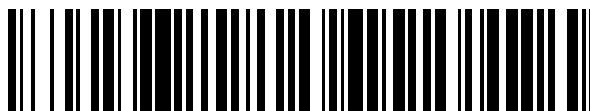


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 732 560**

51 Int. Cl.:

**G10L 19/028** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.01.2014 PCT/EP2014/051649**

87 Fecha y número de publicación internacional: **07.08.2014 WO14118192**

96 Fecha de presentación y número de la solicitud europea: **28.01.2014 E 14701567 (1)**

97 Fecha y número de publicación de la concesión europea: **27.03.2019 EP 2951816**

54 Título: **Llenado de ruido sin información secundaria para codificadores tipo CELP**

30 Prioridad:

**29.01.2013 US 201361758189 P**

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

**25.11.2019**

73 Titular/es:

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

72 Inventor/es:

**FUCHS, GUILLAUME;  
HELMRICH, CHRISTIAN;  
JANDER, MANUEL;  
SCHUBERT, BENJAMIN y  
YOKOTANI, YOSHIKAZU**

74 Agente/Representante:

**SALVÀ FERRER, Joan**

**ES 2 732 560 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

Llenado de ruido sin información secundaria para codificadores tipo CELP.

5 Campo técnico

**[0001]** Las realizaciones de la invención se refieren a un decodificador de audio para proporcionar una información de audio decodificada basándose en una información de audio codificada que comprende coeficientes de predicción lineal (LPC), a un procedimiento para proporcionar una información de audio decodificada basándose en una información de audio codificada que comprende coeficientes de predicción lineal (LPC), a un programa informático para ejecutar dicho procedimiento, en el que el programa informático se ejecuta en un ordenador, y a una señal de audio o un medio de almacenamiento que tiene almacenada dicha señal de audio, habiendo sido tratada la señal de audio con dicho procedimiento.

15 Antecedentes de la invención

**[0002]** Los codificadores de voz digitales de baja velocidad binaria basados en el principio de codificación de predicción lineal excitada por código (CELP) en general padecen artefactos por dispersión de señal cuando la velocidad binaria cae por debajo de aproximadamente 0,5 a 1 bit por muestra, lo que conduce a un sonido algo artificial, metálico. En especial, cuando la voz de entrada tiene ruido ambiental en el fondo, los artefactos de baja velocidad son claramente audibles: el ruido de fondo se atenúa durante secciones de voz activa. La presente invención describe un esquema de inserción de ruido para codificadores del tipo (A)CELP tales como AMR-WB [1] y G.718 [4, 7] los cuales, de forma análoga a las técnicas de llenado de ruido usadas en codificadores basados en transformadas tales como xHE-AAC [5, 6], añade la salida de un generador de ruido aleatorio a la señal de voz decodificada para reconstruir el ruido de fondo.

**[0003]** La publicación internacional WO 2012/110476 A1 muestra un concepto de codificación que se basa en predicción lineal y usa la conformación de ruido en el dominio espectral. Se usa una descomposición espectral de una señal de audio de entrada en un espectrograma que comprende una secuencia de espectros tanto para el cálculo de coeficientes de predicción lineal como para la introducción de conformación en el dominio de la frecuencia basándose en los coeficientes de predicción lineal. Según el documento citado, un codificador de audio comprende un analizador de predicción lineal para analizar una señal de entrada con el fin de deducir los coeficientes de predicción lineal. Un conformador en el dominio de la frecuencia de un codificador de audio está configurado para formar espectralmente un espectro actual de la secuencia de espectros del espectrograma basándose en los coeficientes de predicción lineal proporcionados por el analizador de predicción lineal. Se inserta un espectro cuantificado y conformado espectralmente en un tren de datos junto con información sobre los coeficientes de predicción lineal usados en la conformación espectral de modo que, en el lado de decodificación, se puede realizar la desconformación y la descuantificación. También puede haber un módulo de conformación de ruido temporal para realizar una conformación de ruido temporal.

**[0004]** El artículo "ITU-T Recommendation G.729 Annex B: A Silence Compression Scheme for Use with G.729 Optimized for V.70 Digital Simultaneous Voice and Data Applications" de A. Benyassine y col. describe el Anexo B de ITU-T G.729. El Anexo B define un esquema de compresión de silencio de baja velocidad binaria definido y optimizado para que actúe junto con la versión completa de G.729 y con su Anexo A de baja complejidad. Para conseguir compresión de silencios de baja velocidad binaria y alta calidad es esencial un detector de actividad de voz basado en tramas robusto para detectar tramas de voz inactivas, también denominadas tramas de silencio o de ruido de fondo. Para estas tramas de voz inactiva detectadas, un módulo de transmisión discontinua mide los cambios con el tiempo de las características de la señal de voz inactiva y decide si debe enviarse una nueva trama de descriptor de información de silencio para mantener la calidad de reproducción del ruido de fondo en el extremo receptor. Si se necesita dicha trama, se codifican de manera eficiente los parámetros de espectro y energía que describen las características perceptuales del ruido de fondo y se transmiten usando 15b/trama.

En el extremo receptor, el módulo de generación de ruido de confort regenera el ruido de fondo de salida usando datos transmitidos o parámetros disponibles previamente. El ruido de fondo sintetizado se obtiene mediante filtrado de predicción lineal de una señal de excitación de seudoblancos generada localmente de un nivel controlado. Este procedimiento permite conseguir ahorros de velocidad binaria para una voz codificada a velocidades medias de apenas 4 kb/s durante una conversación de habla normal a la vez que se mantiene la calidad de reproducción.

**[0005]** A la vista de la técnica anterior, sigue existiendo una demanda de un decodificador de audio mejorado, un procedimiento mejorado, un programa informático mejorado para ejecutar dicho procedimiento y una señal de audio o un medio de almacenamiento que tenga dicha señal de audio almacenada, habiendo sido la señal de audio tratada con dicho procedimiento. Más específicamente, es deseable encontrar soluciones que mejoren la calidad de sonido de la información de audio transferida en el tren de bits codificado.

65

Resumen de la invención

**[0006]** Las realizaciones según la presente invención se definen por las reivindicaciones adjuntas.

5 **[0007]** Los signos de referencia en las realizaciones y en la descripción detallada de las realizaciones de la invención se añadieron simplemente para mejorar la legibilidad y en modo alguno pretenden ser limitativos.

**[0008]** La invención sugiere un decodificador de audio para proporcionar una información de audio decodificada basándose en una información de audio codificada que comprende coeficientes de predicción lineal (LPC),  
 10 comprendiendo el decodificador de audio un estimador de nivel de ruido configurado para estimar un nivel de ruido para una trama actual usando un coeficiente de predicción lineal de al menos una trama anterior para obtener una información de ruido y un dispositivo de inserción de ruido configurado para añadir un ruido a la trama actual dependiendo de la información de nivel de ruido obtenida por el estimador de nivel de ruido. Además, el objetivo de la invención se resuelve mediante un procedimiento para proporcionar una información de audio decodificada basándose  
 15 en una información de audio codificada que comprende coeficientes de predicción lineal (LPC), comprendiendo el procedimiento la estimación de un nivel de ruido para una trama actual que usa coeficientes de predicción lineal de al menos una trama anterior para obtener una información de nivel de ruido y la adición de un ruido a la trama actual dependiendo de la información de nivel de ruido proporcionada por la estimación de nivel de ruido. Adicionalmente, el objetivo de la invención se resuelve mediante un programa informático para ejecutar dicho procedimiento, en el que  
 20 el programa informático se ejecuta en un ordenador, y una señal de audio o un medio de almacenamiento que tiene dicha señal de audio almacenada, habiendo sido la señal de audio tratada con dicho procedimiento.

Las soluciones sugeridas evitan tener que proporcionar una información secundaria en el tren de bits de CELP para  
 25 ajustar el ruido proporcionado en el lado del decodificador durante un proceso de llenado de ruido. Esto significa que la cantidad de datos que se transportarán con el tren de bits puede reducirse mientras que la calidad del ruido insertado puede aumentarse simplemente basándose en coeficientes de predicción lineal de tramas decodificadas actualmente o previamente. En otras palabras, se puede omitir información secundaria concerniente al ruido que aumentaría la cantidad de datos que se transferirán con el tren de bits. La invención permite proporcionar un codificador digital de  
 30 baja velocidad binaria y un procedimiento que puede consumir menos anchura de banda relativa al tren de bits y proporcionar una calidad mejorada del ruido de fondo en comparación con soluciones de la técnica anterior.

**[0009]** En una realización de la invención, el decodificador de audio comprende un estimador de nivel de ruido configurado para estimar un nivel de ruido para una trama actual usando un coeficiente de predicción lineal de al  
 35 menos una trama anterior para obtener una información de nivel de ruido, y un dispositivo de inserción de ruido configurado para añadir un ruido a la trama actual dependiendo de la información de nivel de ruido proporcionada por el estimador de nivel de ruido. De este modo, se puede mejorar la calidad del ruido de fondo y por ende la calidad de toda la transmisión de audio, ya que el ruido que se añadirá a la trama actual puede ajustarse según el nivel de ruido que probablemente está presente en la trama actual. Por ejemplo, si se espera un alto nivel de ruido en la trama actual porque se estimó un alto nivel de ruido de las tramas anteriores, el dispositivo de inserción de ruido puede estar  
 40 configurado de manera que aumente el nivel del ruido que se añadirá a la trama actual antes de añadirlo a la trama actual. Así, el ruido que se añadirá puede ser ajustado de manera que no sea demasiado silencioso ni demasiado alto en comparación con el nivel de ruido esperado en la trama actual. Este ajuste, de nuevo, no se basa en información secundaria específica en el tren de bits sino que usa simplemente información de datos necesarios transferidos en el tren de bits, en este caso, un coeficiente de predicción lineal que también proporciona información acerca de un nivel  
 45 de ruido en una trama anterior.

**[0010]** En algunas realizaciones, el nivel de ruido que se añadirá a la trama actual se ajusta también cuando la trama actual es de un tipo de audio general, por ejemplo de un tipo TCX o DTX.

50 **[0011]** Preferentemente, el decodificador de audio comprende un determinador de tipo de trama para determinar un tipo de trama de la trama actual, estando el determinador de tipo de trama configurado para identificar si el tipo de trama de la trama actual es voz o audio general, de manera que la estimación de nivel de ruido se pueda realizar dependiendo del tipo de trama de la trama actual. Por ejemplo, el determinador de tipo de trama puede estar configurado para detectar si la trama actual es una trama CELP o ACELP, que es un tipo de trama de voz, o una trama  
 55 TCX/MDCT o DTX, que son tipos de tramas de audio generales. Como estos formatos de codificación siguen principios diferentes, es conveniente determinar el tipo de trama antes de realizar la estimación de nivel de ruido de manera que se pueden elegir cálculos adecuados, dependiendo del tipo de trama.

**[0012]** En algunas realizaciones de la invención el decodificador de audio está adaptado para calcular una  
 60 primera información que representa una excitación espectralmente no conformada de la trama actual y para calcular una segunda información con relación al ajuste a escala de la trama actual con el fin de calcular un cociente de la primera información y la segunda información para obtener la información de nivel de ruido. Con esto, se puede obtener la información de nivel de ruido sin hacer uso de ninguna información secundaria. Así, se puede mantener baja la velocidad binaria.

65

- 5 **[0013]** Preferentemente, el decodificador de audio está adaptado para decodificar una señal de excitación de la trama actual y para calcular su media cuadrática  $e_{rms}$  a partir de la representación en el dominio del tiempo de la trama actual como la primera información para obtener la información de nivel de ruido con la condición de que la trama actual es de tipo voz. Se prefiere para esta realización que el decodificador de audio esté adaptado para actuar de forma consiguiente si la trama es un tipo CELP o ACELP. La señal de excitación aplanada espectralmente (en el dominio perceptual) es decodificada a partir de un tren de bits y se usa para actualizar una estimación de nivel de ruido. La media cuadrática  $e_{rms}$  de la señal de excitación para la trama actual se calcula después de haber leído el tren de bits. Este tipo de cálculo puede no necesitar una alta potencia de cálculo y así puede ser realizado incluso por decodificadores de audio con baja potencia de cálculo.
- 10 **[0014]** En una realización preferida el decodificador de audio está adaptado para calcular un nivel de pico  $p$  de una función de transferencia de un filtro LPC de la trama actual como una segunda información, usando así un coeficiente de predicción lineal para obtener la información de nivel de ruido con la condición de que la trama actual sea de tipo voz. De nuevo, se prefiere que la trama actual sea del tipo CELP o ACELP. Calcular el nivel de pico  $p$  es más bien poco costoso, y reutilizando coeficientes de predicción lineal de la trama actual, que se usan también para decodificar la información de audio contenida en esa trama, se puede omitir información secundaria y todavía se puede mejorar el ruido de fondo sin aumentar la velocidad de transmisión de datos del tren de bits.
- 15 **[0015]** En una realización preferida de la invención, el decodificador de audio está adaptado para calcular un mínimo espectral  $m_f$  de la trama de audio actual calculando el cociente de la media cuadrática  $e_{rms}$  y el nivel de pico  $p$  para obtener la información de nivel de ruido con la condición de que la trama actual sea de tipo voz. Este cálculo es bastante sencillo y puede proporcionar un valor numérico que puede ser útil para estimar el nivel de ruido en un intervalo de múltiples tramas de audio. Así, se puede usar el mínimo espectral  $m_f$  de una serie de tramas de audio actuales para estimar el nivel de ruido durante el periodo de tiempo cubierto por esa serie de tramas de audio. Esto puede permitir obtener una buena estimación de un nivel de ruido de una trama actual mientras que mantiene razonablemente baja la complejidad. El nivel de pico  $p$  se calcula preferentemente usando la fórmula  $p = \sum |a_k|$ , en la que  $a_k$  son coeficientes de predicción lineal con  $k = 0 \dots 15$ , preferentemente. Así, si la trama comprende 16 coeficientes de predicción lineal, en algunas realizaciones,  $p$  se calcula sumando las amplitudes de los 16  $a_k$  preferentemente.
- 20 **[0016]** Preferentemente, el decodificador de audio está adaptado para decodificar una excitación MDCT no conformada de la trama actual y para calcular su media cuadrática  $e_{rms}$  a partir de la representación en el dominio espectral de la trama actual para obtener la información de nivel de ruido como la primera información si la trama actual es de un tipo de audio general. Esta es la realización preferida de la invención, siempre que la trama actual no sea una trama de voz sino una trama de audio general. Una representación en el dominio espectral en tramas MDCT o DTX es en gran parte equivalente a la representación en el dominio espectral en tramas de voz, por ejemplo, tramas CELP o (A)CELP. Una diferencia reside en que MDCT no tiene en cuenta el teorema de Parseval. Así, preferentemente, la media cuadrática  $e_{rms}$  para una trama de audio general se calcula de una manera similar a la media cuadrática  $e_{rms}$  para tramas de voz. Entonces se prefiere calcular los coeficientes LPC equivalentes de la trama de audio general como se presenta en el documento WO 2012/110476 A1, por ejemplo usando un espectro de potencia de MDCT que se refiere al cuadrado de valores de MDCT en una escala de Bark. En una realización alternativa, las bandas de frecuencia del espectro de potencia MDCT pueden tener una anchura constante de modo que la escala del espectro corresponda a una escala lineal. Con dicha escala lineal los equivalentes de coeficientes LPC calculados son similares a un coeficiente LPC en la representación en el dominio del tiempo de la misma trama, como, por ejemplo, calculados para una trama ACELP o CELP. Además, se prefiere que, si la trama actual es de un tipo de audio general, el nivel de pico  $p$  de la función de transferencia de un filtro LPC de la trama actual que está siendo calculado a partir de la trama MDCT como se presenta en el documento WO 2012/110476 A1, sea calculado como una segunda información, usando así un coeficiente de predicción lineal para obtener la información de nivel de ruido con la condición de que la trama actual sea de un tipo de audio general. A continuación, si la trama actual es de un tipo de audio general, se prefiere calcular el mínimo espectral de la trama de audio actual calculando el cociente de la media cuadrática  $e_{rms}$  y el nivel de pico  $p$  para obtener la información de nivel de ruido con la condición de que la trama actual sea del tipo de audio general. Así, se puede obtener un cociente que describe el mínimo espectral  $m_f$  de una trama de audio actual con independencia de si la trama actual es de tipo voz o de un tipo de audio general.
- 25 **[0017]** En una realización preferida, el decodificador de audio está adaptado para poner en cola el cociente obtenido a partir de la trama de audio actual en el estimador de nivel de ruido con independencia del tipo de trama, comprendiendo el estimador de nivel de ruido un almacenamiento de nivel de ruido para dos o más cocientes obtenidos a partir de diferentes tramas de audio. Esto puede ser ventajoso si el decodificador de audio está adaptado para conmutar entre decodificación de tramas de voz y decodificación de tramas de audio general, por ejemplo, cuando se aplica una decodificación de bajo retardo unificada de voz y audio (LD-USAC, EVS). Con esto, se puede obtener un nivel de ruido promedio sobre múltiples tramas, con independencia del tipo de trama. Preferentemente, un almacenamiento de nivel de ruido puede contener diez o más cocientes obtenidos a partir de diez o más tramas de audio anteriores. Por ejemplo, el almacenamiento de nivel de ruido puede contener espacio para los cocientes de 30 tramas. Así, se puede calcular el nivel de ruido durante un tiempo extendido anterior a la trama actual. En algunas realizaciones, el cociente sólo puede ser puesto en cola en el estimador de nivel de ruido cuando se detecta que la
- 30  
35  
40  
45  
50  
55  
60  
65

trama actual es de tipo voz. En otras realizaciones, el cociente sólo puede ser puesto en cola en el estimador de nivel de ruido cuando se detecta que la trama actual es de un tipo de audio general.

**[0018]** Se prefiere que el estimador de nivel de ruido esté adaptado para estimar el nivel de ruido basándose en el análisis estadístico de dos o más cocientes de tramas de audio diferentes. En una realización de la invención, el decodificador de audio está adaptado para usar un rastreo de densidad espectral de potencia de ruido basado en un error cuadrático medio mínimo para analizar estadísticamente los cocientes. Este rastreo se describe en la publicación de Hendriks, Heusdens y Jensen [2]. Si se va a aplicar el procedimiento según [2], el codificador de audio está adaptado para usar una raíz cuadrada de un valor de rastreo en el análisis estadístico, ya que en el presente caso se investiga directamente el espectro de amplitud. En otra realización de la invención, se usa estadística de mínimos como se conoce a partir de [3] para analizar los dos o más cocientes de diferentes tramas de audio.

**[0019]** En una realización preferida, el decodificador de audio comprende un núcleo configurado para decodificar una información de audio de la trama actual usando un coeficiente de predicción lineal de la trama actual para obtener una señal de salida de codificador de núcleo decodificada y el dispositivo de inserción de ruido añade el ruido dependiendo de un coeficiente de predicción lineal usado para decodificar la información de audio de la trama actual y/o usado cuando se decodifica la información de audio de una o más tramas anteriores. Así, el dispositivo de inserción de ruido hace uso de algunos coeficientes de predicción lineal que son usados para decodificar la información de audio de la trama actual. Se puede omitir la información secundaria para dar instrucciones al dispositivo de inserción de ruido.

**[0020]** Preferentemente, el decodificador de audio comprende un filtro de desacentuación para desacentuar la trama actual, estando el decodificador de audio adaptado para aplicar el filtro de desacentuación a la trama actual después de que el dispositivo de inserción de ruido añadió el ruido a la trama actual. Como la desacentuación es un IIR de primer orden que estimula bajas frecuencias, permite filtrado de paso alto de IIR pronunciado de baja complejidad del ruido añadido evitando artefactos de ruido audibles a bajas frecuencias.

**[0021]** Preferentemente, el decodificador de audio comprende un generador de ruido, estando el generador de ruido adaptado para generar el ruido que se añadirá a la trama actual mediante el dispositivo de inserción de ruido. Tener un generador de ruido incluido en el decodificador de audio puede proporcionar un decodificador de audio más conveniente ya que no se necesita un generador de ruido externo. Como alternativa, el ruido puede ser suministrado por un generador de ruido externo, que puede ser conectado al decodificador de audio mediante una interfaz. Por ejemplo, se pueden aplicar tipos especiales de generadores de ruido, dependiendo del ruido de fondo que se ha de mejorar en la trama actual.

**[0022]** Preferentemente, el generador de ruido está configurado para generar un ruido blanco aleatorio. Dicho ruido se asemeja a ruidos de fondo comunes adecuadamente y dicho generador de ruido puede ser proporcionado fácilmente.

**[0023]** En una realización preferida de la invención, el dispositivo de inserción de ruido está configurado para añadir el ruido a la trama actual con la condición de que la velocidad binaria de la información de audio codificada sea menor que 1 bit por muestra. Preferentemente, la velocidad binaria de la información de audio codificada es menor que 0,8 bits por muestra. Se prefiere más todavía que el dispositivo de inserción de ruido esté configurado para añadir el ruido a la trama actual con la condición de que la velocidad binaria de la información de audio codificada sea menor que 0,5 bits por muestra.

**[0024]** En una realización preferida, el decodificador de audio está configurado para usar un codificador basado en uno o más de los codificadores AMR-WB, G.718 o LD-USAC (EVS) para decodificar la información de audio codificada. Esos son codificadores (A)CELP muy conocidos y ampliamente difundidos en los cuales el uso adicional de dicho procedimiento de llenado de ruido puede ser altamente ventajoso.

#### Breve descripción de los dibujos

**[0025]** A continuación se describen realizaciones de la presente invención con respecto a las figuras.

la fig. 1 muestra una primera realización de un decodificador de audio según la presente invención;

la fig. 2 muestra un primer procedimiento para realizar decodificación de audio según la presente invención que puede ser ejecutado por un decodificador de audio según la fig. 1;

la fig. 3 muestra una segunda realización de un decodificador de audio según la presente invención;

la fig. 4 muestra un segundo procedimiento para realizar decodificación de audio según la presente invención que puede ser ejecutado por un decodificador de audio según la fig. 3;

la fig. 5 muestra una tercera realización de un decodificador de audio según la presente invención;

la fig. 6 muestra un tercer procedimiento para realizar decodificación de audio según la presente invención que puede ser ejecutado por un decodificador de audio según la fig. 5;

5

la fig. 7 muestra una ilustración de un procedimiento para calcular los mínimos espectrales  $m_f$  para las estimaciones de nivel de ruido;

la fig. 8 muestra un diagrama que ilustra una inclinación deducida de los coeficientes LPC; y

10

la fig. 9 muestra un diagrama que ilustra cómo se determinan equivalentes de filtro LPC a partir de un espectro de potencia MDCT.

#### Descripción detallada de realizaciones de la invención

15

**[0026]** A continuación se describe la invención en detalle con relación a las figuras 1 a 9. En ningún modo la invención se limita a las realizaciones mostradas y descritas.

**[0027]** La fig. 1 muestra una primera realización de un decodificador de audio según un ejemplo. El decodificador de audio está adaptado para proporcionar una información de audio decodificada basándose en una información de audio codificada. El decodificador de audio está configurado de manera que use un codificador que puede funcionar en AMR-WB, G.718 y LD-USAC (EVS) para decodificar la información de audio codificada. La información de audio codificada comprende coeficientes de predicción lineal (LPC), que pueden ser designados individualmente como coeficientes  $a_k$ . El decodificador de audio comprende un ajustador de inclinación configurado para ajustar una inclinación de un ruido usando coeficientes de predicción lineal de una trama actual para obtener una información de inclinación y un dispositivo de inserción de ruido configurado para añadir el ruido a la trama actual dependiendo de la información de inclinación obtenida por el calculador de inclinación. El dispositivo de inserción de ruido está configurado para añadir el ruido a la trama actual con la condición de que la velocidad binaria de la información de audio codificada sea menor que 1 bit por muestra. Asimismo, el dispositivo de inserción de ruido puede estar configurado para añadir el ruido a la trama actual con la condición de que la trama actual sea una trama de voz. Así, el ruido puede añadirse a la trama actual para mejorar la calidad de sonido global de la información de audio decodificada que puede estar deteriorada debido a artefactos de codificación, en especial con relación a ruido de fondo de la información de voz. Cuando se ajusta la inclinación del ruido en vista de la inclinación de la trama de audio actual, la calidad de sonido global puede mejorarse sin depender de información secundaria en el tren de bits. Así, se puede reducir la cantidad de datos que se transferirán con el tren de bits.

35

**[0028]** La fig. 2 muestra un primer procedimiento para realizar decodificación de audio según la presente invención que puede ser ejecutado por un decodificador de audio según la fig. 1. Se describen detalles técnicos del decodificador de audio representado en la fig. 1 junto con las características del procedimiento. El decodificador de audio está adaptado para leer el tren de bits de la información de audio codificada. El decodificador de audio comprende un determinador de tipo de trama para determinar un tipo de trama de la trama actual, estando el determinador de tipo de trama configurado para activar el ajustador de inclinación con el fin de ajustar la inclinación del ruido cuando se detecta que el tipo de trama de la trama actual es de tipo voz. Así, el decodificador de audio determina el tipo de trama de la trama de audio actual aplicando el determinador de tipo de trama. Si la trama actual es una trama ACELP, el determinador de tipo trama activa el ajustador de inclinación. El ajustador de inclinación está configurado para usar un resultado del análisis de primer orden de los coeficientes de predicción lineal de la trama actual con el fin de obtener la información de inclinación. Más específicamente, el ajustador de inclinación calcula una ganancia  $g$  usando la fórmula  $g = \sum [a_k \cdot a_{k+1}] / \sum [a_k \cdot a_k]$  como un análisis de primer orden, en el que  $a_k$  son coeficientes LPC de la trama actual. La fig. 8 muestra diagramas que ilustran una inclinación derivada a partir de los coeficientes LPC. La fig. 8 muestra dos tramas de la palabra "see". Para la letra "s", que tiene una cantidad elevada de altas frecuencias, la inclinación apunta hacia arriba. Para las letras "ee", que tienen una cantidad elevada de bajas frecuencias, la inclinación apunta hacia abajo. La inclinación espectral mostrada en la fig. 8 es la función de transferencia del filtro de forma directa  $x(n) - g \cdot x(n-1)$ , estando  $g$  definida según se indica anteriormente. Así, el ajustador de inclinación hace uso de los coeficientes LPC proporcionados en el tren de bits y usados para codificar la información de audio codificada. Por consiguiente, se puede omitir información secundaria lo cual puede reducir la cantidad de datos que se transferirán con el tren de bits. Asimismo, el ajustador de inclinación está configurado para obtener la información de inclinación usando un cálculo de una función de transferencia de filtro de forma directa  $x(n) - g \cdot x(n-1)$ . Por consiguiente, el ajustador de inclinación calcula la inclinación de la información de audio en la trama actual calculando la función de transferencia del filtro de forma directa  $x(n) - g \cdot x(n-1)$  usando la ganancia  $g$  calculada previamente. Después de obtener la información de inclinación, el ajustador de inclinación ajusta la inclinación del ruido que se añadirá a la trama actual dependiendo de la información de inclinación de la trama actual. Después de eso, se añade el ruido ajustado a la trama actual. Además, lo que no se muestra en la fig. 2, el decodificador de audio comprende un filtro de desacentuación para desacentuar la trama actual, estando el decodificador de audio adaptado para aplicar el filtro de desacentuación a la trama actual después de que el dispositivo de inserción de ruido añadió el ruido a la trama actual. Después de desacentuar la trama, lo cual también sirve como un filtrado de paso alto IIR

65

pronunciado de baja complejidad del ruido añadido, el decodificador de audio proporciona la información de audio decodificada. Así, el procedimiento según la fig. 2 permite mejorar la calidad de sonido de una información de audio ajustando la inclinación de un ruido que se añadirá a una trama actual para mejorar la calidad de un ruido de fondo.

5 **[0029]** La fig. 3 muestra una segunda realización de un decodificador de audio según la presente invención. El decodificador de audio nuevamente está adaptado para proporcionar una información de audio decodificada basándose en una información de audio codificada. El decodificador de audio nuevamente está configurado para usar un codificador que funcione en AMR-WB, G.718 y LD-USAC (EVS) para decodificar la información de audio codificada. Nuevamente la información de audio codificada comprende coeficientes de predicción lineal (LPC), que pueden ser designados individualmente como coeficientes  $a_k$ . El decodificador de audio según la segunda realización comprende un estimador de nivel de ruido configurado para estimar un nivel de ruido para una trama actual usando un coeficiente de predicción lineal de al menos una trama anterior para obtener una información de nivel de ruido, y un dispositivo de inserción de ruido configurado para añadir un ruido a la trama actual dependiendo de la información de nivel de ruido proporcionada por el estimador de nivel de ruido. El dispositivo de inserción de ruido está configurado de manera que añade el ruido a la trama actual con la condición de que la velocidad binaria de la información de audio codificada sea menor que 0,5 bits por muestra. Asimismo, el dispositivo de inserción de ruido está configurado de manera que añade el ruido a la trama actual con la condición de que la trama actual sea una trama de voz. Así, nuevamente, el ruido puede ser añadido a la trama actual para mejorar la calidad de sonido global de la información de audio decodificada la cual puede estar deteriorada debido a artefactos de codificación, en especial con relación a ruido de fondo de la información de voz. Cuando se ajusta el nivel del ruido en vista del nivel de ruido de al menos una trama de audio previa, la calidad de sonido global puede mejorarse sin depender de información secundaria en el tren de bits. Así, se puede reducir la cantidad de datos que se transferirán con el tren de bits.

25 **[0030]** La fig. 4 muestra un segundo procedimiento para realizar decodificación de audio según la presente invención, que puede ser ejecutado por un decodificador de audio según la fig. 3. Los detalles técnicos del decodificador de audio representado en la fig. 3 se describen junto con las características del procedimiento. Según la fig. 4, el decodificador de audio está configurado para leer el tren de bits con el fin de determinar el tipo de trama de la trama actual. Además, el decodificador de audio comprende un determinador de tipo de trama para determinar un tipo de trama de la trama actual, estando el determinador de tipo de trama configurado para identificar si el tipo de trama de la trama actual es voz o audio general, de modo que la estimación de nivel de ruido se pueda realizar dependiendo del tipo de trama de la trama actual. En general, el decodificador de audio está adaptado para calcular una primera información que representa una excitación espectralmente no conformada de la trama actual y para calcular una segunda información con relación al ajuste a escala de la trama actual para calcular un cociente de la primera información y la segunda información con el fin de obtener la información de nivel de ruido. Por ejemplo, si el tipo de trama es ACELP, que es un tipo de trama de voz, el decodificador de audio decodifica una señal de excitación de la trama actual y calcula su media cuadrática  $e_{rms}$  para la trama actual  $f$  a partir de la representación en el dominio del tiempo de la señal de excitación. Esto significa que el decodificador de audio está adaptado para decodificar una señal de excitación de la trama actual y para calcular su media cuadrática  $e_{rms}$  a partir de la representación en el dominio del tiempo de la trama actual como la primera información para obtener la información de nivel de ruido con la condición de que la trama actual sea de tipo voz. En otro caso, si el tipo de trama es MDCT o DTX, que es un tipo de trama de audio general, el decodificador de audio decodifica una señal de excitación de la trama actual y calcula su media cuadrática  $e_{rms}$  para la trama actual  $f$  a partir de la representación en el dominio del tiempo equivalente de la señal de excitación. Esto significa que el decodificador de audio está adaptado para decodificar una excitación MDCT no conformada de la trama actual y para calcular su media cuadrática  $e_{rms}$  a partir de la representación en el dominio espectral de la trama actual como la primera información para obtener la información de nivel de ruido con la condición de que la trama actual sea de un tipo de audio general. En el documento WO 2012/110476 A1 se describe en detalle cómo se hace esto. Además, la fig. 9 muestra un diagrama que ilustra cómo se determina un equivalente de filtro LPC a partir de un espectro de potencia MDCT. Aunque la escala representada es una escala de Bark, los equivalentes de coeficientes LPC también pueden obtenerse a partir de una escala lineal. En especial, cuando se obtienen a partir de una escala lineal, los equivalentes de coeficientes LPC son muy similares a los calculados a partir de la representación en el dominio del tiempo de la misma trama, por ejemplo, cuando se codifica en ACELP.

55 **[0031]** Además, el decodificador de audio según la fig. 3, como se ilustra mediante el diagrama de procedimiento de la fig. 4, está adaptado para calcular un nivel de pico  $p$  de una función de transferencia de un filtro LPC de la trama actual como una segunda información, usando así un coeficiente de predicción lineal para obtener la información de nivel de ruido con la condición de que la trama actual sea de tipo voz.

60 **[0032]** Eso significa que el decodificador de audio calcula el nivel de pico  $p$  de la función de transferencia del filtro de análisis LPC de la trama actual  $f$  según la fórmula  $p = \sum |a_k|$ , en la que  $a_k$  es un coeficiente de predicción lineal con  $k = 0 \dots 15$ . Si la trama es una trama de audio general, los equivalentes de coeficientes LPC se obtienen a partir de la representación en el dominio espectral de la trama actual, como se muestra en la fig. 9 y se describe en el documento WO 2012/110476 A1 y anteriormente. Como se aprecia en la fig. 4, después de calcular el nivel de pico  $p$ , se calcula un mínimo espectral  $m_f$  de la trama actual  $f$  dividiendo  $e_{rms}$  por  $p$ . Así, el decodificador de audio está adaptado para calcular una primera información que representa una excitación espectralmente no conformada de una trama actual, en esta realización  $e_{rms}$ , y una segunda información con relación al cambio de escala espectral de la trama

actual, en esta realización el nivel de pico  $p$ , para calcular un cociente de la primera información y la segunda información con el fin de obtener la información de nivel de ruido. A continuación se pone en cola el mínimo espectral de la trama actual en el estimador de nivel de ruido, estando el decodificador de audio adaptado para poner en cola el cociente obtenido a partir de la trama de audio actual en el estimador de nivel de ruido con independencia del tipo de trama y el estimador de nivel de ruido que comprende un almacenamiento de nivel de ruido para dos o más cocientes, en este caso mínimos espectrales  $m_r$ , obtenidos a partir de diferentes tramas de audio. Más específicamente, el almacenamiento de nivel de ruido puede almacenar cocientes de 50 tramas para estimar el nivel de ruido. Además, el estimador de nivel de ruido está adaptado para estimar el nivel de ruido basándose en análisis estadístico de dos o más cocientes de diferentes tramas de audio, y así una colección de mínimos espectrales  $m_r$ . Las etapas para calcular el cociente  $m_r$  están representadas en detalle en la fig. 7 que ilustra las etapas de cálculo necesarias. En la segunda realización, el estimador de nivel de ruido funciona basándose en estadísticas de mínimos como se conoce de [3]. Se cambia de escala el ruido según el nivel de ruido estimado de la trama actual basándose en estadísticas de mínimos y después de ello se añade a la trama actual si la trama actual es una trama de voz. Finalmente, la trama actual se somete a desacentuación (no mostrado en la fig. 4). Así, esta segunda realización también permite omitir información secundaria para llenado de ruido, permitiendo reducir la cantidad de datos que se transferirán con el tren de bits. Por consiguiente, la calidad de sonido de la información de audio puede mejorarse mejorando el ruido de fondo durante la etapa de decodificación sin aumentar la velocidad de datos. Debe observarse que como no se necesitan transformaciones tiempo/frecuencia y como el estimador de nivel de ruido sólo se ejecuta una vez por trama (no en múltiples subbandas), el llenado de ruido descrito muestra muy baja complejidad a la vez que puede mejorar la codificación de baja velocidad binaria de voz con ruido.

**[0033]** La fig. 5 muestra una tercera realización de un decodificador de audio según la presente invención. El decodificador de audio está adaptado para proporcionar una información de audio decodificada basándose en una información de audio codificada. El decodificador de audio está configurado para usar un codificador basado en LD-USAC para decodificar la información de audio codificada. La información de audio codificada comprende coeficientes de predicción lineal (LPC), los cuales pueden ser designados individualmente como coeficientes  $a_k$ . El decodificador de audio comprende un ajustador de inclinación configurado para ajustar una inclinación de un ruido usando coeficientes de predicción lineal de una trama actual para obtener una información de inclinación y un estimador de nivel de ruido configurado para estimar un nivel de ruido para una trama actual usando un coeficiente de predicción lineal de al menos una trama anterior para obtener una información de nivel de ruido. Además, el decodificador de audio comprende un dispositivo de inserción de ruido configurado para añadir el ruido a la trama actual dependiendo de la información de inclinación obtenida por el calculador de inclinación y dependiendo de la información de nivel de ruido proporcionada por el estimador de nivel de ruido. Así, se puede añadir ruido a la trama actual para mejorar la calidad de sonido global de la información de audio decodificada, la cual puede deteriorarse debido a artefactos de codificación, especialmente con relación al ruido de fondo de información de voz, dependiendo de la información de inclinación obtenida por el calculador de inclinación y dependiendo de la información de nivel de ruido proporcionada por el estimador de nivel de ruido. En esta realización, un generador de ruido aleatorio (no mostrado), que está formado por el decodificador de audio, genera un ruido espectralmente blanco, al que después se aplica un cambio de escala según la información de nivel de ruido y es conformado usando la inclinación deducida de  $g$ , como se describió anteriormente.

**[0034]** La fig. 6 muestra un tercer procedimiento para realizar decodificación de audio según la presente invención que puede ser realizado por un decodificador de audio según la fig. 5. Se lee el tren de bits y un determinador de tipo de trama, llamado detector de tipo de trama, determina si la trama actual es una trama de voz (ACELP) o una trama de audio general (TCX/MDCT). Con independencia del tipo de trama, se decodifica el encabezamiento de trama y se decodifica la señal de excitación no conformada, aplanada espectralmente en el dominio perceptual. En caso de una trama de voz, esta señal de excitación es una excitación del dominio del tiempo, como se describió anteriormente. Si la trama es una trama de audio general, se decodifica el dominio MDCT residual (dominio espectral). La representación en el dominio del tiempo y la representación en el dominio espectral se usan respectivamente para estimar el nivel de ruido como se ilustra en la fig. 7 y se describió anteriormente, usando coeficientes LPC también usados para decodificar el tren de bits en lugar de usar información secundaria o coeficientes LPC adicionales. La información de ruido de los dos tipos de tramas se pone en cola para ajustar la inclinación y el nivel de ruido del ruido que se añadirá a la trama actual con la condición de la trama actual sea una trama de voz. Después de añadir el ruido a la trama de voz ACELP (Aplicar llenado de ruido ACELP) la trama de voz ACELP se somete a desacentuación mediante IIR y las tramas de voz y las tramas de audio general se combinan en una señal temporal, que representa la información de audio decodificada. El efecto de paso alto pronunciado de la desacentuación sobre el espectro del ruido añadido está representado por las pequeñas figuras I, II, y III insertadas en la fig. 6.

**[0035]** En otras palabras, según la fig. 6, el sistema de llenado de ruido ACELP descrito anteriormente se implementó en el decodificador LD-USAC (EVS), una variante de bajo retardo del xHE-AAC [6] que puede conmutar entre codificación ACELP (voz) y MDCT (música / ruido) sobre un base de trama. A continuación se resume el proceso de inserción según la fig. 6:

1. Se lee el tren de bits, y se determina si la trama actual es una trama ACELP o MDCT o DTX. Con independencia del tipo de trama, se decodifica la señal de excitación aplanada espectralmente (en dominio perceptual) y se usa para



actualizar la estimación de nivel de ruido como se describe en detalle más adelante. A continuación se reconstruye totalmente la señal hasta la desacentuación, que es la última etapa.

2. Si la trama se codifica con ACELP, se calcula la inclinación (forma espectral global) para la inserción de ruido mediante análisis LPC de primer orden de los coeficientes de filtro LPC. Se deduce la inclinación a partir de la ganancia g de los 16 coeficientes LPC  $a_k$ , que vienen dados por  $g = \sum [a_k \cdot a_{k+1}] / \sum [a_k \cdot a_k]$ .

3. Si la trama se codifica con ACELP, se emplea el nivel de conformación de ruido y la inclinación para realizar la adición de ruido en la trama decodificada: un generador de ruido aleatorio genera la señal de ruido espectralmente blanca, que a continuación se cambia de escala y se conforma usando la inclinación obtenida de g.

4. Se añade la señal de ruido conformada y nivelada para la trama ACELP en la señal decodificada justo antes de finalizar la etapa de filtrado de desacentuación final. Como la desacentuación es un IIR de primer orden que estimula bajas frecuencias, permite filtrado de paso alto de IIR pronunciado del ruido añadido, como en la fig. 6, lo que evita artefactos de ruido audibles a bajas frecuencias.

**[0036]** La estimación de nivel de ruido en la etapa 1 se realiza calculando la media cuadrática  $e_{rms}$  de la señal de excitación para la trama actual (o en caso de una excitación del dominio MDCT, el equivalente del dominio del tiempo, que significa la  $e_{rms}$  que sería calculada para esa trama si fuera una trama ACELP) y a continuación dividiendo por el nivel de pico p de la función de transferencia del filtro de análisis LPC. Esto produce el nivel  $m_f$  del mínimo espectral de la trama f como en la fig. 7. Finalmente,  $m_f$  es puesto en cola en el estimador de nivel de ruido actuando basándose, por ejemplo, en estadísticas de mínimos [3]. Debe observarse que como no se necesitan transformaciones tiempo/frecuencia y como el estimador de nivel sólo se ejecuta una vez por trama (no en múltiples subbandas), el sistema de llenado de ruido CELP descrito muestra muy baja complejidad mientras que puede mejorar la codificación de baja velocidad binaria de voz con ruido.

**[0037]** Aunque se han descrito algunos aspectos en el contexto de un decodificador de audio, está claro que estos aspectos también representan una descripción del procedimiento correspondiente, en el que un bloque o dispositivo corresponde a una etapa de procedimiento o a una característica de una etapa de procedimiento. Análogamente, los aspectos descritos en el contexto de una etapa de procedimiento también representan una descripción de un bloque o componente o característica correspondiente de un decodificador de audio correspondiente. Algunas o todas las etapas de procedimiento pueden llevarse a cabo con (o usando) un aparato de hardware, como por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, alguna o varias de las etapas de procedimiento más importantes pueden llevarse a cabo mediante dicho aparato.

**[0038]** La señal de audio codificada de la invención puede ser almacenada en un medio de almacenamiento digital o puede ser transmitida a través de un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión físico como Internet.

**[0039]** Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden ser implementadas en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un CD, un Blu-Ray, una memoria de sólo lectura, una PROM, una EEPROM o una memoria FLASH, que tienen señales de control legibles electrónicamente almacenadas en los mismos, que cooperan (o son capaces de cooperar) con un sistema informático programable de manera que se ejecute el procedimiento respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

**[0040]** Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control legibles electrónicamente, las cuales son capaces de cooperar con un sistema informático programable, de manera que se lleve a cabo uno de los procedimientos descritos en la presente memoria.

**[0041]** En general, realizaciones de la presente invención pueden implementarse en forma de un programa informático con un código de programa, estando el código de programa operativo para ejecutar uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede almacenarse, por ejemplo, en un soporte legible por una máquina.

**[0042]** Otras realizaciones comprenden el programa informático para ejecutar uno de los procedimientos descritos en la presente memoria, almacenado en un soporte legible por una máquina.

**[0043]** En otras palabras, una realización del procedimiento de la invención es, por lo tanto, un programa informático que tiene un código de programa para ejecutar uno de los procedimientos descritos en la presente memoria, cuando el programa informático se ejecuta en un ordenador.

**[0044]** 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 el mismo, el

programa informático para ejecutar uno de los procedimientos descritos en la presente memoria. El soporte de datos, el medio de almacenamiento digital o el medio de registro son normalmente tangibles y/o no son transitorios.

5 **[0045]** Una realización adicional del procedimiento de la invención es, por lo tanto, un tren de datos o una secuencia de señales que representan el programa informático para ejecutar uno de los procedimientos descritos en la presente memoria. El tren de datos o la secuencia de señales pueden configurarse, por ejemplo, para su transferencia mediante una conexión de comunicación de datos, por ejemplo, por Internet.

10 **[0046]** Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado o adaptado para ejecutar uno de los procedimientos descritos en la presente memoria.

15 **[0047]** Una realización adicional comprende un ordenador que tiene instalado en él el programa informático para ejecutar uno de los procedimientos descritos en la presente memoria.

**[0048]** Una realización adicional según la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, por medios ópticos o electrónicos) un programa informático para ejecutar uno de los procedimientos descritos en la presente memoria a un receptor. El receptor puede ser, por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o sistema puede comprender, por ejemplo, un servidor de archivo para transferir el programa informático al receptor.

25 **[0049]** En algunas realizaciones se puede usar un dispositivo de lógica programable (por ejemplo, una matriz de puertas programables de campo) para realizar algunas o todas las funcionalidades de los procedimientos descritos en la presente memoria. En algunas realizaciones, la matriz de puertas programables de campo puede cooperar con un microprocesador con el fin de realizar uno de los procedimientos descritos en la presente memoria. En general, los procedimientos son realizados preferentemente mediante algún aparato de hardware.

30 **[0050]** El aparato descrito en la presente memoria puede ser implementado usando un aparato de hardware, o usando un ordenador, o usando una combinación de un aparato de hardware y un ordenador.

**[0051]** Los procedimientos descritos en la presente memoria pueden ser ejecutados usando un aparato de hardware, o usando un ordenador, o usando una combinación de un aparato de hardware y un ordenador.

35 **[0052]** Las realizaciones descritas anteriormente son puramente ilustrativas para los principios de la presente invención. Debe entenderse que para los expertos en la materia serán evidentes las modificaciones y variaciones posibles de las disposiciones y de los detalles descritos en la presente memoria. Por lo tanto, se pretende que la invención esté limitada sólo por el alcance de las siguientes reivindicaciones de patente y no por los detalles específicos presentados a modo de descripción y la explicación de las realizaciones en la presente memoria.

40 Lista de bibliografía citada, que no es de patentes

**[0053]**

[1] B. Bessette y col., "The Adaptive Multi-rate Wideband Speech Codec (AMR-WB)", IEEE Trans. On Speech and Audio Processing, Vol. 10, nº 8, Nov. 2002.

45 [2] R. C. Hendriks, R. Heusdens y J. Jensen, "MMSE based noise PSD tracking with low complexity", in IEEE Int. Conf. Acoust., Speech, Signal

Processing, pág. 4266 - 4269, March 2010.

50 [3] R. Martin, "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics", IEEE Trans. On Speech and Audio Processing, Vol. 9, nº 5, Jul. 2001.

[4] M. Jelinek and R. Salami, "Wideband Speech Coding Advances in VMR-WB Standard", IEEE Trans. On Audio, Speech, and Language Processing, Vol. 15, nº 4, May 2007.

55 [5] J. Mäkinen y col., "AMR-WB+: A New Audio Coding Standard for 3rd Generation Mobile Audio Services", in Proc. ICASSP 2005, Filadelfia, EE.UU., Mar. 2005.

60 [6] M. Neuendorf y col., "MPEG Unified Speech and Audio Coding - The ISO/MPEG Standard for High- Efficiency Audio Coding of All Content Types", in Proc. 132da. Convención de AES, Budapest, Hungría, abril de 2012. También aparece en el Journal of the AES, 2013.

[7] T. Vaillancourt y col., "ITU-T EV-VBR: A Robust 8 - 32 kbit/s Scalable Coder for Error Prone Telecommunications Channels", in Proc. EUSIPCO 2008, Lausana, Suiza, agosto de 2008.

## REIVINDICACIONES

1. Un decodificador de audio para proporcionar información de audio basándose en una información de audio codificada que comprende coeficientes de predicción lineal (LPC), comprendiendo el decodificador de audio:
- 5 - un estimador de nivel de ruido configurado para estimar un nivel de ruido para una trama actual que usa una pluralidad de coeficientes de predicción lineal de al menos una trama anterior para obtener una información de nivel de ruido; y
- un dispositivo de inserción de ruido configurado para añadir un ruido a la trama actual dependiendo de la información
- 10 de nivel de ruido proporcionada por el estimador de nivel de ruido,
- en el que** el decodificador de audio está adaptado para decodificar una señal de excitación de la trama actual y calcular su media cuadrática  $e_{rms}$  ;
- 15 **en el que** el decodificador de audio está adaptado para calcular un nivel de pico  $p$  de una función de transferencia de un filtro LPC de la trama actual;
- en el que** el decodificador de audio está adaptado para calcular un mínimo espectral  $m_f$  de la trama de audio actual calculando el cociente de la media cuadrática  $e_{rms}$  y el nivel de pico  $p$  para obtener la información de nivel de ruido;
- 20 **en el que** el estimador de nivel de ruido está adaptado para estimar el nivel de ruido basándose en dos o más cocientes de tramas de audio diferentes;
- en el que** el decodificador de audio comprende un núcleo de decodificador configurado para decodificar una
- 25 información de audio de la trama actual usando coeficientes de predicción lineal de la trama actual para obtener una señal de salida de codificador de núcleo decodificada y en el que el dispositivo de inserción de ruido añade el ruido dependiendo de un coeficiente de predicción lineal usado para decodificar la información de audio de la trama actual y usado cuando se decodifica la información de audio de una o más tramas anteriores.
- 30 2. El decodificador de audio según la reivindicación 1, **en el que** el decodificador de audio comprende un determinador de tipo trama para determinar un tipo de trama de la trama actual, estando el determinador de tipo trama configurado para identificar si el tipo de trama actual es voz o audio general, de manera que la estimación del nivel de ruido pueda realizarse dependiendo del tipo de trama de la trama actual.
- 35 3. El decodificador de audio según la reivindicación 1 o 2, **en el que** el decodificador de audio está adaptado para calcular la media cuadrática  $e_{rms}$  de la trama actual de la representación en el dominio del tiempo de la trama actual con el fin de obtener la información nivel de ruido con la condición de que la trama actual sea de tipo voz.
4. El decodificador de audio según una de las reivindicaciones 1 a 3, **en el que** el decodificador de audio
- 40 está adaptado para decodificar una excitación de MDCT no conformada de la trama actual y para calcular la media cuadrática  $e_{rms}$  de la representación en el dominio espectral de la trama actual con el fin de obtener la información de nivel de ruido si la trama actual es del tipo de audio general.
5. El decodificador de audio según cualquiera de las reivindicaciones 1 a 4, **en el que** el decodificador de
- 45 audio está adaptado para poner en cola el cociente obtenido a partir de la trama de audio actual en el estimador de nivel de ruido independientemente del tipo de trama, comprendiendo el estimador de nivel de ruido un almacenamiento de nivel de ruido para dos o más cocientes obtenidos a partir de diferentes tramas de audio.
6. El decodificador de audio según cualquiera de las reivindicaciones 1 a 5, **en el que** el estimador de nivel
- 50 de ruido está adaptado para estimar el nivel de ruido basándose en el análisis estadístico de dos o más cocientes de diferentes tramas de audio.
7. El decodificador de audio según cualquiera de las reivindicaciones anteriores, **en el que** el decodificador de audio comprende un filtro de desacentuación para desacentuar la trama actual, estando el decodificador de audio
- 55 adaptado para aplicar el filtro de desacentuación de la trama actual después de que el dispositivo de inserción de ruido añade el ruido a la trama actual.
8. El decodificador de audio según cualquiera de las reivindicaciones anteriores, **en el que** el decodificador de audio comprende un generador de ruido, estando el generador de ruido adaptado para generar el ruido que es
- 60 añadido a la trama actual por el dispositivo de inserción de ruido.
9. El decodificador de audio según cualquiera de las reivindicaciones anteriores, **en el que el decodificador de audio comprende** un generador de ruido configurado para generar aleatoriamente ruido blanco.
- 65 10. El decodificador de audio según cualquiera de las reivindicaciones anteriores, **en el que** el decodificador

de audio está configurado para usar un decodificador basándose en uno o más decodificadores AMR-WB, G.718 o LD-USAC (EVS) con el fin de decodificar la información de audio codificada.

11. El decodificador de audio según una de las reivindicaciones 1 a 10, en el que el decodificador de audio está configurado para calcular el nivel pico  $p$  según  $p = \sum |a_k|$ , en el que  $a_k$  son coeficientes de predicción lineal.

12. Un procedimiento para proporcionar una información de audio decodificada basándose en una información de audio codificada que comprende coeficientes de predicción lineal (LPC),

10 comprendiendo el procedimiento:

- la estimación del nivel de ruido para una trama actual que usa una pluralidad de coeficientes de predicción lineal de al menos una trama anterior para obtener una información de nivel de ruido; y

15 - la adición de un ruido a la trama actual dependiendo de la información de nivel de ruido proporcionada por el estimador de nivel de ruido,

**en el que** se decodifica una señal de excitación de la trama actual y en el que se calcula la media cuadrática  $e_{rms}$ ;

20 **en el que** se calcula un nivel de pico  $p$  de una función de transferencia de un filtro LPC de la trama actual;

**en el que** se calcula un mínimo espectral  $m_f$  de la trama de audio actual calculando el cociente de la media cuadrática  $e_{rms}$  y el nivel de pico  $p$  para obtener la información del nivel de ruido;

25 **en el que** el nivel de ruido se estima basándose en dos o más cocientes de tramas de audio diferentes;

**en el que** el procedimiento comprende la decodificación de una información de audio de la trama actual usando coeficientes de predicción lineal de la trama actual para obtener una señal de salida de codificador de núcleo decodificada y

30

en el que el procedimiento comprende la adición del ruido dependiendo de los coeficientes de predicción lineal usados en la decodificación de la información de audio de la trama actual y usados en la decodificación de la información de audio de una o más tramas anteriores.

35 13. El procedimiento según la reivindicación 12, en el que el nivel de pico  $p$  se calcula según  $p = \sum |a_k|$ , en el que  $a_k$  son coeficientes de predicción lineal.

14. Un programa informático para realizar un procedimiento según la reivindicación 12 o 13, **en el que** el programa informático se ejecuta en un ordenador.

40

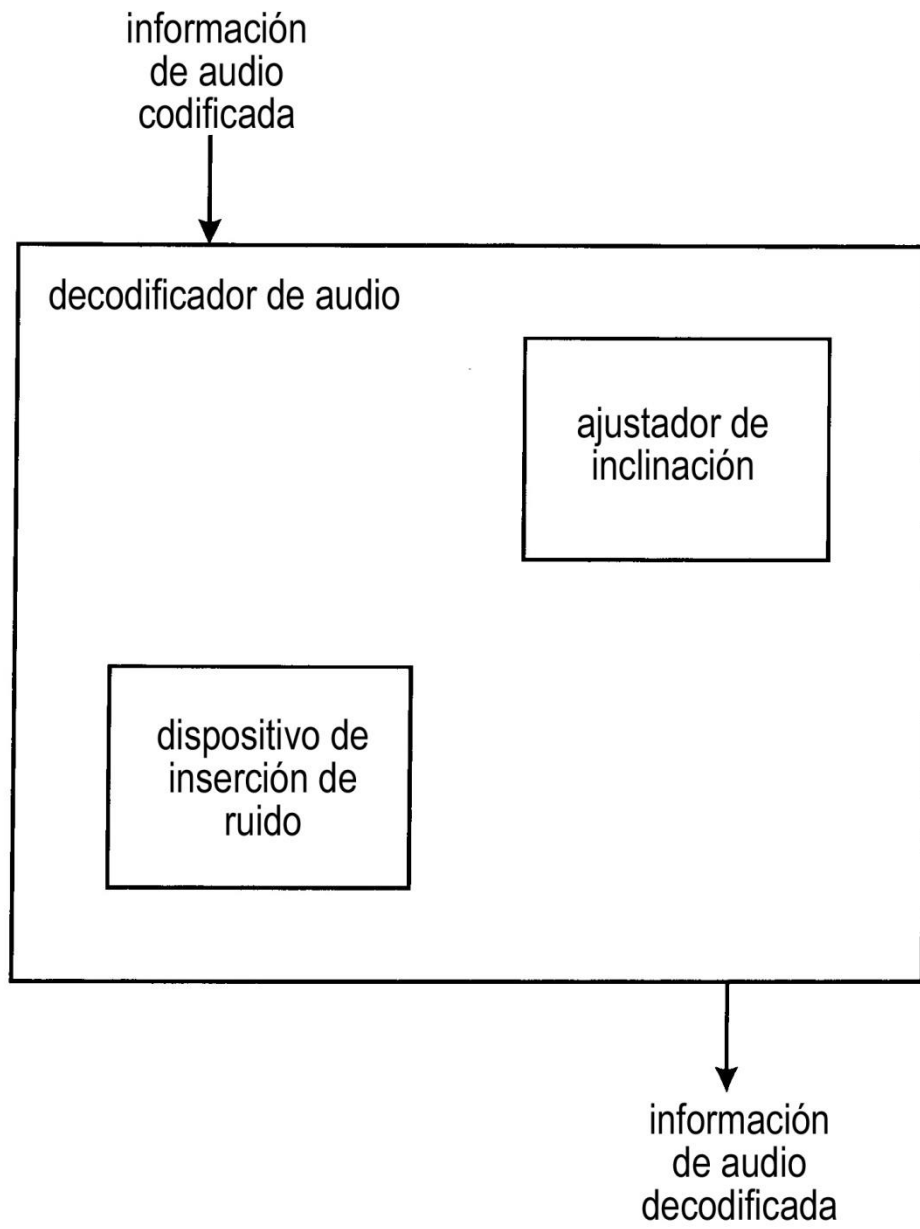


FIGURA 1

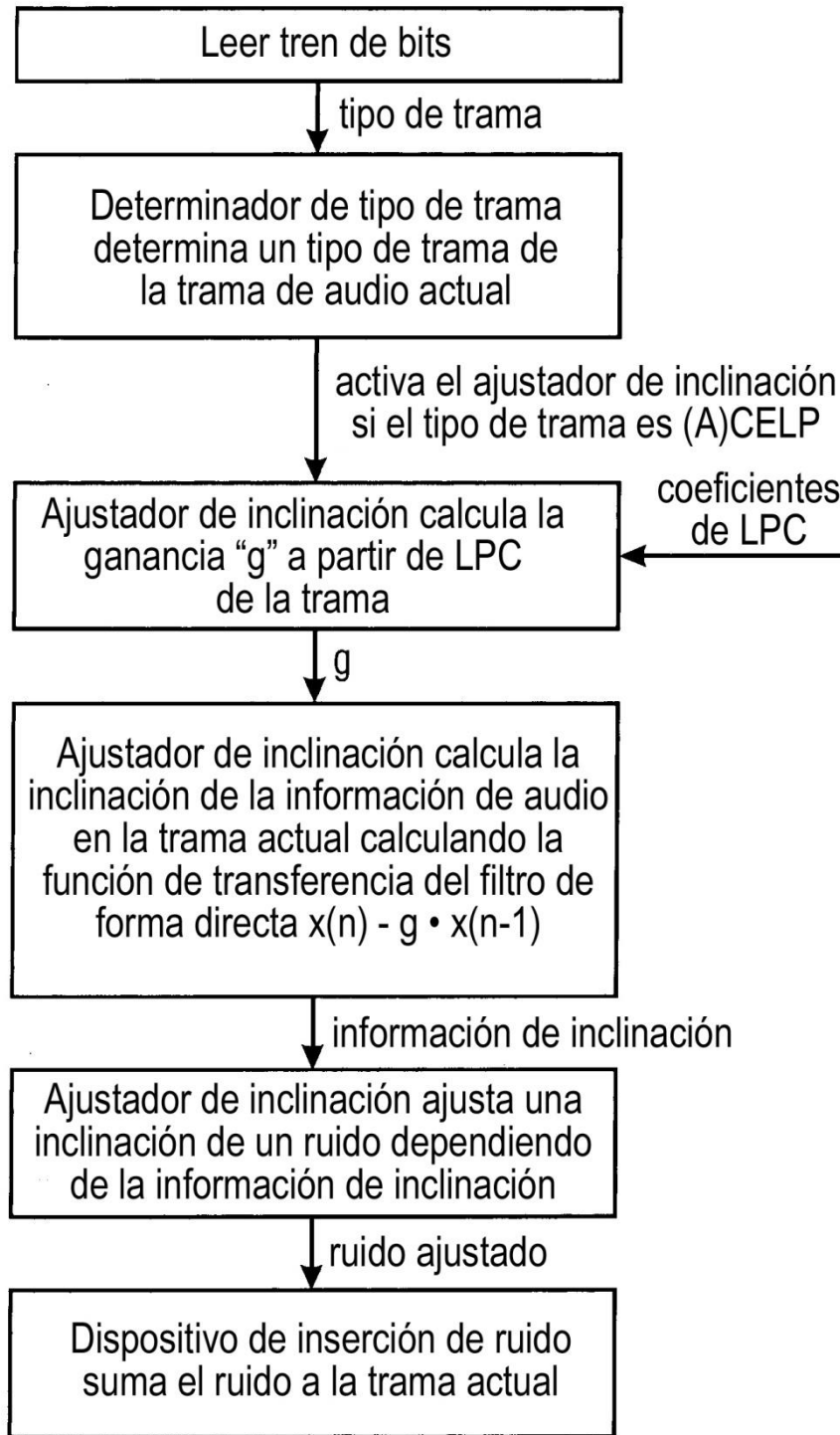


FIGURA 2

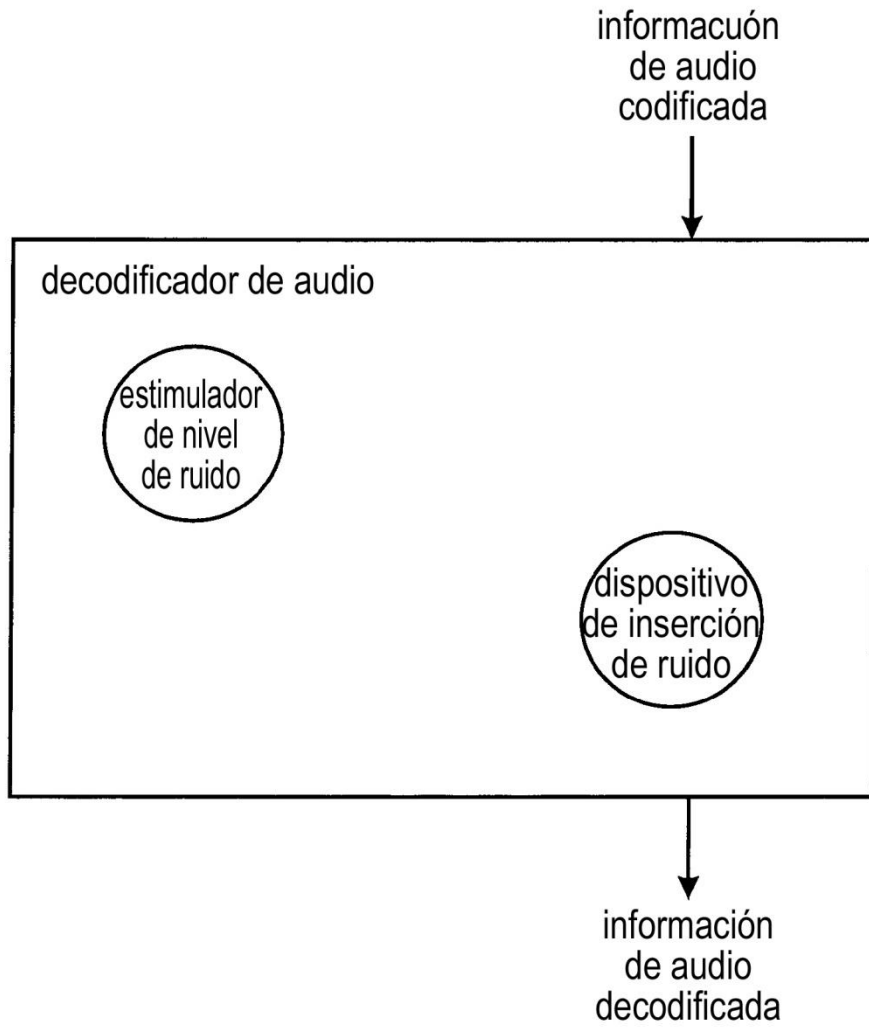


FIGURA 3

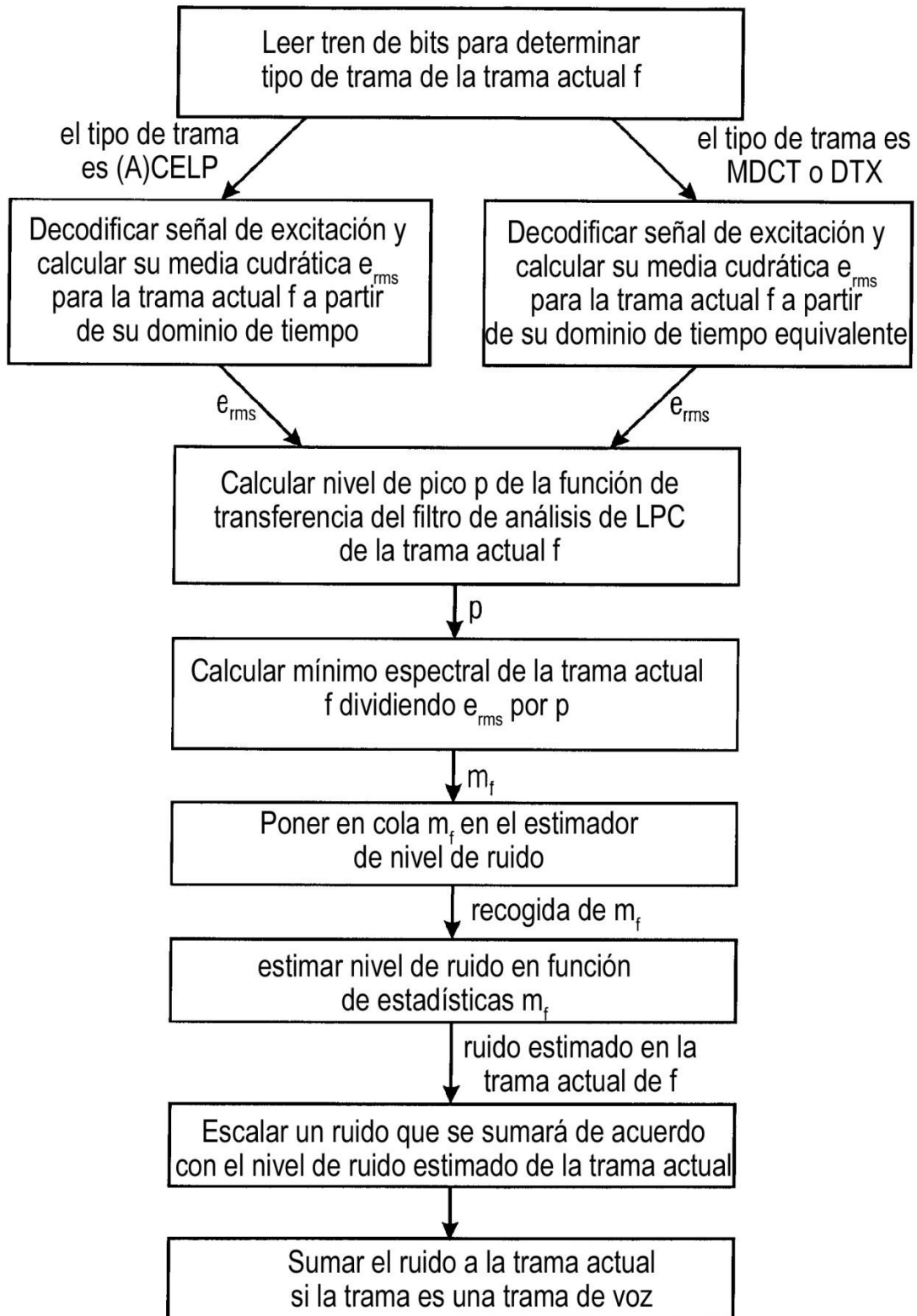


FIGURA 4



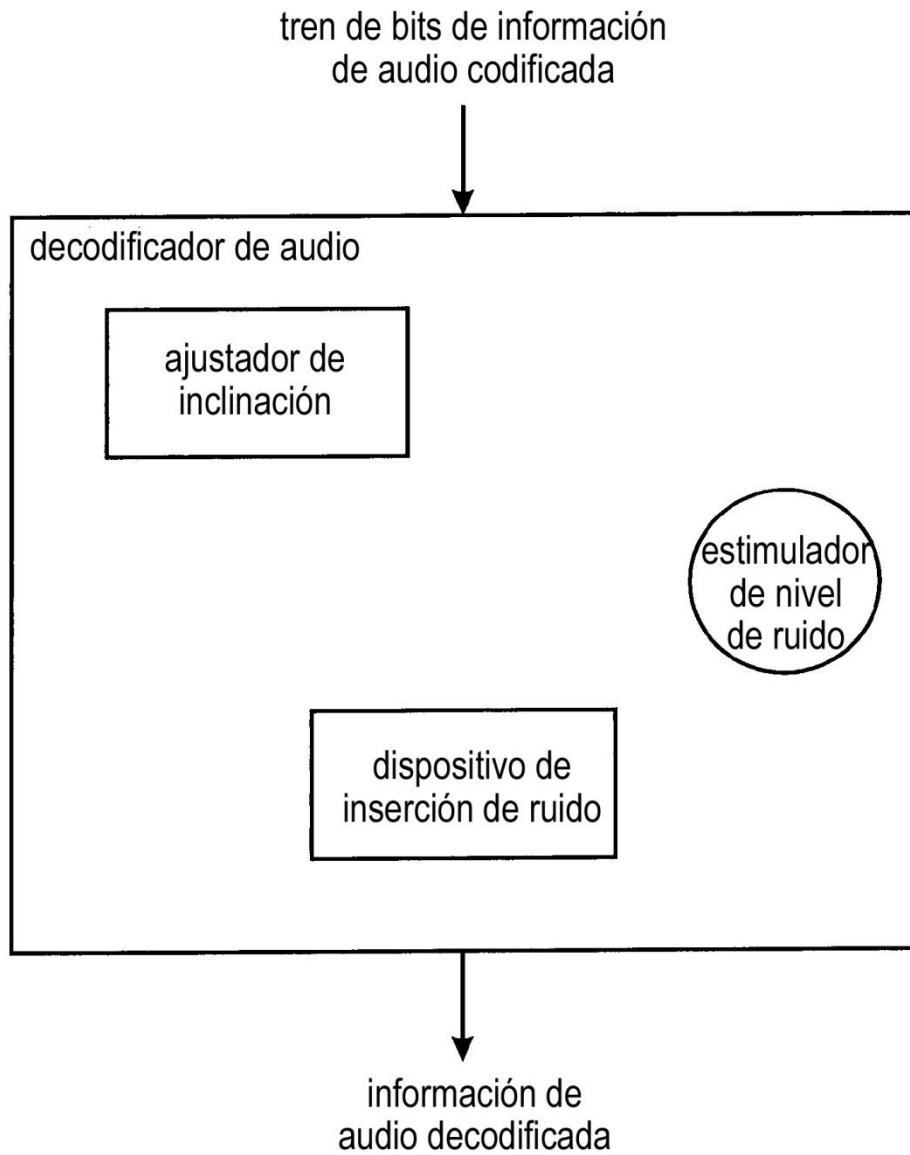


FIGURA 5



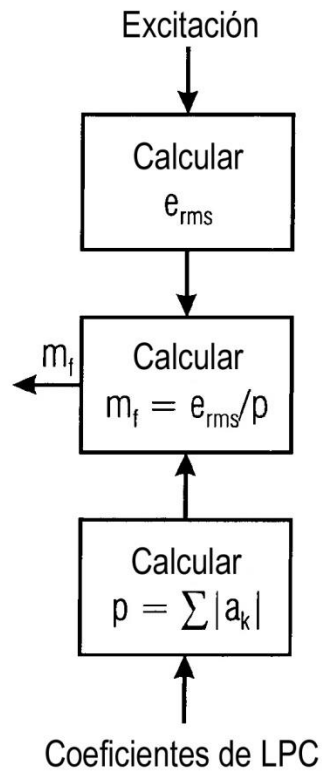


FIGURA 7A

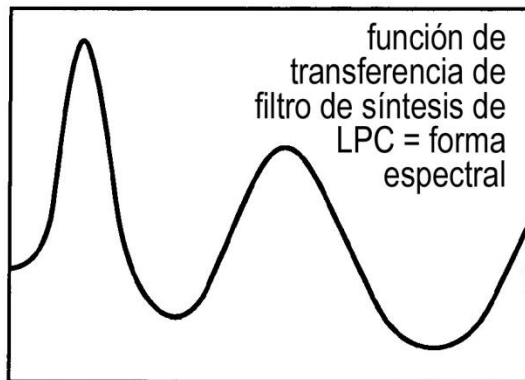


FIGURA 7B

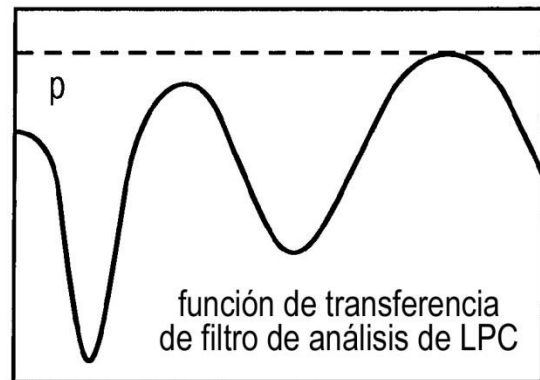


FIGURA 7C

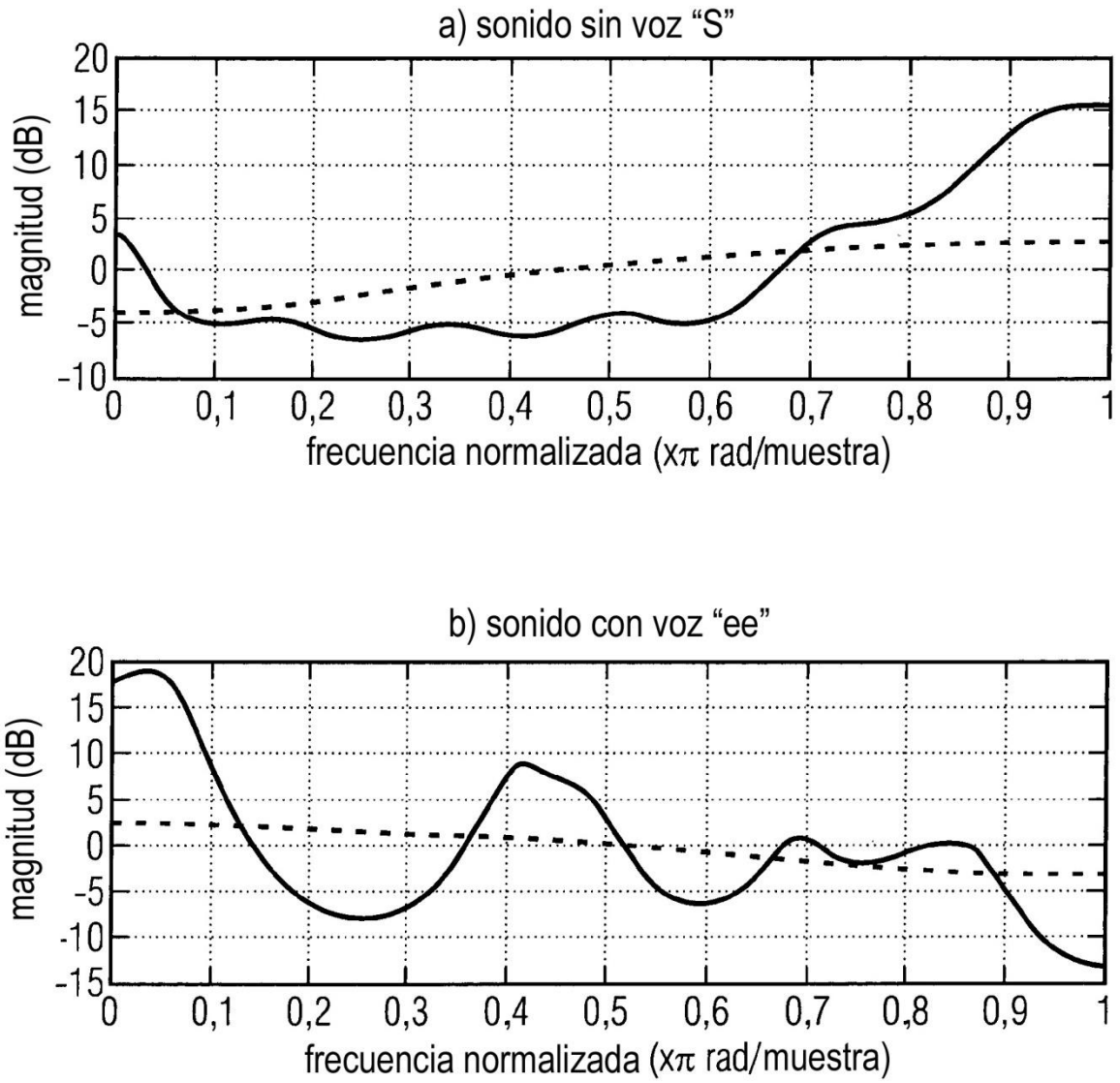
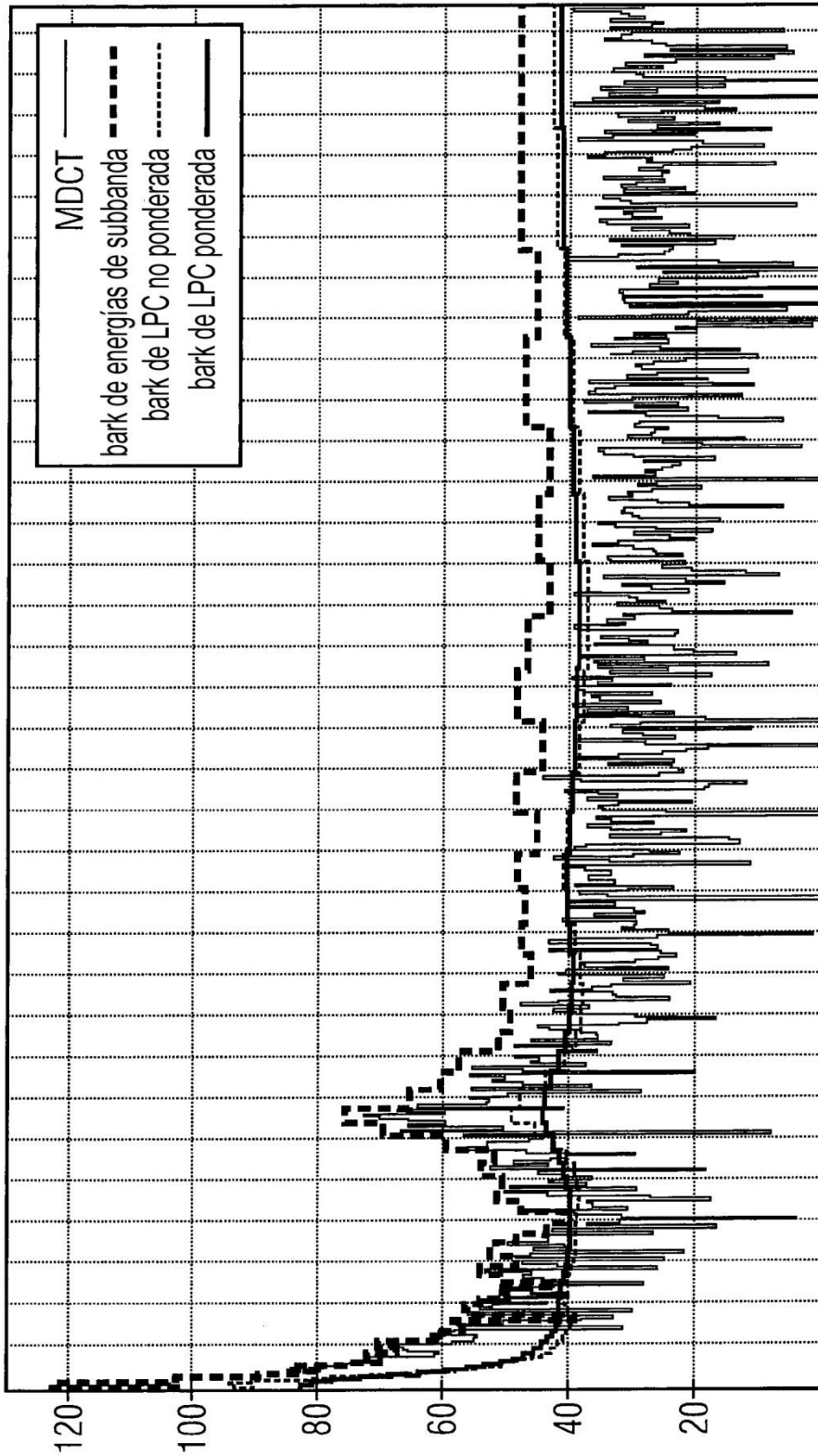


FIGURA 8



determinación de filtro de LPC a partir de espectro de potencia MDCT (valores al cuadrado de MDCT)

FIGURA 9