

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 676 584**

51 Int. Cl.:

**G10L 19/26** (2013.01)

**G10L 21/02** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **24.07.2015 PCT/EP2015/066998**

87 Fecha y número de publicación internacional: **04.02.2016 WO16016121**

96 Fecha de presentación y número de la solicitud europea: **24.07.2015 E 15742238 (7)**

97 Fecha y número de publicación de la concesión europea: **20.06.2018 EP 3175454**

54 Título: **Aparato y método para procesar una señal de audio usando un filtro posterior de armónicos**

30 Prioridad:

**28.07.2014 EP 14178820**

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

**23.07.2018**

73 Titular/es:

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

72 Inventor/es:

**RAVELLI, EMMANUEL;  
HELMRICH, CHRISTIAN;  
MARKOVIC, GORAN;  
NEUSINGER, MATTHIAS;  
DISCH, SASCHA;  
JANDER, MANUEL y  
DIETZ, MARTIN**

74 Agente/Representante:

**ARIZTI ACHA, Monica**

ES 2 676 584 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).

Aparato y método para procesar una señal de audio usando un filtro posterior de armónicos

**DESCRIPCIÓN**

5 La presente invención está relacionada con el procesamiento de audio y, particularmente, con el procesamiento de audio usando un filtro posterior de armónicos.

Los códecs de audio basados en transformada por lo general introducen ruido interarmónico cuando procesan señales de audio armónicas, particularmente a tasas de bits bajas.

10 Este efecto empeora aún más cuando el códec de audio basado en transformada opera a bajo retardo, debido a la peor resolución de frecuencia y/o selectividad introducida por un tamaño de transformada más corto y/o una peor respuesta de frecuencia ventana.

15 Este ruido interarmónico generalmente se percibe como un artefacto muy molesto, reduciendo de forma significativa el rendimiento del códec de audio basado en transformada cuando se evalúa de forma subjetiva en el material de audio altamente tonal.

20 Existen varias soluciones para mejorar la calidad subjetiva de los códecs de audio basados en transformada en señales de audio armónicas. Todas ellas están basadas en técnicas basadas en predicción, ya sea en el dominio de transformada o en el dominio de tiempo.

Los ejemplos de enfoques de dominio de transformada son:

- 25
- [1] H. Fuchs, "Improving MPEG Audio Coding by Backward Adaptive Linear Stereo Prediction", 99ª Convención AES, Nueva York 1995, Edición preliminar 4086.
  - [2] L. Yin, M. Suonio, M. Väänänen, "A New Backward Predictor for MPEG Audio Coding", 103ª Convención AES, Nueva York 1997, Edición preliminar 4521
  - [3] Juha Ojanperä, Mauri Väänänen, Lin Yin, "Long Term Predictor for Transform Domain Perceptual Audio Coding", 107ª Convención AES, Nueva York 1999, Edición preliminar 5036.
- 30

Los ejemplos de enfoques de dominio de tiempo son:

- 35
- [4] Philip J. Wilson, Harprit Chhatwal, "Adaptive transform coder having long term predictor", Patente de Estados Unidos 5.012.517, 30 de abril de 1991.
  - [5] Jeongook Song, Chang-Heon Lee, Hyen-O Oh, Hong-Goo Kang, "Harmonic Enhancement in Low Bitrate Audio Coding Using and Efficient Long-Term Predictor", EURASIP Journal on Advances in Signal Processing 2010.
  - [6] Juin-Hwey Chen, "Pitch-based pre-filtering and post-filtering for compression of audio signals", Patente de Estados Unidos 8.738.385, 27 de mayo de 2014.
- 40

Un objetivo de la presente invención es proporcionar un concepto mejorado para el procesamiento de una señal de audio.

45 Este objetivo se logra mediante un aparato para procesar una señal de audio de acuerdo con la reivindicación 1, un método para el procesamiento de una señal de audio de acuerdo con la reivindicación 12, un sistema de acuerdo con la reivindicación 13, un método para operar un sistema de acuerdo con la reivindicación 15 o un programa informático de acuerdo con la reivindicación 16. La presente invención se basa en el hallazgo de que la calidad subjetiva de una señal de audio se puede mejorar sustancialmente usando un filtro posterior de armónicos que tiene una función de transferencia que comprende un numerador y un denominador. El numerador de la función de transferencia comprende un valor de ganancia indicado por una información de ganancia transmitida y el denominador comprende una parte entera de un desfase de tono indicado por una información de desfase de tono y un filtro de derivaciones múltiples que depende de una parte fraccionaria del desfase de tono.

50 Por lo tanto, es posible eliminar el ruido interarmónico introducido por un decodificador de audio de cambio de dominio típico como un artefacto. Este filtro posterior de armónicos es particularmente útil en que se basa en la información transmitida, es decir, la ganancia de tono y el desfase de tono que están disponibles de todos modos en un decodificador, ya que esta información se recibe de un codificador correspondiente a través de una señal de entrada del decodificador. Además, el filtrado posterior es de una precisión específica debido al hecho de que no solo se tiene en cuenta la parte entera del desfase de tono, sino que, además, se tiene en cuenta la parte fraccionaria del desfase de tono. La parte fraccionaria del desfase de tono se puede introducir particularmente en el filtro posterior a través de un filtro de derivaciones múltiples que tiene coeficientes de filtro que en realidad dependen de la parte fraccionaria del desfase de tono. Este filtro se puede implementar como un filtro FIR o también se puede implementar como cualquier otro filtro tal como un filtro IIR o una implementación de filtro diferente. Cualquier cambio de dominio tal como un cambio de tiempo a frecuencia o un cambio de LPC a tiempo o un cambio de tiempo

55

60

a LPC o un cambio de frecuencia a tiempo se puede mejorar ventajosamente mediante el concepto de filtro posterior de la invención. Preferentemente, sin embargo, el cambio de dominio es un cambio de dominio de frecuencia a tiempo.

5 Por lo tanto, las realizaciones de la presente invención reducen el ruido interarmónico introducido por un códec de audio de transformada basado en un predictor a largo plazo que trabaja en el dominio de tiempo. Contrariamente a [04] a [6], donde se usan tanto un prefiltro antes de la codificación de transformada como un filtro posterior después la decodificación de transformada, preferentemente la presente invención aplica únicamente un filtro posterior.

10 Además, se ha observado que el prefiltro empleado en [04] a [6] tiene la tendencia a introducir inestabilidades en la señal de entrada dada al codificador de transformada. Estas inestabilidades se deben a cambios en la ganancia y/o desfase de tono de trama a trama. El codificador de transformada tiene dificultades en la codificación de dichas inestabilidades, particularmente en tasas de bits bajas, y a veces se introducirá aún más ruido en la señal decodificada en comparación con una situación sin ningún prefiltro o filtro posterior.

15 Preferentemente, la presente invención no emplea ningún prefiltro en absoluto y, por lo tanto, evita completamente los problemas implicados con un prefiltro.

20 Además, la presente invención se basa en un filtro posterior que se aplica en la señal decodificada después de la codificación de transformada. Este filtro posterior se basa en un filtro de predicción a largo plazo que tiene en cuenta la parte entera y la parte fraccionaria del desfase de tono que reduce el ruido interarmónico introducido por el códec de audio de transformada.

25 Para una mejor robustez, los parámetros del filtro posterior desfase de tono y ganancia de tono se estiman en el lado del codificador y transmiten en la secuencia de bits. Sin embargo, en otras implementaciones, el desfase de tono y la ganancia de tono también se pueden estimar en el lado del decodificador basándose en la señal de audio decodificada obtenida por un decodificador de audio que comprende un conversor de frecuencia-tiempo para convertir una representación de frecuencia de la señal de audio en una representación de dominio de tiempo de la señal de audio.

30 En una realización preferida, el numerador comprende adicionalmente un filtro de derivaciones múltiples para una parte fraccionaria cero del desfase de tono con el fin de compensar una inclinación espectral introducida mediante el filtro de derivaciones múltiples en el denominador, que depende de la parte fraccionaria del desfase de tono.

35 Preferentemente, el filtro posterior se configura para suprimir una cantidad de energía entre armónicos en una trama, en la que la cantidad de energía suprimida es menor del 20 % de una energía total de la representación de dominio de tiempo en la trama.

40 En una realización adicional, el denominador comprende un producto entre el filtro de derivaciones múltiples y el valor de ganancia.

45 En una realización adicional, el numerador de filtro además comprende un producto de un primer valor escalar y un segundo valor escalar, en el que el denominador comprende únicamente el segundo valor escalar en lugar del primer valor escalar. Estos valores escalares se establecen en valores predeterminados y tienen valores mayores de 0 y menores de 1; y, adicionalmente, el segundo valor escalar es menor que el primer valor escalar. Por lo tanto, es posible establecer de una manera muy eficaz las características de eliminación de energía que típicamente no son deseadas y establecer adicionalmente la intensidad del filtro, es decir, con qué intensidad atenúa el filtro artefactos interarmónicos en una señal de salida del decodificador del dominio de transformada.

50 El aparato además comprende, en una realización preferida, un controlador de filtro para establecer al menos el segundo valor escalar dependiendo de una tasa de bits de modo que se establece un valor más alto para una tasa de bits menor y viceversa.

55 Además, el controlador de filtro se configura para seleccionar, dependiendo de la parte fraccionaria del desfase de tono, el correspondiente filtro de derivaciones múltiples de una manera dependiente de la señal con el fin de fijar la señal en forma adaptativa del filtro posterior de armónicos, es decir, dependiente del valor de la parte fraccionaria realmente proporcionado del desfase de tono.

60 Posteriormente, se analizan las realizaciones preferidas de la presente invención en el contexto de los dibujos que acompañan, en los que:

la Figura 1 ilustra una realización de un aparato inventivo para procesar una señal de audio;

la Figura 2 ilustra una implementación preferida del filtro posterior de armónicos representado como funciones de

transferencia en el dominio  $z$ ;

- 5 la Figura 3 ilustra una realización preferida adicional para el filtro posterior de armónicos representado por una función de transferencia en el dominio  $z$ ;
- la Figura 4 ilustra una implementación preferida de un codificador para generar una señal codificada a decodificar mediante un decodificador de audio de dominio de transformada ilustrado en la Figura 1;
- 10 la Figura 5 ilustra una implementación preferida del filtro de derivaciones múltiples como un filtro FIR controlado por un controlador de filtro;
- la Figura 6 ilustra una cooperación entre el controlador de filtro y una memoria que tiene ponderaciones de derivaciones almacenados previamente dependiendo de la parte fraccionaria;
- 15 la Figura 7a ilustra una respuesta de frecuencia de un filtro que tiene un valor  $\alpha$  cero;
- la Figura 7b ilustra una respuesta de frecuencia de un filtro posterior de armónicos preferido que tiene un valor  $\alpha$  igual a 1;
- 20 la Figura 7c ilustra una respuesta de frecuencia de un filtro posterior de armónicos preferido que tiene un valor  $\alpha$  de 0,8;
- la Figura 8a ilustra una realización preferida de un filtro posterior de armónicos que tiene un valor  $\beta$  igual a 0,4; y
- 25 la Figura 8b ilustra una respuesta de frecuencia de un filtro posterior de armónicos que tiene un valor  $\beta$  de 0,2.

La Figura 1 ilustra un aparato para procesar una señal de audio que tiene asociada con la misma una información de desfase de tono y una información de ganancia. Esta información de ganancia se puede transmitir a un decodificador 100 a través de una entrada del decodificador 102 que recibe una señal codificada o, como alternativa, esta información se puede calcular en el propio decodificador, cuando esta información no está disponible. Sin embargo, para una operación más robusta, se prefiere calcular la información de desfase de tono y la información de ganancia de tono en el lado del codificador.

35 El decodificador 100 comprende, por ejemplo, un conversor de frecuencia-tiempo para convertir una representación de tiempo-frecuencia de la señal de audio en una representación de dominio de tiempo de la señal de audio. Por lo tanto, el decodificador no es un códec de voz de dominio de tiempo puro, sino que comprende un decodificador de dominio de transformada puro o un decodificador de dominio de transformada mixto o cualquier otro codificador que opera en un dominio diferente de un dominio de tiempo. Además, se prefiere que el segundo dominio sea el dominio de tiempo.

40 El aparato además comprende un filtro posterior de armónicos 104 para filtrar la representación de dominio de tiempo de la señal de audio, y este filtro posterior de armónicos se basa en una función de transferencia que comprende un numerador y un denominador. Particularmente, el numerador comprende un valor de ganancia indicado por la información de ganancia y el denominador comprende una parte entera de un desfase de tono indicado por la información de desfase de tono y, de forma importante, además comprende un filtro de derivaciones múltiples que depende de una parte fraccionaria del desfase de tono.

Una implementación preferida de este filtro posterior de armónicos con una función de transferencia  $H(z)$  se ilustra en la Figura 2. Este filtro recibe la señal de salida del decodificador 106 y somete esta señal de salida decodificada a una operación de filtrado posterior para obtener una señal de salida post filtrada 108. Esta señal de salida post filtrada se puede extraer como la señal procesada o además se puede procesar por cualquier procedimiento para eliminar cualquier discontinuidad introducida mediante la operación de filtrado posterior que, por supuesto, es dependiente de la señal, es decir, puede variar de trama a trama. Esta operación de eliminación de discontinuidad puede ser cualquiera de las operaciones de eliminación de discontinuidad bien conocidas tal como desvanecimiento cruzado, lo que significa que se desvanece una trama anterior y, al mismo tiempo, aparece gradualmente una nueva trama y, preferentemente, la característica de desvanecimiento es para que los factores de desvanecimiento se sumen a uno durante toda la operación de desvanecimiento cruzado. Sin embargo, también se puede aplicar otra eliminación de discontinuidad tal como un filtrado de paso bajo o filtrado LPC.

60 El aparato para procesar una señal de audio ilustrado en la Figura 1 además comprende un almacenamiento 112 de información de filtro de derivaciones múltiples y un controlador de filtro 114. Particularmente, el controlador de filtro 114 recibe la información lateral 116 del decodificador 100, y esta información lateral puede, por ejemplo, ser la información de ganancia de tono  $g$  y la información de desfase de tono, es decir, la información en la parte entera  $T_{int}$  del desfase de tono y la parte fraccionaria  $T_{fr}$  del desfase de tono. Esta información es útil para establecer el filtro

posterior de armónicos de trama a trama y, adicionalmente, para seleccionar una información del filtro de derivaciones múltiples  $B(z, T_{fr})$ . Además, la información adicional tal como la tasa de bits aplicada por el decodificador o la tasa de muestreo que subyace a la señal decodificada también se puede usar por el control de filtro 114 con el fin de fijar particularmente los valores escalares  $\alpha$ ,  $\beta$  para una cierta configuración de codificador y/o decodificador con respecto a la tasa de bits y la tasa de muestreo.

La Figura 2 ilustra una representación de polo/cero de una función de transferencia de filtro  $H(z)$  en el dominio  $z$  de acuerdo con lo conocido en la técnica. Naturalmente, hay numerosas otras representaciones del filtro posterior de armónicos, que son todas representaciones de filtro, que se pueden convertir en el tipo de representación de polo/cero en el dominio  $z$ . Por lo tanto, la presente invención es aplicable para cada filtro, que se puede describir de cualquier manera mediante tal función de transferencia de acuerdo con lo ilustrado en la memoria descriptiva.

La Figura 3 ilustra una realización preferida del filtro posterior de armónicos de nuevo descrito como una función de transferencia en la notación de polo/cero en el dominio  $z$ .

El filtro se puede describir de la siguiente manera:

$$H(z) = \frac{1 - \alpha\beta g B(z, 0)}{1 - \beta g B(z, T_{fr}) z^{-T_{int}}}$$

con  $g$  la ganancia decodificada,  $T_{int}$  y  $T_{fr}$  la parte entera y fraccionaria del desfase de tono decodificada,  $\alpha$  y  $\beta$  dos escalares que pesan la ganancia, y  $B(z, T_{fr})$  un filtro FIR de paso bajo cuyos coeficientes dependen de la parte fraccionaria del desfase de tono decodificada.

Obsérvese que  $B(z, 0)$  en el numerador de  $H(z)$  se usa para compensar la inclinación introducida por  $B(z, T_{fr})$ .

$\beta$  se usa para controlar la intensidad del filtro posterior. Una  $\beta$  igual a 1 produce efectos completos, suprimiendo la cantidad de energía máxima posible entre los armónicos. Una  $\beta$  igual a 0 deshabilita el filtro posterior. Por lo general, se usa un valor bastante bajo para no suprimir demasiada energía entre los armónicos. El valor también puede depender de la tasa de bits con un valor más alto a una tasa de bits más baja, por ejemplo 0,4 a una tasa de bits baja y 0,2 a una tasa de bits alta.

$\alpha$  se usa para agregar una ligera inclinación a la respuesta de frecuencia de  $H(z)$ , con el fin de compensar la ligera pérdida en energía en las frecuencias bajas. El valor de  $\alpha$  por lo general se elige próximo a 1, por ejemplo 0,8.

Se proporciona un ejemplo de  $B(z, T_{fr})$  en la Figura 6. El orden y los coeficientes de  $B(z, T_{fr})$  también pueden depender de la tasa de bits y la tasa de muestreo de salida. Se puede diseñar y sintonizar una respuesta de frecuencia diferente para cada combinación de la tasa de bits y la tasa de muestreo de salida.

Particularmente, se ha descubierto que son útiles incluso valores para  $\alpha$  entre 0,6 y menores de 1,0 y que, adicionalmente, se ha demostrado que los valores para  $\beta$  entre 0,1 y 0,5 también son útiles.

Además, el filtro de derivaciones múltiples puede tener un número variable de derivaciones. Se ha descubierto que para ciertas implementaciones, cuatro derivaciones son suficientes, en las que una pulsación es  $z^{+1}$ . Sin embargo, filtros más pequeños con únicamente dos derivaciones o filtros aún más grandes con más de cuatro derivaciones son útiles para ciertas implementaciones.

La Figura 6 ilustra una implementación preferida de filtros  $B(z)$  para diferentes valores fraccionarios del desfase de tono y, particularmente, para una resolución de desfase de tono de  $\frac{1}{4}$ . Para esta implementación, se ilustran cuatro descripciones de filtro diferentes para el filtro de derivaciones múltiples en el denominador de la función de transferencia del filtro posterior de armónicos. Sin embargo, se ha descubierto que los coeficientes de filtro no necesariamente tienen que indicar exactamente los valores ilustrados en la Figura 6, pero ciertas variaciones de  $\pm 0,05$  pueden ser útiles también en otras implementaciones.

Particularmente, de acuerdo con lo ilustrado en la Figura 1, las ponderaciones de las derivaciones ilustrados en la Figura 6 se almacenan dentro de la memoria 112 para la información del filtro de derivaciones múltiples. El controlador de filtro 114 recibe la parte fraccionaria  $T_{fr}$  desde la línea 116 de la Figura 1 y, en respuesta a este valor, se dirige a la memoria 112 con el fin de recuperar, a través de una línea de recuperación 200 la información de filtro específica para la parte fraccionaria específica del desfase de tono. Esta información se reenvía a continuación a través de una línea de salida 202 al filtro posterior de armónicos 104 de modo que el filtro posterior de armónicos se establezca correctamente. Una cierta implementación del filtro FIR de derivaciones múltiples se ilustra en la Figura 5. La indicación de ponderación  $w_1$  a  $w_4$  corresponde a la notación en la Figura 6 y el controlador de filtro 114 se

aplica, en respuesta a la parte fraccionaria actual del desfase de tono las ponderaciones correspondientes para una cierta trama de audio. Las otras porciones tales como las porciones de retardo 501, 502, 503 y el combinador 505 se pueden implementar de acuerdo con lo ilustrado. En este contexto, se enfatiza que el valor de retardo 501 es, en la notación  $z$  un valor de retardo negativo, ya que se ha descubierto que es particularmente útil una representación de filtro FIR que tiene un valor de retardo negativo además de un valor de retardo positivo tal como 503 y 504.

Posteriormente, en la Figura 4 se ilustra una implementación de codificador preferida que tiene ciertos bloques funcionales y que opera sin ningún prefiltro. La porción de filtro ilustrada en la Figura 4 comprende un estimador de tono 402, un refinador de tono 404, un estimador de parte fraccionaria 406, un detector de transitorio 408, un estimador de ganancia 410 y un cuantificador de ganancia 412. La información proporcionada por el cuantificador de ganancia 412, el estimador de parte fraccionaria 406, el refinador de tono 404 y el bit de decisión generado por el detector de transitorio 408 se introducen en un formador de señal codificada 414. El formador de señal codificada proporciona una señal codificada 102, que luego se ingresa en el decodificador 100 ilustrado en la Figura 1. La señal codificada 102 comprenderá información de señal adicional no ilustrada en la Figura 4.

Posteriormente, se describe la funcionalidad del estimador de tono 402.

Se estima un desfase de tono (parte entera + parte fraccionaria) por trama (tamaño de trama por ejemplo 20 ms). Esto se hace en 3 etapas para reducir la complejidad y mejora la precisión de la estimación.

Se usa un algoritmo de análisis de tono que produce un contorno de evolución de tono suave (por ejemplo, Análisis de tono de bucle abierto descrito en Rec. ITU-T G.718, sección 6.6). Este análisis por lo general se lleva a cabo sobre una base de subtrama (tamaño de subtrama por ejemplo 10 ms), y produce una estimación de desfase de tono por subtrama. Obsérvese que estas estimaciones de desfase de tono no tienen ninguna parte fraccionaria y por lo general se estiman en una señal submuestreada (tasa de muestreo por ejemplo 6400 Hz). La señal usada puede ser cualquier señal de audio, por ejemplo, una señal de audio ponderada LPC de acuerdo con lo descrito en Rec. ITU-T G.718, sección 6.5.

El refinador de tono opera de la siguiente manera:  
 La parte entera final del desfase de tono se estima en una señal de audio  $x[n]$  que se ejecuta a la tasa de muestreo de codificador central, que por lo general es mayor que la tasa de muestreo de la señal submuestreada usada en a. (por ejemplo 12,8 kHz, 16 kHz, 32 kHz...). La señal  $x[n]$  puede ser cualquier señal de audio, por ejemplo, una señal de audio ponderada LPC.

La parte entera del desfase de tono luego es el desfase  $d_m$  que maximiza la función de autocorrelación

$$C(d) = \sum_{n=0}^N x[n]x[n-d]$$

con  $d$  alrededor de un desfase de tono  $T$  estimado en la etapa 1.a.

$$T - \delta_1 \leq d \leq T + \delta_2$$

El estimador de parte fraccionaria 406 opera de la siguiente manera:  
 La parte fraccionaria se halla mediante la interpolación de la función de autocorrelación  $C(d)$  calculada en la etapa 2.b. y la selección el desfase de tono fraccionario que maximiza la función de autocorrelación interpolada. La interpolación se puede realizar usando un filtro FIR de paso bajo de acuerdo con lo descrito en por ejemplo Rec. ITU-T G.718, sección 6.6.7.

El detector de transitorio 408 ilustrado en la Figura 4 se configura para generar un bit de decisión.

Si la señal de audio de entrada no contiene ningún contenido armónico, entonces no se codifica ningún parámetro en la secuencia de bits. Únicamente se envía 1 bit de manera tal que el decodificador sepa si tiene que decodificar o no los parámetros del filtro posterior. La decisión se toma basándose en diversos parámetros:

a. Correlación normalizada en el desfase de tono entero estimado en la etapa 1.b.

$$corr. \text{ norm.} = \frac{\sum_{n=0}^N x[n]x[n-d_m]}{\sqrt{\sum_{n=0}^N x[n]x[n]} \sqrt{\sum_{n=0}^N x[n-d_m]x[n-d_m]}}$$

La correlación normalizada es 1 si la señal de entrada es perfectamente predecible por el desfase de tono entero, y 0 si no es predecible en absoluto. Un valor alto (próximo a 1) entonces indicará una señal armónica. Para una decisión más robusta, la correlación normalizada de la trama pasada también se puede usar en la decisión, por ejemplo:

5 Si (corr. norm(actual)\*corr. norm.(anterior)) > 0,25, entonces la trama actual contiene algún contenido armónico (bit=1)

10 b. Características calculadas por un detector de transitorio (por ejemplo, Medida de planitud temporal, Cambio de energía máxima), para evitar la activación del filtro posterior en una señal que contiene un transitorio. Por ejemplo, Si (Planitudtemp > 3,5 o cambio de Energía máx. > 3,5) entonces establecer bit=0 y no enviar ningún parámetro

Además, el estimador de ganancia 410 calcula una ganancia a introducir en el cuantificador de ganancia 412.

15 La ganancia se estima por lo general sobre la señal de audio de entrada en la tasa de muestreo de codificador central, pero puede ser también cualquier señal de audio como la señal de audio ponderada LPC. Esta señal se indica  $y[n]$  y puede ser la misma o diferente de  $x[n]$ .

20 La predicción  $y_p[n]$  de  $y[n]$  se halla primero filtrando  $y[n]$  con el siguiente filtro

$$P(z) = B(z, T_{fr})z^{-T_{int}}$$

con  $T_{int}$  la parte entera del desfase de tono (estimada en 1.b.) y  $B(z, T_{fr})$  un filtro FIR de paso bajo cuyos coeficientes dependen de la parte fraccionaria del desfase de tono  $T_{fr}$  (estimada en 1.c.).

25 Un ejemplo de  $B(z)$  cuando la resolución de desfase de tono es  $\frac{1}{4}$ :

$$T_{fr} = \frac{0}{4} \quad B(z) = 0,0000z^{-2} + 0,2325z^{-1} + 0,5349z^0 + 0,2325z^1$$

$$T_{fr} = \frac{1}{4} \quad B(z) = 0,0152z^{-2} + 0,3400z^{-1} + 0,5094z^0 + 0,1353z^1$$

$$T_{fr} = \frac{2}{4} \quad B(z) = 0,0609z^{-2} + 0,4391z^{-1} + 0,4391z^0 + 0,0609z^1$$

$$T_{fr} = \frac{3}{4} \quad B(z) = 0,1353z^{-2} + 0,5094z^{-1} + 0,3400z^0 + 0,0152z^1$$

30 La ganancia  $g$  se calcula a continuación de la siguiente manera:

$$g = \frac{\sum_{n=0}^{N-1} y[n]y_p[n]}{\sum_{n=0}^{N-1} y_p[n]y_p[n]}$$

y se limita entre 0 y 1.

35 Por último, la ganancia se cuantifica por ejemplo en 2 bits, usando por ejemplo cuantificación uniforme.

Si la ganancia se cuantifica a 0, entonces no se codifica ningún parámetro en la secuencia de bits, únicamente el único bit de decisión (bit=0).

40 De acuerdo con lo señalado anteriormente, el filtro posterior se aplica en la señal de audio de salida después del decodificador de transformada. Procesa la señal en una base de trama por trama, con el mismo tamaño de trama que el usado en el lado del codificador tal como 20 ms. De acuerdo con lo ilustrado, se basa en un filtro de predicción a largo plazo  $H(z)$  cuyos parámetros se determinan a partir de los parámetros estimados en el lado del codificador y se decodifican de la secuencia de bits. Esta información comprende el bit de decisión, el desfase de tono y la ganancia. Si el bit de decisión es 0, entonces el desfase de tono y la ganancia no se decodifican y se  
45 asume que son 0 no escritos en absoluto en la secuencia de bits.

De acuerdo con lo analizado, si los parámetros de filtro son diferentes de una trama a la siguiente trama, se puede introducir una discontinuidad en la frontera entre las dos tramas. Para evitar discontinuidad, se aplica un eliminador de discontinuidad tal como un desvanecedor cruzado o cualquier otra implementación para ese propósito.

5 Además, en las Figuras 7a a 8b se ilustran diferentes maneras de establecer el filtro posterior de armónicos. Los gráficos ilustran la función de transferencia de dominio de frecuencia. El eje horizontal está relacionado con la frecuencia normalizada 1 y el eje vertical es la magnitud de la respuesta del filtro en dB. Se enfatiza que en todas las ilustraciones excepto en la Figura 7b, el filtro introduce una amplificación para frecuencias bajas, es decir, un cierto valor de magnitud dB positivo.

10 Particularmente, la Figura 7a ilustra una función de transferencia, que implementa el filtro en la Figura 3, con los ciertos valores de los parámetros de acuerdo con lo indicado anteriormente. Además, el valor  $\alpha$ , es decir, el primer valor escalar se establece a 0. La Figura 7b ilustra una situación similar, pero ahora con un valor  $\alpha$  igual a 1. Los otros parámetros son idénticos a la Figura 7a.

15 La Figura 7c ilustra una implementación adicional en la que  $\alpha$  es igual a 0,8 que tiene una ligera inclinación y una potenciación de las frecuencias más bajas. De nuevo, la Figura 7 tiene los mismos otros parámetros de acuerdo con lo indicado en la Figura 7a. Es evidente que  $\alpha$  igual a 1 elimina la inclinación y todas las frecuencias armónicas tienen una ganancia de 1. El inconveniente de esta configuración es una pérdida de energía en las frecuencias entre los armónicos. Por lo tanto, se prefiere un valor de  $\alpha$  igual a 0,8 como en la Figura 7c. Este valor agrega una ligera inclinación en comparación con la situación de  $\alpha$  igual a 1 de la Figura 7b. Con el fin de compensar la pérdida de las energías en las frecuencias entre los armónicos, se usa preferentemente esta ligera inclinación.

25 Además, las Figuras 8a y 8b ilustran configuraciones de filtros para un valor de  $\alpha$  igual a 0,8 y diferentes valores  $\beta$ , es decir, un valor  $\beta$  de 0,4 en la Figura 8a y un valor  $\beta$  de 0,2 en la Figura 8b. Es evidente que un valor  $\beta$  de 0,4 tiene un efecto más intenso de filtrado posterior en comparación con un valor  $\beta$  de 0,2 y, por lo tanto, se usa un valor  $\beta$  de 0,4 en tasas de bits más bajas con el fin de eliminar el ruido interarmónico introducido por una tasa de bits baja de este tipo.

30 Por otra parte,  $\beta$  igual a 0,2 tiene un efecto menos intenso para suprimir energía entre los armónicos y, por lo tanto, se prefiere este valor  $\beta$  para tasas de bits altas debido al hecho de que en tales tasas de bits más altas no existe tanto ruido interarmónico.

35 Si bien se han descrito algunos aspectos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del método correspondiente, en el que un bloque o dispositivo corresponde a una etapa del método o una característica de una etapa del método. De forma análoga, los aspectos descritos en el contexto de una etapa del método también representan una descripción de un bloque correspondiente o un elemento o característica de un aparato correspondiente. Algunas o todas las etapas del método se pueden ejecutar mediante (o usando) 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 más importantes del método se pueden ejecutar por un aparato de este tipo.

45 La señal transmitida o codificada inventiva se puede almacenar en un medio de almacenamiento digital o se puede transmitir en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable tal como la Internet.

50 Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención se pueden implementar en hardware o en software. La aplicación se puede realizar usando un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control legibles electrónicamente almacenadas en el mismo, que cooperan (o son capaces de cooperar) con un sistema informático programable de tal manera que se lleve a cabo el método respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

55 Algunas realizaciones de acuerdo con la invención comprenden un portador de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable, de tal manera que se realiza uno de los métodos descritos en este documento.

60 Por lo general, las realizaciones de la presente invención se pueden implementar como un producto de programa informático con un código de programa, siendo el código de programa operativo para realizar uno de los métodos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa se puede almacenar, por ejemplo, en un portador legible por máquina.

Otras realizaciones comprenden el programa informático para realizar uno de los métodos descritos en este documento, almacenado en un portador legible por máquina.

En otras palabras, una realización del método inventivo es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los métodos descritos en este documento, cuando el programa informático se ejecuta en un ordenador.

5 Una realización adicional del método inventivo es, por lo tanto, un portador de datos (o un medio de almacenamiento no transitorio tal como un medio de almacenamiento digital o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para realizar uno de los métodos descritos en este documento. El portador de datos, el medio de almacenamiento digital o el medio grabado son típicamente tangibles y/o no transitorios.

10 Una realización adicional del método de la invención es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para realizar uno de los métodos descritos en este documento. El flujo de datos o la secuencia de señales se pueden configurar, por ejemplo, para transferirse a través de una conexión de comunicación de datos, por ejemplo, a través de la internet.

15 Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador o un dispositivo lógico programable, configurado para, o adaptado para, realizar uno de los métodos descritos en este documento.

20 Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para realizar uno de los métodos descritos en este documento.

25 Una realización adicional de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa informático para realizar uno de los métodos descritos en este documento a un receptor. El receptor puede, por ejemplo, ser un ordenador, un dispositivo móvil, un dispositivo de memoria o similares. El aparato o sistema puede, por ejemplo, comprender un servidor de archivos para transferir el programa informático al receptor.

30 En algunas realizaciones, se puede usar un dispositivo lógico programable (por ejemplo, un campo de matriz de puertas programables) para realizar algunas o todas de las funcionalidades de los métodos descritos en este documento. En algunas realizaciones, un campo de matriz de puertas programables puede cooperar con un microprocesador con el fin de realizar uno de los métodos descritos en este documento. Por lo general, los métodos se realizan preferentemente mediante cualquier aparato de hardware.

35 Las realizaciones descritas anteriormente son meramente ilustrativas para los principios de la presente invención. Se comprende que serán evidentes modificaciones y variaciones de las disposiciones y los detalles descritos en este documento para otros expertos en la técnica. Se pretende, por lo tanto, limitarse únicamente por el alcance de las reivindicaciones de la patente inminentes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en este documento.

40

**REIVINDICACIONES**

1. Aparato para procesar una señal de audio que tiene asociada con la misma una información de desfase de tono y una información de ganancia, que comprende:

5 un conversor de dominio (100) para convertir una primera representación de dominio de la señal de audio en una segunda representación de dominio de la señal de audio; y  
 un filtro posterior de armónicos (104) para filtrar la segunda representación de dominio de la señal de audio, en el que el filtro posterior se basa en una función de transferencia que comprende un numerador y un denominador,  
 10 en el que el numerador comprende un valor de ganancia indicado por la información de ganancia, y en el que el denominador comprende una parte entera de un desfase de tono indicado por la información de desfase de tono y un filtro de derivaciones múltiples que depende de una parte fraccionaria del desfase de tono.

15 2. Aparato de la reivindicación 1, en el que la función de transferencia del filtro posterior comprende, en el numerador, un filtro FIR de derivaciones múltiples adicional para una parte fraccionaria cero del desfase de tono.

3. Aparato de la reivindicación 1 o 2, en el que el denominador comprende un producto entre el filtro de derivaciones múltiples y el valor de ganancia.

20 4. Aparato de una de las reivindicaciones anteriores, en el que el numerador comprende además un producto de un primer valor escalar y un segundo valor escalar, en el que el denominador comprende el segundo valor escalar y no el primer valor escalar, en el que el primer y segundo valores escalares están predeterminados y tienen valores mayores de 0 y menores de, y en el que el segundo valor escalar es menor que el primer valor escalar.

25 5. Aparato de acuerdo con la reivindicación 4, comprendiendo además:  
 un controlador de filtro (114) configurado para establecer el segundo valor escalar dependiendo de una tasa de bits, mediante la que se opera el conversor de frecuencia-tiempo (100), en el que el segundo valor escalar se establece a un primer valor, cuando la tasa de bits tiene un primer valor, en el que el segundo valor escalar se establece a un segundo valor, cuando la tasa de bits tiene un segundo valor, en el que el segundo valor de la tasa de bits es menor  
 30 que el primer valor de la tasa de bits, y en el que el segundo valor del segundo valor escalar es mayor que el primer valor del segundo valor escalar.

35 6. Aparato de acuerdo con la reivindicación 4 o 5, en el que el primer valor escalar se establece entre 0,6 y 1,0 y en el que el segundo valor escalar se establece entre 0,1 y 0,5.

7. Aparato de una de las reivindicaciones anteriores, en el que el filtro posterior tiene la función de transferencia H(z) en una representación de polo cero basándose en la siguiente ecuación:

$$H(z) = \frac{1 - \alpha\beta g B(z, 0)}{1 - \beta g B(z, T_{fr}) z^{-T_{int}}}$$

40 en la que  $\alpha$  es un primer valor escalar, en la que  $\beta$  es un segundo valor escalar, en la que B(z,0) es un filtro de derivaciones múltiples para un desfase de tono de parte fraccionaria cero, en la que B(z,T<sub>fr</sub>) es un filtro de derivaciones múltiples que depende de la parte fraccionaria del desfase de tono, en la que T<sub>int</sub> es la parte entera del desfase de tono, en la que T<sub>fr</sub> es la parte fraccionaria del desfase de tono, y en la que g es el valor de ganancia indicado por la información de ganancia z es una variable en un plano z.

45 8. Aparato de una de las reivindicaciones anteriores, en el que el filtro de derivaciones múltiples es un filtro de respuesta finita al impulso (FIR) y tiene al menos tres derivaciones.

50 9. Aparato de una de las reivindicaciones anteriores, en el que el filtro de derivaciones múltiples en el denominador comprende cuatro derivaciones, en el que, para una parte fraccionaria cero, la primera pulsación está entre 0,0 y 0,1, la segunda pulsación está entre 0,2 y 0,3, la tercera pulsación está entre 0,5 y 0,6 y la cuarta pulsación está entre 0,2 y 0,3,  
 55 en el que el filtro de derivaciones múltiples comprende, para una primera parte fraccionaria, cuatro derivaciones de filtro, en el que la primera pulsación está entre 0,0 y 0,1, la segunda pulsación está entre 0,3 y 0,4, la tercera pulsación está entre 0,45 y 0,55 y la cuarta pulsación está entre 0,1 y 0,2,  
 en el que el filtro de derivaciones múltiples comprende, para una segunda parte fraccionaria, cuatro derivaciones de filtro, en el que la primera pulsación está entre 0,0 y 0,1, la segunda pulsación está entre 0,35 y 0,45, la tercera pulsación está entre 0,35 y 0,45 y la cuarta pulsación está entre 0,0 y 0,1,  
 60 en el que el filtro de derivaciones múltiples comprende, para una tercera parte fraccionaria, cuatro derivaciones de

filtro, en el que la primera pulsación está entre 0,1 y 0,2, la segunda pulsación está entre 0,45 y 0,55, la tercera pulsación está entre 0,3 y 0,4 y la cuarta pulsación está entre 0,0 y 0,1, en el que la tercera parte fraccionaria es mayor que la segunda parte fraccionaria, y en el que la segunda parte fraccionaria es mayor que la primera parte fraccionaria.

5 10. Aparato de una de las reivindicaciones anteriores, en el que el filtro posterior se configura para tener una inclinación espectral negativa para compensar una pérdida de energía por el filtro posterior de armónicos, o  
 10 en el que el filtro posterior está configurado para suprimir una cantidad de energía entre armónicos en una trama, en el que la cantidad de energía suprimida es menor del 20 % de una energía total de la representación de dominio de tiempo en la trama.

11. Aparato de una de las reivindicaciones precedentes, en el que el conversor de dominio es un conversor de frecuencia-tiempo, en el que el primer dominio es un dominio de frecuencia y el segundo dominio es un dominio de tiempo, o  
 15 en el que el conversor de dominio es un convertidor de tiempo residual LPC, en el que el primer dominio es un dominio residual LPC y el segundo dominio es un dominio de tiempo.”.

12. Método para procesamiento de una señal de audio que tiene asociada con la misma una información de desfase de tono y una información de ganancia, que comprende:

convertir (100) una representación de frecuencia de la señal de audio en una representación de dominio de tiempo de la señal de audio; y  
 25 filtrar la representación de dominio de tiempo de la señal de audio mediante un filtro posterior de armónicos (104), en el que el filtro posterior se basa en una función de transferencia que comprende un numerador y un denominador, en el que el numerador comprende un valor de ganancia indicado por la información de ganancia, y en el que el denominador comprende una parte entera de un desfase de tono indicado por la información de desfase de tono y un filtro de derivaciones múltiples que depende de una parte fraccionaria del desfase de tono.

30 13. Sistema para procesar una señal de audio que comprende un codificador para codificar una señal de audio y un decodificador que comprende un procesador, comprendiendo el procesador:

un conversor de dominio (100) para convertir una representación de frecuencia de la señal de audio en una representación de dominio de tiempo de la señal de audio; y  
 35 un filtro posterior de armónicos (104) para filtrar la representación de dominio de tiempo de la señal de audio, en el que el filtro posterior se basa en una función de transferencia que comprende un numerador y un denominador, en el que el numerador comprende un valor de ganancia indicado por una información de ganancia, y en el que el denominador comprende una parte entera de un desfase de tono indicado por una información de desfase de tono y un filtro de derivaciones múltiples que depende de una parte fraccionaria del desfase de tono.  
 40

14. Sistema de la reivindicación 13, en el que el codificador comprende un calculador de desfase de tono (402, 404, 406) para calcular una parte entera y una parte fraccionaria del desfase de tono y un calculador de ganancia (410, 412) para calcular el valor de ganancia, y un formador de señal codificada (414) para generar una señal codificada (102) que comprende la información de desfase de tono y la información de ganancia.  
 45

15. Método de procesamiento de una señal de audio que comprende un método de codificación de una señal de audio y un método de decodificación que comprende:

50 convertir (100) una representación de frecuencia de la señal de audio en una representación de dominio de tiempo de la señal de audio; y  
 filtrar la representación de dominio de tiempo de la señal de audio usando un filtro posterior de armónicos (104), en el que el filtro posterior se basa en una función de transferencia que comprende un numerador y un denominador, en el que el numerador comprende un valor de ganancia indicado por una información de ganancia, y en el que el denominador comprende una parte entera de un desfase de tono indicado por una información de desfase de tono y un filtro de derivaciones múltiples que depende de una parte fraccionaria del desfase de tono.  
 55

16. Programa informático para realizar un método de la reivindicación 12 o la reivindicación 15, cuando el programa informático se está ejecutando en un ordenador o un procesador.  
 60

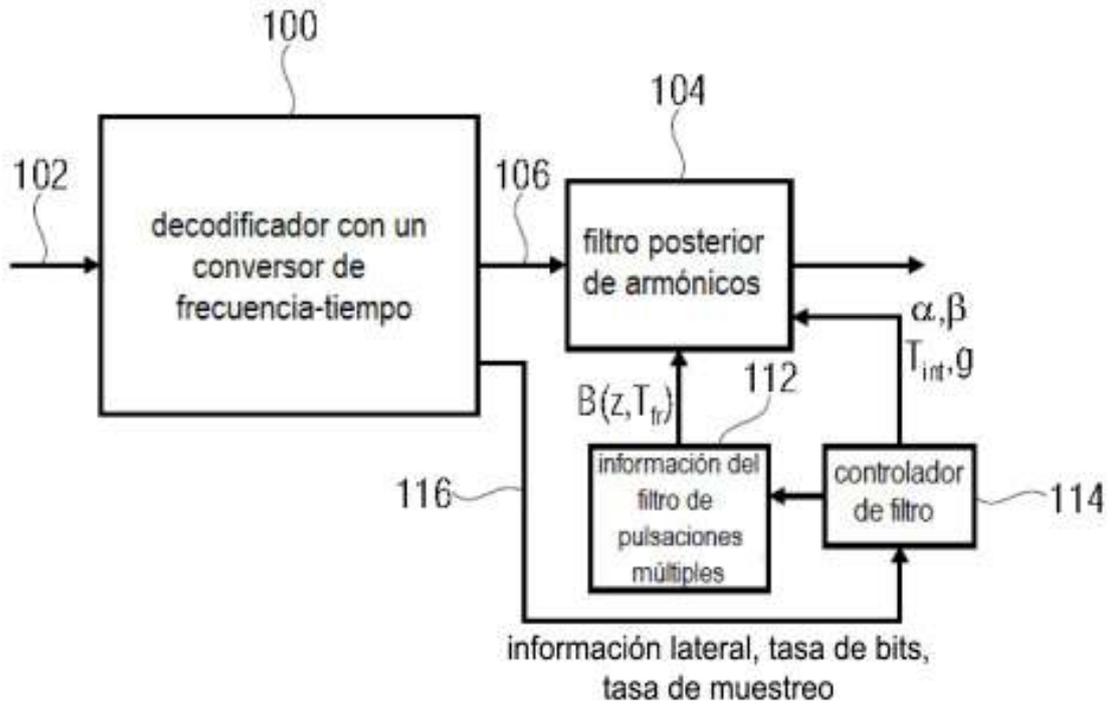


FIG 1

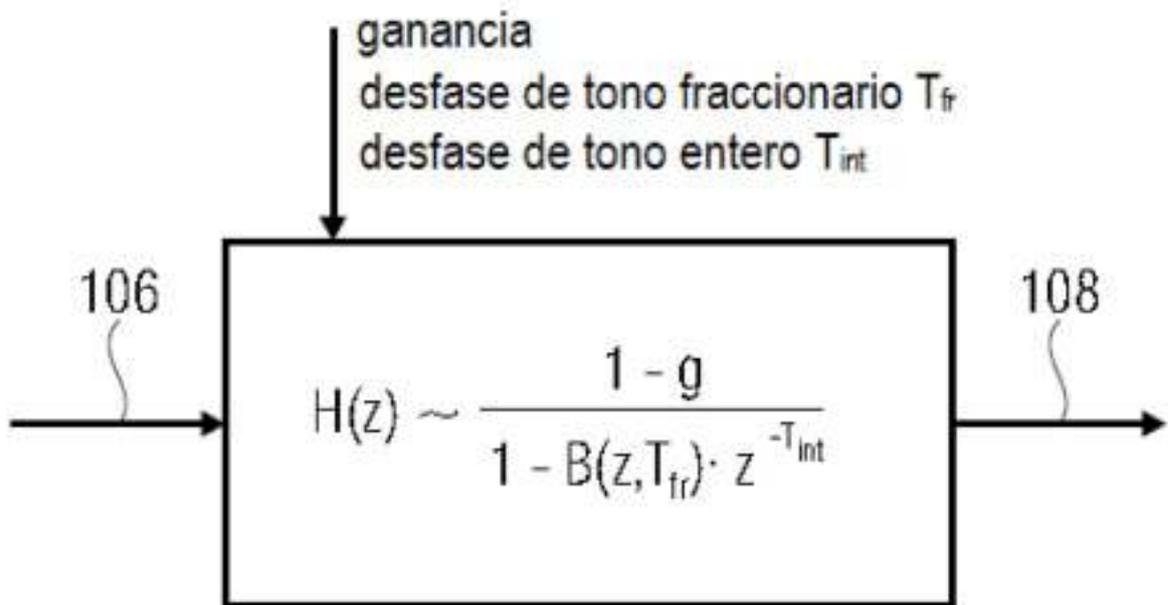


FIG 2

$$H(z) = \frac{1 - \alpha\beta g B(z, 0)}{1 - \beta g B(z, T_{fr}) z^{-T_{int}}}$$

$g$  : valor de ganancia

$T_{fr}$  : parte fraccionaria del desfase de tono

$T_{int}$  : parte entera del desfase de tono

$B(z, T_{fr})$  : filtro FIR de derivaciones múltiples dependiendo de parte fraccionaria

$B(z, 0)$  : filtro FIR de derivaciones múltiples para una parte fraccionaria cero

$\alpha$  : primer valor escalar  $0,6 \leq \alpha < 1,0$

$\beta$  : segundo valor escalar  $0,1 \leq \beta \leq 0,5$

**FIG 3**

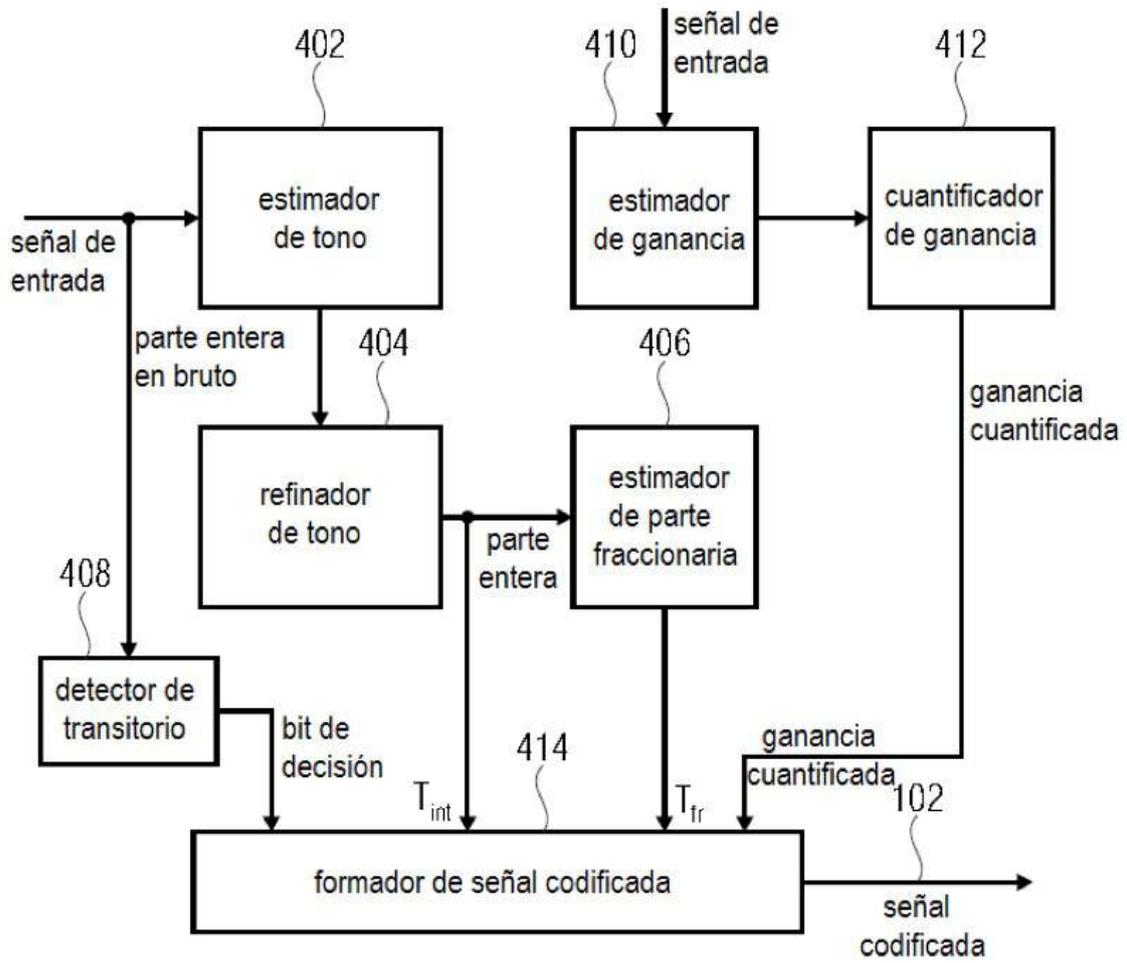


FIG 4

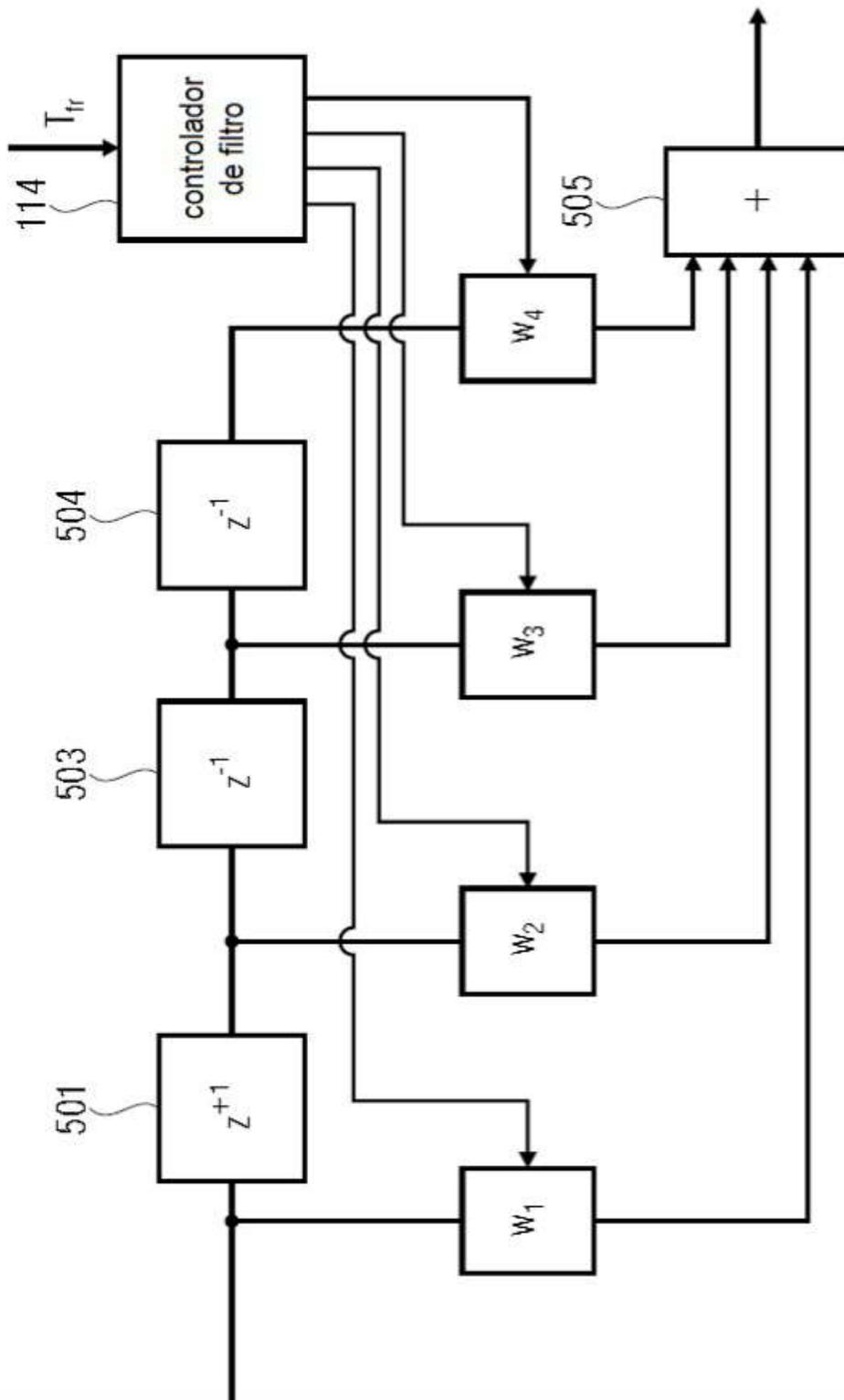


FIG 5

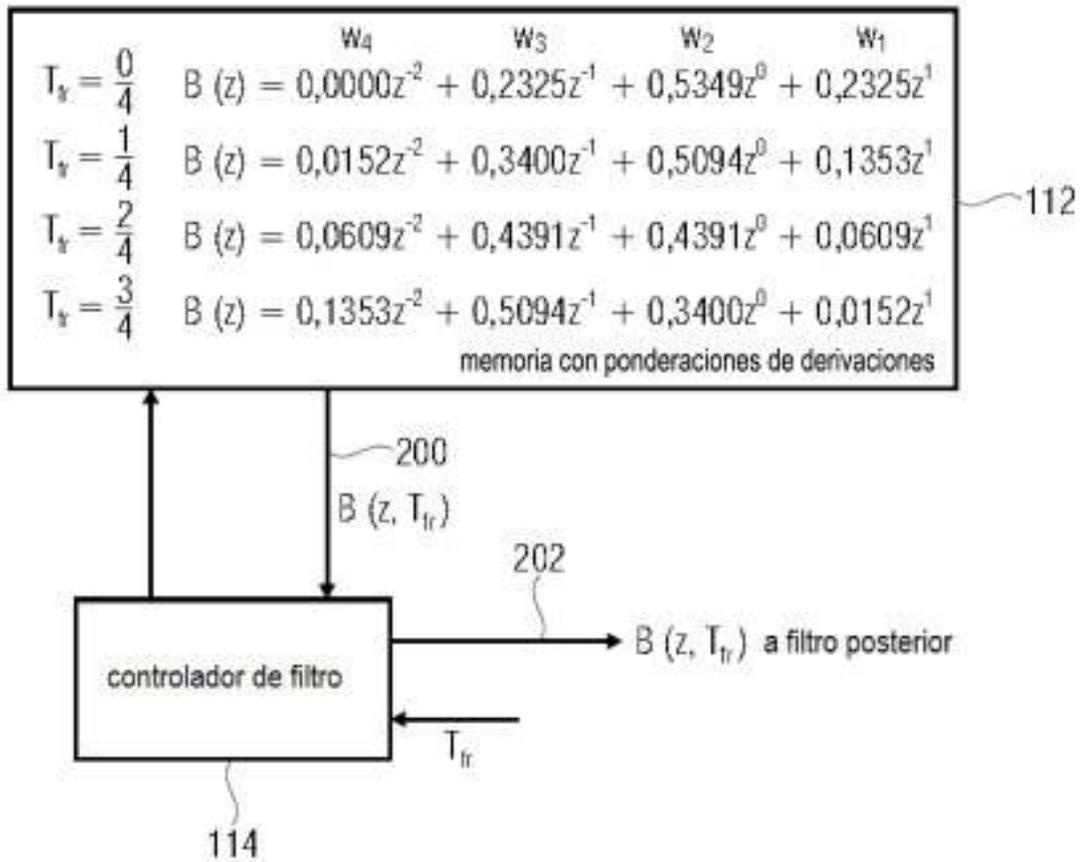


FIG 6

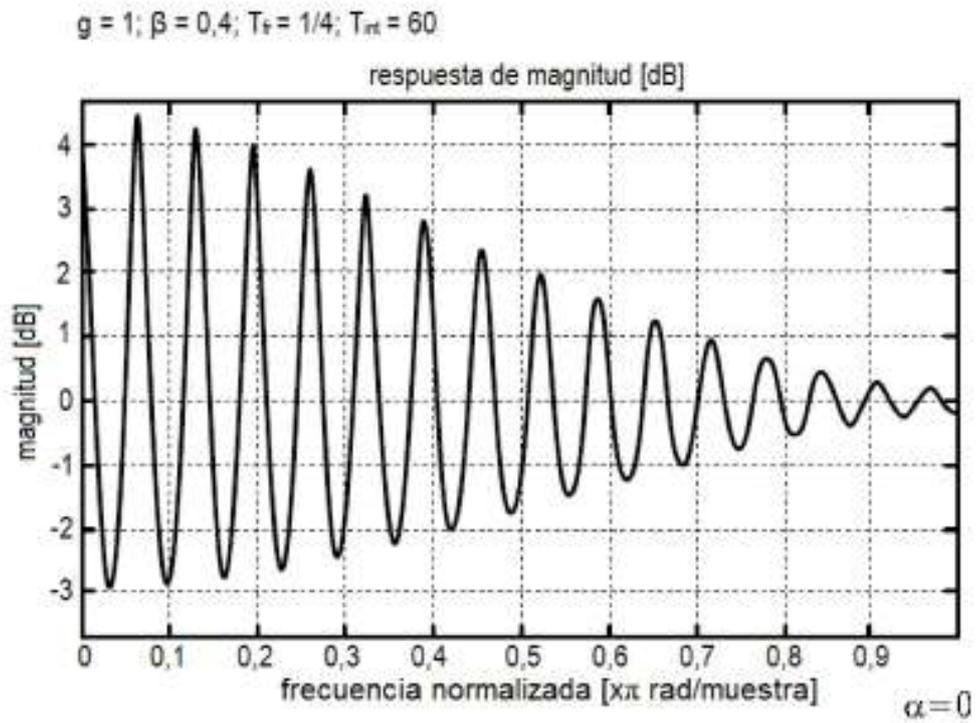


FIG 7A

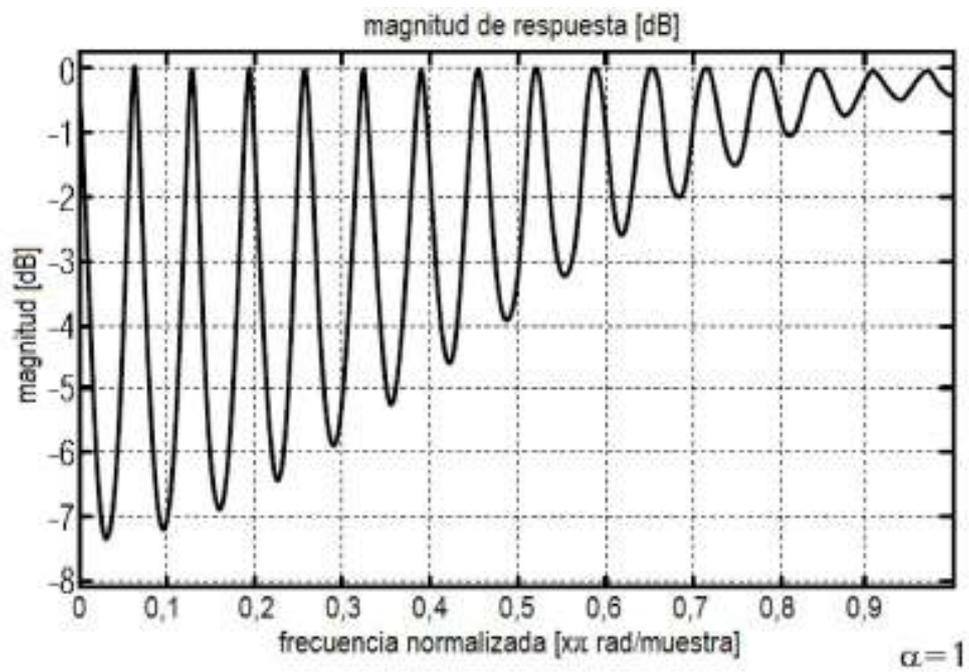


FIG 7B

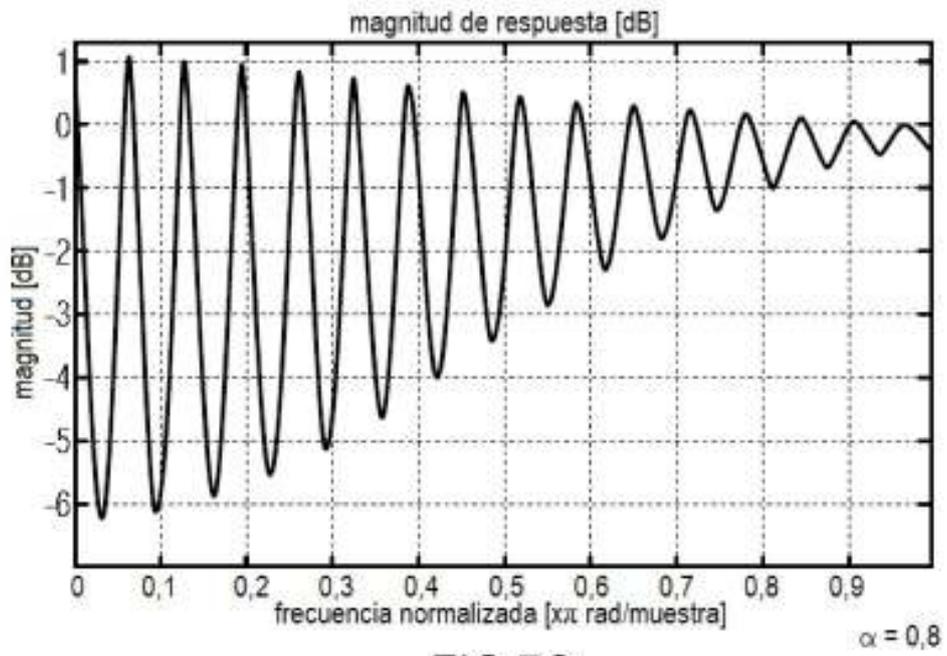


FIG 7C

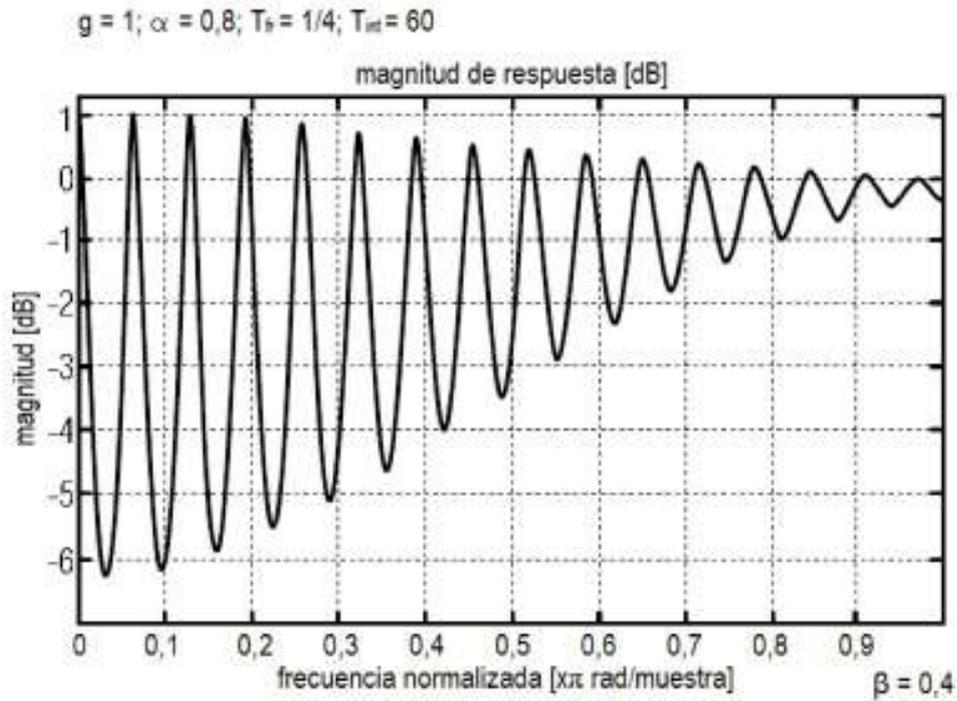


FIG 8A

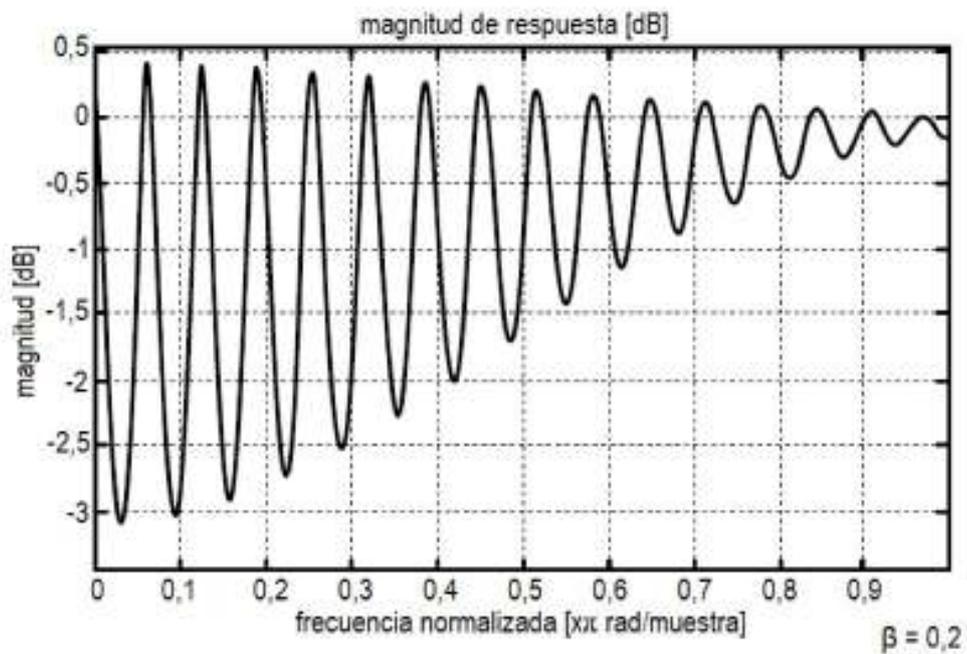


FIG 8B