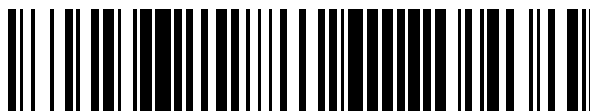


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 703 327**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.10.2013 PCT/EP2013/070533**

87 Fecha y número de publicación internacional: **10.04.2014 WO14053537**

96 Fecha de presentación y número de la solicitud europea: **02.10.2013 E 13773699 (7)**

97 Fecha y número de publicación de la concesión europea: **31.10.2018 EP 2904609**

54 Título: **Codificador, decodificador y métodos para codificación espacial de objetos de audio de multiresolución retrocompatible**

30 Prioridad:

05.10.2012 US 201261710128 P
13.05.2013 EP 13167485

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
08.03.2019

73 Titular/es:

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

72 Inventor/es:

DISCH, SASCHA;
FUCHS, HARALD;
PAULUS, JOUNI;
TERENTIV, LEON;
HELLMUTH, OLIVER y
HERRE, JÜRGEN

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 703 327 T3

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

DESCRIPCIÓN

Codificador, decodificador y métodos para codificación espacial de objetos de audio de multirresolución retrocompatible

5

Campo de la invención

La presente invención se refiere a la codificación de señales de audio, decodificación de señales de audio y procesamiento de señales de audio, y en particular, a un codificador, un decodificador y métodos para la codificación espacial de objetos de audio (SAOC) de multirresolución retrocompatible.

10

Antecedentes de la invención

En los sistemas de audio digital modernos, existe una tendencia principal para permitir modificaciones relacionadas con objetos 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 reposicionamiento espacial de los objetos de audio dedicados en caso de reproducción multicanal mediante altavoces espacialmente distribuidos. Esto puede lograrse distribuyendo de manera individual diferentes partes del contenido de audio a los diferentes altavoces.

15

En otras palabras, en la técnica de procesamiento de audio, transmisión de audio y almacenamiento de audio existe un deseo cada vez mayor de permitir la interacción del usuario en la reproducción del contenido de audio orientado a objetos y también una demanda para utilizar las posibilidades extendidas de la reproducción multicanal para presentar de manera individual contenidos de audio o partes del mismo para mejorar la impresión de la audición. Mediante esto, el uso de contenido de audio multicanal trae consigo mejoras significativas al usuario. Por ejemplo, puede obtenerse una impresión de audición tridimensional, que trae consigo una satisfacción de usuario mejorada en aplicaciones de entretenimiento. Sin embargo, el contenido de audio multicanal también es útil en ambientes profesionales, por ejemplo, en aplicaciones de conferencia telefónica, debido a que la inteligibilidad del hablante puede mejorarse utilizando una reproducción de audio multicanal. Otra aplicación posible es ofrecer a un oyente una pieza musical para ajustar individualmente el nivel de reproducción y/o la posición espacial de diferentes partes (también denominadas como "objetos de audio") o pistas, tal como una parte vocal o diferentes instrumentos. El usuario puede realizar tal ajuste por razones de gusto personal para transcripción más fácil de una o más partes de la pieza musical, propósitos educativos, karaoke, ensayo, etc.

20

25

30

La transmisión discreta directa de todo el contenido de audio multicanal o multiobjeto digital, por ejemplo, en forma de datos de modulación por impulsos codificados (PCM) o incluso formatos de audio comprimido, exige tasas de bits muy altas. Sin embargo, también es deseable transmitir y almacenar datos de audio de una forma eficiente de tasas de bits. Por lo tanto, se está dispuesto a aceptar un intercambio razonable entre calidad de audio y requisitos de tasa de bits para evitar una carga de recursos excesiva provocada por aplicaciones multicanal/multiobjeto.

35

Recientemente, en el campo de codificación de audio, se han introducido técnicas para métricas para la transmisión/almacenamiento eficiente de tasa de bits de señales de audio multicanal/multiobjeto, por ejemplo, por el Grupo de Expertos en Imágenes en Movimiento (MPEG) y otros. Un ejemplo es MPEG Envolvente (MPS) como un enfoque orientado a canal [MPS, BCC], o Codificación Espacial de Objetos de Audio (SAOC) de MPEG, como un enfoque orientado a objetos [JSC, SAOC, SAOC1, SAOC2]. Otro enfoque orientado objeto se denomina como "separación de fuente 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 en función de una mezcla descendente de canales/objetos y la información secundaria adicional que describe la escena de audio transmitida/almacenada y/o los objetos de fuente de audio en la escena de audio.

40

45

La estimación y la aplicación de información secundaria relacionada con canal/objeto en tales sistemas se hace de una forma selectiva de 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 filtro similares a bancos de Filtro de Espejo en Cuadratura (QMF), etc. El principio básico de tales sistemas se representa en la Figura 4, que utiliza el ejemplo de SAOC de MPEG.

50

55

En el caso de STFT, la dimensión temporal se representa por el número de bloque de tiempo y la dimensión espectral se captura por el número de coeficiente espectral ("contenedor"). En el caso de QMF, la dimensión temporal se representa por el número de intervalo de tiempo y la dimensión espectral se captura por el número de subbanda. Si la resolución espectral de QMF se mejora por la aplicación subsiguiente de una segunda fase de filtro, todo el banco de filtro se denomina QMF híbrido y las subbandas de resolución finas se denominan como subbandas híbridas.

60

Como ya se menciona anteriormente, en SAOC, el procesamiento general se lleva a cabo de una forma selectiva de tiempo-frecuencia y puede describirse como sigue dentro de cada banda de frecuencia:

- 5 - N señales de objeto de audio de entrada $s_1...s_N$ se mezclan de forma descendente en P canales $x_1...x_P$ como parte del procesamiento de codificador que utiliza una matriz de mezclado descendente que consiste en los elementos $d_{1,1}...d_{N,P}$. Además, el codificador extrae la información secundaria que describe las características de los objetos de audio de entrada (módulo de Estimador de Información Secundaria (SIE)). Para SAOC de MPEG, las relaciones de las potencias de objeto con respecto a cada una tienen la forma más básica de tal información secundaria.
- 10 - Señales de mezcla descendente y la información secundaria se transmiten/almacenan. Para este fin, las señales de audio de mezcla descendente pueden comprimirse, por ejemplo, utilizando codificadores de audio perceptivos bien conocidos tales como Capa II o III de MPEG-1/2 (también conocida como .mp3), Codificación Avanzada de Audio de MPEG-2/4 (AAC) etc.
- 15 - En el extremo de recepción, el decodificador trata de forma conceptual de restablecer las señales de objeto original ("separación de objeto") de las señales de mezcla descendente (decodificada) utilizando la información secundaria transmitida. Estas señales de objeto aproximadas $\hat{s}_1... \hat{s}_N$ se mezclan a continuación en una escena objetivo representada por M canales de salida de audio $\hat{y}_1... \hat{y}_M$ utilizando una matriz de conversión descrita por los coeficientes $r_{1,1}...r_{N,M}$ en la Figura 4. La escena objetivo deseada puede ser, en el caso extremo, la conversión de una sola señal de fuente 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 de un solo canal, estéreo de 2 canales o escena objetivo multicanal 5.1.

25 Los sistemas basados en tiempo-frecuencia pueden utilizar una transformada de tiempo-frecuencia (t/f) con resolución temporal y de frecuencia estática. Elegir una cierta rejilla de resolución t/f fija típicamente involucra un intercambio entre resolución de tiempo y frecuencia.

30 El efecto de una resolución fija de t/f puede demostrarse en el ejemplo de las señales de objeto típico en una mezcla de señales de audio. Por ejemplo, los espectros de sonidos tonales muestran una estructura armónicamente relacionada con frecuencia fundamental y varios armónicos. La energía de tales señales se concentra en ciertas regiones de frecuencia. Para tales señales, una alta resolución de frecuencia de la representación utilizada de t/f es beneficiosa para separar las regiones espectrales tonales de banda estrecha de una mezcla de señal. Por el contrario, las señales transitorias, como sonidos de tambor, con frecuencia tienen una estructura temporal distinta: la energía sustancial solo está presente durante periodos cortos de tiempo y se propaga sobre un amplio margen de frecuencias. Para estas señales, una alta resolución temporal de la representación utilizada de t/f es ventajosa para separar la porción de señal transitoria de la mezcla de señal.

40 La resolución de frecuencia obtenida a partir de la representación estándar de SAOC se limita al número de bandas paramétricas, que tienen el valor máximo de 28 en el SAOC estándar. Se obtienen a partir de un banco híbrido de QMF que consiste en análisis de QMF de 64 bandas con una fase de filtración híbrida adicional en las bandas más bajas que divide adicionalmente éstas en hasta 4 subbandas complejas. Las bandas de frecuencia obtenidas se agrupan en bandas paramétricas que imitan la resolución de banda crítica del sistema auditivo humano. La agrupación permite reducir la tasa de datos de información secundaria requerida en un tamaño que pueda tratarse de manera eficiente en aplicaciones prácticas.

45 Los esquemas de codificación de objetos de audio actuales ofrecen solo una variabilidad limitada en la selectividad de tiempo-frecuencia del procesamiento de SAOC. Por ejemplo, SAOC de MPEG [SAOC] [SAOC1] [SAOC2] se limita a la resolución de tiempo-frecuencia que puede obtenerse por el uso del denominado Banco de Filtro de Espejo en Cuadratura Híbrido (QMF Híbrido) y su agrupación subsiguiente en bandas paramétricas. Por lo tanto, la restauración de objetos en el SAOC estándar con frecuencia sufre de la resolución de frecuencia aproximada del QMF Híbrido que conduce a diafonía modulada audible de los otros objetos de audio (por ejemplo, artefactos de doble conversación en voz o artefactos de desigualdad auditiva en música).

55 Otro esquema de codificación de objeto de audio que genera codificación de objeto base y codificación de objeto mejorada con una mayor resolución de frecuencia se divulga mediante la publicación de patente US 2010/0087938 A1 (8 de abril de 2010).

60 El sistema existente produce una calidad de separación razonable dada a la tasa de datos razonablemente baja. El problema principal es la resolución de frecuencia insuficiente para una separación limpia de sonidos tonales. Esto se muestra como un "halo" de otros objetos que rodean los componentes tonales de un objeto. De manera perceptiva, esto se observa como desigualdad o un artefacto tipo vocoder. El efecto perjudicial de este halo puede reducirse incrementando la resolución de frecuencia paramétrica. Se observó que una resolución igual o superior a 512 bandas (en una tasa de muestreo de 44,1 kHz) es suficiente para producir una separación significativamente mejorada de manera perceptiva en señales de prueba. El problema con tal alta resolución paramétrica es que la cantidad de información secundaria necesaria incrementa de forma considerable, en cantidades poco prácticas.

Además, la compatibilidad con los sistemas existentes de SAOC estándar puede perderse.

Por lo tanto, se aprecia altamente si pueden proporcionarse conceptos que muestren cómo superar las restricciones antes descritas del estado de la técnica.

5 El objeto de la presente invención es proporcionar tales conceptos mejorados para codificación de objetos de audio. El objeto de la presente invención se resuelve mediante un decodificador de acuerdo con la reivindicación 1, mediante un codificador de acuerdo con la reivindicación 9, mediante una señal de audio codificada de acuerdo con la reivindicación 14, mediante un sistema de acuerdo con la reivindicación 15, mediante un método para decodificar de acuerdo con la reivindicación 16, mediante un método para codificar de acuerdo con la reivindicación 17 y
10 mediante un programa informático de acuerdo con la reivindicación 18.

Al contrario que el SAOC del estado de la técnica, realizaciones de la presente invención proporcionan una parametrización espectral, de manera que

- 15 - los flujos de bits de parámetros de SAOC que se originan a partir de un codificador estándar de SAOC aún pueden decodificarse por un decodificador mejorado con una calidad perceptiva comparable con la obtenida con un decodificador estándar,
- 20 - los flujos de bits de parámetros de SAOC mejorados pueden decodificarse con un decodificador estándar de SAOC con una calidad comparable con la obtenida con flujos de bits de SAOC estándar,
- los flujos de bits de parámetros de SAOC mejorados pueden decodificarse con calidad óptima con el decodificador mejorado,
- 25 - el decodificador de SAOC mejorado puede ajustar de manera dinámica el nivel de mejora, por ejemplo, dependiendo de los recursos de cálculo disponibles,
- los flujos de bits de parámetros de SAOC estándar y mejoradas pueden mezclarse, por ejemplo, en un escenario de unidad de control de varios puntos (MCU), en un flujo de bits común que puede decodificarse con un
30 decodificador estándar o mejorado con la calidad proporcionada por el decodificador, y
- la parametrización adicional es compacta.

35 Para las propiedades mencionadas anteriormente, se prefiere tener una parametrización que se entienda por el decodificador estándar de SAOC, pero que también permita una distribución eficiente de la información en la mayor resolución de frecuencia. La resolución de la representación de tiempo-frecuencia subyacente determina el rendimiento máximo de las mejoras. La invención aquí define un método para distribuir la información de alta frecuencia mejorada de una forma que sea compacta y permita una decodificación retrocompatible.

40 Una calidad perceptiva de SAOC mejorada puede obtenerse, por ejemplo, al adaptar de manera dinámica la resolución de tiempo-frecuencia del banco de filtro o transformada que se emplea para estimar o se utiliza para sintetizar señales de objetos de audio en propiedades específicas del objeto de audio de entrada. Por ejemplo, si el objeto de audio es cuasi-estacionario durante un cierto intervalo de tiempo, la estimación de parámetros y la síntesis se realiza de manera beneficiosa en una resolución de tiempo aproximada y en una resolución de frecuencia fina. Si
45 el objeto de audio contiene elementos transitorios o no estacionarios durante un cierto intervalo de tiempo, la estimación de parámetros y síntesis se realiza de manera ventajosa utilizando una resolución de tiempo fina y una resolución de frecuencia aproximada. Por consiguiente, la adaptación dinámica del banco de filtro o transformada permite

- 50 - una selectividad de alta frecuencia en la separación espectral de señales cuasi-estacionarias para evitar diafonía entre objetos, y
- precisión temporal para inicios de objeto o eventos transitorios para minimizar los ecos previos y posteriores.

55 Al mismo tiempo, la calidad tradicional de SAOC puede obtenerse correlacionando datos de SAOC estándar en la rejilla de tiempo-frecuencia proporcionada por la transformada adaptable de señal retrocompatible inventiva que depende de la información secundaria que describe las características de señales objeto.

60 Al poder decodificar los datos de SAOC estándar y mejorados, utilizando una transformada común, permite la retrocompatibilidad directa para aplicaciones que abarcan la mezcla de datos de SAOC estándar y mejorados novedosos. También permite una mejora selectiva de tiempo-frecuencia sobre la calidad estándar.

Las realizaciones proporcionadas no se limitan a ninguna transformada de tiempo-frecuencia específica, pero pueden aplicarse con cualquier transformada que proporcione una resolución de frecuencia suficientemente alta. El documento describe la aplicación a un banco de filtro basado en Transformada Discreta de Fourier (DFT) con

resolución de tiempo-frecuencia conmutada. En este enfoque, las señales de dominio de tiempo se subdividen en bloques más cortos, los cuales también pueden solaparse. La señal en cada bloque más corto se pondera por una función de formación de ventanas (que normalmente tiene valores grandes en la parte media y en ambos extremos se reducen a cero). Finalmente, la señal ponderada se transforma en el dominio de frecuencia por la transformada seleccionada, en este punto, por la aplicación de la DFT.

Se proporciona un decodificador para generar una señal de audio no mezclada que comprende una pluralidad de canales de audio no mezclados. El decodificador comprende un determinador de información de segregación para determinar información de segregación recibiendo la primera información secundaria paramétrica en al menos una señal de objeto de audio y la segunda información secundaria paramétrica en al menos una señal de objeto de audio, en el que la resolución de frecuencia de la segunda información secundaria paramétrica es superior a la resolución de frecuencia de la primera información secundaria paramétrica. Además, el decodificador comprende un módulo de segregación para aplicar la información de segregación en una señal de mezcla descendente, que indica una mezcla descendente de al menos una señal de objeto de audio, para obtener una señal de audio no mezclada que comprende la pluralidad de canales de audio no mezclados. El determinador de información de segregación se configura para determinar la información de segregación modificando la primera información paramétrica y la segunda información paramétrica para obtener información paramétrica modificada, de manera que la información paramétrica modificada tenga una resolución de frecuencia que es superior que la primera resolución de frecuencia.

Además, se proporciona un codificador para codificar una o más señales de objeto de audio de entrada. El codificador comprende una unidad de mezclado descendente para mezclar de manera descendente una o más señales de objetos de audio de entrada para obtener una o más señales de mezclado descendente. Además, el codificador comprende un generador de información secundaria paramétrico para generar primera información secundaria paramétrica en al menos una señal de objeto de audio y la segunda información secundaria paramétrica en al menos una señal de objeto de audio, de manera que la resolución de frecuencia de la segunda información secundaria paramétrica es superior que la resolución de frecuencia de la primera información secundaria paramétrica.

Además, se proporciona una señal de audio codificada. La señal de audio codificada comprende una porción de mezcla descendente, que indica una mezcla descendente de una o más señales de objeto de audio de entrada, y una porción de información secundaria paramétrica que comprende primera información secundaria paramétrica en al menos una señal de objeto de audio y segunda información secundaria paramétrica en al menos una señal de objeto de audio. La resolución de frecuencia de la segunda información secundaria paramétrica es superior que la resolución de frecuencia de la primera información secundaria paramétrica.

Además, se proporciona un sistema. El sistema comprende un codificador como se describe anteriormente y un decodificador como se describe anteriormente. El codificador se configura para codificar una o más señales de objeto de audio de entrada obteniendo una o más señales de mezcla descendente que indican una mezcla descendente de una o más señales de objeto de audio de entrada obteniendo la primera información secundaria paramétrica en al menos una señal de objeto de audio y obteniendo la segunda información secundaria paramétrica en al menos una señal de objeto de audio, en el que la resolución de frecuencia de la segunda información secundaria paramétrica es superior que la resolución de frecuencia de la primera información secundaria paramétrica. El decodificador se configura para generar una señal de audio no mezclada basándose en una o más señales de mezcla descendente y basándose en la primera información secundaria paramétrica y la segunda información secundaria paramétrica.

El codificador se configura para codificar una o más señales de objeto de audio de entrada obteniendo una o más señales de mezcla descendente que indican una mezcla descendente de una o más señales de objeto de audio de entrada obteniendo la primera información secundaria paramétrica en al menos una señal de objeto de audio y obteniendo la segunda información secundaria paramétrica en al menos una señal de objeto de audio, en el que la resolución de frecuencia de la segunda información secundaria paramétrica es superior que la resolución de frecuencia de la primera información secundaria paramétrica. El decodificador se configura para generar una señal de salida de audio basándose en al menos una o más señales de mezcla descendente y basándose en la primera información secundaria paramétrica y la segunda información secundaria paramétrica.

Además, se proporciona un método para generar una señal de audio no mezclada que comprende una pluralidad de canales de audio no mezclados. El método comprende:

- Determinar información de segregación recibiendo la primera información secundaria paramétrica en al menos una señal de objeto de audio y la segunda información secundaria paramétrica en al menos una señal de objeto de audio, en el que la resolución de frecuencia de la segunda información secundaria paramétrica es superior que la resolución de frecuencia de la primera información secundaria paramétrica. Y:
- Aplicar la información de segregación en una señal de mezclado descendente, que indica una mezcla

descendente de al menos una señal de objeto de audio, para obtener una señal de audio no mezclada que comprende la pluralidad de canales de audio no mezclados.

5 Determinar la información de segregación comprende modificar la primera información paramétrica y la segunda información paramétrica para obtener información paramétrica modificada, de manera que la información paramétrica modificada tiene una resolución de frecuencia que es mayor que la primera resolución de frecuencia.

Además, se proporciona un método para codificar una o más señales de objeto de audio de entrada. El método comprende:

- 10
- Mezclar de manera descendente una o más señales de objetos de audio de entrada para obtener una o más señales de mezclado descendente. Y:
 - 15 - Generar primera información secundaria paramétrica en al menos una señal de objeto de audio y la segunda información secundaria paramétrica en al menos una señal de objeto de audio, de manera que la resolución de frecuencia de la segunda información secundaria paramétrica es superior que la resolución de frecuencia de la primera información secundaria paramétrica.

20 Además, se proporciona un programa informático para implementar uno de los métodos anteriormente descritos cuando se ejecuta en un ordenador o procesador de señales.

Realizaciones preferidas se proporcionarán en las reivindicaciones dependientes.

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

- 30 La Figura 1A ilustra un decodificador de acuerdo con una realización,
 La Figura 1B ilustra un decodificador de acuerdo con otra realización,
 La Figura 2A ilustra un codificador de acuerdo con una realización,
 La Figura 2B ilustra un codificador de acuerdo con otra realización,
 La Figura 2C ilustra una señal de audio codificada de acuerdo con una realización,
 La Figura 3 ilustra un sistema de acuerdo con una realización,
 La Figura 4 muestra un diagrama de bloques esquemático de una descripción general conceptual de un sistema de SAOC,
 35 La Figura 5 muestra un diagrama esquemático e ilustrativo de una representación temporal espectral de una señal de audio de un solo canal,
 La Figura 6 muestra un diagrama de bloques esquemático de un cálculo selectivo de tiempo-frecuencia de la información secundaria dentro del codificador de SAOC,
 La Figura 7 ilustra representación retrocompatible de acuerdo con realizaciones,
 40 La Figura 8 ilustra la curva de diferencia entre el valor de parámetro verdadero y el valor de media de baja resolución de acuerdo con una realización,
 La Figura 9 representa una ilustración de alto nivel del codificador mejorado que proporciona un flujo de bits retrocompatible con mejoras de acuerdo con una realización
 La Figura 10 ilustra un diagrama de bloques de un codificador de acuerdo con una realización particular que implementa la trayectoria paramétrica de un codificador,
 45 La Figura 11 representa un diagrama de bloques de alto nivel de un codificador mejorado de acuerdo con una realización que es compatible para decodificar los flujos de bits estándar y mejoradas,
 La Figura 12 ilustra un diagrama de bloques que ilustra una realización de la unidad de decodificación de PSI mejorada,
 50 La Figura 13 representa un diagrama de bloques para decodificar flujos de bits SAOC estándar con el decodificador de SAOC mejorado de acuerdo con una realización,
 La Figura 14 representa los bloques funcionales principales del decodificador de acuerdo con una realización,
 La Figura 15 ilustra una señal tonal y de ruido y, en particular, el espectro de potencia de alta resolución y las reconstrucciones aproximadas correspondientes,
 55 La Figura 16 ilustra la modificación de ambas señales ilustrativas, en particular los factores de corrección para las señales ilustrativas,
 La Figura 17 ilustra los factores de corrección originales y las aproximaciones basadas en predicción lineal de orden reducido para las señales ilustrativas, y
 60 La Figura 18 ilustra el resultado de aplicar los factores de corrección modelados en las reconstrucciones aproximadas.

Antes de describir realizaciones de la presente invención, se proporcionan más antecedentes sobre los sistemas de SAOC del estado de la técnica.

La Figura 4 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 codificador 10 comprende un mezclador descendente 16 que recibe las señales de audio s_1 a s_N y mezcla descendentemente parte de la señal de mezcla descendente 18. Alternativamente, la mezcla descendente puede proporcionarse externamente (“mezcla descendente artística”) y el sistema estima información secundaria adicional para hacer la mezcla descendente proporcionada con relación en la mezcla descendente calculada. En la Figura 4, la señal de mezcla descendente se muestra que es una señal de canal P . De esta manera, es concebible cualquier configuración de señal de mezcla descendente de mono ($P=1$), estéreo ($P=2$) o multicanal ($P>2$).

En el caso de una mezcla descendente estéreo, los canales de la señal de mezcla descendente 18 se indican $L0$ y $R0$, en caso de una misma mezcla descendente mono se indica simplemente $L0$. Para permitir que el decodificador de SAOC 12 recupere los objetos individuales s_1 a s_N , el estimador de información secundaria 17 proporciona al decodificador de SAOC 12 con la información secundaria que incluye los parámetros de SAOC. Por ejemplo, en el caso de una mezcla descendente estéreo, los parámetros de SAOC comprenden diferencias de nivel de objeto (OLD), correlaciones entre objetos (IOC) (parámetros de correlación cruzada entre objetos), valores de ganancia de mezcla descendente (DMG) y diferencias de nivel de canal de mezcla descendente (DCLD). Incluyendo la información secundaria 20 los parámetros de SAOC, junto con la señal de mezcla descendente 18, forma el flujo 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 secundaria 20 para recuperar y presentar las señales de audio \hat{s}_i y \hat{s}_N sobre cualquier conjunto seleccionado por el usuario de canales \hat{y}_i a \hat{y}_M , prescribiéndose la conversión convirtiendo la información 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 dominio de tiempo o espectral. En caso de que las señales de audio s_1 a s_N se alimenten en el codificador 10 en el dominio de tiempo, tal como codificadas por PCM, el codificador 10 puede utilizar un banco de filtro, tal como el banco de QMF híbrido, para transferir las señales hacia un dominio espectral, en el cual las señales de audio se representan en varias subbandas asociadas con diferentes porciones espectrales, en una resolución específica de banco de filtro. Si las señales de audio s_1 a s_N ya se encuentran en la representación esperada por el codificador 10, algunas no tienen que realizar la descomposición espectral.

La Figura 5 muestra una señal de audio en el dominio espectral recién mencionado. Como puede observarse, la señal de audio se representa como una pluralidad de señales de subbanda. Cada señal 30_1 a 30_k de subbanda consiste en una secuencia temporal de valores de subbanda indicados por los recuadros 32 pequeños. Como puede observarse, los valores de subbanda 32 de las señales 30_1 a 30_k de subbanda se sincronizan entre sí en tiempo de modo que, para cada uno de los intervalos 34 de tiempo consecutivos de banco de filtro, cada 30_1 a 30_k comprende un valor de subbanda 32 exacto. Como se ilustra por el eje de frecuencia 36, las señales de subbanda 30_1 o 30_k se asocian con diferentes regiones de frecuencia, y como se ilustra por el eje de tiempo 38, los intervalos de tiempo de banco de filtro se disponen de manera consecutiva en tiempo.

Como se indica anteriormente, el extractor de información secundaria 17 de la Figura 4 calcula los parámetros de SAOC de las señales de audio s_1 a s_N de entrada. De acuerdo con la norma de SAOC y actualmente implementado, el codificador 10 realiza este cálculo en una resolución de tiempo/frecuencia que puede disminuirse con respecto a la resolución original de tiempo/frecuencia que se determina por los intervalos 34 de tiempo de banco de filtro y la descomposición de subbandas, por una cierta cantidad, con esta cierta cantidad siendo señalada en el lado del decodificador dentro de la información secundaria 20. Grupos de intervalos 34 de tiempo consecutivos de banco de filtro pueden formar una trama de SAOC 41. También, el número de bandas de parámetros dentro de la trama de SAOC 41 se transporta dentro de la información secundaria 20. Por lo tanto, el dominio de tiempo/frecuencia se divide en cintas de tiempo/frecuencia ilustrados en la Figura 5 por líneas 42 discontinuas. En la Figura 5, las bandas de parámetros se distribuyen de la misma manera en las diversas tramas de SAOC 41 representadas de modo que una disposición regular de cintas de tiempo/frecuencia se obtiene. En general, sin embargo, las bandas de parámetros pueden variar de una trama de SAOC 41 a la subsiguiente, dependiendo de las diferentes necesidades de resolución espectral en las respectivas tramas de SAOC 41. Además, la duración de las tramas de SAOC 41 puede variar, también. Como consecuencia, la disposición de cintas de tiempo/frecuencia puede ser irregular. No obstante, las cintas de tiempo/frecuencia dentro de una trama de SAOC 41 particular típicamente tienen la misma duración y se alinean en la dirección de tiempo, es decir, todas las cintas de t/f en dicha trama de SAOC 41 comienzan al inicio de la trama de SAOC 41 dada y terminan al final de dicha trama de SAOC 41.

El extractor de información secundaria 17 representado en la Figura 4 calcula los parámetros de SAOC de acuerdo con las siguientes formulas. En particular, el extractor de información secundaria 17 calcula las diferencias de nivel de objeto para cada objeto i como

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

en la que las sumas y los índices n y k , respectivamente, experimentan todos los índices 34 temporales, y todos los índices 30 espectrales que pertenecen a una cierta cinta 42 de tiempo/frecuencia, referenciados por los índices l

5 para la trama de SAOC (o intervalo de tiempo de procesamiento) y m para la banda de parámetro, y $x_i^{n,k*}$ es el conjugado complejo de $x_i^{n,k}$. Por consiguiente, las energías de todos los valores x_i de subbanda de una señal de audio u objeto i se suman y normalizan en el valor de energía más alto de esa cinta entre todos los objetos o señales de audio.

10 Además, el extractor de información secundaria de SAOC 17 es capaz de calcular una medida de similitud de las cintas de tiempo/frecuencia correspondientes de pares de diferentes objetos de entrada s_1 a s_N . Aunque el extractor de información secundaria 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 secundaria de SAOC 17 también puede suprimir la señalización de las medidas de similitud o restringir el cálculo de las medidas de similitud en objetos de audio s_1 a s_N , que forman canales izquierdo o derecho de un canal estéreo común. En cualquier caso, la medida de similitud se denomina
15 como el parámetro de correlación cruzada entre objetos $IOC_{i,j}^{l,m}$. El cálculo es como se indica a continuación

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

20 de nuevo con los índices de ganancia n y k experimentando todos los valores de subbanda que pertenecen a una cierta cinta 42 de tiempo/frecuencia, indicando i y j un cierto par de objetos de audio s_1 a s_N , e indicando $\text{Re}\{\}$ la operación de retener solo la parte real (es decir, descartar la parte imaginaria) del argumento con valor complejo.

25 El mezclador descendente 16 de la Figura 4 mezcla de forma descendente los objetos s_1 a s_N por el uso de factores de ganancia aplicados a cada objeto s_1 a s_N . Es decir, un factor de ganancia de d_i se aplica al objeto i y después todos los objetos s_1 a s_N de esta manera ponderados se suman para obtener una señal de mezcla descendente mono, que se ilustra en la Figura 4 si $P=1$. En otro caso ilustrativo de una señal de mezcla descendente de dos canales, representada en la Figura 4 si $P=2$, un factor de ganancia $d_{1,i}$ se aplica al objeto i y después todos los objetos de ganancia amplificada se suman para obtener el canal de mezcla descendente izquierdo $L0$, y los factores de ganancia $d_{2,i}$ se aplican al objeto y después los objetos de esta manera amplificados con ganancia se suman para obtener el canal de mezcla descendente derecho $R0$. Un procesamiento que es análogo al anterior se aplicará en caso de una mezcla descendente multicanal ($P>2$).

35 La prescripción de mezcla descendente se señala en el lado de decodificador por medio de ganancias de mezcla descendente DMG_i y, en caso de una señal de mezcla descendente estéreo, las diferencias $DCLD_i$ y nivel de canal de mezcla descendente.

Las ganancias de mezcla descendente se calculan de acuerdo con:

$$40 \quad \begin{aligned} DMG_i &= 20 \log_{10} (d_i + \varepsilon) & , \text{ (mezcla descendente mono),} \\ DMG_i &= 10 \log_{10} (d_{1,i}^2 + d_{2,i}^2 + \varepsilon) & , \text{ (mezcla descendente estéreo),} \end{aligned}$$

en la que ε es un número pequeño tal como 10^{-9} .

Para los DCLD, se aplica la siguiente formula:

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

5

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

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

para una mezcla descendente mono, o

10

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

para una mezcla descendente estéreo, respectivamente.

15

De esta manera, en las fórmulas antes mencionadas, los parámetros OLD e IOC son una función de las señales de audio y los parámetros DMG y DCLD son funciones de los coeficientes de mezcla descendente d . Por cierto, se observa que d puede variar en tiempo y frecuencia.

20

De esta manera, en el modo normal, el mezclador descendente 16 mezcla todos los objetos s_1 a s_N sin ninguna preferencia, es decir, con el tratamiento de todos los objetos s_1 a s_N de manera equitativa.

25

En el lado de decodificador, el mezclador ascendente realiza la inversión del procedimiento de mezcla descendente y la implementación de la "información de conversión" 26 representada por una matriz \mathbf{R} (en la bibliografía algunas veces también denominada como \mathbf{A}) en una etapa de cálculo, particularmente, en caso de una matriz descendente de dos canales

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

30

en la que la matriz de \mathbf{E} es una función de los parámetros OLD e IOC y la matriz \mathbf{D} contiene los coeficientes de mezcla descendente como

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

35

y en la que \mathbf{D}^* indica la transposición compleja de \mathbf{D} . La matriz de \mathbf{E} es una matriz de covarianza estimada de los objetos de audio s_1 a s_N . En implementaciones actuales de SAOC, el cálculo de la matriz de \mathbf{E} de covarianza estimada típicamente se realiza en la resolución espectral/temporal de los parámetros de SAOC, es decir, para cada (l,m) , de manera que la matriz de covarianza estimada pueda escribirse como $\mathbf{E}^{l,m}$. La matriz de covarianza estimada $\mathbf{E}^{l,m}$ es de tamaño $N \times N$ con coeficientes siendo definidos como

40

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

De esta manera la matriz $\mathbf{E}^{l,m}$ 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 de nivel de objeto, es decir, $e_{i,j}^{l,m} = OLD_i^{l,m}$ para $i=j$, puesto que

5 $OLD_i^{l,m} = OLD_j^{l,m}$ y $IOC_{i,j}^{l,m} = 1$ para $i=j$. Fuera de su diagonal, la matriz **E** de covarianza estimada tiene coeficientes de matriz que representan la media geométrica de las diferencias de nivel de objeto de los objetos i y j , respectivamente, ponderados con la medida de correlación cruzada entre objetos $IOC_{i,j}^{l,m}$.

La Figura 6 despliega un principio posible de implementación sobre el ejemplo del Estimador de Información Secundaria (SIE) como parte de un codificador de SAOC 10. El codificador de SAOC 10 comprende el mezclador 16 y el estimador 17 de estimación secundaria (SIE). El SIR consiste de manera conceptual en dos módulos: un módulo 10 45 para calcular una representación de t/f basada en corto tiempo (por ejemplo, STFT o QMF) de cada señal. La representación de t/f de corto tiempo calculada se alimenta en el segundo módulo 46, el módulo de Estimación de Información Secundaria selectivo de t/f (t/f-SIE). El módulo 46 de t/f-SIE calcula la información secundaria para cada t/f-cinta. En implementaciones actuales de SAOC, la transformada de tiempo/frecuencia se fija y es idéntica para 15 todos los objetos de audio s_1 a s_N . Además, los parámetros de SAOC se determinan sobre tramas de SAOC que son las mismas 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 , sin tomar en cuenta de esta manera las necesidades específicas de objeto para la resolución temporal fina en algunos casos o la resolución espectral fina en otros casos.

20 A continuación, se describen realizaciones de la presente invención.

La Figura 1A ilustra un decodificador para generar una señal de audio no mezclada que comprende una pluralidad de canales de audio no mezclados de acuerdo con una realización.

25 El decodificador comprende un determinador de información de segregación 112 para determinar información de segregación recibiendo la primera información secundaria paramétrica en al menos una señal de objeto de audio y la segunda información secundaria paramétrica sobre al menos una señal de objeto de audio, en el que la resolución de frecuencia de la segunda información secundaria paramétrica es superior a la resolución de frecuencia de la primera información secundaria paramétrica.

30 Además, el decodificador comprende un módulo de segregación 113 para aplicar la información de segregación en una señal de mezcla descendente, que indica una mezcla descendente de al menos una señal de objeto de audio, para obtener una señal de audio no mezclada que comprende la pluralidad de canales de audio no mezclados.

35 El determinador de información de segregación 112 se configura para determinar la información de segregación modificando la primera información paramétrica y la segunda información paramétrica para obtener información paramétrica modificada, de manera que la información paramétrica modificada tenga una resolución de frecuencia que es superior que la primera resolución de frecuencia.

40 La Figura 1B ilustra un decodificador para generar una señal de audio no mezclada que comprende una pluralidad de canales de audio no mezclados de acuerdo con otra realización. El decodificador de la Figura 1B además comprende una primera unidad 111 de transformada para transformar una entrada de mezcla descendente, que se representa en un dominio de tiempo, para obtener la señal de mezcla descendente, que se representa en un dominio de tiempo-frecuencia. Además, el decodificador de la Figura 1B comprende una segunda unidad 114 de 45 transformada para transformar la señal de audio no mezclada del dominio de tiempo-frecuencia en el dominio de tiempo.

La Figura 2A ilustra un codificador para codificar una o más señales de objeto de audio de entrada de acuerdo con una realización.

50 El codificador comprende una unidad de mezcla descendente 91 para mezclar de forma descendente una o más señales de objetos de audio de entrada para obtener una o más señales de mezcla descendente.

55 Además, el codificador comprende un generador de información secundaria paramétrica 93 para generar la primera información secundaria paramétrica sobre al menos una señal de objeto de audio y la segunda información

secundaria paramétrica sobre al menos una señal de objeto de audio, de manera que la resolución de frecuencia de la segunda información secundaria paramétrica es superior a la resolución de frecuencia de la primera información secundaria paramétrica.

5 La Figura 2B ilustra un codificador para codificar una o más señales de objeto de audio de acuerdo con otra realización. El codificador de la Figura 2B además comprende una unidad de transformada 92 para transformar una o más señales de objeto de audio de entrada desde un dominio de tiempo hasta un dominio de tiempo-frecuencia para obtener una o más señales de objetos de audio transformadas. En la realización de la Figura 2B, el generador de información secundaria paramétrica 93 se configura para generar la primera información secundaria paramétrica y la segunda información secundaria paramétrica basándose en una o más señales de objeto de audio transformadas.

15 La Figura 2C ilustra una señal de audio codificada de acuerdo con una realización. La señal de audio codificada comprende una porción de mezcla descendente 51, que indica una mezcla descendente de una o más señales de objeto de audio de entrada, y una porción de información secundaria paramétrica 52 que comprende la primera información secundaria paramétrica sobre al menos una señal de objeto de audio y la segunda información secundaria paramétrica sobre al menos una señal de objeto de audio. La resolución de frecuencia de la segunda información secundaria paramétrica es superior a la resolución de frecuencia de la primera información secundaria paramétrica.

20 La Figura 3 ilustra un sistema de acuerdo con una realización, El sistema comprende un codificador 61 como se describe anteriormente y un decodificador 62 como se describe anteriormente.

25 El codificador 61 se configura para codificar una o más señales de objeto de audio de entrada obteniendo una o más señales de mezcla descendente que indican una mezcla descendente de una o más señales de objeto de audio de entrada, obteniendo la primera información secundaria paramétrica sobre al menos una señal de objeto de audio y obteniendo la segunda información secundaria paramétrica sobre al menos una señal de objeto de audio, en el que la resolución de frecuencia de la segunda información secundaria paramétrica es superior a la resolución de frecuencia de la primera información secundaria paramétrica.

30 El decodificador 62 se configura para generar una señal de audio no mezclada basándose en una o más señales de mezcla descendente, y basándose en la primera información secundaria paramétrica y la segunda información secundaria paramétrica.

35 A continuación, se describe SAOC mejorado utilizando mejora de resolución de frecuencia retrocompatible.

La Figura 7 ilustra representación retrocompatible de acuerdo con realizaciones. La propiedad de señal que se representa, por ejemplo, el envolvente espectral de potencia 71, varía sobre la frecuencia. El eje de frecuencia se divide en bandas paramétricas, y un único conjunto de descriptores de señal se asigna para cada subbanda. Utilizando los mismos en lugar de distribuir la descripción para cada contenedor de frecuencia por separado permite ahorros en la cantidad de información secundaria requerida sin una pérdida significativa en la calidad perceptual. En el SAOC estándar, el descriptor único para cada banda es el valor medio 72, 73, 74 de los descriptores de contenedores. Como puede entenderse, esto puede introducir una pérdida de información cuya magnitud depende de las propiedades de señal. En la Figura 7, las bandas $k-1$ y K tienen un error bastante grande, mientras que en la banda $k+1$ el error es mucho menor.

50 La Figura 8 ilustra la curva de diferencia 81 entre el valor de parámetro verdadero y el valor medio de baja resolución de acuerdo con una realización, por ejemplo, la información de estructura fina perdida en la parametrización de SAOC estándar. Se describe un método para parametrizar y transmitir las curvas de diferencia 81 entre los valores medios 72, 73, 74 (por ejemplo, el descriptor estándar de SAOC) y los valores de resolución fina verdaderos de una forma eficiente que permita aproximarse a la estructura de resolución fina en el decodificador.

55 Cabe señalar que agregar la información de mejora a un solo objeto en una mezcla no solo mejora la calidad resultante de ese objeto específico, sino también la calidad de todos los objetos que comparten la ubicación espacial aproximada y tienen cierto solapamiento espectral.

A continuación, se describe la codificación mejorada de SAOC retrocompatible con un codificador mejorado, en particular, un codificador mejorado de SAOC que produce un flujo de bits que contiene una porción de información secundaria retrocompatible y mejoras adicionales. La información agregada puede insertarse en el flujo de bits de SAOC estándar de tal manera que los decodificadores que cumplen con el estándar antiguo simplemente ignoran los datos agregados mientras los decodificadores mejorados hacen uso del mismo. Los decodificadores de SAOC estándar existentes pueden decodificar la porción retrocompatible de la información secundaria paramétrica (PSI) y producen reconstrucciones de los objetos, mientras la información agregada utilizada por el decodificador mejorado de SAOC mejora la calidad perceptiva de las reconstrucciones en la mayoría de los casos. Adicionalmente, si el

5 decodificador de SAOC mejorado se ejecuta en recursos limitados, las mejoras pueden ignorarse y aún se obtiene una reconstrucción de calidad básica. Cabe señalar que las reconstrucciones de los decodificadores de SAOC estándar y de SAOC mejorado que utilizan solo la PSI compatible con SAOC estándar difieren, pero se juzgan muy similares de manera perceptiva (la diferencia es de naturaleza similar como en la decodificación de flujos de bits de SAOC estándar con un decodificador de SAOC mejorado).

La Figura 9 representa una ilustración de alto nivel del codificador mejorado que proporciona un flujo de bits retrocompatible con mejoras de acuerdo con una realización.

10 El codificador comprende una unidad de mezcla descendente 91 para mezclar de forma descendente una pluralidad de señales de objetos de audio para obtener una o más señales de mezcla descendente. Por ejemplo, las señales de objeto de audio (por ejemplo, los objetos individuales (de audio)) se utilizan por una unidad de mezcla descendente 91 para crear una señal de mezcla descendente. Esto puede suceder en el dominio de tiempo, dominio de frecuencia, o incluso puede utilizarse una mezcla descendente proporcionada de manera externa.

15 En la trayectoria de PSI, las señales de objeto (audio) se transforman por una unidad de transformada 92 de un dominio de tiempo a un dominio de frecuencia, un dominio de tiempo-frecuencia o un dominio espectral (por ejemplo, por una unidad de transformada 92 que comprende una o más subunidades de transformada de t/f 921, 922).

20 Además, el codificador comprende un generador de información secundaria paramétrica 93 para generar información secundaria paramétrica. En la realización de la Figura 9, el generador de información secundaria paramétrica 93 por ejemplo, puede comprender una unidad de extracción de PSI 94 y un divisor de PSI 95: de acuerdo con una realización de este tipo, en el dominio de frecuencia, la PSI se extrae por la unidad de extracción de PSI 94. El divisor de PSI 95 entonces se lleva a cabo para dividir la PSI en dos partes: la parte de resolución de frecuencia estándar que puede decodificarse con cualquier decodificador de SAOC que cumple con el estándar, y la parte de resolución de frecuencia mejorada. Esta última puede estar "oculta" de los elementos de flujo de bits, de manera que estos se ignorarán por los decodificadores estándar pero se utilizarán por los decodificadores mejorados.

30 La Figura 10 ilustra un diagrama de bloques de un codificador de acuerdo con una realización particular que implementa la trayectoria paramétrica del codificador descrito anteriormente. Los bloques funcionales en negrita (102, 105, 106, 107, 108, 109) indican los componentes principales del procesamiento inventivo. En particular, la Figura 10 ilustra un diagrama de bloques de codificación de dos fases que produce el flujo de bits retrocompatible con mejoras para decodificadores más capaces. El codificador se configura para producir PSI que puede decodificarse con ambas versiones del decodificador. La unidad de transformada 92 de la Figura 9 se implementa por una unidad de detección transitoria 101, por una unidad de secuencia de creación de ventana 102, y por una unidad de análisis de t/f 103 en la Figura 10. Las otras unidades 104, 105, 106, 107, 108, 109 en la Figura 10 implementan el generador de información secundaria paramétrica 93 (por ejemplo, las unidades 104, 105, 106, 107, 108, 109 pueden implementar la funcionalidad de la combinación de la unidad de extracción de PSI 94 y el divisor de PSI 95).

45 En primer lugar, la señal se subdivide en tramas de análisis, que a continuación se transforman en el dominio de frecuencia. Múltiples tramas de análisis se agrupan en una trama de parámetros de longitud fija, por ejemplo, en longitudes de SAOC estándar de 16 y 32 tramas de análisis son comunes. Se supone que las propiedades de señales siguen siendo cuasi-estacionarias durante la trama de parámetros y de esta manera pueden caracterizarse con solo un conjunto de parámetros. Si las características de señales cambian dentro de la trama de parámetros, se sufre de un error de modelado, y puede ser beneficioso subdividir la trama de parámetros más larga en partes en las que se cumple nuevamente con la suposición cuasi-estacionaria. Para este propósito, se necesita una detección transitoria.

50 En una realización, la unidad de transformada 92 se configura para transformar una o más señales de objeto de audio de entrada desde el dominio de tiempo hasta el dominio de tiempo-frecuencia dependiendo de una longitud de ventana de un bloque de transformada de señales que comprende valores de señales de al menos una de una o más señales de objeto audio de entrada. La unidad de transformada 92 comprende una unidad de detección transitoria 101 para determinar un resultado de detección transitoria que indica si se encuentra presente un elemento transitorio en una o más de al menos una de las señales de objeto de audio, en el que un elemento transitorio indica un cambio de señal en una o más de al menos una de las señales de objeto de audio. Además, la unidad de transformada 92 comprende adicionalmente una unidad de secuencia de ventana 102 para determinar la longitud de ventana dependiendo del resultado de detección transitoria.

60 Por ejemplo los elementos transitorios pueden detectarse por la unidad de detección transitoria 101 a partir de todos los objetos de entrada por separado, y cuando existe un evento transitorio en únicamente uno de los objetos cuya ubicación se declara como ubicación transitoria global. La información de las ubicaciones transitorias se utiliza para

construir una secuencia de formación de ventanas adecuada. La construcción puede basarse, por ejemplo, en la siguiente lógica:

- 5 - Establecer una longitud de ventana por defecto, es decir, la longitud de un bloque de transformada de señal por defecto, por ejemplo, 2048 muestras.
- 10 - Establecer la longitud de trama de parámetros, por ejemplo, 4096 muestras, que corresponden a 4 ventanas por defecto con un solapamiento del 50 %. Las tramas de parámetros agrupan múltiples ventanas juntas y se utiliza un único conjunto de descriptores de señales para todo el bloque en lugar de tener descriptores para cada ventana por separado. Esto permite reducir la cantidad de PSI.
- 15 - Si no se ha detectado ningún elemento transitorio, utilizar las ventanas por defecto y toda la longitud de trama de parámetros.
- 20 - Si se detecta un elemento transitorio, adaptar la formación de ventanas para proporcionar una mejor resolución temporal en la ubicación del elemento transitorio.

La unidad de secuencia de creación de ventana 102 construye la secuencia de formación de ventanas. Al mismo tiempo, también crea subtramas de parámetros a partir de una o más ventanas de análisis. Cada subconjunto se analiza como entidad y solo un conjunto de parámetros de PSI se transmite para cada subbloque. Para proporcionar una PSI compatible con SAOC estándar, la longitud de bloque de parámetro definida se utiliza como la longitud de bloque de parámetro principal, y los elementos transitorios ubicados posibles dentro de ese bloque definen los subconjuntos de parámetros.

25 La secuencia de ventana construida se produce para el análisis de tiempo-frecuencia de las señales de audio de entrada llevadas a cabo por la unidad de análisis de t/f 103 y se transmiten en la porción de mejora de SAOC mejorado de la PSI.

30 La PSI consiste en conjuntos de diferencias de nivel de objeto (OLD), correlaciones entre objetos (IOC) e información de la matriz de **D** de mezcla descendente utilizada para crear la señal de mezcla descendente a partir de los objetos individuales en el codificador. Cada conjunto de parámetros se asocia con un limitador de parámetros que define la región temporal a la que se asocian los parámetros.

35 Los datos espectrales de cada ventana de análisis se utilizan por la unidad de estimación de PSI 104 para estimar la PSI para la parte de SAOC estándar. Esto se hace agrupando los contenedores espectrales en las bandas paramétricas de SAOC estándar y estimar los IOC, OLD y energías de objetos absolutas (NRG) en las bandas. Siguiendo ligeramente la notación de SAOC estándar, el producto normalizado de dos espectros de objeto es $\mathbf{S}_i(f, n)$ y $\mathbf{S}_j(f, n)$ en una cinta de parametrización se define como

$$nrg_{i,j}(b) = \frac{\sum_{n=0}^{N-1} \sum_{f=0}^{F_n-1} \mathbf{K}(b, f, n) \mathbf{S}_i(f, n) \mathbf{S}_j^*(f, n)}{\sum_{n=0}^{N-1} \sum_{f=0}^{F_n-1} \mathbf{K}(b, f, n)},$$

40

en la que la matriz de $\mathbf{K}(b, f, n) : \mathbb{R}^{B \times F_n \times N}$ define la correlación de los F_n contenedores de representación t/f en la trama n en B bandas paramétricas mediante

$$\mathbf{K}(b, f, n) = \begin{cases} 1, & \text{if } f \in b \\ 0, & \text{en caso contrario} \end{cases}.$$

45

La resolución espectral puede variar entre las tramas dentro de un solo bloque paramétrico, de manera que la matriz de correlación convierte los datos en una base de resolución común. La energía de objeto máxima en este cinta de parametrización se define por ser la energía de objeto máxima $NRG(b) = \max_i(nrg_{i,i}(b))$. Teniendo este valor, los OLD se definen a continuación para ser las energías de objeto normalizado.

50

$$OLD_i(b) = \frac{nrg_{i,i}(b)}{NRG(b)}$$

Y finalmente, la IOC puede obtenerse a partir de las potencias cruzadas como

$$IOC_{i,j}(b) = \operatorname{Re} \left\{ \frac{nrg_{i,j}(b)}{\sqrt{nrg_{i,i}(b)nrg_{j,j}(b)}} \right\}$$

5

Esto concluye la estimación de las partes compatibles con SAOC estándar del flujo de bits.

10 Una unidad de reconstrucción de espectro de potencia aproximada 105 se configura para utilizar los OLD y los NRG para reconstruir una estimación aproximada del envolvente espectral en el bloque de análisis de parámetros. El envolvente se construye en la resolución más alta de frecuencia utilizada en ese bloque.

15 El espectro original de cada ventana de análisis se utiliza por una unidad de estimación de espectro de potencia 106 para calcular el espectro de potencia en esa ventana.

20 Los espectros de potencia obtenidos se transforman en una representación de alta resolución de frecuencia común por una unidad de adaptación de resolución de frecuencia 107. Esto puede hacerse, por ejemplo, interpolando los valores espectrales de potencia. Después, el perfil espectral de potencia promedio se calcula promediando los espectros dentro del bloque de parámetros. Este corresponde aproximadamente a la estimación de OLD que omite la adición de bandas paramétricas. El perfil espectral obtenido se considera como el OLD de resolución fina.

25 El codificador además comprende una unidad de estimación delta 108 para estimar una pluralidad de factores de corrección dividiendo cada uno de la pluralidad de OLD de una de al menos una señal de objeto de audio por un valor de una reconstrucción de espectro de potencia de dicha una de al menos una señal objeto de audio para obtener la segunda información secundaria paramétrica, en el que dicha pluralidad de OLD tiene una mayor resolución de frecuencia que dicha reconstrucción de espectro de potencia.

30 En una realización, la unidad de estimación delta 108 se configura para estimar una pluralidad de factores de corrección dependiendo de una pluralidad de valores paramétricos que depende de al menos una señal de objeto de audio para obtener la segunda información secundaria paramétrica. Por ejemplo, la unidad de estimación delta 108 puede configurarse para estimar un factor de corrección, "delta", por ejemplo, dividiendo el OLD de resolución fina por la reconstrucción de espectro de potencia aproximada. Como resultado, esto proporciona a cada contenedor de frecuencia un factor de corrección (por ejemplo, multiplicativo) que puede utilizarse para aproximarse al OLD de resolución fina dados los espectros aproximados.

35

Finalmente, una unidad de modelado delta 109 se configura para modelar el factor de corrección estimado de una forma eficiente para transmisión. Una posibilidad para modelar utilizando los Coeficientes de Predicción Lineal (LPC) se describe posteriormente a continuación.

40 Efectivamente, las modificaciones de SAOC que están mejoradas consisten en agregar la información de secuencia de formación de ventanas y los parámetros para transmitir la "delta" al flujo de bits.

A continuación, se describe un decodificador mejorado.

45 La Figura 11 representa un diagrama de bloques de alto nivel de un decodificador mejorado de acuerdo con una realización que es capaz de decodificar flujos de bits estándar y mejoradas. En particular, la Figura 11 ilustra un diagrama de bloques operacional de un decodificador mejorado capaz de decodificar ambos flujos de bits estándar así como flujos de bits que incluyen mejoras de resolución de frecuencia.

50 La señal de mezcla descendente de entrada se transforma en el dominio de frecuencia por una unidad de transformada t/ff 111.

La matriz de segregación estimada se aplica en la señal de mezcla descendente transformada por una unidad de segregación 110 para generar una salida de segregación.

55

Adicionalmente, una trayectoria de decorrelación se incluye para permitir un mejor control espacial de los objetos en el no mezclado. Una unidad de decorrelación 119 lleva a cabo la decorrelación en la señal de mezcla descendente transformada y el resultado de la decorrelación se alimenta en la unidad de segregación 110. La unidad de segregación 110 utiliza el resultado de decorrelación para generar la salida no mezclada.

La salida no mezclada entonces se transforma nuevamente en el dominio de tiempo por una unidad de transformada f/t 114.

- 5 La trayectoria de procesamiento paramétrico puede tomar la PSI de resolución estándar como la entrada, en cuyo caso la PSI decodificada, que se genera por una unidad de decodificación de PSI estándar 115, se adapta por una unidad de conversión de resolución de frecuencia 116 en la resolución de frecuencia utilizada en las transformadas de t/f .
- 10 Una entrada alternativa combina la parte de resolución de frecuencia estándar de la PSI con la parte de resolución de frecuencia mejorada y los cálculos incluyen la información de resolución de frecuencia mejorada. En mayor detalle, una unidad de decodificación de PSI mejorada 117 genera la PSI decodificada que muestra resolución de frecuencia mejorada.
- 15 Un generador de matriz de segregación 118 genera una matriz de segregación basándose en la PSI decodificada recibida de la unidad de conversión de resolución de frecuencia 116 o de la unidad de decodificación de PSI mejorada 117. El generador de matriz de segregación 118 también puede generar la matriz de segregación basándose en la información de conversión, por ejemplo, basándose en una matriz de conversión. La unidad de segregación 110 se configura para generar la salida no mezclada aplicando esta matriz no mezclada, que se genera por el generador de matriz de segregación 118, en la señal de mezcla descendente transformada.
- 20

La Figura 12 ilustra un diagrama de bloques que ilustra una realización de la unidad de decodificación de PSI mejorada 117 de la Figura 11.

- 25 La primera información paramétrica comprende una pluralidad de primeros valores de parámetros, en el que la segunda información paramétrica comprende una pluralidad de segundos valores de parámetros. El determinador de información de segregación 112 comprende una subunidad de conversión de resolución de frecuencia 122 y un combinador 124. La unidad de conversión de resolución de frecuencia 112 se configura para generar los valores de parámetros adicionales, por ejemplo, replicando los primeros valores de parámetros, en la que los primeros valores de parámetros y los valores de parámetros adicionales juntos forman una pluralidad de primeros valores de parámetros procesados. El combinador 124 se configura para combinar los primeros valores de parámetros procesados y los segundos valores de parámetros para obtener una pluralidad de valores de parámetros modificados como la información paramétrica modificada.
- 30
- 35 De acuerdo con una realización, la parte de resolución de frecuencia estándar se decodifica por una subunidad de decodificación 121 y se convierte por una subunidad de conversión de resolución de frecuencia 122 en la resolución de frecuencia utilizada por la parte de mejora. La parte de mejora decodificada, generada por una subunidad de decodificación de PSI mejorada 123, se combina por un combinador 124 con la parte de resolución estándar convertida.
- 40 A continuación, se describen en mayor detalle los dos modos de decodificación con posibles implementaciones.

En primer lugar, se describe la decodificación de flujos de bits de SAOC estándar con un decodificador mejorado:

- 45 El decodificador de SAOC mejorado se diseña de manera que sea capaz de decodificar flujos de bits a partir de codificadores de SAOC estándar con una buena calidad. La decodificación se limita a la reconstrucción paramétrica únicamente y se ignoran posibles flujos residuales.

- 50 La Figura 13 representa un diagrama de bloques para decodificar flujos de bits de SAOC estándar con el decodificador de SAOC mejorado que ilustra el proceso de decodificación de acuerdo con una realización. Los bloques funcionales en negrita (131, 132, 133, 135) indican la parte principal del procesamiento inventivo.

- 55 Un calculador de matriz de segregación 131, un interpolador temporal 132 y una unidad de adaptación de resolución de frecuencia de ventana 133 implementan la funcionalidad de la unidad de decodificación de PSI estándar 115, de la unidad de conversión de resolución de frecuencia 116 y del generador de matriz de segregación 118 de la Figura 11. Un generador de secuencia de ventana 134 y el módulo de análisis de t/f 135 implementan la unidad de transformada de t/f 111 de la Figura 11.

- 60 Normalmente, los contenedores de frecuencia de la representación de tiempo/frecuencia subyacentes se agrupan en bandas paramétricas. El espaciado de las bandas se asemeja al de las bandas críticas en el sistema auditivo humano. Además, múltiples tramas de representación de t/f pueden agruparse en una trama de parámetros. Ambas de estas operaciones proporcionan una reducción en la cantidad de información secundaria requerida con el coste de imprecisiones de modelado.

Como se describe en la norma de SAOC, los OLD e IOC se utilizan para calcular la matriz de segregación $\mathbf{G} = \mathbf{ED}^*\mathbf{J}$,

$$\mathbf{E}(i, j) = \text{IOC}_{i,j} \sqrt{\text{OLD}_i \text{OLD}_j}$$

en la que los elementos de \mathbf{E} se definen a medida que se aproxima a la matriz de correlación cruzada de objeto, i y j son índices de objeto, $\mathbf{J} \approx (\mathbf{DED}^*)^{-1}$. El calculador de matriz de segregación 131 puede llevarse a cabo para calcular la matriz de segregación.

5 La matriz de segregación entonces se interpola linealmente por el interpolador temporal 132 a partir de la matriz de segregación de la trama precedente sobre la trama de parámetros hasta el límite de parámetro en el cual se alcanzan los valores estimados, según el SAOC estándar. Esto resulta en matrices de segregación para cada ventana de análisis de tiempo/frecuencia y banda paramétrica.

10 La resolución de frecuencia de banda paramétrica de las matrices de segregación se expande a la resolución de la representación de tiempo/frecuencia en esa ventana de análisis por la unidad de adaptación de resolución de frecuencia de ventana 133. Cuando la matriz de segregación interpolada para la banda paramétrica b en una trama de tiempo se define como $G(b)$, los mismos coeficientes de segregación se utilizan para todos los contenedores de frecuencia dentro de esa banda paramétrica.

20 El generador de secuencia de ventana 134 se configura para utilizar la información de intervalo de conjunto de parámetros de la PSI para determinar una secuencia de formación de ventanas adecuada para analizar la señal de audio de mezcla descendente de entrada. El requisito principal es que cuando exista un limitador de conjunto de parámetros en la PSI, el punto de cruce entre las ventanas de análisis consecutivas debe coincidir. La formación de ventanas determina también la resolución de frecuencia de los datos dentro de cada ventana (utilizada en la expansión de datos de segregación, como se describe anteriormente).

25 Los datos en ventanas entonces se transforman por el módulo de análisis de t/f 135 en una representación de dominio de frecuencia que utiliza una transformada de tiempo-frecuencia adecuada, por ejemplo, Transformada Discreta de Fourier (DFT) Transformada Discreta de Coseno Modificada Compleja (CMDCT) o Transformada Discreta de Fourier de Apilada Aleatoriamente (ODFT).

30 Finalmente, una unidad de segregación 136 aplica las matrices de segregación de contenedor por trama, por frecuencia de la representación espectral de la señal \mathbf{X} de mezcla descendente para obtener las representaciones paramétricas \mathbf{Y} . El canal de salida j es una combinación lineal de los canales de mezcla descendente

$$\mathbf{Y}_j = \sum_i \mathbf{G}_{j,i} \mathbf{X}_i .$$

35 La calidad que puede obtenerse con este proceso es para la mayor parte de los propósitos indistinguible de manera perceptiva a partir del resultado obtenido con un decodificador de SAOC estándar.

40 Cabe señalar que el texto anterior describe la reconstrucción de objetos individuales, pero en SAOC estándar, la conversión se incluye en la matriz de segregación, es decir, se incluye en interpolación paramétrica. Como operación lineal, el orden de las operaciones no importa, pero la diferencia sí.

A continuación, se describe la decodificación de flujos de bits de SAOC mejorado con un decodificador mejorado.

45 La funcionalidad principal del decodificador de SAOC mejorado ya se describe previamente en la decodificación de flujos de bits de SAOC estándar. Esta sección detallará como las mejoras de SAOC mejorado presentadas en la PSI pueden utilizarse para obtener una mejor calidad perceptiva.

50 La Figura 14 representa los bloques funcionales principales del decodificador de acuerdo con una realización que ilustra la decodificación de las mejoras de resolución de frecuencia. Los bloques funcionales en negrita (141, 142, 143) indican la parte principal del procesamiento inventivo. Una unidad de expansión de valor sobre banda 141, una unidad de recuperación de función delta 142, una unidad de aplicación delta 143, un calculador de matriz de segregación 131, un interpolador temporal 132 y una unidad de adaptación de resolución de frecuencia de ventana 133 implementan la funcionalidad de la unidad de decodificación de PSI mejorada 117 y el generador de matriz de segregación 118 de la Figura 11.

55 El decodificador de la Figura 14 comprende un determinador de información de segregación 112. Entre otros, el determinador información de segregación 112, comprende la unidad de recuperación de función delta 142 y la unidad de aplicación delta 143. La primera información paramétrica comprende una pluralidad de valores paramétricos dependiendo de al menos una señal de objeto de audio, por ejemplo, los valores de diferencia de nivel de objeto. La segunda información paramétrica comprende una parametrización de factor de corrección. La unidad de recuperación de función delta 142 se configura para invertir la parametrización de factor de corrección para obtener una función delta. La unidad de aplicación delta 143 se configura para aplicar la función delta en los valores

paramétricos, por ejemplo, en los valores de diferencia de nivel de objeto, para determinar la información de segregación. En una realización, la parametrización de factor de corrección comprende una pluralidad de coeficientes de predicción lineal, y la unidad de recuperación de función delta 142 se configura para invertir la parametrización de factor de corrección generando una pluralidad de factores de corrección dependiendo de la pluralidad de coeficientes de predicción lineal, y se configura para generar la función delta basándose en la pluralidad de factores de corrección.

Por ejemplo, en primer lugar, la unidad de expansión de valor sobre banda 141 adapta los valores de OLD e IOC para cada banda paramétrica en la resolución de frecuencia utilizada en las mejoras, por ejemplo, en 1024 contenedores. Esto se hace replicando el valor sobre los contenedores de frecuencia que corresponden a la banda paramétrica. Esto resulta en nuevos OLD $OLD_i^{mej}(f) = \mathbf{K}(f, b)OLD_i(b)$ e IOC $IOC_{i,j}^{mej}(f) = \mathbf{K}(f, b)IOC_{i,j}(b)$. $\mathbf{K}(f, b)$ es una matriz de núcleo que define la asignación de contenedores de frecuencia f en bandas paramétricas b .

Paralela a esta, la unidad de recuperación de función delta 142 invierte la parametrización de factor de corrección para obtener la función delta $\mathbf{C}_i^{rec}(f)$ del mismo tamaño que el OLD e IOC expandidos.

A continuación, la unidad de aplicación delta 143 aplica la delta en los valores de OLD expandidos, y los valores de OLD de resolución fina obtenidos se obtienen mediante $OLD_i^{fina}(f) = \mathbf{C}_i^{rec}(f)OLD_i^{mej}(f)$.

En una realización particular, el cálculo de las matrices de segregación, por ejemplo, puede realizarse por el calculador de matriz de segregación 131 con la decodificación del flujo de bits de SAOC estándar:

$\mathbf{G}(f) = \mathbf{E}(f)\mathbf{D}^*(f)\mathbf{J}(f)$, con $\mathbf{E}_{i,j}(f) = IOC_{i,j}^{mej}(f)\sqrt{OLD_i^{fina}(f)OLD_j^{fina}(f)}$, y $\mathbf{J}(f) \approx (\mathbf{D}(f)\mathbf{E}(f)\mathbf{D}^*(f))^{-1}$. Si se desea, la matriz de conversión puede multiplicarse en la matriz de segregación $\mathbf{G}(f)$. La interpolación temporal por el interpolador temporal 132 sigue de acuerdo con el SAOC estándar.

Como la resolución de frecuencia en cada ventana puede ser diferente (menor) de la resolución de frecuencia alta nominal, la unidad de adaptación de resolución de frecuencia de ventana 133 necesita adaptar las matrices de segregación para correlacionar la resolución de los datos espectrales de audio para permitir que los aplique. Esto puede hacerse por ejemplo, al volver a muestrear los coeficientes sobre el eje de frecuencia en la resolución correcta. O si las resoluciones son múltiples enteros, simplemente promediar a partir de los datos de alta resolución

$$\mathbf{G}^{baja}(b) = 1/\|b\| \sum_{f \in b} \mathbf{G}(f).$$

los índices que corresponden a un contenedor de frecuencia en la resolución más baja

La información de secuencia de formación de ventanas del flujo de bits puede utilizarse para obtener un análisis de tiempo-frecuencia completamente complementario al utilizado en el codificador, o la secuencia de formación de ventanas puede construirse basándose en los limitadores de parámetros, como se hace en la decodificación de flujo de bits de SAOC estándar. Para esto, puede emplearse un generador de secuencia de ventana 134.

El análisis de tiempo-frecuencia del audio en la mezcla descendente entonces se lleva a cabo por un módulo de análisis de tf 135 utilizando las ventanas dadas.

Finalmente, las matrices de segregación temporalmente interpoladas y espectralmente adaptadas (posiblemente) se aplican por una unidad de segregación 136 en la representación de tiempo-frecuencia del audio de entrada, y el canal de salida j puede obtenerse como una combinación lineal de los canales de entrada $\mathbf{Y}_j(f) = \sum_i \mathbf{G}_{j,i}^{baja}(f)\mathbf{X}_i(f)$.

A continuación, se describen aspectos particulares de las realizaciones.

En una realización, la unidad 109 de modelado delta de la Figura 10 se configura para determinar coeficientes de predicción lineal a partir de una pluralidad de factores de corrección (*delta*) realizando una predicción lineal.

Ahora, se describe el proceso de estimación del factor de corrección $\mathbf{G}^{low}(b) = 1/\|b\| \sum_{f \in b} \mathbf{G}(f)$. alternativo utilizando coeficientes de predicción lineal (LPC) de acuerdo con una real

En primer lugar, se describe la estimación delta de acuerdo con una realización.

La entrada en la estimación consiste en los perfiles espectrales de potencia de resolución fina estimados sobre el bloque de parámetros y a partir de la reconstrucción aproximada del perfil de espectral de potencia basándose en los parámetros de OLD y NRG. Los perfiles de espectro de potencia fina se calculan de la siguiente manera. $\mathbf{S}_i(f, n)$

es el espectro complejo del ^{ésimo} objeto siendo f el índice de contenedor de frecuencia y siendo $0 \leq n \leq N-1$ el índice de ventana temporal en el bloque de modelado de la longitud N . El espectro de potencia de resolución fina entonces es

$$P_i(f) = \frac{1}{N} \sum_{n=0}^{N-1} S_i(f, n) S_i^*(f, n).$$

5 La reconstrucción aproximada se calcula a partir de los OLD y NRG (decuantificado) mediante

$$Z_i(f) = K(f, b) OLD_i(b) NRG_i(b),$$

10 en la que $K(f, b)$ es la matriz de núcleo que define la asignación de contenedores de frecuencia f en las bandas paramétricas b .

15 Se utilizarán dos señales con diferentes propiedades espectrales como ejemplos en esta sección: la primera es ruido (rosa) con espectro prácticamente plano (ignorando la inclinación espectral) y el segundo es un tono del instrumento carrillón que tiene un espectralmente tonal, es decir, en pico.

20 La Figura 15 ilustra los espectros de potencia de una señal tonal y de ruido. Sus espectros de potencia de alta resolución ("orig") y las reconstrucciones aproximadas correspondientes basadas en OLD y NRG ("recon"). En particular, la Figura 15 ilustra los espectros de potencia fina y aproximada de ambas señales. Más particularmente, se muestran los espectros de potencia de una señal tonal 151 original y una señal de ruido original, y los espectros de potencia reconstruidos de la señal tonal 153 y la señal de ruido 154. Cabe señalar que, en las siguientes figuras, para las señales 153 y 154 en lugar de los factores de escala (parámetros de espectros de potencia reconstruidos) y no las señales completamente reconstruidas se representan.

25 Puede observarse rápidamente, la diferencia promedio entre el valor fino y aproximado es más bien pequeña en el caso de la señal de ruido, mientras las diferencias son muy grandes en la señal tonal. Estas diferencias provocan degradaciones perceptivas en la reconstrucción paramétrica de todos los objetos.

30 El factor de corrección se obtiene dividiendo la curva de resolución fina por la curva de reconstrucción aproximada:

$$C_i(f) = P_i(f) / Z_i(f).$$

35 Esto permite recuperar un factor multiplicativo que puede aplicarse en la reconstrucción aproximada para obtener la curva de resolución fina:

$$P_i^{rec}(f) = Z_i(f) C_i(f).$$

40 La Figura 16 ilustra la modificación para ambas señales ilustrativas, en particular, los factores de corrección para las señales ilustrativas. En particular, se muestran los factores de corrección para la señal tonal 151 y la señal 152 de ruidos.

A continuación, se describe el modelado delta.

45 La curva C de corrección se asigna en uno más bloques de modelado sobre el eje de frecuencia. Una alternativa natural es utilizar las mismas definiciones de banda de parámetro que se utilizan para la PSI de SAOC estándar. El modelado entonces se hace para cada bloque por separado con las siguientes etapas:

50 1. El factor C de corrección espectral se transforma en la secuencia de auto correlación de dominio de tiempo con Transformada Discreta de Fourier Inversa (IDFT).
Cuando la longitud del bloque de modelado es impar, el pseudo espectro que se transforma se define como

$$R(l) = \begin{cases} C(l), & \text{cuando } 0 \leq l \leq N-1 \\ C(2N-2-l), & \text{cuando } N \leq l \leq 2N-3 \end{cases}$$

55 Cuando el bloque de modelado es par, el pseudo espectro se define como

$$R(l) = \begin{cases} C(l), & \text{cuando } 0 \leq l \leq N-1 \\ C(2N-1-l), & \text{cuando } N \leq l \leq 2N-2 \end{cases}$$

El resultado de transformada entonces es $r(t) = \text{IDFT}(\mathbf{R}(l))$.

2. El resultado se trunca en la primera mitad:

5

$$\begin{cases} 0 \leq t \leq N-2, & \text{cuando } N \text{ es impar} \\ 0 \leq t \leq N-1, & \text{cuando } N \text{ es par} \end{cases}$$

3. Las recursiones de Levinson-Durbin se aplican en la secuencia de auto correlación $r(t)$ para obtener los coeficientes de reflexión \mathbf{k} y las varianzas residuales de modelado \mathbf{e} para incrementar los órdenes de modelo.

10

4. Opcional: basándose en la varianza residual de modelado \mathbf{e} , omitir el modelado (ya que no se obtuvo ninguna ganancia) o seleccionar un orden adecuado.

5. Los parámetros de modelos se cuantifican para transmisión.

15

Es posible tomar una decisión si la delta debe transmitirse por cada cinta de t-f (banda paramétrica estándar que define el margen de frecuencia y el bloque de parámetros del margen temporal) en forma independiente. La decisión puede tomarse basándose, por ejemplo, en

20 - Inspeccionar la energía residual de modelado delta. Si la energía residual de modelado no excede un cierto umbral, no se transmite la información de mejora.

25 - Medir la "angularidad"/desigualdad de la descripción paramétrica modelada en resolución fina, el modelado delta, o el envolvente espectral de potencia de la señal de objeto de audio. Dependiendo del valor medido, los parámetros de modelado delta, que describen la resolución espectral fina, se transmiten o no, o se calculan en absoluto dependiendo de la desigualdad del envolvente espectral de potencia de la señal objeto de audio). Medidas adecuadas por ejemplo, son el factor de cresta espectral, la medida de planitud espectral, o la relación mínima-máxima.

30 - La calidad perceptiva de la reconstrucción obtenida. El codificador calcula las reconstrucciones de conversión con y sin las mejoras y determina la ganancia de calidad para cada mejora. Después, se ubica el punto de equilibrio adecuado entre la complejidad de modelado y la ganancia de calidad, y las mejoras indicadas se transmiten. Por ejemplo, una relación de distorsión a señal perceptiblemente ponderada o medidas perceptivas mejoradas pueden utilizarse para la decisión. La decisión puede tomarse para cada banda paramétrica (aproximada) por separado (es decir, optimización de calidad local), pero también bajo consideración de bandas adyacentes para justificar las distorsiones de señal provocadas por la manipulación variante de tiempo y frecuencia de los coeficientes de tiempo-frecuencia (es decir, optimización de calidad global).

35

Ahora, se describe la reconstrucción y aplicación delta.

40

La reconstrucción de la curva de corrección sigue las etapas:

1. Los coeficientes de reflexión recibidos \mathbf{k} (un vector de la longitud $L-1$) se decuantifican y transmiten en coeficientes de filtro IIR \mathbf{a} de la longitud L , en sintaxis de pseudo código (donde la función $X = \text{diag}(x)$ produce una matriz X con los elementos diagonales de X siendo x y todos los elementos no diagonales de X siendo nulos):

45

```

A = diag(k)
for ii=1 to L
    for l=1 to ii-1
        A(l, ii) = A(l, ii-1) + k(ii)*A(ii-1, ii-1)
    end
end
a = [1; A(1 to end, end)]

```

50

55

$$\mathbf{h}(n) = 1 / \sum_{l=0}^{L-1} \mathbf{a}(l) \exp(-in2\pi / N), \quad \text{donde } i$$

2. La respuesta de frecuencia $\mathbf{h}(n)$ del filtro resultante \mathbf{a} se calcula con $i = \sqrt{-1}$. indica la unidad imaginaria

3. La reconstrucción de función de corrección se obtiene a partir de esto mediante $\mathbf{C}^{raw}(n) = \mathbf{h}(n)\mathbf{h}^*(n)$.

4. La respuesta se normaliza para tener una media unitaria, de manera que la energía general del bloque

$$C^{rec}(n) = C^{raw}(n) / \sum_{n=0}^{N-1} C^{raw}(n).$$

modelado no cambia

5. Los factores de corrección se aplican en OLD, que se han extendido a la resolución fina $OLD_i^{fina}(f) = C_i^{rec}(f)K(f,b)OLD_i(b)$. Obsérvese que en las energías absolutas pueden ignorarse ya que pueden cancelarse en los cálculos adicionales.

La Figura 17 ilustra los factores de corrección originales y las aproximaciones basadas en LPC de orden reducido (después del modelado) para ambas de las señales ilustrativas. En particular, se muestran los factores de corrección originales de la señal tonal 151, la señal de ruido original 152 y las estimaciones de factor de corrección reconstruidas de la señal tonal 153 y la señal de ruido 154.

La Figura 18 ilustra el resultado de aplicar los factores de corrección modelada en las reconstrucciones aproximadas ilustradas en la Figura 15. En particular, se muestran los espectros de potencia de la señal tonal original 151 y la señal de ruido original 152, y las estimaciones de espectros de potencia reconstruidos de la señal tonal 153 y la señal de ruido 154. Estas curvas ahora pueden utilizarse en lugar de OLD en los siguientes cálculos, en particular, los espectros de potencia de resolución fina reconstruidos después de aplicar los factores de corrección modelada. Aquí, la información de energía absoluta se incluye para hacer la comparación mejor visible, pero el mismo principio funciona también sin las mismas.

El método inventivo y aparato superan las ventajas antes mencionadas del procesamiento de SAOC del estado de la técnica utilizando un banco de filtro o transformada de tiempo-frecuencia con una alta resolución de frecuencia y proporcionan una parametrización eficiente de la información adicional. Además, es posible transmitir esta información adicional de tal manera que los decodificadores de SAOC estándar puedan decodificar la porción retrocompatible de la información en una calidad obtenible comparable con una obtenida utilizando un codificador de SAOC que cumple con el estándar, y aún permitir que los decodificadores mejorados utilicen la información adicional para una mejor calidad perceptiva. Más importante, la información adicional puede representarse de una forma muy compacta para transmisión o almacenamiento eficiente.

El método inventivo presentado puede aplicarse a cualquier esquema de SAOC. Puede combinarse con cualquier formato de audio actual y también futuro. El método inventivo permite una calidad de audio perceptiva mejorada en aplicaciones de SAOC por una representación de dos niveles de información secundaria espectral. La misma idea puede utilizarse también junto con MPEG envolvente cuando se sustituye el concepto de OLD con diferencias de nivel de canal (CLD).

Se proporciona un codificador de audio o método de codificación de audio o programa informático relacionado como se describe anteriormente. Además, se proporciona un codificador de audio o método de decodificación de audio o programa informático relacionado como se describe anteriormente. Además, se proporciona una señal de audio codificada o medio de almacenamiento que tiene almacenada la señal de audio codificada como se describe anteriormente.

Aunque algunos aspectos se han descrito en el contexto de un aparato, es claro que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a una etapa de método o característica de una etapa de método. De forma análoga, aspectos descritos en el contexto de una etapa de método también representan una descripción de un bloque correspondiente o artículo o característica de un aparato correspondiente.

La señal descompuesta inventiva puede almacenarse en un medio de almacenamiento digital o puede transmitirse en un medio de transmisión tal como un medio de transmisión inalámbrica o un medio de transmisión alámbrica tal como la Internet.

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

Algunas realizaciones de acuerdo con la invención comprenden una portadora de datos no transitorios que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema informático programable, de manera que se realiza uno de los métodos descritos en este documento.

5 Generalmente, 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 de programa por ejemplo, puede almacenarse en una portadora legible por máquina.

Otras realizaciones comprenden el programa informático para realizar uno de los métodos descritos en este documento, almacenado en una portadora legible por máquina.

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

15 Una realización adicional de los métodos inventivos por lo tanto, es una portadora de datos (o un medio de almacenamiento digital o un medio legible por ordenador) que comprende, grabado en la misma, el programa informático para realizar uno de los métodos descritos en este documento.

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

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

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

30 En algunas realizaciones, un dispositivo lógico programable (por ejemplo, un campo de matrices de puertas programables) puede utilizarse para realizar algunas o todas las funcionalidades de los métodos descritos en este documento. En algunas realizaciones, un campo de matrices de puertas programables puede cooperar con un microprocesador para realizar uno de los métodos descritos en este documento. Generalmente, los métodos se realizan preferentemente mediante cualquier aparato de hardware.

35 Las realizaciones descritas anteriormente son solamente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en este documento serán aparentes para los expertos en la técnica. Por lo tanto, la intención es que se limite únicamente por el alcance de las reivindicaciones de patente a continuación y no por los detalles específicos presentados por medio de descripción y explicación de las realizaciones en este documento.

40

Referencias

- [BCC] C. Faller and 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.
- 45 [JSC] C. Faller, "Parametric Joint-Coding of Audio Sources", 120th AES Convention, París, 2006.
- [SAOC1] J. Herre, S. Disch, J. Hilpert, O. Hellmuth: "From SAC To SAOC - Recent Developments in Parametric Coding of Spatial Audio", 22nd Regional UK AES Conference, Cambridge, Reino Unido, abril de 2007.
- [SAOC2] J. Engdegård, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Hölzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers and W. Oomen: " Spatial Audio Object Coding (SAOC) - The Upcoming MPEG Standard on Parametric Object Based Audio Coding", 124th AES Convention, Ámsterdam, 2008.
- 50 [SAOC] ISO/IEC, "MPEG audio technologies - Part 2: Spatial Audio Object Coding (SAOC)," ISO/IEC JTC1/SC29/WG11 (MPEG) Norma Internacional 23003-2:2010.
- [AAC] M .Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, "ISO/IEC MPEG-2 Advanced Audio Coding", J. Audio Eng. Soc, vol 45, n.º 10, pp. 789-814, 1997.
- 55 [1551] M. Parvaix and L. Girin: "Informed Source Separation of underdetermined instantaneous Stereo Mixtures using Source Index Embedding", IEEE ICASSP, 2010.
- [ISS2] M. Parvaix, L. Girin, J.-M. Brossier: "A watermarking-based method for informed source separation of audio signals with a single sensor", IEEE Transactions on Audio, Speech and Language Processing, 2010.
- [ISS3] A. Liutkus and J. Pinel and R. Badeau and L. Girin and G. Richard: "Informed source separation through spectrogram coding and data embedding", Signal Processing Journal, 2011.
- 60 [ISS4] A. Ozerov, A. Liutkus, R. Badeau, G. Richard: "Informed source separation: source coding meets source separation", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2011.
- [ISS5] S. Zhang and L. Girin: "An Informed Source Separation System for Speech Signals", INTERSPEECH, 2011.

[ISS6] L. Girin and J. Pinel: "Informed Audio Source Separation from Compressed Linear Stereo Mixtures", AES 42nd International Conference: Semantic Audio, 2011.

[ISS7] A. Nesbit, E. Vincent, and M. D. Plumbley: "Benchmarking flexible adaptive time-frequency transforms for underdetermined audio source separation", IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 37-40, 2009.

5

REIVINDICACIONES

1. Un decodificador para generar una señal de audio no mezclada que comprende una pluralidad de canales de audio no mezclados, en el que el decodificador comprende:
- 5 un determinador de información de segregación (112) para determinar información de segregación recibiendo primera información secundaria paramétrica en la al menos una señal de objeto de audio y segunda información secundaria paramétrica en la al menos una señal de objeto de audio, en el que la resolución de frecuencia de la segunda información secundaria paramétrica es mayor que la resolución de frecuencia de la primera información secundaria paramétrica, y
- 10 un módulo de segregación (113) para aplicar la información de segregación en una señal de mezcla descendente, indicando una mezcla descendente de al menos una señal de objeto de audio, para obtener una señal de audio no mezclada que comprende la pluralidad de canales de audio no mezclados, en el que el determinador de información de segregación (112) se configura para determinar la información de segregación modificando la primera información paramétrica y la segunda información paramétrica para obtener
- 15 información paramétrica modificada, de tal forma que la información paramétrica modificada tiene una resolución de frecuencia que es mayor que la primera resolución de frecuencia.
2. Un decodificador de acuerdo con la reivindicación 1, en el que el decodificador comprende además una primera unidad de transformada (111) para transformar una entrada de mezcla descendente, que se representa en un dominio de tiempo, para obtener la señal de mezcla descendente, que se representa en un dominio de tiempo-frecuencia, y en el que el decodificador comprende una segunda unidad de transformada (114) para transformar la señal de audio no mezclada a partir del dominio de tiempo-frecuencia al dominio del tiempo.
- 20 3. Un decodificador de acuerdo con la reivindicación 1 o 2, en el que el determinador de información de segregación (112) se configura para determinar la información de segregación combinando la primera información paramétrica y la segunda información paramétrica para obtener la información paramétrica modificada, de tal forma que la información paramétrica modificada tiene una resolución de frecuencia que es igual a la segunda resolución de frecuencia.
- 25 30 4. Un decodificador de acuerdo con una de las reivindicaciones anteriores, en el que la primera información paramétrica comprende una pluralidad de primeros valores de parámetro, en el que la segunda información paramétrica comprende una pluralidad de segundos valores de parámetro, en el que el determinador de información de segregación (112) comprende una subunidad de conversión de resolución de frecuencia (122) y un combinador (124),
- 35 en el que la unidad de conversión de resolución de frecuencia (112) se configura para generar valores de parámetro adicionales, en el que los primeros valores de parámetro y los valores de parámetro adicionales juntos forman una pluralidad de primeros valores de parámetro procesador, y en el que el combinador (124) se configura para combinar los primeros valores de parámetro procesador y los
- 40 segundos valores de parámetro para obtener una pluralidad de valores de parámetro modificados como la información paramétrica modificada.
5. Un decodificador de acuerdo con una de las reivindicaciones anteriores, en el que el determinador de información de segregación (112) comprende una unidad de recuperación de función delta (142) y una unidad de aplicación delta (143),
- 45 en el que la primera información paramétrica comprende una pluralidad de valores paramétricos dependiendo de la al menos una señal de objeto de audio, y en el que la segunda información paramétrica comprende una parametrización de factor de corrección, en el que la unidad de recuperación de función delta (142) se configura para invertir la parametrización de factor de corrección para obtener una función delta, y
- 50 en el que la unidad de aplicación delta (143) se configura para aplicar la función delta en los valores paramétricos para determinar la información de segregación.
6. Un decodificador de acuerdo con la reivindicación 5, en el que la parametrización de factor de corrección comprende una pluralidad de coeficientes de predicción lineal, en el que la unidad de recuperación de función delta (142) se configura para invertir la parametrización de factor de corrección generando una pluralidad de factores de corrección dependiendo de la pluralidad de coeficientes de predicción lineal, y
- 55 en el que la unidad de recuperación de función delta (142) se configura para generar la función delta basándose en la pluralidad de factores de corrección.
- 60 7. Un decodificador de acuerdo con una de las reivindicaciones anteriores, en el que el decodificador comprende además un generador de matriz de segregación (118) para generar una matriz de segregación dependiendo de la primera información secundaria paramétrica, dependiendo de la segunda

información secundaria paramétrica y dependiendo de información de conversión, y en el que el módulo de segregación (113) se configura para aplicar la matriz de segregación en la mezcla descendente transformada para obtener la señal de audio no mezclada.

- 5 8. Un decodificador de acuerdo con una de las reivindicaciones anteriores, en el que el módulo de segregación (113) comprende una unidad de decorrelación (119) y una unidad de segregación (110),
 en el que la unidad de decorrelación (119) se configura para realizar decorrelación en la mezcla descendente transformada para obtener un resultado de decorrelación, y
 10 en el que la unidad de segregación (110) se configura para emplear el resultado de decorrelación para obtener la señal de audio no mezclada.

9. Un codificador para codificar una o más señales de objeto de audio, que comprende:

- 15 una unidad de mezcla descendente (91) para mezclar de forma descendente la una o más señales de objeto de audio para obtener una o más señales de mezcla descendente, y
 un generador de información secundaria paramétrica (93) para generar primera información secundaria paramétrica en la al menos una señal de objeto de audio y segunda información secundaria paramétrica en la al menos una señal de objeto de audio, de tal forma que la resolución de frecuencia de la segunda información
 20 secundaria paramétrica es mayor que la resolución de frecuencia de la primera información secundaria paramétrica,
 en el que el generador de información secundaria paramétrica (93) se configura para generar la primera información secundaria paramétrica y la segunda información secundaria paramétrica de modo que la una o más
 25 señales de objeto de audio se puede decodificar a partir de la una o más señales de mezcla descendente usando la primera información secundaria paramétrica junto con la segunda información secundaria paramétrica.

10. Un codificador de acuerdo con la reivindicación 9, en el que el codificador comprende además una unidad de transformada (92) para transformar la una o más señales de objeto de audio a partir de un dominio de tiempo a un dominio de tiempo-frecuencia para obtener una o más
 30 señales de objeto de audio transformadas, y
 en el que el generador de información secundaria paramétrica (93) se configura para generar la primera información secundaria paramétrica y la segunda información secundaria paramétrica basándose en la una o más señales de objeto de audio transformadas.

- 35 11. Un codificador de acuerdo con la reivindicación 10, en el que la unidad de transformada (92) se configura para transformar la una o más señales de objeto de audio a partir del dominio del tiempo al dominio de tiempo-frecuencia dependiendo de una longitud de ventana de un bloque de transformación de señal que comprende valores de señal de al menos una de la una o más señales de objeto de audio,
 40 en el que la unidad de transformada (92) comprende una unidad de detección transitoria (101) para determinar un resultado de detección transitorio indicando si un transitorio está presente en uno o más de la al menos una señal de objeto de audio, en el que un transitorio indica un cambio de señal en uno o más de la al menos una señal de objeto de audio, y
 45 en el que la unidad de transformada (92) comprende además una unidad de secuencia de ventana (102) para determinar la longitud de ventana dependiendo del resultado de detección transitorio.

12. Un codificador de acuerdo con una de las reivindicaciones 9 a 11, en el que el codificador comprende además una unidad de estimación de secuencia (108) para estimar una pluralidad de factores de corrección basándose en una pluralidad de valores paramétricos dependiendo de la al menos una señal de objeto de audio para obtener la
 50 segunda información secundaria paramétrica.

13. Un codificador de acuerdo con la reivindicación 12, en el que el codificador comprende además una unidad de modelado delta (109) para determinar coeficientes de predicción lineal a partir de la pluralidad de factores de corrección realizando una predicción lineal.
 55

14. Un sistema que comprende:

- un codificador (61) de acuerdo con una de las reivindicaciones 9 a 13 para codificar una o más señales de objeto de audio obteniendo una o más señales de mezcla descendente indicando una mezcla descendente de una o
 60 más señales de objeto de audio, obteniendo primera información secundaria paramétrica en la al menos una señal de objeto de audio y obteniendo segunda información secundaria paramétrica en la al menos una señal de objeto de audio, en el que la resolución de frecuencia de la segunda información secundaria paramétrica es mayor que la resolución de frecuencia de la primera información secundaria paramétrica, y
 un decodificador (62) de acuerdo con una de las reivindicaciones 1 a 8 para generar una señal de audio no

mezclada basándose en la una o más señales de mezcla descendente y basándose en la primera información secundaria paramétrica y la segunda información secundaria paramétrica.

5 15. Un método para generar una señal de audio no mezclada que comprende una pluralidad de canales de audio no mezclados, en el que el método comprende:

10 determinar información de segregación recibiendo primera información secundaria paramétrica en la al menos una señal de objeto de audio y segunda información secundaria paramétrica en la al menos una señal de objeto de audio, en el que la resolución de frecuencia de la segunda información secundaria paramétrica es mayor que la resolución de frecuencia de la primera información secundaria paramétrica, y
15 aplicar la información de segregación en una señal de mezcla descendente, indicando una mezcla descendente de al menos una señal de objeto de audio, para obtener una señal de audio no mezclada que comprende la pluralidad de canales de audio no mezclados,
en el que determinar la información de segregación comprende modificar la primera información paramétrica y la segunda información paramétrica para obtener información paramétrica modificada, de tal forma que la información paramétrica modificada tiene una resolución de frecuencia que es mayor que la primera resolución de frecuencia.

20 16. Un método para codificar una o más señales de objeto de audio, que comprende:

mezclar de forma descendente la una o más señales de objeto de audio para obtener una o más señales de mezcla descendente, y
25 generar primera información secundaria paramétrica en la al menos una señal de objeto de audio y segunda información secundaria paramétrica en la al menos una señal de objeto de audio, de tal forma que la resolución de frecuencia de la segunda información secundaria paramétrica es mayor que la resolución de frecuencia de la primera información secundaria paramétrica,
en el que la primera información secundaria paramétrica y la segunda información secundaria paramétrica se generan de modo que la una o más señales de objeto de audio se puede decodificar a partir de la una o más
30 señales de mezcla descendente usando la primera información secundaria paramétrica junto con la segunda información secundaria paramétrica.

17. Un programa informático para implementar el método de la reivindicación 15 o 16 cuando se ejecuta en un ordenador o procesador de señal.

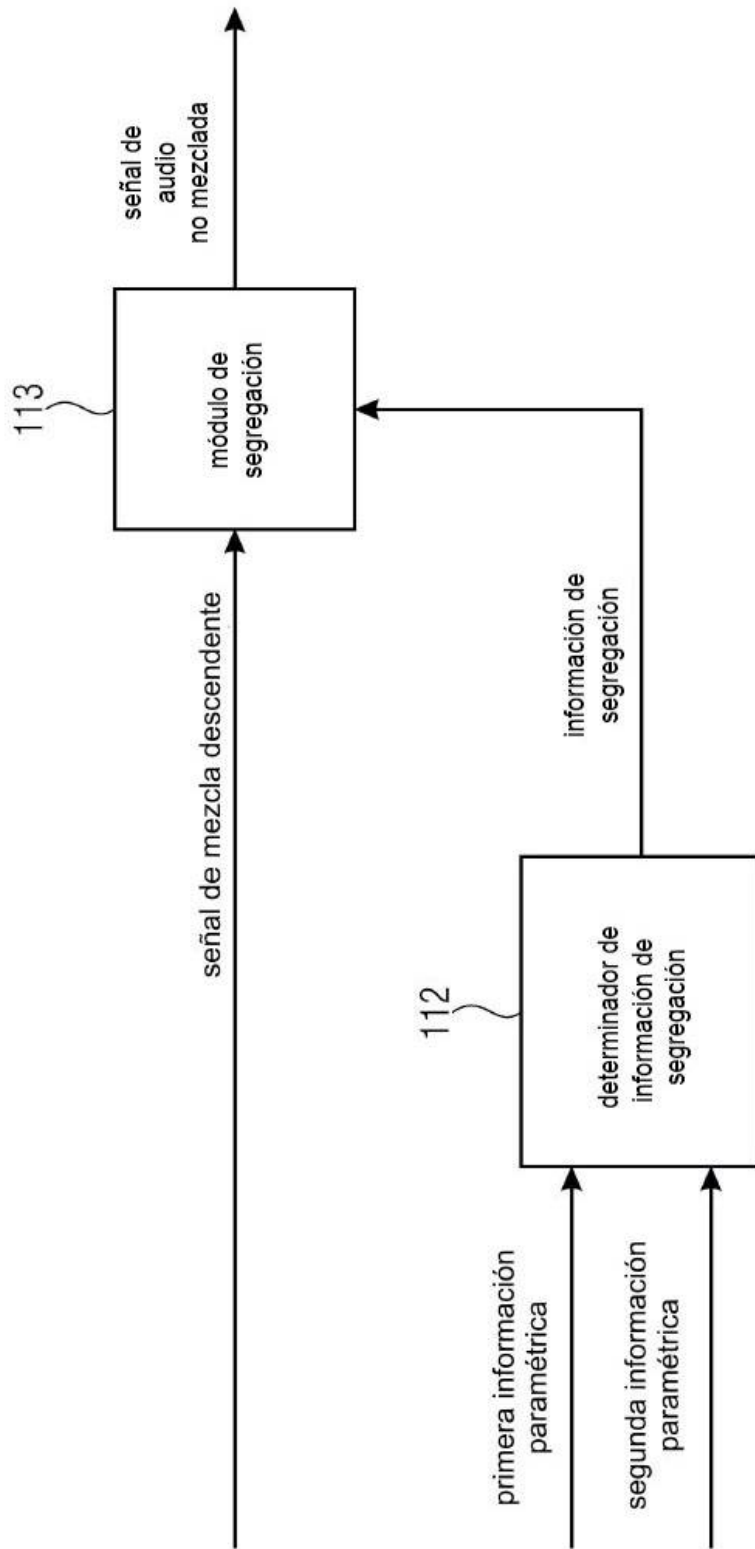


FIGURA 1A

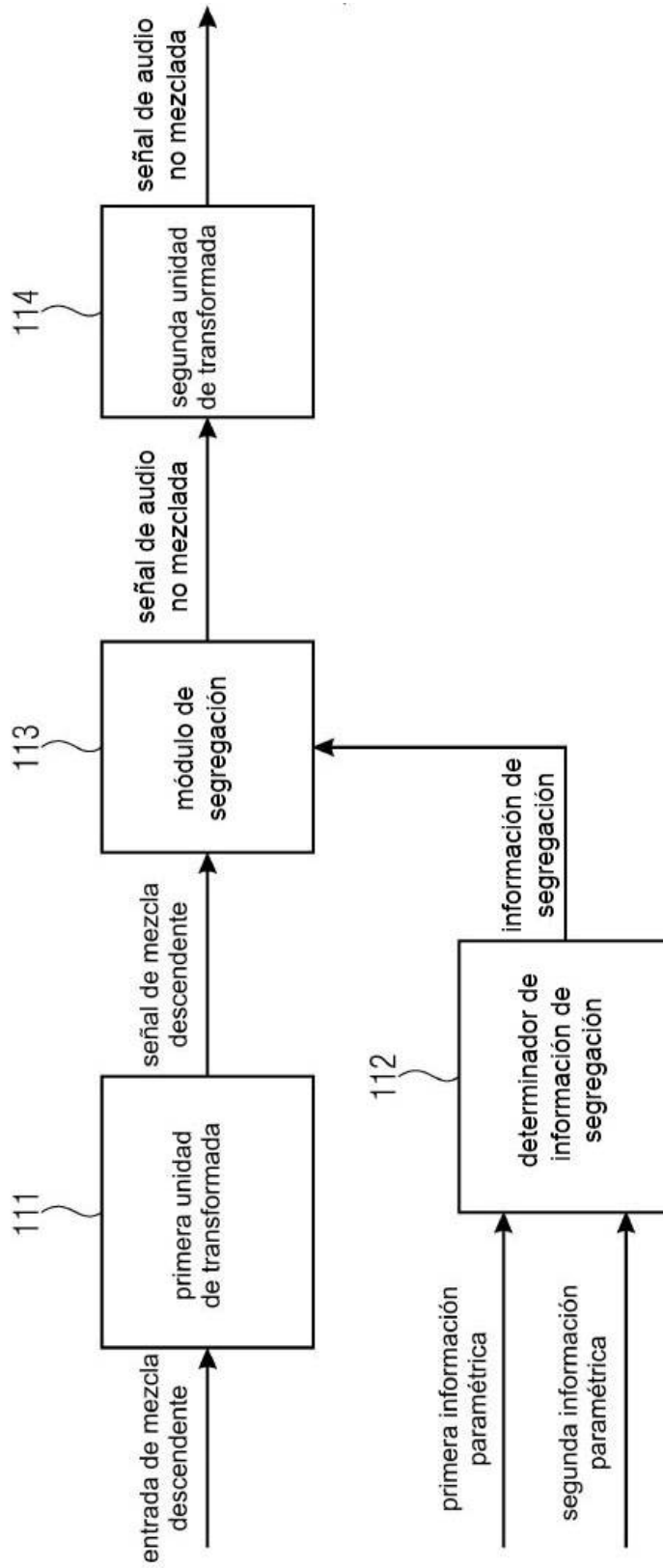


FIGURA 1B

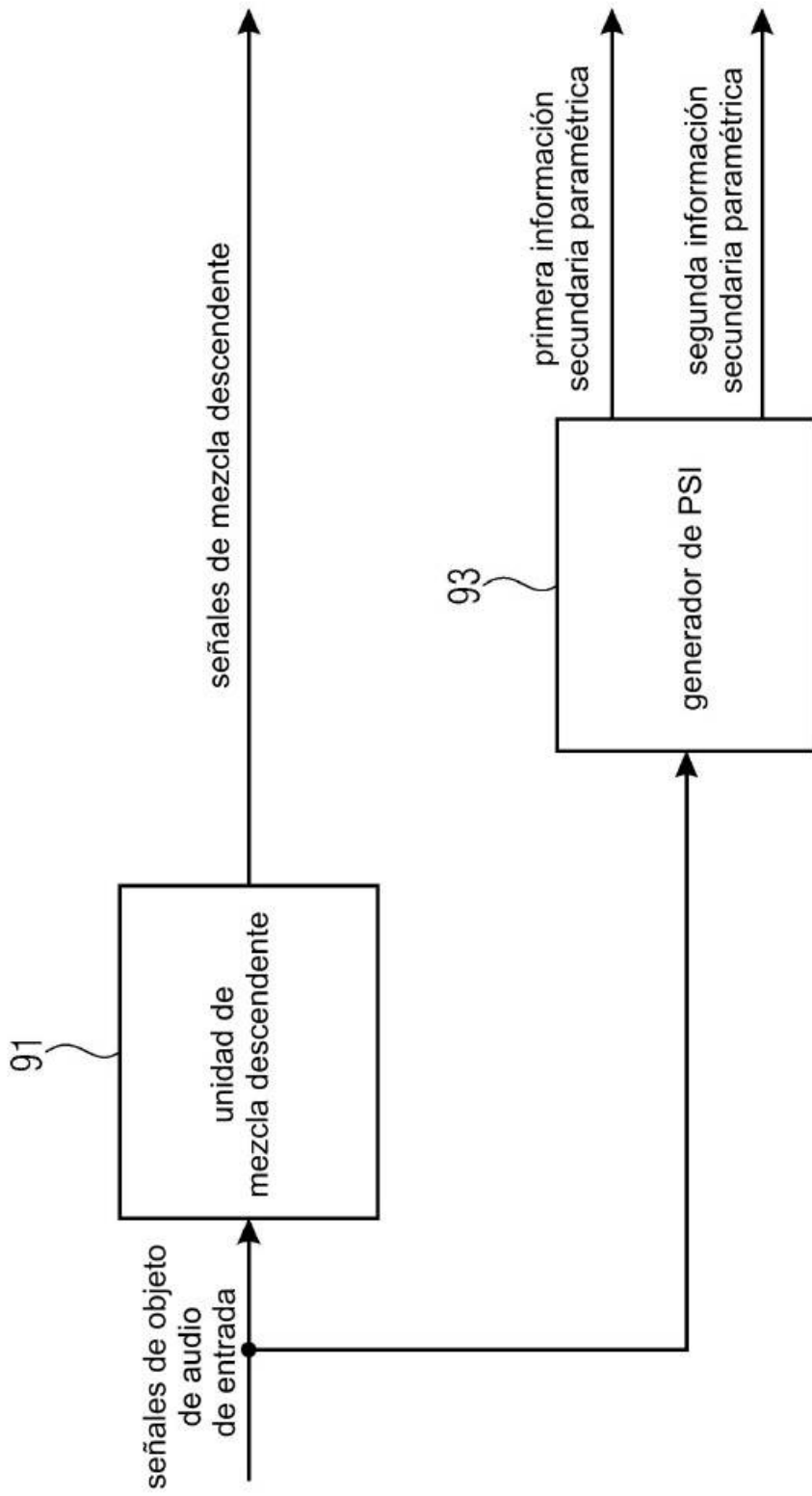


FIGURA 2A

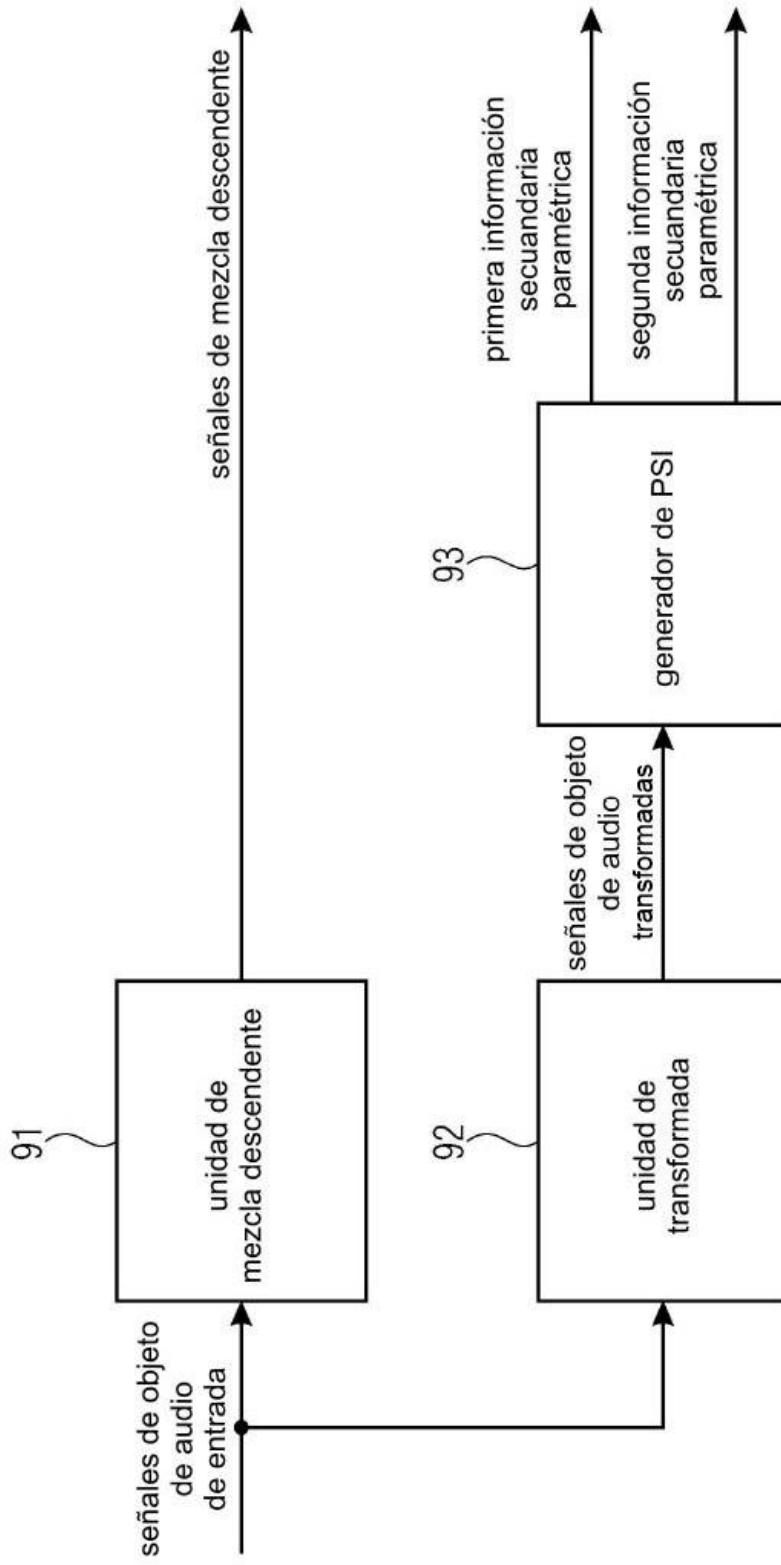


FIGURA 2B

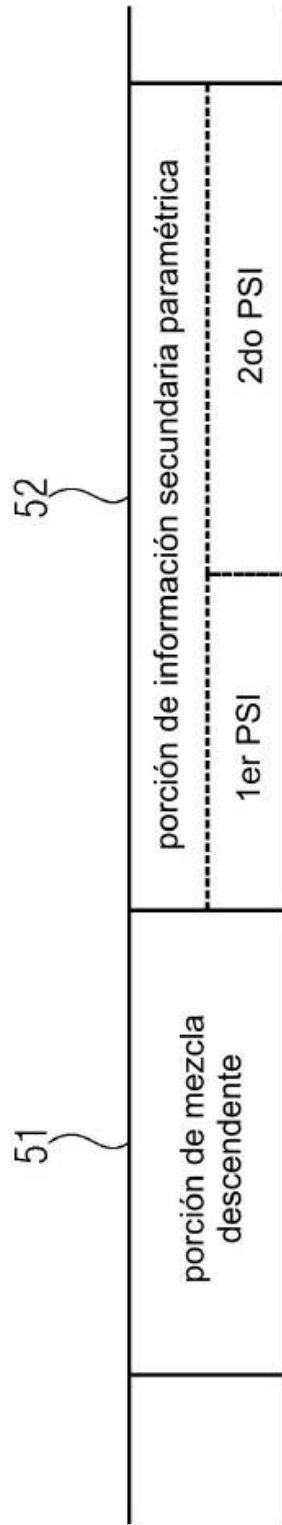


FIGURA 2C

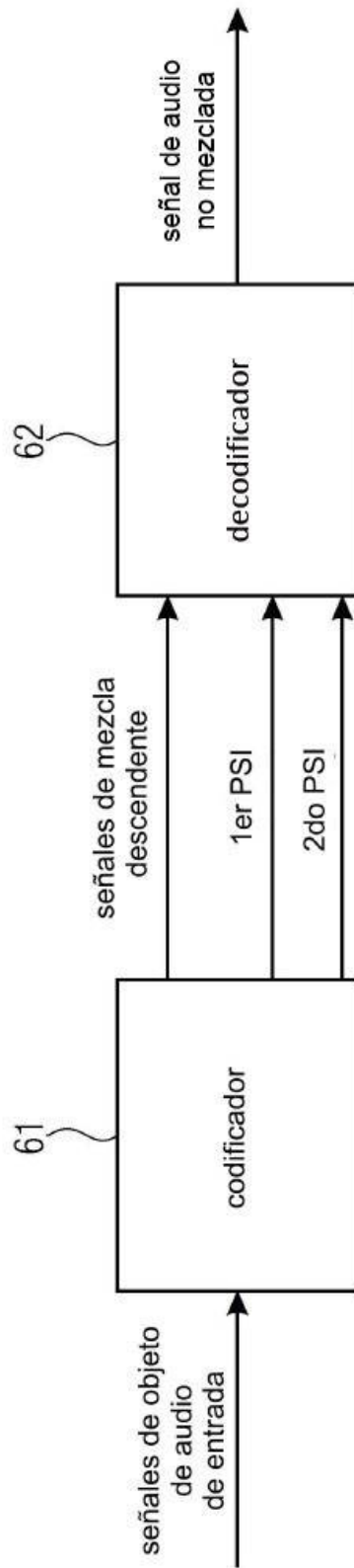


FIGURA 3

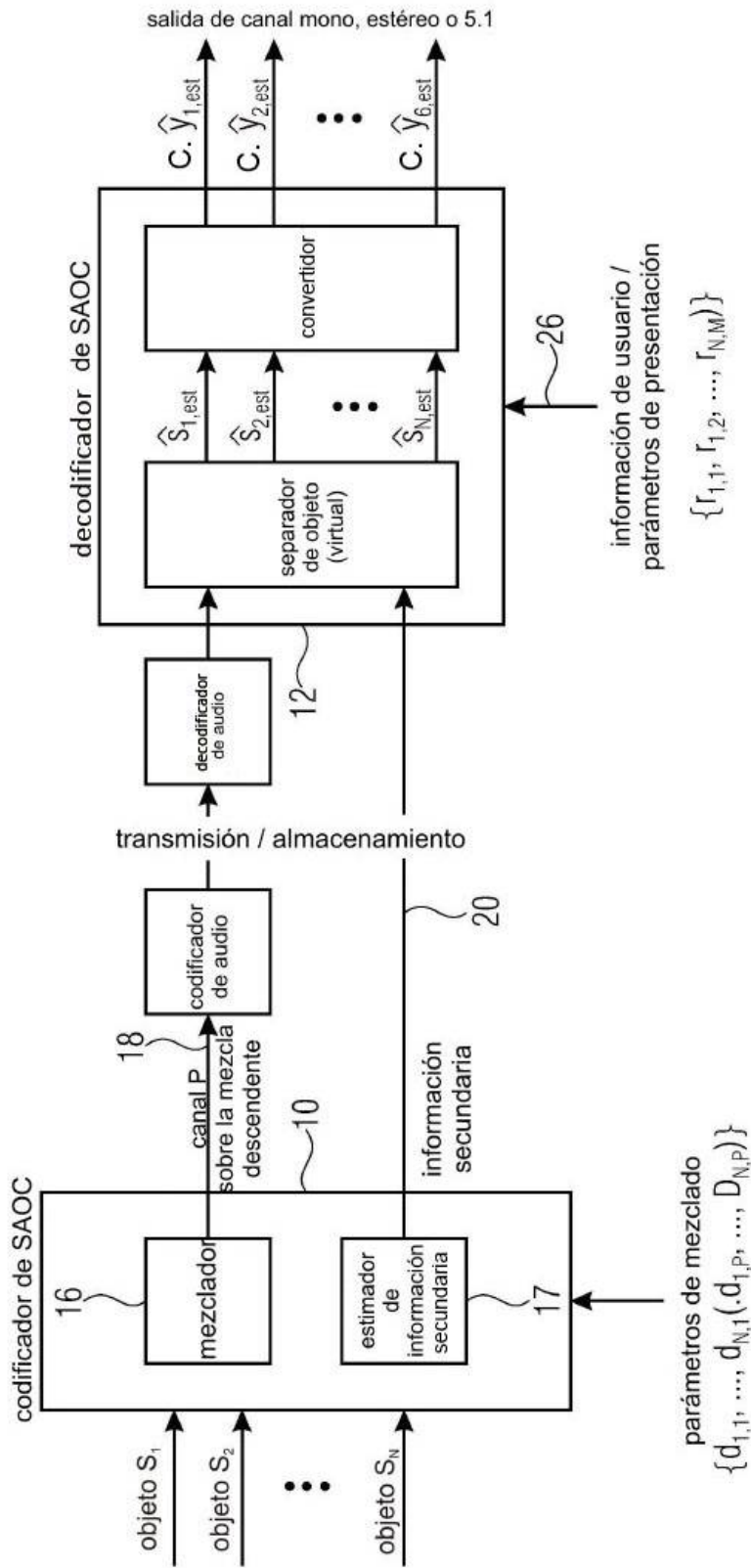


FIGURA 4

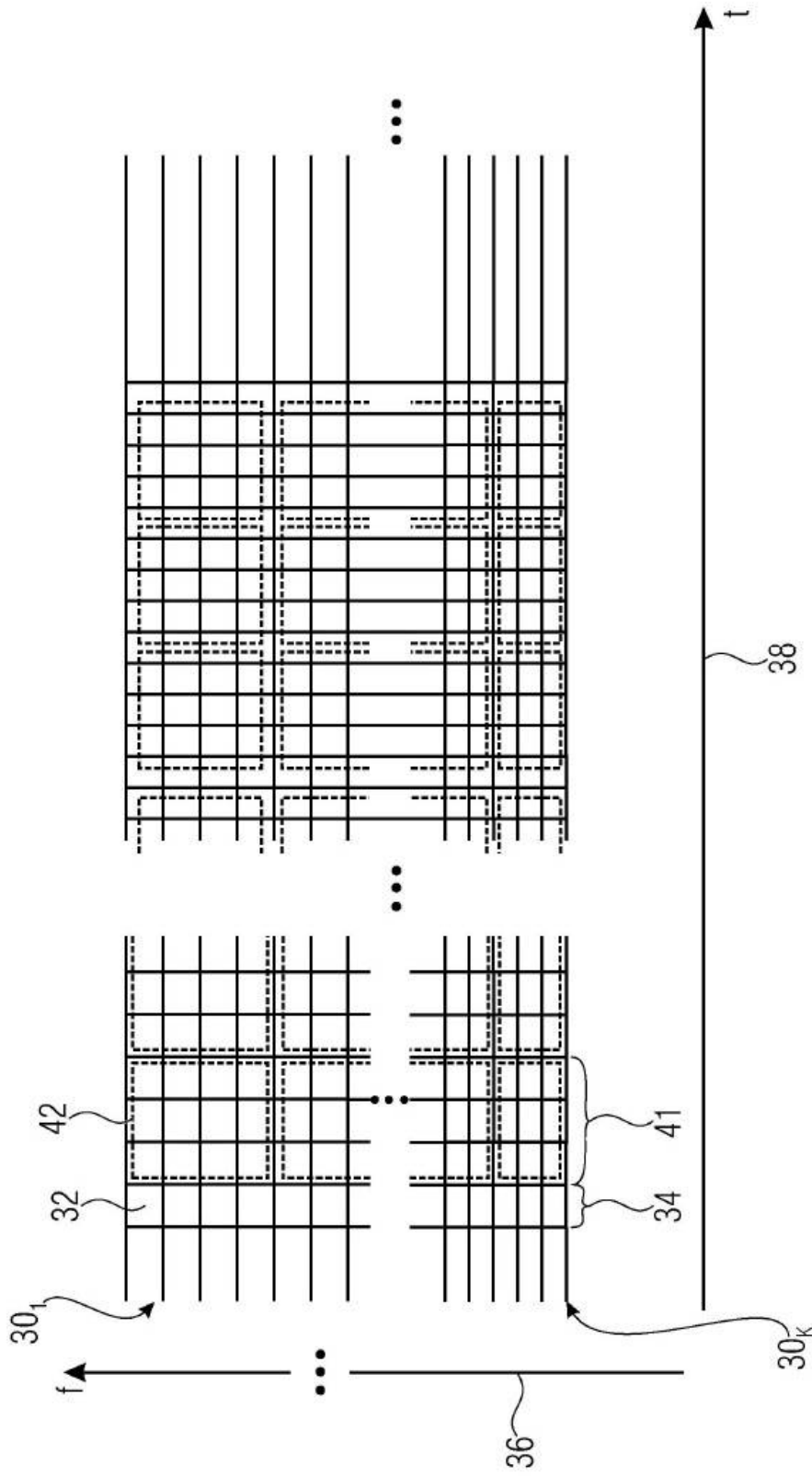


FIGURA 5

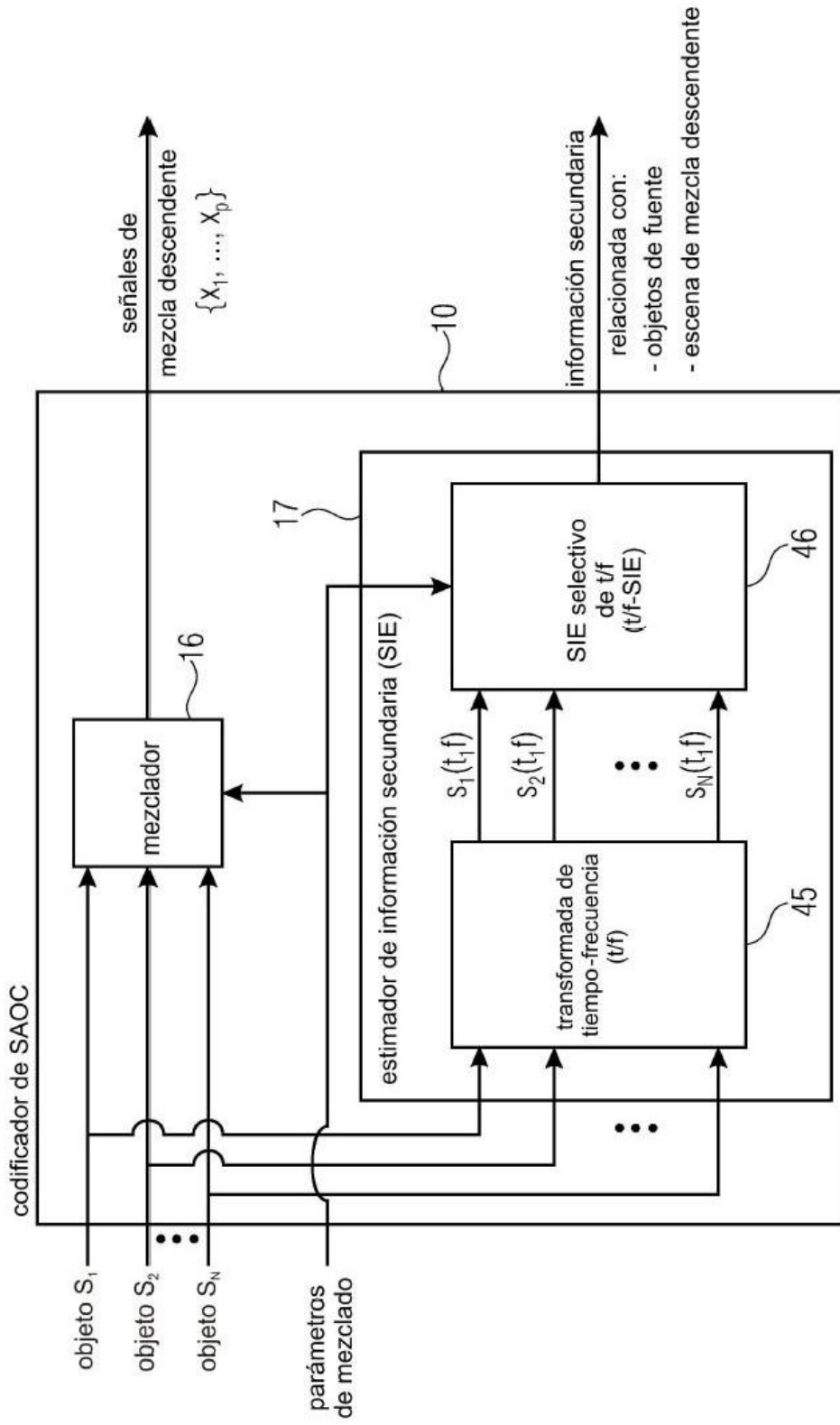


FIGURA 6

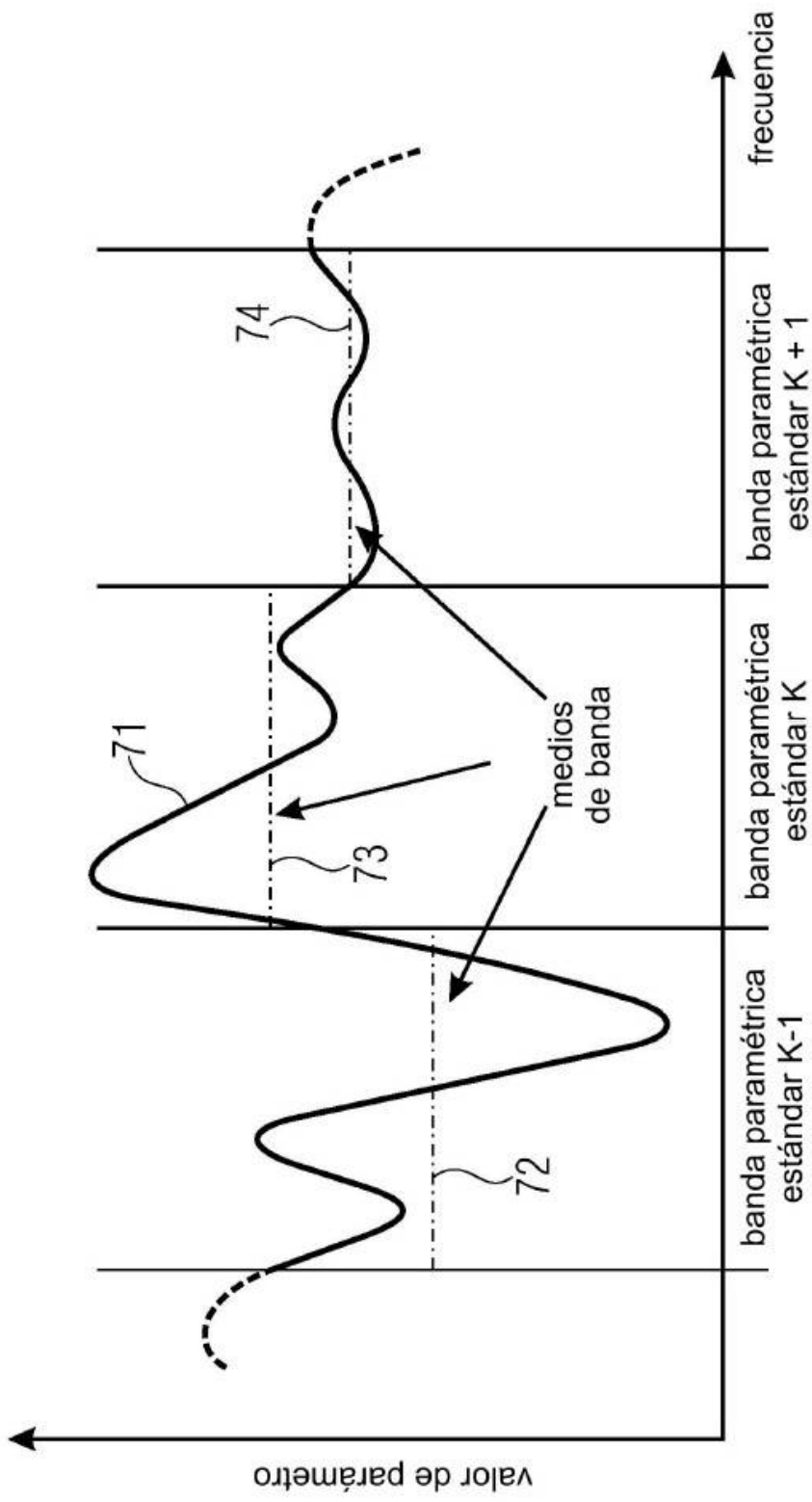


FIGURA 7

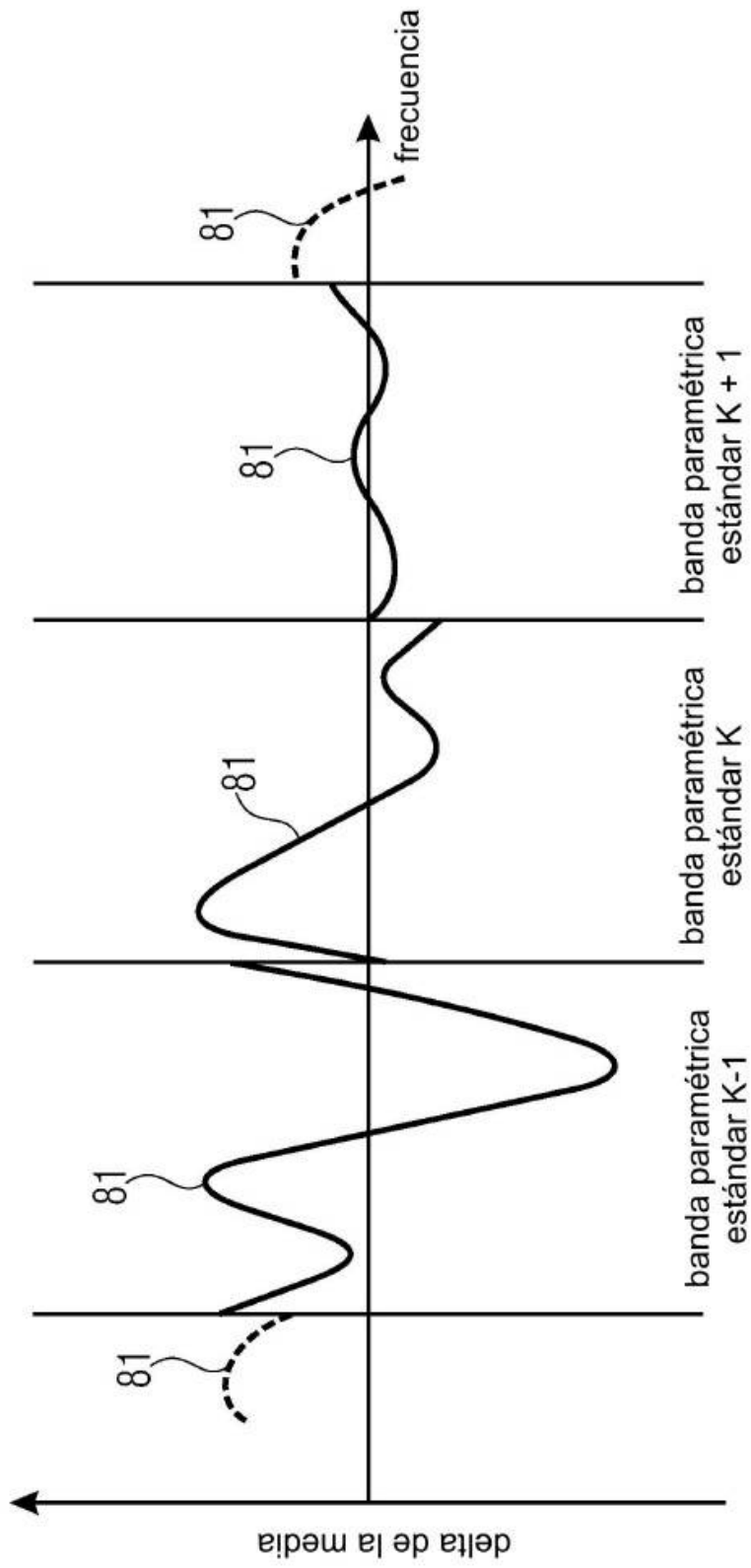


FIGURA 8

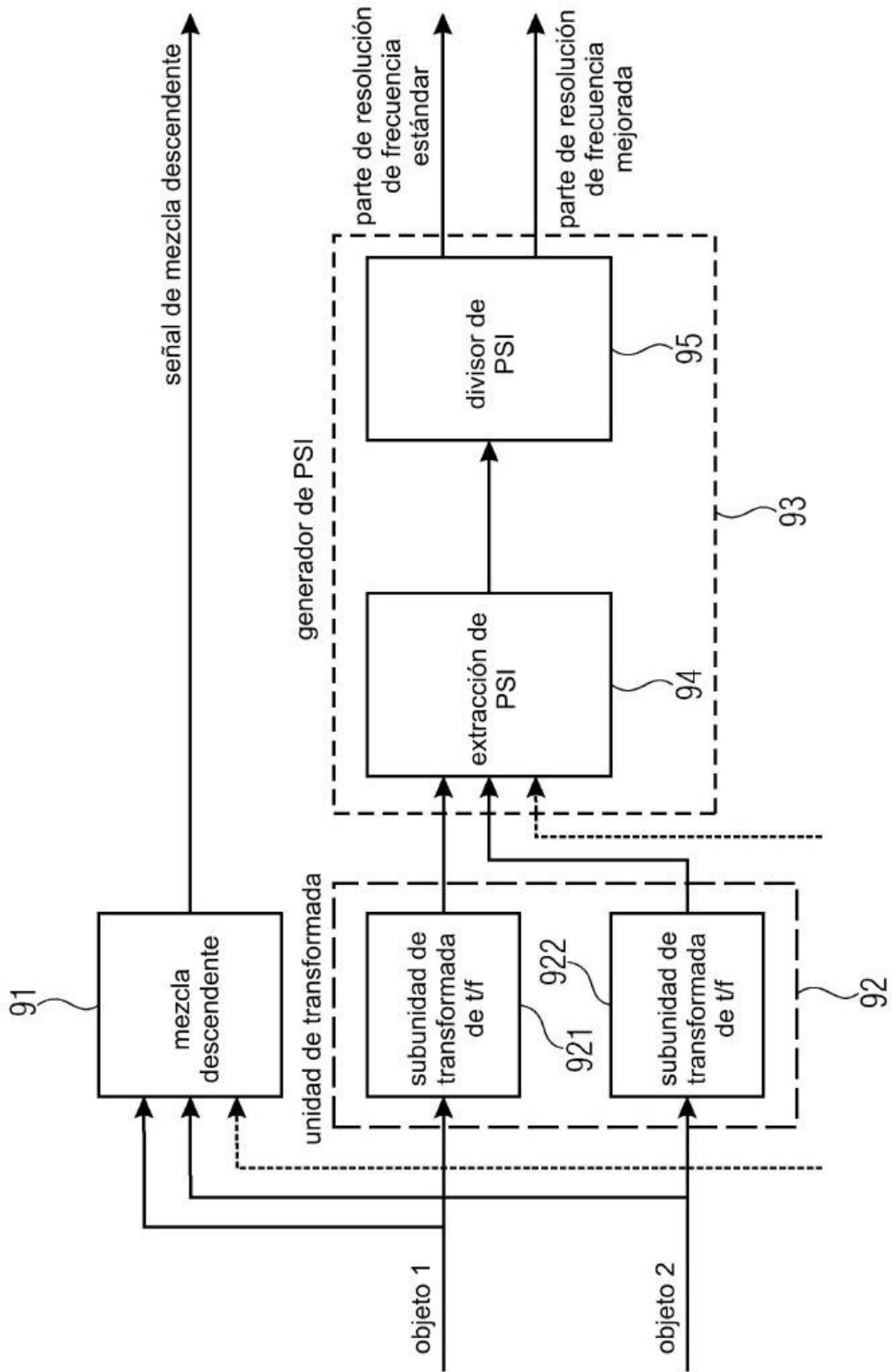


FIGURA 9

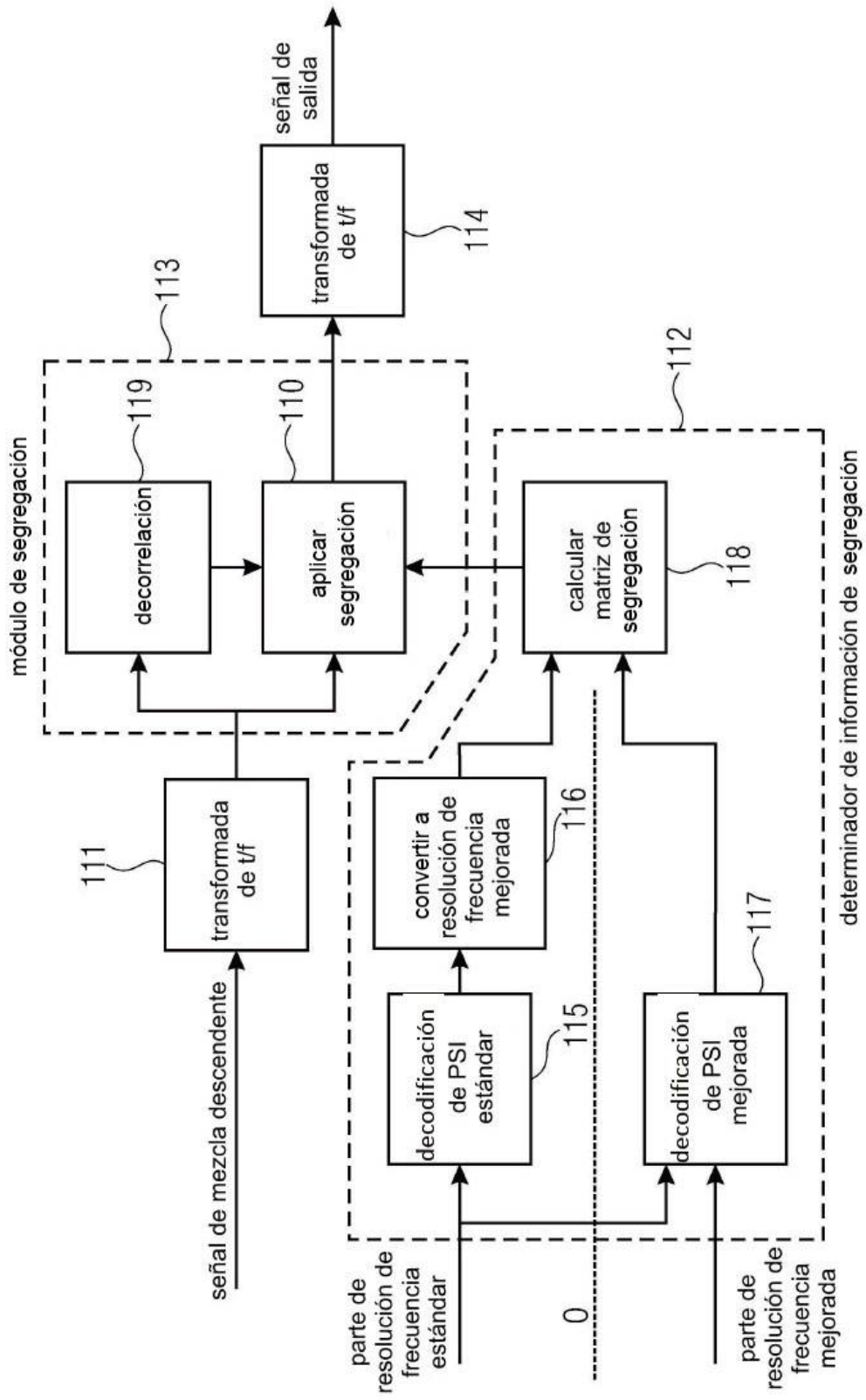


FIGURA 11

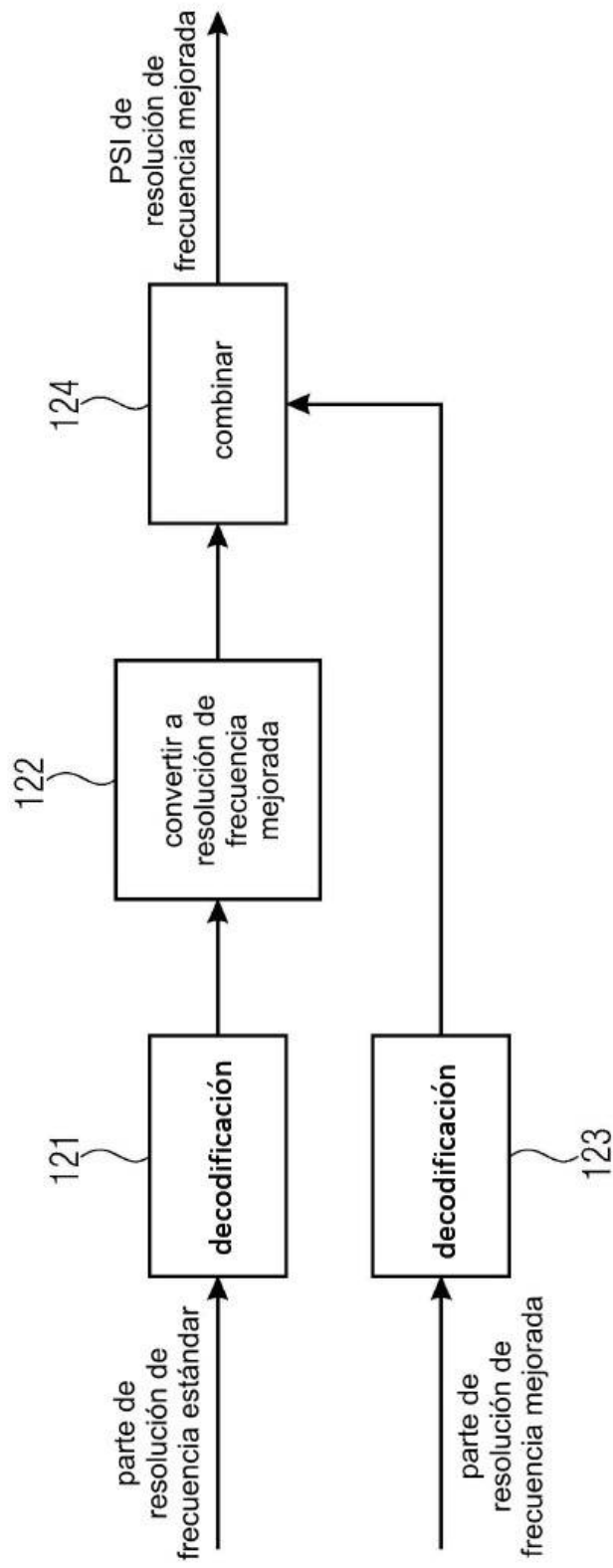


FIGURA 12

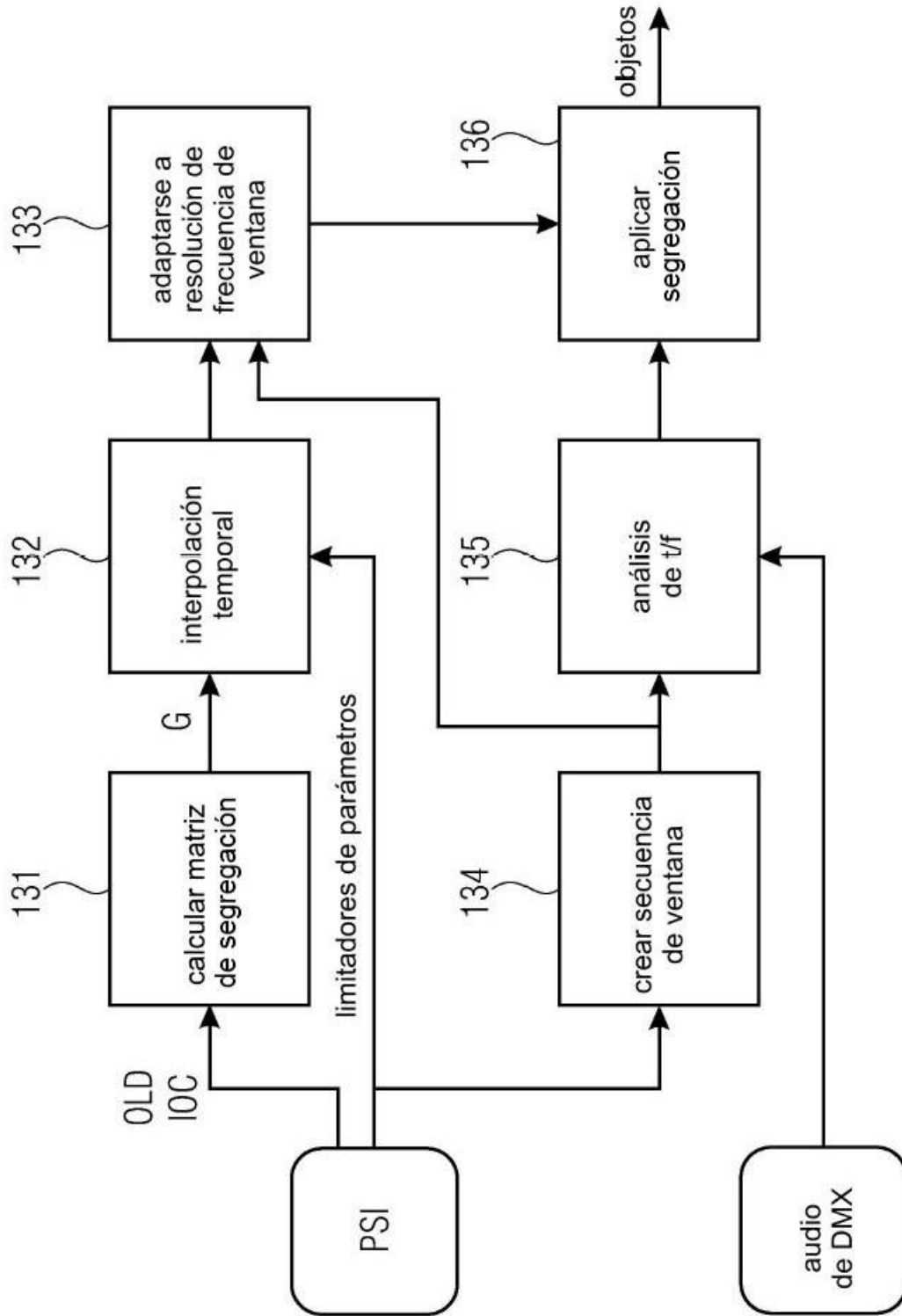


FIGURA 13

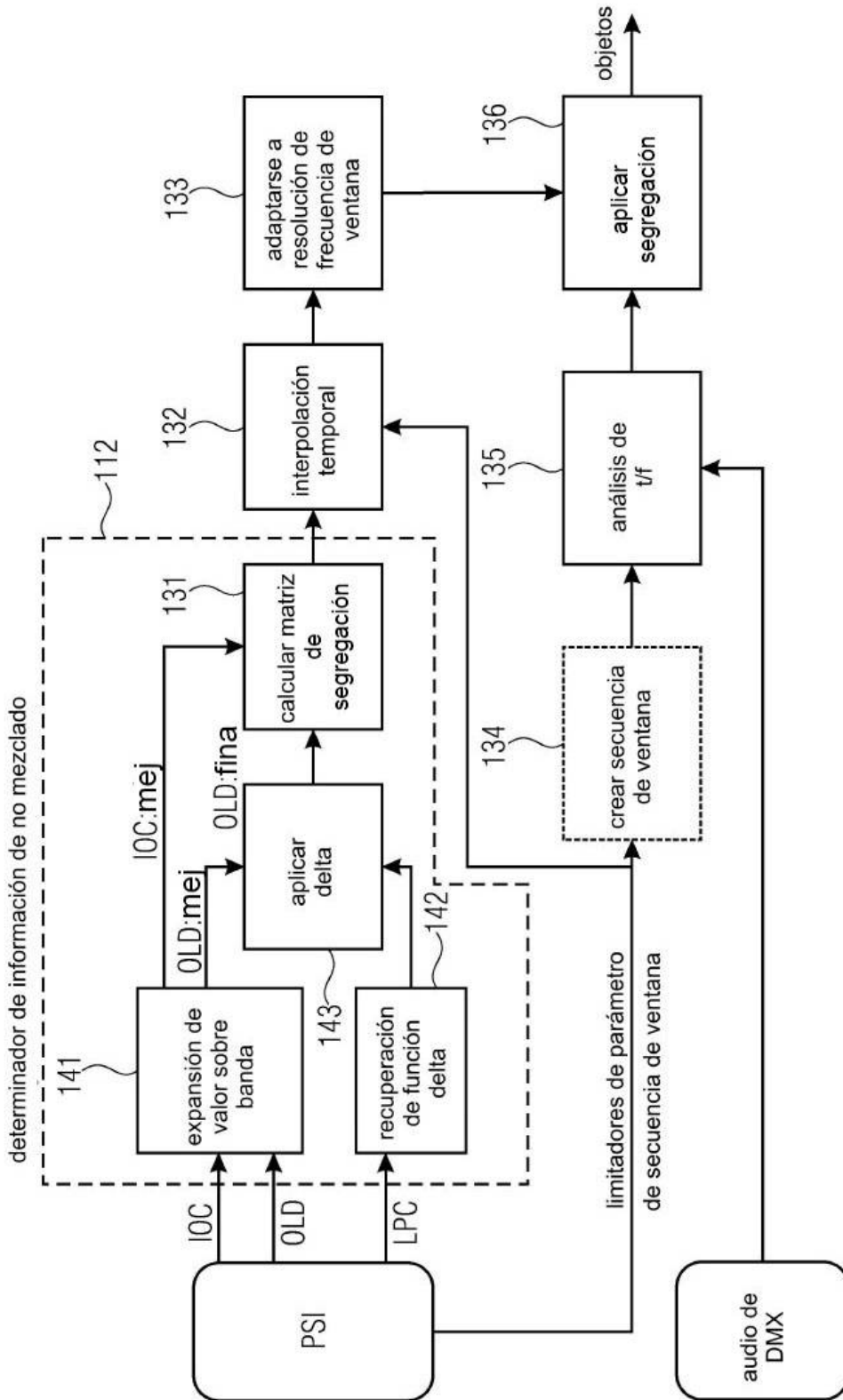


FIGURA 14

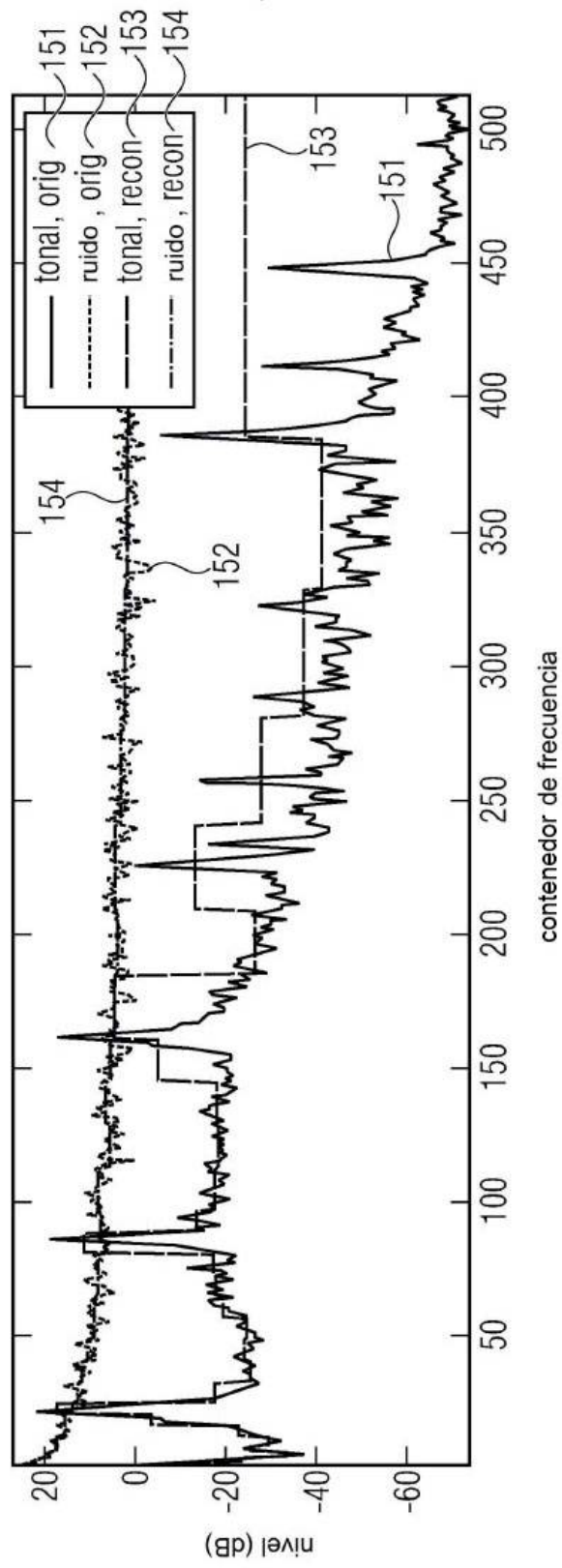


FIGURA 15

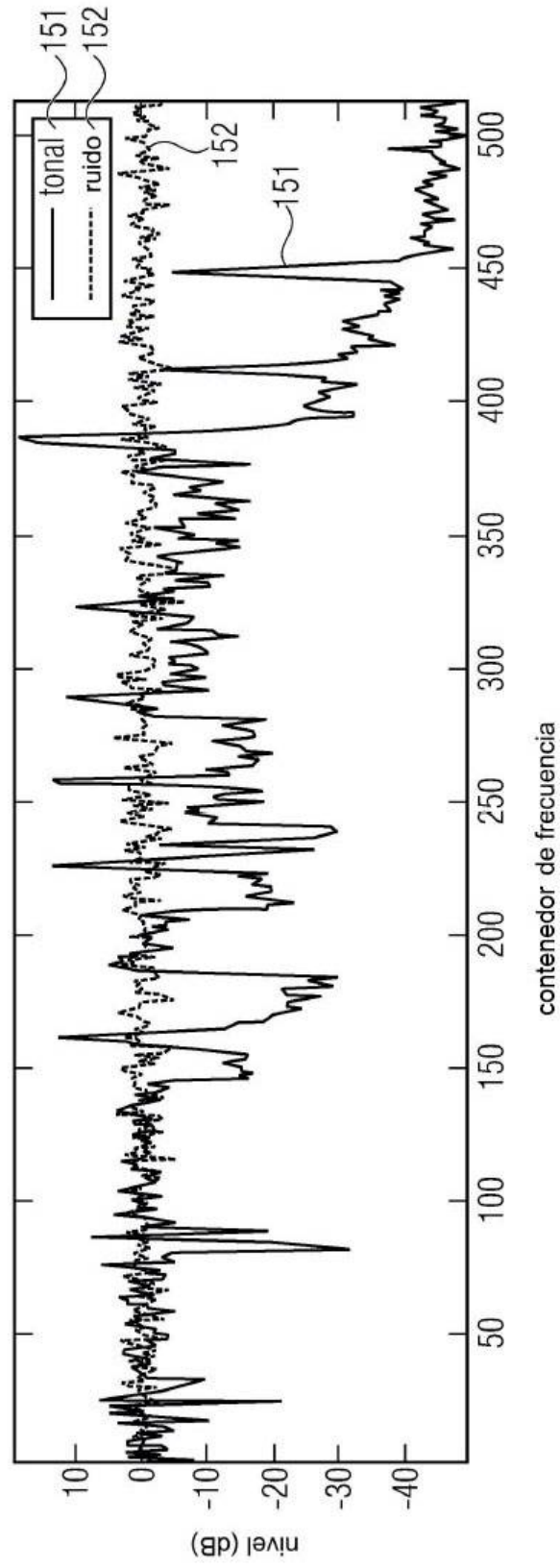


FIGURA 16

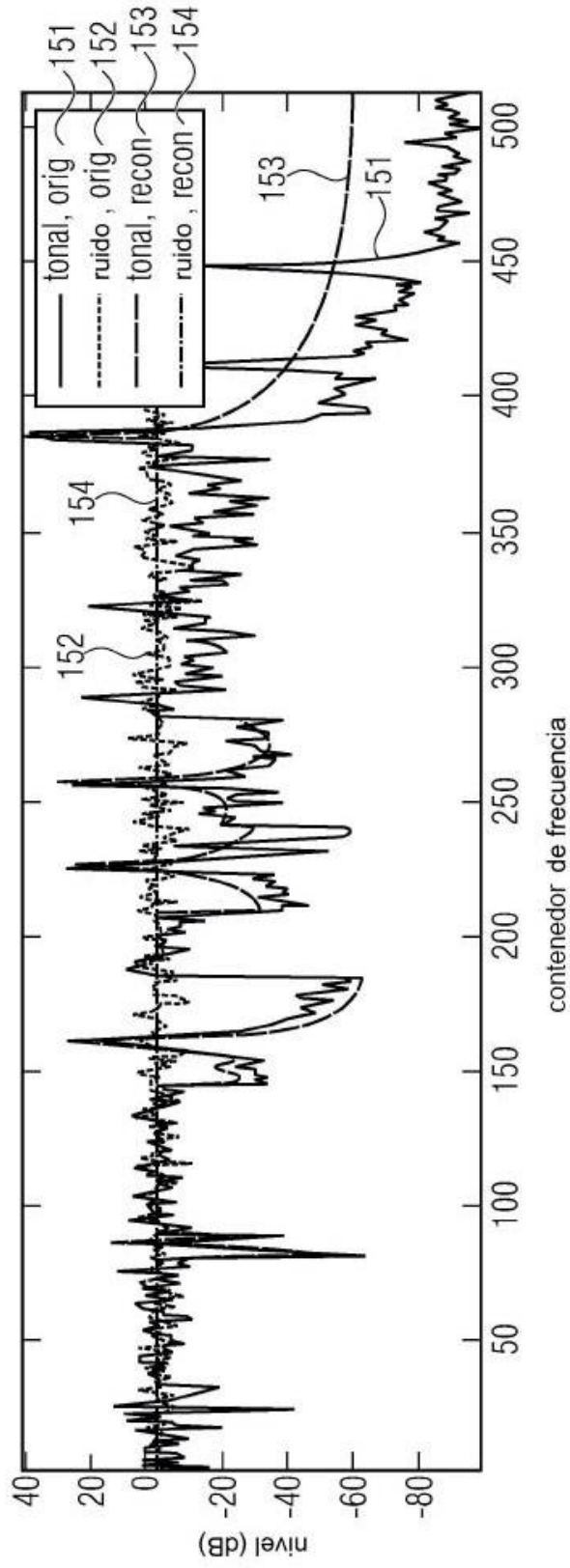


FIGURA 17

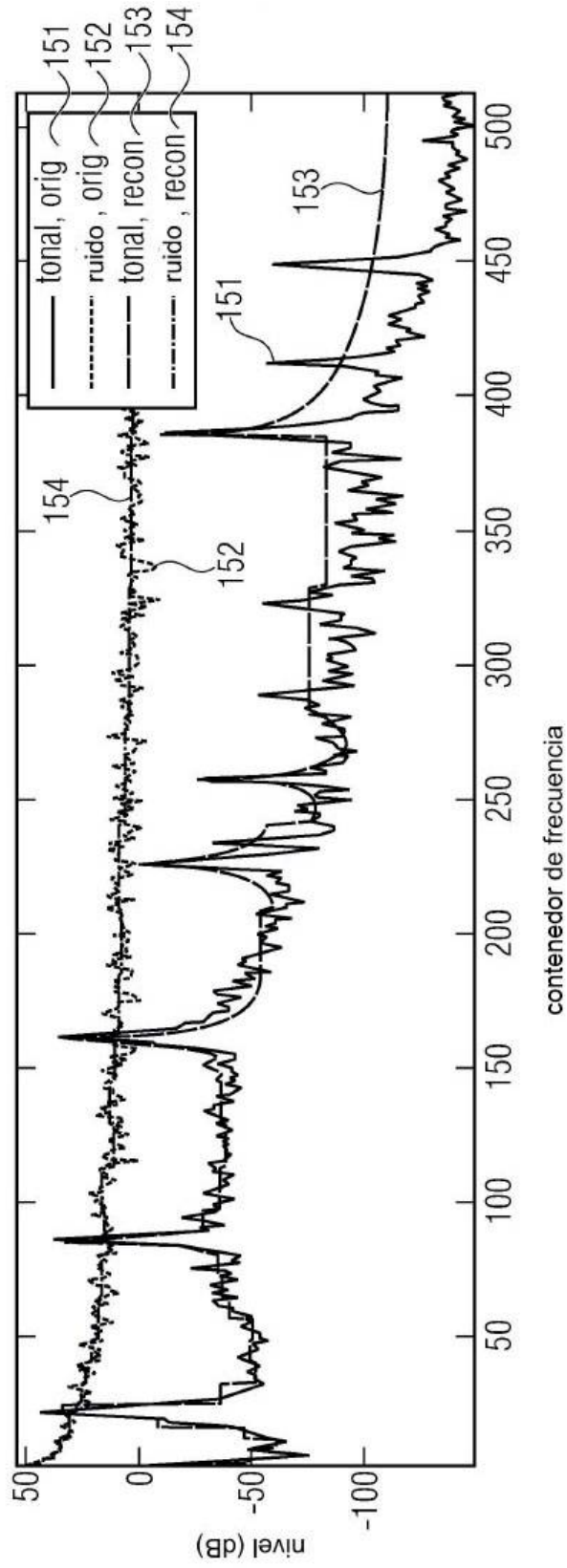


FIGURA 18