

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 638 498**

51 Int. Cl.:

G10L 21/0388 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **15.07.2014 PCT/EP2014/065112**

87 Fecha y número de publicación internacional: **29.01.2015 WO15010950**

96 Fecha de presentación y número de la solicitud europea: **15.07.2014 E 14741264 (7)**

97 Fecha y número de publicación de la concesión europea: **21.06.2017 EP 3025344**

54 Título: **Aparato y procedimiento para decodificar una señal de audio codificada mediante un filtro de cruce en torno a una frecuencia de transición**

30 Prioridad:

22.07.2013 EP 13177346

22.07.2013 EP 13177350

22.07.2013 EP 13177353

22.07.2013 EP 13177348

18.10.2013 EP 13189389

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

23.10.2017

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)**

**Hansastraße 27c
80686 München, DE**

72 Inventor/es:

**DISCH, SASCHA;
GEIGER, RALF;
HELMRICH, CHRISTIAN;
NAGEL, FREDERIK;
NEUKAM, CHRISTIAN;
SCHMIDT, KONSTANTIN y
FISCHER, MICHAEL**

74 Agente/Representante:

SALVA FERRER, Joan

ES 2 638 498 T3

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

DESCRIPCIÓN

Aparato y procedimiento para decodificar una señal de audio codificada mediante un filtro de cruce en torno a una frecuencia de transición.

5

CAMPO TÉCNICO

[0001] La presente invención se refiere a la codificación/decodificación de audio y, en especial, a la codificación de audio mediante el relleno inteligente de espacios (IGF, por sus siglas en inglés).

10

[0002] La codificación de audio es el campo de la compresión de señales que se ocupa de explotar la redundancia y la irrelevancia en señales de audio utilizando conocimientos de psicoacústica. En la actualidad, los códecs suelen necesitar aproximadamente 60 kbps/canal para llevar a cabo una codificación perceptualmente transparente de casi cualquier tipo de señal de audio. Los códecs más recientes tienen como objetivo reducir la tasa de bits de la codificación explotando las semejanzas espectrales en la señal mediante técnicas como la ampliación del ancho de banda (BWE, por sus siglas en inglés). Un esquema de BWE utiliza un conjunto de parámetros de baja tasa de bits para representar los componentes de alta frecuencia (HF, por sus siglas en inglés) de una señal de audio. El espectro de HF se rellena con contenido espectral de las zonas de baja frecuencia (LF, por sus siglas en inglés) y la forma espectral, inclinación y continuidad temporal se ajustan para mantener el timbre y el color de la señal original. Dichos procedimientos de BWE permiten que los códecs de audio mantengan una alta calidad incluso a bajas tasas de bits de aproximadamente 24 kbps/canal.

15

20

[0003] El sistema de codificación de audio de la invención codifica de forma eficiente señales de audio arbitrarias a una amplia variedad de tasas de bits. Mientras que, para altas tasas de bits, el sistema de la invención converge hacia la transparencia, para bajas tasas de bits se minimiza la molestia perceptual. Por lo tanto, la parte principal de la tasa de bits disponible se utiliza para codificar como forma de onda tan solo la estructura de mayor relevancia perceptiva de la señal en el codificador, y los espacios espectrales resultantes se rellenan en el decodificador con un contenido de señal que se aproxima más o menos al espectro original. Se consume un volumen de bits muy limitado para controlar el denominado relleno inteligente de espacios (IGF, por sus siglas en inglés) espectral basado en parámetros mediante información secundaria específica transmitida del codificador al decodificador.

25

30

[0004] El almacenamiento o la transmisión de señales de audio a menudo están sometidos a estrictas restricciones en lo que referente a la tasa de bits. Antes, los codificadores se veían obligados a reducir drásticamente el ancho de banda transmitido cuando solo estaba disponible una tasa de bits muy baja.

35

[0005] En la actualidad, los códecs de audio modernos son capaces de codificar señales de banda ancha mediante procedimientos de ampliación del ancho de banda (BWE) [1]. Estos algoritmos recurren a una representación paramétrica del contenido de alta frecuencia (HF), que se genera a partir de la parte de baja frecuencia (LF) codificada como forma de onda de la señal decodificada por medio de la transposición en la zona espectral de HF ("interconexión") y la aplicación de un posprocesamiento basado en parámetros. En esquemas de BWE, a menudo la reconstrucción de la zona espectral de HF por encima de una frecuencia dada denominada frecuencia de cruce se basa en la interconexión espectral. Normalmente, la zona de HF está compuesta por múltiples interconexiones adyacentes y cada una de estas interconexiones procede de zonas de paso de banda (BP, por sus siglas en inglés) del espectro de LF por debajo de la frecuencia de cruce dada. Los sistemas del estado de la técnica llevan a cabo la interconexión de manera eficiente dentro de una representación de banco de filtros, por ejemplo, banco de filtros espejo de cuadratura (QMF, por sus siglas en inglés), copiando un conjunto de coeficientes de subbanda contiguos desde un origen hasta la zona de destino.

45

[0006] Otra técnica empleada en los actuales códecs de audio que incrementa la eficiencia de la compresión y, por tanto, permite la ampliación del ancho de banda de audio a bajas tasas de bits es la sustitución sintética basada en parámetros de partes adecuadas de los espectros de audio. Por ejemplo, las porciones de la señal de tipo ruido de la señal de audio original se pueden sustituir por ruido artificial generado en el decodificador, sin que se produzca una pérdida considerable de la calidad subjetiva, y se puede ajustar la escala mediante los parámetros de la información secundaria. Un ejemplo de ello es la herramienta de sustitución de ruido perceptual (PNS, por sus siglas en inglés) contenida en la codificación avanzada de audio (AAC, por sus siglas en inglés) MPEG-4 [5].

50

55

[0007] Otro recurso que también permite la ampliación del ancho de banda a bajas tasas de bits es la técnica de relleno de ruido incluida en la codificación unificada de voz y audio (USAC, por sus siglas en inglés) de MPEG-D

[7]. Los espacios espectrales (ceros) inferidos por la zona muerta del cuantificador debidos a una cuantificación demasiado gruesa se rellenan posteriormente con ruido artificial en el decodificador y se ajusta la escala mediante un posprocesamiento basado en parámetros.

5 **[0008]** Otro sistema del estado de la técnica se denomina sustitución espectral precisa (ASR, por sus siglas en inglés) [2-4]. Además de un códec de forma de onda, la ASR emplea una etapa específica de síntesis de señal que restituye porciones sinusoidales perceptualmente importantes de la señal en el decodificador. Además, un sistema descrito en [5] recurre al modelado sinusoidal en la zona de HF de un codificador de forma de onda para permitir la ampliación del ancho de banda de audio con una calidad perceptual aceptable a bajas tasas de bits.

10 Todos estos procedimientos conllevan la transformación de los datos en un segundo dominio aparte de la transformada de coseno discreta modificada (MDCT, por sus siglas en inglés), además de etapas de análisis/síntesis bastante complejas para la conservación de los componentes sinusoidales de HF.

[0009] Otro sistema del estado de la técnica se denomina reconstrucción de alta frecuencia y sus principios se describen en el documento WO 2010/136459 A1.

[0010] La fig. 13 ilustra un diagrama esquemático de un codificador de audio para una tecnología de ampliación de ancho de banda como la utilizada, por ejemplo, en la codificación avanzada de audio de alta eficiencia (HE-AAC, por sus siglas en inglés). Una señal de audio en la línea 1300 se introduce en un sistema de filtros que comprende un paso bajo 1302 y un paso alto 1304. La señal de salida del filtro de paso alto 1304 se introduce en un codificador/extractor de parámetros 1306. El codificador/extractor de parámetros 1306 está configurado para calcular y codificar parámetros tales como un parámetro de envolvente espectral, un parámetro de adición de ruido, un parámetro de armónicos faltantes o un parámetro de filtrado inverso, por ejemplo. Estos parámetros extraídos se introducen en un multiplexor de corrientes de bits 1308. La señal de salida de paso bajo se introduce en un procesador que normalmente comprende la funcionalidad de un muestreador descendente 1310 y un codificador central 1312. El paso bajo 1302 limita el ancho de banda que se va a codificar a un ancho de banda considerablemente más pequeño que el de la señal de audio de entrada original 1300. Esto proporciona una considerable ganancia de codificación debido al hecho de que la totalidad de las funcionalidades que se dan en el codificador central solo tienen que actuar sobre una señal con un ancho de banda reducido. Cuando, por ejemplo, el ancho de banda de la señal de audio en la línea 1300 es de 20 kHz y cuando el filtro de paso bajo 1302 posee, ejemplarmente, un ancho de banda de 4 kHz, para cumplir el teorema del muestreo, teóricamente es suficiente con que la señal posterior al muestreador descendente tenga una frecuencia de muestreo de 8 kHz, lo que constituye una reducción considerable en la tasa de muestreo requerida para la señal de audio 1300, que debe ser de al menos 40 kHz.

35 **[0011]** La fig. 13b ilustra un diagrama esquemático de un correspondiente decodificador de ampliación de ancho de banda. El decodificador comprende un multiplexor de corrientes de bits 1320. El demultiplexor de corrientes de bits 1320 extrae una señal de entrada para un decodificador central 1322 y una señal de entrada para un decodificador de parámetros 1324. Una señal de salida del decodificador central posee, en el anterior ejemplo, una tasa de muestreo de 8 kHz y, por lo tanto, un ancho de banda de 4 kHz, mientras que, para una reconstrucción completa del ancho de banda, la señal de salida de un reconstructor de alta frecuencia 1330 debe estar a 20 kHz, lo que requiere una tasa de muestreo de al menos 40 kHz. Para que esto sea posible, se necesita un procesador de decodificador con la funcionalidad de un muestreador ascendente 1325 y un banco de filtros 1326. El reconstructor de alta frecuencia 1330 recibe entonces la señal de salida de baja frecuencia analizada en frecuencia producida por el banco de filtros 1326 y reconstruye el rango de frecuencias definido por el filtro de paso alto 1304 de la fig. 13a mediante la representación paramétrica de la banda de alta frecuencia. El reconstructor de alta frecuencia 1330 posee varias funcionalidades, tales como la regeneración del rango de frecuencia superior mediante la utilización del rango de origen en el rango de baja frecuencia, un ajuste de envolvente espectral, una funcionalidad de adición de ruido y una funcionalidad para introducir armónicos faltantes en el rango de frecuencia superior y, si se aplica y calcula en el codificador de la fig. 13a, una operación de filtrado inverso para tener en cuenta el hecho de que normalmente el rango de frecuencia superior no es tan tonal como el rango de frecuencia inferior. En la HE-AAC, los armónicos faltantes se resintetizan en el lado del decodificador y se colocan exactamente en el medio de una banda de reconstrucción. Por lo tanto, ninguna de las líneas armónicas faltantes que se han determinado en una cierta banda de reconstrucción se coloca en los valores de frecuencia en los que se situaba en la señal original. Esas líneas armónicas faltantes se colocan, en cambio, en frecuencias situadas en el centro de dicha cierta banda. De este modo, cuando una línea armónica faltante de la señal original se coloca muy cerca del límite de la banda de reconstrucción en la señal original, el error en la frecuencia introducido al colocar esta línea de armónicos faltantes en el centro de la banda en la señal reconstruida es de cerca del 50% de la banda de reconstrucción individual, para la que se han generado y transmitido parámetros.

[0012] Además, aunque los típicos codificadores centrales de audio actúan en el dominio espectral, el decodificador central genera una señal de dominio temporal que, de nuevo, se convierte después en un dominio espectral mediante la funcionalidad del banco de filtros 1326. Esto introduce retardos de procesamiento adicionales, puede introducir fallas debido al procesamiento en tándem consistente en transformar en primer lugar el dominio espectral en dominio de frecuencia y transformar de nuevo, normalmente, en un dominio de frecuencia diferente y, por supuesto, también requiere un grado considerable de complejidad computacional y, por tanto, energía eléctrica, lo cual supone un problema específico cuando se aplica la tecnología de ampliación de ancho de banda en dispositivos portátiles tales como teléfonos móviles, tabletas, ordenadores portátiles, etc.

[0013] Los actuales códecs de audio llevan a cabo una codificación de audio a baja tasa de bits utilizando la BWE como parte integral del esquema de codificación. No obstante, las técnicas de BWE tienen la limitación de sustituir únicamente contenidos de alta frecuencia (HF). Además, no permiten codificar como forma de onda contenidos perceptualmente importantes por encima de una frecuencia de cruce dada. Por lo tanto, los actuales códecs de audio pierden o bien detalle o bien timbre en HF cuando se implementa la BWE, ya que en la mayoría de sistemas no se tiene en cuenta la alineación exacta de los armónicos tonales de la señal.

[0014] Otro inconveniente de los actuales sistemas BWE del estado de la técnica es la necesidad de transformar la señal de audio en un nuevo dominio para la implementación de la BWE (por ejemplo, transformar un dominio MDCT en QMF). Esta necesidad da lugar a complicaciones en la sincronización, una complejidad computacional añadida y mayores requisitos de memoria.

[0015] El almacenamiento o la transmisión de señales de audio a menudo se ven sometidos a estrictas restricciones en lo que respecta a la tasa de bits. Anteriormente, los codificadores se veían obligados a reducir de manera drástica el ancho de banda de audio transmitido cuando solo estaba disponible una tasa de bits muy baja. En la actualidad, los modernos códecs de audio son capaces de codificar señales de banda ancha mediante procedimientos de ampliación de ancho de banda (BWE) [1-2]. Estos algoritmos recurren a una representación paramétrica del contenido de alta frecuencia (HF), que se genera a partir de la parte de baja frecuencia (LF) codificada como forma de onda de la señal decodificada por medio de la transposición en la zona espectral de HF ("Interconexión") y la aplicación de un posprocesamiento basado en parámetros.

[0016] En esquemas BWE, la reconstrucción de la zona espectral de HF por encima de una, así denominada, frecuencia de cruce dada se basa a menudo en la interconexión espectral. Otros esquemas que poseen la funcionalidad para rellenar espacios espectrales, por ejemplo, el rellenado inteligente de espacios (IGF), utilizan unos, así denominados, recuadros espectrales contiguos para regenerar partes de los espectros de HF de una señal de audio. Normalmente, la zona de HF está compuesta por múltiples interconexiones o recuadros adyacentes y cada una de estas interconexiones o recuadros se obtiene a partir de zonas de paso de banda (BP) del espectro de LF por debajo de la frecuencia de cruce dada. Los sistemas del estado de la técnica llevan a cabo de manera eficiente la interconexión o formación de recuadros dentro de una representación de banco de filtros copiando un conjunto de coeficientes de subbanda adyacentes desde un origen a la zona de destino. Si bien, para algunos contenidos de la señal, el ensamblado de la señal reconstruida a partir de la banda de LF e interconexiones adyacentes dentro de la banda de HF puede producir batimiento, disonancia y aspereza auditiva.

[0017] Por lo tanto, en [19], el concepto de filtrado en banda de guarda de la disonancia se presenta en el contexto de un sistema BWE basado en un banco de filtros. Se sugiere aplicar de manera eficaz un filtro de muesca con un ancho de banda de aproximadamente 1 bark en la frecuencia de cruce entre LF y HF regenerada mediante BWE para evitar la posibilidad de disonancia y sustituir el contenido espectral por ceros o ruido.

[0018] Sin embargo, la solución propuesta en [19] presenta algunos inconvenientes: en primer lugar, la sustitución estricta de contenido espectral por ceros o ruido también puede afectar a la calidad perceptual de la señal. Además, el procesamiento propuesto no es adaptable a la señal y, por tanto, en algunos casos, puede deteriorar la calidad perceptual. Por ejemplo, si la señal contiene transitorios, pueden producirse preecos y posecos.

[0019] En segundo lugar, las disonancias también pueden producirse en transiciones entre interconexiones de HF consecutivas. La solución propuesta en [19] solo resulta funcional para remediar las disonancias que se producen en la frecuencia de cruce entre LF y HF regenerada por BWE.

[0020] Por último, al contrario que los sistemas basados en bancos de filtros como el propuesto en [19], los sistemas BWE también se pueden llevar a la práctica en implementaciones basadas en transformadas, como, por

ejemplo, la transformada de coseno discreta modificada (MDCT). Las transformadas de tipo MDCT son muy propensas a sufrir las denominadas fallas de zumbido o trino [20] que se producen si se copian zonas de paso de banda de coeficientes espectrales o si los coeficientes espectrales se fijan en cero, tal como se propone en [19].

5 **[0021]** En particular, la patente de EE. UU. 8,412,365 describe el uso, en una traslación o plegado basados en banco de filtros, de las denominadas bandas de guarda que se insertan y están hechas de uno o varios canales de subbanda fijados en cero. Se utilizan varios canales del banco de filtros como bandas de guarda, y un ancho de banda de una banda de guarda debería ser de 0,5 barks. Esas bandas de guarda de disonancia se reconstruyen parcialmente mediante señales de ruido blanco aleatorio, es decir, se suministra ruido blanco a las subbandas en
10 lugar de que estas sean cero. Las bandas de guarda se insertan independientemente de la señal actual que se vaya a procesar.

[0022] Los sistemas de ampliación de ancho de banda resultan particularmente problemáticos cuando se llevan a la práctica en implementaciones basadas en transformadas como, por ejemplo, la transformada de coseno discreta modificada (MDCT). Las transformadas de tipo MDCT, así como otras transformadas, son muy propensas a
15 sufrir el denominado trino que se explica en [3] y a fallas de zumbido que se producen si se copian las zonas de paso de banda de coeficientes espectrales o si los coeficientes espectrales se fijan en cero como se propone en [2].

[0023] El objeto de la presente invención consiste en proporcionar un aparato y un procedimiento mejorados
20 para decodificar una señal de audio codificada.

[0024] Este objeto se logra mediante un aparato para decodificar una señal de audio codificada de la reivindicación 1, un procedimiento de decodificación de una señal de audio codificada de la reivindicación 14 o un programa informático de acuerdo con la reivindicación 15. De acuerdo con la presente invención, un aparato para
25 decodificar una señal de audio codificada comprende un decodificador central, un generador de recuadros para generar uno o más recuadros espectrales con frecuencias no incluidas en la señal central decodificada, a partir de una porción espectral de la señal central decodificada y un filtro de cruce para llevar a cabo el filtrado de cruce espectral de la señal central decodificada y un recuadro de una primera frecuencia con frecuencias que se amplían desde una frecuencia de relleno de espacios hasta una primera frecuencia de fin de recuadro o para el filtrado de
30 cruce espectral de un recuadro y otro recuadro de frecuencia, teniendo este otro recuadro de frecuencia una frecuencia de límite inferior adyacente en frecuencia a una frecuencia de límite superior del recuadro de frecuencia.

[0025] Preferentemente, este procedimiento está destinado a su aplicación dentro de una ampliación de ancho de banda basada en una transformada como la MDCT. No obstante, la presente invención se puede aplicar
35 de forma general y, especialmente en una situación de ampliación de ancho de banda en la que se recurre a un banco de filtros espejo de cuadratura (QMF), especialmente si el sistema se muestrea críticamente, por ejemplo cuando hay una representación de QMF con valores reales como conversión de tiempo en frecuencia o como conversión de frecuencia en tiempo.

40 **[0026]** La presente invención resulta especialmente útil para señales de tipo transitorio, ya que para dichas señales de tipo transitorio, el zumbido constituye una falla audible y molesta. Las fallas de zumbido de filtro son provocadas por la denominada característica de "pared de ladrillo" (*brick-wall*) de un filtro en la banda de transición, es decir, una transición brusca desde un banda de paso hasta una banda de fin a una frecuencia de corte. Dichos filtros se pueden implementar de manera eficiente fijando un coeficiente o grupo de coeficientes en cero en un
45 dominio de frecuencia de una transformada de tiempo/frecuencia. Por lo tanto, la presente invención recurre a un filtro de cruce en cada frecuencia de transición entre interconexiones/recuadros o entre una banda central y una primera interconexión/recuadro para reducir esta falla de zumbido. El filtro de cruce se implementa preferentemente mediante la ponderación espectral en el dominio de la transformada empleando funciones de ganancia adecuadas.

50 **[0027]** Preferentemente, el filtro de cruce es adaptable a la señal y consiste en dos filtros: un filtro de disminución progresiva del volumen, que se aplica a la zona espectral inferior, y un filtro de aumento progresivo del volumen, que se aplica a la zona espectral superior. Los filtros pueden ser simétricos o asimétricos, según la implementación concreta.

55 **[0028]** En otra realización, un recuadro de frecuencia o interconexión de frecuencia no solo se ve sometido al filtrado de cruce, sino que el generador de recuadros lleva a cabo, antes de llevar a cabo el filtrado de cruce, una adaptación de interconexiones que comprende una fijación de límites de frecuencia en los mínimos espectrales locales y una eliminación o atenuación de porciones tonales que permanecen en rangos de transición en torno a las frecuencias de transición.

[0029] En esta realización, se lleva a cabo un análisis de la señal en el lado del decodificador mediante un analizador, para analizar la señal central decodificada antes o después de llevar a cabo una operación de regeneración de frecuencia para proporcionar un resultado del análisis. Después, este resultado del análisis es empleado por un regenerador de frecuencia para regenerar porciones espectrales no incluidas en la señal central decodificada.

[0030] De este modo, a diferencia de una configuración de decodificador fija, en la que la interconexión o formación de recuadros de frecuencia se llevan a cabo de forma fija, es decir, donde se toma un cierto rango de origen a partir de la señal central y se aplican unos ciertos límites de frecuencia fijos para fijar la frecuencia entre el rango de origen y el rango de reconstrucción, o el límite de frecuencia entre dos interconexiones o recuadros de frecuencia adyacentes dentro del rango de reconstrucción, se lleva a cabo una interconexión o formación de recuadros dependientes de la señal, en las que, por ejemplo, se puede analizar la señal central para encontrar mínimos locales en la señal central y, después, se selecciona el rango central de manera que los límites de frecuencia del rango central coincidan con mínimos locales en el espectro de la señal central.

[0031] Como otra posibilidad o de forma añadida, se puede llevar a cabo un análisis de señal con una señal regenerada preliminar o señal preliminar con interconexiones de frecuencia o divididas en recuadros, en la que, tras el primer procedimiento de regeneración preliminar, se analiza el límite entre el rango central y el rango de reconstrucción con el fin de detectar todas las porciones de la señal que generan fallas, como, por ejemplo, porciones tonales que resultan problemáticas debido a que se encuentran lo bastante próximas entre sí como para generar una falla de batimiento cuando se están reconstruyendo. Como otra posibilidad o de forma añadida, también se pueden examinar los límites de manera que se detecte un recorte intermedio de una porción tonal y este recorte de una porción tonal también generaría una falla cuando se estuviera reconstruyendo tal como es. Para evitar estos procedimientos, el límite de frecuencia del rango de reconstrucción y/o el rango de origen y/o entre dos recuadros o interconexiones individuales en el rango de reconstrucción se pueden modificar mediante un manipulador de señal con el fin de volver a llevar a cabo una reconstrucción con los límites recién fijados.

[0032] De forma añadida o como otra posibilidad, la regeneración de frecuencia es una regeneración basada en el resultado del análisis, ya que los límites de frecuencia se dejan tal como están y se efectúa una eliminación o al menos una atenuación de porciones tonales problemáticas cerca de los límites de frecuencia entre el rango de origen y el rango de reconstrucción o entre dos recuadros de frecuencia o interconexiones individuales dentro del rango de reconstrucción. Dichas porciones tonales pueden ser tonos cercanos que darían lugar a una falla de batimiento o podrían ser porciones tonales recortadas.

[0033] En concreto, cuando se utiliza una transformada sin conservación de potencia, tal como una MDCT, un tono individual no se corresponde directamente con una línea espectral individual. En lugar de ello, un tono individual se corresponderá con un grupo de líneas espectrales con ciertas amplitudes dependiendo de la fase del tono. Cuando una operación de interconexión recorta esta porción tonal, se producirá una falla tras la reconstrucción, aunque se aplique una reconstrucción perfecta, como en un reconstructor MDCT. Esto se debe al hecho de que el reconstructor MDCT requeriría el patrón tonal completo para un tono con el fin de reconstruir finalmente este tono de manera correcta. Debido al hecho de que antes se ha producido un recorte, esto deja ya de ser posible y, por tanto, se generará una falla de trino variable en el tiempo. Basándose en el análisis de acuerdo con la presente invención, el regenerador de frecuencia evitará esta situación atenuando la porción tonal completa que genera una falla o, como se explica anteriormente, cambiando las correspondientes frecuencias límite o aplicando ambas medidas, o incluso reconstruyendo la porción recortada basándose en cierto conocimiento previo sobre dichos patrones tonales.

[0034] El enfoque de la invención está destinado principalmente para su aplicación dentro de una BWE basada en una transformada como la MDCT. No obstante, las enseñanzas de la invención se pueden aplicar en general, por ejemplo, de forma análoga dentro de un sistema basado en un banco de filtros de espejo de cuadratura (QMF), especialmente si el sistema se muestrea críticamente, por ejemplo, una representación QMF con valores reales.

[0035] A continuación se exponen realizaciones preferidas haciendo referencia a los dibujos adjuntos, en los que:

la fig. 1a ilustra un aparato para codificar una señal de audio;

la fig. 1b ilustra un decodificador para decodificar una señal de audio, correspondiente al codificador de la fig. 1a;

- la fig. 2a ilustra una implementación preferida del decodificador;
 la fig. 2b ilustra una implementación preferida del codificador;
 la fig.3a ilustra una representación esquemática de un espectro generado por el decodificador de dominio espectral de la fig. 1b;
- 5 la fig.3b ilustra una tabla que indica la relación entre factores de ajuste de escala para bandas de factor de escala y potencias para bandas de reconstrucción e información de relleno de ruido para una banda de relleno de ruido;
 la fig. 4a ilustra la funcionalidad del codificador de dominio espectral para aplicar la selección de porciones espectrales en el primer y el segundo conjunto de porciones espectrales;
 la fig. 4b ilustra una implementación de la funcionalidad de la fig. 4a;
- 10 la fig. 5a ilustra una funcionalidad de un codificador MDCT;
 la fig. 5b ilustra una funcionalidad del decodificador con una tecnología MDCT;
 la fig. 5c ilustra una implementación del regenerador de frecuencia;
 la fig. 6a es un aparato para decodificar una señal de audio codificada de acuerdo con una implementación;
 la fig. 6b ilustra otra realización de un aparato para decodificar una señal de audio codificada;
- 15 la fig. 7a ilustra una implementación preferida del regenerador de frecuencia de la fig. 6a o 6b;
 la fig. 7b ilustra otra implementación de una cooperación entre el analizador y el regenerador de frecuencia;
 la fig. 8a ilustra otra implementación del regenerador de frecuencia;
 la fig. 8b ilustra otra realización de la invención;
 la fig. 9a ilustra un decodificador con tecnología de regeneración de frecuencia que utiliza valores de potencia para el rango de frecuencia de regeneración;
- 20 la fig. 9b ilustra una implementación más detallada del regenerador de frecuencia de la fig. 9a;
 la fig. 9c ilustra un esquema que ilustra la funcionalidad de la fig. 9b;
 la fig. 9d ilustra otra implementación del decodificador de la fig. 9a;
 la fig. 10a ilustra un diagrama de bloques de un codificador que se corresponde con el decodificador de la fig. 9a;
- 25 la fig. 10b ilustra un diagrama de bloques para ilustrar otra funcionalidad de la calculadora de parámetros de la fig. 10a;
 la fig. 10c ilustra un diagrama de bloques que ilustra otra funcionalidad de la calculadora paramétrica de la fig. 10a;
 la fig. 10d ilustra un diagrama de bloques que ilustra otra funcionalidad de la calculadora paramétrica de la fig. 10a;
 la fig. 11a ilustra un espectro de un zumbido de filtro que rodea un transitorio;
- 30 la fig. 11b ilustra un espectrograma de un transitorio tras aplicar la ampliación de ancho de banda;
 la fig. 11c ilustra un espectrograma de un transitorio tras aplicar una ampliación de ancho de banda con reducción de zumbido de filtro;
 la fig. 12a ilustra un diagrama de bloques de un aparato para decodificar una señal de audio codificada;
 la fig. 12b ilustra unos espectros de magnitud (estilizados) de una señal tonal, una copia sin adaptación de interconexión/recuadro; una copia con límites de frecuencia cambiados y una eliminación adicional de porciones tonales que generan fallas;
- 35 la fig. 12c ilustra un ejemplo de función de atenuación cruzada;
 la fig. 13a ilustra un codificador de la técnica anterior con ampliación de ancho de banda; y
 la fig. 13b ilustra un decodificador de la técnica anterior con ampliación de ancho de banda;
- 40 La fig. 14a ilustra otro aparato para decodificar una señal de audio codificada mediante un filtro de cruce;
 la fig. 14b ilustra una ilustración más detallada de un filtro de cruce ejemplar.

[0036] En la fig. 6a se ilustra un aparato para decodificar una señal de audio codificada que comprende una señal central codificada y unos datos paramétricos. El aparato comprende un decodificador central 600 para
 45 decodificar la señal central codificada con el fin de obtener una señal central decodificada, un analizador 602 para analizar la señal central decodificada antes o después de llevar a cabo una operación de regeneración de frecuencia. El analizador 602 está configurado para proporcionar un resultado de análisis 603. El regenerador de frecuencia 604 está configurado para regenerar porciones espectrales no incluidas en la señal central decodificada a partir de una porción espectral de la señal central decodificada, datos de envolvente 605 para las porciones
 50 espectrales faltantes y el resultado del análisis 603. De este modo, a diferencia de las antiguas implementaciones, la regeneración de frecuencia en el lado del decodificador no se lleva a cabo independientemente de la señal, sino que se lleva a cabo dependiendo de la señal. Esto ofrece la ventaja de que, cuando no existe ningún problema, la regeneración de frecuencia se lleva a cabo tal cual, pero cuando existen porciones de señal problemáticas, se detectan mediante el resultado del análisis 603 y entonces el regenerador de frecuencia 604 lleva a cabo una forma
 55 adaptada de regeneración de frecuencia que, por ejemplo, puede consistir en el cambio de un límite de frecuencia inicial entre la zona central y la banda de reconstrucción o el cambio de un límite de frecuencia entre dos recuadros/interconexiones individuales dentro de la banda de reconstrucción. Al contrario que la implementación de las bandas de guarda, esta ofrece la ventaja de que los procedimientos concretos solo se llevan a cabo cuando es necesario y no, tal como sucede en la implementación de la banda de guarda, durante todo el tiempo sin depender

en modo alguno de la señal.

[0037] Preferentemente, el decodificador central 600 se implementa como una etapa de decodificación y decuantificación por entropía 612 (por ejemplo, un decodificador de tipo Huffman o aritmético), tal como se ilustra en la fig. 6b. El decodificador central 600 genera como salida un espectro de señal central y el espectro se analiza mediante el analizador espectral 614, que es bastante similar al analizador 602 de la fig. 6a implementado como un analizador espectral en lugar de como un analizador arbitrario cualquiera que podría, tal como se ilustra en la fig. 6a, analizar también una señal en dominio temporal. En la realización de la fig. 6b, el analizador espectral está configurado para analizar la señal espectral de manera que se determinen los mínimos locales en la banda de origen y/o en una banda de destino, es decir, en las interconexiones de frecuencia o recuadros de frecuencia. Después, el regenerador de frecuencia 604 lleva a cabo, tal como se ilustra en 616, una regeneración de frecuencia en la que los límites de las interconexiones se sitúan en los mínimos en la banda de origen y/o la banda de destino.

[0038] Posteriormente, se explica la fig. 7a con el fin de describir una implementación preferida del regenerador de frecuencia 604 de la fig. 6a. Un regenerador de señal preliminar 702 recibe, como entrada, datos de origen procedentes de la banda de origen y, además, información de interconexión preliminar, como, por ejemplo, frecuencias de límite preliminares. Después se genera una señal regenerada preliminar 703, que es detectada por el detector 704 para detectar los componentes tonales dentro de la señal reconstruida preliminar 703. Como otra posibilidad o de forma añadida, los datos de origen 705 también se pueden analizar mediante el detector correspondiente al analizador 602 de la fig. 6a. Por tanto, la etapa de regeneración de señal preliminar no sería necesaria. Cuando hay una correspondencia bien definida entre los datos de origen y los datos de reconstrucción, los mínimos o porciones tonales se pueden detectar incluso teniendo en cuenta únicamente los datos de origen, tanto si hay porciones tonales cercanas al límite superior del rango central, como en un límite de frecuencia entre dos recuadros de frecuencia generados de manera individual, tal como se explicará más adelante con respecto a la fig. 12b.

[0039] En el caso de que se hayan descubierto componentes tonales problemáticos cerca de los límites de frecuencia, un regulador de frecuencia de transición 706 lleva a cabo un ajuste de una frecuencia de transición o frecuencia de cruce o frecuencia de inicio de relleno de espacios entre la banda central y la banda de reconstrucción o entre porciones de frecuencia individuales generadas por los mismos datos de origen en la banda de reconstrucción. La señal de salida del bloque 706 se envía a un eliminador 708 de componentes tonales en los límites. El eliminador está configurado para eliminar los componentes tonales restantes que permanecen allí tras el ajuste de la frecuencia de transición mediante el bloque 706. El resultado del eliminador 708 se envía a continuación a un filtro de cruce 710 con el fin de abordar el problema del zumbido del filtro y el resultado del filtro de cruce 710 se introduce después en un bloque de modelado de la envolvente espectral 712, que lleva a cabo un modelado de la envolvente espectral en la banda de reconstrucción.

[0040] Tal como se explica en el contexto de la fig. 7a, la detección de componentes tonales en el bloque 704 se puede llevar a cabo tanto con unos datos de origen 705 como con una señal reconstruida preliminar 703. Esta realización se ilustra en la fig. 7b, en la que se crea una señal regenerada preliminar, como se muestra en el bloque 718. La señal correspondiente a la señal 703 de la fig. 7a se envía después a un detector 720 que detecta los componentes que generan fallas. Aunque el detector 720 puede estar configurado para ser un detector que detecte componentes tonales en los límites de frecuencia, tal como se ilustra en 704 en la fig. 7a, el detector también se puede implementar para detectar otros componentes generadores de fallas. Dichos componentes espectrales pueden ser incluso componentes distintos de los componentes tonales y se puede llevar a cabo una detección que indique si se ha generado una falla, probando diferentes regeneraciones y comparando los diferentes resultados de las regeneraciones con el fin de descubrir cuál de ellas ha producido componentes que generan fallas.

[0041] El detector 720 controla ahora un manipulador 722 para manipular la señal, es decir, la señal regenerada preliminar. Esta manipulación se puede realizar procesando realmente la señal regenerada preliminar 723 o llevando de nuevo a cabo una regeneración, pero ahora con, por ejemplo, las frecuencias de transición corregidas, como se ilustra con la línea 724.

[0042] Una implementación del procedimiento de manipulación consiste en que la frecuencia de transición se ajusta del modo que se ilustra en 706 de la fig. 7a. En la fig. 8a se ilustra otra implementación, que se puede llevar a cabo en lugar del bloque 706 o junto con el bloque 706 de la fig. 7a. Se proporciona un detector 802 para detectar frecuencias de inicio y final de una porción tonal problemática. A continuación, se configura un interpolador 804 para realizar una interpolación y, preferentemente, una interpolación compleja entre el inicio y el final de la porción tonal dentro del rango espectral. A continuación, tal como se ilustra en la fig. 8a mediante el bloque 806, se sustituye la

porción tonal por el resultado de la interpolación.

[0043] Los bloques 808 y 810 de la fig. 8a ilustran una implementación alternativa. En lugar de llevar a cabo una interpolación, se lleva a cabo una generación aleatoria de líneas espectrales 808 entre el inicio y el final de la porción tonal. Después, se lleva a cabo un ajuste de potencia de las líneas espectrales generadas al azar, tal como se ilustra en 810, y la potencia de las líneas espectrales generadas al azar se fija de manera que la potencia sea similar a la de las partes espectrales no tonales adyacentes. Después, la porción tonal se sustituye por líneas espectrales generadas aleatoriamente y ajustadas por envolvente. Las líneas espectrales pueden ser generadas al azar o de forma pseudoaleatoria con el fin de proporcionar una señal de sustitución que, en la medida de lo posible, carezca de fallas.

[0044] En la fig. 8b se ilustra otra implementación. En el bloque 820 se ilustra un generador de recuadros de frecuencia que se encuentra dentro del generador de frecuencia 604 de la fig. 6a. El generador de recuadros de frecuencia utiliza unos límites de frecuencia predeterminados. El analizador analiza entonces la señal generada por el generador de recuadros de frecuencia, y el generador de recuadros de frecuencia 820 se configura preferentemente para llevar a cabo múltiples operaciones de formación de recuadros para generar múltiples recuadros de frecuencia. Después el manipulador 824 de la fig. 8b manipula el resultado del generador de recuadros de frecuencia de acuerdo con el resultado del análisis producido como salida por el analizador 822. La manipulación puede consistir en el cambio de los límites de frecuencia o la atenuación de porciones individuales. Después, un regulador de envolvente espectral 826 lleva a cabo un ajuste de envolvente espectral empleando la información paramétrica 605, como se explica anteriormente en el contexto de la fig. 6a.

[0045] Después, la señal ajustada espectralmente generada como salida por el bloque 826 se introduce en un convertidor de frecuencia-tiempo que, además, recibe las primeras porciones espectrales, es decir una representación espectral de la señal de salida del decodificador central 600. La salida del convertidor de frecuencia-tiempo 828 se puede utilizar después para almacenarla o transmitirla a un altavoz para reproducir el audio.

[0046] La presente invención se puede aplicar a procedimientos de regeneración de frecuencia como los que se ilustran en las figs. 13a, 13b o se pueden aplicar preferentemente dentro del contexto del relleno inteligente de espacios, que se describe a continuación haciendo referencia a las figs. 1a a 5b y 9a a 10d.

[0047] La fig. 1a ilustra un aparato para codificar una señal de audio 99. La señal de audio 99 se introduce en un convertidor de espectro temporal 100 para convertir una señal de audio con una tasa de muestreo en una representación espectral 101 generada como salida por el convertidor de espectro temporal. El espectro 101 se introduce en un analizador espectral 102 para analizar la representación espectral 101. El analizador espectral 101 está configurado para determinar un primer conjunto de primeras porciones espectrales 103 que se codificarán con una primera resolución espectral y un segundo conjunto diferente de segundas porciones espectrales 105 que se codificarán con una segunda resolución espectral. La segunda resolución espectral es más pequeña que la primera resolución espectral. El segundo conjunto de segundas porciones espectrales 105 se introduce en una calculadora de parámetros o un codificador paramétrico 104 para calcular una información de envolvente espectral con una segunda resolución espectral. Además, se proporciona un codificador de audio de dominio espectral 106 para generar una primera representación codificada 107 del primer conjunto de porciones espectrales con la primera resolución espectral. Además, la calculadora de parámetros/codificador paramétrico 104 están configurados para generar una segunda representación codificada 109 del segundo conjunto de segundas porciones espectrales. La primera representación codificada 107 y la segunda representación codificada 109 se introducen en un multiplexor de corriente de bits o formador de corriente de bits 108 y el bloque 108 finalmente genera como salida la señal de audio codificada para transmitirla o almacenarla en un dispositivo de almacenamiento.

[0048] Normalmente, una primera porción espectral como, por ejemplo, 306 de la fig. 3a estará rodeada por dos porciones espectrales como, por ejemplo, 307a, 307b. Este no es el caso en la HE AAC, en la que el rango de frecuencia del codificador central es de banda limitada.

[0049] La fig. 1 ilustra un decodificador que se corresponde con el codificador de la fig. 1a. La primera representación codificada 107 se introduce en un decodificador de audio de dominio espectral 112 para generar una primera representación decodificada de un primer conjunto de primeras porciones espectrales, y la representación decodificada presenta una primera resolución espectral. Además, la segunda representación codificada 109 se introduce en un decodificador paramétrico 114 para generar una segunda representación decodificada de un segundo conjunto de segundas porciones espectrales con una segunda resolución espectral que es más baja que la primera resolución espectral.

[0050] El decodificador también comprende un regenerador de frecuencia 116 para regenerar una segunda porción espectral reconstruida que presenta la primera resolución espectral, utilizando una primera porción espectral. El regenerador de frecuencia 116 lleva cabo una operación de relleno de recuadros, es decir, utiliza un recuadro o
 5 porción del primer conjunto de primeras porciones espectrales y copia este primer conjunto de primeras porciones espectrales en el rango de reconstrucción o banda de reconstrucción que contiene la segunda porción espectral y normalmente lleva a cabo un modelado de la envolvente espectral u otra operación según lo indique la segunda representación decodificada generada como salida por el decodificador paramétrico 114, es decir, utilizando la información del segundo conjunto de segundas porciones espectrales. El primer conjunto decodificado de primeras
 10 porciones espectrales y el segundo conjunto reconstruido de porciones espectrales según se indica en la salida del regenerador de frecuencia 117 en la línea 117 se introduce en un convertidor de espectro-tiempo 118 configurado para convertir la primera representación decodificada y la segunda porción espectral reconstruida en una representación temporal 119, y la representación temporal presenta una cierta tasa de muestreo elevada.

[0051] La fig. 2b ilustra una implementación del codificador de la fig. 1a. Se introduce una señal de entrada de audio 99 en un banco de filtros de análisis 220 correspondiente al convertidor de tiempo-espectro 100 de la fig. 1a. Después, se lleva a cabo una operación de modelado temporal de ruido (TNS, por sus siglas en inglés) en el bloque de TNS 222. Por lo tanto, la entrada en el analizador espectral 102 de la fig. 1a correspondiente a un enmascaramiento tonal del bloque 226 de la fig. 2b puede consistir en valores espectrales completos, cuando no se
 20 aplica la operación de modelado temporal de recuadros/modelado temporal de ruido, o bien en valores espectrales residuales, cuando se aplica la operación de TNS ilustrada en la fig. 2b, bloque 222. Además, para señales de dos canales o señales multicanal, se puede llevar a cabo una codificación conjunta de canales 228, de manera que el codificador de dominio espectral 106 de la fig. 1a puede comprender el bloque de codificación conjunta de canales 228. Además, se proporciona un codificador por entropía 232 para llevar a cabo una compresión sin pérdidas de los
 25 datos, que también forma parte del codificador de dominio espectral 106 de la fig. 1a.

[0052] El analizador espectral/enmascaramiento tonal 226 divide la salida del bloque de TNS 222 entre la banda central y componentes tonales correspondientes al primer conjunto de porciones espectrales 103 y los componentes residuales correspondientes al segundo conjunto de porciones espectrales 105 de la fig. 1a. El bloque
 30 224 indicado como codificación de extracción de parámetros de IGF corresponde al codificador paramétrico 104 de la fig. 1a y el multiplexor de corriente de bits 230 corresponde al multiplexor de corriente de bits 108 de la fig. 1a.

[0053] Preferentemente, el banco de filtros de análisis 222 se implementa como una MDCT (banco de filtros de transformada de coseno discreta modificada) y la MDCT se utiliza para transformar la señal 99 en un dominio de
 35 tiempo-frecuencia con la transformada de coseno discreta modificada actuando como herramienta de análisis de frecuencia.

[0054] El analizador espectral 226 aplica preferentemente un enmascaramiento de tonalidad. Esta etapa de estimación de enmascaramiento de tonalidad se utiliza para separar los componentes tonales de los componentes
 40 de tipo ruido de la señal. Esto permite al codificador central 228 codificar todos los componentes tonales con un módulo psicoacústico. La etapa de estimación del enmascaramiento de tonalidad se puede implementar de muchas maneras distintas y se implementa preferentemente con una funcionalidad similar a la de la etapa de estimación de la pista sinusoidal utilizada en el modelado de seno y ruido para la codificación de voz/audio [8, 9] o un codificador de audio basado en el modelo HILN descrito en [10]. Preferentemente, se utiliza una implementación que resulta
 45 fácil de implementar sin necesidad de mantener trayectorias de nacimiento-muerte, pero también se puede utilizar cualquier otro detector de tonalidad o ruido.

[0055] El módulo de IGF calcula la semejanza existente entre una zona de origen y una zona de destino. La zona de destino estará representada por el espectro obtenido de la zona de origen. La medida de la semejanza
 50 entre las zonas de origen y de destino se realiza mediante un enfoque de correlación cruzada. La zona de destino se divide en $nTar$ recuadros de frecuencia sin solapamiento. Para cada recuadro de la zona de destino, se crean $nSrc$ recuadros de origen a partir de una frecuencia de inicio fija. Estos recuadros de origen se solapan en un factor de entre 0 y 1, en el que 0 significa un 0% de solapamiento y 1 significa un 100% de solapamiento. Cada uno de estos recuadros de origen se correlaciona con el recuadro de destino a diferentes retardos para hallar el recuadro de
 55 origen que mejor se corresponda con el recuadro de destino. El número del recuadro que mejor se corresponde se almacena en $tileNum[idx_tar]$, el retardo con el que se obtiene la mejor correlación con el destino se almacena en $xcorr_lag[idx_tar][idx_src]$ y el signo de la correlación se almacena en $xcorr_sign[idx_tar][idx_src]$. En el caso de que la correlación sea altamente negativa, es necesario multiplicar el recuadro de origen por -1 antes del proceso de relleno de recuadros en el decodificador. El módulo de IGF también se encarga de no reemplazar los componentes

tonales del espectro, ya que los componentes tonales se conservan mediante el enmascaramiento de tonalidad. Se utiliza un parámetro de potencia por bandas para almacenar la potencia de la zona de destino, lo que nos permite reconstruir el espectro de manera precisa.

5 **[0056]** Este procedimiento ofrece ciertas ventajas frente a la clásica SBR [1], ya que el codificador central conserva la rejilla armónica de una señal multitono, al tiempo que solo los huecos entre las sinusoides se rellenan con el "ruido modelado" que mejor se corresponde, procedente de la zona de origen. Otra ventaja de este sistema en comparación con la ASR (sustitución espectral precisa) [2-4] es la ausencia de una etapa de síntesis de señal que crea las porciones importantes de la señal en el decodificador. En lugar de ello, el codificador central se hace
10 cargo de esta tarea, lo que permite conservar componentes importantes del espectro. Otra ventaja del sistema propuesto es la escalabilidad continua que ofrece gracias a sus características. La simple utilización de un $tileNum[idx_tar]$ y $xcorr_lag = 0$, para cada recuadro se denomina correspondencia de granularidad gruesa y se puede utilizar para bajas tasas de bits, al tiempo que la utilización de un $xcorr_lag$ variable para cada recuadro nos permite establecer una mejor correspondencia entre los espectros de destino y de origen.

15 **[0057]** Además, se propone una técnica de estabilización de elección de recuadro que elimina fallas en el dominio de frecuencia tales como el gorjeo y el ruido musical.

[0058] En el caso de los pares de canales estéreo, se aplica un procesamiento estéreo conjunto adicional.
20 Este procesamiento es necesario debido a que, para un cierto rango de destino, la señal puede ser una fuente de sonido panoramizada que presenta una alta correlación. En el caso de las zonas de origen escogidas para esta zona concreta no están bien correlacionadas, aunque las potencias se corresponden con las zonas de destino, la imagen espacial puede verse deteriorada debido a la falta de correlación de las zonas de origen. El codificador analiza la banda de potencia de cada zona de destino, normalmente, llevando a cabo una correlación cruzada de los valores
25 espectrales y, si se supera un cierto umbral, establece un indicador conjunto para esta banda de potencia. En el decodificador, se tratan de manera individual las bandas de potencia del canal izquierdo y el derecho si no se establece este indicador de estéreo conjunto. En el caso de que se establezca el indicador de estéreo conjunto, tanto las potencias como las interconexiones se llevan a cabo en el dominio de estéreo conjunto. La información de estéreo conjunto para las zonas de IGF se señala de manera similar a la información de estéreo conjunto para la
30 codificación central, incluido un indicador que indica, en caso de predicción, si la dirección de la predicción va desde mezcla reductora hasta residual o viceversa.

[0059] Las potencias se pueden calcular a partir de las potencias transmitidas en el dominio L/R (izquierda/derecha).

35

$$midNrg[k] = leftNrg[k] + rightNrg[k];$$

$$sideNrg[k] = leftNrg[k] - rightNrg[k];$$

40

siendo k el índice de frecuencia en el dominio de la transformada.

[0060] Otra solución consiste en calcular y transmitir las potencias directamente en el dominio de estéreo conjunto para bandas en las que el estéreo conjunto está activado, de manera que no se necesita una
45 transformación de potencia adicional en el lado del decodificador.

[0061] Los recuadros de origen siempre se crean conforme a la matriz Mid/Side (medio/lateral):

50

$$midTile[k] = 0,5 \cdot (leftTile[k] + rightTile[k])$$

$$sideTile[k] = 0,5 \cdot (leftTile[k] - rightTile[k])$$

[0062] Ajuste de potencia:

55

$$midTile[k] = midTile[k] * midNrg[k];$$

$$sideTile[k] = sideTile[k] * sideNrg[k];$$

Transformación estéreo conjunto → LR:

[0063] Si no se codifica ningún parámetro de predicción adicional:

5

$$\text{leftTile}[k] = \text{midTile}[k] + \text{sideTile}[k]$$

$$\text{rightTile}[k] = \text{midTile}[k] - \text{sideTile}[k]$$

10 **[0064]** Si se codifica un parámetro de predicción adicional y si la dirección señalada va del medio al lateral:

$$\text{sideTile}[k] = \text{sideTile}[k] - \text{predictionCoeff} \cdot \text{midTile}[k]$$

$$\text{leftTile}[k] = \text{midTile}[k] + \text{sideTile}[k]$$

15

$$\text{rightTile}[k] = \text{midTile}[k] - \text{sideTile}[k]$$

[0065] Si la dirección señalada va del lateral al medio:

20

$$\text{midTile}[k] = \text{midTile}[k] - \text{predictionCoeff} \cdot \text{sideTile}[k]$$

$$\text{leftTile}[k] = \text{midTile}[k] - \text{sideTile}[k]$$

$$\text{rightTile}[k] = \text{midTile}[k] + \text{sideTile}[k]$$

25

[0066] Este procesamiento garantiza que a partir de los recuadros utilizados para regenerar zonas de destino con una alta correlación y zonas de destino panoramizadas, los canales izquierdo y derecho resultantes siguen representando una fuente de sonido correlacionada y panoramizada aunque las zonas de origen no estén correlacionadas, y se conserva la imagen estéreo para dichas zonas.

30

[0067] Expresado de otro modo, en la corriente de bits, se transmiten indicadores de estéreo conjunto que indican si se utilizará L/R o M/S como ejemplo para la codificación general de estéreo conjunto. En el decodificador, en primer lugar, se decodifica la señal central según indican los indicadores estéreo para las bandas centrales. En segundo lugar, la señal central se almacena tanto en una representación L/R como M/S. Para el relleno de recuadros IGF, se escoge la representación de los recuadros de origen de manera que se ajuste a la representación de los recuadros de destino, tal como indica la información de estéreo conjunto para las bandas de IGF. El modelado temporal de ruido (TNS) es una técnica estándar y forma parte de la AAC [11-13]. El TNS se puede considerar como una ampliación del esquema básico de un codificador perceptual, en la que se intercala una etapa de procesamiento opcional entre el banco de filtros y la etapa de cuantificación. La tarea principal del módulo de TNS consiste en ocultar el ruido producido por la cuantificación en la zona de enmascaramiento temporal de señales de tipo transitorio y, por tanto, da lugar a un esquema de codificación más eficiente. En primer lugar, el TNS calcula un conjunto de coeficientes de predicción mediante una "predicción hacia delante" en el dominio de la transformada, por ejemplo, MDCT. Estos coeficientes se usan más tarde para aplanar la envolvente temporal de la señal. Debido a que la cuantificación afecta al espectro filtrado mediante el TNS, el ruido de la cuantificación también es temporalmente plano. Al aplicar el filtrado TNS inverso en el lado del decodificador, el ruido de cuantificación se modela de acuerdo con la envolvente temporal del filtro TNS y, por lo tanto, el ruido de cuantificación queda enmascarado por el transitorio.

35

40

45

[0068] El IGF se basa en una representación de tipo MDCT. Para lograr una codificación eficiente, se deben utilizar preferentemente bloques de aproximadamente 20 ms. Si la señal contenida en un bloque con semejante longitud contiene transitorios, se producen preecos y posecos audibles en las bandas espectrales de IGF a causa del relleno de recuadros. La fig. 7c muestra un típico efecto de preeco anterior a la aparición del transitorio a causa del IGF. En el lado izquierdo, se muestra el espectrograma de la señal original y, en el lado derecho, se muestra el espectrograma de la señal de ancho de banda ampliado sin filtrado TNS.

55

[0069] Este efecto de preeco se reduce mediante el TNS en el contexto del IGF. Aquí, el TNS se utiliza como

una herramienta de modelado de recuadros temporales (TTS, por sus siglas en inglés) cuando la regeneración espectral en el decodificador se lleva a cabo con la señal residual del TNS. Los coeficientes de predicción de TTS necesarios se calculan y se aplican utilizando el espectro completo en el codificador, como es habitual. Las frecuencias de inicio y fin de TNS/TTS no se ven afectadas por la frecuencia de inicio del IGF $f_{IGFstart}$ de la herramienta de IGF. En comparación con el TNS legado, la frecuencia de fin de TTS se incrementa hasta la frecuencia de fin de la herramienta IGF, que es mayor que $f_{IGFstart}$. En el lado del decodificador, los coeficientes TNS/TTS se aplican de nuevo sobre la totalidad del espectro, es decir, el espectro central más el espectro regenerado, más los componentes tonales obtenidos a partir del mapa de tonalidades (véase la fig. 7e). La aplicación del TTS es necesaria para formar la envolvente temporal del espectro regenerado de forma que se corresponda de nuevo con la envolvente de la señal original. De este modo, se reducen los preecos. Además, continúa modelando el ruido de cuantificación en la señal por debajo de $f_{IGFstart}$, como es habitual con el TNS.

[0070] En decodificadores legados la interconexión espectral en una señal de audio corrompe la correlación espectral en los límites de la interconexión y, de ese modo, altera la envolvente temporal de la señal de audio introduciendo dispersión. Por tanto, otra ventaja de llevar a cabo el relleno de recuadros IGF en la señal residual consiste en que, tras la aplicación del filtro de modelado, los límites de los recuadros se correlacionan a la perfección, lo que da lugar a una reproducción temporal más fiel de la señal.

[0071] En un codificador de la invención, el espectro que ha sido sometido al filtrado TNS/TTS, procesamiento de enmascaramiento de tonalidad y estimación de parámetros IGF no presenta ninguna señal por encima de la frecuencia de inicio de IGF. Este espectro disperso se codifica ahora mediante el codificador central aplicando principios de codificación aritmética y codificación predictiva. Estos componentes codificados, junto con los bits de señalización, forman la corriente de bits del audio.

[0072] La fig. 2a ilustra la implementación del correspondiente decodificador. La corriente de bits de la fig. 2a correspondiente a la señal de audio codificada se introduce en el demultiplexor/decodificador que estaría conectado, con respecto a la fig. 1b, con los bloques 112 y 114. El demultiplexor de corriente de bits divide la señal de audio de entrada entre la primera representación codificada 107 de la fig. 1b y la segunda representación codificada 109 de la fig. 1b. La primera representación codificada que contiene el primer conjunto de primeras porciones espectrales se introduce en el bloque de decodificación de canales conjuntos 204, correspondiente al dominio espectral del decodificador 112 de la fig. 1b. La segunda representación codificada se introduce en el decodificador paramétrico 114 que no se ilustra en la fig. 2a, y después se introduce en el bloque de IGF 202 correspondiente al regenerador de frecuencia 116 de la fig. 1b. El primer conjunto de primeras porciones espectrales necesario para la regeneración de frecuencia se introduce en el bloque de IGF 202 a través de la línea 203. Además, posteriormente a la decodificación de canales conjuntos 204, se aplica la decodificación central específica en el bloque de enmascaramiento tonal 206, de manera que la salida del enmascaramiento tonal 206 corresponde a la salida del decodificador de dominio espectral 112. Después, se lleva a cabo una combinación mediante el combinador 208, es decir, una construcción de cuadros en la que la salida del combinador 208 contiene ahora el espectro de rango completo, pero aún en el dominio filtrado mediante TNS/TTS. Después, en el bloque 210, se lleva a cabo una operación de TNS/TTS inversa a partir de la información del filtro TNS/TTS proporcionada a través de la línea 109, es decir, la información secundaria de TTS se incluye preferentemente en la primera representación codificada generada por el codificador de dominio espectral 106 que, por ejemplo, puede ser un codificador central de tipo AAC o USAC, o también puede incluirse en la segunda representación codificada. En la salida del bloque 210, se proporciona un espectro completo hasta la frecuencia máxima, que consiste en la frecuencia de rango completo definida por la tasa de muestreo de la señal de entrada original. Después, se lleva a cabo una conversión de espectro/tiempo en el banco de filtros de síntesis 212 para obtener finalmente la señal de audio de salida.

[0073] La fig. 3a ilustra una representación esquemática del espectro. El espectro se subdivide en bandas de factor de escala SCB y presenta siete bandas de factor de escala SCB1 a SCB7 en el ejemplo ilustrado de la fig. 3a. Las bandas de factor de escala pueden ser bandas de factor de escala AAC que se definen en la norma AAC y poseen un ancho de banda creciente para frecuencias superiores, tal como se ilustra de manera esquemática en la fig. 3a. Es preferible llevar a cabo el relleno inteligente de espacios, no desde el mismo comienzo del espectro, es decir, a bajas frecuencias, sino iniciando la operación de IGF en una frecuencia de inicio de IGF ilustrada en 309. Por lo tanto, la banda de frecuencia central se extiende desde la frecuencia más baja hasta la frecuencia de inicio de IGF. Por encima de la frecuencia de inicio de IGF, se aplica el análisis de espectro para separar los componentes espectrales de alta resolución 304, 305, 306 y 307 (el primer conjunto de primeras porciones espectrales) de los componentes de baja resolución representados por el segundo conjunto de porciones espectrales. La fig. 3a ilustra un espectro que se introduce, ejemplarmente, en el codificador de dominio espectral 106 o el codificador de canales conjuntos 228, es decir, el codificador central actúa sobre el rango completo, pero codifica una cantidad considerable

de valores espectrales cero, es decir, estos valores espectrales cero se cuantifican a cero o se fijan en cero antes de la cuantificación o después de la cuantificación. En cualquier caso, el codificador central actúa en el rango completo, es decir, como si el espectro fuera como el que se ilustra, es decir, no es necesario que el decodificador central tenga conocimiento de cualquier relleno inteligente de espacios o codificación del segundo conjunto de segundas 5 porciones espectrales con una resolución espectral más baja.

[0074] Preferentemente, la alta resolución se define mediante una codificación por líneas de líneas espectrales tales como líneas MDCT, mientras que la segunda resolución o baja resolución se define, por ejemplo, mediante el cálculo de un valor espectral individual por cada banda de factor de escala, en el que una banda de 10 factor de escala abarca varias líneas de frecuencia. De este modo, la segunda resolución es, con respecto a su resolución espectral, mucho más baja que la primera o alta resolución definida mediante la codificación por líneas que se aplica normalmente mediante el codificador central, por ejemplo un codificador central de tipo AAC o USAC.

[0075] En lo que respecta al cálculo de los factores de ajuste de escala o potencias, la situación se ilustra en 15 la fig. 3b. Debido al hecho de que el codificador es un codificador central y debido al hecho de que puede haber, aunque no necesariamente, componentes del primer conjunto de porciones espectrales en cada banda, el codificador central calcula un factor de ajuste de escala para cada banda, no solo en el rango central por debajo de la frecuencia de inicio de IGF 309, sino también por encima de la frecuencia de inicio de IGF hasta la frecuencia máxima $f_{IGFstop}$, que es menor o igual que la mitad de la frecuencia de muestreo, es decir, $f_s/2$. De este modo, las 20 porciones tonales codificadas 302, 304, 305, 306, 307 de la fig. 3a y, en esta realización, junto con los factores de escala SCB1 a SCB7, corresponden a los datos espectrales de alta resolución. Los datos espectrales de baja resolución se calculan comenzando desde la frecuencia de inicio de IGF y corresponden a los valores de información de potencia E_1 , E_2 , E_3 y E_4 , que se transmiten junto con los factores de ajuste de escala SF4 a SF7.

[0076] En particular, cuando el codificador central se encuentra en un estado de baja tasa de bits, se puede 25 aplicar una operación adicional de relleno de ruido en la banda central, es decir con una frecuencia más baja que la frecuencia de inicio de IGF, es decir, en las bandas de factor de escala SCB1 a SCB3. En el relleno de ruido, existen varias líneas espectrales adyacentes que se han cuantificado a cero. En el lado del decodificador, se resintetizan estos valores espectrales cuantificados a cero, y los valores espectrales resintetizados se ajustan en cuanto a su 30 magnitud empleando una potencia de relleno de ruido tal como NF_2 ilustrada en 308 de la fig. 3b. La potencia de relleno de ruido, que se puede dar en términos absolutos o en términos relativos, particularmente con respecto al factor de ajuste de escala como en USAC, corresponde a la potencia del conjunto de valores espectrales cuantificados a cero. Estas líneas de relleno de ruido también se pueden considerar como un tercer conjunto de 35 terceras porciones espectrales que se regeneran mediante una sencilla síntesis de relleno de ruido sin que ninguna operación de IGF recurra a la regeneración de la potencia a partir de recuadros de frecuencia procedentes de otras frecuencias para reconstruir recuadros de frecuencia a partir de valores espectrales procedentes de un rango de origen y la información de la potencia E_1 , E_2 , E_3 y E_4 .

[0077] Preferentemente, las bandas para las que se calcula la información de la potencia coinciden con las 40 bandas de factor de escala. En otras realizaciones se aplica un agrupamiento de valores de información de potencia, de manera que, por ejemplo, para las bandas de factor de escala 4 y 5, solo se transmite un valor individual de información de potencia, pero incluso en esta realización, los límites de las bandas de reconstrucción agrupadas coinciden con los límites de las bandas de factor de escala. Si se aplican diferentes separaciones de bandas, entonces se pueden aplicar ciertos recálculos o cálculos de sincronización, y esto puede tener sentido dependiendo 45 de la implementación concreta.

[0078] Preferentemente, el codificador de dominio espectral 106 de la fig. 1a es un codificador basado en la psicoacústica tal como se ilustra en la fig. 4a. Normalmente, como se ilustra, por ejemplo, en la norma MPEG2/4 AAC o MPEG1/2, capa 3, la señal de audio que se codificará tras haberla transformado en el rango espectral (401 50 en la fig. 4a) se envía a una calculadora de factor de ajuste de escala 400. La calculadora de factor de ajuste de escala está controlada por un modelo psicoacústico que recibe además la señal de audio que se desea cuantificar o recibe, como en la norma MPEG1/2 capa 3 o MPEG AAC, una representación espectral compleja de la señal de audio. El modelo psicoacústico calcula, para cada banda de factor de escala, un factor de ajuste de escala que representa el umbral psicoacústico. Además, los factores de ajuste de escala se ajustan después, mediante la 55 cooperación de los ya conocidos bucles de iteración interno y externo o mediante cualquier otro procedimiento de codificación adecuado, de manera que se cumplan ciertas condiciones en cuanto a la tasa de bits. Después, los valores espectrales que se van a cuantificar, por una parte, y los factores de ajuste de escala calculados, por otra, se introducen en un procesador cuantificador 404. En la sencilla operación realizada por el codificador de audio, los valores espectrales que se van a cuantificar se ponderan mediante factores de ajuste de escala y los valores

ponderados se introducen después en un cuantificador fijo que normalmente incluye una funcionalidad de compresión hasta rangos superiores de amplitud. Entonces, en la salida del procesador cuantificador, existen unos índices de cuantificación que se envían después a un codificador por entropía que normalmente ofrece una codificación específica y muy eficiente para un conjunto de índices de cuantificación a cero para valores de frecuencia adyacentes o, como también se denominan en la técnica, una “racha” de valores cero.

[0079] No obstante, en el codificador de audio de la fig. 1a, normalmente, el procesador cuantificador recibe información sobre las segundas porciones espectrales procedente del analizador espectral. Así, el procesador cuantificador 404 se asegura de que, en la salida del procesador cuantificador 404, las segundas porciones espectrales identificadas por el analizador espectral 102 son cero o tienen una representación reconocida por un codificador o un decodificador como una representación cero que se puede codificar de manera muy eficiente, concretamente cuando existen “rachas” de valores cero en el espectro.

[0080] La fig. 4b ilustra una implementación del procesador cuantificador. Los valores espectrales MDCT se pueden introducir en un bloque de fijado en cero 410. Entonces, las segundas porciones espectrales están ya fijadas en cero antes de que se lleve a cabo una ponderación mediante los factores de ajuste de escala en el bloque 418 posterior al bloque de ponderación 412. En otra implementación, la operación de fijado en cero también se puede llevar a cabo en un bloque de fijado en cero 422 posterior a una cuantificación en el bloque cuantificador 420. En esta implementación, los bloques 410 y 418 no estarían presentes. En general, se proporciona al menos uno de los bloques 410, 418 y 422 según la implementación concreta.

[0081] Después, en la salida del bloque 422, se obtiene un espectro cuantificado correspondiente a lo que se muestra en la ilustración de la fig. 3a. Este espectro cuantificado se introduce después en un codificador por entropía como el 232 de la fig. 2b, que puede ser un codificador de tipo Huffman o un codificador aritmético, tal como se define, por ejemplo, en la norma USAC.

[0082] Los bloques de fijado en cero 410, 418 y 422, que se proporcionan como distintas alternativas o en paralelo, se controlan mediante el analizador espectral 424. El analizador espectral comprende preferentemente cualquier implementación del ya conocido detector de tonalidad o comprende cualquier tipo diferente de detector que sea capaz de dividir un espectro entre componentes que se codificarán con una alta resolución y componentes que se codificarán con una baja resolución. Otros algoritmos de este tipo implementados en el analizador espectral pueden ser un detector de actividad vocal, un detector de ruido, un detector de habla o cualquier otro detector que decide en función de la información espectral o metadatos asociados sobre los requisitos de resolución para las diferentes porciones espectrales.

[0083] La fig. 5a ilustra una implementación preferida del convertidor de espectro 100 de la fig. 1a tal como, por ejemplo, se implementa en AAC o USAC. El convertidor de tiempo-espectro 100 comprende un divisor de ventanas 502 controlado por un detector de transitorios 504. Cuando el detector de transitorios 504 detecta un transitorio, el divisor de ventanas recibe una indicación de cambio de ventanas largas a ventanas cortas. El divisor de ventanas 502 calcula entonces, para bloques de solapamiento, cuadros con ventana, en los que cada cuadro con ventana posee normalmente dos valores N como, por ejemplo, 2048 valores. Después se lleva a cabo una transformación dentro de un transformador de bloques 506, y este transformador de bloques normalmente proporciona, además, una eliminación, de manera que se lleva a cabo una combinación de eliminación/transformada para obtener un cuadro espectral con valores N como, por ejemplo, valores espectrales MDCT. De este modo, para una operación de ventana larga, el cuadro en la entrada del bloque 506 comprende dos valores N como, por ejemplo, 2048 valores y, entonces, un cuadro espectral posee 1024 valores. No obstante, después se lleva a cabo un cambio a bloques cortos, cuando se llevan a cabo ocho bloques cortos en los que cada bloque corto posee valores en dominio temporal en ventanas de 1/8 en comparación con una ventana larga y cada bloque espectral posee valores espectrales de 1/8 en comparación con un bloque largo. De este modo, cuando esta eliminación se combina con una operación del divisor de ventanas con un solapamiento del 50%, el espectro es una versión muestreada críticamente de la señal de audio en dominio temporal 99.

[0084] A continuación, se hace referencia a la fig. 5b que ilustra una implementación concreta del regenerador de frecuencia 116 y el convertidor de espectro-tiempo 118 de la fig. 1b, o de la operación combinada de los bloques 208 y 212 de la fig. 2a. En la fig. 5b, se considera una reconstrucción de banda concreta como la banda de factor de escala 6 de la fig. 3a. La primera porción espectral en esta banda de reconstrucción, es decir la primera porción espectral 306 de la fig. 3a, se introduce en el bloque del constructor de cuadros/regulador 510. Además, también se introduce en el constructor de cuadros/regulador 510 una segunda porción espectral reconstruida para la banda de factor de escala 6. Además, también se introduce en el constructor de cuadros/regulador 510 una

información de potencia tal como E_3 de la fig. 3b, para una banda de factor de escala 6. La segunda porción espectral reconstruida en la banda de reconstrucción ya se ha generado mediante el relleno de recuadros de frecuencia utilizando un rango de origen, y la banda de reconstrucción corresponde entonces al rango de destino. Ahora se lleva a cabo un ajuste de potencia del cuadro para obtener finalmente el cuadro reconstruido completo con los valores N, por ejemplo, obtenidos en la salida del combinador 208 de la fig. 2a. Después, en el bloque 512, se lleva a cabo una transformada/interpolación inversa de bloques para obtener 248 valores de dominio temporal para los, por ejemplo, 124 valores espectrales en la entrada del bloque 512. Después, se lleva a cabo una operación de división de ventana de síntesis en el bloque 514, que, de nuevo, se controla mediante una indicación de ventana larga/ventana corta transmitida como información secundaria en la señal de audio codificada. Después, en el bloque 516, se lleva a cabo una operación de solapamiento/agregación con un cuadro de tiempo anterior. Preferentemente, la MDCT aplica un solapamiento del 50%, de manera que, para cada nuevo cuadro temporal de valores 2N, finalmente se generan como salida valores N en dominio temporal. Se prefiere, con diferencia, un solapamiento del 50% debido al hecho de que proporciona un muestreo crítico y un cruce continuo de un cuadro al siguiente a causa de la operación de solapamiento/agregación del bloque 516.

15 **[0085]** Como se ilustra en 301 de la fig. 3a, se puede aplicar, además, una operación de relleno de ruido, no solo por debajo de la frecuencia de inicio de IGF, sino también por encima de la frecuencia de inicio de IGF, al igual que para la banda de reconstrucción contemplada que coincide con la banda de factor de escala 6 de la fig. 3a. Después, también se pueden introducir valores espectrales de relleno de ruido en el constructor de cuadros/regulador 510, y dentro de este bloque también se puede aplicar el ajuste de los valores espectrales de relleno de ruido, o los valores espectrales de relleno de ruido pueden estar ya ajustados a partir de la potencia de relleno de ruido antes de introducirlos en el constructor de cuadros/regulador 510.

20 **[0086]** Preferentemente, una operación de IGF, es decir una operación de relleno de recuadros de frecuencia a partir de valores espectrales procedentes de otras porciones, se puede aplicar en el espectro completo. De este modo, se puede aplicar una operación de relleno de recuadros espectrales, no solo en la banda alta por encima de una frecuencia de inicio de IGF, sino también en la banda baja. Además, también se puede aplicar el relleno de ruido sin relleno de recuadros de frecuencia, no solo por debajo de la frecuencia de inicio de IGF, sino también por encima de la frecuencia de inicio de IGF. No obstante, se ha descubierto que se puede obtener una codificación de audio de alta calidad y eficiencia cuando la operación de relleno de ruido se limita al rango de frecuencias situado por debajo de la frecuencia de inicio de IGF y cuando la operación de relleno de recuadros de frecuencia se restringe al rango de frecuencias situado por encima de la frecuencia de inicio de IGF, tal como se ilustra en la fig. 3a.

30 **[0087]** Preferentemente, los recuadros de destino (TT) (con frecuencias superiores a la frecuencia de inicio de IGF) están delimitados a los límites de la banda de factor de escala del codificador de tasa completa. Los recuadros de origen (ST), a partir de los cuales se toma la información, es decir, para frecuencias inferiores a la frecuencia de inicio de IGF, no están delimitados por los límites de la banda de factor de escala. El tamaño de los ST debería corresponder al tamaño de los TT a los que están asociados. Esto se ilustra mediante el siguiente ejemplo. TT[0] tiene una longitud de 10 bins de MDCT. Esto corresponde exactamente a la longitud de dos posteriores SCB (por ejemplo, 4 +6). Entonces, todos los posibles ST que se van a correlacionar con TT[0] también tienen una longitud de 10 bins. Un segundo recuadro de destino TT[1] que es adyacente a TT[0] tiene una longitud de 15 bins (SCB con una longitud de 7 + 8). Entonces, los ST para el mismo tienen una longitud de 15 bins, en lugar de los 10 bins para el caso de TT[0].

45 **[0088]** Si se diera el caso de que no se pudiera encontrar un TT para un ST con la longitud del recuadro de destino (cuando, por ejemplo, la longitud del TT es mayor que el rango de origen disponible), entonces no se calcula la correlación y el rango de origen se copia varias veces en este TT (la copias se realizan una después de otra, de manera que una línea de frecuencia para la frecuencia más baja de la segunda copia sigue inmediatamente, en frecuencia, a la línea de frecuencia para la frecuencia más alta de la primera copia), hasta que el recuadro de destino TT esté completamente lleno.

50 **[0089]** A continuación, se hace referencia a la fig. 5c que ilustra otra realización preferida del regenerador de frecuencia 116 de la fig. 1b o el bloque de IGF 202 de la fig. 2a. El bloque 522 es un generador de recuadros de frecuencia que recibe, no solo una identificación (ID) de banda de destino, sino además una ID de banda de origen. Ejemplarmente, se ha determinado en el lado del codificador que la banda de factor de escala 3 de la fig. 3a resulta muy adecuada para la reconstrucción de la banda de factor de escala 7. De este modo, la ID de la banda de origen sería 2 y la ID de la banda de destino sería 7. Basándose en esta información, el generador de recuadros de frecuencia 522 aplica una operación de copiado o relleno de recuadros armónicos o cualquier otro tipo de operación de relleno de recuadros para generar la segunda porción en bruto de componentes espectrales 523. La segunda

porción en bruto de componentes espectrales posee una resolución de frecuencia idéntica a la resolución de frecuencia incluida en el primer conjunto de primeras porciones espectrales.

- [0090]** Después, la primera porción espectral de la banda de reconstrucción, por ejemplo, 307 de la fig. 3a, se introduce en un constructor de cuadros 524 y la segunda porción en bruto 523 también se introduce en el constructor de cuadros 524. Después, el cuadro reconstruido se ajusta mediante el regulador 526 empleando un factor de ganancia para la banda de reconstrucción calculado por la calculadora del factor de ganancia 528. No obstante es importante que la primera porción espectral del cuadro no resulte afectada por el regulador 526, sino que solo la segunda porción en bruto para el cuadro de reconstrucción resulte afectada por el regulador 526. Para lograrlo, la calculadora del factor de ganancia 528 analiza la banda de origen o la segunda porción en bruto 523 y, además, analiza la primera porción espectral de la banda de reconstrucción para hallar finalmente el factor de ganancia corregido 527, de manera que la potencia del cuadro ajustado generado como salida por el regulador 526 tenga la potencia E₄ cuando se contempla una banda de factor de escala 7.
- [0091]** En este contexto, es muy importante evaluar la precisión en la reconstrucción de altas frecuencias que ofrece la presente invención en comparación con la HE-AAC. Esto se explica con respecto a la banda de factor de escala 7 de la fig. 3a. Se supone que un codificador de la técnica anterior como el que se ilustra en la fig. 13a detectaría la porción espectral 307 que se va a codificar con una alta resolución como “armónicos faltantes”. Entonces, la potencia de este componente espectral se transmitiría al decodificador junto con una información de envolvente espectral para la banda de reconstrucción, por ejemplo, la banda de factor de escala 7. Después, el decodificador recrearía el armónico faltante. No obstante, el valor espectral al que el armónico faltante 307 sería reconstruido por el decodificador de la técnica anterior de la fig. 13 estaría en el medio de la banda 7 a una frecuencia indicada por la frecuencia de reconstrucción 390. De este modo, la presente invención evita un error de frecuencia 391 que el decodificador de la técnica anterior de la fig. 13d sí introduciría.
- [0092]** En una implementación, el analizador espectral también se implementa para calcular semejanzas entre las primeras porciones espectrales y las segundas porciones espectrales y para determinar, basándose en las semejanzas calculadas, para una segunda porción espectral en un rango de reconstrucción una primera porción espectral que se corresponde lo más posible con la segunda porción espectral. A continuación, en esta implementación de rango de origen/rango de destino variable, el codificador paramétrico introducirá además un intervalo de origen que se corresponda. En el lado del decodificador, esta información sería utilizada después por un generador de recuadros de frecuencia 522 de la fig. 5c que ilustra una generación de una segunda porción en bruto 523 basándose en una ID de banda de origen y una ID de banda de destino.
- [0093]** Además, tal como se ilustra en la fig. 3a, el analizador espectral está configurado para analizar la representación espectral hasta una frecuencia máxima de análisis que es solo un poco menos de la mitad de la frecuencia de muestreo y, preferentemente, es de al menos un cuarto de la frecuencia de muestreo o, normalmente, más alta.
- [0094]** Tal como se ilustra, el codificador actúa sin muestreo descendente y el decodificador actúa sin muestreo ascendente. Expresado de otro modo, el codificador de audio de dominio espectral está configurado para generar una representación espectral con una frecuencia de Nyquist definida por la tasa de muestreo de la señal de audio introducida originalmente.
- [0095]** Además, tal como se ilustra en la fig. 3a, el analizador espectral está configurado para analizar la representación comenzando con una frecuencia de inicio de relleno de espacios y terminando con una frecuencia máxima representada por una frecuencia máxima incluida en la representación espectral, en la que una porción espectral que se extiende desde una frecuencia mínima hasta la frecuencia de inicio de relleno de espacios pertenece al primer conjunto de porciones espectrales y en la que otra porción espectral, por ejemplo, 304, 305, 306 o 307, con valores de frecuencia superiores a la frecuencia de relleno de espacios se incluye también en el primer conjunto de primeras porciones espectrales.
- [0096]** Tal como se ha explicado en líneas generales, el decodificador de audio de dominio espectral 112 está configurado de manera que una frecuencia máxima representada por un valor espectral en la primera representación decodificada es igual a una frecuencia máxima incluida en la representación temporal que presenta la tasa de muestreo en la que el valor espectral para la frecuencia máxima en el primer conjunto de primeras porciones espectrales es cero o distinta de cero. En cualquier caso, para esta frecuencia máxima en el primer conjunto de primeras porciones espectrales, existe un factor de ajuste de escala para la banda de factor de escala, que se genera y transmite independientemente de si todos los valores espectrales en esta banda de factor de escala están

fijados en cero o no, tal como se explica en el contexto de las figs. 3a y 3b.

[0097] Por lo tanto, la invención resulta ventajosa con respecto a otras técnicas paramétricas para aumentar la eficiencia de la compresión, por ejemplo, la sustitución de ruido y el relleno de ruido (estas técnicas sirven exclusivamente para la representación eficiente de un contenido de señal local de tipo ruido), ya que la invención permite obtener una representación precisa de la frecuencia de componentes tonales. Hasta la fecha, ninguna técnica del estado de la técnica aborda la representación paramétrica eficiente de un contenido de señal arbitrario mediante el relleno de espacios espectrales sin la restricción de una división fijada *a priori* en la banda baja (LF) y la banda alta (HF).

10

[0098] Las realizaciones del sistema de la invención mejoran los enfoques del estado de la técnica y, de este modo, proporcionan una alta eficiencia de compresión, sin ninguna molestia perceptual, o solo un pequeño grado de molestia, y un ancho de banda de audio completo, incluso para bajas tasas de bits.

15 **[0099]** El sistema general consiste en:

- codificación central de banda completa
- relleno inteligente de espacios (relleno de recuadros o relleno de ruido)
- partes tonales dispersas en centro seleccionadas mediante enmascaramiento tonal
- 20 • codificación de par estéreo conjunto para banda completa, incluido el relleno de recuadros
- TNS en recuadro
- blanqueamiento espectral en rango de IGF

[0100] Un primer paso para lograr un sistema más eficiente consiste en suprimir la necesidad de transformar datos espectrales en un segundo dominio de transformada distinto al del codificador central. Debido a que la mayoría de los códecs de audio, como, por ejemplo, AAC, utilizan la MDCT como transformada básica, resulta útil llevar a cabo la BWE también en el dominio de la MDCT. Un segundo requisito para el sistema BWE sería la necesidad de conservar la rejilla tonal, con lo que se conservan incluso los componentes tonales de HF y, por tanto, la calidad del audio codificado es superior a la de los sistemas existentes. Para atender a ambos requisitos para un esquema BWE mencionados anteriormente, se propone un nuevo sistema denominado relleno inteligente de espacios (IGF). La fig. 2b muestra el diagrama de bloques del sistema propuesto en el lado del codificador y la fig. 2a muestra el sistema en el lado del decodificador.

[0101] La fig. 9a ilustra un aparato para decodificar una señal de audio codificada que comprende una representación codificada de un primer conjunto de primeras porciones espectrales y una representación codificada de unos datos paramétricos que indican potencias espectrales para un segundo conjunto de segundas porciones espectrales. El primer conjunto de primeras porciones espectrales se indica en 901a de la fig. 9a, y la representación codificada de los datos paramétricos se indica en 901b en la fig. 9a. Se proporciona un decodificador de audio 900 para decodificar la representación codificada 901a del primer conjunto de primeras porciones espectrales para obtener un primer conjunto de primeras porciones espectrales decodificado 904 y para decodificar la representación codificada de los datos paramétricos para obtener unos datos paramétricos decodificados 902 para el segundo conjunto de segundas porciones espectrales que indican potencias individuales para bandas de reconstrucción individuales, en la que las segundas porciones espectrales están situadas en las bandas de reconstrucción. Además, se proporciona un regenerador de frecuencia 906 para reconstruir valores espectrales de una banda de reconstrucción que comprende una segunda porción espectral. El regenerador de frecuencia 906 utiliza una primera porción espectral del primer conjunto de primeras porciones espectrales y una información de potencia individual para la banda de reconstrucción, en la que la banda de reconstrucción comprende una primera porción espectral y la segunda porción espectral. El regenerador de frecuencia 906 comprende una calculadora 912 para determinar una información de la potencia de conservación que comprende una potencia acumulada de la primera porción espectral con frecuencias en la banda de reconstrucción. Además, el regenerador de frecuencia 906 comprende una calculadora 912 para determinar una información de la potencia de conservación que comprende una potencia acumulada de la primera porción espectral con frecuencias en la banda de reconstrucción. Además, el regenerador de frecuencia 906 comprende una calculadora 918 para determinar una información de la potencia de recuadros de otras porciones espectrales de la banda de reconstrucción y para valores de frecuencia que son diferentes de la primera porción espectral, en la que estos valores de frecuencia tienen frecuencias en la banda de reconstrucción, en la que se van a generar las otras porciones espectrales mediante la regeneración de frecuencia a partir de una primera porción espectral diferente de la primera porción espectral de la banda de reconstrucción.

[0102] El regenerador de frecuencia 906 también comprende una calculadora 914 para una potencia faltante

en la banda de reconstrucción, y la calculadora 914 actúa empleado la potencia individual para la banda de reconstrucción y la potencia de conservación generada por el bloque 912. Además, el regenerador de frecuencia 906 comprende un regulador de la envolvente espectral 916 para ajustar las otras porciones espectrales de la banda de reconstrucción basándose en la información de la potencia faltante y la información de la potencia de recuadros generada por el bloque 918.

[0103] Se hace referencia a la fig. 9c que ilustra una cierta banda de reconstrucción 920. La banda de reconstrucción comprende una primera porción espectral de la banda de reconstrucción, como la primera porción espectral 306 de la fig. 3a ilustrada de manera esquemática en 921. Además, el resto de los valores espectrales en la banda de reconstrucción 920 se generarán a partir de una zona de origen, por ejemplo, desde la banda de factor de escala 1, 2, 3 por debajo de la frecuencia de inicio de relleno inteligente de espacios 309 de la fig. 3a. El regenerador de frecuencia 906 está configurado para generar valores espectrales en bruto para las segundas porciones espectrales 922 y 923. Después, se calcula un factor de ganancia g tal como se ilustra en la fig. 9c para ajustar finalmente los valores espectrales en bruto en las bandas de frecuencia 922 y 923 con el fin de obtener las segundas porciones espectrales reconstruidas y ajustadas en la banda de reconstrucción 920 que ahora presentan la misma resolución espectral, es decir la misma distancia entre líneas que la primera porción espectral 921. Es importante entender que la primera porción espectral de la banda de reconstrucción que se ilustra en 921 de la fig. 9c se decodifica mediante el decodificador de audio 900 y no se ve afectada por el ajuste de envolvente llevado a cabo en el bloque 916 de la fig. 9b. En lugar de ello, la primera porción espectral en la banda de reconstrucción indicada en 921 se deja tal y como está, ya que esta primera porción espectral es generada como salida por el decodificador de audio de ancho de banda completo o tasa completa 900 a través de la línea 904.

[0104] A continuación, se expone un ejemplo concreto con números reales. La potencia de conservación restante calculada por el bloque 912 es, por ejemplo, cinco unidades de potencia y esta potencia es la potencia de las cuatro líneas espectrales de la primera porción espectral 921 indicadas ejemplarmente.

[0105] Además, el valor de potencia E_3 para la banda de reconstrucción correspondiente a la banda de factor de escala 6 de la fig. 3b o la fig. 3a es igual a 10 unidades. Un aspecto importante consiste en que el valor de potencia no solo comprende la potencia de las porciones espectrales 922, 923, sino la potencia completa de la banda de reconstrucción 920 calculada en el lado del codificador, es decir, antes de llevar a cabo el análisis espectral mediante, por ejemplo, el enmascaramiento de tonalidad. Por lo tanto, las diez unidades de potencia abarcan la primera y la segunda porción espectral de la banda de reconstrucción. Entonces, se supone que la potencia de los datos del rango de origen para los bloques 922, 923 o para los datos de destino en bruto para el bloque 922, 923 es igual a ocho unidades de potencia. De este modo, se calcula una potencia faltante de cinco unidades.

[0106] Basándose en la potencia faltante dividida por la potencia de recuadro tE_k , se calcula un factor de ganancia de 0,79. Después, se multiplican las líneas espectrales en bruto para las segundas porciones espectrales 922, 923 por el factor de ganancia calculado. De este modo, solo se ajustan los valores espectrales para las segundas porciones espectrales 922, 923, y las líneas espectrales para la primera porción espectral 921 no se ven afectadas por este ajuste de la envolvente. Tras multiplicar los valores espectrales en bruto para las segundas porciones espectrales 922, 923, se ha calculado una banda de reconstrucción completa consistente en las primeras porciones espectrales de la banda de reconstrucción, y consistente en líneas espectrales en las segundas porciones espectrales 922, 923 de la banda de reconstrucción 920.

[0107] Preferentemente, el rango de origen para generar los datos espectrales en bruto en las bandas 922, 923 está, con respecto a la frecuencia, por debajo de la frecuencia de inicio de IGF 309 y la banda de reconstrucción 920 está por encima de la frecuencia de inicio de IGF 309.

[0108] Además, es preferible que los límites de la banda de reconstrucción coincidan con los límites de la banda de factor de escala. De este modo, una banda de reconstrucción tiene, en una realización, el tamaño de las correspondientes bandas de factor de escala del decodificador de audio central o están dimensionadas de manera que, cuando se aplica un emparejamiento de potencia, un valor de potencia para una banda de reconstrucción proporciona la potencia de dos bandas de factor de escala o un número entero superior. De este modo, cuando se supone que se lleva a cabo la acumulación de potencia para la banda de factor de escala 4, la banda de factor de escala 5 y la banda de factor de escala 6, el límite de frecuencia inferior de la banda de reconstrucción 920 es igual al límite inferior de la banda de factor de escala 4 y el límite de frecuencia superior de la banda de reconstrucción 920 coincide con el límite superior de la banda de factor de escala 6.

[0109] A continuación, se explica la fig. 9d para mostrar otras funcionalidades del decodificador de la fig. 9a. El decodificador de audio 900 recibe los valores espectrales decuantificados correspondientes a las primeras porciones espectrales del primer conjunto de porciones espectrales y, además, en un bloque de ajuste de escala inverso 940, se proporcionan factores de ajuste de escala para bandas de factor de escala como las ilustradas en la fig. 3b. El bloque de ajuste de escala inverso 940 proporciona todos los primeros conjuntos de primeras porciones espectrales por debajo de la frecuencia de inicio de IGF 309 de la fig. 3a y, además, las primeras porciones espectrales por encima de la frecuencia de inicio de IGF, es decir, las primeras porciones espectrales 304, 305, 306, 307 de la fig. 3a que están todas situadas en una banda de reconstrucción como la que se ilustra en 941 de la fig. 9d. Además, las primeras porciones espectrales de la banda de origen utilizadas para el relleno de recuadros de frecuencia en la banda de reconstrucción se proporcionan al regulador de envolvente/calculadora 942 y este bloque recibe además la información de la potencia para la banda de reconstrucción proporcionada como información secundaria paramétrica a la señal de audio codificada, tal como se ilustra en 943 de la fig. 9d. El regulador de envolvente/calculadora 942 proporciona entonces las funcionalidades de las figs. 9b y 9c y finalmente genera como salida valores espectrales ajustados para las segundas porciones espectrales de la banda de reconstrucción. Estos valores espectrales ajustados 922, 923 para las segundas porciones espectrales de la banda de reconstrucción y las primeras porciones espectrales 921 de la banda de reconstrucción indicadas en la línea 941 de la fig. 9d representan conjuntamente la representación espectral completa de la banda de reconstrucción.

[0110] A continuación, se hace referencia a las figs. 10a a 10b para explicar realizaciones preferidas de un codificador de audio para codificar una señal de audio para proporcionar o generar una señal de audio codificada. El codificador comprende un convertidor de tiempo/espectro 1002 que alimenta un analizador espectral 1004, y el analizador espectral 1044 está conectado a una calculadora de parámetros 1006, por una parte, y a un codificador de audio 1008, por la otra. El codificador de audio 1008 proporciona la representación codificada de un primer conjunto de primeras porciones espectrales y no abarca el segundo conjunto de segundas porciones espectrales. Por otra parte, la calculadora de parámetros 1006 proporciona información de potencia para una banda de reconstrucción que abarca las primeras y las segundas porciones espectrales. Además, el codificador de audio 1008 está configurado para generar una primera representación codificada del primer conjunto de primeras porciones espectrales con la primera resolución espectral, en la que el codificador de audio 1008 proporciona factores de ajuste de escala para todas las bandas de la representación espectral generada por el bloque 1002. Además, como se ilustra en la fig. 3b, el codificador proporciona información de potencia al menos para las bandas de reconstrucción situadas, con respecto a la frecuencia, por encima de la frecuencia de inicio de IGF 309, tal como se ilustra en la fig. 3a. De este modo, para las bandas de reconstrucción que coinciden preferentemente con bandas de factor de escala o con grupos de bandas de factor de escala, se dan dos valores, es decir, el correspondiente factor de ajuste de escala procedente del codificador de audio 1008 y, además, la información de potencia generada como salida por la calculadora de parámetros 1006.

[0111] El codificador de audio presenta preferentemente bandas de factor de escala con diferentes anchos de banda de frecuencia, es decir, con un número diferente de valores espectrales. Por lo tanto, la calculadora paramétrica comprende un normalizador 1012 para normalizar las potencias para el diferente ancho de banda con respecto al ancho de banda de la banda de reconstrucción concreta. Con este fin, el normalizador 1012 recibe, como entradas, una potencia en la banda y varios valores espectrales en la banda y el normalizador 1012 genera, entonces, como salida una potencia normalizada por cada banda de reconstrucción/factor de escala.

[0112] Además, la calculadora paramétrica 1006a de la fig. 10 comprende una calculadora de valores de potencia que recibe información de control procedente del codificador central o de audio 1008, tal como se ilustra mediante la línea 1007 de la fig. 10a. Esta información de control puede comprender información sobre bloques largos/cortos utilizados por el codificador de audio y/o información de agrupamiento. Por tanto, mientras que la información sobre bloques largos/cortos y la información de agrupamiento sobre ventanas cortas se refieren a un agrupamiento "temporal", la información de agrupamiento también puede hacer referencia a un agrupamiento espectral, es decir, el agrupamiento de dos bandas de factor de escala en una única banda de reconstrucción. Por tanto, la calculadora de valores de potencia 1014 genera como salida un único valor de potencia para cada banda agrupada que abarca una primera y una segunda porción espectral cuando solo se han agrupado las porciones espectrales.

[0113] La fig. 10d ilustra otra realización para implementar el agrupamiento espectral. Con este fin, el bloque 1016 está configurado para calcular valores de potencia para dos bandas adyacentes. Después, en el bloque 1018, se comparan los valores de potencia y, cuando los valores de potencia no son muy diferentes o tienen una diferencia menor que la definida, por ejemplo, por un umbral, entonces se genera un valor individual (normalizado) para ambas bandas, tal como se indica en el bloque 1020. Tal como se ilustra mediante la línea 1019, el bloque 1018 se puede

saltar. Además, la generación de un valor individual para dos o más bandas llevada a cabo por el bloque 1020 se puede controlar mediante un control de tasa de bits del codificador 1024. De este modo, cuando se va a reducir una tasa de bits, el control de tasa de bits del codificador 1024 controla el bloque 1020 para generar un valor normalizado individual para dos o más bandas aunque no se hubiera permitido que la comparación del bloque 1018 agrupara los valores de información de la potencia.

[0114] En el caso de que el codificador de audio lleve a cabo el agrupamiento de dos o más ventanas cortas, este agrupamiento también se aplica para la información de la potencia. Cuando el codificador central lleva a cabo un agrupamiento de dos o más bloques cortos, entonces, para estos dos o más bloques, solo se calcula y transmite un único conjunto de factores de ajuste de escala. En el lado del decodificador, el decodificador de audio aplica después el mismo grupo de factores de ajuste de escala para ambas ventanas agrupadas.

[0115] En referencia al cálculo de la información de la potencia, los valores espectrales de la banda de reconstrucción se acumulan a lo largo de dos o más ventanas cortas. Expresado de otro modo, esto significa que los valores espectrales en una cierta banda de reconstrucción para un bloque corto y para el posterior bloque corto se acumulan conjuntamente y solo se transmite un único valor individual de información de potencia para esta banda de reconstrucción que abarca dos bloques cortos. Después, en el lado del decodificador, el ajuste de la envolvente que se explica con respecto a las figs. 9a a 9d no se lleva a cabo de manera individual para cada bloque corto, sino que se lleva a cabo de forma conjunta para este conjunto de ventanas cortas agrupadas.

[0116] Después, se vuelve a aplicar la correspondiente normalización, de manera que, aunque se haya llevado a cabo cualquier agrupamiento en frecuencia o agrupamiento en tiempo, la normalización permite fácilmente que, para el cálculo de la información del valor de potencia en el lado del decodificador, solo sea necesario conocer el valor de información de potencia, por una parte, y la cantidad de líneas espectrales en la banda de reconstrucción o en el conjunto de bandas de reconstrucción agrupadas.

[0117] Además, se hace hincapié en que una información sobre potencias espectrales, una información sobre potencias individuales o una información de potencia individual, una información sobre una potencia de conservación o una información de potencia de conservación, una información sobre una potencia de recuadro o una información de potencia de recuadro, o una información sobre potencia faltante o una información de potencia faltante puede comprender, no solo un valor de potencia, sino también un valor de amplitud (por ejemplo, absoluta), un valor de nivel o cualquier otro valor, a partir del cual se puede obtener un valor de potencia final. Por tanto, la información sobre una potencia puede comprender, por ejemplo, el propio valor de potencia y/o un valor de un nivel y/o de una amplitud y/o de una amplitud absoluta.

[0118] La fig. 12a ilustra otra implementación del aparato para decodificar. Una corriente de bits es recibida por un decodificador central 1200 que puede ser, por ejemplo, un decodificador AAC. El resultado se configura en una etapa para llevar a cabo una interconexión o formación de recuadros de ampliación de ancho de banda 1202 correspondiente al regenerador de frecuencia 604, por ejemplo. Después, se lleva a cabo un procedimiento de adaptación de interconexión/recuadro y posprocesamiento, y, cuando se ha llevado a cabo una adaptación de interconexión, se controla el regenerador de frecuencia 1202 para llevar a cabo otra regeneración de frecuencia, pero ahora con, por ejemplo, límites de frecuencia ajustados. Además, cuando se lleva a cabo un procesamiento de interconexión, por ejemplo, mediante eliminación o atenuación de líneas tonales, el resultado se envía después al bloque 1206 que lleva a cabo el modelado de envolvente de ancho de banda basado en parámetros, tal como se expone, por ejemplo, en el contexto del bloque 712 o 826. El resultado se envía después a un bloque de transformada de síntesis 1208 para llevar a cabo una transformada en el dominio de salida final que, por ejemplo, es un dominio de salida PCM, tal como se ilustra en la fig. 12a.

[0119] Las principales características de las realizaciones de la invención son las siguientes:

La realización preferida se basa en la MDCT que presenta las fallas de trino citadas anteriormente si la desafortunada elección de frecuencia de cruce y/o márgenes de interconexión “poda” zonas espectrales tonales, o hay componentes tonales que acaban situados demasiado cerca en los límites de la interconexión.

[0120] La fig. 12b muestra el modo en que la técnica propuesta de forma novedosa reduce las fallas que se encuentran en los procedimientos de BWE del estado de la técnica. En la fig. 12, panel (2), se muestra el espectro de magnitud estilizado de la salida de un procedimiento de BWE contemporáneo. En este ejemplo, la señal resulta afectada perceptualmente por el batimiento causado por dos tonos próximos, y también por la división de un tono. Cada una de ambas zonas espectrales problemáticas está marcada con un círculo.

- [0121]** Para superar estos problemas, la nueva técnica detecta en primer lugar la ubicación espectral de los componentes tonales contenidos en la señal. Después, de acuerdo con un aspecto de la invención, se intentan ajustar las frecuencias de transición entre LF y todas las interconexiones mediante desplazamientos individuales (dentro de unos límites dados), de manera que se minimiza la división o el batimiento de componentes tonales. Con este fin, la frecuencia de transición, preferentemente, debe corresponderse con un mínimo espectral local. Este paso se muestra en la fig. 12b, panel (2) y panel (3), en los que la frecuencia de transición f_{x2} se desplaza hacia frecuencias más altas, lo que da lugar a f'_{x2} .
- 10 **[0122]** De acuerdo con otro aspecto de la invención, si continúa habiendo contenidos espectrales problemáticos en zonas de transición, se elimina al menos uno de los componentes tonales mal colocados para reducir o bien la falla de batimiento en las frecuencias de transición o bien el trino. Esto se realiza a través de una extrapolación espectral o interpolación/filtrado, tal como se muestra en la figura 2, panel (3). De este modo, se elimina un componente tonal de punto de base a punto de base, es decir, desde su mínimo local izquierdo hasta su
- 15 mínimo local derecho. El espectro resultante tras la aplicación de la tecnología de la invención se muestra en la fig. 12b, panel (4).
- [0123]** Expresado de otro modo, la fig. 12b ilustra, en la esquina superior izquierda, es decir, en el panel (1), la señal original. En la esquina superior derecha, es decir, en el panel (2), se muestra una señal comparativa de ancho de banda ampliada con zonas problemáticas marcadas con elipses 1220 y 1221. En la esquina inferior izquierda, es decir, en el panel (3), se ilustran dos características de procesamiento de recuadro de interconexión o frecuencia de recuadros. La división de porciones tonales se ha abordado aumentando el límite de frecuencia f'_{x2} , de manera que deja de aparecer allí un recorte de la correspondiente porción tonal. Además, se aplican funciones de ganancia 1030 para eliminar la porción tonal 1031 y 1032 o, como otra posibilidad, se indica una interpolación
- 25 ilustrada por 1033. Por último, la esquina inferior derecha de la fig. 12b, es decir, el panel (4), ilustra la señal mejorada resultante de una combinación del ajuste de frecuencia de interconexión/recuadro, por una parte, y la eliminación o al menos la atenuación de porciones tonales problemáticas.
- [0124]** El panel (1) de la fig. 12b ilustra, tal como se explica anteriormente, el espectro original, y el espectro
- 30 original tiene un rango de frecuencia central hasta la frecuencia de cruce o relleno de espacios f_{x1} .
- [0125]** De este modo, una frecuencia f_{x1} ilustra una frecuencia límite 1250 entre el rango de origen 1252 y un rango de reconstrucción 1254 que se extiende entre la frecuencia límite 1250 y una frecuencia máxima que es más pequeña o igual que la frecuencia de Nyquist $f_{Nyquist}$. En el lado del codificador, se supone que una señal está
- 35 limitada en ancho de banda en f_{x1} o, cuando se aplica la tecnología relativa al relleno inteligente de espacios, se supone que f_{x2} corresponde a la frecuencia de inicio de relleno de espacios 309 de la fig. 3a. Dependiendo de la tecnología, el rango de reconstrucción por encima de f_{x1} estará vacío (en el caso de la implementación de las figs. 13a y 13b) o comprenderá ciertas primeras porciones espectrales que se codificarán con una alta resolución, tal como se explica en el contexto de la fig. 3a.
- 40 **[0126]** La fig. 12b, panel (2), ilustra una señal regenerada preliminar, por ejemplo, generada por el bloque 702 de la fig. 7a, que presenta dos porciones problemáticas. Una porción problemática se ilustra en 1220. La distancia de frecuencia entre la porción tonal incluida en la zona central ilustrada en 1220a y la porción tonal situada en el comienzo del recuadro de frecuencia ilustrado en 1220b es demasiado pequeña como para que se cree una
- 45 falla de batimiento. El otro problema consiste en que, en el límite superior del primer recuadro de frecuencia generado por la primera operación de interconexión u operación de formación de recuadros de frecuencia ilustrada en 1225 es una porción tonal con un recorte intermedio o dividida 1226. Cuando esta porción tonal 1226 se compara con las otras porciones tonales de la fig. 12b, se pone de manifiesto que el ancho es más pequeño que el ancho de una típica porción tonal y esto significa que esta porción tonal ha quedado dividida al fijar el límite de frecuencia
- 50 entre el primer recuadro de frecuencia 1225 y el segundo recuadro de frecuencia 1227 en el lugar equivocado del rango de origen 1252. Para abordar este problema, se ha modificado la frecuencia de límite f_{x2} para hacerla un poco mayor que la ilustrada en el panel (3) de la fig. 12b, de manera que no se produzca un recorte de esta porción tonal.
- [0127]** Por otra parte, este procedimiento, en el que se ha cambiado f'_{x2} no aborda de manera eficaz el problema del batimiento que, por tanto, se aborda mediante la eliminación de los componentes tonales mediante filtrado o interpolación o cualquier otro procedimiento, tal como se explica en el contexto del bloque 708 de la fig. 7a. Así, la fig. 12b ilustra una aplicación secuencial del ajuste de la frecuencia de transición 706 y la eliminación de componentes tonales en los límites se ilustra en 708.

[0128] Otra opción habría consistido en fijar el límite de transición f_{x1} , para que fuera un poco más bajo, de manera que la porción tonal 1220a dejara de estar situada en el rango central. Después, también se ha retirado o eliminado la porción tonal 1220a fijando la frecuencia de transición f_{x1} en un valor más bajo.

5 **[0129]** Este procedimiento también habría funcionado para abordar el problema con el componente tonal problemático 1032. Al fijar un f_{x2} incluso más alto, la porción espectral en la que está situada la porción tonal 1032 se podría haber regenerado dentro de la primera operación de interconexión 1225 y, por lo tanto, no habrían aparecido dos porciones tonales adyacentes o contiguas.

10 **[0130]** Básicamente, el problema de batimiento depende de las amplitudes y la distancia en frecuencia de las porciones tonales adyacentes. El detector 704, 720 o, expresado de forma más general, el analizador 602 está configurado preferentemente de manera que también se realiza un análisis de la porción espectral inferior situada en la frecuencia que se encuentra por debajo de la frecuencia de transición, como por ejemplo f_{x1} , f_{x2} , f_{x2} con el fin de detectar un componente tonal. Además, también se analiza el rango espectral situado por encima de la frecuencia de transición con el fin de detectar un componente tonal. Cuando la detección da lugar a dos componentes tonales, uno a la izquierda de la frecuencia de transición con respecto a la frecuencia, y otro a la derecha (con respecto a la frecuencia ascendente), se activa el eliminador de componentes tonales en los límites que se ilustra en 708 de la fig. 7a. La detección de componentes tonales se lleva a cabo en un cierto rango de detección que se extiende en ambas direcciones, desde la frecuencia de transición, al menos un 20% con respecto al ancho de banda de la banda correspondiente y, preferentemente, solo se extiende hasta un 10% hacia abajo, a la izquierda de la frecuencia de transición, y hacia arriba, a la derecha de la frecuencia de transición en relación con el ancho de banda correspondiente, es decir, el ancho de banda del rango de origen, por una parte, y el rango de reconstrucción, por otra, o, cuando la frecuencia de transición es la frecuencia de transición entre dos recuadros de frecuencia 1225, 1227, una cantidad correspondiente de un 10% del recuadro de frecuencia correspondiente. En otra realización, el ancho de banda de detección predeterminado es un bark. Deberían poderse eliminar porciones tonales dentro de un rango de un bark alrededor de un límite interconexión, de manera que el rango completo de detección sea 2 bark, es decir, un bark en la banda inferior y un bark en la banda superior, en las que el bark de la banda inferior es inmediatamente adyacente a un bark de la banda superior.

30 **[0131]** De acuerdo con otro aspecto de la invención, para reducir la falla de zumbido de filtrado, se aplica un filtro de cruce en el dominio de frecuencia a dos zonas espectrales consecutivas, es decir, entre la banda central y la primera interconexión o entre dos interconexiones. Preferentemente, el filtro de cruce es adaptable a la señal.

[0132] El filtro de cruce consiste en dos filtros, un filtro de disminución progresiva de volumen h_{out} , que se aplica a la zona espectral inferior, y un filtro de aumento progresivo de volumen h_{in} , que se aplica a la zona espectral superior.

[0133] Cada uno de los filtros tiene una longitud N.

40 **[0134]** Además, la pendiente de ambos filtros se caracteriza por un valor adaptable a la señal denominado $Xbias$ que determina la característica de muesca del filtro de cruce, con $0 \leq Xbias \leq N$:

Si $Xbias = 0$, entonces la suma de ambos filtros es igual a 1, es decir, no hay característica de filtro de muesca en el filtro resultante.

Si $Xbias = N$, entonces ambos filtros son completamente cero.

45

[0135] El diseño básico de los filtros de cruce está limitado por las siguientes ecuaciones:

$$h_{out}(k) = h_{in}(N-1-k), \forall Xbias$$

50

$$h_{out}(k) + h_{in}(k) = 1, Xbias = 0$$

siendo $k = 0, 1, \dots, N-1$ el índice de frecuencia. La fig. 12c muestra un ejemplo de dicho filtro de cruce.

[0136] En este ejemplo, se utiliza la siguiente ecuación para crear el filtro h_{out} :

55

$$h_{out}(k) = 0,5 + 0,5 \cdot \cos\left(\frac{k}{N-1-Xbias} \cdot \pi\right), k = 0, 1, \dots, N-1-Xbias$$

[0137] La siguiente ecuación describe el modo en que se aplican después los filtros h_{in} y h_{out} ,

$$Y(k_t - (N - 1) + k) = LF(k_t - (N - 1) + k)h_{out}(k) + HF(k_t - (N - 1) + k)h_{in}(k), k = 0, 1, \dots, N - 1$$

5

en la que Y denota el espectro ensamblado, k_t es la frecuencia de transición, LF es el contenido de baja frecuencia y HF es el contenido de alta frecuencia.

[0138] A continuación, se presentará la prueba de la ventaja que ofrece esta técnica. La señal original en los siguientes ejemplos es una señal de tipo transitorio, en particular una versión de la misma filtrada con un filtro de paso bajo, con una frecuencia de corte de 22 kHz. En primer lugar, este transitorio está limitado en banda a 6 kHz en el dominio de la transformada. A continuación, el ancho de banda de la señal original filtrada con un filtro de paso bajo se amplía a 24 kHz. La ampliación de ancho de banda se logra copiando la banda LF tres veces para rellenar completamente el rango de frecuencias que se halla disponible por encima de 6 kHz dentro de la transformada.

15

[0139] La fig. 11a muestra el espectro de esta señal, que se puede considerar como un espectro típico de una falla de zumbido del filtro que rodea espectralmente el transitorio debido a dicha característica de "pared de ladrillo" o *brick-wall* de la transformada (picos de habla 1100). Aplicando el enfoque de la invención, se reduce el zumbido del filtro en aproximadamente 20 dB en cada frecuencia de transición (picos de habla reducidos).

20

[0140] En las figs. 11b y 11c, se muestra el mismo efecto, aunque en una ilustración diferente. La fig. 11b muestra el espectrograma de la citada señal de tipo transitorio con la falla de zumbido del filtro que precede y sucede temporalmente al transitorio tras aplicar la técnica BWE descrita anteriormente sin ninguna reducción del zumbido del filtro. Cada una de las líneas horizontales representa el zumbido del filtro en la frecuencia de transición entre interconexiones consecutivas. La fig. 6 muestra la misma señal tras aplicar el enfoque de la invención dentro de la BWE. Mediante la aplicación de la reducción del zumbido, se reduce el zumbido del filtro en aproximadamente 20 dB en comparación con la señal que se muestra en la figura anterior.

25

[0141] A continuación, se explican las figs. 14a y 14b con el fin de ilustrar aún más el aspecto del filtro de cruce de la invención que ya se ha explicado en el contexto de la característica del analizador. No obstante, el filtro de cruce 710 también se puede implementar independientemente de la invención que se explica en el contexto de las figs. 6a-7b.

30

[0142] La fig. 14a ilustra un aparato para decodificar una señal de audio codificada que comprende una señal central codificada e información sobre datos paramétricos. El aparato comprende un decodificador central 1400 para decodificar la señal central codificada a fin de obtener una señal central decodificada. La señal central decodificada puede estar limitada en el ancho de banda en el contexto de la implementación de la fig. 13a y la fig. 13b, o el decodificador central puede ser un codificador de rango de frecuencia completo o de tasa completa en el contexto de las figs. 1 a 5c o 9a-10d.

35

[0143] Además, hay un generador de recuadros 1404 para regenerar uno o más recuadros espectrales con frecuencias no incluidas en la señal central decodificada que se generan a partir de una porción espectral de la señal central decodificada. Los recuadros pueden ser unas segundas porciones espectrales reconstruidas dentro de una banda de reconstrucción como se ilustra, por ejemplo, en el contexto de la fig. 3a, o que puede incluir unas primeras porciones espectrales que se reconstruirán con una alta resolución, pero, como otra posibilidad, los recuadros espectrales también pueden comprender bandas de frecuencia completamente vacías cuando el codificador ha llevado a cabo una limitación estricta de banda, tal como se ilustra en la fig. 13a.

40

[0144] Además, se proporciona un filtro de cruce 1406 para llevar a cabo un filtrado de cruce espectral con la señal central decodificada y un primer recuadro de frecuencia con frecuencias que se extienden desde una frecuencia de relleno de espacios 309 hasta una primera frecuencia de fin de recuadros o para el filtrado de cruce espectral de un primer recuadro de frecuencia 1225 y un segundo recuadro de frecuencia 1221, en el que el segundo recuadro de frecuencia presenta una frecuencia de límite inferior que es adyacente en frecuencia a una frecuencia de límite superior del primer recuadro de frecuencia 1225.

45

[0145] En otra implementación, la señal de salida del filtro de cruce 1406 se introduce en un regulador de envolvente 1408 que aplica la información de la envolvente espectral paramétrica incluida en una señal de audio codificada como información secundaria paramétrica para obtener finalmente una señal regenerada y con envolvente ajustada. Los elementos 1404, 1406 y 1408 se pueden implementar como un regenerador de frecuencia,

50

como se ilustra, por ejemplo, en la fig. 13b, la fig. 1b o la fig. 6a.

[0146] La fig. 14b ilustra otra implementación del filtro de cruce 1406. El filtro de cruce 1406 comprende un subfiltro de disminución progresiva de volumen que recibe una primera señal de entrada IN1, y un segundo subfiltro de aumento progresivo de volumen 1422 que recibe una segunda entrada IN2 y los resultados o salidas de ambos filtros 1420 y 1422 se ponen a disposición de un combinador 1424 que, por ejemplo, es un sumador. El sumador o combinador 1424 genera como salida los valores espectrales para los bins de frecuencia. La fig. 12c ilustra un ejemplo de función de atenuación cruzada que comprende la característica de subfiltro de disminución progresiva de volumen 1420a y la característica de subfiltro de aumento progresivo de volumen 1422a. Ambos filtros presentan un cierto solapamiento en el ejemplo de la fig. 12c igual a 21, es decir $N = 21$. De este modo, otros valores de frecuencia o, por ejemplo, la zona de origen 1252 no se ven afectados. Solo los bins de frecuencia más alta 21 del rango de origen 1252 se ven afectados por la función de disminución progresiva de volumen 1420a.

[0147] Por otra parte, solo las líneas de frecuencia más baja 21 del primer recuadro de frecuencia 1225 se ven afectadas por la función de aumento progresivo de volumen 1420a.

[0148] Además, se pone de manifiesto por las funciones de atenuación cruzada que las líneas de frecuencia entre 9 y 13 se ven afectadas, pero la función de aumento progresivo de volumen en realidad no afecta a las líneas de frecuencia entre 1 y 9 y la función de disminución progresiva de volumen 1420a no afecta a las líneas de frecuencia entre 13 y 21. Esto significa que solo sería necesario un solapamiento entre las líneas de frecuencia 9 y 13, y la frecuencia de cruce, como por ejemplo f_{x1} , se colocaría en la muestra de frecuencia o el bin de frecuencia 11. De este modo, solo se requeriría un solapamiento de dos bins de frecuencia o valores de frecuencia entre el rango de origen y el primer recuadro de frecuencia para implementar la función de cruce o de atenuación cruzada.

[0149] Dependiendo de la implementación concreta, se puede aplicar un solapamiento mayor o menor y, además, se pueden utilizar otras funciones de atenuación aparte de la función coseno. Además, como se ilustra en la fig. 12c, es preferible aplicar una cierta muesca en el rango de cruce. Expresado de un modo diferente, la potencia en los rangos de límite se reducirá debido al hecho de que ambas funciones de filtro no suman una unidad, como sería el caso en una función de atenuación sin muesca. Esta pérdida de potencia para los límites del recuadro de frecuencia, es decir, el primer recuadro de frecuencia, se atenuará en el límite inferior y el límite superior, y las potencias se concentrarán más en el medio de las bandas. No obstante, debido al hecho de que el ajuste de la envolvente espectral tiene lugar posteriormente al procesamiento realizado por el filtro de cruce, la frecuencia total no se toca, pero se define por los datos de la envolvente espectral, tales como los factores de ajuste de escala correspondientes, tal como se explica en el contexto de la fig. 3a. Expresado de otro modo, la calculadora 918 de la fig. 9b calcularía entonces el "rango de destino en bruto ya generado", que consiste en la salida del filtro de cruce. Además, también se compensaría la pérdida de potencia debida a la eliminación de una porción tonal mediante interpolación, debido al hecho de que esta eliminación da lugar entonces a una potencia de recuadro más baja y el factor de ganancia para la banda de reconstrucción completa será más alto. No obstante, por otra parte, la frecuencia de cruce da lugar a una concentración de potencia más hacia el medio de un recuadro de frecuencia y esto, al final, reduce de manera eficaz las fallas, causadas especialmente por los transitorios, como se explica en el contexto de las figs. 11a-11c.

[0150] La fig. 14b ilustra diferentes combinaciones de entradas. Para un filtrado en el límite entre el rango de frecuencia de origen y el recuadro de frecuencia, la entrada 1 es la porción espectral superior del rango central y la entrada 2 es la porción espectral inferior del primer recuadro de frecuencia o del recuadro de frecuencia individual, cuando solo existe un recuadro de frecuencia individual. Además, la entrada puede ser el primer recuadro de frecuencia y la frecuencia de transición puede ser el límite de frecuencia superior del primer recuadro y la entrada en el subfiltro 1422 será la porción inferior del segundo recuadro de frecuencia. Cuando también existe un tercer recuadro de frecuencia, otra frecuencia de transición constituirá el límite de frecuencia entre el segundo recuadro de frecuencia y el tercer recuadro de frecuencia y la entrada en el subfiltro de disminución progresiva de volumen 1421 constituirá el rango espectral superior del segundo recuadro de frecuencia determinado por el parámetro de filtrado, cuando se utiliza la característica de la fig. 12c, y la entrada en el subfiltro de aumento progresivo de volumen 1422 será la porción inferior del tercer recuadro de frecuencia y, en el ejemplo de la fig. 12c, las líneas espectrales más bajas 21.

[0151] Como se ilustra en la fig. 12c, es preferible que el parámetro N sea igual para el subfiltro de disminución progresiva de volumen, sin embargo, no es necesario. Los valores para N pueden variar y, entonces, el resultado será que la "muesca" del filtro será asimétrica entre el rango inferior y el superior. Además, las funciones de disminución progresiva de volumen/aumento progresivo de volumen no tienen que estar necesariamente en la

misma característica que en la fig. 12c. En lugar de ello, también se pueden utilizar características asimétricas.

[0152] Además, es preferible hacer que la característica del filtro de cruce sea adaptable a la señal. Por lo tanto, basándose en un análisis de la señal, se adapta la característica del filtro. Debido al hecho de que el filtro de cruce resulta especialmente útil para las señales transitorias, se detecta si aparecen señales transitorias. Cuando aparecen señales transitorias, se puede utilizar una característica de filtro como la que se ilustra en la fig. 12c. Sin embargo, cuando se detecta una señal no transitoria, es preferible cambiar la característica del filtro para reducir la influencia del filtro de cruce. Esto se podría obtener, por ejemplo, fijando N en cero o fijando X_{bias} en cero, de manera que la suma de ambos filtros sea igual a 1, es decir, que en el filtro resultante no hay ninguna característica de filtro de muesca. Otra posibilidad consiste, simplemente, en saltarse el filtro de cruce 1406 en el caso de señales no transitorias. No obstante, preferentemente, se prefiere una característica de filtro de cambio relativamente lento mediante el cambio de los parámetros N , X_{bias} con el fin de evitar fallas creadas por el cambio rápido de las características del filtro. Además, es preferible un filtro de paso bajo para permitir únicamente dichos cambios relativamente pequeños en las características del filtro aunque la señal esté cambiando más rápidamente de lo que detecta cierto detector de transitorios/tonalidad. El detector se ilustra en 1405 de la fig. 14a. Puede recibir una señal de entrada en un generador de recuadros o una señal de salida del generador de recuadros 1404 o incluso se puede conectar al decodificador central 1400 para obtener una información de transitorio/no transitorio como, por ejemplo, una indicación de bloque corto procedente, por ejemplo, de la decodificación AAC. Como es natural, también se puede utilizar cualquier otro filtro de cruce diferente al que se muestra en la fig. 12c.

[0153] Después, basándose en la detección de transitorios, o basándose en una detección de tonalidad, o basándose en cualquier otra detección de características de la señal, se cambia la característica del filtro de cruce 1406 tal como se explica.

[0154] Aunque algunos aspectos se han descrito en el contexto de un aparato para codificar o decodificar, es evidente que estos aspectos también representan una descripción del procedimiento correspondiente, en el que un bloque o dispositivo corresponde a un paso del procedimiento o a una característica de un paso del procedimiento. De forma análoga, los aspectos descritos en el contexto de un paso del procedimiento también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente. Algunos o todos los pasos se pueden ejecutar mediante (o empleando) un aparato de *hardware* como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, se pueden ejecutar uno o más de los pasos más importantes del procedimiento mediante dicho aparato.

[0155] Dependiendo de ciertos requisitos de implementación, las realizaciones de la presente invención se pueden implementar en *hardware* o en *software*. La implementación se puede llevar a cabo empleando un medio de almacenamiento no transitorio, como, por ejemplo, un medio de almacenamiento digital, por ejemplo un disco flexible, una unidad de disco duro (HDD), un DVD, un BluRay, un CD, una ROM, una PROM, y una memoria EPROM, EEPROM o *flash*, con unas señales de control legibles electrónicamente almacenadas en las mismas, que cooperan (o son capaces de cooperar) con un sistema informático programable, de manera que se lleve cabo el respectivo procedimiento. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

[0156] Algunas realizaciones de acuerdo con la invención comprenden un soporte de datos provisto de señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable, de manera que se lleva a cabo uno de los procedimientos descritos en la presente memoria.

[0157] En general, las realizaciones de la presente invención se pueden implementar como producto de programa informático con un código de programa, siendo el código de programa capaz de llevar a cabo uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código del programa puede almacenarse, por ejemplo, en un soporte legible por máquina.

[0158] Otras realizaciones comprenden el programa informático para llevar a cabo uno de los procedimientos descritos en la presente memoria, almacenado en un soporte legible por máquina.

[0159] Expresado de otro modo, una realización del procedimiento de la invención consiste, por tanto, en un programa informático con un código de programa para llevar a cabo uno de los procedimientos descritos en la presente memoria, cuando el programa informático se ejecuta en un ordenador.

[0160] Por tanto, otra realización del procedimiento de la invención consiste en un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el

programa informático para llevar a cabo uno de los procedimientos descritos en la presente memoria. Normalmente, el soporte de datos, el medio de almacenamiento digital o el medio grabado son tangibles y no transitorios.

5 **[0161]** Otra realización de la invención consiste, por tanto, en una corriente de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los procedimientos descritos en la presente memoria. La corriente de datos o la secuencia de señales puede, por ejemplo, estar configurada para ser transferida a través de una conexión de comunicación de datos, por ejemplo, a través de Internet.

10 **[0162]** Otra realización comprende unos medios de procesamiento, por ejemplo, un ordenador o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los procedimientos descritos en la presente memoria.

15 **[0163]** Otra realización comprende un ordenador en el que está instalado el programa informático para llevar a cabo uno de los procedimientos descritos en la presente memoria.

20 **[0164]** Otra realización de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, de forma electrónica u óptica) a un receptor un programa informático para llevar a cabo uno de los procedimientos descritos en la presente memoria. El receptor puede ser, por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o sistema puede comprender, por ejemplo, un servidor de archivos para transferir el programa informático al receptor.

25 **[0165]** En algunas realizaciones, se puede utilizar un dispositivo lógico programable (por ejemplo, una matriz de puertas programables *in situ*) para llevar a cabo algunas de las funcionalidades, o todas, de los procedimientos descritos en la presente memoria. En algunas realizaciones, una matriz de puertas programables *in situ* puede cooperar con un microprocesador con el fin de llevar a cabo uno de los procedimientos descritos en la presente memoria. Por lo general, los procedimientos se llevan a cabo preferentemente mediante cualquier aparato de *hardware*.

30 **[0166]** Las realizaciones descritas anteriormente tienen un carácter meramente ilustrativo de los principios de la presente invención. Se entiende que para los expertos en la materia resultarán evidentes modificaciones y variaciones de las disposiciones y los detalles descritos en la presente memoria. Por lo tanto, se pretende que solo quede limitada por el alcance de las siguientes reivindicaciones de patente y no por los detalles concretos presentados a modo de descripción y explicación de las realizaciones de la presente memoria.

REIVINDICACIONES

1. Aparato para decodificar una señal de audio codificada que comprende una señal central codificada (1), que comprende:
- 5 un decodificador central (1400) configurado para decodificar la señal central codificada (1401) para obtener una señal central decodificada;
 0 de un rango de cruce que se extiende a lo largo de al menos tres valores de frecuencia o para llevar a cabo una adición ponderada por frecuencia (1424) de al menos una parte de un primer recuadro de frecuencia filtrado
 10 mediante el subfiltro de disminución progresiva de volumen (1420) y al menos una parte de un segundo recuadro de frecuencia filtrada mediante el subfiltro de aumento progresivo de volumen (1422) dentro de un rango de cruce que se extiende a lo largo de al menos tres valores de frecuencia.
2. Aparato de la reivindicación 1,
- 15 en el que una porción espectral de la señal central decodificada, una porción espectral del primer recuadro de frecuencia o una porción espectral del segundo recuadro de frecuencia afectada por el filtro de cruce (1406) tienen un tamaño inferior al 30% de la porción espectral abarcada por una banda espectral total de la banda de frecuencia central decodificada o una banda espectral total del primer o el segundo recuadro de frecuencia y es mayor o igual
 20 que una banda definida por al menos 5 valores de frecuencia adyacentes.
3. Aparato de las reivindicaciones 1 y 2,
- en el que el filtro de cruce (1406) está configurado para aplicar una característica de filtro de tipo coseno para la
 25 disminución progresiva de volumen y el aumento progresivo de volumen.
4. Aparato de acuerdo con una de las reivindicaciones anteriores que comprende un regulador de envolvente (1408) configurado para ajustar la envolvente de una señal espectral filtrada con un filtro de cruce en un rango espectral definido por unos rangos espectrales de los uno o más recuadros espectrales a partir de una
 30 información de envolvente espectral paramétrica (1407) incluida en la señal de audio codificada.
5. Aparato de acuerdo con una de las reivindicaciones anteriores,
- que además comprende un convertidor de frecuencia-tiempo (828) configurado para convertir una señal con
 35 envolvente ajustada junto con la señal central decodificada en una representación temporal.
6. Aparato de acuerdo con la reivindicación 5, en el que el convertidor de frecuencia-tiempo está configurado para aplicar una transformada de coseno discreta modificada (512, 514, 516) que comprende un procesamiento de solapamiento/adición (516) de un cuadro actual con un cuadro de un tiempo anterior.
 40
7. Aparato de acuerdo con una de las reivindicaciones anteriores, en el que el filtro de cruce es un filtro controlable,
- 45 en el que el aparato comprende además un detector de características de señales (1405), y
 en el que el detector de características de señales (1405) está configurado para controlar una característica de filtro del filtro de cruce (1406) de acuerdo con un resultado de detección obtenido a partir de la señal central decodificada.
8. Aparato de la reivindicación 7,
- 50 en el que el detector de características de la señal (1405) es un detector de transitorios, y en el que el detector de transitorios (1405) está configurado para controlar el filtro de cruce de manera que, para una porción de señal más transitoria, el filtro de cruce produce un primer efecto en una señal de entrada de filtro de cruce, y el filtro de cruce (1406) produce un segundo efecto sobre la señal de entrada del filtro de cruce para una porción de señal menos
 55 transitoria, en las que el primer efecto es mayor que el segundo efecto.
9. Aparato de acuerdo con una de las reivindicaciones anteriores,
- en el que una característica del filtro de cruce (1406) se define por una característica de subfiltro de disminución

progresiva de volumen (1420a) y una característica de subfiltro de aumento progresivo de volumen (1422a),

en el que la característica de subfiltro de aumento progresivo de volumen $h_{in}(k)$ y la característica de subfiltro de disminución progresiva de volumen $h_{out}(k)$ se definen basándose en las siguientes ecuaciones:

5

$$h_{out}(k) = h_{in}(N-1-k), \forall Xbias$$

$$h_{out}(k) + h_{in}(k) = 1, Xbias = 0$$

10

$$h_{out}(k) = 0,5 + 0,5 \cdot \cos\left(\frac{k}{N-1-Xbias} \cdot \pi\right), k = 0, 1, \dots, N-1-Xbias$$

15 en las que $Xbias$ es un número entero que define una pendiente de ambos filtros que se extiende entre cero y un número entero N , en las que k es un índice de frecuencia que se extiende entre cero y $N-1$, y en el que N es otro número entero, y en las que diferentes valores para N y $Xbias$ dan lugar a diferentes características del filtro de cruce.

10. Aparato de la reivindicación 9,

20 en el que $Xbias$ se fija entre 2 y 20 y en el que N se fija entre 10 y 50.

11. Aparato de acuerdo con una de las reivindicaciones anteriores,

25 en el que el generador de recuadros (1404) está configurado para generar un recuadro de frecuencia preliminar (703), en el que un analizador (702) está configurado para analizar el primer recuadro de frecuencia preliminar, en el que el generador de recuadros también está configurado para generar una señal regenerada con porciones tonales atenuadas o eliminadas en relación con el recuadro de frecuencia preliminar, en el que el generador de recuadros está configurado para eliminar o atenuar las porciones tonales cerca de los límites del recuadro de frecuencia (708) para obtener una señal de entrada en el filtro de cruce (1406).

30 12. Aparato de la reivindicación 11, en el que el generador de recuadros está configurado para detectar y eliminar o atenuar porciones espectrales tonales dentro de un rango de detección inferior al 20% de un ancho de banda de un recuadro de frecuencia o un rango de origen para la regeneración.

35 13. Aparato de una de las reivindicaciones anteriores, en el que el filtro de cruce (1406) está configurado para realizar un filtrado de cruce dentro de un rango de solapamiento, y el rango de solapamiento comprende una porción de frecuencia superior de la señal central decodificada y una porción de frecuencia inferior del primer recuadro de frecuencia o,

40 en el que el filtro de cruce (1406) está configurado para realizar un filtrado de cruce dentro de un rango de solapamiento y el rango de solapamiento comprende una porción de frecuencia superior de un primer recuadro de frecuencia y una porción de frecuencia inferior de un segundo recuadro de frecuencia.

45 14. Procedimiento de decodificación de una señal de audio codificada que comprende una señal central codificada (1), que comprende:

decodificación (1400) de la señal central codificada (1401) para obtener una señal central decodificada; generación (1404) de uno o más recuadros espectrales con frecuencias no incluidas en la señal central decodificada a partir de una porción espectral de la señal central decodificada; y filtrado de cruce espectral, mediante un filtro de cruce (1406), de la señal central decodificada y un primer recuadro de frecuencia con frecuencias que se extienden desde una frecuencia de relleno de espacios (309) hasta una frecuencia de límite superior o para realizar un filtrado de cruce espectral de un primer recuadro de frecuencia y un segundo recuadro de frecuencia,

50 en el que el filtro de cruce (1406) está configurado para llevar a cabo una adición ponderada por frecuencia (1424) de la señal central decodificada filtrada mediante un subfiltro de disminución progresiva de volumen (1420) y al menos una porción del primer recuadro de frecuencia filtrada por un subfiltro de aumento progresivo de volumen (1422) dentro de un rango de cruce que se extiende a lo largo de al menos tres valores de frecuencia o para llevar a

cabo una adición ponderada por frecuencia (1424) de al menos una parte de un primer recuadro de frecuencia filtrado mediante el subfiltro de disminución progresiva de volumen (1420) y al menos una parte de un segundo recuadro de frecuencia filtrada mediante el subfiltro de aumento progresivo de volumen (1422) dentro de un rango de cruce que se extiende a lo largo de al menos tres valores de frecuencia.

5

15. Programa informático para llevar a cabo, cuando se ejecuta en un ordenador o procesador, el procedimiento de la reivindicación 14.

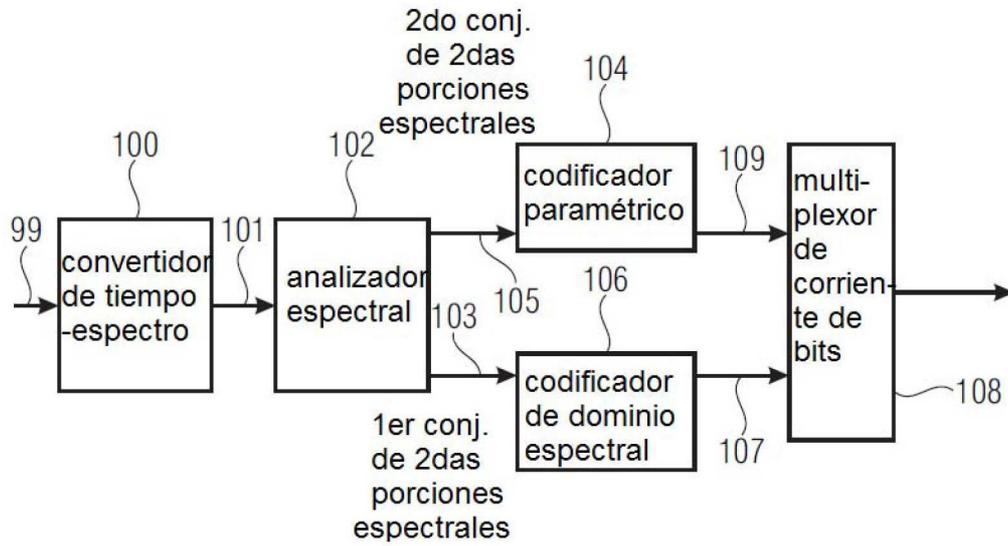


FIG 1A

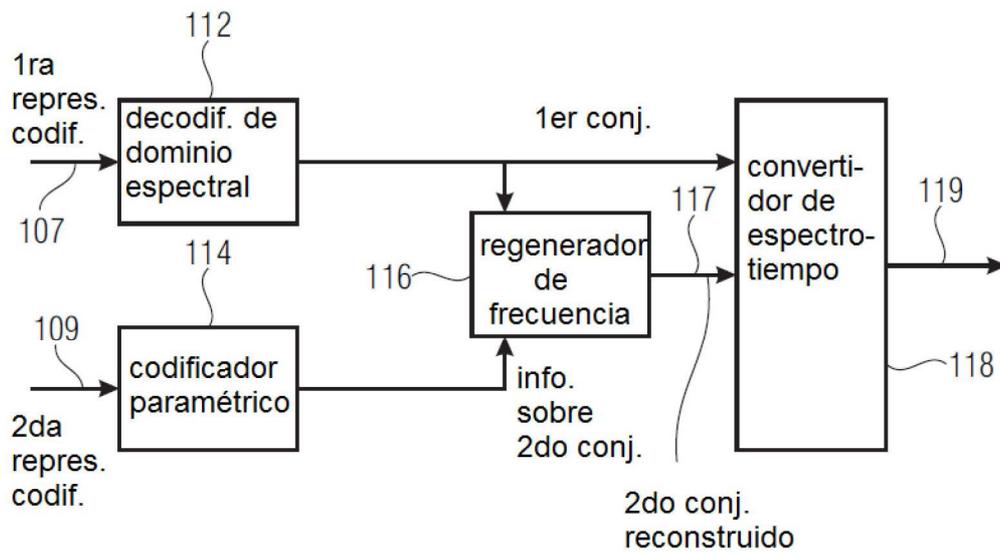


FIG 1B

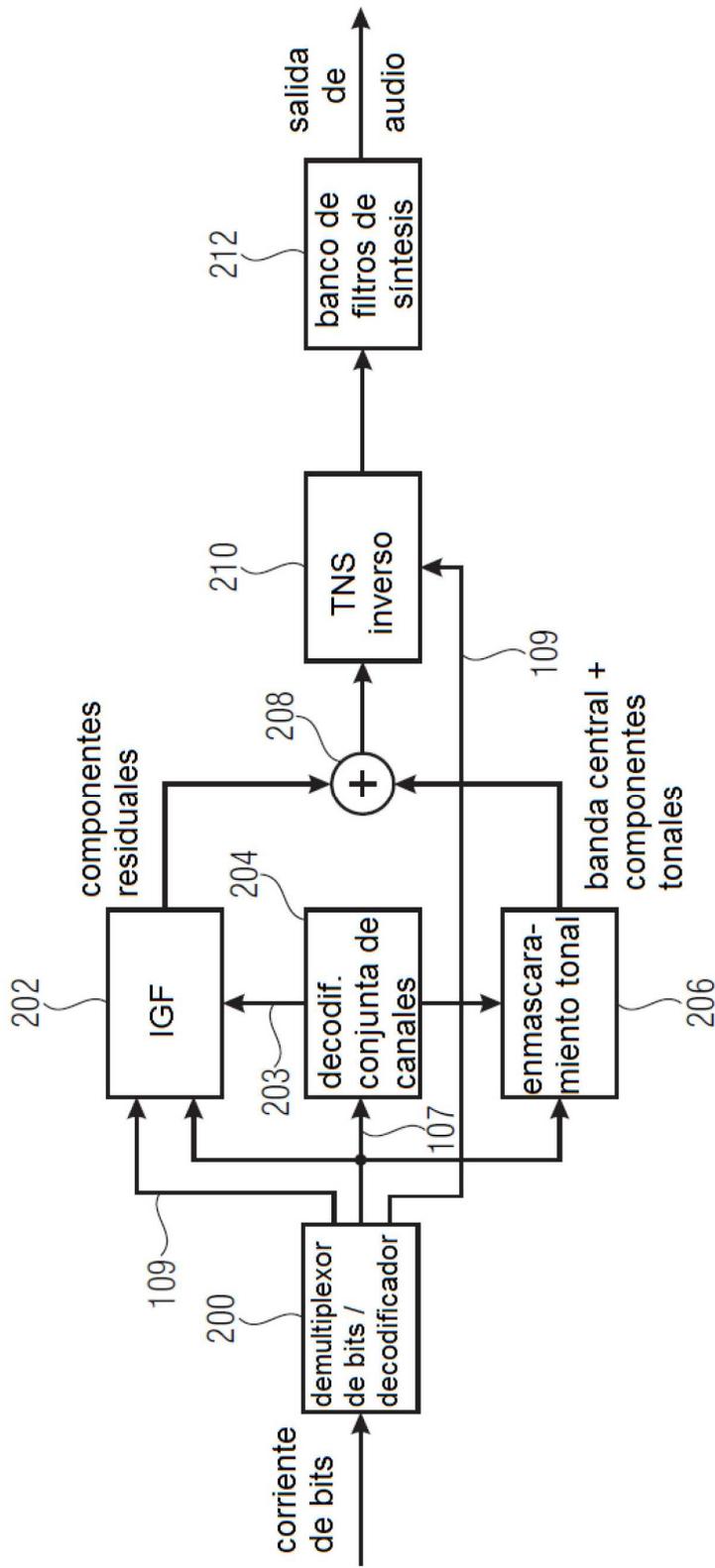


FIG 2A

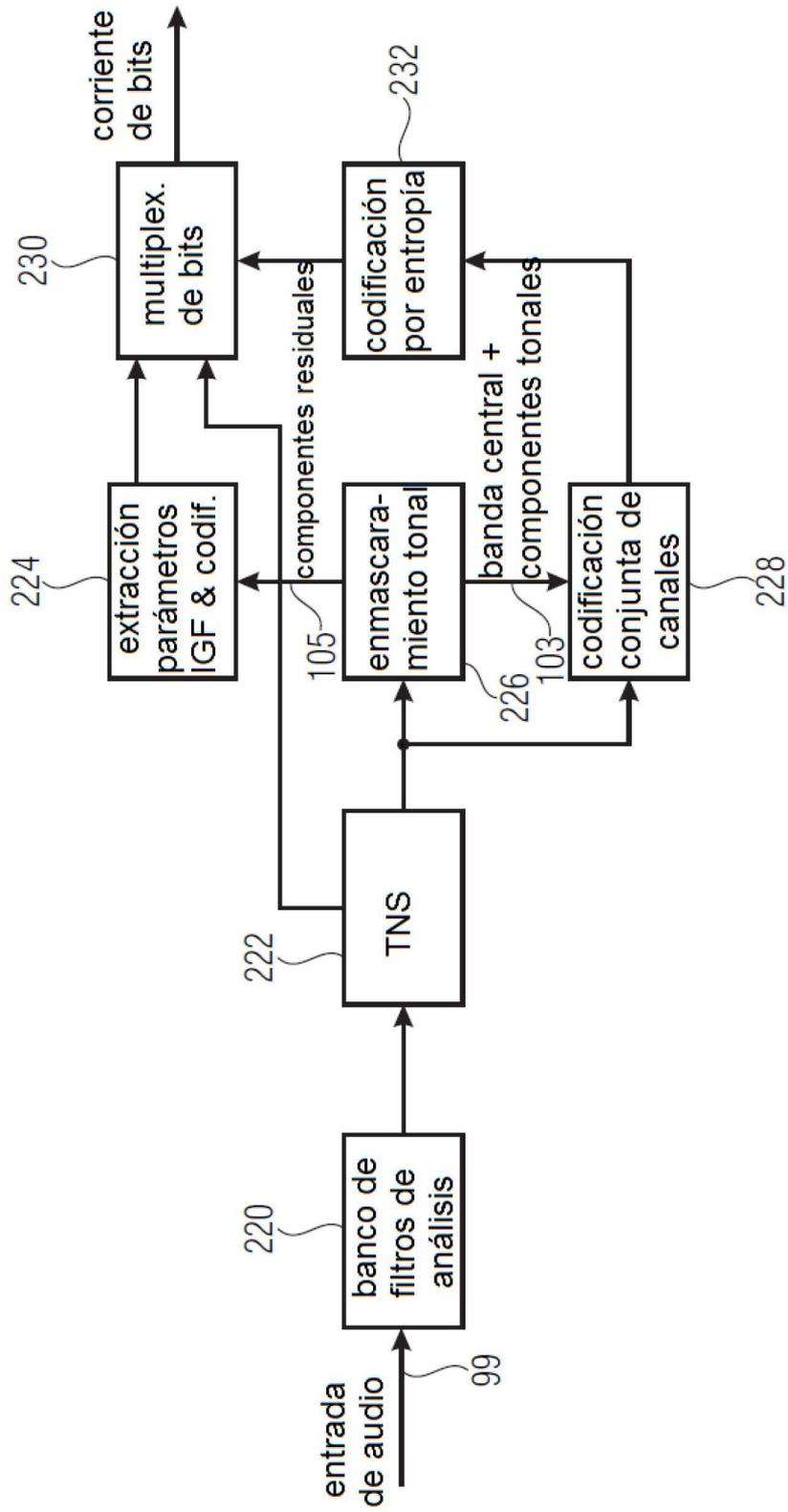


FIG 2B

- 1ra resolución (alta resolución) para la "envolvente" del 1er conjunto (codificación por líneas);
- 2da resolución (baja resolución) para la "envolvente" del 2do conjunto (factor de ajuste de escala por bandas de factor de escala (SCB));

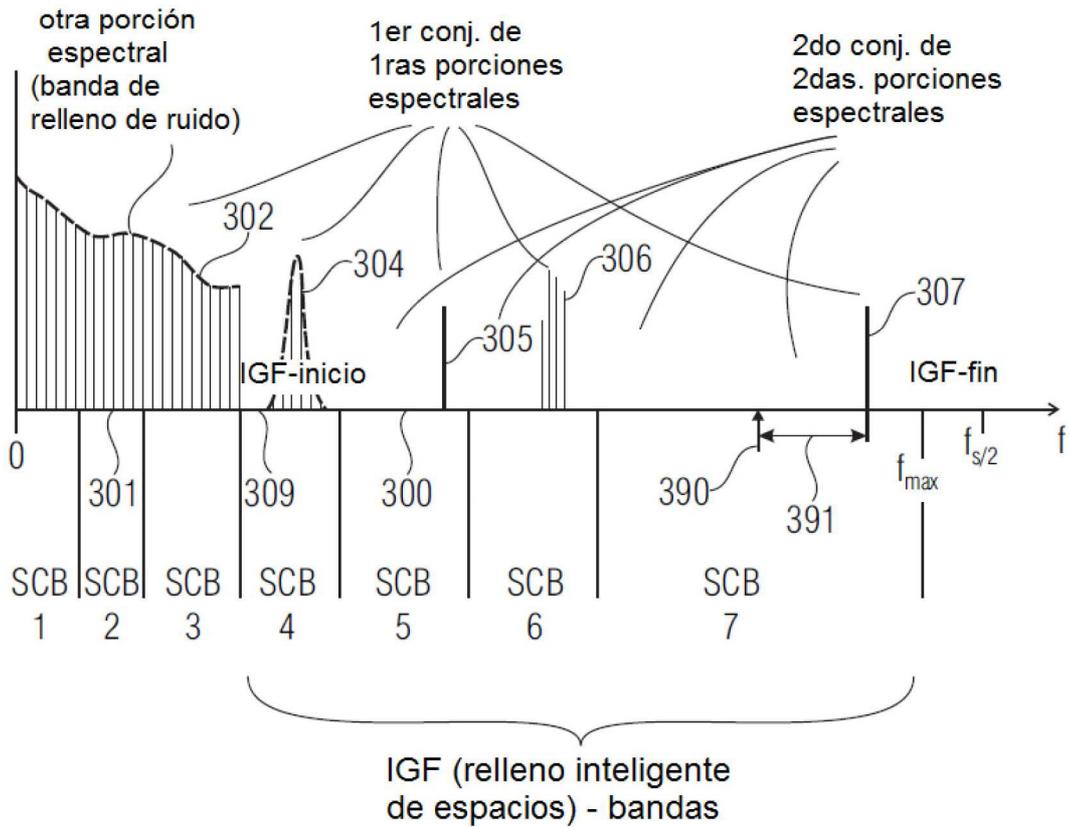


FIG 3A

SCB: Banda de Factor de Escala

SF: Factor de Ajuste de Escala

| | | | | | | |
|------|-----------------|------|----------------|----------------|----------------|----------------|
| SCB1 | SCB2 | SCB3 | SCB4 | SCB5 | SCB6 | SCB7 |
| SF1 | SF2 | SF3 | SF4 | SF5 | SF6 | SF7 |
| | | | E ₁ | E ₂ | E ₃ | E ₄ |
| | NF ₂ | | | | | |

308
310
312

FIG 3B

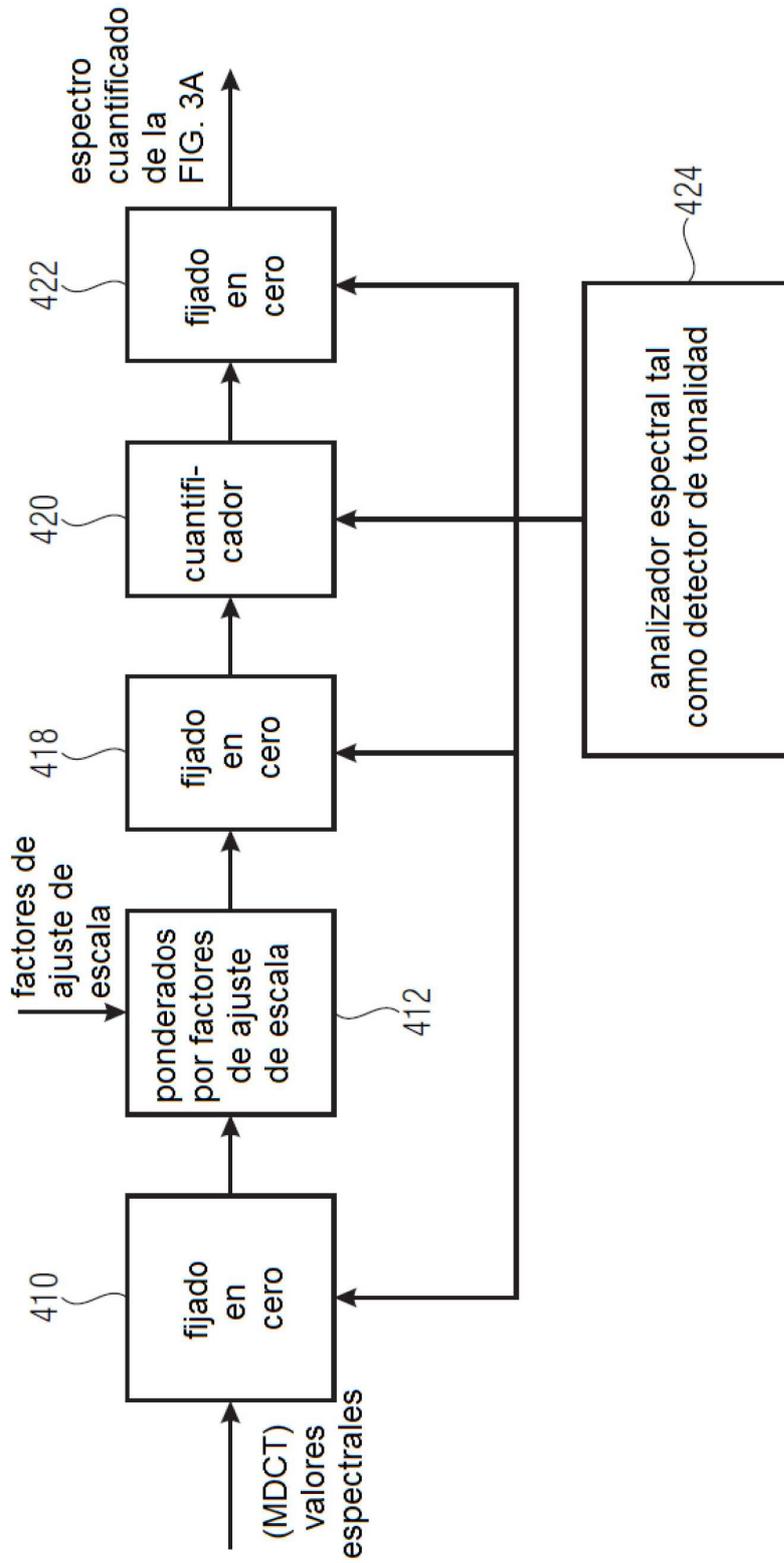


FIG 4B
(PROCESADOR CUANTIFICADOR)

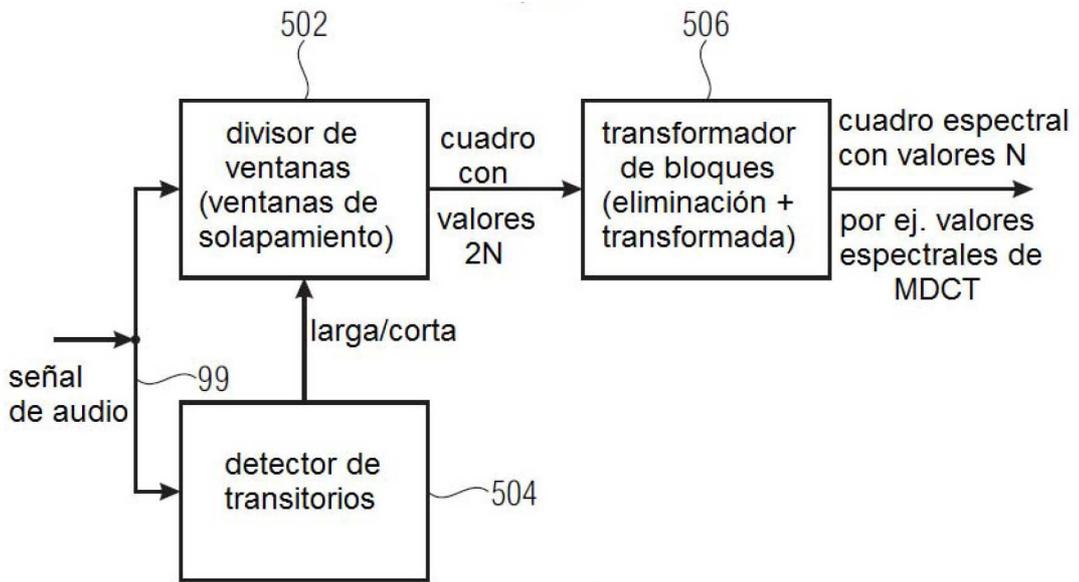


FIG 5A
(OTRAS PORCIONES ESPECTRALES)

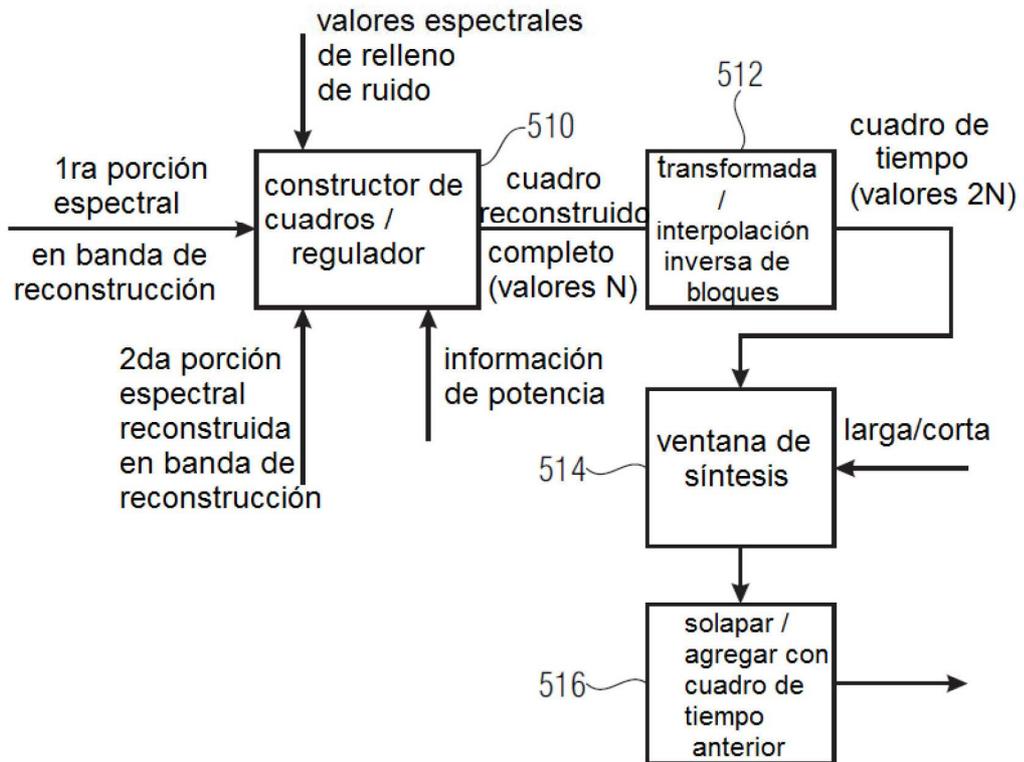


FIG 5B

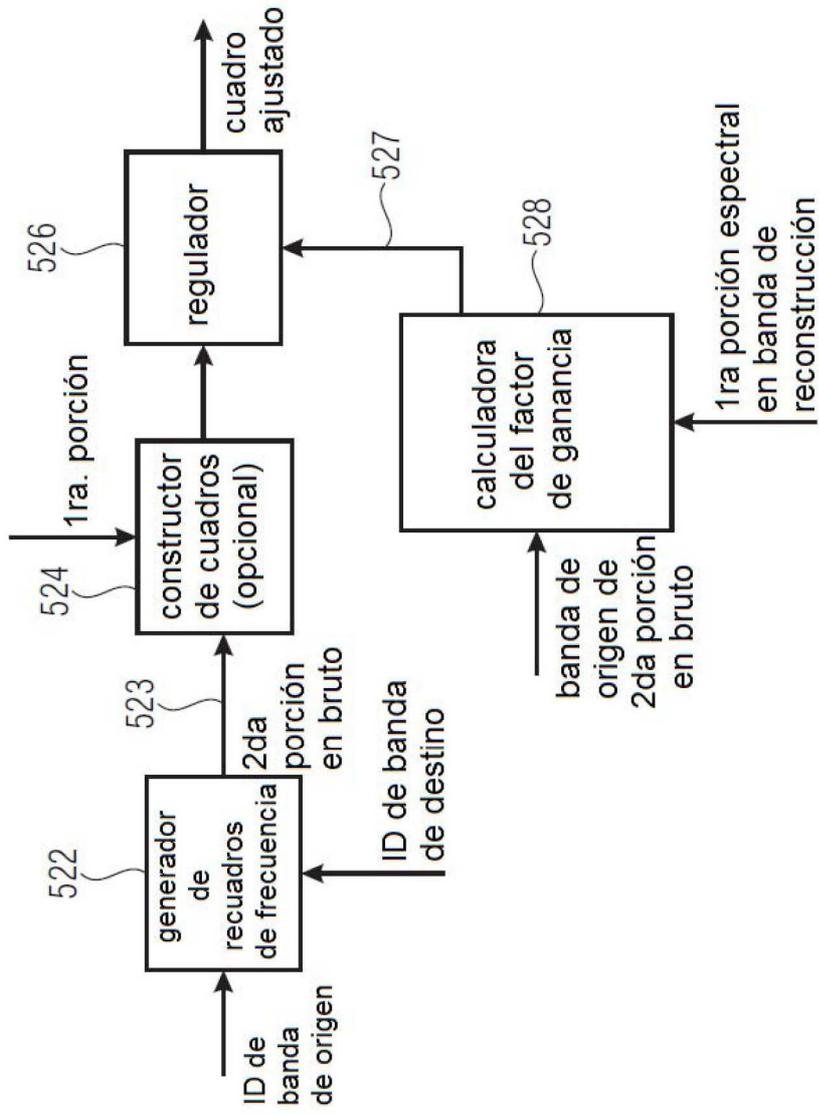


FIG 5C

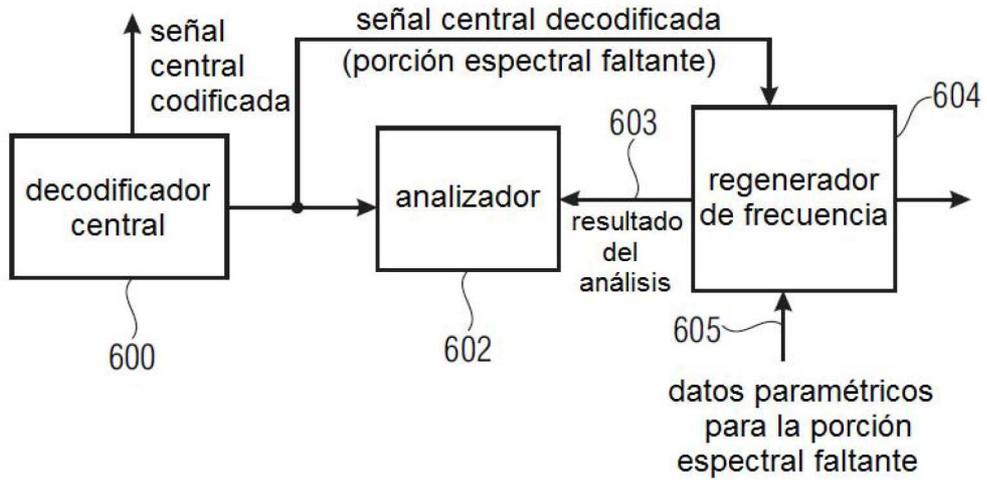


FIG 6A

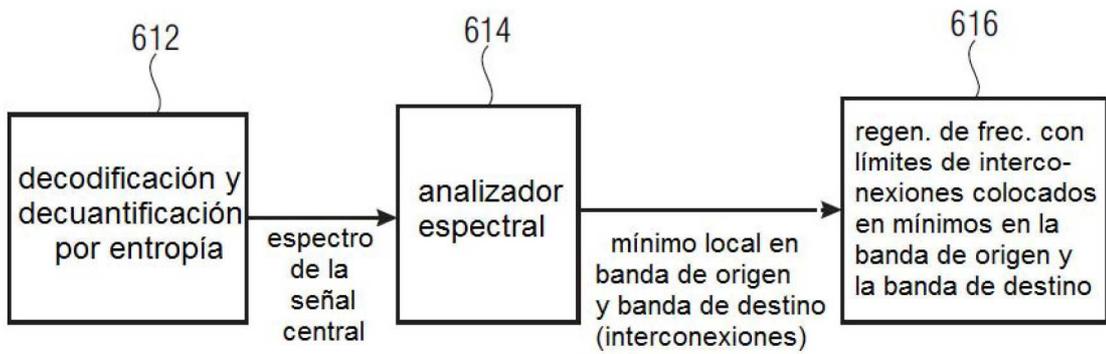


FIG 6B

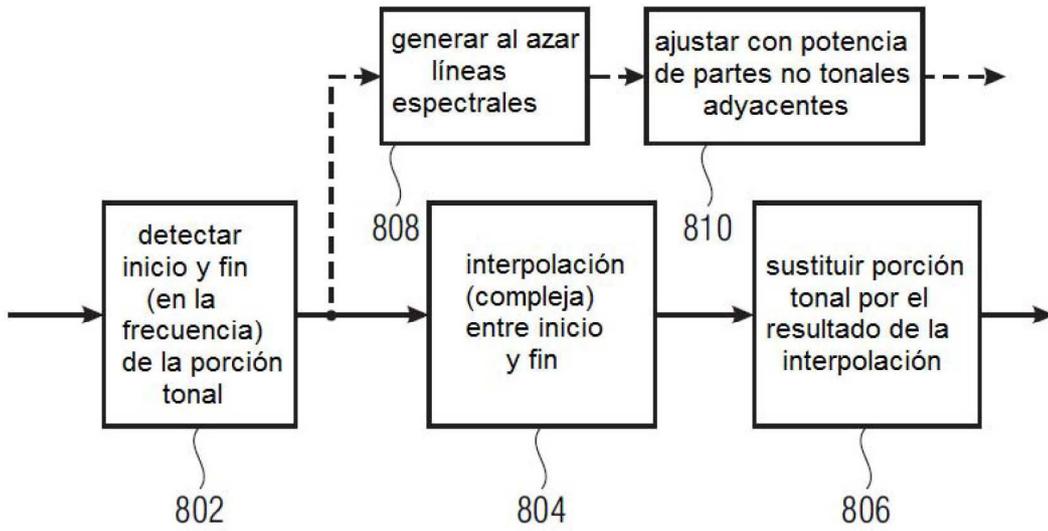


FIG 8A

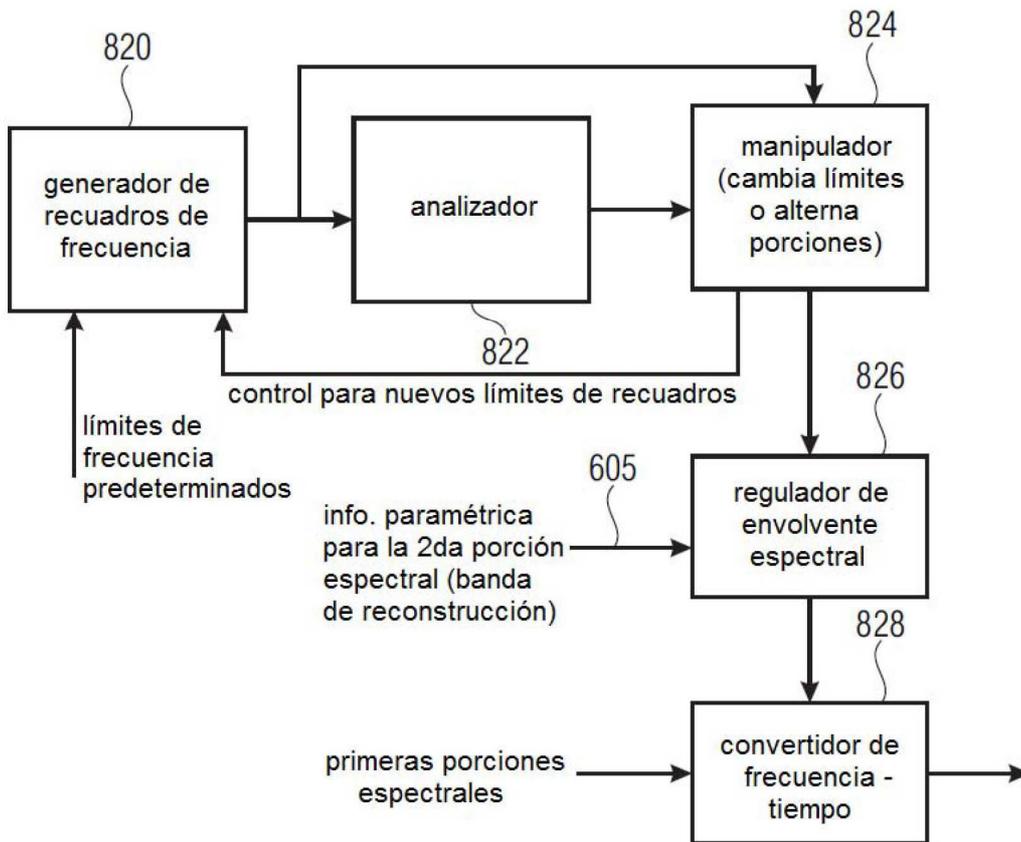


FIG 8B

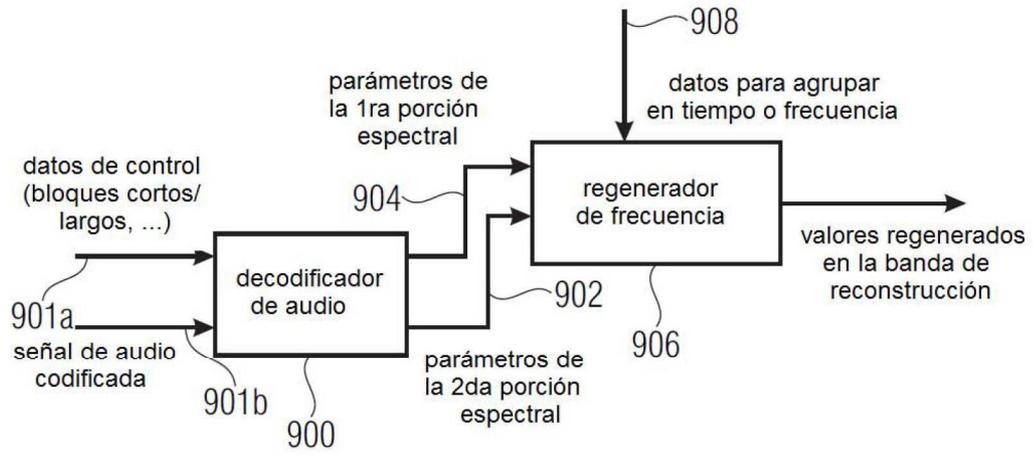


FIG 9A

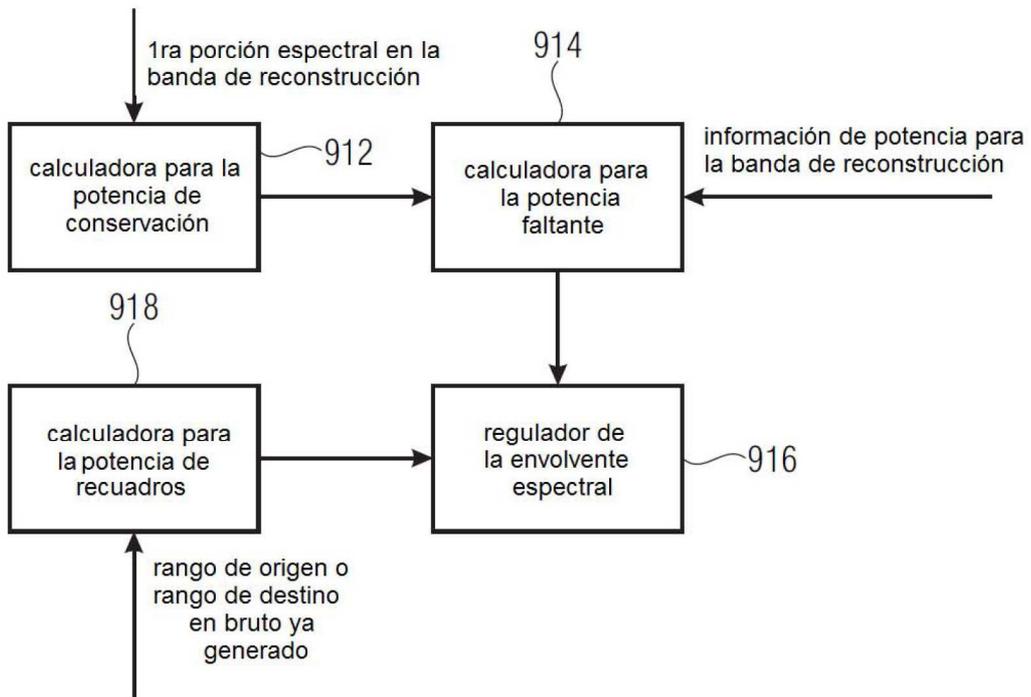
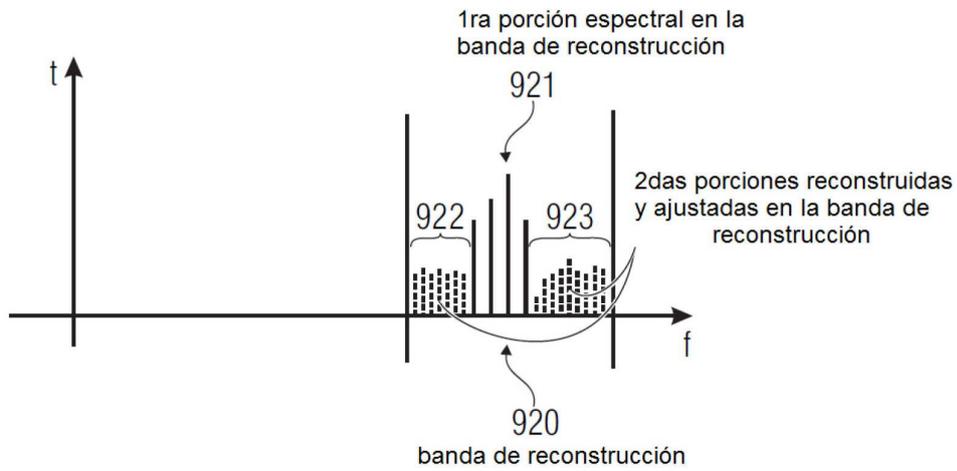


FIG 9B



por ej.

- potencia de conservación: 5 unidades
- valor de la potencia para la banda de reconstr.: 10 unidades (abarca 1ras y 2das porciones espectrales en la banda de reconstrucción)
- potencia de datos del rango de origen o datos del rango de destino en bruto: 8 unidades
- potencia faltante: 5 unidades
- factor de ganancia: $g := \sqrt{\frac{mE_k}{pE_k}} = 0.79$

→ sólo se ajustan valores espectrales para las 2das porciones espectrales

→ la 1ra porción espectral no está influenciada por el ajuste de envolvente

FIG 9C

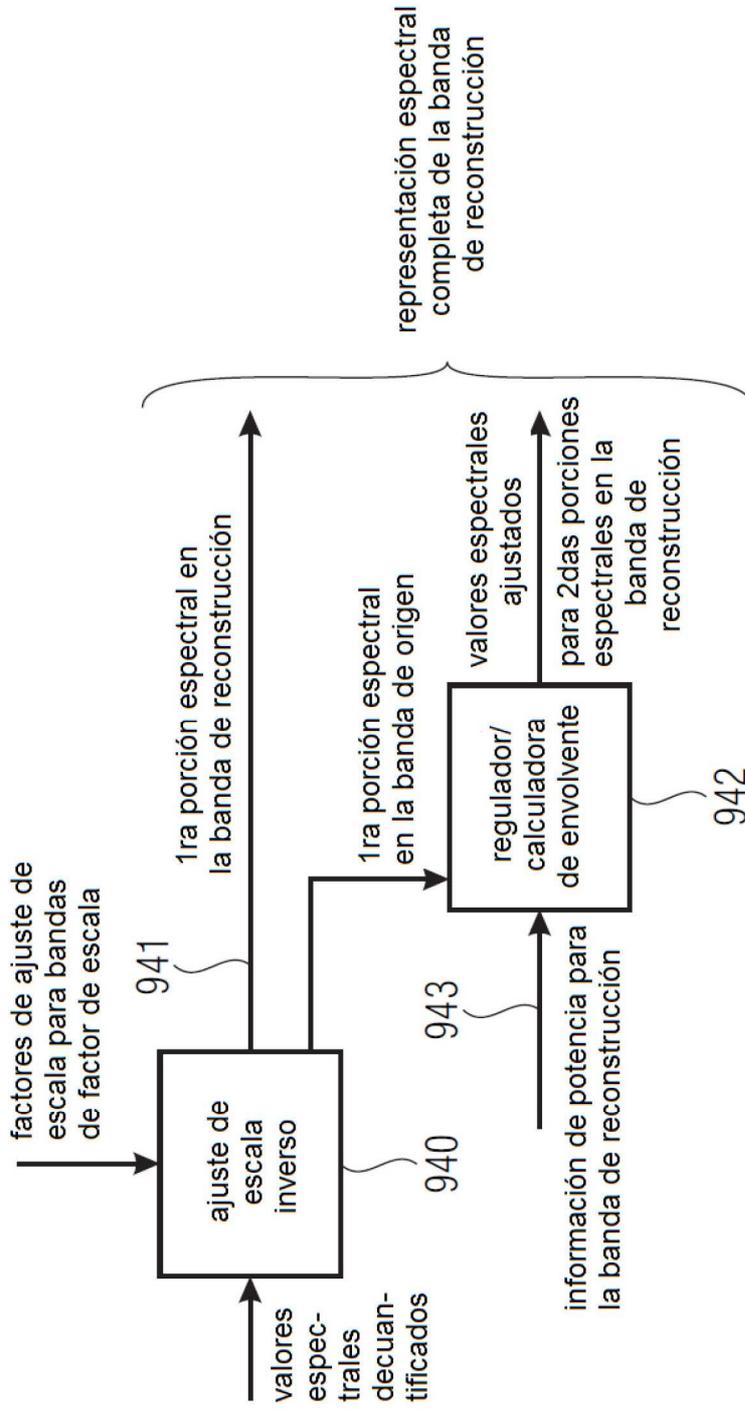


FIG 9D

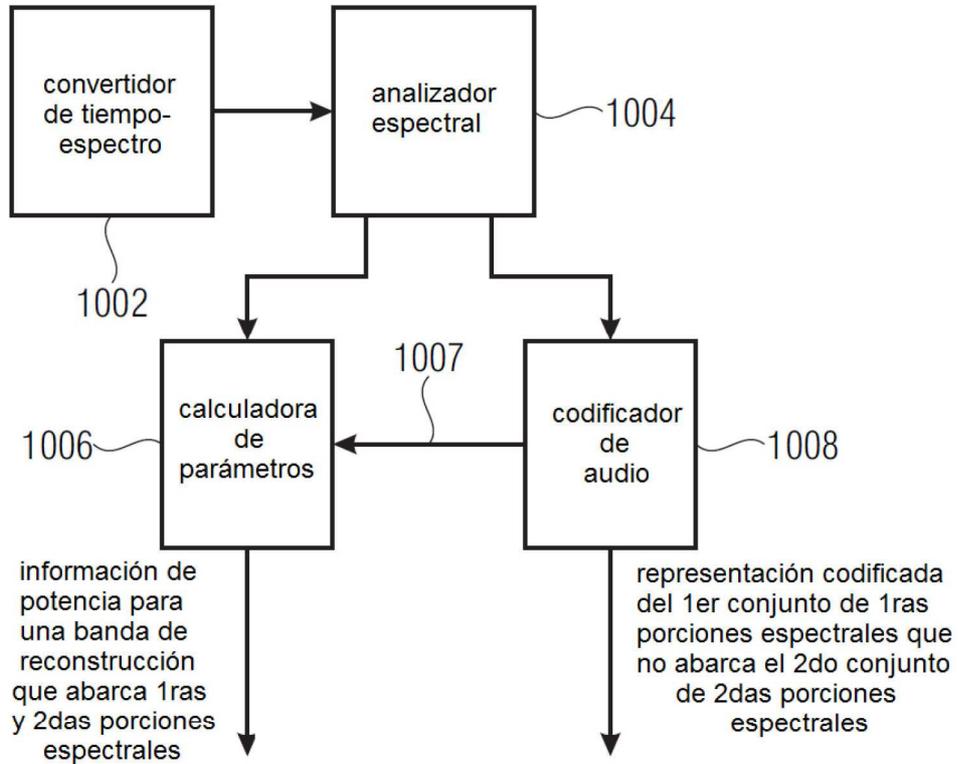


FIG 10A

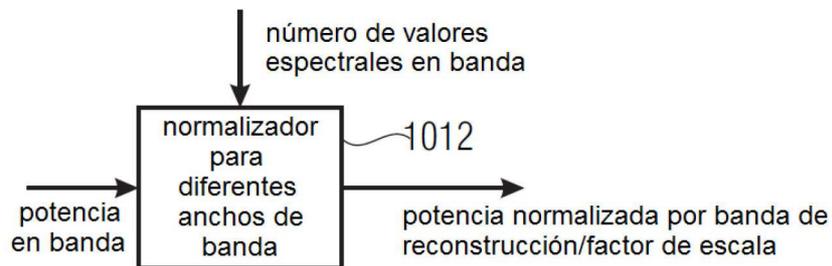


FIG 10B

info de control del codificador central

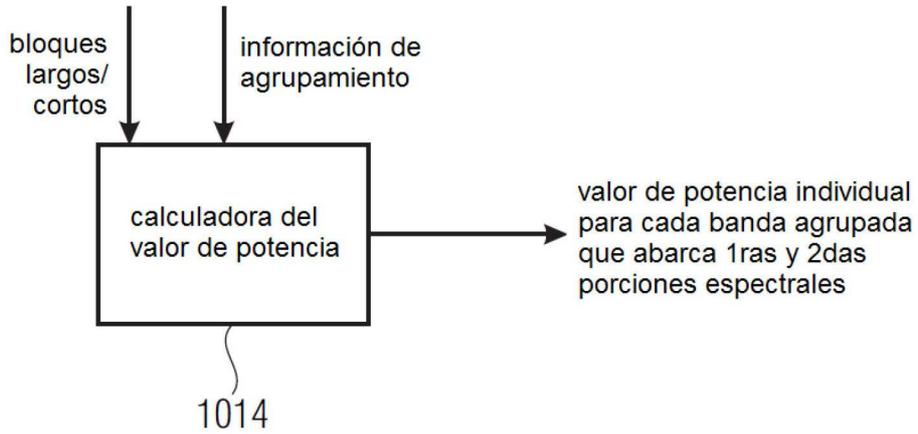


FIG 10C

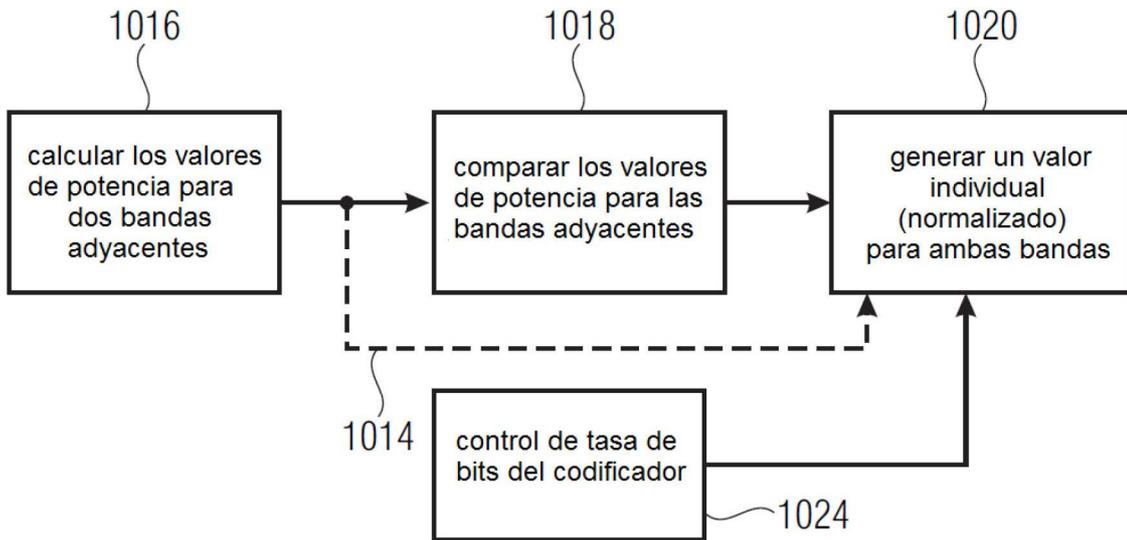


FIG 10D

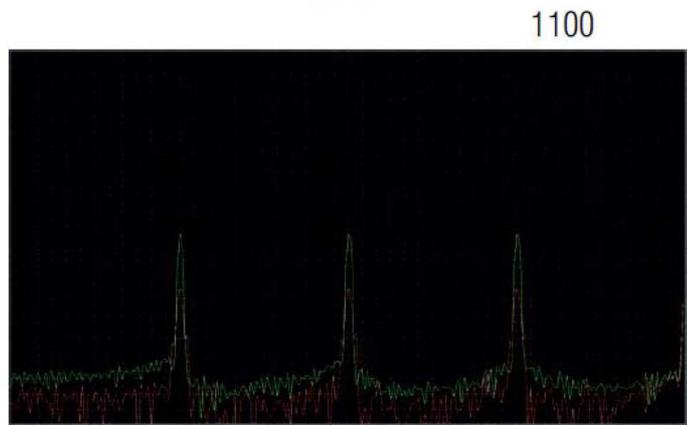
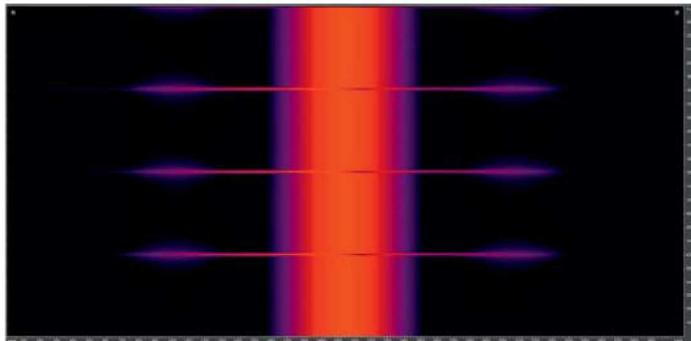
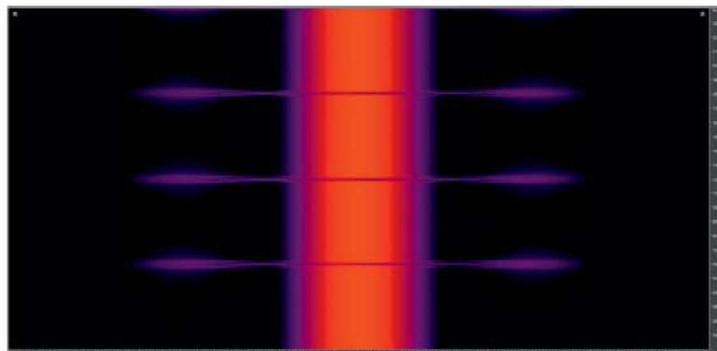


FIG 11A



Espectrograma de un transitorio luego de aplicar BWE.
El eje x representa el tiempo, el eje y representa la frecuencia.

FIG 11B



Espectrograma de un transitorio luego de aplicar BWE.
El eje x representa el tiempo, el eje y representa la frecuencia.
La aplicación de la reducción de zumbido del filtro reduce el
zumbido del filtro en aproximadamente 20dB.

FIG 11C

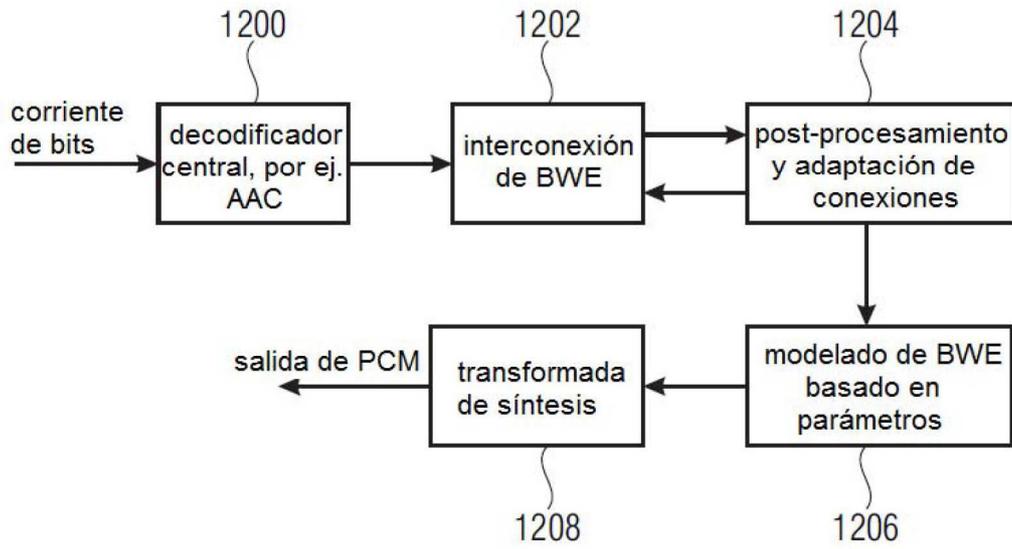
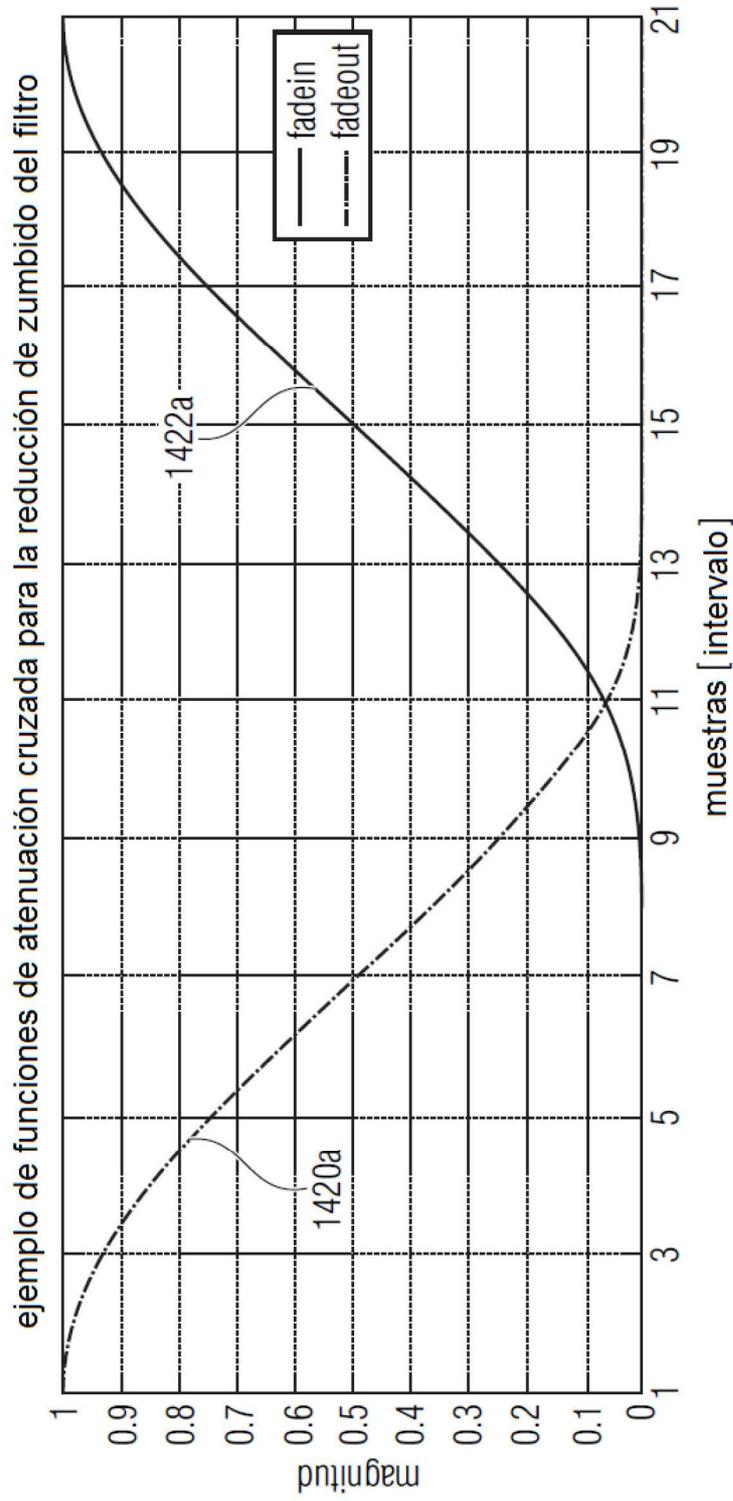


FIG 12A



ejemplo de funciones de atenuación cruzada, $N = 21$, $X_{bias} = 8$.
 fade in: aumento progresivo del volumen; fade out: disminución progresiva del volumen

FIG 12C

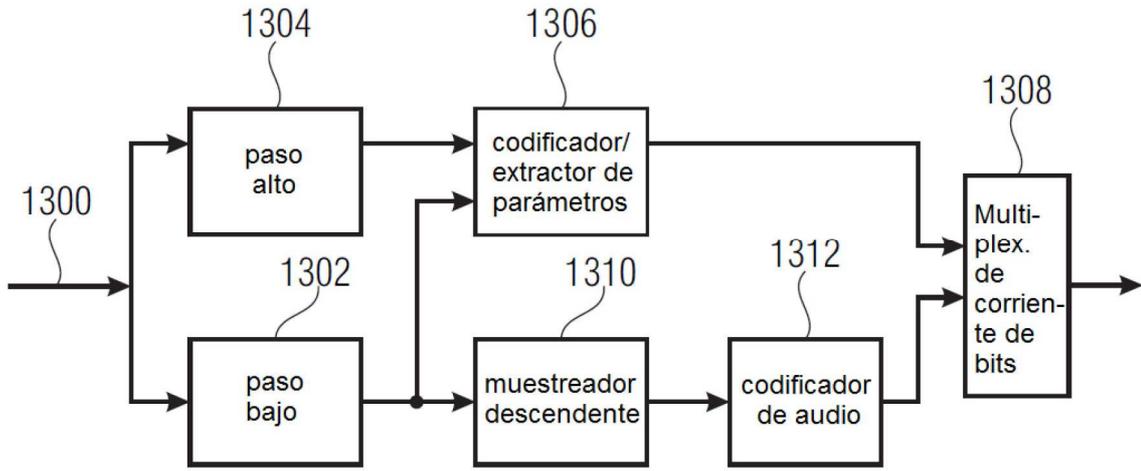


FIG 13A
(TÉCNICA ANTERIOR)

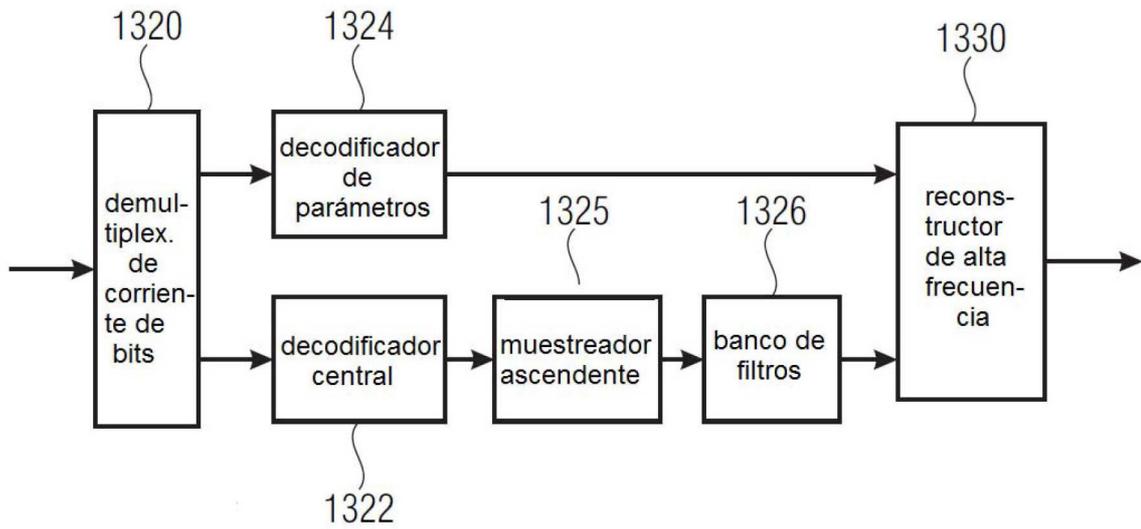


FIG 13B
(TÉCNICA ANTERIOR)

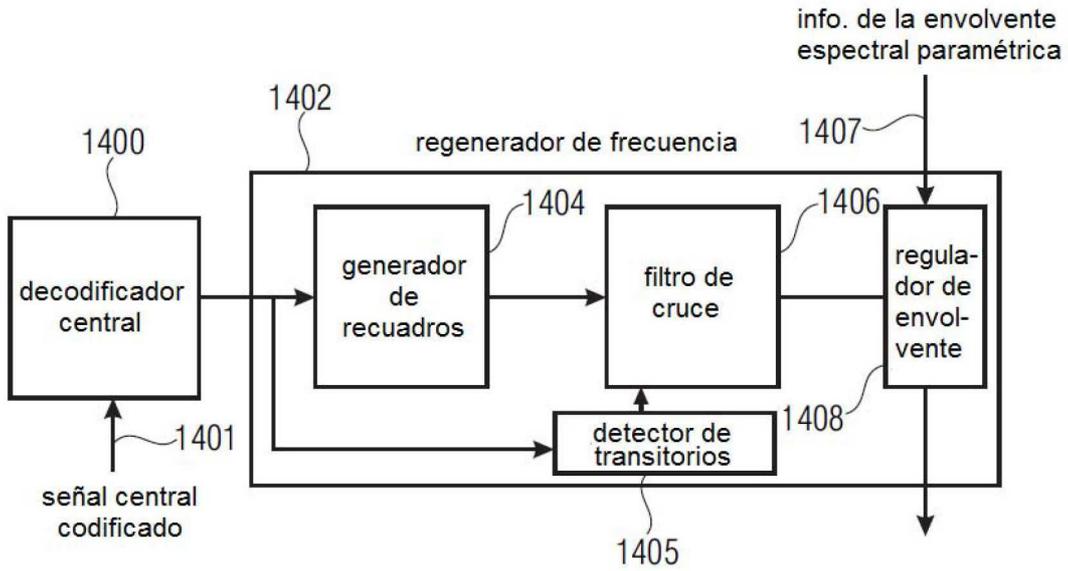


FIG 14A

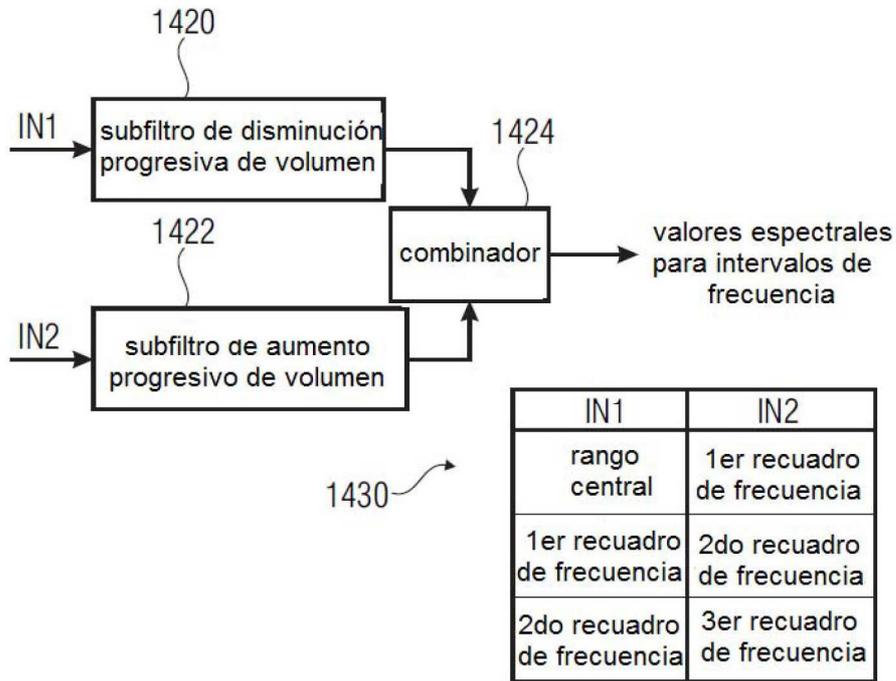


FIG 14B