

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 572 083**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/20 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.04.2010 E 14180279 (3)**

97 Fecha y número de publicación de la concesión europea: **23.03.2016 EP 2816555**

54 Título: **Codificador de señal de audio, flujo de bits de audio, método y programa informático que utiliza información paramétrica relacionada con el objeto**

30 Prioridad:

28.04.2009 US 173456 P

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

30.05.2016

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (33.0%)**

**Hansastraße 27c
80686 München, DE;**

**DOLBY INTERNATIONAL AB (33.0%) y
FRIEDRICH-ALEXANDER-UNIVERSITÄT
ERLANGEN-NÜRNBERG (33.0%)**

72 Inventor/es:

**HERRE, JÜRGEN;
HÖLZER, ANDREAS;
TERENTIV, LEON;
FALCH, CORNELIA;
PURNHAGEN, HEIKO;
ENGDEGARD, JONAS;
RIDDERBUSCH, FALKO y
KASTNER, THORSTEN**

74 Agente/Representante:

ARIZTI ACHA, Mónica

ES 2 572 083 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).

Codificador de señal de audio, flujo de bits de audio, método y programa informático que utiliza información paramétrica relacionada con el objeto

DESCRIPCIÓN

- 5 **Campo técnico**
- Las realizaciones de acuerdo con la invención se relacionan con un codificador de señal de audio, un método correspondiente y un flujo de bits de audio.
- 10 Algunas realizaciones más se relacionan con programas informáticos correspondientes.
- Solicitud divisional del documento EP10716830.4
- 15 **Antecedentes de la Invención**
- En la técnica del procesamiento de audio, transmisión de audio y almacenamiento de audio, existe un deseo creciente de manejar el contenido de múltiples canales a fin de mejorar la impresión auditiva. El uso del contenido de audio multicanal acarrea considerables mejoras para el usuario. Por ejemplo, se puede obtener una impresión
- 20 auditiva tridimensional, lo que acarrea una mayor satisfacción del usuario en aplicaciones de entretenimiento. Sin embargo, los contenidos de audio multicanal también son útiles en entornos profesionales, por ejemplo en aplicaciones de conferencias telefónicas, ya que se puede mejorar la inteligibilidad del hablante utilizando una reproducción de audio en múltiples canales.
- 25 Sin embargo, también es conveniente contar con una buena relación entre calidad de audio y requisitos de flujo de bits a fin de evitar una carga de recursos excesiva causada por aplicaciones multicanal.
- Últimamente, se han propuesto técnicas paramétricas para la transmisión con eficiencia de flujo de bits y/o almacenamiento de escenas de audio que contienen múltiples objetos de audio, por ejemplo, la Binaural Cue Coding (Codificación Binaural de Monitorización de Señales) (Tipo I) (véase, por ejemplo la referencia [BCC]), Joint Source Coding (Codificación Conjunta de Canales Fuente) (véase, por ejemplo, la referencia [JSC]), y Codificación de
- 30 Objetos de Audio Espacial MPEG (MPEG Spatial Audio Object Coding) (SAOC) (véase, por ejemplo, las referencias [SAOC1], [SAOC2]).
- 35 Estas técnicas tienen por objeto reconstruir perceptualmente la escena de audio de salida en lugar de hacerlo por la paridad de una forma de onda.
- La Fig. 8 ilustra una reseña de sistema de ese tipo de sistema (en este caso: MPEG SAOC). El sistema MPEG SAOC 800 expuesto en la Fig. 8 comprende un codificador SAOC 810 y un decodificador SAOC 820. El codificador SAOC 810 recibe una pluralidad de señales de objeto x_1 a x_N , que pueden estar representadas, por ejemplo, en forma de señales en el dominio del tiempo o en forma de señales en el dominio de tiempo-frecuencia (por ejemplo, en forma de una serie de coeficientes de transformada del tipo transformada de Fourier, o en forma de señales de subbandas QMF). El codificador SAOC 810 también recibe, por lo general, coeficientes de mezcla descendente d_1 a d_N , que están asociados a las señales de objeto x_1 a x_N . Pueden existir series separadas de coeficientes de mezcla descendente para cada canal de la señal de mezcla descendente. El codificador SAOC 810 está configurado típicamente para obtener un canal de la señal de mezcla descendente mediante la combinación de las señales de objeto x_1 a x_N de acuerdo con los coeficientes de mezcla descendente asociados d_1 a d_N . Por lo general hay menos canales de mezcla descendente que señales de objeto x_1 a x_N . Para dar lugar (por lo menos aproximadamente) a la separación (o tratamiento separado) de las señales de objeto en el lado del decodificador SAOC 820, el codificador SAOC 810 suministra tanto una o más señales de mezcla descendente (designadas como canales de mezcla descendente) 812 como una información complementaria 814. La información complementaria 814 describe características de las señales de objeto x_1 a x_N , para dar lugar al procesamiento específico de objetos en el lado del decodificador.
- 50 Un enfoque para especificar información complementaria puede encontrarse por ejemplo en el documento US 2008/0140426 A1.
- 60 El decodificador SAOC 820 está configurado para recibir tanto dichas una o más señales de mezcla descendente 812 como la información complementaria 814. Además, el decodificador SAOC 820 está configurado típicamente para recibir una información de interacción con el usuario y/o una información de control del usuario 822, que describe una configuración pretendida de presentación. Por ejemplo, la información de interacción con el usuario/control del usuario 822 puede describir una configuración de los hablantes y la colocación espacial conveniente de los objetos que producen las señales de objeto x_1 a x_N .

El decodificador SAOC 820 está configurado para producir, por ejemplo, una pluralidad de señales de canales decodificadas de mezcla descendente \hat{y}_1 a \hat{y}_M . Las señales de canales de mezcla ascendente pueden estar asociadas, por ejemplo a hablantes individuales de una disposición de presentación de múltiples hablantes. El decodificador SAOC 820 puede comprender, por ejemplo, un separador de objetos 820a, que está configurado para
 5 reconstruir, por lo menos aproximadamente, las señales de objeto x_1 a x_N sobre la base de dichas una o más señales de mezcla descendente 812 y la información complementaria 814, a fin de obtener así las señales de objetos reconstruidas 820b. Sin embargo, las señales de objetos reconstruidas 820b se pueden desviar hasta cierto punto de las señales de objetos originales x_1 a x_N , por ejemplo, porque la información complementaria 814 no es bastante suficiente para una reconstrucción perfecta debido a restricciones de velocidad de transmisión de bits. El
 10 decodificador SAOC 820 puede comprender además un mezclador 820c, que puede estar configurado para recibir las señales de objetos reconstruidas 820b y la información de interacción con el usuario/información de control de usuarios 822, y para suministrar, sobre la base de éstas, las señales de canales de mezcla ascendente \hat{y}_1 a \hat{y}_M . El mezclador 820 puede estar configurado para usar la información de interacción con el usuario/información de control de usuarios 822 para determinar la contribución de las señales de objetos reconstruidas 820b individuales a las
 15 señales de canales de mezcla ascendente \hat{y}_1 a \hat{y}_M . La información de interacción con el usuario/información de control de usuarios 822 puede comprender, por ejemplo, parámetros de presentación (que también se denominan coeficientes de presentación), que determinan la contribución de las señales de objetos reconstruidas 822 individuales a las señales de canales de mezcla ascendente \hat{y}_1 a \hat{y}_M .

20 Sin embargo, cabe señalar que, en muchas realizaciones, la separación de los objetos, que está indicada por el separador de objetos 820a de la Fig. 8, y la mezcla, que está indicada por el mezclador 820c de la Fig. 8, se llevan a cabo en un solo paso. Para este fin, se pueden calcular los parámetros generales que describen un mapeo directo de dichas una o más señales de mezcla descendente 812 sobre las señales de canales de mezcla ascendente \hat{y}_1 a \hat{y}_M . Estos parámetros se pueden calcular sobre la base de la información complementaria y la información de
 25 interacción con el usuario/ información de control de usuarios 820.

Tomando, ahora, como referencia las Figs. 9a, 9b y 9c, se describen diferentes aparatos para obtener una representación de la señal de mezcla ascendente sobre la base de una representación de señal de mezcla descendente e información complementaria relacionada con objetos. La Fig. 9a ilustra un diagrama esquemático de
 30 bloques de un sistema MPEG SAOC 900 que comprende un decodificador SAOC 920. El decodificador SAOC 920 comprende, como bloques funcionales separados, un decodificador de objetos 922 y un mezclador/presentador 926. El decodificador de objetos 922 produce una pluralidad de señales de objetos reconstruidas 924 que depende de la representación de la señal de mezcla descendente (por ejemplo, en forma de una o más señales de mezcla descendente representadas en el dominio del tiempo o en el dominio del tiempo-frecuencia) e información
 35 complementaria relacionada con objetos (por ejemplo, en forma de metadatos de objeto). El mezclador/presentador 924 recibe las señales de objetos reconstruidas 924 asociadas a una pluralidad de objetos N y produce, sobre la base de éstas, una o más señales de canales de mezcla ascendente 928. En el decodificador SAOC 920, la extracción de las señales de objeto 924 se realiza en forma independiente de la mezcla/presentación, lo que da lugar a una separación de la funcionalidad de decodificación de objetos de la funcionalidad de mezcla/presentación, aunque trae aparejada una complejidad computacional relativamente elevada.

Haciendo referencia, ahora, a la Fig. 9b, se describe brevemente otro sistema MPEG SAOC 930, que comprende un decodificador SAOC 950. El decodificador SAOC 950 produce una pluralidad de señales de canales de mezcla ascendente 958 que depende de una representación de señal de mezcla descendente (por ejemplo, en forma de
 45 una o más señales de mezcla descendente) e información complementaria relacionada con un objeto (por ejemplo, en forma de metadatos de objeto). El decodificador SAOC 950 comprende una combinación de decodificador de objeto y mezclador/presentador, que está configurado para obtener las señales de canales de mezcla ascendente 958 en un proceso de mezclado conjunto sin separación de la decodificación de objetos y la mezcla/presentación, donde los parámetros para dicho proceso conjunto de mezcla ascendente dependen tanto de la información
 50 complementaria relacionada con objetos y la información sobre la presentación. El proceso conjunto de mezcla ascendente depende asimismo de la información sobre mezcla descendente, que se considera parte de la información complementaria relacionada con objetos.

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

Haciendo referencia, ahora, a la Fig. 9c, se describe un sistema MPEG SAOC 960. El sistema SAOC 960 comprende un transcodificador de SAOC a MPEG Envoltente (Surround) 980, en lugar de un decodificador SAOC.

60 El transcodificador de SAOC a MPEG Envoltente comprende un transcodificador de información complementaria 982, que está configurado para recibir la información complementaria relacionada con objetos (por ejemplo, en forma de metadatos de objeto) u, opcionalmente, información de dicha una o más señales de mezcla descendente y la información sobre la presentación. El transcodificador de información complementaria también está configurado para proporcionar una información complementaria sobre MPEG Envoltente (por ejemplo, en forma de flujo de bits de

MPEG Envoltente) sobre la base de ciertos datos recibidos. En consecuencia, el transcodificador de información complementaria 982 está configurado para transformar una información complementaria relacionada con objetos (paramétrica), que se emite por un codificador de objetos, en una información complementaria (paramétrica) relacionada con los canales, teniendo en cuenta la información sobre la presentación y, opcionalmente, la información sobre el contenido de dichas una o más señales de mezcla descendente.

Opcionalmente, el transcodificador de SAOC a MPEG Envoltente 980 puede estar configurado para manipular dichas una o más señales de mezcla descendente, descritas, por ejemplo, por la representación de la señal de mezcla descendente, para obtener una representación de señal de mezcla descendente manipulada 988. Sin embargo, se puede omitir el manipulador de señal de mezcla descendente 986, por lo que la representación de señal de mezcla descendente 988 de salida del transcodificador de SAOC a MPEG Envoltente 980 es idéntica a la representación de señal de mezcla descendente de entrada del transcodificador de SAOC a MPEG Envoltente. El manipulador de señales de mezcla descendente 986 se puede utilizar, por ejemplo, en caso de que la información complementaria de MPEG Envoltente relacionada con los canales 984 no permita la producción de una impresión auditiva adecuada sobre la base de la representación de señal de mezcla descendente de entrada del transcodificador de SAOC a MPEG Envoltente 980, lo que puede ocurrir en algunas constelaciones de presentación.

En consecuencia, el transcodificador de SAOC a MPEG Envoltente 980 da origen a la representación de la señal de mezcla descendente 988 y el flujo de bits de MPEG Envoltente 984 razón por la cual se puede generar una pluralidad de señales de canales de mezcla ascendente, que representan los objetos de audio de acuerdo con la información sobre la presentación introducida en el transcodificador de SAOC a MPEG Envoltente 980 utilizando un decodificador de MPEG Envoltente que recibe el flujo de bits de MPEG Envoltente 984 y la representación de la señal de mezcla descendente 988.

Para resumir lo anterior, se pueden emplear diferentes conceptos para decodificar señales de audio codificadas por SAOC. En algunos casos, se utiliza un decodificador SAOC, que produce señales de canales de mezcla ascendente (por ejemplo, las señales de canales de mezcla ascendente 928, 958) que dependen de la representación de la señal de mezcla descendente y la información complementaria paramétrica relacionada con el objeto. Se pueden ver ejemplos de este concepto en las Figs. 9a y 9b. Por otro lado, se puede transcodificar la información de audio codificada por SAOC para obtener una representación de señal de mezcla descendente (por ejemplo, una representación de señal de mezcla descendente 988) y una información complementaria relacionada con los canales (por ejemplo, el flujo de bits de MPEG Envoltente relacionado con los canales 984), que puede utilizarse por un decodificador de MPEG Envoltente para producir las pretendidas señales de canales de mezcla ascendente.

En el sistema MPEG SAOC 800, una reseña general del sistema el cual se presenta en la Fig. 8, se lleva a cabo el procesamiento general en forma selectiva en frecuencia y se puede describir de la siguiente manera dentro de cada banda de frecuencias:

- Se realiza la mezcla descendente de N señales de audio de entrada de objeto x_1 a x_N como parte del procesamiento del codificador SAOC. Para una mezcla descendente mono, los coeficientes están indicados por d_1 a d_N . Además, el codificador SAOC 810 extrae información complementaria 814 que describe las características de los objetos de audio de entrada. En el caso de la MPEG SAOC, las relaciones de las potencias de objeto entre sí son la forma más básica de dicha información complementaria.
- La señal (o señales) de mezcla descendente 812 y la información complementaria 814 se transmiten y/o almacenan. Para este fin, se puede comprimir la señal de audio de mezcla descendente empleando codificadores de audio perceptuales muy conocidos tales como MPEG-1 Capa II o III (también conocido como ".mp3"), Codificación de Audio Avanzada MPEG (AAC), o cualquier otro codificador de audio.
- En el lado del receptor, el decodificador SAOC 820 intenta conceptualmente restablecer la señal de objeto original ("separación de objetos") utilizando la información complementaria transmitida 814 (y, naturalmente, la una o más señales de mezcla descendente 812). Estas señales de objeto aproximadas (que también se denominan señales de objetos reconstruidas 820b) se mezclan a continuación en una escena objetivo representada por M canales de salida de audio (que pueden estar representados, por ejemplo, por las señales de canales de mezcla ascendente \hat{y}_1 a \hat{y}_M) utilizando una matriz de presentación. Para una salida mono, los coeficientes de la matriz de presentación están expresados por r_1 a r_N .
- Efectivamente, raramente se ejecuta la separación de las señales de objeto (o incluso nunca se ejecutan), puesto que tanto el paso de separación (indicado por el separador de objetos 820a) como el paso de mezcla (indicado por el mezclador 820c) se combinan para obtener un solo paso de transcodificación, que con frecuencia da lugar a una enorme reducción de la complejidad computacional.

Se ha encontrado que ese tipo de esquema es tremendamente eficiente, tanto en términos de velocidad de

transmisión de bits (sólo es necesario transmitir unos pocos canales de mezcla descendente más cierta información complementaria en lugar de N señales discretas de audio de objeto o un sistema discreto) y complejidad computacional (la complejidad de procesamiento se relaciona principalmente con el número de canales de salida en lugar del número de objetos de audio). Otras ventajas para el usuario en el lado de la recepción incluyen la libertad de elegir una configuración de presentación de su elección (mono, estéreo, envolvente, reproducción virtualizada con auriculares y demás) y la característica de interactividad con el usuario: se puede ajustar la matriz de presentación y, por consiguiente, la escena de salida, puede ajustarse y cambiarse cambiada interactivamente por el usuario según su voluntad, preferencias personales u otros criterios. Por ejemplo, es posible localizar los interlocutores de un grupo juntos en un área espacial para maximizar la discriminación de las demás personas que conversan. Esta interactividad se obtiene produciendo una interfaz de usuario del decodificador:

Por cada objeto de sonido transmitido, se puede ajustar su nivel relativo y (en el caso de la presentación no mono) la posición espacial de la presentación. Esto puede ocurrir en tiempo real al cambiar el usuario la posición de los dispositivos deslizantes de interfaz gráfica asociada (GUI) (por ejemplo: nivel de objeto = +5 dB, posición de objeto = -30 grados).

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

En vista de esta situación, el objetivo de la presente invención consiste en crear un concepto que da lugar a la reducción o incluso a la eliminación de la distorsión audible al suministrar una representación de la señal de mezcla ascendente (por ejemplo, en forma de señales de canales de mezcla ascendente \hat{y}_1 a \hat{y}_M).

Sumario de la invención

Este problema se soluciona mediante un codificador de señal de audio de acuerdo con la reivindicación 1, un método de acuerdo con la reivindicación 4, un flujo de bits de audio de acuerdo con la reivindicación 5 y un programa informático de acuerdo con la reivindicación 6.

Una realización de acuerdo con la invención se refiere a un codificador de señales de audio para producir una representación de señal de mezcla descendente y una información paramétrica relacionada con el objeto sobre la base de una pluralidad de señales de objeto. El codificador de audio comprende un mezclador descendente configurado para producir una o más señales de mezcla descendente que dependen de coeficientes de mezcla descendente asociados con las señales de objeto, por lo que dichas una o más señales de mezcla descendente comprenden una superposición de una pluralidad de señales de objeto. El codificador de audio también comprende un productor de información complementaria configurado para producir una información complementaria de relación entre objetos que describen diferencias de nivel y las características de correlación de señales de objeto y una información complementaria de los objetos individuales que describe una o más propiedades individuales de las señales de objeto individuales, en el que la información complementaria de objetos individuales comprende una información de tonalidad de señal de objeto que describe tonalidades de las señales de objeto individuales. Se ha descubierto que la provisión tanto de información complementaria relacionada con objeto y una información complementaria de objetos individuales por medio de un codificador de señales de audio permite reducir con eficiencia, o incluso evitar, las distorsiones audibles en el lado del decodificador de señales de audio en múltiples canales. Si bien la información complementaria de relación entre objetos se utiliza para separar las señales de objetos en el lado del decodificador, la información complementaria de objetos individuales se puede utilizar para determinar si las características individuales de las señales de objetos se mantienen en el lado del decodificador lo que indica que las distorsiones están dentro de tolerancias aceptables.

Se ha hallado que la tonalidad de los objetos individuales es una cantidad importante desde el punto de vista psicoacústico, que permite la limitación de distorsiones en el lado del decodificador.

Otra realización de acuerdo con la invención se refiere a un método correspondiente.

Otra realización de acuerdo con la invención se refiere a un flujo de bits de audio que representa una pluralidad de señales de objeto (audio) en forma codificada. El flujo de bits de audio comprende una representación de señal de mezcla descendente que representa una o más señales de mezcla descendente, donde por lo menos una de las señales de mezcla descendente comprende una superposición de una pluralidad de señales de objeto (de audio). El flujo de bits de audio también comprende una información complementaria de relación entre objetos que describe diferencias de nivel y características de correlación de señales de objeto y una información complementaria de objetos individuales que describe una o más propiedades individuales de las señales de objeto individuales, en el que la información complementaria de objetos individuales comprende una información de tonalidad de señal de objetos que describe tonalidades de las señales de objetos individuales.

Como se ha analizado anteriormente, dicho flujo de bits de audio permite la reconstrucción de la señal de audio de múltiples canales, donde las distorsiones audibles que serían causadas por una configuración inapropiada de los parámetros de presentación, pueden reconocerse y reducirse o incluso eliminarse.

- 5 Otras realizaciones de acuerdo con la invención se refieren a un programa de informático para implementar el método anteriormente descrito.

Breve descripción de las figuras

- 10 A continuación se describen realizaciones de acuerdo con la invención tomando como referencia las figuras adjuntas, en las cuales:

- La Figura 1 ilustra un diagrama esquemático de bloques de un aparato para suministrar uno o más parámetros ajustados para la producción de una representación de señal de mezcla ascendente sobre la base de una representación de señal de mezcla descendente y una información paramétrica relacionada con el objeto;
- 15 La Figura 2 ilustra un diagrama esquemático de bloques de un sistema MPEG SAOC, de acuerdo con una realización de la invención;
- 20 La Figura 3 ilustra un diagrama esquemático de bloques de un sistema MPEG SAOC, de acuerdo con otra realización de la invención;
- La Figura 4 ilustra una representación esquemática de la contribución de señales de objeto a la señal de mezcla descendente y a una señal mezclada;
- 25 La Figura 5a ilustra un diagrama esquemático de bloques de un transcodificador de SAOC a MPEG Envolvente basado en mezcla descendente mono de acuerdo con una realización de la invención;
- 30 La Figura 5b ilustra un diagrama esquemático de bloques de un transcodificador SAOC a MPEG Envolvente basado en mezcla descendente estéreo, de acuerdo con una realización de la invención;
- La Figura 6 ilustra una representación esquemática de bloques de un codificador de señales de audio de acuerdo con una realización de la invención;
- 35 La Figura 7 ilustra una representación esquemática de un flujo de bits de audio de acuerdo con una realización de la invención;
- La Figura 8 ilustra un diagrama esquemático de bloques de un sistema MPEG SAOC de referencia;
- 40 La Figura 9a ilustra un diagrama esquemático de bloques de un sistema SAOC de referencia que utiliza un decodificador y mezclador separados;
- La Figura 9b ilustra un diagrama esquemático de bloques de un sistema SAOC de referencia que utiliza un decodificador y mezclador integrado; y
- 45 La Figura 9c ilustra un diagrama esquemático de bloques de un sistema SAOC de referencia que utiliza un transcodificador de SAOC a MPEG.

50 Descripción detallada de las realizaciones

1. Aparato para suministrar uno o más parámetros ajustados, de acuerdo con la Figura 1.

A continuación, se describe un aparato 100 para suministrar uno o más parámetros ajustados para la producción de una representación de señal de mezcla ascendente sobre la base de una representación de señal de mezcla descendente y una información paramétrica relacionada con los objetos tomando como referencia la Figura 1. La Figura 1 ilustra un diagrama esquemático de bloques de ese aparato 100, que está configurado para recibir uno o más parámetros de entrada 110. Los parámetros de entrada 110 pueden ser, por ejemplo, parámetros de presentación convenientes. El aparato 100 también está configurado para suministrar, sobre la base de estos, uno o más parámetros ajustados 120. Los parámetros ajustados pueden ser, por ejemplo, parámetros de presentación ajustados. El aparato 100 está configurado además para recibir una información paramétrica relacionada con los objetos 130. La información paramétrica relacionada con los objetos 130 puede ser, por ejemplo, una información de diferencia de niveles entre objetos y/o una información de correlación entre objetos que describe una pluralidad de objetos. El aparato 100 comprende un ajustador de parámetros 140, que está configurado para recibir uno o más

parámetros de entrada 110, y para suministrar sobre la base de estos, uno o más parámetros ajustados 120. El ajustador de parámetros 140 está configurado para producir dichos uno o más parámetros ajustados 120 dependiendo de dichos uno o más parámetros de entrada 110 y la información paramétrica relacionada con los objetos 130, por lo que se reduciría la distorsión de una señal de mezcla ascendente que sería causada por el uso de parámetros no óptimos (por ejemplo, dichos uno o más parámetros de entrada 110) en un aparato para suministrar una representación de señal de mezcla ascendente sobre la base una representación de señal de mezcla descendente y la información paramétrica relacionada con los objetos 130, al menos en el caso de los parámetros de entrada 110 que se desvían de los parámetros óptimos en más de una desviación predeterminada.

En consecuencia, el aparato 100 recibe dichos uno o más parámetros de entrada 110 y produce, sobre la base de éstos, dichos uno o más parámetros ajustados 120. Al suministrar dichos uno o más parámetros ajustados 120, el aparato 100 determina, de manera explícita o implícita, si el uso inalterado de uno o más parámetros de entrada 110 causaría distorsiones inadmisiblemente elevadas si se utilizara el uno o más parámetros de entrada 110 para controlar una provisión de una señal de mezcla ascendente sobre la base de una representación de una señal de mezcla descendente y la información paramétrica relacionada con los objetos 130. De esta manera, los parámetros ajustados 120 son por lo general más adecuados para ajustar dicho aparato para la provisión de la representación de señal de mezcla ascendente que dichos uno o más parámetros de entrada 110, al menos si dichos uno o más parámetros de entrada 110 se eligen de manera no ventajosa.

En consecuencia, el aparato 100 mejora por lo general la impresión perceptual de una representación de señal de mezcla ascendente, que se proporciona por un productor de representación de señal de mezcla ascendente dependiendo del uno o más parámetros ajustados 120. El uso de la información paramétrica relacionada con los objetos para ajustar los uno o más parámetros de entrada, para derivar los uno o más parámetros ajustados, han mostrado traer consigo buenos resultados, puesto que la calidad de la representación de señal de mezcla ascendente es típicamente satisfactoria si dichos uno o más parámetros ajustados 120 corresponden a la información paramétrica relacionada con los objetos 130, mientras que los parámetros que violan la relación conveniente con la información paramétrica relacionada con los objetos 130 por lo general dan lugar a distorsiones audibles. La información paramétrica relacionada con los objetos puede comprender, por ejemplo, parámetros de mezcla descendente, que describen una contribución de señales de objeto (de una pluralidad de objetos de audio) a dichas una o más señales de mezcla descendente. La información paramétrica relacionada con los objetos puede comprender también, por otro lado o además, parámetros de diferencias de niveles de los objetos y/o parámetros de correlación entre objetos, que describen las características de las señales de objetos. Se ha encontrado que tanto los parámetros que describen un procesamiento en el lado del codificador de las señales de objeto, como los parámetros que describen características de los objetos de audio por sí mismos, pueden considerarse información útil para usar por el ajustador de parámetros 120. Sin embargo, el aparato 100 puede utilizar de lo contrario o además información paramétrica relacionada con los objetos 130

Sin embargo, se debe tener en cuenta que el ajustador de parámetros 140 puede utilizar información adicional para suministrar dichos uno o más parámetros ajustados 120 sobre la base de dichos uno o más parámetros de entrada 110. Por ejemplo, el ajustador de parámetros 140 puede evaluar opcionalmente coeficientes de mezcla descendente, una o más señales de mezcla descendente o cualquier información adicional para mejorar aún la provisión de uno o más parámetros ajustados 120.

2. Sistema de acuerdo con la figura 2

A continuación se describirá en detalle el sistema MPEG SAOC 200 de la Figura 2.

Para ofrecer una buena comprensión del sistema MPEG SAOC 200, se presenta una reseña general de las especificaciones deseadas del sistema y las consideraciones de diseño. Seguidamente, se presenta una reseña estructural del sistema. Además, se describe una pluralidad de métricas de distorsión SAOC, y la aplicación de estas métricas de distorsión SAOC para la limitación de las distorsiones. Además, se describen otras extensiones del sistema 200.

2.1 Consideraciones de diseño del sistema

Como se ha analizado anteriormente, las técnicas paramétricas para la transmisión eficiente en cuanto a la velocidad de transmisión de bits escenas de almacenamiento de audio que contienen múltiples objetos de audio son por lo general eficientes, tanto en términos de velocidad de transmisión de bits como en complejidad computacional. Otras ventajas para el usuario de dicho sistema en el lado de la recepción, incluyen la libertad de elegir una configuración de presentación de su elección (mono, estéreo, envolvente, reproducción por auriculares virtualizada, y así sucesivamente) y la característica de interactividad del usuario: la matriz de presentación, y de esa manera la escena de salida, puede ajustarse y cambiarse interactivamente de acuerdo con la voluntad, preferencias personales u otros criterios. Por ejemplo, es posible ubicar interlocutores de un grupo juntos en un área espacial para maximizar la discriminación de otros interlocutores restantes. Esta interactividad se obtiene incluyendo una

interfaz del usuario del decodificador.

Por cada objeto de sonido transmitido, se puede ajustar su nivel relativo y (en el caso de la presentación no mono) la posición espacial de presentación. Esto puede tener lugar en tiempo real al cambiar el usuario la posición de los dispositivos de deslizamiento de interfaz gráfica del usuario (GUI) asociados (por ejemplo: nivel de objeto = +5 dB, posición de objeto = -30 grados). Sin embargo, se ha descubierto que debido al enfoque paramétrico basado en separación de mezcla descendente/mezcla, la calidad subjetiva de la salida de audio presentada depende de la configuración de los parámetros de presentación. Se ha encontrado que los cambios del nivel relativo de los objetos afectan la calidad de audio final más que los cambios de posición de la presentación espacial (“re-panorámica”). Se ha encontrado también que las configuraciones extremas para los parámetros relativos (por ejemplo, +20 dB) pueden llevar incluso a una calidad de salida inadmisibles. Si bien simplemente esto es el resultado de violación de algunos de los supuestos preceptuales que subyacen en este esquema, sigue siendo inadmisibles que un producto comercial produzca sonido deficiente y perturbaciones dependiendo de la configuración de la interfaz del usuario. En consecuencia, las realizaciones de acuerdo con la invención, como, por ejemplo, el sistema 200, abordan este problema de evitar estas degradaciones inadmisibles independientemente de la configuración de la interfaz del usuario (configuración de la interfaz del usuario que se debe considerar como “parámetros de entrada”).

A continuación, se describen algunos detalles con respecto a los enfoques para evitar las distorsiones SAOC. El enfoque para limitar las distorsiones SAOC presentadas en la presente se basa en los siguientes conceptos:

- Aparecen distorsiones SAOC importantes por elecciones inapropiadas de coeficientes de presentación (que se pueden considerar parámetros de entrada). Esta elección es realizada habitualmente por el usuario de manera interactiva (por ejemplo, por medio de una interfaz gráfica del usuario en tiempo real (GUI) para aplicaciones interactivas). Por lo tanto, se introduce un paso adicional de procesamiento que modifica los coeficientes de presentación que se proporcionan por el usuario (por ejemplo, los limita sobre la base de ciertos cálculos) y utiliza estos coeficientes modificados para el motor de presentación SAOC. Por ejemplo, los coeficientes de presentación que se proporcionaron por el usuario se pueden considerar como parámetros de entrada, y los coeficientes modificados para el motor de presentación SAOC se pueden considerar parámetros modificados.
- Para controlar la degradación excesiva de la salida de audio SAOC producida, es conveniente desarrollar una medida computacional de degradación perceptual (también designada medida de distorsión DM). Se ha hallado que esta medida de la distorsión debe cumplir ciertos criterios:
 - La medida de la distorsión debe ser fácilmente calculable a partir de parámetros internos del motor de decodificación SAOC. Por ejemplo, es conveniente que no sea necesario el cálculo extra de bancos de filtro para obtener la medida de distorsión.
 - El valor de medida de la distorsión se debe correlacionar con la calidad del sonido percibido subjetivamente (degradación perceptual), es decir, debe estar en línea con los principios básicos de la psicoacústica. Para este fin, el cálculo de la medida de la distorsión puede realizarse preferentemente en una manera selectiva en frecuencia, y por lo general se la conoce como la codificación y el procesamiento de audio perceptual.

Se ha descubierto que se pueden definir y calcular una cantidad de medidas de distorsión SAOC. Sin embargo, se ha encontrado que las medidas de distorsión SAOC deben considerar preferentemente ciertos factores básicos a fin de llegar a una evaluación correcta de una calidad SAOC presentada y por consiguiente (aunque no necesariamente) tienen ciertos puntos en común.

- Se consideran los coeficientes de mezcla descendente. Estos determinan las fracciones de mezcla relativas de cada objeto de audio dentro de las unas o más señales de mezcla descendente. Como información de fondo, se debe indicar que se ha encontrado que la distorsión SAOC que se produce depende de la relación entre la mezcla descendente y los coeficientes de presentación. Si la contribución relativa del objeto definida por los coeficientes de presentación es sustancialmente diferente de la contribución relativa del objeto dentro de la mezcla descendente, a continuación el motor de decodificación SAOC (que utiliza los parámetros modificados) tiene que ejecutar un considerable ajuste de la señal de mezcla descendente para convertirla en la salida presentada. Se ha encontrado que esto da lugar a la distorsión SAOC.
- Se consideran los coeficientes de presentación. Estos determinan la potencia de salida relativa de cada objeto de audio a cada una de las unas o más señales de salida presentadas. Como información de fondo, cabe señalar que se ha encontrado que la distorsión SAOC que se produce también depende de la relación mutua de las potencias de los objetos. Si un objeto en cierto punto de tiempo tiene una potencia mucho mayor que otros objetos (y si el coeficiente de mezcla descendente de este objeto no es demasiado pequeño) entonces este objeto domina la mezcla descendente y se reproduce muy favorablemente en la señal de salida presentada. Por el contrario, los objetos débiles solo se representan muy débilmente en la mezcla descendente y por consiguiente no pueden llevarse a niveles de salida elevados sin distorsiones significativas.

- Considerando la potencia de objeto/nivel (relativo) de cada objeto con respecto al otro. Esta información se describe, por ejemplo, en términos de diferencias de niveles de objetos SAOC (OLD). Como información de fondo, se debe tener en cuenta que se ha encontrado que la distorsión SAOC que se produce depende además de las propiedades de las señales de objetos individuales. Por ejemplo, la amplificación de un objeto de una naturaleza tonal en la salida presentada a niveles superiores (mientras que los demás objetos pueden ser más de del tipo ruido) da lugar a una distorsión percibida considerable.
- Además de esto, se puede considerar otra información acerca de las propiedades de las señales de objeto originales. Estas pueden transmitirse por el codificador SAOC como parte de la información complementaria SAOC. Por ejemplo, la información acerca de la tonalidad o el ruido de cada elemento de objeto se puede transmitir como parte de la información complementaria SAOC y utilizar con el fin de limitar las distorsiones.

2.2 reseña general del sistema

Basándose en las consideraciones antes expuestas, se presenta ahora una reseña general del sistema MPEG SAOC 200 para proporcionar un buen entendimiento de la presente invención. Debe indicarse que el sistema SAOC 200 de acuerdo con la Figura 2 es una versión extendida del sistema MPEG SAOC 800 de acuerdo con la Figura 8, por lo que también se aplica la explicación anterior. Además, debe indicarse que el sistema MPEG SAOC 200 puede modificarse de acuerdo con las alternativas de implementación 900, 930, 960 expuestas en las Figuras 9a, 9b y 9c, donde el codificador de objetos corresponde al codificador SAOC, donde la información de interacción con el usuario/información de control de usuario 822 corresponde a la información de control de presentación/coeficiente de presentación.

Adicionalmente, el decodificador SAOC del sistema MPEG SAOC 100 puede reemplazarse por la disposición separada del decodificador de objetos y mezclador/presentador 920, por la disposición integrada de decodificador de objetos y mezclador/presentador 930 o el transcodificador de SAOC a MPEG Envoltante 980.

Tomando ahora como referencia la Figura 2, se puede apreciar que el sistema MPEG SAOC 200 comprende un codificador SAOC 210, que está configurado para recibir una pluralidad de señales de objeto x_1 a x_N , asociadas con una pluralidad de objetos numerados de 1 a N. El codificador SAOC 210 está también configurado para recibir (o de otro modo obtener) coeficientes de mezcla descendente d_1 a d_N . Por ejemplo, el codificador SAOC 210 puede obtener una serie de coeficientes de mezcla descendente d_1 a d_N por cada canal de la señal de mezcla descendente 212 proporcionada por el codificador SAOC 210. El codificador SAOC 210 puede estar configurado, por ejemplo, para obtener una combinación ponderada de las señales de objeto x_1 a x_N a fin de obtener una señal de mezcla descendente, donde cada una de las señales de objeto x_1 a x_N se pondera con su coeficiente de mezcla descendente asociado d_1 a d_N . El codificador SAOC 210 está configurado asimismo para obtener información de relación entre objetos, que describe una relación entre diferentes señales de objetos. Por ejemplo, la información de relación entre objetos puede comprender información de diferencia del nivel de objeto, por ejemplo, en forma de parámetros OLD e información de correlación entre objetos, por ejemplo en forma de parámetros IOC. En consecuencia el codificador SAOC 200 está configurado entonces para suministrar una o más señales de mezcla descendente 212, cada una de las cuales comprende una combinación ponderada de una o más señales de objetos, ponderados de acuerdo con una serie de parámetros de mezcla descendente asociados con la respectiva señal de mezcla descendente (o un canal de la señal de mezcla descendente de múltiples canales 212). El codificador SAOC 210 también está configurado para proporcionar información complementaria 214, donde la información complementaria 214 comprende la información de relación entre objetos (por ejemplo, en forma de parámetros de diferencia de nivel entre objetos y parámetros de correlación entre objetos). La información complementaria 214 también comprende una información de parámetros de mezcla descendente, por ejemplo, en forma de parámetros de ganancia de mezcla descendente y parámetros de diferencias de nivel de los canales de mezcla descendente. La información complementaria 214 puede comprender además una información complementaria opcional sobre propiedades de los objetos, que puede representar las propiedades individuales de los objetos. A continuación se describen los detalles con respecto a la información complementaria opcional sobre propiedades de los objetos.

El sistema MPEG SAOC 200 comprende además un decodificador SAOC 220, que puede comprender la funcionalidad del decodificador SAOC 820. En consecuencia, el decodificador SAOC 220 recibe la una o más señales de mezcla descendente 212 e información complementaria 214 así como también los coeficientes de presentación modificados ("o ajustados", o "reales") 222 y suministra, sobre la base de estos, una o más señales con mezcla ascendente \hat{y}_1 a \hat{y}_N .

El sistema MPEG SAOC 200 comprende además un aparato 240 para suministrar uno o más parámetros modificados (o "ajustados" o "reales"), es decir los coeficientes de presentación modificados 222, dependiendo de uno o más parámetros de entrada, es decir los parámetros de entrada que describen una información de control de presentación o coeficientes de presentación 242. El aparato 240 está configurado para recibir también al menos una parte de la información complementaria 214. Por ejemplo, El aparato 240 está configurado para recibir los parámetros 214a que describen potencias de los objetos (por ejemplo, potencias de las señales de objeto x_1 a x_N).

Por ejemplo, los parámetros 214a pueden comprender los parámetros de diferencia de nivel de los objetos (también designados OLD). El aparato 240 también recibe preferentemente parámetros 214b de la información complementaria 214 que describen los coeficientes de mezcla descendente. Por ejemplo, los parámetros 214b describen los coeficientes de mezcla descendente d_1 a d_N . Opcionalmente, el aparato 240 puede recibir además parámetros adicionales 214c, que constituyen una información complementaria sobre propiedades de los objetos individuales.

El aparato 240 está configurado en general para proporcionar los coeficientes de presentación modificados 222 sobre la base de los coeficientes de presentación de entrada 242 (que pueden recibirse, por ejemplo, desde una interfaz del usuario, o pueden calcularse, por ejemplo, dependiendo de la entrada del usuario o proporcionarse como información predeterminada), por lo que se reduce la distorsión de la representación de la señal de mezcla ascendente que sería causada por el uso de parámetros de presentación no óptimos por el decodificador SAOC 220. En otras palabras, los coeficientes de presentación modificados 222 son una versión modificada de los coeficientes de presentación de entrada 242, donde los cambios se realizan dependiendo de los parámetros 214a, 214b, de modo tal que se reducen o limitan todas las distorsiones audibles en las señales de canales de mezcla ascendente \hat{y}_1 a \hat{y}_N (que forman la representación de señal de mezcla ascendente).

El aparato 240 para suministrar el uno o más parámetros ajustados 242 puede comprender por ejemplo, un ajustador de coeficientes de presentación 250, que recibe los coeficientes de presentación de entrada 242 y suministra, sobre la base de estos, los coeficientes de presentación modificados 222. Para este fin, el ajustador de coeficientes de presentación 250 puede recibir una medida de la distorsión 252, que describe las distorsiones que serían causadas por el uso de los coeficientes de presentación de entrada 242. La medida de la distorsión 252 puede proporcionarse, por el ejemplo, por el calculador de distorsión 260 dependiendo de los parámetros 214a, 214b y los coeficientes de presentación 242.

Sin embargo, las funcionalidades del ajustador de coeficientes de presentación 250 y del calculador de distorsión 260 también pueden estar integradas en una unidad funcional única, de tal manera que los coeficientes de presentación modificados 222 se proporcionen sin cálculo explícito de una medida de distorsión 252. Por el contrario, se pueden aplicar mecanismos implícitos para reducir o limitar la medida de la distorsión.

Con respecto a la funcionalidad del sistema MPEG SAOC 200, debe indicarse que la representación de señal de mezcla ascendente, que se produce como salida en forma de señales de canales de mezcla ascendente \hat{y}_1 a \hat{y}_N , se genera con buena calidad perceptual puesto que las distorsiones audibles que serían causadas por una elección incorrecta de la información de interacción con el usuario/información de control del usuario 822 en el sistema de referencia 800, se evitan mediante la modificación o ajuste de los coeficientes de presentación. La modificación o ajuste se ejecuta por el aparato 240 de tal manera que se evitan las degradaciones graves de la impresión perceptual, o de tal manera que al menos se reduzcan las degradaciones de la impresión perceptual en comparación con un caso en el cual el decodificador SAOC 220 utilice directamente (sin modificación o ajuste) los coeficientes de presentación de entrada 242.

A continuación, se resume brevemente la funcionalidad del concepto de la invención. Dada una medida de la distorsión (DM), se puede evitar la distorsión excesiva en la salida de audio calculando el valor medida de la distorsión correspondiente a las señales dadas, y modificando el algoritmo de decodificación SAOC (limitando los coeficientes de presentación 212 utilizados en realidad) de tal manera que el valor medida de la distorsión no exceda un determinado umbral. En la Figura 2 se ilustra un sistema 200 de acuerdo con este concepto y que se ha explicado con cierto detalle anteriormente.

Con referencia al sistema 200, se pueden realizar las siguientes observaciones:

- Los coeficientes de presentación deseados 242 se introducen por el usuario u otra interfaz.
- Antes de aplicarse en el motor de decodificación de SAOC 220, los coeficientes de presentación 242 se modifican por un ajustador de coeficientes de presentación 250, que hace uso de una o más medidas de la distorsión calculadas 252, que se suministran por un calculador de distorsión 260.
- El calculador de distorsión 260 evalúa la información (por ejemplo, parámetros 214a, 214b) de la información complementaria 214 (por ejemplo la potencia relativa de los objetos/OLD, coeficientes de mezcla descendente, y - opcionalmente - información sobre propiedades de señales de objeto). Además, se basa en la entrada de coeficientes de presentación 242 deseada.

En una realización preferida, el aparato 240 está configurado para modificar los coeficientes de presentación sobre la base de una medida de la distorsión. Preferentemente, los coeficientes de presentación se ajustan en forma selectiva en frecuencia, usando por ejemplo, ponderación selectiva en frecuencia.

La modificación de los coeficientes de presentación se puede basar en esta trama (por ejemplo, en una trama actual), o los coeficientes de presentación pueden ajustarse en el tiempo no solo en forma de trama a trama, sino también procesado/controlado en el tiempo (por ejemplo, suavizados en el tiempo) donde se pueden aplicar posiblemente diferentes constantes de ataque/decaimiento, o en el caso de un compresor/limitador de rangos dinámicos.

En algunas realizaciones, la medida de la distorsión puede ser selectiva en todas las frecuencias.

En algunas realizaciones, la medida de la distorsión puede considerar una o más de las siguientes características.

- Potencia/energía/nivel de cada objeto;
- Coeficientes de mezcla descendente;
- Coeficientes de presentación; y/o
- Información complementaria adicional sobre propiedades de los objetos, en caso de aplicarse.

En algunas realizaciones, la medida de la distorsión se puede calcular por cada objeto y combinar para llegar a una distorsión total.

En algunas realizaciones, se puede evaluar una información complementaria adicional sobre propiedades de los objetos 214c. La información complementaria adicional sobre propiedades de los objetos 214c puede extraerse en un codificador SAOC mejorado, por ejemplo, en el codificador SAOC 210. La información complementaria adicional sobre propiedades de los objetos puede incluirse, por ejemplo, en un flujo de bits SAOC potenciado, lo que se describe con referencia a la Figura 7. Además, la información complementaria adicional sobre propiedades de los objetos puede utilizarse para limitar las distorsiones por un codificador SAOC mejorado.

En un caso especial, se puede utilizar el ruido/tonalidad como una propiedad del objeto descrito por la información complementaria adicional sobre propiedades de los objetos. En este caso, el ruido/tonalidad puede transmitirse con una resolución de frecuencia mucho más basta que otros parámetros del objeto (por ejemplo, OLD) para guardar en la información complementaria. En un caso extremo, se puede transmitir la información complementaria sobre propiedades de objetos respecto a ruido/tonalidad con solo una información por objeto (por ejemplo, con características de banda ancha).

2.3 Métrica de distorsión SAOC

A continuación, se describe una pluralidad de medidas de distorsión diferentes, que se pueden obtener, por ejemplo, utilizando el calculador de distorsión 260. Los detalles con respecto a la aplicación de estas medidas de distorsión para la limitación de los coeficientes de presentación se describen más adelante en la sección 2.4.

En otras palabras, esta sección resume varias medidas de distorsión. Estas se pueden utilizar individualmente o se las puede combinar para formar una métrica compuesta, más compleja de distorsión, por ejemplo, mediante la adición ponderada de valores de métrica de distorsión individuales. Cabe señalar aquí que los términos "medida de distorsión" y "métrica de distorsión" designan cantidades similares y en la mayoría de los casos no es necesario distinguirlas.

A continuación, se describe una pluralidad de métricas de distorsión que pueden evaluarse por el calculador de distorsión 260 y que pueden utilizarse por el ajustador de coeficientes de presentación 250 a fin de obtener los coeficientes de presentación modificados 222 sobre la base de los coeficientes de presentación de entrada 242.

2.3.1 Medida de distorsión N.º 1

A continuación, se describe una primera medida de distorsión (también designada medida de distorsión N.º 1).

Para mayor simplicidad conceptual, se considera un sistema N-1-1 SAOC (por ejemplo, una señal de mezcla descendente mono (212) y un canal de mezcla ascendente único (señal)). Se realiza la mezcla descendente de N objetos de audio de entrada en una señal mono y se presentan en una salida mono. Como se presenta en la Figura 8 los coeficientes de mezcla descendente se designan $d_1...d_N$ y los coeficientes de presentación están indicados por $r_1...r_N$. En las siguientes fórmulas, se han omitido los índices de tiempo para simplificar. Del mismo modo, se han omitido los índices de frecuencia, señalando que las ecuaciones se relacionan con las señales de subbanda. En algunas de las ecuaciones siguientes, las letras minúsculas indican coeficientes o señales, y las letras mayúsculas indican las potencias correspondientes, que se pueden observar desde el contexto de las ecuaciones. También, cabe señalar que las señales en ocasiones están representadas por correspondientes coeficientes en el dominio de tiempo-frecuencia, en lugar en el dominio del tiempo.

Supongamos que el objeto N.º m (índice de objeto auditivo m) es un objeto de interés, por ejemplo el objeto más

dominante que se incrementa en su nivel relativo y, de esa manera, limita la calidad general del sonido. A continuación la señal de salida deseada ideal (señal de canal de mezcla ascendente) está dada por

$$\hat{y}_1 = [x_m \cdot r_m] + \left[\sum_{i=1; i \neq m}^N x_i \cdot r_i \right] \quad (1)$$

5 En este caso, el primer término es la contribución deseada del objeto de interés a la señal de salida, mientras que el segundo término indica las contribuciones de todos los demás objetos ("interferencia").

En realidad, sin embargo, debido al proceso de mezcla descendente, la señal de salida está dada por

$$y_1 = t \cdot \sum_{i=1}^N x_i \cdot d_i = [x_m \cdot t \cdot d_m] + \left[\sum_{i=1; i \neq m}^N x_i \cdot t \cdot d_i \right] \quad (2)$$

10 es decir, la señal de mezcla descendente se escala seguidamente por un coeficiente de transcodificación, t , correspondiente a la matriz "m2" de un decodificador MPEG Envolvente. Una vez más, ésta se puede dividir en un primer término (contribución real de la señal de objeto a la señal de salida) y un segundo término ("interferencia" real
15 por otras señales de objeto). En este caso, el sistema SAOC (por ejemplo, el decodificador SAOC 220, y, opcionalmente, también el aparato 240) determina dinámicamente el coeficiente de transcodificación, t , de tal manera que la potencia de la señal de salida presentada en realidad se equipare a la potencia de la señal ideal:

$$\hat{Y}_1 = Y_1 \Rightarrow t^2 = \frac{\sum_{i=1}^N r_i^2 \cdot X_i}{\sum_{i=1}^N d_i^2 \cdot X_i} \quad (3)$$

20 Una medida de la distorsión (DM) se puede definir calculando la relación entre la contribución de potencia real del objeto $N.^o$ m y su contribución de potencia real:

$$dm_1(m) = \frac{P_{ideal}}{P_{real}} = \frac{r_m^2}{d_m^2 \cdot t^2} = \frac{r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i}{d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i} \quad (4)$$

25 NN

En este caso, $\sum_{i=1}^N r_i^2 \cdot X_i$ indica la potencia de la señal presentada en última instancia, y $\sum_{i=1}^N d_i^2 \cdot X_i$ es la potencia de la señal de mezcla descendente. Obsérvese que, en la implementación real, los valores X_i pueden reemplazarse directamente por los valores correspondientes de *Diferencia de Nivel de los Objetos (OLD_i)* que se transmiten como parte de la información complementaria SAOC 214.

30 Para una mejor interpretación de la dm_1 , se puede reformular su definición de la siguiente manera:

$$dm_1(m) = \frac{r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i}{d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i} = \frac{\frac{r_m^2 \cdot X_m}{\sum_{i=1}^N r_i^2 \cdot X_i}}{\frac{d_m^2 \cdot X_m}{\sum_{i=1}^N d_i^2 \cdot X_i}} \quad (4a)$$

En efecto, esto significa que la métrica de distorsión es la relación de la contribución de potencia relativa de los

objetos en la señal presentada de manera ideal (salida) frente la señal de mezcla descendente (entrada). Esto coincide con el hallazgo de que el esquema SAOC da mejor resultado cuando no tiene que alterar las potencias relativas de los objetos en factores elevados.

5 El aumento de los valores de dm_1 indica la reducción de la calidad del sonido con respecto al objeto de sonido N.º m. Se ha descubierto que el valor de dm_1 se mantiene constante si todos los coeficientes de presentación se escalan por un factor común, o si todos los coeficientes de mezcla descendente se escalan de igual manera. También se ha descubierto que aumentando el coeficiente de presentación correspondiente al objeto N.º m (incrementando su nivel relativo) se da lugar a una mayor distorsión. Los valores de dm_1 se pueden interpretar de la siguiente manera:

- 10
- Un valor de 1 indica una calidad ideal con respecto al objeto N.º m;
 - Los valores dm_1 por encima de 1 indican el descenso de la calidad;
 - Los valores de dm_1 por debajo de 1 no aumentan más la calidad con respecto al objeto N.º m.
- 15 En consecuencia, la medida general de la calidad del sonido (es decir la calidad de todos los objetos) se puede calcular de la siguiente manera:

$$DM_1 = \frac{\sum_{m=1}^N w(m) \cdot \max[dm_1(m), 1]}{\sum_{m=1}^N w(m)} \quad (5)$$

20 En esta ecuación, $w(m)$ indica un factor de ponderación del objeto N.º m que se relaciona con la significancia y sensibilidad del objeto específico dentro de la escena de audio. Por ejemplo, se podría elegir entonces $w(m)$ dependiendo de la potencia / volumen del objeto $w(m) = (r_m^2 X_m)^\alpha$ donde α puede elegirse típicamente como 0,25 para emular, a grandes rasgos, el crecimiento del volumen psicoacústico correspondiente a este objeto. Adicionalmente, $w(m)$ podría tener en cuenta fenómenos de tonalidad y enmascaramiento. Por otro lado, se puede

25 ajustar $w(m)$ a 1, lo que facilita el cálculo de DM_1 .

2.3.2 Medida de distorsión N.º 2

30 Se puede construir una medida de la distorsión alternativa a partir de la ecuación (4) para formar una medida perceptual al estilo de una Relación Ruido a Máscara (NMR), es decir, calcular la relación entre ruido/interferencia y el umbral de enmascaramiento:

$$dm_2(m) = \frac{P_{Ruido}}{Máscara} = \frac{P_{real} - P_{real}}{msr \cdot P_{total}} = \frac{(r_m^2 - d_m^2 \cdot t^2) \cdot X_m}{msr \cdot \sum_{i=1}^N r_i^2 \cdot X_i} = \frac{\left(r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i - d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i \right) \cdot X_m}{msr \cdot \left(\sum_{i=1}^N r_i^2 \cdot X_i \right) \cdot \left(\sum_{i=1}^N d_i^2 \cdot X_i \right)}$$

(6)

35 En esta ecuación, msr es la Relación Máscara a Señal de la señal de audio total que depende de su tonalidad. El aumento de los valores de dm_2 indica una mayor distorsión con respecto al objeto de sonido N.º m. Una vez más, el valor de dm_2 se mantiene constante si todos los coeficientes de presentación se escalan por un factor común, o si todos los coeficientes de mezcla descendente se escalan del mismo modo. El rango de valores de dm_2 se puede

40 interpretar de la siguiente manera:

- Un valor de 0 indica la calidad ideal con respecto al objeto N.º m;
- El aumento de los valores dm_2 por encima de 1 indica degradaciones audibles progresivas;
- Los valores de dm_2 por debajo de 1 indican calidad no distinguible con respecto al objeto N.º m.

45 En consecuencia, una medida general de la calidad de escena sonora (es decir la calidad correspondiente a todos los objetos) se puede calcular de la siguiente manera:

$$DM_2 = \frac{\sum_{m=1}^N w(m) \cdot \max[dm_2(m), 1]}{\sum_{m=1}^N w(m)} \quad (7)$$

Una vez más, $w(m)$ indica un factor de ponderación del objeto N.º m que se relaciona con la significancia / nivel / volumen del objeto específico dentro de la escena de audio, que típicamente se selecciona en términos de $w(m) = (r_m^2 X_m)^\alpha$ con $\alpha = 0,25$.

5 La medida de la distorsión en la ecuación (6) calcula la distorsión en términos de diferencia de las potencias (esto corresponde a una medición de "NMR con la diferencia espectral"). Por otro lado, se puede calcular la distorsión sobre la base de una forma de onda que lleva a la siguiente medida que incluye un término de producto mezclado adicional:

$$dm_2'(m) = \frac{P_{\text{Ruido}}}{\text{Máscara}} = \frac{E\{|y_{m;\text{ideal}} - \hat{y}_{m;\text{real}}|^2\}}{MSF \cdot P_{\text{total}}} =$$

$$\frac{\left| r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i + d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i - 2 \cdot d_m r_m \cdot \sqrt{\left(\sum_{i=1}^N r_i^2 \cdot X_i\right) \cdot \left(\sum_{i=1}^N d_i^2 \cdot X_i\right)} \right| \cdot X_m}{MSF \cdot \left(\sum_{i=1}^N r_i^2 \cdot X_i\right) \cdot \left(\sum_{i=1}^N d_i^2 \cdot X_i\right)} \quad (8)$$

2.3.3 Medida de distorsión N.º 3

15 Se presenta una tercera medida de la distorsión que describe la coherencia entre la señal de mezcla descendente y la señal presentada. La mayor coherencia da lugar a una mejor calidad subjetiva del sonido. Además, se puede tener en cuenta la correlación de los objetos de audio si los datos de IOC están presentes en el decodificador SAOC.

20 A partir de los parámetros de SAOC (por ejemplo, los parámetros 214a, que pueden comprender los parámetros de diferencia de nivel de los objetos y los parámetros de correlación entre objetos) se puede determinar un modelo de la covarianza de los objetos

$$\mathbf{E} = \sqrt{\mathbf{OLD}^T \cdot \mathbf{OLD} \cdot \mathbf{IOC}}$$

25 Para calcular la medida de la distorsión se integra una Matriz \mathbf{M} que contiene los coeficientes de presentación y mezcla descendente (\mathbf{M} se puede interpretar como una matriz de presentación correspondiente a un sistema N-1-2 SAOC)

$$\mathbf{M} = \begin{pmatrix} r_1 & r_2 & \cdots & r_N \\ d_1 & d_2 & \cdots & d_N \end{pmatrix}$$

30 La covarianza entre la señal de mezcla descendente y presentada \mathbf{C} es entonces

$$\mathbf{C} = \mathbf{M} \cdot \mathbf{E} \cdot \mathbf{M}^* = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

Una medida de la distorsión DM_3 se define como

$$DM_3 = 1 - \min\left(\frac{|c_{12}|}{\sqrt{c_{11} \cdot c_{22}}}, 1\right)$$

Los valores de DM_3 se pueden interpretar de la siguiente manera:

- 40 • Los valores están en el rango [0 .. 1] e indican la coherencia entre la señal de mezcla descendente y la

presentada.

- Un valor de 0 indica la calidad ideal.
- El aumento de los valores DM₃ indica un descenso de la calidad.

5 2.3.4 Medida de distorsión N.º 4

2.3.4.1 reseña general

10 Este enfoque propone el uso, como medida de la distorsión, de la relación ponderada promediada entre la energía de presentación objetivo (UPMIX) y la energía de mezcla descendente óptima (calculada a partir de una mezcla descendente dada DMX).

15 Para obtener detalles, se hace referencia asimismo a la Fig. 4, que ilustra una representación gráfica de la mezcla descendente (DMX), la energía de mezcla descendente óptima (DMX_{opt}) y la energía de presentación objetivo (UPMIX).

2.3.4.2 Nomenclatura

20	$ch = \{1, 2, \dots, N_{ch}\}$ $dx = \{1, 2\}$ $ob = \{1, 2, \dots, N_{ob}\}$ $pb = \{1, 2, \dots, N_{pb}\}$ $r_{ch,ob,pb} = r(ch, ob, pb)$	índice correspondiente a los canales de mezcla ascendente índice correspondiente a los canales de mezcla descendente índice correspondiente a objetos de audio índice correspondiente a bandas de parámetros matriz de presentación correspondiente al canal ch, objeto de audio ob y banda de parámetros pb
25	$d_{dx,ob,pb} = d(dx, ob, pb)$ $w_{ob,pb} = w(ob, pb)$ $NRG_{pb} = NRG(pb)$	matriz de mezcla descendente correspondiente al canal de mezcla descendente dx, objeto de audio ob y banda de parámetros pb factor de ponderación que representa la significancia / nivel / volumen del objeto de audio ob correspondiente a la banda de parámetros pb energía absoluta del objeto de audio con la energía más elevada correspondiente a la banda de frecuencias pb
30	$OLD_{ob,pb} = OLD(ob, pb)$	diferencia de nivel de los objetos, que describe las diferencias de intensidad entre un objeto de audio ob y el objeto con la energía más elevada de la correspondiente banda de frecuencias pb
35	$IOC_{ob,ab,pb} = IOC(ob_i, ob_j, pb)$	correlación entre objetos, que describe la correlación entre dos canales de objetos de audio.

2.3.4.3 Algoritmo

40 A continuación se describen brevemente los pasos de un algoritmo para obtener la medida de la distorsión N.º 4:

- Cálculo de las energías relativas de mezcla ascendente y mezcla descendente:

$$\hat{r}_{ch,ob,pb}^2 = OLD_{ob,pb} \cdot r_{ch,ob,pb}^2, \quad \hat{d}_{dx,ob,pb}^2 = OLD_{ob,pb} \cdot d_{dx,ob}^2$$

$$\sum_{ob=1}^{N_{ob}} \tilde{r}_{ch,ob,pb}^2 = 1 \quad \text{y} \quad \sum_{ob=1}^{N_{ob}} \tilde{d}_{dm,ob,pb}^2 = 1$$

- Normalización de las energías, de tal manera que

$$\tilde{r}_{ch,ob,pb}^2 = \frac{\hat{r}_{ch,ob,pb}^2}{\sum_{ob=1}^{N_{ob}} \hat{r}_{ch,ob,pb}^2}, \quad \tilde{d}_{dm,ob,pb}^2 = \frac{\hat{d}_{dm,ob,pb}^2}{\sum_{ob=1}^{N_{ob}} \hat{d}_{dm,ob,pb}^2}$$

- Construcción de la mezcla descendente óptima $d_{ch,ob,pb}^{2(opt)}$ por cada canal y banda de mezcla ascendente:

$$d_{ch,ob,pb}^{2(opt)} = \alpha_{ch,ob,pb} \cdot \tilde{d}_{1,ob,pb}^2 + \beta_{ch,ob,pb} \cdot \tilde{d}_{2,ob,pb}^2$$

Las constantes de multiplicación $\alpha_{ch,ob,pb}$, $\beta_{ch,ob,pb}$ se calculan resolviendo el sistema sobredefinido de ecuaciones

lineales para satisfacer la siguiente condición: $\left\| d_{ch,ob,pb}^{2(opt)} - \tilde{r}_{ch,ob,pb}^2 \right\|_{\alpha, \beta} \rightarrow 0$

- Cálculo de la medida de la distorsión:

$$DM_4 = \sum_{ob=1}^{N_{ob}} \sum_{ch=1}^{N_{ch}} \left| 1 - \frac{\tilde{r}_{ch,ob,pb}^2}{d_{ch,ob,pb}^{2(opt)}} \right| w_{ob,pb} \hat{r}_{ch,ob,pb}^2$$

2.3.4.4 Control de distorsión

5 Se logra controlar la distorsión limitando uno o más coeficientes de presentación que dependen de la medida de la distorsión DM4.

Se apreciará que (i) la medida sólo es pertinente para el caso de la mezcla descendente estéreo, y (ii) puede reducirse a DM1 en el caso de N.º dx=1 y N.º ch=1.

10

2.3.4.5 Propiedades

A continuación se resumen brevemente las propiedades del concepto de cálculo de la medida de la distorsión número 4. El concepto

15

- presume la transcodificación ideal
- puede tratar la mezcla descendente estéreo y
- permite la generalización a una presentación multicanal.

2.3.5 Medida de distorsión N.º 5

20

Se sugiere un cálculo alternativo del coeficiente de transcodificación t . Se lo puede interpretar en términos de extensión de t y lleva a la matriz de transcodificación T que se caracteriza por la incorporación de la coherencia entre objetos (IOC) y al mismo tiempo se extiende a la métrica actual DM N.º 1 y DM N.º 2 a mezcla descendente estéreo y mezcla ascendente multicanal. La implementación actual del coeficiente de transcodificación t considera la paridad de la potencia de la señal de salida presentada realmente y la potencia de la señal presentada ideal, es decir

25

$$t^2 = \frac{\sum_{i=1}^N r_i^2 X_i}{\sum_{i=1}^N d_i^2 X_i}$$

La incorporación de la matriz de covarianza E produce una formulación modificada para t , es decir la matriz de transcodificación T , que tiene en cuenta, también, la coherencia entre objetos. Los elementos de E se calculan a partir de los parámetros SAOC 214, como

30

$$e_{ij} = \sqrt{OLD_i OLD_j} IOC_{ij}$$

35 La matriz de transcodificación representa la conversión de la señal de mezcla descendente a la de salida presentada, de tal manera que $TDx \approx Rx$. Esto se obtiene mediante la minimización del error cuadrático medio, para dar

$$T = RED^* (DED^*)^{-1}$$

40 Con

$$H = RED^* \circ h_{ij} = \sum_{l=1}^N \sum_{m=1}^N r_{il} d_{jm} e_{lm}$$

y

$$V = DED^* \circ v_{ij} = \sum_{l=1}^N \sum_{m=1}^N d_{il} d_{jm} e_{lm}$$

45

la medida de la distorsión al estilo de dm_1 aunque en este caso para cada combinación de mezcla

descendente/presentación (n, k) del objeto m está dada por

$$dm_5^n(m, n, k) = \frac{r_{m,k}^2 v_{n,n}}{d_{m,n}^2 h_{k,n}}$$

Considerando $dm_1(m)$ por separado para los canales de mezcla descendente izquierdo y derecho lleva a

$$dm_L(m, k) = \frac{r_{m,k}^2 v_{1,1}}{d_{m,1}^2 h_{k,1}} \quad y \quad dm_R(m, k) = \frac{r_{m,k}^2 v_{2,2}}{d_{m,2}^2 h_{k,2}}$$

5 Se puede suponer que el mejor de los dos trayectos de mezcla descendente/mezcla ascendente es relevante para la calidad de la salida presentada, por lo que la medida corresponde al valor mínimo, es decir

$$dm_5'(m, k) = \min[dm_L, dm_R]$$

10 Una medida general de todos los canales de salida, designada por el índice k , se puede calcular de la siguiente manera

$$dm_5(m) = \frac{\sum_{k=1}^{N_{Ch}} dm_5'(m, k) r_{m,k}^2 X_m}{\sum_{k=1}^{N_{Ch}} r_{m,k}^2 e_{k,k}}$$

15 La medida general de todos los objetos se puede obtener según

$$DM_5 = \frac{\sum_{m=1}^N w(m) \max[dm_5(m), 1]}{\sum_{m=1}^N w(m)} \quad \text{con } w(m) = [r_m^2 X_m]^\alpha \text{ como antes.}$$

20 Es posible una extensión similar de t a T en el caso de dm_2 y dm_2' .

2.3.6. medida de distorsión N.º 6

A continuación se describe una sexta medida de la distorsión.

25 Supongamos que $e_i(t)$ es el cuadrado de la envolvente de Hilbert de la señal de objeto N.º i y P_i la potencia de la señal de objeto N.º i (ambas típicamente dentro de una subbanda), entonces se puede obtener una medida N de la semejanza tonalidad/ruido de un cálculo normalizado de la varianza de la envolvente de Hilbert como

$$N_i = \frac{\text{var}\{e_i\}}{P_i^2}$$

30 Por otro lado, también se puede utilizar la potencia / varianza de señal con diferencia de envolvente de Hilbert en lugar de la varianza de la envolvente de Hilbert en sí. En cualquier caso, la medida describe la fuerza de la fluctuación de la envolvente en el tiempo.

35 Esta medida de la semejanza tonalidad/ruido, N , se puede determinar tanto con respecto a la mezcla de señales presentadas idealmente como a la mezcla de sonido presentada SAOC real y se puede calcular una medida de la distorsión a partir de la diferencia entre ambas, por ejemplo:

$$DM_6 = |N_{ideal} - N_{real}|^p$$

40

donde β es un parámetro (por ejemplo $\beta = 2$).

5 2.3.7. Cálculo de las energías de las imágenes de las señales de origen correspondientes a la escena de referencia y la escena presentada por SAOC

Para calcular las energías de objeto de la imagen de origen en la escena de referencia y la presentada por SAOC utilizadas para las medidas de la distorsión, hay que tener en cuenta la matriz de transcodificación T correspondiente a la escena presentada por SAOC como se hace en "Medida de distorsión 5", aunque también la correlación de las señales de origen tanto de la escena de referencia como de la escena presentada.

10 *Comentario:* La notación de las señales en letras mayúsculas refleja en este caso la notación de matriz de las señales, no las energías de las señales como en los capítulos anteriores

En el caso de un origen arbitrario x_m las partes de señal de x_m de todas las fuentes x_i se pueden calcular de la siguiente manera:

Se dividen todas las señales de origen x_i en una parte de señal $x_{i||m}$ que se correlaciona con el objeto de interés x_m y una parte $x_{i\perp m}$ que no se correlaciona con x_m . Esto se puede lograr mediante la proyección subespacial de x_m sobre todas las señales x_i , es decir $x_i = x_{i||m} + x_{i\perp m}$. La parte correlacionada está dada por

$$x_{i||m} = \frac{x_m^T x_i}{x_m^T x_m} x_m = \frac{IOC_{i,m}}{\|x_m\|^2} x_m = g_{i,m} x_m$$

20 2.3.7.1 Cálculo de P_{ideal,x_m} a partir de la imagen de origen y_{x_m} en la escena de referencia Y :

Donde $Y = RX$ y $X = X_{\perp m} + X_{||m}$, la imagen y_{x_m} de la fuente x_m correspondiente a todos los canales presentados se puede calcular mediante $Y_{x_m} = RX_{||m}$ donde

$$X_{||m} = \begin{pmatrix} x_{1||m}^T \\ x_{2||m}^T \\ \vdots \\ x_{N||m}^T \end{pmatrix} = \begin{pmatrix} g_{1,m} x_m^T \\ g_{2,m} x_m^T \\ \vdots \\ g_{N,m} x_m^T \end{pmatrix}$$

Y_{x_m} se puede calcular según

$$Y_{x_m} = RX_{||m} = \begin{pmatrix} r_{ch_1,x_1} & r_{ch_1,x_2} & \vdots & r_{ch_1,x_N} \\ r_{ch_2,x_1} & r_{ch_2,x_2} & \vdots & r_{ch_2,x_N} \\ \dots & \dots & \ddots & r_{ch_{N-1},x_N} \\ r_{N_{ch},x_1} & r_{N_{ch},x_2} & r_{N_{ch},x_{N-1}} & r_{N_{ch},x_N} \end{pmatrix} \begin{pmatrix} g_{1,m} x_m^T \\ g_{2,m} x_m^T \\ \vdots \\ g_{N,m} x_m^T \end{pmatrix}$$

Por lo tanto, la energía P_{ideal,x_m} de la imagen de origen Y_{x_m} en la escena de referencia será:

$$P_{ideal, x_m} = \begin{pmatrix} \|r_{ch_1, x_1} g_{1,m} + r_{ch_1, x_2} g_{2,m} + \dots + r_{ch_1, x_N} g_{N,m}\|^2 \|x_m\|^2 \\ \dots \\ \|r_{N_{ch}, x_1} g_{1,m} + r_{N_{ch}, x_2} g_{2,m} + \dots + r_{N_{ch}, x_N} g_{N,m}\|^2 \|x_m\|^2 \end{pmatrix}$$

2.3.7.2 Cálculo de P_{real, x_m} de la imagen de origen \hat{y}_{x_m} en la escena presentada por SAOC \hat{y} :

Esto se puede realizar de manera igual a la correspondiente a P_{ideal, x_m} donde T es la matriz de transcodificación y

5 D la matriz de mezcla descendente, \hat{y}_{x_m} para todos los canales de la escena presentada es:

$$\hat{Y}_{x_m} = T^{0.5} D X_{\parallel m}$$

Utilizando $D = \begin{pmatrix} d_{11} & \dots & d_{1N} \\ d_{21} & \dots & d_{2N} \end{pmatrix}$ y $T = \begin{pmatrix} t_{11} & t_{12} \\ \vdots & \vdots \\ t_{N_{ch}1} & t_{N_{ch}2} \end{pmatrix}$

$$\hat{Y}_{x_m} = \begin{pmatrix} \sqrt{t_{11}} d_{11} + \sqrt{t_{12}} d_{21} & \sqrt{t_{11}} d_{12} + \sqrt{t_{12}} d_{22} & \dots & \sqrt{t_{11}} d_{1N} + \sqrt{t_{12}} d_{2N} \\ \sqrt{t_{21}} d_{11} + \sqrt{t_{22}} d_{21} & \sqrt{t_{21}} d_{12} + \sqrt{t_{22}} d_{22} & \dots & \sqrt{t_{21}} d_{1N} + \sqrt{t_{22}} d_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{t_{N_{ch}1}} d_{11} + \sqrt{t_{N_{ch}2}} d_{21} & \sqrt{t_{N_{ch}1}} d_{12} + \sqrt{t_{N_{ch}2}} d_{22} & \dots & \sqrt{t_{N_{ch}1}} d_{1N} + \sqrt{t_{N_{ch}2}} d_{2N} \end{pmatrix} \begin{pmatrix} g_{1,m} x_m^T \\ g_{2,m} x_m^T \\ \vdots \\ g_{N,m} x_m^T \end{pmatrix}$$

10

Por lo tanto, la energía P_{real, x_m} de la imagen de origen \hat{Y}_{x_m} de la escena de referencia ha de ser:

$$P_{real, x_m} = \begin{pmatrix} \|g_{1,m} (\sqrt{t_{11}} d_{11} + \sqrt{t_{12}} d_{21}) + g_{2,m} (\sqrt{t_{11}} d_{12} + \sqrt{t_{12}} d_{22}) + \dots + g_{N,m} (\sqrt{t_{11}} d_{1N} + \sqrt{t_{12}} d_{2N})\|^2 \|x_m\|^2 \\ \dots \\ \|g_{1,m} (\sqrt{t_{N_{ch}1}} d_{11} + \sqrt{t_{N_{ch}2}} d_{21}) + g_{2,m} (\sqrt{t_{N_{ch}1}} d_{12} + \sqrt{t_{N_{ch}2}} d_{22}) + \dots + g_{N,m} (\sqrt{t_{N_{ch}1}} d_{1N} + \sqrt{t_{N_{ch}2}} d_{2N})\|^2 \|x_m\|^2 \end{pmatrix}$$

15

2.3.7.3. Cálculo de la medida de la distorsión

La medida de la distorsión al estilo de dm_1 se puede calcular para cada objeto m y canal de presentación de salida k de la siguiente manera:

$$dm_1^i(m, k) = \frac{P_{ideal}}{P_{real}} = \frac{\|r_{k1} IOC_{1m} + \dots + r_{kN} IOC_{Nm}\|^2}{\|(\sqrt{t_{k1}} d_{11} + \sqrt{t_{k2}} d_{21}) IOC_{1m} + \dots + (\sqrt{t_{k1}} d_{1N} + \sqrt{t_{k2}} d_{2N}) IOC_{Nm}\|^2}$$

20

$$dm_7(m) = \frac{\sum_{k=1}^{N_{Ch}} dm_7'(m,k) r_{m,k}^2 \|x_m\|^2}{\sum_{k=1}^{N_{Ch}} r_{m,k}^2 e_{k,k}}$$

$$DM_7 = \frac{\sum_{m=1}^N w(m) \max[dm_7(m), 1]}{\sum_{m=1}^N w(m)} \quad \text{con } w(m) = [r_m^2 X_m]^\alpha \text{ como antes.}$$

2.3.8 Propiedades de señales de objeto

5 A continuación se describe un ejemplo de propiedades de las señales de objeto que pueden utilizarse, por ejemplo, por el aparato 250 o el dispositivo reductor de perturbaciones 320 para obtener una medida de la distorsión.

10 En el procesamiento SAOC, se realiza la mezcla descendente de varias señales de objetos de audio para obtener una señal de mezcla descendente que después se utiliza para generar la salida presentada final. Si se mezcla una señal de objeto tonal con una segunda señal de objeto más del tipo ruido de igual potencia de señal, el resultado tiene a ser más de tipo ruido. Lo mismo ocurre si la segunda señal de objeto tiene una potencia más elevada. Únicamente, si la segunda señal de objeto tiene una potencia sustancialmente menor a la de la primera, el resultado tiende a ser tonal. Del mismo modo, la semejanza de tonalidad /ruido de la señal de salida SAOC presentada está determinada principalmente por la semejanza de tonalidad / ruido de la señal de mezcla descendente independientemente de los coeficientes de presentación aplicados. Para obtener una calidad de salida subjetiva favorable, también la semejanza de tonalidad/ruido de la señal presentada en realidad debe acercarse a la semejanza de tonalidad /ruido de la señal presentada idealmente. Para utilizar este concepto en la medida de la distorsión, es necesario transmitir la información acerca de la semejanza de tonalidad/ruido de cada objeto como parte del flujo de bits. A continuación se puede estimar la semejanza de tonalidad/ruido N de la salida presentada idealmente en el decodificador SAOC en función de la semejanza de tonalidad/ruido de cada objeto N_i y su potencia de objeto P_i , es decir

$$N = f(N_1, P_1, N_2, P_2, N_3, P_3, \dots)$$

25 y compararla con la semejanza de tonalidad/ruido de la señal de salida presentada en realidad para calcular una medida de la distorsión. Por ejemplo, se puede utilizar la siguiente función f():

$$N = \frac{\sum_i N_i \cdot P_i^\alpha}{(\sum_i P_i)^\alpha}$$

30 que combina los valores de semejanza de tonalidad/ruido del objeto y las potencias de los objetos en una sola salida estimando el valor de semejanza de tonalidad/ruido de las mezcla de las señales. Se puede elegir el parámetro α para optimizar la precisión del procedimiento de cálculo para una medida de semejanza de tonalidad/ruido (por ejemplo $\alpha=2$). Se describe una métrica de distorsión adecuada basada en la semejanza de tonalidad/ruido en la Sección 2.3.6 como medida de distorsión N.º 6.

2.4 Esquemas de limitación de distorsiones

2.4.1 Reseña general de los esquemas de limitación de distorsiones

40 A continuación se presenta una breve reseña de una pluralidad de esquemas de limitación de las distorsiones. Como se ha analizado anteriormente, el ajustador de coeficientes de presentación 250 recibe los coeficientes de presentación de entrada 242 y produce, sobre la base de éstos, un coeficiente de presentación 222 modificado para utilizar por el decodificador SAOC 220.

45 Se pueden distinguir diferentes conceptos para la provisión de coeficientes de presentación modificados, donde los conceptos también se pueden combinar en algunas realizaciones. De acuerdo con el primer concepto, se obtiene uno o más valores límite de los parámetros de presentación en un primer paso dependiendo de uno o más parámetros de información complementaria 214 (es decir, que depende de la información paramétrica relacionada

con el objeto 214). Seguidamente, se obtienen los coeficientes de presentación reales “(modificados o ajustados)” 222 que dependen del parámetro de presentación deseado 242 y del uno o más valores límite de los parámetros de presentación, de manera que los parámetros de presentación reales respeten los límites definidos por los valores límite de los parámetros de presentación. En consecuencia, dichos parámetros de presentación que exceden los valores límite de los parámetros de presentación, se ajustan (modifican) para adaptarse a los valores límite de los parámetros de presentación. Este primer concepto es fácil de implementar, aunque en ocasiones conlleva una satisfacción levemente degradada del usuario, puesto que la elección del usuario de los parámetros de presentación deseados 242 queda fuera de consideración si los parámetros de presentación deseados 242 definidos por el usuario exceden los valores límite de los parámetros de presentación.

De acuerdo con el segundo concepto, el ajustador de parámetros calcula una combinación entre el cuadrado de un parámetro de presentación deseado y el cuadrado de un parámetro de presentación óptimo, para obtener el parámetro de presentación real. En este caso, el ajustador de parámetros está configurado para determinar una contribución del parámetro de presentación deseado y del parámetro de presentación óptimo a la combinación de lineal dependiendo de un parámetro umbral predeterminado y una métrica de distorsión (descrita anteriormente).

Además, se puede distinguir si la medida de la distorsión (métrica de distorsión) se calcula utilizando las propiedades de relaciones entre objetos y/o las propiedades de los objetos individuales. En algunas realizaciones, sólo se evalúan las propiedades de la relación entre objetos, dejando a las propiedades de los objetos individuales (que se relacionan sólo con un objeto) fuera de consideración. En algunas realizaciones adicionales, solo se tienen en cuenta las propiedades de los objetos individuales, dejando las propiedades de relación entre objetos fuera de consideración. Sin embargo, en algunas realizaciones, se evalúa una combinación de propiedades de relación entre objetos y propiedades de los objetos individuales.

Basándose en las consideraciones anteriores, y también basándose en la descripción anterior de las diferentes medidas de distorsión, se pasa a definir un número de esquemas para limitar la distorsión, resumidos en las siguientes subsecciones. Estos esquemas para limitar la distorsión pueden aplicarse por el ajustador de coeficientes de presentación 250 para obtener los coeficientes de presentación modificados dependiendo de los coeficientes de presentación de entrada 242.

2.4.2 Esquema de limitación de distorsiones N.º 1

En la subsección 2.3.1 se define una sencilla medida de distorsión calculando la relación entre la contribución de potencia ideal del objeto N.º m y su contribución de potencia real (ecuación 4):

$$dm_1(m) = \frac{P_{ideal}}{P_{real}} = \frac{r_m^2}{d_m^2 \cdot t^2} = \frac{r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i}{d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i} \quad (4)$$

En esta ecuación, las únicas variables que están bajo el control del presentador SAOC son los coeficientes de presentación que se utilizan en el proceso de transcodificación. Por lo tanto, si la métrica de distorsión obtenida no excede un determinado valor umbral, T, esto impone una condición al correspondiente coeficiente de matriz de presentación:

$$dm_1(m) = \frac{r_m^2 \cdot \sum_{i=1}^N d_i^2 \cdot X_i}{d_m^2 \cdot \sum_{i=1}^N r_i^2 \cdot X_i} \leq T \Leftrightarrow r_m^2 \leq \hat{r}_m^2 = T \cdot \frac{d_m^2 \cdot \sum_{i=1, i \neq m}^N r_i^2 \cdot X_i}{\left| \sum_{i=1}^N d_i^2 \cdot X_i - T \cdot d_m^2 \cdot X_m \right|} \quad (6.1.a)$$

Para encontrar una solución para todos los \hat{r}_m^2 se puede establecer una serie de ecuaciones lineales $Ax = b$ donde

$$\mathbf{x} = \begin{bmatrix} \hat{r}_1^2 \\ \hat{r}_2^2 \\ \vdots \\ \hat{r}_N^2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \sum_{i=1}^N r_i^2 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} -c_1 & d_1^2 X_2 & \cdots & d_1^2 X_N \\ d_2^2 X_1 & -c_2 & \cdots & d_2^2 X_N \\ \vdots & \vdots & \ddots & \vdots \\ d_N^2 X_1 & d_N^2 X_2 & \cdots & -c_N \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

con

$$c_m = \frac{1}{T} \left(\sum_{i=1}^N d_i^2 \cdot X_i - T \cdot d_m^2 \cdot X_m \right)$$

5 Las primeras N filas de A se obtienen directamente de la ecuación (6.1.a). Además, se agrega una restricción de manera que la energía de los nuevos coeficientes (limitados) de presentación sea igual a la energía de los coeficientes especificados por el usuario. Se obtiene entonces una solución correspondiente a \hat{r}_m^2 (que se considera como valores límite de los parámetros de presentación) de la siguiente manera:

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$$

10 A partir de esto, se puede ver un primer esquema de limitación de las distorsiones, como sigue: en lugar de usar los coeficientes de la matriz de presentación 242 tal como se introducen al decodificador SAOC desde la interfaz del usuario, el coeficiente de presentación utilizado efectivamente r_m' , 222 para el objeto N.º m se modifica / limita (por ejemplo, por el ajustador de coeficientes de presentación 240 trama a trama antes de utilizarlo para el proceso de decodificación SAOC:

$$r_m'^2 = \min(r_m^2, \hat{r}_m^2)$$

20 Obsérvese que el proceso de limitación depende de las energías de los objetos individuales en cada trama específica. El enfoque es sencillo y tiene los siguientes inconvenientes leves:

- No tiene en cuenta el volumen del objeto ni el enmascaramiento perceptual y
- Sólo captura los efectos de incremento de ganancia de un objeto específico, pero no captura los efectos atenuando las ganancias de objeto. Esto podría solucionarse estableciendo asimismo un límite más bajo para el valor dm.

2.4.3 Esquema de limitación N.º 2

2.4.3.1 Reseña general del esquema de limitación

30 Esta sección describe una función de limitación que tiene en cuenta los siguientes aspectos:

- la medida de la distorsión está restringida por un umbral limitante,
- la derivación de la matriz de presentación limitada se basa en la función de limitación y en su distancia de la matriz de presentación inicial.

40 Esta función de limitación (o esquema de limitación) puede ejecutarse, por ejemplo, por el ajustador de coeficientes de presentación 250 en combinación con el calculador de distorsiones 260.

La medida de la distorsión es una función de la matriz de presentación, por lo que

- una matriz de presentación inicial (descrita, por ejemplo, por los coeficientes de presentación de entrada 242) da una medida de la distorsión inicial,
- la medida de distorsión óptima produce una matriz de presentación óptima, pero la distancia de esta matriz de presentación óptima a la matriz de presentación inicial puede no ser óptima,
- la medida de la distorsión es inversamente proporcional lineal a la distancia de una matriz de presentación con la matriz de presentación inicial,

- para un determinado umbral, se deriva la matriz de presentación limitada (descrita, por ejemplo, por los coeficientes de presentación ajustados o modificados 222) mediante interpolación (por ejemplo, interpolación lineal) entre el punto de trabajo inicial y el óptimo.

Además, se puede suponer que la potencia de la señal presentada en cada punto de trabajo es aproximadamente constante, por lo que

$$\sum_{i=1}^{N_{ob}} r_i^2 X_i \approx \sum_{i=1}^{N_{ob}} r_{lim,i}^2 X_i \approx \sum_{i=1}^{N_{ob}} r_{opt,i}^2 X_i$$

Se puede utilizar el esquema de limitación N.º 2 en combinación con diferentes medidas de distorsión, como se describe a continuación.

2.4.3.2 Limitación de la medida de distorsión N.º 1

Por cada banda de parámetros, la medida de la distorsión $dm_1(m)$ correspondiente a un objeto de interés m se define como

$$dm_1(m) = \frac{r_m^2 \sum_{i=1}^{N_{ob}} d_i^2 X_i}{d_m^2 \sum_{i=1}^{N_{ob}} r_i^2 X_i}$$

La matriz de presentación óptima se produce cuando se ajusta $dm_1(m)$ en su valor óptimo, es decir $dm_{1,opt}(m) = 1$

$$r_{opt,m}^2 = d_m^2 \frac{\sum_{i=1}^{N_{ob}} r_i^2 X_i}{\sum_{i=1}^{N_{ob}} d_i^2 X_i}$$

En consecuencia, se pueden obtener los valores óptimos de la matriz de presentación $r_{opt,m}^2$ utilizando un sistema de ecuaciones, donde r_i^2 se reemplaza por $r_{opt,i}^2$.

Dentro del umbral predefinido T correspondiente a $dm_1(m)$ la matriz de presentación limitada está dada por

$$r_{lim,m}^2 = \frac{T-1}{dm_1(m)} (r_m^2 - r_{opt,m}^2) + r_{opt,m}^2$$

2.4.3.3 Limitación de la medida de distorsión N.º 2a

La medida de distorsión $dm_{2a}(m)$, que en ocasiones se designa también, brevemente, " $dm_2(m)$ ", se define como

$$dm_{2a}(m) = \frac{\left(r_m^2 \sum_{i=1}^{N_{ob}} d_i^2 X_i - d_m^2 \sum_{i=1}^{N_{ob}} r_i^2 X_i \right) X_m}{msr \sum_{i=1}^{N_{ob}} r_i^2 X_i \sum_{i=1}^{N_{ob}} d_i^2 X_i} = \frac{\frac{r_m^2 X_m}{\sum_{i=1}^{N_{ob}} r_i^2 X_i} - \frac{d_m^2 X_m}{\sum_{i=1}^{N_{ob}} d_i^2 X_i}}{msr}$$

en el caso del objeto m y cada banda de parámetros. Con respecto a ciertas bandas de parámetros pb la relación

máscara a señal $msr(pb)$ es una función de la potencia de la señal presentada

$$msr(pb) = \left[\sum_{i=1}^{N_{ob}} r_i^2 X_i M_k \right]_{k=\max(pb)} = \left[\sum_{i=1}^{N_{ob}} r_i^2 X_i \right]_{k=\max(pb)} [M_k]_{k=\max(pb)}$$

- 5 El valor óptimo de la medida de la distorsión es cero, es decir $dm_{2a,opt}(m) = 0$. Esto corresponde a un proceso de transcodificación perfecto que no introduce error alguno. Por lo tanto, la matriz de presentación óptima produce

$$r_{opt,m}^2 = d_m^2 \frac{\sum_{i=1}^{N_{ob}} r_i^2 X_i}{\sum_{i=1}^{N_{ob}} d_i^2 X_i}$$

- 10 donde $dm_{2a}(m) = T$ la matriz de la matriz de presentación limitada, que puede describirse por los coeficientes de presentación modificados 222, se convierte en

$$r_{lim,m}^2 = \frac{T-1}{dm_{2a}(m)} (r_m^2 - r_{opt,m}^2) + r_{opt,m}^2$$

15 2.4.3.4 Limitación de la medida de distorsión N.º 2b

La medida de la distorsión $dm_{2b}(m)$, la que en ocasiones también se denomina en forma breve $dm_2(m)$, también puede utilizarse por el aparato 240 para obtener la matriz de presentación limitada, que puede describirse por los coeficientes de presentación modificados 222, que dependen de los coeficientes de presentación de entrada 242.

20 2.4.3.5 Limitación de la medida de distorsión N.º 4

La medida de la distorsión $dm_4(m)$ se define de la siguiente manera:

$$dm_4(m) = \left| 1 - \frac{r_m^2 \sum_{i=1}^{N_{ob}} d_i^2 X_i}{d_m^2 \sum_{i=1}^{N_{ob}} r_i^2 X_i} \right|$$

- 25 en el caso del objeto m y cada banda de parámetros y su valor óptimo es $dm_{4,opt}(m) = 0$. En consecuencia, las matrices de presentación óptima y limitada dan como resultado

$$r_{opt,m}^2 = d_m^2 \frac{\sum_{i=1}^{N_{ob}} r_i^2 X_i}{\sum_{i=1}^{N_{ob}} d_i^2 X_i}$$

y

$$r_{lim,m}^2 = \frac{T-1}{dm_4(m)} (r_m^2 - r_{opt,m}^2) + r_{opt,m}^2$$

- 30 En consecuencia, el aparato 240 puede suministrar los coeficientes de presentación modificados 222 dependiendo de los coeficientes de presentación de entrada 242 y dependiendo asimismo de la medida de la distorsión 252, que

puede ser igual a la cuarta medida de distorsión $dm_4(m)$.

2.4.4 Esquema de limitación N.º 3

5 Con respecto a la fórmula (6.1.a) se puede calcular el coeficiente de presentación limitado correspondiente al objeto m para la medida de la distorsión N.º 3 de la siguiente manera. Con las abreviaturas

$$c_1 = \sum_{i=1}^N \sum_{j=1}^N d_i d_j e_{ij} \quad , \quad c_2 = \sum_{i=1, i \neq m}^N r_i e_{im} \quad , \quad c_3 = \sum_{i=1, i \neq m}^N \sum_{j=1, j \neq m}^N r_i r_j e_{ij} \quad , \quad c_4 = \sum_{i=1}^N d_i e_{mi}$$

10 y

$$c_5 = \sum_{i=1, i \neq m}^N \sum_{j=1}^N r_i d_j e_{ij}$$

se establece una ecuación cuadrática

$$15 \quad \hat{r}_m^2 \left((1-T)^2 \cdot c_1 e_{mm} - c_4^2 \right) + \hat{r}_m \cdot 2 \cdot \left((1-T)^2 \cdot c_1 c_2 - c_4 c_5 \right) + (1-T)^2 \cdot c_1 c_3 - c_5^2 = a \cdot \hat{r}_m^2 + b \cdot \hat{r}_m + c = 0$$

cuya solución (positiva) es

$$\hat{r}_m = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad (6.2.a)$$

20 En consecuencia, el aparato 240 puede comprender valores límite de los parámetros de presentación \hat{r}_m , y puede limitar los coeficientes de presentación ajustados (o modificados) 222 de acuerdo con dichos valores límite de los parámetros de presentación.

2.4.5 Otras mejoras opcionales

25 El concepto antes descrito para limitar los coeficientes de presentación 222, que se ejecuta, en forma individual o en combinación, por el aparato 240, puede mejorarse aún. Por ejemplo, se puede ejecutar una generalización para la presentación de M canales. Para este fin, se puede utilizar la suma de los cuadrados/potencia de los coeficientes de presentación en lugar de un solo coeficiente de presentación.

30 Además, se puede ejecutar una generalización a una mezcla descendente estéreo. Para este fin, se puede utilizar una suma de los cuadrados/potencias de los coeficientes de mezcla descendente en lugar de un solo coeficiente de mezcla descendente.

35 En algunas realizaciones, se puede combinar la mezcla de distorsión a través de la frecuencia para obtener una sola que se utiliza para controlar la degradación. Por otro lado, puede ser mejor (y más sencillo) en algunos casos, realizar el control de la distorsión independientemente por cada banda de frecuencia.

40 Se pueden aplicar diferentes conceptos para efectuar realmente el control de la distorsión. Por ejemplo, se puede limitar dichos uno o más coeficientes de presentación. Por otro lado, o además, se puede limitar un coeficiente de matriz m_2 (por ejemplo de una decodificación de MPEG Envoltente). Por otro lado, o además, se puede limitar una ganancia relativa del objeto.

3. Realización de acuerdo con la Fig. 3

45 A continuación se describe otra realización de un decodificador SAOC tomando como referencia la Fig. 3. Para facilitar la comprensión, en primer lugar se presenta una breve explicación de las consideraciones subyacentes. La salida de un sistema de "codificación espacial de objetos de audio" (SAOC) (como el que se ha normalizado como ISO/IEC 23003-2) puede mostrar perturbaciones que dependen de las propiedades del objeto de audio y la relación entre la matriz de presentación y la matriz de mezcla descendente. Para explicar este problema, se considera aquí el caso en que las matrices de mezcla descendente y presentación tienen la misma dimensión sin pérdida de generalidad. Se aplican consideraciones correspondientes si el número de canales de la escena de mezcla descendente y presentada son diferentes.

Se ha descubierto que, en general, el riesgo de perturbaciones se incrementa cuando la matriz de presentación se vuelve significativamente diferente de la matriz de mezcla descendente. Se pueden distinguir diferentes tipos de perturbaciones:

- 5 1. Imperfecciones de la presentación, es decir que la matriz de presentación "efectiva" difiere de la matriz de presentación deseada que se introduce como entrada al decodificador SAOC (la atenuación lograda en efecto o la ganancia de un objeto es diferente de la especificada en la matriz de presentación). Este es típicamente el efecto de la superposición de objetos en ciertas bandas de parámetros.
- 10 2. Cambios desfavorables y, posiblemente, temporalmente variantes del timbre de un objeto. Esta perturbación es especialmente grave cuando la "fuga" mencionada en 1. sólo tiene lugar localmente en caso de una sola banda de parámetros.
- 15 3. Perturbaciones, como señales de objeto moduladas, tonos musicales o ruido modulado, causadas por el procesamiento de señales de tiempo y frecuencia variantes en el decodificador SAOC.

Se ha descubierto que es ventajoso minimizar todo tipo de perturbaciones.

20 Una estrategia generalizada para abordar este problema y minimizar las perturbaciones consiste en emplear un post-procesamiento variante en tiempo-frecuencia de la matriz de presentación deseada antes de enviarla al decodificador SAOC. Esta estrategia está expuesta en la Fig. 3.

25 La Fig. 3 ilustra un diagrama esquemático de bloques de una disposición de decodificador SAOC 300. El decodificador SAOC 300 puede designarse, en pocas palabras, como un decodificador de señales de audio. El decodificador de señales de audio 300 comprende un núcleo de decodificador SAOC 310, que está configurado para recibir una representación de señal de mezcla descendente 312 y un flujo de bits SAOC 314 y para producir, sobre la base de estos, una descripción 316 de una escena presentada, por ejemplo, en forma de representación de una pluralidad de canales de audio de mezcla ascendente.

30 El decodificador de señales de audio 300 comprende asimismo un dispositivo reductor de perturbaciones 320, que se puede presentar, por ejemplo, en forma de aparato para suministrar uno o más parámetros ajustados dependiendo de uno o más parámetros de entrada. El dispositivo reductor de perturbaciones 320 está configurado para recibir información 322 acerca de una matriz de presentación deseada. La información 322 puede tomar la forma, por ejemplo, de una pluralidad de parámetros de presentación deseados, que pueden constituir los parámetros de entrada del dispositivo reductor de perturbaciones. El dispositivo reductor de perturbaciones 320 está configurado además para recibir la representación de la señal de mezcla descendente 312 y el flujo de bits SAOC 314, donde el flujo de bits SAOC 314 puede llevar información paramétrica relacionada con el objeto. El dispositivo reductor de perturbaciones 320 está configurado además para producir una matriz de presentación modificada 324 (por ejemplo, en forma de una pluralidad de parámetros de presentación ajustados) dependiendo de la información 322 acerca de la matriz de presentación deseada.

45 En consecuencia, el núcleo del decodificador SAOC 310 puede estar configurado para transmitir la representación 316 de la escena presentada dependiendo de la representación de la señal de mezcla descendente 312, el flujo de bits SAOC 314 y la matriz de presentación modificada 324.

50 A continuación se presentan algunos detalles con respecto a la funcionalidad del decodificador de señales de audio. Se ha descubierto que, para evaluar el riesgo de perturbaciones debidas a las capacidades de separación potencialmente limitadas del sistema SAOC con respecto a una matriz de presentación deseada, es conveniente tener en cuenta tanto la señal de mezcla descendente (descrita por la representación de la señal de mezcla descendente 312) y el flujo de bits SAOC 314. Contando con esta información, es posible intentar mitigar estas perturbaciones, por ejemplo, mediante la modificación de la matriz de presentación. Esto se ejecuta por el dispositivo reductor de perturbaciones 320. Las estrategias de mitigación avanzadas tienen en cuenta tanto las limitaciones (superposición) de la selectividad en tiempo y frecuencia del sistema SAOC como los efectos perceptuales, es decir, que deberían tratar de lograr que la señal presentada sea lo más similar posible a la señal de salida deseada y a la vez tener la menor cantidad posible de perturbaciones audibles.

60 Un enfoque preferido para la reducción de perturbaciones, que se utiliza en el decodificador de señales de audio 300 expuesto en la Fig. 3, se basa en una medida de distorsión global que es una combinación ponderada de las medidas de distorsión que evalúan los diferentes tipos de perturbaciones antes enumeradas. Estas ponderaciones determinan una relación adecuada entre los diferentes tipos de perturbaciones antes enumeradas. Cabe señalar que las ponderaciones para estos tipos de perturbaciones pueden depender de la aplicación en la cual se utiliza el sistema SAOC.

En otras palabras, el dispositivo reductor de perturbaciones 320 puede estar configurado para obtener medidas de

distorsión correspondientes a una pluralidad de tipos de perturbaciones. Por ejemplo, el dispositivo reductor de perturbaciones 320 puede aplicar algunas de las medidas de la distorsión dm_1 a dm_6 descritas anteriormente. Por otro lado, o además, el dispositivo reductor de perturbaciones 320 puede utilizar otras medidas de distorsión que describen otros tipos de perturbaciones, como se describe dentro de esta sección. Asimismo, el dispositivo reductor de perturbaciones puede estar configurado para obtener la matriz de presentación modificada 324 sobre la base de la matriz de presentación deseada 322 empleando uno o más de los esquemas de limitación de distorsiones que ya ha sido descrito en párrafos anteriores (por ejemplo, en las secciones 2.4.2, 2.4.3 y 2.4.4), o esquemas de limitación de perturbaciones comparables.

4. Transcodificadores de señales de audio de acuerdo con las Figs. 5a y 5b

4.1 Transcodificador de señales de audio de acuerdo con la Fig. 5a

Debe indicarse que los conceptos antes descritos se pueden aplicar tanto a un decodificador de señales de audio como un transcodificador de señales de audio. Tomando como referencia las Figs. 2 y 3, se ha descrito el concepto en combinación con decodificadores de audio. A continuación se describe en forma breve el concepto de la presente invención en combinación con transcodificadores de señales de audio.

Con respecto a este tema, se debe tener en cuenta que ya se han descrito las similitudes entre los decodificadores de señales de audio y los transcodificadores de señales de audio con referencia a las Figs. 9a, 9b y 9c, razón por la cual las explicaciones presentadas con respecto a las Figs. 9a, 9b y 9c son aplicables al concepto de la invención.

La Fig. 5a ilustra un diagrama esquemático de bloques de un transcodificador de señales de audio 500 en combinación con un decodificador MPEG Envoltante 510. Como se puede apreciar, el transcodificador de señales de audio 500, que puede consistir en un transcodificador de SAOC a MPEG Envoltante, está configurado para recibir un flujo de bits SAOC 520 y para producir, sobre la base de éste, un flujo de bits MPEG Envoltante 522 sin afectar (ni modificar) la representación de una señal de mezcla descendente 524. El transcodificador de señales de audio 500 comprende un analizador SAOC 530, que está configurado para recibir el flujo de bits SAOC 520 y para extraer los parámetros SAOC deseados del flujo de bits SAOC 530. El transcodificador de señales de audio 500 comprende además un motor de presentación de escenas 540, que está configurado para recibir los parámetros SAOC proporcionados por el analizador SAOC 530 e información de la matriz de presentación 542, que puede considerarse información de (matriz) presentación real, y que puede estar representada, por ejemplo, en forma de una pluralidad de parámetros de presentación ajustados (o modificados). El motor de presentación de escenas 540 está configurado para transmitir el flujo de bits de MPEG Envoltante 522 que depende de dichos parámetros SAOC y la matriz de presentación 542. Para este fin, el motor de presentación de escenas 540 está configurado para calcular los parámetros del flujo de bits MPEG Envoltante 522, que son parámetros relativos a los canales (también designado como información paramétrica). Por consiguiente, el motor de presentación de escenas 540 está configurado para transformar (o "transcodificar") los parámetros del flujo de bits SAOC 520, lo que constituye información paramétrica relacionada con el objeto, en los parámetros del flujo de bits MPEG Envoltante, lo que constituye una información paramétrica relacionada con los canales, que depende de la matriz de presentación 542.

El transcodificador de señales de audio 500 comprende además la generación de una matriz de presentación 550, que está configurada para recibir una información acerca de una matriz de presentación deseada, por ejemplo, en forma de información 552 acerca de una configuración de reproducción y una información 554 acerca de posiciones de objetos. Por otro lado, la generación de matrices de presentación 550 puede recibir información acerca de los parámetros de presentación (por ejemplo, entradas a la matriz de presentación). El generador de matrices de presentación también está configurado para recibir el flujo de bits SAOC 520 (o, por lo menos, una subserie de información paramétrica relacionada con el objeto representado por el flujo de bits SAOC 520). El generador de matrices de presentación 550 también está configurado para producir la matriz de presentación real (ajustada o modificada) 542 sobre la base de la información recibida. En tal medida, el generador de matrices de presentación 550 puede asumir la funcionalidad del aparato 100 o del aparato 240.

El decodificador MPEG Envoltante 510 está configurado típicamente para obtener una pluralidad de señales de canales de mezcla ascendente sobre la base de la información de señal de mezcla descendente 524 y el flujo de bits MPEG Envoltante 522 proporcionado por el motor de presentación de escenas 540.

Para resumir, el transcodificador de señales de audio 500 está configurado para producir el flujo de bits MPEG Envoltante 522 de tal manera que el flujo de bits MPEG Envoltante 522 da lugar a la producción de una representación de señal de mezcla ascendente sobre la base de la representación de la señal de mezcla descendente 524, donde la representación de señal de mezcla ascendente proporcionada en realidad por el decodificador MPEG Envoltante 510. El generador de matrices de presentación 550 ajusta la matriz de presentación 542 utilizada por el motor de presentación de escenas 540 de manera que la representación de señal de mezcla ascendente generada por el decodificador MPEG Envoltante 510 no comprenda una distorsión audible inadmisibles.

4.2 Transcodificador de señales de audio de acuerdo con la Fig. 5b

La Fig. 5b ilustra otra disposición de un transcodificador de señales de audio 560 y un decodificador MPEG Envolvente 510. Cabe señalar que la disposición de la Fig. 5b es muy similar a la disposición de la Fig. 5a, razón por la cual se designan los medios y señales idénticos con idénticos números de referencia. El transcodificador de señales de audio 560 difiere del transcodificador de señales de audio 500 en que el transcodificador de señales de audio 560 comprende un transcodificador de mezcla descendente 570, que está configurado para recibir la representación de mezcla descendente de entrada 524 y para producir una representación de mezcla descendente modificada 574, que es alimentada al decodificador MPEG Envolvente 510. La modificación de la representación de la señal de mezcla descendente se realiza para obtener más flexibilidad en la definición del resultado de audio buscado. Esto se debe a que el flujo de bits de MPEG Envolvente 522 no puede representar ciertos mapeos de la señal de entrada del decodificador MPEG Envolvente 510 sobre las señales de canales de mezcla ascendente producidas como salida por el decodificador MPEG Envolvente 510. En consecuencia, la modificación de la representación de la señal de mezcla descendente utilizando el transcodificador de mezcla descendente 570 puede traer aparejada un aumento de la flexibilidad.

Una vez más, el generador de matrices de presentación 550 puede asumir la funcionalidad del aparato 100 o el aparato 240, garantizando así que las distorsiones audibles en la representación de la señal de mezcla ascendente proporcionada por el decodificador MPEG Envolvente 510 se mantengan suficientemente bajas.

5. Codificador de Señales de Audio de acuerdo con la Fig. 6

A continuación se describe un codificador de señales de audio 600 tomando como referencia la Fig. 6, que ilustra un diagrama esquemático de bloques de ese tipo de codificador de señales de audio. El codificador de señales de audio 600 está configurado para recibir una pluralidad de señales de objeto 612a, 612N (también designadas x_1 a x_N) y para producir, sobre la base de éstas, una representación de señal de mezcla descendente 614 e información paramétrica relacionada con el objeto 616. El codificador de señales de audio 600 comprende un dispositivo de mezcla descendente 620 configurado para producir una o más señales de mezcla descendente (que constituyen la representación de la señal de mezcla descendente 614) dependiendo de los coeficientes de mezcla descendente d_1 a d_N asociados a las señales de objeto, de tal manera que dichas una o más señales de mezcla descendente comprendan una superposición de una pluralidad de señales de objeto. El codificador de señales de audio 600 comprende además un proveedor de información complementaria 630, que está configurado para producir información complementaria de relación entre objetos que describe las diferencias de nivel y características de correlación de dos o más señales de objeto 612a a 612N. El proveedor de información complementaria 630 también está configurado para suministrar una información complementaria sobre objetos individuales que describe una o más propiedades individuales de las señales de objeto individuales. El codificador de señales de audio 600 produce entonces la información paramétrica relacionada con el objeto 616 por lo que la información paramétrica relacionada con el objeto comprende tanto la información complementaria de relación entre objetos como información complementaria de objetos individuales.

Se ha descubierto que dicha información paramétrica relacionada con el objeto, que describe tanto una relación entre señales de objeto como las características individuales de señales de objetos individuales permite la provisión de una señal de audio multicanal en un decodificador de señales de audio, como se ha analizado anteriormente. La información complementaria de relación entre objetos puede aprovecharse por el decodificador de señales de audio para recibir la información paramétrica relacionada con el objeto 616 a fin de extraer, por lo menos aproximadamente, señales de objetos individuales de la representación de la señal de mezcla descendente. La información complementaria sobre objetos individuales, que también está incluida en la información paramétrica relacionada con el objeto 614, puede utilizarse por el decodificador de señales de audio para verificar si el proceso de mezcla ascendente conlleva distorsiones de señal demasiado fuertes, por lo que es necesario ajustar los parámetros de mezcla ascendente (por ejemplo, los parámetros de presentación).

Preferentemente, el proveedor de información complementaria 630 está configurado para suministrar la información complementaria sobre objetos individuales, por lo que la información complementaria sobre objetos individuales describe una tonalidad de las señales de objetos individuales. Se ha descubierto que se puede emplear una información de tonalidad como criterio seguro para evaluar si el proceso de mezcla ascendente trae consigo, o no, distorsiones significativas.

Debe indicarse asimismo que el codificador de señales de audio 600 puede complementarse por medio de cualquiera de las características y funcionalidades descritas en este documento con respecto a los codificadores de señales de audio, y que la representación de la señal de mezcla descendente 614 y la información paramétrica relacionada con el objeto 616 pueden proporcionarse por el codificador de señales de audio 600 de tal manera que comprendan las características descritas con respecto al decodificador de señales de audio de la invención.

6. Flujo de bits de audio de acuerdo con la Fig. 7

Una realización de acuerdo con la invención da origen a un flujo de bits de audio 700, una representación esquemática del cual está expuesta en la Fig. 7. El flujo de bits representa una pluralidad de señales de objeto en forma codificada.

5 El flujo de bits 700 comprende una representación de señal de mezcla descendente 710 que representa una o más señales de mezcla descendente, donde por lo menos una de las señales de mezcla descendente comprende una superposición de una pluralidad de señales de objeto. El flujo de bits de audio 700 comprende además una información complementaria de relación entre objetos 720 que describe diferencias de nivel y características de correlación de
10 señales de objeto. El flujo de bits de audio comprende asimismo una información complementaria sobre objetos individuales 730 que describe una o más propiedades individuales de las señales de objetos individuales (que constituyen la base de la representación de la señal de mezcla descendente 710).

15 La información complementaria de relación entre objetos y la información de objetos individuales pueden considerarse, en su conjunto, como información complementaria paramétrica relacionada con el objeto.

De acuerdo con la invención, la información complementaria sobre objetos individuales describe tonalidades de las señales de objetos individuales.

20 Naturalmente, como el flujo de bits de audio 700 se proporciona por lo general por un codificador de señales de audio de acuerdo con lo descrito en este documento, y se evalúa por un decodificador de señales de audio, que se describe en el presente documento. El flujo de bits de audio puede comprender las características descritas con respecto al codificador de señales de audio y el decodificador de señales de audio. En consecuencia, el flujo de bits de audio 700 puede ser muy adecuado para la producción de una señal de audio multicanal utilizando un
25 decodificador de señales de audio, como se describe aquí.

7. Conclusión

30 Las realizaciones de acuerdo con la invención ofrecen soluciones para reducir o evitar el problema de distorsiones antes explicado, que se originan por el hecho de que no se pueden reconstruir las señales de objetos únicos originales perfectamente a partir de las pocas señales de mezcla descendente transmitidas. Hay otras soluciones sencillas que se aplican para solucionar este problema:

35 • Una estrategia sencilla consistiría en limitar el rango de ganancia relativa de los objetos a, por ejemplo +/-12 dB. Aunque es cierto que la configuración de grandes ganancias de objeto pueden llevar a degradaciones audibles (por ejemplo: incrementar la ganancia un objeto en 20 dB dejando, al mismo tiempo, los demás niveles de objeto a 0 dB): esto no es necesario, sin embargo. Por ejemplo, si se incrementa la ganancia de todos los niveles relativos de los objetos en el mismo factor se produce una salida del sistema sin alteraciones.

40 • Un enfoque más elaborado consistiría en considerar las diferencias de los niveles relativos de los objetos. Para la presentación de dos objetos de audio, la diferencia de los niveles relativos de ambos objetos ofrece un gancho para las posibles degradaciones de la salida presentada. Sin embargo, no está claro cómo se generaliza esta idea a más de dos objetos de audio presentados.

45 En vista de esta situación, las realizaciones de acuerdo con la presente invención ofrecen medios para abordar este problema y, de esa manera, prevenir una experiencia desfavorable del usuario. Algunas realizaciones de acuerdo con la presente invención pueden dar lugar a soluciones aún más elaboradas que las descritas en la sección anterior.

50 En consecuencia, se puede obtener una buena impresión auditiva utilizando la presente invención, aun en el caso en que un usuario proporcione parámetros de presentación inapropiados.

55 En términos generales, las realizaciones de acuerdo con la invención se relacionan con un aparato, un método o un programa informático para codificar una señal de audio o para decodificar una señal de audio codificada, o con una señal de audio codificada (por ejemplo, en forma de flujo de bits de audio) de acuerdo con lo descrito anteriormente.

8. Alternativas de Implementación

60 Si bien se han descrito ciertos aspectos en el contexto de un aparato, es claro que estos aspectos representan asimismo una descripción del método correspondiente, donde un bloque o dispositivo corresponde a un paso del método o una característica de un paso del método. De manera análoga, los aspectos descritos en el contexto de un paso del método también representan una descripción de un bloque o elemento o característica correspondiente de un aparato respectivo. Algunos o todos los pasos del método pueden ejecutarse (o utilizarse) por un aparato de hardware, como por ejemplo un microprocesador, un ordenador programable o un circuito electrónico. En algunas

realizaciones, uno o más de los pasos más importantes del método pueden ejecutarse por ese aparato.

La señal de audio codificada o el flujo de bits de audio de la invención pueden almacenarse en un medio de almacenamiento digital o pueden transmitirse en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión alámbrico tal como internet.

Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación se puede ejecutar 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 FLASH, con señales de control legibles electrónicamente almacenadas en el mismo, que cooperan (o pueden cooperar) con un sistema informático programable a fin de ejecutar el método respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

Algunas realizaciones de acuerdo con la invención comprenden un portador de datos que cuenta con señales de control legibles electrónicamente, con capacidad para cooperar con un sistema informático programable para la ejecución de los métodos aquí descritos.

En general, las realizaciones de la presente invención pueden implementarse en forma de producto programa informático con un código de programa, donde el código de programa es operativo para realizar uno de los métodos al ejecutarse el programa informático en un ordenador. El código de programa puede almacenarse, por ejemplo, en un portador legible por una máquina.

Otras realizaciones comprenden el programa informático para ejecutar uno de los métodos aquí descritos, almacenados en un portador legible por una máquina.

En otras palabras, una realización del método de la invención consiste, por lo tanto, en un programa informático que consta de un código de programa para realizar uno de los métodos aquí descritos al ejecutarse el programa informático en un ordenador.

Otra realización del método de la invención consiste, por lo tanto, en un portador de datos (o un medio de almacenamiento digital, o un medio legible por computadora) que comprende, grabado en el mismo, el programa informático para ejecutar uno de los métodos descritos en la presente.

Otra realización del método de la invención consiste, por lo tanto, en un flujo de datos o una secuencia de señales que representan el programa informático para ejecutar uno de los métodos aquí descritos. El flujo de datos o la secuencia de señales pueden estar configurados, por ejemplo, para transferirse a través de una conexión de comunicación de datos, por ejemplo a través de Internet.

Otra realización comprende un medio de procesamiento, por ejemplo un ordenador o un dispositivo lógico programable, configurado o adaptado para ejecutar uno de los métodos aquí descritos.

Otra realización comprende un ordenador en la cual se ha instalado el programa informático para ejecutar uno de los métodos aquí descritos.

En algunas realizaciones, se puede utilizar un dispositivo lógico programable (por ejemplo, un campo de matrices de puertas programables) para ejecutar algunas o todas las funcionalidades de los métodos aquí descritos. En algunas realizaciones, un campo de matrices de puertas programables puede cooperar con un microprocesador para ejecutar uno de los métodos aquí descritos. En general, los métodos se ejecutan preferentemente mediante cualquier aparato de hardware.

Las realizaciones anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y detalles descritos en este documento han de ser evidentes para los expertos en la materia. Por lo tanto, sólo se pretende limitarse al alcance de las reivindicaciones de patente siguientes y no a los detalles específicos presentados a título de descripción y explicaciones de las presentes realizaciones.

Referencias

[BCC] C. Faller y F. Baumgarte, "Binaural Cue Coding - Part II: Schemes and applications", IEEE Trans. on Speech and Audio Proc., vol. 11, n.º 6, noviembre de 2003

[JSC] C. Faller, "Parametric Joint-Coding of Audio Sources", 120a Convención de AES, París, 2006, preimpresión 6752

[SAOC1] J. Herre, S. Disch, J. Hilpert, O. Hellmuth: "From SAC To SAOC - Recent Developments in Parametric

Coding of Spatial Audio”, 22ª Conferencia Regional en el RU de AES, Cambridge, RU, abril de 2007

5 [SAOC2] J. Engdegård, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Hölzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers y W. Oomen: “Spatial Objeto de audio Coding (SAOC) - The Upcoming MPEG Standard on Parametric Object Based Audio Coding”, 124ª Convención de AES, Ámsterdam 2008, Preimpresión 7377

REIVINDICACIONES

1. Un codificador de señales de audio (600) para proporcionar una representación de señal de mezcla descendente (614) e información paramétrica relacionada con el objeto (616) sobre la base de una pluralidad de señales de objeto (5 x_1 a x_N), en el que el codificador de audio comprende:
- un dispositivo de mezcla descendente (620) configurado para proporcionar una o más señales de mezcla descendente dependiendo de coeficientes de mezcla descendente (d_1 a d_N) asociados a las señales de objeto (x_1 a x_N), por lo que dichas una o más señales de mezcla descendente comprenden una superposición de una pluralidad de señales de objeto;
- un proveedor de información complementaria (630) configurado para proporcionar una información complementaria de relación entre objetos (OLD, IOC) que describe las diferencias de nivel y las características de correlación de las señales de objeto (x_1 a x_N) y una información complementaria sobre objetos individuales que describe una o más propiedades individuales de las señales de objeto individuales (x_1 a x_N), **caracterizado por que** la información complementaria sobre objetos individuales comprende una información de tonalidad de señal de objeto (N_i) que describe tonalidades de las señales de objetos individuales.
2. El codificador de señal de audio de acuerdo con la reivindicación 1, en el que el codificador de señal de audio está configurado para transmitir la información de tonalidad de señal de objeto con una resolución de frecuencia mucho más basta que otros parámetros de objetos.
3. El codificador de señal de audio de acuerdo con la reivindicación 1 o 2, en el que el codificador de señal de audio está configurado para transmitir la información de tonalidad de señal de objeto con solo una información por objeto.
4. Un método para proporcionar una representación de señal de mezcla descendente e información paramétrica relacionada con el objeto sobre la base de una pluralidad de señales de objeto, en el que las señales de objeto son señales de objeto de audio, comprendiendo el método:
- proporcionar una o más señales de mezcla descendente dependiendo de coeficientes de mezcla descendente asociados a las señales de objeto, de tal manera que dichas una o más señales de mezcla descendente comprendan una superposición de una pluralidad de señales de objeto; y
- proporcionar una información complementaria de relación entre objetos que describe las diferencias de nivel y características de correlación de las señales de objeto; y
- proporcionar información complementaria sobre objetos individuales que describe una o más propiedades individuales de las señales de objetos individuales, **caracterizado por que** la información complementaria sobre objetos individuales comprende una información de tonalidad de señal de objeto (N_i) que describe tonalidades de las señales de objetos individuales.
5. Un flujo de bits de audio (700) que representa una pluralidad de señales de objetos (x_1 a x_N) en forma codificada, comprendiendo el flujo de bits de audio:
- una representación de señal de mezcla descendente (710) que representa una o más señales de mezcla descendente, en el que al menos una de las señales de mezcla descendente comprende una superposición de una pluralidad de señales de objetos; y
- una información complementaria de relación entre objetos (720) que describe las diferencias de nivel y características de correlación de las señales de objeto; y
- una información complementaria sobre objetos individuales (730) que describe una o más propiedades individuales de las señales de objetos individuales;
- caracterizado por que** la información complementaria sobre objetos individuales comprende una información de tonalidad de señal de objeto (N_i) que describe tonalidades de las señales de objetos individuales.
6. Un programa informático adaptado para realizar el método de acuerdo con la reivindicación 4.

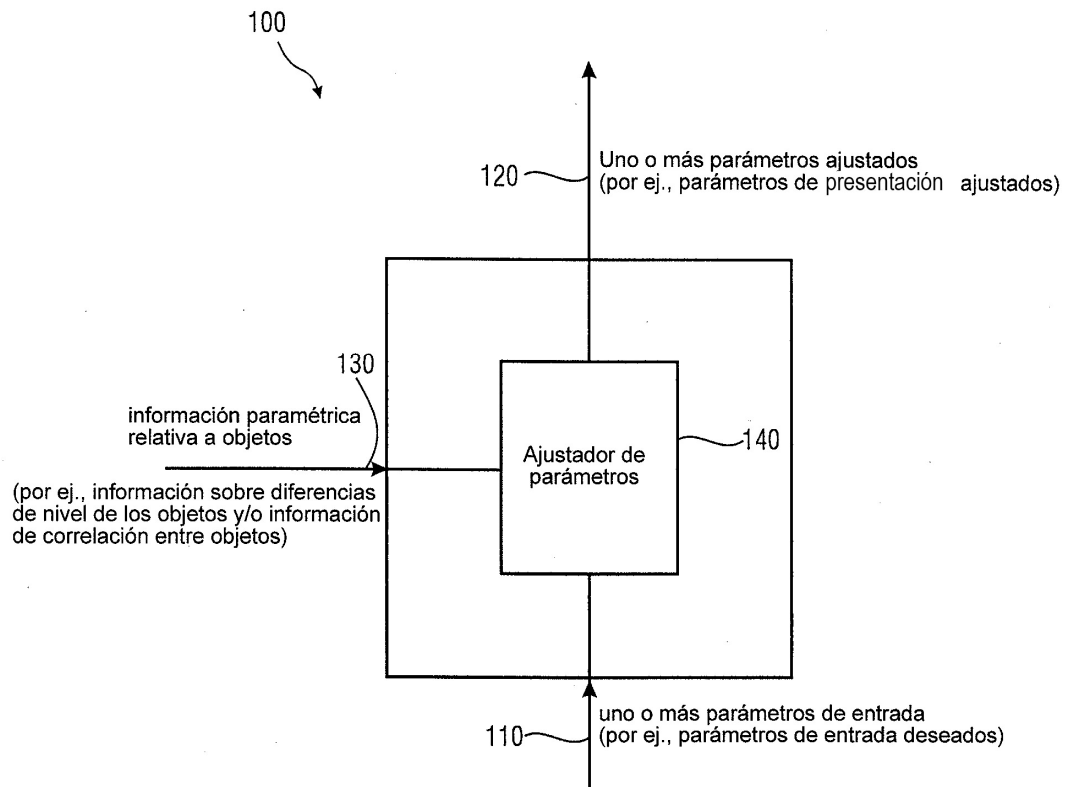


FIG. 1

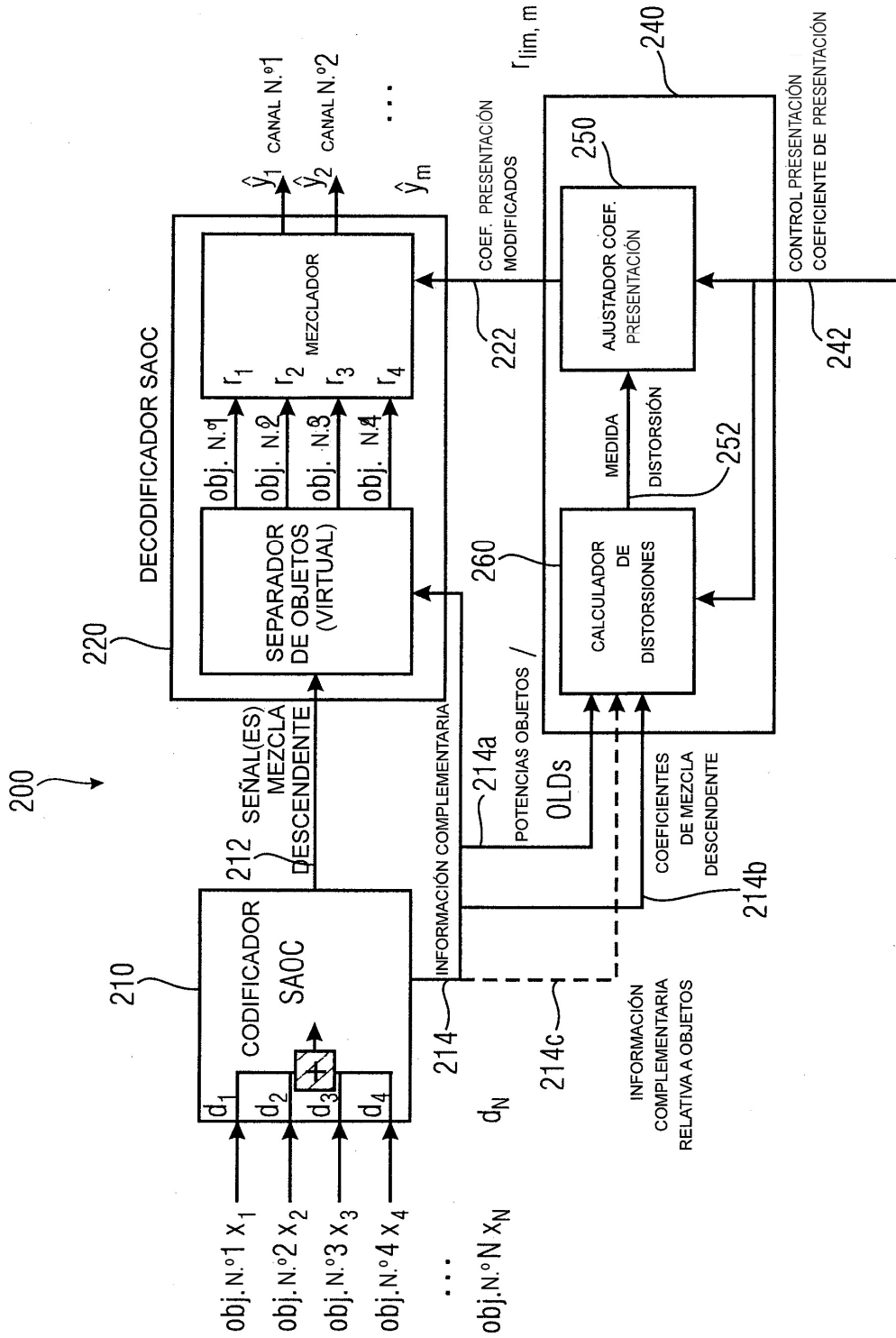


FIG. 2

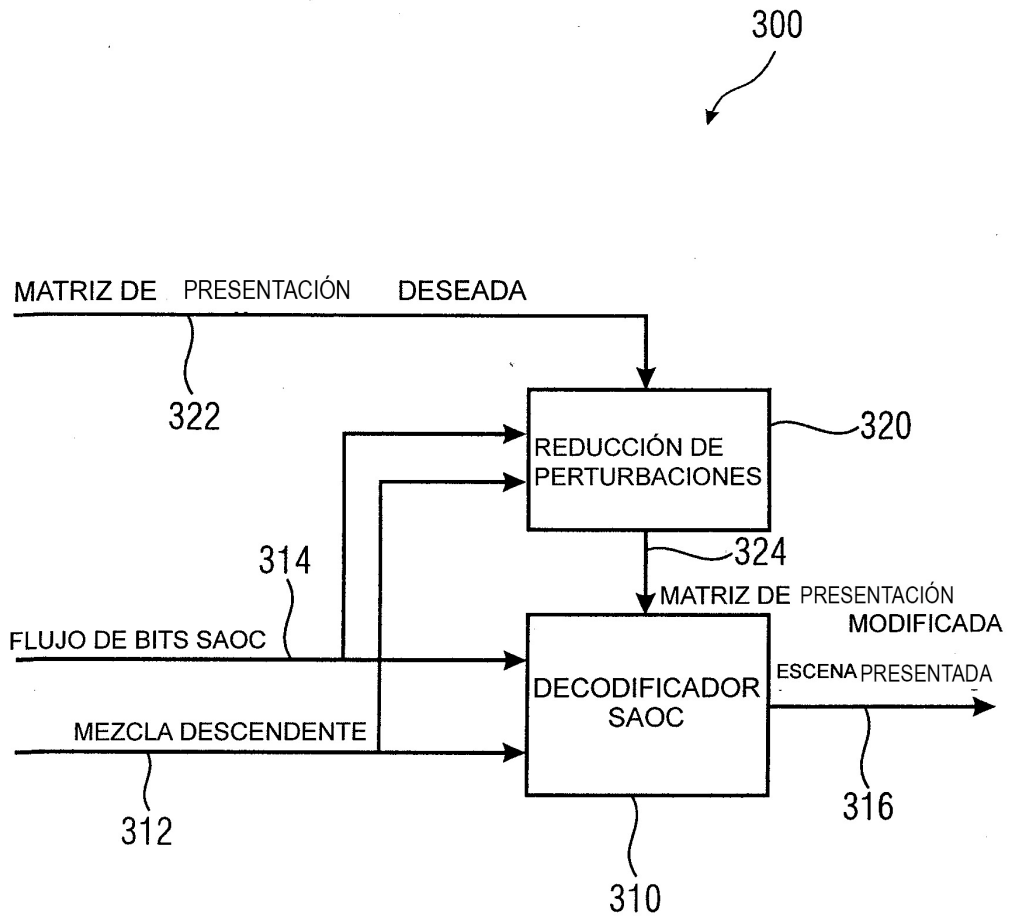


FIG. 3

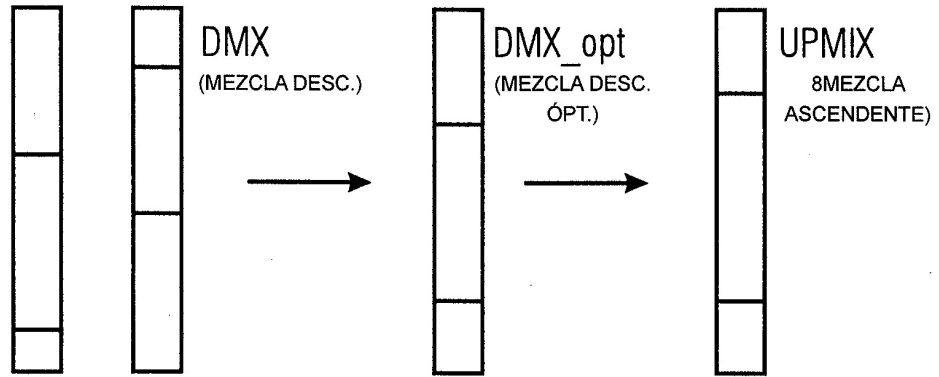
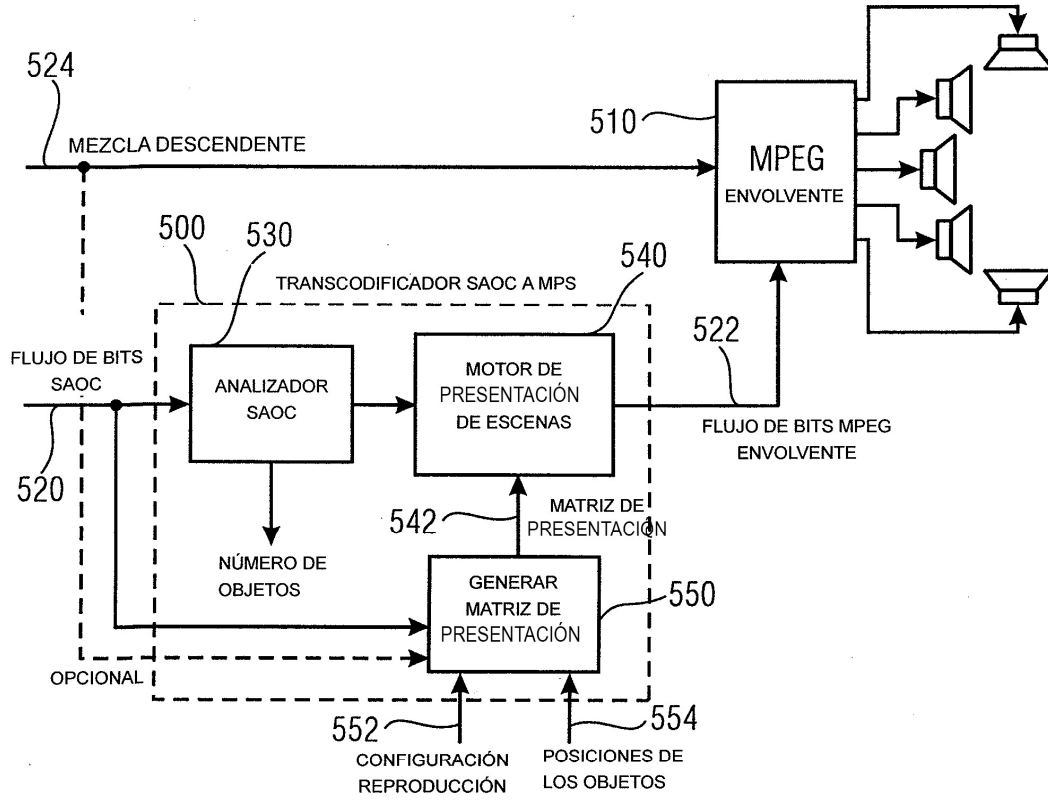
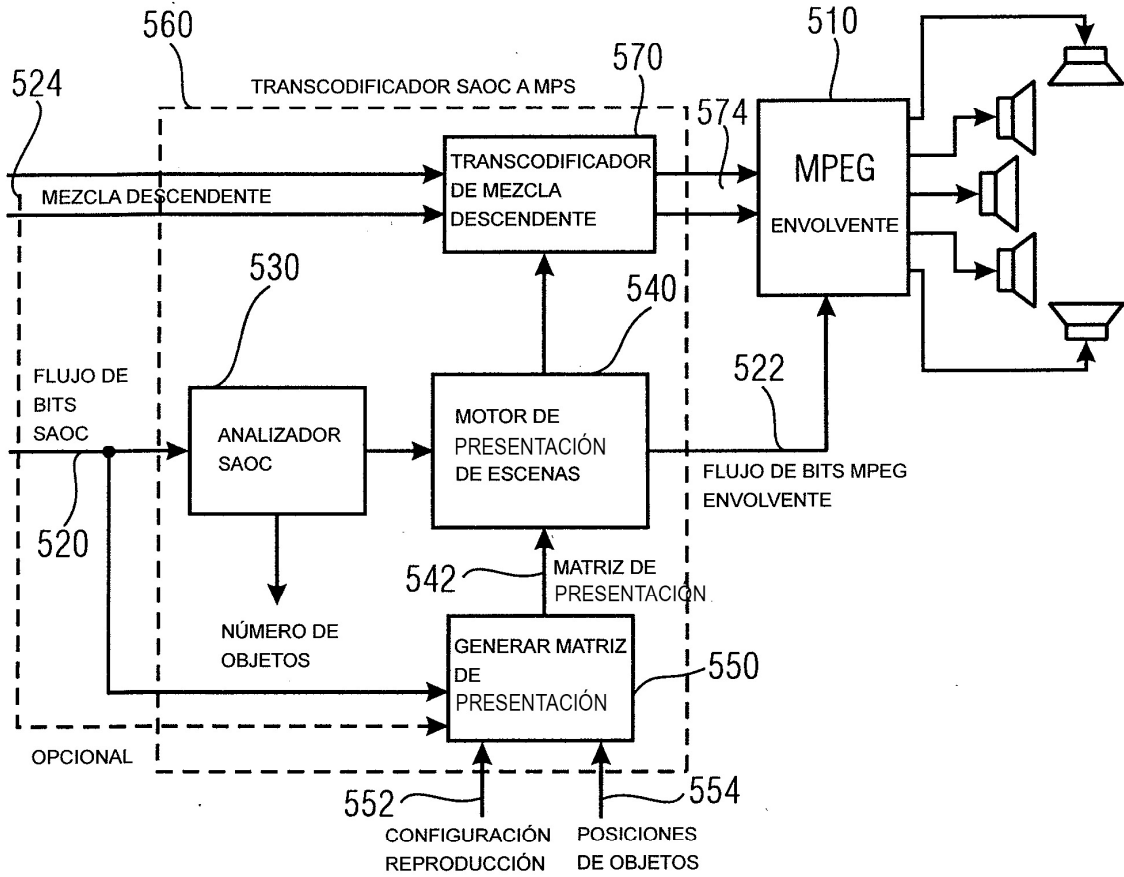


FIG. 4



(A) TRANSCODIFICADOR MONO BASADO EN MEZCLA DESCENDENTE

FIG. 5 A



(B) TRANSCODIFICADOR ESTEREO BASADO EN MEZCLA DESCENDENTE

FIG. 5 B

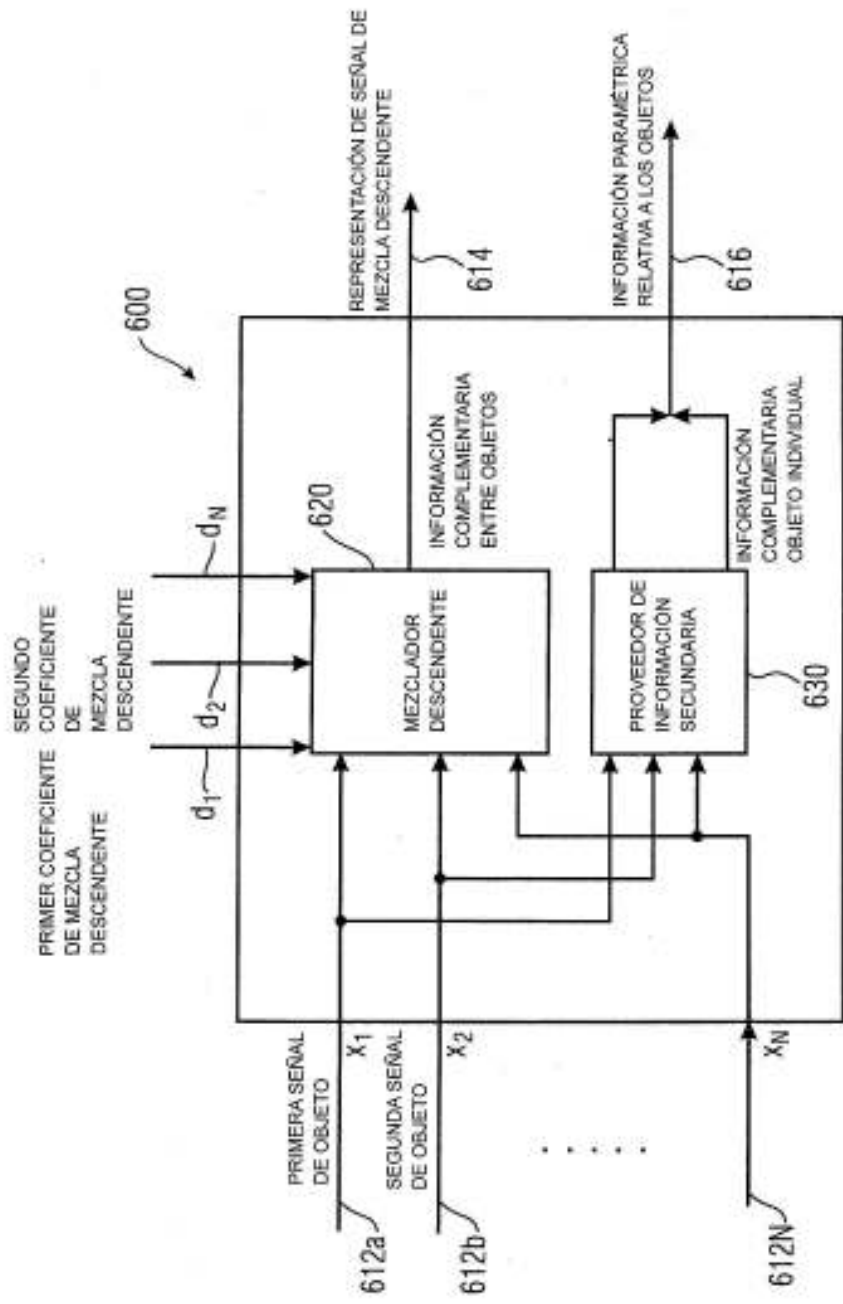


FIG. 6

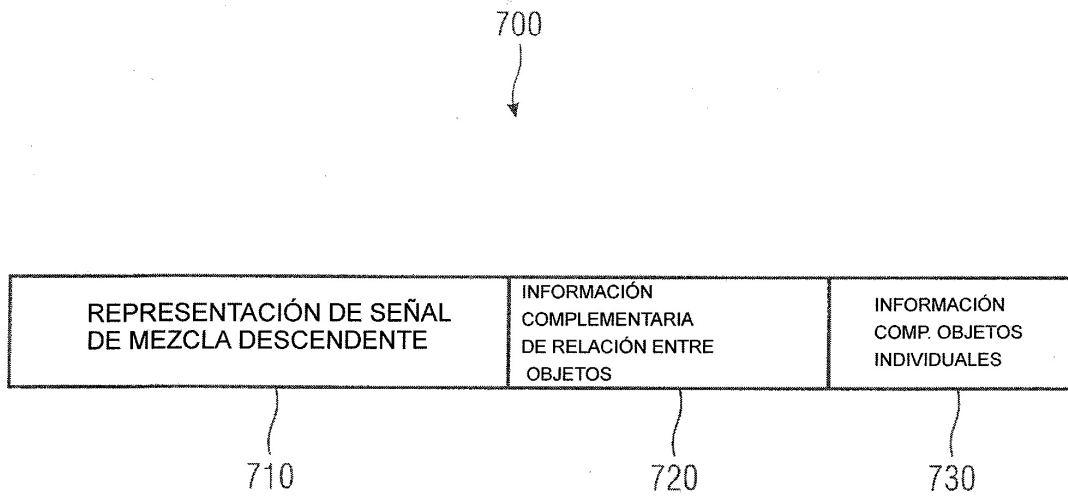
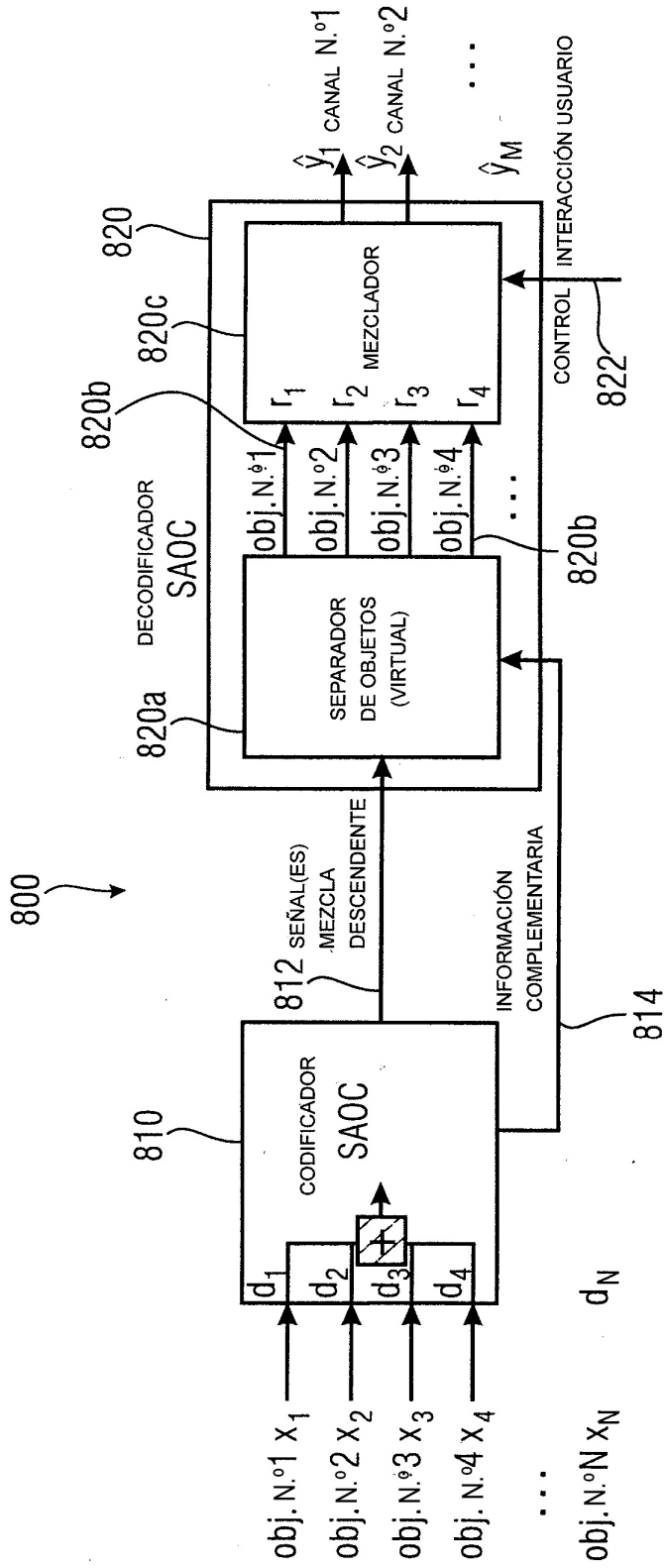
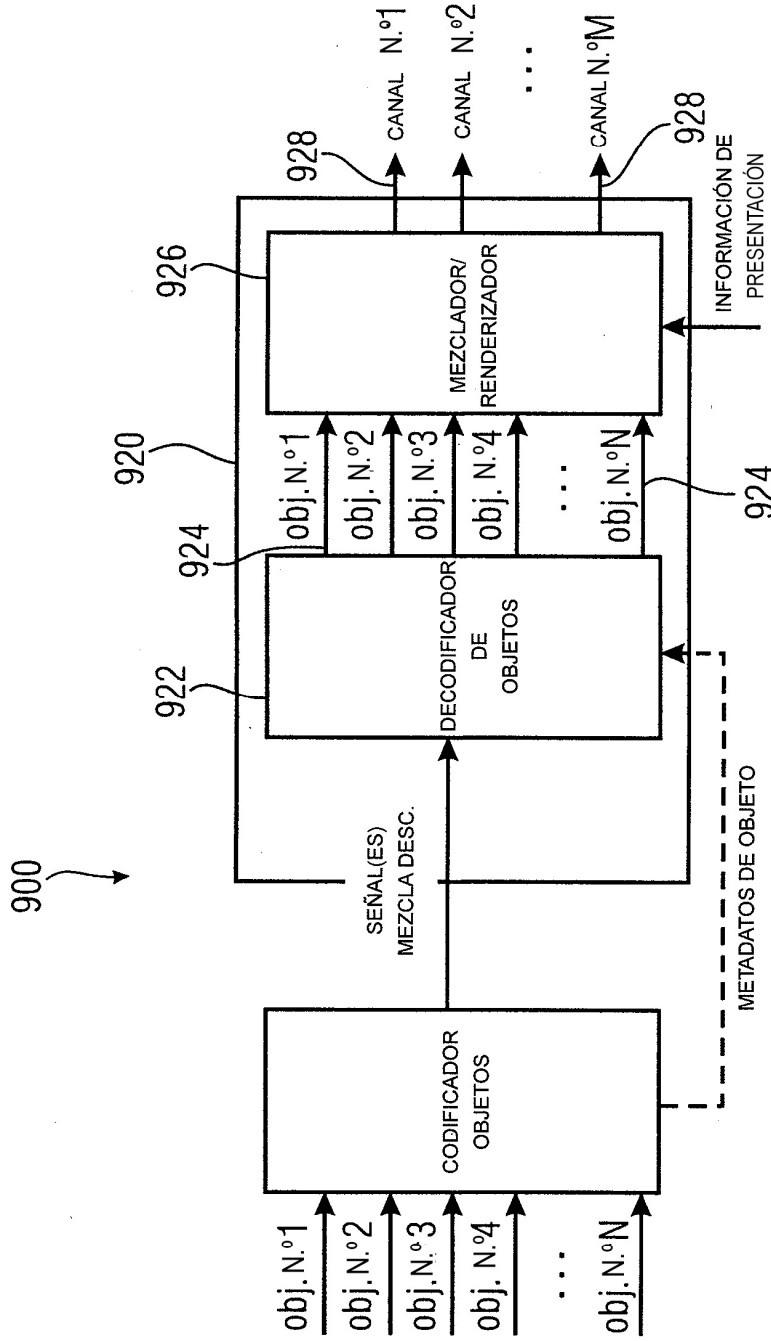


FIG. 7



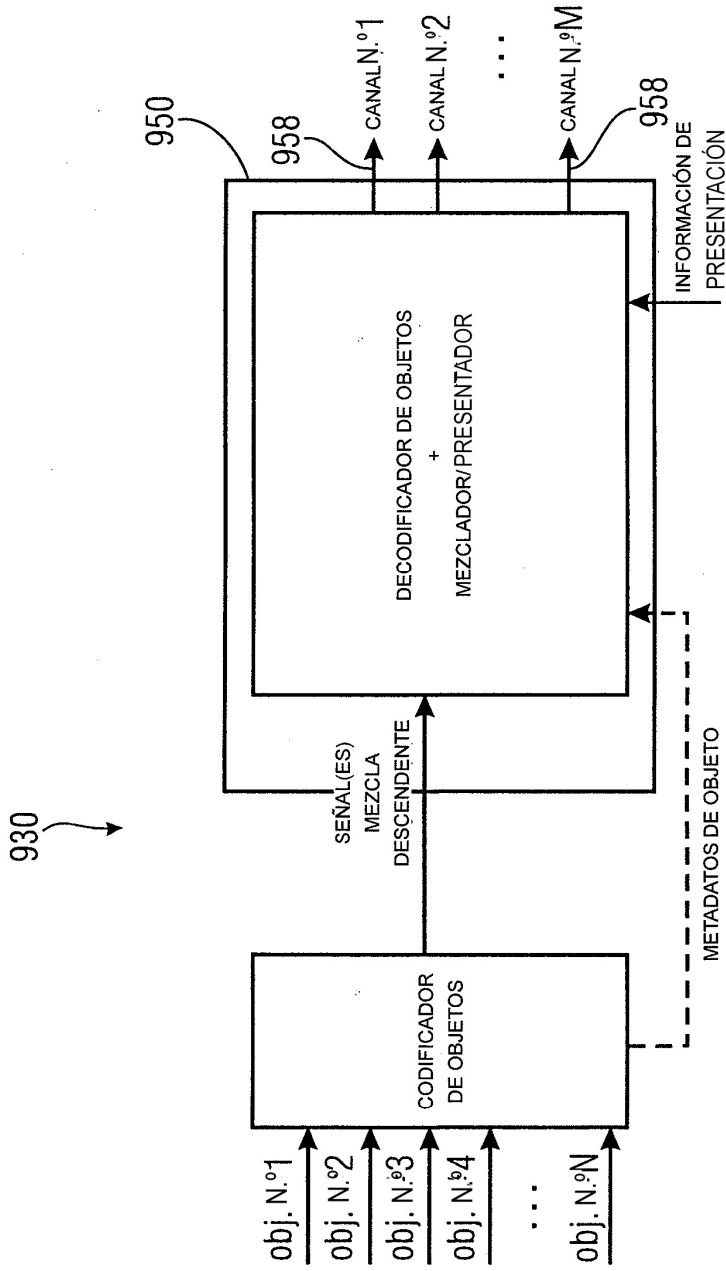
RESEÑA GENERAL SISTEMA MPEG SAOC

FIG. 8



RESEÑA GENERAL SISTEMA MPEG SAOC

FIG. 9 A



DECODIFICADOR Y MEZCLADOR INTEGRADOS

FIG. 9 B

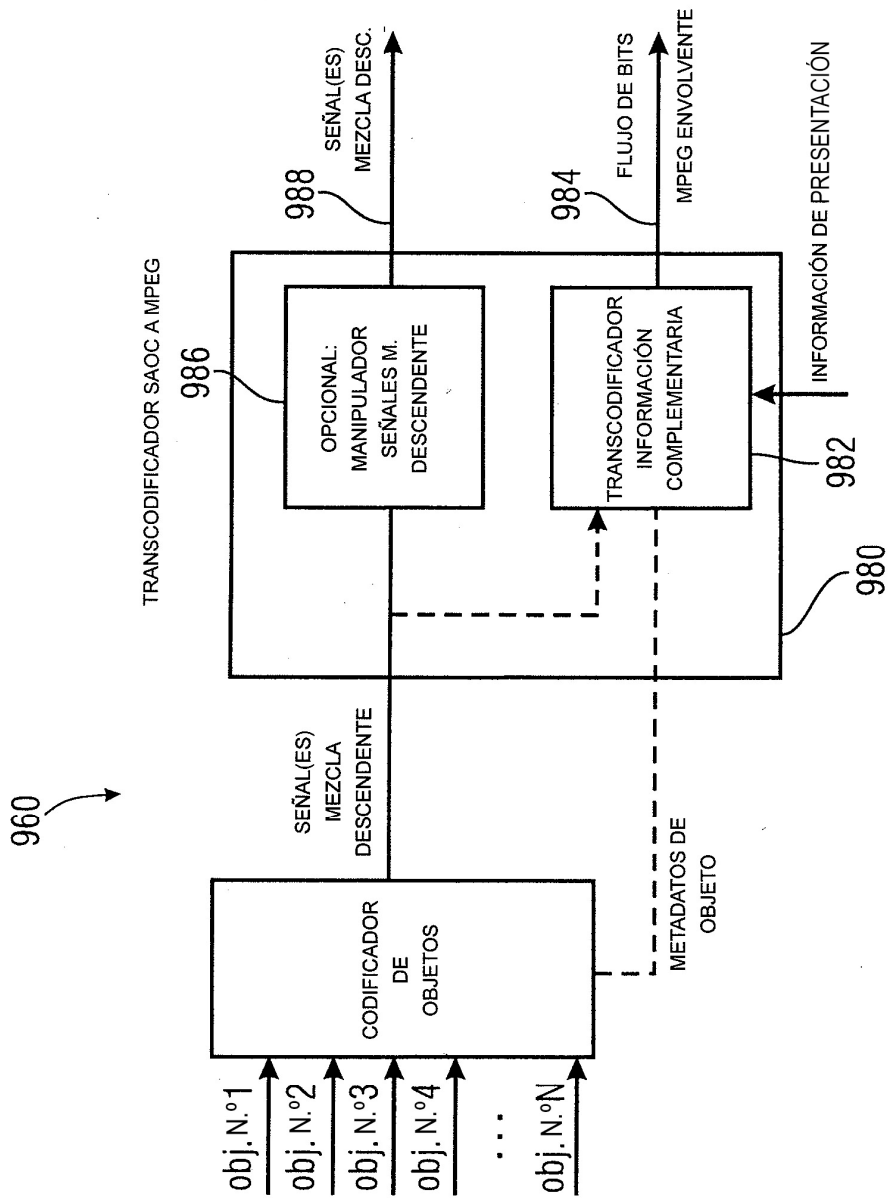


FIG. 9 C