

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 501 840**

51 Int. Cl.:

G10L 19/26 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.04.2011** **E 11780883 (2)**

97 Fecha y número de publicación de la concesión europea: **11.06.2014** **EP 2569767**

54 Título: **Procedimiento y disposición para el procesamiento de señales de audio**

30 Prioridad:

11.05.2010 US 333498 P

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

02.10.2014

73 Titular/es:

**TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)
(100.0%)
164 83 Stockholm, SE**

72 Inventor/es:

**GRANCHAROV, VOLODYA y
SVERRISSON, SIGURDUR**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 501 840 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

Procedimiento y disposición para el procesamiento de señales de audio

Campo técnico

5 La invención se refiere al procesamiento de señales de audio, en particular a un procedimiento y una disposición para mejorar la calidad perceptual mediante un post-filtrado.

Antecedentes

10 La codificación de audio a tasas de bits bajas o moderadas se usa ampliamente para reducir la carga de la red. Sin embargo, la reducción de la tasa de bits conduce inevitablemente a una disminución de la calidad debida a una mayor cantidad de ruido de cuantificación. Una manera de minimizar el impacto perceptual del ruido de cuantificación es usando un post-filtro. Un post-filtro funciona en el decodificador y afecta a los parámetros de la señal reconstruida o, directamente, a la forma de onda de la señal. El uso de un post-filtro tiene por objeto atenuar los valles del espectro, donde el ruido de cuantificación es más audible y conseguir, de esta manera, una calidad perceptual mejorada.

15 Ambos post-filtros de tono y formantes se usan para mejorar la calidad en los códecs de voz denominados ACELP (Predicción lineal excitada por código algebraico, Algebraic Code Excited Linear Prediction). Estos filtros funcionan en el dominio del tiempo y, típicamente, se basan en el modelo de voz usado en el códec ACELP [1]. Sin embargo, esta familia de post-filtros no es adecuada para su uso con códecs de audio con transformación, tales como por ejemplo, G.719 [2].

20 Otro ejemplo de post-procesador de voz para mejora una señal de voz dividida en una pluralidad de bandas en el dominio de la frecuencia se describe en la referencia [3].

De esta manera, hay una necesidad de mejorar la calidad perceptual de las señales de audio que han sido sometidas a una codificación de audio con transformación.

Sumario

25 Sería deseable conseguir una calidad perceptual de audio mejorada de las señales de audio que han sido sometidas a una codificación de audio con transformación. Un objeto de la invención es mejorar la calidad perceptual de una señal de audio que ha sido sometida a una codificación de audio con transformación. Además, un objeto de la invención es proporcionar un procedimiento y una disposición para el post-filtrado de una señal de audio que ha sido sometida a una codificación de audio con transformación. Estos objetos pueden ser conseguidos mediante un procedimiento y un aparato según las reivindicaciones independientes adjuntas. Las realizaciones se exponen en las reivindicaciones dependientes.

30 Según un primer aspecto, se proporciona un procedimiento en un decodificador según se reivindica en la reivindicación 1. El procedimiento implica obtener un vector d , que comprende coeficientes cuantificados de dominio MDCT de un segmento de tiempo de una señal de audio. Además, se deriva un vector \hat{d} procesado aplicando un post-filtro directamente al vector d . El post-filtro está configurado para tener una función de transferencia H que es una versión comprimida de la envolvente del vector d . Además, se deriva una forma de onda de la señal aplicando una transformada MDCT inversa al vector \hat{d} procesado.

35 Según un segundo aspecto, se proporciona un decodificador según se reivindica en la reivindicación 8. El decodificador comprende una unidad funcional adaptada para obtener un vector d , que comprende los coeficientes cuantificados en el dominio MDCT de un segmento de tiempo de una señal de audio. El decodificador comprende además una unidad funcional, adaptada para derivar un vector \hat{d} procesado mediante la aplicación de un post-filtro directamente al vector d . El post-filtro está configurado para tener una función de transferencia H que es una versión comprimida de la envolvente del vector d . El decodificador comprende además una unidad funcional adaptada para derivar una forma de onda de la señal aplicando una transformada MDCT inversa al vector \hat{d} procesado.

45 El procedimiento y la disposición indicados anteriormente, que implican un post-filtro MDCT, pueden ser usados para mejorar la calidad de los sistemas de codificación de audio con tasas de bits moderada y baja. Cuando el post-filtro es usado en un códec MDCT, la complejidad adicional es muy baja, ya que el post-filtro opera directamente sobre el vector MDCT.

El procedimiento y la disposición indicados anteriormente pueden ser implementados en diferentes realizaciones.

En algunas realizaciones, el denominador de la función de transferencia H está configurado para comprender un máximo del vector $|d|$, que puede ser una estimación obtenida realizando un seguimiento recursivo de máximo sobre el vector $|d|$. En algunas realizaciones, la función de transferencia H está configurada para comprender un componente de énfasis, configurado para controlar la agresividad del post-filtro sobre el espectro MDCT. Por ejemplo, el componente de énfasis podría depender de la frecuencia o podría ser constante. Además, la energía del vector \hat{d} procesado puede ser normalizada a la energía del vector d .

En algunas realizaciones, el vector \hat{d} procesado es derivado sólo cuando se determina que el segmento de tiempo de la señal de audio comprende voz. Además, la función de transferencia H podría ser limitada o suprimida cuando se determina que el segmento de tiempo de la señal de audio consiste principalmente en uno o más de entre, por ejemplo, voz sorda, ruido de fondo y música.

Las realizaciones anteriores se han descrito principalmente en términos de un procedimiento. Sin embargo, la descripción anterior pretende abarcar también realizaciones del decodificador, adaptadas para permitir la realización de las características descritas anteriormente. Las diferentes características de las realizaciones ejemplares anteriores pueden combinarse de diferentes maneras según las necesidades, los requisitos o las preferencias.

Breve descripción de los dibujos

La invención se describirá ahora más detalladamente por medio de realizaciones ejemplares y con referencia a los dibujos adjuntos, en los que:

La Figura 1 muestra un diagrama de un factor $a(k)$ de énfasis ejemplar, que disminuye (para limitar el efecto del post-filtro) conforme aumenta la frecuencia, según una realización ejemplar.

La Figura 2 muestra un diagrama que ilustra el efecto del post-filtro sobre un espectro de la señal, donde la línea de puntos delgada representa el espectro de la señal antes del post-filtro, y la línea continua representa el espectro de la señal después del post-filtro, según una realización ejemplar.

La Figura 3 muestra el resultado de un ensayo de escucha MUSHRA que compara un códec de audio MDCT con y sin post-filtro, según una realización ejemplar.

La Figura 4 es un diagrama de flujo que ilustra las acciones de un procedimiento realizado en un decodificador, según una realización ejemplar.

Las Figuras 5-7 son diagramas de bloques que ilustran una disposición respectiva en un decodificador y una entidad de gestión de audio, según realizaciones ejemplares.

Descripción detallada

En breve, se proporciona un decodificador que comprende un post-filtro, cuyo post-filtro está diseñado para trabajar con códecs con transformadas de tipo MDCT (Modified Discrete Cosine Transform, Transformada Discreta de Coseno Modificada), tales como por ejemplo, G.719 [2]. El post-filtro sugerido opera directamente en el dominio MDCT, y no requiere una transformación adicional de la señal de audio al dominio DFT o al dominio del tiempo, lo que mantiene la complejidad computacional baja. La mejora de la calidad debida al post-filtro se confirma en los ensayos de escucha.

El concepto de codificación mediante transformación es convertir, o transformar, una señal de audio a ser codificada al dominio de la frecuencia y, a continuación, cuantificar los coeficientes de frecuencia, que son almacenados o suministrados, a continuación, a un decodificador. El decodificador usa los coeficientes de frecuencia (cuantificados) recibidos para reconstruir la forma de onda de la señal de audio, aplicando la transformada de frecuencia inversa. La motivación detrás de este esquema de codificación es que los coeficientes en el dominio de la frecuencia pueden ser cuantificados de manera más eficiente que los coeficientes en el dominio del tiempo.

En un codificador de transformada de tipo MDCT, una forma de onda $x(n)$ de un bloque de la señal es transformada a un vector $d^*(k)$. La longitud "L" de dicho vector corresponde a segmentos de voz de 20-40 ms. La transformada MDCT puede definirse como:

$$d^*(k) = \sum_{n=0}^{L-1} \text{sen} \left[\left(n + \frac{1}{2} \right) \frac{\pi}{2} \right] \cos \left[\left(n + \frac{1}{2} \right) \left(k + \frac{1}{2} \right) \frac{\pi}{L} \right] x(n)$$

5 Los coeficientes MDCT son cuantificados, formando de esta manera un vector $d(k) = Q(d^*(k))$ de coeficientes MDCT cuantificados, que deben ser decodificados por un decodificador MDCT.

El post-filtro puede ser aplicado directamente al vector $d(k)$ recibido en el decodificador, y derivar de esta manera el vector \hat{d} post-filtrado como

$$\hat{d}(k) = H(k)d(k)$$

10 La función de transferencia o función de filtro, $H(k)$, es una versión comprimida de la envolvente del espectro MDCT:

$$H(k) = \left(\frac{\text{abs}[d(k)]}{\max[\text{abs}(d)]} \right)^{a(k)} \quad (1)$$

15 El parámetro $a(k)$ puede ser configurado para controlar la "agresividad" o "cantidad de énfasis" del post-filtro sobre el espectro MDCT. La Figura 1 muestra un diagrama de un ejemplo de cómo $a(k)$ puede ser configurado como un vector dependiente de la frecuencia. Sin embargo, $a(k)$ también podría ser constante a lo largo del espectro. El efecto del post-filtro sobre el espectro de la señal se ilustra en la Figura 2. Tal como puede verse en la Figura 2, los valles del espectro son más profundos después del post-filtrado.

Preferentemente, la energía de la salida del post-filtro puede ser normalizada a la energía de la entrada del post-filtro:

$$\hat{d}_{(normalizado)}(k) = \frac{\text{std}(d)}{\text{std}(\hat{d})} \hat{d}(k)$$

25 Aquí, $\text{std}(d)$ es la desviación estándar del vector d , que comprende los coeficientes MDCT cuantificados, antes de la operación de post-filtrado; y $\text{std}(\hat{d})$ es la desviación estándar del vector \hat{d} procesado, es decir, del vector d después de la operación de post-filtrado.

30 Además, el ruido de cuantificación audible debido a la codificación es más audible en la voz sonora, por ejemplo, en comparación con la música. De esta manera, por ejemplo, el uso del post-filtro sugerido es más eficaz para reducir el ruido de cuantificación audible en las señales de voz, en lugar de en las señales de música. De esta manera, cuando sea adecuado, el post-filtro podría ser desactivado, o suprimido, en tramas o segmentos de trama para los cuales se considera que el post-filtro es menos eficaz. Por ejemplo, el post-filtro podría ser desactivado, o suprimido, en tramas o segmentos de trama para los cuales se determina que consisten principalmente en voz sorda, ruido de fondo y/o la música. El post-filtro podría ser usado en combinación, por ejemplo, con un discriminador de voz-música, y/o un módulo de estimación de ruido de fondo, para determinar el contenido de una trama. Sin embargo, cabe señalar que el post-filtro no causa ninguna degradación, por ejemplo, en segmentos sordos.

40 El efecto percibido del uso del post-filtro ha sido ensayado en un ensayo denominado MUSHRA, cuyo resultado se ilustra en la Figura 3. "MUSHRA" hace referencia a "MULTiple Stimuli with Hidden Reference and Anchor" (múltiples estímulos con referencia oculta y base), y es una metodología para la evaluación subjetiva de la calidad de audio, usada típicamente para evaluar la calidad percibida de la salida desde los algoritmos de compresión de audio con pérdidas. Cuantos más puntos MUSHURA recibe una señal, mejor será la calidad de audio percibida. En la Figura 45 1, la primera barra (#1) representa una señal MDCT decodificada e la que no se usó ningún post-filtro en el

procedimiento de decodificación. La segunda barra (#2) representa una señal MDCT decodificada, en la que se usó el post-filtro sugerido en el procedimiento de decodificación. La tercera barra (#3) representa una señal de voz original, que no ha sido sometida a codificación y, de esta manera, recibe la máxima cantidad de puntos/puntuación. Tal como puede verse en la Figura 3, el uso del post-filtro proporciona un aumento considerable de la calidad de audio percibida.

Procedimiento ejemplar Figura 4

Ahora, se describirá una realización ejemplar del procedimiento de decodificación de una señal de audio con codificación MDCT, con referencia a la Figura 4. El procedimiento podría ser realizado en una entidad de gestión de audio, tal como por ejemplo un nodo en un sistema de teleconferencia y/o un nodo o terminal en un sistema de comunicación inalámbrica o por cable, un nodo implicado en una radiodifusión de audio, o una entidad o dispositivo usado en la producción de música.

Un vector d , que comprende los coeficientes MDCT cuantificados de un segmento de tiempo de una señal de audio, es obtenido en una acción 402. Se supone que el vector de coeficientes es producido por un codificador MDCT, y se supone que es recibido desde otro nodo o entidad, o, es recuperado, por ejemplo, desde una memoria.

Un vector \hat{d} procesado es derivado en una acción 406, aplicando directamente un post-filtro al vector d , cuyo post-filtro está configurado para tener una función de transferencia H , que es una versión comprimida de la envolvente del vector d . Además, se deriva una forma de onda de la señal reconstruida en una acción 408 aplicando una transformada MDCT inversa al vector \hat{d} procesado

El denominador de la función de transferencia H puede estar configurado para comprender un máximo del vector d . Dicho máximo podría ser el coeficiente más grande (valor absoluto) de $|d|$ o, por ejemplo, una estimación obtenida mediante un seguimiento recursivo de máximo sobre el vector $|d|$.

La función de transferencia H puede estar configurada además para comprender un componente de énfasis, configurado para controlar la agresividad o la cantidad de énfasis del post-filtro, sobre el espectro MDCT. Este componente se denota "a" en la Figura 1 y la Ecuación 1. El componente "a" podría ser, por ejemplo un vector dependiente de la frecuencia, o una constante.

La energía de la salida del post-filtro, es decir, el vector \hat{d} procesado, puede ser normalizada a la energía de la entrada del post-filtro, es decir, a la energía del vector d . Además, podría determinarse el contenido del segmento de señal de audio, y el post-filtro podría ser aplicado según dicho contenido. Por ejemplo, el vector \hat{d} procesado podría ser derivado, por ejemplo, sólo cuando se determina que el segmento de tiempo de la señal de audio comprende voz. Además, la función de transferencia H del post-filtro podría ser limitada o suprimida cuando se determina que el segmento de tiempo de la señal de audio consiste principalmente, por ejemplo, voz sorda, ruido de fondo o música. Estas acciones condicionales se ilustran como las acciones 404 y 410 en la Figura 4. El contenido del segmento de señal de audio puede ser determinado en base al vector d , o podría ser determinado en el codificador, en base a la forma de onda de la señal de audio y, a continuación, la información relacionada con el contenido podría ser señalizada, de una manera adecuada, desde el codificador al decodificador.

Disposiciones ejemplares, Figuras 5 y 6

A continuación, se describirá un decodificador 501 ejemplar, adaptado para permitir la realización del procedimiento descrito anteriormente relacionado con la decodificación de una señal, con referencia a la Figura 5.

El decodificador 501 comprende una unidad 502 de obtención, que está adaptada para obtener un vector d , que comprende coeficientes de dominio MDCT cuantificados de un segmento de tiempo de una señal de audio. El vector d podría ser recibido, por ejemplo, desde otro nodo, o podría ser recuperado, por ejemplo, de una memoria.

El decodificador comprende además una unidad 504 de filtro, que está adaptada para derivar un vector \hat{d} procesado, aplicando un post-filtro directamente al vector d obtenido. El post-filtro debería estar configurado para tener una función de transferencia H , que es una versión comprimida de la envolvente del vector d obtenido. Además, el decodificador comprende una unidad 506 configurada para derivar una forma de onda de la señal, es decir, una estimación o reconstrucción de la forma de onda de la señal comprendida en el segmento de tiempo de la señal de audio, aplicando una transformada MDCT inversa al vector \hat{d} procesado.

La disposición 500 es adecuada para su uso en un decodificador, y podría ser implementada, por ejemplo por uno o más de entre: un procesador o un microprocesador y un software adecuado, un dispositivo lógico programable (Programmable Logic Device, PLD) u otro componente o componentes electrónicos.

El decodificador puede comprender además otras unidades 508 funcionales regulares, tales como una o más unidades de almacenamiento.

La Figura 6 ilustra un decodificador 601 similar al 501, ilustrado en la Figura 5. El decodificador 601 se ilustra como situado o comprendido en una entidad 602 de gestión de audio en un sistema de comunicación. La entidad de gestión de audio podría ser, por ejemplo un nodo o terminal en un sistema de comunicación inalámbrica o por cable, un nodo o terminal en un sistema de teleconferencia y/o un nodo implicado en la radiodifusión de audio. La entidad 602 de gestión de audio y el decodificador 601 se ilustran adicionalmente de manera que se comunican con otras entidades a través de una unidad 603 de comunicación, que puede considerarse que comprende medios convencionales para la comunicación inalámbrica y/o por cable. La disposición 600 y las unidades 604-610 corresponden a la disposición 500 y las unidades 502-508 en la Figura 5. La entidad 602 de gestión de audio podría comprender además unidades 614 funcionales regulares adicionales y una o más unidades 612 de almacenamiento.

Disposición ejemplar, Figura 7

La Figura 7 ilustra una implementación de un decodificador o disposición 700 adecuado para su uso en una entidad de gestión de audio, donde un programa 710 de ordenador está contenido en un producto 708 de programa de ordenador, conectado a un procesador 706. El producto 708 de programa de ordenador comprende unos medios legibles por ordenador en los que se almacena el programa 710 de ordenador. El programa 710 de ordenador puede estar configurado como un código de programa de ordenador estructurado en módulos de programa de ordenador. Por lo tanto, en la realización ejemplar descrita, los medios de código en el programa 710 de ordenador comprenden un módulo 710a de obtención para obtener un vector d que comprende coeficientes de dominio MDCT cuantificados de un segmento de tiempo de una señal de audio. El programa de ordenador comprende además un módulo 710b de filtro para derivar un vector \hat{d} procesado. El programa 710 de ordenador comprende además un módulo 710c de conversión para derivar una estimación del segmento de tiempo de la señal de audio. El programa de ordenador puede comprender módulos adicionales, por ejemplo, 710a-d para proporcionar otra funcionalidad de decodificador.

Los módulos 710a-d podrían realizar esencialmente las acciones del flujo ilustrado en la Figura 4, para emular el decodificador ilustrado en la Figura 5. En otras palabras, cuando los diferentes módulos 710a-d se ejecutan en la unidad 706 de procesamiento, corresponden a la funcionalidad respectiva de las unidades 502-508 de la Figura 5. Por ejemplo, el producto de programa de ordenador puede ser una memoria flash, una memoria RAM (Random-Access Memory, memoria de acceso aleatorio), una memoria ROM (Read-Only Memory, memoria de solo lectura) o una memoria EEPROM (Electrically Erasable Programmable ROM, ROM programable y borrable eléctricamente) y, en realizaciones alternativas, los módulos 710a-d de programa de ordenador podrían ser distribuidos en diferentes productos de programa de ordenador en la forma de memorias dentro del decodificador 601 y/o la entidad 602 de gestión de audio. Las unidades 702 y 704 conectadas al procesador representan unidades de comunicación, por ejemplo, entrada y salida. La unidad 702 y la unidad 704 pueden disponerse como una entidad integrada.

Aunque los medios de código en la realización descrita anteriormente en conjunción con la Figura 7 se implementan como módulos de programa de ordenador que, cuando se ejecutan en la unidad de procesamiento, hacen que el decodificador y/o la entidad de gestión de audio realicen las acciones descritas anteriormente en conjunción con las figuras indicadas anteriormente, al menos uno de los medios de código puede ser implementado, en realizaciones alternativas, al menos en parte, como circuitos de hardware.

Cabe señalar que la elección de las unidades o módulos interactivos, así como la denominación de las unidades sólo tienen propósitos ejemplares, y los nodos de red adecuados para ejecutar cualquiera de los procedimientos descritos anteriormente pueden estar configurados en una pluralidad de maneras alternativas con el fin de poder ejecutar las acciones de procesamiento sugeridas.

Cabe señalar también que las unidades o los módulos descritos en esta descripción deben ser considerados como entidades lógicas y no necesariamente como entidades físicas separadas.

Abreviaturas

ACELP - Algebraic Code Excited Linear Prediction, Predicción lineal excitada por código algebraico

MDCT - Modified Discrete Cosine Transform, Transformada discreta de coseno modificada

DFT - Discrete Fourier Transform, Transformada discreta de Fourier

MUSHRA - Multiple Stimuli with Hidden Reference and Anchor, Múltiples estímulos con referencia oculta y base

Referencias

[1] J.-H. Chen y A. Gersho, "Adaptive postfiltering for quality enhancement of coded speech" IEEE Trans. Speech, Audio Processing, vol. 3, pp. 59-71, 1995

[2] ITU-T Rec. G.719, "Low-complexity full-band audio coding for high-quality conversational applications," 2008

5 [3] Documento de patente US 2007/219785 A1 (GAO YANG).

REIVINDICACIONES

1. Procedimiento en un decodificador, en el que el procedimiento comprende:

– obtener(402) un vector $d(k)$, que comprende coeficientes de dominio MDCT cuantificados de un segmento de tiempo de una señal de audio,

5 – derivar (404) un vector $\hat{d}(k)$ procesado aplicando de un post-filtro directamente al vector $d(k)$, cuyo post-filtro está configurado para tener una función de transferencia $H(k)$,

$$H(k) = \left(\frac{\text{abs}[d(k)]}{\text{max}[\text{abs}(d)]} \right)^{a(k)}$$

10

que es una versión comprimida de la envolvente del vector $d(k)$, en la que k varía de 1 al número de coeficientes de dominio MDCT del segmento de tiempo; y $a(k)$ es un componente de énfasis, configurado para controlar la agresividad del post-filtro sobre el espectro MDCT; y

15 – derivar (406) una forma de onda de la señal aplicando una transformada MDCT inversa al vector $\hat{d}(k)$ procesado.

2. Procedimiento según la reivindicación 1, en el que el valor máximo en el denominador de la función de transferencia $H(k)$ es el coeficiente de $|d|$ que tiene la magnitud más grande.

20 3. Procedimiento según la reivindicación 1, en el que el valor máximo en el denominador de la función de transferencia $H(k)$ es una estimación del máximo del vector $|d|$, obtenido mediante seguimiento recursivo de máximo sobre el vector $|d|$.

4. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que el componente $a(k)$ de énfasis depende de la frecuencia.

5. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que la energía del vector $\hat{d}(k)$ procesado es normalizado a la energía del vector d .

25 6. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que el vector $\hat{d}(k)$ procesado es derivado sólo cuando se determina que el segmento de tiempo de la señal de audio comprende voz.

7. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que la función de transferencia $H(k)$ es limitada o suprimida cuando se determina que el segmento de tiempo de la señal de audio consiste principalmente en uno o más de entre:

- 30
- voz sorda,
 - ruido de fondo,
 - música.

8. Decodificador que comprende:

35 – una unidad (502) de obtención, adaptada para obtener un vector $d(k)$, que comprende coeficientes de dominio MDCT cuantificados de un segmento de tiempo de una señal de audio,

– una unidad (504) de filtro, adaptada para derivar un vector $\hat{d}(k)$ procesado aplicando un post-filtro directamente al vector $d(k)$ obtenido, cuyo post-filtro está configurado para tener una función de transferencia $H(k)$

$$H(k) = \left(\frac{\text{abs}[d(k)]}{\text{max}[\text{abs}(d)]} \right)^{a(k)}$$

40

que es una versión comprimida de la envolvente del vector $d(k)$ obtenido, donde k está comprendido entre 1 y el número de coeficientes de dominio MDCT del segmento de tiempo; y $a(k)$ es un componente de énfasis, configurado para controlar la agresividad del post-filtro sobre el espectro MDCT; y

- 5 – una unidad (506) de conversión configurada para derivar una forma de onda de la señal aplicando una transformada MDCT inversa al vector $\hat{d}(k)$ procesado.
9. Decodificador según la reivindicación 8, en el que la función de transferencia H está configurada para comprender el coeficiente de $|d|$ que tiene la magnitud más grande.
10. Decodificador según la reivindicación 8 o 9, en el que la función de transferencia $H(k)$ está configurada para comprender una estimación de un máximo del vector $|d|$ en el denominador, cuya estimación se obtiene mediante un seguimiento recursivo de máximo sobre el vector $|d|$.
11. Decodificador según cualquiera de las reivindicaciones 8-10, en el que el componente $a(k)$ de énfasis depende de la frecuencia.
12. Decodificador según cualquiera de las reivindicaciones 8-11, adaptado además para normalizar la energía del vector $\hat{d}(k)$ procesado a la energía del vector $d(k)$.
- 15 13. Decodificador según cualquiera de las reivindicaciones 8-12, adaptado además para derivar $\hat{d}(k)$ sólo cuando se determina que el segmento de tiempo de la señal de audio comprende voz.
14. Decodificador según cualquiera de las reivindicaciones 8 a 13, adaptado además para limitar o suprimir la función de transferencia $H(k)$ cuando se determina que el segmento de tiempo de la señal de audio consiste principalmente en uno o más de entre:
- 20 – voz sorda,
 - ruido de fondo,
 - música
15. Entidad (601) de gestión de audio que comprende un decodificador según cualquiera de las reivindicaciones 8-14.

25

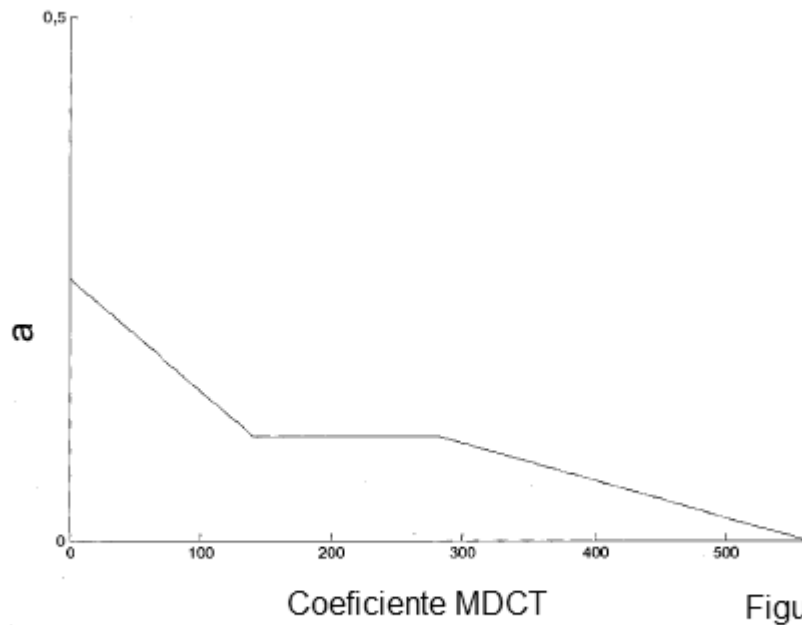


Figura 1

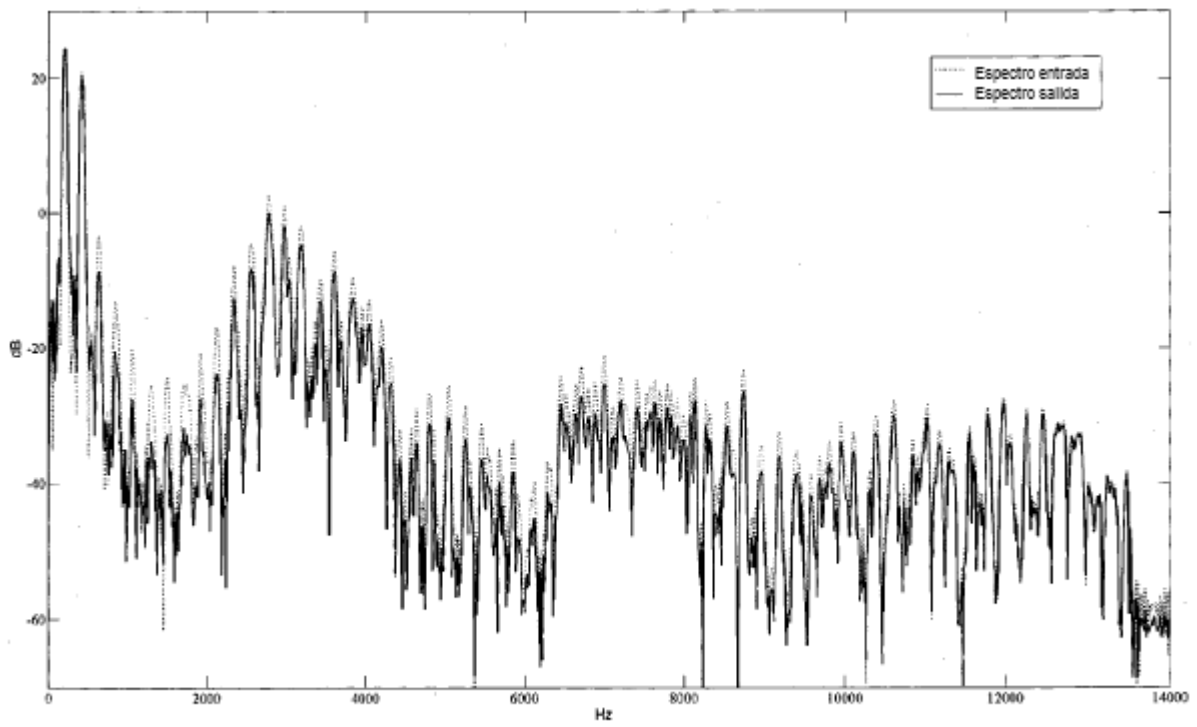


Figura 2

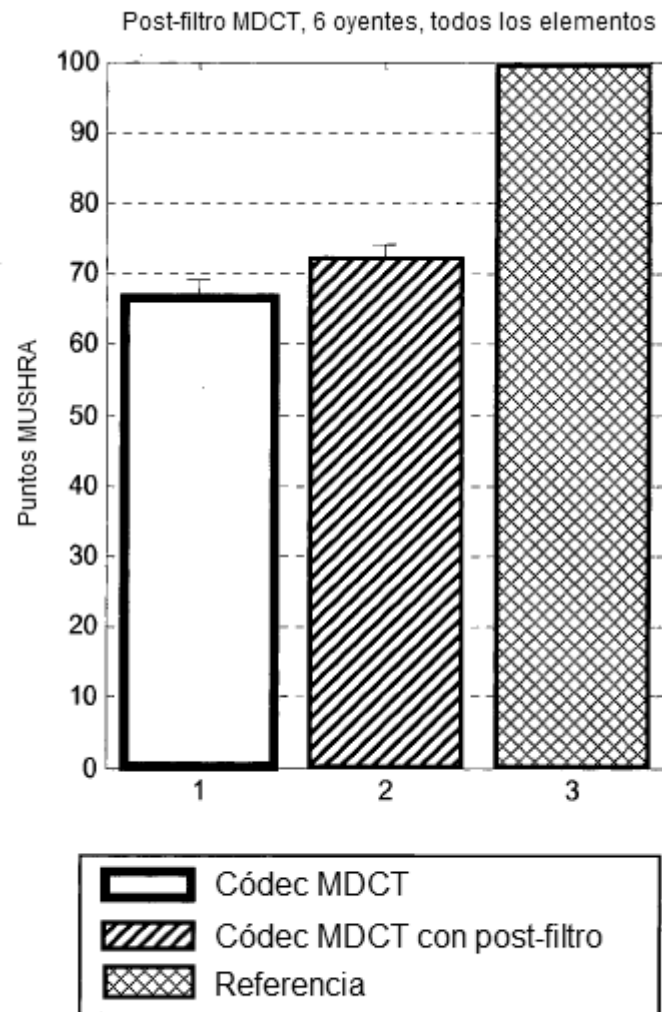


Figura 3

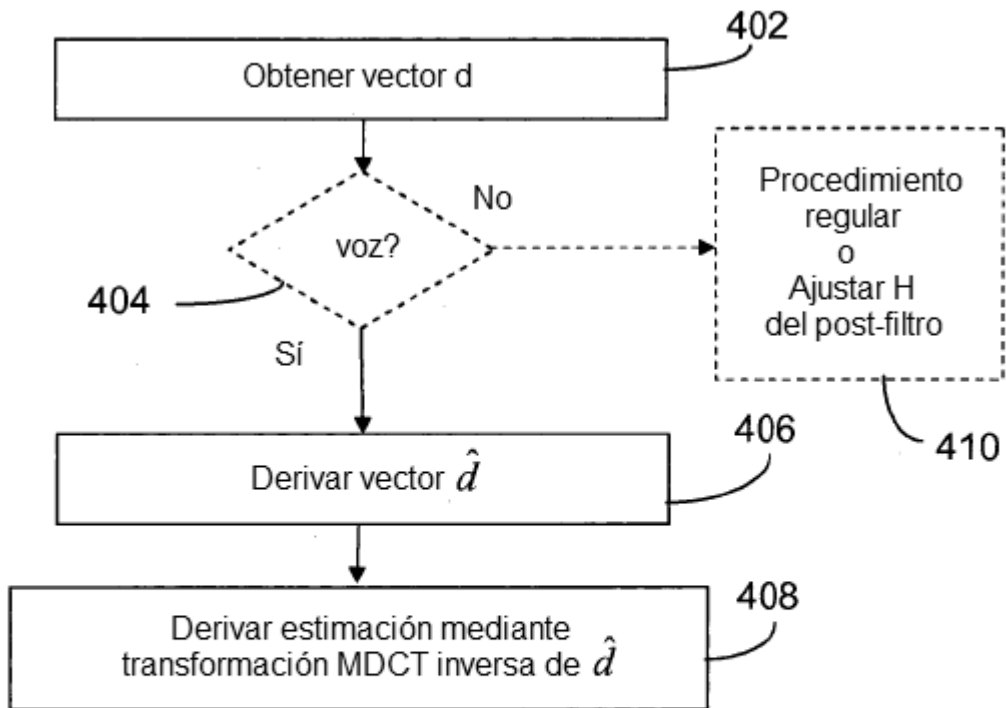


Figura 4

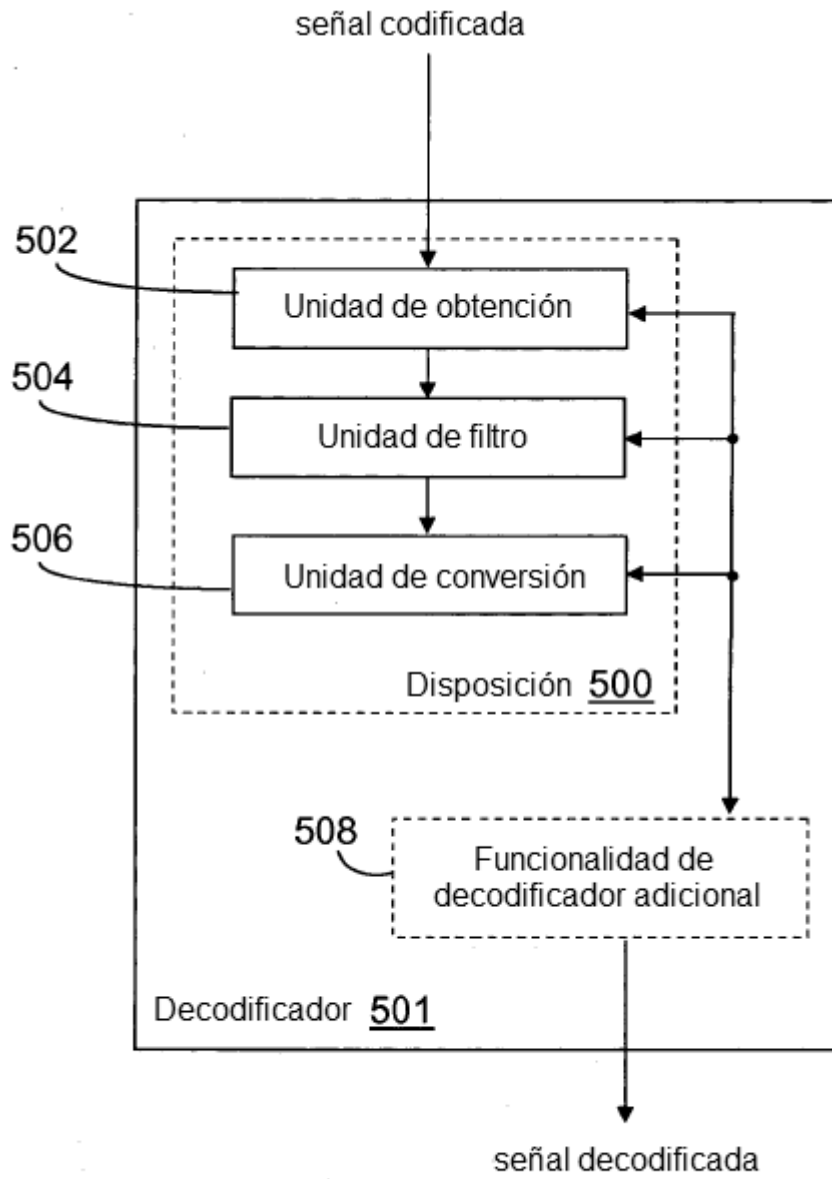


Figura 5

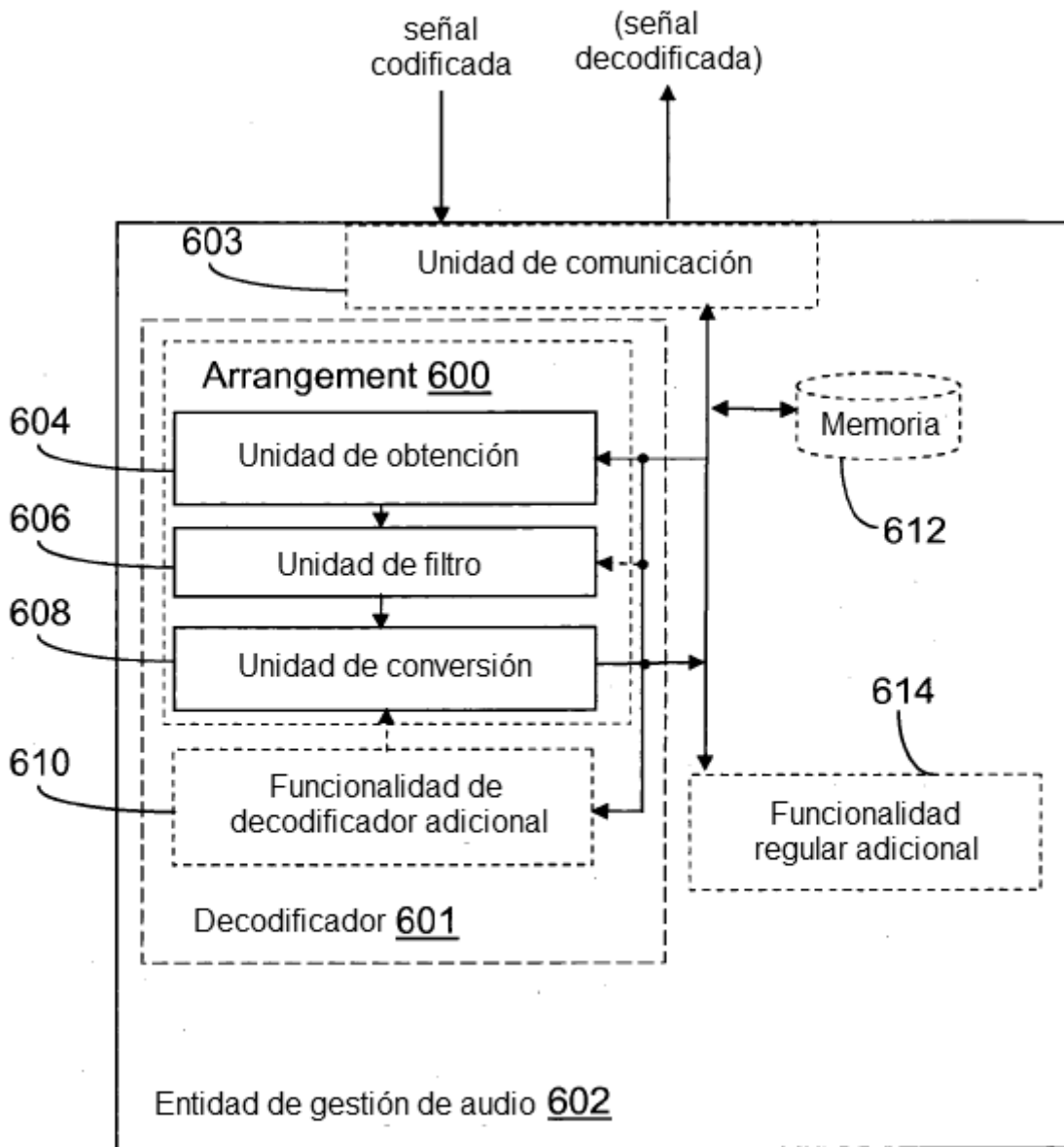


Figura 6

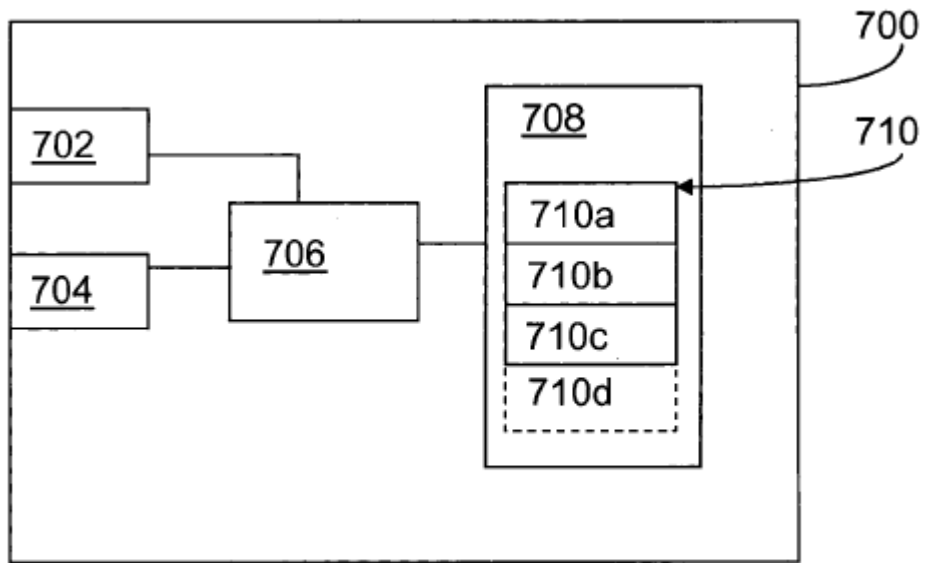


Figura 7