

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 595 220**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/16 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.06.2013** **PCT/EP2013/063703**

87 Fecha y número de publicación internacional: **13.02.2014** **WO14023477**

96 Fecha de presentación y número de la solicitud europea: **28.06.2013** **E 13732189 (9)**

97 Fecha y número de publicación de la concesión europea: **03.08.2016** **EP 2883226**

54 Título: **Aparato y métodos para adaptar información de audio a codificación de objeto de audio espacial**

30 Prioridad:

10.08.2012 US 201261681732 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.12.2016

73 Titular/es:

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

72 Inventor/es:

**KASTNER, THORSTEN;
HERRE, JÜRGEN;
TERENTIV, LEON;
HELLMUTH, OLIVER;
PAULUS, JOUNI y
RIDDERBUSCH, FALKO**

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 595 220 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).

Aparato y métodos para adaptar información de audio a codificación de objeto de audio espacial

DESCRIPCIÓN

5 La presente invención se refiere a la decodificación de señales de audio y el procesamiento de señales de audio y, en particular, a un decodificador y unos métodos para adaptar información de audio a la codificación de objetos de audio espacial (SAOC).

10 En los sistemas de audio digitales modernos, hay una tendencia importante a permitir modificaciones relacionadas con el objeto de audio del contenido transmitido en el lado del receptor. Estas modificaciones incluyen modificaciones de ganancia de partes seleccionadas de la señal de audio y/o el re-posicionamiento espacial de objetos de audio dedicados en caso de reproducción multicanal a través de altavoces distribuidos espacialmente. Esto puede lograrse entregando individualmente diferentes partes del contenido de audio a los diferentes altavoces.

15 En otras palabras, en la técnica del procesamiento de audio, la transmisión de audio y el almacenamiento de audio, existe un deseo en aumento de permitir la interacción del usuario en la reproducción de contenidos de audio orientada a objetos y también una demanda para utilizar las posibilidades extendidas de la reproducción multicanal para presentar individualmente contenidos de audio o partes de los mismos, con el fin de mejorar la impresión de audición. Mediante esto, el uso del contenido de audio multicanal trae consigo mejoras significativas para el usuario. Por ejemplo, puede obtenerse una impresión de audición tridimensional, lo que trae consigo una mejora en la satisfacción del usuario en las aplicaciones de entretenimiento. Sin embargo, el contenido de audio multicanal también es útil en entornos profesionales, por ejemplo, en aplicaciones de conferencia telefónica, debido a que la inteligibilidad del hablante puede mejorarse usando una reproducción de audio multicanal. Otra aplicación posible es
20 ofrecer a un oyente de una pieza musical la posibilidad de ajustar individualmente el nivel de reproducción y/o posición espacial de diferentes partes (también denominadas "objetos de audio") o pistas, tales como una parte vocal o diferentes instrumentos. El usuario puede realizar tal ajuste por razones de gusto personal, para transcribir más fácilmente una o más partes de la pieza musical, con fines educativos, karaoke, ensayo, etc.

30 La transmisión discreta directa de todo el contenido de audio multicanal o multi-objeto digital, por ejemplo en forma de datos de modulación de código de impulso (PCM) o incluso formatos de audio comprimidos, exige velocidades de bits muy altas. Sin embargo, también es deseable transmitir y almacenar datos de audio de una manera eficiente en velocidad de bits. Por lo tanto, se está dispuesto a aceptar una solución intermedia razonable entre la calidad de audio y los requisitos de velocidad de bits con el fin de evitar una carga de recursos excesiva provocada por las aplicaciones multicanal/multiobjeto.

Recientemente, en el campo de la codificación de audio, se han introducido técnicas paramétricas para la transmisión/almacenamiento eficiente en velocidad de bits de señales de audio multicanal/multi-objeto, por ejemplo, por el Moving Picture Experts Group (MPEG) y otros. Un ejemplo es el Surround de MPEG (MPS) como un enfoque orientado al canal [MPS, BCC] o la codificación de objeto de audio espacial (SAOC) de MPEG como un enfoque orientado al objeto [JSC, SAOC, SAOC1, SAOC2]. Otro enfoque orientado al objeto se denomina "separación de fuentes informada" [ISS1, ISS2, ISS3, ISS4, ISS5, ISS6]. Estas técnicas tienen como objetivo reconstruir una escena de audio de salida deseada o un objeto de fuente de audio deseada basándose en una mezcla descendente de canales/objetos e información lateral adicional que describe la escena de audio transmitida/almacenada y/o los
45 objetos de fuente de audio en la escena de audio.

La estimación y aplicación de la información lateral relacionada con el canal/objeto en tales sistemas se hace de manera selectiva por tiempo-frecuencia. Por lo tanto, tales sistemas emplean transformadas de tiempo-frecuencia tales como la transformada discreta de Fourier (DFT), la transformada de Fourier de tiempo corto (STFT) o bancos de filtros como los bancos de filtros de espejo en cuadratura (QMF), etc. El principio básico de tales sistemas se representa en la figura 3, usando el ejemplo de SAOC de MPEG.

En el caso de la STFT, la dimensión temporal se representa por el número de tiempo-bloque y la dimensión espectral se captura por el número de coeficiente espectral ("bin"). En caso de QMF, la dimensión temporal se
55 representa por el número de tiempo-segmento y la dimensión espectral se captura por el número de sub-banda. Si la resolución espectral del QMF se mejora por la aplicación subsiguiente de una segunda etapa de filtro, todo el banco de filtros se denomina QMF híbrido y las sub-bandas de resolución fina se denominan sub-bandas híbridas.

Como ya se ha mencionado anteriormente, en SAOC el procesamiento general se realiza de manera selectiva por tiempo-frecuencia y puede describirse de la siguiente manera dentro de cada banda de frecuencia, como se representa en la figura 3:

- N señales de objeto de audio de entrada $s_1 \dots s_N$ se mezclan con los P canales $x_1 \dots x_P$ como parte del procesamiento de codificador, usando una matriz de mezcla descendente que consiste en los elementos $d_{1,1} \dots$

$d_{N,P}$. Además, el codificador extrae información lateral que describe las características de los objetos de audio de entrada (módulo estimador de información lateral (SIE)). Para SAOC DE MPEG, las relaciones de las potencias de objeto w.r.t. entre sí son la forma más básica de tal información lateral.

- La o las señales de mezcla descendente y la información lateral se transmiten/se almacenan. Para este fin, la o las señales de audio de mezcla descendente pueden comprimirse, por ejemplo, usando codificadores de audio perceptual bien conocidos, tales como MPEG-1/2 Capa II o III (también conocido como .mp3), codificación de audio avanzada de MPEG-2/4 (AAC), etc.
- En el extremo receptor, el decodificador intenta, conceptualmente, restablecer las señales de objeto originales ("separación de objeto") de las señales de mezcla descendente (decodificadas) usando la información lateral transmitida. Estas señales de objeto aproximadas $\hat{s}_1 \dots \hat{s}_N$ se mezclan, a continuación, en una escena objetivo representada por M canales de salida de audio $\hat{y}_1 \dots \hat{y}_M$ usando una matriz de presentación descrita por los coeficientes $r_{1,1} \dots r_{N,M}$ en la figura 3. La escena objetivo deseada puede ser, en caso extremo, la presentación de solo una señal fuente fuera de la mezcla (escenario de separación de fuente), pero también cualquier otra escena acústica arbitraria que consiste en los objetos transmitidos. Por ejemplo, la salida puede ser una escena objetivo de un solo canal, de 2 canales estéreo o 5.1 multicanal.

La figura 6 representa esquemáticamente el principio de un esquema de codificación/decodificación de audio. En particular, la figura 6 es una descripción del principio de una cadena de codificación/decodificación de audio.

En el lado de codificación, la señal de audio se comprime por un esquema de codificación de audio (que normalmente aprovecha efectos perceptuales) y se calcula la información lateral paramétrica (PSI) (véase codificador 601). El flujo de bits resultante que consiste en la señal de audio codificada y la PSI se almacena en (o se transmite a) el lado del decodificador, en el que pueden decodificarse por varias instancias de decodificador 620, 621, 622, marcadas como "A", "B", etc. en la figura 6. Estas instancias de decodificador pueden diferir entre sí (por ejemplo, diferentes niveles de complejidad en la especificación convencional, restricciones de aplicación o de implementación, etc.) [SAOC. SAOC1, SAOC2].

Los esquemas de codificación del estado de la técnica no son capaces de adaptar la PSI a un escenario o plataforma de aplicación objetivo específica de manera eficiente. Esto puede conducir a una complejidad computacional más alta (de lo necesario) en el lado del decodificador o puede dar como resultado problemas de compatibilidad.

El objeto de la presente invención es proporcionar conceptos mejorados para la codificación de objetos de audio. El objeto de la presente invención se resuelve por un aparato para adaptar información de audio de entrada de acuerdo con la reivindicación 1, por un método para adaptar información de audio de entrada de acuerdo con la reivindicación 11 y por un programa informático de acuerdo con la reivindicación 13.

Se proporciona un aparato para adaptar información de audio de entrada, que codifica uno o más objetos de audio, para obtener información de audio adaptada. La información de audio de entrada comprende dos o más canales de mezcla descendente de audio de entrada y comprende además información lateral paramétrica de entrada. La información de audio adaptada comprende uno o más canales de mezcla descendente de audio adaptados y comprende además información lateral paramétrica adaptada.

El aparato comprende un modificador de señales de mezcla descendente para adaptar, dependiendo de la información de adaptación, los dos o más canales de mezcla descendente de audio de entrada para obtener el uno o más canales de mezcla descendente de audio adaptados.

Además, el aparato comprende un adaptador de información lateral paramétrica para adaptar, dependiendo de la información de adaptación, la información lateral paramétrica de entrada para obtener la información lateral paramétrica adaptada.

La información de adaptación comprende una matriz de adaptación (\mathbf{D}_{dmx}^{DSM}).

El modificador de señal de mezcla descendente está configurado para adaptar, dependiendo de la matriz de adaptación (\mathbf{D}_{dmx}^{DSM}), los dos o más canales de mezcla descendente de audio de entrada (\mathbf{X}_{dmx}^{ENC}) para obtener el uno o más canales de mezcla descendente de audio adaptados (\mathbf{X}_{dmx}^{DSM}).

El adaptador de información lateral paramétrica está configurado para adaptar, dependiendo de la matriz de adaptación (\mathbf{D}_{dmx}^{DSM}), la información lateral paramétrica de entrada (\mathbf{D}_{dmx}^{ENC}) para obtener la información lateral

paramétrica adaptada (\mathbf{D}_{dmx}^{PSI}).

De acuerdo con una realización, el modificador de señal de mezcla descendente puede estar configurado para adaptar los dos o más canales de mezcla descendente de audio de entrada dependiendo de la información de adaptación, de tal manera que el número del uno o más canales de mezcla descendente de audio adaptados es menor que el número de los dos o más canales de mezcla descendente de audio de entrada.

En una realización, la información de adaptación puede depender de una instancia de decodificador. El modificador de señal de mezcla descendente puede estar configurado para adaptar los dos o más canales de mezcla descendente de audio de entrada dependiendo de la instancia de decodificador. En este caso y en lo sucesivo, las expresiones "decodificador" e "instancia de decodificador" tienen el mismo significado.

De acuerdo con una realización, la instancia de decodificador puede ser capaz de decodificar a lo sumo un número máximo de canales de mezcla descendente. La información de adaptación puede depender de dicho número máximo de canales de mezcla descendente. Además, el modificador de señal de mezcla descendente puede estar configurado para adaptar los dos o más canales de mezcla descendente de audio de entrada, dependiendo de la información de adaptación para obtener el uno o más canales de mezcla descendente de audio adaptados, de tal manera que el número del uno o más canales de mezcla descendente adaptados es igual a dicho número máximo de canales de mezcla descendente.

De acuerdo con una realización, el modificador de señal de mezcla descendente puede estar configurado para adaptar, dependiendo de la matriz de adaptación \mathbf{D}_{dmx}^{DSM} , los dos o más canales de mezcla descendente de audio de entrada \mathbf{X}_{dmx}^{ENC} para obtener el uno o más canales de mezcla descendente de audio adaptados \mathbf{X}_{dmx}^{DSM} aplicando la fórmula:

$$\mathbf{X}_{dmx}^{DSM} = \mathbf{D}_{dmx}^{DSM} \mathbf{X}_{dmx}^{ENC}.$$

De acuerdo con una realización, el adaptador de información lateral paramétrica puede estar configurado para adaptar, dependiendo de la matriz de adaptación \mathbf{D}_{dmx}^{DSM} , la información lateral paramétrica de entrada \mathbf{D}_{dmx}^{ENC} para obtener la información lateral paramétrica adaptada \mathbf{D}_{dmx}^{PSI} aplicando la fórmula:

$$\mathbf{D}_{dmx}^{PSI} = \mathbf{D}_{dmx}^{DSM} \mathbf{D}_{dmx}^{ENC}.$$

En una realización, la información lateral paramétrica de entrada (\mathbf{D}_{dmx}^{enc}) puede indicar una matriz de mezcla descendente inicial, de tal manera que aplicando la matriz de mezcla descendente inicial (\mathbf{D}_{dmx}^{enc}) sobre el uno o más objetos de audio (S), se obtienen los dos o más canales de mezcla descendente de audio de entrada (\mathbf{X}_{dmx}^{enc}). El adaptador de información lateral paramétrica puede estar configurado para determinar una matriz de mezcla descendente adaptada (\mathbf{D}_{dmx}^{PSI}) como la información lateral paramétrica adaptada, de tal manera que aplicando la matriz de mezcla descendente adaptada (\mathbf{D}_{dmx}^{PSI}) sobre uno o más objetos de audio (S), se obtienen el uno o más canales de mezcla descendente de audio adaptados (\mathbf{X}_{dmx}^{DSM}).

Además, de acuerdo con una realización, se proporciona un aparato para generar uno o más canales de audio a partir de la información de audio que codifican uno o más objetos de audio.

El aparato para generar el uno o más canales de audio comprende un aparato de acuerdo con una de las realizaciones descritas anteriormente para adaptar la información de audio de entrada para obtener información de audio adaptada, en el que la información de audio de entrada comprende dos o más canales de mezcla descendente de audio de entrada y comprende, además, información lateral paramétrica de entrada, en el que la información de audio adaptada comprende uno o más canales de mezcla descendente de audio adaptados y comprende, además, información lateral paramétrica adaptada.

Además, el aparato para generar el uno o más canales de audio comprende una instancia de decodificador, para decodificar, dependiendo de la información lateral paramétrica adaptada, el uno o más canales de mezcla descendente de audio adaptados para obtener el uno o más canales de audio.

5 De acuerdo con una realización, el adaptador de información lateral paramétrica del aparato para adaptar la información de audio de entrada puede estar configurado para recibir un flujo de bits de entrada que comprende la información lateral paramétrica de entrada. El adaptador de información lateral paramétrica del aparato para adaptar información de audio de entrada puede estar configurado para adaptar la información lateral paramétrica de entrada para obtener la información lateral paramétrica adaptada y para introducir la información lateral paramétrica adaptada en la instancia de decodificador. La instancia de decodificador puede estar configurada para decodificar el uno o más canales de mezcla descendente de audio adaptados dependiendo de la información lateral paramétrica adaptada.

15 En otra realización, el adaptador de información lateral paramétrica del aparato para adaptar información de audio de entrada puede estar configurado para recibir un flujo de bits de entrada que comprende la información lateral paramétrica de entrada. El adaptador de información lateral paramétrica del aparato para adaptar información de audio de entrada puede estar configurado para sustituir la información lateral paramétrica de entrada dentro del flujo de bits de entrada por la información lateral paramétrica adaptada para obtener un flujo de bits modificado. El adaptador de información lateral paramétrica del aparato para adaptar la información de audio de entrada puede estar configurado para introducir el flujo de bits modificado en la instancia de decodificador. Además, la instancia de decodificador puede estar configurada para decodificar el uno o más canales de mezcla descendente de audio adaptados dependiendo del flujo de bits modificado.

25 Además, se proporciona un método para adaptar información de audio de entrada, que codifica uno o más objetos de audio, para obtener información de audio adaptada. La información de audio de entrada comprende dos o más canales de mezcla descendente de audio de entrada y comprende además información lateral paramétrica de entrada. La información de audio adaptada comprende uno o más canales de mezcla descendente de audio adaptados y comprende además información lateral paramétrica adaptada. El método comprende las etapas de la reivindicación 11.

30 Además, se proporciona un programa informático para implementar el método descrito anteriormente, cuando se ejecuta por un ordenador o un procesador de señales.

Las realizaciones preferidas se proporcionarán en las reivindicaciones dependientes.

35 A continuación, se describen en más detalle realizaciones de la presente invención con referencia a las figuras, en las que:

40 Figura 1 ilustra un aparato para adaptar información de audio de entrada, que codifica uno o más objetos de audio, para obtener información de audio adaptada de acuerdo con una realización.

Figura 2 ilustra un aparato para adaptar información de audio de entrada, que codifica uno o más objetos de audio, para obtener información de audio adaptada de acuerdo con otra realización.

45 Figura 3 muestra un diagrama de bloques esquemático de una visión general conceptual de un sistema de SAOC.

Figura 4 muestra un diagrama esquemático e ilustrativo de una representación temporal-espectral de una señal de audio de un solo canal.

50 Figura 5 muestra un diagrama de bloques esquemático de un cálculo selectivo de tiempo-frecuencia de información lateral dentro de un codificador de SAOC.

Figura 6 representa esquemáticamente el principio de un esquema de codificación/decodificación de audio.

55 Figura 7 ilustra un aparato para generar uno o más canales de audio a partir de una información de audio de entrada que codifica uno o más objetos de audio de acuerdo con una realización.

Figura 8 ilustra una aplicación de PSIA junto con un esquema de codificación/decodificación de acuerdo con una realización.

60 Figura 9 ilustra una aplicación de PSIA por separado dentro de un esquema de codificación/decodificación de acuerdo con una realización.

Antes de describir las realizaciones de la presente invención, se proporcionan más antecedentes del estado de la

técnica de los sistemas de SAOC.

La figura 3 muestra una disposición general de un codificador de SAOC 10 y un decodificador de SAOC 12. El codificador de SAOC 10 recibe como entrada N objetos, es decir, señales de audio s_1 a s_N . En particular, el
 5 codificador 10 comprende un mezclador descendente 16 que recibe las señales de audio s_1 a s_N y mezcla de manera descendente las mismas con una señal de mezcla descendente 18. Como alternativa, la mezcla descendente puede proporcionarse externamente ("mezcla descendente artística") y el sistema estima la información lateral adicional para hacer que la mezcla descendente proporcionada coincida con la mezcla descendente calculada. En la figura 3, se muestra que la señal de mezcla descendente es una señal de canal P. Por
 10 lo tanto, es concebible cualquier configuración de señal de mezcla descendente monofónica ($P = 1$), estereofónica ($P = 2$) o multicanal ($P > 2$).

En el caso de una mezcla descendente estereofónica, los canales de la señal de mezcla descendente 18 se indican como L0 y R0, en el caso de una mezcla descendente monofónica, lo mismo se indica simplemente como L0. Con el
 15 fin de permitir que el decodificador de SAOC 12 recupere los objetos individuales s_1 a s_N , el estimador de información lateral 17 provee al decodificador de SAOC 12 de una información lateral que incluye unos parámetros de SAOC. Por ejemplo, en el caso de una mezcla descendente estereofónica, los parámetros de SAOC comprenden diferencias a nivel de objeto (OLD), correlaciones inter-objeto (COI) (parámetros de correlación cruzada de inter-objeto), valores de ganancia de mezcla descendente (DMG) y diferencias de nivel de canal de mezcla descendente (DCLD). La información lateral 20, que incluye los parámetros de SAOC, junto con la señal de mezcla descendente
 20 18, forma la corriente de datos de salida de SAOC recibida por el decodificador de SAOC 12.

El decodificador de SAOC 12 comprende un mezclador ascendente que recibe la señal de mezcla descendente 18, así como la información lateral 20 con el fin de recuperar y presentar las señales de audio \hat{s}_1 y \hat{s}_N sobre cualquier
 25 conjunto de canales seleccionados por el usuario \hat{y}_1 a \hat{y}_M , prescribiéndose la presentación por la información de presentación 26 introducida en el decodificador de SAOC 12.

Las señales de audio s_1 a s_N pueden introducirse en el codificador 10 en cualquier dominio de codificación, tal como en un dominio de tiempo o espectral. En caso de que las señales de audio s_1 a s_N se introduzcan en el codificador
 30 10 en el dominio de tiempo, tal como PCM codificadas, el codificador 10 puede usar un banco de filtros, tal como un banco de QMF híbrido, con el fin de transferir las señales a un dominio espectral, en el que las señales de audio se representan en varias sub-bandas asociadas con diferentes porciones espectrales, en una resolución de banco de filtros específica. Si las señales de audio s_1 a s_N ya están en la representación esperada por el codificador 10, el mismo no tiene que realizar la descomposición espectral.

La figura 4 muestra una señal de audio en el dominio espectral recién mencionado. Como puede verse, la señal de audio se representa como una pluralidad de señales de sub-banda. Cada señal de sub-banda 30_1 a 30_k consiste en una secuencia temporal de valores de sub-banda indicados por las cajas pequeñas 32. Como puede verse, los
 40 valores de sub-banda 32 de las señales de sub-banda 30_1 a 30_k se sincronizan entre sí en el tiempo, de manera que, por cada uno de los segmentos de tiempo de banco de filtros consecutivos 34, cada sub-banda 30_1 a 30_k comprende un valor de sub-banda exacto 32. Como se ilustra por el eje de frecuencia 36, las señales de sub-banda 30_1 a 30_k se asocian con diferentes regiones de frecuencia y como se ilustra por el eje del tiempo 38, los segmentos de tiempo de banco de filtros 34 se disponen consecutivamente en el tiempo.

Como se ha descrito anteriormente, el extractor de información lateral 17 de la figura 3 calcula los parámetros de SAOC a partir de las señales de audio de entrada a s_1 a s_N . De acuerdo con la norma de SAOC implementada actualmente, el codificador 10 realiza este cálculo en una resolución de tiempo/frecuencia que puede reducirse en
 45 relación con la resolución de tiempo/frecuencia original como se determinó por los segmentos de tiempo de banco de filtros 34 y por la descomposición de sub-banda, en una cierta cantidad, señalándose esta cierta cantidad en el lado del decodificador dentro de la información lateral 20. Los grupos de segmentos de tiempo de banco de filtros consecutivos 34 pueden formar un marco de SAOC 41. Además, el número de bandas de parámetro dentro del marco de SAOC 41 se transporta dentro de la información lateral 20. Por lo tanto, el dominio de tiempo/frecuencia se divide en mosaicos de tiempo/frecuencia ejemplificados en la figura 4 por las líneas discontinuas 42. En la figura 4 las bandas de parámetros se distribuyen de la misma manera en los diversos marcos de SAOC ilustrados 41, de tal
 50 manera que se obtiene una disposición regular de mosaicos de tiempo/frecuencia. En general, sin embargo, las bandas de parámetros pueden variar de un marco de SAOC 41 al siguiente, dependiendo de las diferentes necesidades de resolución espectral en los marcos de SAOC respectivos 41. Además, la longitud de los marcos de SAOC 41 también puede variar. Como consecuencia, la disposición de los mosaicos de tiempo/frecuencia puede ser irregular. Sin embargo, los mosaicos de tiempo/frecuencia dentro de un marco de SAOC específico 41 tienen
 55 habitualmente la misma duración y están alineados en la dirección del tiempo, es decir, todos los mosaicos de tiempo/frecuencia en dicho marco de SAOC 41 empiezan al inicio del marco de SAOC dado 41 y terminan al final de dicho marco de SAOC 41.

El extractor de información lateral 17 representado en la figura 3 calcula los parámetros de SAOC de acuerdo con

las siguientes fórmulas. En particular, el extractor de información lateral 17 calcula diferencias a nivel de objeto para cada objeto i como

$$OLD_i^{l,m} = \frac{\sum_{n \in l} \sum_{k \in m} x_i^{n,k} x_i^{n,k*}}{\max_j \left(\sum_{n \in l} \sum_{k \in m} x_j^{n,k} x_j^{n,k*} \right)}$$

5 en donde las sumas y los índices n y k , respectivamente, avanzan a través de todos los índices temporales 34 y todos los índices espectrales 30 que pertenecen a un cierto mosaico de tiempo/frecuencia 42, referenciado por los índices l para el marco de SAOC (o segmento de tiempo de procesamiento) y m para la banda de parámetro. Por lo tanto, las energías de todos los valores de sub-banda x_i de una señal u objeto de audio i se suman hasta, y se normalizan con respecto a, el valor de energía más alto de ese mosaico entre todos los objetos o señales de audio.

10 $x_i^{n,k*}$ indica el conjugado complejo de $x_i^{n,k}$.

Además, el extractor de información lateral de SAOC 17 es capaz de calcular una medida de similitud de los mosaicos de tiempo/frecuencia correspondientes de pares de objetos de entrada diferentes s_1 a s_N . Aunque el extractor de información lateral de SAOC 17 puede calcular la medida de similitud entre todos los pares de objetos de entrada s_1 a s_N , el extractor de información lateral 17 también puede suprimir la señalización de las medidas de similitud o restringir el cálculo de las medidas de similitud a los objetos de audio s_1 a s_N , que forman los canales izquierdo o derecho de un canal estereofónico común. En cualquier caso, la medida de similitud se denomina parámetro de correlación cruzada de inter-objetos $IOC_{i,j}^{l,m}$. El cálculo es de la siguiente manera:

$$20 \quad IOC_{i,j}^{l,m} = IOC_{j,i}^{l,m} = \text{Re} \left\{ \frac{\sum_{n \in l} \sum_{k \in m} x_i^{n,k} x_j^{n,k*}}{\sqrt{\sum_{n \in l} \sum_{k \in m} x_i^{n,k} x_i^{n,k*} \sum_{n \in l} \sum_{k \in m} x_j^{n,k} x_j^{n,k*}}} \right\}$$

avanzando de nuevo los índices n y k a través de todos los valores de sub-banda pertenecientes a un determinado mosaico de tiempo/frecuencia 42, indicando i y j un determinado par de objetos de audio s_1 a s_N e indicando $\text{Re}\{\}$ la operación de descartar la parte imaginaria del argumento complejo.

25 El mezclador descendente 16 de la figura 3 mezcla de manera descendente los objetos s_1 a s_N mediante el uso de factores de ganancia aplicados a cada objeto s_1 a s_N . Es decir, se aplica un factor de ganancia d_i al objeto i y, a continuación, todos los objetos así ponderados s_1 a s_N se suman hasta obtener una señal de mezcla descendente monofónica, que se ejemplifica en la figura 3 si $P = 1$. En otro caso a modo de ejemplo de una señal de mezcla descendente de dos canales, representado en la figura 3, si $P = 2$, se aplica un factor de ganancia $d_{1,i}$ al objeto i y, a continuación, todos estos objetos amplificados de ganancia se suman con el fin de obtener el canal de mezcla descendente izquierdo $L0$ y se aplican los factores de ganancia $d_{2,i}$ al objeto i y, a continuación, los objetos amplificados de ganancia se suman con el fin de obtener el canal de mezcla descendente derecho $R0$. Se aplicará un procesamiento que es análogo al anterior en el caso de una mezcla descendente multicanal ($P > 2$).

35 Esta prescripción de mezcla descendente se señala al lado del decodificador por medio de ganancias de mezcla descendente DMG_i y en el caso de una señal de mezcla descendente estereofónica, diferencias de nivel de canal de mezcla descendente $DCLD_i$.

40 Las ganancias de mezcla descendente se calculan de acuerdo con:

$$DMG_i = 20 \log_{10} (d_i + \varepsilon) \quad , \text{ (mezcla descendente monofónica),}$$

$$DMG_i = 10 \log_{10} (d_{1,i}^2 + d_{2,i}^2 + \varepsilon) \quad , \text{ (mezcla descendente estereofónica),}$$

45 donde ε es un número pequeño tal como 10^{-9} .

Para las DCLD se aplica la siguiente fórmula:

$$DCLD_i = 20 \log_{10} \left(\frac{d_{1,i}}{d_{2,i} + \varepsilon} \right).$$

En el modo normal, el mezclador descendente 16 genera la señal de mezcla descendente de acuerdo con:

$$(L0) = (d_i) \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix}$$

para una mezcla descendente monofónica o

$$\begin{pmatrix} L0 \\ R0 \end{pmatrix} = \begin{pmatrix} d_{1,i} \\ d_{2,i} \end{pmatrix} \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix}$$

para una mezcla descendente estereofónica, respectivamente.

Por lo tanto, en las formulas mencionadas anteriormente, los parámetros *OLD* e *IOC* son una función de las señales de audio y los parámetros *DMG* y *DCLD* son una función de *d*. Por cierto, cabe señalar que *d* puede ser variable en tiempo y en frecuencia.

Por lo tanto, en el modo normal, el mezclador descendente 16 mezcla todos los objetos s_1 a s_N sin preferencias, es decir, manejando todos los objetos s_1 a s_N por igual.

En el lado del decodificador, el mezclador ascendente realiza la inversión del procedimiento de mezcla descendente y la implementación de la "información de presentación" 26 representada por una matriz **R** (en la literatura también denominada algunas veces **A**) en una etapa de cálculo, es decir, en el caso de una mezcla descendente de dos canales

$$\begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_M \end{pmatrix} = \mathbf{RED}^* (\mathbf{DED}^*)^{-1} \begin{pmatrix} L0 \\ R0 \end{pmatrix},$$

donde la matriz **E** es una función de los parámetros *OLD* e *IOC* y la matriz **D** contiene los coeficientes de mezcla descendente como

$$\mathbf{D} = \begin{pmatrix} d_{1,1} & \cdots & d_{1,N} \\ \vdots & \ddots & \vdots \\ d_{P,1} & \cdots & d_{P,N} \end{pmatrix}.$$

La matriz **E** es una matriz de covarianza estimada de los objetos de audio s_1 a s_N . En implementaciones de SAOC actuales, el cálculo de la matriz de covarianza estimada **E** se realiza habitualmente en la resolución espectral/temporal de los parámetros de SAOC, esto es, para cada (l, m) , de tal manera que la matriz de covarianza estimada puede escribirse como $E^{l,m}$. La matriz de covarianza estimada $E^{l,m}$ es de tamaño $N \times N$ definiéndose sus coeficientes como

$$e_{i,j}^{l,m} = \sqrt{OLD_i^{l,m} OLD_j^{l,m}} IOC_{i,j}^{l,m}.$$

Por lo tanto, la matriz $E^{l,m}$ con

$$E^{l,m} = \begin{pmatrix} e_{1,1}^{l,m} & \cdots & e_{1,N}^{l,m} \\ \vdots & \ddots & \vdots \\ e_{N,1}^{l,m} & \cdots & e_{N,N}^{l,m} \end{pmatrix}$$

tiene a lo largo de su diagonal las diferencias a nivel de objeto, es decir, $e_{i,j}^{l,m} = OLD_i^{l,m}$ para $i=j$, debido a que $OLD_i^{l,m} = OLD_j^{l,m}$ e $IOC_{i,j}^{l,m} = 1$ para $i=j$. Fuera de su diagonal, la matriz de covarianza estimada **E** tiene coeficientes de matriz que representan la media geométrica de las diferencias a nivel de objeto de los objetos i y j , respectivamente, ponderados con la medida de correlación cruzada de inter-objeto $IOC_{i,j}^{l,m}$.

La figura 5 muestra un posible principio de implementación en el ejemplo del estimador de información lateral (SIE) como parte de un codificador de SAOC 10. El codificador de SAOC 10 comprende el mezclador 16 y el estimador de información lateral (SIE) 17. El SIE consiste conceptualmente en dos módulos: un módulo 45 para calcular una representación de tiempo/frecuencia basada en tiempo corto (por ejemplo, STFT o QMF) de cada señal. La representación de tiempo/frecuencia de tiempo corto calculada se introduce en el segundo módulo 46, el módulo de estimación de información lateral selectivo en tiempo/frecuencia (t/f-SIE). El módulo de t/f-SIE 46 calcula la información lateral para cada mosaico de tiempo/frecuencia. En implementaciones de SAOC actuales, la transformada de tiempo/frecuencia es fija e idéntica para todos los objetos de audio s_1 a s_N . Además, los parámetros de SAOC se determinan sobre marcos de SAOC que son los mismos para todos los objetos de audio y tienen la misma resolución de tiempo/frecuencia para todos los objetos de audio s_1 a s_N , independientemente por lo tanto de las necesidades de objeto-específico para la resolución temporal fina en algunos casos o la resolución espectral fina en otros casos.

A continuación, se describen realizaciones de la presente invención.

La figura 1 ilustra un aparato para adaptar información de audio de entrada, que codifica uno o más objetos de audio, para obtener información de audio adaptada de acuerdo con una realización.

La información de audio de entrada comprende dos o más canales de mezcla descendente de audio de entrada y comprende, además, información lateral paramétrica de entrada. La información de audio adaptada comprende uno o más canales de mezcla descendente de audio adaptados y comprende, además, información lateral paramétrica adaptada.

El aparato comprende un modificador de señal de mezcla descendente (DSM) 110 para adaptar, dependiendo de la información de adaptación, los dos o más canales de mezcla descendente de audio de entrada para obtener el uno o más canales de mezcla descendente de audio adaptados.

Además, el aparato comprende un adaptador de información lateral paramétrica (PSIA) 120 para adaptar, dependiendo de la información de adaptación, la información lateral paramétrica de entrada para obtener la información lateral paramétrica adaptada.

La figura 2 ilustra un aparato para adaptar información de audio de entrada, que codifica uno o más objetos de audio, para obtener información de audio adaptada de acuerdo con otra realización.

En una realización, la información de adaptación puede depender de una instancia de decodificador y el modificador de señal de mezcla descendente 110 puede estar configurado para adaptar los dos o más canales de mezcla descendente de audio de entrada dependiendo de la instancia de decodificador.

Por ejemplo, el modificador de señal de mezcla descendente 110 de la figura 2 adapta la mezcla descendente a las capacidades de la instancia de decodificador específica.

De acuerdo con una realización, el modificador de señal de mezcla descendente 110 puede estar configurado para adaptar los dos o más canales de mezcla descendente de audio de entrada dependiendo de la información de adaptación, de tal manera que el número del uno o más canales de mezcla descendente de audio adaptados es menor que el número de los dos o más canales de mezcla descendente de audio de entrada.

Por ejemplo, en la realización de la figura 2, el modificador de señal de mezcla descendente 110 reduce el número de canales de transporte/mezcla descendente.

Por ejemplo, 22,2 canales de mezcla descendente de audio de entrada (= 24 canales de mezcla descendente de

audio de entrada) pueden reducirse a 7,1 canales de mezcla descendente de audio adaptados (= 8 canales de mezcla descendente de audio adaptados).

O, por ejemplo, 5,1 canales de mezcla descendente de audio de entrada (= 6 canales de mezcla descendente de audio de entrada) se reducen a 2,0 canales de mezcla descendente de audio adaptados (= 2 canales de mezcla descendente de audio adaptados).

O, por ejemplo, 2 canales de mezcla descendente de audio de entrada se reducen a 1 canal de mezcla descendente de audio adaptado.

Otras diversas combinaciones de canales de mezcla descendente de audio de entrada y canales de mezcla descendente de audio adaptados son posibles.

De acuerdo con una realización, la instancia de decodificador puede ser capaz de decodificar a lo sumo un número máximo de canales de mezcla descendente. La información de adaptación puede depender de dicho número máximo de canales de mezcla descendente. Además, el modificador de señal de mezcla descendente 110 puede estar configurado para adaptar los dos o más canales de mezcla descendente de audio de entrada dependiendo de la información de adaptación para obtener el uno o más canales de mezcla descendente de audio adaptados, de tal manera que el número del uno o más canales de mezcla descendente adaptados es igual a dicho número máximo de canales de mezcla descendente.

Por ejemplo, el modificador de señal de mezcla descendente 110 de la figura 2 convierte la mezcla descendente a la señal de audio que corresponde a la configuración de canal de salida soportada máxima de la instancia de codificador específica.

De acuerdo con la invención, la información de adaptación comprende una matriz de adaptación (\mathbf{D}_{dmx}^{DSM}).

El adaptador de información lateral paramétrica 120 puede, por ejemplo, adaptar la PSI para que corresponda a la mezcla descendente modificada, con el fin de disminuir la complejidad computacional del decodificador y para reducir el tamaño de flujo de bits de datos/velocidad de bits correspondiente sin producir una influencia negativa sobre la calidad de audio de salida del decodificador.

Por ejemplo, la PSIA 120 modifica el flujo de bits de PSI correspondiente sustituyendo la información que representa la matriz de mezcla descendente inicial por la información actualizada que describe la mezcla descendente resultante (tomando en cuenta las modificaciones de DSM) para corresponder a la especificación específica del decodificador.

Por ejemplo, un codificador de SAOC proporciona la señal de mezcla descendente estereofónica \mathbf{X}_{dmx}^{ENC} resultante de la aplicación de la matriz de mezcla descendente de codificador \mathbf{D}_{dmx}^{ENC} a las señales de objeto de audio de entrada S:

$$\mathbf{X}_{dmx}^{ENC} = \mathbf{D}_{dmx}^{ENC} \mathbf{S}.$$

De acuerdo con la invención, el modificador de señal de mezcla descendente 110 está configurado para adaptar, dependiendo de la matriz de adaptación \mathbf{D}_{dmx}^{DSM} , los dos o más canales de mezcla descendente de audio de entrada \mathbf{X}_{dmx}^{ENC} para obtener el uno o más canales de mezcla descendente de audio adaptados \mathbf{X}_{dmx}^{DSM} . En una realización, esto se realiza, por ejemplo, aplicando la fórmula $\mathbf{X}_{dmx}^{DSM} = \mathbf{D}_{dmx}^{DSM} \mathbf{X}_{dmx}^{ENC}$.

Por ejemplo, en una realización, donde se supone que la instancia de decodificador de SAOC específica soporta solamente una mezcla descendente monofónica (por ejemplo, perfil de bajo retardo de SAOC/nivel 1). En este caso, el DSM 110 convierte la mezcla descendente estéreo \mathbf{X}_{dmx}^{ENC} a la señal monofónica \mathbf{X}_{dmx}^{DSM} usando una matriz de mezcla descendente predefinida \mathbf{D}_{dmx}^{DSM} de la siguiente manera:

$$\mathbf{X}_{dmx}^{DSM} = \mathbf{D}_{dmx}^{DSM} \mathbf{X}_{dmx}^{ENC}.$$

De acuerdo con la invención, el adaptador de información lateral paramétrica 120 puede estar configurado para

adaptar, dependiendo de la matriz de adaptación \mathbf{D}_{dmx}^{DSM} , la información lateral paramétrica de entrada \mathbf{D}_{dmx}^{ENC} para obtener la información lateral paramétrica adaptada \mathbf{D}_{dmx}^{PSI} . En una realización, esto puede realizarse, por ejemplo, aplicando la fórmula:

$$\mathbf{D}_{dmx}^{PSI} = \mathbf{D}_{dmx}^{DSM} \mathbf{D}_{dmx}^{ENC}$$

Por ejemplo, de acuerdo con una realización, el PSIA 120 analiza el flujo de bits de PSI correspondiente; extrae información que describe la matriz de mezcla descendente \mathbf{D}_{dmx}^{ENC} ; sustituye estos datos por información actualizada que describe la nueva matriz de mezcla descendente \mathbf{D}_{dmx}^{PSI} :

$$\mathbf{D}_{dmx}^{PSI} = \mathbf{D}_{dmx}^{DSM} \mathbf{D}_{dmx}^{ENC}$$

Por lo tanto, de acuerdo con una realización, la información lateral paramétrica de entrada (\mathbf{D}_{dmx}^{enc}) puede indicar una matriz de mezcla descendente inicial, de tal manera que al aplicar la matriz de mezcla descendente inicial (\mathbf{D}_{dmx}^{enc}) sobre el uno o más objetos de audio (S), se obtienen los dos o más canales de mezcla descendente de

audio de entrada (\mathbf{X}_{dmx}^{enc}). El adaptador de información lateral paramétrica puede estar configurado para determinar una matriz de mezcla descendente adaptada (\mathbf{D}_{dmx}^{PSI}) como la información lateral paramétrica adaptada, de tal manera que al aplicar la matriz de mezcla descendente adaptada (\mathbf{D}_{dmx}^{PSI}) sobre el uno o más objetos de audio (S), se obtienen el uno o más canales de mezcla descendente de audio adaptados (\mathbf{X}_{dmx}^{DSM}).

En una realización, el PSIA formatea el nuevo flujo de bits modificado o hace pasar directamente estos parámetros al decodificador.

Este proceso de codificación y decodificación realizado por el PSIA también puede incluir la conversión de diferentes formatos de representación de matriz de mezcla descendente (por ejemplo, sistema de coordenadas polar a cartesiano, etc.).

Esta función descrita del PSIA puede resolver posibles problemas de compatibilidad y reducir el tamaño del flujo de bits correspondiente.

La figura 7 ilustra un aparato 700 para generar uno o más canales de audio a partir de información de audio de entrada que codifica uno o más objetos de audio de acuerdo con una realización.

El aparato 700 para generar el uno o más canales de audio comprende un aparato 710 de acuerdo con una de las realizaciones descritas anteriormente para adaptar la información de audio de entrada para obtener información de audio adaptada. La información de audio de entrada comprende dos o más canales de mezcla descendente de audio de entrada y comprende, además, información lateral paramétrica de entrada. La información de audio adaptada comprende uno o más canales de mezcla descendente de audio adaptados y comprende, además, información lateral paramétrica adaptada.

El aparato 710 de acuerdo con una de las realizaciones descritas anteriormente para adaptar la información de audio de entrada comprende un modificador de señal de mezcla descendente 110 y un adaptador de información lateral paramétrica 120.

Además, el aparato 700 para generar el uno o más canales de audio comprende una instancia de decodificador 720, para decodificar, dependiendo de la información lateral paramétrica adaptada, el uno o más canales de mezcla descendente de audio adaptados para obtener el uno o más canales de audio.

De acuerdo con una realización, el adaptador de información lateral paramétrica 120 del aparato 710 para adaptar la información de audio de entrada puede estar configurado para recibir un flujo de bits de entrada que comprende la información lateral paramétrica de entrada. El adaptador de información lateral paramétrica 120 del aparato 710 para adaptar la información de audio de entrada puede estar configurado para adaptar la información lateral paramétrica de entrada para obtener la información lateral paramétrica adaptada y para introducir la información lateral paramétrica adaptada en la instancia de decodificador 720. La instancia de decodificador 720 puede estar configurada para

decodificar el uno o más canales de mezcla descendente de audio adaptados dependiendo de la información lateral paramétrica adaptada.

En otra realización, el adaptador de información lateral paramétrica 120 del aparato 710 para adaptar la información de audio de entrada puede estar configurado para recibir un flujo de bits de entrada que comprende la información lateral paramétrica de entrada. El adaptador de información lateral paramétrica 120 del aparato 710 para adaptar la información de audio de entrada puede estar configurado para sustituir la información lateral paramétrica de entrada dentro del flujo de bits de entrada por la información lateral paramétrica adaptada para obtener un flujo de bits modificado. El adaptador de información lateral paramétrica 120 del aparato 710 para adaptar la información de audio de entrada puede estar configurado para introducir el flujo de bits modificado en la instancia de decodificador 720. Además, la instancia de decodificador 720 puede estar configurada para decodificar el uno o más canales de mezcla descendente de audio adaptados dependiendo de la corriente de bits modificada.

Las figuras 8 y 9 ilustran dos posibilidades para incorporar el aparato para adaptar información de audio de entrada a la cadena de procesamiento de decodificación.

En particular, la figura 8 ilustra una aplicación de PSIA conjunta dentro de un esquema de codificación/decodificación de acuerdo con una realización.

La figura 8 ilustra una pluralidad de aparatos 800, 801, 802 para generar uno o más canales de audio a partir de información de audio de entrada que codifica uno o más objetos de audio, comprendiendo el aparato 800 para generar uno o más canales de audio un aparato 810 para adaptar información de audio de entrada y una instancia de decodificador 820, comprendiendo el aparato 801 para generar uno o más canales de audio un aparato 811 para adaptar información de audio de entrada y una instancia de decodificador 821 y comprendiendo el aparato 802 para generar uno o más canales de audio un aparato 812 para adaptar información de audio de entrada y una instancia de decodificador 822. Cabe señalar que, por ejemplo, el aparato 800 para generar uno o más canales de audio, que comprende el aparato 810 para adaptar información de audio de entrada y la instancia de decodificador 820, no tiene que realizarse como una sola unidad de hardware 800, sino que en lugar de esto puede realizarse por dos unidades separadas 810, 820 que se conectan por un alambre o que se conectan de manera inalámbrica.

La implementación conjunta (integrada) del aparato para adaptar información de audio de entrada puede realizarse con el fin de reducir la complejidad computacional para la decodificación (véase la figura 8). Además, esto permite implementar una interfaz no cuantificada (no codificada) entre el aparato para adaptar información de audio de entrada y el decodificador. Esto puede ser relevante, en particular para dispositivos de aplicaciones móviles para reducir el consumo de energía.

La figura 9 ilustra la aplicación PSIA disjunta en un esquema de codificación/decodificación de acuerdo con una realización.

En particular, la figura 9 ilustra una pluralidad de aparatos 900, 901, 902 para generar uno o más canales de audio a partir de información de audio de entrada que codifica uno o más objetos de audio, comprendiendo el aparato 900 para generar uno o más canales de audio un aparato 910 para adaptar información de audio de entrada y una instancia de decodificador 920, comprendiendo el aparato 901 para generar uno o más canales de audio un aparato 911 para adaptar información de audio de entrada y una instancia de decodificador 921 y comprendiendo el aparato 902 para generar uno o más canales de audio un aparato 912 para adaptar información de audio de entrada y una instancia de decodificador 922. Cabe señalar que, por ejemplo, el aparato 900 para generar uno o más canales de audio, que comprende el aparato 910 para adaptar información de audio de entrada y la instancia del decodificador 920, no tiene que realizarse como una sola unidad de hardware 900, sino que puede realizarse por dos unidades separadas 910, 920 que se conectan por un alambre o que se conectan de manera inalámbrica.

La implementación disjunta (separada) del aparato para adaptar información de audio de entrada puede realizarse con el fin de reducir el tamaño de flujo de bits de datos/velocidad de bits correspondiente, véase la figura 9. Esto puede ser relevante, en particular, para dispositivos de aplicación móvil con capacidad de almacenamiento y transmisión limitada y sistemas de unidad de control de multipuntos (MCU) con canales de transición de datos estrechos.

Aunque algunos aspectos se han descrito en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del método correspondiente, correspondiendo un bloque o dispositivo a una etapa de método o una característica de una etapa de método. Análogamente, los aspectos descritos en el contexto de una etapa de método también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente.

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

Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo, un disquete, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control legibles electrónicamente almacenadas en el mismo, que cooperan (o son capaces de cooperar) con un sistema informático programable, de tal manera que se realiza el método respectivo.

Algunas realizaciones de acuerdo con la invención comprenden un portador de datos no transitorio que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable, de tal manera que se realiza uno de los métodos descritos en el presente documento.

En general, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, siendo el código de programa operativo para realizar uno de los métodos, cuando el producto de programa informático se ejecuta en un ordenador. El código del programa puede almacenarse, por ejemplo, en un portador legible por máquina.

Otras realizaciones comprenden el programa informático para realizar uno de los métodos descritos en el presente documento, almacenado en un portador legible por máquina.

En otras palabras, una realización del método de la invención es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los métodos descritos en el presente documento, cuando el programa informático se ejecuta en un ordenador.

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

Una realización adicional del método de la invención es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para realizar uno de los métodos descritos en el presente documento. El flujo de datos o la secuencia de señales pueden, por ejemplo, configurarse para transferirse a través de una conexión de comunicación de datos, por ejemplo, a través de internet.

Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador o un dispositivo lógico programable, configurado o adaptado para realizar uno de los métodos descritos en el presente documento.

Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para realizar uno de los métodos descritos en el presente documento.

En algunas realizaciones, un dispositivo lógico programable (por ejemplo, una disposición de compuertas programable en el campo) puede usarse para realizar algunas o todas las funcionalidades de los métodos descritos en el presente documento. En algunas realizaciones, una disposición de compuertas programable en el campo puede cooperar con un microprocesador con el fin de realizar uno de los métodos descritos en el presente documento. En general, los métodos se realizan, preferentemente, por cualquier aparato de hardware.

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

Referencias

[MPS] ISO/IEC 23003-1:2007, MPEG-D (tecnologías de audio MPEG), Parte 1: MPEG Surround, 2007.

[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, Nov. 2003.

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

[SAOC1] J. Herre, S. Disch, J. Hilpert, O. Hellmuth: " From SAC To SAOC - Recent Developments in Parametric Coding of Spatial Audio", 22ª conferencia de AES UK Regional, Cambridge, Reino Unido, abril 2007.

[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 Audio Object Coding (SAOC) - The Upcoming MPEG Standard on

Parametric Object Based Audio Coding”, 124ª Convención de AES, Amsterdam 2008.

[SAOC] ISO/IEC, “MPEG audio technologies - Part 2: Spatial Audio Object Coding (SAOC)”, norma internacional 23003-2 ISO/IEC JTC1/SC29/WG11 (MPEG).

5 [ISS1] M. Parvaix y L. Girin: “Informed Source Separation of underdetermined instantaneous Stereo Mixtures using Source Index Embedding”, IEEE ICASSP, 2010

10 [ISS2] M. Parvaix, L. Girin, J.-M. Brossier: “A watermarking-based method for informed source separation of audio signals with a single sensor”, operaciones del IEEE en procesamiento de audio, habla y lenguaje, 2010.

[ISS3] A. Liutkus, J. Pinel, R. Badeau, L. Girin y G. Richard: “Informed source separation through spectrogram coding and data embedding”, Signal Processing Journal, 2011.

15 [ISS4] A. Ozerov, A. Liutkus, R. Badeau, G. Richard: “Informed source separation: source coding meets source separation”, Taller del IEEE en cuanto a aplicaciones de procesamiento de Señal a audio y acústica, 2011.

20 [ISS5] Shuhua Zhang y Laurent Girin: “An Informed Source Separation System for Speech Signals”, INTERSPEECH, 2011.

[ISS6] L. Girin y J. Pinel: “Informed Audio Source Separation from Compressed Linear Stereo Mixtures”, AES 42ª Conferencia internacional: Semántica de audio, 2011.

REIVINDICACIONES

1. Un aparato para adaptar información de audio de entrada, que codifica uno o más objetos de audio, para obtener información de audio adaptada, en el que la información de audio de entrada comprende dos o más canales de mezcla descendente de audio de entrada y comprende, además, información lateral paramétrica de entrada, en el que la información de audio adaptada comprende uno o más canales de mezcla descendente de audio adaptados y comprende, además, información lateral paramétrica adaptada, en el que el aparato comprende:
 - un modificador de señal de mezcla descendente (110) para adaptar, dependiendo de la información de adaptación, los dos o más canales de mezcla descendente de audio de entrada para obtener el uno o más canales de mezcla descendente de audio adaptados, y
 - un adaptador de información lateral paramétrica (120) para adaptar, dependiendo de la información de adaptación, la información lateral paramétrica de entrada para obtener la información lateral paramétrica adaptada,
- en el que la información de adaptación comprende una matriz de adaptación (\mathbf{D}_{dmx}^{DSM}), en el que el modificador de señal de mezcla descendente (110) está configurado para adaptar, dependiendo de la matriz de adaptación (\mathbf{D}_{dmx}^{DSM}), los dos o más canales de mezcla descendente de audio de entrada (\mathbf{X}_{dmx}^{ENC}) para obtener el uno o más canales de mezcla descendente de audio adaptados (\mathbf{X}_{dmx}^{DSM}), en el que el adaptador de información lateral paramétrica (120) está configurado para adaptar, dependiendo de la matriz de adaptación (\mathbf{D}_{dmx}^{DSM}), la información lateral paramétrica de entrada (\mathbf{D}_{dmx}^{ENC}) para obtener la información lateral paramétrica adaptada (\mathbf{D}_{dmx}^{PSI}).
2. Un aparato de acuerdo con la reivindicación 1, en el que la información lateral paramétrica de entrada (\mathbf{D}_{dmx}^{enc}) indica una matriz de mezcla descendente inicial, de tal manera que aplicando la matriz de mezcla descendente inicial (\mathbf{D}_{dmx}^{enc}) sobre el uno o más objetos de audio (S), se obtienen los dos o más canales de mezcla descendente de audio de entrada (\mathbf{X}_{dmx}^{enc}), y en el que el adaptador de información lateral paramétrica (120) está configurado para determinar una matriz de mezcla descendente adaptada (\mathbf{D}_{dmx}^{PSI}) como la información lateral paramétrica adaptada, de tal manera que aplicando la matriz de mezcla descendente adaptada (\mathbf{D}_{dmx}^{PSI}) sobre el uno o más objetos de audio (S), se obtienen el uno o más canales de mezcla descendente de audio adaptados (\mathbf{X}_{dmx}^{DSM}).
3. Un aparato de acuerdo con la reivindicación 1 o 2, en el que el modificador de señal de mezcla descendente (110) está configurado para adaptar los dos o más canales de mezcla descendente de audio de entrada dependiendo de la información de adaptación, de tal manera que el número del uno o más canales de mezcla descendente de audio adaptados es menor que el número de los dos o más canales de mezcla descendente de audio de entrada.
4. Un aparato de acuerdo con una de las reivindicaciones 1 a 3, en el que la información de adaptación depende de una instancia de decodificador y en el que el modificador de señal de mezcla descendente (110) está configurado para adaptar los dos o más canales de mezcla descendente de audio de entrada dependiendo de la instancia de decodificador.
5. Un aparato de acuerdo con la reivindicación 4, en el que la instancia de decodificador es capaz de decodificar a lo sumo un número máximo de canales de mezcla descendente, en el que la información de adaptación depende de dicho número máximo de canales de mezcla descendente, y en el que el modificador de señal de mezcla descendente (110) está configurado para adaptar los dos o más canales de mezcla descendente de audio de entrada dependiendo de la información de adaptación para obtener el uno o más canales de mezcla descendente de audio adaptados, de tal manera que el número del uno o más canales de mezcla descendente adaptados es igual a dicho número máximo de canales de mezcla descendente.
6. Un aparato de acuerdo con una de las reivindicaciones anteriores, en el que el modificador de señal de mezcla descendente (110) está configurado para adaptar, dependiendo de la matriz de adaptación \mathbf{D}_{dmx}^{DSM} , los dos o más

canales de mezcla descendente de audio de entrada \mathbf{X}_{dmx}^{ENC} para obtener el uno o más canales de mezcla descendente de audio adaptados \mathbf{X}_{dmx}^{DSM} aplicando la fórmula:

$$\mathbf{X}_{dmx}^{DSM} = \mathbf{D}_{dmx}^{DSM} \mathbf{X}_{dmx}^{ENC}$$

7. Un aparato de acuerdo con una de las reivindicaciones anteriores, en el que el adaptador de información lateral paramétrica (120) está configurado para adaptar, dependiendo de la matriz de adaptación \mathbf{D}_{dmx}^{DSM} , la información lateral paramétrica de entrada \mathbf{D}_{dmx}^{ENC} para obtener la información lateral paramétrica adaptada \mathbf{D}_{dmx}^{PSI} aplicando la fórmula:

$$\mathbf{D}_{dmx}^{PSI} = \mathbf{D}_{dmx}^{DSM} \mathbf{D}_{dmx}^{ENC}$$

8. Un aparato (700; 800, 801, 802; 900, 901, 902) para generar uno o más canales de audio a partir de una información de audio de entrada que codifica uno o más objetos de audio, en el que el aparato comprende:

un aparato (710; 810, 811, 812; 910, 911, 912) de acuerdo con una de las reivindicaciones 1 a 6 para adaptar la información de audio de entrada para obtener información de audio adaptada, en el que la información de audio de entrada comprende dos o más canales de mezcla descendente de audio de entrada y comprende, además, información lateral paramétrica de entrada, en el que la información de audio adaptada comprende uno o más canales de mezcla descendente de audio adaptados y comprende, además, información lateral paramétrica adaptada, y
una instancia de decodificador (720; 820, 821, 822; 920, 921, 922) para decodificar, dependiendo de la información lateral paramétrica adaptada, el uno o más canales de mezcla descendente de audio adaptados para obtener el uno o más canales de audio.

9. Un aparato (700; 800, 801, 802) de acuerdo con la reivindicación 8, en el que el adaptador de información lateral paramétrica (120) del aparato (710; 810, 811, 812) de acuerdo con una de las reivindicaciones 1 a 7 está configurado para recibir un flujo de bits de entrada que comprende la información lateral paramétrica de entrada, en el que el adaptador de información lateral paramétrica (120) del aparato (710; 810, 811, 812) de acuerdo con una de las reivindicaciones 1 a 7 está configurado para adaptar la información lateral paramétrica de entrada para obtener la información lateral paramétrica adaptada, y para introducir la información lateral paramétrica adaptada en la instancia de decodificador (720; 820, 821, 822) y en el que la instancia de decodificador (720; 820, 821, 822) está configurada para decodificar el uno o más canales de mezcla descendente de audio adaptados dependiendo de la información lateral paramétrica adaptada.

10. Un aparato (700; 900, 901, 902) de acuerdo con la reivindicación 8, en el que el adaptador de información lateral paramétrica (120) del aparato (710; 910, 911, 912) de acuerdo con una de las reivindicaciones 1 a 7 está configurado para recibir un flujo de bits de entrada que comprende la información lateral paramétrica de entrada, en el que el adaptador de información lateral paramétrica (120) del aparato (710; 910, 911, 912) de acuerdo con una de las reivindicaciones 1 a 7 está configurado para sustituir la información lateral paramétrica de entrada dentro del flujo de bits de entrada por la información lateral paramétrica adaptada para obtener un flujo de bits modificado, en el que el adaptador de información lateral paramétrica (120) del aparato (710; 910, 911, 912) de acuerdo con una de las reivindicaciones 1 a 7 está configurado para introducir el flujo de bits modificado en la instancia de decodificador (720; 920, 921, 922) y en el que la instancia de decodificador (720; 920, 921, 922) está configurada para decodificar el uno o más canales de mezcla descendente de audio adaptados dependiendo del flujo de bits modificado.

11. Un método para adaptar información de audio de entrada, que codifica uno o más objetos de audio, para obtener información de audio adaptada, en el que la información de audio de entrada comprende dos o más canales de mezcla descendente de audio de entrada y comprende, además, información lateral paramétrica de entrada, en el que la información de audio adaptada comprende uno o más canales de mezcla descendente de audio adaptados y comprende, además, información lateral paramétrica adaptada, en el que el método comprende:

adaptar, dependiendo de la información de adaptación, los dos o más canales de mezcla descendente de audio de entrada para obtener el uno o más canales de mezcla descendente de audio adaptados y
adaptar, dependiendo de la información de adaptación, la información lateral paramétrica de entrada para

- obtener la información lateral paramétrica adaptada,
 en el que la información de adaptación comprende una matriz de adaptación (\mathbf{D}_{dmx}^{DSM}),
 en el que la etapa de adaptar los dos o más canales de mezcla descendente de audio de entrada comprende
 adaptar, dependiendo de la matriz de adaptación (\mathbf{D}_{dmx}^{DSM}), los dos o más canales de mezcla descendente de audio
 5 de entrada (\mathbf{X}_{dmx}^{ENC}) para obtener el uno o más canales de mezcla descendente de audio adaptados (\mathbf{X}_{dmx}^{DSM}),
 en el que la etapa de adaptar la información lateral paramétrica de entrada comprende adaptar, dependiendo de la
 matriz de adaptación (\mathbf{D}_{dmx}^{DSM}), la información lateral paramétrica de entrada (\mathbf{D}_{dmx}^{ENC}) para obtener la información
 lateral paramétrica adaptada (\mathbf{D}_{dmx}^{PSI}).
- 10 12. Un método de acuerdo con la reivindicación 11,
 en el que la información lateral paramétrica de entrada (\mathbf{D}_{dmx}^{enc}) indica una matriz de mezcla descendente inicial, de
 tal manera que aplicando la matriz de mezcla descendente inicial (\mathbf{D}_{dmx}^{enc}) sobre el uno o más objetos de audio (S),
 se obtienen los dos o más canales de mezcla descendente de audio de entrada (\mathbf{X}_{dmx}^{enc}) y
 en el que la etapa de adaptar la información lateral paramétrica de entrada comprende determinar una matriz de
 15 mezcla descendente adaptada (\mathbf{D}_{dmx}^{PSI}) como la información lateral paramétrica adaptada, de tal manera que
 aplicando la matriz de mezcla descendente adaptada (\mathbf{D}_{dmx}^{PSI}) sobre el uno o más objetos de audio (S), se obtienen
 el uno o más canales de mezcla descendente de audio adaptados (\mathbf{X}_{dmx}^{DSM}).
- 20 13. Un programa informático configurado para implementar el método de la reivindicación 11 o 12 cuando se ejecuta
 por un ordenador o un procesador de señales.

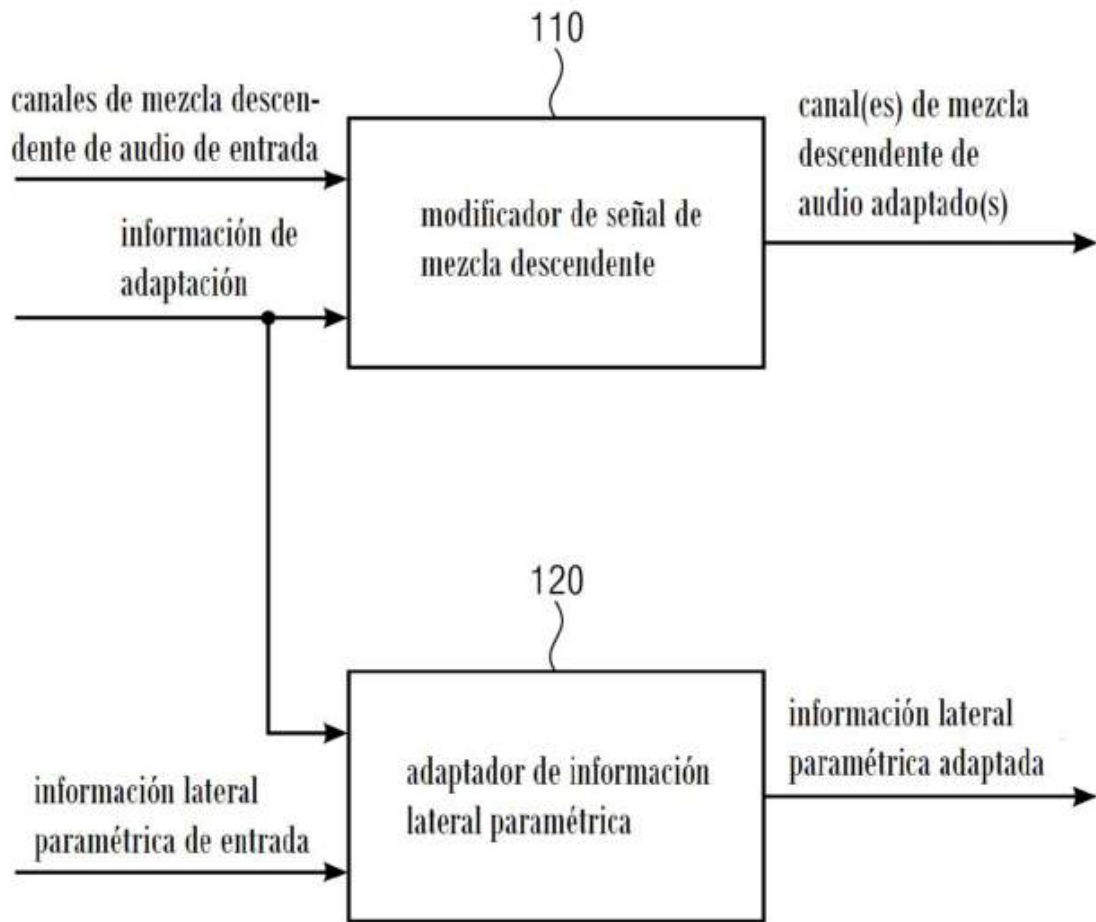


FIGURA 1

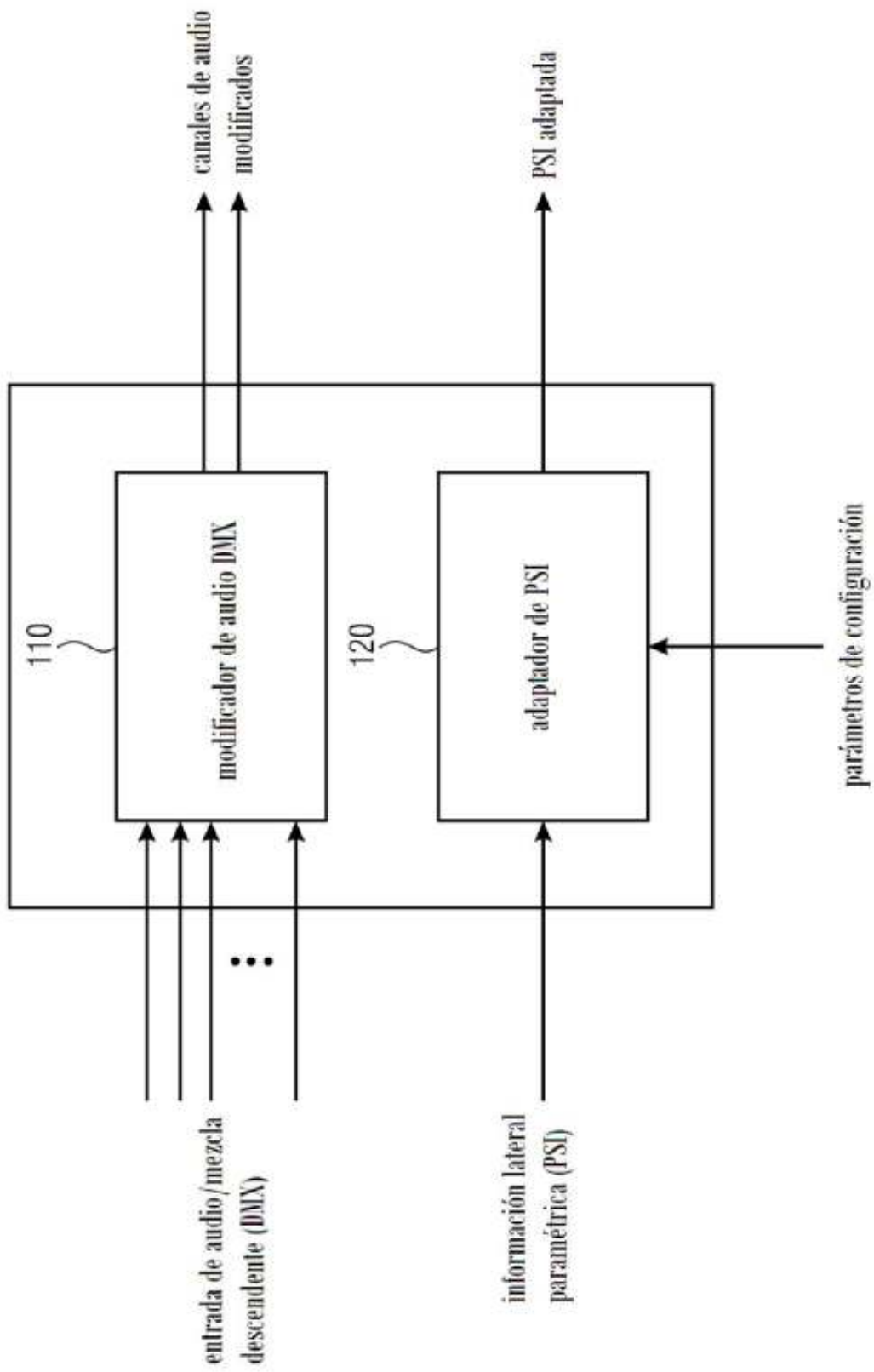


FIGURA 2

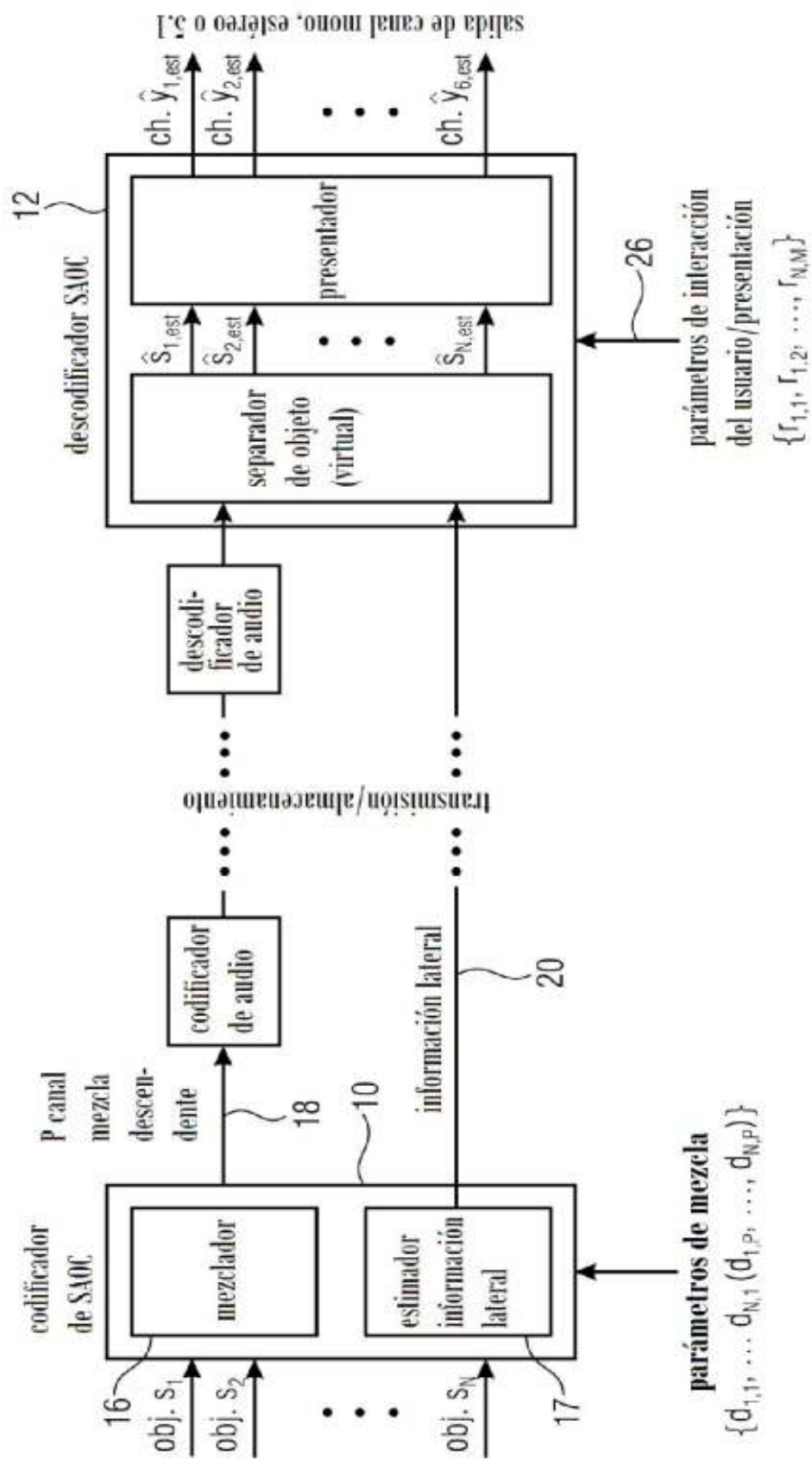


FIGURA 3

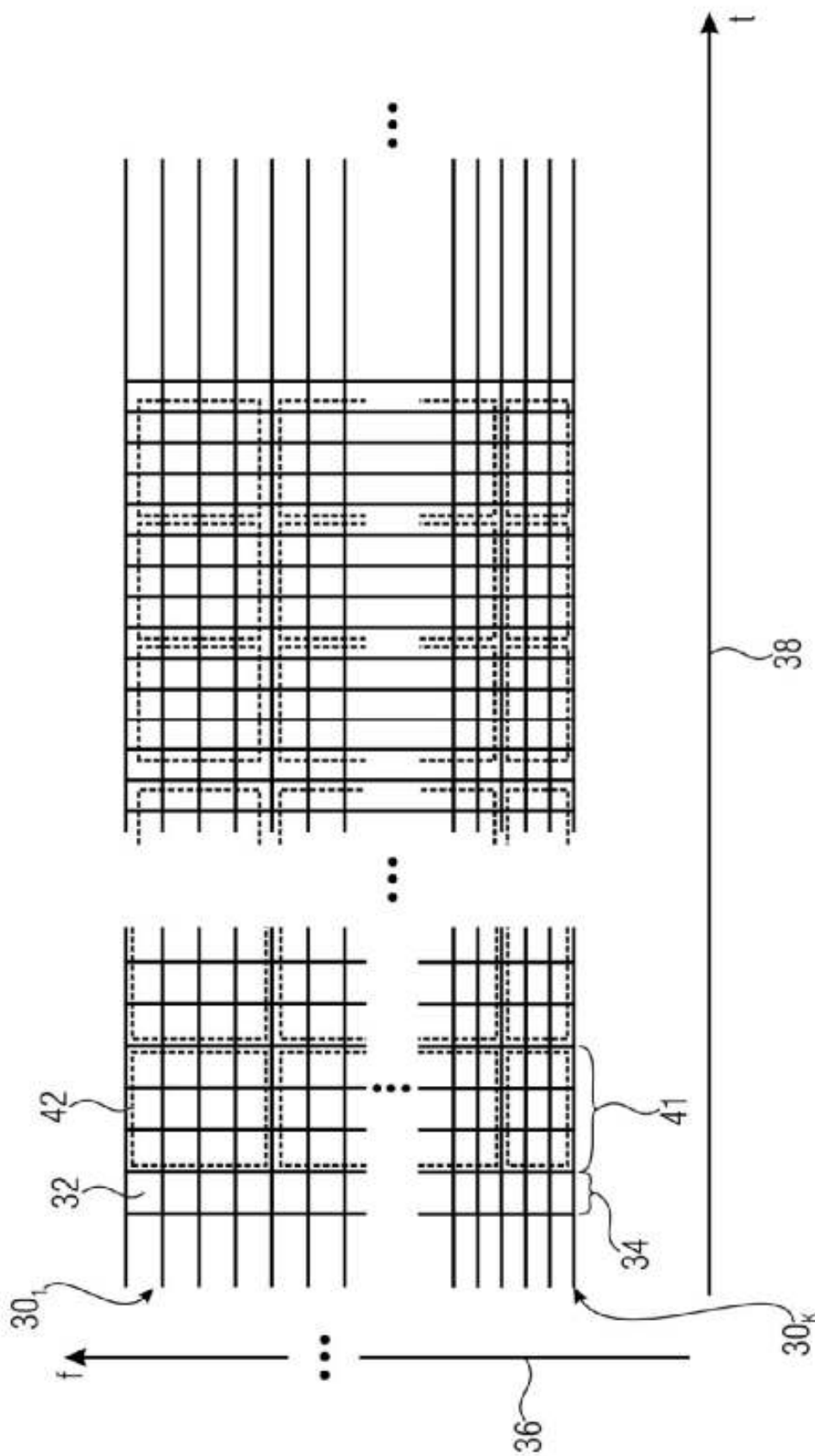


FIGURA 4

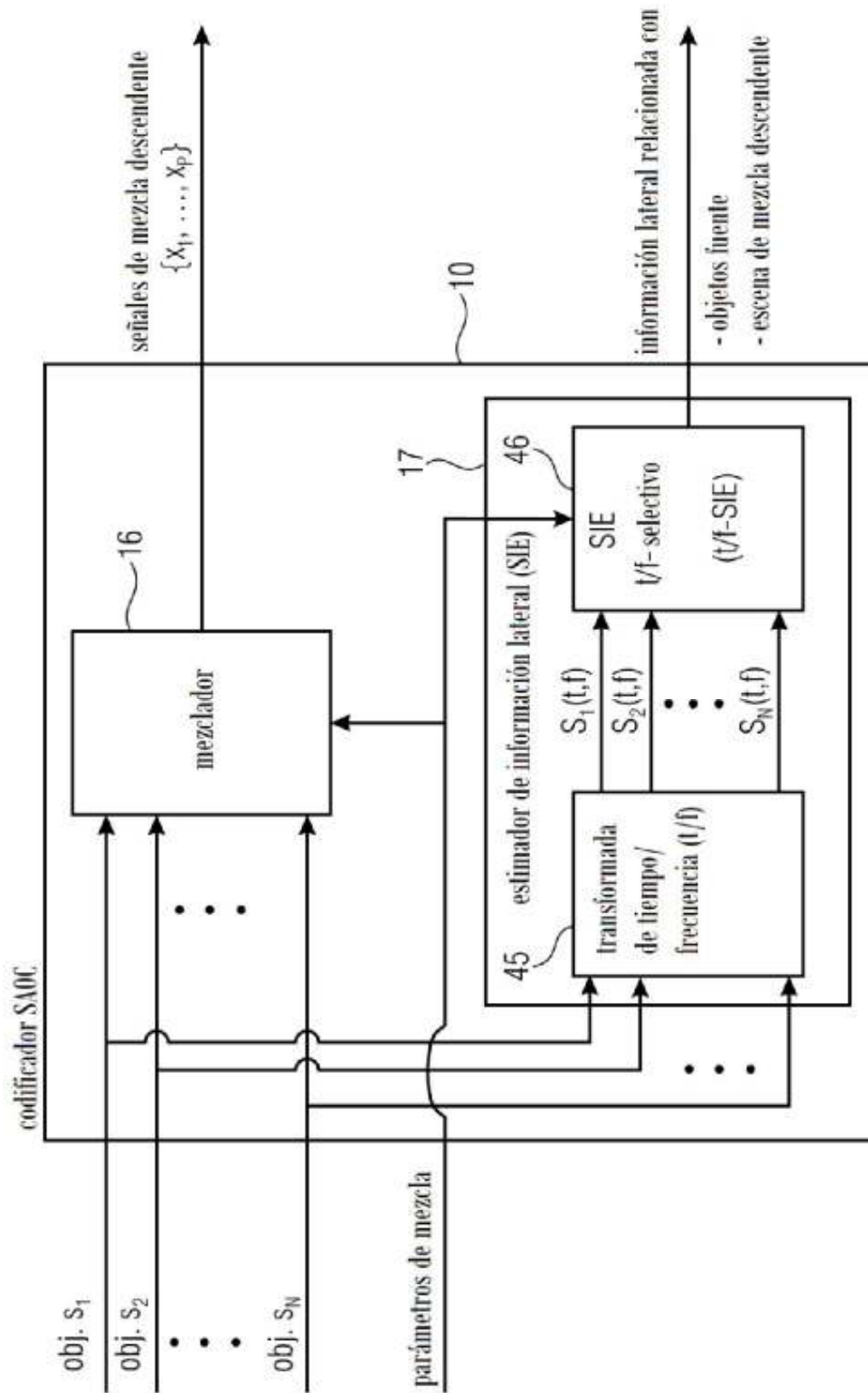


FIGURA 5

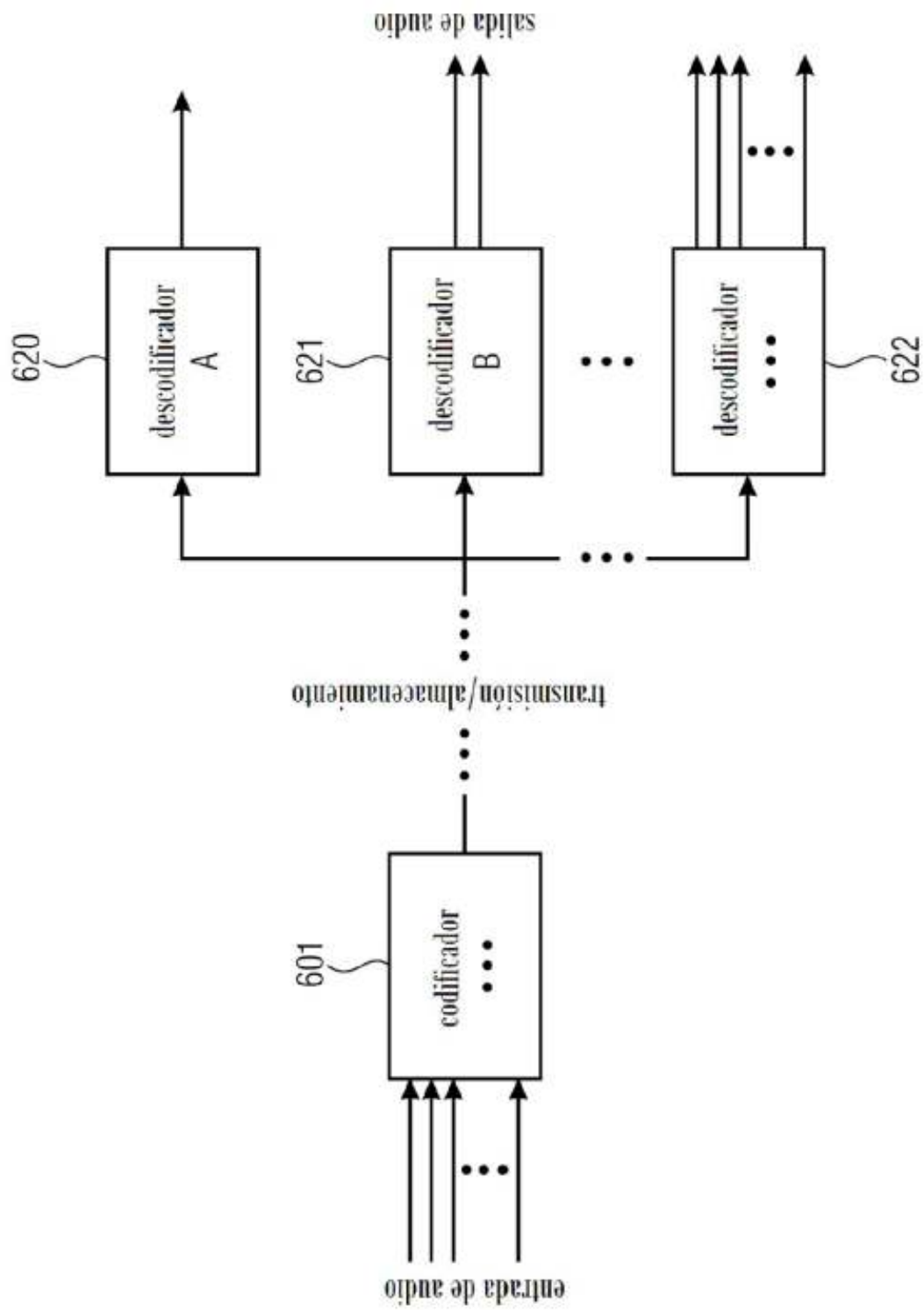


FIGURA 6

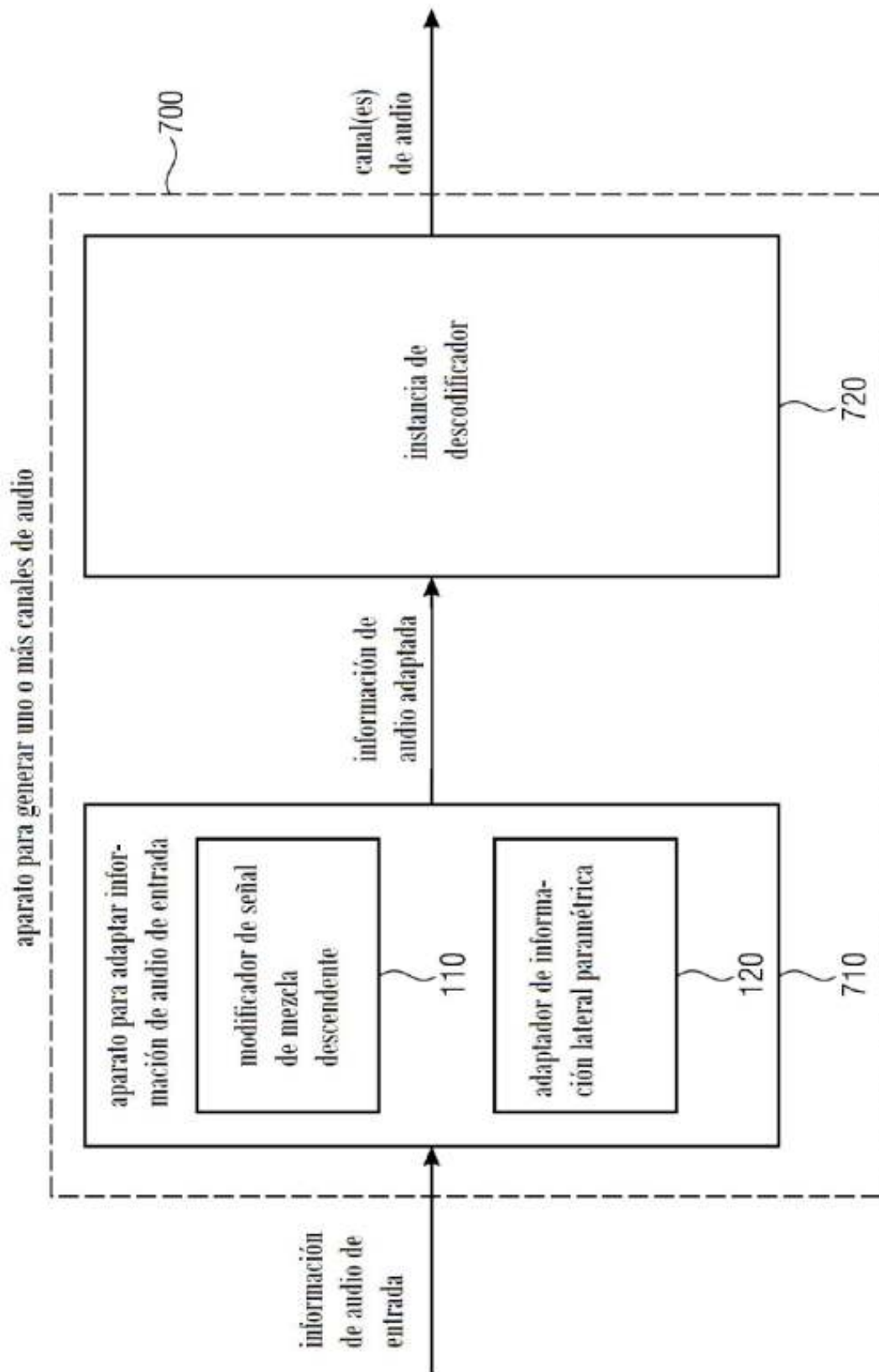


FIGURA 7

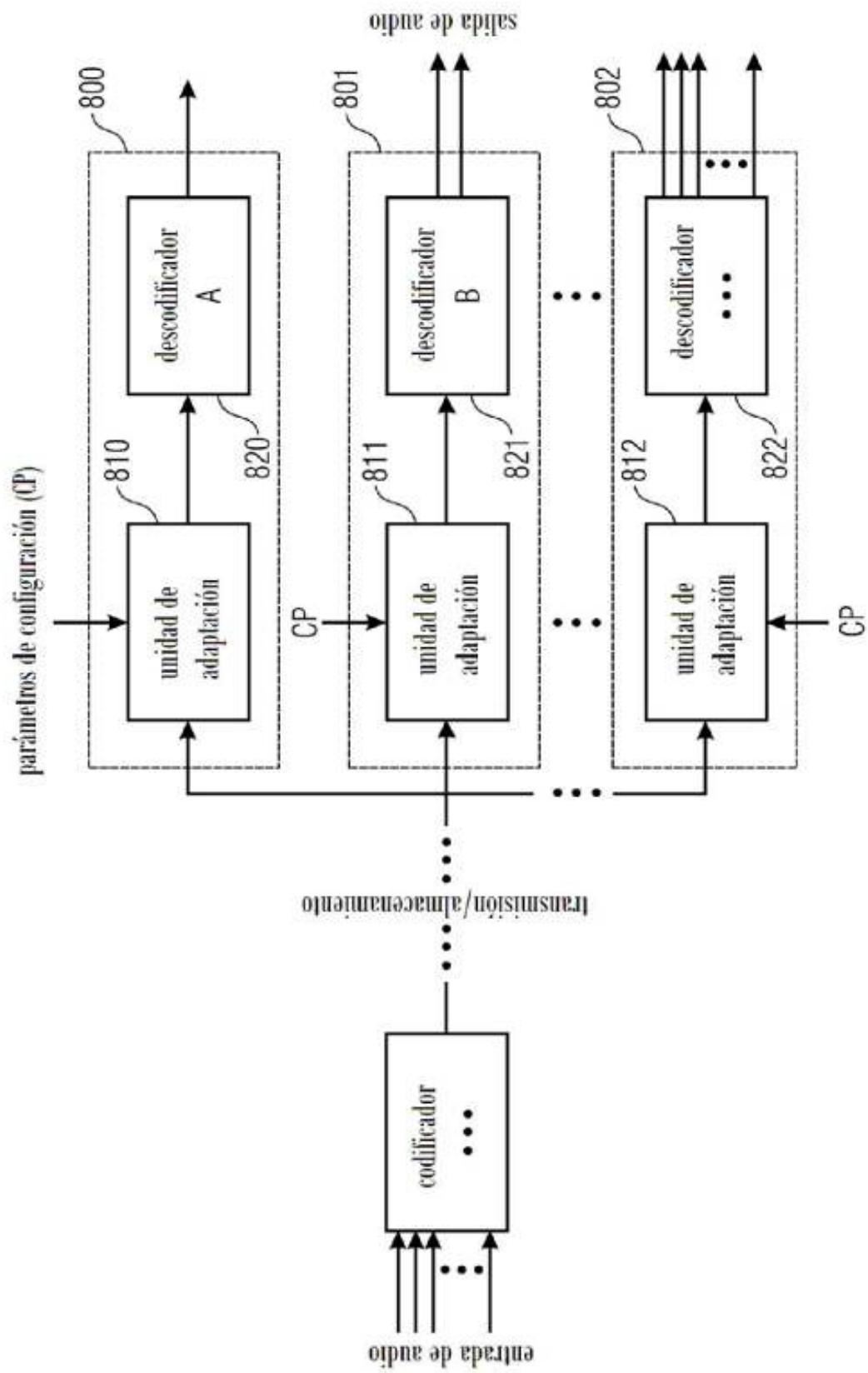


FIGURA 8

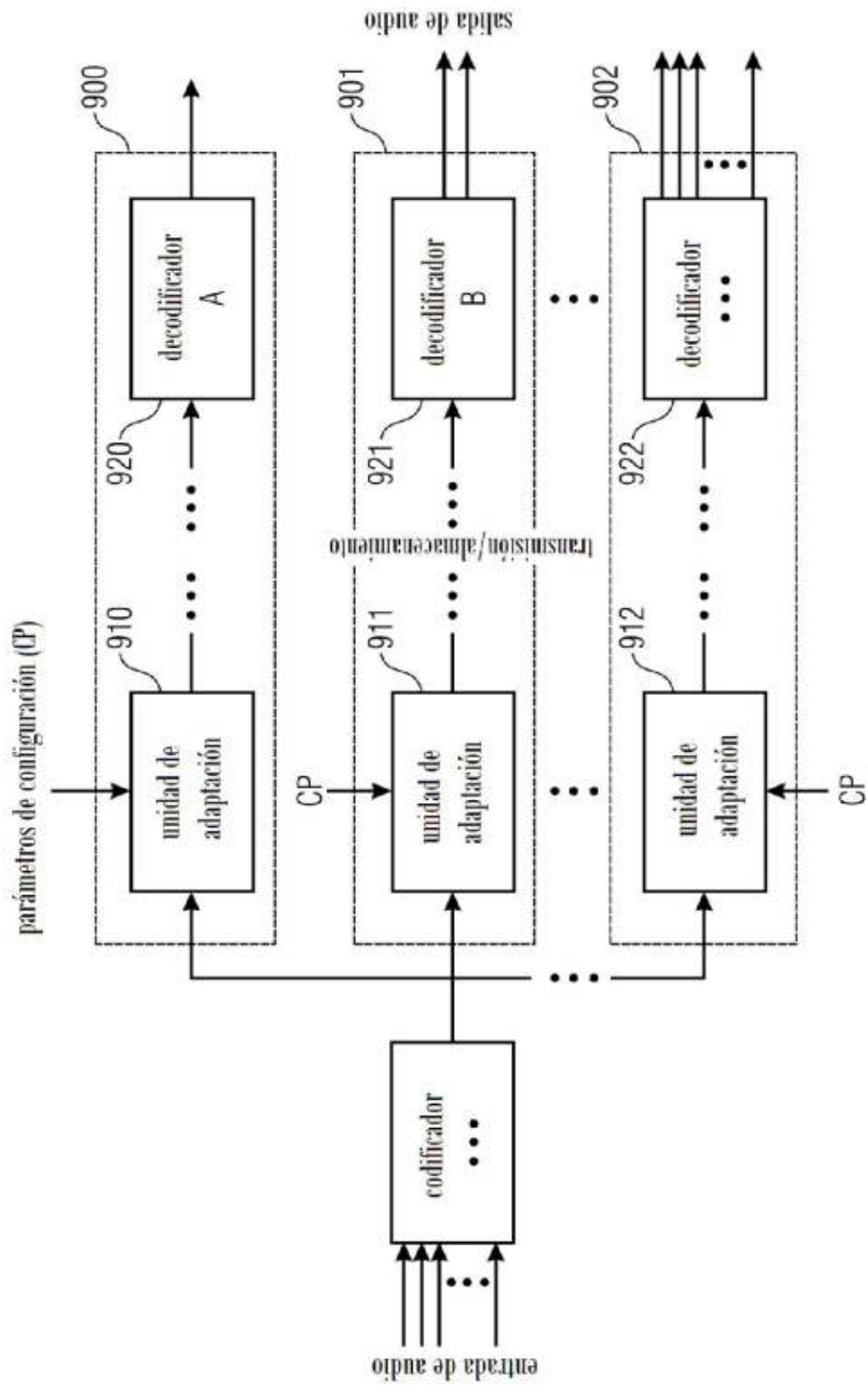


FIGURA 9