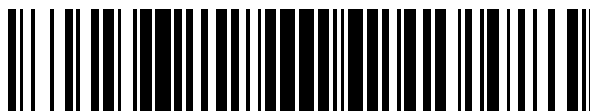


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 799 773**

51 Int. Cl.:

**G10L 19/028** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.01.2014** E 16176505 (2)

97 Fecha y número de publicación de la concesión europea: **18.03.2020** EP 3121813

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:

**21.12.2020**

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 2 799 773 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 en función de una información de audio codificada que comprende coeficientes de predicción lineal (LPC), a un procedimiento para proporcionar una información de audio decodificada en función de una información de audio codificada que comprende coeficientes de predicción lineal (LPC), a un programa informático para ejecutar tal 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 tal señal de audio, habiendo sido tratada la señal de audio con tal procedimiento.

15 Antecedentes de la invención

**[0002]** Los codificadores de voz digitales de baja velocidad de bits basados en el principio de codificación de predicción lineal excitada por código (CELP) en general padecen artefactos por dispersión de señal cuando la velocidad de bits 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 la 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 en función de 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 derivar de allí, los coeficientes de predicción lineal. Un conformador en el dominio de la frecuencia de un codificador de audio está configurado para conformar espectralmente un espectro actual de la secuencia de espectros del espectrograma en función de los coeficientes de predicción lineal proporcionados por el analizador de predicción lineal. Se inserta un espectro cuantificado y espectralmente conformado en un flujo 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 estar presente un módulo de conformación de ruido temporal para realizar una conformación de ruido temporal.

**[0004]** El documento US 6.691.085 B1 describe un procedimiento y un sistema para estimar la señal artificial de banda alta en códec de voz mediante el uso de información de actividad de voz. Dicho documento describe un procedimiento y un sistema para codificar y decodificar una señal de entrada, en el que la señal de entrada se divide en una banda de frecuencia más alta y una banda de frecuencia más baja en los procedimientos de codificación y decodificación. La decodificación de la banda de frecuencia más alta se lleva a cabo mediante el uso de una señal artificial junto con parámetros relacionados con la voz obtenidos de la banda de frecuencia más baja. En particular, la señal artificial se escala antes de transformarse en una señal de banda ancha artificial que contiene ruido de color tanto en la banda de frecuencia más baja como en la más alta. Además, la información de actividad de voz se usa para definir períodos de voz y períodos sin voz de la señal de entrada. En función de la información de actividad de la voz, se utilizan diferentes factores de ponderación para escalar la señal artificial en períodos de voz y períodos sin voz.

**[0005]** El documento US 2012/046955 describe un sistema para codificar vectores de señal para almacenamiento o transmisión, que comprende un algoritmo de inyección de ruido para ajustar adecuadamente la ganancia, la forma espectral y/u otras características del ruido inyectado para maximizar la calidad perceptiva y minimizar al mismo tiempo la cantidad de información que se va a transmitir.

**[0006]** En vista de la técnica anterior, sigue existiendo una demanda de un decodificador de audio mejorado, un procedimiento mejorado, un programa informático mejorado para ejecutar tal procedimiento y una señal de audio o un medio de almacenamiento que tenga tal señal de audio almacenada, habiendo sido tratada la señal de audio con tal procedimiento. Más específicamente, es deseable encontrar soluciones que mejoren la calidad de sonido de la información de audio transferida en la corriente de bits codificada.

Resumen de la invención

65

**[0007]** Los signos de referencia en las reivindicaciones 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 es como se define por las reivindicaciones adjuntas.

5

Breve descripción de los dibujos

**[0009]** A continuación, se describen realizaciones de la presente invención con respecto a las figuras.

- 10 La figura 1 muestra una primera realización de un decodificador de audio según la presente invención;  
 La figura 2 muestra un primer procedimiento para realizar decodificación de audio según la presente invención el cual puede ser ejecutado por un decodificador de audio según la figura 1;  
 La figura 3 muestra una segunda realización de un decodificador de audio según la presente invención;  
 15 La figura 4 muestra un segundo procedimiento para realizar decodificación de audio según la presente invención, el cual puede ser ejecutado por un decodificador de audio según la figura 3;  
 La figura 5 muestra una tercera realización de un decodificador de audio según la presente invención;  
 La figura 6 muestra un tercer procedimiento para realizar decodificación de audio según la presente invención, el cual puede ser ejecutado por un decodificador de audio según la figura 5;  
 20 La figura 7 muestra una ilustración de un procedimiento para calcular los mínimos espectrales  $m_f$  para las estimaciones de nivel ruido;  
 La figura 8 muestra un diagrama que ilustra una inclinación derivada a partir de los coeficientes LPC; y  
 La figura 9 muestra un diagrama que ilustra cómo se determinan equivalentes de filtro LPC a partir de un espectro de potencia MDCT.

25 Descripción detallada de las realizaciones de la invención

**[0010]** A continuación, se describe la invención en detalle con respecto a las figuras 1 a 9. De ninguna manera la invención está destinada a ser limitada a las realizaciones mostradas y descritas.

30 **[0011]** Todas las apariciones siguientes de la palabra "realización(es)", si se refieren a combinaciones de características diferentes de las definidas por las reivindicaciones independientes, se refieren a ejemplos que se depositaron originalmente pero que no representan a las realizaciones de la invención reivindicada actualmente; estos ejemplos se muestran aún solo con fines ilustrativos.

35 **[0012]** La figura 1 muestra una primera realización de un decodificador de audio según la presente invención. El decodificador de audio está adaptado para proporcionar una información de audio decodificada en función de una información de audio codificada. El decodificador de audio está configurado para usar 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 individualmente designados  
 40 como coeficientes  $a_k$ . El decodificador de audio comprende un ajustador de inclinación configurado para ajustar una inclinación del ruido mediante el uso de 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 bajo la condición de que la velocidad de transmisión de  
 45 bits 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 bajo 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 respecto a  
 50 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 ser mejorada sin depender de información secundaria en la corriente de bits. Así, se puede reducir la cantidad de datos que van a ser transferidos con la corriente de bits.

**[0013]** La figura 2 muestra un primer procedimiento para realizar decodificación de audio según la presente invención el cual puede ser ejecutado por un decodificador de audio según la figura 1. Se describen detalles técnicos  
 55 del decodificador de audio representado en la figura 1 junto con las características del procedimiento. El decodificador de audio está adaptado para leer la corriente 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 para 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  
 60 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 de 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 para obtener la información de inclinación. Más específicamente, el ajustador de inclinación calcula una ganancia  $g$  mediante el uso de la fórmula  $g = \frac{\sum[a_k \cdot a_{k+1}]}{\sum[a_k \cdot a_k]}$  como un análisis de primer orden, en el que  $a_k$  son coeficientes  
 65 LPC de la trama actual. La figura 8 muestra diagramas que ilustran una inclinación derivada a partir de los coeficientes

LPC. la figura 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 figura 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 la corriente 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 van a transferir con la corriente de bits. Asimismo, el ajustador de inclinación está configurado para obtener la información de inclinación mediante el uso de 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  $x(n) - g \cdot x(n-1)$  mediante el uso de la ganancia  $g$  calculada anteriormente. Después de obtener la información de inclinación, el ajustador de inclinación ajusta la inclinación del ruido que se va a añadir 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 figura 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 pronunciado, de baja complejidad del ruido añadido, el decodificador de audio proporciona la información de audio decodificada. Así, el procedimiento según la figura 2 permite mejorar la calidad de sonido de una información de audio ajustando la inclinación de un ruido que va a ser añadido a una trama actual para mejorar la calidad de un ruido de fondo.

**[0014]** La figura 3 muestra una segunda realización de un decodificador de audio según la presente invención. El decodificador de audio está adaptado de nuevo para proporcionar una información de audio decodificada en función de una información de audio codificada. El decodificador de audio está configurado de nuevo para usar un codificador que funcione en AMR-WB, G.718 y LD-USAC (EVS) para decodificar la información de audio codificada. La información de audio codificada comprende de nuevo coeficientes de predicción lineal (LPC), los cuales pueden ser individualmente designados 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 mediante el uso de 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 para añadir el ruido a la trama actual con la condición de que la velocidad de transmisión de bits de la información de audio codificada sea menor de 0,5 bits por muestra. Asimismo, el dispositivo de inserción de ruido está configurado para añadir el ruido a la trama actual con la condición de que la trama actual sea una trama de voz. Así, de nuevo, 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 ser deteriorada debido a artefactos de codificación, en especial, con respecto 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 anterior, la calidad de sonido global puede ser mejorada sin depender de información secundaria en la corriente de bits. Así, se puede reducir la cantidad de datos que van a ser transferidos con la corriente de bits.

**[0015]** La figura 4 muestra un segundo procedimiento para realizar decodificación de audio según la presente invención, el cual puede ser ejecutado por un decodificador de audio según la figura 3. Se describen detalles técnicos del decodificador de audio representado en la figura 3 junto con las características del procedimiento. Según la figura 4, el decodificador de audio está configurado para leer la corriente 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 de 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 respecto 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, el cual 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 del 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 un tipo de 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 del 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 figura 9 muestra un diagrama que ilustra cómo se determina un equivalente de filtro LPC a partir de un espectro de potencia MDCT. Mientras la escala representada es una escala 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.

**[0016]** Además, el decodificador de audio según la figura 3, como se ilustra mediante el diagrama de procedimiento de la figura 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 un tipo de voz.

**[0017]** 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 figura 9 y se describe en el documento WO 2012/110476 A1 y anteriormente. Como se ve en la figura 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 respecto a la puesta en 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 para 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, el decodificador de audio que está adaptado para poner en cola el cociente obtenido a partir de la trama de audio actual en el estimador de nivel de ruido sin importar el 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  $m_f$ , 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 en función de análisis estadístico de dos o más cocientes de diferentes tramas de audio, y así una colección de mínimos espectrales  $m_f$ . Las etapas para calcular el cociente  $m_f$  se representan en detalle en la figura 7, que ilustra las etapas de cálculo necesarias. En la segunda realización, el estimador de nivel de ruido funciona basado en estadística de mínimo como se conoce de [3]. Se pone a escala el ruido según el nivel de ruido estimado de la trama actual en función de la estadística de mínimo y, después de esto, se añade a la trama actual si la trama actual es una trama de voz. Finalmente, la trama actual se somete a desacentuación (no mostrada en la figura 4). Así, esta segunda realización permite también omitir información secundaria para el llenado de ruido, permitiendo reducir la cantidad de datos que van a ser transferidos con la corriente de bits. Por consiguiente, la calidad de sonido de la información de audio puede ser mejorada mejorando el ruido de fondo durante la etapa de decodificación sin aumentar la tasa de transmisión de datos. Cabe observar que como no se necesitan transformaciones tiempo/frecuencia y como el estimador de nivel de ruido solo se ejecuta una vez por trama (no en múltiples subbandas), el llenado de ruido descrito exhibe muy baja complejidad mientras que puede mejorar la codificación de baja velocidad de transmisión de bits de voz con ruido.

**[0018]** La figura 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 en función de una información de audio codificada. El decodificador de audio está configurado para usar un codificador basado en LD-USAC con el fin de decodificar la información de audio codificada. La información de audio codificada comprende coeficientes de predicción lineal (LPC), los cuales pueden ser individualmente designados como coeficientes  $a_k$ . El decodificador de audio comprende un ajustador de inclinación configurado para ajustar una inclinación de un ruido mediante el uso de 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 mediante el uso de 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, en especial, con respecto 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, el cual es puesto a escala, a continuación, según la información de nivel de ruido y se forma mediante el uso de la inclinación derivada de  $g$ , como se ha descrito anteriormente.

**[0019]** La figura 6 muestra un tercer procedimiento para realizar decodificación de audio según la presente invención el cual puede ser ejecutado por un decodificador de audio según la figura 5. La corriente de bits es leída y un determinador de tipo de trama, denominado detector de tipo de trama, determina si la trama actual es una trama de voz (ACELP) o una trama de audio general (TCX/MDCT). Sin importar el 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 figura 7 y se describió anteriormente, mediante el uso de coeficientes LPC también usados para decodificar la corriente de bits en lugar de usar cualquier información secundaria o coeficientes LPC adicionales. La información de ruido de ambos tipos de tramas se pone en cola para ajustar la inclinación y el nivel de ruido del ruido que se va a añadir 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 son combinadas 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 figura 6. En otras palabras, según la figura 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 la codificación ACELP (voz) y MDCT (música/ruido) sobre un base de trama. El procedimiento de inserción según la figura 6 se resume del siguiente modo:

1. La corriente de bits es leída, y se determina si la trama actual es una trama ACELP o MDCT o DTX. Independientemente del tipo de trama, la señal de excitación aplanada espectralmente (en dominio perceptual) se decodifica y se usa para actualizar la estimación de nivel de ruido como se describe más abajo en detalle. A continuación, la señal es totalmente reconstruida 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 deriva la inclinación a partir de la ganancia  $g$  de los 16 coeficientes LPC  $a_k$ , que vienen dados por  $g = \frac{\sum [a_k \cdot a_{k+1}]}{\sum [a_k \cdot a_k]}$ .
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 sobre la trama decodificada: un generador de ruido aleatorio genera la señal de ruido espectralmente blanco, la cual se pone a escala a continuación y se conforma mediante el uso de la inclinación derivada de  $g$ .
4. Se añade la señal de ruido conformada y nivelada para la trama ACELP sobre la señal decodificada justo antes de finalizar la etapa de filtrado de desacentuación final. Dado que la desacentuación es un IIR de primer orden que estimula bajas frecuencias, permite el filtrado de paso alto de IIR pronunciado del ruido añadido, como en la figura 6, lo que evita artefactos de ruido audibles a bajas frecuencias.

**[0020]** La estimación de nivel de ruido en la etapa 1 se realiza mediante el cálculo de 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 figura 7. Finalmente,  $m_f$  se pone en cola en el estimador de nivel de ruido operando basándose en, por ejemplo, la estadística de mínimo [3]. Cabe observar que como no se necesitan transformaciones tiempo/frecuencia y como el estimador de nivel solo se ejecuta una vez por trama (no en múltiples subbandas), el sistema de llenado de ruido CELP descrito exhibe muy baja complejidad mientras que puede mejorar la codificación de baja velocidad de transmisión de bits de voz con ruido.

**[0021]** Aunque se han descrito algunos aspectos en el contexto de un decodificador de audio, es claro que estos aspectos también representan una descripción del procedimiento correspondiente, donde 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 representan también una descripción de un bloque o elemento o característica correspondiente de un decodificador de audio correspondiente. Algunas o todas las etapas de procedimiento pueden ser ejecutadas por (o mediante el uso de) un aparato de hardware como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, alguna o más de las etapas de procedimiento más importantes se pueden ejecutar mediante tal aparato.

**[0022]** La señal de audio codificada inventiva 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 tal como Internet.

**[0023]** Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden ser implementadas en hardware o en software. La implementación puede ser realizada mediante el uso de un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un CD, un Blu-Ray, una memoria de solo lectura, una PROM, una EEPROM o una memoria FLASH, teniendo señales de control legibles electrónicamente almacenadas en las mismas, las cuales cooperan (o son capaces de cooperar) con un sistema informático programable de tal forma que se ejecute el respectivo procedimiento. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

**[0024]** 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 tal forma que se ejecute uno de los procedimientos descritos en esta invención.

**[0025]** En general, realizaciones de la presente invención pueden ser implementadas como un programa

informático con un código de programa, siendo 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 ser almacenado, por ejemplo, en un soporte legible por máquina.

5 **[0026]** Otras realizaciones comprenden el programa informático para ejecutar uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por máquina.

**[0027]** En otras palabras, una realización del procedimiento inventivo es, por lo tanto, un programa informático que tiene un código de programa para ejecutar uno de los procedimientos descritos en esta invención, cuando el  
10 programa informático se ejecuta en un ordenador.

**[0028]** Una realización adicional de los procedimientos inventivos 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 esta invención. El soporte de datos, el  
15 medio de almacenamiento digital o el medio de registro son típicamente tangibles y/o no transitorios.

**[0029]** Una realización adicional del procedimiento inventivo es, por lo tanto, una corriente de datos o una secuencia de señales que representan el programa informático para ejecutar uno de los procedimientos descritos en esta invención. La corriente de datos o la secuencia de señales pueden estar configuradas, por ejemplo, para ser  
20 transferidas a través de una conexión de comunicación de datos, por ejemplo, a través de Internet.

**[0030]** Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado para o adaptado para ejecutar uno de los procedimientos descritos en esta invención.  
25

**[0031]** Una realización adicional comprende un ordenador que tiene instalado en él el programa informático para ejecutar uno de los procedimientos descritos en esta invención.

**[0032]** Una realización adicional según la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa informático para ejecutar uno de los procedimientos descritos en esta invención, 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.  
30

**[0033]** 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 esta invención. En algunas realizaciones, la matriz de puertas programables de campo puede cooperar con un microprocesador para realizar uno de los procedimientos descritos en esta invención. En general, los procedimientos de preferencia, se realizan por medio de algún aparato de hardware.  
35

**[0034]** El aparato descrito en esta invención puede ser implementado mediante el uso de un aparato de hardware, o mediante el uso de un ordenador, o por medio de una combinación de un aparato de hardware y un ordenador.  
40

**[0035]** Los procedimientos descritos en esta invención se pueden ejecutar mediante el uso de un aparato de hardware, o mediante el uso de un ordenador, o por medio de una combinación de un aparato de hardware y un ordenador.  
45

**[0036]** Las realizaciones descritas anteriormente son meramente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y variaciones posibles de las disposiciones y de los detalles descritos en esta invención serán evidentes para los expertos en la materia. Por lo tanto, se pretende que la invención esté limitada solo 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 esta invención.  
50

55 Lista de bibliografía citada, que no es de patentes

**[0037]**

[1] B. Bessette y col., "El códec de banda ancha para voz multitasa (AMR-WB)" (*The Adaptive Multi-rate Wideband Speech Codec (AMR-WB)*), IEEE Trans. Procesamiento de Voz y Audio (*Speech and Audio Processing*), vol. 10, N. ° 8, noviembre de 2002.  
60

[2] R. C. Hendriks, R. Heusdens y J. Jensen, "Rastreo PSD de ruido basado en MMSE con baja complejidad" (*MMSE based noise PSD tracking with low complexity*), en Conf. Internac. de IEEE Procesamiento de señal, acústica, voz, pp. 4266 – 4269, marzo de 2010.  
65

- 5 [3] R. Martin, "Estimación de densidad espectral de potencia de ruido basada en estadística de alisamiento óptimo y mínimo" (*"Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics,"*) Trans. de IEEE sobre Procesamiento de voz y audio, Vol. 9, N. ° 5, julio de 2001.
- [4] M. Jelinek y R. Salami, "Avances de codificación de voz de banda ancha en Estándar VMR—WB (*"Wideband Speech Coding Advances in VMR—WB Standard,"*) Trans. de IEEE sobre Procesamiento de audio, voz y lenguaje, Vol. 15, N. ° 4, mayo de 2007.
- 10 [5] J. Mäkinen y col., "AMR—WB+: un nuevo estándar de codificación de audio para servicios de audio móvil de 3ra. generación" (*"AMR—WB+: A New Audio Coding Standard for 3<sup>rd</sup> Generation Mobile Audio Services,"*) en Proc. ICASSP 2005, Filadelfia, EE.UU., marzo de 2005.
- 15 [6] M. Neuendorf y col., "Codificación de voz y audio unificada MPEG — El estándar ISO/MPEG para codificación de audio de alta eficacia para todos los tipos de contenidos" (*"MPEG Unified Speech and Audio Coding – The ISO/MPEG Standard for High—Efficiency Audio Coding of All Content Types,"*) en Proc. 132da. Convención de AES, Budapest, Hungría, abril de 2012. También aparece en el Journal of the AES, 2013.
- 20 [7] T. Vaillancourt y col., "ITU—T EV—VBR: un codificador escalable robusto 8 — 32 kbit/s para canales de telecomunicaciones propensos a error" (*"ITU—T EV—VBR: A Robust 8 – 32 kbit/s Scalable Coder for Error Prone Telecommunications Channels,"*) en Proc. EUSIPCO 2008, Lausanne, Suiza, agosto de 2008.



**REIVINDICACIONES**

1. Un decodificador de audio para proporcionar una información de audio decodificada sobre la base de una información de audio codificada que comprende coeficientes de predicción lineal (LPC), comprendiendo el decodificador de audio:

- un ajustador de inclinación configurado para ajustar una inclinación de un ruido de fondo dependiendo de una información de inclinación, en el que el ajustador de inclinación está configurado para usar coeficientes de predicción lineal de una trama actual para obtener una información de inclinación; y
- un núcleo de decodificador configurado para decodificar una información de audio de la trama actual mediante el uso de los coeficientes de predicción lineal de la trama actual para obtener una señal de salida del codificador de núcleo decodificado; y
- un dispositivo de inserción de ruido configurado para añadir el ruido de fondo ajustado a la trama actual para realizar un llenado de ruido;

**caracterizado porque**

el ajustador de inclinación está configurado para usar un resultado de un análisis de primer orden de los coeficientes de predicción lineal de la trama actual para obtener la información de inclinación, y

**en el que** el ajustador de inclinación está configurado para obtener la información de inclinación mediante el uso de un cálculo de una ganancia  $g$  de los coeficientes de predicción lineal de la trama actual como el análisis de primer orden, en el que

$$g = \frac{\sum(a_k \cdot a_{k+1})}{\sum(a_k \cdot a_k)},$$

en el que  $a_k$  es un coeficiente de predicción lineal de la trama actual, ubicado en el índice LPC  $k$ .

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 activar el ajustador de inclinación para que ajuste la inclinación del ruido de fondo cuando el tipo de trama de la trama actual se detecta para que sea del tipo de voz.

3. El decodificador de audio según cualquiera de las reivindicaciones anteriores, **en el que** el decodificador de audio comprende además:

- un estimador de nivel de ruido configurado para estimar un nivel de ruido para una trama actual que utiliza una pluralidad de coeficientes de predicción lineal de al menos una trama anterior para obtener la información de nivel de ruido;
- en el que el dispositivo de inserción de ruido se configura para añadir el ruido de fondo a la trama actual dependiendo de la información 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 la media cuadrática  $e_{rms}$ ;

**en el que** el decodificador de audio está adaptado para calcular un nivel de pico  $p$  de una función de transferencia de un filtro de 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 mediante el cálculo del cociente de la media cuadrática  $e_{rms}$  y el nivel de pico  $p$  para obtener la información de nivel de ruido;

**en el que** el estimador de nivel de ruido está adaptado para estimar el nivel de ruido basándose en dos o más cocientes de diferentes tramas de audio.

4. 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 adaptado para aplicar el filtro de desacentuación a la trama actual después de que el dispositivo de inserción de ruido añada el ruido a la trama actual.

5. 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 va a ser añadido a la trama actual por el dispositivo de inserción de ruido.

6. 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 ruido blanco aleatorio.**

7. El decodificador de audio según cualquiera de las reivindicaciones anteriores, **en el que** el decodificador

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

8. Un procedimiento para proporcionar una información de audio decodificada en función de una información de audio codificada que comprende coeficientes de predicción lineal (LPC), comprendiendo el procedimiento:

- ajustar una inclinación de un ruido de fondo dependiendo de una información de inclinación, en el que los coeficientes de predicción lineal de una trama actual se usan para obtener la información de inclinación; y
- 10 - decodificar una información de audio de la trama actual mediante el uso de coeficientes de predicción lineal de la trama actual para obtener una señal de salida de codificador de núcleo decodificado; y
- añadir el ruido de fondo ajustado a la trama actual para realizar un llenado de ruido;

**caracterizado por que**

15 un resultado de un análisis de primer orden de los coeficientes de predicción lineal de la trama actual se usa para obtener la información de inclinación, y

**en el que** la información de inclinación se obtiene mediante el uso de un cálculo de una ganancia  $g$  de los coeficientes de predicción lineal de la trama actual como el análisis de primer orden, en el que

20

$$g = \frac{\sum(a_k \cdot a_{k+1})}{\sum(a_k \cdot a_k)},$$

en el que  $a_k$  es un coeficiente de predicción lineal de la trama actual, ubicado en el índice LPC  $k$ .

25 9. Un programa informático para realizar un procedimiento según la reivindicación 8, **en el que** el programa informático se ejecuta en un ordenador.

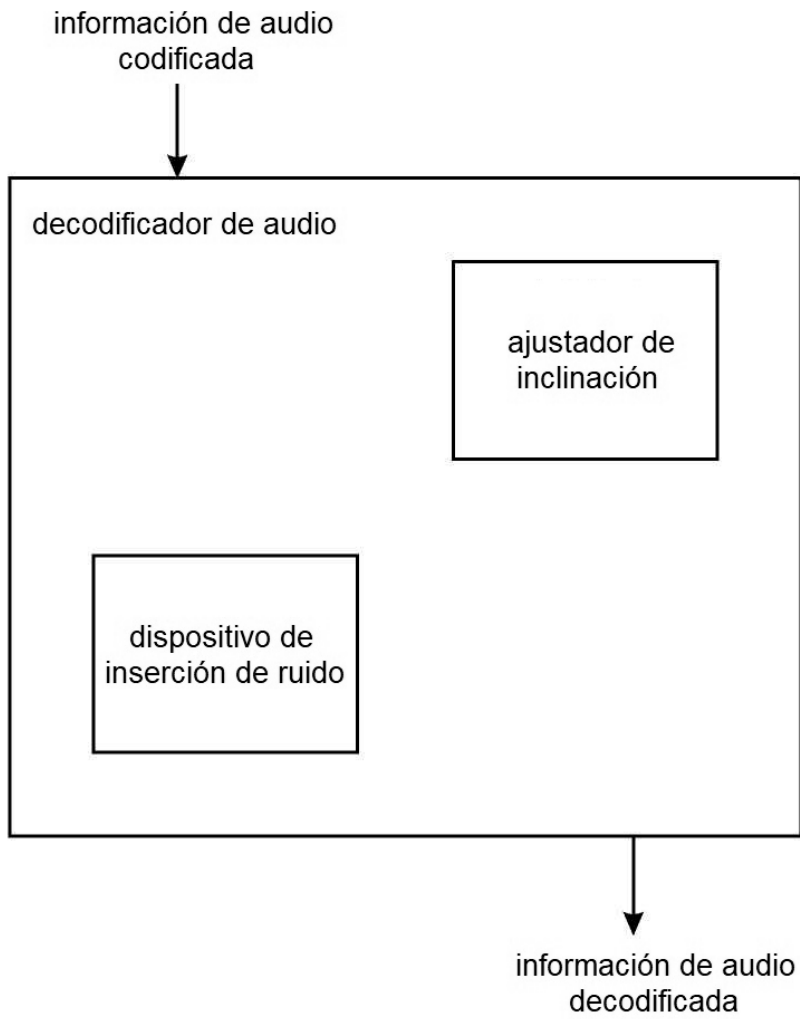


FIGURA 1

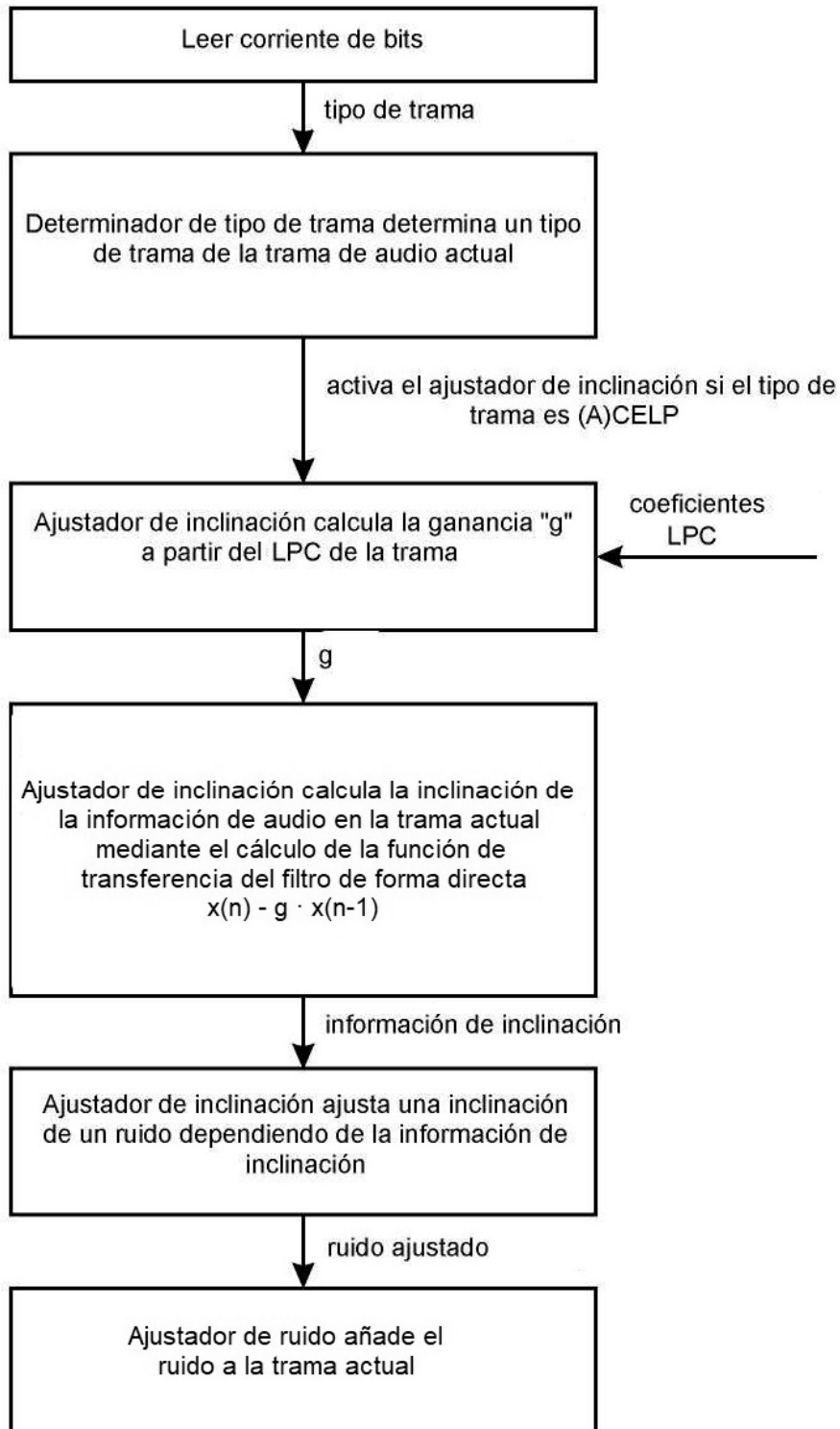


FIGURA 2

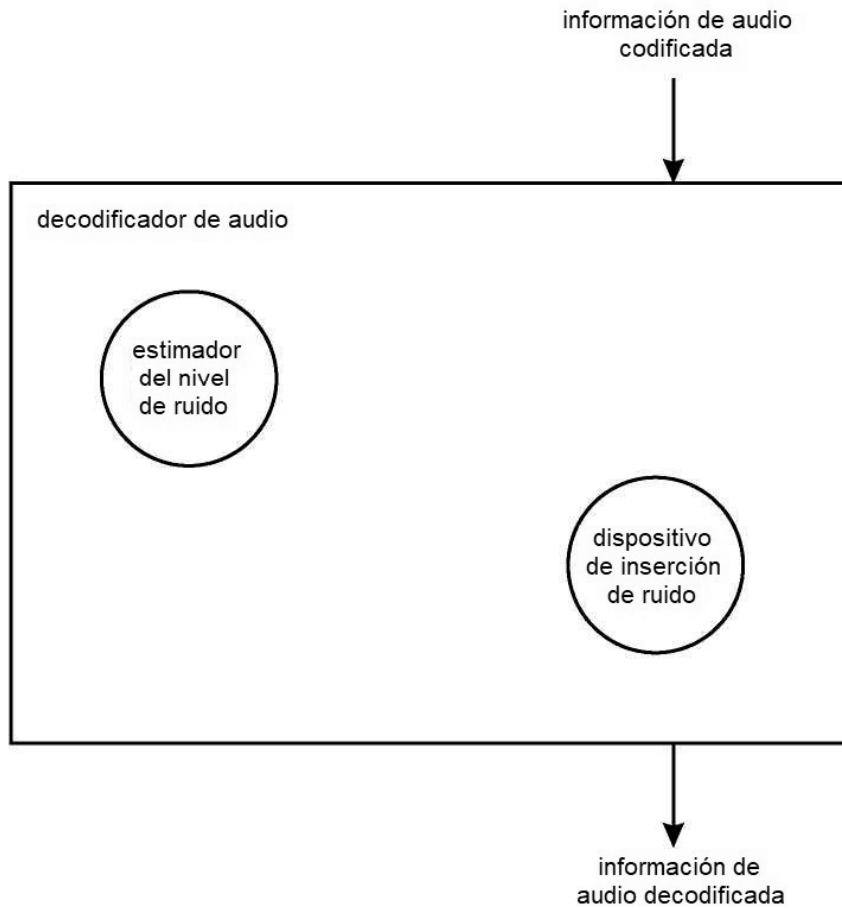


FIGURA 3

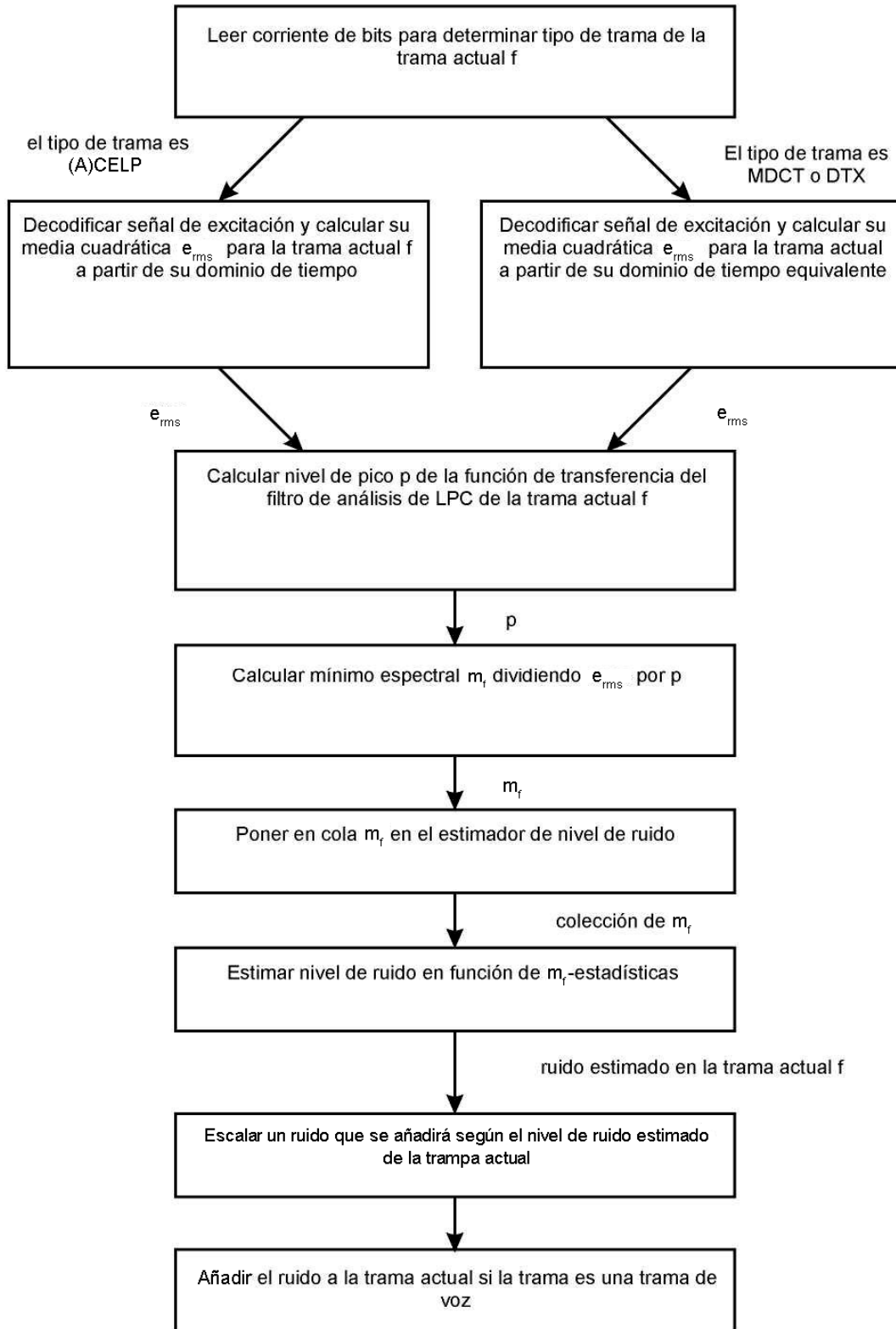


FIGURA 4

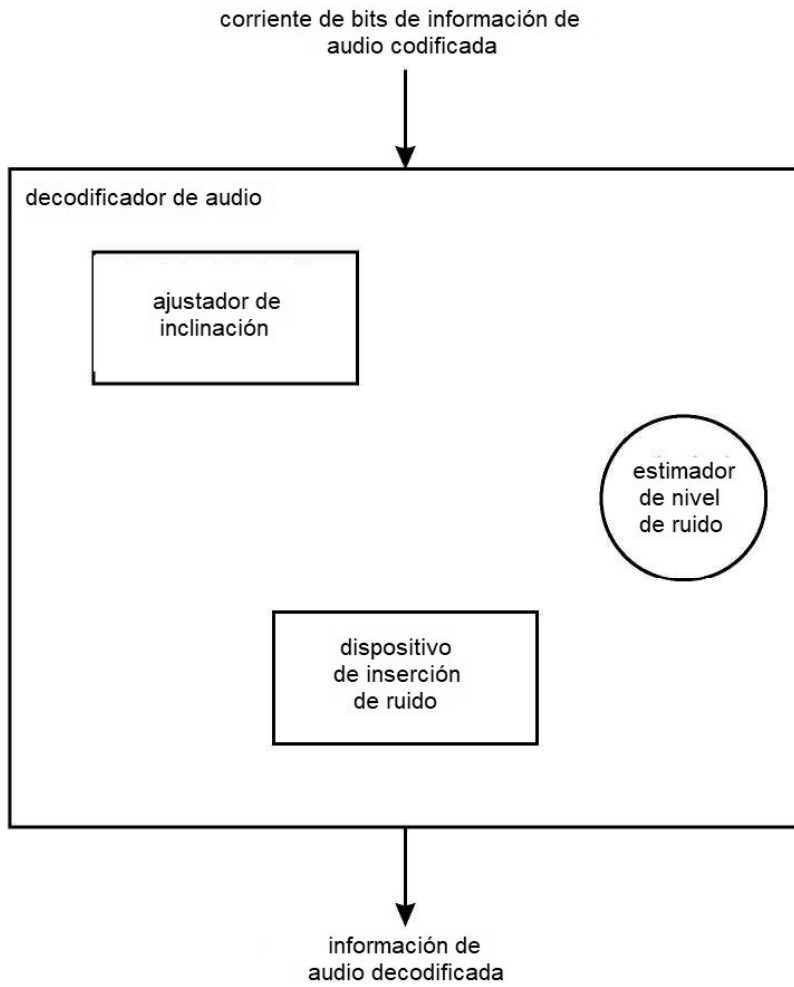


FIGURA 5

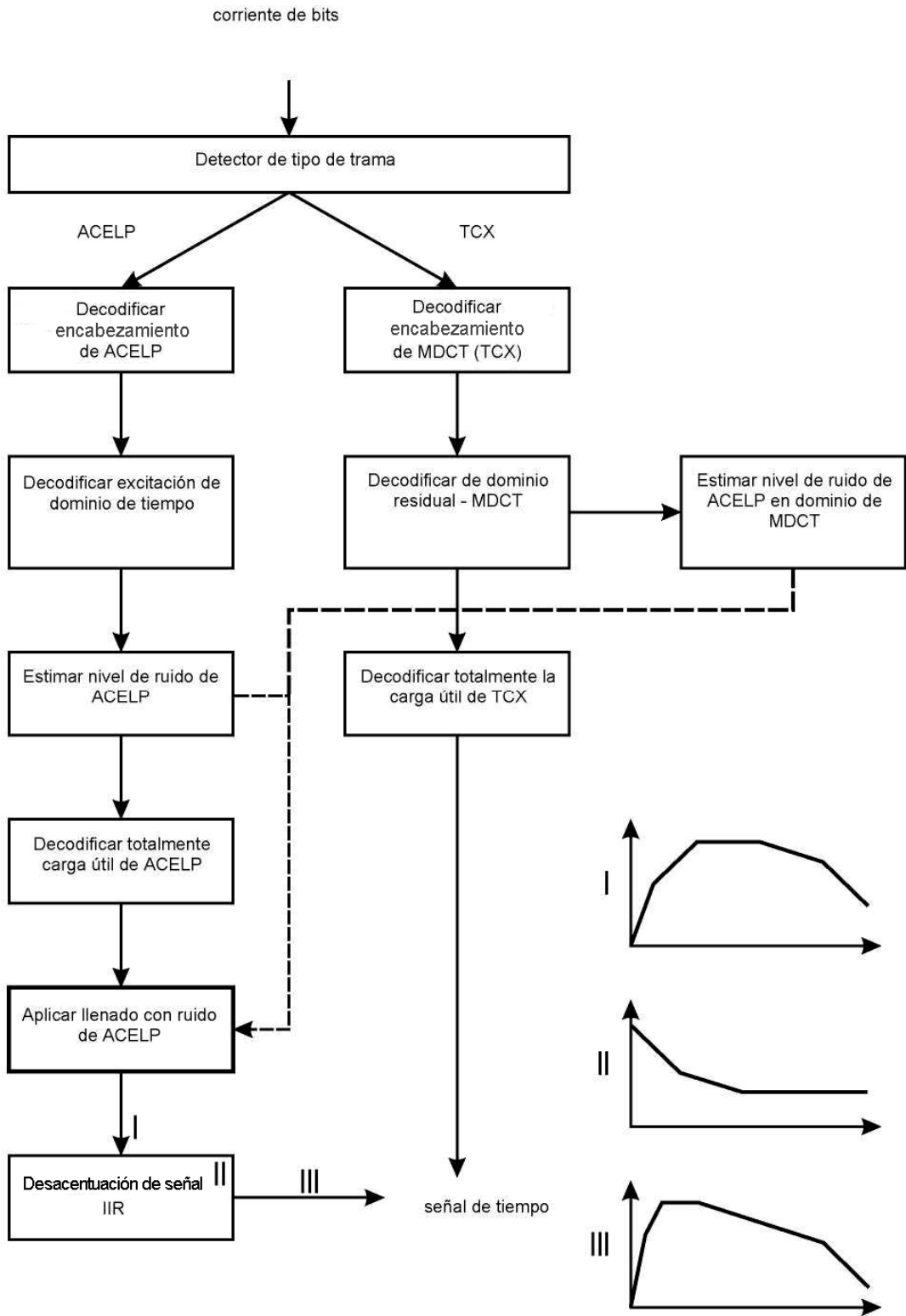


FIGURA 6



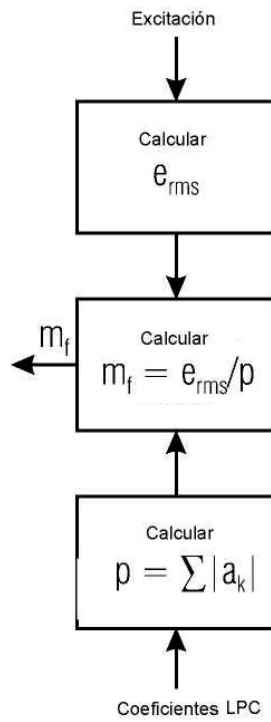


FIGURA 7A

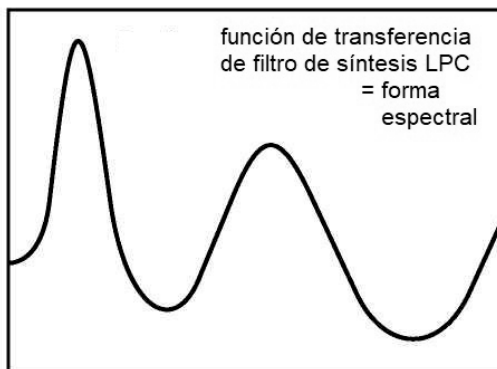


FIGURA 7B

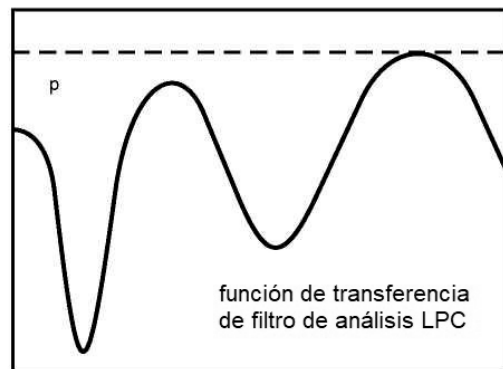


FIGURA 7C

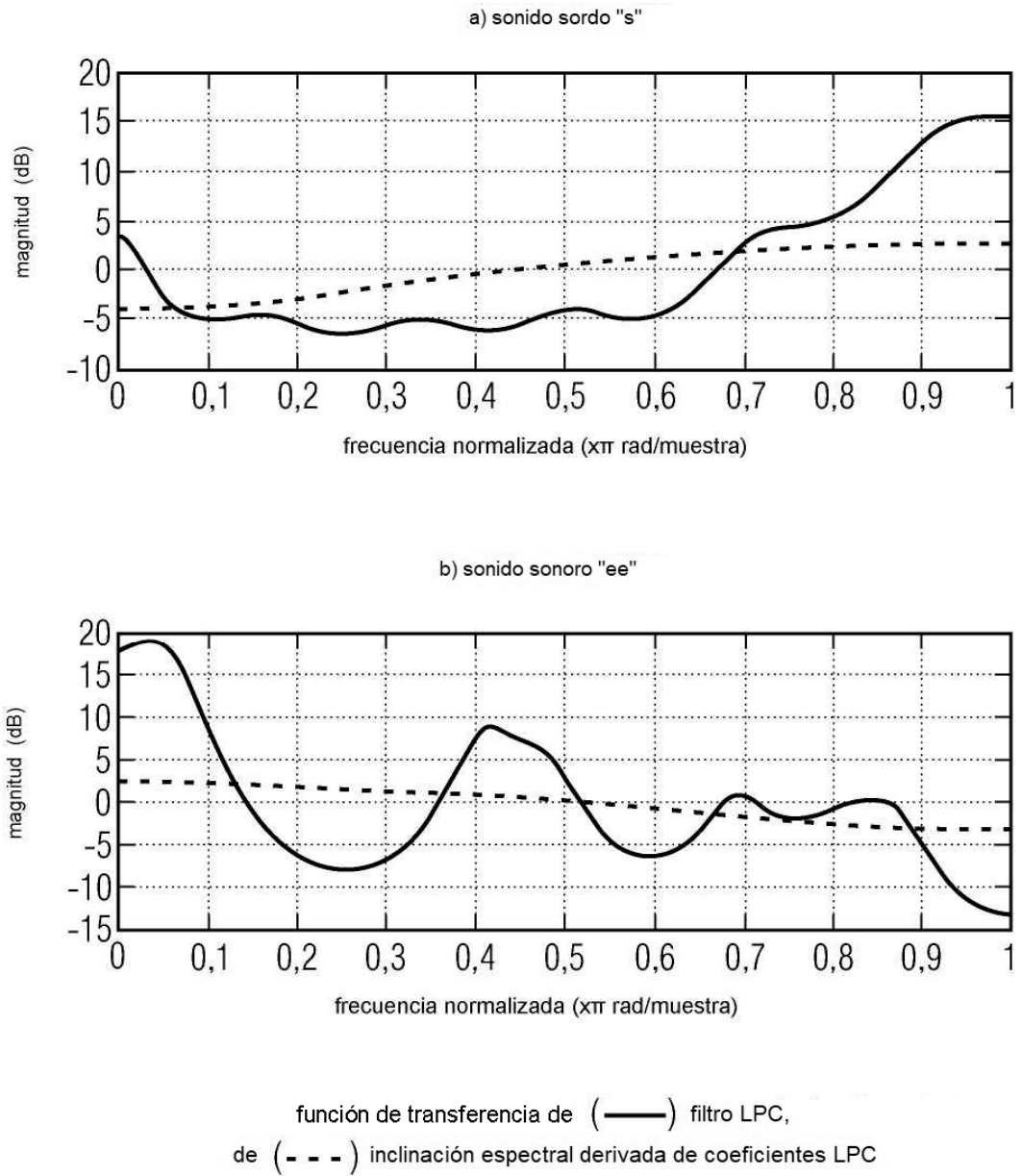
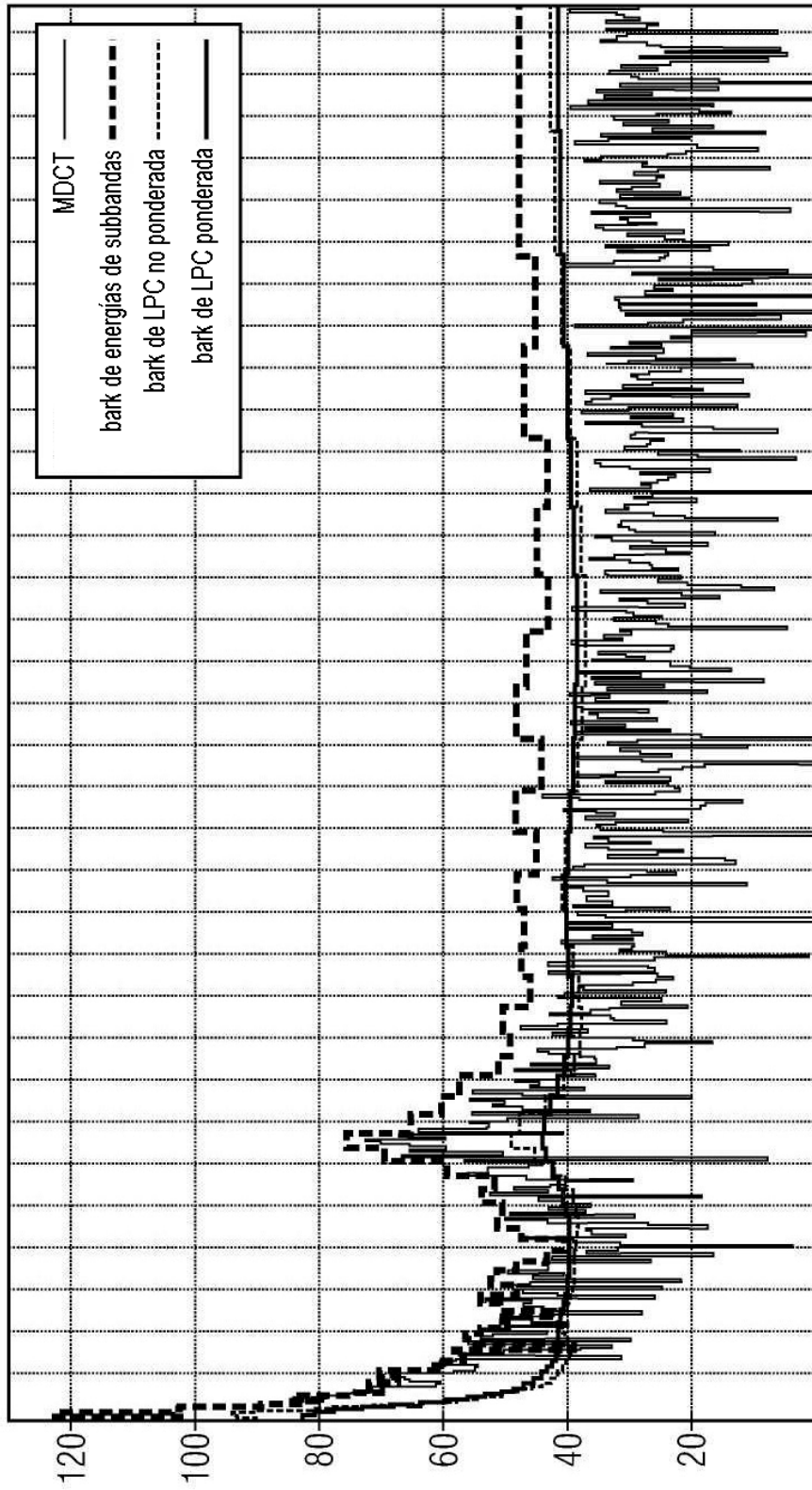


FIGURA 8



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

FIGURA 9