matriz de reproducción por r_1 a r_N .

40 • Efectivamente, la separación de las señales del objeto raramente se ejecuta (o incluso nunca se ejecuta), ya que tanto el paso de separación (indicado por el separador del objeto 820a) y el paso de mezcla (indicado por el mezclador 820c) se combinan en un único paso de transcodificación, que a menudo se traduce en una enorme reducción de la complejidad computacional.

45 Se ha encontrado que tal sistema es tremendamente eficiente, tanto en términos de velocidad de bits de transmisión (solamente es necesario para transmitir algunos pocos canales de mezcla descendente más alguna información lateral en lugar de N señales de audio del objeto discretas o un sistema discreto) y la complejidad computacional (la complejidad del procesamiento se refiere principalmente al número de canales de salida en lugar del número de objetos de audio). Otras ventajas para el usuario en el extremo receptor incluyen la libertad de elegir una configuración de la reproducción de su elección (mono, estéreo, surround, playback virtualizado de auriculares, y así sucesivamente) y la función de la interactividad del usuario: la matriz de reproducción, y por lo tanto la escena de salida, se pueden establecer y cambiar de forma interactiva por el usuario de acuerdo a su voluntad, a la preferencia personal u otros criterios. Por ejemplo, es posible localizar los conversadores de un grupo juntos en un área espacial para maximizar la discriminación de otros conversadores restantes. Esta interactividad se logra al proporcionar una interfaz de usuario descodificadora:

5 Para cada objeto de sonido transmitido, se puede ajustar su nivel relativo (para la reproducción no mono) y la posición espacial de la reproducción. Esto puede ocurrir en tiempo real mientras el usuario cambia la posición de las deslizaderas de interfaz de usuario de gráfica asociada (GUI) (por ejemplo: nivel del objeto = +5 dB, posición del objeto = -30 grados).

10 Sin embargo, se ha encontrado que la elección del lado del descodificador de parámetros para la provisión de la representación de señal de mezcla ascendente (por ejemplo, las señales del canal de mezcla ascendente \hat{y}_1 a \hat{y}_M) trae consigo degradaciones audibles en algunos casos.

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

15 **BREVE DESCRIPCIÓN DE LA INVENCION**

20 Una realización de acuerdo con la invención crea un aparato para proporcionar una representación de señal de mezcla ascendente con base en una representación de señal de mezcla descendente y una información paramétrica relacionada con el objeto, que se incluyen en una representación de flujo de bits de un contenido de audio, y que dependen de una matriz de reproducción especificada por el usuario. El aparato comprende un limitador de distorsión configurado para obtener una matriz de reproducción modificada utilizando una combinación lineal de una matriz de reproducción especificada por el usuario y una matriz de reproducción objetivo, que dependen de un parámetro de combinación lineal. El aparato comprende también un procesador de señales configurado para obtener la representación de señal de mezcla ascendente con base en la representación de señal de mezcla descendente y la información paramétrica relacionada con el objeto utilizando la matriz de reproducción modificada. El aparato está configurado para evaluar un elemento del flujo de bits que representa el parámetro de combinación lineal con el fin de obtener el parámetro de combinación lineal.

30 Esta realización de acuerdo con la invención se basa en la idea fundamental de que las distorsiones audibles de la representación de señal de mezcla ascendente pueden reducirse o incluso evitarse con baja complejidad computacional mediante la realización de una combinación lineal de una matriz de reproducción especificada por el usuario y la matriz de reproducción objetivo que depende de un parámetro de combinación lineal, el cual se extrae de la representación del flujo de bits del contenido de audio, porque una combinación lineal puede realizarse de manera eficiente, y porque la ejecución de la tarea solicitada de determinar el parámetro de combinación lineal se puede realizar en el lado del codificador de señal de audio donde hay típicamente más recursos computacionales disponibles que en el lado del descodificador de señal de audio (aparato para proporcionar una representación de señal de mezcla ascendente).

40 En consecuencia, el concepto anteriormente discutido permite obtener una matriz de reproducción modificada, que da por resultado distorsiones audibles reducidas incluso para una elección inapropiada de la matriz de reproducción especificada por el usuario, sin agregar ninguna complejidad significativa al aparato para proporcionar una representación de señal de mezcla ascendente. En particular, incluso puede ser innecesario modificar el procesador de señales cuando se compara con un aparato sin un limitador de distorsión, porque la matriz de reproducción modificada constituye una cantidad de entrada para el procesador de señales y solamente reemplaza la matriz de reproducción especificada por el usuario.

50 Además, el concepto inventivo trae consigo la ventaja de que un codificador de señal de audio puede ajustar el esquema de limitación de distorsión, que se aplica en el lado del descodificador de señal de audio, de acuerdo con los requisitos especificados en el lado del codificador, simplemente definiendo el parámetro de combinación lineal, que está incluido en la representación del flujo de bits del contenido de audio. En consecuencia, el codificador de señal de audio gradualmente puede proporcionar más o menos libertad con respecto a la elección de la matriz de reproducción al usuario del descodificador (aparato para proporcionar una representación de señal de mezcla ascendente) al elegir adecuadamente el parámetro de combinación lineal. Esto permite la adaptación del descodificador de señal de audio a las expectativas del usuario para un servicio dado, ya que para algunos servicios un usuario puede esperar una calidad máxima (lo que implica reducir la posibilidad del usuario para ajustar arbitrariamente la matriz de reproducción), mientras que para otros servicios el usuario típicamente puede esperar un grado máximo de libertad (lo que implica aumentar el impacto de la matriz de reproducción especificada por el usuario, en el resultado de la combinación lineal).

5 Para resumir lo anterior, el concepto inventivo combina alta eficiencia computacional en el lado del descodificador, que puede ser particularmente importante para los descodificadores de audio portátiles, con la posibilidad de una implementación simple, sin provocar la necesidad de modificar el procesador de señales, y también proporciona un alto grado de control a un codificador de señal de audio, que puede ser importante para cumplir con las expectativas del usuario para diferentes tipos de servicios de audio.

10 En una realización preferida, el limitador de distorsión está configurado para obtener la matriz de reproducción objetivo de manera que la matriz de reproducción objetivo es una matriz de reproducción objetivo sin distorsión. Esto trae consigo la posibilidad de tener un escenario de playback en el que no hay distorsiones o al menos casi ninguna distorsión causada por la elección de la matriz de reproducción. También, se ha encontrado que la computación de una matriz de reproducción objetivo sin distorsión se puede realizar de una manera muy simple en algunos casos. Además, se ha encontrado que una matriz de reproducción, que se elige entre una matriz de reproducción especificada por el usuario y una matriz de reproducción objetivo sin distorsión, típicamente da por resultado una impresión de buena audición.

20 En una realización preferida, el limitador de distorsión está configurado para obtener la matriz de reproducción objetivo de manera que la matriz de reproducción objetivo es una matriz de reproducción objetivo similar a la mezcla descendente. Se ha encontrado que el uso de una matriz de reproducción objetivo similar a enlace descendente trae consigo un grado muy bajo o incluso mínimo de distorsiones. También, tal matriz de reproducción objetivo similar a la mezcla descendente puede obtenerse con un esfuerzo computacional muy bajo, debido a que la matriz de reproducción objetivo similar a la mezcla descendente puede obtenerse mediante el escalamiento de las entradas de la matriz de mezcla descendente con un factor de escala común y la adición de algunas entradas cero adicionales.

25 En una realización preferida, el limitador de distorsión está configurado para escalar una matriz de mezcla descendente ampliada utilizando un valor escalar de normalización de energía, para obtener la matriz de reproducción objetivo, en donde la matriz de mezcla descendente ampliada es una versión ampliada de la matriz de mezcla descendente (una fila de la matriz de mezcla descendente describe contribuciones de una pluralidad de señales del objeto de audio a uno o más canales de la representación de señal de mezcla descendente), ampliada por filas de elementos cero, de manera que un número de filas de la matriz de mezcla descendente ampliada es idéntico a una constelación de reproducción descrita por la matriz de reproducción especificada por el usuario. Así, la matriz de mezcla descendente ampliada se obtiene utilizando una copia de los valores de la matriz de mezcla descendente en la matriz de mezcla descendente ampliada, una adición de entradas cero de la matriz, y una multiplicación escalar de todos los elementos de la matriz con el mismo valor escalar de energía de normalización. Todas estas operaciones se pueden realizar de forma muy eficiente, de tal manera que la matriz de reproducción objetivo se puede obtener rápidamente, incluso en un descodificador de audio muy sencillo.

40 En una realización preferida, el limitador de distorsión está configurado para obtener la matriz de reproducción objetivo de manera que la matriz de reproducción objetivo es una matriz de reproducción objetivo de mejor esfuerzo. A pesar de que este procedimiento es computacionalmente un poco más exigente que el uso de una matriz de reproducción objetivo similar a la mezcla descendente, el uso de una matriz de reproducción objetivo de mejor esfuerzo de representación ofrece una mejor consideración del escenario de reproducción deseada de un usuario. Utilizando la matriz de reproducción objetivo de mejor esfuerzo, la definición de un usuario de la matriz de reproducción deseada se toma en cuenta al determinar la matriz de reproducción objetivo en la medida en que sea posible sin introducir distorsiones o distorsiones significativas. En particular, la matriz de reproducción objetivo de mejor esfuerzo tiene en cuenta la sonoridad deseada del usuario para una pluralidad de altavoces (o canales de la representación de señal de mezcla ascendente). En consecuencia, cuando se utiliza la matriz de reproducción objetivo de mejor esfuerzo puede dar por resultado la impresión de una mejor audición.

55 En una realización preferida, el limitador de distorsión está configurado para obtener la matriz de reproducción objetivo de manera que la matriz de reproducción objetivo depende de una matriz de mezcla descendente y la matriz de reproducción especificada por el usuario. En consecuencia, la matriz de reproducción objetivo está relativamente cerca de las expectativas del usuario, pero todavía proporciona una reproducción de audio sustancialmente sin distorsión. Así, el parámetro de combinación lineal determina un equilibrio entre una aproximación de la reproducción deseada del usuario y la minimización de las distorsiones audibles, en donde la consideración de la matriz de reproducción especificada por el usuario

para la computación de la matriz de reproducción objetivo proporciona una buena satisfacción de los deseos del usuario, incluso si el parámetro de combinación lineal indica que la matriz de reproducción objetivo debe dominar la combinación lineal.

5 En una realización preferida, el limitador de distorsión está configurado para computar una matriz que comprende valores de normalización individuales por canal para una pluralidad de canales de salida de audio del aparato para proporcionar una representación de señal de mezcla ascendente, de tal manera que un valor de normalización de energía para un canal de salida dado del aparato describe, al menos aproximadamente, una proporción entre una suma de valores de reproducción de energía asociados con el canal de salida dado en la matriz de reproducción especificada por el usuario para una pluralidad de objetos de audio, y una suma de valores de mezcla descendente de energía para la pluralidad de objetos de audio. En consecuencia, la expectativa de un usuario con respecto a la sonoridad de los diferentes canales de salida del aparato puede ser satisfecha en algún grado.

15 En este caso, el limitador de distorsión está configurado para escalar un conjunto de valores de mezcla descendente utilizando un valor de normalización de energía individual del canal asociado, para obtener un conjunto de valores de reproducción de la matriz de reproducción objetivo asociada con el canal de salida dado. Por consiguiente, la contribución relativa de un objeto de audio dado para un canal de salida del aparato es idéntica a la contribución relativa del objeto de audio dado para la representación de señal de mezcla descendente, lo que permite evitar sustancialmente distorsiones audibles que podrían ser causadas por una modificación de las contribuciones relativas de los objetos de audio. Por consiguiente, cada uno de los canales de salida del aparato está sustancialmente sin distorsión. No obstante, las expectativas del usuario con respecto a una distribución de sonoridad sobre una pluralidad de altavoces (o canales de la representación de señal de mezcla ascendente) se toma en consideración, aunque los detalles donde colocar cuál objeto de audio y/o cómo cambiar intensidades relativas de los objetos de audio con respecto a los otros se dejaron sin considerar (al menos en cierto grado) con el fin de evitar distorsiones que posiblemente podrían ser causadas por una separación espacial excesivamente brusca de los objetos de audio o una modificación excesiva de las intensidades relativas de objetos de audio.

30 Así, la evaluación de la proporción entre una suma de valores de reproducción de energía (por ejemplo, cuadrados de los valores de reproducción de magnitud) asociada con un canal de salida dado en la matriz de reproducción especificada por el usuario para una pluralidad de objetos de audio y una suma de valores de mezcla descendente de energía para la pluralidad de objetos de audio permite considerar todos los canales de salida de audio, aunque la representación de la señal de mezcla descendente puede comprender menos canales, mientras que aun así se evitarían distorsiones que podrían ser causadas por una redistribución espacial de los objetos de audio o por un cambio excesivo de la sonoridad relativa de los diferentes objetos de audio.

40 En una realización preferida, el limitador de distorsión está configurado para computar una matriz que describe una normalización de energía individual del canal para una pluralidad de canales de salida de audio del aparato para proporcionar una representación de señal de mezcla ascendente que depende de la matriz de reproducción especificada por el usuario y una matriz de mezcla descendente. En este caso, el limitador de distorsión está configurado para aplicar la matriz que describe la normalización de energía individual del canal para obtener un conjunto de coeficientes de reproducción de la matriz de reproducción objetivo asociada con el canal de salida dado del aparato como una combinación lineal de conjuntos de valores de mezcla descendente (es decir, valores que describen una escala aplicada a las señales de audio de diferentes objetos de audio para obtener un canal de la señal de mezcla descendente) asociados con diferentes canales de la representación de señal de mezcla descendente. Al usar este concepto, una matriz de reproducción objetivo, que está bien adaptada a la matriz de reproducción especificada por el usuario, deseada, puede obtenerse incluso si la representación de señal de mezcla descendente comprende más de un canal de audio, mientras que aun así sustancialmente se evitarían distorsiones. Se ha encontrado que la formación de una combinación lineal de conjuntos de valores de mezcla descendente da por resultado un conjunto de coeficientes de reproducción que típicamente causa solamente pequeñas distorsiones audibles. No obstante, se ha encontrado que es posible aproximarse a la expectativa de un usuario utilizando este procedimiento para derivar la matriz de reproducción objetivo.

En una realización preferida, el aparato está configurado para leer un valor de índice que representa el parámetro de combinación lineal de la reproducción del flujo de bits del contenido de audio, y mapear el valor de índice sobre el parámetro de combinación lineal utilizando una tabla de cuantificación de

5 parámetros. Se ha encontrado que este es un concepto en particular computacionalmente eficiente para derivar el parámetro de combinación lineal. También se ha encontrado que este procedimiento trae consigo un mejor equilibrio entre la satisfacción del usuario y la complejidad computacional cuando se compara con otros conceptos posibles en los que se llevan a cabo cómputos complicados, en lugar de la evaluación de una tabla de mapeo unidimensional.

10 En una realización preferida, la tabla de cuantificación describe una cuantificación no uniforme, en la que pequeños valores del parámetro de combinación lineal, que describen una mayor contribución de la matriz de reproducción especificada por el usuario sobre la matriz de reproducción modificada, se cuantifican con resolución comparativamente alta y valores más grandes del parámetro de combinación lineal, que describen una menor contribución de la matriz de reproducción especificada por el usuario sobre la matriz de reproducción modificada se cuantifican con una resolución comparativamente más baja. Se ha encontrado que en muchos casos solamente ajustes extremos de la matriz de reproducción traen significativas distorsiones audibles. En consecuencia, se ha encontrado que un ajuste fino del parámetro de combinación lineal es más importante en la región de una mayor contribución de la matriz de reproducción especificada por el usuario sobre la matriz de reproducción objetivo, con el fin de obtener un ajuste que permite un equilibrio óptimo entre el cumplimiento de la expectativa de reproducción de un usuario y una minimización de las distorsiones audibles.

20 En una realización preferida, el aparato está configurado para evaluar un elemento del flujo de bits que describe un modo de limitación de distorsión. En este caso, el limitador de distorsión preferentemente está configurado para obtener selectivamente la matriz de reproducción objetivo de manera que la matriz de reproducción objetivo es una matriz de reproducción objetivo similar a la mezcla descendente o tal que la matriz de reproducción objetivo es una matriz de reproducción objetivo de mejor esfuerzo. Se ha encontrado que tal concepto conmutable ofrece una posibilidad eficiente para obtener un buen equilibrio entre el cumplimiento de las expectativas de reproducción de un usuario y una minimización de las distorsiones audibles para un gran número de diferentes piezas de audio. Este concepto también permite un buen control de un codificador de señal de audio sobre la reproducción real en el lado del descodificador. Por consiguiente, se pueden cumplir los requisitos de una gran variedad de diferentes servicios de audio.

30 Otra realización de acuerdo con la invención crea un aparato para proporcionar un flujo de bits que representa una señal de audio multicanal.

35 El aparato comprende un mezclador descendente configurado para proporcionar una señal de mezcla descendente con base en una pluralidad de señales del objeto de audio. El aparato comprende también un proveedor de información lateral configurado para proporcionar una información lateral paramétrica relacionada con el objeto, que describe las características de las señales del objeto de audio y los parámetros de mezcla descendente, y un parámetro de combinación lineal que describe las contribuciones de una matriz de reproducción especificada por el usuario y de una matriz de reproducción objetivo a una matriz de reproducción modificada. El aparato para proporcionar un flujo de bits comprende también un formateador del flujo de bits configurado para proporcionar un flujo de bits que comprende una reproducción de la señal de mezcla descendente, la información lateral paramétrica relacionada con el objeto y el parámetro de combinación lineal.

45 Este aparato para proporcionar un flujo de bits que representa una señal de audio multicanal está bien adaptado para la cooperación con el aparato anteriormente discutido para proporcionar una representación de señal de mezcla ascendente. El aparato para proporcionar un flujo de bits que representa una señal de audio multicanal permite proporcionar el parámetro de combinación lineal que depende de su conocimiento de las señales del objeto de audio. En consecuencia, el codificador de audio (es decir, el aparato para proporcionar un flujo de bits que representa una señal de audio multicanal) puede tener un fuerte impacto sobre la calidad de reproducción proporcionada por un descodificador de audio (es decir, el aparato anteriormente discutido para proporcionar una representación de señal de mezcla ascendente) que evalúa el parámetro de combinación lineal. Así, el aparato para proporcionar el flujo de bits que representa una señal de audio multicanal tiene un nivel muy alto de control sobre el resultado de reproducción, que proporciona una satisfacción mejorada del usuario en muchos escenarios diferentes. En consecuencia, el codificador de audio de un proveedor de servicios es el que ofrece la guía, al utilizar el parámetro de combinación lineal, si al usuario se le debe permitir o no usar la configuración de reproducción extrema en riesgo de distorsiones audibles. Así, la decepción del usuario, junto con las correspondientes consecuencias económicas negativas, pueden evitarse mediante el uso del codificador de audio descrito anteriormente.

5 Otra realización de acuerdo con la invención crea un método para proporcionar una representación de señal de mezcla ascendente con base en una representación de señal de mezcla descendente y una información de parámetro relacionado con el objeto, que se incluyen en una reproducción del flujo de bits del contenido de audio, que dependen de una matriz de reproducción especificada por el usuario. Este método se basa en la misma idea clave como el aparato descrito anteriormente.

10 Otro método de acuerdo con la invención crea un método para proporcionar un flujo de bits que representa una señal de audio multicanal. Tal método se basa en el hallazgo mismo que el aparato anteriormente descrito.

Otra realización de acuerdo con la invención crea un programa informático para realizar los métodos anteriores.

15 Otra realización de acuerdo con la invención crea un flujo de bits que representa una señal de audio multicanal. El flujo de bits comprende una representación de una señal de mezcla descendente que combina las señales de audio de una pluralidad de objetos de audio en una información lateral paramétrica relacionada con el objeto que describe las características de los objetos de audio. El flujo de bits también comprende un parámetro de combinación lineal que describe las contribuciones de una matriz de reproducción especificada por el usuario y de una matriz de reproducción objetivo para una matriz de reproducción modificada. El flujo de bits permite algún grado de control sobre los parámetros de reproducción del lado del descodificador desde el lado del codificador de señal de audio.

25 BREVE DESCRIPCIÓN DE LAS FIGURAS

Las realizaciones de acuerdo con la presente invención se describirán posteriormente haciendo referencia a las figuras anexas, en las cuales:

30 La Figura 1a muestra un diagrama esquemático de bloques de un aparato para proporcionar una representación de señal de mezcla ascendente, de acuerdo con una realización de la invención;

La Figura 1b muestra un diagrama esquemático de bloques de un aparato para proporcionar un flujo de bits que representa una señal de audio multicanal, de acuerdo con una realización de la invención;

35 La figura 2 muestra un diagrama esquemático de bloques de un aparato para proporcionar una representación de señal de mezcla ascendente, de acuerdo con otra realización de la invención;

La Figura 3a muestra una representación esquemática de un flujo de bits que representa una señal de audio multicanal, de acuerdo con una realización de la invención;

40 La Figura 3b muestra una representación detallada de la sintaxis de una información de configuración específica de SAOC, de acuerdo con una realización de la invención;

45 La Figura 3c muestra una representación detallada de la sintaxis de una información de cuadro de SAOC, de acuerdo con una realización de la invención;

La Figura 3d muestra una representación esquemática de una codificación de un modo de control de distorsión en un elemento del flujo de bits "bsDcuMode" que puede ser utilizado en un flujo de bits de SAOC;

50 La Figura 3e muestra una representación de la tabla de una asociación entre un índice de flujo de bits idx y un valor de un parámetro de combinación lineal "DcuParam[idx]", que puede ser utilizado para la codificación de una información de combinación lineal en un flujo de bits de SAOC;

55 La Figura 4 muestra un diagrama esquemático de bloques de un aparato para proporcionar una representación de señal de mezcla ascendente, de acuerdo con otra realización de la invención;

La Figura 5a muestra una representación de la sintaxis de una información de configuración específica de SAOC, de acuerdo con una realización de la invención;

La Figura 5b muestra una representación de la tabla de una asociación entre un índice de flujo de bits idx y un parámetro de combinación lineal $Param[idx]$, que puede ser utilizado para codificar el parámetro de combinación lineal en un flujo de bits de SAOC;

5 La Figura 6a muestra una tabla que describe las condiciones de prueba de audición;

La Figura 6b muestra una tabla que describe los elementos de audio de las pruebas de audición;

10 La Figura 6c muestra una tabla que describe las condiciones probadas de mezcla descendente/reproducción para un escenario de descodificación de SAOC estéreo a estéreo;

La Figura 7 muestra una representación gráfica de los resultados de pruebas de audición de la unidad de control de distorsión (DCU) para un escenario SAOC estéreo a estéreo;

15 La Figura 8 muestra un diagrama esquemático de bloques de un sistema de referencia MPEG SAOC;

La Figura 9a muestra un diagrama esquemático de bloques de un sistema SAOC de referencia utilizando un descodificador y mezclador separados;

20 La Figura 9b muestra un diagrama esquemático de bloques de un sistema SAOC de referencia utilizando un descodificador y mezclador integrados; y

La Figura 9c muestra un diagrama esquemático de bloques de un sistema de referencia SAOC utilizando un transcodificador SAOC a MPEG.

25

DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES

30 1. Aparato para proporcionar una representación de señal de mezcla ascendente, de acuerdo a la Figura 1a

La Figura 1a muestra un diagrama esquemático de bloques de un aparato para proporcionar una representación de señal de mezcla ascendente, de acuerdo con una realización de la invención.

35

El aparato 100 está configurado para recibir una representación de señal de mezcla descendente 110 y una información paramétrica relacionada con el objeto 112. El aparato 100 también está configurado para recibir un parámetro de combinación lineal 114. La representación de señal de mezcla descendente 110, la información paramétrica relacionada con el objeto 112 y el parámetro de combinación lineal 114 se incluyen en una representación de flujo de bits de un contenido de audio. Por ejemplo, el parámetro de combinación lineal 114 se describe por un elemento del flujo de bits dentro de la representación del flujo de bits. El aparato 100 también está configurado para recibir una información de reproducción 120, que define una matriz de reproducción especificada por el usuario.

40

45 El aparato 100 está configurado para proporcionar una representación de señal de mezcla ascendente 130, por ejemplo, las señales de canales individuales o una señal de mezcla descendente surround MPEG en combinación con una información lateral surround MPEG.

45

El aparato 100 comprende un limitador de distorsión 140 que está configurado para obtener una matriz de reproducción modificada 142 utilizando una combinación lineal de una matriz de reproducción especificada por el usuario 144 (que se describe, directa o indirectamente, por la información de reproducción 120) y una matriz de reproducción objetivo que depende de un parámetro de combinación lineal 146, que puede, por ejemplo, ser designado con g_{DCU} .

50

55 El aparato 100, por ejemplo, puede ser configurado para evaluar un elemento del flujo de bits 114 que representa el parámetro de combinación lineal 146 con el fin de obtener el parámetro de combinación lineal.

El aparato 100 comprende también un procesador de señales 148 que está configurado para obtener la representación de señal de mezcla ascendente 130 con base en la representación de señal de mezcla

descendente 110 y la información paramétrica relacionada con el objeto 112 utilizando la matriz de reproducción modificada 142.

5 En consecuencia, el aparato 100 es capaz de proporcionar la representación de señal de mezcla ascendente con buena calidad de reproducción utilizando, por ejemplo, un procesador de señales SAOC 148, o cualquier otro procesador de señales relacionado con el objeto 148. La matriz de reproducción modificada 142 se adapta por el limitador de distorsión 140 de tal manera que se logra una impresión de audiencia suficientemente buena con distorsiones suficientemente pequeñas, en la mayoría o en todos los casos. La matriz de reproducción modificada típicamente se encuentra “entre” la matriz de reproducción especificada por el usuario (deseada) y la matriz de reproducción objetivo, en donde se determina un grado de similitud de la matriz de reproducción modificada a la matriz de reproducción especificada por el usuario y la matriz de reproducción objetivo es determinada por el parámetro de combinación lineal, que en consecuencia permite un ajuste de una calidad de reproducción alcanzable y/o de un nivel de distorsión máxima de la representación de señal de mezcla ascendente 130.

15 El procesador de señales 148, por ejemplo, puede ser un procesador de señales de SAOC. En consecuencia, el procesador de señales 148 puede ser configurado para evaluar la información paramétrica relacionada con el objeto 112 para obtener los parámetros que describen las características de los objetos de audio representados, en una forma de mezcla descendente, por la representación de la señal de mezcla descendente 110. Además, el procesador de señales 148 puede obtener (por ejemplo, recibir) los parámetros que describen el procedimiento de mezcla descendente, que se utiliza en el lado de un codificador de audio proporcionando la representación del flujo de bits del contenido de audio con el fin de derivar la representación de señal de mezcla descendente 110 mediante la combinación de las señales del objeto de audio de una pluralidad de objetos de audio. Así, el procesador de señales 148, por ejemplo, puede evaluar una información de diferencia de nivel de objeto OLD que describe una diferencia de nivel entre una pluralidad de objetos de audio para un cuadro de audio dado y una o más bandas de frecuencia, y una información de correlación entre objetos IOC que describe una correlación entre señales de audio de una pluralidad de pares de objetos de audio para un cuadro de audio dado y para una o más bandas de frecuencia. Además, el procesador de señales 148 también puede evaluar una información de mezcla descendente DMG, DCLD que describe una mezcla descendente, que se realiza en el lado de un codificador de audio proporcionando la representación del flujo de bits del contenido de audio, por ejemplo, en la forma de uno o más parámetros de ganancia de mezcla descendente DMG y uno o más parámetros de diferencia de nivel de los canales de mezcla descendente DCLD.

35 Además, el procesador de señales 148 recibe la matriz de reproducción modificada 142, que indica que los canales de audio de la representación de señal de mezcla ascendente 130 deben comprender un contenido de audio de los diferentes objetos de audio. En consecuencia, el procesador de señales 148 está configurado para determinar las contribuciones de los diferentes objetos de audio a la representación de señal de mezcla descendente 110 utilizando su conocimiento (obtenido a partir de la información OLD y la información IOC) de los objetos de audio así como su conocimiento del proceso de mezcla descendente (obtenido a partir de la información DMG y la información DCLD). Además, el procesador de señales proporciona la representación de señal de mezcla ascendente tal que se considera la matriz de reproducción modificada 142.

45 En consecuencia, el procesador de señales 148 cumple la funcionalidad del descodificador SAOC 820, en donde la representación de señal de mezcla descendente 110 toma el lugar de una o más señales de mezcla descendente 812, en donde la información paramétrica relacionada con el objeto 112 toma el lugar de la información lateral 814, y en donde la matriz de reproducción modificada 142 toma el lugar de la información la interacción/control del usuario 822. Las señales del canal \hat{y}_1 a \hat{y}_M toman el papel de la representación de señal de mezcla ascendente 130. En consecuencia, se hace referencia a la descripción del descodificador SAOC 820.

55 De manera similar, el procesador de señales 148 puede tomar el papel del descodificador/mezclador 920, en donde la representación de señal de mezcla descendente 110 toma el papel de una o más señales de mezcla descendente, en donde la información paramétrica relacionada con el objeto 112 toma el papel de los metadatos del objeto, en donde la matriz de reproducción modificada 142 toma el papel de la entrada de información de reproducción al mezclador/reproductor 926, y en donde la señal del canal 928 toma el papel de la representación de señal de mezcla ascendente 130.

5 Alternativamente, el procesador de señales 148 puede realizar la funcionalidad del descodificador y mezclador 950 integrados, en donde la representación de señal de mezcla descendente 110 puede tomar el papel de una o más señales de mezcla descendente, en donde la información paramétrica relacionada con el objeto 112 puede tomar el papel de los metadatos del objeto, en donde la matriz de reproducción modificada 142 puede tomar el papel de la entrada de información de reproducción para el descodificador del objeto más mezclador/reproductor 950, y en donde las señales del canal 958 pueden tomar el papel de la representación de señal de mezcla ascendente 130.

10 Alternativamente, el procesador de señales 148 puede realizar la funcionalidad del transcodificador surround de SAOC a MPEG 980, en donde la representación de señal de mezcla descendente 110 puede tomar el papel de una o más señales de mezcla descendente, en donde la información paramétrica relacionada con el objeto 112 puede tomar el papel de los metadatos del objeto, en donde la matriz de reproducción modificada 142 puede tomar el papel de la información de reproducción, y en donde una o más señales de mezcla descendente 988 en combinación con el flujo de bits surround MPEG 984 puede tomar el papel de la representación de señal de mezcla ascendente 130.

20 Por consiguiente, para obtener más detalles con respecto a la funcionalidad del procesador de señales 148, se hace referencia a la descripción del descodificador SAOC 820, del descodificador y mezclador 920 separados, del descodificador y mezclador integrados 950, y del transcodificador surround de SAOC a MPEG 980. También se hace referencia, por ejemplo, a los documentos [3] y [4] con respecto a la funcionalidad del procesador de señales 148, en donde la matriz de reproducción modificada 142, en lugar de la matriz de reproducción especificada por el usuario 120, toma el papel de la información de reproducción de entrada en las realizaciones de acuerdo con la invención.

25 Más detalles acerca de la funcionalidad del limitador de distorsión 140 se describirán más adelante.

2. Aparato para proporcionar un flujo de bits que representa una señal de audio multicanal, de acuerdo con la Figura 1b

30 La Figura 1b muestra un diagrama esquemático de bloques de un aparato 150 para proporcionar un flujo de bits que representa una señal de audio multicanal.

35 El aparato 150 está configurado para recibir una pluralidad de señales del objeto de audio 160a a 160N. El aparato 150 está configurado además para proporcionar un flujo de bits 170 que representa la señal de audio multicanal, que se describe por las señales del objeto de audio 160a a 160N.

40 El aparato 150 comprende un mezclador descendente 180 que está configurado para proporcionar una señal de mezcla descendente 182 con base en la pluralidad de señales del objeto de audio 160a a 160N. El aparato 150 comprende también un proveedor de información lateral 184 que está configurado para proporcionar una información lateral paramétrica relacionada con el objeto 186 que describe las características de las señales del objeto de audio 160a a 160N y los parámetros de mezcla descendente utilizados por el mezclador descendente 180. El proveedor de información lateral 184 también está configurado para proporcionar un parámetro de combinación lineal 188 que describe una contribución deseada de una matriz de reproducción especificada (deseada) por el usuario y de una matriz de reproducción objetivo (baja distorsión) a una matriz de reproducción modificada.

50 La información lateral paramétrica relacionada con el objeto 186, por ejemplo, puede comprender una información de diferencia de nivel de objeto (OLD), que describe las diferencias al nivel de objeto de las señales del objeto de audio 160a a 160N (por ejemplo, de una manera relacionada con la banda). La información lateral paramétrica relacionada con el objeto también puede comprender una información de correlación entre objetos (IOC) que describe las correlaciones entre las señales del objeto de audio 160a a 160N. Además, la información lateral paramétrica relacionada con el objeto puede describir la ganancia de mezcla descendente (por ejemplo, en una forma relacionada con el objeto), en donde los valores de ganancia de mezcla descendente son utilizados por el mezclador descendente 180 con el fin de obtener la señal de mezcla descendente 182 combinando las señales del objeto de audio 160a a 160N. La información lateral paramétrica relacionada con el objeto 186 puede comprender una información de diferencia de nivel del canal de mezcla descendente (DCLD), que describe las diferencias entre los niveles de mezcla descendente para múltiples canales de la señal de mezcla descendente 182 (por ejemplo, si la señal de mezcla descendente 182 es una señal multicanal).

5 El parámetro de combinación lineal 188 puede ser por ejemplo un valor numérico entre 0 y 1, que describe utilizar solamente una matriz de mezcla descendente especificada por el usuario (por ejemplo, para un valor de parámetro de 0), solamente una matriz de reproducción objetivo (por ejemplo, para un valor de parámetro de 1) o cualquier combinación dada de la matriz de reproducción especificada por el usuario y la matriz de reproducción objetivo entre estos extremos (por ejemplo, para valores de parámetros entre 0 y 1).

10 El aparato 150 comprende también un formateador del flujo de bits 190 que está configurado para proporcionar el flujo de bits 170 de tal manera que el flujo de bits comprende una representación de la señal de mezcla descendente 182, la información lateral paramétrica relacionada con el objeto 186 y el parámetro de combinación lineal 188.

15 En consecuencia, el aparato 150 realiza la funcionalidad del codificador SAOC 810 de acuerdo con la Figura 8 o del codificador objeto de acuerdo con las Figuras 9a-9c. Las señales del objeto de audio 160a a 160N son equivalentes a las señales del objeto x_1 a x_N recibidas, por ejemplo, por el codificador SAOC 810. La señal de mezcla descendente 182, por ejemplo, puede ser equivalente a una o más señales de mezcla descendente 812. La información lateral paramétrica relacionada con el objeto 186, por ejemplo, puede ser equivalente a la información lateral 814 o a los metadatos del objeto. Sin embargo, además de una señal de mezcla descendente de 1 canal o una señal de mezcla descendente multicanal 182 y la información lateral paramétrica relacionada con el objeto 186, el flujo de bits 170 también puede codificar el parámetro de combinación lineal 188.

20 En consecuencia, el aparato 150, que puede ser considerado como un codificador de audio, tiene un impacto sobre un manejo del lado del descodificador del esquema de control de distorsión, que es realizado por el limitador de distorsión 140, al establecer apropiadamente el parámetro de combinación lineal 188, de tal manera que el aparato 150 espera una calidad de reproducción suficiente, provista por un descodificador de audio (por ejemplo, un aparato 100) que recibe el flujo de bits 170.

30 Por ejemplo, el proveedor de información lateral 184 puede establecer el parámetro de combinación lineal que depende de una información de requisito de calidad, que se recibe de una interfaz de usuario opcional 199 del aparato 150. Alternativamente, o además, el proveedor de información lateral 184 también puede tener en consideración las características de las señales del objeto de audio 160a a 160N, y de los parámetros de mezcla descendente del mezclador descendente 180. Por ejemplo, el aparato 150 puede estimar un grado de distorsión, que se obtiene en un descodificador de audio bajo la suposición de una o más matrices de reproducción especificadas por el usuario del peor caso, y puede ajustar el parámetro de combinación lineal 188 de tal manera que se espera obtener una calidad de reproducción por el descodificador de señal de audio bajo la consideración de este parámetro de combinación lineal, todavía se considera como suficiente por el proveedor de información lateral 184. Por ejemplo, el aparato 150 puede establecer el parámetro de combinación lineal 188 a un valor que permita un impacto fuerte del usuario (influencia de la matriz de reproducción especificada por el usuario) sobre la matriz de reproducción modificada, si el proveedor de información lateral 184 encuentra que una calidad de audio de una representación de señal de mezcla ascendente no se degradaría severamente, incluso en presencia de ajustes extremos de reproducción especificados por el usuario. Este puede ser el caso, por ejemplo, si las señales del objeto de audio 160a a 160N son suficientemente similares. Por el contrario, el proveedor de información lateral 184 puede establecer el parámetro de combinación lineal 188 a un valor que permita un impacto comparativamente pequeño del usuario (o de la matriz de reproducción especificada por el usuario), si el proveedor de información lateral 184 encuentra que los ajustes de la reproducción extrema podrían provocar fuertes distorsiones audibles. Este puede ser el caso, por ejemplo, si las señales del objeto de audio 160a a 160N son significativamente diferentes, de tal manera que es difícil una clara separación de los objetos de audio en el lado del descodificador de audio (o conectados con distorsiones audibles).

55 Se debe señalar aquí que el aparato 150 puede usar el conocimiento para el ajuste del parámetro de combinación lineal 188 que solamente está disponible en el lado del aparato 150, pero no en el lado de un descodificador de audio (por ejemplo, el aparato 100) tal como, por ejemplo, una entrada de información de calidad de reproducción deseada al aparato 150 por medio de una interfaz de usuario o el conocimiento detallado respecto a los objetos separados de audio, representados por las señales del objeto de audio 160a y 160N.

En consecuencia, el proveedor de información lateral 184 puede proporcionar el parámetro de combinación

lineal 188 de una manera muy significativa.

3. Sistema SAOC con Unidad de Control de distorsión (DCU), de acuerdo con la Figura 2

5 3.1 Estructura del Descodificador SAOC

En lo siguiente, se describirá tomando como referencia la Figura 2, un proceso realizado por una unidad de control de distorsión (procesamiento DCU) que muestra un diagrama esquemático de bloques de un sistema SAOC 200. Específicamente, la Figura 2 ilustra la unidad de control de distorsión DCU dentro del sistema SAOC global.

Tomando como referencia la Figura 2, el decodificador SAOC 200 está configurado para recibir una representación de señal de mezcla descendente 210 que representa, por ejemplo, una señal de mezcla descendente de 1 canal o una señal de mezcla descendente de 2 canales, o incluso una señal de mezcla descendente que tenga más de dos canales. El decodificador SAOC 200 está configurado para recibir un flujo de bits SAOC 212, que comprende una información lateral paramétrica relacionada con el objeto, tal como, por ejemplo, una información de diferencia de nivel de objeto OLD, una información de correlación entre objetos IOC, una información de ganancia de mezcla descendente DMG, y, opcionalmente, una información de diferencia de nivel de canal de mezcla descendente DCLD. El decodificador SAOC 200 también está configurado para obtener un parámetro de combinación lineal 214, que también se designa con g_{DCU} .

Típicamente, la representación de señal de mezcla descendente 210, el flujo de bits SAOC 212 y el parámetro de combinación lineal 214 se incluyen en una representación del flujo de bits de un contenido de audio.

El decodificador SAOC 200 también está configurado para recibir, por ejemplo, de una interfaz de usuario, una entrada de matriz de reproducción 220. Por ejemplo, el decodificador SAOC 200 puede recibir una entrada de matriz de reproducción 220 en forma de una matriz \mathbf{M}_{ren} , que define la contribución (especificada por el usuario, deseada) de una pluralidad de \mathbf{N}_{obj} objetos de audio a 1, 2, o incluso más canales de señal de salida de audio (de la representación de mezcla ascendente). La matriz de reproducción \mathbf{M}_{ren} , por ejemplo, puede ser la entrada de una interfaz de usuario, en donde la interfaz de usuario puede traducir una diferente forma de representación especificada por el usuario de configuración deseada de reproducción en parámetros de la matriz de reproducción \mathbf{M}_{ren} . Por ejemplo, la interfaz de usuario puede traducir una entrada en la forma de valores de deslizadera de nivel y una información de la posición del objeto de audio en una matriz de reproducción especificada por el usuario \mathbf{M}_{ren} utilizando algún mapeo.

Se debe señalar aquí que a lo largo de la presente descripción, los índices l que definen un segmento de tiempo de parámetro y m que definen una banda de procesamiento algunas veces se omiten en aras de la claridad. No obstante, debe tenerse en cuenta que el procesamiento puede realizarse de forma individual para una pluralidad de ranuras de tiempo de parámetros posteriores que tienen índices l y para una pluralidad de bandas de frecuencia que tienen índices de banda de frecuencia m .

El decodificador SAOC 200 también comprende una unidad de control de distorsión DCU 240 que está configurada para recibir la matriz de reproducción especificada por el usuario \mathbf{M}_{ren} , al menos una parte de la información del flujo de bits SAOC 212 (como se describirá en detalle más adelante) y el parámetro de combinación lineal 214. La unidad de control de distorsión 240 proporciona la matriz de reproducción modificada $\mathbf{M}_{ren,lim}$.

El decodificador de audio 200 comprende también una unidad de descodificación/transcodificación SAOC 248, que puede ser considerada como un procesador de señal, y que recibe la representación de señal de mezcla descendente 210, el flujo de bits SAOC 212 y la matriz de reproducción modificada $\mathbf{M}_{ren,lim}$. La Unidad de descodificación/transcodificación SAOC 248 proporciona una representación 230 de uno o más canales de salida, que puede considerarse como una representación de señal de mezcla ascendente. La representación 230 de uno o más canales de salida, por ejemplo, puede tomar la forma de una representación del dominio de frecuencia de canales de señales de audio individuales, de una representación de dominio de tiempo de canales de audio individuales o de una representación multicanal paramétrica. Por ejemplo, la representación de señal de mezcla ascendente 230 puede tomar la forma de una representación surround MPEG que comprende una señal de mezcla descendente surround MPEG y

una información lateral surround MPEG.

5 Se debe señalar que la unidad de descodificación/transcodificación SAOC 248 puede comprender la misma funcionalidad que un procesador de señales 148, y puede ser equivalente al descodificador SAOC 820, al codificador y mezclador separados 920, al descodificador y mezclador integrados 950 y al transcodificador surround de SAOC a MPEG 980.

3.2 Introducción en la operación del Descodificador SAOC

10 A continuación, se dará una breve introducción en la operación del descodificador 200 SAOC.

15 Dentro del sistema SAOC global, la unidad de control de distorsión (DCU) se incorpora en la cadena de procesamiento del descodificador/transcodificador SAOC entre la interfaz de reproducción (por ejemplo, una interfaz de usuario en donde se puede derivar la matriz de reproducción especificada por el usuario, o una información de la matriz de reproducción especificada por el usuario, es la entrada) y la unidad de descodificación/transcodificación SAOC real.

20 La unidad de control de distorsión 240 proporciona una matriz de reproducción modificada $\mathbf{M}_{ren,lim}^{l,m}$ usando la información de la interfaz de reproducción (por ejemplo, la entrada de la matriz de reproducción especificada por el usuario, directamente o indirectamente, por medio de la interfaz de reproducción o interfaz de usuario) y los datos SAOC (por ejemplo, los datos del flujo de bits SAOC 212). Para más detalles, se hace referencia a la Figura 2. A la matriz de reproducción modificada $\mathbf{M}_{ren,lim}^{l,m}$ se puede acceder mediante la aplicación (por ejemplo, la unidad de descodificación/transcodificación SAOC 248), que refleja las configuraciones de reproducción realmente eficaces.

25 Con base en el escenario de la reproducción especificada por el usuario, representado por la matriz de reproducción (especificada por el usuario) $\mathbf{M}_{ren}^{l,m}$ con los elementos $m_{i,j}^{l,m}$, la DCU evita ajustes extremos de reproducción mediante la producción de una matriz modificada $\mathbf{M}_{ren,lim}^{l,m}$ que comprende coeficientes de reproducción limitada, que serán utilizados por la máquina de reproducción SAOC. Para todos los modos operativos de SAOC, los coeficientes de reproducción finales (procesados por la DCU) se calcularán de acuerdo a:

$$\mathbf{M}_{ren,lim}^{l,m} = (1 - g_{DCU}) \mathbf{M}_{ren}^{l,m} + g_{DCU} \mathbf{M}_{ren,tar}^{l,m}$$

35 El parámetro $g_{DCU} \in [0,1]$, que también se diseña como un parámetro de combinación lineal, se utiliza para definir el grado de transición de la matriz de reproducción especificada por el usuario $\mathbf{M}_{ren}^{l,m}$ hacia la matriz objetivo sin distorsión $\mathbf{M}_{ren,tar}^{l,m}$.

40 El parámetro g_{DCU} se deriva del elemento del flujo de bits "bsDcuParam" de acuerdo con:

$$g_{DCU} = \text{DcuParam}[\text{bsDcuParam}].$$

45 En consecuencia, se forma una combinación lineal entre la matriz de reproducción especificada por el usuario \mathbf{M}_{ren} y la matriz de reproducción objetivo sin distorsión $\mathbf{M}_{ren,tar}$ que depende del parámetro de combinación lineal g_{DCU} . El parámetro de combinación lineal g_{DCU} se deriva de un elemento del flujo de bits, de tal manera que no se requiere computación difícil del parámetro de combinación lineal g_{DCU} (al menos en el lado del descodificador). También, al derivar el parámetro de combinación lineal g_{DCU} del flujo de bits, incluyendo la representación de señal de mezcla descendente 210, el flujo de bits SAOC 212 y el elemento del flujo de bits que representa el parámetro de combinación lineal, da a un codificador de señal de audio la oportunidad de controlar parcialmente el mecanismo de control de distorsión, que se realiza en el lado del descodificador SAOC.

50 Hay dos versiones posibles de la matriz objetivo sin distorsión $\mathbf{M}_{ren,tar}^{l,m}$, adecuadas para aplicaciones

diferentes. Ésta es controlada por el elemento del flujo de bits “bsDcuMode”:

• (“bsDcuMode” = 0): La reproducción “similar a mezcla descendente”, en donde $\mathbf{M}_{\text{ren,tar}}^{l,m}$ corresponde a la matriz de mezcla descendente normalizada de energía.

• (“bsDcuMode” = 1): La reproducción de “mejor esfuerzo”, en donde $\mathbf{M}_{\text{ren,tar}}^{l,m}$ se define como una función tanto de matriz de reproducción de mezcla descendente como especificada por el usuario.

En resumen, hay dos modos de control de distorsión llamados reproducción “similar a mezcla descendente” y reproducción de “mejor esfuerzo”, que se pueden seleccionar de acuerdo con los elementos del flujo de bits “bsDcuMode”. Estos dos modos difieren en la manera en que se computa su matriz de reproducción objetivo. A continuación, se describirán con detalle aspectos sobre la computación de la matriz de reproducción objetivo para los dos modos de reproducción “similar a mezcla descendente” y reproducción de “mejor esfuerzo”.

3.3 Reproducción “similar a mezcla descendente”

3.3.1 Introducción

El método de reproducción “similar a mezcla descendente” se puede utilizar típicamente en los casos en que la mezcla descendente es una importante referencia de alta calidad artística. La matriz de reproducción “similar a mezcla descendente” $\mathbf{M}_{\text{ren,DS}}^l$ se computa como

$$\mathbf{M}_{\text{ren,DS}}^l = \mathbf{M}_{\text{ren,tar}}^l = \sqrt{N_{DS}^l} \mathbf{D}_{DS}^l,$$

en donde N_{DS}^l representa un valor escalar de normalización de energía (para cada ranura de parámetro l) y \mathbf{D}_{DS}^l es la matriz de mezcla descendente \mathbf{D}^l extendida por filas de elementos cero de manera que el número y el orden de las filas corresponden a la constelación de $\mathbf{M}_{\text{ren}}^{l,m}$.

Por ejemplo, en el modo estéreo SAOC a transcodificación multicanal $N_{MPS} = 6$. Por consiguiente \mathbf{D}_{DS}^l es de tamaño $N_{MPS} \times N$ (en donde N representa el número de objetos de audio de entrada) y sus filas representan los canales de salida frontal izquierdo y derecho iguales a \mathbf{D}^l (o corresponden a las filas de \mathbf{D}^l).

Para facilitar la comprensión de lo anterior, se deben considerar las siguientes definiciones de la matriz de reproducción y de la matriz de mezcla descendente.

La matriz de reproducción (modificada) $\mathbf{M}_{\text{ren,lim}}$ aplicada a los objetos de audio de entrada \mathbf{S} determina la salida reproducida objetivo como $\mathbf{Y} = \mathbf{M}_{\text{ren,lim}}$. La matriz de reproducción (modificada) $\mathbf{M}_{\text{ren,lim}}$ con elementos $m_{i,j}$ mapea todos los objetos de entrada i (es decir, objetos de entrada que tienen índice de objeto i) para los canales de salida deseados j (es decir, los canales de salida que tienen índice de canal j). La matriz de reproducción (modificada) $\mathbf{M}_{\text{ren,lim}}$ se da por

5

$$\mathbf{M}_{ren,lim} = \begin{pmatrix} m_{0,Lf} & \dots & m_{N-1,Lf} \\ m_{0,Rf} & \dots & m_{N-1,Rf} \\ m_{0,C} & \dots & m_{N-1,C} \\ m_{0,Lfe} & \dots & m_{N-1,Lfe} \\ m_{0,Ls} & \dots & m_{N-1,Ls} \\ m_{0,Rs} & \dots & m_{N-1,Rs} \end{pmatrix}, \text{ para la configuración de salida 5.1,}$$

10

$$\mathbf{M}_{ren,lim} = \begin{pmatrix} m_{0,L} & \dots & m_{N-1,L} \\ m_{0,R} & \dots & m_{N-1,R} \end{pmatrix}, \text{ para la configuración de salida estéreo,}$$

15

$$\mathbf{M}_{ren,lim} = (m_{0,C} \dots m_{N-1,C}), \text{ para la configuración de salida mono.}$$

Típicamente las mismas dimensiones se aplican también a la matriz de reproducción especificada por el usuario \mathbf{M}_{ren} y la matriz de reproducción objetivo $\mathbf{M}_{ren,tar}$.

La matriz de mezcla descendente \mathbf{D} aplicada a los objetos de audio de entrada \mathbf{S} (en un descodificador de audio) determina la señal de mezcla descendente como $\mathbf{X} = \mathbf{D}\mathbf{S}$.

Para el caso de mezcla descendente estéreo, la matriz de mezcla descendente \mathbf{D} de tamaño $2 \times N$ (también designada con \mathbf{D}^l , para mostrar una posible dependencia del tiempo) con elementos $d_{i,j}$ se obtiene ($i = 0, 1; j = 0, \dots, N - 1$) (en un descodificador de audio) de los parámetros DMG y DCLD como

$$d_{0,j} = 10^{0.05DMG_j} \sqrt{\frac{10^{0.1DCLD_j}}{1 + 10^{0.1DCLD_j}}}, \quad d_{1,j} = 10^{0.05DMG_j} \sqrt{\frac{1}{1 + 10^{0.1DCLD_j}}}.$$

Para el caso de la mezcla descendente mono, se obtiene la matriz de mezcla descendente \mathbf{D} de tamaño $1 \times N$ con elementos $d_{i,j}$ ($i = 0; j = 0, \dots, N - 1$) (en un descodificador de audio) a partir de los parámetros DMG como

$$d_{0,j} = 10^{0.05DMG_j}.$$

Los parámetros de mezcla descendente DMG y DCLD se obtienen del flujo de bits SAOC 212.

3.3.2 Computación del Valor Escalar de Normalización de Energía para todos los modos SAOC de descodificación/transcodificación

Para todos los modos SAOC de descodificación/transcodificación el valor escalar de normalización de energía N_{DS}^l se computa utilizando la siguiente ecuación:

$$N_{DS}^l = \frac{\text{traza} \left(\mathbf{M}_{ren}^{l,m} \left(\mathbf{M}_{ren}^{l,m} \right)^* \right) + \varepsilon}{\text{traza} \left(\mathbf{D}^l \left(\mathbf{D}^l \right)^* \right) + \varepsilon}$$

3.4 Reproducción de “mejor esfuerzo”

3.4.1 Introducción

5 El método de reproducción de “mejor esfuerzo” se puede utilizar típicamente en los casos en que la reproducción objetivo es una referencia importante.

10 La matriz de reproducción de “mejor esfuerzo” describe una matriz de reproducción objetivo, que depende de la mezcla descendente y la información de reproducción. La normalización de energía se representa por una matriz $\mathbf{N}_{BE}^{l,m}$ de tamaño $N_{MPS} \times M$, por lo que proporciona valores individuales para cada canal de salida. Esto requiere diferentes cálculos de $\mathbf{N}_{BE}^{l,m}$ para los diferentes modos de operación SAOC, que se exponen más adelante. La matriz de reproducción de “mejor esfuerzo” se computa como

$$15 \quad \mathbf{M}_{ren,BE}^l = \mathbf{M}_{ren,tar}^l = \sqrt{\mathbf{N}_{BE}^l} \mathbf{D}^l, \text{ para los siguientes modos SAOC “x-1-1/2/5/b”, “x-2-1/b”,}$$

$$\mathbf{M}_{ren,BE}^l = \mathbf{M}_{ren,tar}^l = \mathbf{N}_{BE}^l \mathbf{D}^l, \text{ para los siguientes modos SAOC “x-2-2/5”.$$

Aquí \mathbf{D}^l es la matriz de mezcla descendente y $\mathbf{N}_{BE}^{l,m}$ representa la matriz de la normalización de energía.

20 El operador raíz cuadrada en la ecuación anterior designa una formación de raíz cuadrada en forma de elemento.

25 A continuación, la computación del valor \mathbf{N}_{BE}^l , que puede ser un valor escalar de normalización de energía en el caso de un modo de descodificación SAOC mono a mono, y que puede ser una matriz de normalización de energía en el caso de otros modos de descodificación o modos de transcodificación, serán analizados en detalle.

3.4.2 Modo de descodificación SAOC mono a mono (“x-1-1”)

30 Para el modo SAOC “x-1-1” en donde se descodifica una señal de mezcla descendente mono para obtener una señal de salida mono (como una representación de señal de mezcla ascendente), el valor escalar de normalización de energía $\mathbf{N}_{BE}^{l,m}$ se calcula utilizando la siguiente ecuación

$$\mathbf{N}_{BE}^{l,m} = \frac{\sum_{j=0}^{N-1} (m_{j,0}^{l,m})^2 + \varepsilon}{\sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon}$$

3.4.3 Modo de descodificación SAOC mono a estéreo (“x-1-2”)

40 Para el modo SAOC “x-1-2”, en donde se descodifica una señal de mezcla descendente mono para obtener una salida estéreo (2 canales) (como una representación de señal de mezcla ascendente), la matriz de normalización de energía $\mathbf{N}_{BE}^{l,m}$ de tamaño 2×1 se computa utilizando la siguiente ecuación

$$45 \quad \mathbf{N}_{BE}^{l,m} = \left(\frac{\sum_{j=0}^{N-1} (m_{j,0}^{l,m})^2 + \varepsilon}{\sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon}, \frac{\sum_{j=0}^{N-1} (m_{j,1}^{l,m})^2 + \varepsilon}{\sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon} \right)^T.$$

3.4.4 Modo de decodificación SAOC mono a binaural (“x-1-b”)

5 Para el modo SAOC “x-1-b”, en donde se decodifica una señal de mezcla descendente mono para obtener una señal de salida reproducida binaural (como una representación de señal de mezcla ascendente), la matriz de normalización de energía $\mathbf{N}_{BE}^{l,m}$ de tamaño 2×1 se computa utilizando la siguiente ecuación

10
$$\mathbf{N}_{BE}^{l,m} = \left(\begin{array}{c} \sum_{j=0}^{N-1} a_{j,1}^{l,m} (a_{j,1}^{l,m})^* + \varepsilon \\ \sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon \end{array} , \begin{array}{c} \sum_{j=0}^{N-1} a_{j,2}^{l,m} (a_{j,2}^{l,m})^* + \varepsilon \\ \sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon \end{array} \right)^T$$

15 Los elementos $a_{x,y}^{l,m}$ comprenden (o se toman de) la matriz de reproducción binaural objetivo $\mathbf{A}^{l,m}$.

3.4.5 Modo de decodificación SAOC estéreo a mono (“x-2-1”)

20 Para el modo SAOC “x-2-1”, en donde se decodifica una señal de mezcla descendente (estéreo) de dos canales para obtener una señal de salida (mono) de un canal (como una representación de señal de mezcla ascendente), la matriz de normalización de energía $\mathbf{N}_{BE}^{l,m}$ de tamaño 1×2 se computa usando la siguiente ecuación

25
$$\mathbf{N}_{BE}^{l,m} = \mathbf{M}_{\text{ren}}^{l,m} (\mathbf{D}^l)^* \mathbf{J}^l$$
,

en donde $\mathbf{M}_{\text{ren}}^{l,m}$ es la matriz de reproducción mono de tamaño $1 \times N$.

3.4.6 Modo de decodificación SAOC estéreo a estéreo (“x-2-2”)

30 Para el modo SAOC “x-2-2”, en donde se decodifica una señal de mezcla descendente estéreo para obtener una señal de salida estéreo (como una representación de señal de mezcla ascendente), la matriz de normalización de energía $\mathbf{N}_{BE}^{l,m}$ de tamaño 2×2 se computa usando la siguiente ecuación

35
$$\mathbf{N}_{BE}^{l,m} = \mathbf{M}_{\text{ren}}^{l,m} (\mathbf{D}^l)^* \mathbf{J}^l$$
,

en donde $\mathbf{M}_{\text{ren}}^{l,m}$ es la matriz de reproducción estéreo de tamaño $2 \times N$.

40 3.4.7 Modo de decodificación SAOC estéreo a binaural (“x-2-b”)

Para el modo SAOC “x-2-b”, en donde se decodifica una señal de mezcla descendente estéreo para obtener una señal de salida reproducida binaural (como una representación de señal de mezcla ascendente), la matriz de normalización de energía $\mathbf{N}_{BE}^{l,m}$ de tamaño 2×2 se computa usando la siguiente ecuación

45
$$\mathbf{N}_{BE}^{l,m} = \mathbf{A}^{l,m} (\mathbf{D}^l)^* \mathbf{J}^l$$
,

en donde $\mathbf{A}^{l,m}$ es una matriz de reproducción binaural de tamaño $2 \times N$.

3.4.8 Modo de transcodificación SAOC mono a multicanal ("x-1-5")

5 Para el modo SAOC "x-1-5", en donde se transcodifica una señal de mezcla descendente mono para obtener una señal de salida de 5 canales o de 6 canales (como una representación de señal de mezcla ascendente), la matriz de normalización de energía $\mathbf{N}_{BE}^{l,m}$ de tamaño $N_{MPS} \times 1$ se computa usando la siguiente ecuación

$$10 \quad \mathbf{N}_{BE}^{l,m} = \left(\frac{\sum_{j=0}^{N-1} (m_{j,0}^{l,m})^2 + \varepsilon}{\sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon}, \dots, \frac{\sum_{j=0}^{N-1} (m_{j,N_{MPS}-1}^{l,m})^2 + \varepsilon}{\sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon} \right)^T .$$

20 3.4.9 Modo de transcodificación SAOC estéreo a multicanal ("x-2-5")

Para el modo SAOC "x-2-5", en donde se transcodifica una señal de mezcla descendente estéreo para obtener una señal de salida de 5 canales o de 6 canales (como una representación de señal de mezcla ascendente), la matriz de normalización de energía $\mathbf{N}_{BE}^{l,m}$ de tamaño $N_{MPS} \times 2$ se computa usando la siguiente ecuación

$$25 \quad \mathbf{N}_{BE}^{l,m} = \mathbf{M}_{ren}^{l,m} (\mathbf{D}^l)^* \mathbf{J}^l .$$

3.4.10 Computación de \mathbf{J}^l

30 Para evitar problemas numéricos cuando se calcula el término, $\mathbf{J}^l = (\mathbf{D}^l (\mathbf{D}^l)^*)^{-1}$ en los incisos 3.4.5, 3.4.6, 3.4.7 y 3.4.9, \mathbf{J}^l se modifica en algunas realizaciones. En primer lugar los valores propios $\lambda_{1,2}$ de \mathbf{J}^l se calculan, resolviendo $\det(\mathbf{J} - \lambda_{1,2} \mathbf{I}) = 0$.

35 Los valores propios se ordenan en orden descendente ($\lambda_1 \geq \lambda_2$) y el vector propio correspondiente al valor propio más grande se calcula de acuerdo con la ecuación anterior. Se debe asegurar de que se encuentren en el plano x positivo (el primer elemento tiene que ser positivo). El segundo vector propio se obtiene del primero por una rotación de -90 grados:

$$40 \quad \mathbf{J} = (\mathbf{v}_1 \mathbf{v}_2) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} (\mathbf{v}_1 \mathbf{v}_2)^* .$$

45 3.4.1 Aplicación de la Unidad de Control de Distorsión (DCU) para objetos de audio mejorados (EAO)

A continuación, se describirán algunas extensiones opcionales con respecto a la aplicación de la unidad de control de distorsión, que pueden ser implementadas en algunas realizaciones de acuerdo con la invención.

50 Para descodificadores SAOC que descodifican datos de codificación residuales y así soportan el manejo de EAOs, puede ser significativo el proporcionar una segunda parametrización de la DCU, que permite aprovechar al máximo la calidad de audio mejorada proporcionada por el uso de EAOs. Esto se logra

5 mediante la descodificación y el uso de un segundo conjunto alternativo de parámetros de la DCU (es decir bsDcuMode2 y bsDcuParam2) que adicionalmente se transmite como parte de las estructuras de datos que contienen los datos residuales (es decir SAOCExtensionConfigData() y SAOCExtensionFrameData()). Una aplicación puede hacer uso de este segundo conjunto de parámetros, si se descodifican los datos residuales de codificación y opera en modo estricto EAO que está definido por la condición de que solamente los EAOs pueden ser modificados arbitrariamente mientras que todos los no EAOs solamente pasan por una modificación común única. Específicamente, este modo estricto de EAO requiere el cumplimiento de dos condiciones siguientes:

10 La matriz de mezcla descendente y la matriz de reproducción tienen las mismas dimensiones (lo que implica que el número de canales de reproducción es igual al número de canales de mezcla descendente).

15 La aplicación solamente emplea coeficientes de reproducción para cada uno de los objetos regulares (es decir, no EAOs) que se relacionan con sus correspondientes coeficientes de mezcla descendente por un factor de escala común única.

4. Flujo de bits de acuerdo con la Figura 3a

20 En lo siguiente, se describirá un flujo de bits que representa una señal de audio multicanal tomando como referencia la Figura 3a que muestra una representación gráfica del flujo de bits 300.

25 El flujo de bits 300 comprende una representación de señal de mezcla descendente 302, que es una representación (por ejemplo, una representación codificada) de una señal de mezcla descendente combinando las señales de audio de una pluralidad de objetos de audio. El flujo de bits 300 comprende también una información lateral paramétrica relacionada con el objeto 304 que describe características del objeto de audio y, típicamente, también características de un mezcla descendente realizada en un codificador de audio. La información paramétrica relacionada con el objeto 304 comprende preferentemente una información de diferencia de nivel de objeto OLD, una información de correlación entre objetos IOC, una información de ganancia de mezcla descendente DMG y una información diferente de nivel de canal de mezcla descendente DCLD. El flujo de bits 300 también comprende un parámetro de combinación lineal 306 que describe contribuciones deseadas de una matriz de reproducción especificada por el usuario y de una matriz de reproducción objetivo a una matriz de reproducción modificada (para ser aplicado por un descodificador de señal de audio).

35 Más detalles opcionales acerca de este flujo de bits 300, que puede ser proporcionado por el aparato 150 como el flujo de bits 170, y que puede ser la entrada en el aparato 100 para obtener la representación de señal de mezcla descendente 110, la información paramétrica relacionada con el objeto 112 y el parámetro de combinación lineal 140, o en el aparato 200 para obtener la información de mezcla descendente 210, la información del flujo de bits SAOC 212 y el parámetro de combinación lineal 214, se describirán en lo siguiente haciendo referencia a las Figuras 3b y 3c.

5. Detalles de Sintaxis del Flujo de Bits

5.1. Sintaxis de Configuración Específica de SAOC

45 La Figura 3b muestra una representación detallada de la sintaxis de una información de configuración específica de SAOC.

50 La configuración específica de SAOC 310 de acuerdo con la Figura 3b, por ejemplo, puede formar parte de un encabezado del flujo de bits 300 de acuerdo con la Figura 3a.

55 La configuración específica de SAOC, por ejemplo, puede comprender una configuración de frecuencia de muestreo que describe una frecuencia de muestreo a ser aplicada por un descodificador SAOC. La configuración específica de SAOC comprende también una configuración de modo de bajo retardo que describe si se debe utilizar un modo de bajo retardo o un modo de alto retardo del procesador de señales 148 o de la unidad de descodificación/transcodificación SAOC 248. La configuración específica de SAOC comprende también una configuración de resolución de frecuencia que describe una resolución de frecuencia para ser utilizada por el procesador de señales 148 o por la unidad de descodificación/transcodificación SAOC 248. Además, la configuración específica de SAOC puede

comprender una configuración de longitud de cuadro que describe una longitud de cuadros de audio para ser utilizados por el procesador de señales 148, o por la unidad de descodificación/transcodificación SAOC 248. Además, la configuración específica de SAOC comprende típicamente una configuración del número de objetos que describe un número de objetos de audio para ser procesados por el procesador de señales 148, o por la unidad de descodificación/transcodificación SAOC 248. La configuración del número de objetos también describe una serie de parámetros relacionados con los objetos incluidos en la información paramétrica relacionada con el objeto 112, o en el flujo de bits SAOC 212. La configuración específica de SAOC puede comprender una configuración de relación de objetos, que designa a los objetos que tienen una información paramétrica común relacionada con el objeto. La configuración específica de SAOC también puede comprender una configuración de transmisión de energía absoluta, que indica si una información de energía absoluta se transmite desde un codificador de audio hacia un descodificador de audio. La configuración específica de SAOC puede comprender también una configuración del número de canales de mezcla descendente, que indica si solamente hay un canal de mezcla descendente, si hay dos canales de mezcla descendente, o si existen, opcionalmente, más de dos canales de mezcla descendente. Además, la configuración específica de SAOC puede comprender información de configuración adicional en algunas realizaciones.

La configuración específica de SAOC puede comprender también información de configuración de ganancia de mezcla descendente post-procesamiento “bsPdgFlag”, que define si se transmite una ganancia de mezcla descendente post-procesamiento para un post-procesamiento opcional.

La configuración específica de SAOC comprende también un indicador “bsDcuFlag” (que, por ejemplo, puede ser un indicador de 1 bit), que define si los valores “bsDcuMode” y “bsDcuParam” se transmiten en el flujo de bits. Si este indicador “bsDcuFlag” toma el valor de “1”, otro indicador que está marcado “bsDcuMandatory” y un indicador “bsDcuDynamic” se incluyen en la configuración específica de SAOC 310. El indicador “bsDcuMandatory” describe si el control de distorsión debe ser aplicado por un descodificador de audio. Si el indicador “bsDcuMandatory” es igual a 1, entonces la unidad de control de distorsión debe ser aplicada usando los parámetros “bsDcuMode” y “bsDcuParam” tal como se transmiten en el flujo de bits. Si el indicador “bsDcuMandatory” es igual a “0”, entonces los parámetros de la unidad de control de distorsión “bsDcuMode” y “bsDcuParam” transmitidos en el flujo de bits solamente son valores recomendados, y también se pueden usar otros ajustes de la unidad de control de distorsión. En otras palabras, un codificador de audio puede activar el indicador “bsDcuMandatory” con el fin de hacer cumplir el uso del mecanismo de control de distorsión en un descodificador de audio compatible con el estándar, y puede desactivar el indicador con el fin de dejar la decisión de aplicar o no la unidad control de distorsión, y si es así, cuáles parámetros utilizar para la unidad de control de distorsión, al descodificador de audio.

El indicador “bsDcuDynamic” permite una señalización dinámica de los valores “bsDcuMode” y “bsDcuParam”. Si el indicador “bsDcuDynamic” está desactivado, los parámetros “bsDcuMode” y “bsDcuParam” se incluyen en la configuración específica de SAOC, y en caso contrario, los parámetros “bsDcuMode” y “bsDcuParam” se incluyen en los marcos de SAOC, o, al menos, en algunos de los marcos de SAOC, como se discutirá más adelante. En consecuencia, un codificador de señal de audio puede cambiar entre una señalización de una sola vez (por pieza de audio que comprende una sola configuración específica de SAOC y, típicamente, una pluralidad de cuadros de SAOC) y una transmisión dinámica de los parámetros dentro de algunos o todos los cuadros de SAOC. El parámetro “bsDcuMode” define el tipo de matriz objetivo sin distorsión para la unidad de control de distorsión (DCU) de acuerdo con la tabla de la Figura 3d.

El parámetro “bsDcuParam” define el valor del parámetro para el algoritmo de la unidad de control de distorsión (DCU) de acuerdo con la tabla de la Figura 3e. En otras palabras, el parámetro de 4 bits “BsDcuParam” define un valor de índice idx, que puede ser mapeado por un descodificador de señal de audio en un valor de combinación lineal g_{DCU} (también designado con “DcuParam[ind]” o “DcuParam[idx]”). Así, el parámetro “bsDcuParam” representa, de una manera cuantificada, el parámetro de combinación lineal.

Como puede verse en la Figura 3b, los parámetros “bsDcuMandatory”, “bsDcuDynamic”, “bsDcuMode” y “bsDcuParam” se establecen en un valor por defecto de “0”, si el indicador “bsDcuFlag” toma el valor de “0”, lo que indica que no se transmiten parámetros de la unidad de control de distorsión.

La configuración específica de SAOC comprende también, opcionalmente, uno o más bits de alineación de bytes "ByteAlign()" para llevar la configuración específica de SAOC a una longitud deseada.

5 Además, la configuración específica de SAOC puede comprender opcionalmente una Configuración de extensión de SAOC "SAOCExtensionConfig()", que comprende parámetros de configuración adicionales. Sin embargo, los parámetros de configuración no son relevantes para la presente invención, de tal manera que aquí se omite una discusión en aras de la brevedad.

10 5.2. Sintaxis de Cuadro de SAOC

En lo siguiente, la sintaxis de un cuadro de SAOC se describirá tomando como referencia la Figura 3c.

15 El cuadro de SAOC "SAOCFrame" típicamente comprende valores de diferencia de nivel de objetos codificados OLD como se mencionó anteriormente, los cuales pueden ser incluidos en los datos del cuadro de SAOC para una pluralidad de bandas de frecuencia ("modo de banda") y para una pluralidad de objetos de audio (por objeto de audio).

20 El cuadro de SAOC, opcionalmente, también comprende valores de energía absoluta codificada NRG que pueden ser incluidos para una pluralidad de bandas de frecuencia (modo de banda).

El cuadro de SAOC puede comprender también valores de correlación entre objetos codificados IOC, que se incluyen en los datos de cuadro de SAOC para una pluralidad de combinaciones de objetos de audio. Los valores IOC se incluyen típicamente en una forma de banda.

25 El cuadro de SAOC comprende también valores de ganancia de mezcla descendente codificados DMG, en donde típicamente existe un valor de ganancia de mezcla descendente por objeto de audio por cuadro de SAOC.

30 El cuadro de SAOC comprende también, opcionalmente, diferencias de nivel de canal de mezcla descendente codificadas DCLD, en donde típicamente existe un valor de diferencia de nivel de canal de mezcla descendente por objeto de audio y por cuadro de SAOC.

También, el cuadro de SAOC típicamente comprende, opcionalmente, valores de ganancia de mezcla descendente codificados post-procesamiento PDG.

35 Además, un cuadro de SAOC puede comprender también, en algunas circunstancias, uno o más parámetros de control de distorsión. Si el indicador "bsDcuFlag", que se incluye en la sección de configuración específica de SAOC, es igual a "1", que indica el uso de información de la unidad de control de distorsión en el flujo de bits, y si el indicador "bsDcuDynamic" en la configuración específica de SAOC también toma el valor de "1", indicando el uso de una información de la unidad de control de distorsión dinámica (en forma de cuadro), la información de control de distorsión se incluye en el cuadro de SAOC, siempre que el cuadro de SAOC sea un llamado cuadro de SAOC "independiente", para que el indicador "bsIndependencyFlag" se active, o que el indicador "bsDcuDynamicUpdate" se active.

45 Debe señalarse aquí que el indicador "bsDcuDynamicUpdate" solamente se incluye en el cuadro de SAOC si el indicador "bsIndependencyFlag" está inactivo y que el indicador "bsDcuDynamicUpdate" define si los valores "bsDcuMode" y "bsDcuParam" están actualizados. Más precisamente, "bsDcuDynamicUpdate" = 1 significa que los valores "bsDcuMode" y "bsDcuParam" se actualizan en el cuadro actual, mientras que "bsDcuDynamicUpdate" = 0 significa que los valores transmitidos previamente se mantienen.

50 Por consiguiente, los parámetros "bsDcuMode" y "bsDcuParam", que se han explicado anteriormente, se incluyen en el cuadro de SAOC si la transmisión de los parámetros de la unidad de control de distorsión está activada y una transmisión dinámica de los datos de la unidad de control de distorsión también está activada y el indicador "bsDcuDynamicUpdate" está activado. Además, los parámetros "bsDcuMode" y "bsDcuParam" también se incluyen en el cuadro de SAOC si el cuadro de SAOC es un cuadro de SAOC "independiente", la transmisión de datos de la unidad de control de distorsión está activada y la transmisión dinámica de los datos de la unidad de control de distorsión también está activada.

El cuadro de SAOC comprende también, opcionalmente, datos de relleno "byteAlign()" para llenar el cuadro

de SAOC a una longitud deseada.

Opcionalmente, el cuadro de SAOC puede comprender información adicional, la cual se designa como "SAOCExt o ExtensionFrame()". Sin embargo, esta información adicional opcional del cuadro de SAOC no es relevante para la presente invención y, en aras de la brevedad, por lo tanto no se discutirá aquí.

Para completar, hay que señalar que el indicador "bsIndependencyFlag" indica si la codificación sin pérdida del cuadro de SAOC actual se hace independientemente del cuadro de SAOC previo, es decir, si el cuadro de SAOC actual puede ser descodificado sin el conocimiento del cuadro de SAOC previo.

6. Descodificador/transcodificador SAOC de acuerdo con la Figura 4

En lo siguiente, se describirán realizaciones adicionales de esquemas limitantes del coeficiente de reproducción para el control de distorsión en SAOC.

6.1 Perspectiva general

La Figura 4 muestra un diagrama esquemático de bloques de un descodificador de audio 400, de acuerdo con una realización de la invención.

El descodificador de audio 400 está configurado para recibir una señal de mezcla descendente 410, un flujo de bits de SAOC 412, un parámetro de combinación lineal 414 (también designado con Λ), y una información de matriz de reproducción 420 (también designada con R). El descodificador de audio 400 está configurado para recibir una representación de señal de mezcla ascendente, por ejemplo, en la forma de una pluralidad de canales de salida 130a a 130M. El descodificador de audio 400 comprende una unidad de control de distorsión 440 (también designada con DCU) que recibe al menos una parte de la información del flujo de bits de SAOC del flujo de bits de SAOC 412, el parámetro de combinación lineal 414 y la información de la matriz de reproducción 420. La unidad de control de distorsión proporciona una información de reproducción modificada R_{lim} que puede ser una información de la matriz de reproducción modificada.

El descodificador de audio 400 también comprende un descodificador SAOC y/o transcodificador SAOC 448, que recibe la señal de mezcla descendente 410, el flujo de bits de SAOC 412 y la información de reproducción modificada R_{lim} y proporciona, con base en el mismo, los canales de salida 130a a 130M.

En lo siguiente, la funcionalidad del descodificador de audio 400, que utiliza uno o más esquemas limitantes del coeficiente de reproducción de acuerdo con la presente invención, se discutirá en detalle.

El general procesamiento de SAOC se lleva a cabo de una manera selectiva de tiempo/frecuencia y se puede describir como sigue. El codificador SAOC (por ejemplo, el codificador SAOC 150) extrae las características psicoacústicas (por ejemplo, las relaciones y correlaciones de poder del objeto) de varias señales de objeto de audio de entrada y luego realiza una mezcla descendente de ellos en un canal mono combinado o estéreo (por ejemplo, la señal de mezcla descendente 182 o la señal de mezcla descendente 410). Esta señal de mezcla descendente y la información lateral extraída (por ejemplo, la información lateral paramétrica relacionada con el objeto o la información del flujo de bits de SAOC 412 se transmiten (o almacenan) en formato comprimido usando los codificadores de audio de percepción bien conocidos. En el extremo receptor, el descodificador SAOC 418 conceptualmente intenta restaurar las señales de los objetos originales (es decir, objetos mezclados descendentemente separados) usando la información lateral transmitida 412. Estas señales aproximadas de objetos después se mezclan en una escena objetivo utilizando una matriz de reproducción. La matriz de reproducción, por ejemplo, R o R_{lim} se compone de los Coeficientes de Reproducción (RCs) especificados para cada objeto de audio transmitido y altavoz de configuración de mezcla ascendente. Estos RCs determinan las ganancias y las posiciones espaciales de todos los objetos separados/reproducidos.

Efectivamente, la separación de las señales de objetos raramente o incluso nunca se ejecuta, ya que la separación y la mezcla se realizan en un solo paso de procesamiento combinado que da por resultado una enorme reducción de la complejidad computacional. Este esquema es tremendamente eficiente, tanto en términos de velocidad de transmisión de bits (solamente necesita transmitir uno o dos canales de mezcla descendente 182, 410, además de alguna información lateral 186, 188, 412, 414, en lugar de varias señales de objetos de audio individuales) y la complejidad computacional (la complejidad del procesamiento se

refiere principalmente al número de canales de salida en lugar del número de objetos de audio). El descodificador SAOC transforma (en un nivel paramétrico) las ganancias de objeto y otra información lateral directamente en los Coeficientes de Transcodificación (TCs) que se aplican a la señal de mezcla descendente 182, 414 para crear las señales correspondientes 130a a 130M para la escena de salida de audio reproducida (o la señal de mezcla descendente preprocesada por una operación de descodificación adicional, es decir, típicamente reproducción Surround MPEG multicanal).

La calidad de audio percibida subjetivamente de la escena de salida reproducida se puede mejorar mediante la aplicación de una unidad de control de distorsión DCU (por ejemplo, una unidad de modificación de matriz de reproducción), como se describe en [6]. Este mejoramiento se puede conseguir por el precio de aceptar una modificación dinámica moderada de las configuraciones de reproducción objetivo. La modificación de la información de reproducción se puede hacer con la variante del tiempo y la frecuencia, que en circunstancias específicas puede dar por resultado coloraciones sonoras no naturales y/o artefactos de fluctuación temporal.

Dentro del sistema SAOC global, la DCU se puede incorporar en la cadena de procesamiento del descodificador/transcodificador SAOC de manera sencilla. Específicamente, se coloca en el extremo frontal de la SAOC mediante el control de los RCs R , véase la Figura 4.

6.2 Hipótesis subyacente

La hipótesis subyacente del método de control indirecto considera una relación entre el nivel de distorsión y las desviaciones de los RCs desde su nivel de los objetos correspondientes en la mezcla descendente. Esto se basa en la observación de que la atenuación/elevación más específicas se aplican por los RCs a un objeto en particular con respecto a los otros objetos, la modificación más agresiva de la señal transmitida de mezcla descendente se va a realizar por el descodificador/transcodificador SAOC. En otras palabras: cuanto mayor es la desviación de los valores de "ganancia del objeto" que son relativos entre sí, mayor será la probabilidad de que se produzca distorsión inaceptable (suponiendo idénticos coeficientes de mezcla descendente).

6.3 Cálculo de los coeficientes de reproducción limitada

Con base en el escenario de la reproducción especificada por el usuario, representado por los coeficientes (los CRs) de una matriz R de tamaño $N_{ch} \times N_{ob}$ (es decir, las filas corresponden a los canales de salida 130a a 130M, las columnas a los objetos de audio de entrada), la DCU evita ajustes extremos de reproducción mediante la producción de una matriz modificada R_{lim} que comprende coeficientes de reproducción limitada, que son realmente utilizados por la máquina de reproducción SAOC 448. Sin pérdida de generalidad, en la descripción posterior, se supone que los RCs son de frecuencia invariante para simplificar la notación. Para todos los modos operativos de SAOC, los coeficientes de reproducción limitada pueden derivarse como

$$R_{lim} = (1 - \Lambda) R + \Lambda \tilde{R}$$

Esto significa que al incorporar el parámetro de desvanecimiento $\Lambda \in [0, 1]$ (también designado como un parámetro de combinación lineal), se puede realizar una mezcla de la matriz de reproducción (especificada por el usuario) R hacia una matriz objetivo \tilde{R} . En otras palabras, la matriz limitada R_{lim} representa una combinación lineal de la matriz de reproducción R y una matriz objetivo. Por un lado, la matriz de reproducción objetivo podría ser la matriz de mezcla descendente (es decir, los canales de mezcla descendente se pasan a través del transcodificador 448) con un factor de normalización u otra matriz estática que da por resultado una matriz de transcodificación estática. Esta "reproducción similar a mezcla descendente" asegura que la matriz de reproducción objetivo no introduzca ningún artefacto de procesamiento SAOC y por lo tanto representa un punto de reproducción óptima en términos de calidad de audio no obstante ser totalmente independiente de los coeficientes de reproducción inicial.

No obstante, si una aplicación exige un escenario de reproducción específica o un valor establecido alto por el usuario en su configuración de reproducción inicial (especialmente, por ejemplo, la posición espacial de uno o más objetos), la reproducción similar a mezcla descendente no sirve como punto objetivo. Por otro lado, un punto tal puede ser interpretado como "la reproducción de mejor esfuerzo" cuando se tiene en cuenta los coeficientes tanto de mezcla descendente como de reproducción inicial (por ejemplo, la matriz de reproducción especificada por el usuario). El objetivo de esta segunda definición de la matriz de reproducción objetivo es preservar el escenario de reproducción especificado (por ejemplo, definido por la matriz de reproducción especificada por el usuario) de una mejor manera posible, pero al mismo tiempo manteniendo la degradación audible debido a la excesiva manipulación del objeto en un nivel mínimo.

6.4 Reproducción Similar a Mezcla Descendente

6.4.1 Introducción

La matriz de mezcla descendente D de tamaño $N_{dmx} \times N_{ob}$ se determina por el codificador (por ejemplo, el codificador de audio 150) y comprende información sobre cómo se combinan linealmente los objetos de entrada en la señal de mezcla descendente que se transmite al descodificador. Por ejemplo, con una señal de mezcla descendente mono, D se reduce a un vector de una sola fila, y en el caso de mezcla descendente estéreo $N_{dmx} = 2$.

La matriz de "reproducción similar a mezcla descendente" R_{DS} se computa como

$$\tilde{R}(= R_{DS}) = N_{DS} D_R$$

en donde N_{DS} representa el valor escalar de normalización de energía y D_R es la matriz de mezcla descendente extendida por filas de elementos cero de modo que el número y el orden de las filas de D_R corresponden a la constelación de R . Por ejemplo, en el modo de transcodificación SAOC estéreo a multicanal (x-2-5) $N_{dmx} = 2$ y $N_{ch} = 6$. Por consiguiente D_R es de tamaño $N_{ch} \times N_{ob}$ y sus filas que representan los canales de salida frontal izquierdo y derecho son iguales a D .

6.4.2 Todos los modos de descodificación/transcodificación SAOC

Para todos los modos de descodificación/transcodificación SAOC el valor escalar de normalización de energía N_{DS} se puede computar utilizando la siguiente ecuación

$$N_{DS} = \frac{\text{traza}(RR^*) + \mathcal{E}}{\text{traza}(DD^*) + \mathcal{E}}$$

en donde el operador $\text{traza}(X)$, implica la suma de todos los elementos diagonales de la matriz X . El (*) implica el operador de transposición conjugado complejo.

6.5 Reproducción de mejor esfuerzo

6.5.1 Introducción

El método reproducción de mejor esfuerzo describe una matriz de reproducción objetivo, que depende de la información de mezcla descendente y de reproducción. La normalización de energía está representada por una matriz N_{BE} de tamaño $N_{ch} \times N_{dmx}$, por lo que proporciona valores individuales para cada canal de salida (siempre que exista más de un canal de salida). Esto requiere diferentes cálculos de N_{BE} para los diferentes modos de operación de SAOC, que se exponen en las secciones siguientes.

La matriz de "reproducción de mejor esfuerzo" se computa como

$$\tilde{R}(= R_{BE}) = N_{BE} D ,$$

en donde D es la matriz de mezcla descendente y N_{BE} representa la matriz de normalización de energía.

5

6.5.2 Modo de descodificación SAOC mono a mono ("x-1-1")

Para el modo SAOC "x-1-1" el valor escalar de normalización de energía N_{BE} se puede computar usando la siguiente ecuación

10

$$N_{BE} = \frac{\sum_{j=1}^{N_{ob}} r_{1,j}^2 + \varepsilon}{\sum_{j=1}^{N_{ob}} d_{1,j}^2 + \varepsilon}$$

15

20

6.5.3 Modo de descodificación SAOC mono a estéreo ("x-1-2")

Para el modo SAOC "x-1-2" la matriz de normalización de energía N_{BE} de tamaño 2×1 se puede computar usando la siguiente ecuación

25

$$N_{BE} = \begin{bmatrix} \frac{\sum_{j=1}^{N_{ob}} r_{1,j}^2 + \varepsilon}{\sum_{j=1}^{N_{ob}} d_{1,j}^2 + \varepsilon} & \frac{\sum_{j=1}^{N_{ob}} r_{2,j}^2 + \varepsilon}{\sum_{j=1}^{N_{ob}} d_{1,j}^2 + \varepsilon} \end{bmatrix}^T$$

30

35

6.5.4 Modo de descodificación SAOC mono a binaural ("x-1-b")

Para el modo SAOC "x-1-b" la matriz de normalización de energía N_{BE} de tamaño 2×1 se puede computar usando la siguiente ecuación

40

$$N_{BE} = \begin{bmatrix} \frac{\sum_{j=1}^{N_{ob}} r_{1,j}^2 + \varepsilon}{\sum_{j=1}^{N_{ob}} d_{1,j}^2 + \varepsilon} & \dots & \frac{\sum_{j=1}^{N_{ob}} r_{2,j}^2 + \varepsilon}{\sum_{j=1}^{N_{ob}} d_{1,j}^2 + \varepsilon} \end{bmatrix}^T$$

45

Se debe señalar además que aquí r_1 y r_2 consideran/incorporan información de parámetros de HRTF binaurales.

50

También se debe señalar que, para todas las 3 ecuaciones anteriores, se debe tomar la raíz cuadrada de N_{BE} , es decir

$$\tilde{R}(= R_{BE}) = \sqrt{N_{BE}} D$$

55

(véase la descripción anterior).

6.5.5 Modo de descodificación SAOC estéreo a mono ("x-2-1")

5 Para el modo SAOC "x-2-1" la matriz de normalización de energía N_{BE} de tamaño 1×2 se puede computar usando la siguiente ecuación

$$N_{BE} = R_1 D^* (DD^*)^{-1} ,$$

10 en donde la matriz de reproducción mono R_1 de tamaño $1 \times N_{ob}$ se define como

$$R_1 = \begin{bmatrix} r_{1,1} & \dots & r_{1,N_{ob}} \end{bmatrix} .$$

15 6.5.6 Modo de descodificación SAOC estéreo a estéreo ("x-2-2")

Para el modo SAOC "x-2-2" la matriz de normalización de energía N_{BE} de tamaño 2×2 se puede computar usando la siguiente ecuación

20
$$N_{BE} = R_2 D^* (DD^*)^{-1} ,$$

en donde la matriz de reproducción estéreo R_2 de tamaño $2 \times N_{ob}$ se define como

25
$$R_2 = \begin{bmatrix} r_{1,1} & \dots & r_{1,N_{ob}} \\ r_{2,1} & \dots & r_{2,N_{ob}} \end{bmatrix} .$$

30 6.5.7 Modo de descodificación SAOC mono a binaural ("x-2-b")

Para el modo SAOC "x-2-b" la matriz de normalización de energía N_{BE} de tamaño 2×2 se puede computar usando la siguiente ecuación

35
$$N_{BE} = R_2 D^* (DD^*)^{-1} ,$$

en donde la matriz de reproducción binaural R_2 de tamaño $2 \times N_{ob}$ se define como

40
$$R_2 = \begin{bmatrix} r_{1,1} & \dots & r_{1,N_{ob}} \\ r_{2,1} & \dots & r_{2,N_{ob}} \end{bmatrix} .$$

Se debe señalar además que aquí $r_{1,n}$ y $r_{2,n}$ consideran/incorporan información de parámetros de HRTF binaurales.

45

6.5.8 Modo de transcodificación SAOC mono a multicanal ("x-1-5")

50 Para el modo de SAOC "x-1-5" la matriz de normalización de energía N_{BE} de tamaño $N_{ch} \times 1$ se puede computar usando la siguiente ecuación

5

$$N_{BE} = \left[\begin{array}{c} \sum_{j=1}^{N_{ob}} r_{1,j}^2 + \varepsilon \\ \sum_{j=1}^{N_{ob}} d_{1,j}^2 + \varepsilon \end{array} , \dots , \left[\begin{array}{c} \sum_{j=1}^{N_{ob}} r_{N_{ch},j}^2 + \varepsilon \\ \sum_{j=1}^{N_{ob}} d_{N_{ch},j}^2 + \varepsilon \end{array} \right]^T \right]$$

De nuevo, tomar la raíz cuadrada para cada elemento, se recomienda o incluso se requiere en algunos casos.

10 6.5.9 Modo de transcodificación SAOC estéreo a multicanal (“x-2-5”)

Para el modo de SAOC “x-2-5” la matriz de normalización de energía N_{BE} de tamaño $N_{ch} \times 2$ se puede computar usando la siguiente ecuación

15

$$N_{BE} = RD^* (DD^*)^{-1}$$

20 6.5.10 Cálculo del $(DD^*)^{-1}$

Para el cálculo del término $(DD^*)^{-1}$ se pueden aplicar métodos de regularización para prevenir los malos resultados de la matriz planteada.

25 6.6 Control de los esquemas limitantes de coeficientes de reproducción

6.6.1 Ejemplo de sintaxis del flujo de bits

30 En lo siguiente se describirá una representación sintáctica de una configuración específica de SAOC tomando como referencia la Figura 5a. La configuración específica de SAOC “SAOCSpecificConfig()” comprende la información de configuración convencional de SAOC. Además, la configuración específica de SAOC comprende una adición específica de DCU 510, que se describirá con mayor detalle más adelante. La configuración específica de SAOC también comprende uno o más bits de relleno “ByteAlign()”, que pueden ser utilizados para ajustar la longitud de la configuración específica de SAOC. Además, la configuración específica de SAOC puede comprender opcionalmente la configuración de extensión de SAOC, que comprende más parámetros de configuración.

35 La adición específica de DCU 510 de acuerdo con la Figura 5a, para el elemento de sintaxis del flujo de bits “SAOCSpecificConfig()” es un ejemplo de señalización de flujo de bits para el esquema de DCU propuesto. Esto se relaciona con la sintaxis descrita en la subcláusula “5.1 campos de información para SAOC” del proyecto de Estándar de SAOC de acuerdo con la referencia [8].

40 A continuación, se dará la definición de algunos de los parámetros.

5

“bsDcuFlag”	Define si las configuraciones de la DCU son determinadas por el codificador o descodificador/transcodificador SAOC. Más precisamente, “bsDcuFlag” = 1 significa que los valores “bsDcuMode” y “bsDcuParam” especificados en la SAOCSpecificConfig() por el codificador SAOC se aplican a la DCU, mientras que “bsDcuFlag” = 0 significa que las variables “bsDcuMode” y “bsDcuParam” (inicializadas por los valores por defecto) pueden ser modificadas por la aplicación de descodificador/transcodificador SAOC o usuario.
“bsDcuMode”	Define el modo de la DCU. Más precisamente, “bsDcuMod” = 0 significa que el modo de reproducción “similar a mezcla descendente” es aplicado por la DCU, mientras que “bsDcuMode” = 1 que el modo de reproducción de “mejor esfuerzo” es aplicado por el algoritmo de la DCU.
“bsDcuParam”	Define el valor del parámetro de mezcla para el algoritmo de la DCU, en donde la tabla de la Figura 5b muestra una tabla de cuantificación para los parámetros “bsDcuParam” .

Los posibles valores “**bsDcuParam**” son en este ejemplo, parte de una tabla con 16 entradas representadas por 4 bits. Por supuesto, se podría utilizar cualquier tabla, más grande o más pequeña. El espaciamiento entre los valores puede ser logarítmico a fin de corresponder a la separación máxima del objeto en decibeles. Pero los valores también podrían ser linealmente espaciados, o una combinación híbrida de logarítmicos y lineales, o cualquier otro tipo de escala.

El parámetro “**bsDcuMode**” en el flujo de bits hace posible en el lado del codificador elegir, de acuerdo con la situación, el algoritmo óptimo de DCU. Esto puede ser muy útil, ya que algunas aplicaciones o contenidos podrían beneficiarse del modo de reproducción “similar a mezcla descendente”, mientras que otras podrían beneficiarse del modo de reproducción de “mejor esfuerzo”.

Típicamente, el modo de reproducción “similar a mezcla descendente” puede ser el método deseado para aplicaciones en donde la compatibilidad hacia atrás/hacia adelante es importante y la mezcla descendente tiene importantes cualidades artísticas que necesitan ser preservadas. Por otro lado, el modo de reproducción de “mejor esfuerzo” puede tener un mejor desempeño en los casos en que este no sea el caso.

Estos parámetros de la DCU relacionados con la presente invención, por supuesto, podrían ser transmitidos en cualquier otra parte del flujo de bits de SAOC. Una ubicación alternativa sería utilizar el contenedor “SAOCExtensionConfig()” en donde se podría utilizar un ID de cierta extensión. Ambas secciones se encuentran en el encabezado de SAOC, asegurando mínima sobrecarga a la velocidad de datos.

Otra alternativa consiste en transmitir los datos de la DCU en los datos de campos de información (es decir, en SAOCFrame()). Esto permitiría la señalización variante del tiempo (por ejemplo, control adaptable de señal).

Un procedimiento flexible es definir la señalización del flujo de bits de los datos de la DCU tanto para encabezado (es decir, señalización estática) como en los datos de campos de información (es decir señalización dinámica). A continuación, un codificador SAOC es libre de elegir uno de los dos métodos de señalización.

6.7 Estrategia de Procesamiento

En el caso de que las configuraciones de la DCU (por ejemplo, el modo de DCU “**bsDcuMode**” y la configuración del parámetro de mezcla “**bsDcuParam**”) se especifiquen explícitamente por el codificador SAOC (por ejemplo, “**bsDcuFlag**” = 1), el descodificador/transcodificador SAOC aplica estos valores directamente a la DCU. Si las configuraciones de la DCU no se especifican explícitamente (por ejemplo, “**bsDcuFlag**” = 0) el descodificador/transcodificador SAOC utiliza los valores por defecto y permite la aplicación del descodificador/transcodificador SAOC o que el usuario los modifique. El primer índice de cuantificación (por ejemplo idx = 0) se puede utilizar para deshabilitar la DCU. Alternativamente, el valor por defecto de la DCU (“**bsDcuParam**”) puede ser “0”, es decir deshabilitar la DCU o “1”, es decir limitación completa.

7. Evaluación del Desempeño

7.1 Diseño de la prueba de audición

Se llevó a cabo una prueba de audición subjetiva para evaluar el desempeño de percepción del concepto de DCM propuesto y fue comparada con los resultados del procesamiento de descodificación/transcodificación RM SAOC. En comparación con otras pruebas de audición, la tarea de esta prueba consiste en considerar la mejor calidad de reproducción posible en situaciones extremas de reproducción (“objetos solistas”, “objetos mudos”) en relación con dos aspectos de la calidad:

1. alcanzar el objetivo de la reproducción (una buena atenuación/elevación de los objetos objetivo)
2. la calidad del sonido de la escena en general (considerando las distorsiones, artefactos, elementos antinaturales, etc.).

Por favor, tenga en cuenta que un procesamiento sin modificar SAOC puede cumplir el aspecto # 1, pero no el aspecto # 2, mientras que el simple uso de la señal transmitida de mezcla descendente puede cumplir el

aspecto # 2, pero no el aspecto # 1.

5 La prueba de audición se llevó a cabo al presentar solamente opciones verdaderas al oyente, es decir, solamente material que realmente está disponible como una señal en el lado del descodificador. Así, las señales presentadas son la señal de salida del descodificador SAOC regular (sin procesar por la DCU), demostrando el desempeño de línea de base de la SAOC y la salida de SAOC/DCU. Además, el caso de la reproducción trivial, que corresponde a la señal de mezcla descendente, se presenta en la prueba de audición.

10 La tabla de la Figura 6a describe las condiciones de la prueba de audición.

Ya que la DCU propuesta opera utilizando los datos regulares de la SAOC y los mezcla descendentemente y no se basa en información residual, no se aplicó codificador de núcleo a las correspondientes señales de mezcla descendente de la SAOC.

15 7.2 Elementos de prueba de audición

Los siguientes elementos, junto con la reproducción extrema y crítica han sido elegidos para la prueba de audición actual del material de prueba de audición CfP.

20 a tabla de la Figura 6b describe los elementos de audio de las pruebas de audición.

7.3 Ajustes de mezcla descendente y reproducción

25 Las ganancias de los objetos de reproducción que se describen en la tabla de la Figura 6c se han aplicado para los escenarios considerados de mezcla ascendente.

7.4 Instrucciones de prueba de audición

30 Las pruebas subjetivas de audición se llevaron a cabo en una sala de audición acústicamente aislada que está diseñada para permitir la audición de alta calidad. El playback se hizo usando auriculares (STAX SR Lambda Pro con Convertidor Lake-People D/A y Monitor STAX SRM).

35 El método de prueba siguió el procedimiento utilizado en las pruebas de verificación de audio espacial, similares al método de "Estímulo Múltiple con Referencia Oculta y Anclas" (MUSHRA, por sus siglas en inglés) para la evaluación subjetiva de la calidad de audio intermedia [2]. El método de prueba se ha modificado como se describió anteriormente con el fin de evaluar el rendimiento de percepción de la DCU propuesta. Los oyentes fueron instruidos para cumplir con las siguientes instrucciones de prueba de audición:

40 "Escenario de aplicación: Imagínese que usted es el usuario de un sistema de remezcla de música interactiva que le permite realizar remezclas dedicadas de material musical. El sistema proporciona controles deslizantes de mezcla de estilo de escritorio para cada instrumento, para cambiar su nivel, posición espacial, etc. Debido a la naturaleza del sistema, algunas mezclas de sonidos extremos pueden conducir a la distorsión que degrada la calidad general del sonido. Por otro lado, las mezclas de sonido con los niveles de instrumentos similares tienden a producir una mejor calidad de sonido.

50 El objetivo de esta prueba es evaluar diferentes algoritmos de procesamiento con respecto a su impacto en la fuerza de la modificación del sonido y una calidad de sonido.

¡No hay ninguna "Señal de referencia" en esta prueba! Más bien, una descripción de las mezclas de sonido deseadas se da a continuación.

55 Para cada elemento de audio, por favor:

- primero lea la descripción de las mezclas de sonido que desee, como a un usuario del sistema le gustaría lograr

5	Elemento "CaféNegro":	Sección de metales suaves dentro de la mezcla de sonido
10	Elemento "VozSobreMúsica":	Música de fondo suave
15	Elemento "Audición":	Sonido vocal fuerte y música suave
20	Elemento "AmorPop":	Sección de cuerdas suaves dentro de la mezcla de sonido

- 25
- posteriormente, clasificar las señales usando un grado común para describir tanto
 - alcanzar el objetivo de reproducción de la mezcla deseada del sonido como
 - la calidad de sonido de la escena global (considerando las distorsiones, artefactos, elementos antinaturales, distorsiones espaciales, ...)".
- 30

35

Un total de 8 oyentes participaron en cada una de las pruebas realizadas. Todos los sujetos pueden ser considerados como oyentes experimentados. Las condiciones de prueba fueron aleatorizadas automáticamente para cada elemento de prueba y para cada oyente. Las respuestas subjetivas fueron registradas por un programa informático de prueba de audición basada en una escala que osciló de 0 a 100, con cinco intervalos etiquetados de la misma manera como en la escala MUSHRA. Se permitió un cambio instantáneo entre los elementos a prueba.

40

7.5 Resultados de la prueba de audición

45

Las gráficas mostradas en la representación gráfica de la Figura 7 muestran el puntaje promedio por elemento sobre todos los oyentes y el valor medio estadístico sobre todos los elementos evaluados junto con los intervalos de confianza asociados del 95%.

50

Se pueden hacer las siguientes observaciones con base en los resultados de las pruebas realizadas de audición: Para la prueba realizada de audición, las puntuaciones obtenidas de MUSHRA demostraron que la funcionalidad de la DCU propuesta ofrece un desempeño significativamente mejor en comparación con el sistema regular de RM SAOC en el sentido de los valores medios estadísticos globales. Hay que señalar que la calidad de todos los elementos producidos por el descodificador regular SAOC (que muestran fuertes artefactos de audio para las condiciones de reproducción extremas consideradas) se clasifica tan baja como la calidad de las configuraciones de reproducción idéntica a mezcla descendente que no cumple del todo con el escenario de la reproducción deseada. Por lo tanto, se puede concluir que los métodos propuestos de DCU conducen a un mejoramiento considerable de la calidad de la señal subjetiva para todos los escenarios de prueba de audición considerados.

55

8. Conclusiones

5 Para resumir la discusión anterior, se describieron esquemas que limitan el coeficiente de reproducción para el control de distorsión en SAOC. Las realizaciones de acuerdo con la invención pueden utilizarse en combinación con técnicas paramétricas para la transmisión/almacenamiento eficiente en cuanto a flujo de bits de escenas de audio que contienen múltiples objetos de audio, que recientemente se han propuesto (por ejemplo, véanse las referencias [1], [2], [3], [4] y [5]).

10 En combinación con la interactividad del usuario en el lado receptor, tales técnicas pueden convencionalmente (sin el uso de los esquemas que limitan el coeficiente de reproducción de la invención) conducir a una baja calidad de las señales de salida si se lleva a cabo la reproducción extrema del objeto (véase, por ejemplo, la referencia [6]).

15 La presente especificación se enfoca en la Codificación del objeto de Audio Espacial (SAOC), que proporciona los medios para que una interfaz de usuario seleccione la configuración de playback deseada (por ejemplo, mono, estéreo, 5.1, etc.) y la modificación interactiva en tiempo real de la escena de reproducción de salida deseada mediante el control de la matriz de reproducción de acuerdo con las preferencias personales u otros criterios. Sin embargo, la invención también es aplicable para técnicas paramétricas en general.

20 Debido al procedimiento paramétrico basado en la mezcla descendente/separación/mezcla, la calidad subjetiva de la salida de audio reproducida depende de las configuraciones de los parámetros de reproducción. La libertad de seleccionar las configuraciones de la reproducción de la elección del usuario conlleva el riesgo de que el usuario seleccione opciones de reproducción de objetos inapropiados, tales como manipulaciones de ganancia extrema de un objeto dentro de la escena del sonido global.

25 Para un producto comercial, es inaceptable por todos los medios el producir una mala calidad de sonido y/o artefactos de audio para cualquier configuración de la interfaz de usuario. Con el fin de controlar el deterioro excesivo de la salida de audio producida por SAOC, se han descrito varias medidas computacionales, que se basan en la idea de computar una medida de la calidad de percepción de la escena reproducida, y dependiendo de esta medida (y, opcionalmente, otra información), modificar los coeficientes de reproducción realmente aplicados (véase, por ejemplo, la referencia [6]).

30 El presente documento describe ideas alternativas para salvaguardar la calidad del sonido subjetivo de la escena reproducida de SAOC para la cual se lleva a cabo todo el procesamiento completo dentro del descodificador/transcodificador SAOC, y que no implica el cálculo explícito de sofisticadas medidas de calidad de audio percibida de la escena de sonido reproducida.

35 Por lo tanto, estas ideas se pueden implementar en una forma estructuralmente simple y extremadamente eficiente dentro del ambiente de trabajo del descodificador/transcodificador SAOC. El algoritmo de la Unidad de Control de Distorsión (DCU) propuesto tiene como objetivo limitar los parámetros de entrada del descodificador SAOC, es decir, los coeficientes de reproducción.

40 Para resumir lo anterior, las realizaciones de acuerdo con la invención crean un codificador de audio, un descodificador de audio, un método de codificación, un método de descodificación, y programas informáticos para la codificación o descodificación, o señales codificadas de audio como se describió anteriormente.

9. Alternativas de implementación

45 Aunque algunos aspectos han sido descritos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del método correspondiente, en donde un bloque o dispositivo corresponde a un paso del método o una característica de un paso del método. Análogamente, los aspectos descritos en el contexto de un paso del método también representan una descripción de un bloque o un elemento o característica correspondiente de un aparato correspondiente. Algunos o todos los pasos del método pueden ser ejecutados por (o usando) un aparato de hardware, como por ejemplo, un microprocesador, una computadora programable o un circuito electrónico.

En algunas realizaciones, alguno o más de los pasos del método más importantes pueden ser ejecutados por tal aparato.

La señal de audio codificada inventiva puede almacenarse en un medio de almacenamiento digital o puede transmitirse en un medio de transmisión tal como un medio de transmisión inalámbrica o un medio de transmisión por cable como Internet.

5

Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden ser implementadas en hardware o en software. La implementación se puede realizar utilizando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un Blue-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria instantánea (FLASH), que tengan señales de control legibles electrónicamente almacenadas en el mismo, que cooperen (o sean capaces de cooperar) con un sistema informático programable de tal manera que se realice el método respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible en computadora.

10

Algunas realizaciones de acuerdo con la invención comprenden un portador de datos que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema informático programable, de tal manera que se realiza uno de los métodos descritos en la presente.

15

Generalmente, las realizaciones de la presente invención pueden ser implementadas como un producto de programa informático con un código de programa, el código de programa es operativo para llevar a cabo uno de los métodos cuando el producto de programa informático se ejecuta en una computadora. El código de programa puede por ejemplo ser almacenado en un portador legible en máquina.

20

Otras realizaciones incluyen el programa informático para llevar a cabo uno de los métodos descritos en la presente, almacenado en un portador legible en máquina.

25

En otras palabras, una realización del método inventivo es, por lo tanto, un programa informático que tiene un código de programa para llevar a cabo uno de los métodos descritos en la presente, cuando el programa informático se ejecuta en una computadora.

30

Una realización adicional de los métodos inventivos es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible en computadora) que comprende, registrado en él, el programa informático para llevar a cabo uno de los métodos descritos en la presente. El portador de datos, el medio de almacenamiento digital o el medio registrado son típicamente tangibles y/o no transicionales.

35

Una realización adicional del método inventivo es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los métodos descritos en la presente. El flujo de datos o la secuencia de señales, por ejemplo puede ser configurado para ser transferido por medio de una conexión de comunicación de datos, por ejemplo a través de Internet.

40

Una realización adicional comprende un medio de procesamiento, por ejemplo una computadora, o un dispositivo lógico programable, configurado para adaptarse para realizar uno de los métodos descritos en la presente.

45

Una realización adicional comprende una computadora que tiene instalado en ella el programa informático para llevar a cabo uno de los métodos descritos en la presente.

50

En algunas realizaciones, se puede utilizar un dispositivo lógico programable (por ejemplo un conjunto de puertas programables de campo) para realizar todas o algunas de las funcionalidades de los métodos descritos en la presente. En algunas realizaciones, un conjunto de puertas programables de campo puede cooperar con un microprocesador con el fin de realizar uno de los métodos descritos en la presente. Generalmente, los métodos se realizan preferentemente por cualquier aparato de hardware.

55

Las realizaciones descritas anteriormente son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en la presente serán evidentes para otros expertos en la técnica. Es la intención, por lo tanto, que se limiten solamente por el alcance de las reivindicaciones de patente siguientes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en la presente.

REFERENCIAS

- 5
- [1] C. Faller and F. Baumgarte, "Binaural Cue Coding - Part IT Schemes and applications", IEEE Trans, on Speech and Audio Proc., vol. 11, no. 6, Nov. 2003.
- [2] C. Faller, "Parametric Joint-Coding of Audio Sources", 120ª Convención de AES, París, 2006, Preimpresión 6752.
- 10
- [3] J. Herre, S. Disch, J. Hilpert, O. Hellmuth: "From SAC To SAOC - Recent Developments in Parametric Coding of Spatial Audio", 22ª Conferencia Regional AES Reino Unido, Cambridge, Reino Unido, Abril 2007.
- [4] J. Engdegard, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Holzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers y W. Oomen: "Spatial Audio Object Coding (SAOC) - The Upcoming MPEG Standard on Parametric Object Based Audio Coding", 124ª Convención AES, Amsterdam 2008, Preimpresión 7377.
- 15
- [5] ISO/IEC, "MPEG audio technologies - Part 2: Spatial Audio Object Coding (SAOC)," ISO/IEC JTC1/SC29/WG11 (MPEG) FCD 23003-2.
- 20
- [6] Solicitud de patente de los Estados Unidos 61/173,456, METHODS, APPARATUS, AND COMPUTER PROGRAMS FOR DISTORTION AVOIDING AUDIO SIGNAL PROCESSING
- [7] EBU Technical recommendation: "MUSHRA-EBU Method for Subjective Listening Tests of Intermediate Audio Quality", Doc. B/AIM022, Octubre 1999.
- 25
- [8] ISO/IEC JTC1/SC29/WG11 (MPEG), Documento N10843, "Study on ISO/IEC 23003-2:200x Spatial Audio Object Coding (SAOC)", 89ª Junta de MPEG, Londres, Reino Unido, Julio 2009.
- [9] WO2008/100067 A1 (LG ELECTRONICS INC [KR]; OH HYEN O [KR]; JUNG YANG WON [KR]), 21 DE Agosto del 2008 (21-08-2008).
- 30
- [10] HERRE JÜRGEN ET AL.: "MPEG Surround-The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding", JAES, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, vol. 56, no. 11, 1 de Noviembre del 2008 (01-11-2008), páginas 932-955, XP040508729.
- 35

REIVINDICACIONES

1. Un aparato de procesamiento de audio (100; 200) para proporcionar una representación de señal de mezcla ascendente (130; 230) con base en una representación de señal de mezcla descendente (110; 210) y una información paramétrica relacionada con el objeto, que se incluyen en una representación de flujo de bits (300) de un contenido de audio, y en dependencia de una matriz de reproducción especificada por el usuario (144, \mathbf{M}_{ren}), el aparato está caracterizado porque comprende:

un limitador de distorsión (140; 240) configurado para obtener una matriz de reproducción modificada ($\mathbf{M}_{ren,lim}$) utilizando una combinación lineal de una matriz de reproducción especificada por el usuario (\mathbf{M}_{ren}) y una matriz de reproducción objetivo sin distorsión ($\mathbf{M}_{ren,tar}$) en dependencia de un parámetro de combinación lineal (146; g_{DCU}); y

un procesador de señales (148; 248) configurado para obtener la representación de señal de mezcla ascendente con base en la representación de señal de mezcla descendente y la información paramétrica relacionada con el objeto utilizando la matriz de reproducción modificada;

en donde el aparato está configurado para evaluar un elemento del flujo de bits (306; bsDcuParameter) que representa el parámetro de combinación lineal (146; g_{DCU}) para obtener el parámetro de combinación lineal.

2. El aparato (100; 200) de conformidad con la reivindicación 1, en donde el limitador de distorsión está configurado para obtener la matriz de reproducción objetivo ($\mathbf{M}_{ren,tar}$) de manera que la matriz de reproducción objetivo es una matriz de reproducción objetivo sin distorsión.

3. El aparato (100; 200) de conformidad con la reivindicación 1 o reivindicación 2, en donde el limitador de distorsión está configurado para obtener la matriz de reproducción modificada $\mathbf{M}_{ren,lim}^{l,m}$ de acuerdo con:

$$\mathbf{M}_{ren,lim}^{l,m} = (1 - g_{DCU}) \mathbf{M}_{ren}^{l,m} + g_{DCU} \mathbf{M}_{ren,tar}^{l,m}$$

en donde g_{DCU} designa el parámetro de combinación lineal, un valor del cual está en un intervalo [0,1];

en donde $\mathbf{M}_{ren}^{l,m}$ designa la matriz de reproducción especificada por el usuario; y

en donde $\mathbf{M}_{ren,tar}^{l,m}$ designa la matriz de reproducción objetivo.

4. El aparato (100; 200) de conformidad con una de las reivindicaciones 1 a 3, en donde el limitador de distorsión está configurado para obtener la matriz de reproducción objetivo ($\mathbf{M}_{ren,tar}$) de modo que la matriz de reproducción objetivo es una matriz de reproducción objetivo similar a mezcla descendente.

5. El aparato (100; 200) de conformidad con una de las reivindicaciones 1 a 4, en donde el limitador de distorsión está configurado para escalar una matriz de mezcla descendente ampliada (\mathbf{D}'_{DS}) utilizando un valor escalar de normalización de energía ($\sqrt{N_{DS}^l}$), para obtener la matriz de reproducción objetivo ($\mathbf{M}_{ren,tar}$), en donde la matriz de mezcla descendente ampliada es una versión ampliada de una matriz de mezcla descendente, una o más filas de cuya matriz de mezcla descendente describen contribuciones de una pluralidad de señales del objeto de audio a uno o más canales de la representación de señal de mezcla descendente, ampliada por filas de elementos cero, de manera que un número de filas de la matriz de mezcla descendente ampliada es idéntico a una constelación de reproducción descrita por la matriz de reproducción especificada por el usuario (\mathbf{M}_{ren}).

6. El aparato (100; 200) de conformidad con una de las reivindicaciones 1 a 3, en donde el limitador de distorsión está configurado para obtener la matriz de reproducción objetivo ($\mathbf{M}_{ren,tar}$), de manera que la matriz de reproducción objetivo es una matriz de reproducción objetivo de mejor esfuerzo.

7. El aparato (100; 200) de conformidad con una de las reivindicaciones 1 a 4 o 6, en donde el limitador de distorsión está configurado para obtener la matriz de reproducción objetivo ($\mathbf{M}_{ren,tar}$), de manera que la matriz de reproducción objetivo depende de una matriz de mezcla descendente (\mathbf{D}) y de la matriz de reproducción especificada por el usuario (\mathbf{M}_{ren}).

8. El aparato (100; 200) de conformidad con una de las reivindicaciones 1 a 3, 6 o 7, en donde el limitador de distorsión está configurado para computar una matriz (\mathbf{N}_{BE}) que comprende valores de normalización de energía individual del canal para una pluralidad de canales de audio de salida del aparato para proporcionar una representación de señal de mezcla ascendente, de tal manera que un valor de normalización de energía para un canal de audio de salida dado del aparato describe, al menos aproximadamente, una proporción entre una suma de valores de reproducción de energía asociados con el canal de audio de salida dado en la matriz de reproducción especificada por el usuario para una pluralidad de objetos de audio y una suma de valores de mezcla descendente de energía para la pluralidad de objetos de audio; y

5

en donde el limitador de distorsión está configurado para escalar un conjunto de valores de mezcla descendente utilizando un valor de normalización de energía individual del canal, para obtener un conjunto de valores de reproducción de la matriz de reproducción objetivo ($\mathbf{M}_{ren,tar}$) asociada con el canal de salida dado.

10

9. El aparato (100; 200) de conformidad con una de las reivindicaciones 1 a 3 y 6 a 8, en donde el limitador de distorsión está configurado para computar una matriz ($\mathbf{N}_{BE}^{l,m}$) que comprende valores de normalización de energía individual del canal para una pluralidad de canales de audio de salida de acuerdo con:

$$\mathbf{N}_{BE}^{l,m} = \begin{pmatrix} \frac{\sum_{j=0}^{N-1} (m_{j,0}^{l,m})^2 + \varepsilon}{\sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon} & \frac{\sum_{j=0}^{N-1} (m_{j,1}^{l,m})^2 + \varepsilon}{\sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon} \\ \dots & \dots \end{pmatrix}^T$$

15

para el caso de una representación de señal de mezcla descendente de 1 canal y una señal de salida de 2 canales del aparato; o

de acuerdo con:

$$\mathbf{N}_{BE}^{l,m} = \begin{pmatrix} \frac{\sum_{j=0}^{N-1} a_{j,1}^{l,m} (a_{j,1}^{l,m})^* + \varepsilon}{\sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon} & \dots & \frac{\sum_{j=0}^{N-1} a_{j,2}^{l,m} (a_{j,2}^{l,m})^* + \varepsilon}{\sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon} \end{pmatrix}^T$$

20

para el caso de una representación de señal de mezcla descendente de 1 canal y una señal de salida reproducida binaural del aparato; o

de acuerdo con:

$$\mathbf{N}_{BE}^{l,m} = \begin{pmatrix} \frac{\sum_{j=0}^{N-1} (m_{j,0}^{l,m})^2 + \varepsilon}{\sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon} & \dots & \frac{\sum_{j=0}^{N-1} (m_{j,N_{MPS}-1}^{l,m})^2 + \varepsilon}{\sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon} \end{pmatrix}^T$$

25

para el caso de una representación de señal de mezcla descendente de 1 canal y una señal de salida de canal \mathbf{N}_{MPS} del aparato;

30

en donde $m_{j,0}^{l,m}$ designa coeficientes de reproducción de la matriz de reproducción especificada por el usuario ($\mathbf{M}_{ren,lim}^{l,m}$) que describe una contribución deseada de un objeto de audio que tiene índice de objeto j a un primer canal de audio de salida del aparato;

en donde $m_{j,1}^{l,m}$ designa coeficientes de reproducción de la matriz de reproducción especificada por el usuario ($\mathbf{M}_{ren}^{l,m}$) que describe una contribución deseada de un objeto de audio que tiene índice de objeto j a un segundo canal de audio de salida del aparato;

5 en donde $a_{j,1}^{l,m}$ y $a_{j,2}^{l,m}$ designan los coeficientes de reproducción de la matriz de reproducción especificada por el usuario ($\mathbf{M}_{ren,lim}^{l,m}$) que describe una contribución deseada de un objeto de audio que tiene índice de objeto j a un primero y segundo canal de audio de salida del aparato, y tomando en consideración información de HRTF paramétrica;

10 en donde d_j^l designa un coeficiente de mezcla descendente que describe una contribución de un objeto de audio que tiene un índice de objeto j a la representación de la señal de mezcla descendente; y

en donde ε designa una constante aditiva para evitar la división por cero; y

15 en donde el limitador de distorsión está configurado para computar la matriz de reproducción objetivo $[\mathbf{M}_{ren,tar}^l]$ de acuerdo con:

$$\mathbf{M}_{ren,BE}^l = \mathbf{M}_{ren,tar}^l = \sqrt{\mathbf{N}_{BE}^l} \mathbf{D}^l,$$

20 en donde \mathbf{D}^l designa una matriz de mezcla descendente que comprende el coeficiente de mezcla descendente d_j .

10. El aparato (100; 200) de conformidad con una de las reivindicaciones 1 a 3 o 6 a 7, en donde el limitador de distorsión está configurado para computar una matriz que describe una normalización de energía individual del canal para una pluralidad de canales de audio de salida del aparato en dependencia de la matriz de reproducción especificada por el usuario (\mathbf{M}_{ren}), y una matriz de mezcla descendente \mathbf{D} ; y

25 en donde el limitador de distorsión está configurado para aplicar la matriz que describe la normalización de energía individual del canal para obtener un conjunto de coeficientes de reproducción de la matriz de reproducción objetivo ($\mathbf{M}_{ren,tar}$) asociada con un canal de audio de salida dado del aparato como una combinación lineal de conjuntos de valores de mezcla descendente asociados con diferentes canales de la representación de señal de mezcla descendente.

30 11. El aparato (100; 200) de conformidad con una de las reivindicaciones 1 a 3 o 6 a 7, o 10, en donde el limitador de distorsión está configurado para computar una matriz $\mathbf{N}_{BE}^{l,m}$ que describe la normalización de energía individual del canal para una pluralidad de canales de audio de salida de acuerdo con:

$$\mathbf{N}_{BE}^{l,m} = \mathbf{M}_{ren}^{l,m} (\mathbf{D}^l)^* \mathbf{J}^l$$

para el caso de una representación de señal de mezcla descendente de 2 canales y una señal de audio de salida multicanal del aparato;

40 en donde $\mathbf{M}_{ren}^{l,m}$ designa la matriz de reproducción especificada por el usuario que describe contribuciones deseadas especificadas por el usuario de una pluralidad de señales del objeto de audio a la señal de audio de salida multicanal del aparato;

45 en donde \mathbf{D}^l designa una matriz de mezcla descendente que describe contribuciones de una pluralidad de señales del objeto de audio a la representación de señal de mezcla descendente;

en donde

$$\mathbf{J}^l = (\mathbf{D}^l (\mathbf{D}^l)^*)^{-1}; \text{ y}$$

50 en donde el limitador de distorsión está configurado para computar la matriz de reproducción objetivo

$\mathbf{M}_{ren,tar}^l$ de acuerdo con

$$\mathbf{M}_{ren,BE}^l = \mathbf{M}_{ren,tar}^l = \mathbf{N}_{BE}^l \mathbf{D}^l .$$

5 12. El aparato (100; 200) de conformidad con las reivindicaciones 1 a 3 o 6 a 7, o 10, en donde el limitador de distorsión está configurado para computar una matriz $\mathbf{N}_{BE}^{l,m}$ de acuerdo con

$$\mathbf{N}_{BE}^{l,m} = \mathbf{M}_{ren}^{l,m} (\mathbf{D}^l)^* \mathbf{J}^l$$

para el caso de una representación de señal de mezcla descendente de 2 canales y una señal de audio de salida de 1 canal del aparato, o

10 de acuerdo con

$$\mathbf{N}_{BE}^{l,m} = \mathbf{A}^{l,m} (\mathbf{D}^l)^* \mathbf{J}^l$$

15 para el caso de una representación de señal de mezcla descendente de 2 canales y una señal de audio de salida reproducida binauralmente del aparato;

en donde $\mathbf{M}_{ren}^{l,m}$ designa la matriz de reproducción especificada por el usuario que describe contribuciones deseadas especificadas por el usuario de una pluralidad de señales del objeto de audio a la señal de salida del aparato;

20 en donde \mathbf{D}^l designa una matriz de mezcla descendente que describe contribuciones de una pluralidad de señales del objeto de audio a la representación de señal de mezcla descendente;

en donde $\mathbf{A}^{l,m}$ designa una matriz de reproducción binaural que está basada en la matriz de reproducción especificada por el usuario y parámetros de una función de transferencia relacionada con el encabezado.

25

13. El aparato (100; 200) de conformidad con una de las reivindicaciones 1 a 3 o 6 a 7, en donde el limitador de distorsión está configurado para computar un valor escalar de normalización de energía $N_{BE}^{l,m}$ de acuerdo con

$$N_{BE}^{l,m} = \frac{\sum_{j=0}^{N-1} (m_{j,0}^{l,m})^2 + \varepsilon}{\sum_{j=0}^{N-1} (d_j^l)^2 + \varepsilon}$$

30 en donde $m_{j,0}^{l,m}$ designa un coeficiente de reproducción de la matriz de reproducción especificada por el usuario ($\mathbf{M}_{ren}^{l,m}$) que describe una contribución deseada de un objeto de audio que tiene índice de objeto j a una señal de audio de salida del aparato;

35 en donde d_j^l designa un coeficiente de mezcla descendente que describe una contribución de un objeto de audio que tiene índice de objeto j a la representación de señal de mezcla descendente; y

en donde ε designa una constante aditiva para evitar la división por cero.

40 14. El aparato (100; 200) de conformidad con una de las reivindicaciones 1 a 13, en donde el aparato está configurado para leer un valor de índice (idx) que representa el parámetro de combinación lineal (g_{DCU}) de la representación del flujo de bits del contenido de audio y para mapear el valor del índice sobre el parámetro de combinación lineal (g_{DCU}) utilizando una tabla de cuantificación del parámetro.

45 15. El aparato (100; 200) de conformidad con la reivindicación 14, en donde la tabla de cuantificación describe una cuantificación no uniforme, en donde los valores más pequeños del parámetro de combinación lineal (g_{DCU}), que describen una contribución mayor de la matriz de reproducción especificada por el usuario (\mathbf{M}_{ren}) sobre la matriz de reproducción modificada ($\mathbf{M}_{ren,lim}$), se cuantifican con mayor resolución.

16. El aparato (100; 200) de conformidad con una de las reivindicaciones 1 a 15, en donde el aparato está

- configurado para evaluar un elemento del flujo de bits (bsDcuMode) que describe un modo de limitación de distorsión, y en donde el limitador de distorsión está configurado para obtener selectivamente la matriz de reproducción objetivo de modo que la matriz de reproducción objetivo es una matriz de reproducción objetivo similar a la mezcla descendente, o de modo que la matriz de reproducción objetivo es una matriz de reproducción objetivo de mejor esfuerzo.
- 5
17. Un aparato (150) para proporcionar un flujo de bits (170) que representa una señal de audio multicanal, el aparato comprende:
- 10 un mezclador descendente (180) configurado para proporcionar una señal de mezcla descendente (182) con base en una pluralidad de señales del objeto de audio (160a-160N);
- 15 un proveedor de información lateral (184) configurado para proporcionar una información lateral paramétrica relacionada con el objeto (186) que describe características de las señales del objeto de audio (160a-160N) y parámetros de mezcla descendente, y un parámetro de combinación lineal (188) que describe contribuciones deseadas de una matriz de reproducción especificada por el usuario (M_{ren}) y de una matriz de reproducción objetivo ($M_{ren,tar}$) a una matriz de reproducción modificada ($M_{ren,lim}$) a ser utilizada por un aparato (100; 200) para proporcionar una representación de señal de mezcla ascendente con base en el flujo de bits; y
- 20 un formateador de flujo de bits (190) configurado para proporcionar un flujo de bits (170) que comprende una representación de la señal de mezcla descendente, de la información lateral paramétrica relacionada con el objeto y del parámetro de combinación línea.
- 25 18. Un método para proporcionar una representación de señal de mezcla ascendente con base en una representación de señal de mezcla descendente y una información paramétrica relacionada con el objeto, que se incluyen en una representación del flujo de bits de un contenido de audio, y en dependencia de una matriz de reproducción especificada por el usuario, el método comprende:
- 30 evaluar un elemento del flujo de bits que representa un parámetro de combinación lineal, para obtener el parámetro de combinación lineal;
- 35 obtener una matriz de reproducción modificada utilizando una combinación lineal de una matriz de reproducción especificada por el usuario y una matriz de reproducción objetivo en dependencia del parámetro de combinación lineal; y
- 40 obtener la representación de señal de mezcla ascendente con base en la representación de señal de mezcla descendente y la información paramétrica relacionada con el objeto utilizando la matriz de reproducción modificada.
- 45 19. Un método para proporcionar un flujo de bits que representa una señal de audio multicanal, el método comprende:
- proporcionar una señal de mezcla descendente con base en una pluralidad de señales del objeto de audio;
- proporcionar una información lateral paramétrica relacionada con el objeto que describe características de las señales del objeto de audio y parámetros de mezcla descendente, y un parámetro de combinación lineal que describe contribuciones deseadas de una matriz de reproducción especificada por el usuario y de una matriz de reproducción objetivo a una matriz de reproducción modificada; y
- 50 proporcionar un flujo de bits que comprende una representación de la señal de mezcla descendente, de la información lateral paramétrica relacionada con el objeto y el parámetro de combinación lineal.
- 55 20. Un programa informático para realizar un método de acuerdo con la reivindicación 18 o 19 cuando el programa informático se ejecuta en un ordenador.
21. Un flujo de bits (300) que representa una señal de audio multicanal, el flujo de bits comprende:
- una representación (302) de una señal de mezcla descendente combinando señales de audio de una

pluralidad de objetos de audio;

una información paramétrica relacionada con el objeto (304) que describe características de los objetos de audio; y

5

un parámetro de combinación lineal (306) que describe contribuciones deseadas de una matriz de reproducción especificada por el usuario y de una matriz de reproducción objetivo a una matriz de reproducción modificada.

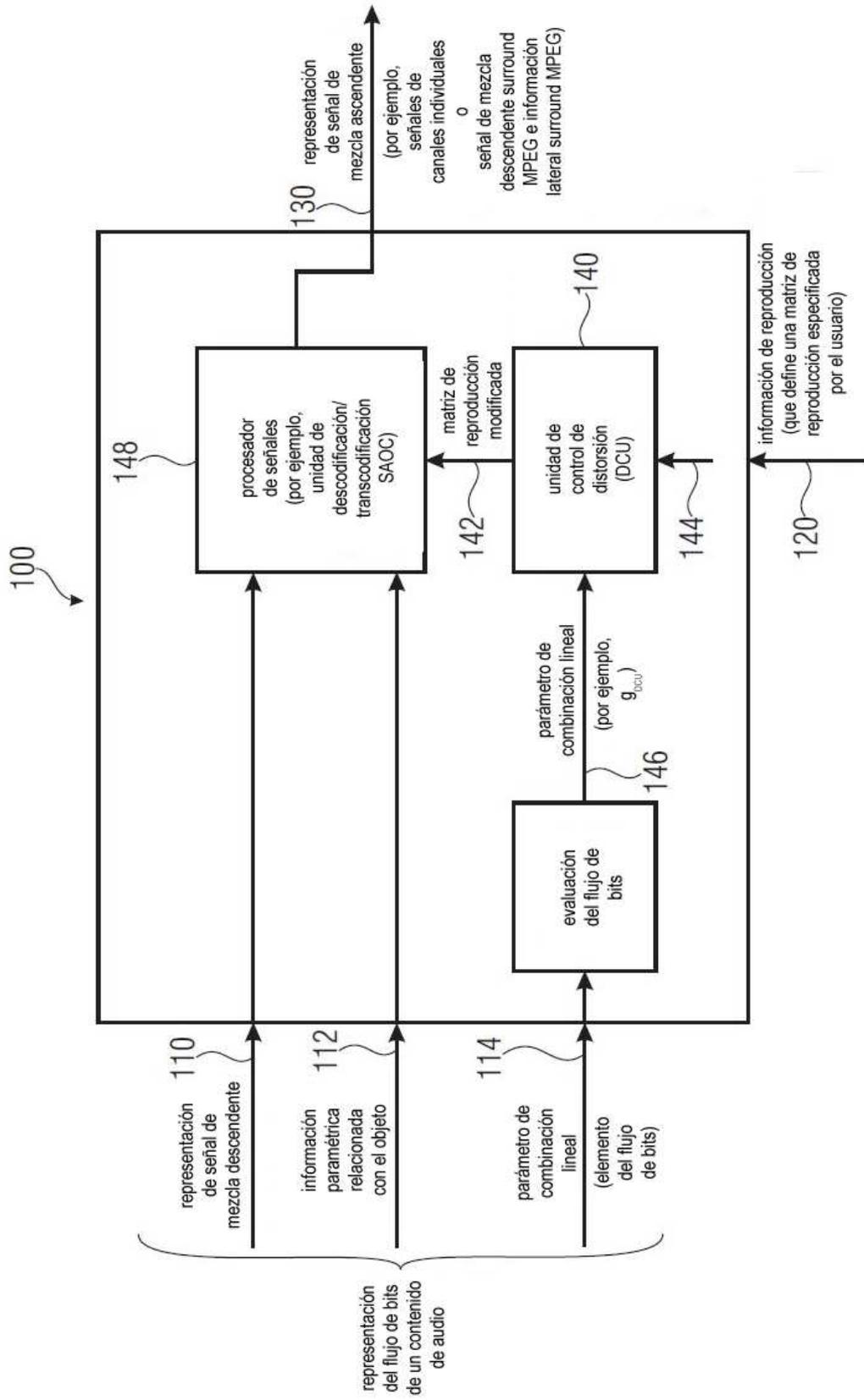


FIGURA 1A

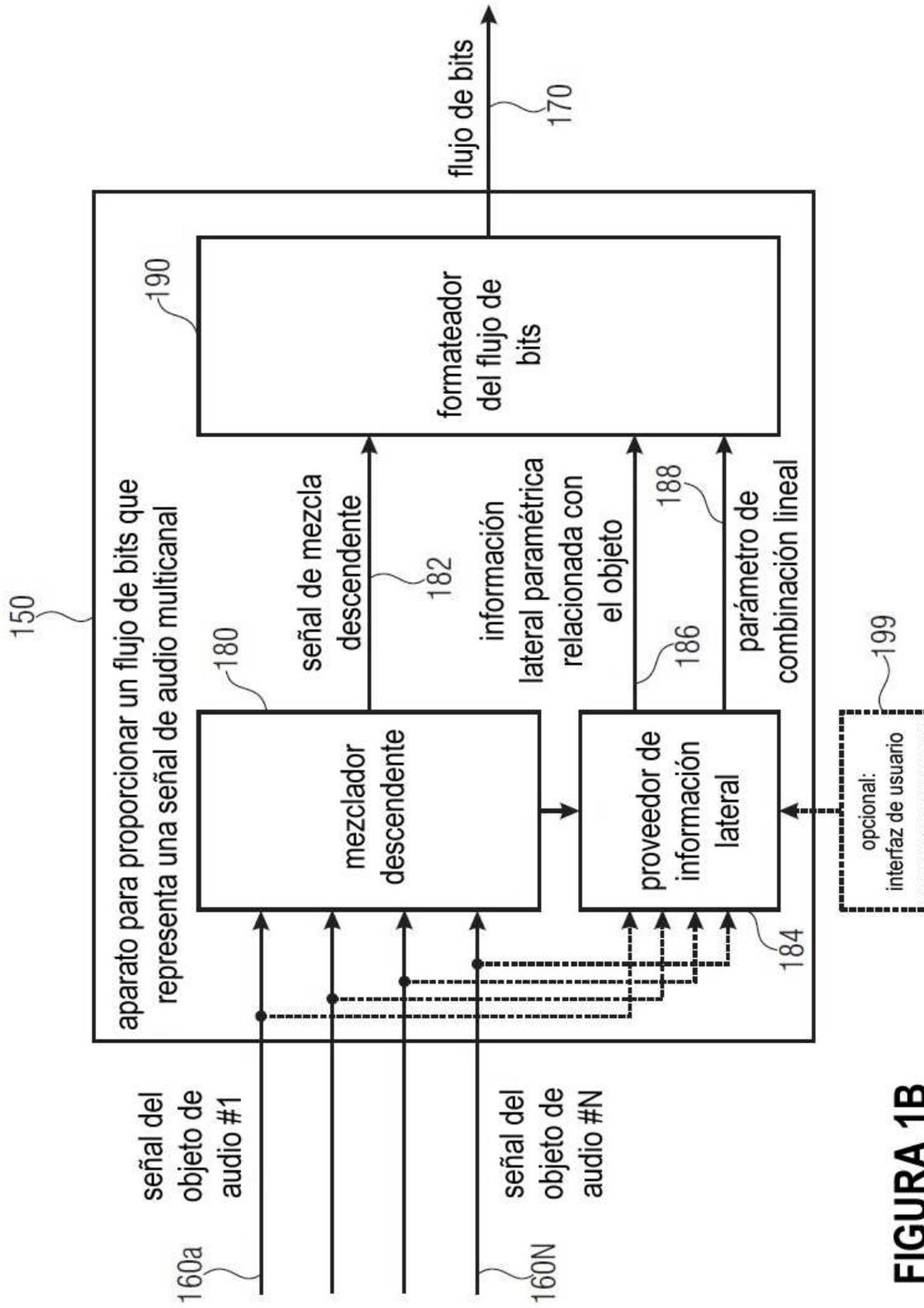


FIGURA 1B

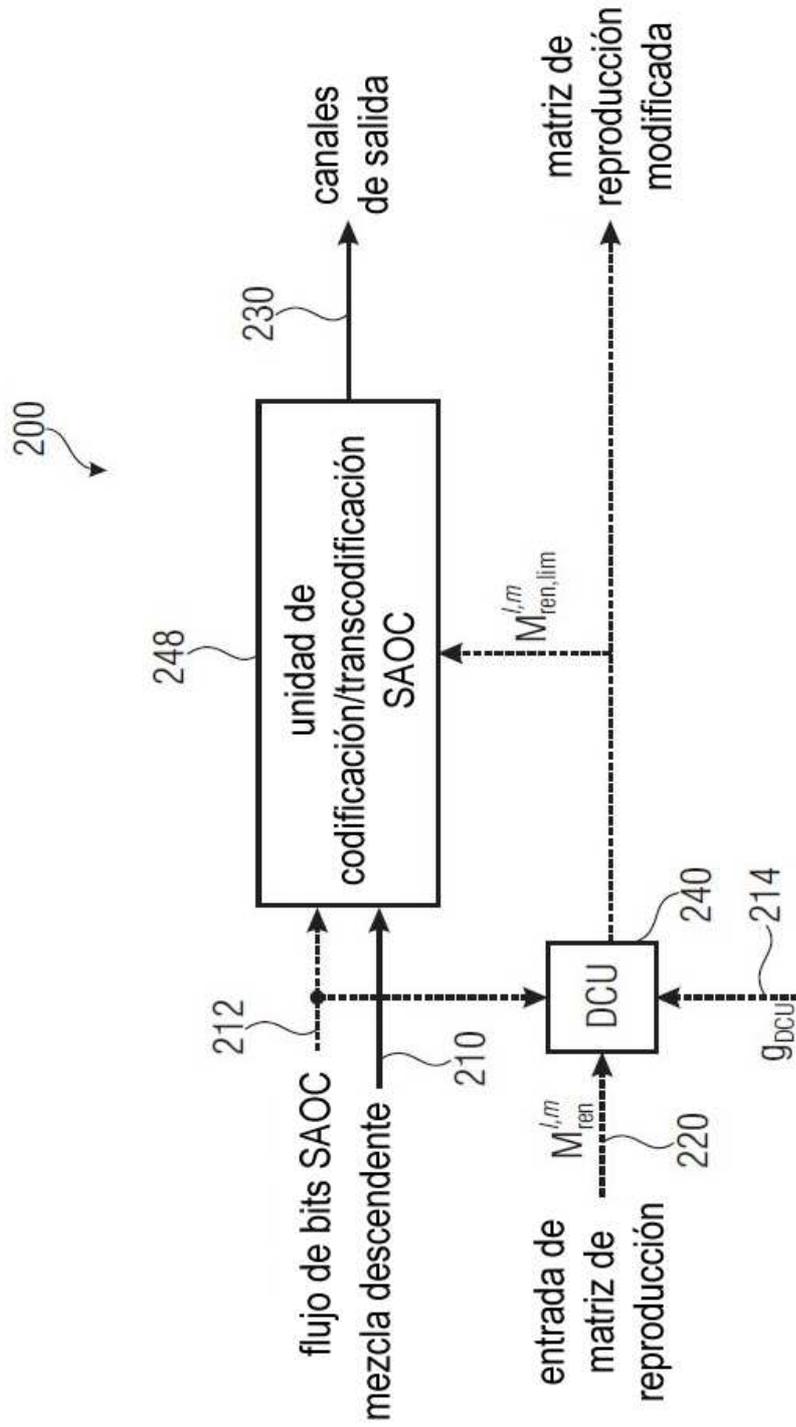


FIGURA 2

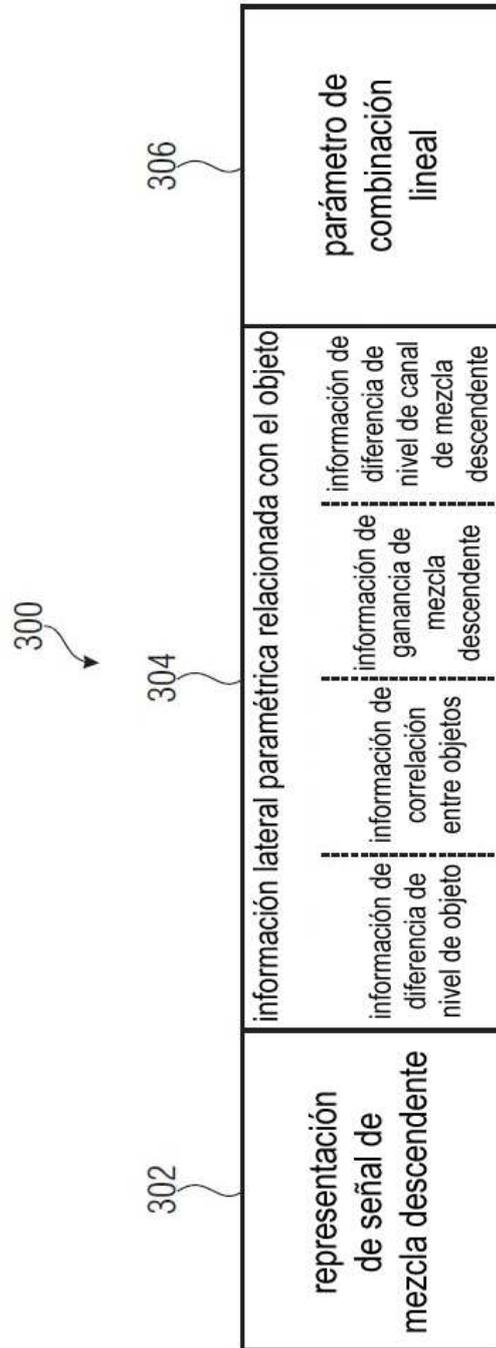


FIGURA 3A

310
 ↘
 Sintaxis de SAOCSpecificConfig()

Sintaxis	No. de bits	Mnemotécnico
SAOCSpecificConfig() {		
Configuración de frecuencia de muestreo;		
Configuración de modo de bajo retardo;		
Configuración de resolución de frecuencia;		
Configuración de longitud de cuadro;		
Configuración de número de objeto;		
Configuración de relación de objeto;		
Configuración de transmisión de energía absoluta;		
Configuración de número de canal de mezcla descendente;		
Información de configuración adicional;		
bsDdgFlag;	1	
bsDcuFlag;	1	uimsbf
if (bsDcuFlag == 1) {		
bsDcuMandatory;	1	uimsbf
bsDcuDynamic;	1	uimsbf
if (bsDcuDynamic == 0) {		
bsDcuMode;	1	uimsbf
bsDcuParam	4	uimsbf
}		
} else {		
bsDcuMandatory = 0;		
bsDcuDynamic = 0;		
bsDcuMode = 0;		
bsDcuParam = 0;		
}		
ByteAlign();		
SAOCExtensionConfig();		
}		

FIGURA 3B

bsDcuMode

bsDcuMode	Significado
0	matriz de destino similar a mezcla descendente
1	matriz de destino de esfuerzo máximo

FIGURA 3D

Tabla 39 - tabla de cuantificación de parámetros bsDcuParam

idx	0	1	2	3	4	5	6	7
DcuParam[idx]	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35
idx	8	9	10	11	12	13	14	15
DcuParam[idx]	0.40	0.45	0.50	0.60	0.70	0.80	0.90	1.00

FIGURA 3E

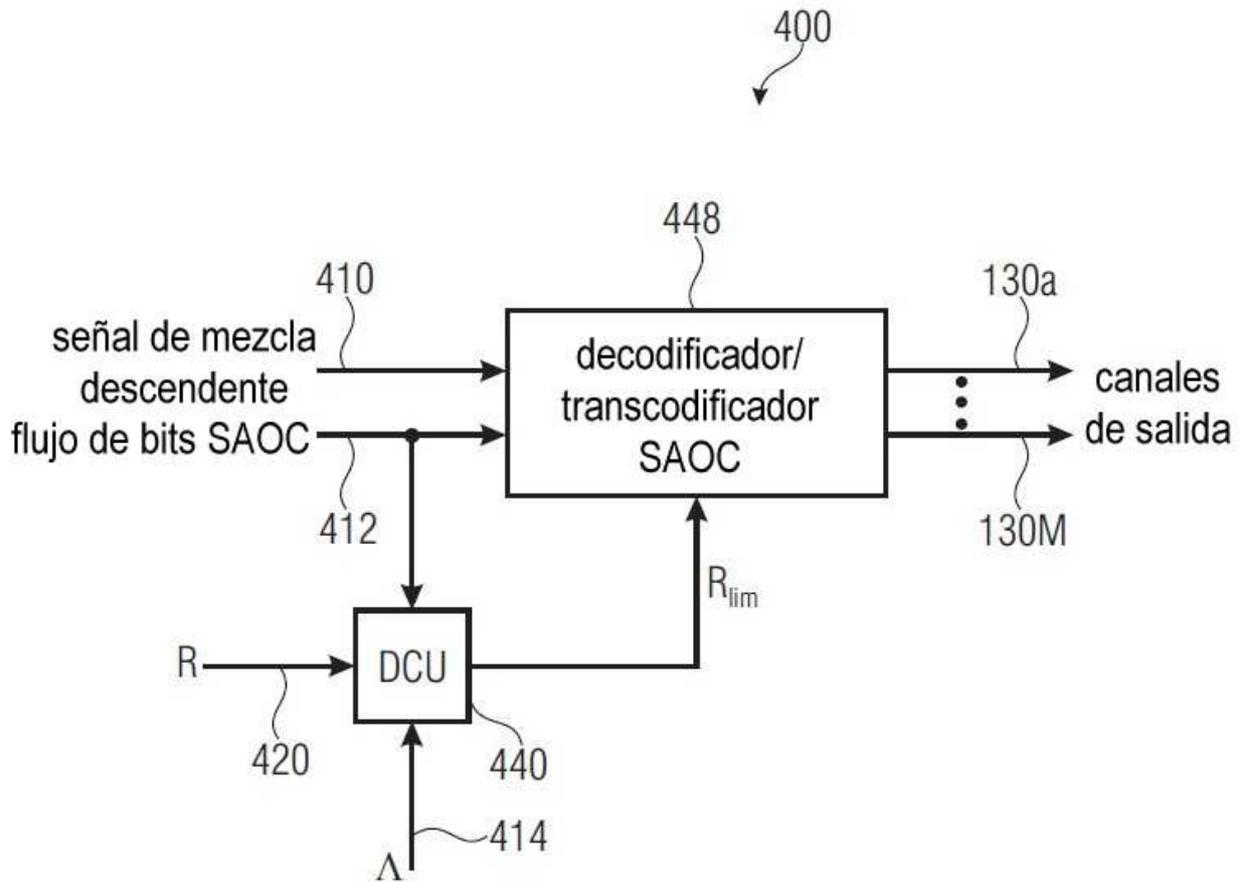


FIGURA 4

Tabla 4 - Sintaxis de SAOCSpecificConfig()

Sintaxis	No. de bits	Nemotécnico
SAOCSpecificConfig() { Información de Configuración SAOC; bsPdgFlag; bsDcuFlag; if (bsDcuFlag == 1) { bsDcuMode; bsDcuParam; } else { bsDcuMode = 0 ; bsDcuParam = 6; } } ByteAlign(); SAOCExtensionConfig(); }		
	1	uimsbf
	1	uimsbf
	1	uimsbf
	4	uimsbf
	} 510	

FIGURA 5A

Tabla de cuantificación de parámetros bsDcuParam

idx	0	1	2	3	4	5	6	7
DcuParam[idx]	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35
idx	8	9	10	11	12	13	14	15
DcuParam[idx]	0.40	0.45	0.50	0.60	0.70	0.80	0.90	1.00

FIGURA 5B

Condiciones de prueba de audición

Nombre codificador	Descripción
"_DMX"	señal de reproducción similar a mezcla descendente trivial del descodificador SAOC regular
"_noLim"	salida del descodificador SAOC regular (sin procesar por la DCU)
"_Rlim"	salida del descodificador SAOC con DCU usando la "reproducción similar a mezcla descendente"
"_RlimBE"	salida del descodificador SAOC con DCU usando la "reproducción del máximo esfuerzo"

FIGURA 6A

Elementos de audio de la prueba de audición

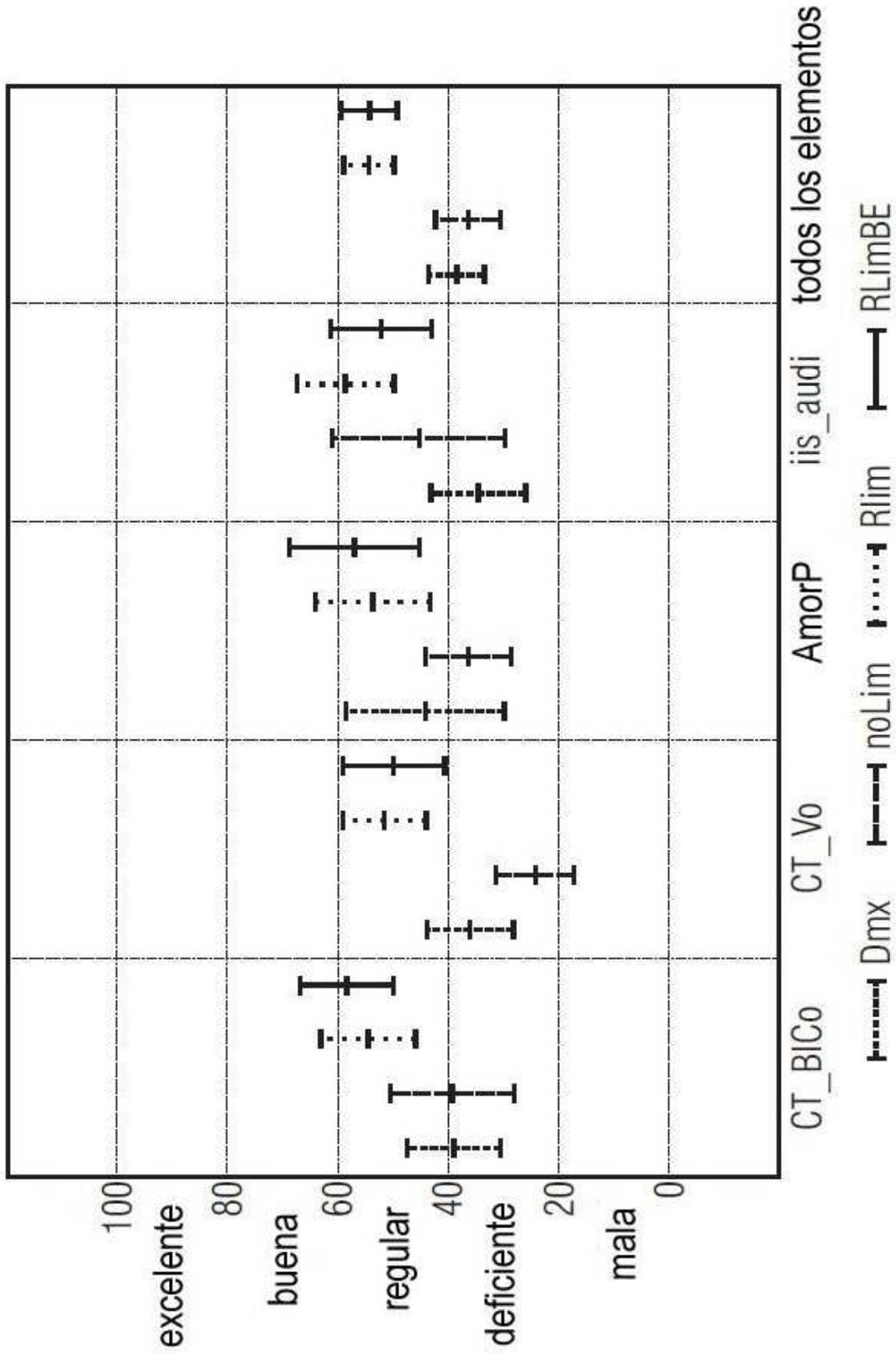
Elementos de audición	Descripción	Duración
"CaféNegro"	Sección de metales suaves dentro de la mezcla de sonido	7 segundos
"Fanta4"	Tambor fuerte dentro de la mezcla de sonido	7 segundos
"AmorPop"	Sonido vocal fuerte y música suave	7 segundos
"Audición"	Sección de cuerdas suaves dentro de la mezcla de sonido	13 segundos

FIGURA 6B

Condiciones de prueba de mezcla descendente/reproducción para un escenario de descodificación de SAOC estéreo a estéreo

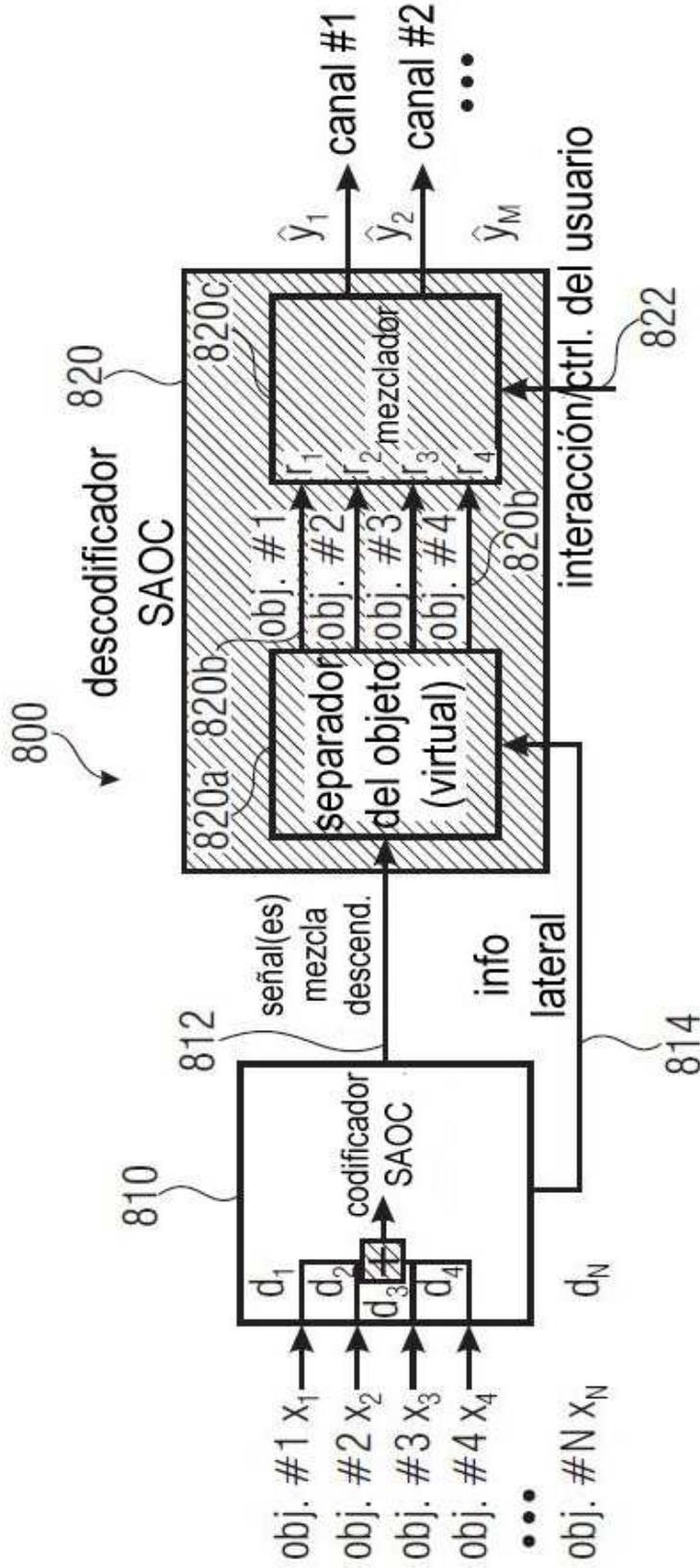
Elementos de audición	Obj	Descripción de objetos de audio	Mezcla descendente (L/R)		Reproducción (L/R)	
"CaféNegro"	1	Metales 1	1.0	0.6	0.0001	0.0001
	2	Metales 2	0.5	1.0	0.0001	0.0001
	3	Órgano	0.5778	0.8162	1.2628	0.6329
	4	Tambores y bajo	0.8162	0.5778	0.9988	0.9988
	5	Percusión	0.4481	0.8940	0.9988	0.9988
"VozSobre Música"	1	Música de fondo (L)	1.0	0.0	0.01	0.0
	2	Música de fondo (R)	0.0	1.0	0.0	0.01
	3	Vocal	0.7071	0.7071	0.7071	0.7071
"AmorPop"	1	Tambores	0.7071	0.7071	0.4462	0.4462
	2	Bajo	0.7071	0.7071	0.4462	0.4462
	3	Guitarra eléctrica	0.7071	0.7071	0.5150	0.3646
	4	Guitarra acústica	0.7071	0.7071	0.1105	0.6213
	5	Cuerdas	0.7071	0.7071	0.0	0.0
"Audición"	1	Voces de fondo	2.7335	0.6866	0.4860	0.1221
	2	Bajo	1.9929	1.9929	0.3543	0.3543
	3	Re toque tambores	1.9929	1.9929	0.3543	0.3543
	4	Bombo	1.9929	1.9929	0.3543	0.3543
	5	Guitarra solista	1.0425	2.6185	0.1853	0.4656
	6	2 voces de fondo	0.8498	2.6872	2.6873	8.4978
	7	Voces principales	1.9929	1.9929	6.3022	6.3022
	8	Tambores aéreos	2.3003	1.6285	0.4090	0.2895
	9	Guitarras rítmicas	2.5197	1.2628	0.4480	0.2245
	10	Eco aplausos	2.7749	0.4934	8.7749	1.5604
	11	Tambor de tirantes	1.6285	2.3003	0.2895	0.4090
	12	Tam-tam	0.6866	2.7335	0.1221	0.4860

FIGURA 6C



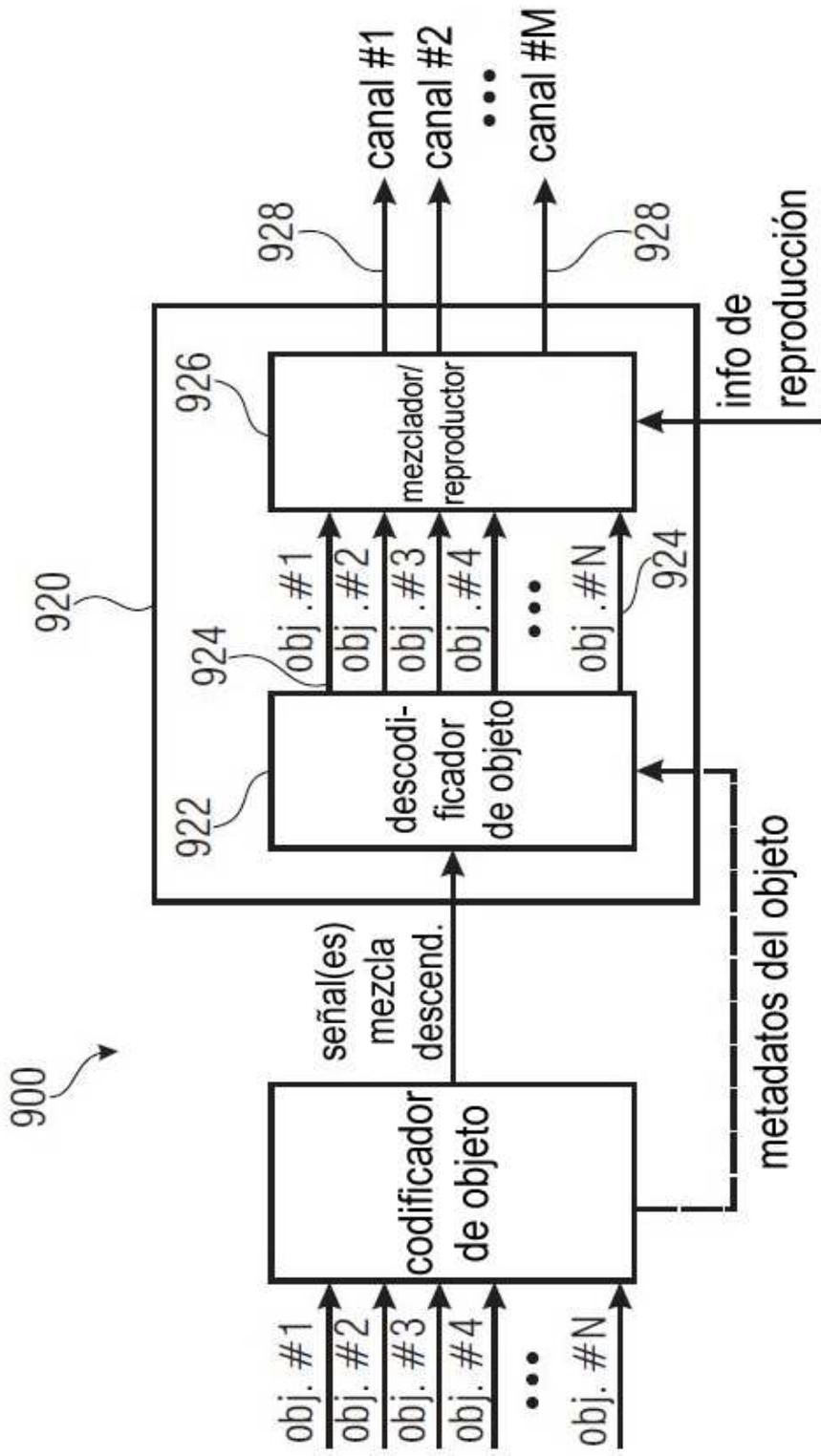
Resultados de pruebas de audición de la DCU para escenarios de descodificación SAOC estéreo a estéreo

FIGURA 7



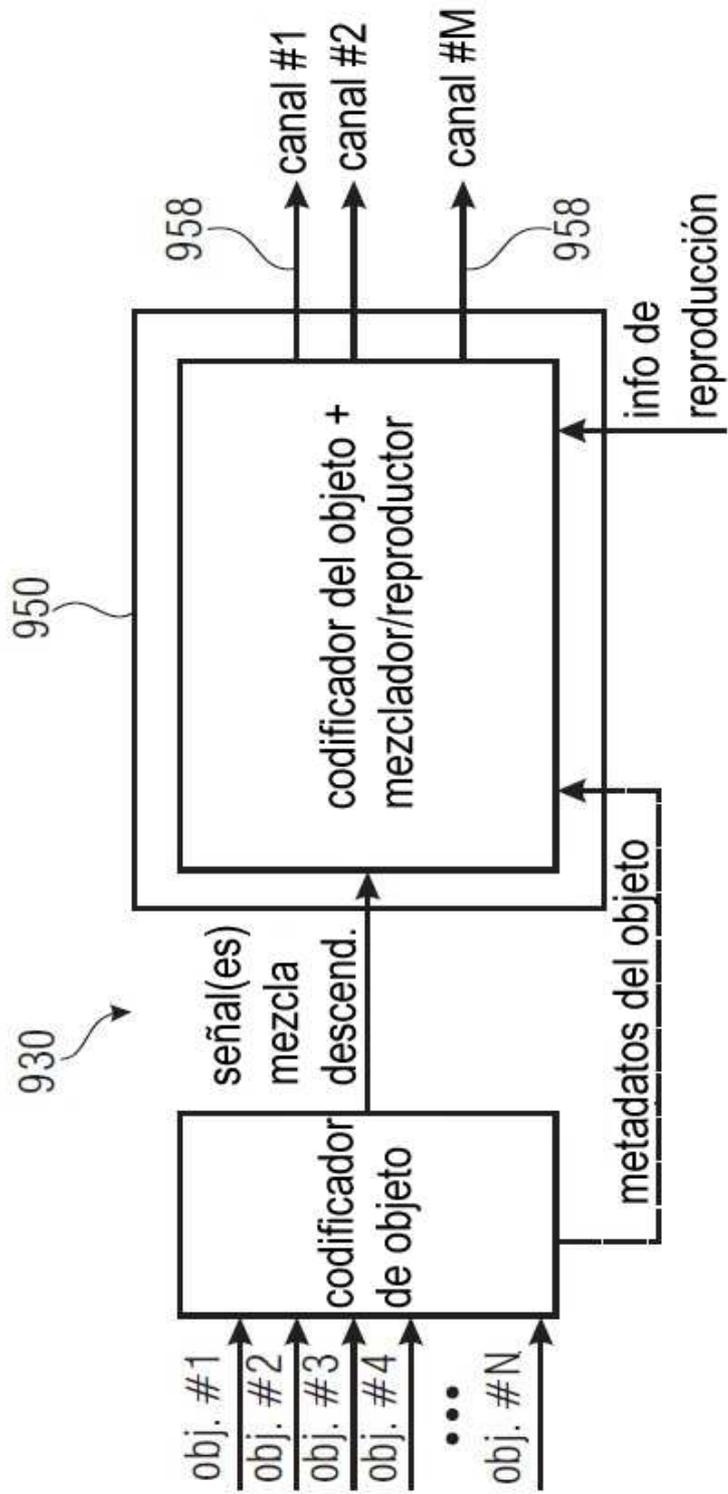
Vista general de sistema MPEG SAOC

FIGURA 8



Descodificador y mezclador separados

FIGURA 9A



Descodificador y mezclador integrados
FIGURA 9B

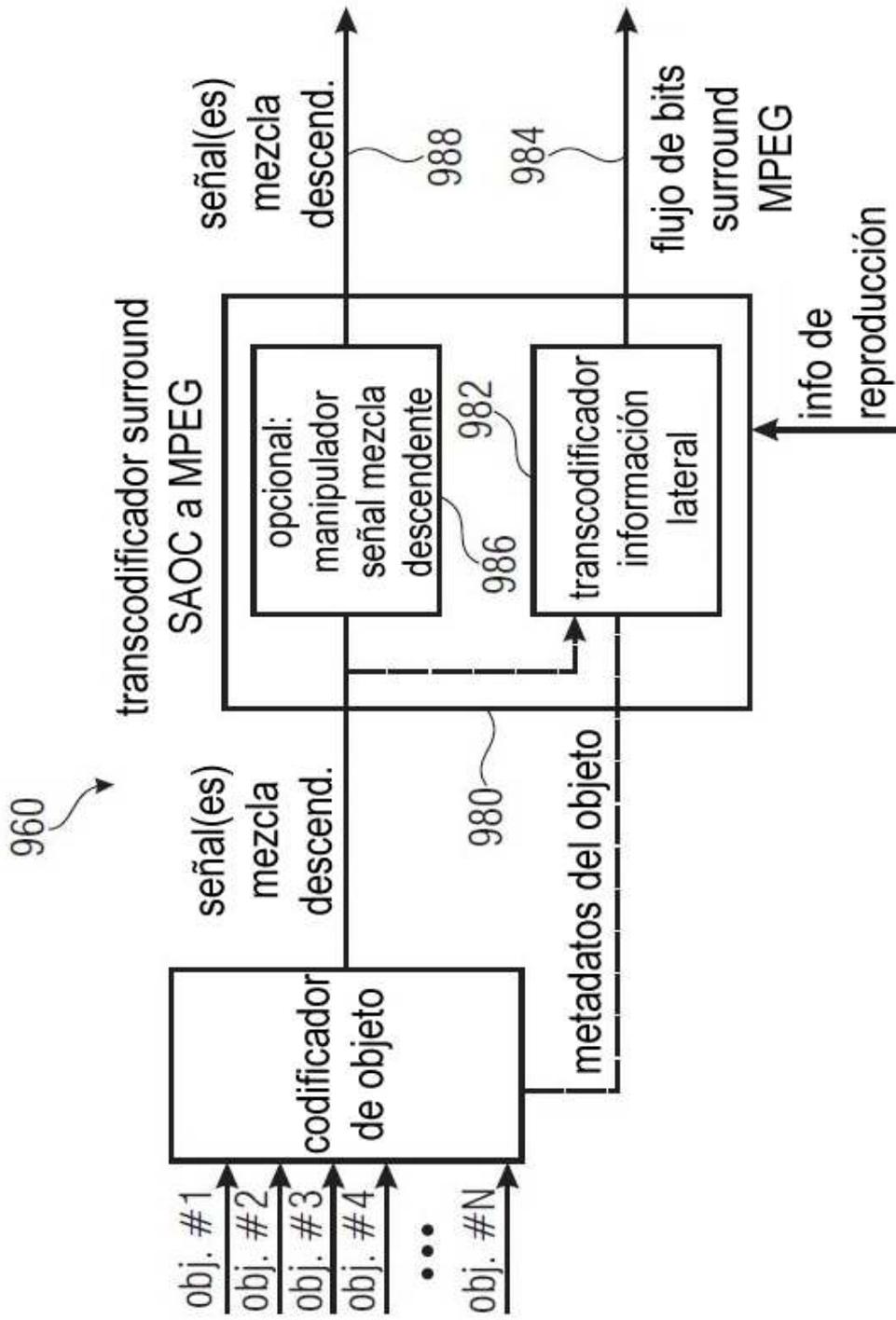


FIGURA 9C