

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 735 279**

51 Int. Cl.:

**G10L 19/022** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.06.2012 PCT/FR2012/051463**

87 Fecha y número de publicación internacional: **03.01.2013 WO13001226**

96 Fecha de presentación y número de la solicitud europea: **26.06.2012 E 12734996 (7)**

97 Fecha y número de publicación de la concesión europea: **15.05.2019 EP 2727107**

54 Título: **Ventanas de ponderación en codificación/decodificación por transformada con recubrimiento, optimizadas en retardo**

30 Prioridad:

**28.06.2011 FR 1155769**

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

**17.12.2019**

73 Titular/es:

**ORANGE (100.0%)  
78, rue Olivier de Serres  
75015 Paris, FR**

72 Inventor/es:

**FAURE, JULIEN y  
PHILIPPE, PIERRICK**

74 Agente/Representante:

**ISERN JARA, Jorge**

**ES 2 735 279 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

Ventanas de ponderación en codificación/decodificación por transformada con recubrimiento, optimizadas en retardo

5 La invención se sitúa en el campo de la codificación/decodificación de señales de audio digitales, más específicamente en el campo denominado de la "codificación/decodificación de audio por transformada con recubrimiento".

10 La "codificación por transformada" consiste en codificar unas señales temporales en el campo transformado (de frecuencia). Esta transformación permite, en concreto, utilizar las características de frecuencia de las señales de audio (musicales, de voz u otras), con el fin de optimizar y mejorar los rendimientos de codificación. Por ejemplo, se utiliza el hecho de que un sonido armónico está representado en el campo de frecuencia por un número finito y reducido de líneas espectrales que, de este modo, pueden codificarse de manera concisa. Por ejemplo, también se utilizan, ventajosamente, unos efectos de enmascaramiento de frecuencia para dar forma al ruido de codificación, de manera que sea lo menos audible posible.

15 Una técnica habitual de codificación por transformada se puede resumir como sigue.

20 El flujo de audio digital (muestreado a una frecuencia de muestreo  $F_s$  dada), que hay que codificar, se secciona en tramas (o "bloques", más generalmente) de número de muestras  $2M$  finitas. Cada trama se recubre convencionalmente en un 50 % con la trama anterior. Una ventana de ponderación  $h_a$  (llamada "ventana de análisis") se aplica a cada trama.

25 A continuación, se aplica una transformación a la señal. En el caso de una transformación denominada "MDCT" (para "Modified Discrete Cosine Transform", "Transformada de Coseno Discreta Modificada" en español) y en una realización particular, la trama ponderada se "repliega" según una transformación  $2M$  muestras hacia  $M$  muestras. A continuación, se aplica una transformada DCT de tipo IV a la trama replegada, con el fin de obtener una trama de tamaño  $M$  en el campo transformado.

30 La trama en el campo transformado se cuantifica, entonces, utilizando un cuantificador adaptado. La cuantificación permite reducir el tamaño de los datos, pero introduce un ruido (audible o no) en la trama original. Cuanto más elevada es la velocidad del codificador, más se reduce este ruido y más cercana está la trama cuantificada de la trama original.

35 En la decodificación, una transformación MDCT inversa se aplica, entonces, a la trama cuantificada. La trama cuantificada de tamaño  $M$  se convierte en trama de tamaño  $M$  en el campo temporal utilizando una DCT de tipo IV inversa. Entonces, se aplica una segunda transformación de "despliegue" de  $M$  hacia  $2M$  a la trama temporal de tamaño  $M$ .

Unas ventanas de ponderación  $h_s$  denominadas de "síntesis" se aplican, a continuación, a las tramas de tamaños  $2M$ .

40 El flujo de audio decodificado se sintetiza, entonces, sumando las partes en recubrimiento.

Para una ventana de síntesis y un recubrimiento dado, se determina una ventana de análisis que permite obtener una reconstrucción perfecta de la señal que hay que codificar (en ausencia de cuantificación).

45 Una ventana utilizada convencionalmente en codificación por transformada es una ventana de tipo sinusoidal idéntica a la vez en el análisis y en la síntesis. En esta configuración, el plazo algorítmico mínimo introducido por el sistema de codificación es de  $2M/F_s$  segundos.

50 Para reducir este plazo, se pueden imponer unos ceros al inicio de ventana de síntesis y al final de ventana de análisis. Como el resultado de una multiplicación de la señal por "0" se conoce de antemano, se puede desplazar el ritmo de tramas con respecto a la posición de las ventanas. Estas ventanas, simétricas, están compuestas, por ejemplo:

- por un cierto número de ceros  $Mz$  que se extienden sobre un intervalo correspondiente a la mitad del retardo algorítmico que se quiere salvaguardar,
- 55 - por una subida en seno de longitud  $M-2Mz$ ,
- por una sección de  $2Mz$  valores a 1,
- siendo, por último, la segunda mitad de la ventana lo simétrico de la primera, como se ilustra en la figura 1 adjunta.

60 Estas ventanas tienen un plazo algorítmico de  $(2M-2Mz)/F_s$  segundos y, de este modo, permiten reducir el retardo en  $2Mz/F_s$  segundos.

65 Sin embargo, una técnica de este tipo, si permite reducir el plazo, tiende, cuando la reducción del plazo aumenta, a parecerse a una ventana rectangular. Una forma de ventana de este tipo es poco selectiva en frecuencias y, finalmente, hacer bajar de manera drástica la calidad de audio de la señal codificada. Además, restringe fuertemente la ventana, ya que se imponen  $4Mz$  muestras durante su construcción. No se dispone de muchos grados de libertad para proponer unas ventanas eficaces para la codificación, en concreto, para ofrecer una selectividad de frecuencia importante.

En el documento WO-2009/081003 se ha propuesto utilizar unas ventanas asimétricas para mitigar este problema. Estas ventanas, en el análisis, están compuestas por 0 únicamente al final de la ventana de análisis. Con el fin de limitar el espacio de almacenamiento requerido, la ventana de síntesis se elige como el retorno temporal de la ventana de análisis. Esta técnica permite, en concreto, reducir el plazo en la codificación, así como el plazo en la decodificación. Para un número de ceros total  $M_z$  dos veces menos importante que el de las ventanas simétricas presentadas anteriormente, la ganancia de plazo es la misma. Teniendo en cuenta el número reducido de ceros, la selectividad de frecuencia de unas ventanas asimétricas de este tipo es superior a la de las ventanas simétricas. Se mejora con ello la calidad de audio de la señal decodificada.

Más particularmente, el documento WO-2009/081003 presenta una ventana de análisis  $h_a(n)$  compuesta por dos partes  $h_{a1}$  y  $h_{a2}$  a partir de una ventana inicial  $h(n)$  dada por:

$$h(n) = \text{sen} \left[ \frac{\pi}{2M - M_z} \left( n + \frac{1}{2} \right) \right]$$

para  $0 \leq n < 2M - M_z$

y  $h(n) = 0$  si no (es decir, para  $2M - M_z \leq n < 2M$ ).

y por un factor de corrección  $\Delta(n)$  que permite tener la condición de reconstrucción perfecta, dado por:

$$\Delta(n) = \sqrt{h(n)h(2M - 1 - n) + h(n + M)h(M - 1 - n)}$$

La ventana de análisis  $h_a$  está dado por:

$$h_{a1}(n + M) = h(n + M) / \Delta(n) \text{ y } h_{a2}(n) = h(n) / \Delta(n)$$

para  $0 \leq n < M$

La ventana de síntesis  $h_s(n)$  es el retorno temporal de la ventana de análisis:

$$h_s(2M - 1 - n) = h_a(n),$$

para  $0 \leq n < 2M$

Unas ventanas de este tipo son, por una misma ganancia de plazo, de mejor calidad que unas ventanas simétricas, por el hecho de su mejor selectividad de frecuencia.

Sin embargo, incluso si la técnica anterior es ventajosa y propone una mejora de la calidad con respecto a las técnicas anteriores, cuando se busca una solución con una ganancia de plazo más importante, por ejemplo, con un número de ceros  $M_z$  superior a  $M/4$  (donde  $M$  es una duración de trama), se observa, aplicando unas ventanas de este tipo, una degradación audible que puede explicarse, en concreto, por el hecho de que una porción de la ventana toma unos valores elevados, bastante superiores a 1, como se ilustra en la figura 2. Ahora bien, es preferible generalmente, en procesamiento de señal digital, utilizar unas ponderaciones con valores inferiores a 1 en valor absoluto, por el hecho de la aplicación de coma fija.

La presente invención viene a mejorar la situación.

Para ello, propone un procedimiento de codificación de una señal digital constituida por bloques sucesivos de muestras, siendo la codificación del tipo por transformada con recubrimiento y comprendiendo, en el análisis, la aplicación de una ventana de ponderación sobre dos bloques de  $M$  muestras sucesivas. En particular, la ventana de ponderación mencionada anteriormente es asimétrica e incluye cuatro porciones distintas que se extienden sucesivamente sobre los dos bloques, con:

- una primera porción, creciente sobre un primer intervalo de muestras,
- una segunda porción, constante a un valor de 1 sobre un segundo intervalo,
- una tercera porción, decreciente sobre un tercer intervalo y
- una cuarta porción, constante a un valor de 0 sobre un cuarto intervalo.

Como se verá esto más adelante, según una de las ventajas que procura la invención, la constancia de la ventana a 1 durante el segundo intervalo permite reducir la complejidad de la utilización de las ventanas de la técnica anterior, tales como las descritas en el documento WO-2009/081003, conservando al mismo tiempo sus ventajas de escaso

retardo y mejorando, además, la calidad de la reproducción de audio, como se verá esto más adelante con referencia a la figura 3.

5 Por otra parte, es ventajoso que los intervalos primero, segundo y tercero se calculen al menos en función de la duración del cuarto intervalo (número de 0 al final de ventana), que fija el retardo mencionado anteriormente, que puede llevarse ventajosamente a un retardo algorítmico como se verá esto en los ejemplos de realización a continuación. Asimismo, los frentes de subida y de descenso de las porciones respectivamente creciente y decreciente se pueden optimizar en función, en concreto, de la duración del cuarto intervalo. El cuarto intervalo (indicado como  $M_z$  a continuación) se expresa él mismo en función del número de muestras  $M$  por bloque y, por supuesto, en función del retardo máximo deseado.

15 La sucesión de las ventanas de análisis presenta unas características similares con la descrita en el documento WO-2009/081003 mencionado anteriormente. En particular, estando dos bloques de muestras de mismo tamaño  $2M$  ponderados respectivamente por una primera y una segunda ventana de análisis, el frente ascendente de la segunda ventana puede ser diferente del frente descendente de la primera ventana retornada temporalmente. Estos frentes no son, por lo tanto, simétricos *a priori*.

20 Por otra parte, en la implementación de la invención, en la codificación, varias mismas ventanas de análisis que tienen las características mencionadas anteriormente (subida, constancia a 1, descenso y constancia a 0) se pueden aplicar sucesivamente a una pluralidad de bloques sucesivos de  $2M$  muestras, que se recubren con  $M$  muestras, como se ilustra en la figura 10. En términos genéricos, el procedimiento de la invención incluye, entonces, la aplicación de una pluralidad de ventanas sucesivas, asimétricas y que incluye las cuatro porciones mencionadas anteriormente, a una sucesión de pares de bloques. De este modo, como en el documento WO-2009/081003, no se prevé en el presente documento tampoco una ventana de transición particular cuando se aplica la ventana asimétrica.

25 La ventana varía sobre el primer intervalo, indicado como  $R_1$ , como una función  $w_1$  de tipo:

$$w_1(n) = \frac{W_1(n)}{\Delta_n(n)}, \text{ con } W_1(n) = \text{sen} \left( \frac{\pi n+1}{2 R_1} \right)^{C_1} \text{ y}$$

30  $\Delta_n(n) = \sqrt{h_i(n) \cdot h_i(2M - n - 1) + h_i(n + M) \cdot h_i(M - n - 1)}$ ,  $n \in [0; M - 1]$  con  $C_1$  y  $R_1$ , unas constantes superiores a 0, siendo  $h_i$  una ventana de inicialización.

35 Por optimización, por ejemplo, se encuentra que el término  $C_1$  puede estar comprendido entre 3 y 5, para una gama de retardos del orden de 15 a 30 ms. En un ejemplo de realización particular,  $C_1 = 4,8425$ .

El primer intervalo, indicado como  $R_1$ , es, en este ejemplo de realización, de duración dada por:

$$R_1 = \left\lceil \frac{2M \times 404}{1280} + 0,5 \right\rceil, \text{ donde } M \text{ corresponde a la duración de un bloque,}$$

40 designando la indicación  $\lceil x \rceil$  el número entero más pequeño o igual a  $x$  y el más cercano de  $x$ .

En lo que se refiere a la segunda porción de la ventana, esta última varía sobre el tercer intervalo, indicado como  $R_2$ , como una función  $w_3$  de tipo:

45 
$$w_3(n) = \frac{W_3(n)}{\Delta_n(n)}, \text{ con } W_3(n) = \cos \left( \frac{\pi n}{2 R_2} \right)^{C_2} \text{ y:}$$

$\Delta_n(n) = \sqrt{h_i(n) \cdot h_i(2M - n - 1) + h_i(n + M) \cdot h_i(M - n - 1)}$ ,  $n \in [0; M - 1]$  con  $C_2$  y  $R_2$ , unas constantes superiores a 0.

50 De acuerdo con unas pruebas de optimización realizadas, el término  $C_2$  está preferentemente comprendido entre 0,85 y 1,05 y en el ejemplo de realización particular,  $C_2 = 0,9659$ .

El tercer intervalo, indicado como  $R_2$ , es, en este ejemplo de realización, de duración preferencial dada por:

55  $R_2 = \left\lceil \frac{2M \times 196}{1280} + 0,5 \right\rceil$  donde  $M$  corresponde a la duración de un bloque, designando la indicación  $\lceil x \rceil$  el número entero más pequeño o igual a  $x$  y el más cercano de  $x$ .

En este ejemplo de realización, el cuarto intervalo, indicado como  $M_z$ , se elige inicialmente tal que es de duración dada por

$$M_z = \left[ \frac{2M \times 7}{40} \right]$$

donde la indicación [x] designa el número entero mayor o igual a x y el más cercano de x, correspondiendo  $M$  a la duración de un bloque.

5 Por supuesto, se pueden aplicar otros retardos y, de este modo, se pueden prever otras duraciones del cuarto intervalo  $M_z$ . La tabla 1 dada en los ejemplos de realización detallados a continuación indica unos valores de los parámetros  $C_1$ ,  $C_2$  y unos intervalos  $R_1$ ,  $R_2$ , para diferentes valores de retardos y, por lo tanto, para diferentes duraciones de cuarto intervalo  $M_z$ .

10 De este modo, el procedimiento puede incluir ventajosamente una etapa previa de optimización de la forma de la ventana, basándose esta optimización en al menos una estimación de función de coste, para llegar a los parámetros óptimos  $C_1$ ,  $C_2$  y/o también  $R_1$ ,  $R_2$  presentados anteriormente.

15 Globalmente, es ventajoso que el número total de "1" en la ventana sea aproximadamente dos veces mayor que el número de 0, de acuerdo con las optimizaciones realizadas, en concreto, para una buena calidad de reproducción de audio.

20 De este modo, en otro ejemplo de realización particular, el primer intervalo, indicado como  $R_1$  es, por ejemplo, de duración dada por:

$R_1 = [M - M_z + 1 + 0,5]$ , donde  $M$  corresponde a la duración de un bloque y  $M_z$  corresponde a la duración del cuarto intervalo.

En este ejemplo, el tercer intervalo  $R_2$  puede ser de duración dada por:

25  $R_2 = [M - 2M_z + 1 + 0,5]$  donde  $M$  corresponde a la duración de un bloque y  $M_z$  corresponde a la duración del cuarto intervalo.

La tabla 2 dada en los ejemplos de realización detallados a continuación indica unos valores de los parámetros  $C_1$  y  $C_2$  cuando los intervalos  $R_1$  y  $R_2$  se han fijado de este modo, para diferentes valores de retardos y, por lo tanto, para diferentes duraciones de cuarto intervalo  $M_z$ .

30 De este modo, la invención propone, sobre un principio de ventana asimétrica con reconstrucción perfecta con un número elegido de ceros del tipo presentado en el documento WO-2009/081003, la utilización de ventanas optimizadas, de análisis y de síntesis, que permiten tener una buena reproducción sonora, asegurando al mismo tiempo una implementación eficaz. En la figura 3, se ha ilustrado una comparación de los rendimientos en términos de calidad de audio entre una ventana según la invención (referencia INV), una ventana según el documento WO-2009/081003 (referencia AA) y una ventana sinusoidal convencional (SEN) para un plazo a 26 ms. Los otros puntos, sin referencia, se refieren a otras técnicas del estado de la técnica. Aparece que la invención permite conservar una calidad de audio equivalente a la lograda con unas ventanas convencionales (SEN), ofreciendo al mismo tiempo una ganancia de plazo comparable a la lograda por la implementación descrita en el documento WO-2009/081003.

35 40 Entonces, se retiene en la descripción a continuación, como criterio de apreciación, la calidad perceptiva, resultando un error de reconstrucción de la transformada insignificante, perceptiblemente, con respecto a un umbral de percepción predeterminado.

45 La presente invención tiene como propósito un procedimiento de decodificación de una señal digital, codificada por la implementación del procedimiento de anteriormente, siendo la decodificación del tipo por transformada con recubrimiento y comprendiendo, en la síntesis, la aplicación de una ventana de ponderación sobre dos bloques de muestras codificadas sucesivas. En particular, la ventana de ponderación, en la síntesis, es idéntica a la ventana de análisis utilizada en la codificación, retornada temporalmente, por ejemplo, para asegurar una propiedad de reconstrucción perfecta, como se verá esto más adelante.

50 La presente invención también tiene como propósito un programa informático destinado a ser almacenado en memoria de un dispositivo de codificación o de decodificación y que incluye, en particular, unas instrucciones para la implementación del procedimiento de codificación de anteriormente o del procedimiento de decodificación mencionado anteriormente, cuando las instrucciones son ejecutadas por un procesador del dispositivo. En particular, puede tratarse de un programa informático que incluye una primera parte específica de la codificación y una segunda parte específica de la decodificación. La figura 7 que se comenta más adelante es un ejemplo de diagrama de flujo del algoritmo general de un programa de este tipo y la figura 8 ilustra unas etapas previas que puede incluir, en una realización particular, el algoritmo de este programa, en particular, si este programa incluye unas instrucciones para la construcción previa de las ventanas de análisis y de síntesis que utilizan los procedimientos de codificación/decodificación en el sentido de la invención.

55 60

Como se representa esquemáticamente en la figura 9, la presente invención también tiene como propósito un dispositivo de codificación COD de una señal, que incluye unos medios de almacenamiento MEM y/o de cálculo  $\mu P$  de datos de una ventana de ponderación, de análisis, para la implementación del procedimiento de codificación según

la invención. Asimismo, la presente invención también tiene como propósito un dispositivo de decodificación DECOD de una señal, que incluye unos medios de almacenamiento MEM' y/o de cálculo  $\mu P$  de datos de una ventana de ponderación, de síntesis, para la implementación del procedimiento de decodificación según la invención. En particular, se puede, para la codificación como para la decodificación, partir de valores preinicializados de ventanas de análisis y/o síntesis (almacenados en memorias MEM y/o MEM') y, posiblemente, optimizar estos valores sobre la marcha (por la implementación de los medios de cálculos  $\mu P$ ).

Otras ventajas y características de la invención aparecerán en la lectura de la descripción detallada, dada a continuación a título de ejemplos de realización no limitativos y de los dibujos adjuntos en los que:

- la figura 1 ilustra la apariencia de una ventana simétrica convencional con escaso plazo, que incluye unos ceros al inicio y al final de ventana,
- la figura 2 ilustra la apariencia de una ventana asimétrica tal como se describe en el documento WO-2009/081003, con una ganancia de plazo de 15 ms para una trama 2M de 40 ms,
- la figura 3 ilustra una comparación de la calidad de audio entre una ventana en el sentido de la invención INV, una ventana sinusoidal SEN según la figura 1 y una ventana AA según la figura 2, para un plazo de 26 ms,
- la figura 4 ilustra un primer ejemplo de ventana de inicialización  $h_i$  para la construcción de una ventana de análisis en el sentido de la invención,
- la figura 5 ilustra un segundo ejemplo de ventana de inicialización  $h_i$ , para la construcción de una ventana de análisis en el sentido de la invención,
- la figura 6 ilustra un ejemplo de ventana de análisis en el sentido de la invención, obtenida a partir de una ventana de inicialización del tipo representada en la figura 5,
- la figura 7 presenta esquemáticamente las etapas de un ejemplo de procedimiento de codificación/decodificación por transformada, para ilustrar el contexto de la presente invención,
- la figura 8 presenta esquemáticamente las etapas de un ejemplo de construcción de las ventanas de análisis y de síntesis para una implementación del procedimiento en el sentido de la invención,
- la figura 9 presenta esquemáticamente unos dispositivos de codificación y de decodificación para la implementación de la invención,
- la figura 10 ilustra una sucesión de ventanas de análisis en el sentido de la invención, aplicadas con recubrimiento.

En primer lugar, se hace referencia a la figura 7 para describir, en calidad de marco general de la implementación de la invención, un procedimiento de codificación/decodificación de una señal digital (dada en la etapa 70), por transformada con recubrimiento. El procedimiento comprende una etapa 71 de corte de la señal de audio digital en tramas de longitud  $M$ . A continuación, cada trama se agrupa con la anterior, lo que corresponde a un recubrimiento de un 50 %. En la siguiente etapa 72, una ventana de análisis  $h_a$  se aplica a este conjunto de dos tramas. En la figura 6, se ha ilustrado un ejemplo de ventana de análisis  $h_a$  en el sentido de la invención (amplitud de 0 a 1 en función del número de muestras "N.º Mues"). Se observará que la ventana de análisis  $h_a$  se termina con  $Mz$  valores a 0 y es de longitud total  $2M$ .

La siguiente etapa 73 consiste en aplicar una transformada, por ejemplo, de tipo MDCT, al bloque de dos tramas ponderado de este modo por la ventana  $h_a$ , luego en cuantificar (etapa 74) los valores en el campo transformado.

En la decodificación, una transformada inversa permite, en la etapa 75, reconvertir los valores en el campo temporal. Si la transformada en la codificación es, por ejemplo, de tipo MDCT, se puede aplicar una transformada inversa iMDCT a la decodificación.

En la etapa 76, para asegurar una condición de reconstrucción perfecta, una ventana de síntesis  $h_s$  se aplica a las dos tramas decodificadas. Como se describe en el documento WO-2009/081003, esta condición impone un factor de corrección  $1/\Delta_n$  (descrito con más detalle a continuación) aplicado a una ventana de inicialización  $h_i$ , para obtener la ventana de análisis  $h_a$ . La ventana de síntesis  $h_s$  se presenta como el retorno temporal de la ventana de análisis  $h_a$ , con:

$$h_s(n) = h_a(2M - n - 1)$$

En la etapa 77, los bloques que se recubren se suman para suministrar, finalmente, la señal decodificada, asegurando una reconstrucción perfecta.

Se hace referencia de nuevo a la figura 6, para describir con más detalle la apariencia de la ventana de análisis  $h_a$  utilizada en un procedimiento de este tipo, en particular, en la codificación. La ventana de análisis está constituida por cuatro segmentos distintos  $w_1, w_2, w_3, w_4$ . Sucede lo mismo para la ventana de síntesis  $h_s$  que se presenta como el retorno temporal de la ventana de análisis.

Sobre un primer intervalo indicado, a continuación, como  $R_1$ , la ventana de análisis varía como una función  $w_1$  de tipo:

$$w_1(n) = \frac{W_1(n)}{\Delta_n(n)} \text{ con, } W_1(n) = \text{sen} \left( \frac{\pi n+1}{2 R_1} \right)^{C_1} \text{ y:}$$

$$\Delta_n(n) = \sqrt{h_i(n) \cdot h_i(2M - n - 1) + h_i(n + M) \cdot h_i(M - n - 1)}, n \in [0; M - 1]$$

A continuación, la ventana de análisis es constante y vale  $w_2=1$  sobre un segundo intervalo.

Sobre un tercer intervalo indicado, a continuación, como  $R_2$ , la ventana varía como una función  $w_3$  de tipo:

$$w_3(n) = \frac{W_3(n)}{\Delta_n(n)} \text{ con, } W_3(n) = \text{cos} \left( \frac{\pi n}{2 R_2} \right)^{C_2} \text{ y:}$$

$$\Delta_n(n) = \sqrt{h_i(n) \cdot h_i(2M - n - 1) + h_i(n + M) \cdot h_i(M - n - 1)} \text{ con:}$$

$$n \in [0; M - 1]$$

Por último, sobre un cuarto intervalo  $M_z$ , la ventana de análisis es constante y vale  $w_4=0$ .

A continuación, se muestra que, con respecto a la técnica anterior WO-2009/081003, la forma de unas ventanas de este tipo permite asegurar una calidad de audio óptima para una ganancia de plazo importante (que puede llegar, por ejemplo, hasta un 40 %). La utilización de unas ventanas de este tipo (de variaciones generales: creciente, constante a 1, decreciente y constante a 0) no se limita a un tipo de transformada MDCT. También se pueden utilizar, por ejemplo, en el caso en que la codificación no se refiera únicamente más que a una parte solamente de la señal que hay que codificar (por ejemplo, solamente una parte de la banda de audio, por ejemplo, para procesar un error de codificación con respecto a otra técnica de codificación u otros).

En este momento, se describe con referencia a la figura 8 cómo se obtiene la ventana  $h_a$  utilizada en un procedimiento tal como se ha descrito anteriormente con referencia a la figura 7.

Como continuación a una etapa de inicialización 80, se definen unos parámetros de inicialización  $R_1$ ,  $R_2$ ,  $C_1$  y  $C_2$  en la etapa 81. En la siguiente etapa 82, se calcula una ventana de inicialización  $h_i$  a partir de estos parámetros de inicialización. Como se ilustra en las figuras 4 y 5, esta ventana de inicialización  $h_i$  está constituida por cuatro segmentos  $W_1$ ,  $W_2$ ,  $W_3$ ,  $W_4$ .

El primer segmento  $W_1$  de longitud  $R_1$  se define por la ecuación:

$$W_1(n_1) = \text{sen} \left( \frac{\pi n_1+1}{2 R_1} \right)^{C_1}, n_1 \in [0; R_1 - 1]$$

donde  $C_1$  es una constante  $>0$ .

El segundo segmento  $W_2$  de longitud  $L=2M-R_1-R_2-M_z$  está compuesto por valores a 1.

El tercer segmento  $W_3$  de longitud  $R_2$  se define por la ecuación:

$$W_3(n_2) = \text{cos} \left( \frac{\pi n_2}{2 R_2} \right)^{C_2}, n_2 \in [0; R_2 - 1]$$

donde  $C_2$  es una constante  $>0$ .

El cuarto segmento  $W_4$  de longitud  $M_z$  está compuesto por valores todos a 0.

La ventana de inicialización  $h_i(n)$  está dada, por lo tanto, por:

$$h_i(n) = \begin{cases} W_1(n) & \text{para } n \in [0; R_1 - 1] \\ 1 & \text{para } n \in [R_1; 2M - R_2 - M_z - 1] \\ W_3(n - 2M + R_2 + M_z) & \text{para } n \in [2M - R_2 - M_z; 2M - M_z - 1] \\ 0 & \text{para } n \in [2M - M_z; 2M - 1] \end{cases}$$

En la etapa 83, se aplica a la ventana de inicialización  $h_i$ , un factor  $1/\Delta_n$  que permite asegurar una condición de reconstrucción perfecta y esto, suponiendo que la ventana de síntesis es el retorno temporal de la ventana de análisis y que la transformada en la decodificación será dual (por ejemplo, una transformada iMDCT, si una transformada MDCT se ha aplicado en la codificación).

5 El término  $\Delta_n$  está dado por:

$$\Delta_n(n) = \Delta_n(n + M) = \sqrt{h_i(n) \cdot h_i(2M - n - 1) + h_i(n + M) \cdot h_i(M - n - 1)}$$

$n \in [0; M - 1]$

10 La ventana de análisis  $h_a$  con reconstrucción perfecta está, entonces, dada por

$$h_a(n) = \frac{h_i(n)}{\Delta_n(n)}, \quad n \in [0; 2M - 1]$$

15 Se observará que los valores estrictamente a "1" no necesitan ninguna multiplicación de ponderación. De este modo, la ganancia esperada en complejidad es importante con respecto al estado de la técnica, como lo muestra la tabla a continuación, según otra ventaja que procura la invención.

Ventana (512 puntos)	Número de multiplicaciones	Ganancia
Ventana SEN de la figura 1	512	0 %
Ventana AA de la figura 2	416	19 %
Ventana INV de la figura 6	224	56 %

20 De este modo, para una ganancia de plazo dado (un número  $M_z$  de ceros dado), existe una combinación óptima de los cuatro parámetros  $R_1, R_2, C_1, C_2$  que definen la forma precisa de las ventanas de análisis y de síntesis que permite tener la mejor calidad de audio posible. Estos parámetros se pueden obtener por optimización con iteraciones sucesivas, como se describe más adelante con referencia de nuevo a la figura 8.

25 Por ejemplo, una ganancia de plazo de un 35 % (correspondiente a 14 ms para unas tramas de 40 ms en el ejemplo descrito en el presente documento) simplemente corresponde a un plazo algorítmico del codificador. Este plazo, de 26 ms en el ejemplo descrito, impone un número de ceros  $M_z$  dado por:

$$30 \quad M_z = \left\lceil \frac{2M \times 7}{40} \right\rceil$$

donde  $\lceil x \rceil$  indica el número entero mayor o igual a  $x$  el más cercano de  $x$ . Entonces, se encuentra  $M_z = 14$  ms, para  $M = 40$  ms.

35 Con esta elección, los valores de los parámetros de las ventanas de análisis y de síntesis pueden ser, por ejemplo, dados por:

$$R_1 = \left\lceil \frac{2M \times 404}{1280} + 0,5 \right\rceil$$

$$40 \quad R_2 = \left\lceil \frac{2M \times 196}{1280} + 0,5 \right\rceil$$

$$C_1 = 4.8425$$

$$45 \quad C_2 = 0,9659$$

donde  $\lfloor x \rfloor$  indica el número entero menor o igual a  $x$  el más cercano de  $x$ .

Una realización de este tipo corresponde a la ventana de inicialización representada en la figura 4.

Otro ejemplo de enfoque permite reducir el número de parámetros de la ventana de inicialización que hay que buscar. Este enfoque consiste en considerar que el número de "1" en la ventana  $h_i$  debe ser el doble del número de "0" y que el eje de simetría de la ventana  $h_i$  corresponde a la mitad del segmento  $W_2$ , como se ilustra en la figura 5. En estas condiciones, los valores de los intervalos  $R_2$  y  $R_1$  dependen únicamente de la extensión del intervalo  $M_z$  (es decir, del número elegido de "0" al final de la ventana de inicialización).

En estas condiciones:

$$M_z = \left[ \frac{2M \times \frac{G_{\text{plazo}}}{2}}{FL} \right]$$

$$R_2 = [M - 2M_z + 1 + 0,5]$$

$$R_1 = M - M_z + 1 + 0,5$$

donde  $G_{\text{plazo}}$  es la ganancia deseada sobre el plazo, en ms, FL es la longitud de la trama 2M en ms (estando la frecuencia de muestreo dada por  $F_s = 2M/FL \times 1000$ , si se cuenta en milisegundos o "ms").

A continuación, se busca en la etapa 84 de la figura 8 determinar las constantes  $C_1$  y  $C_2$  (y eventualmente  $R_1$  y  $R_2$ ), por ejemplo, por la implementación de una optimización iterativa (como, por ejemplo, la técnica denominada "de los gradientes").

Partiendo, por ejemplo, en la etapa 81 con una elección de parámetros iniciales para los valores de  $R_1$ ;  $R_2$ ;  $C_1$  y  $C_2$  respectivamente de  $2M_z$ ;  $M_z$ ; 1 y 1, se calcula en la etapa 84 una función de coste aplicada a:

- la ventana de análisis que se deriva de la ventana de inicialización que tiene una forma del tipo presentado anteriormente y
- su dual de síntesis, que asegura una reconstrucción perfecta.

La función de coste puede estar compuesta por un conjunto de criterios que hay que verificar en la prueba 85. Se puede citar, por ejemplo y de manera no exhaustiva:

- el análisis de la derivada primera de la ventana:

$$D1 = \log_{10} \left( \frac{1}{2M} \sum_{n=0}^{n=2M-1} D'(n) \right)$$

con

$$D'(n) = \begin{cases} |W(0)| & \text{si } n = 0 \\ |W(n+1) - W(n)| & \text{si } 0 < n < 2M - 1 \\ |W(2M-1)| & \text{si } n = 2M - 1 \end{cases}$$

- el análisis de la derivada segunda de la ventana:

$$D2 = \log_{10} \left( \frac{1}{2M} \sum_{n=0}^{n=2M-1} D''(n) \right)$$

con

$$D''(n) = \begin{cases} D'(0) & \text{si } n = 0 \\ |D'(n+1) - D'(n)| & \text{si } 0 < n < 2M - 1 \\ |D'(2M-1)| & \text{si } n = 2M - 1 \end{cases}$$

- la búsqueda de ganancia de codificación tal como se define, por ejemplo, en el documento WO-2009/081003,

- la maximización de la relación señal a ruido "SNR", expresada en logaritmo entre una señal no codificada y el error introducido por la codificación, para una muestra de audio dada aplicando a la vez la ventana de análisis y la ventana de síntesis en el sentido de la invención.

En la prueba 85, se puede verificar si se han logrado los criterios de final de optimización (valor de función de coste, variación del valor de los parámetros u otros). Llegado el caso (flecha OK), los valores de parámetros optimizados se encuentran en la etapa de final 87. Si no, en el caso contrario (flecha KO), los valores de  $R_1$ ,  $R_2$ ,  $C_1$  y  $C_2$  se modifican, de manera que se reduzca la función de coste.

De este modo, partiendo de una elección de parámetros iniciales  $R_1$ ;  $R_2$ ;  $C_1$  y  $C_2$  respectivamente a 2Mz; Mz; 1 y 1, se encuentra, después de optimización:

$$R_1 = \left[ \frac{2M \times 404}{1280} + 0,5 \right]$$

$$R_2 = \left[ \frac{2M \times 196}{1280} + 0,5 \right]$$

$$C_1 = 4.8425$$

$$C_2 = 0,9659$$

si se toma la opción de optimizar a la vez estos cuatro parámetros, o, alternativamente:

- $C_1$  con un valor cercano de 5 y
- $C_2$  con un valor cercano de 1,

si se toma la opción de no optimizar más que los dos parámetros  $C_1$  y  $C_2$ .

Por lo tanto, en esta segunda opción, se puede limitar la búsqueda de valores óptimos a  $C_1$  y  $C_2$  en el caso en que se considera que el número de 1 es el doble del número de 0 y que el eje de simetría de la ventana corresponde a la mitad del segmento  $W_2$ , como se muestra en la figura 5.

Cabe indicar que con la primera opción mencionada anteriormente (optimizando los cuatro parámetros a la vez), para un retardo de 28 ms (en lugar de 26 ms como anteriormente), se encuentra  $C_1=4,1582$  y  $C_2=0,9181$  con esta técnica de optimización.

Más generalmente, la tabla 1 más abajo presenta un conjunto de parámetros  $R_1$ ,  $R_2$ ,  $C_1$  y  $C_2$  optimizados, según diferentes opciones de retardo y, por lo tanto, de número de muestras a "0" (Mz). El número de muestras por trama M se da para una frecuencia de muestreo a 48.000 Hz y unas tramas de 20 ms:

Tabla 1: optimización de los cuatro parámetros a la vez  $R_1$ ,  $C_1$ ,  $R_2$ ,  $C_2$

Retardo (ms)	M	Mz	$R_1$	$R_2$	$C_1$	$C_2$
26	960	336	$404 \cdot M/640=606$	$196 \cdot M/640=294$	4,8425	0,9659
28	960	288	$M-Mz=672$	$M-2 \cdot Mz=384$	4,1582	0,9181
30	960	240	$480 \cdot M/640=720$	$320 \cdot M/640=480$	3,8945	0,8502
26	960	336	$398 \cdot M/640=597$	$190 \cdot M/640=285$	5,0304	0,9777

Los datos en esta tabla se expresan, en particular, en función del número de muestras a "0" en el cuarto intervalo Mz, como sigue:

Retardo (ms)	M	Mz	$R_1$	$R_2$	$C_1$	$C_2$
26	960	336	$404 \cdot (20/7) \cdot Mz/640=606$	$196 \cdot (20/7) \cdot Mz/640=294$	4,8425	0,9659
28	960	288	$M-Mz=672$	$M-2 \cdot Mz=384$	4,1582	0,9181
30	960	240	$M-Mz=720$	$M-2 \cdot Mz=480$	3,8945	0,8502
26	960	336	$398 \cdot (20/7) \cdot Mz/640=597$	$190 \cdot (20/7) \cdot Mz/640=285$	5,0304	0,9777

La tabla 2 más abajo presenta un conjunto de parámetros  $C_1$  y  $C_2$  optimizados, según la segunda opción mencionada

anteriormente, fijando los intervalos  $R_1$  y  $R_2$  en función del número de ceros  $M_z$ . El número de muestras por trama  $M$  se da para una frecuencia de muestreo a 48.000 Hz y unas tramas de 20 ms:

Tabla 2: optimización de los dos parámetros  $C_1$  y  $C_2$  después de fijación de los intervalos  $R_1$  y  $R_2$

5

Retardo (ms)	M	$M_z$	$R_1=M-M_z+1$	$R_1=M-2*M_z+1$	$C_1$	$C_2$
26	960	336	625	289	4,4265	1,0256
28	960	288	673	385	4,1406	0,9163
30	960	240	721	481	3,7382	0,8699

Por supuesto, la presente invención no se limita a la forma de realización descrita anteriormente a título de ejemplo; se extiende a otras variantes.

10 Por ejemplo, el método de optimización por iteración puede ser como se ha descrito anteriormente, un método de tipo denominado "de los gradientes" pero, por supuesto, son posibles otras variantes. También se puede, por ejemplo, dentro del proceso de optimización, fijar el máximo de la ventana  $h_a$  a 1, de manera que se procesen las muestras sonoras con un peso relativamente constante.

15 De forma más general, los valores de las ventanas de análisis y de síntesis se pueden determinar previamente (por optimización, como se ha descrito anteriormente), luego, ser almacenados en masa en una memoria de un dispositivo de codificación o de decodificación, para ser utilizados en un procedimiento de codificación o de decodificación. En una variante, es posible definir unas ventanas de inicialización, almacenarlas en memoria de los dispositivos mencionados anteriormente y, a continuación, optimizarlas según, por ejemplo, unas condiciones de comunicación.

20

**REIVINDICACIONES**

1. Procedimiento de codificación de una señal digital, estando dicha señal constituida por bloques sucesivos de muestras, siendo la codificación del tipo por transformada con recubrimiento y comprendiendo, en el análisis, la aplicación de una ventana de ponderación sobre dos bloques de  $M$  muestras sucesivas, siendo dicha ventana de ponderación asimétrica e incluyendo cuatro porciones distintas que se extienden sucesivamente sobre dichos dos bloques, con:

- una primera porción ( $w_1$ ), creciente sobre un primer intervalo de muestras,
- una segunda porción ( $w_2$ ), constante a un valor de 1 sobre un segundo intervalo,
- una tercera porción ( $w_3$ ), decreciente sobre un tercer intervalo y
- una cuarta porción ( $w_4$ ), constante a un valor de 0 sobre un cuarto intervalo.

caracterizado por que la ventana varía sobre el primer intervalo, indicado como  $R_1$ , como una función  $w_1$  de tipo:

$$w_1(n) = \frac{W_1(n)}{\Delta_n(n)}, \text{ con } W_1(n) = \text{sen} \left( \frac{\pi n+1}{2 R_1} \right)^{C_1} \text{ y}$$

$$\Delta_n(n) = \sqrt{h_i(n) \cdot h_i(2M - n - 1) + h_i(n + M) \cdot h_i(M - n - 1)}, n \in [0; M - 1]$$

con  $C_1$  y  $R_1$ , unas constantes superiores a 0, siendo  $h_i$  una ventana de inicialización,

y por que la ventana varía sobre el tercer intervalo, indicado como  $R_2$ , como una función  $w_3$  de tipo:

$$w_3(n) = \frac{W_3(n)}{\Delta_n(n)}, \text{ con } W_3(n) = \text{cos} \left( \frac{\pi n}{2 R_2} \right)^{C_2},$$

con  $C_2$  y  $R_2$ , unas constantes superiores a 0.

2. Procedimiento según la reivindicación 1, caracterizado porque los intervalos primero, segundo y tercero se calculan al menos en función del cuarto intervalo ( $M_z$ ).

3. Procedimiento según la reivindicación 1, caracterizado por que el término  $C_1$  está comprendido entre 3 y 5.

4. Procedimiento según la reivindicación 1, caracterizado por que el término  $C_2$  está comprendido entre 0,85 y 1,05.

5. Procedimiento según una de las reivindicaciones anteriores, caracterizado por que la duración del cuarto intervalo, indicado como  $M_z$ , se elige como que es:

$$M_z = \left[ \frac{2M \times 7}{40} \right]$$

donde la indicación  $[x]$  designa el número entero mayor o igual a  $x$  y el más cercano de  $x$ , correspondiendo  $M$  a la duración de un bloque.

6. Procedimiento según una de las reivindicaciones anteriores, caracterizado por que el primer intervalo, indicado como  $R_1$ , es de duración dada por:

$$R_1 = \left[ \frac{2M \times 404}{1280} + 0,5 \right], \text{ donde } M \text{ corresponde a la duración de un bloque,}$$

designando la indicación  $[x]$  el número entero más pequeño o igual a  $x$  y el más cercano de  $x$ .

7. Procedimiento según una de las reivindicaciones anteriores, caracterizado por que el tercer intervalo, indicado como  $R_2$ , es de duración dada por:

$$R_2 = \left[ \frac{2M \times 196}{1280} + 0,5 \right], \text{ donde } M \text{ corresponde a la duración de un bloque,}$$

designando la indicación  $[x]$  el número entero más pequeño o igual a  $x$  y el más cercano de  $x$ .

8. Procedimiento según una de las reivindicaciones anteriores, caracterizado por que el primer intervalo, indicado como  $R_1$ , es de duración dada por:

$$R_1 = [M - M_z + 1 + 0,5]$$

donde  $M$  corresponde a la duración de un bloque y  $M_z$  corresponde a la duración del cuarto intervalo.

5 9. Procedimiento según la reivindicación 8, caracterizado por que el tercer intervalo, indicado como  $R_2$ , es de duración dada por:

$$10 \quad R_2 = [M - 2M_z + 1 + 0,5], \text{ donde } M \text{ corresponde a la duración de un bloque y } M_z \text{ corresponde a la duración del cuarto intervalo.}$$

15 10. Procedimiento según una de las reivindicaciones anteriores, caracterizado por que incluye la aplicación de una pluralidad de ventanas sucesivas, asimétricas y que incluye dichas cuatro porciones, a una sucesión de pares de bloques.

20 11. Procedimiento de decodificación de una señal digital, codificada por la implementación del procedimiento según una de las reivindicaciones anteriores, siendo la decodificación del tipo por transformada con recubrimiento y comprendiendo, en la síntesis, la aplicación de una ventana de ponderación sobre dos bloques de muestras codificadas sucesivas, caracterizado por que dicha ventana de ponderación, en la síntesis, es idéntica a la ventana de análisis utilizada en la codificación, retornada temporalmente.

25 12. Programa informático destinado a ser almacenado en memoria de un dispositivo de codificación o de decodificación, caracterizado por que incluye unas instrucciones para la implementación del procedimiento de codificación según una de las reivindicaciones 1 a 10 o del procedimiento de decodificación según la reivindicación 11, cuando las instrucciones son ejecutadas por un procesador del dispositivo.

30 13. Dispositivo de codificación de una señal, caracterizado por que incluye unos medios de almacenamiento (MEM) y/o de cálculo ( $\mu P$ ) de datos de una ventana de ponderación, de análisis, para la implementación del procedimiento de codificación según una de las reivindicaciones 1 a 10.

35 14. Dispositivo de decodificación de una señal, caracterizado por que incluye unos medios de almacenamiento (MEM') y/o de cálculo ( $\mu P$ ) de datos de una ventana de ponderación, de síntesis, para la implementación del procedimiento de decodificación según la reivindicación 11.

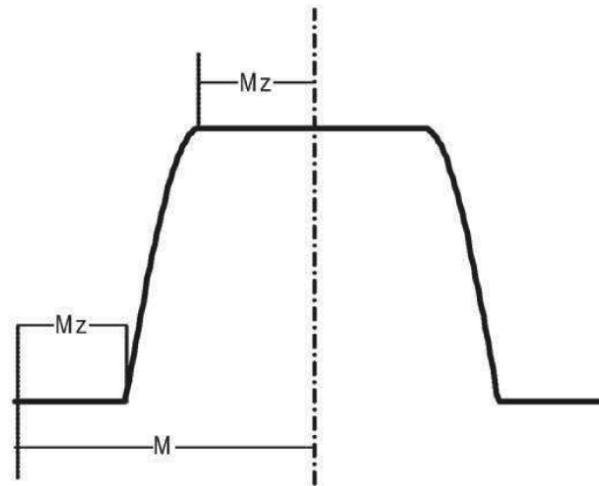


FIG. 1

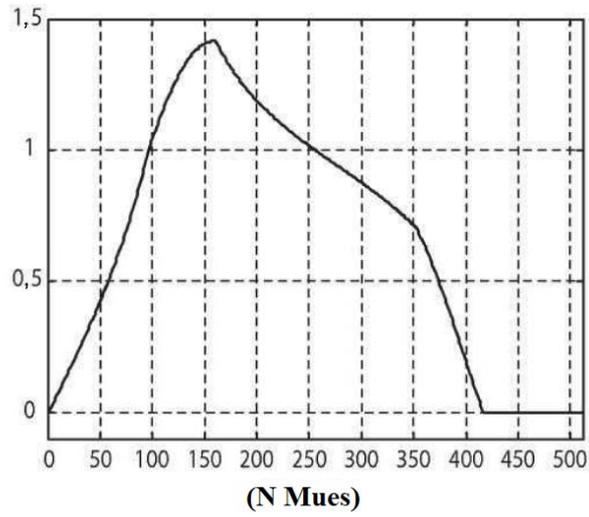


FIG. 2

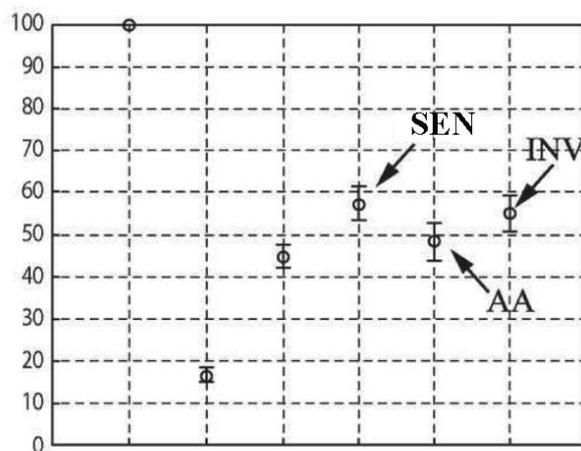


FIG. 3

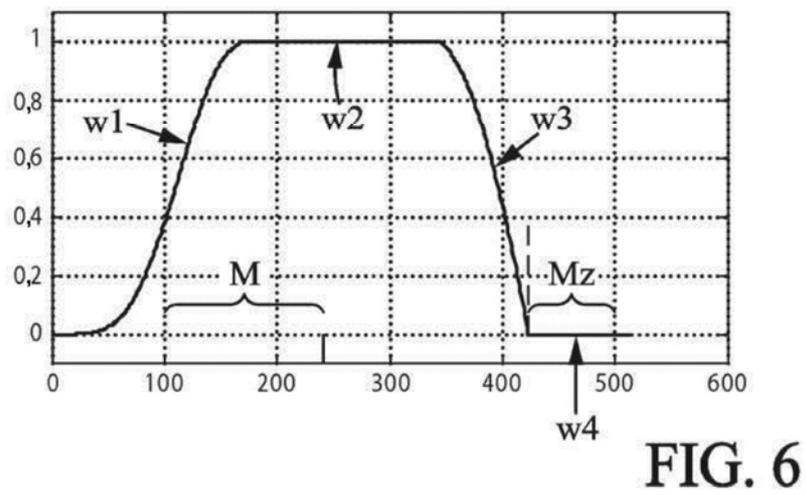
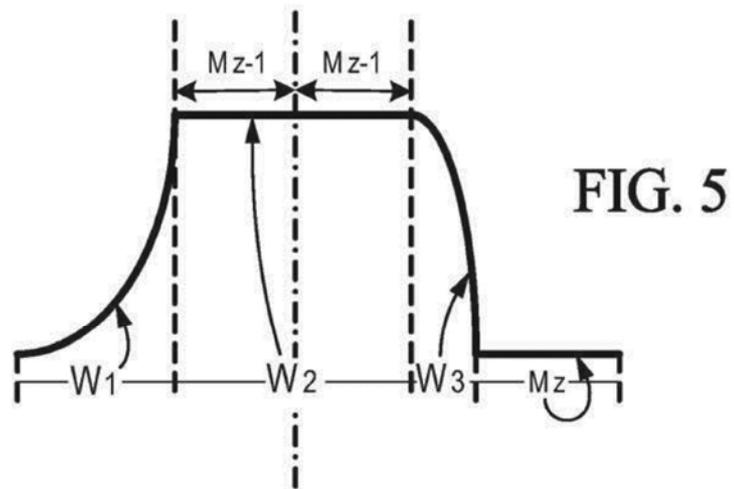
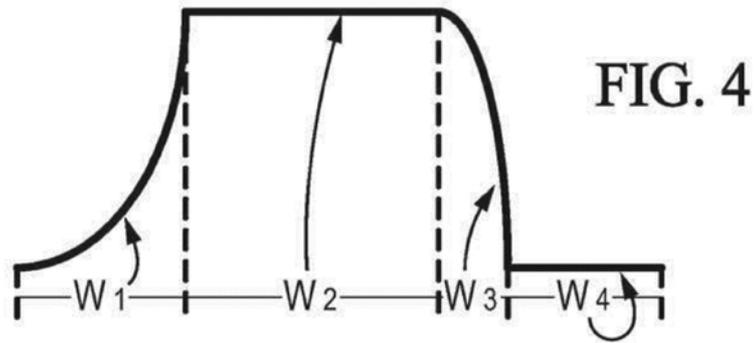


FIG. 7

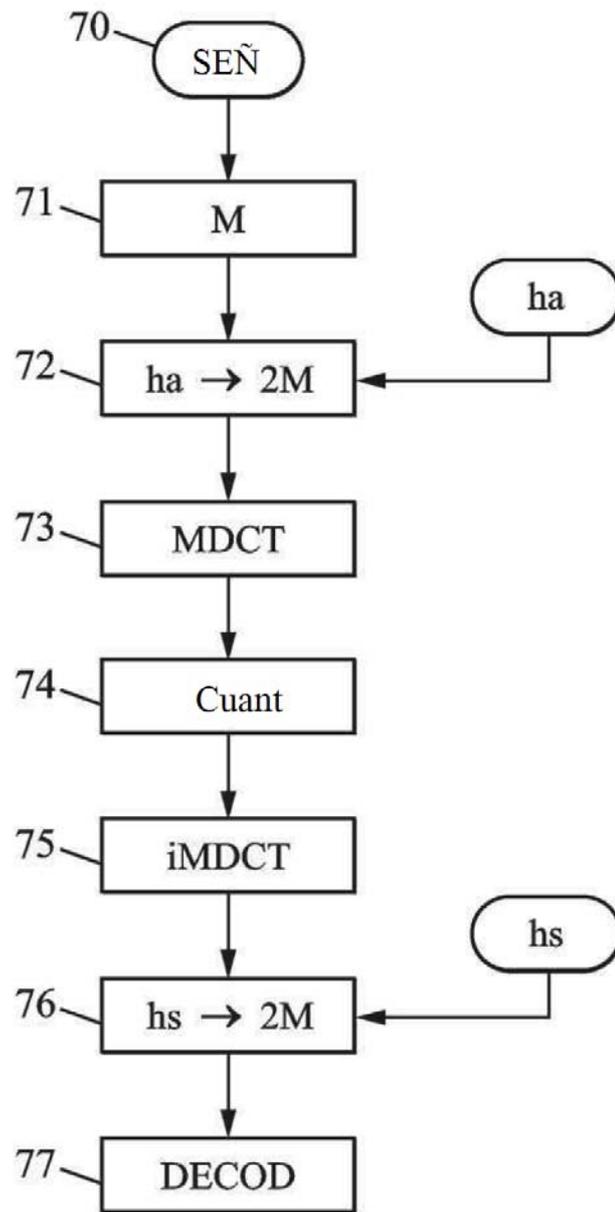
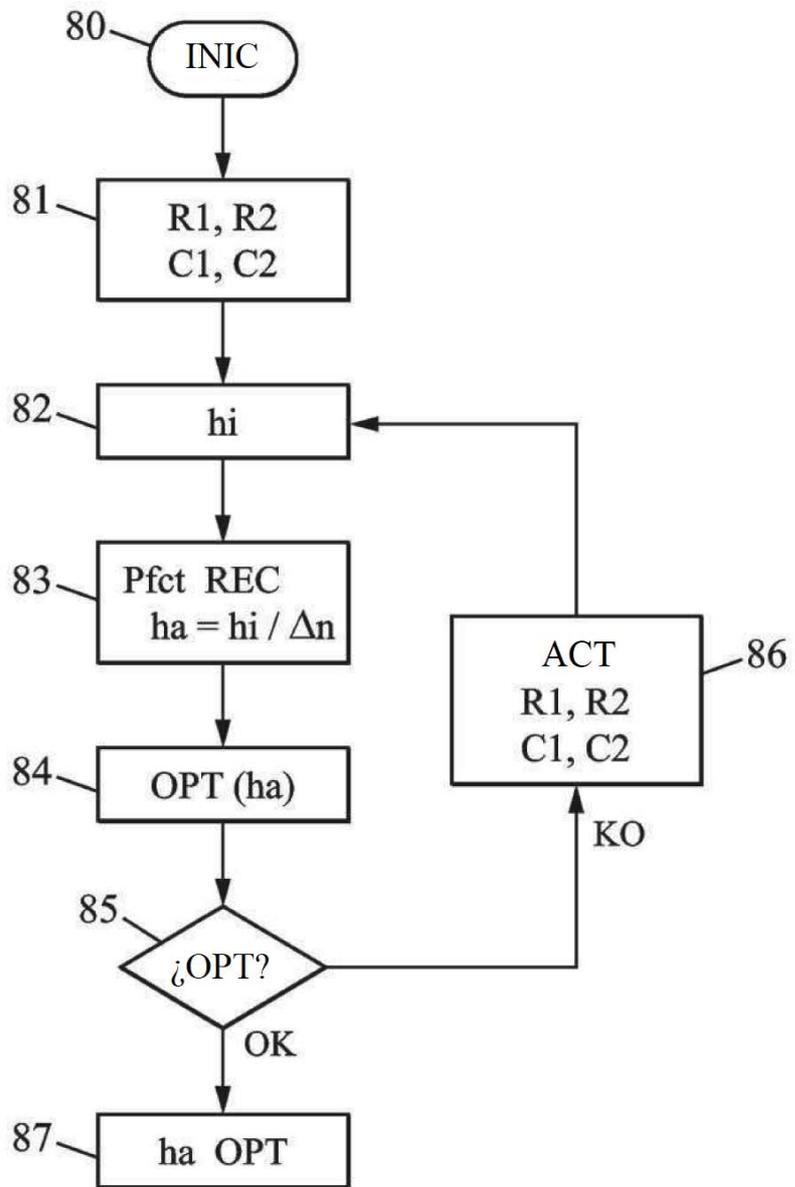


FIG. 8



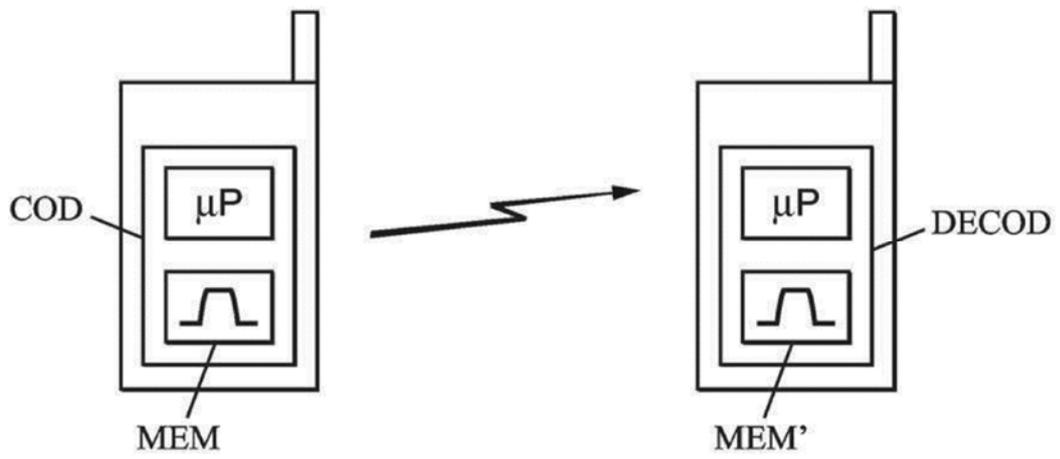


FIG. 9

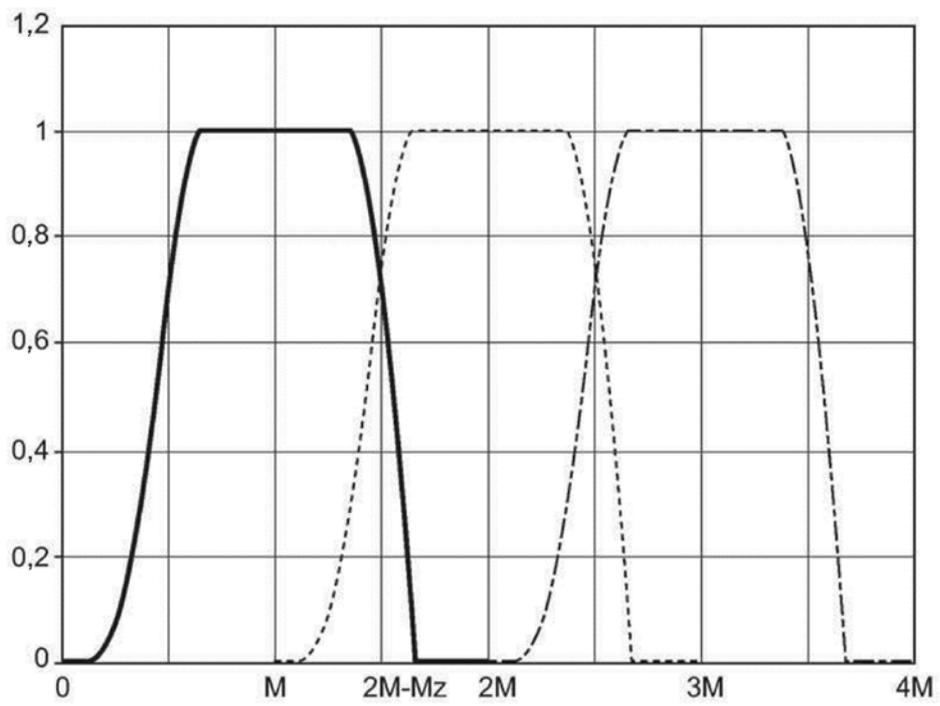


FIG. 10