

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 539 304**

51 Int. Cl.:

G10L 19/02 (2013.01)

G10L 19/025 (2013.01)

G10L 21/038 (2013.01)

G10L 19/20 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.06.2009 E 09776809 (7)**

97 Fecha y número de publicación de la concesión europea: **08.04.2015 EP 2301027**

54 Título: **Un aparato y un método para generar datos de salida por ampliación de ancho de banda**

30 Prioridad:

11.07.2008 US 79841

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

29.06.2015

73 Titular/es:

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

72 Inventor/es:

**NEUENDORF, MAX;
GRILL, BERNHARD;
KRAEMER, ULRICH;
MULTRUS, MARKUS;
POPP, HARALD;
RETELBACH, NIKOLAUS;
NAGEL, FREDERIK;
LOHWASSER, MARKUS;
GAYER, MARC;
JANDER, MANUEL y
BACIGALUPO, VIRGILIO**

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 539 304 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

Un aparato y un método para generar datos de salida por ampliación de ancho de banda

5 **[0001]** La presente invención se relaciona con un aparato y un método para generar datos de salida por ampliación de ancho de banda (BWE, por sus siglas en inglés) y un codificador de audio.

10 **[0002]** La codificación de audio y codificación de voz natural constituyen dos grandes clases de codificaciones para las señales de audio. La codificación de audio natural se utiliza comúnmente para señales de música o señales arbitrarias en tasa de bits media y generalmente ofrece amplios anchos de banda de audio. Los codificadores de voz están básicamente limitados a la reproducción de voz y pueden utilizarse a baja tasa de bits. Una banda ancha de voz ofrece una gran mejora de calidad subjetiva en comparación con la banda estrecha de voz. Además, debido al tremendo crecimiento del campo de la multimedia, la transmisión de señales de música y otras señales diferentes a las de voz así como el almacenamiento y por ejemplo, la transmisión para radio/TV a alta calidad por sistemas telefónicos es una característica deseable.

15 **[0003]** Para reducir de manera drástica la tasa de bits, se puede realizar una codificación de fuente utilizando codificadores de audio perceptivos de banda dividida. Estos codificadores de audio naturales explotan la irrelevancia perceptiva y redundancia estadística en la señal. En caso que la explotación de lo anteriormente mencionado por sí sola no sea suficiente con respecto a la limitación de tasa de bits dada, la velocidad de muestreo se reduce. Es también común disminuir la cantidad de niveles de composición, permitiendo en forma ocasional una distorsión de cuantización auditiva y emplear una degradación del campo estéreo a través de una codificación estéreo conjunta o codificación paramétrica de dos o más canales. El uso excesivo de dichos métodos da por resultado una degradación perceptiva engorrosa. Con el fin de mejorar el desempeño de codificación, se utiliza la replicación de banda espectral (SBR, por su sigla en inglés) como método eficiente para generar señales de frecuencia alta en un codificador basado en la reconstrucción de frecuencia alta (HFR, por su sigla en inglés).

20 **[0004]** Al grabar y transmitir señales acústicas, un umbral de ruido como el ruido de fondo se encuentra siempre presente. Con el fin de generar una señal acústica auténtica del lado del decodificador, el umbral de ruido debería ser transmitido o generado. En este último caso, el umbral de ruido en la señal de audio original debería ser determinado. En la replicación de banda espectral, esto se realiza por medio de herramientas SBR o módulos relacionados con SBR, que generan parámetros que caracterizan (entre otras cosas) el umbral de ruido y que se transmiten al decodificador para reconstruir el umbral de ruido.

25 **[0005]** En WO 00/45379, se describe una herramienta de umbral de ruido adaptativa, que provee suficientes contenidos de ruido en los componentes de frecuencia de banda alta sintetizados. Sin embargo, se generan artefactos que perturban los componentes de frecuencia de banda alta si, en la banda base, ocurren fluctuaciones de energía o los llamados transitorios, a corto plazo. Estos artefactos no son perceptivamente aceptables y no se proporciona en técnicas previas una solución aceptable (especialmente si el ancho de banda es limitado).

30 **[0006]** EP 2056294 A2 describe una ampliación de ancho de banda que incluye la medida de un umbral de ruido a ser utilizada para reconstrucción de banda de alta frecuencia.

35 **[0007]** Un objetivo de la presente invención consiste, por lo tanto, en proveer un aparato, que permita una eficiente codificación sin artefactos que puedan percibirse, especialmente para señales de voz.

40 **[0008]** Este objetivo se logra mediante un aparato para generar datos de salida por ampliación de ancho de banda de acuerdo con la reivindicación 1, el codificador de acuerdo con la reivindicación 3, un método para generar datos de salida por ampliación de ancho de banda de acuerdo con la reivindicación 6 y un programa de ordenador de acuerdo con la reivindicación 7.

45 **[0009]** La presente invención se basa en el principio que una adaptación de un umbral de ruido medido que depende de la distribución de energía de la señal de audio dentro de una porción de tiempo puede mejorar la calidad perceptivo de una señal de audio sintetizada del lado del decodificador. Aunque desde el punto de vista teórico no es necesaria una adaptación o manipulación del umbral de ruido medido, las técnicas convencionales para generar el umbral de ruido muestran una cantidad de inconvenientes. Por otro lado, la estimación del umbral de ruido basada en una medición de tonalidad, desarrollada por métodos convencionales, es dificultosa y no siempre precisa. Por otro lado, el objetivo del umbral de ruido es reproducir la correcta impresión de tonalidad del lado del decodificador. Aunque la impresión de

tonalidad subjetiva para la señal de audio original y la señal decodificada sea la misma, existe aún la posibilidad de artefactos generados; por ejemplo para señales de voz.

5 **[00010]** Los tests subjetivos muestran que diferentes tipos de señales de voz deberían ser tratadas en forma diferente. En señales de discurso vocalizadas una disminución del umbral de ruido calculado produce una mayor calidad perceptiva al compararla con el umbral de ruido calculado original. En consecuencia la voz suena menos reverberante en este caso. En caso que la señal de audio comprenda sibilantes un aumento artificial de umbral de ruido podrá encubrir los inconvenientes en el método de implementación de parches ("patching") relacionado con los sibilantes. Por ejemplo, las fluctuaciones de energía a corto plazo (transitorios) producen artefactos perturbadores al ser cambiadas o transformadas en la banda de frecuencia mayor y un aumento en el umbral de ruido podrá también encubrir estas fluctuaciones de energía.

10 **[00011]** Dichos transitorios pueden definirse como porciones dentro de señales convencionales s, donde un fuerte aumento en la energía aparece dentro de un corto período de tiempo, el que puede o no estar limitado a una región de frecuencia específica. Transitorios son golpes de castañuelas y de instrumentos de percusión, pero también ciertos sonidos de la voz humana como, por ejemplo, las letras: P, T, K, La detección de este tipo de transitorios es implementada siempre de la misma forma o por el mismo algoritmo (utilizando un valor umbral de transitorios), independiente de la señal, ya sea clasificado como voz o música. Además, una posible distinción entre el discurso vocalizado y no vocalizado no tiene influencia en el mecanismo de detección de transitorios convencional o clásico.

15 **[00012]** En consecuencia, las formas de realización proveen una disminución del umbral de ruido para una señal como un discurso vocalizado y un aumento del umbral de ruido para una señal s que comprende, por ejemplo, sibilantes.

20 **[00013]** Para distinguir las diferentes señales, las formas de realización utilizan datos de distribución de energía (por ejemplo un parámetro de sibilancia) que mide si la energía se encuentra mayormente en frecuencias mayores o frecuencias menores, o en otras palabras, si la representación espectral de la señal de audio muestra una pendiente en aumento o en disminución hacia frecuencias mayores. Otras formas de realización también utilizan el primer coeficiente LPC (LPC = sigla en inglés correspondiente a la traducción codificación predictiva lineal) para generar el parámetro de sibilancia.

25 **[00014]** Existen dos posibilidades para cambiar el umbral de ruido. La primera posibilidad consiste en transmitir dicho parámetro de sibilancia para que el decodificador pueda utilizar el parámetro de sibilancia con el fin de ajustar el umbral de ruido (por ejemplo aumentar o disminuir el umbral de ruido además del umbral de ruido calculado). Este parámetro de sibilancia puede ser transmitido además del parámetro de umbral de ruido calculado por métodos convencionales o calculado del lado del decodificador. Una segunda posibilidad consiste en cambiar el umbral de ruido transmitido utilizando el parámetro de sibilancia (o la distribución de datos de energía) para que el codificador transmita los datos del umbral de ruido modificados al decodificador y no se necesitan modificaciones del lado del decodificador –puede utilizarse el mismo decodificador. Por lo tanto, la manipulación del umbral de ruido puede en principio ser realizada del lado del codificador como así también del lado del decodificador.

30 **[00015]** La replicación de banda espectral como ejemplo de ancho de banda se basa en cuadros SBR que definen una porción de tiempo en la cual la señal de audio está separada en componentes en la primera banda de frecuencia y la segunda banda de frecuencia. El umbral de ruido puede medirse y/o cambiarse para todo la trama SBR. De manera alternativa, es también posible que el cuadro SBR se divida en envolventes de ruido, para que para cada envolvente de ruido, se realizar un ajuste para el umbral de ruido. En otras palabras, la resolución temporal resolución de las herramientas del umbral de ruido es determinada por las llamadas envolventes de ruido dentro de la trama SBR. De acuerdo con la Norma (ISO/IEC 14496-3), cada trama SBR comprende un máximo de dos envolventes de ruido, para que se pueda realizar un ajuste del umbral de ruido teniendo en cuenta los cuadros parciales SBR. Para algunas aplicaciones, esto podría ser suficiente. Sin embargo, también es posible incrementar la cantidad de envolventes de ruido para mejorar el modelo para la variación temporal de tonalidad.

35 **[00016]** En consecuencia, las formas de realización comprenden un aparato para generar datos de salida BWE para una señal de audio, donde la señal de audio comprende componentes en una primera banda de frecuencia y una segunda banda de frecuencia y los datos de salida BWE se adapta para controlar una síntesis de los componentes en la segunda banda de frecuencia. El aparato comprende un medidor de umbral de ruido para medir los datos del umbral de ruido de la segunda banda de frecuencia para una porción de tiempo de la señal de audio. Como el umbral de ruido medido ejerce influencia en la tonalidad de la señal de audio, el medidor del umbral de ruido puede comprender un medidor de tonalidad. En forma alternativa, el medidor del umbral de ruido puede ser implementado para medir el ruido de una señal para obtener el umbral de ruido. El aparato además comprende un caracterizador de energía de señal para derivar los datos de distribución de energía, donde los datos de distribución de energía caracterizan una distribución de energía en

un espectro de la porción de tiempo de la señal de audio y, finalmente, el aparato comprende un procesador para combinar los datos del umbral de ruido y los datos de distribución de energía para obtener datos de salida de BWE.

5 **[00017]** En otras formas de realización, el caracterizador de energía de señal se adapta para usar el parámetro de sibilancia como datos de distribución de energía y el parámetro de sibilancia puede, por ejemplo, ser el primer coeficiente LPC. En otras formas de realización, el procesador se adapta para agregar los datos de distribución de energía a la secuencia de datos de los dador de audio codificados o, en forma alternativa, el procesador se adapta para ajustar el parámetro de umbral de ruido de modo que el umbral de ruido aumente o disminuya dependiendo de los datos de distribución de energía (dependiente de señal). En esta forma de realización, el medidor de umbral de ruido medirá
10 primero el umbral de ruido para generar datos del umbral de ruido, que serán ajustados o cambiados por el procesador más adelante.

15 **[00018]** En otras formas de realización, la porción de tiempo es una trama SBR y el caracterizador de energía de señal se adapta para generar una cantidad de envolventes de umbral de ruido por trama SBR. En consecuencia, el medidor de umbral de ruido como el caracterizador de energía de señal pueden adaptarse para medir los datos del umbral de ruido como también los datos de la distribución de energía derivados para cada envolvente de umbral de ruido. La cantidad de envolvente de umbral de ruido puede, por ejemplo ser 1, 2, 4, ... por trama SBR.

20 **[00019]** Otros ejemplos comprenden además una herramienta de replicación de banda espectral utilizada en un decodificador para generar componentes en una segunda banda de frecuencia de la señal de audio. En esta generación se utilizan datos de salida de replicación de banda espectral y representación espectral de señal cruda para los componentes en la segunda banda de frecuencia. La herramienta de replicación de banda espectral comprende una unidad de cálculo del umbral de ruido, configurada para calcular un umbral de ruido de acuerdo con los datos de distribución de energía, y un combinador para combinar la representación espectral de señal con el umbral de ruido
25 calculado para generar componentes en la segunda banda de frecuencia con el umbral de ruido calculado.

30 **[00020]** Una ventaja de las formas de realización consiste en la combinación de una decisión externa (voz/audio) con un detector interno de discurso vocalizado o detector interno de sibilante (un caracterizador de energía de señal) que controla el evento de ruido adicional señalizado al decodificador o ajustando el umbral de ruido calculado. Para señales de discurso no vocalizado, se ejecuta el cálculo usual de umbral de ruido. Para señales de discurso vocalizado (derivadas de la decisión de conmutación externa) se realiza un análisis de discurso adicional para determinar la vocalización de la señal real. La cantidad de ruido a agregar en el decodificador o codificador es escalada dependiendo del grado de sibilancia (contrario al de vocalización) de la señal. El grado de sibilancia puede ser determinado, por ejemplo, midiendo la pendiente espectral de las partes de la señal corta.
35

Breve Descripción de los Dibujos

40 **[00021]** La presente invención será descrita a continuación a modo de ejemplos ilustrativos. Las características de la invención serán apreciadas con mayor facilidad y mejor comprensión teniendo en cuenta la siguiente descripción detallada, que debería ser considerada con referencia a los dibujos adjuntos, donde:

- La Fig. 1 muestra un diagrama de bloque de un aparato para generar datos de salida BWE de acuerdo con las formas de realización de la presente invención;
- 45 La Fig. 2a ilustra una pendiente espectral negativa de una señal no sibilante;
- La Fig. 2b ilustra una pendiente espectral positiva para una señal del tipo sibilante;
- La Fig. 2c explica el cálculo de la pendiente espectral m basada en orden bajo de parámetros LPC;
- 50 La Fig. 3 muestra un diagrama de bloque de un codificador;
- La Fig. 4 muestra diagramas de bloque para procesar la secuencia de audio codificada para emitir muestras PCM (sigla en inglés correspondiente a la traducción modulación por código de pulso) del lado del decodificador;
- 55 La Fig. 5a, b muestra una comparación de una herramienta de cálculo convencional de umbral de ruido con una herramienta de cálculo modificada de umbral de ruido de acuerdo con unos ejemplos; y
- 60 La Fig. 6 ilustra la partición de una trama SBR en una cantidad predeterminada de porciones de tiempo.

5 [00022] La Fig. 1 muestra un aparato 100 para generar datos de salida por ampliación de ancho de banda (BWE, por su sigla en inglés) 102 para una señal de audio 105. La señal de audio 105 comprende componentes en una primera banda de frecuencia 105a y componentes de una segunda banda de frecuencia 105b. Los datos de salida BWE 102 se adapta para controlar una síntesis de los componentes en la segunda banda de frecuencia 105b. El aparato 100 comprende un medidor de umbral de ruido 110, un caracterizador de energía de señal 120 y un procesador 130. El medidor de umbral de ruido 110 se adapta para medir o determinar los datos del umbral de ruido 115 de la segunda banda de frecuencia 105b para una porción de tiempo de la señal de audio 105. En detalle, el umbral de ruido puede ser determinado al comparar el ruido medido de la banda base con el ruido medido de la banda superior, para que la cantidad de ruido necesario después de la implementación del parche ("patch") para reproducir una impresión de tonalidad natural pueda ser determinada. El caracterizador de energía de señal 120 deriva los datos de distribución de energía 125 caracterizando una distribución de energía en un espectro de porción de tiempo de la señal de audio 105. Por ende, el medidor de umbral de ruido 110 recibe, por ejemplo, la primera y/o segunda banda de frecuencia 105a,b y el caracterizador de energía de señal 120 recibe, por ejemplo, la primera y/o segunda banda de frecuencia 105a, b. El procesador 130 recibe los datos del umbral de ruido 115 y los datos de distribución de energía 125 y los combina para obtener los datos de salida BWE 102. La replicación de banda espectral comprende un ejemplo para la ampliación de ancho de banda, donde los datos de salida BWE 102 se transforman en datos de salida SBR. Las siguientes formas de realización describen principalmente el ejemplo de SBR, pero el aparato/método de invención no queda restringido a este ejemplo.

20 [00023] Los datos de distribución de energía 125 indican una relación entre la energía contenida dentro de la segunda banda de frecuencia en comparación con la energía contenida en la primera banda de frecuencia. En el caso más simple los datos de distribución de energía surgen de un bit que indica si hay más energía almacenada dentro de la banda base en comparación con la banda SBR (banda superior) o vice versa. La banda SBR (banda superior) puede, por ejemplo, definirse como componentes de frecuencia sobre un valor umbral, que puede estar dado, por ejemplo, por 4 kHz y la banda base (banda inferior) puede ser los componentes de la señal, que se encuentren debajo de este valor umbral de frecuencia (por ejemplo, debajo de 4 kHz u otra frecuencia). Ejemplos para este valor umbral de frecuencias serían 5 kHz o 6 kHz.

30 [00024] Las Figs. 2a y 2b muestran dos distribuciones de energía en el espectro dentro de una porción de tiempo de la señal de audio 105. Las distribuciones de energía mostradas por un nivel P como una función de la frecuencia F como señal análoga, que puede también ser una envolvente de una señal dada por una pluralidad de muestras o líneas (transformada en dominio de frecuencia). Los gráficos mostrados están más simplificados para visualizar el concepto de pendiente espectral. La banda de frecuencia superior e inferior pueden definirse como frecuencias por debajo o sobre el valor umbral de frecuencia F_0 (frecuencia de cruce, por ejemplo 500 Hz, 1 kHz o 2 kHz).

40 [00025] La Fig. 2a muestra una distribución de energía que exhibe una pendiente espectral en caída (disminuyendo con frecuencias mayores). En otras palabras, en este caso, hay más energía almacenada en los componentes de baja frecuencia que en los componentes de alta frecuencia. Por ende, el nivel P disminuye para frecuencias mayores implicando una pendiente espectral negativa (función disminuyente). Por lo tanto, un nivel P comprende una pendiente espectral negativa si el nivel de señal P indica que hay menos energía en la banda superior ($F > F_0$) que en la banda inferior ($F < F_0$). Este tipo de señal ocurre, por ejemplo, para una señal de audio que comprende una baja sibilancia o nada de sibilancia.

45 [00026] La Fig. 2b muestra el caso, donde el nivel P aumenta con las frecuencias F implicando una pendiente espectral positiva (una función en aumento del nivel P dependiendo de las frecuencias). Por lo tanto, el nivel P comprende una pendiente espectral positiva si el nivel de señal P indica que hay más energía en la banda superior ($F > F_0$) comparado con la banda inferior ($F < F_0$). Dicha distribución de energía es generada si la señal de audio 105 comprende, por ejemplo, dichos sibilantes.

50 [00027] La Fig. 2a ilustra un espectro de potencia de una señal con pendiente espectral negativa. Una pendiente espectral negativa significa una pendiente en caída espectral. En contraposición, la Fig. 2b ilustra un espectro de potencia de una señal con pendiente espectral positiva. En otras palabras, esta pendiente espectral posee una pendiente en ascenso. Naturalmente, cada espectro como el espectro ilustrado en la Fig. 2a o el espectro ilustrado en la Fig. 2b sufrirá variaciones en una escala local que posee inclinaciones diferentes de la pendiente espectral.

60 [00028] La pendiente espectral puede ser obtenida, cuando, por ejemplo, una línea recta es ajustada al espectro de potencia minimizando las diferencias cuadráticas entre esta línea recta y el espectro real. Ajustar línea recta es ajustada al espectro puede constituir una de las formas para calcular la pendiente espectral de un espectro de tiempo corto. Sin embargo, es preferible calcular la pendiente espectral utilizando coeficientes LPC.

[00029] La publicación "Cálculo eficiente de la pendiente espectral desde varios parámetros LPC" por V. Goncharoff, E. Von Colln y R. Morris, Naval Command, Control and Ocean Surveillance Center (NCCOSC), RDT y E Division, San Diego, CA 92152-52001, 23 de Mayo 23, 1996 divulga varias maneras de calcular la pendiente espectral.

5

[00030] En una realización, la pendiente espectral se define como la inclinación de una función lineal de cuadrados mínimos contenida dentro espectral de potencia logarítmico. Sin embargo, las funciones lineales al espectro de potencia no logarítmicos o al espectro de amplitud u otro tipo de espectro pueden también ser aplicados. Esto es específicamente cierto en el contexto de la presente invención, donde, en la forma de realización preferida, uno está principalmente interesado en el signo de la pendiente espectral, es decir, donde el resultado de del desplazamiento del función lineal es positivo o negativo. El valor real de la pendiente espectral, sin embargo, no radica gran importancia en la alta eficiencia de la forma de realización de la presente invención, pero el valor real puede ser importante en formas de realización más elaboradas.

10

[00031] Cuando la codificación predictiva lineal (LPC, por su sigla en inglés) de voz se utiliza para modelar su espectro a tiempo corto, es computacionalmente más eficiente calcular la pendiente espectral directamente del modelo de parámetros LPC en vez espectral de potencia logarítmico. Fig. 2c ilustra un ecuación para los coeficientes cepstrales c_k correspondientes al espectro de potencia logarítmico que contiene todos los polos de orden n . En esta ecuación, k es un índice entero, p_n es el polo de orden n en la representación de todos los polos de la función de transferencia del dominio z $H(z)$ del filtro LPC. La próxima ecuación en la Fig. 2c es la pendiente espectral en términos de coeficientes cepstrales. Específicamente, m es la pendiente espectral, k y n son enteros y N es el mayor polo de orden del modelo de todos los polos para $H(z)$. La próxima ecuación en la Fig. 2c define espectro de potencia logarítmico $S(\omega)$ del filtro LPC de orden N . G es la constante de ganancia y α_k son los coeficientes de predicción lineal, y ω es igual a $2\pi xf$, donde f es la frecuencia. La ecuación menor en la Fig. 2c directamente resulta en los coeficientes cepstrales como función de los coeficientes LPC α_k . Los coeficientes cepstrales c_k son utilizados para calcular la pendiente espectral. Generalmente, este método será más eficiente computacionalmente que la factorización polinomial de LPC para obtener los valores de polo, y para resolver la pendiente espectral utilizando ecuaciones del polo. En consecuencia, después de haber calculado los coeficientes LPC α_k , uno puede calcular los coeficientes cepstrales c_k utilizando la ecuación en la parte inferior de la Fig. 2c y, luego, se pueden calcular los polos p_n desde los coeficientes cepstrales utilizando la primera ecuación de la Fig. 2c. Luego, teniendo en cuenta los polos, se puede calcular la pendiente espectral m como se define en la segunda ecuación de la Fig. 2c.

15

20

25

30

[00032] Se ha descubierto que el coeficiente LPC α_1 de primer orden es suficiente para obtener una buena estimación para el signo de pendiente espectral. α_1 es, en consecuencia, una buena estimación para c_1 . En consecuencia, c_1 es una buena estimación para p_1 . Cuando p_1 es insertado en la ecuación para la pendiente espectral m , resulta obvio, que debido al signo menos en la segunda ecuación en la Fig. 2c, el signo de pendiente espectral m es inverso al signo del primer coeficiente LPC α_1 en la definición de coeficiente LPC de la Fig. 2c.

35

[00033] Preferentemente, el caracterizador de energía de señal 120 está configurado para generar, como datos de distribución de energía, una indicación en un signo de la pendiente espectral de la señal de audio en una porción de tiempo actual de la señal de audio.

40

[00034] Preferentemente, el caracterizador de energía de señal 120 está configurado para generar, como datos de distribución de energía, datos derivados del análisis LPC de una porción de tiempo de la señal de audio para estimar uno o más coeficientes LPC de bajo orden y derivar los datos de distribución de energía de los uno o más coeficientes LPC de bajo orden.

45

[00035] Preferentemente, el caracterizador de energía de señal 120 está configurado sólo para calcular el primer coeficiente LPC y para no calcular coeficientes LPC adicionales y para derivar los datos de distribución de energía de un signo del primer coeficiente LPC.

50

[00036] Preferentemente, el caracterizador de energía de señal 120 está configurado para determinar la pendiente espectral como pendiente espectral negativa, donde una energía espectral disminuye de frecuencias menores a frecuencias mayores, cuando el primer coeficiente LPC posee un signo positivo, y para detectar la pendiente espectral como pendiente espectral positiva, donde la energía espectral aumenta desde frecuencias menores a frecuencias mayores, cuando el primer coeficiente LPC posee un signo negativo.

55

[00037] En otras formas de realización, el detector de pendiente espectral detector o caracterizador de energía de señal 120 está configurado no sólo para calcular los coeficientes LPC de primer orden sino para calcular varios coeficientes LPC de bajo orden como los coeficientes LPC hasta el orden de 3 ó 4 ó aún mayor. En dicha forma de realización, la

60

pendiente espectral se calcula hasta tal alta precisión que no sólo se puede indicar el signo como parámetro de sibilancia, sino también un valor que depende de la pendiente, que posee más de dos valores como en la forma de realización relativa al signo.

5 **[00038]** Como se explicó previamente la sibilancia comprende una gran cantidad de energía en la región de frecuencia superior, mientras que para partes sin o con poca sibilancia (por ejemplo, vocales) la energía es mayormente distribuida dentro de la banda base (banda de frecuencia baja). Esta observación puede utilizarse con el fin de determinar si o hasta donde una parte de señal de voz comprende un sibilante o no.

10 **[00039]** Por lo tanto, el medidor del umbral de ruido 110 (detector) puede utilizar la pendiente espectral para la decisión sobre la cantidad de sibilancia o para otorgar el grado de sibilancia dentro de una señal. La pendiente espectral puede básicamente ser obtenida de un simple análisis LPC de la distribución de energía. Por ejemplo, puede ser suficiente calcular el primer coeficiente LPC para determinar el parámetro de pendiente espectral (parámetro de sibilancia), ya que desde el primer coeficiente LPC se puede inferir el comportamiento espectral (función de aumento o disminución). Este análisis puede realizarse dentro del caracterizador de energía de señal 20. En caso que el codificador de audio use LPC para la decodificación de la señal de audio, no habrá necesidad de transmitir el parámetro de sibilancia, ya que el primer coeficiente LPC puede ser utilizado como datos de distribución de energía del lado del decodificador.

15 **[00040]** En algunas formas de realización el procesador 130 puede estar configurado para cambiar los datos del umbral de ruido 115 de acuerdo con los datos de distribución de energía 125 (pendiente espectral) para obtener datos modificados del umbral de ruido, y el procesador 130 puede estar configurado para agregar los datos modificados del umbral de ruido a una secuencia de bits que comprende datos de salida de BWE 102. El cambio de datos del umbral de ruido 115 puede ser tal que el umbral de ruido modificado aumenta para una señal de audio 105 que comprende mayor sibilancia (Fig. 2b) en comparación con una señal de audio 105 que comprende menor sibilancia (Fig. 2a).

20 **[00041]** El aparato 100 para generar datos de salida por ampliación de ancho de banda (BWE, por su sigla en inglés) 102 puede ser parte de un codificador 300. La Fig. 3 muestra una forma de realización para el codificador 300, que comprende módulos relacionados con BWE 310 (que pueden, por ejemplo comprender módulos relacionados con SBR), un banco de análisis QMF (sigla en inglés correspondiente a la traducción filtro espejo en cuadratura) 320, un filtro paso bajo (filtro LP) 330, un codificador de núcleo AAC 340 y un formateador de secuencia de bits de carga útil 350. Además, el codificador 300 comprende el calculador de datos de la envolvente 210. El codificador 300 comprende una entrada para muestras PCM (señal de audio 105; PCM = modulación por código de pulso), conectada al banco de análisis QMF 320, y a los módulos relacionados con BWE 310 y al filtro LP 330. El banco de análisis QMF 320 puede comprender un filtro de paso alto para separar la segunda banda de frecuencia 105b y se conecta al calculador de datos de la envolvente 210, que a la vez se conecta con el formateador de secuencia de bits de carga útil 350. El filtro LP 330 puede comprender filtro de paso bajo para separar la primera banda de frecuencia de núcleo 105a y se conecta con el codificador de núcleo AAC 340, que a la vez se conecta con el formateador de secuencia de bits de carga útil 350. Finalmente, el módulo relacionado con BWE 310 se conecta al calculador de datos de la envolvente 210 y al codificador de núcleo AAC 340.

25 **[00042]** En consecuencia, el codificador 300 realiza un sub-muestreo de la señal de audio 105 para generar componentes en la banda de frecuencia de núcleo 105a (en el filtro LP 330), que son ingresados en el codificador de núcleo AAC 340, que codifica la señal de audio en la banda de frecuencia de núcleo y envía la señal codificada 355 al formateador de secuencia de bits de carga útil 350 en el cual la señal de audio codificada 355 de la banda de frecuencia de núcleo es agregada a la secuencia de audio codificada 345 (secuencia de bits). Por otro lado, la señal de audio 105 se analiza mediante el banco de análisis QMF 320, que extrae los componentes de frecuencia de la banda de frecuencia alta 105b e ingresa las señales en el calculador de datos de la envolvente 210 para generar datos BWE 375. Por ejemplo, un banco 64 de sub-banda QMF 320 realiza el filtrado de sub-banda de la señal de entrada. La salida desde el banco de filtro (es decir las muestras de sub-bandas) es valuada en forma compleja y por lo tanto, sometida a un sobre-muestreo por un factor de dos en comparación con un banco QMF regular.

30 **[00043]** El módulo relacionado con BWE 310 puede por ejemplo, comprender el aparato 100 para generar datos de salida de BWE 102 y controla el calculador de datos de la envolvente 210 proporcionando, por ejemplo, datos de salida BWE 102 (parámetro de sibilancia) al calculador de datos de la envolvente 210. Utilizando los componentes de audio 105b generados por el banco de análisis QMF 320, el calculador de datos de la envolvente 210 calcula los datos BWE 375 y envía los datos BWE 375 al formateador de secuencia de bits de carga útil 350, que combina los datos BWE 375 con los componentes 355 codificados por el codificador de núcleo 340 en la secuencia de audio codificada 345. Además, el calculador de datos de la envolvente 210 puede por ejemplo usar el parámetro de sibilancia 125 para ajustar el umbral de ruidos dentro de las envolventes de ruido.

35

40

[00044] En forma alternativa, el aparato 100 para generar datos de salida BWE data 102 puede también ser parte del calculador de datos de la envolvente 210 y el procesador puede también ser parte del formateador de secuencia de bits de carga útil 350. En consecuencia, los diferentes componentes del aparato 100 pueden ser parte de los diferentes componentes del codificador de la Fig. 3.

5

[00045] La Fig. 4 muestra un ejemplo para un decodificador 400, donde la secuencia de audio codificada 345 es ingresada en un desformateador de secuencia de bits de carga útil 357, que separa la señal de audio codificada 355 de los datos BWE 375. La señal de audio codificada 355 es ingresada, por ejemplo, en un decodificador de núcleo AAC 360, que genera la señal de audio decodificada 105a en la primera banda de frecuencia. La señal de audio 105a (componentes en la primera banda de frecuencia) es ingresada en un banco QMF de análisis de banda 32 370, que genera, por ejemplo, sub-bandas de frecuencia $32 \cdot 105_{32}$ desde la señal de audio 105a en la primera banda de frecuencia. Las sub-bandas de frecuencia de la señal de audio 105_{32} ingresan en el generador del parche 410 para generar una representación espectral de señal cruda 425 (parche), que ingresa en una herramienta BWE 430a. La herramienta BWE 430a puede, por ejemplo, comprender una unidad de cálculo de umbral de ruido para generar un umbral de ruido. Además, la herramienta BWE 430a puede reconstruir armónicos perdidos o realizar un paso de filtrado inverso. La herramienta BWE 430a puede implementar métodos conocidos de replicación de banda espectral a utilizar en la salida de datos espectrales QMF del generador de parche ("patching") 410. El algoritmo de implementación de parches ("patching") usado en el dominio de frecuencia podría, por ejemplo, emplear el método de espejo copiado simple de los datos espectrales dentro del dominio de frecuencia.

10

15

20

[00046] Por otro lado, los datos BWE 375 (por ejemplo que comprende datos de salida BWE 102) ingresan a un verificador de validez de secuencia de bits 380, que analiza los datos BWE 375 para obtener diferente sub-información 385 y los ingresan en por ejemplo, una decodificación Huffman y unidad de descuantización 390 que, por ejemplo, extrae la información de control 412 y parámetros de replicación de banda espectral 102. La información de control 412 controla el generador de parche 430 (por ejemplo usar un algoritmo de implementación de parches ("patching") específico) y el parámetro BWE 102 comprende, por ejemplo, también los datos de distribución de energía 125 (por ejemplo el parámetro de sibilancia). La información de control 412 ingresa a la herramienta BWE 430a y los parámetros de replicación de banda espectral 102 ingresan a la herramienta BWE 430a y a un ajustador de envolvente 430b. El ajustador de envolvente 430b es operativo para ajustar la envolvente para el parche generado. En consecuencia, el ajustador de envolvente 430b genera la señal cruda ajustada 105b para la segunda banda de frecuencia y la ingresa en un banco QMF de síntesis QMF 440, que combina los componentes de la segunda banda de frecuencia 105b con la señal de audio en dominio de frecuencia 105_{32} . El banco QMF de síntesis 440 puede, por ejemplo, comprender 64 bandas de frecuencia y genera combinando ambas señales (los componentes en la segunda banda de frecuencia 105b y la señal de audio en dominio de frecuencia 105_{32}) la señal de audio de síntesis 105 (por ejemplo, salida de muestras PCM, PCM = modulación de código de pulso).

25

30

35

[00047] El banco QMF de síntesis 440 puede, por ejemplo, comprender un combinador, que combina la señal en dominio de frecuencia 105_{32} con la segunda banda de frecuencia 105b antes de ser transformada en el dominio de tiempo y antes de ser emitida como señal de audio 105. Opcionalmente, el combinador puede emitir la señal de audio 105 en dominio de frecuencia.

40

[00048] Las herramientas BWE 430a pueden comprender una herramienta convencional de umbral de ruido, que agrega ruido adicional al espectro con parche implementado (representación espectral de señal cruda 425), para que los componentes espectrales 105a que han sido transmitidos por un codificador de núcleo 340 y son usados para sintetizar los componentes de la segunda banda de frecuencia 105b exhiben la tonalidad de la segunda banda de frecuencia 105b de la señal original. Especialmente en vías de discurso vocalizado, sin embargo, el ruido adicional agregado por la herramienta convencional de umbral de ruido puede dañar la calidad percibida de la señal reproducida.

45

[00049] De acuerdo con unos ejemplos la herramienta del umbral de ruido puede ser modificada para que la herramienta del umbral de ruido tome en cuenta los datos de distribución de energía 125 (parte de los datos BWE 102) para cambiar el umbral de ruido según el grado de sibilancia detectado (ver Fig. 2). De modo alternativo, como se describe antes el decodificador no puede que no sea modificado y en cambio el codificador puede cambiar los datos del umbral de ruido según el grado de sibilancia detectado.

50

[00050] La Fig. 5 muestra una comparación de una herramienta de cálculo convencional del umbral de ruido con una herramienta de cálculo modificada del umbral de ruido de acuerdo con unos ejemplos.

55

[00051] Esta herramienta de cálculo modificada del umbral de ruido puede ser parte de la herramienta BWE 430.

5 [00052] La Fig. 5a muestra la herramienta de cálculo convencional del umbral de ruido que comprende un calculador 433, que utiliza los parámetros espectrales de replicación de banda 102 y la representación espectral de señal cruda 425 para calcular líneas espectrales crudas y líneas espectrales de ruido. Los datos BWE 375 pueden comprender datos de la envolvente y datos del umbral de ruido, are transmitidos desde el codificador como parte de la secuencia de audio
 10 codificada 345. La representación espectral de señal cruda 425 se obtiene, por ejemplo, de un generador de parche, que genera componentes de la señal de audio en la banda de frecuencia superior (componentes sintetizados en la segunda banda de frecuencia 105b). Las líneas espectrales crudas y líneas espectrales de ruido serán además procesadas, lo cual incluirá un filtrado inverso, ajuste de envolvente, adición de armonía perdida etc. Finalmente, un combinador 434 combina las líneas espectrales crudas con las líneas espectrales de ruido calculadas a los componentes en la segunda banda de frecuencia 105b.

15 [00053] La Fig. 5b muestra una herramienta de cálculo del umbral de ruido de acuerdo con unos ejemplos. Además, de la herramienta de cálculo convencional del umbral de ruido de la Fig. 5a, los ejemplos comprenden una unidad modificadora del umbral de ruido 431 configurada, por ejemplo, para modificar los datos transmitidos del umbral de ruido basados en los datos de distribución de energía 125 antes de ser procesados en la herramienta de cálculo del umbral de ruido 433. Los datos de distribución de energía 125 pueden también ser transmitidos desde el codificador como parte de o además de os datos de BWE 375. La modificación de los datos transmitidos del umbral de ruido comprende, por ejemplo, un aumento de pendiente espectral positiva (ver Fig. 2a) o disminución de pendiente espectral negativa (ver Fig. 2b) del nivel del umbral de ruido, por ejemplo, un aumento de 3 dB o disminución de 3 dB u otro valor discreto (por ejemplo +/- 1 dB o +/- 2 dB). El valor discreto puede ser un valor entero dB o valor no entero dB. Puede haber además una dependencia funcional (por ejemplo una relación lineal) entre la disminución/aumento y la pendiente espectral.

25 [00054] Teniendo en cuenta estos datos modificados del umbral de ruido la herramienta de cálculo del umbral de ruido 433 calcula nuevamente las líneas espectrales crudas y las líneas espectrales de ruido modificado basada en la representación espectral de señal cruda 425, que nuevamente puede obtenerse desde un generador de parche. La herramienta de replicación de banda espectral 430 de la Fig. 5b comprende además un combinador 434 para combinar las líneas espectrales crudas con el umbral de ruido calculado (con la modificación desde la unidad modificadora 431) para generar los componentes en la segunda banda de frecuencia 105b.

30 [00055] Los datos de distribución de energía 125 pueden indicar en el caso más simple una modificación en el nivel de datos transmitidos del umbral de ruido. Como se explicó anteriormente, se puede utilizar también el primer coeficiente LPC como datos de distribución de energía 125. En consecuencia, si la señal de audio 105 fue codificada utilizando LPC, otros ejemplos usan el primer coeficiente LPC, ya transmitido por la secuencia de audio codificada 345, como datos de distribución de energía 125. En este caso no hay necesidad de transmitir además los datos de distribución de energía
 35 125.

40 [00056] En forma alternativa una modificación del umbral de ruido puede realizarse también después del cálculo dentro del calculador 433 para que la unidad modificadora del umbral de ruido 431 pueda disponerse después del procesador 433. En otros ejemplos los datos de distribución de energía 125 pueden ingresarse directamente en el calculador 433 modificando directamente el cálculo del umbral de ruido como parámetro de cálculo. Por lo tanto, la unidad modificadora del umbral de ruido 431 y el calculador/procesador 433 pueden combinarse con la herramienta de modificadora del umbral de ruido 433, 431.

45 [00057] En otro ejemplo la herramienta BWE 430 que comprende la herramienta de cálculo del piso comprende un conmutador, donde el conmutador está configurado para conmutar entre un nivel alto para el umbral de ruido (pendiente espectral positiva) y un nivel bajo para el umbral de ruido (pendiente espectral negativa). El nivel alto puede, por ejemplo, corresponder al caso donde el nivel transmitido para el ruido es duplicado (o multiplicado por un factor), mientras que el nivel bajo corresponde al caso donde el nivel transmitido disminuye por factor. El conmutador puede ser controlado por un bit en la secuencia de bits de la señal de audio codificada 345 indicando una pendiente espectral positiva o negativa de la señal de audio. En forma alternativa el conmutador puede activarse mediante un análisis de la señal de audio decodificada 105a (componentes en la primera banda de frecuencia) o de la señal de audio de sub-bandas de frecuencia 105₃₂, por ejemplo con respecto a la pendiente espectral (si la pendiente espectral es positiva o negativa). En forma alternativa el conmutador puede ser también controlado por el primer coeficiente LPC, ya que este coeficiente indica la pendiente espectral (ver arriba).
 50
 55

[00058] Aunque algunas de las Figs. 1, 3 a 5 se ilustran como diagramas de bloque de los aparatos, estas figuras simultáneamente son ilustraciones de un método, donde las funcionalidades del bloque corresponden a los pasos del método.

[00059] Como se explicó anteriormente, una unidad de tiempo SBR (trama SBR) o porción de tiempo pueden dividirse en varios bloques de datos, llamadas envolventes. Esta partición puede ser uniforme en todo la trama SBR y permite ajustar flexibilidad a la síntesis de la señal de audio dentro de la trama SBR.

5 **[00060]** La Fig. 6 ilustra dicha partición para la trama SBR en una cantidad n de envolventes. La trama SBR cubre un periodo de tiempo o porción de tiempo T entre el tiempo inicial t_0 y el tiempo final t_n . La porción de tiempo T es, por ejemplo, dividida en ocho porciones de tiempo, una primera porción de tiempo T_1 , una segunda porción de tiempo T_2 , ..., una octava porción de tiempo T_8 . En este ejemplo, el número máximo de envolventes coincide con el número de porciones de tiempo y está dado por $n = 8$. Las 8 porciones de tiempo T_1, \dots, T_8 están separadas por 7 bordes, lo que implica que un borde 1 separa la primera y segunda porción de tiempo T_1, T_2 , un borde 2 se encuentra entre la segunda porción T_2 y una tercera porción T_3 , y así hasta que un borde 7 separa la séptima porción T_7 y la octava porción T_8 .

10 **[00061]** En otras formas de realización, la trama SBR se divide en cuatro envolventes de ruido ($n = 4$) o en dos envolventes de ruido ($n = 2$). En la forma de realización de la Fig. 6, todas las envolventes comprenden la misma longitud temporal, que puede ser diferente en otras formas de realización para que las envolventes de ruido cubran diferentes longitudes de tiempo. En detalle, el caso con dos envolventes de ruido ($n = 2$) comprende una primera envolvente que se extiende desde el tiempo t_0 por las primeras cuatro porciones de tiempo (T_1, T_2, T_3 y T_4) y la segunda envolvente de ruido que cubre de la quinta a la octava porción de tiempo (T_5, T_6, T_7 y T_8). Debido a la Norma ISO/IEC 14496-3, la cantidad máxima de envolventes se restringe a dos. Pero las formas de realización pueden usar cualquier cantidad de envolventes (por ejemplo dos, cuatro u ocho envolventes).

15 **[00062]** En otras formas de realización el calculador de datos de envolvente 210 está configurado para cambiar la cantidad de envolventes dependiendo del cambio de los datos medidos del umbral de ruido 115. Por ejemplo, si los datos medidos del umbral de ruido 115 indican un umbral de ruido variado (por ejemplo sobre el valor umbral) la cantidad de envolventes puede aumentar mientras en caso que los datos del umbral de ruido 115 indican un ruido constante en el umbral de ruido la cantidad de envolventes puede disminuir.

20 **[00063]** En otras formas de realización, caracterizador de energía de la señal 120 puede basarse en información lingüística para detectar sibilantes en la voz. Cuando, por ejemplo, una señal de voz ha asociado meta información como la ortografía de fonética internacional, un análisis de esta meta información proveerá también una detección de sibilante de la porción de voz. En este contexto, se analiza la porción de meta datos de la señal de audio.

25 **[00064]** Aunque algunos aspectos han sido descritos en el contexto de un aparato, queda claro que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a un paso del método o rasgo de un paso del método. En forma análoga, los aspectos descritos en el contexto de un paso del método también representan una descripción de un bloque o ítem correspondiente o rasgo de un aparato correspondiente.

30 **[00065]** La señal de audio codificada puede almacenarse en un medio de almacenamiento digital o transmitida en un medio de transmisión como un medio de transmisión inalámbrico o por cable como la Internet.

35 **[00066]** Dependiendo de ciertos requisitos de implementación, las formas de realización de la invención pueden implementarse en hardware o software. La implementación puede realizarse utilizando un medio de almacenamiento digital por ejemplo un disquete, DVD, CD, memoria ROM, PROM, EPROM, EEPROM o FLASH, con señales de control capaces de ser leídas en forma electrónica almacenados en ellos, que cooperan (o son capaces de cooperar) con un sistema de ordenador programable tal que se lleve a cabo el método respectivo

40 **[00067]** Algunas formas de realización de acuerdo con la invención comprenden un portador de datos con señales de control capaces de ser leídas en forma electrónica, capaces de cooperar con un sistema de ordenador programable tal que se lleve a cabo uno de los métodos descritos.

45 **[00068]** Generalmente, las formas de realización de la presente invención pueden implementarse como un producto de programa de ordenador con un código de programa, el código de programa es operativo para desarrollar uno de los métodos cuando el producto de programa de ordenador es utilizado en un ordenador. El código de programa puede por ejemplo ser almacenado en un portador capaz de ser leído por una máquina.

50 **[00069]** Otras formas de realización comprenden el programa de ordenador para desarrollar uno de los métodos descritos en la presente, almacenado en un portador capaz de ser leído por una máquina.

- [00070] En otras palabras, una forma de realización del método de invención consiste, por lo tanto, en un programa de ordenador con un código de programa para desarrollar uno de los métodos descritos en la presente, cuando el programa de ordenador es utilizado en un ordenador.
- 5 [00071] Otra forma de realización del método de invención consiste, por lo tanto, un portador de datos (o medio almacenador digital, o medio capaz de ser leído por un ordenador) que comprende, el programa de ordenador grabado en los mismos, para desarrollar uno de los métodos descritos en la presente.
- 10 [00072] Otra forma de realización del método de invención consiste, por lo tanto, en una secuencia de datos o secuencia de señales que representan el programa de ordenador para desarrollar uno de los métodos descritos en la presente. La secuencia de datos o secuencia de señales puede por ejemplo ser configurada para ser transferida a través de una conexión de comunicación de datos, por ejemplo por Internet.
- 15 [00073] Otra forma de realización comprende un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado o adaptado para desarrollar uno de los métodos descritos en la presente.
- [00074] Otra forma de realización comprende un ordenador con programa de ordenador instalado en la misma, para desarrollar uno de los métodos descritos en la presente.
- 20 [00075] En algunas formas de realización, un dispositivo lógico programable (por ejemplo una matriz de puertas programables por campo) puede utilizarse para desarrollar una de las funcionalidades de los métodos descritos en la presente. En algunas formas de realización, una matriz de puertas programables por campo puede cooperar con un microprocesador con el fin de desarrollar uno de los métodos descritos en la presente. Generalmente, los métodos son preferentemente desarrollados por cualquier aparato de hardware.
- 25 [00076] Las formas de realización descritas anteriormente son meramente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en la presente resultarán evidentes para los expertos en la técnica. Por lo tanto, se intenta limitar sólo por el alcance de las reivindicaciones de la patente inminentes y no por los detalle específicos presentados a modo de descripción y explicación de las formas de realización de la presente.
- 30

REIVINDICACIONES

1. Un aparato (100) para generar datos de salida por ampliación de ancho de banda (102) para una señal de audio (105), la señal de audio (105) comprende componentes en una primera banda de frecuencia (105a) y componentes en una segunda banda de frecuencia (105b), donde los datos de salida por ampliación del ancho de banda (102) están adaptados para controlar una síntesis de los componentes en la segunda banda de frecuencia (105b), el aparato comprende:
- 5 un medidor de umbral de ruido (110) para medir un umbral de ruido para generar datos de umbral de ruido (115) de la segunda banda de frecuencia (105b) para una porción de tiempo (T) de la señal de audio (105);
- 10 un caracterizador de energía de señal (120) para derivar un parámetro de sibilancia o un parámetro de pendiente espectral como datos de distribución de energía (125), donde el caracterizador de energía de señal está adaptado para recibir la primera banda de frecuencia (105a) y la segunda banda de frecuencia (105b), los datos de distribución de energía (125) caracterizan una distribución de energía en un espectro de la porción de tiempo (T) de la señal de audio (105), el parámetro de sibilancia o el parámetro de pendiente espectral identifican un nivel en aumento o disminución de la señal de audio (105) con la frecuencia (F); y
- 15 un procesador (130) para combinar los datos de umbral de ruido (115) y los datos de distribución de energía (125) para obtener los datos de salida por ampliación de ancho de banda (102),
- 20 donde el procesador (130) está configurado para cambiar los datos de umbral de ruido (115) según los datos de distribución de energía (125) para obtener datos de umbral de ruido modificados, los datos de umbral de ruido modificados indicando un umbral de ruido modificado que aumenta o disminuye, dependiendo de los datos de distribución de energía, con respecto al umbral de ruido indicado por los datos de umbral de ruido,
- 25 donde el cambio de los datos de umbral de ruido (115) es tal que el umbral de ruido modificado se incrementa para una señal de audio (105) que comprende un primer grado de sibilancia en comparación con una señal de audio (105) que comprende un segundo grado de sibilancia, donde el segundo grado es menor que el primer grado,
- 30 donde el aparato (100) para generar datos de salida por ampliación de ancho de banda (102) está configurado para ejecutar una decisión externa para determinar si la porción de tiempo (T) de la señal de audio (105) es una señal de voz o una señal que no es de voz,
- 35 donde los datos de umbral de ruido medidos por el medidor de umbral de ruido (110) son utilizados como los datos de salida por ampliación de ancho de banda, cuando la porción de tiempo (T) de la señal de audio (105) es una señal que no es de voz, y
- 40 donde el caracterizador de energía de señal (120) está configurado para realizar, cuando la porción de tiempo (T) de la señal de audio (105) es una señal de voz, un análisis de discurso adicional, para determinar un grado de sibilancia de la señal de voz, y donde el procesador (130) está configurado para agregar los datos de umbral de ruido modificados a una secuencia de bits como los datos de salida por ampliación de ancho de banda (102), cuando la porción de tiempo (T) de la señal de audio (105) es una señal de voz.
- 45
2. El aparato (100) de la reivindicación 1, donde el caracterizador de energía de señal (120) está configurado para utilizar el primer coeficiente de codificación predictiva lineal como parámetro de sibilancia.
- 50
3. Un codificador (300) para codificar una señal de audio (105), la señal de audio (105) comprende componentes en una primera banda de frecuencia (105a) y componentes en una segunda banda de frecuencia (105b), el codificador comprende:
- 55 un codificador de núcleo (340) para codificar los componentes en la primera banda de frecuencia (105a);
- un aparato (100) para generar datos de salida por ampliación de ancho de banda (102) de acuerdo con una de las reivindicaciones 1 a 2; y

un calculador de datos de la envolvente (210) para calcular datos de ampliación de ancho de banda (375) basado en los componentes en la segunda banda de frecuencia (105b), donde los datos de ampliación de ancho de banda calculados (375) comprenden los datos de salida por ampliación de ancho de banda (102).

5 4. El codificador (300) de la reivindicación 3, donde la porción de tiempo (T) cubre una trama SBR (replicación de banda espectral), la trama SBR comprende una pluralidad de envolventes de ruido, y donde el calculador de datos de la envolvente de ruido (210) está configurado para calcular diferentes datos de ampliación de ancho de banda (375) para diferentes envolventes de ruido de la pluralidad de envolventes de ruido.

10 5. El codificador (300) de la reivindicación 3 o reivindicación 4, donde el calculador de datos de la envolvente (210) está configurado para cambiar un número de envolventes en función de un cambio de los datos medidos del umbral de ruido (115).

15 6. Un método para generar datos de salida por ampliación de ancho de banda (102) para una señal de audio (105), la señal de audio (105) comprende componentes en una primera banda de frecuencia (105a) y componentes en una segunda banda de frecuencia (105b), donde los datos de salida por ampliación de ancho de banda (102) están adaptados para controlar una síntesis de los componentes en la segunda banda de frecuencia (105b), el método comprende:

20 la medición de un umbral de ruido para generar datos de umbral de ruido (115) de la segunda banda de frecuencia (105b) para una porción de tiempo (T) de la señal de audio (105);

25 la derivación de un parámetro de sibilancia o de un parámetro de pendiente espectral como datos de distribución de energía (125), donde por tanto se reciben la primera banda de frecuencia (105a) y la segunda banda de frecuencia (105b), los datos de distribución de energía (125) caracterizan una distribución de energía en un espectro de la porción de tiempo (T) de la señal de audio (105), el parámetro de sibilancia o el parámetro de pendiente espectral identifican un nivel de aumento o disminución de la señal de audio (105) con la frecuencia (F); y

30 la combinación de los datos de umbral de ruido (115) y los datos de distribución de energía (125) para obtener los datos de salida por ampliación de ancho de banda (102),

35 donde, en la etapa de combinación, los datos de umbral de ruido (115) cambian según los datos de distribución de energía (125) para obtener datos de umbral de ruido modificados, los datos de umbral de ruido modificados indican un umbral de ruido modificado que aumenta o disminuye, dependiendo de los datos de distribución de energía, con respecto al umbral de ruido indicado por los datos de umbral de ruido,

40 donde el cambio de los datos de umbral de ruido (115) es tal que el umbral de ruido modificado se incrementa para una señal de audio (105) que comprende un primer grado de sibilancia en comparación con una señal de audio (105) que comprende un segundo grado de sibilancia, donde el segundo grado es menor que el primer grado,

45 donde el método para generar datos de salida por ampliación de ancho de banda (102) ejecuta una decisión externa para determinar si la porción de tiempo (T) de la señal de audio (105) es una señal de voz o una señal que no es de voz,

50 donde los datos de umbral de ruido medidos por el medidor de umbral de ruido (110) son utilizados como los datos de salida por ampliación de ancho de banda, cuando la porción de tiempo (T) de la señal de audio (105) es una señal que no es de voz, y

55 donde, cuando la porción de tiempo (T) de la señal de audio (105) es una señal de voz, se realiza un análisis de discurso adicional para determinar un grado de sibilancia de la señal de voz, y donde los datos de umbral de ruido modificados se agregan a una secuencia de bits como los datos de salida por ampliación de ancho de banda (102), cuando la porción de tiempo (T) de la señal de audio (105) es una señal de voz.

7. Un programa de ordenador adaptado para implementar, al ejecutarse en un ordenador, el método de la reivindicación 6.

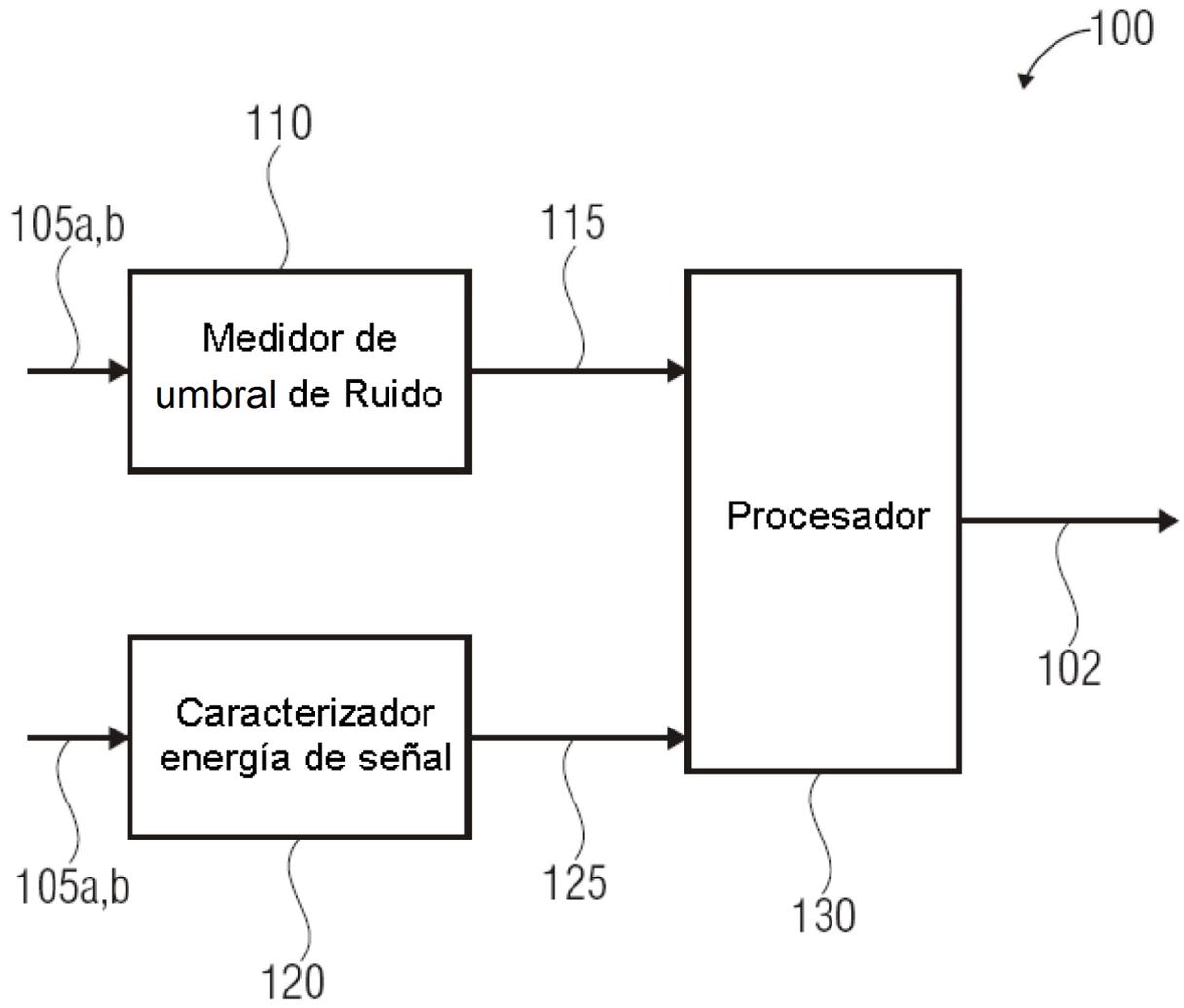


FIG 1

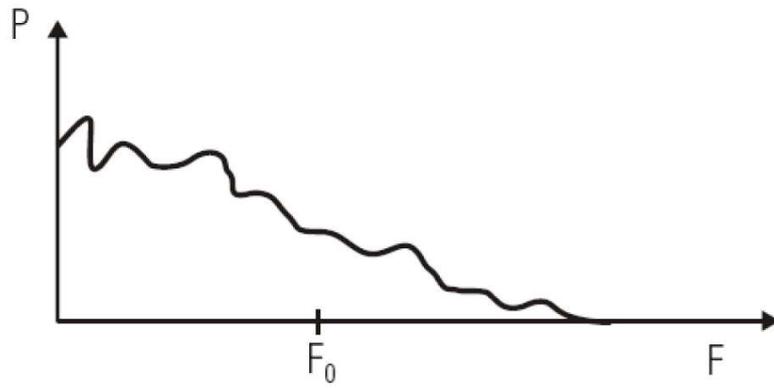


FIG 2A

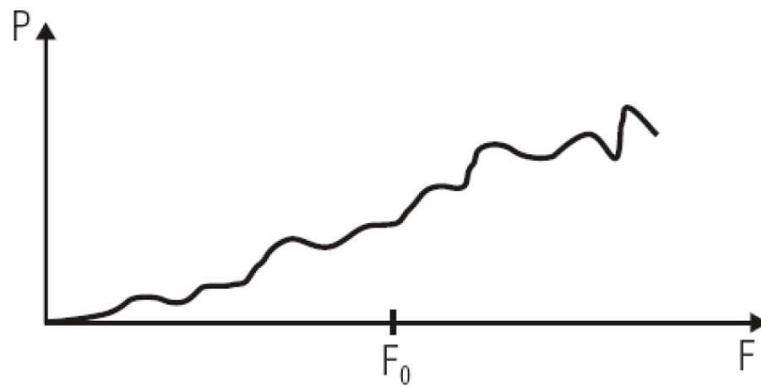


FIG 2B

$$c_k = \frac{1}{k} \sum_{n=1}^N (p_n)^k$$

Coefficiente cepstrales que corresponden al espectro de potencia logarítmico que contiene los polos de orden N.

$$m = \frac{-48}{\pi^3} \sum_{k=1,3,5,\dots}^{\infty} \left\{ \frac{1}{k^3} \sum_{n=1}^N (p_n)^k \right\}$$

Desplazamiento espectral en términos de coeficientes cepstrales

$$S(\omega) = \ln |H(e^{j\omega})|^2 = \ln G^2 - \ln \left| 1 - \sum_{k=1}^N a_k e^{-j\omega k} \right|^2$$

Espectro de potencia log. de filtro LPC de orden N

$$c_k = \begin{cases} \alpha_k + \frac{1}{k} \sum_{n=1}^{k-1} n c_n \alpha_{k-n}, & 1 \leq k \leq N; \\ \frac{1}{k} \sum_{n=k-N}^{k-1} n c_n \alpha_{k-n}, & k > N. \end{cases}$$

Coefficientes cepstrales c_k en dependencia de coeficientes LPC α_k α_1 : primer coeficiente con signo positivo ó negativo

FIG 2C

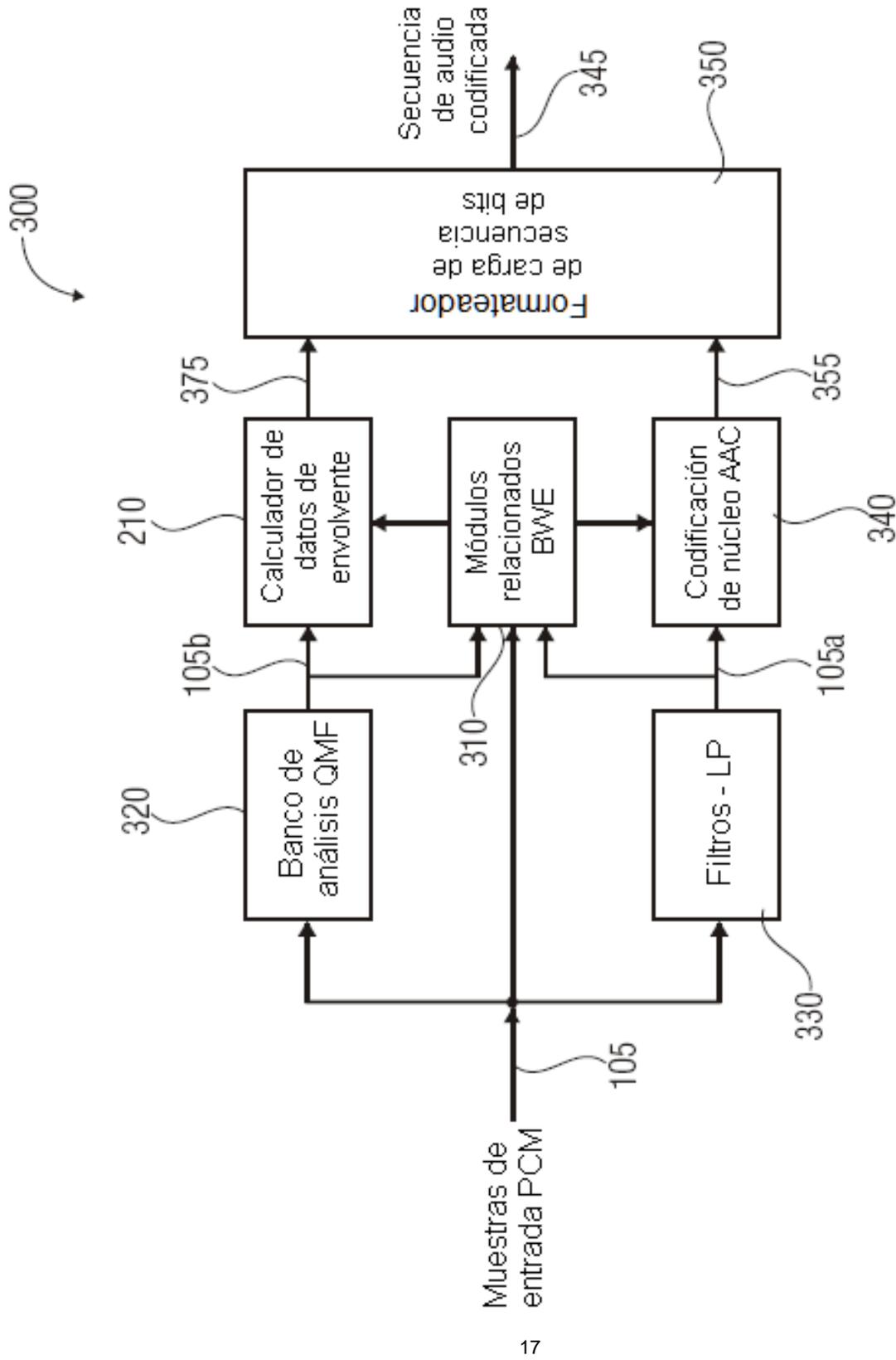


FIG 3

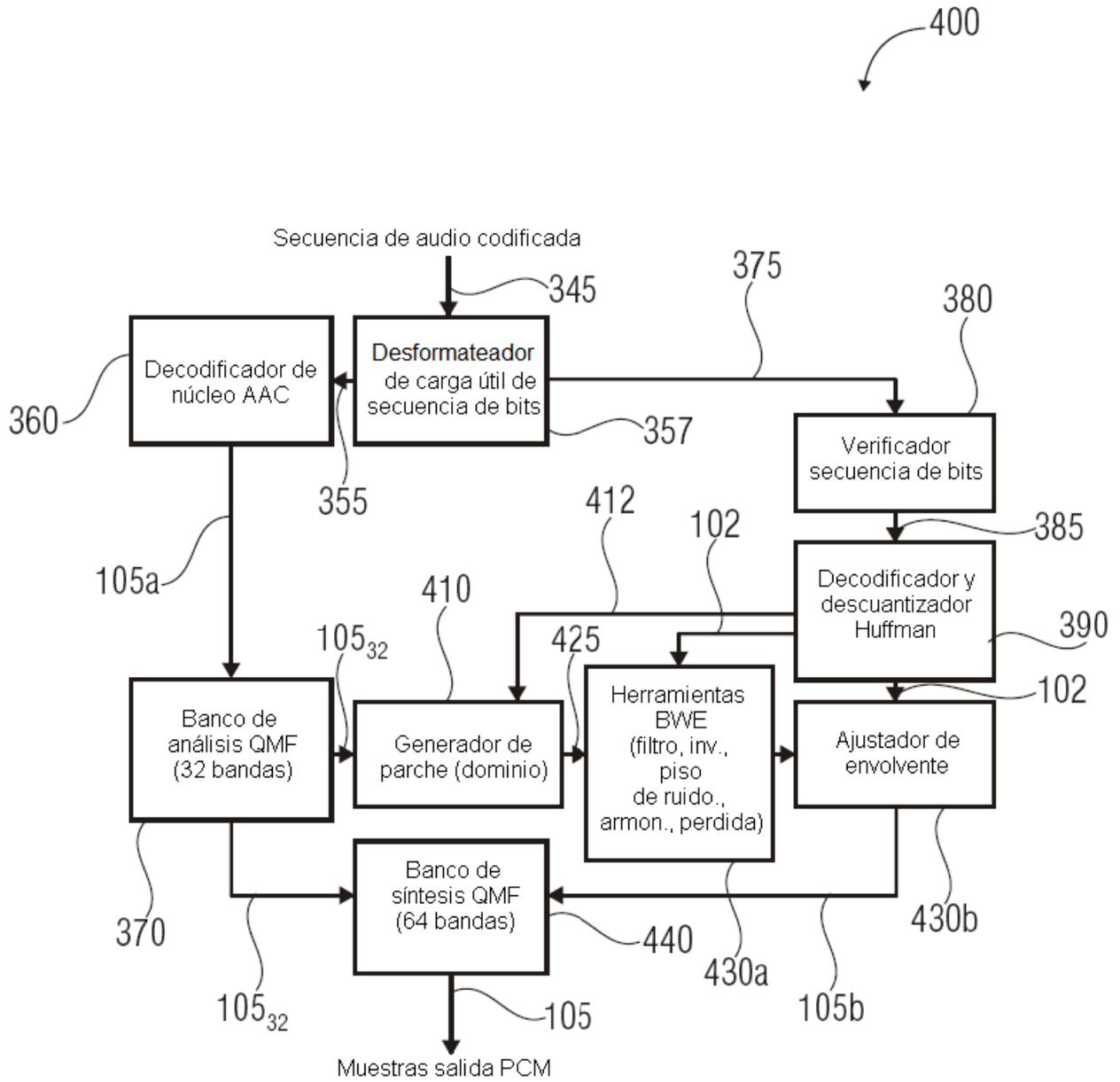


FIG 4

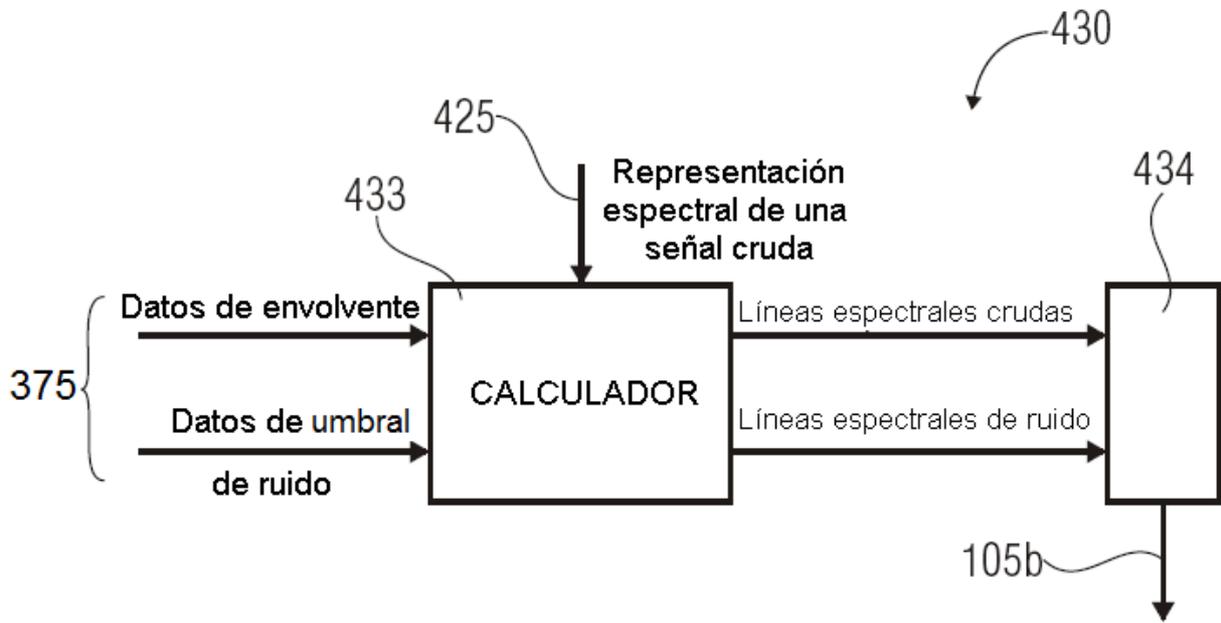


FIG 5A

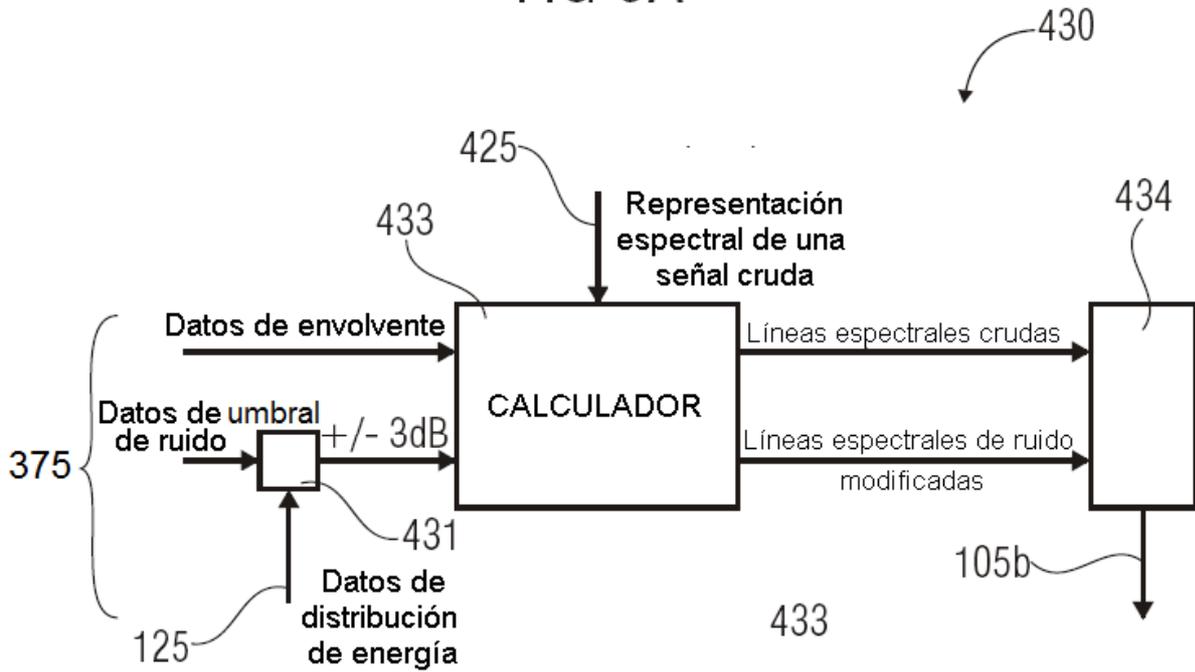


FIG 5B

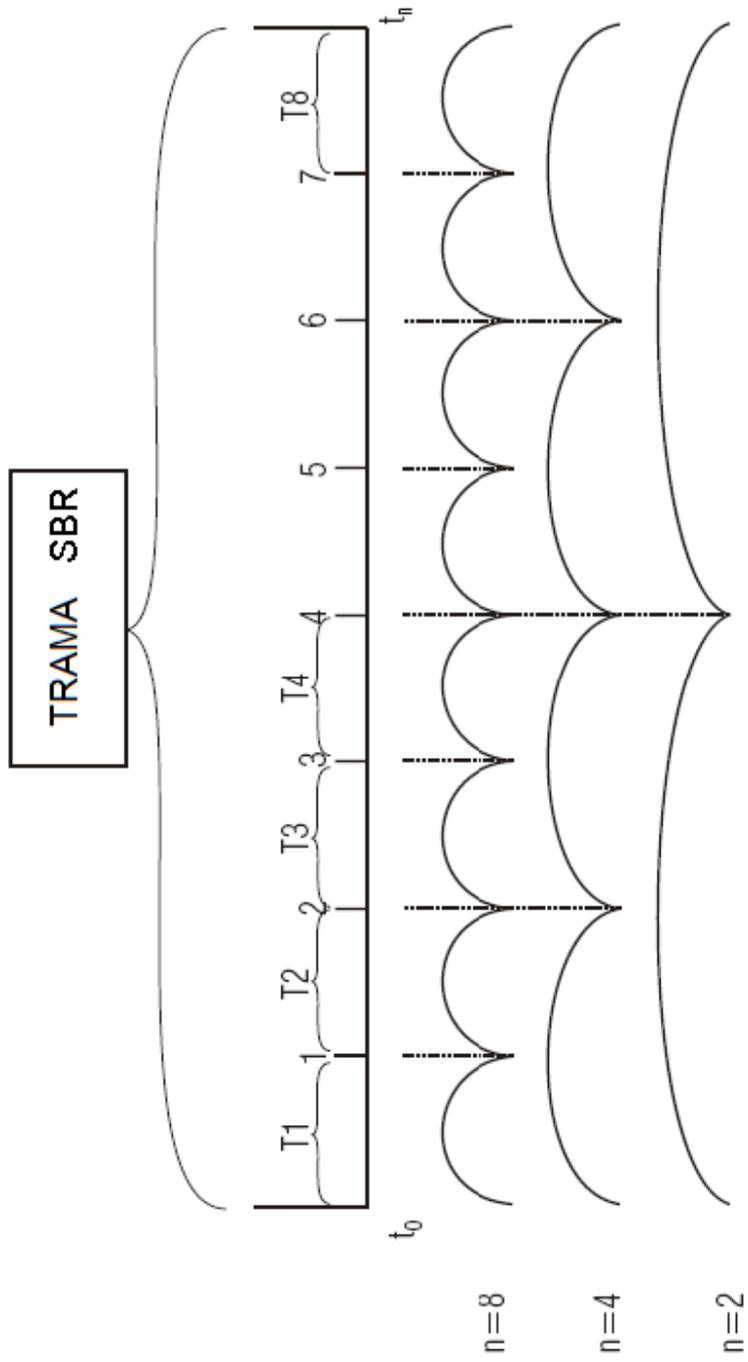


FIG 6