

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 797 525**

51 Int. Cl.:

G10L 19/00 (2013.01)

G10L 19/032 (2013.01)

G10L 19/18 (2013.01)

G10L 19/26 (2013.01)

G10L 19/02 (2013.01)

G10L 21/0208 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **15.10.2010** **PCT/CA2010/001649**

87 Fecha y número de publicación internacional: **21.04.2011** **WO11044700**

96 Fecha de presentación y número de la solicitud europea: **15.10.2010** **E 10822970 (9)**

97 Fecha y número de publicación de la concesión europea: **20.05.2020** **EP 2489041**

54 Título: **Conformación simultánea de ruido en el dominio del tiempo y el dominio de la frecuencia para transformaciones TDAC**

30 Prioridad:

15.10.2009 US 272644 P

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

02.12.2020

73 Titular/es:

VOICEAGE CORPORATION (100.0%)
750 Lucerne Road, Suite 250
Town of Mount Royal, Quebec H3R 2H6, CA

72 Inventor/es:

BESSETTE, BRUNO

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 797 525 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

Conformación simultánea de ruido en el dominio del tiempo y el dominio de la frecuencia para transformaciones TDAC

5 Campo de la invención

La presente invención se refiere a un método y dispositivo de conformación de ruido en el dominio de la frecuencia para interpolar una forma espectral y una envolvente en el dominio del tiempo de un ruido de cuantificación en una señal de audio en ventana y codificada por transformación.

10

Antecedentes

La codificación de transformación especializada produce importantes ahorros en la velocidad de bits al representar señales digitales como el audio. Las transformaciones tal como la Transformada discreta de Fourier (DFT) y la Transformada discreta de coseno (DCT) proporcionan una representación compacta de la señal de audio al condensar la mayor parte de la energía de la señal en relativamente pocos coeficientes espectrales, en comparación con las muestras en el dominio del tiempo donde la energía se distribuye en todas las muestras. Esta propiedad de compactación de energía de las transformaciones puede conducir a una cuantificación eficiente, por ejemplo, a través de la asignación de bits adaptativa, y la minimización de la distorsión percibida, por ejemplo, mediante el uso de modelos de enmascaramiento de ruido. Una reducción de datos adicional puede lograrse mediante el uso de transformaciones solapadas y la cancelación de solapamiento en el dominio del tiempo (TDAC). La DCT modificada (MDCT) es un ejemplo de tales transformaciones solapadas, en las que los bloques adyacentes de muestras de la señal de audio a procesar se solapan entre sí para evitar artefactos de discontinuidad mientras se mantiene el muestreo crítico (N muestras de la señal de audio de entrada producen N coeficientes de transformación). La propiedad TDAC de la MDCT proporciona esta ventaja adicional en la compactación de energía.

Los modelos recientes de codificación de audio usan un enfoque multimodo. En este enfoque, pueden usarse varias herramientas de codificación para codificar de manera más eficiente cualquier tipo de señal de audio (voz, música, mixta, etc.). Estas herramientas comprenden transformaciones tal como la MDCT y predictores tales como los predictores de tono y filtros de codificación predictiva lineal (LPC) usados en la codificación de voz. Cuando se opera un códec multimodo, las transiciones entre los diferentes modos de codificación se procesan cuidadosamente para evitar artefactos audibles debido a la transición. En particular, la conformación del ruido de cuantificación en los diferentes modos de codificación se realiza típicamente mediante el uso de diferentes procedimientos. En las tramas que usan codificación de transformación, el ruido de cuantificación se forma en el dominio de transformación (es decir, al cuantificar los coeficientes de transformación), al aplicar varias etapas de cuantificación que se controlan por factores de escala derivados, por ejemplo, de la energía de la señal de audio en diferentes bandas espectrales. Por otro lado, en las tramas que usan un modelo predictivo en el dominio del tiempo (que típicamente involucra predictores a largo plazo y predictores a corto plazo), el ruido de cuantificación se forma mediante el uso de un llamado filtro de ponderación cuya función de transferencia en el dominio de transformación $W(z)$ se denota $W(z)$. La conformación de ruido se aplica entonces al filtrar primero las muestras en el dominio del tiempo de la señal de audio de entrada a través del filtro de ponderación $W(z)$ para obtener una señal ponderada, y luego codificar la señal ponderada en este llamado dominio ponderado. La forma espectral, o respuesta de frecuencia, del filtro de ponderación $W(z)$ se controla de manera que el ruido de codificación (o cuantificación) quede enmascarado por la señal de audio de entrada. Típicamente, el filtro de ponderación $W(z)$ se deriva del filtro LPC, que modela la envolvente espectral de la señal de audio de entrada.

45

Un ejemplo de un códec de audio multimodo es el Códec unificado de voz y audio (USAC) del Grupo de expertos en imágenes en movimiento (MPEG). Este códec integra herramientas que incluyen la codificación de transformación y la codificación predictiva lineal, y puede cambiar entre diferentes modos de codificación en dependencia de las características de la señal de audio de entrada. Hay tres (3) modos de codificación básicos en el USAC:

50

1) Un modo de codificación basado en la codificación de audio avanzada (AAC), que codifica la señal de audio de entrada mediante el uso de la MDCT y la cuantificación derivada perceptivamente de los coeficientes de la MDCT;

2) Un modo de codificación basado en la predicción lineal excitada de código algebraico (ACELP), que codifica la señal de audio de entrada como una señal de excitación (una señal en el dominio del tiempo) procesada a través de un filtro de síntesis; y

55

3) Un modo de codificación basado en excitación codificada por transformación (TCX) que es una especie de híbrido entre los dos modos anteriores, en donde la excitación del filtro de síntesis del segundo modo se codifica en el dominio de la frecuencia; en realidad, esta es una señal objetivo o la señal ponderada que se codifica en el dominio de transformación.

En el USAC, el modo de codificación basado en TCX y el modo de codificación basado en AAC usan una transformación similar, por ejemplo, la MDCT. Sin embargo, en su forma estándar, la AAC y la TCX no aplican el mismo mecanismo para controlar la forma espectral del ruido de cuantificación. La AAC controla explícitamente el ruido de cuantificación en el dominio de la frecuencia en las etapas de cuantificación de los coeficientes de transformación. La TCX, sin embargo, controla la forma espectral del ruido de cuantificación mediante el uso del filtrado en el dominio del tiempo y, más específicamente, mediante el uso de un filtro de ponderación $W(z)$ como se describió anteriormente. Para facilitar la conformación del ruido de cuantificación en un códec de audio multimodo, existe la necesidad de un dispositivo y un método para la conformación simultánea del ruido en el dominio del tiempo y el dominio de la frecuencia para las

65

transformaciones TDAC. Jeremie Lecomte y otros: "Efficient Cross-Fade Windows for Transitions between LPC-Based and Non-LPC Based Audio Coding", CONVENCIÓN AES 126; MAYO 2009 describe el cambio entre ventanas de diferentes formas en los bordes AAC y TCX.

5 Resumen de la invención

El alcance de la invención se establece solamente por las reivindicaciones adjuntas. De acuerdo con un primer aspecto, la presente invención se refiere a un método de conformación de ruido en el dominio de la frecuencia para interpolar una forma espectral y una envolvente en el dominio del tiempo de un ruido de cuantificación en una señal de audio en ventana y codificada por transformación, que comprende dividir los coeficientes de transformación de la señal de audio en ventana y codificada por transformación en una pluralidad de bandas espectrales. El método de conformación de ruido en el dominio de la frecuencia comprende además, para cada banda espectral: calcular una primera ganancia que representa, junto con las ganancias correspondientes calculadas para las otras bandas espectrales, una forma espectral del ruido de cuantificación en una primera transición entre una primera ventana de tiempo y una segunda ventana de tiempo; calcular una segunda ganancia que representa, junto con las ganancias correspondientes calculadas para las otras bandas espectrales, una forma espectral del ruido de cuantificación en una segunda transición entre la segunda ventana de tiempo y una tercera ventana de tiempo; y filtrar los coeficientes de transformación de la segunda ventana de tiempo en base a las primera y segunda ganancias, para interpolar entre las primera y segunda transiciones la forma espectral y la envolvente en el dominio del tiempo del ruido de cuantificación.

De acuerdo con un segundo aspecto, la presente invención se refiere a un dispositivo de conformación de ruido en el dominio de la frecuencia para interpolar una forma espectral y una envolvente en el dominio del tiempo de un ruido de cuantificación en una señal de audio en ventana y codificada por transformación, que comprende: un divisor de los coeficientes de transformación de la señal de audio en ventana y codificada por transformación en una pluralidad de bandas espectrales; una calculadora, para cada banda espectral, de una primera ganancia que representa, junto con las ganancias correspondientes calculadas para las otras bandas espectrales, una forma espectral del ruido de cuantificación en una primera transición entre una primera ventana de tiempo y una segunda ventana de tiempo, y de una segunda ganancia que representa, junto con las ganancias correspondientes calculadas para las otras bandas espectrales, una forma espectral del ruido de cuantificación en una segunda transición entre la segunda ventana de tiempo y una tercera ventana de tiempo; y un filtro de los coeficientes de transformación de la segunda ventana de tiempo en base a las primera y segunda ganancias, para interpolar entre las primera y segunda transiciones la forma espectral y la envolvente en el dominio del tiempo del ruido de cuantificación.

En la presente descripción y las reivindicaciones adjuntas, el término "ventana de tiempo" designa un bloque de muestras en el dominio del tiempo, y el término "señal en ventana" designa una ventana del dominio del tiempo después de la aplicación de una ventana no rectangular.

Los anteriores y otros objetivos, ventajas y características de la presente invención resultarán más evidentes tras la lectura de la siguiente descripción no restrictiva de una modalidad ilustrativa de la misma, que se muestra a manera de ejemplo solo con referencia a las figuras acompañantes.

Breve descripción de las figuras

En las figuras adjuntas:

La Figura 1 es un diagrama de bloques esquemático que ilustra el principio general de la Conformación de ruido temporal (TNS);

La Figura 2 es un diagrama de bloques esquemático de un dispositivo de conformación de ruido en el dominio de la frecuencia para interpolar una forma espectral y una envolvente en el dominio del tiempo del ruido de cuantificación;

La Figura 3 es un diagrama de flujo que describe las operaciones de un método de conformación de ruido en el dominio de la frecuencia para interpolar la forma espectral y la envolvente en el dominio del tiempo del ruido de cuantificación;

La Figura 4 es un diagrama esquemático de las posiciones relativas de la ventana para transformaciones y ganancias de ruido, que considera el cálculo de las ganancias de ruido para la ventana 1;

La Figura 5 es un gráfico que ilustra el efecto de la interpolación de la forma del ruido, tanto en la forma espectral como en la envolvente en el dominio del tiempo del ruido de cuantificación;

La Figura 6 es un gráfico que ilustra una $m^{\text{ésima}}$ envolvente en el dominio del tiempo, que puede verse como la forma de ruido en una $m^{\text{ésima}}$ banda espectral que evoluciona en el tiempo desde el punto A al punto B;

La Figura 7 es un diagrama de bloques esquemático de un codificador capaz de conmutar entre un modo de codificación en el dominio de la frecuencia mediante el uso de, por ejemplo, MDCT y un modo de codificación en el dominio del tiempo mediante el uso de, por ejemplo, ACELP, el codificador que aplica la Conformación del ruido en el dominio de la frecuencia (FNDS) para codificar un bloque de muestras de una señal de audio de entrada; y

La Figura 8 es un diagrama de bloques esquemático de un decodificador que produce un bloque de señal de síntesis mediante el uso de FDNS, en donde el decodificador puede cambiar entre un modo de codificación en el dominio de la frecuencia mediante el uso de, por ejemplo, MDCT y un modo de codificación en el dominio del tiempo mediante el uso de, por ejemplo, ACELP.

Descripción detallada

Primero se discutirá brevemente el principio básico de la Conformación de ruido temporal (TNS), mencionada en la siguiente descripción.

La TNS es una técnica conocida por los expertos en la técnica de la codificación de audio para conformar el ruido de codificación en el dominio del tiempo. Con referencia a la Figura 1, un sistema TNS 100 comprende:

- Un procesador de transformación 101 para someter un bloque de muestras de una señal de audio de entrada $x[n]$ a una transformación, por ejemplo, la Transformada discreta del coseno (DCT) o la DCT modificada (MDCT), y producir los coeficientes de transformación $X[k]$;

- Un único filtro 102 aplicado a todas las bandas espectrales, más específicamente a todos los coeficientes de transformación $X[k]$ del procesador de transformación 101 para producir los coeficientes de transformación filtrados $X_f[k]$;

- Un procesador 103 para cuantificar, codificar, transmitir a un receptor o almacenar en un dispositivo de almacenamiento, decodificar y cuantificar inversamente los coeficientes de transformación filtrados $X_f[k]$ para producir los coeficientes de transformación cuantificados $Y_f[k]$;

- Un único filtro inverso 104 para procesar los coeficientes de transformación cuantificados $Y_f[k]$ para producir los coeficientes de transformación decodificados $Y[k]$; y finalmente,

- Un procesador de transformación inversa 105 para aplicar una transformación inversa a los coeficientes de transformación decodificados $Y[k]$ para producir un bloque decodificado de muestras de salida en el dominio del tiempo $y[n]$.

Dado que, en el ejemplo de la Figura 1, el procesador de transformación 101 usa la DCT o MDCT, la transformación inversa aplicada en el procesador de transformación inversa 105 es la DCT inversa o MDCT inversa. El filtro único 102 de la Figura 1 se deriva de un filtro de predicción óptimo para los coeficientes de transformación. Esto da como resultado, en TNS, la modulación del ruido de cuantificación con una envolvente en el dominio del tiempo que sigue a la envolvente en el dominio del tiempo de la señal de audio para la trama actual.

Con referencia a las Figuras 2 y 3, la siguiente descripción describe simultáneamente un dispositivo de conformación de ruido en el dominio de la frecuencia 200 y el método 300 para interpolar la forma espectral y la envolvente en el dominio del tiempo del ruido de cuantificación. Más específicamente, en el dispositivo 200 y el método 300, la forma espectral y la amplitud en el dominio del tiempo del ruido de cuantificación en la transición entre dos bloques codificados por transformación superpuestos se interpolan simultáneamente. Los bloques adyacentes codificados por transformación pueden ser de naturaleza similar, tal como dos bloques consecutivos de codificación de audio avanzada (AAC) producidos por un codificador AAC o dos bloques consecutivos de excitación codificada por transformación (TCX) producidos por un codificador TCX, pero también pueden ser de diferente naturaleza tal como un bloque AAC seguido de un bloque TCX, o viceversa, en donde dos codificadores distintos se usan consecutivamente. Tanto la forma espectral como la envolvente en el dominio del tiempo del ruido de cuantificación evolucionan suavemente (o se interpolan continuamente) en la unión entre dos de estos bloques codificados por transformación.

Operación 301 (Figura 3) - Transformación

La señal de audio de entrada $x[n]$ de las Figuras 2 y 3 es un bloque de N muestras en el dominio del tiempo de la señal de audio de entrada que cubre la longitud de un bloque de transformación. Por ejemplo, la señal de entrada $x[n]$ abarca la longitud de la ventana en el dominio del tiempo 1 de la Figura 4.

En la operación 301, la señal de entrada $x[n]$ se transforma a través de un procesador de transformación 201 (Figura 2). Por ejemplo, el procesador de transformación 201 puede implementar una MDCT que incluye una ventana en el dominio del tiempo (por ejemplo, la ventana 1 de la Figura 4) que multiplica la señal de entrada $x[n]$ antes de calcular los coeficientes de transformación $X[k]$. Como se ilustra en la Figura 2, el procesador de transformación 201 genera los coeficientes de transformación $X[k]$. En el ejemplo no limitativo de una TCMD, los coeficientes de transformación $X[k]$ comprenden N coeficientes espectrales, que es el mismo que el número de muestras en el dominio del tiempo que forman la señal de audio de entrada $x[n]$.

Operación 302 (Figura 3) - División de banda

En la operación 302, un divisor de banda 202 (Figura 2) divide los coeficientes de transformación $X[k]$ en M bandas espectrales. Más específicamente, los coeficientes de transformación $X[k]$ se dividen en la bandas espectrales $B_1[k]$, $B_2[k]$, $B_3[k]$, ..., $B_M[k]$. La concatenación de las bandas espectrales $B_1[k]$, $B_2[k]$, $B_3[k]$, ..., $B_M[k]$ da el conjunto completo de coeficientes de transformación, específicamente $B[k]$. El número de bandas espectrales y el número de coeficientes de transformación por banda espectral pueden variar en dependencia de la resolución de frecuencia deseada.

Operación 303 (Figura 3) - Filtrado 1, 2, 3, ..., M

Después de la división de banda 302, en la operación 303, cada banda espectral $B_1[k]$, $B_2[k]$, $B_3[k]$, ..., $B_M[k]$ se filtra a través de un filtro específico de banda (Filtros 1, 2, 3, ..., M en la Figura 2). Los Filtros 1, 2, 3, ..., M pueden ser diferentes para cada banda espectral, o puede usarse el mismo filtro para todas las bandas espectrales. En una modalidad, los

Filtros 1, 2, 3, ..., M de la Figura 2 son diferentes para cada bloque de muestras de la señal de audio de entrada $x[n]$. La operación 303 produce las bandas filtradas $B_{1f}[k]$, $B_{2f}[k]$, $B_{3f}[k]$, ..., $B_{Mf}[k]$ de las Figuras 2 y 3.

Operación 304 (Figura 3) - Cuantificación, codificación, transmisión o almacenamiento, decodificación, cuantificación inversa

En la operación 304, las bandas filtradas $B_{1f}[k]$, $B_{2f}[k]$, $B_{3f}[k]$, ..., $B_{Mf}[k]$ de los Filtros 1, 2, 3, ..., M pueden cuantificarse, codificarse, transmitirse a un receptor (no mostrado) y/o almacenarse en cualquier dispositivo de almacenamiento (no mostrado). La cuantificación, codificación, transmisión a un receptor y/o almacenamiento en un dispositivo de almacenamiento se realizan y/o controlan mediante un procesador Q de la Figura 2. El procesador Q puede conectarse además a y controlar un transceptor (no mostrado) para transmitir las bandas cuantificadas, codificadas y filtradas $B_{1f}[k]$, $B_{2f}[k]$, $B_{3f}[k]$, ..., $B_{Mf}[k]$ al receptor. De la misma manera, el procesador Q puede conectarse a y controlar el dispositivo de almacenamiento para almacenar las bandas cuantificadas, codificadas y filtradas $B_{1f}[k]$, $B_{2f}[k]$, $B_{3f}[k]$, ..., $B_{Mf}[k]$.

En la operación 304, las bandas cuantificadas, codificadas y filtradas $B_{1f}[k]$, $B_{2f}[k]$, $B_{3f}[k]$, ..., $B_{Mf}[k]$ también pueden recibirse por el transceptor o recuperarse desde el dispositivo de almacenamiento, decodificarse y cuantificarse inversamente por el procesador Q. Estas operaciones de recepción (a través del transceptor) o recuperación (desde el dispositivo de almacenamiento), decodificación y cuantificación inversa producen las bandas espectrales cuantificadas $C_{1f}[k]$, $C_{2f}[k]$, $C_{3f}[k]$, ..., $C_{Mf}[k]$ en la salida del procesador Q.

Cualquier tipo de cuantificación, codificación, transmisión (y/o almacenamiento), recepción, decodificación y cuantificación inversa pueden usarse en la operación 304 sin pérdida de generalidad.

Operación 305 (Figura 3) - Filtrado inverso 1, 2, 3, ..., M

En la operación 305, las bandas espectrales cuantificadas $C_{1f}[k]$, $C_{2f}[k]$, $C_{3f}[k]$, ..., $C_{Mf}[k]$ se procesan a través de filtros inversos, más específicamente el Filtro inverso 1, el Filtro inverso 2, el Filtro inverso 3, ..., el Filtro inverso M de la Figura 2, para producir las bandas espectrales decodificadas $C_1[k]$, $C_2[k]$, $C_3[k]$, ..., $C_M[k]$. El Filtro inverso 1, el Filtro inverso 2, el Filtro inverso 3, ..., el Filtro inverso M tienen funciones de transferencia inversas de las funciones de transferencia del Filtro 1, el Filtro 2, el Filtro 3, ..., el Filtro M , respectivamente.

Operación 306 (Figura 3) - Concatenación de banda espectral

En la operación 306, las bandas espectrales decodificadas $C_1[k]$, $C_2[k]$, $C_3[k]$, ..., $C_M[k]$ se concatenan después en un concatenador de banda 203 de la Figura 2, para producir los coeficientes espectrales decodificados $Y[k]$ (espectro decodificado).

Operación 307 (Figura 3) - Transformación inversa

Finalmente, en la operación 307, un procesador de transformación inversa 204 (Figura 2) aplica una transformación inversa a los coeficientes espectrales decodificados $Y[k]$ para producir un bloque decodificado de las muestras de salida en el dominio del tiempo $y[n]$. En el caso del ejemplo no limitativo anterior que usa la MDCT, el procesador de transformación inversa 204 aplica la MDCT inversa (IMDCT) a los coeficientes espectrales decodificados $Y[k]$.

Operación 308 (Figura 3) - Cálculo de las ganancias de ruido $g_1[m]$ y $g_2[m]$

En la Figura 2, el Filtro 1, el Filtro 2, el Filtro 3, ..., el Filtro M y el Filtro inverso 1, el Filtro inverso 2, el Filtro inverso 3, ..., los parámetros del Filtro inverso M (ganancias de ruido) $g_1[m]$ y $g_2[m]$ como entrada. Estas ganancias de ruido representan las formas espectrales del ruido de cuantificación y se describirán adicionalmente en la presente descripción más abajo. Además, los Filtrados 1, 2, 3, ..., M de la Figura 3 pueden ser secuenciales; el Filtro 1 puede aplicarse antes que el Filtro 2, luego el Filtro 3, y así sucesivamente hasta el Filtro M (Figura 2). Los Filtrados inversos 1, 2, 3, ..., M también pueden ser secuenciales; el Filtro inverso 1 puede aplicarse antes que el Filtro inverso 2, luego el Filtro inverso 3, y así sucesivamente hasta el Filtro inverso M (Figura 2). Como tal, cada filtro y filtro inverso pueden usar como estado inicial el estado final del filtro o filtro inverso anterior. Esta operación secuencial puede garantizar la continuidad en el proceso de filtrado de una banda espectral a la siguiente. En una modalidad, esta restricción de continuidad en los estados de filtro de una banda espectral a la siguiente puede no aplicarse.

La Figura 4 ilustra cómo puede usarse la conformación del ruido en el dominio de la frecuencia para interpolar la forma espectral y la envolvente en el dominio del tiempo del ruido de cuantificación al procesar una señal de audio segmentada por ventanas superpuestas (ventana 0, ventana 1, ventana 2 y ventana 3) en bloques de transformación adyacentes superpuestos (bloques de muestras de la señal de audio de entrada). Cada ventana de la Figura 4, es decir, la ventana 0, la ventana 1, la ventana 2 y la ventana 3, muestra el intervalo de tiempo de un bloque de transformación y la forma de la ventana aplicada por el procesador de transformación 201 de la Figura 2 a ese bloque de muestras de la señal de audio de entrada. Como se describió anteriormente, el procesador de transformación 201 de la Figura 2 implementa tanto la visualización en ventanas de la señal de audio de entrada $x[n]$ como la aplicación de la transformación para producir los

coeficientes de transformación $X[k]$. La forma de las ventanas (ventana 0, ventana 1, ventana 2 y ventana 3) mostradas en la Figura 4 puede cambiarse sin pérdida de generalidad.

En la Figura 4, se considera el procesamiento de un bloque de muestras de la señal de audio de entrada $x[n]$ desde el principio hasta el final de la ventana 1. El bloque de muestras de la señal de audio de entrada $x[n]$ se suministra al procesador de transformación 201 de la Figura 2. En la operación de cálculo 308 (Figura 3), la calculadora 205 (Figura 2) calcula dos conjuntos de ganancias de ruido $g_1[m]$ y $g_2[m]$ usados para las operaciones de filtrado (Filtros 1 al M y Filtros inversos 1 al M). Estos dos conjuntos de ganancias de ruido en realidad representan los niveles deseados de ruido en las M bandas espectrales en una posición dada en el tiempo. Por lo tanto, las ganancias de ruido $g_1[m]$ y $g_2[m]$ cada una representa la forma espectral del ruido de cuantificación en tal posición en el eje del tiempo. En la Figura 4, las ganancias de ruido $g_1[m]$ corresponden a algún análisis centrado en el punto A en el eje de tiempo, y las ganancias de ruido $g_2[m]$ corresponden a otro análisis más arriba en el eje del tiempo, en la posición B. Para un funcionamiento óptimo, los análisis de estas ganancias de ruido se centran en el punto medio de la superposición entre las ventanas adyacentes y los bloques de muestras correspondientes. En consecuencia, con referencia a la Figura 4, el análisis para obtener las ganancias de ruido $g_1[m]$ para la ventana 1 se centra en el punto medio de la superposición (o transición) entre la ventana 0 y la ventana 1 (vea el punto A en el eje de tiempo). Además, el análisis para obtener las ganancias de ruido $g_2[m]$ para la ventana 1 se centra en el punto medio de la superposición (o transición) entre la ventana 1 y la ventana 2 (vea el punto B en el eje de tiempo).

La calculadora 205 (Figura 2) puede usar una pluralidad de diferentes procedimientos de análisis para obtener los conjuntos de ganancias de ruido $g_1[m]$ y $g_2[m]$, siempre que tal procedimiento de análisis conduzca a un conjunto de ganancias de ruido adecuadas en el dominio de la frecuencia para cada una de las M bandas espectrales $B_1[k]$, $B_2[k]$, $B_3[k]$, ..., $B_M[k]$ de las Figuras 2 y 3. Por ejemplo, puede aplicarse una Codificación predictiva lineal (LPC) a la señal de audio de entrada $x[n]$ para obtener un predictor a corto plazo del cual se deriva un filtro de ponderación $W(z)$. El filtro de ponderación $W(z)$ luego se mapea en el dominio de la frecuencia para obtener las ganancias de ruido $g_1[m]$ y $g_2[m]$. Este sería un procedimiento de análisis típico que se puede usar cuando el bloque de muestras de la señal de entrada $x[n]$ en la ventana 1 de la Figura 4 está codificado en el modo TCX. Otro enfoque para obtener las ganancias de ruido $g_1[m]$ y $g_2[m]$ de las Figuras 2 y 3 podría ser como en AAC, donde el nivel de ruido en cada banda de frecuencia se controla por factores de escala (derivados de un modelo psicoacústico) en el dominio MDCT.

Al haber procesado a través del procesador de transformación 201 de la Figura 2, el bloque de muestras de la señal de entrada $x[n]$ que abarca la longitud de la ventana 1 de la Figura 4, y al haber obtenido los conjuntos de ganancias de ruido $g_1[m]$ y $g_2[m]$ en las posiciones A y B en el eje de tiempo de la Figura 4 mediante el uso de la calculadora 205, se realizan las operaciones de filtrado para cada banda espectral $B_1[k]$, $B_2[k]$, $B_3[k]$, ..., $B_M[k]$ de la Figura 2. El objetivo de las operaciones de filtrado (y filtrado inverso) es lograr una forma espectral deseada del ruido de cuantificación en las posiciones A y B en el eje del tiempo, y también asegurar una transición o interpolación suave de esta forma espectral o la envolvente de esta forma espectral del punto A al punto B, en una base de muestra por muestra. Esto se muestra en la Figura 5, en la cual una ilustración de las ganancias de ruido $g_1[m]$ se muestra en el punto A y una ilustración de las ganancias de ruido $g_2[m]$ se muestra en el punto B. Si cada una de las bandas espectrales $B_1[k]$, $B_2[k]$, $B_3[k]$, ..., $B_M[k]$ simplemente se multiplicó por una función de las ganancias de ruido $g_1[m]$ y $g_2[m]$, por ejemplo, al tomar una suma ponderada de $g_1[m]$ y $g_2[m]$ y multiplicar por este resultado los coeficientes en la banda espectral $B_m[k]$, m que toma uno de los valores 1, 2, 3, ..., M , entonces las curvas de ganancia interpoladas mostradas en la Figura 5 serían constantes (horizontales) desde el punto A hasta el punto B. Para obtener curvas de ganancia de ruido que varían suavemente de la ganancia $g_1[m]$ a la ganancia $g_2[m]$ para cada banda espectral como se muestra en la Figura 5, el filtrado puede aplicarse a cada banda espectral $B_m[k]$. Por la propiedad de dualidad de muchas transformaciones lineales, en particular la DCT y la MDCT, una operación de filtrado (o convolución) en un dominio da como resultado una multiplicación en el otro dominio. En consecuencia, filtrar los coeficientes de transformación en una banda espectral $B_m[k]$ da como resultado la interpolación y la aplicación de una envolvente en el dominio del tiempo (multiplicación) al ruido de cuantificación en esa banda espectral. Esta es la base de TNS, cuyo principio se presenta brevemente en la descripción anterior de la Figura 1.

Sin embargo, existen diferencias fundamentales entre TNS y la interpolación propuesta en la presente descripción. Como una primera diferencia entre TNS y la técnica descrita en la presente descripción, el objetivo y el procesamiento son diferentes. En la técnica descrita en la presente descripción, el objetivo es imponer, durante la duración de una ventana dada (por ejemplo, la ventana 1 de la Figura 4), una envolvente en el dominio del tiempo para el ruido de cuantificación en una banda dada $B_m[k]$ que varía suavemente de la ganancia de ruido $g_1[m]$ calculada en el punto A a la ganancia de ruido $g_2[m]$ calculada en el punto B. La Figura 6 muestra un ejemplo de la envolvente interpolada en el dominio del tiempo de la ganancia de ruido, para la banda espectral $B_m[k]$. Hay varias posibilidades para tal curva interpolada, y el filtro en el dominio de la frecuencia correspondiente para esa banda espectral $B_m[k]$. Por ejemplo, puede usarse una estructura de filtro recursivo de primer orden para cada banda espectral. Son posibles muchas otras estructuras de filtro, sin pérdida de generalidad.

Dado que el objetivo es conformar, a través del filtrado, el ruido de cuantificación en cada banda espectral $B_m[k]$, la primera preocupación se dirige a los Filtros inversos 1 al M de la Figura 2, que es la operación de filtrado inverso que formará el ruido de cuantificación introducido por el procesador Q (Figura 2).

Si consideramos entonces que los coeficientes de transformación cuantificados $Y_l[k]$ de la banda espectral $C_m[k]$ se filtran de la siguiente manera

$$C_m[k] = aC_{mf}[k] + bC_m[k-1] \quad (1)$$

mediante el uso de los parámetros de filtro a y b . La ecuación (1) representa un filtro recursivo de primer orden, aplicado a los coeficientes de transformación de la banda espectral $C_m[k]$. Como se indicó anteriormente, está dentro del alcance de la presente invención usar otras estructuras de filtro.

Para comprender el efecto, en el dominio del tiempo, del filtro de la Ecuación (1) aplicado en el dominio de la frecuencia, se hace uso de una propiedad de dualidad de las transformadas de Fourier que se aplica en particular a la MDCT. Esta propiedad de dualidad establece que una convolución (o filtrado) de una señal en un dominio es equivalente a una multiplicación (o en realidad, una modulación) de la señal en el otro dominio. Por ejemplo, si el siguiente filtro se aplica a una señal en el dominio del tiempo $x[n]$:

$$y[n] = ax[n] + by[n-1] \quad (2)$$

donde $x[n]$ es la entrada del filtro y $y[n]$ es la salida del filtro, entonces esto es equivalente a multiplicar la transformación de la entrada $x[n]$, que puede notarse $X(e^{j\theta})$, por:

$$H(e^{j\theta}) = \frac{a}{1 - be^{-j\theta}} \quad (3)$$

En la Ecuación (3), θ es la frecuencia normalizada (en radianes por muestra) y $H(e^{j\theta})$ es la función de transferencia del filtro recursivo de la Ecuación (2). Lo que se usa es el valor de $H(e^{j\theta})$ al principio ($\theta = 0$) y al final ($\theta = \pi$) de la escala del dominio de la frecuencia. Es fácil demostrar que, para la Ecuación (3),

$$H(e^{j0}) = \frac{a}{1 - b} \quad (4)$$

$$H(e^{j\pi}) = \frac{a}{1 + b} \quad (5)$$

las Ecuaciones (4) y (5) representan los valores iniciales y finales de la curva descrita por la Ecuación (3). Entre esos dos puntos, la curva evolucionará suavemente entre los valores iniciales y finales. Para la Transformada discreta de Fourier (DFT), que es una transformación de valor complejo, esta curva tendrá valores complejos. Pero para otras transformaciones de valor real tales como la DCT y la MDCT, esta curva exhibirá solo valores reales.

Ahora, debido a la propiedad de dualidad de la transformada de Fourier, si el filtrado de la Ecuación (2) se aplica en el dominio de la frecuencia como en la Ecuación (1), entonces esto tendrá el efecto de multiplicar la señal en el dominio del tiempo por una envolvente suave con valores iniciales y finales como en las Ecuaciones (4) y (5). Esta envolvente en el dominio del tiempo tendrá una forma que podría parecerse a la curva de la Figura 6. Además, si el filtrado en el dominio de la frecuencia como en la Ecuación (1) se aplica solo a una banda espectral, entonces la envolvente en el dominio del tiempo producida solo se relaciona con esa banda espectral. Los otros filtros entre el Filtro inverso 1, el Filtro inverso 2, el Filtro inverso 3, ..., el Filtro inverso M de las Figuras 2 y 3 producirán diferentes envolventes en el dominio del tiempo para las bandas espectrales correspondientes, tales como las que se muestran en la Figura 5.

Se recuerda que estas envolventes en el dominio del tiempo de cada banda espectral se hacen iguales, al principio y al final de un bloque de muestras de la señal de entrada $x[n]$ (por ejemplo, la ventana 1 de la Figura 4), a las ganancias de ruido $g_1[m]$ y $g_2[m]$ calculadas en estos instantes de tiempo. Para la m -ésima banda espectral, la ganancia de ruido al comienzo del bloque de muestras de la señal de entrada $x[n]$ (trama) es $g_1[m]$ y la ganancia de ruido al final del bloque de muestras de la señal de entrada $x[n]$ (trama) es $g_2[m]$. Entre los puntos de inicio (A) y final (B), se hacen las envolventes en el dominio del tiempo (una por banda espectral), más específicamente se interpolan para variar suavemente en el tiempo, de manera que la ganancia de ruido en cada banda espectral evolucione suavemente en la señal en el dominio del tiempo. De esta manera, la forma espectral del ruido de cuantificación evoluciona suavemente en el tiempo, desde el punto A hasta el punto B. Esto se muestra en la Figura 5. La forma espectral punteada en el instante de tiempo C representa la forma espectral instantánea del ruido de cuantificación en algún instante de tiempo entre el comienzo y el final del segmento (puntos A y B).

Para el caso específico del filtro en el dominio de la frecuencia de la Ecuación (1), esto implica las siguientes restricciones para determinar los parámetros a y b en la ecuación de filtro de las ganancias de ruido $g_1[m]$ y $g_2[m]$:

$$g_1[m] = \frac{a}{1 - b} \quad (6)$$

$$g_2[m] = \frac{a}{1+b} \quad (7)$$

5 Para simplificar la notación, establezcamos $g_1 = g_1[m]$ y $g_2 = g_2[m]$, y recordemos que esto es solo para la banda espectral $B_m[k]$. Se obtienen las siguientes relaciones:

$$g_1 = \frac{a}{1-b} \quad (8)$$

$$g_2 = \frac{a}{1+b} \quad (9)$$

15 De las Ecuaciones (8) y (9), es sencillo, para cada Filtro inverso 1, 2, 3, ..., M , calcular los coeficientes del filtro a y b como una función de g_1 y g_2 . Se obtienen las siguientes relaciones:

$$a = -2 \left(\frac{g_1 g_2}{g_1 + g_2} \right) \quad (10)$$

$$b = \frac{g_1 - g_2}{g_1 + g_2} \quad (11)$$

25 Para resumir, los coeficientes a y b en las Ecuaciones (10) y (11) son los coeficientes a usar en el filtrado en el dominio de la frecuencia de la Ecuación (1) para conformar temporalmente el ruido de cuantificación en esa $m^{\text{ésima}}$ banda espectral de manera que siga la envolvente en el dominio del tiempo que se muestra en la Figura 6. En el caso especial de la MDCT usada como la transformada en el procesador de transformación 201 de la Figura 2, los signos de las Ecuaciones (10) y (11) se invierten, es decir, los coeficientes de filtro a usar en la Ecuación (1) se convierten en:

$$a = 2 \left(\frac{g_1 g_2}{g_1 + g_2} \right) \quad (12)$$

$$b = \frac{g_2 - g_1}{g_1 + g_2} \quad (13)$$

40 Esta inversión en el dominio del tiempo de la Cancelación de solapamiento en el dominio del tiempo (TDAC) es específica para el caso especial de la MDCT.

Ahora, el filtrado inverso de la Ecuación (1) conforma tanto el ruido de cuantificación como la señal misma. Para garantizar un proceso reversible, más específicamente para garantizar que $y[n] = x[n]$ en las Figuras 2 y 3 si el ruido de cuantificación es cero, un filtrado a través del Filtro 1, el Filtro 2, el Filtro 3, ..., el Filtro M también se aplica a cada banda espectral $B_m[k]$ antes de la cuantificación en el procesador Q (Figura 2). El Filtro 1, el Filtro 2, el Filtro 3, ..., el Filtro M de la Figura 2 forman prefiltros (es decir, filtros previos a la cuantificación) que en realidad son el "inverso" del Filtro inverso 1, el Filtro inverso 2, el Filtro inverso 3, ..., el Filtro inverso M . En el caso específico de la Ecuación (1) que representa la función de transferencia del Filtro inverso 1, el Filtro inverso 2, el Filtro inverso 3, ..., el Filtro inverso M , los filtros antes de la cuantificación, más específicamente el Filtro 1, el Filtro 2, el Filtro 3, ..., el Filtro M de la Figura 2 se definen por:

$$B_{mf}[k] = aB_m[k] - bB_m[k-1] \quad (14)$$

55 En la Ecuación (14), los coeficientes a y b calculados para los Filtros 1, 2, 3, ..., M son los mismos que en las Ecuaciones (10) y (11), o las Ecuaciones (12) y (13) para el caso especial de la TCMD. La Ecuación (14) describe el inverso del filtro recursivo de la Ecuación (1). Nuevamente, si se usa otro tipo o estructura de filtro diferente al de la Ecuación (1), entonces se usa el inverso de este otro tipo o estructura de filtro en lugar de ese de la Ecuación (14).

60 Otro aspecto es que el concepto puede generalizarse a cualquier forma de ruido de cuantificación en los puntos A y B de las ventanas de la Figura 4, y no se limita a formas de ruido que tengan siempre la misma resolución (el mismo número de bandas espectrales M y el mismo número de coeficientes espectrales $X[k]$ por banda). En la descripción anterior, se supuso que el número M de bandas espectrales $B_m[k]$ es el mismo en las ganancias de ruido $g_1[m]$ y $g_2[m]$, y que cada banda espectral tiene el mismo número de coeficientes de transformación $X[k]$. Pero en realidad, esto puede generalizarse de la siguiente manera: cuando se aplican los filtros en el dominio de la frecuencia como en las Ecuaciones (1) y (14), los coeficientes del filtro (por ejemplo, coeficientes a y b) pueden recalcularse siempre que la ganancia de ruido en un contenedor de frecuencia k cambie en cualquiera de las descripciones de forma de ruido en el punto A o el punto B. Como

un ejemplo, si en el punto A de la Figura 4, la forma de ruido es constante (solo una ganancia para todo el eje de la frecuencia) y en el punto B de la Figura 5 hay tantas ganancias de ruido diferentes como el número N de coeficientes de transformación $X[k]$ (señal de entrada $x[n]$ después de la aplicación de una transformada en el procesador de transformación 201 de la Figura 2). Entonces, al aplicar los filtros del dominio de la frecuencia de las Ecuaciones (1) y (14), los coeficientes de los filtros volverían a calcularse en cada componente de frecuencia, aunque la descripción del ruido en el punto A no cambie sobre todos los coeficientes. Las ganancias de ruido interpoladas de la Figura 5 comenzarían todas desde la misma amplitud (ganancia de ruido constante en el punto A) y convergerían hacia las diferentes ganancias de ruido individuales en las diferentes frecuencias en el punto B.

Tal flexibilidad permite el uso del dispositivo de conformación de ruido en el dominio de la frecuencia 200 y el método 300 para interpolar la forma espectral y la envolvente en el dominio del tiempo del ruido de cuantificación en un sistema en el que la resolución de la forma del ruido espectral cambia con el tiempo. Por ejemplo, en un códec de velocidad de bits variable, puede haber suficientes bits en algunas tramas (punto A o punto B en las Figuras 4 y 5) para refinar la descripción de las ganancias de ruido al agregar más bandas espectrales o cambiar la resolución de la frecuencia para seguir mejor las llamadas bandas espectrales críticas, o el uso de una cuantificación en varias etapas de las ganancias de ruido, y así sucesivamente. Los filtros y los filtros inversos de las Figuras 2 y 3, descritos anteriormente como que operan por banda espectral, en realidad pueden verse como un único filtro (o un único filtro inverso) de un componente de frecuencia a la vez de manera que los coeficientes de los filtros se actualizan siempre que el punto de inicio o el punto final de la envolvente de ruido deseada cambia en una descripción del nivel de ruido.

En la Figura 7 se ilustra un codificador 700 para codificar señales de audio, cuyo principio puede usarse, por ejemplo, en el Códec unificado de voz y audio (USAC) del Grupo de expertos en imágenes en movimiento (MPEG). Más específicamente, el codificador 700 es capaz de conmutar entre un modo de codificación en el dominio de la frecuencia mediante el uso de, por ejemplo, MDCT y un modo de codificación en el dominio del tiempo mediante el uso de, por ejemplo, ACELP. En este ejemplo particular, el codificador 700 comprende: un codificador ACELP que incluye un cuantificador LPC que calcula, codifica y transmite los coeficientes LPC a partir de un análisis LPC; y un codificador basado en transformación que usa un modelo perceptual (o modelo psicoacústico) y factores de escala para conformar el ruido de cuantificación de los coeficientes espectrales. El codificador basado en transformación comprende un dispositivo como se describió anteriormente, para conformar simultáneamente en el dominio del tiempo y el dominio de la frecuencia el ruido de cuantificación del codificador basado en transformación entre dos límites de trama del codificador basado en transformación. En el que las ganancias del ruido de cuantificación pueden describirse ya sea solo por la información de los coeficientes LPC, o solo por la información de los factores de escala, o por cualquier combinación de ambos. Un selector (no mostrado) elige entre el codificador ACELP que usa el modo de codificación en el dominio del tiempo y el codificador basado en transformación que usa el modo de codificación en el dominio de transformación al codificar una ventana de tiempo de la señal de audio, en dependencia, por ejemplo, del tipo de señal de audio a codificar y/o el tipo de modo de codificación que se usará para ese tipo de señal de audio.

Todavía con referencia a la Figura 7, las operaciones de ventanas se aplican primero en el procesador de ventanas 701 a un bloque de muestras de una señal de audio de entrada. De esta manera, se producen versiones en ventanas de la señal de audio de entrada en las salidas del procesador de ventanas 701. Estas versiones en ventanas de la señal de audio de entrada tienen posiblemente longitudes diferentes en dependencia de los procesadores posteriores en los que se usarán como entrada en la Figura 7.

Como se describió anteriormente, el codificador 700 comprende un codificador ACELP que incluye un cuantificador LPC que calcula, codifica y transmite los coeficientes LPC a partir de un análisis LPC. Más específicamente, con referencia a la Figura 7, el codificador ACELP del codificador 700 comprende un analizador LPC 704, un cuantificador LPC 706, una calculadora de objetivos ACELP 708 y un codificador de excitación 712. El analizador LPC 704 procesa una primera versión en ventana de la señal de audio de entrada del procesador 701 para producir los coeficientes LPC. Los coeficientes LPC del analizador LPC 704 se cuantifican en un cuantificador LPC 706 en cualquier dominio adecuado para la cuantificación de esta información. En una trama ACELP, la conformación de ruido se aplica como conocen bien los expertos en la técnica como un filtro en el dominio del tiempo, mediante el uso de un filtro de ponderación derivado del filtro LPC (coeficientes LPC). Esto se realiza en la calculadora de objetivos ACELP 708 y el codificador de excitación 712. Más específicamente, la calculadora 708 usa una segunda versión en ventana de la señal de audio de entrada (mediante el uso de típicamente una ventana rectangular) y produce en respuesta a los coeficientes LPC cuantificados del cuantificador 706 las llamadas señales objetivo en la codificación ACELP. A partir de las señales objetivo producidas por la calculadora 708, el codificador 712 aplica un procedimiento para codificar la excitación del filtro LPC para el bloque actual de muestras de la señal de audio de entrada.

Como se describió anteriormente, el sistema 700 de la Figura 7 comprende además un codificador basado en transformación que usa un modelo perceptual (o modelo psicoacústico) y factores de escala para conformar el ruido de cuantificación de los coeficientes espectrales, en donde el codificador basado en transformación comprende un dispositivo para conformar simultáneamente en el dominio del tiempo y el dominio de la frecuencia el ruido de cuantificación del codificador basado en transformación. El codificador basado en transformación comprende, como se ilustra en la Figura 7, un procesador MDCT 702, un procesador FDNS inverso 707 y un cuantificador de espectro procesado 711, en donde el dispositivo para conformar simultáneamente en el dominio del tiempo y el dominio de la frecuencia el ruido de cuantificación del codificador basado en transformación comprende el procesador FDNS inverso 707. El procesador

MDCT 702 procesa una tercera versión en ventana de la señal de audio de entrada del procesador de ventanas 701 para producir los coeficientes espectrales. El procesador TCMD 702 es un caso específico del procesador más general 201 de la Figura 2 y se entiende que representa la TCMD (Transformada discreta de coseno modificada). Antes de cuantificarse y codificarse (en cualquier dominio adecuado para la cuantificación y codificación de esta información) para su transmisión por el cuantificador 711, los coeficientes espectrales del procesador MDCT 702 se procesan a través del procesador FDNS inverso 707. El funcionamiento del procesador FDNS inverso 707 es como en la Figura 2, que comienza con los coeficientes espectrales $X[k]$ (Figura 2) como entrada al procesador FDNS 707 y termina antes del procesador Q (Figura 2). El procesador FDNS inverso 707 requiere como entrada los conjuntos de ganancias de ruido $g_1[m]$ y $g_2[m]$ como se describió en la Figura 2. Las ganancias de ruido se obtienen del sumador 709, que suma dos entradas: la salida de un cuantificador de factores de escala 705 y la salida de una calculadora de ganancias de ruido 710. Son posibles cualquier combinación de factores de escala, por ejemplo, de un modelo psicoacústico, y ganancias de ruido, por ejemplo, de un modelo LPC, desde el uso de solo factores de escala hasta el uso de solo ganancias de ruido, hasta cualquier combinación o proporción de los factores de escala y las ganancias de ruido. Por ejemplo, los factores de escala del modelo psicoacústico pueden usarse como un segundo conjunto de ganancias o factores de escala para refinar o corregir las ganancias de ruido del modelo LPC. En consecuencia con otra alternativa, la combinación de las ganancias de ruido y los factores de escala comprende la suma de las ganancias de ruido y los factores de escala, donde los factores de escala se usan como una corrección de las ganancias de ruido. Para producir los factores de escala cuantificados en la salida del cuantificador 705, una cuarta versión en ventana de la señal de entrada del procesador 701 se procesa por un analizador psicoacústico 703 que produce factores de escala no cuantificados que luego se cuantifican por el cuantificador 705 en cualquier dominio adecuado para la cuantificación de esta información. De manera similar, para producir las ganancias de ruido en la salida de la calculadora 710, se suministra a una calculadora de ganancias de ruido 710 los coeficientes LPC cuantificados del cuantificador 706. En un bloque de señal de entrada donde el codificador 700 cambiaría entre una trama ACELP y una trama MDCT, el FDNS solo se aplica a las muestras codificadas con MDCT.

El multiplexor de bits 713 recibe como entrada los coeficientes espectrales cuantificados y codificados del cuantificador de espectro procesado 711, los factores de escala cuantificados del cuantificador 705, los coeficientes LPC cuantificados del cuantificador LPC 706 y la excitación codificada del filtro LPC del codificador 712 y produce en respuesta a estos parámetros codificados una cadena de bits para la transmisión o almacenamiento.

En la Figura 8 se ilustra un decodificador 800 que produce un bloque de señal de síntesis mediante el uso de FDNS, en donde el decodificador puede cambiar entre un modo de decodificación en el dominio de la frecuencia mediante el uso de, por ejemplo, IMDCT y un modo de decodificación en el dominio del tiempo mediante el uso de, por ejemplo, ACELP. Un selector (no mostrado) elige entre el decodificador ACELP que usa el modo de decodificación en el dominio del tiempo y el decodificador basado en transformación que usa el modo de codificación en el dominio de transformación cuando decodifica una ventana de tiempo de la señal de audio de codificación, en dependencia del tipo de codificación de esta señal de audio.

El decodificador 800 comprende un demultiplexor 801 que recibe como entrada la cadena de bits del multiplexor de bits 713 (Figura 7). La cadena de bits recibida se demultiplexa para recuperar los coeficientes espectrales cuantificados y codificados del cuantificador de espectro procesado 711, los factores de escala cuantificados del cuantificador 705, los coeficientes LPC cuantificados del cuantificador LPC 706 y la excitación codificada del filtro LPC del codificador 712.

Los coeficientes LPC cuantificados recuperados (ventana codificada por transformación de la señal de audio en ventana) del demultiplexor 801 se suministran a un decodificador LPC 804 para producir los coeficientes LPC decodificados. La excitación codificada recuperada del filtro LPC del demultiplexor 301 se suministra a y decodificada por un decodificador de excitación ACELP 805. Un filtro de síntesis ACELP 806 responde a los coeficientes LPC decodificados del decodificador 804 y a la excitación decodificada del decodificador 805 para producir una señal de audio decodificada por ACELP.

Los factores de escala cuantificados recuperados se suministran a y decodifican por un decodificador de factores de escala 803.

Los coeficientes espectrales cuantificados y codificados recuperados se suministran a un decodificador de coeficientes espectrales 802. El decodificador 802 produce los coeficientes espectrales decodificados que se usan como entrada por un procesador FDNS 807. El funcionamiento del procesador FDNS 807 como se describió en la Figura 2, comienza después del procesador Q y termina antes del procesador 204 (procesador de transformación inversa). El procesador FDNS 807 se suministra con los coeficientes espectrales decodificados del decodificador 802, y una salida del sumador 808 que produce conjuntos de ganancias de ruido, por ejemplo los conjuntos de ganancias de ruido descritos anteriormente $g_1[m]$ y $g_2[m]$ resultantes de la suma de los factores de escala decodificados del decodificador 803 y las ganancias de ruido calculadas por la calculadora 809. La calculadora 809 calcula las ganancias de ruido de los coeficientes LPC decodificados producidos por el decodificador 804. Como en el codificador 700 (Figura 7), son posibles cualquier combinación de factores de escala (de un modelo psicoacústico) y ganancias de ruido (de un modelo LPC), desde el uso de solo factores de escala hasta el uso de solo ganancias de ruido, hasta cualquier proporción de factores de escala y ganancias de ruido. Por ejemplo, los factores de escala del modelo psicoacústico pueden usarse como un segundo conjunto de ganancias o factores de escala para refinar o corregir las ganancias de ruido del modelo LPC. En consecuencia con otra alternativa, la combinación de las ganancias de ruido y los factores de escala comprende la suma

de las ganancias de ruido y los factores de escala, donde los factores de escala se usan como una corrección de las ganancias de ruido. Los coeficientes espectrales resultantes en la salida del procesador FDNS 807 se someten a un procesador IMDCT 810 para producir una señal de audio decodificada por transformación.

- 5 Finalmente, un procesador de ventanas y superposición/adición 811 combina la señal de audio decodificada por ACELP del filtro de síntesis ACELP 806 con la señal de audio decodificada por transformación del procesador IMDCT 810 para producir una señal de audio de síntesis.

REIVINDICACIONES

1. Un método de conformación de ruido en el dominio de la frecuencia para interpolar una forma espectral y una envolvente en el dominio del tiempo del ruido de cuantificación en una señal de audio en ventana y codificada por transformación, caracterizado porque comprende:
- procesar (305) las bandas espectrales cuantificadas $C_{1f}[k]$, $C_{2f}[k]$, $C_{3f}[k]$, ..., $C_{Mf}[k]$ de la señal de audio en ventana y codificada por transformación a través de los filtros inversos respectivos (*Filtro inverso 1, Filtro inverso 2, Filtro inverso 3, ..., Filtro inverso M*) para producir las bandas espectrales decodificadas $C_1[k]$, $C_2[k]$, $C_3[k]$, ..., $C_M[k]$; concatenar (306) las bandas espectrales decodificadas $C_1[k]$, $C_2[k]$, $C_3[k]$, ..., $C_M[k]$ para producir los coeficientes espectrales decodificados ($Y[k]$); y transformación inversa (307) de los coeficientes espectrales decodificados ($Y[k]$) para producir un bloque decodificado de muestras en el dominio del tiempo ($y[n]$) de la señal de audio;
- en donde el procesamiento (305) de las bandas espectrales cuantificadas $C_{1f}[k]$, $C_{2f}[k]$, $C_{3f}[k]$, ..., $C_{Mf}[k]$ comprende, para cada banda espectral cuantificada $C_{1f}[k]$, $C_{2f}[k]$, $C_{3f}[k]$, ..., $C_{Mf}[k]$:
- calcular (308) las ganancias de ruido $g_1[m]$ y $g_2[m]$ que representan las formas espectrales del ruido de cuantificación respectivamente en una primera transición (A) entre una ventana de procesamiento de transformación actual (ventana 1) y una ventana de procesamiento de transformación anterior (ventana 0) y en una segunda transición (B) entre la ventana de procesamiento de transformación actual (ventana 1) y una ventana de procesamiento de transformación posterior (ventana 2); y filtrar los coeficientes espectrales cuantificados ($Y_f[k]$) de la banda espectral cuantificada mediante el uso de la relación:

$$C_m[k] = aC_{mf}[k] + bC_m[k-1]$$

donde a y b son los parámetros del filtro y m identifica la banda espectral, y donde

$$a = 2 ((g_1[m]g_2[m]) / (g_1[m] + g_2[m]))$$

$$b = ((g_2[m] - g_1[m]) / (g_1[m] + g_2[m])).$$

2. Un dispositivo de conformación de ruido en el dominio de la frecuencia para interpolar una forma espectral y una envolvente en el dominio del tiempo del ruido de cuantificación en una señal de audio en ventana y codificada por transformación, caracterizado porque comprende:
- medios para procesar las bandas espectrales cuantificadas $C_{1f}[k]$, $C_{2f}[k]$, $C_{3f}[k]$, ..., $C_{Mf}[k]$ de la señal de audio en ventana y codificada por transformación a través de los filtros inversos respectivos (*Filtro inverso 1, Filtro inverso 2, Filtro inverso 3, ..., Filtro inverso M*) para producir las bandas espectrales decodificadas $C_1[k]$, $C_2[k]$, $C_3[k]$, ..., $C_M[k]$; medios (203) para concatenar las bandas espectrales decodificadas $C_1[k]$, $C_2[k]$, $C_3[k]$, ..., $C_M[k]$ para producir los coeficientes espectrales decodificados ($Y[k]$); y medios (204) para la transformación inversa de los coeficientes espectrales decodificados ($Y[k]$) para producir un bloque decodificado de muestras en el dominio del tiempo ($y[n]$) de la señal de audio;
- en donde los medios para procesar las bandas espectrales cuantificadas $C_{1f}[k]$, $C_{2f}[k]$, $C_{3f}[k]$, ..., $C_{Mf}[k]$ comprenden, para cada banda espectral cuantificada $C_{1f}[k]$, $C_{2f}[k]$, $C_{3f}[k]$, ..., $C_{Mf}[k]$:
- medios (205) para calcular las ganancias de ruido $g_1[m]$ y $g_2[m]$ que representan las formas espectrales del ruido de cuantificación respectivamente en una primera transición (A) entre una ventana de procesamiento de transformación actual (ventana 1) y una ventana de procesamiento de transformación anterior (ventana 0) y en una segunda transición (B) entre la ventana de procesamiento de transformación actual (ventana 1) y una ventana de procesamiento de transformación posterior (ventana 2); y medios para filtrar los coeficientes espectrales cuantificados ($Y_f[k]$) de la banda espectral cuantificada mediante el uso de la relación:

$$C_m[k] = aC_{mf}[k] + bC_m[k-1]$$

donde a y b son los parámetros del filtro y m identifica la banda espectral, y donde

$$a = 2 ((g_1[m]g_2[m]) / (g_1[m] + g_2[m]))$$

$$b = ((g_2[m] - g_1[m]) / (g_1[m] + g_2[m])).$$

Figura 1

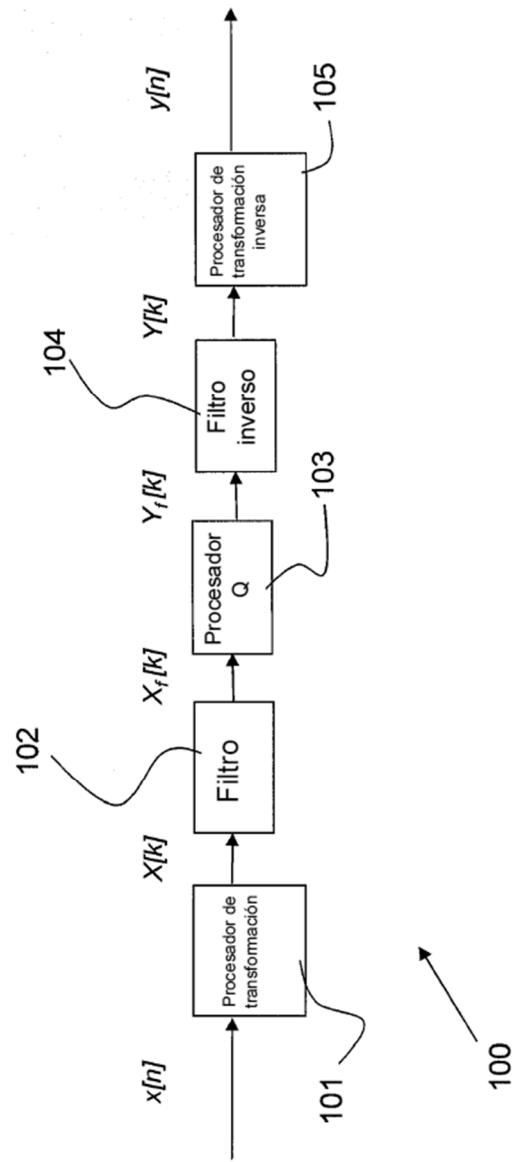


Figura 2

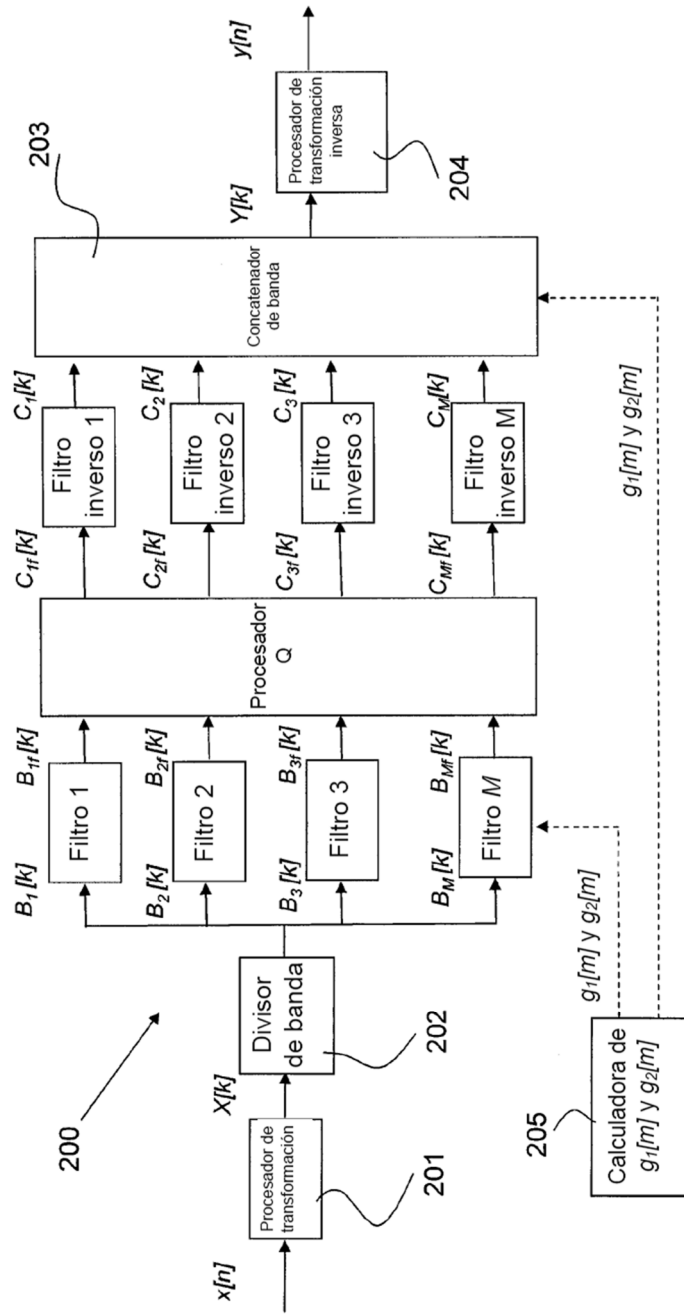


Figura 3

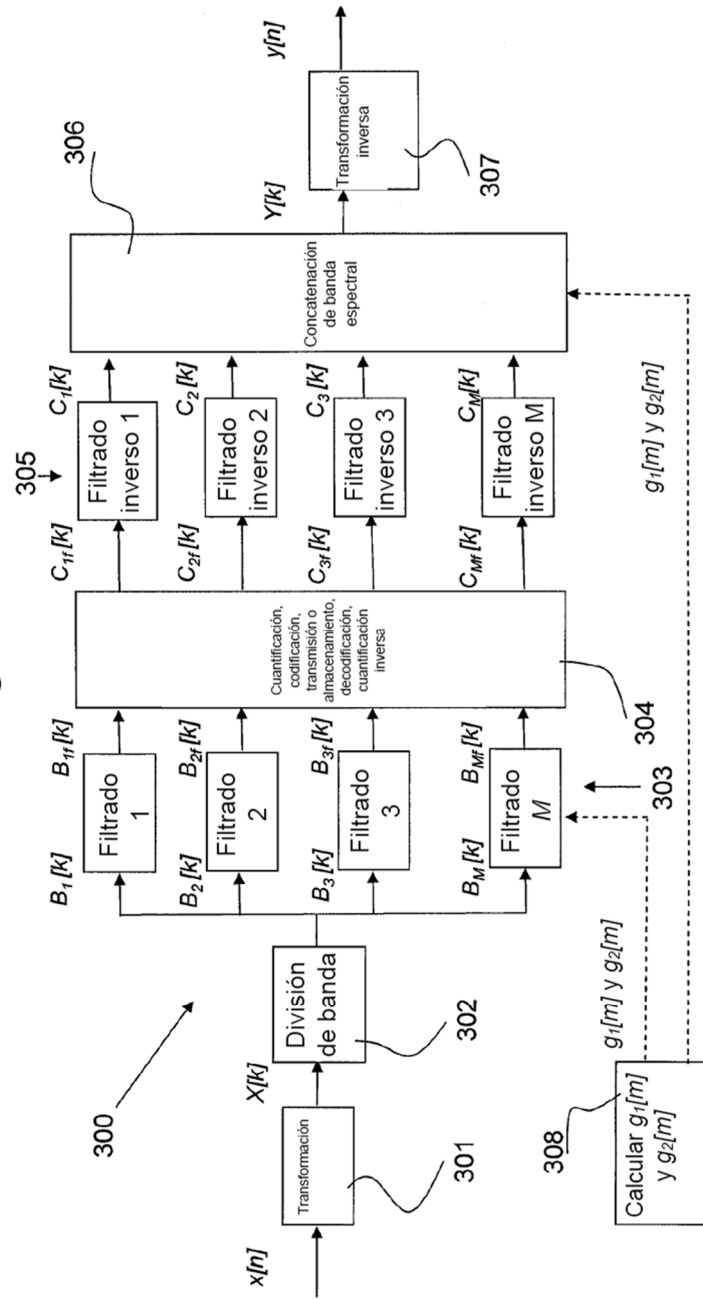


Figura 4

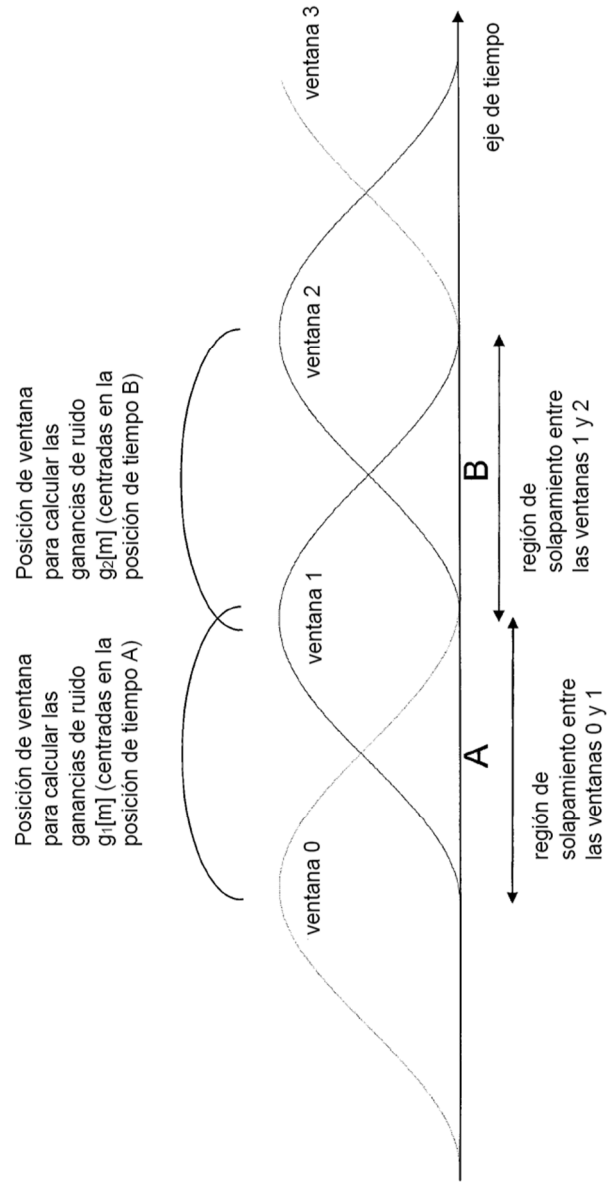


Figura 5

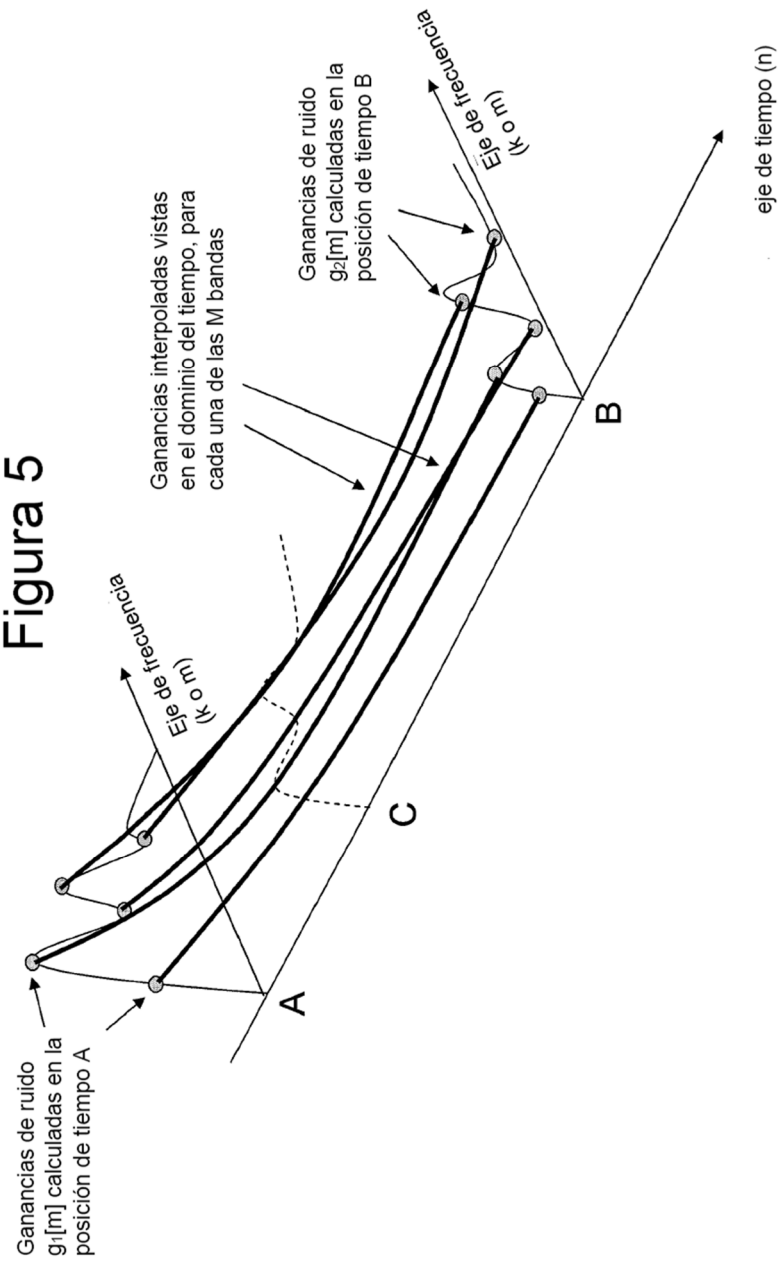


Figura 6

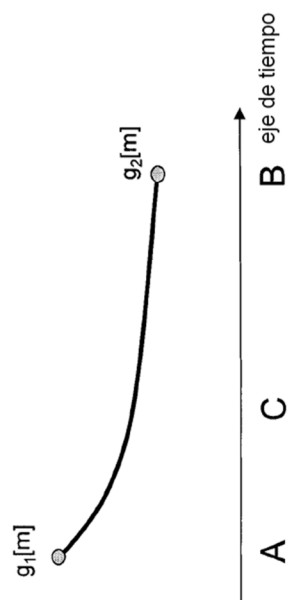


Figura 7

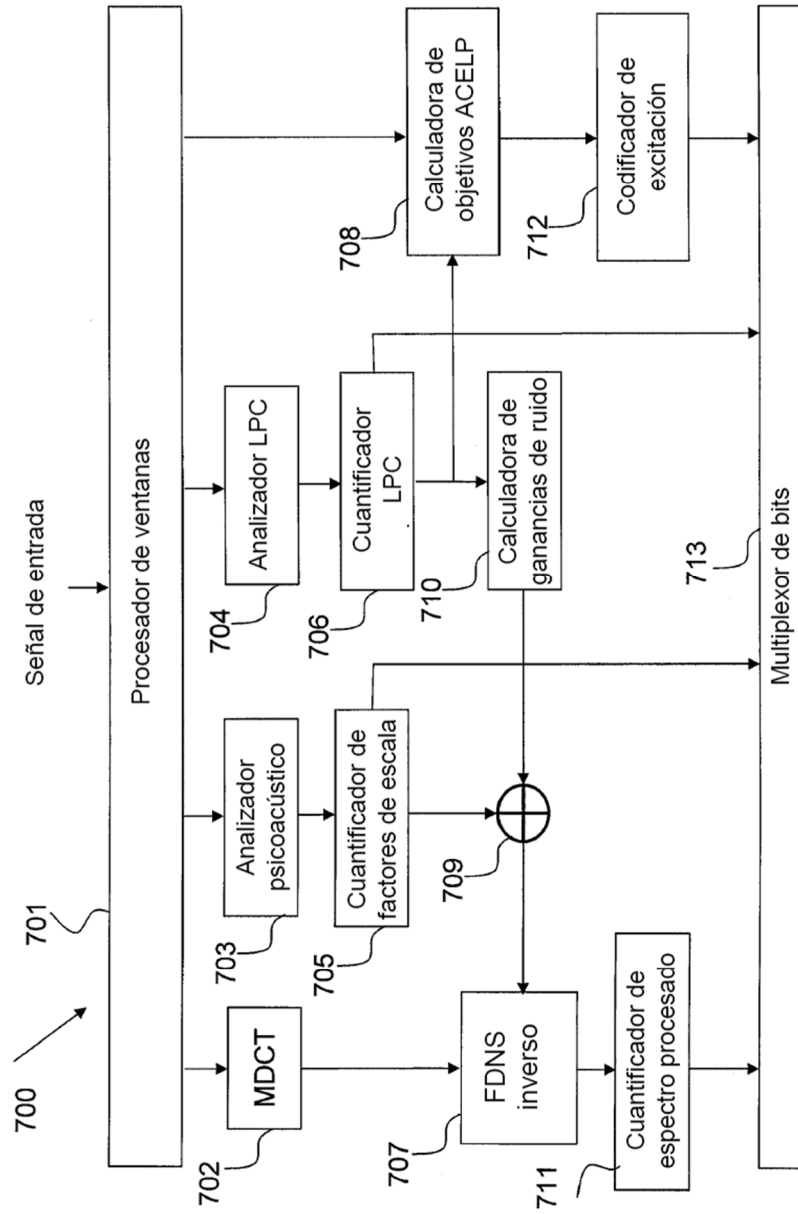


Figura 8

