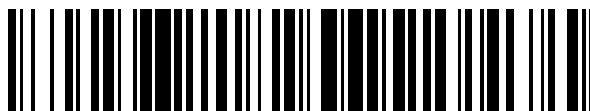


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 844 930**

51 Int. Cl.:

G10L 21/038 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **21.08.2017** **PCT/EP2017/071048**

87 Fecha y número de publicación internacional: **01.03.2018** **WO18036972**

96 Fecha de presentación y número de la solicitud europea: **21.08.2017** **E 17752404 (8)**

97 Fecha y número de publicación de la concesión europea: **16.12.2020** **EP 3504707**

54 Título: **Aparato y procedimiento para codificar una señal de audio usando un valor de compensación**

30 Prioridad:

23.08.2016 EP 16185398

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

23.07.2021

73 Titular/es:

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

72 Inventor/es:

**DISCH, SASCHA;
REUTELHUBER, FRANZ;
BÜTHE, JAN;
MULTRUS, MARKUS y
EDLER, BERND**

74 Agente/Representante:

PONTI & PARTNERS, S.L.P.

ES 2 844 930 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 codificar una señal de audio usando un valor de compensación

5 **[0001]** La presente invención está dirigida a la codificación y decodificación de audio y, especialmente, a la codificación/decodificación de audio usando tecnologías que mejoran el espectro tales como extensión de ancho de banda o SBR (por sus siglas en inglés: *spectral band replication*/replicación de banda espectral) o IGF (por sus siglas en inglés: *intelligent gap filling*/cobertura inteligente de intervalos).

10 **[0002]** El almacenamiento o la transmisión de señales de audio a menudo están sujetos a estrictas restricciones en cuanto a tasa de bits. En el pasado, los codificadores se veían drásticamente forzados a reducir el ancho de banda de audio transmitido cuando solo estaba disponible una tasa de bits muy baja. Hoy en día, los modernos códecs son capaces de codificar señales de banda ancha usando procedimientos BWE (por sus siglas en inglés: *bandwidth extension*/extensión ancho de banda) [1-2]. Estos algoritmos están basados en una representación paramétrica del
15 contenido de alta frecuencia (HF) - que se genera desde la parte de baja frecuencia (LF) codificada por forma de onda de la señal decodificada por medio de transposición en la región espectral HF ("emparchado") y la aplicación de un post-procesamiento realizado mediante parametrización. No obstante, si, por ejemplo, la estructura fina del espectro en un parche copiado en alguna región objetivo es extremadamente diferente de la estructura fina del espectro del contenido original, podrían obtenerse fallos molestos y se podría degradar la calidad perceptual de la señal de audio
20 decodificada.

[0003] En los esquemas BWE, la reconstrucción de la región espectral HF sobre una frecuencia dada, así llamada frecuencia de cruce, a menudo está basada en el emparchado espectral. Generalmente, la región HF está compuesta por múltiples parches adyacentes y cada uno de estos parches está provisto de regiones de paso de banda
25 (BP) del espectro LF por debajo de la frecuencia de cruce dada. Los sistemas de última generación realizan el emparchado eficazmente dentro de una representación de banco de filtros mediante el copiado de un conjunto de coeficientes sub-banda adyacentes desde una fuente a la región objetivo. En una etapa siguiente, la envolvente espectral se ajusta de tal forma que represente de la forma más cercana posible la envolvente de la señal HF original que se ha medido en el codificador y transmitido en el flujo de bits como información lateral.

30 **[0004]** No obstante, a menudo existe una discordancia en la estructura fina espectral que podría llevar a la percepción de fallos. Una discordancia comúnmente conocida está relacionada con la tonalidad. Si la HF original incluye un tono con un contenido de energía bastante dominante y el parche que se va a copiar en la ubicación espectral del tono tiene una característica ruidosa, este ruido de paso de banda se puede escalar de tal forma que se
35 vuelva audible como una molesta ráfaga de ruido.

[0005] La SBR (por sus siglas en inglés: *spectral band replication*/replicación de banda espectral) es una BWE muy conocida empleada en los códecs de audio contemporáneos [1]. En la SBR, el problema de la discordancia de tonalidad se soluciona mediante la inserción de sinusoides de reemplazo artificiales. No obstante, esto requiere que
40 se transmita información lateral adicional al decodificador aumentando la demanda de bits de datos BWE. Además, los tonos insertados pueden llevar a la inestabilidad a lo largo del tiempo si la inserción del tono conmuta encendido/apagado para bloques subsiguientes.

[0006] La cobertura inteligente de intervalos (IGF) indica una técnica de codificación semi-paramétrica dentro
45 de códecs modernos como Audio MPEG (por sus siglas en inglés: *Motion Picture Experts Group*/Grupo de Expertos en Imágenes en Movimiento)-H 3D o el códec EVS (por sus siglas en inglés: *Enhanced Voice Services*/Servicios de Voz Mejorados) 3 gpp (por sus siglas en inglés: *3rd. Generation Partnership Project*/Proyecto Asociación de Tercera Generación). La IGF se puede aplicar para cubrir agujeros espectrales introducidos por el procedimiento de cuantificación en el codificador debido a restricciones en cuanto a baja tasa de bits. Generalmente, si el presupuesto
50 limitado de bits no permite una codificación transparente, los agujeros espectrales emergen primero en la región de alta frecuencia (HF) de la señal y afectan de modo creciente al intervalo espectral superior completo para tasas de bits más bajas. En el lado del decodificador, tales agujeros espectrales se sustituyen mediante IGF usando el contenido HF sintético generado en forma semi-paramétrica del contenido de baja frecuencia (LF), y post-procesamiento controlado mediante información lateral paramétrica adicional.

55 **[0007]** Dado que la IGF está basada fundamentalmente en la cobertura del espectro de alta frecuencia mediante la copia de partes espectrales (los así llamados elementos de imagen) de frecuencias más bajas y en el ajuste de las energías aplicando un factor de ganancia, esto puede resultar problemático si en la señal original, el intervalo de frecuencia usado como la fuente del procedimiento de copiado difiere de su destino en términos de
60 estructura fina espectral.

[0008] Uno de esos casos que puede tener un fuerte impacto perceptual es una diferencia en la tonalidad. Esta discordancia en la tonalidad puede ocurrir de dos formas diferentes: o bien un intervalo de frecuencia con fuerte tonalidad se copia a una región espectral que se supone que es similar a ruido en estructura, o por el contrario, con
65 ruido que reemplaza un componente tonal en la señal original. En la IGF el primer caso, que es más común ya que la

mayoría de las señales de audio usualmente se vuelven más similares a ruido hacia frecuencias más altas, se maneja mediante la aplicación de blanqueado espectral donde se transmiten parámetros al decodificador que señalan cuánto blanqueado es necesario, en caso de serlo. En el último caso, la tonalidad podría corregirse usando la capacidad de codificación de banda completa del codificador de núcleo para preservar las líneas tonales en la banda HF mediante codificación por forma de onda. Estas así llamadas "líneas supervivientes" podrían seleccionarse sobre la base de la tonalidad fuerte. La codificación por forma de onda es muy demandante en términos de tasa de bits y en escenarios con baja tasa de bits es muy probable que no se pueda abordar. Además, se tiene que evitar conmutar de trama en trama entre codificar y no codificar un componente tonal ya que esto podría causar fallos molestos.

10 **[0009]** La tecnología de cobertura inteligente de intervalos se da a conocer y describe adicionalmente en la solicitud de patente Europea EP 2830054 A1. La tecnología IGF soluciona los problemas relacionados con la separación de la extensión de ancho de banda por una parte, y con la decodificación de núcleo por la otra, al realizar la extensión de ancho de banda en el mismo dominio espectral en el cual opera el decodificador de núcleo. Por lo tanto, se proporciona un codificador/decodificador de núcleo de velocidad completa que codifica y decodifica el
15 intervalo completo de la señal de audio. No se requiere la necesidad de un sub-muestreador en el lado del codificador y un sobre-muestreador en el lado del decodificador. En cambio, el procesamiento total se realiza en la tasa de muestreo completa o el dominio de ancho de banda completo. A fin de obtener una alta ganancia de codificación, la señal de audio se analiza con el fin de encontrar el primer conjunto de primeras porciones espectrales que tiene que ser codificado con una alta resolución, donde este primer conjunto de primeras porciones espectrales puede incluir,
20 en una realización, porciones tonales de la señal de audio. Por otra parte, los componentes no tonales o ruidosos en la señal de audio que constituyen un segundo conjunto de segundas porciones espectrales se codifican paramétricamente con baja resolución espectral. A continuación, la señal de audio codificada solo requiere el primer conjunto de primeras porciones espectrales codificado de la manera que conserva la forma de onda con una alta resolución espectral, y además el segundo conjunto de segundas porciones espectrales codificado paramétricamente
25 con una baja resolución usando "elementos de imagen" de frecuencia proporcionados por el primer conjunto. Del lado del decodificador, el decodificador de núcleo, que es un decodificador de banda completa, reconstruye el primer conjunto de primeras porciones espectrales en una manera que conserva la forma de onda, es decir, sin ningún conocimiento de que existe regeneración de frecuencia adicional alguna. No obstante, el espectro generado tiene muchos intervalos espectrales. Estos intervalos son posteriormente cubiertos con la tecnología de cobertura inteligente
30 de intervalos (IGF) de la invención usando una regeneración de frecuencia que aplica datos paramétricos por una parte y, por otra parte, usando un intervalo espectral fuente, es decir, las primeras porciones espectrales reconstruidas por el decodificador de audio de velocidad completa.

[0010] La tecnología IGF también se incluye y se describe en 3GPP TS 26.445 V13.2.0 (2016-06), Proyecto
35 Asociación de Tercera Generación; Servicios de Grupo de Especificación Técnica y Aspecto del Sistema; Códec para Servicios de Voz Mejorados (EVS); Descripción Algorítmica Detallada (publicación 13). En particular, se hace referencia a la sección 5.3.3.2.11 "Cobertura Inteligente de Intervalos" de esta referencia con respecto a un lado del codificador, y se hace referencia adicional a la sección 6 y en particular a la sección 6.2.2.3.8 "Aplicación de IGF" y otros pasajes relacionados con IGF, tal como la sección 6.2.2.2.9 "Lector de flujo de bits IGF" o la sección 6.2.2.3.11
40 "Aplanamiento Temporal IGF" con respecto a la implementación del lado del decodificador.

[0011] El documento de patente EP 2301027 B1 describe un aparato y un procedimiento para generar datos de salida de extensión de ancho de banda. En señales de voz habladas, el descenso del ruido de fondo calculado produce una calidad perceptualmente superior en comparación con el ruido de fondo calculado originalmente. Como
45 resultado, la voz suena menos reverberante en este caso. En el caso en que las señales de audio comprendan sibilantes, un aumento artificial de ruido de fondo puede cubrir las desventajas en el procedimiento de emparchado en relación con los sibilantes. Por lo tanto, la referencia describe el suministro de una disminución del ruido de fondo para señales tales como la voz hablada y un aumento del ruido de fondo para señales que comprenden, por ejemplo, sibilantes. Para distinguir las diferentes señales, las realizaciones usan los datos de distribución de energía (por
50 ejemplo, un parámetro de sibilancia) que mide si la energía está principalmente ubicada en frecuencias más altas o en una frecuencia más alta, o, en otras palabras, si la representación espectral de la señal de audio muestra una inclinación decreciente o creciente hacia frecuencias más altas. Implementaciones adicionales usan también el primer coeficiente LPC (LPC igual a codificación de predicción lineal) para generar el parámetro de sibilancia.

55 **[0012]** El documento US 2008/0027733 A1 describe otra estrategia para la extensión del ancho de banda.

[0013] Es un objeto de la presente invención proporcionar un concepto mejorado para codificación de audio o procesamiento de audio.

60 **[0014]** Este objeto se logra mediante un aparato para codificar una señal de audio de la reivindicación 1, un procedimiento para codificar una señal de audio de la reivindicación 22, un sistema para procesar una señal de audio de la reivindicación 23, un procedimiento para procesar una señal de audio de la reivindicación 24 o un programa informático de la reivindicación 25.

65 **[0015]** Por lo tanto, la presente invención se basa en el hallazgo de que, con el fin de encontrar si una

reconstrucción usando un cierto parámetro en el lado del decodificador resuelve una cierta característica requerida por la señal de audio, la primera banda espectral, que es generalmente la banda fuente, se analiza para obtener el primer resultado de análisis. De forma análoga, la segunda banda espectral, que es generalmente la banda objetivo, y que se reconstruye en el lado del decodificador usando la primera banda espectral, es decir, la banda fuente, es
5 analizada adicionalmente por el analizador para obtener el segundo resultado de análisis. Por lo tanto, se calcula un resultado de análisis independiente para la banda fuente así como para la banda objetivo.

[0016] A continuación, sobre la base de estos dos resultados de análisis, un compensador calcula un valor de compensación para cambiar un determinado parámetro que se habría obtenido sin ninguna compensación a un valor
10 modificado. En otras palabras, la presente invención se aparta del procedimiento típico, en el cual se calcula un parámetro para la segunda banda espectral desde la señal de audio original y se transmite al decodificador de modo que la segunda banda espectral se reconstruye usando el parámetro calculado, y resulta en cambio en un parámetro compensado calculado desde la banda objetivo por una parte, y el valor de compensación que depende del primer y del segundo resultado de análisis, por la otra.

[0017] El parámetro compensado se puede calcular al calcular primeramente el parámetro no compensado y, a continuación, este parámetro no compensado se puede combinar con el valor de compensación para obtener el parámetro compensado, o el parámetro compensado se puede calcular directamente, sin el parámetro sin compensar como resultado intermedio. El parámetro compensado se puede transmitir, a continuación, desde el codificador al
20 decodificador y, a continuación, el decodificador aplica una determinada tecnología que mejora el ancho de banda tal como una replicación de banda espectral o cobertura inteligente de intervalos o cualquier otro procedimiento usando el valor del parámetro compensado. Por lo tanto, la estricta obediencia a un cierto algoritmo de cálculo de parámetros independientemente de si el parámetro proporciona el resultado deseado de mejoramiento de banda espectral se supera de modo flexible al realizar, además del cálculo de parámetros, el análisis de la señal en la banda fuente y la
25 banda objetivo y el cálculo subsiguiente de un valor de compensación sobre la base del resultado desde la banda fuente y el resultado desde la banda objetivo, es decir, desde la primera banda espectral y la segunda banda espectral, respectivamente.

[0018] Preferentemente, el analizador y/o el compensador aplican una especie de modelo psicoacústico que
30 determina una discordancia psicoacústica. Por lo tanto, en una realización, el cálculo del valor de compensación está basado en la detección de una discordancia psicoacústica de ciertos parámetros de la señal tal como la tonalidad y se aplica una estrategia de compensación para minimizar las molestias perceptuales generales a través de la modificación de otros parámetros de la señal tales como factores de ganancia de banda espectral. Así, realizando un intercambio entre diferentes tipos de fallos, se obtiene un resultado perceptualmente bien equilibrado.

[0019] Contrariamente a las estrategias de la técnica existente "que tratan de fijar la tonalidad a cualquier
35 precio", las realizaciones enseñan más bien a remediar los fallos mediante la aplicación del amortiguamiento de las partes problemáticas del espectro donde se detecta la discordancia de tonalidad, por lo cual se realiza un intercambio de discordancia de la envolvente de energía espectral versus una discordancia de tonalidad.

[0020] Ante la entrada de diversos parámetros de la señal, la estrategia de compensación que contiene un
40 modelo de molestia perceptual puede decidir una estrategia para obtener el mejor ajuste perceptual en vez de un mero ajuste de parámetros de la señal.

[0021] La estrategia consiste en ponderar el significado perceptual de fallos potenciales y elegir una
45 combinación de parámetros para minimizar la reducción de la calidad en general.

[0022] Esta estrategia está principalmente destinada a su aplicación dentro de una BWE basada en una transformada como la MDCT (por sus siglas en inglés: *Modified Discrete Cosine Transform*/Transformada Discreta de
50 Coseno Modificada). No obstante, las enseñanzas de la invención tienen aplicación general, por ejemplo, de forma análoga dentro de un sistema basado en QMF (por sus siglas en inglés: *Quadrature Mirror Filter*/Filtro Espejo en Cuadratura).

[0023] Un escenario posible en el cual se puede aplicar esta técnica es la detección y subsiguiente
55 amortiguamiento de bandas de ruido en el contexto de la cobertura inteligente de intervalos (IGF).

[0024] Las realizaciones gestionan una posible discordancia de tonalidad a través de la detección de su
aparición y la reducción de su efecto mediante la atenuación del factor de escala correspondiente. Por una parte, esto puede llevar a una desviación de la envolvente de energía espectral original, pero, por otra parte, a una reducción de
60 ruidos de HF que contribuye a un aumento general en la calidad perceptual.

[0025] Por lo tanto, las realizaciones mejoran la calidad perceptual a través de una novedosa técnica de
compensación paramétrica, generalmente controlada por un modelo de molestia perceptual, particularmente en casos en los que, por ejemplo, existe una discordancia en la estructura espectral fina entre la fuente o la primera banda
65 espectral y el objetivo o segunda banda espectral.

[0026] Las realizaciones preferidas se describen a continuación en el contexto de las figuras adjuntas, en las cuales:

5 La Fig. 1 ilustra un diagrama de bloques de un aparato para codificar una señal de audio de acuerdo con una realización;

La Fig. 2 ilustra un diagrama de bloques de un aparato para codificar con un enfoque en el detector de compensación;

10 La Fig. 3a ilustra una representación esquemática de un espectro de audio que tiene un intervalo fuente y una IGF o intervalo de extensión de banda y un mapeo asociado entre las bandas fuente y de destino;

La Fig. 3b ilustra un espectro de una señal de audio donde el codificador de núcleo aplica la tecnología IGF y donde existen líneas sobrevivientes en la segunda banda espectral;

15 La Fig. 3c ilustra una representación de una primera información de audio simulada en la primera banda espectral que se va a usar para calcular el primer resultado de análisis;

La Fig. 4 ilustra una representación más detallada del compensador;

20 La Fig. 5 ilustra una representación más detallada del calculador de parámetros;

La Fig. 6 ilustra un diagrama de flujo que ilustra la funcionalidad del detector de compensación en una realización;

25 La Fig. 7 ilustra una funcionalidad del calculador de parámetros para calcular un factor de ganancia no compensado;

La Fig. 8a ilustra la implementación de un codificador que tiene un decodificador de núcleo para calcular el primer resultado de análisis desde una primera banda espectral codificada y decodificada;

30 La Fig. 8b ilustra un diagrama de bloques de un codificador en una realización, en la cual se aplica un simulador de parche para generar una primera línea de ancho de banda espectral desplazada desde la segunda banda espectral para obtener el primer resultado de análisis;

35 La Fig. 9 ilustra un efecto de una discordancia de tonalidad en una implementación de cobertura inteligente de intervalos;

La Fig. 10 ilustra, en una realización, la implementación del codificador paramétrico; y

40 Las Figs. 11a-11c ilustran los resultados de pruebas de escucha obtenidos de la codificación de información de audio usando valores de parámetro compensado.

[0027] La Fig. 1 ilustra un aparato para codificar una señal de audio 100 en una realización de la presente invención. El aparato comprende un codificador de núcleo 110 y un codificador paramétrico 120. Asimismo, el codificador de núcleo 110 y el codificador paramétrico 120 están conectados, en sus lados de entrada, con un analizador espectral 130 y están conectados, en sus lados de salida, con una interfaz de salida 140. La interfaz de salida 140 genera una señal de audio codificada 150. La interfaz de salida 140 recibe, por una parte, una señal de núcleo codificada 160 y al menos un parámetro para la segunda banda espectral y, generalmente, una representación de parámetros completa que comprende el parámetro para una segunda banda espectral en la línea de entrada 170. Asimismo, el analizador espectral 130 separa la señal de audio 100 en una primera banda espectral 180 y una segunda banda espectral 190. En particular, el calculador de parámetros comprende un analizador 121 que se ilustra como un analizador de señal en la Fig. 1 para analizar la primera información de audio en la primera banda espectral 180 para obtener un primer resultado de análisis 122 y para analizar la segunda información de audio en la segunda banda espectral 190 para obtener un segundo resultado de análisis 123. Tanto el primer resultado de análisis 122 como el segundo resultado de análisis 123 se proporcionan a un compensador 124 para calcular un valor de compensación. Por lo tanto, el compensador 124, está configurado para usar el primer resultado de análisis 122 y el segundo resultado de análisis 123 para calcular un valor de compensación. A continuación, el valor de compensación 125, por una parte, y al menos la segunda información de audio de la segunda banda espectral 190 (se puede usar también la primera información espectral de la primera banda espectral) se proporcionan a un calculador de parámetros 126 para calcular un parámetro 170 de la segunda información de audio en la segunda banda espectral usando el valor de compensación 125.

[0028] El analizador espectral 130 en la Fig. 1 puede ser, por ejemplo, un convertidor directo tiempo/frecuencia para obtener bandas espectrales individuales o líneas MDCT. En esta implementación, por lo tanto, el analizador espectral 130 implementa una transformada de coseno discreta modificada (MDCT) para obtener datos espectrales. A continuación, esta información espectral se analiza ulteriormente con el fin de separar información para el codificador

de núcleo 110 por una parte, e información para el codificador paramétrico 120, por la otra. La información para el codificador de núcleo 110 comprende al menos la primera banda espectral. Asimismo, la información del núcleo puede comprender adicionalmente información fuente adicional cuando el codificador de núcleo tiene que codificar más de una banda fuente.

[0029] Por lo tanto, el codificador de núcleo puede recibir, como información de entrada que se va a codificar en el núcleo, el ancho de banda completo por debajo de una frecuencia de cruce en el caso de tecnologías de replicación de banda espectral, mientras que el codificador paramétrico recibe entonces toda la información de audio por encima de esta frecuencia de cruce.

[0030] En el caso de una estructura de cobertura inteligente de intervalos, no obstante, el codificador de núcleo 110 puede recibir adicionalmente líneas espectrales por encima de una frecuencia inicial IGF, que también son analizadas por el analizador espectral 130, de modo que el analizador espectral 130 determina adicionalmente datos inclusive por encima de la frecuencia inicial IGF donde esta información por encima de la frecuencia inicial IGF es codificada adicionalmente por el codificador de núcleo. A tal fin, el analizador espectral 130 puede implementarse también como una "máscara tonal" sobre la que, por ejemplo, también se expone en la sección 5.3.3.2.11.5 "Máscara Tonal IGF" como se describe en 3GPP TS 26.445 V13.0.0(12). Por lo tanto, con el fin de determinar qué componente espectral deberá transmitirse con el codificador de núcleo, la máscara tonal se calcula mediante el analizador espectral 130. Por lo tanto, se identifica todo el contenido espectral importante, mientras que el contenido que es adecuado para codificación paramétrica mediante IGF es cuantizado a cero por la máscara tonal. No obstante, el analizador espectral 130 envía el contenido espectral que es adecuado para codificación paramétrica al codificador paramétrico 120, y esta información puede ser, por ejemplo, la información que se ha ajustado a cero por el procesamiento de la máscara tonal.

[0031] En una realización, ilustrada en la Fig. 2, el codificador paramétrico 120 está configurado adicionalmente para codificar paraméricamente una tercera información de audio en una tercera banda espectral para obtener un parámetro adicional 200 para esta tercera banda espectral. En este caso, el analizador 121 está configurado para analizar la tercera información de audio en la tercera banda espectral 202 para obtener un tercer resultado de análisis 204 además del primer resultado de análisis 122 y el segundo resultado de análisis 123.

[0032] Asimismo, el codificador paramétrico 120 de la Fig. 1 comprende adicionalmente un detector de compensación 210 para detectar, usando al menos el tercer resultado de análisis 204, si se tiene que compensar o no la tercera banda espectral. El resultado de esta detección es emitido por una línea de control 212 que indica una situación de compensación para la tercera banda espectral o no. El calculador de parámetros 126 está configurado para calcular el parámetro adicional 200 para la tercera banda espectral sin ningún valor de compensación, cuando el detector de compensación detecta que la tercera banda espectral no se ha de compensar tal como lo establece la línea de control 212. No obstante, si el detector de compensación detecta que la tercera banda espectral se tiene que compensar, entonces el calculador de parámetros está configurado para calcular el parámetro adicional 200 para la tercera banda espectral con un valor de compensación adicional calculado por el compensador 124 a partir del tercer resultado de análisis 200.

[0033] En una realización preferida, donde se aplica una compensación cuantitativa, el analizador 121 está configurado para calcular, como el primer resultado de análisis, un primer valor cuantitativo 122 y como el segundo resultado de análisis un segundo valor cuantitativo 123. A continuación, el compensador 124 está configurado para calcular un valor de compensación cuantitativo 125 a partir del primer valor cuantitativo y a partir del segundo valor cuantitativo. Finalmente, el calculador de parámetros está configurado para calcular el parámetro cuantitativo usando el valor de compensación cuantitativo.

[0034] No obstante, la presente invención es aplicable también cuando solo se obtienen resultados de análisis cualitativo. En esta situación, se calcula un valor de compensación cualitativo, que controla a continuación el calculador de parámetros para disminuir o aumentar un determinado parámetro no compensado en un cierto grado. Por lo tanto, ambos resultados de análisis en conjunto pueden dar como resultado un determinado aumento o disminución de un parámetro, el determinado aumento o disminución es fijo y por lo tanto no depende de ningún resultado cuantitativo. No obstante, los resultados cuantitativos se prefieren sobre incrementos fijos de aumento/disminución, aunque los últimos cálculos requieren menos esfuerzo computacional.

[0035] Preferentemente, el analizador de señal 121 analiza una primera característica de la información de audio para obtener el primer resultado de análisis y analiza adicionalmente la misma primera característica de la segunda información de audio en la segunda banda espectral para obtener el segundo resultado de análisis. Contrariamente, el calculador de parámetros está configurado para calcular el parámetro de la segunda información de audio en la segunda banda espectral mediante la evaluación de una segunda característica donde la segunda característica es diferente de la primera característica.

[0036] A modo de ejemplo, la Fig. 2 ilustra la situación en la que la primera característica es una estructura fina espectral o una distribución de energía dentro de una determinada banda tal como la primera, la segunda o cualquier

otra banda. Contrariamente, la segunda característica aplicada por el calculador de parámetros o determinada por el calculador de parámetros es una medición de envolvente espectral, una medición de energía o una medición de potencia o generalmente una medición relacionada con la amplitud que ofrece una medición absoluta o relativa de la potencia/energía en una banda, tal como, por ejemplo, un factor de ganancia. No obstante, otros parámetros que

5 miden una característica diferente de una característica de factor de ganancia se pueden calcular también con el calculador de parámetros. Asimismo, otras características para la banda fuente individual, por una parte, y la banda de destino por la otra, es decir, la primera banda espectral y la segunda banda espectral, respectivamente, se pueden aplicar y analizar mediante el analizador 121.

10 **[0037]** Asimismo, el analizador 121 está configurado para calcular el primer resultado de análisis 122 sin usar la segunda información de audio en la segunda banda espectral 190 y para calcular además el segundo resultado de análisis 123 sin usar la primera información de audio en la primera banda espectral 180 donde, en esta realización, la primera banda espectral y la segunda banda espectral son entre sí mutuamente excluyentes, es decir no se superponen entre sí.

15 **[0038]** Asimismo, el analizador espectral 130 está configurado adicionalmente para construir tramas de la señal de audio o para generar ventanas de un flujo de entrada de muestras de audio para obtener tramas de muestras de audio, donde las muestras de audio en tramas vecinas están superpuestas entre sí. En el caso de un solapamiento de 50 %, por ejemplo, una segunda porción de una trama anterior tiene muestras de audio que derivan de las mismas

20 muestras de audio originales incluidas en la primera mitad de la trama subsiguiente, donde las muestras de audio dentro de una trama se derivan de las muestras de audio originales mediante formación de ventana.

[0039] En este caso, cuando la señal de audio comprende una secuencia temporal de tramas como, por ejemplo, la adicionalmente proporcionada por el bloque 130 de la Fig. 1, que además tiene una funcionalidad

25 generadora de trama, el compensador 124 está configurado para calcular un valor de compensación actual para una trama actual usando un valor de trama de compensación previo para una trama anterior. Generalmente, esto da como resultado una especie de operación de suavizado.

[0040] Como se señalará más adelante, el detector de compensación 210 ilustrado en la Fig. 2 puede, adicional

30 o alternativamente, a partir de otras características en la Fig. 2, comprender una entrada de espectro de energía y una entrada transitoria ilustrada en 221, 223, respectivamente.

[0041] En particular, el detector de compensación 210 está configurado solo para dar instrucciones de una compensación que será utilizada por el calculador de parámetros 126, cuando un espectro de energía de la señal de

35 audio original 100 de la Fig. 1 está disponible. Este dato, es decir, si el espectro de energía está disponible, está señalado por un elemento de información o bandera determinado.

[0042] Además, el detector de compensación 210 está configurado para permitir solo una operación de compensación mediante la línea de control 212, cuando una línea de información transitoria 223 señala que, para la

40 trama actual, no está presente un transitorio. Así, cuando la línea 223 señala que un transitorio está presente, la operación de compensación completa se deshabilita, independientemente de cualquier resultado del análisis. Naturalmente, esto se aplica a la tercera banda espectral, cuando una compensación ha sido señalada para la segunda banda espectral. No obstante, esto se aplica también a la segunda banda espectral en una trama determinada, cuando para esta trama se detecta una situación tal como una situación transitoria. A continuación, puede ocurrir y ocurrirá la

45 situación en que, para una cierta trama de tiempo, ninguna compensación de parámetros tiene lugar en absoluto.

[0043] La Fig. 3a ilustra una representación de un espectro de amplitudes $A(f)$ o amplitudes al cuadrado $A^2(f)$. En particular, se ilustra una frecuencia inicial XOVER o IGF.

50 **[0044]** Asimismo, se ilustra un conjunto de bandas fuente superpuestas, donde las bandas fuente comprenden la primera banda espectral 180, una banda fuente adicional 302 e incluso una banda fuente adicional más 303. Adicionalmente, las bandas de destino por encima de la frecuencia IGF o XOVER son la segunda banda espectral 190, una banda de destino adicional 305, e incluso una banda de destino adicional más 307 y la tercera banda espectral 202, por ejemplo.

55 **[0045]** Generalmente, el mapeo de funciones dentro de la IGF o la estructura de la extensión de ancho de banda definen un mapeo entre las bandas fuente individuales 180, 302, 303 y las bandas de destino individuales 305, 190, 307, 202. Este mapeo se puede fijar como es el caso en 3GPP TS 26.445 o se puede determinar adaptativamente mediante cierto algoritmo codificador de IGF. En cualquier caso, la Fig. 3a ilustra, en la tabla inferior, el mapeo entre

60 una banda de destino y la banda fuente para el caso de bandas de destino sin solapamiento y bandas fuente con solapamiento, independientemente de si este mapeo es fijo o está determinado adaptativamente y realmente se ha determinado adaptativamente para una cierta trama, estando el espectro ilustrado en la porción superior de la Fig. 3a.

[0046] La Fig. 4 ilustra una implementación más detallada del compensador 124. El compensador 124 recibe,

65 en esta implementación, además del primer resultado de análisis 122, que puede ser una medición de aplanamiento

espectral, un factor de cresta, un valor de inclinación espectral o cualquier otra clase de información paramétrica para la primera banda espectral, un análisis de resultado 123 para la segunda banda espectral. Este resultado de análisis puede ser, una vez más, una medición de aplanamiento espectral para la segunda banda espectral, un factor de cresta para la segunda banda espectral o un valor de inclinación, es decir, un valor de inclinación espectral limitado a la segunda banda espectral mientras que el valor de inclinación o el valor de inclinación espectral para la primera banda espectral está limitado también para la primera banda espectral. Además, el compensador 124 recibe una información espectral sobre la segunda banda espectral tal como una línea de detención de la segunda banda espectral. Por lo tanto, en la situación en la cual el calculador de parámetros 126 de la Fig. 2 está configurado para codificar paramétricamente una tercera información de audio en la tercera banda espectral 202, la tercera banda espectral comprende frecuencias más altas que la segunda banda espectral. Esto se ilustra también en el ejemplo de la Fig. 3a, donde la tercera banda espectral está a frecuencias más altas que la segunda banda espectral, es decir, donde la banda 202 tiene frecuencias más altas que la banda 190. En esta situación, el compensador 124 está configurado para usar un valor de ponderación para calcular el valor de compensación para la tercera banda espectral, donde este tercer valor de ponderación es diferente de un valor de ponderación usado para calcular el valor de compensación para la segunda banda espectral. Por lo tanto, en general, el compensador 124 ejerce influencia en el cálculo del valor de compensación 125, de modo que, para los mismos otros valores de entrada, el valor de compensación es menor para frecuencias más altas.

[0047] El valor de ponderación puede ser, por ejemplo, un exponente aplicado en el cálculo del valor de compensación sobre la base del primer y el segundo resultado de análisis tal como el exponente á, como se describirá más adelante, o puede ser, por ejemplo, un valor multiplicador o inclusive un valor que se va a sumar o restar de modo que se obtiene una influencia diferente para frecuencias más altas en comparación con la influencia obtenida cuando el parámetro se va a calcular para frecuencias más bajas.

[0048] Además, como se ilustra en la Fig. 4, el compensador recibe una relación tonal/ruido para la segunda banda espectral con el fin de calcular el valor de compensación dependiente de la relación tonal/ruido de la segunda información de audio en la segunda banda espectral. Así, se obtiene un primer valor de compensación para una primera relación tonal/ruido o se obtiene un segundo valor de compensación para una segunda relación tonal/ruido, donde el primer valor de compensación es mayor que el segundo valor de compensación cuando la primera relación tonal/ruido es mayor que la segunda relación tonal/ruido.

[0049] Como se estableció, el compensador 124 está configurado para determinar generalmente el valor de compensación mediante la aplicación de un modelo psicoacústico, donde el modelo psicoacústico está configurado para evaluar la discordancia psicoacústica entre la primera información de audio y la segunda información de audio usando el primer resultado de análisis y el segundo resultado de análisis para obtener el valor de compensación. Este modelo psicoacústico que evalúa la discordancia psicoacústica se puede implementar como un cálculo con prealimentación como se expondrá más adelante en el contexto de los siguientes cálculos SFM (por sus siglas en inglés: *Spectral Flatness Measure*/Medición de Aplanamiento Espectral) o puede, alternativamente, ser un módulo de cálculo con retroalimentación que aplica una clase de análisis mediante procedimientos de síntesis. Asimismo, el modelo psicoacústico puede ser implementado también como una red neuronal o una estructura similar que es automáticamente drenada por ciertos datos de entrenamiento para decidir en qué caso es necesaria una compensación y en qué caso no lo es.

[0050] A continuación, se ilustra la funcionalidad del detector de compensación 210 ilustrado en la Fig.2 o, en general, un detector incluido en el calculador de parámetros 120.

[0051] La funcionalidad del detector de compensación está configurada para detectar una situación de compensación cuando una diferencia entre el primer resultado de análisis y el segundo resultado de análisis tiene una característica predeterminada como se ilustra, por ejemplo, en 60 y 602 en la Fig. 6. El bloque 600 está configurado para calcular una diferencia entre el primer y el segundo resultado de análisis y el bloque 602 determina a continuación si la diferencia tiene una característica predeterminada o un valor predeterminado. Si se determina que la característica predeterminada no está allí, entonces se determina mediante el bloque 602 que no se va a realizar compensación como se ilustra en 603. No obstante, si se determina que existe la característica predeterminada, entonces el control prosigue mediante la línea 604. Asimismo, el detector está configurado para alternativa o adicionalmente determinar si el segundo resultado de análisis tiene un cierto valor predeterminado o una cierta característica predeterminada. Si se determina que la característica no existe, entonces la línea 605 señala que no se va a realizar ninguna compensación. No obstante, si se determina que el valor predeterminado está allí, el control prosigue mediante la línea 606. En algunas realizaciones, las líneas 604 y 606 pueden ser suficientes para determinar si existe o no una compensación. No obstante, en la realización ilustrada en la Fig. 6, se realizan determinaciones adicionales basadas en la inclinación espectral de la segunda información de audio para la segunda banda espectral 190 de la Fig. 1, como se describe más adelante.

[0052] En una realización, el analizador está configurado para calcular una medición de aplanamiento espectral, un factor de cresta o un cociente de la medición de aplanamiento espectral y el factor de cresta para la primera banda espectral como el primer resultado de análisis y para calcular una medición de aplanamiento espectral

o un valor de cresta o un cociente de la medición de aplanamiento espectral y el factor de cresta de la segunda información de audio como el segundo resultado de análisis.

[0053] En tal realización, el calculador de parámetros 126 está configurado además para calcular, a partir de la segunda información de audio, una información de envolvente espectral o un factor de ganancia. Asimismo, en tal realización, el compensador 124 está configurado para calcular el valor de compensación 125 de modo que, para una primera diferencia entre el primer resultado de análisis y el segundo resultado de análisis, se obtiene un primer valor de compensación, y para una diferencia entre el primer resultado de análisis y el segundo resultado de análisis, se calcula un segundo valor de compensación, donde la primera diferencia es mayor que la segunda diferencia, cuando el primer valor de compensación es mayor que el segundo valor de compensación.

[0054] A continuación, la descripción de la Fig. 6 continuará mediante el ejemplo de la determinación adicional opcional acerca de si se va a detectar o no una situación de compensación.

[0055] En el bloque 608, se calcula una inclinación espectral a partir de la segunda información de audio. Cuando se determina que esta inclinación espectral está por debajo de un umbral como se ilustra en 610, entonces se afirma positivamente una situación de compensación como se ilustra en 612. Cuando, sin embargo, se determina que la inclinación espectral no está por debajo del umbral predeterminado, sino que está por encima del umbral, entonces esta situación se señala mediante una línea 614. En el bloque 616, se determina si un componente tonal está cerca de un borde de la segunda banda espectral 190. Cuando se determina que existe un componente tonal cerca del borde como se ilustra en el elemento 618, entonces una vez más se afirma positivamente una situación de compensación. Cuando, sin embargo, se determina que no existe ningún componente tonal cerca de un borde, entonces se cancela cualquier compensación, es decir, se desconecta como se ilustra mediante la línea 620. La determinación en el bloque 616, es decir, la determinación de si un componente tonal está cerca de un borde o no, se hace realizando, en cualquier realización, un cálculo de SFM desplazado. Cuando existe un marcado declive en la pendiente según se determina en el bloque 608, entonces la región de la frecuencia, para la cual se calcula la SFM, se desplazará hacia abajo la mitad del ancho de la correspondiente banda de factor de escala (SFB) o la segunda banda espectral. Para un marcado ascenso, la región de frecuencia, para la cual se calcula la SFM se desplaza hacia arriba la mitad del ancho de la segunda banda espectral. De esta forma, los componentes tonales que se supone están amortiguados, aún se pueden detectar correctamente debido a una baja SFM mientras que para valores SFM más altos, no se aplicará amortiguamiento.

[0056] A continuación, se expone con mayor detalle la Fig. 5. Particularmente, el calculador de parámetros 126 puede comprender el calculador 501 para calcular el parámetro no compensado de la información de audio para la segunda banda espectral, es decir, la banda de destino, y el calculador de parámetros 126 comprende además un combinador 503 para combinar el parámetro no compensado 502 y el valor de compensación 125. Esta combinación puede ser, por ejemplo, una multiplicación, cuando el parámetro no compensado 502 es un valor de ganancia y el valor de compensación 105 es un valor de compensación cuantitativo. No obstante, la combinación realizada por el combinador 503 también puede ser, alternativamente, una operación de ponderación usando el valor de compensación como un exponente o una modificación aditiva cuando se usa el valor de compensación como un valor de suma o resta.

[0057] Además, cabe destacar que la realización ilustrada en la Fig. 5, donde se calcula el parámetro no compensado y, a continuación, se realiza una combinación subsiguiente con el valor de combinación, es tan solo una forma de realización. En realizaciones alternativas, el valor de compensación ya puede estar introducido en el cálculo del parámetro compensado de modo que el resultado intermedio con un parámetro no compensado explícito no se produce. En cambio, se realiza una sola operación, donde, como resultado de esta "sola operación", el parámetro compensado se calcula usando el valor de compensación y usando un algoritmo de cálculo que daría como resultado el parámetro no compensado, cuando el valor de compensación 125 no fuera introducido en tal cálculo.

[0058] La Fig. 7 ilustra un procedimiento para su aplicación mediante el calculador 501 para calcular el parámetro no compensado. La representación en la Fig. 7 "cálculo del factor de escala IGF" se corresponde a grandes rasgos con la sección 5.3.3.2.11.4 de 3gpp TS 26.445 V13.3.3 (2015/12). Cuando está disponible un espectro de energía P TCX (por sus siglas en inglés: *Transformed Coded Excitation*/Excitación de Codificación por Transformada) "complejo", entonces el calculador 501 para calcular el parámetro no compensado de la Fig. 5 realiza el cálculo de una medición relacionada con la amplitud para la segunda banda espectral a partir del espectro de energía P como se ilustra en 700. Además, el calculador 501 realiza el cálculo de una medición relacionada con la amplitud para la primera banda espectral a partir del espectro complejo P como se ilustra en 702. Adicionalmente, el calculador 501 realiza un cálculo de una medición relacionada con la amplitud a partir de la parte real de la primera banda espectral, es decir, la banda fuente como se ilustra en 704, de modo que se obtienen tres mediciones relacionadas con la amplitud E_{cplx} , objetivo, E_{cplx} , fuente, E_{real} , fuente y se introducen en una funcionalidad adicional de cálculo de factor de ganancia 706 para finalmente obtener un factor de ganancia que es una función del cociente entre E_{real} , fuente y E_{cplx} , fuente multiplicado por E_{cplx} , objetivo.

[0059] Cuando, alternativamente, el espectro de energía TCX complejo no está disponible, entonces la

medición relacionada con la amplitud solo se calcula a partir de la segunda banda espectral real como se ilustra en la parte inferior de la Fig. 7.

[0060] Asimismo, cabe destacar que el espectro de energía TCX P se calcula, por ejemplo, como se ilustra en la subcláusula 5.3.3.2.11.1.2 sobre la base de la siguiente ecuación:

$$P(sb)=R^2(sb)+I^2(sb), \quad sb=0,1,2,\dots,n-1.$$

[0061] Aquí, n es la longitud real de la ventana TCX, R es el vector que contiene la parte real valuada (transformada de coseno) del espectro TCX actual, e I es el vector que contiene la parte imaginaria (transformada de seno) del espectro TCX actual. En particular, el término "TCX" está relacionado con la terminología 3gpp, pero en general, menciona los valores espectrales en la primera banda espectral o la segunda banda espectral según los proporciona el analizador espectral 130 al codificador de núcleo 110 o al codificador paramétrico 120 de la Fig. 1.

[0062] La Fig. 8a ilustra una realización preferida, donde el analizador de señal 121 comprende además un decodificador de núcleo 800 para calcular una primera banda espectral codificada y decodificada de nuevo y para calcular, naturalmente, la información de audio en la primera banda espectral codificada/decodificada.

[0063] A continuación, el decodificador de núcleo 800 alimenta con la primera banda espectral codificada/decodificada un calculador de resultado de análisis 801 incluido en el analizador de señal 821 para calcular el primer resultado de análisis 122. Asimismo, el analizador de señal comprende un calculador de segundo resultado de análisis 802 incluido en el analizador de señal 121 de la Fig. 1 para calcular el segundo resultado de análisis calculado 123. Por lo tanto, el analizador de señal 121 está configurado de tal forma que el primer resultado de análisis real 122 se calcula usando la primera banda espectral codificada y decodificada de nuevo mientras que el segundo resultado de análisis se calcula a partir de la segunda banda espectral original. Por lo tanto, la situación en el lado del decodificador está mejor simulada en el lado de codificador, dado que la entrada en el calculador de resultado de análisis 801 ya tiene incluidos todos los errores de cuantificación en la primera información de audio decodificada para la primera banda espectral disponible en el decodificador.

[0064] La Fig. 8b ilustra una implementación preferida adicional del analizador de señal que tiene, ya sea alternativamente al procedimiento de la Fig. 8a, o adicionalmente al procedimiento de la Fig. 8a, un simulador de parche 804. El simulador de parche 804 ratifica específicamente la funcionalidad del codificador IGF, es decir, que pueden existir líneas o al menos una línea dentro de la segunda banda de destino que está realmente codificada por el codificador de núcleo.

[0065] Particularmente, esta situación se ilustra en la Fig. 3b.

[0066] La Fig. 3b ilustra, al igual que la Fig. 3a, la porción superior de la primera banda espectral 180 y la segunda banda espectral 190. No obstante, además de lo que se ha expuesto en la Fig. 3a, la segunda banda espectral comprende líneas específicas 351, 352 incluidas dentro de la segunda banda espectral que han sido determinadas por el analizador espectral 130 como líneas que están adicionalmente codificadas por el codificador de núcleo 110 además de la primera banda espectral 180.

[0067] Esta codificación específica de ciertas líneas por encima de la frecuencia inicial IGF 310 refleja la situación de que el codificador de núcleo 110 es un codificador de banda completa que tiene una frecuencia Nyquist hasta $f_{\max}$ 354 que es más alta que la frecuencia inicial IGF. Esto se contrapone con las implementaciones relacionadas con la tecnología SBR donde la frecuencia de cruce es también la frecuencia máxima, y, por lo tanto, la frecuencia Nyquist del codificador de núcleo 110.

[0068] El simulador de ensayo 804 recibe la primera banda espectral 180 o la primera banda espectral decodificada del decodificador de núcleo 800 y, además, información del analizador espectral 130 o del codificador de núcleo 110 de que existen realmente líneas en la segunda banda espectral que están incluidas en la señal de salida del codificador de núcleo. Esto es señalado por el analizador espectral 130, mediante una línea 806 o es señalado por el codificador de núcleo mediante una línea 808. El simulador de parche 804 simula ahora la primera información de audio para la primera banda espectral usando la primera información de audio directa para las cuatro bandas espectrales e insertando las líneas 351, 352 de la segunda banda espectral en la primera banda espectral al desplazar estas líneas a la primera banda espectral. Por lo tanto, las líneas 351' y 352' representan líneas espectrales obtenidas mediante el desplazamiento de las líneas 351, 352 de la Fig. 3b desde la segunda banda espectral hacia la primera banda espectral. Preferentemente, las líneas espectrales 351, 352 están generadas de tal forma para la primera banda espectral que la ubicación de estas líneas dentro de los bordes de la banda son idénticas en ambas bandas, es decir, la diferencia de frecuencia entre una línea y el borde de la banda es idéntico a la segunda banda espectral 190 y la primera banda espectral 180.

[0069] Por lo tanto, el simulador de parche emite una información simulada 808 ilustrada en la Fig. 3c que tiene una información directa de la primera banda espectral y, además, tiene las líneas desplazadas desde la segunda

banda espectral hacia la primera banda espectral. Ahora, el calculador de resultado de análisis 801 calcula el primer resultado de análisis 102 usando la información específica 808 mientras que el calculador de resultado de análisis 802 calcula el segundo resultado de análisis 123 de la segunda información de audio original en la segunda banda espectral, es decir, la información de audio original que incluye las líneas 351, 352 ilustradas en la Fig. 3b.

5

[0070] Este procedimiento con el simulador de parche 804 tiene la ventaja de que no es necesario poner ciertas condiciones en las líneas adicionales 351, 352 tal como alta tonalidad o algo más. En cambio, queda totalmente a criterio del analizador espectral 130 o del codificador de núcleo 110 decidir si ciertas líneas en la segunda banda espectral van a ser codificadas por el codificador de núcleo. No obstante, el resultado de esta operación, es tenido en cuenta automáticamente mediante el uso de estas líneas como una entrada adicional para el cálculo del primer resultado de análisis 122 como se ilustra en la Fig. 8b.

10

[0071] A continuación, se ilustra el efecto de una discordancia de tonalidad dentro de una estructura de cobertura inteligente de intervalos.

15

[0072] Con el fin de detectar fallos en la banda de ruido se tiene que determinar la diferencia en tonalidad entre las bandas de factor de escala (SFB) fuente y objetivo. Para el cálculo de la tonalidad se puede utilizar la medición de aplanamiento espectral (SFM). Si se encuentra una discordancia de tonalidad, donde la banda fuente es mucho más ruidosa que la banda objetivo, se deberá aplicar una cierta cantidad de amortiguamiento. Esta situación se ilustra en la Fig. 9 sin la aplicación del procesamiento de la invención.

20

[0073] También es sensato aplicar algo de suavizado a los factores de amortiguamiento con el fin de evitar un comportamiento abrupto de conexión/desconexión de la herramienta. A continuación se proporciona una descripción detallada de las etapas necesarias para aplicar amortiguamiento en los lugares correctos. (Cabe observar que solo se aplicará amortiguamiento si el espectro de energía TCX P está disponible y si la trama no es transitoria (bandera *esTransitoria* inactivada)).

25

[0074] Detección de discordancia de tonalidad: Parámetros

30

[0075] En una primera etapa, se tienen que identificar aquellas SFB, donde una discordancia de tonalidad podría provocar fallos en la banda de ruido. A tal fin, se tienen que determinar la tonalidad en cada SFB del intervalo de IGF y las bandas correspondientes que se usan para copiar. Una medición adecuada para calcular la tonalidad es la medición de aplanamiento espectral (SFM) que se basa en una división de la media geométrica de un espectro mediante su media aritmética y está comprendida entre 0 y 1. Los valores cercanos a 0 indican una fuerte tonalidad mientras que un valor que se aproxima a 1 es un signo de un espectro muy ruidoso. La fórmula se proporciona como

35

$$sfm(P, b, e) = 2^{\left(\frac{1}{2}+p\right)} \left(\frac{1}{e-b} \left(1 + \sum_{sb=b}^{e-1} P(sb) \right) \right)^{-1}$$

donde P es el espectro de energía TCX, b es la línea de inicio y e es la línea de detención de la SFB actual mientras que p se define como

40

$$p = \frac{1}{e-b} \sum_{sb=b}^{e-1} \lfloor \max(0, \log_2(P(sb))) \rfloor$$

[0076] Además de la SFM, se calcula el factor de cresta el cual proporciona también una indicación de cómo está distribuida la energía dentro de un espectro al dividir la energía máxima por la energía media de todos los intervalos de frecuencia en el espectro. La división de la SFM por el factor de cresta da como resultado una medición de la tonalidad de una SFB para la trama actual. El factor de cresta se calcula mediante

45

$$cresta(P, b, e) = \max \left(1, E_{\max} \left(\frac{1}{e-b} \sum_{sb=b}^{e-1} \lfloor \max(0, \log_2(P(sb))) \rfloor^2 \right)^{-\frac{1}{2}} \right)$$

50

donde P es el espectro de energía TCX, b es la línea de inicio y e es la línea de detención de la SFB actual mientras que $E_{\max}$ se define como

$$E_{\max} = \left\lfloor \max_{sb \in [b, e] \subset N} (0, \log_2(P(sb))) \right\rfloor$$

[0077] Sin embargo, resulta sensato usar también los resultados de tramas anteriores para lograr una estimación de tonalidad uniforme. Por lo tanto, la estimación de la tonalidad se realiza con la siguiente fórmula:

$$SFM = \min \left(2.7, \frac{sfm}{crest} + \frac{sfm_{prev}}{crest_{prev}} + 0.5 * SFM_{prev} \right),$$

donde sfm indica el resultado del cálculo real de aplanamiento espectral, mientras que la variable SFM incluye la división por el factor de cresta así como suavizado.

[0078] Ahora se calcula la diferencia de tonalidad entre la fuente y el destino:

$$SFM_{dif} = SFM_{src} - SFM_{dest}$$

[0079] Para los valores positivos de esta diferencia se cumple la condición de que algo que es más ruidoso que el espectro objetivo se use para copiado. Tal SFB se vuelve un probable candidato para amortiguamiento.

[0080] No obstante, un valor bajo de SFM no indica necesariamente una tonalidad fuerte sino que también se puede deber a un súbito declive o ascenso de la energía en una SFB. Esto se aplica particularmente a los elementos en los que existe una limitación de banda en alguna parte en la mitad de una SFB. Esto puede llevar a un amortiguamiento no deseado, creando la impresión de una señal de paso bajo ligeramente filtrada.

[0081] A fin de evitar el amortiguamiento en tales casos, las SFB posiblemente afectadas se determinan mediante el cálculo de la **inclinación espectral** de la energía en todas las bandas con SFM_{dif} positiva, donde una marcada inclinación en una dirección podría indicar una súbita caída que provoca un bajo valor SFM. La inclinación espectral se calcula como una regresión lineal a través de todos los intervalos espectrales en la SFB, con la pendiente de la línea de regresión dada por la siguiente fórmula:

$$pendiente = \left(\sum_i x_i P_i - \frac{1}{e-b} \sum_i x_i \sum_i P_i \right) \left(\sum_i x_i^2 - \frac{1}{e-b} \left(\sum_i x_i \right)^2 \right)^{-1}, \quad i = b, \dots, e-1$$

con x como el número de intervalo, P el espectro de energía TCX, b la línea de inicio y e la línea de detención de la SFB actual.

[0082] No obstante, un componente tonal cercano a un borde de una SFB podría provocar también una inclinación empinada, pero aun así deberá someterse a amortiguamiento. Para separar estos dos casos, se deberá realizar otro cálculo de SFM desplazada para las bandas con inclinación empinada.

[0083] El umbral para el valor de la pendiente se define como

$$umbral_{incl} = \frac{60}{e-b}$$

con la división mediante el ancho de SFB como normalización.

[0084] Si existe un marcado declive pendiente $< -umbral_{incl}$, la región de la frecuencia para la cual se calcula SFM se desplazará hacia abajo la mitad del ancho de la SFB; para un marcado ascenso pendiente $> umbral_{incl}$ se desplazará hacia arriba. De esta forma, los componentes tonales que se supone están amortiguados, aún se pueden detectar correctamente debido a una baja SFM mientras que para valores SFM más altos, no se aplicará amortiguamiento. Aquí el umbral se define como el valor 0,04, donde el amortiguamiento solo se aplica si la SFM desplazada cae por debajo del umbral.

Modelo de molestia perceptual

[0085] No se deberá aplicar amortiguamiento para ninguna SFM_{dif} positiva, pero solo tiene sentido si la SFB objetivo es, en efecto, muy tonal. Si en una SFB específica la señal original está superpuesta con una señal de fondo ruidosa, entonces la diferencia perceptual con una banda aún más ruidosa será pequeña y la monotonía debida a la pérdida de energía por amortiguamiento puede tener más peso que los beneficios.

[0086] Para asegurar la aplicación dentro de límites razonables, el amortiguamiento deberá usarse solo si la

SFB objetivo es, en efecto, muy tonal. Por lo tanto, solo cuando ambas

$$SFM_{dif} > 0$$

5 y

$$SFM_{dest} < 0.1$$

se mantienen, se deberá aplicar amortiguamiento.

10

[0087] Otra materia que deberá considerarse es el fondo de componentes tonales en el espectro IGF. Es probable que la degradación perceptual provocada por fallos en bandas de ruido sea muy evidente siempre que haya escaso o ningún fondo del tipo ruido rodeando al componente tonal original. En este caso, cuando se compara el original con el espectro HF creado por IGF, se percibirá una banda de ruido introducida como algo completamente nuevo y por lo tanto sobresaldrá de manera muy prominente. Si, por otra parte, ya hay una considerable cantidad de ruido de fondo existente, entonces el ruido adicional se combina con el fondo obteniéndose una diferencia perceptual menos estridente. Por lo tanto, la cantidad de amortiguamiento aplicada también dependerá de la proporción tonal/ruido en la SFB afectada.

15

20 **[0088]**

Para el cálculo de esta proporción tonal/ruido se suman los valores espectrales de potencia TCX P al cuadrado de todos los intervalos i en una SFB y se dividen por el ancho de la SFB (dado por la línea de inicio b y la línea de detención e) para obtener la energía promedio de la banda. Este promedio posteriormente se usa para normalizar todas las energías en la banda.

25

$$P_{norm,k} = \sqrt{P_k} \left(\frac{1}{e-b} * \sum_{sb=b}^{e-1} \sqrt{P_{sb}} \right)^{-1}, \quad k = b, \dots, e-1$$

Todos los intervalos con una energía $P_{norm,k}$ normalizada por debajo de 1 se suman, a continuación, y se cuentan como la parte ruidosa P_{ruido} mientras que todo lo que esté por encima del umbral de $1 + adap$ con

30

$$adap = \frac{e-b}{40}$$

se cuenta como la parte tonal P_{tonal} . Este umbral depende del ancho de la SFB de modo que bandas más pequeñas obtienen un umbral más bajo para responder por el promedio más alto debido a la mayor influencia de los intervalos con alta energía del componente tonal. Finalmente, de la parte tonal y de ruido se calcula un logaritmo de la proporción.

35

$$tonal \ a \ ruido = 20 * \log_{10} \left(\frac{P_{tonal}}{P_{ruido}} \right)$$

[0089]

El amortiguamiento depende tanto de la diferencia en SFM entre la fuente y el destino y la SFM de la SFB objetivo donde diferencias mayores y una SFM objetivo más pequeña deberán conducir a un amortiguamiento más marcado. Es razonable que para una diferencia mayor en tonalidad se deba aplicar un amortiguamiento más marcado. Asimismo, la cantidad de amortiguamiento deberá aumentar también más rápidamente si la SFM objetivo es más baja, es decir, la SFB objetivo más tonal. Esto significa que para las SFB extremadamente tonales se aplicará un amortiguamiento más marcado que para las SFB donde la SFM cae justo dentro del intervalo de amortiguamiento.

45 **[0090]**

Además, también se deberá aplicar amortiguamiento más moderado para frecuencias más altas dado que quitar la energía en las bandas más altas podría llevar fácilmente a la impresión perceptual de limitación de banda mientras que la estructura fina de las SFB se vuelve menos importante debido a la sensibilidad decreciente del sistema auditivo humano hacia las frecuencias más altas.

50 **[0091]**

Compensación de discordancia de tonalidad: Cálculo del factor de amortiguamiento

[0092]

Para incorporar todas estas consideraciones en una sola fórmula de amortiguamiento, la proporción entre la SFM objetivo y fuente se toma como la base de la fórmula. De esta manera, una mayor diferencia absoluta en la SFM y un valor más pequeño de SFM objetivo llevarán a un amortiguamiento más pronunciado, lo cual lo hace más adecuado que simplemente tomar la diferencia. Para añadir también dependencias sobre frecuencia y proporción tonal/ruido, se aplican a esta proporción parámetros de ajuste. Por lo tanto, la fórmula de amortiguamiento se puede escribir como

55

$$d_{curr} = \left(\frac{SFM_{dest}}{SFM_{src}} \right)^\alpha + \beta,$$

donde d es el factor de amortiguamiento que se multiplicará por el factor de escala y α y β los parámetros de ajuste de amortiguamiento que se calculan como

$$\alpha = \min\left(\frac{320}{e-1}, 1,25\right)$$

donde e es la línea de detención de la SFB actual y

$$\beta = \begin{cases} 0,1 * ((10 + adap) - tonalaruido), & \text{if } ((10 + adap) - tonalaruido) > 0 \\ 0, & \text{si no} \end{cases}$$

donde $adap$ depende del ancho de la SFB calculado mediante

$$adap = \frac{ancho}{40}$$

[0093] El parámetro α disminuye con la frecuencia a fin de aplicar menos amortiguamiento para altas frecuencias mientras que β se usa para reducir todavía más la fuerza del amortiguamiento si la proporción tonal/ruido de la SFB que se va a amortiguar cae por debajo de un umbral. Cuanto más cae por debajo de este umbral, mayor es la reducción del amortiguamiento.

[0094] Dado que el amortiguamiento solo se activa dentro de ciertas limitaciones, es necesario aplicar suavizado con el fin de prevenir abruptas transiciones de encendido/apagado. Para realizar esto, están activos diversos mecanismos de suavizado.

[0095] Directamente después de un transitorio, un conmutador de núcleo a TCX o un amortiguamiento de trama anterior no amortiguada se aplica solo gradualmente con fuerza plena para evitar caídas de energía extremas después de transitorios de alta energía. Asimismo, se utiliza un factor de olvido en la forma de un filtro IIR para tener en cuenta también los resultados de tramas anteriores.

[0096] Todas las técnicas de suavizado están comprendidas en la siguiente fórmula:

$$d = \min\left(\frac{d_{actual} + d_{prev}}{2} + 0,1 * suavizado, 1\right),$$

donde d_{prev} es el factor de amortiguamiento de la trama anterior. Si el amortiguamiento no está activo en la trama anterior d_{prev} se sobrescribe con d_{actual} pero limitado a un mínimo de 0,1. El suavizado variable es un factor de suavizado adicional que se fijará en 2 durante las tramas transitorias (bandera *esTransitoria* activada) o después de que el núcleo cambie (bandera *esCelpaTCX* activada), a 1 si en la trama anterior el amortiguamiento estaba inactivado. En cada trama con amortiguamiento la variable disminuirá en 1, pero no puede caer por debajo de 0.

[0097] En la etapa final, el factor de amortiguamiento d se multiplica por la ganancia de escala g :

$$g_{amortiguado} = g * d$$

[0098] La Fig. 10 ilustra una implementación preferida de la presente invención.

[0099] La señal de audio, tal como, por ejemplo, emitida por el analizador espectral 130 está disponible como un espectro MDCT o inclusive un espectro complejo como se indica mediante (c) a la izquierda de la Fig. 10

[0100] El analizador de señal 121 se implementa mediante los detectores de tonalidad 801 y 802 en la Fig. 10 para detectar la tonalidad del contenido objetivo mediante el bloque 802 y para detectar la tonalidad del contenido fuente (simulado) en el elemento 801.

[0101] A continuación, se realiza el cálculo del factor de amortiguamiento 124 para obtener el valor de compensación y, a continuación, el compensador 503 opera usando la información obtenida del elemento 501, 700-

706. El elemento 501 y el elemento 700-706 refleja la estimación de la envolvente desde el contenido objetivo y la estimación de la envolvente desde el contenido fuente simulado y el cálculo posterior del factor de escala como se ilustra, por ejemplo, en la Fig.7 en el elemento 700-706.

5 **[0102]** Por lo tanto, el vector de escala no compensado se introduce en el bloque 503 como valor 502 en analogía con lo que se ha expuesto en el contexto de la Fig. 5. Asimismo, en la Fig. 10 se ilustra un modelo de ruido 1000 como un bloque constructor separado, aunque este también se puede incluir directamente dentro del calculador del factor de amortiguamiento 124 como se ha expuesto en el contexto de la Fig. 4.

10 **[0103]** Asimismo, el codificador IGF paramétrico en la Fig. 10 que comprende además un estimador de blanqueo está configurado para calcular niveles de blanqueo como se expone, por ejemplo, en el elemento 5.3.3.2.11.6.4 "Codificación de niveles de blanqueo IGF". Particularmente, los niveles de blanqueo IGF se calculan y transmiten usando uno o dos bits por elemento de imagen. Esta información se introduce en el multiplexor de flujo de bits 140 también con el fin de obtener finalmente la información paramétrica IGF completa.

15 **[0104]** Asimismo, el bloque "dispersar espectro" que puede corresponder al bloque 130 con respecto a la determinación de las líneas espectrales que se van a codificar mediante el codificador de núcleo 110 se proporciona adicionalmente y se ilustra como un bloque separado 1020 en la Fig. 10. Esta información es usada preferentemente por el compensador 503 con el fin de reflejar la situación IGF específica.

20 **[0105]** Asimismo, el término "simulado" a la izquierda del bloque 801 y el bloque "estimación de la envolvente" en el bloque de la Fig. 10 se refiere a la situación ilustrada en la Fig. 8a, donde el "contenido fuente simulado" es la información de audio codificada y decodificada de nuevo en la primera banda espectral.

25 **[0106]** Por otra parte, el contenido fuente "simulado" es la información obtenida por el simulador de parche 804 de la primera información de audio original en la primera banda espectral como se indica mediante la línea 180 o es la primera banda espectral decodificada como se obtiene mediante el decodificador de núcleo 800 enriquecida con las líneas desplazadas desde la segunda banda espectral hacia la primera banda espectral.

30 **[0107]** A continuación se ilustra una realización adicional de la invención que constituye una versión modificada de un códec 3gpp TS 26.445. A continuación se proporciona texto recientemente añadido que especifica el procesamiento de la invención. En esta invención, se hace referencia explícita a ciertas subcláusulas ya contenidas en la especificación 3gpp TS 26.445.

35 5.3.3.2.11.1.9 La función de inclinación espectral SLOPE

[0108] Dado $P \in P^n$ que es el espectro de energía TCX calculado según la subcláusula 5.3.3.2.11.1.2 y b la línea de inicio y e la línea de detención del intervalo de medición de la inclinación espectral.

40 **[0109]** La función *SLOPE*, aplicada con IGF, se define con:

$$SLOPE: P^n \times N \times N \rightarrow P,$$

$$SLOPE(P, b, e) = \left(\sum_{sb=b}^{e-1} x(sb)P(sb) - \frac{1}{e-b} \sum_{sb=b}^{e-1} x(sb) \sum_{sb=b}^{e-1} P(sb) \right) \left(\sum_{sb=b}^{e-1} x(sb)^2 - \frac{1}{e-b} \left(\sum_{sb=b}^{e-1} x(sb) \right)^2 \right)^{-1},$$

donde n es la longitud real de la ventana TCX y x el número de intervalo.

45 5.3.3.2.11.1.10. La función de la proporción tonal/ruido *TNR*

[0110] Dado $P \in P^n$ que es el espectro de energía TCX calculado según la subcláusula 5.3.3.2.11.1.2 y b la línea de inicio y e la línea de detención del intervalo de medición de la proporción tonal/ruido.

50 **[0111]** La función *TNR*, aplicada con IGF, se define con:

$$TNR: P^n \times N \times N \rightarrow P,$$

$$TNR(P, b, e) = 20 * \log_{10} \left(\sum_{\substack{sb=b \\ P_{norm}(sb) > 1+adap}}^{e-1} P_{norm}(sb) \left(\sum_{\substack{sb=b \\ P_{norm}(sb) < 1}}^{e-1} P_{norm}(sb) \right)^{-1} \right),$$

donde n es la longitud real de la ventana TCX, $P_{norm}(sb)$ se define con

$$P_{norm}(sb) = \sqrt{P(sb)} \left(\frac{1}{e-b} \sum_{i=b}^{e-1} \sqrt{P_i} \right)^{-1}$$

y $adap$ se define con

$$adap = \frac{e-b}{40}.$$

Amortiguamiento:

[0112] Para el cálculo del factor de amortiguamiento IGF se necesitan 6 matrices estáticas (*prevObjetivoFIR*, *prevSrcFIR*, *prevObjetivoIIR* y *prevSrcIIR* para el cálculo de SFM en intervalos objetivo y fuente así como *prevAmort* y *amortiguamientoSuavizado*), todas de tamaño nB para mantener estados de filtro sobre las tramas. Además se necesita una bandera estática *eraTransitorio* para guardar la información de la bandera de entrada *esTransitorio* de la trama anterior.

Restablecimiento de estados de filtro

[0113] Todos los vectores *prevObjetivoFIR*, *prevSrcFIR*, *prevObjetivoIIR*, *prevSrcIIR*, y *prevAmortiguamiento* y *amortiguamientoSuavizado* son matrices estáticas de tamaño nB en el módulo IGF y se inicializan como sigue:

$$\text{para } k = 0, 1, \dots, nB - 1 \quad \left\{ \begin{array}{l} prevObjetivoFIR(k) := 0 \\ prevSrcFIR(k) := 0 \\ prevObjetivoIIR(k) := 0 \\ prevSrcIIR(k) := 0 \\ prevAmortiguamiento(k) := -1 \\ amortSuavizado(k) := 2 \end{array} \right.$$

[0114] Esta inicialización se deberá hacer

- Con códec de arranque
- Con cualquier conmutador de tasa de bits
- Con cualquier conmutador tipo códec
- Con transición desde CELP a TCX, por ejemplo, *esCelpaTCX* = verdadero
- Si la trama actual tiene propiedades transitorias, por ejemplo, *esTransitorio* = verdadero
- Si el espectro de energía TCX P no está disponible

Cálculo del factor de amortiguamiento

[0115] Si el espectro de energía TCX P está disponible y *esTransitorio* es falso, calcular

$$tmpObjetivo(k) := \frac{SFM(P, t(k), t(k+1))}{CRESTA(P, t(k), t(k+1))}, k = 0, 1, \dots, nB - 1$$

40 y

$$\text{tmpSrc}(k) := \frac{\text{SFM}\left(P, m(t(k)), m(t(k+1))\right)}{\text{CRESTA}\left(P, m(t(k)), m(t(k+1))\right)}, k = 0, 1, \dots, nB - 1,$$

donde $t(0), t(1), \dots, t(nB)$ ya estará mapeado con la función tF , véase la subcláusula 5.3.3.2.11.1.1, $m: N \rightarrow N$ es la función de mapeo que mapea el intervalo objetivo IGF en el intervalo fuente IGF descrito en la subcláusula 5.3.3.2.11.1.8 y nB es el número de bandas de factor de escala, véase la tabla 94. SFM es una función de medición de aplanamiento espectral, descrita en la subcláusula 5.3.3.2.11.1.3 y CRESTA es una función del factor cresta descrita en la subcláusula 5.3.3.2.11.1.4.

10 **[0116]** Si esCelpaTCX es verdadero o eraTransitorio es verdadero, establecer

$$\text{para } k = 0, 1, \dots, nB - 1 \quad \begin{cases} \text{prevObjetivoFIR}(k) := \text{tmpObjetivo}(k) \\ \text{preSrcFIR}(k) := \text{tmpSrc}(k) \\ \text{prevObjetivoIIR}(k) := \text{tmpObjetivo}(k) \\ \text{prevSrcFIR}(k) := \text{tmpSrc}(k) \end{cases}$$

[0117] Calcular:

$$15 \quad sObjetivo(k) := \min\left(2.7, \text{tmp}(k) + \text{prevObjetivoFIR}(k) + \frac{1}{2}\text{prevObjetivoIIR}(k)\right), k = 0, 1, \dots, nB - 1$$

y

$$20 \quad sSrc(k) := \min\left(2.7, \text{tmp}(k) + \text{prevSrcFIR}(k) + \frac{1}{2}\text{prevSrcIIR}(k)\right), k = 0, 1, \dots, nB - 1.$$

[0118] Con estos vectores calcular:

$$\text{difSFM}(k) := sSrc(k) - sObjetivo(k), k = 0, 1, \dots, nB - 1.$$

25 Si para $k = 0, 1, \dots, nB - 1$

$$\text{difSFM}(k) \leq 0,$$

o

$$sObjetivo(k) > 0.1,$$

Establecer

$$\text{prevAmort}(k) := -1$$

$$\text{amortSuavizado}(k) := 1$$

30 si no, calcular la inclinación espectral con la función SLOPE , descrita en la subcláusula 5.3.3.2.11.1.9:

$$\text{inclinación}(k) := \text{SLOPE}(P, t(k), t(k+1)), k = 0, 1, \dots, nB - 1.$$

35

Si para $k = 0, 1, \dots, nB - 1$

$$\text{inclinación}(k) < -\text{InclinaciónUmbral}$$

o si no, si

40 $\text{inclinación}(k) > \text{inclinaciónUmbral}$ y $k < nB - 1$,
donde inclinaciónUmbral se define como

$$45 \quad \text{inclinaciónUmbral} := \frac{60}{t(k+1) - t(k)},$$

calcular la SFM en un espectro desplazado:

$$sShift(k) := \frac{SFM(P, t(k) + desplazamiento, t(k+1) + desplazamiento)}{CRESTA(P, t(k) + desplazamiento, t(k+1) + desplazamiento)}$$

con desplazamiento definido como

$$5 \quad desplazamiento := \frac{\text{sgn}(\text{tilt}(k))}{2} (t(k+1) - t(k)).$$

Si

$$-inclinaciónUmbral \leq \text{tilt}(k) \leq inclinaciónUmbral$$

establecer

$$10 \quad sDesplazamiento(k) := 0.$$

Si para $k = 0, 1, \dots, nB - 1$

$$sDesplazamiento(k) > 0,04$$

establecer el factor de amortiguamiento de la trama actual *amortActual* en cero en la banda k:

$$amortActual(k) := 0.$$

15

[0119] De lo contrario, calcular *amortActual*(k) como sigue:

$$amortActual(k) := e^{\alpha(k) \ln\left(\frac{sObjetivo(k)}{sFuente(k)}\right)} + \beta(k),$$

20 donde α se define como

$$\alpha(k) := \min\left(\frac{320}{t(k+1)}, 1,25\right)$$

y β se define como

25

$$\beta(k) := \begin{cases} 10 + adap - TNR(P, t(k), t(k+1)), & 10 + adap - TNR(P, t(k), t(k+1)) > 0, \\ 0, & \text{si no} \end{cases}$$

donde *TNR* es la función de la proporción tonal/ruido como se describe en la subcláusula 5.3.3.2.11.1.10 y *adap* se define como

30

$$adap := \frac{t(k+1) - t(k)}{40}.$$

[0120] Si para $k = 0, 1, \dots, nB - 1$

$$amortPrevio(k) = -1,$$

35 Establecer

$$amortPrevio(k) := \max(amortActual(k), 0,1).$$

[0121] Calcular el vector de factores de amortiguamiento *d* de tamaño *nB*:

40

$$d(k) := \min\left(\frac{1}{2}(amortActual(k) + amortPrevio(k)) + 0,1 * amortSuavizado(k), 1\right).$$

[0122] Finalmente, si *esTransitorio* es falso y el espectro de energía *P* está disponible, actualizar los filtros

$$\text{para } k = 0, 1, \dots, nB - 1 \quad \begin{cases} \text{prevObjetivoFIR}(k) := \text{tmpObjetivo}(k) \\ \text{preSrcFIR}(k) := \text{tmpSrc}(k) \\ \text{prevObjetivoIIR}(k) := sObjetivo(k) \\ \text{prevSrcIIR}(k) := sSrc(k) \end{cases}$$

45 **[0123]** Los nombres de los valores/índices/parámetros en la porción anterior son similares a los correspondientes parámetros/índices/valores que se han expuesto a lo largo de toda la memoria descriptiva. Por

consiguiente, diversos resultados de ensayos de escucha se exponen en el contexto de la Fig. 11a a 11c.

[0124] Estos ensayos de escucha se realizaron mostrando el beneficio del amortiguamiento al comparar elementos que se codificaron con la posibilidad de amortiguamiento contra elementos que se codificaron sin amortiguamiento.

[0125] El primer resultado ilustrado en la Fig. 11a es un ensayo de comparación A-B con una tasa de bits de 13,2 kbps y una velocidad de muestra de 32 kHz usando mono-elementos. Los resultados se muestran en la Fig. 11a y muestran el amortiguamiento del ensayo A-B frente a falta de amortiguamiento a 13,2 kbps.

[0126] El segundo ilustrado en la Fig. 11b fue un ensayo MUSHRA a 24,4 kbps y una velocidad de muestra de 32 kHz usando mono-elementos. En este caso, se compararon dos versiones sin amortiguamiento con la nueva versión con amortiguamiento. Los resultados se muestran en la Fig. 11b (puntuación absoluta) y la Fig. 11c (puntuación diferencial).

[0127] La señal de audio codificada con la invención se puede almacenar sobre un medio de almacenamiento digital o un medio de almacenamiento no transitorio o se puede transmitir sobre un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable tal como Internet.

[0128] Aunque se han descrito algunos aspectos en el contexto de un aparato, es claro que estos aspectos también representan una descripción del correspondiente procedimiento, donde un bloque o dispositivo corresponde a una etapa del procedimiento o a una característica de una etapa del procedimiento. De forma análoga, los aspectos descritos en el contexto de una etapa de un procedimiento representan una descripción de un bloque o un elemento o una característica correspondiente de un aparato correspondiente.

[0129] Dependiendo de ciertos requisitos para la implementación, las realizaciones de la invención se pueden implementar en un equipo físico o en un soporte lógico. La implementación se puede realizar usando un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un CD, una memoria ROM, PROM, EPROM, EEPROM o FLASH, que contenga señales de control legibles electrónicamente almacenadas, que cooperen (o sean capaces de cooperar) con un sistema informático programable de modo que se lleve a cabo el respectivo procedimiento.

[0130] Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable, de modo que se lleve a cabo uno de los procedimientos descritos en esta invención.

[0131] Generalmente, las realizaciones de la presente invención se pueden implementar como un producto de programa informático con un código de programa, el código de programa es operativo para realizar uno de los procedimientos cuando el programa informático se ejecuta en un ordenador. El código de programa se puede almacenar por ejemplo en un soporte legible por una máquina.

[0132] Otras realizaciones comprenden el programa informático para realizar uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por una máquina o un medio de almacenamiento no transitorio.

[0133] En otras palabras, una realización del procedimiento de la invención es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los procedimientos descritos en esta invención, cuando el programa informático se ejecuta en un ordenador.

[0134] Una realización adicional de los procedimientos de la invención es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en él, el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención.

[0135] Una realización adicional del procedimiento de la invención es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los procedimientos descritos en él. El flujo de datos o la secuencia de señales pueden, por ejemplo, estar configurados para ser transmitida mediante una conexión para comunicación de datos, por ejemplo, mediante Internet.

[0136] Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los procedimientos descritos en esta invención.

[0137] Una realización adicional comprende un ordenador que tiene instalado en sí el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención.

[0138] En algunas realizaciones, se puede usar un dispositivo lógico programable (por ejemplo una matriz de

puertas programables por campo) para realizar algunas o todas las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, una matriz de puertas programables por campo puede cooperar con un microprocesador con el fin de llevar a cabo uno de los procedimientos descritos en esta invención. Generalmente, los procedimientos se llevan a cabo preferentemente mediante cualquier equipo físico.

5

[0139] Las realizaciones descritas anteriormente son meramente ilustrativas de los principios de la presente invención. Se entiende que modificaciones y variaciones de las disposiciones y los detalles descritos en esta invención se tornarán evidentes a otros expertos en la materia. Por lo tanto, la intención es estar limitados solo por el alcance de las reivindicaciones de la patente próxima a concederse y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en esta invención.

10

REIVINDICACIONES

1. Aparato para codificar una señal de audio, que comprende:
 - 5 un codificador de núcleo (110) para codificar el núcleo de la primera información de audio en una primera banda espectral (180); un codificador paramétrico (120) para codificar paramétricamente la segunda información de audio en una segunda banda espectral (190) que es diferente de la primera banda espectral (180), donde el codificador paramétrico (120) comprende:
 - 10 un analizador (121) para analizar la primera información de audio en la primera banda espectral (180) para obtener un primer resultado de análisis (122) y para analizar la segunda información de audio en la segunda banda espectral (190) para obtener un segundo resultado de análisis (123);
 - un compensador (124) para calcular un valor de compensación (125) usando el primer resultado de análisis (122) y el segundo resultado de análisis (123); y
 - 15 un calculador de parámetros (126) para calcular un parámetro (170) de la segunda información de audio en la segunda banda espectral (190) usando el valor de compensación,donde el codificador paramétrico (120) está configurado para codificar paramétricamente una tercera información de audio en una tercera banda espectral (202);
 - 20 donde el analizador (121) está configurado para analizar la tercera información de audio en la tercera banda espectral (202) para obtener el tercer resultado de análisis (204);
 - donde el codificador paramétrico (120) comprende adicionalmente un detector de compensación (210) para detectar, usando al menos el tercer resultado de análisis (204), si se tiene que compensar o no la tercera banda espectral, y
 - 25 donde el calculador de parámetros (126) está configurado para calcular un parámetro adicional (200) de la información de audio en la tercera banda espectral (204) sin ningún valor de compensación, cuando el detector de compensación (210) detecta que la tercera banda espectral no se va a compensar.
2. Aparato de la reivindicación 1,
 - 30 donde el analizador (121) está configurado para calcular, como primer resultado de análisis (122), un primer valor cuantitativo, y como resultado del segundo análisis (123), un segundo valor cuantitativo,
 - donde el compensador (124) está configurado para calcular un valor de compensación cuantitativo (125) a partir del primer valor cuantitativo y del segundo valor cuantitativo, y
 - donde el calculador de parámetros (126) está configurado para calcular un parámetro cuantitativo utilizando el valor
 - 35 de compensación cuantitativo (125).
3. Aparato de una de las reivindicaciones anteriores,
 - donde el analizador (121) está configurado para analizar una primera característica de la primera información de audio para obtener el primer resultado de análisis (122) y analizar la misma primera característica de la segunda información
 - 40 de audio en la segunda banda espectral para obtener el segundo resultado de análisis (123); y
 - donde el calculador de parámetros (126) está configurado para calcular el parámetro (170) de la segunda información de audio en la segunda banda espectral mediante la evaluación de una segunda característica siendo la segunda característica diferente de la primera característica.
4. Aparato de la reivindicación 3,
 - 45 donde la primera característica es una característica de estructura fina espectral o una característica de distribución de energía dentro de la primera banda espectral (180); o
 - donde la segunda característica es una medición de envolvente o una medición relacionada con la energía o una medición relacionada con la potencia de valores espectrales dentro de la segunda banda espectral (190).
5. Aparato de una de las reivindicaciones anteriores,
 - 50 donde la primera banda espectral (180) y la segunda banda espectral (190) son mutuamente excluyentes entre sí;
 - donde el analizador (121) está configurado para calcular el primer resultado de análisis (122) sin usar la segunda información de audio en la segunda banda espectral (190), y para calcular el segundo resultado de análisis (123) sin
 - 55 usar la primera información de audio en la primera banda espectral (180).
6. El aparato de una de las reivindicaciones anteriores,
 - donde la señal de audio (100) comprende una secuencia de tiempo o tramas;
 - donde el compensador (124) está configurado para calcular un valor de compensación actual para una trama actual
 - 60 usando un valor de compensación anterior para una trama anterior.
7. Aparato de una de las reivindicaciones anteriores,
 - donde el codificador paramétrico (120) está configurado para codificar paramétricamente una tercera información de audio en una tercera banda espectral (202),
 - 65 donde la tercera banda espectral (202) comprende frecuencias más altas que la segunda banda espectral (190), y

donde el compensador (124) está configurado para usar un tercer valor de ponderación (d) para calcular el valor de compensación para la tercera banda espectral (202),
donde el tercer valor de ponderación es diferente de un segundo valor de ponderación usado para calcular el valor de compensación para la segunda banda espectral (190).

5

8. Aparato de una de las reivindicaciones anteriores,
donde el analizador (121) está configurado para calcular adicionalmente una proporción tonal/ruido (400) de la segunda información de audio en la segunda banda espectral; y donde el compensador (124) está configurado para calcular el valor de compensación dependiente de la proporción tonal/ruido (400) de la segunda información de audio,
10 donde se obtiene un primer valor de compensación para una primera proporción tonal/ruido y se obtiene un segundo valor de compensación para una segunda proporción tonal/ruido, donde el primer valor de compensación es mayor que el segundo valor de compensación, y la primera proporción tonal/ruido es mayor que la segunda proporción tonal/ruido.

15 9. Aparato de una de las reivindicaciones anteriores, donde el calculador de parámetros (126) está configurado para calcular un parámetro no compensado (502) de la segunda información de audio y para combinar (503) el parámetro no compensado (502) y el valor de compensación (125) para obtener el parámetro.

10. Aparato de una de las reivindicaciones anteriores,
20 que comprende además una interfaz de salida (140) para emitir información de audio codificada en el núcleo (160) en la primera banda espectral (180) y el parámetro (170).

11. Aparato de una de las reivindicaciones anteriores,
donde el compensador (124) está configurado para determinar el valor de compensación mediante la aplicación de un modelo psicoacústico (1002), donde el modelo psicoacústico (1000) está configurado para evaluar la discordancia psicoacústica entre la primera información de audio y la segunda información de audio usando el primer resultado de análisis (122) y el segundo resultado de análisis (123) para obtener el valor de compensación (125).
25

12. Aparato de una de las reivindicaciones anteriores,
30 donde la señal de audio (100) comprende una secuencia de tiempo o tramas; y
donde el analizador (121) está configurado para analizar la primera información de audio en la primera banda espectral (180) de una trama para obtener el primer resultado de análisis (122) y para analizar la segunda información de audio de la trama en la segunda banda espectral (190) para obtener un segundo resultado de análisis (123) para la trama,
donde el compensador (124) está configurado para calcular un valor de compensación (125) para la trama usando el
35 primer resultado de análisis (122) para la trama y el segundo resultado de análisis (123) para la trama; y donde el calculador de parámetros (126) está configurado para calcular el parámetro (170) de la segunda información de audio en la segunda banda espectral (190) de la trama usando el valor de compensación para la trama, o
donde el detector de compensación (210) está configurado para detectar, sobre la base del primer resultado de análisis (122) y el segundo resultado de análisis (123), si el parámetro para la segunda banda espectral de una trama se tiene
40 que calcular usando el valor de compensación (125) en una situación de compensación o en una situación de no compensación.

13. Aparato de la reivindicación 1-12, donde el detector de compensación (210) está configurado para detectar una situación de compensación, cuando una diferencia entre el primer resultado de análisis (122) y el
45 segundo resultado de análisis (123) tiene una característica predeterminada, o cuando el segundo resultado de análisis tiene una característica predeterminada,
donde el detector de compensación (210) está configurado para detectar que una banda espectral no se tiene que compensar, cuando un espectro de energía no está disponible para el aparato para codificación o cuando se detecta que una trama actual es una trama transitoria (223), o
50 donde el compensador (124) está configurado para calcular el valor de compensación (125) sobre la base de un cociente del primer resultado de análisis y el segundo resultado de análisis.

14. Aparato de una de las reivindicaciones anteriores,
donde el analizador (121) está configurado para calcular una medición de aplanamiento espectral, un factor de cresta
55 o un cociente de la medición de aplanamiento espectral y el factor de cresta para la primera banda espectral (180) como el primer resultado de análisis (122), y para calcular una medición de aplanamiento espectral o un factor de cresta o un cociente de la medición de aplanamiento espectral y el factor de cresta para la segunda banda espectral (190) como el segundo resultado de análisis (123), o
donde el calculador de parámetros (126) está configurado para calcular, a partir de la segunda información de audio,
60 una información de envolvente espectral o un factor de ganancia, o
donde el compensador (124) está configurado para calcular el valor de compensación (125) de modo que, para una primera diferencia entre el primer resultado de análisis (122) y el segundo resultado de análisis (123), se obtiene un primer valor de compensación, y para una diferencia entre el primer resultado de análisis (122) y el segundo resultado de análisis (123), se calcula un segundo valor de compensación, donde la primera diferencia es mayor que la segunda
65 diferencia, y donde el primer valor de compensación es mayor que el segundo valor de compensación.

15. Aparato de la reivindicación 14,
donde el analizador (121) está configurado para calcular una inclinación espectral (608) de la segunda información de audio,
- 5 donde el analizador (121) está configurado para examinar (616) si existe un componente tonal cercano a un borde de la segunda banda espectral, y
donde un detector de compensación (210) del codificador paramétrico (120) está configurado para determinar (612) que el parámetro se tiene que calcular usando el valor de compensación solo cuando la inclinación espectral está por debajo de un umbral predeterminado (610), o (618) cuando la inclinación espectral está por encima de un umbral predeterminado (614) y el examen (616) ha determinado que existe un componente tonal cercano al borde.
- 10 predeterminado (614) y el examen (616) ha determinado que existe un componente tonal cercano al borde.
16. Aparato de una de las reivindicaciones anteriores, que comprende además:
- 15 un decodificador (800) para decodificar la primera información de audio codificada en la primera banda espectral para obtener una primera información de audio codificada y decodificada,
donde el analizador (121) está configurado para calcular (801) el primer resultado de análisis usando la primera información de audio codificada y decodificada, y
para calcular (802) el segundo resultado de análisis de la segunda información de audio de la entrada de señal de audio en el aparato.
- 20
17. Aparato de una de las reivindicaciones anteriores, que comprende además:
- 25 un simulador de parche (804) para simular un resultado de emparchamiento para la segunda banda espectral, comprendiendo el resultado de emparchamiento al menos una línea espectral (351, 352) de la segunda banda espectral incluida en una señal de audio codificada en el núcleo;
donde el analizador (121) está configurado para calcular el primer resultado de análisis usando la primera información de audio y la al menos una línea espectral (351', 352') de la segunda banda espectral; y
para calcular el segundo resultado de análisis de la segunda información de audio de la entrada de señal de audio (100) en el aparato para su codificación.
- 30
18. Aparato de una de las reivindicaciones anteriores,
donde el codificador de núcleo (110) está configurado para codificar la primera información de audio en una secuencia de espectros reales valuados,
donde el analizador (121) está configurado para calcular el primer y el segundo resultado de análisis de una secuencia de espectros de energía,
- 35 donde un espectro de energía se calcula a partir de la entrada de señal de audio (100) en el aparato para su codificación o se deriva de un espectro real valuado usado por el codificador de núcleo.
19. Aparato de una de las reivindicaciones anteriores,
donde el codificador de núcleo (110) está configurado para codificar el núcleo de la señal de audio al menos en una banda núcleo que se extiende hasta una frecuencia inicial de mejora (310),
donde la banda núcleo comprende la primera banda espectral (180) y al menos una banda fuente adicional (302, 303) superponiéndose con la primera banda espectral (180),
donde la señal de audio comprende un intervalo de mejora que se extiende desde la frecuencia inicial de mejora (310)
- 40 hasta una frecuencia máxima (354), donde la segunda banda espectral (190) y al menos una banda objetivo adicional se incluyen en el intervalo de mejora, donde la segunda banda espectral (305, 307, 202) y la banda objetivo adicional no se superponen entre sí.
- 45
20. Aparato de la reivindicación 19,
donde la frecuencia inicial de mejora (310) es una frecuencia de cruce y una señal codificada en el núcleo está limitada en banda a la frecuencia de cruce (310), o
donde la frecuencia inicial de mejora (310) es una frecuencia inicial con cobertura inteligente de intervalos (IGF) y una señal codificada en el núcleo está limitada en banda a la frecuencia máxima (354) que es mayor que la frecuencia inicial de mejora.
- 50
21. Aparato de una de las reivindicaciones anteriores,
donde el calculador de parámetros (126) está configurado
- 55 para calcular un factor de ganancia para la segunda banda espectral sobre la base de la segunda información de audio en la segunda banda espectral,
para calcular un factor de amortiguamiento como el valor de compensación, y
para multiplicar el factor de ganancia para la banda mediante el factor de amortiguamiento para obtener un factor de ganancia compensado como el parámetro, y
- 60
- 65 donde el aparato comprende además una interfaz de salida (140) para emitir información de audio codificada en el

núcleo (160) en la primera banda espectral (180) y el factor de ganancia compensado como el parámetro (170).

22. Procedimiento para codificar una señal de audio, que comprende:

- 5 codificar el núcleo (110) de la primera información de audio en una primera banda espectral (180);
codificar paramétricamente (120) la segunda información de audio en una segunda banda espectral (190) que es
diferente de la primera banda espectral (180), donde la codificación paramétrica comprende:

10 analizar (121) la primera información de audio en la primera banda espectral (180) para obtener un primer
resultado de análisis (122) y para analizar la segunda información de audio en la segunda banda espectral
(190) para obtener un segundo resultado de análisis (123);
calcular (124) un valor de compensación (125) usando el primer resultado de análisis (122) y el segundo
resultado de análisis (123); y
15 calcular (126) un parámetro (170) de la segunda información de audio en la segunda banda espectral (190)
usando el valor de compensación (125),

donde la codificación paramétrica (120) comprende la codificación paramétrica de tercera información de audio en
una tercera banda espectral (202),
donde el análisis (121) comprende analizar la tercera información de audio en la tercera banda espectral (202)
20 para obtener un tercer resultado de análisis (204),
donde la codificación paramétrica (120) comprende además detectar, usando al menos el tercer resultado de
análisis (204), si la tercera banda espectral debe compensarse o no, y
donde el cálculo (126) comprende calcular un parámetro adicional (200) a partir de la tercera información de audio
en la tercera banda espectral (204) sin ningún valor de compensación, cuando la etapa de detección detecta que
25 la tercera banda espectral no debe compensarse.

23. Sistema para procesar una señal de audio, que comprende:

- un aparato para codificar una señal de audio de una de las reivindicaciones 1 a 21; y
30 un decodificador para recibir una señal de audio codificada que comprende una primera información de audio
codificada (160) en la primera banda espectral (180) y un parámetro (170) que representa una segunda información
de audio en la segunda banda espectral (190);
donde el decodificador está configurado para realizar una operación de mejora espectral con el fin de regenerar
información de audio sintetizada para la segunda banda espectral (190) usando el parámetro (170) y la primera
35 información de audio decodificada en la primera banda espectral (180).

24. Procedimiento para procesar una señal de audio, que comprende:

- codificar una señal de audio de acuerdo con la reivindicación 22; y
40 recibir una señal de audio codificada que comprende una primera información de audio codificada (160) en la
primera banda espectral (180) y un parámetro (170) que representa una segunda información de audio en la
segunda banda espectral (190); y
realizar una operación de mejora espectral con el fin de regenerar información de audio sintetizada para la segunda
banda espectral (190) usando el parámetro (170) y la primera información de audio decodificada en la primera
45 banda espectral (180).

25. Programa informático configurado para realizar, cuando se ejecuta en un ordenador o un procesador, el
procedimiento de la reivindicación 22 o 24.

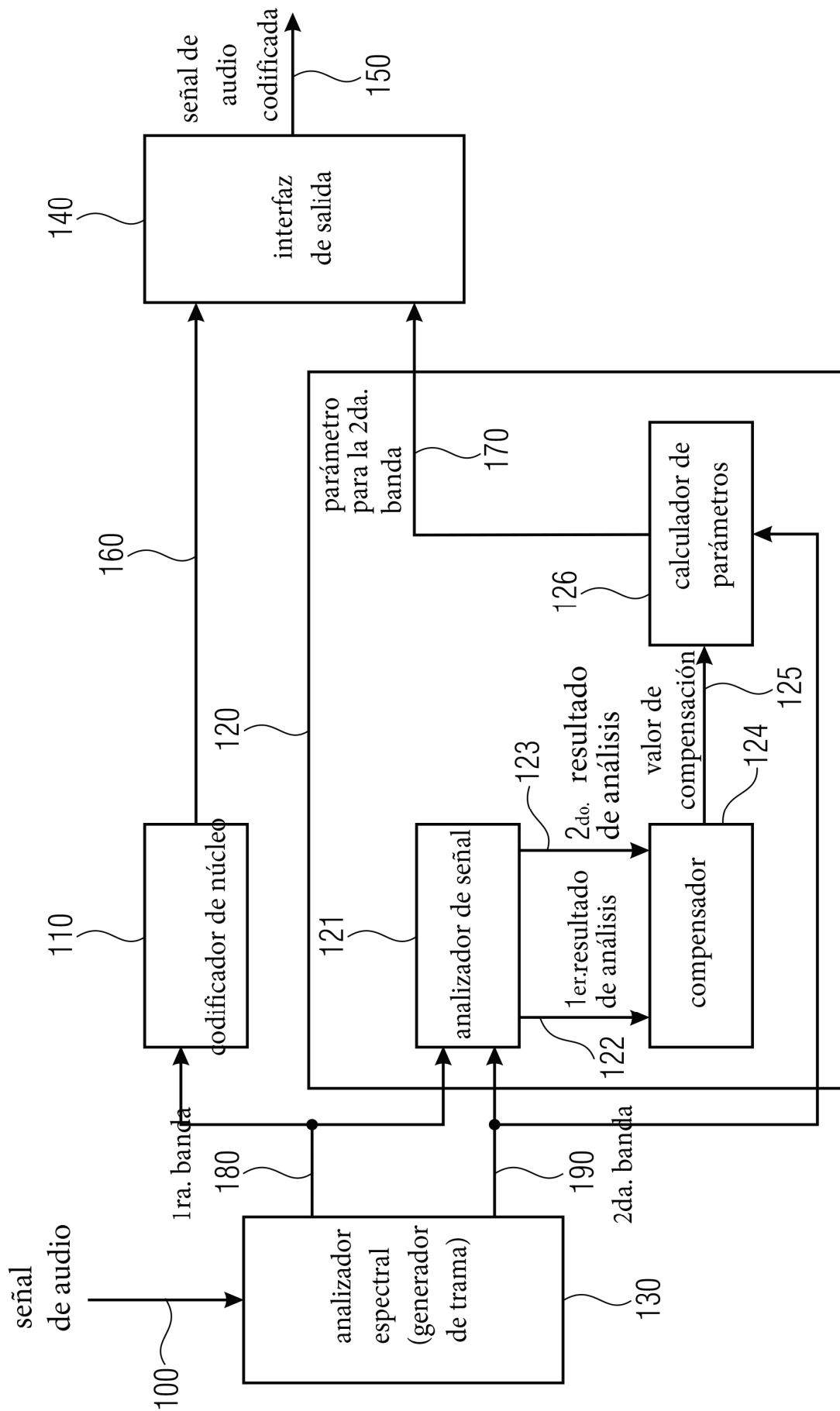


Fig. 1

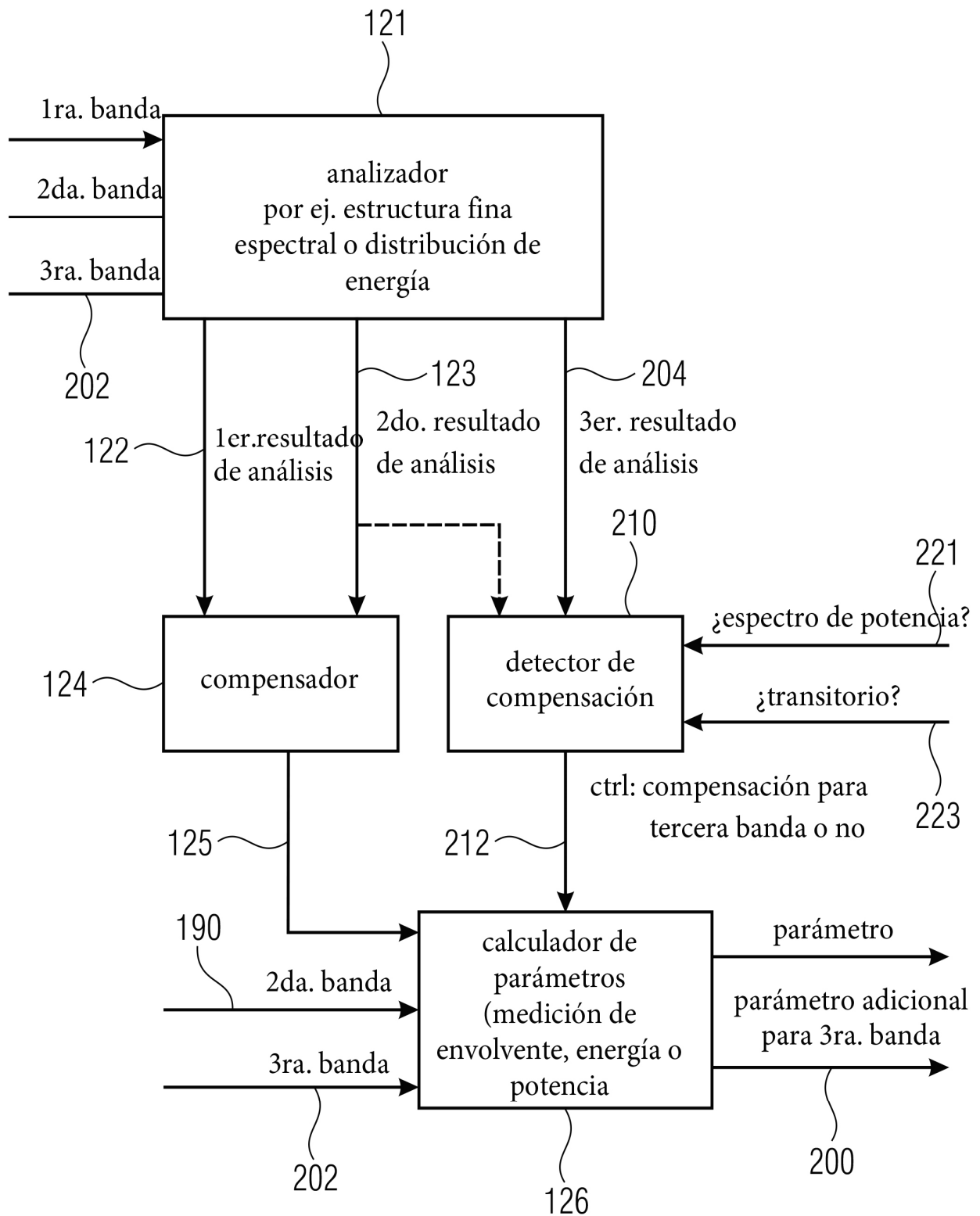
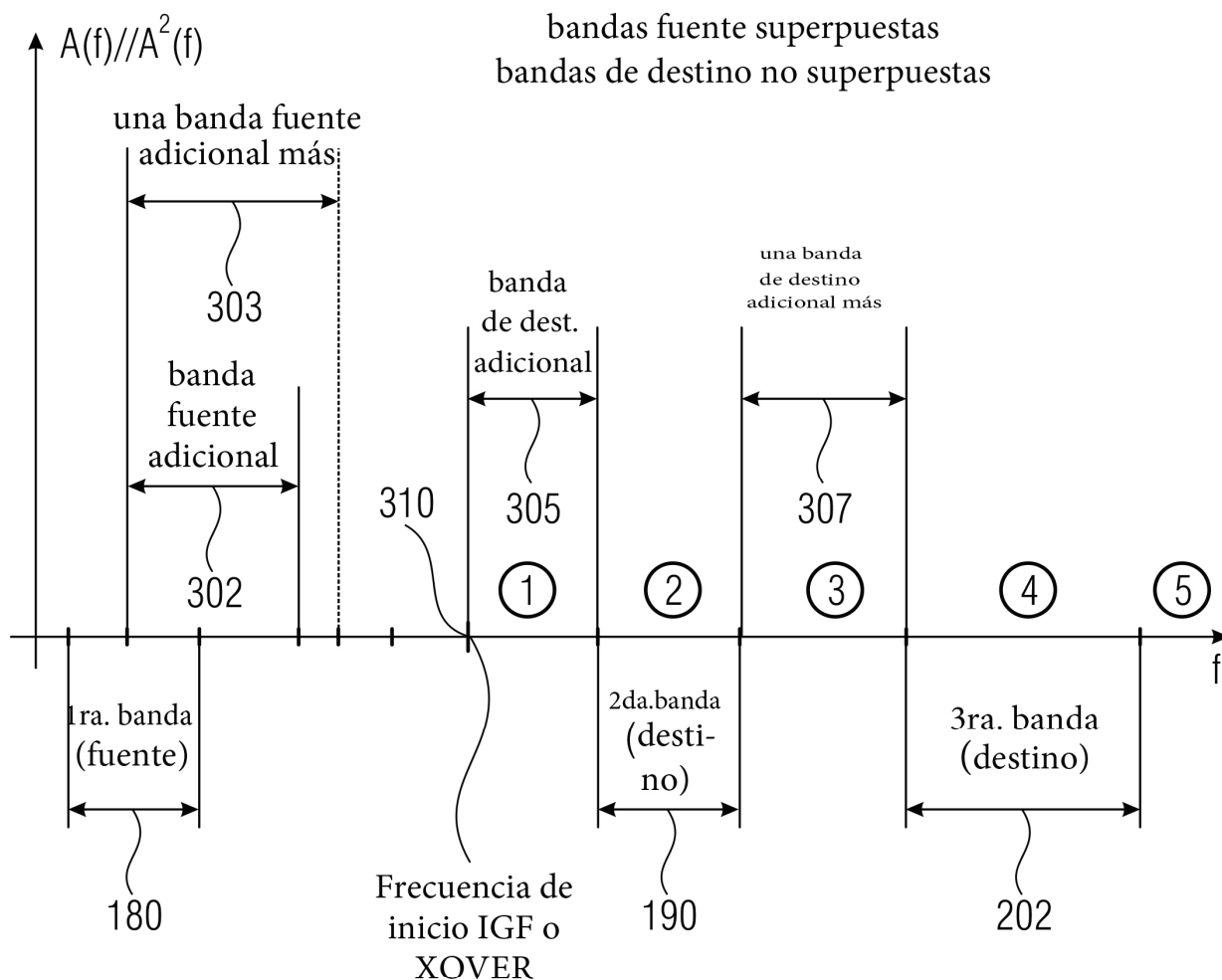


Fig. 2



Nro. de banda	banda de destino	para ser reconstruida en el decodificador mediante banda espectral
①	banda de destino adicional	1ra. banda
②	2da. banda	1ra. banda
③	una banda de destino adicional más	banda fuente adicional
④	3ra. banda	una banda fuente adicional más
⑤	...	...

Fig. 3a

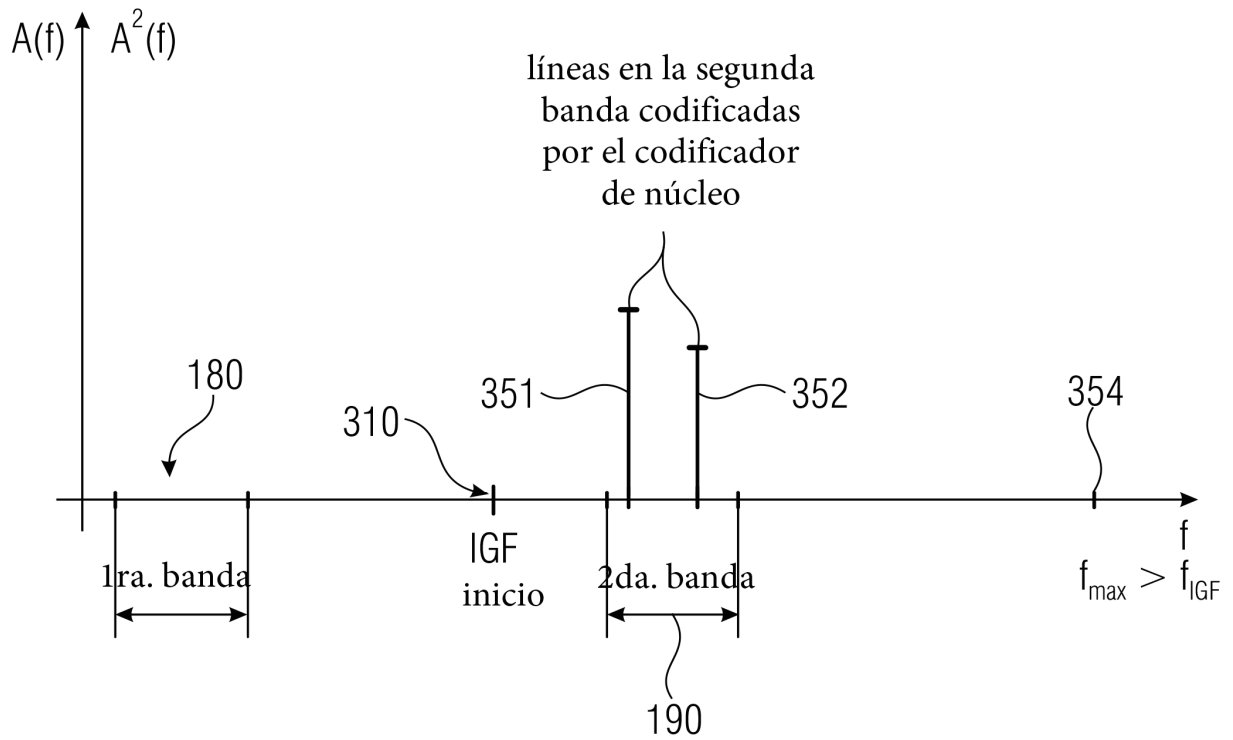


Fig. 3b

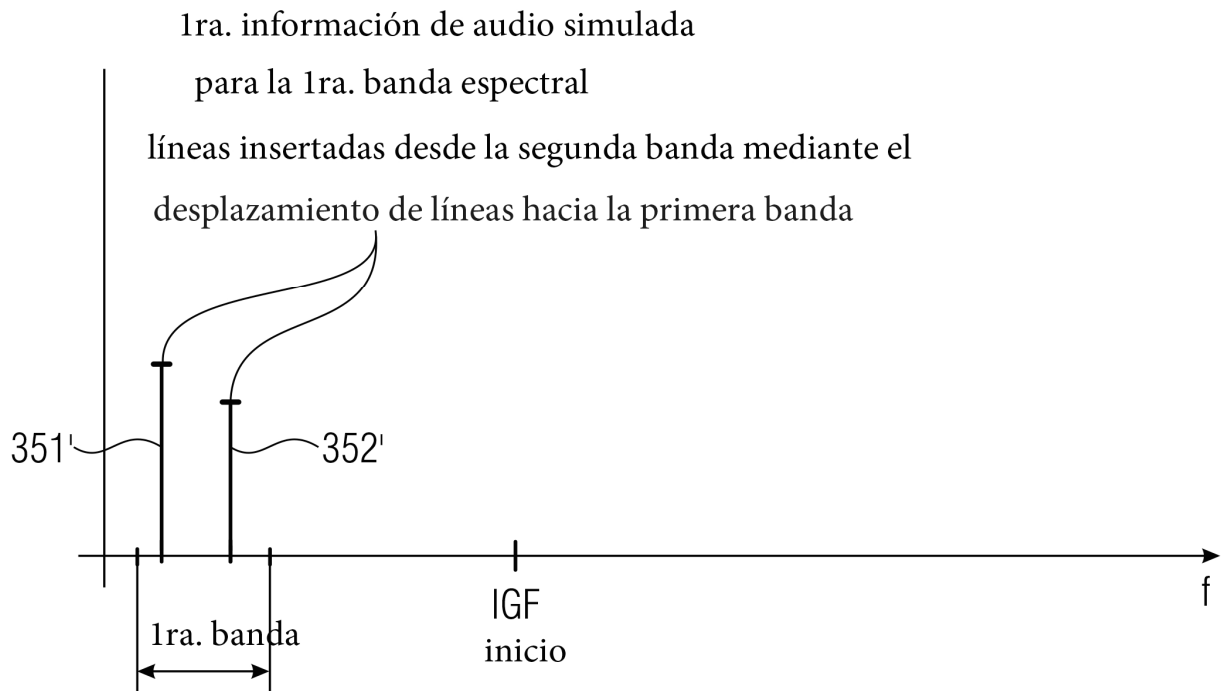


Fig. 3c

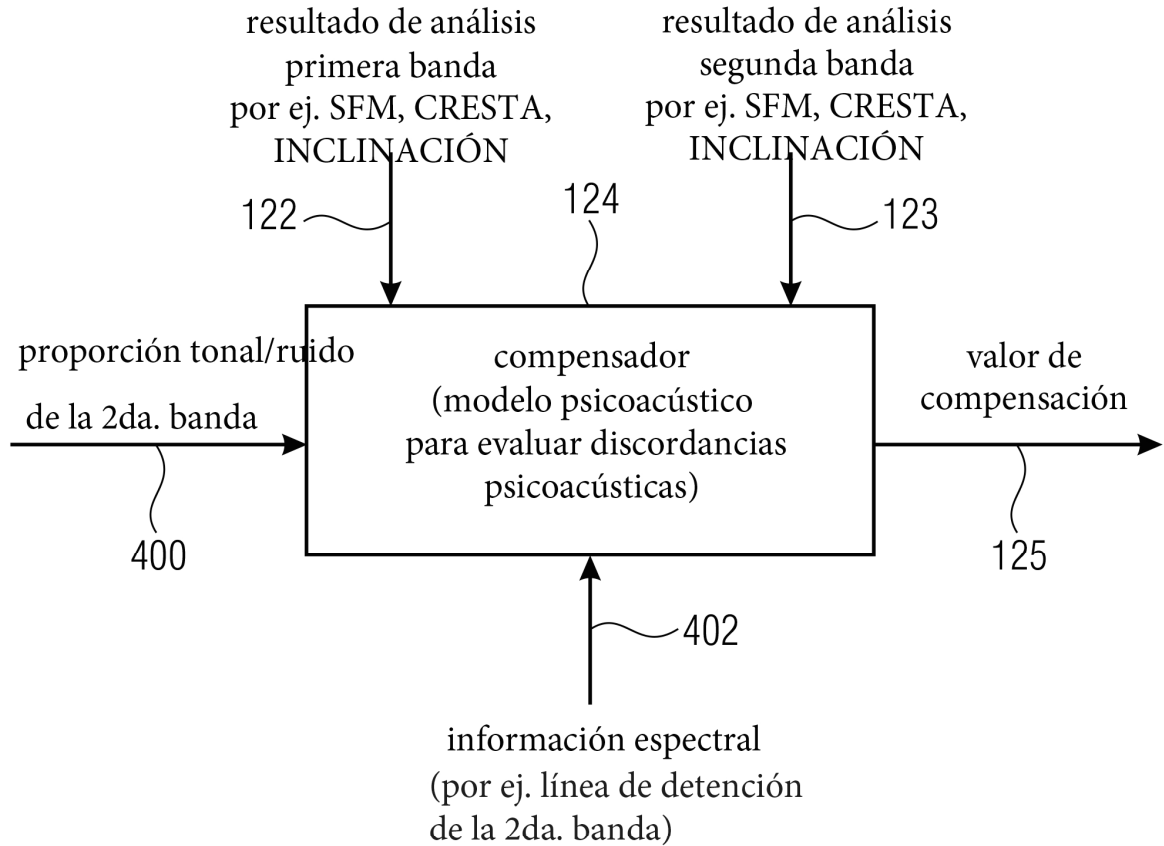


Fig. 4

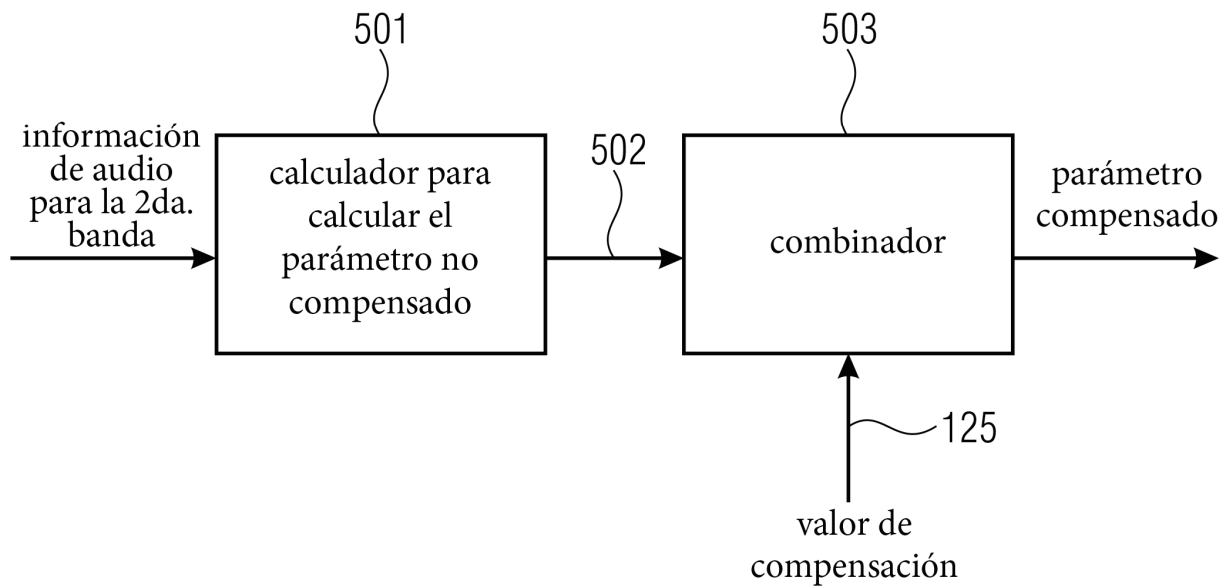


Fig. 5

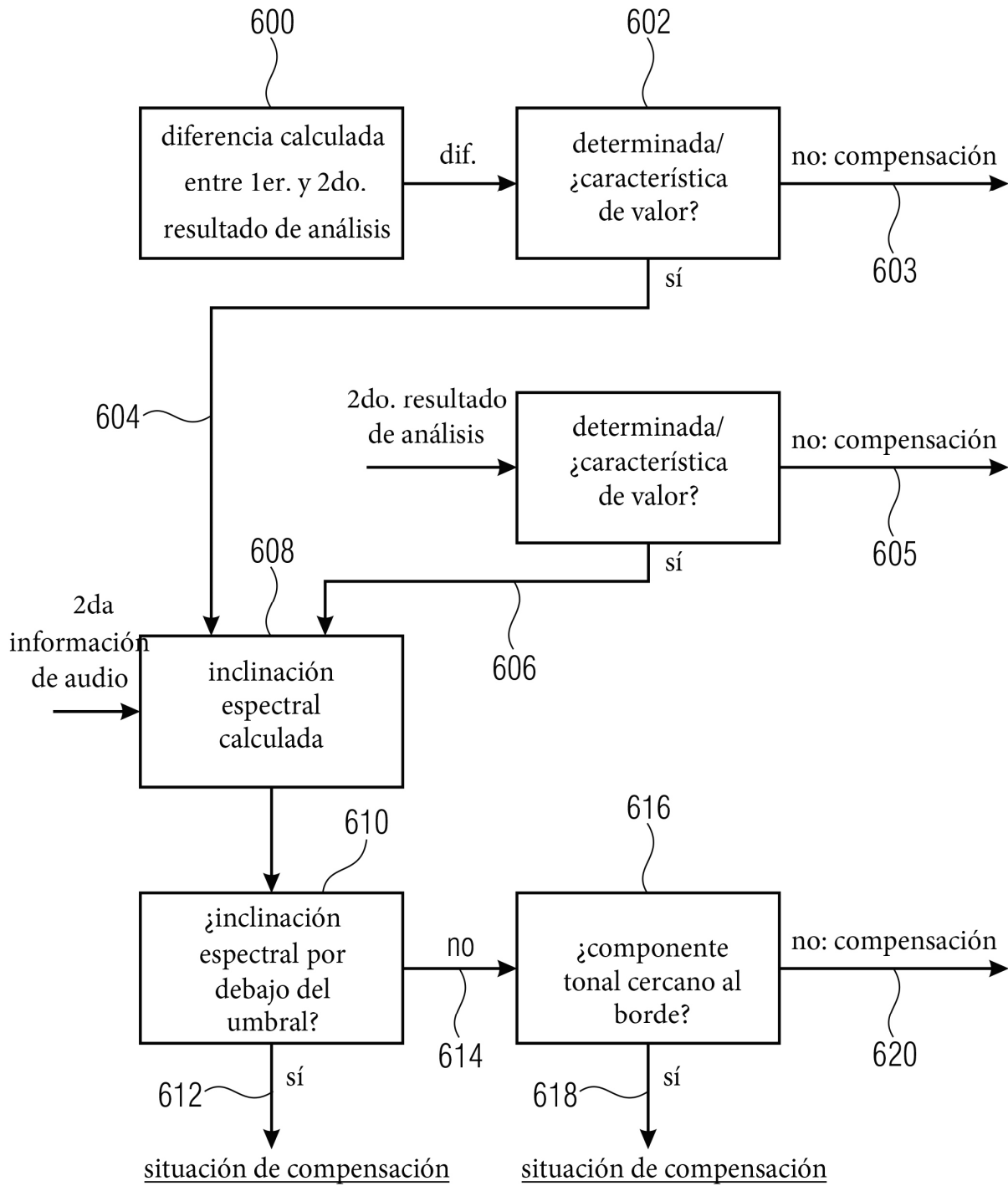
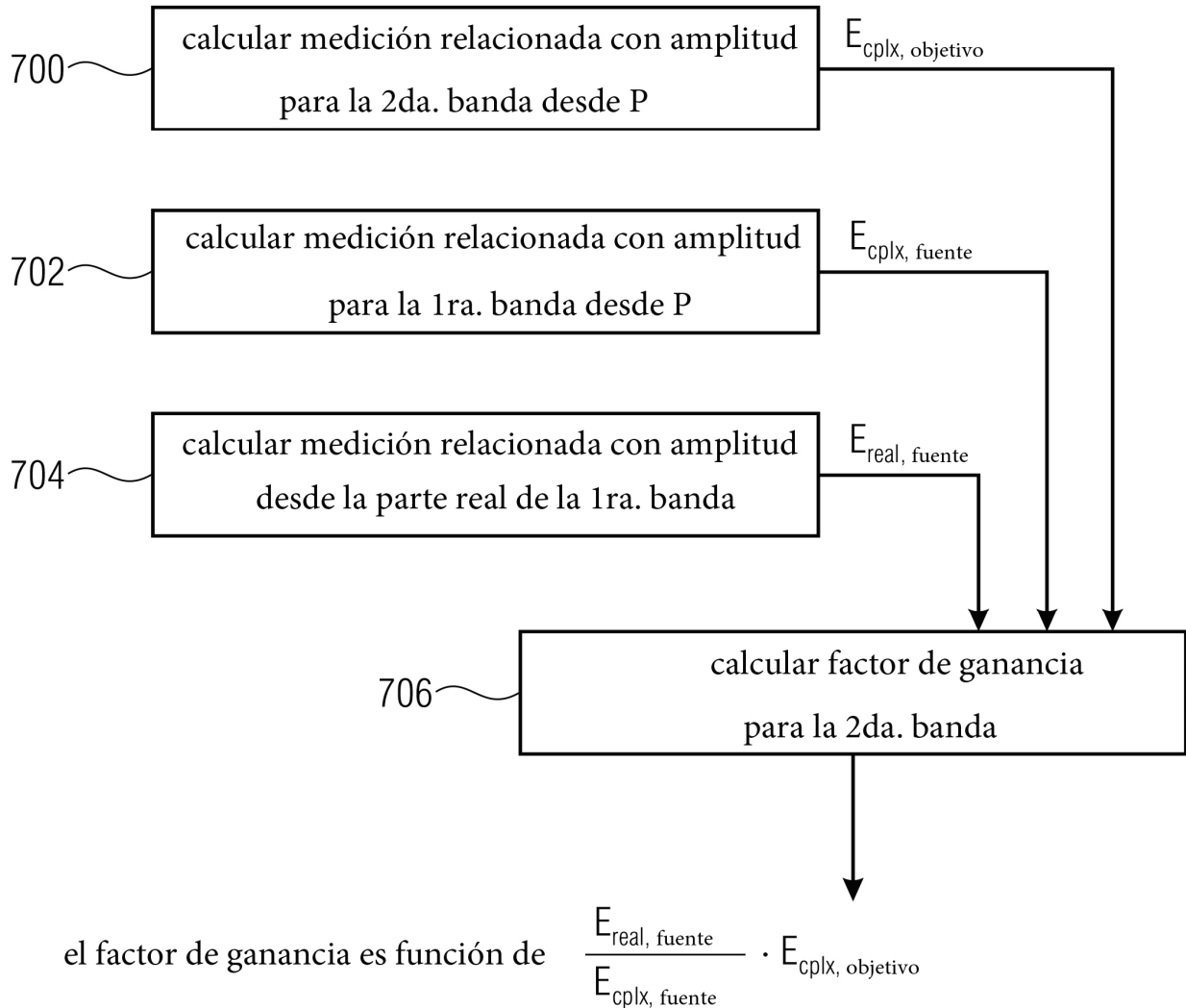
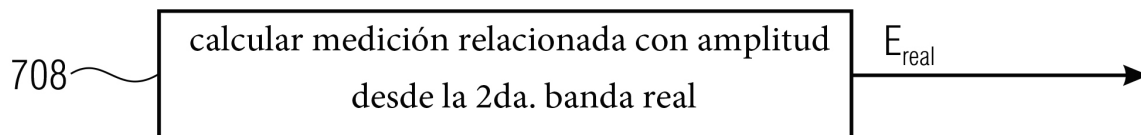


Fig. 6
(FUNCIONALIDAD DEL DETECTOR DE
COMPENSACIÓN)

- espectro de energía P TCX "complejo" disponible



- el espectro de energía TCX no está disponible



el factor de ganancia va en función de E_{real}

Fig. 7

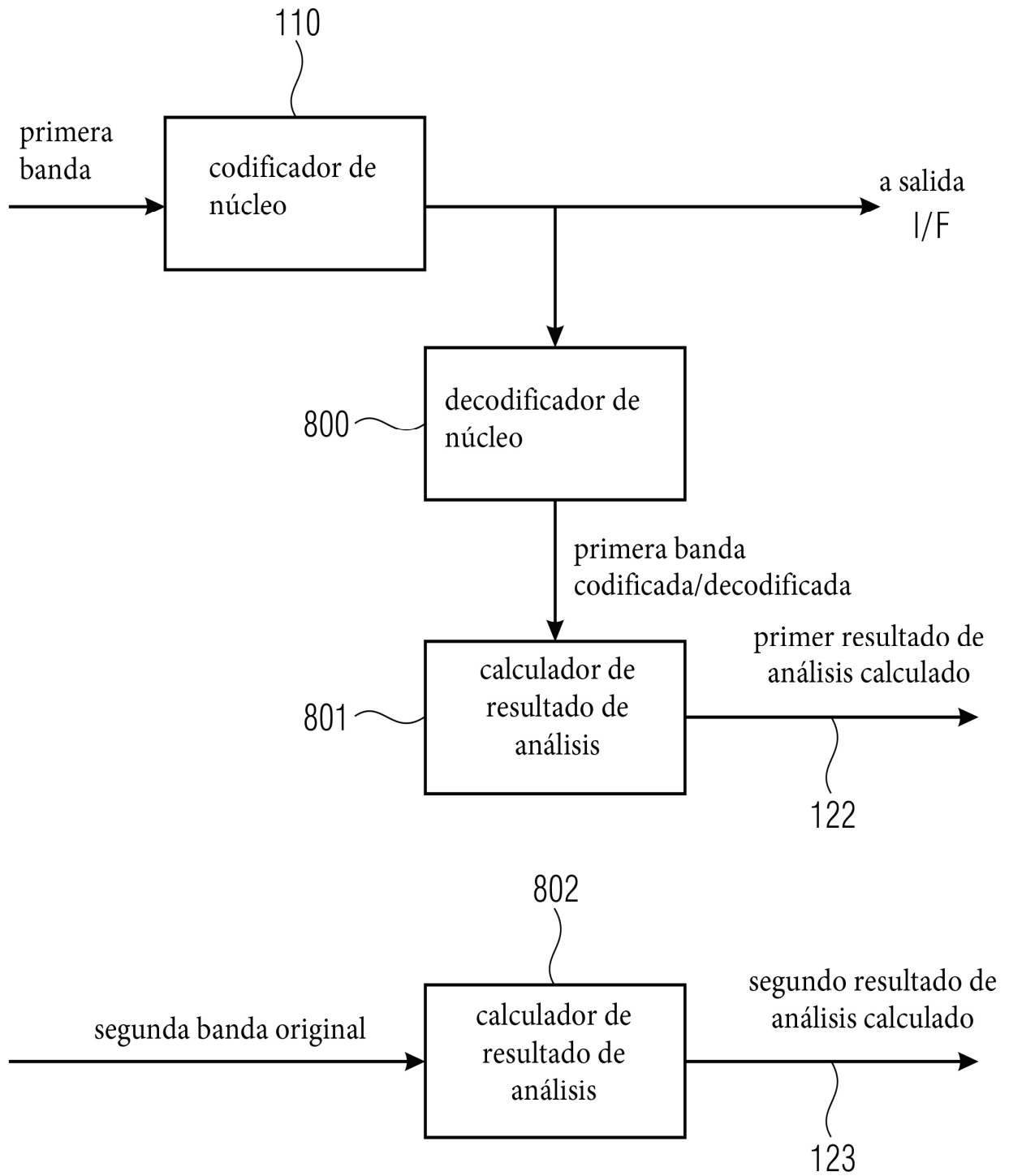


Fig. 8a

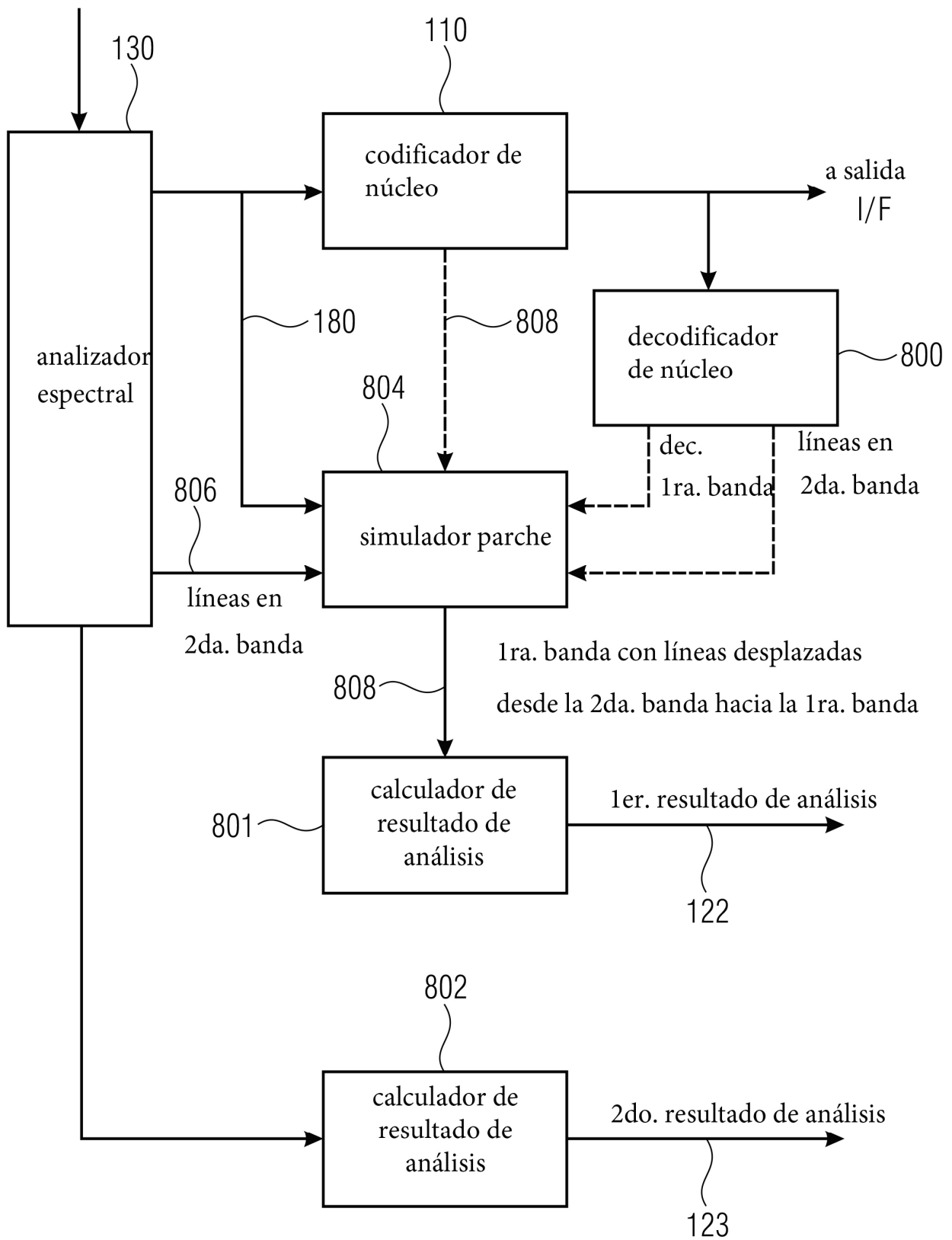


Fig. 8b

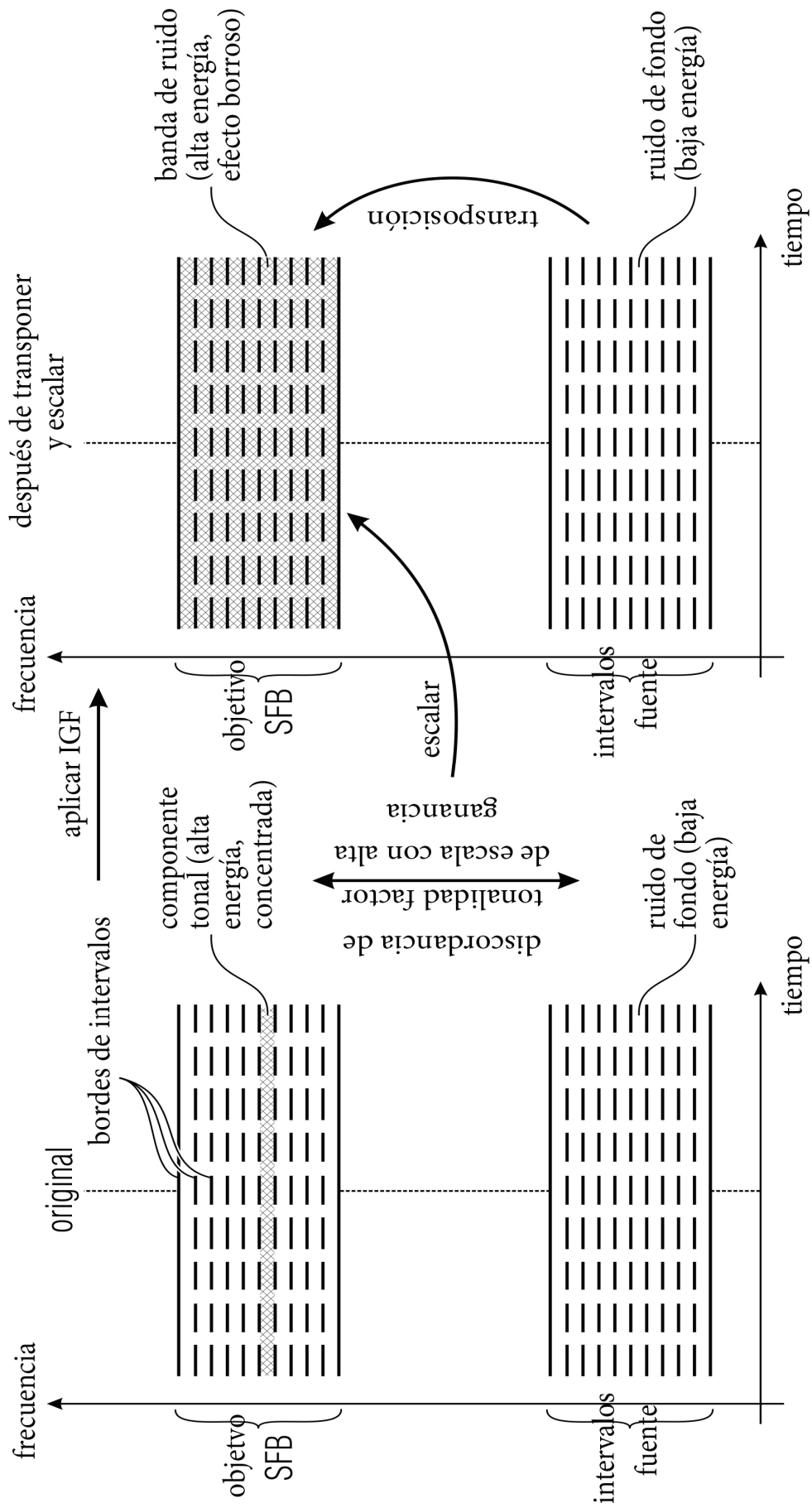


Fig. 9

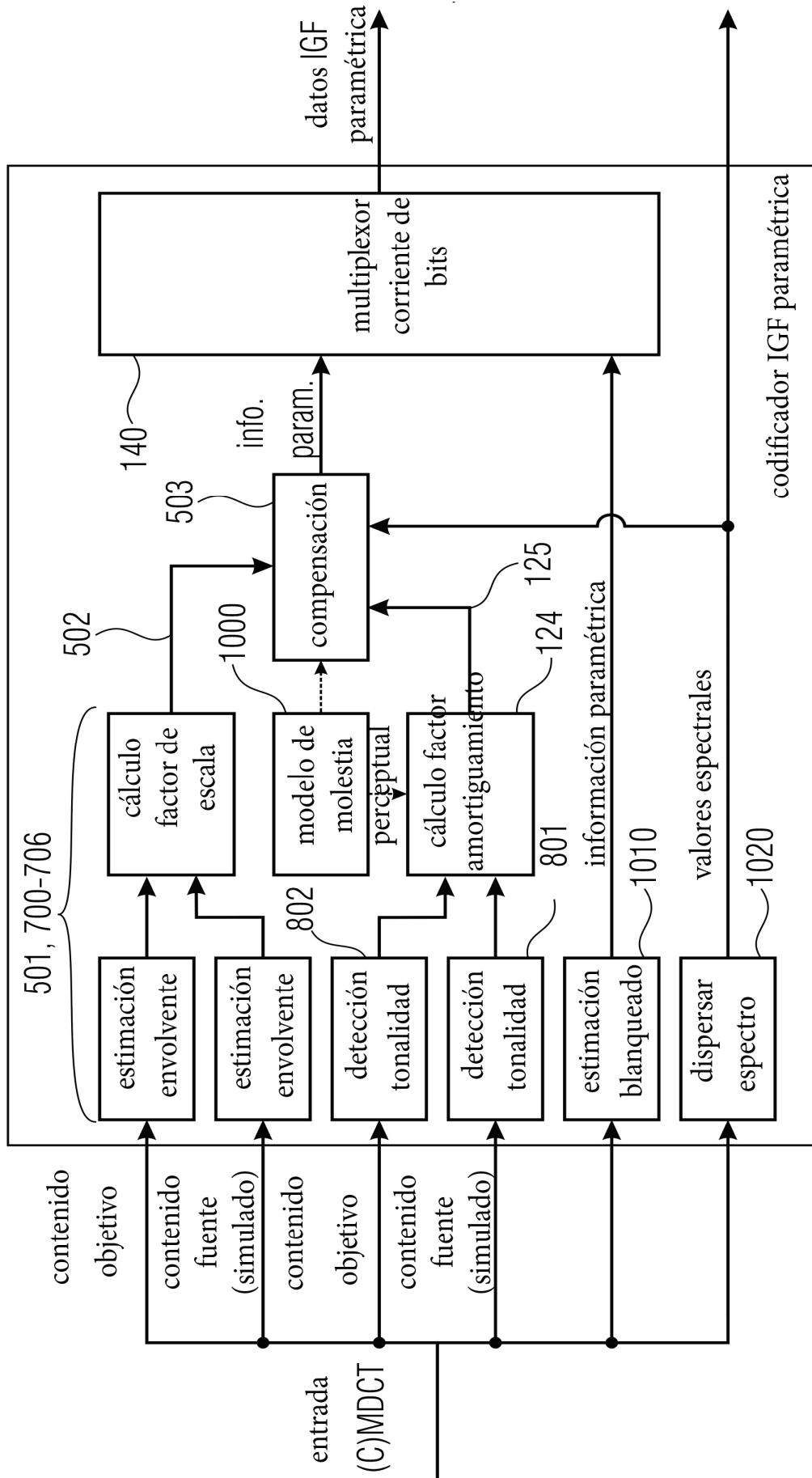


Fig. 10

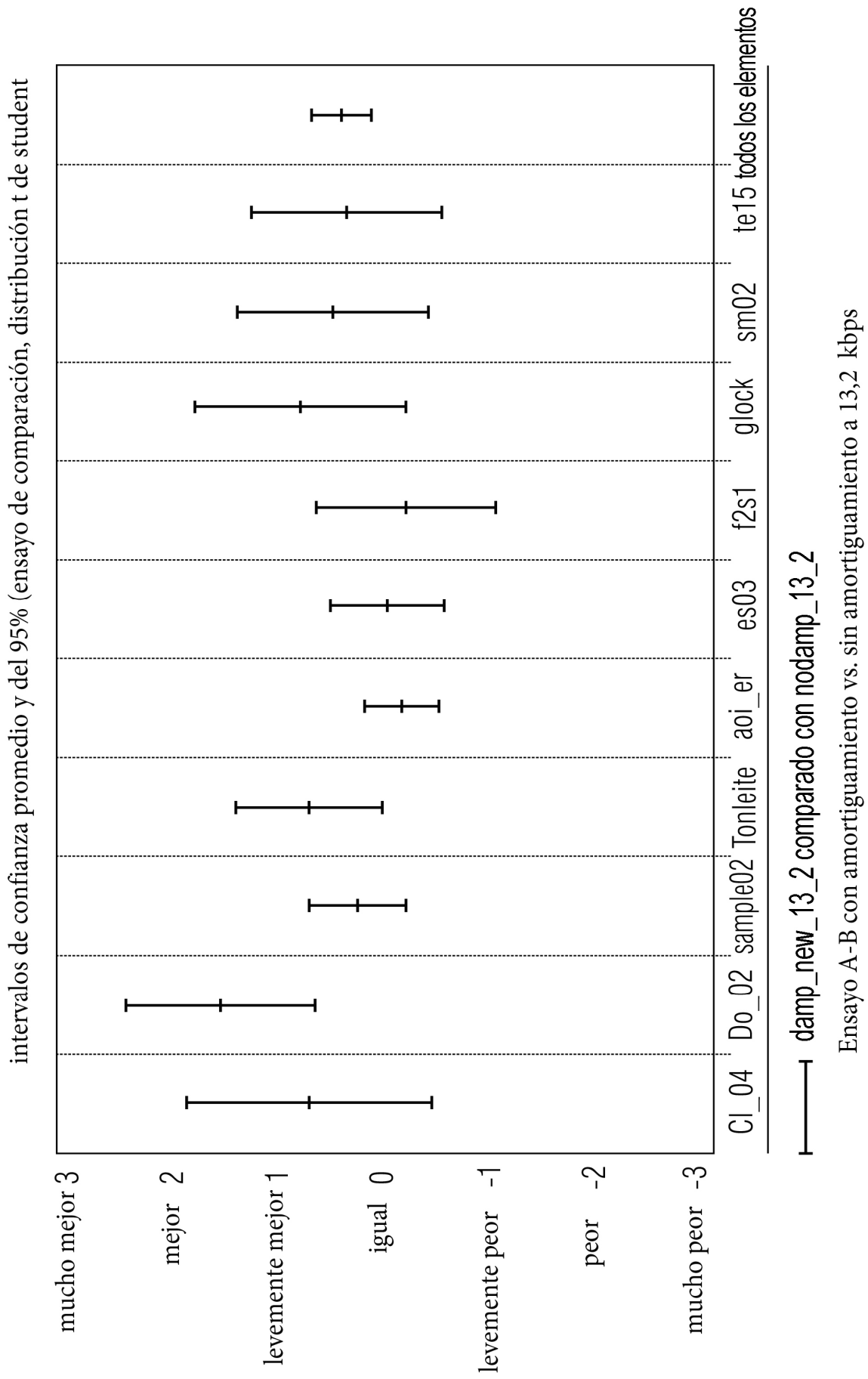


Fig. 11a

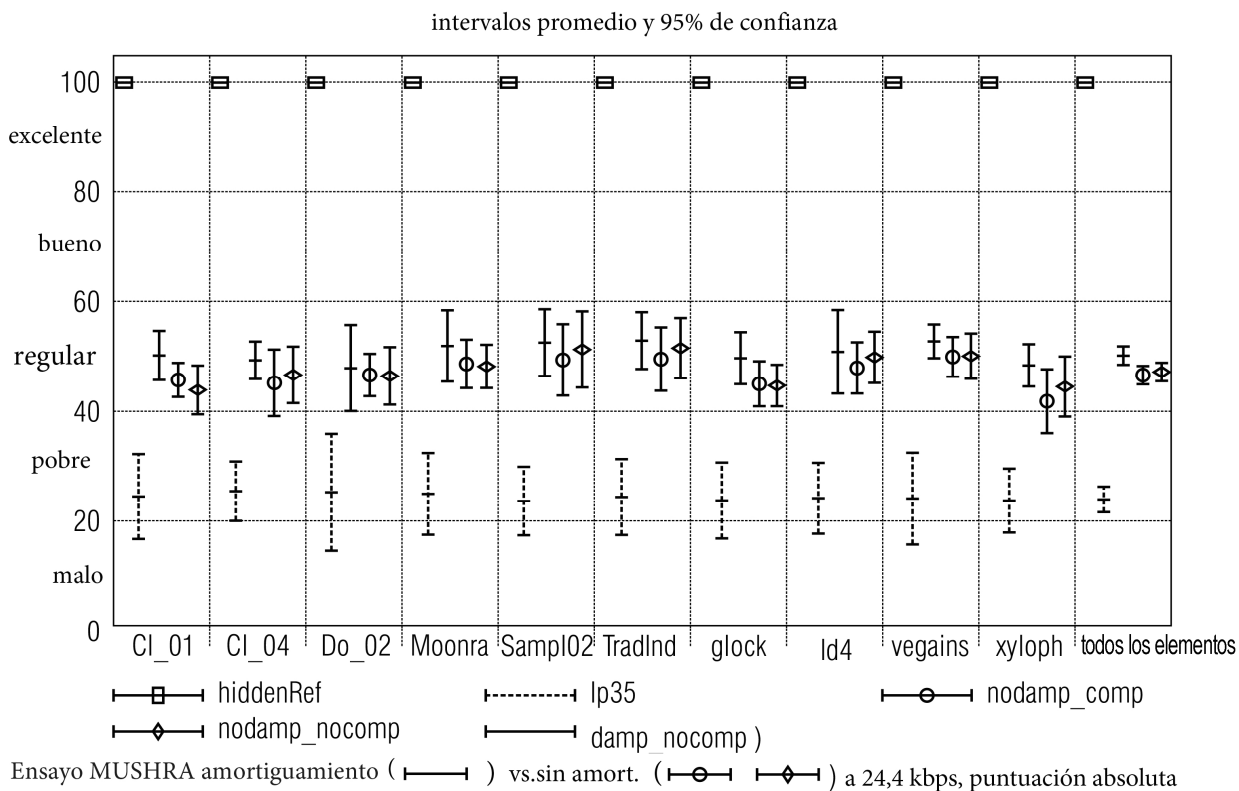


Fig. 11b

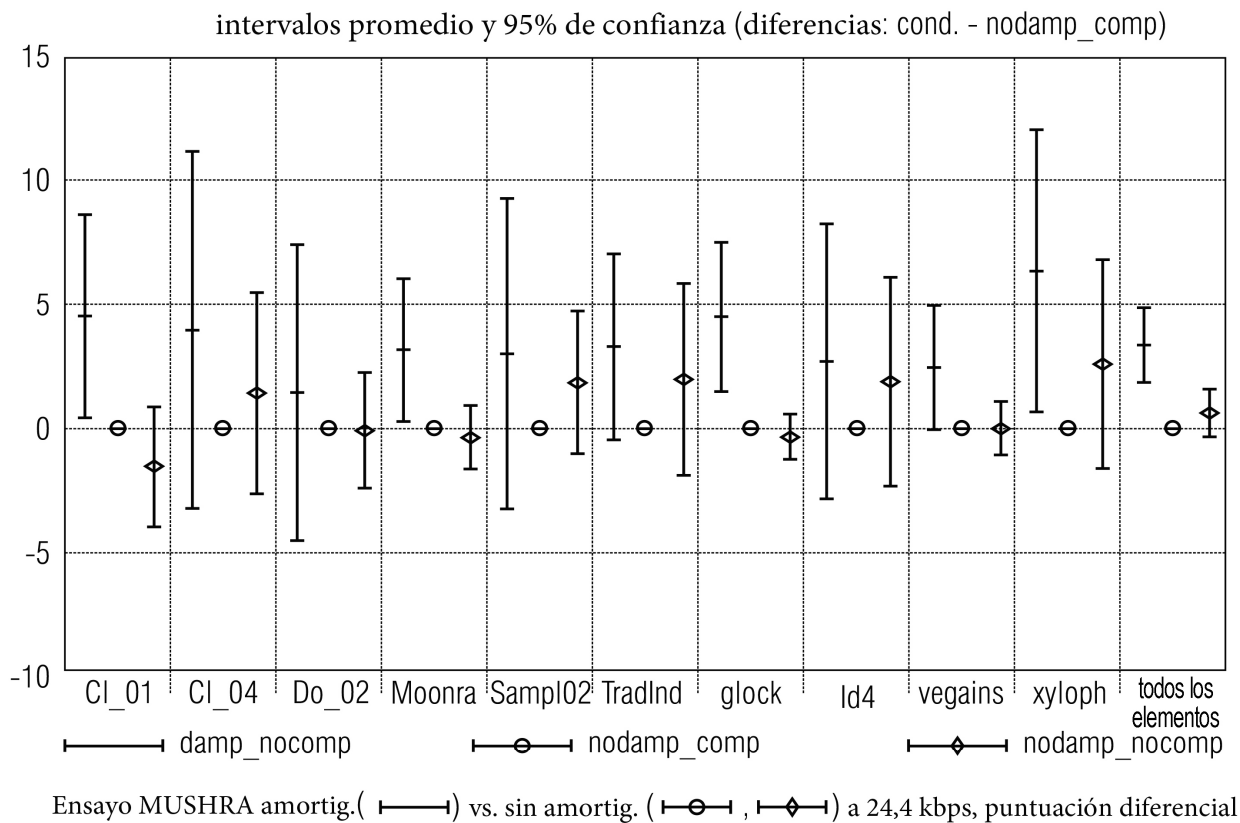


Fig. 11c