

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 671 711**

51 Int. Cl.:

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **18.09.2009 PCT/KR2009/005340**

87 Fecha y número de publicación internacional: **25.03.2010 WO10032992**

96 Fecha de presentación y número de la solicitud europea: **18.09.2009 E 09814808 (3)**

97 Fecha y número de publicación de la concesión europea: **21.03.2018 EP 2339577**

54 Título: **Aparato de codificación y aparato de decodificación para transformar entre codificador basado en transformada de coseno discreta modificada y hetero codificador**

30 Prioridad:

18.09.2008 KR 20080091697

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
08.06.2018

73 Titular/es:

**ELECTRONICS AND TELECOMMUNICATIONS
RESEARCH INSTITUTE (50.0%)
161 Gajeong-dong
Yuseong-gu, Daejeon 305-350, KR y
KWANGWOON UNIVERSITY INDUSTRY-
ACADEMIC COLLABORATION FOUNDATION
(50.0%)**

72 Inventor/es:

**BEACK, SEUNG KWON;
LEE, TAE JIN;
KIM, MIN JE;
JANG, DAE YOUNG;
KANG, KYEONGOK;
HONG, JIN WOO;
PARK, HO CHONG y
PARK, YOUNG-CHEOL**

74 Agente/Representante:

SUGRAÑES MOLINÉ, Pedro

ES 2 671 711 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

Aparato de codificación y aparato de decodificación para transformar entre codificador basado en transformada de coseno discreta modificada y hetero codificador

5

Campo técnico

La presente invención se refiere a un aparato y un método para reducir un artefacto, generado cuando se realiza la transformada entre diferentes tipos de codificadores, cuando se codifica y decodifica una señal de audio combinando un codificador de audio basado en la Transformada de Coseno Discreta Modificada (MDCT) y un codificador de voz/audio diferente.

10

Antecedentes de la técnica

Cuando se aplica de manera diferente un método de codificación/decodificación a una señal de entrada donde se combina una voz y audio dependiendo de una característica de la señal de entrada, puede mejorarse un rendimiento y una calidad de sonido. Por ejemplo, puede aplicarse de manera eficaz un codificador basado en Predicción Lineal con Excitación por Código (CELP) a una señal que tiene una característica similar a una señal de voz, y aplicar un codificador basado en conversión de frecuencia a una señal idéntica a una señal de audio.

15

20

Una Codificación Unificada de Voz y Audio (USAC) puede desarrollarse aplicando los conceptos anteriormente descritos. La USAC puede recibir continuamente una señal de entrada y analizar una característica de la señal de entrada a tiempos particulares. A continuación, la USAC puede codificar la señal de entrada aplicando diferentes tipos de aparatos de codificación a través de conmutación dependiendo de la característica de la señal de entrada.

25

Un artefacto de señal puede generarse durante conmutación de señal en la USAC. Puesto que la USAC codifica una señal de entrada para cada bloque, un artefacto de bloqueo puede generarse cuando se aplican diferentes tipos de codificaciones. Para superar una desventaja de este tipo, la USAC puede realizar una operación de superposición-adición aplicando una ventana a bloques donde se aplican diferentes codificaciones. Sin embargo, puede requerirse información de flujo de bits adicional debido a la superposición, y cuando tiene lugar frecuentemente la conmutación, puede aumentar un flujo de bits adicional para eliminar el artefacto de bloqueo. Cuando aumenta un flujo de bits, puede reducirse una eficacia de codificación.

30

En particular, la USAC puede codificar una señal característica de audio usando un aparato de codificación basado en la Transformada de Coseno Discreta Modificada (MDCT). Un esquema de MDCT puede transformar una señal de entrada de un dominio de tiempo en una señal de entrada de un dominio de frecuencia, y realizar una operación de superposición-adición entre los bloques. En un esquema de MDCT, puede generarse solapando en un dominio de tiempo, mientras que una tasa de bits puede no aumentar incluso cuando se realiza una operación de superposición-adición.

35

40

En este caso, se ha de realizar una operación de superposición-adición del 50 % con un bloque vecino para restaurar una señal de entrada basándose en un esquema de MDCT. Es decir, un bloque actual a emitirse puede decodificarse dependiendo de un resultado de salida de un bloque anterior. Sin embargo, cuando no se decodifica el bloque anterior usando la USAC usando un esquema de MDCT, el bloque actual, codificado usando el esquema de MDCT, no puede decodificarse a través de una operación de superposición-adición puesto que la información de MDCT del bloque anterior no puede usarse. Por consiguiente, la USAC puede requerir adicionalmente la información de MDCT del bloque anterior, cuando se codifica un bloque actual usando un esquema de MDCT después de la conmutación.

45

Cuando la conmutación tiene lugar frecuentemente, la información de MDCT adicional para decodificar puede aumentarse en proporción al número de conmutaciones. En este caso, una tasa de bits puede aumentar debido a la información de MDCT adicional, y una eficacia de codificación puede reducirse de manera significativa. Por consiguiente, se requiere un método que pueda eliminar el artefacto de bloqueo y reducir la información de MDCT adicional durante la conmutación.

50

55

El documento US 2003/009325 A1 describe un método para conmutación controlada de señal entre esquemas de codificación que incluye recibir señales de audio de entrada, clasificar un primer conjunto de las señales de audio de entrada como señales de voz o no de voz, codificar las señales de voz usando un esquema de codificación de dominio de tiempo, y codificar las señales no de voz usando un esquema de codificación de transformada. Un codificador de múltiple código tiene una entrada de señal de audio y un conmutador para recibir las entradas de señal de audio, teniendo la conmutación un codificador de dominio de tiempo, codificador de transformada y un clasificador de señal para clasificar las señales de audio generalmente como de voz o no de voz, dirigiendo el clasificador de señal señales de audio de voz al codificador de dominio del tiempo y señales de audio no de voz al codificador de transformada. Se proporciona también un decodificador de código múltiple.

60

65

Divulgación de la invención

Objetivos técnicos

5 La invención se define en las reivindicaciones independientes 1 y 4. Se definen realizaciones específicas en las reivindicaciones dependientes. Un aspecto proporciona un método y aparato de codificación y un método y aparato de decodificación que pueden eliminar un artefacto de bloqueo y reducir la información de MDCT requerida.

10 De acuerdo con un aspecto, se proporciona una primera unidad de codificación para codificar una señal característica de voz de una señal de entrada de acuerdo con un esquema de hetero codificación diferente de un esquema de codificación basado en la Transformada de Coseno Discreta Modificada (MDCT); y una segunda unidad de codificación para codificar una señal característica de audio de la señal de entrada de acuerdo con el esquema de codificación basado en MDCT. La segunda unidad de codificación puede realizar codificación aplicando una ventana de análisis que no supera un punto de pliegue, cuando tiene lugar el punto de pliegue donde existe la conmutación entre la señal característica de voz y la señal característica de audio en una trama actual de la señal de entrada. En este punto, el punto de pliegue puede ser un área donde las señales de solapamiento se pliegan cuando se realiza una MDCT y una MDCT inversa (IMDCT). Cuando se realiza una MDCT de N puntos, el punto de pliegue puede localizarse en un punto de $N/4$ y $3N/4$. El punto de pliegue puede ser una cualquiera de las características bien conocidas asociadas con una MDCT, y no se describe una base matemática para el punto de pliegue en el presente documento. También, un concepto de la MDCT y el punto de pliegue se describen en detalle con referencia a la Figura 5.

25 También, para facilidad de descripción, cuando una señal de trama anterior es una señal característica de voz y una señal de trama actual es una señal característica de audio, el punto de pliegue, usado cuando se conectan los dos diferentes tipos de señales características, puede denominarse en lo sucesivo como un 'punto de pliegue donde tiene lugar la conmutación'. También, cuando una señal de trama posterior es una señal característica de voz, y una señal de trama actual es una señal característica de audio, el punto de pliegue usado cuando se conectan los dos diferentes tipos de señales características, puede denominarse como un 'punto de pliegue donde tiene lugar la conmutación'.

Soluciones técnicas

35 De acuerdo con un aspecto, se proporciona un aparato de codificación, que incluye: una unidad de procesamiento de ventana para aplicar una ventana de análisis a una trama actual de una señal de entrada; una unidad de MDCT para realizar una MDCT con respecto a la trama actual donde se aplica la ventana de análisis; una unidad de generación de flujo de bits para codificar la trama actual y para generar un flujo de bits de la señal de entrada. La unidad de procesamiento de ventana puede aplicar una ventana de análisis que no supera un punto de pliegue, cuando el punto de pliegue tiene lugar donde existe la conmutación entre una señal característica de voz y una señal característica de audio en la trama actual de la señal de entrada.

40 De acuerdo con un aspecto, se proporciona un aparato de decodificación, que incluye: una primera unidad de decodificación para decodificar una señal característica de voz de una señal de entrada codificada de acuerdo con un esquema de hetero codificación diferente de un esquema de codificación basado en MDCT; una segunda unidad de codificación para decodificar una señal característica de audio de la señal de entrada codificada de acuerdo con el esquema de codificación basado en MDCT; y una unidad de compensación de bloque para realizar compensación de bloque con respecto a un resultado de la primera unidad de decodificación y un resultado de la segunda unidad de decodificación, y para restaurar la señal de entrada. La unidad de compensación de bloque puede aplicar una ventana de síntesis que no supera un punto de pliegue, cuando el punto de pliegue tiene lugar donde existe la conmutación entre la señal característica de voz y la señal característica de audio en una trama actual de la señal de entrada.

50 De acuerdo con un aspecto, se proporciona un aparato de decodificación, que incluye: una unidad de compensación de bloque para aplicar una ventana de síntesis a información adicional extraída desde una señal característica de voz y una trama actual y para restaurar una señal de entrada, cuando tiene lugar un punto de pliegue donde existe conmutación entre la señal característica de voz y la señal característica de audio en la trama actual de la señal de entrada.

Efectos ventajosos

60 De acuerdo con un aspecto, se proporciona un aparato y un método de codificación y un aparato y un método de decodificación que pueden reducir información de MDCT adicional requerida cuando tiene lugar la codificación entre diferentes tipos de codificadores dependiendo de una característica de una señal de entrada, y eliminar un artefacto de bloqueo.

65 También, de acuerdo con un aspecto, se proporciona un aparato y un método de codificación y un aparato y un método de decodificación que pueden reducir información de MDCT adicional requerida cuando tiene lugar la

codificación entre diferentes tipos de codificadores, y puede evitar de esta manera que un flujo de bits aumente y mejore una eficacia de codificación.

Breve descripción de los dibujos

- 5 La Figura 1 es un diagrama de bloques que ilustra un aparato de codificación y un aparato de decodificación de acuerdo con una realización de la presente invención;
 La Figura 2 es un diagrama de bloques que ilustra una configuración de un aparato de codificación de acuerdo con una realización de la presente invención;
 10 La Figura 3 es un diagrama que ilustra una operación de codificación de una señal de entrada a través de una segunda unidad de codificación de acuerdo con una realización de la presente invención;
 La Figura 4 es un diagrama que ilustra una operación de codificación de una señal de entrada a través de procesamiento de ventana de acuerdo con una realización de la presente invención;
 15 La Figura 5 es un diagrama que ilustra una operación de Transformada de Coseno Discreta Modificada (MDCT) de acuerdo con una realización de la presente invención;
 La Figura 6 es un diagrama que ilustra una operación de hetero codificación (C1, C2) de acuerdo con una realización de la presente invención;
 La Figura 7 es un diagrama que ilustra una operación de generación de un flujo de bits en una C1 de acuerdo con una realización de la presente invención;
 20 La Figura 8 es un diagrama que ilustra una operación de codificación de una señal de entrada a través de procesamiento de ventana en una C1 de acuerdo con una realización de la presente invención;
 La Figura 9 es un diagrama que ilustra una operación de generación de un flujo de bits en una C2 de acuerdo con una realización de la presente invención;
 25 La Figura 10 es un diagrama que ilustra una operación de codificación de una señal de entrada a través de procesamiento de ventana en una C2 de acuerdo con una realización de la presente invención;
 La Figura 11 es un diagrama que ilustra información adicional aplicada cuando se codifica la señal de entrada de acuerdo con una realización de la presente invención;
 La Figura 12 es un diagrama de bloques que ilustra una configuración de un aparato de decodificación de acuerdo con una realización de la presente invención;
 30 La Figura 13 es un diagrama que ilustra una operación de decodificación de un flujo de bits a través de una segunda unidad de codificación de acuerdo con una realización de la presente invención;
 La Figura 14 es un diagrama que ilustra una operación de extracción de una señal de salida a través de una operación de superposición-adición de acuerdo con una realización de la presente invención;
 35 La Figura 15 es un diagrama que ilustra una operación de generación de una señal de salida en una C1 de acuerdo con una realización de la presente invención;
 La Figura 16 es un diagrama que ilustra una operación de compensación de bloque en una C1 de acuerdo con una realización de la presente invención;
 La Figura 17 es un diagrama que ilustra una operación de generación de una señal de salida en una C2 de acuerdo con una realización de la presente invención; y
 40 La Figura 18 es un diagrama que ilustra una operación de compensación de bloque en una C2 de acuerdo con una realización de la presente invención.

Mejor modo para llevar a cabo la invención

- 45 Se hará ahora referencia en detalle a las realizaciones de la presente invención, ejemplos de las cuales se ilustran en los dibujos adjuntos, en las que números de referencia similares hacen referencia a elementos similares a lo largo de todo el documento. Las realizaciones se describen a continuación para explicar la presente invención haciendo referencia a las figuras.
- 50 La Figura 1 es un diagrama de bloques que ilustra un aparato de codificación 101 y un aparato de decodificación 102 de acuerdo con una realización de la presente invención.
- El aparato de codificación 101 puede generar un flujo de bits mediante codificación de una señal de entrada para cada bloque. En este caso, el aparato de codificación 101 puede codificar una señal característica de voz y una
 55 señal característica de audio. La señal característica de voz puede tener una característica similar a una señal de voz, y la señal característica de audio puede tener una característica similar a una señal de audio. El flujo de bits con respecto a una señal de entrada puede generarse como resultado de la codificación, y transmitirse al aparato de decodificación 102. El aparato de decodificación 101 puede generar una señal de salida decodificando el flujo de bits, y de esta manera puede restaurar la señal de entrada codificada.
- 60 Específicamente, el aparato de codificación 101 puede analizar un estado de la señal introducida de manera continua, y conmutar para posibilitar que se aplique un esquema de codificación que corresponde a la característica de la señal de entrada de acuerdo con un resultado del análisis. Por consiguiente, el aparato de codificación 101 puede codificar bloques donde se aplica un esquema de hetero codificación. Por ejemplo, el aparato de codificación
 65 101 puede codificar la señal característica de voz de acuerdo con un esquema de Predicción Lineal con Excitación por Código (CELP), y codificar la señal característica de audio de acuerdo con un esquema de Transformada de

Coseno Discreta Modificada (MDCT). A la inversa, el aparato de decodificación 102 puede restaurar la señal de entrada decodificando la señal de entrada, codificada de acuerdo con el esquema de CELP, de acuerdo con el esquema de CELP y decodificando la señal de entrada, codificada de acuerdo con el esquema de MDCT, de acuerdo con el esquema de MDCT.

5 En este caso, cuando la señal de entrada se conmuta a la señal característica de audio desde la señal característica de voz, el aparato de codificación 101 puede codificar conmutando desde el esquema de CELP al esquema de MDCT. Puesto que la codificación se realiza para cada bloque, puede generarse artefacto de bloqueo. En este caso, el aparato de decodificación 102 puede eliminar el artefacto de bloqueo a través de una operación de superposición-
10 adición entre los bloques.

También, cuando se codifica un bloque actual de la señal de entrada de acuerdo con el esquema de MDCT, se requiere información de MDCT de un bloque anterior para restaurar la señal de entrada, sin embargo, cuando el bloque anterior se codifica de acuerdo con el esquema de CELP, puesto que la información de MDCT del bloque anterior no existe, el bloque actual no puede restaurarse de acuerdo con el esquema de MDCT. Por consiguiente, se requiere información de MDCT adicional del bloque anterior. También, el aparato de codificación 101 puede reducir la información de MDCT adicional, y de esta manera puede evitar que aumente una tasa de bits.

La Figura 2 es un diagrama de bloques que ilustra una configuración de un aparato de codificación 101 de acuerdo con una realización de la presente invención.

Haciendo referencia a la Figura 2, el aparato de codificación 101 puede incluir una unidad de retardo de bloque 201, una unidad de análisis de estado 202, una unidad de corte de señal 203, una primera unidad de codificación 204 y una segunda unidad de codificación 205.

25 La unidad de retardo de bloque 201 puede retardar una señal de entrada para cada bloque. La señal de entrada puede procesarse para cada bloque para codificación. La unidad de retardo de bloque 201 puede retardar hacia atrás (-) o retardar hacia adelante (+) del bloque actual introducido.

30 La unidad de análisis de estado 202 puede determinar una característica de la señal de entrada. Por ejemplo, la unidad de análisis de estado 202 puede determinar si la señal de entrada es una señal característica de voz o una señal característica de audio. En este caso, la unidad de análisis de estado 202 puede emitir un parámetro de control. El parámetro de control puede usarse para determinar qué esquema de codificación se usa para codificar el bloque actual de la señal de entrada.

35 Por ejemplo, la unidad de análisis de estado 202 puede analizar la característica de la señal de entrada, y determinar, como la señal característica de voz, un periodo de señal que corresponde a (1) un estado armónico estable (SH) que muestra un componente armónico evidente y estable, (2) un estado armónico estable bajo (LSH) que muestra una característica estable intensa en un ancho de banda de baja frecuencia y que muestra un componente armónico de un periodo relativamente largo, y (3) un estado de ruido estable (SN) que es un estado de ruido blanco. También, la unidad de análisis de estado 202 puede analizar la característica de la señal de entrada, y determinar, como la señal característica de audio, un periodo de señal que corresponde a (4) un estado armónico complejo (CH) que muestra una estructura armónica compleja donde se combinan diversos componentes de tono, y (5) un estado con ruido complejo (CN) que incluye componentes de ruido inestables. En este punto, el periodo de
45 señal puede corresponder a una unidad de bloque de la señal de entrada.

La unidad de corte de señal 203 puede posibilitar que la señal de entrada de la unidad de bloque sea un subconjunto.

50 La primera unidad de codificación 204 puede codificar la señal característica de voz de entre señales de entrada de la unidad de bloque. Por ejemplo, la primera unidad de codificación 204 puede codificar la señal característica de voz en un dominio de tiempo de acuerdo con una codificación predictiva lineal (LPC). En este caso, la primera unidad de codificación 204 puede codificar la señal característica de voz de acuerdo con un esquema de codificación basado en CELP. Aunque se ilustra una única primera unidad de codificación 204 en la Figura 3, puede configurarse una o más primeras unidades de codificación.

La segunda unidad de codificación 205 puede codificar la señal característica de audio de entre las señales de entrada de la unidad de bloque. Por ejemplo, la segunda unidad de codificación 205 puede transformar la señal característica de audio desde el dominio del tiempo al dominio de la frecuencia para realizar codificación. En este caso, la segunda unidad de codificación 205 puede codificar la señal característica de audio de acuerdo con un esquema de codificación basado en MDCT. Un resultado de la primera unidad de decodificación 204 y un resultado de la segunda unidad de codificación 205 puede generarse en un flujo de bits, y el flujo de bits generado en cada una de las unidades de codificación puede controlarse para que sea un único flujo de bits a través de un multiplexor de flujo de bits (MUX).

65 Es decir, el aparato de codificación 101 puede codificar la señal de entrada a través de una cualquiera de la primera

unidad de codificación 204 y la segunda unidad de codificación 205, combinando dependiendo de un parámetro de control de la unidad de análisis de estado 202. También, la primera unidad de codificación 204 puede codificar la señal característica de voz de la señal de entrada de acuerdo con el esquema de hetero codificación diferente del esquema de codificación basado en MDCT. También, la segunda unidad de codificación 205 puede codificar la señal característica de audio de la señal de entrada de acuerdo con el esquema de codificación basado en MDCT.

La Figura 3 es un diagrama que ilustra una operación de codificación de una señal de entrada a través de una segunda unidad de codificación 205 de acuerdo con una realización de la presente invención.

Haciendo referencia a la Figura 3, la segunda unidad de codificación 205 puede incluir una unidad de procesamiento de ventana 301, una unidad de MDCT 302 y una unidad de generación de flujo de bits 303.

En la Figura 3, X(b) puede indicar una unidad de bloque básica de la señal de entrada. La señal de entrada se describe en detalle con referencia a la Figura 4 y la Figura 6. La señal de entrada puede introducirse a la unidad de procesamiento de ventana 301, y también puede introducirse a la unidad de procesamiento de ventana 301 a través de la unidad de retardo de bloque 201.

La unidad de procesamiento de ventana 301 puede aplicar una ventana de análisis a una trama actual de la señal de entrada. Específicamente, la unidad de procesamiento de ventana 301 puede aplicar la ventana de análisis a un bloque actual X(b) y un bloque retardado X(b-2). El bloque actual X(b) puede retardarse hacia atrás al bloque anterior X(b-2) a través de la unidad de retardo de bloque 201.

Por ejemplo, la unidad de procesamiento de ventana 301 puede aplicar una ventana de análisis, que no supera un punto de pliegue, hasta la trama actual, cuando tiene lugar un punto de pliegue cuando existe la conmutación entre una señal característica de voz y una señal característica de audio en la trama actual. En este caso, la unidad de procesamiento de ventana 301 puede aplicar la ventana de análisis que está configurada como una ventana que tiene un valor de 0 y corresponde a un primer subbloque, una ventana que corresponde a un área de información adicional de un segundo subbloque, y una ventana que tiene un valor de 1 y corresponde a un área restante del segundo subbloque basándose en el punto de pliegue. En este punto, el primer subbloque puede indicar la señal característica de voz, y el segundo subbloque puede indicar la señal característica de audio.

Un grado de retardo de bloque realizado por la unidad de retardo de bloque 201, puede variar dependiendo de una unidad de bloque de la señal de entrada. Cuando la señal de entrada pasa a través de la unidad de procesamiento de ventana 301, puede aplicarse la ventana de análisis, y por lo tanto puede extraerse $\{X(b-2), X(b)\} \otimes W_{\text{análisis}}$. Por consiguiente, la unidad de MDCT 302 puede realizar una MDCT con respecto a la trama actual donde se aplica la ventana de análisis. También, la unidad de generación de flujo de bits 303 puede codificar la trama actual y generar un flujo de bits de la señal de entrada.

La Figura 4 es un diagrama que ilustra una operación de codificación de una señal de entrada a través de procesamiento de ventana de acuerdo con una realización de la presente invención.

Haciendo referencia a la Figura 4, la unidad de procesamiento de ventana 301 puede aplicar la ventana de análisis a la señal de entrada, en este caso, la ventana de análisis puede ser en forma de un rectángulo o un seno. Una forma de la ventana de análisis puede variar dependiendo de la señal de entrada.

Cuando se introduce el bloque actual X(b), la unidad de procesamiento de ventana 301 puede aplicar la ventana de análisis al bloque actual X(b) y al bloque anterior X(b-2). En este punto, el bloque anterior X(b-2) puede retardarse hacia atrás por la unidad de retardo de bloque 102. Por ejemplo, el bloque X(b) puede establecerse como una unidad básica de la señal de entrada de acuerdo con la Ecuación 1 dada a continuación. En este caso, pueden establecerse dos bloques como una única trama y codificarse.

[Ecuación 1]

$$X(b) = [s(b-1), s(b)]^T$$

En este caso, s(b) puede indicar un subbloque que configura un único bloque, y puede definirse por,

[Ecuación 2]

$$s(b) = [s((b-1) \cdot N/4), s((b-1) \cdot N/4 + 1), \dots, s((b-1) \cdot N/4 + N/4 - 1)]^T$$

s(n): una muestra de una señal de entrada

En este punto, N puede indicar un tamaño de un bloque de la señal de entrada. Es decir, puede incluirse una pluralidad de bloques en la señal de entrada, y cada uno de los bloques puede incluir dos subbloques. Un número

de subbloques incluidos en un único bloque puede variar dependiendo de una configuración de sistema y de la señal de entrada.

- 5 Por ejemplo, la ventana de análisis puede definirse de acuerdo con la Ecuación 3 dada como a continuación. También, de acuerdo con la Ecuación 2 y la Ecuación 3, un resultado de aplicación de la ventana de análisis a un bloque actual de la señal de entrada puede representarse como la Ecuación 4.

[Ecuación 3]

$$W_{\text{análisis}} = [w_1, w_2, w_3, w_4]^T$$

$$w_i = [w_i(0), \dots, w_i(N/4-1)]^T$$

10 [Ecuación 4]

$$[X(b-2), X(b)]^T \otimes W_{\text{análisis}} = [s((b-2)N/4) \cdot w_1(0), \dots, s((b-1)N/4 + N/4 - 1) \cdot w_4(N/4 - 1)]^T$$

- 15 $W_{\text{análisis}}$ puede indicar la ventana de análisis, y tener una característica simétrica. Como se ilustra en la Figura 4, la ventana de análisis puede aplicarse a dos bloques. Es decir, la ventana de análisis puede aplicarse a cuatro subbloques. También, la unidad de procesamiento de ventana 301 puede realizar multiplicación 'punto por punto' con respecto a unos N puntos de la señal de entrada. Los N puntos pueden indicar un tamaño de MDCT. Es decir, la unidad de procesamiento de ventana 301 puede multiplicar un subbloque con un área que corresponde a un subbloque de la ventana de análisis.

- 20 La unidad de MDCT 302 puede realizar una MDCT con respecto a la señal de entrada donde se procesa la ventana de análisis.

La Figura 5 es un diagrama que ilustra una operación de MDCT de acuerdo con una realización de la presente invención.

- 25 Una señal de entrada configurada como una unidad de bloque y una ventana de análisis aplicada a la señal de entrada se ilustran en la Figura 5. Como se ha descrito anteriormente, la señal de entrada puede incluir una trama que incluye una pluralidad de bloques, y un único bloque puede incluir dos subbloques.

- 30 El aparato de codificación 101 puede aplicar una ventana de análisis $W_{\text{análisis}}$ a la señal de entrada. La señal de entrada puede dividirse en cuatro subbloques $X_1(Z)$, $X_2(Z)$, $X_3(Z)$, $X_4(Z)$ incluidos en una trama actual, y la ventana de análisis puede dividirse en $W_1(Z)$, $W_2(Z)$, $W_2^H(Z)$, $W_1^H(Z)$. También, cuando se aplica una MDCT/cuantificación/MDCT inversa (IMDCT) a la señal de entrada basándose en el punto de pliegue que divide los subbloques, puede aparecer un área original y un área de solapamiento.

- 35 El aparato de decodificación 102 puede aplicar una ventana de síntesis a la señal de entrada codificada, eliminar el solapamiento generado durante la operación de MDCT a través de una operación de superposición-adición, y de esta manera puede extraer una señal de salida.

- 40 La Figura 6 es un diagrama que ilustra una operación de hetero codificación (C1, C2) de acuerdo con una realización de la presente invención.

- 45 En la Figura 6, la C1 (caso de cambio 1) y C2 (caso de cambio 2) pueden indicar un límite de una señal de entrada donde se aplica un esquema de hetero codificación. Los subbloques, $s(b-5)$, $s(b-4)$, $s(b-3)$ y $s(b-2)$, localizados en un lado izquierdo basándose en la C1 pueden indicar una señal característica de voz. Los subbloques, $s(b-1)$, $s(b)$, $s(b+1)$ y $s(b+2)$, localizados en un lado derecho basándose en la C1 pueden indicar una señal característica de audio. También, los subbloques, $s(b+m-1)$ y $s(b+m)$, localizados en un lado izquierdo basándose en la C2 pueden indicar una señal característica de audio, y los subbloques, $s(b+m+1)$ y $s(b+m+2)$, localizados en un lado derecho basándose en la C2 pueden indicar una señal característica de voz.

- 50 En la Figura 2, la señal característica de voz puede codificarse a través de la primera unidad de codificación 204, la señal característica de audio puede codificarse a través de la segunda unidad de codificación 205, y por lo tanto la conmutación puede tener lugar en la C1 y en la C2. En este caso, la conmutación puede tener lugar en un punto de pliegue entre subbloques. También, una característica de la señal de entrada puede ser diferente basándose en la C1 y la C2, y por lo tanto se aplican diferentes esquemas de codificación, y puede tener lugar un artefacto de bloqueo.

- 55 En este caso, la codificación se realiza de acuerdo con un esquema de codificación basado en MDCT, el aparato de decodificación 102 puede eliminar el artefacto de bloqueo a través de una operación de superposición-adición usando tanto un bloque anterior como un bloque actual. Sin embargo, cuando tiene lugar la conmutación entre la señal característica de voz y la señal característica de audio como la C1 y la C2, no puede realizarse una operación

de superposición-adición basada en MDCT. Puede requerirse información adicional para la decodificación basada en MDCT. Por ejemplo, puede requerirse información adicional $S_{oL}(b-1)$ en la C1, y puede requerirse información adicional $S_{hL}(b+m)$ en la C2. De acuerdo con una realización de la presente invención, puede evitarse un aumento en una tasa de bits, y puede mejorarse una eficacia de codificación minimizando la información adicional $S_{oL}(b-1)$ y la información adicional $S_{hL}(b+m)$.

Cuando tiene lugar la conmutación entre la señal característica de voz y la señal característica de audio, el aparato de codificación 101 puede codificar la información adicional para restaurar la señal característica de audio. En este caso, la información adicional puede codificarse por la primera unidad de codificación 204 codificando la señal característica de voz. Específicamente, en la C1, un área que corresponde a la información adicional $S_{oL}(b-1)$ en la señal característica de voz $s(b-2)$ puede codificarse como la información adicional. También, en la C2, un área que corresponde a la información adicional $S_{hL}(b+m)$ en la señal característica de voz $s(b+m+1)$ puede codificarse como la información adicional.

Un método de codificación cuando tiene lugar la C1 y la C2 se describe en detalle con referencia a las Figuras 7 a 11, y un método de decodificación se describe en detalle con referencia a las Figuras 15 a 18.

La Figura 7 es un diagrama que ilustra una operación de generación de un flujo de bits en una C1 de acuerdo con una realización de la presente invención.

Cuando se introduce un bloque $X(b)$ de una señal de entrada, la unidad de análisis de estado 202 puede analizar un estado del bloque correspondiente. En este caso, cuando el bloque $X(b)$ es una señal característica de audio y un bloque $X(b-2)$ es una señal característica de voz, la unidad de análisis de estado 202 puede reconocer que la C1 tiene lugar en un punto de pliegue que existe entre el bloque $X(b)$ y el bloque $X(b-2)$. Por consiguiente, puede transmitirse información de control acerca de la generación de la C1 a la unidad de retardo de bloque 201, a la unidad de procesamiento de ventana 301, y a la primera unidad de codificación 204.

Cuando se introduce el bloque $X(b)$ de la señal de entrada, el bloque $X(b)$ y un bloque $X(b+2)$ pueden introducirse a la unidad de procesamiento de ventana 301, el bloque $X(b+2)$ puede retardarse hacia adelante (+2) a través de la unidad de retardo de bloque 201. Por consiguiente, puede aplicarse una ventana de análisis al bloque $X(b)$ y al bloque $X(b+2)$ en la C1 de la Figura 6. En este punto, el bloque $X(b)$ puede incluir los subbloques $s(b-1)$ y $s(b)$, y el bloque $X(b+2)$ puede incluir los subbloques $s(b+1)$ y $s(b+2)$. Puede realizarse una MDCT con respecto al bloque $X(b)$ y al bloque $X(b+2)$ donde se aplica la ventana de análisis a través de la unidad de MDCT 302. Un bloque donde se realiza la MDCT puede codificarse a través de la unidad de generación de flujo de bits 303, y por lo tanto puede generarse un flujo de bits del bloque $X(b)$ de la señal de entrada.

También, para generar la información adicional $S_{oL}(b-1)$ para una operación de superposición-adición con respecto al bloque $X(b)$, la unidad de retardo de bloque 201 puede extraer un bloque $X(b-1)$ retardando hacia atrás el bloque $X(b)$. El bloque $X(b-1)$ puede incluir los subbloques $s(b-2)$ y $s(b-1)$. También, la unidad de corte de señal 203 puede extraer la información adicional $S_{oL}(b-1)$ desde el bloque $X(b-1)$ a través de corte de señal.

Por ejemplo, la información adicional $S_{oL}(b-1)$ puede determinarse por,

[Ecuación 5]

$$s_{oL}(b-1) = [s((b-2) \cdot N/4), \dots, s((b-2) \cdot N/4 + oL - 1)]^T$$

$$0 < oL \leq N/4$$

En este caso, N puede indicar un tamaño de un bloque para MDCT.

La primera unidad de codificación 204 puede codificar un área que corresponde a la información adicional de la señal característica de voz para superposición entre bloques basándose en el punto de pliegue cuando tiene lugar la conmutación entre la señal característica de voz y la señal característica de audio. Por ejemplo, la primera unidad de codificación 204 puede codificar la información adicional $S_{oL}(b-1)$ que corresponde a un área de información adicional (oL) en el subbloque $s(b-2)$ que es la señal característica de voz. Es decir, la primera unidad de codificación 204 puede generar un flujo de bits de la información adicional $S_{oL}(b-1)$ codificando la información adicional $S_{oL}(b-1)$ extraída por la unidad de corte de señal 203. Es decir, cuando tiene lugar la C1, la primera unidad de codificación 204 puede generar únicamente el flujo de bits de la información adicional $S_{oL}(b-1)$. Cuando tiene lugar la C1, la información adicional $S_{oL}(b-1)$ puede usarse como información adicional para eliminar el artefacto de bloqueo.

Para otro ejemplo, cuando puede obtenerse la información adicional $S_{oL}(b-1)$ cuando se codifica el bloque $X(b-1)$, la primera unidad de codificación 204 no puede codificar la información adicional $S_{oL}(b-1)$.

La Figura 8 es un diagrama que ilustra una operación de codificación de una señal de entrada a través de

análisis ilustrada en la Figura 8.

La Figura 9 es un diagrama que ilustra una operación de que genera un flujo de bits en la C2 de acuerdo con una realización de la presente invención.

5 Cuando se introduce un bloque X(b) de una señal de entrada, la unidad de análisis de estado 202 puede analizar un estado de un bloque correspondiente. Como se ilustra en la Figura 6, cuando el subbloque s(b+m) es una señal característica de audio y un subbloque s(b+m+1) es una señal característica de voz, la unidad de análisis de estado 202 puede reconocer que tenga lugar la C2. Por consiguiente, puede transmitirse información de control acerca de la generación de la C2 a la unidad de retardo de bloque 201, a la unidad de procesamiento de ventana 301, y a la primera unidad de codificación 204.

15 Cuando se introduce un bloque X(b+m-1) de la señal de entrada, el bloque X(b+m-1) y un bloque X(b+m+1), que se retarda hacia delante (+2) a través de la unidad de retardo de bloque 201, pueden introducirse a la unidad de procesamiento de ventana 301. Por consiguiente, la ventana de análisis puede aplicarse al bloque X(b+m+1) y al bloque X(b+m-1) en la C2 de la Figura 6. En este punto, el bloque X(b+m+1) puede incluir los subbloques s(b+m+1) y s(b+m), y el bloque X(b+m-1) puede incluir los subbloques s(b+m-2) y s(b+m-1).

20 Por ejemplo, cuando tiene lugar la C2 en el punto de pliegue entre la señal característica de voz y la señal característica de audio en una trama actual de la señal de entrada, la unidad de procesamiento de ventana 301 puede aplicar la ventana de análisis, que no supera el punto de pliegue, a la señal característica de audio.

25 Puede realizarse una MDCT con respecto a los bloques X(b+m+1) y X(b+m-1) donde se aplica la ventana de análisis a través de la unidad de MDCT 302. Un bloque donde se realiza la MDCT puede codificarse a través de la unidad de generación de flujo de bits 303, y por lo tanto puede generarse un flujo de bits del bloque X(b+m-1) de la señal de entrada.

30 También, para generar la información adicional $S_{hL}(b+m)$ para una operación de superposición-adición con respecto al bloque X(b+m-1), la unidad de retardo de bloque 201 puede extraer un bloque X(b+m) retardando hacia delante (+1) el bloque X(b+m-1). El bloque X(b+m) puede incluir los subbloques s(b+m-1) y s(b+m). También, la unidad de corte de señal 203 puede extraer únicamente la información adicional $S_{hL}(b+m)$ a través de corte de señal con respecto al bloque X(b+m).

35 Por ejemplo, la información adicional $S_{hL}(b+m)$ puede determinarse por,

[Ecuación 7]

$$s_{hL}(b+m) = [s((b+m-1) \cdot N/4), \dots, s((b+m-1) \cdot N/4 + hL-1)]^T$$

$$0 < hL \leq N/4$$

En este caso, N puede indicar un tamaño de un bloque para MDCT.

40 La primera unidad de codificación 204 puede codificar la información adicional $S_{hL}(b+m)$ y generar un flujo de bits de la información adicional $S_{hL}(b+m)$. Es decir, cuando tiene lugar la C2, la primera unidad de codificación 204 puede generar únicamente el flujo de bits de la información adicional $S_{hL}(b+m)$. Cuando tiene lugar la C2, la información adicional $S_{hL}(b+m)$ puede usarse como información adicional para eliminar un artefacto de bloqueo.

45 La Figura 10 es un diagrama que ilustra una operación de codificación de una señal de entrada a través de procesamiento de ventana en la C2 de acuerdo con una realización de la presente invención.

50 En la Figura 10, un punto de pliegue puede localizarse entre el subbloque s(b+m) y el subbloque s(b+m+1) con respecto a la C2. También, el punto de pliegue puede ser un punto de pliegue donde la señal característica de audio se conmuta a la señal característica de voz. Es decir, cuando una trama actual ilustrada en la Figura 10 puede incluir subbloques que tienen un tamaño de N/4, el punto de pliegue puede localizarse en un punto de 3N/4.

55 Por ejemplo, cuando existe un punto de pliegue cuando tiene lugar la conmutación entre la señal característica de audio y la señal característica de voz en la trama actual de la señal de entrada, la unidad de procesamiento de ventana 301 puede aplicar una ventana de análisis que no supera el punto de pliegue a la señal característica de audio, es decir, la unidad de procesamiento de ventana 301 puede aplicar la ventana de análisis al subbloque s(b+m) del bloque X(b+m+1) y X(b+m-1).

60 También, la unidad de procesamiento de ventana 301 puede aplicar la ventana de análisis. La ventana de análisis puede configurarse como una ventana que tiene un valor de 0 y corresponde a un primer subbloque, una ventana que corresponde a un área de información adicional de un segundo subbloque, y una ventana que tiene un valor de 1 y corresponde a un área restante del segundo subbloque basándose en el punto de pliegue. El primer subbloque

puede indicar la señal característica de voz, y el segundo subbloque puede indicar la señal característica de audio. En la Figura 10, el punto de pliegue puede localizarse en un punto de $3N/4$ en la trama actual configurada como subbloques que tienen un tamaño de $N/4$.

- 5 Es decir, la unidad de procesamiento de ventana 301 puede sustituir la ventana de análisis w_z para un valor de cero. En este punto, la ventana de análisis puede corresponder al subbloque $s(b+m+1)$ que es la señal característica de voz. También, la unidad de procesamiento de ventana 301 puede determinar una ventana de análisis \hat{w}_3 que corresponde a el subbloque $s(b+m)$ que es la señal característica de audio de acuerdo con la Ecuación 8.

10 [Ecuación 8]

$$w_3 = [w_{unos}, w_{hL}]^T$$

$$w_{hL} = [w_{hL}(0), \dots, w_{hL}(hL-1)]^T$$

$$w_{unos}^{N/4-hL} = \underbrace{[1, \dots, 1]}_{N/4-hL}^T$$

- 15 Es decir, la ventana de análisis \hat{w}_3 , aplicada al subbloque $s(b+m)$ que indica la señal característica de audio basándose en el punto de pliegue, puede incluir un área de información adicional (hL) y un área restante ($N/4-hL$) del área de información adicional (hL). En este caso, el área restante puede configurarse como 1.

- 20 En este caso, w_{hL} puede indicar una segunda mitad de una ventana seno que tiene un tamaño de $2 \times hL$. Un área de información adicional (hL) puede indicar un tamaño para una operación de superposición-adición entre bloques en la C2, y determinar un tamaño de cada uno de w_{hL} y $s_{hL}(b+m)$. También, una muestra de bloque $X_{c2} = [X_{c2}^i, X_{c2}^a]$ puede definirse para seguir la descripción en una muestra de bloque 1000.

- 25 Por ejemplo, la primera unidad de codificación 204 puede codificar una porción que corresponde al área de información adicional en un subbloque, que es una señal característica de voz, para superposición entre bloques basándose en el punto de pliegue. En la Figura 10, la primera unidad de codificación 204 puede codificar una porción que corresponde al área de información adicional (hL) en el subbloque cero $s(b+m+1)$. Como se ha descrito anteriormente, la primera unidad de codificación 204 puede codificar la porción que corresponde al área de información adicional de acuerdo con el esquema de codificación basado en MDCT y el esquema de hetero codificación.

- 30 Como se ilustra en la Figura 10, la unidad de procesamiento de ventana 301 puede aplicar una ventana de análisis con forma de seno a una señal de entrada. Sin embargo, cuando tiene lugar la C2, la unidad de procesamiento de ventana 301 puede establecer una ventana de análisis, que corresponde a un subbloque localizado detrás en el punto de pliegue, como cero. También, la unidad de procesamiento de ventana 301 puede establecer una ventana de análisis, que corresponde al subbloque $s(b+m)$ localizado delante del punto de pliegue, para configurarse como una ventana de análisis que corresponde al área de información adicional (hL) y una ventana de análisis restante. En este punto, la ventana de análisis restante puede tener un valor de 1. La unidad de MDCT 302 puede realizar una MDCT con respecto a una señal de entrada $\{X(b+m-1), X(b+m+1)\} \otimes W$ donde se aplica la ventana de análisis ilustrada en la Figura 10.

- 40 La Figura 11 es un diagrama que ilustra información adicional aplicada cuando se codifica la señal de entrada de acuerdo con una realización de la presente invención.

- 45 La información adicional 1101 puede corresponder a una porción de un subbloque que indica una señal característica de voz basándose en un punto de pliegue C1, y la información adicional 1102 puede corresponder a una porción de un subbloque que indica una señal característica de voz basándose en un punto de pliegue C2. En este caso, un subbloque que corresponde a una señal característica de audio detrás del punto de pliegue de la C1 puede aplicarse a una ventana de síntesis donde se refleja una primera mitad (oL) de la información adicional 1101. Un área restante ($N/4-oL$) puede sustituirse por 1. También, un subbloque, que corresponde a una señal característica de audio hacia delante del punto de pliegue de la C2, puede aplicarse a una ventana de síntesis donde se refleja una segunda mitad (hL) de la información adicional 1102. Un área restante ($N/4-hL$) puede sustituirse por 1.

La Figura 12 es un diagrama de bloques que ilustra una configuración de un aparato de decodificación 102 de acuerdo con una realización de la presente invención.

5 Haciendo referencia a la Figura 12, el aparato de decodificación 102 puede incluir una unidad de retardo de bloque 1201, una primera unidad de decodificación 1202, una segunda unidad de codificación 1203, y una unidad de compensación de bloque 1204.

10 La unidad de retardo de bloque 1201 puede retardar hacia atrás o hacia delante un bloque de acuerdo con un parámetro de control (C1 y C2) incluido en un flujo de bits introducido.

15 También, el aparato de decodificación 102 puede conmutar un esquema de decodificación dependiendo del parámetro de control del flujo de bits introducido para posibilitar que una cualquiera de la primera unidad de decodificación 1202 y la segunda unidad de decodificación 1203 decodifiquen el flujo de bits. En este caso, la primera unidad de decodificación 1202 puede decodificar una señal característica de voz codificada, y la segunda unidad de decodificación 1203 puede decodificar una señal característica de audio codificada. Por ejemplo, la primera unidad de decodificación 1202 puede decodificar la señal característica de audio de acuerdo con un esquema de codificación basado en CELP, y la segunda unidad de decodificación 1203 puede decodificar la señal característica de voz de acuerdo con un esquema de codificación basado en MDCT.

20 Un resultado de decodificación a través de la primera unidad de decodificación 1202 y la segunda unidad de decodificación 1203 puede extraerse como una señal de entrada final a través de la unidad de compensación de bloque 1204.

25 La unidad de compensación de bloque 1204 puede realizar compensación de bloque con respecto al resultado de la primera unidad de decodificación 1202 y el resultado de la segunda unidad de decodificación 1203 para restaurar la entrada, por ejemplo, cuando tiene lugar un punto de pliegue donde existe conmutación entre la señal característica de voz y la señal característica de audio en una trama actual de la señal de entrada, la unidad de compensación de bloque 1204 puede aplicar una ventana de síntesis que no supera el punto de pliegue.

30 En este caso, la unidad de compensación de bloque 1204 puede aplicar una primera ventana de síntesis a información adicional, y aplicar una segunda ventana de síntesis hasta la trama actual para realizar una operación de superposición-adición. En este punto, la información adicional puede extraerse por la primera unidad de decodificación 1202, y la trama actual puede extraerse por la segunda unidad de decodificación 1203. La unidad de compensación de bloque 1204 puede aplicar la segunda ventana de síntesis hasta la trama actual. La segunda
35 ventana de síntesis puede configurarse como una ventana que tiene un valor de 0 y corresponde a un primer subbloque, una ventana que corresponde a un área de información adicional de un segundo subbloque, y una ventana que tiene un valor de 1 y corresponde a un área restante del segundo subbloque basándose en el punto de pliegue. El primer subbloque puede indicar la señal característica de voz, y el segundo subbloque puede indicar la señal característica de audio. La unidad de compensación de bloque 1204 se describe en detalle con referencia a
40 las Figuras 16 a 18.

La Figura 13 es un diagrama que ilustra una operación de decodificación de un flujo de bits a través de una segunda unidad de codificación 1303 de acuerdo con una realización de la presente invención.

45 Haciendo referencia a la Figura 13, la segunda unidad de decodificación 1203 puede incluir una unidad de restauración de flujo de bits 1301, una unidad de MDCT 1302, una unidad de síntesis de ventana 1303, y una unidad de operación de superposición-adición 1304.

50 La unidad de restauración de flujo de bits 1301 puede decodificar un flujo de bits introducido. También, la unidad de IMDCT 1302 puede transformar una señal decodificada a una muestra en un dominio de tiempo a través de una IMDCT.

55 Un bloque Y(b), transformado a través de la unidad de MDCT 1302, puede retardarse hacia atrás a través de la unidad de retardo de bloque 1201 e introducirse a la unidad de procesamiento de ventana 1303. También, el bloque Y(b) puede introducirse directamente a la unidad de procesamiento de ventana 1303 sin el retardo. En este caso, el bloque Y(b) puede tener un valor de $Y(b) = [\tilde{x}(b-2), \tilde{x}(b)]$. En este caso, el bloque Y(b) puede ser un bloque actual introducido a través de la segunda unidad de codificación 205 en la Figura 3.

60 La unidad de síntesis de ventana 1303 puede aplicar la ventana de síntesis al bloque introducido Y(b) y a un bloque retardado Y(b-2). Cuando no tiene lugar la C1 y C2, la unidad de síntesis de ventana 1303 puede aplicar idénticamente la ventana de síntesis a los bloques Y(b) y Y(b-2).

65 Por ejemplo, la unidad de síntesis de ventana 1303 puede aplicar la ventana de síntesis al bloque Y(b) de acuerdo con la Ecuación 9.

[Ecuación 9]

$$\left[\tilde{X}(b-2), \tilde{X}(b) \right]^T \otimes W_{\text{síntesis}} = \left[s((b-2)N/4) \cdot w_1(0), \dots, s((b-1)N/4 + N/4 - 1) \cdot w_1(N/4 - 1) \right]^T$$

En este caso, la ventana de síntesis $W_{\text{síntesis}}$ puede ser idéntica a una ventana de análisis $W_{\text{análisis}}$.

5 La unidad de operación de superposición-adición 1304 puede realizar una operación de superposición-adición del 50 % con respecto a un resultado de aplicación de la ventana de síntesis a los bloques $Y(b)$ y $Y(b-2)$. Un resultado $X(b-2)$ obtenido por la unidad de operación de superposición-adición 1304 puede proporcionarse por,

[Ecuación 10]

$$\tilde{X}(b-2) = \left(\left[\tilde{X}(b-2) \right]^T \otimes [w_1, w_2]^T \right) \oplus \left(\left[\tilde{X}(b-2) \right]^T \otimes [w_3, w_4]^T \right)$$

En este caso, $\left[\tilde{X}(b-2) \right]^T$ y $\left[\tilde{X}(b-2) \right]^T$ pueden asociarse con el bloque $Y(b)$ y el bloque $Y(b-2)$, respectivamente. Haciendo referencia a la Ecuación 10, $X(b-2)$ puede obtenerse realizando una operación de superposición-adición con respecto a un resultado de combinar $\left[\tilde{X}(b-2) \right]^T$ y una primera mitad $[w_1, w_2]^T$ de la ventana de síntesis, y un resultado de combinar $\left[\tilde{X}(b-2) \right]^T$ y una segunda mitad $[w_3, w_4]^T$ de la ventana de síntesis.

La Figura 14 es un diagrama que ilustra una operación de extracción de una señal de salida a través de una operación de superposición-adición de acuerdo con una realización de la presente invención.

Las ventanas 1401, 1402 y 1403 ilustradas en la Figura 14 pueden indicar una ventana de síntesis. La unidad de operación de superposición-adición 1304 puede realizar una operación de superposición-adición con respecto a los bloques 1405 y 1406 donde se aplica la ventana de síntesis 1402, y con respecto a los bloques 1404 y 1405 donde se aplica la ventana de síntesis 1401, y de esta manera puede emitir un bloque 1405. De manera idéntica, la unidad de operación de superposición-adición 1304 puede realizar una operación de superposición-adición con respecto a los bloques 1405 y 1406 donde se aplica la ventana de síntesis 1402, y con respecto a los bloques 1406 y 1407 donde se aplica la ventana de síntesis 1403, y de esta manera puede emitir el bloque 1406.

Es decir, haciendo referencia a la Figura 14, la unidad de operación de superposición-adición 1304 puede realizar una operación de superposición-adición con respecto a un bloque actual y un bloque anterior retardado, y de esta manera puede extraer un subbloque incluido en el bloque actual. En este caso, cada bloque puede indicar una señal característica de audio asociada con una MDCT.

Sin embargo, cuando el bloque 1404 es la señal característica de voz y el bloque 1405 es la señal característica de audio, es decir, cuando tiene lugar la C1, no puede realizarse una operación de superposición-adición puesto que la información de MDCT no está incluida en el bloque 1404. En este caso, puede requerirse información adicional de MDCT del bloque 1404 para la operación superposición-adición. A la inversa, cuando el bloque 1404 es la señal característica de audio y el bloque 1405 es la señal característica de voz, es decir, cuando tiene lugar la C2, no puede realizarse una operación de superposición-adición puesto que la información de MDCT no está incluida en el bloque 1405. En este caso, puede requerirse la información adicional de MDCT del bloque 1405 para la operación de superposición-adición.

La Figura 15 es un diagrama que ilustra una operación de generación de una señal de salida en la C1 de acuerdo con una realización de la presente invención. Es decir, la Figura 15 ilustra una operación de decodificación de la señal de entrada codificada en la Figura 7.

La C1 puede indicar un punto de pliegue donde se genera la señal característica de audio después de la señal característica de voz en la trama actual 800. En este caso, el punto de pliegue puede localizarse en un punto de $N/4$ en la trama actual 800.

La unidad de restauración de flujo de bits 1301 puede decodificar el flujo de bits introducido. De manera secuencial, la unidad de IMDCT 1302 puede realizar una IMDCT con respecto a un resultado de la decodificación. La unidad de síntesis de ventana 1303 puede aplicar la ventana de síntesis a un bloque \tilde{X}_{c1}^n en la trama actual 800 de la señal de entrada codificada por la segunda unidad de codificación 205. Es decir, la segunda unidad de decodificación 1203 puede decodificar un bloque $s(b)$ y un bloque $s(b+1)$ que no están adyacentes al punto de pliegue en la trama actual 800 de la señal de entrada.

En este caso, de manera diferente de la Figura 13, un resultado de la IMDCT no puede pasar a la unidad de retardo de bloque 1201 en la Figura 15.

El resultado de aplicar la ventana de síntesis al bloque \tilde{X}_{c1}^h puede proporcionarse por,

5

[Ecuación 11]

$$\tilde{X}_{c1}^h = \tilde{X}_{c1}^h \otimes [w_3, w_4]^T$$

El bloque \tilde{X}_{c1}^h puede usarse como una señal de bloque para superponerse con respecto a la trama actual 800.

10

Únicamente puede restaurarse la señal de entrada que corresponde al bloque \tilde{X}_{c1}^h en la trama actual 800 por la segunda unidad de decodificación 1203. Por consiguiente, puesto que únicamente puede existir el bloque \tilde{X}_{c1}^i en la trama actual 800, la unidad de operación de superposición-adición 1304 puede restaurar una señal de entrada que corresponde al bloque \tilde{X}_{c1}^i donde no se realiza la operación de superposición-adición. El bloque \tilde{X}_{c1}^i puede ser un bloque donde no se aplica la ventana de síntesis por la segunda unidad de decodificación 1203 en la trama actual 800. También, la primera unidad de decodificación 1202 puede decodificar información adicional incluida en un flujo de bits, y de esta manera puede emitir un subbloque $\tilde{s}_{oL}(b-1)$.

15

El bloque \tilde{X}_{c1}^i extraído por la segunda unidad de decodificación 1203, y el subbloque $\tilde{s}_{oL}(b-1)$ extraído por la primera unidad de decodificación 1202, pueden introducirse a la unidad de compensación de bloque 1204. Una señal de salida final puede generarse por la unidad de compensación de bloque 1204.

20

La Figura 16 es un diagrama que ilustra una operación de compensación de bloque en la C1 de acuerdo con una realización de la presente invención.

25

La unidad de compensación de bloque 1204 puede realizar compensación de bloque con respecto al resultado de la primera unidad de decodificación 1202 y al resultado de la segunda unidad de decodificación 1203, y de esta manera puede restaurar la entrada, por ejemplo, cuando tiene lugar un punto de pliegue cuando existe la conmutación entre una señal característica de voz y una señal característica de audio en una trama actual de la señal de entrada, la unidad de compensación de bloque 1204 puede aplicar una ventana de síntesis que no supera el punto de pliegue.

30

En la Figura 15, información adicional, es decir, el subbloque $\tilde{s}_{oL}(b-1)$ puede extraerse por la primera unidad de decodificación 1202. La unidad de compensación de bloque 1204 puede aplicar una ventana $w_{oL}^T = [w_{oL}(oL-1), \dots, w_{oL}(0)]^T$ al subbloque $\tilde{s}_{oL}(b-1)$. Por consiguiente, un subbloque $\tilde{s}'_{oL}(b-1)$ donde se aplica la ventana w_{oL}^T al subbloque $\tilde{s}_{oL}(b-1)$ puede extraerse de acuerdo con la Ecuación 12.

35

[Ecuación 12]

$$\tilde{s}'_{oL}(b-1) = \tilde{s}_{oL}(b-1) \otimes w_{oL}^T$$

40

También, el bloque \tilde{X}_{c1}^i extraído por la unidad de operación de superposición-adición 1304, puede aplicarse a una ventana de síntesis 1601 a través de la unidad de compensación de bloque 1204.

Por ejemplo, la unidad de compensación de bloque 1204 puede aplicar una ventana de síntesis hasta la trama actual 800. En este punto, la ventana de síntesis puede configurarse como una ventana que tiene un valor de 0 y corresponde a un primer subbloque, una ventana que corresponde a un área de información adicional de un segundo subbloque, y una ventana que tiene un valor de 1 y corresponde a un área restante del segundo subbloque basándose en el punto de pliegue. El primer subbloque puede indicar la señal característica de voz, y el segundo subbloque puede indicar la característica de audio. El bloque \tilde{X}_{c1}^i donde se aplica la ventana de síntesis 1601 puede representarse como,

45

50

[Ecuación 13]

$$\begin{aligned}\tilde{\mathbf{X}}_{c1}^t &= \tilde{\mathbf{X}}_{c1}^t \otimes [\mathbf{w}_1, \mathbf{w}_2]^T = [\mathbf{0}, \dots, \mathbf{0}, \tilde{\mathbf{s}}(b-1) \otimes \hat{\mathbf{w}}_2^T]^T \\ &= [\mathbf{0}, \dots, \mathbf{0}, \tilde{\mathbf{s}}_{oL}(b-1) \otimes \hat{\mathbf{w}}_{oL}^T, \tilde{\mathbf{s}}_{N/4-oL}(b-1)]^T\end{aligned}$$

Es decir, la ventana de síntesis puede aplicarse al bloque $\tilde{\mathbf{X}}_{c1}^t$. La ventana de síntesis puede incluir un área W_1 de 0, y tener un área que corresponde al subbloque $\tilde{\mathbf{s}}(b-1)$ que es idéntico a $\hat{\mathbf{w}}_2$ en la Figura 8. En este caso, el subbloque $\tilde{\mathbf{s}}(b-1)$ incluido en el bloque $\tilde{\mathbf{X}}_{c1}^t$ puede determinarse por,

[Ecuación 14]

$$\tilde{\mathbf{s}}(b-1) = [\tilde{\mathbf{s}}_{oL}(b-1), \tilde{\mathbf{s}}_{N/4-oL}(b-1)]^T$$

En este punto, cuando la unidad de compensación de bloque 1204 realiza una operación de superposición-adición con respecto a un área W_{oL} en las ventanas de síntesis 1601 y 1602, el subbloque $\tilde{\mathbf{s}}_{oL}(b-1)$ que corresponde a un área (oL) puede extraerse desde el subbloque $\tilde{\mathbf{s}}(b-1)$. En este caso, el subbloque $\tilde{\mathbf{s}}_{oL}(b-1)$ puede determinarse de acuerdo con la Ecuación 15. También, un subbloque $\tilde{\mathbf{s}}_{N/4-oL}(b-1)$ que corresponde a un área restante excluyendo el área (oL) del subbloque $\tilde{\mathbf{s}}(b-1)$ puede determinarse de acuerdo con la Ecuación 16.

[Ecuación 15]

$$\tilde{\mathbf{s}}_{oL}(b-1) = \tilde{\mathbf{s}}'_{oL}(b-1) \oplus \tilde{\mathbf{s}}''_{oL}(b-1)$$

[Ecuación 16]

$$\tilde{\mathbf{s}}_{N/4-oL}(b-1) = [\tilde{\mathbf{s}}((b-2) \cdot N/4 + oL), \dots, \tilde{\mathbf{s}}((b-2) \cdot N/4 + N/4 - 1)]^T$$

Por consiguiente, una señal de salida $\tilde{\mathbf{s}}(b-1)$ puede extraerse por la unidad de compensación de bloque 1204.

La Figura 17 es un diagrama que ilustra una operación de generación de una señal de salida en la C2 de acuerdo con una realización de la presente invención. Es decir, la Figura 17 ilustra una operación de decodificación de la señal de entrada codificada en la Figura 9.

La C2 puede indicar un punto de pliegue donde se genera la señal característica de voz después de la señal característica de audio en la trama actual 1000. En este caso, el punto de pliegue puede localizarse en un punto de $3N/4$ en la trama actual 1000.

La unidad de restauración de flujo de bits 1301 puede decodificar el flujo de bits introducido. De manera secuencial, la unidad de IMDCT 1302 puede realizar una IMDCT con respecto a un resultado de la decodificación. La unidad de síntesis de ventana 1303 puede aplicar la ventana de síntesis a un bloque $\hat{\mathbf{X}}_{c2}^t$ en la trama actual 1000 de la señal de entrada codificada por la segunda unidad de codificación 205. Es decir, la segunda unidad de decodificación 1203 puede decodificar un bloque $s(b+m-2)$ y un bloque $s(b+m-1)$ que no son adyacentes al punto de pliegue en la trama actual 1000 de la señal de entrada.

En este caso, diferente de la Figura 13, un resultado de la IMDCT no puede pasar a la unidad de retardo de bloque 1201 en la Figura 17.

El resultado de aplicar la ventana de síntesis al bloque $\hat{\mathbf{X}}_{c2}^t$ puede proporcionarse por,

[Ecuación 17]

$$\tilde{\mathbf{X}}_{c2}^t = \hat{\mathbf{X}}_{c2}^t \otimes [\mathbf{w}_1, \mathbf{w}_2]^T$$

El bloque $\hat{\mathbf{X}}_{c2}^t$ puede usarse como una señal de bloque para superposición con respecto a la trama actual 1000.

Únicamente puede restaurarse la señal de entrada que corresponde al bloque $\hat{\tilde{X}}_{c2}^l$ en la trama actual 1000 por la segunda unidad de decodificación 1203. Por consiguiente, puesto que únicamente puede existir el bloque \tilde{X}_{c2}^h en la trama actual 1000, la unidad de operación de superposición-adición 1304 puede restaurar una señal de entrada que corresponde al bloque \tilde{X}_{c2}^h donde no se realiza la operación de superposición-adición. El bloque \tilde{X}_{c2}^h puede ser un bloque donde no se aplica la ventana de síntesis por la segunda unidad de decodificación 1203 en la trama actual 1000. También, la primera unidad de decodificación 1202 puede decodificar información adicional incluida en un flujo de bits, y de esta manera puede emitir un subbloque $\tilde{s}_{hL}(b+m)$.

El bloque extraído \tilde{X}_{c2}^h , por la segunda unidad de decodificación 1203, y el subbloque $\tilde{s}_{hL}(b+m)$, extraído por la primera unidad de decodificación 1202, pueden introducirse a la unidad de compensación de bloque 1204. Una señal de salida final puede generarse por la unidad de compensación de bloque 1204.

La Figura 18 es un diagrama que ilustra una operación de compensación de bloque en la C2 de acuerdo con una realización de la presente invención.

La unidad de compensación de bloque 1204 puede realizar compensación de bloque con respecto al resultado de la primera unidad de decodificación 1202 y el resultado de la segunda unidad de decodificación 1203, y de esta manera puede restaurar la señal de entrada. Por ejemplo, cuando tiene lugar un punto de pliegue cuando existe la conmutación entre una señal característica de voz y una señal característica de audio en una trama actual de la señal de entrada, la unidad de compensación de bloque 1204 puede aplicar una ventana de síntesis que no supera el punto de pliegue.

En la Figura 17, información adicional, es decir, el subbloque $\tilde{s}_{hL}(b+m)$ puede extraerse por la primera unidad de decodificación 1202. La unidad de compensación de bloque 1204 puede aplicar una ventana $\mathbf{w}_{hL}^T = [w_{hL}(hL-1), \dots, w_{hL}(0)]^T$ al subbloque $\tilde{s}_{hL}(b+m)$. Por consiguiente, un subbloque $\tilde{s}'_{hL}(b+m)$ donde se aplica la ventana \mathbf{w}_{hL}^T al subbloque $\tilde{s}_{hL}(b+m)$, puede extraerse de acuerdo con la Ecuación 18.

[Ecuación 18]

$$\tilde{s}'_{hL}(b+m) = \tilde{s}_{hL}(b+m) \otimes \mathbf{w}_{hL}^T$$

También, el bloque \tilde{X}_{c2}^h extraído por la unidad de operación de superposición-adición 1304, puede aplicarse a una ventana de síntesis 1801 a través de la unidad de compensación de bloque 1204. Por ejemplo, la unidad de compensación de bloque 1204 puede aplicar una ventana de síntesis hasta la trama actual 1000. En este punto, la ventana de síntesis puede configurarse como una ventana que tiene un valor de 0 y corresponde a un primer subbloque, una ventana que corresponde a un área de información adicional de un segundo subbloque, y una ventana que tiene un valor de 1 y corresponde a un área restante del segundo subbloque basándose en el punto de pliegue. El primer subbloque puede indicar la señal característica de voz, y el segundo subbloque puede indicar la señal característica de audio. El bloque \tilde{X}_{c2}^h donde se aplica la ventana de síntesis 1801 puede representarse como,

[Ecuación 19]

$$\begin{aligned} \tilde{X}_{c2}^h &= \tilde{X}_{c2}^h \otimes [\hat{\mathbf{w}}_3, \mathbf{w}_2]^T = [\tilde{\mathbf{s}}(b+m) \otimes \hat{\mathbf{w}}_3^T, 0, \dots, 0]^T \\ &= [\tilde{\mathbf{s}}_{N/4-hL}(b+m), \tilde{\mathbf{s}}_{hL}(b+m) \otimes \hat{\mathbf{w}}_{hL}^T, 0, \dots, 0]^T \end{aligned}$$

Es decir, la ventana de síntesis 1801 puede aplicarse al bloque \tilde{X}_{c2}^h . La ventana de síntesis 1801 puede incluir un área que corresponde al subbloque $s(b+m)$ de 0, y tener un área que corresponde al subbloque $s(b+m+1)$ que es idéntica a $\hat{\mathbf{w}}_3$ en la Figura 10. En este caso, el subbloque $\tilde{\mathbf{s}}(b+m)$ incluido en el bloque puede determinarse por,

[Ecuación 20]

$$\tilde{\mathbf{s}}(b+m) = [\tilde{\mathbf{s}}_{N/4-hL}(b+m), \tilde{\mathbf{s}}'_{hL}(b+m)]^T$$

5 En este punto, cuando la unidad de compensación de bloque 1204 realiza una operación de superposición-adición con respecto a un área W_{hL} en las ventanas de síntesis 1801 y 1802, el subbloque $\tilde{\mathbf{s}}_{hL}(b+m)$ que corresponde a un área (hL) puede extraerse desde el subbloque $\tilde{\mathbf{s}}(b+m)$. En este caso, el subbloque $\tilde{\mathbf{s}}'_{hL}(b+m)$ puede determinarse de acuerdo con la Ecuación 21. También, un subbloque $\tilde{\mathbf{s}}_{N/4-hL}(b+m)$ que corresponde a un área restante excluyendo el área (hL) del subbloque $\tilde{\mathbf{s}}(b+m)$, puede determinarse de acuerdo con la Ecuación 22.

10 [Ecuación 21]

$$\tilde{\mathbf{s}}_{hL}(b+m) = \tilde{\mathbf{s}}'_{hL}(b+m) \oplus \tilde{\mathbf{s}}'_{hL}(b=m)$$

[Ecuación 22]

15
$$\tilde{\mathbf{s}}_{N/4-hL}(b+m) = [\tilde{\mathbf{s}}((b+m-1) \cdot N/4), \dots, \tilde{\mathbf{s}}((b+m-1) \cdot N/4 + hL - 1)]^T$$

Por consiguiente, una señal de salida $\tilde{\mathbf{s}}(b+m)$ puede extraerse por la unidad de compensación de bloque 1204.

20 Aunque se han mostrado y descrito unas pocas realizaciones de la presente invención, la presente invención no está limitada a las realizaciones descritas. En su lugar, se apreciaría por los expertos en la materia que pueden realizarse cambios a estas realizaciones sin alejarse de la invención, el alcance de la cual se define mediante las reivindicaciones y sus equivalentes.

REIVINDICACIONES

1. Un aparato de codificación (101), que comprende:

5 una primera unidad de codificación (204) adaptada para codificar una señal característica de voz de una señal de entrada de acuerdo con un esquema de hetero codificación diferente de un esquema de codificación basado en Transformada de Coseno Discreta Modificada (MDCT); y
 una segunda unidad de codificación (205) adaptada para codificar una señal característica de audio de la señal de entrada de acuerdo con el esquema de codificación basado en MDCT,
 10 en el que la segunda unidad de codificación (205) está adaptada para realizar codificación aplicando una ventana de análisis basándose en un punto de pliegue, cuando el punto de pliegue tiene lugar donde existe la conmutación entre la señal característica de voz y la señal característica de audio en una trama actual de la señal de entrada,
 en el que la primera unidad de codificación (204) está adaptada para codificar información adicional para restaurar la señal característica de audio de acuerdo con un esquema de codificación basado en MDCT, y
 15 en el que la información adicional corresponde a un área de una porción de la señal característica de voz.

2. El aparato de codificación (101) de la reivindicación 1,
 en el que la segunda unidad de codificación (205) está adaptada para aplicar la ventana de análisis, estando configurada la ventana de análisis como una ventana que tiene un valor de 0 y corresponde a un primer subbloque, una ventana que corresponde al área de información adicional de un segundo subbloque, y una ventana que tiene un valor de 1 y corresponde a un área restante del segundo subbloque basándose en el punto de pliegue, indicando el primer subbloque la señal característica de voz, e indicando el segundo subbloque la señal característica de audio.
 25

3. El aparato de codificación (101) de la reivindicación 2, en el que la primera unidad de codificación (204) está adaptada para codificar una porción que corresponde al área de información adicional en el primer subbloque para superposición entre bloques basándose en el punto de pliegue.

4. Un aparato de decodificación (102), que comprende:

una primera unidad de decodificación (1202) adaptada para decodificar una señal característica de voz de una señal de entrada codificada de acuerdo con un esquema de hetero codificación diferente de un esquema de codificación basado en MDCT;
 35 una segunda unidad de decodificación (1203) adaptada para decodificar una señal característica de audio de la señal de entrada codificada de acuerdo con el esquema de codificación basado en MDCT; y
 una unidad de compensación de bloque (1204) adaptada para realizar compensación de bloque con respecto a un resultado de la primera unidad de decodificación (1202) y un resultado de la segunda unidad de decodificación (1202), y para restaurar la señal de entrada,
 40 en el que la unidad de compensación de bloque (1204) está adaptada para aplicar una ventana de síntesis basándose en un punto de pliegue, cuando el punto de pliegue tiene lugar donde existe la conmutación entre la señal característica de voz y la señal característica de audio en una trama actual de la señal de entrada;
 en el que la primera unidad de decodificación (1202) está adaptada para decodificar información adicional para restaurar la señal característica de audio de acuerdo con el esquema de codificación basado en MDCT,
 45 en el que la unidad de compensación de bloque (1204) está adaptada para realizar compensación de bloque usando la información adicional, y
 en el que la información adicional corresponde a un área de una porción de la señal característica de voz.

5. El aparato de decodificación (102) de la reivindicación 4,
 50 en el que la unidad de compensación de bloque (1204) está adaptada para aplicar una primera ventana de síntesis a la información adicional, y aplicar una segunda ventana de síntesis hasta la trama actual para realizar una operación de superposición-adición, extrayéndose la información adicional por la primera unidad de decodificación (1202), y extrayéndose la trama actual por la segunda unidad de decodificación.

6. El aparato de decodificación (102) de la reivindicación 5, en el que la unidad de compensación de bloque (1204) está adaptada para aplicar la segunda ventana de síntesis, configurándose la segunda ventana de síntesis como una ventana que tiene un valor de 0 y corresponde a un primer subbloque, una ventana que corresponde a un área de información adicional de un segundo subbloque, y una ventana que tiene un valor de 1 y corresponde a un área restante del segundo subbloque basándose en el punto de pliegue, indicando el primer subbloque la señal característica de voz, e indicando el segundo subbloque la señal característica de audio.
 60

7. El aparato de decodificación (102) de la reivindicación 4, en el que la segunda unidad de decodificación (1203) está adaptada para decodificar un bloque que es no es adyacente al punto de pliegue en la trama actual de la señal de entrada, y la unidad de compensación de bloque (1204) está adaptada para aplicar la segunda ventana de síntesis a un subbloque adyacente al punto de pliegue en la trama actual de la señal de entrada; y/o
 65 en el que la primera unidad de decodificación está adaptada para decodificar la información adicional codificada de

acuerdo con el esquema de hetero codificación para restaurar la señal característica de audio en la trama actual de la señal de entrada.

FIG. 1

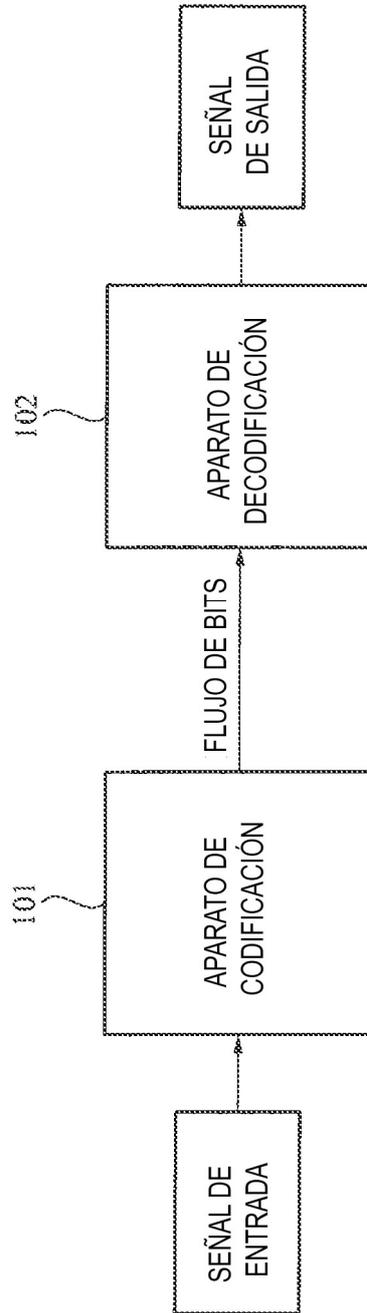


FIG. 2

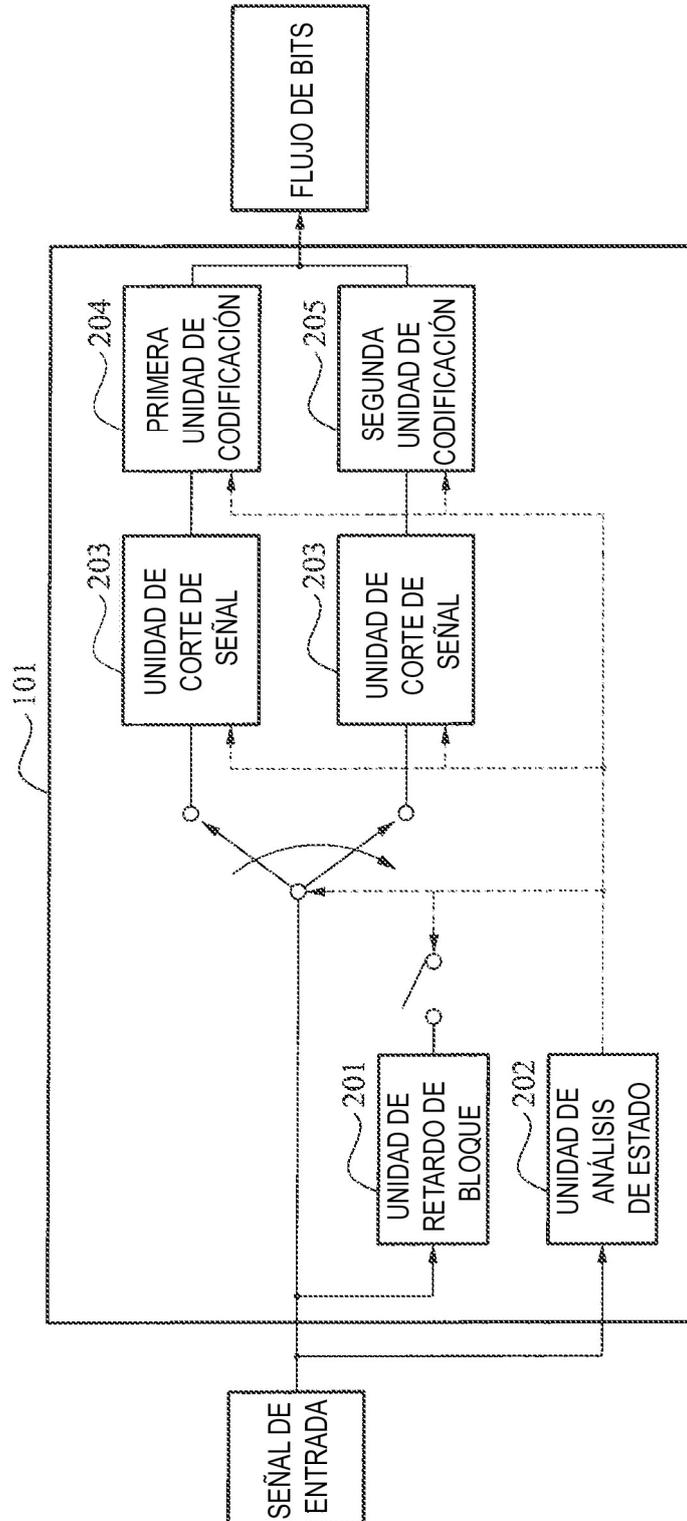


FIG. 3

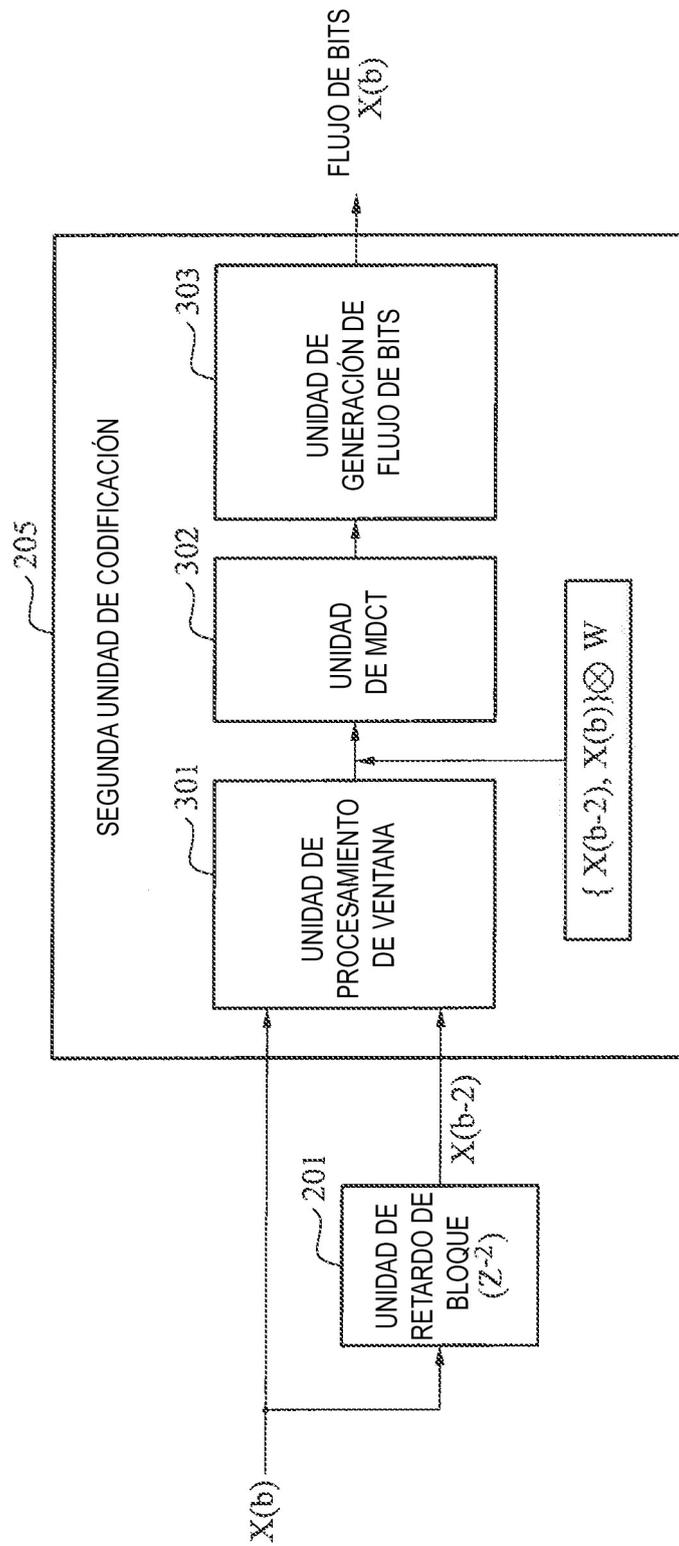
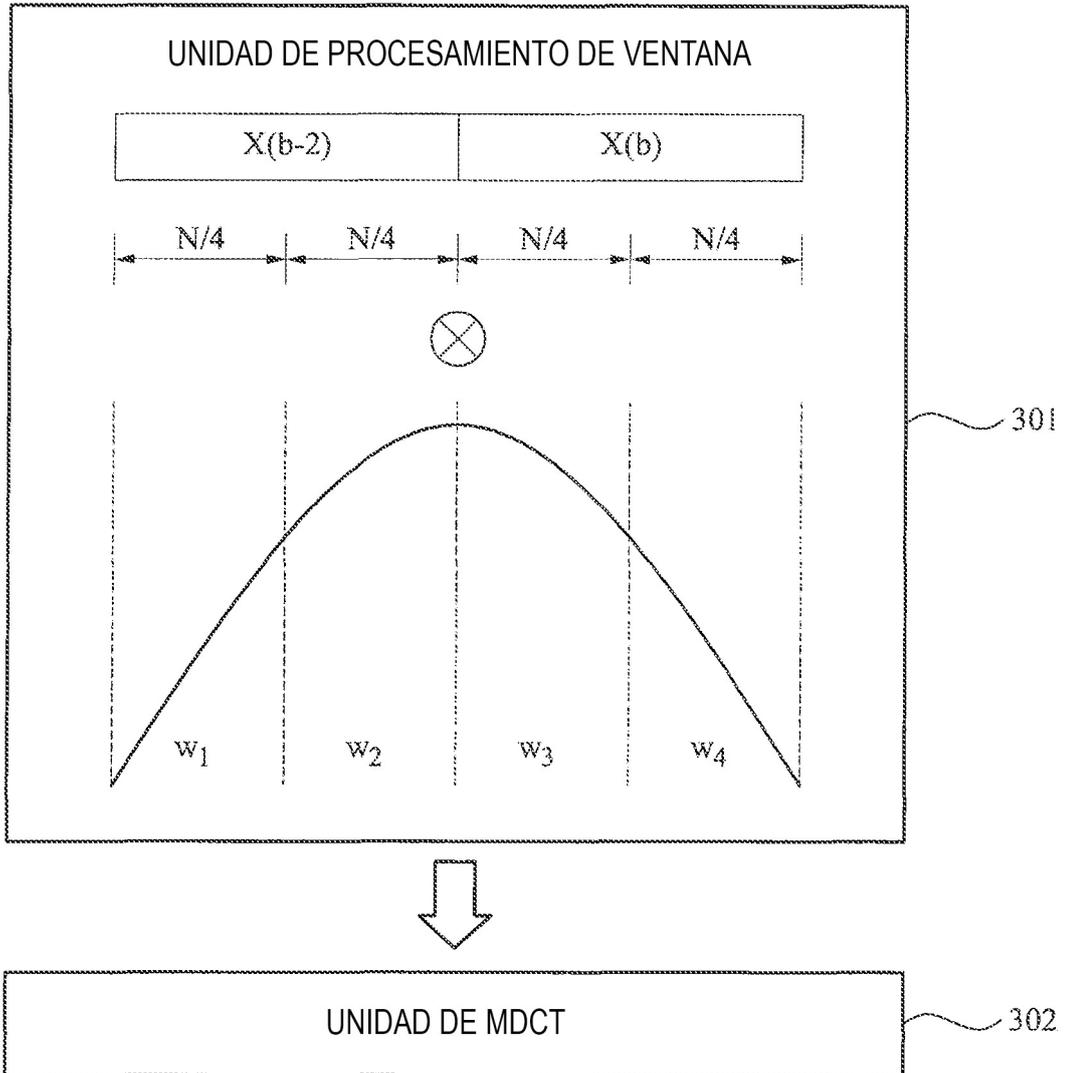


FIG. 4



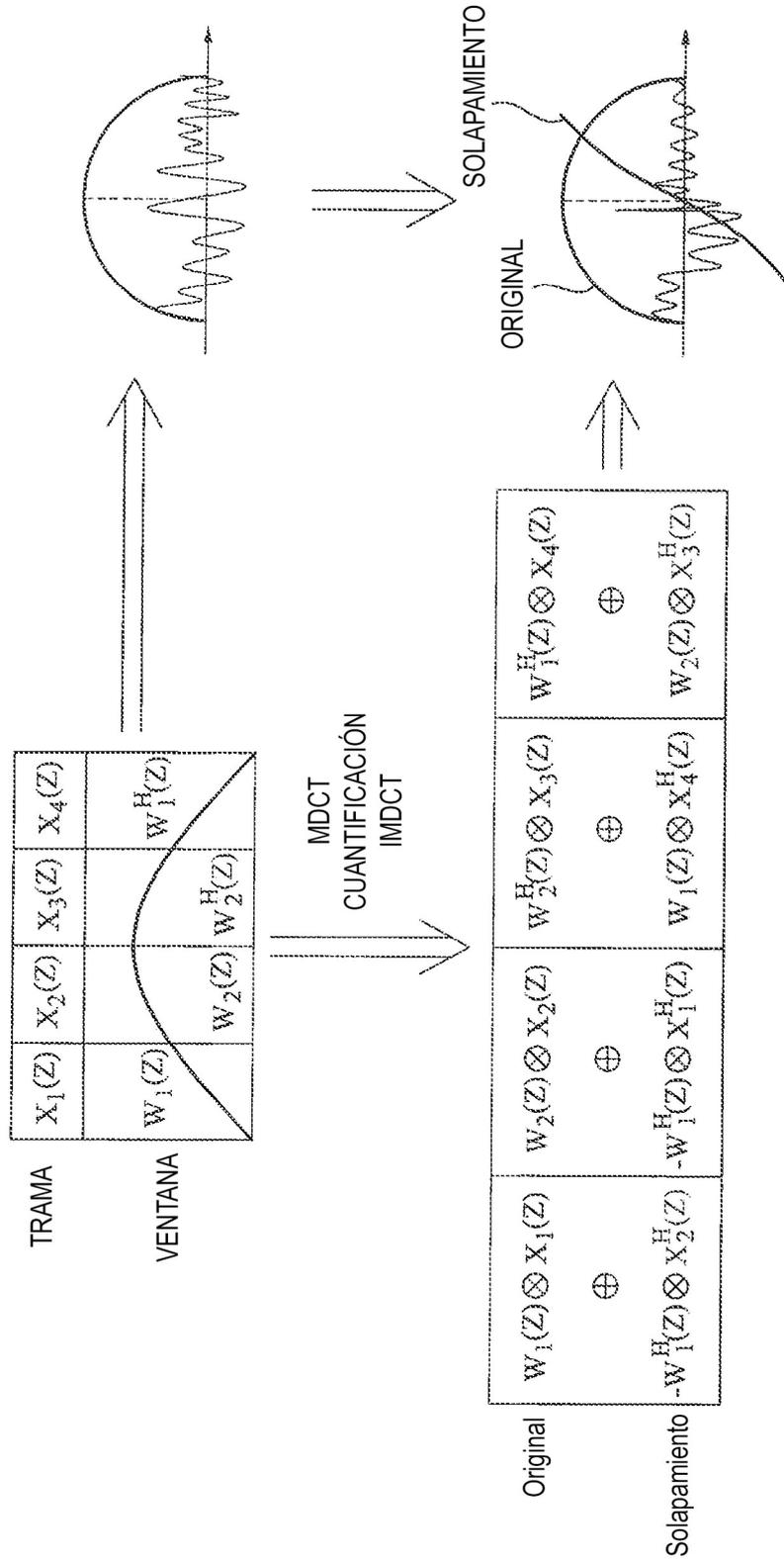


FIG. 6

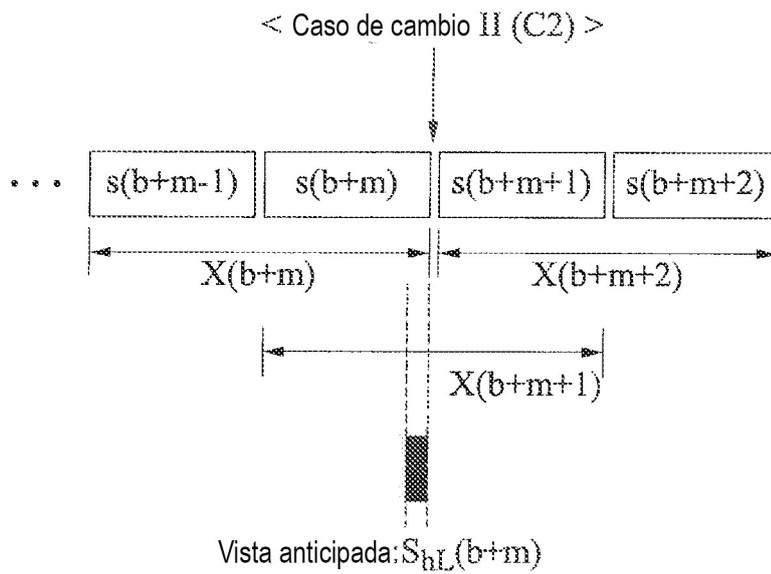
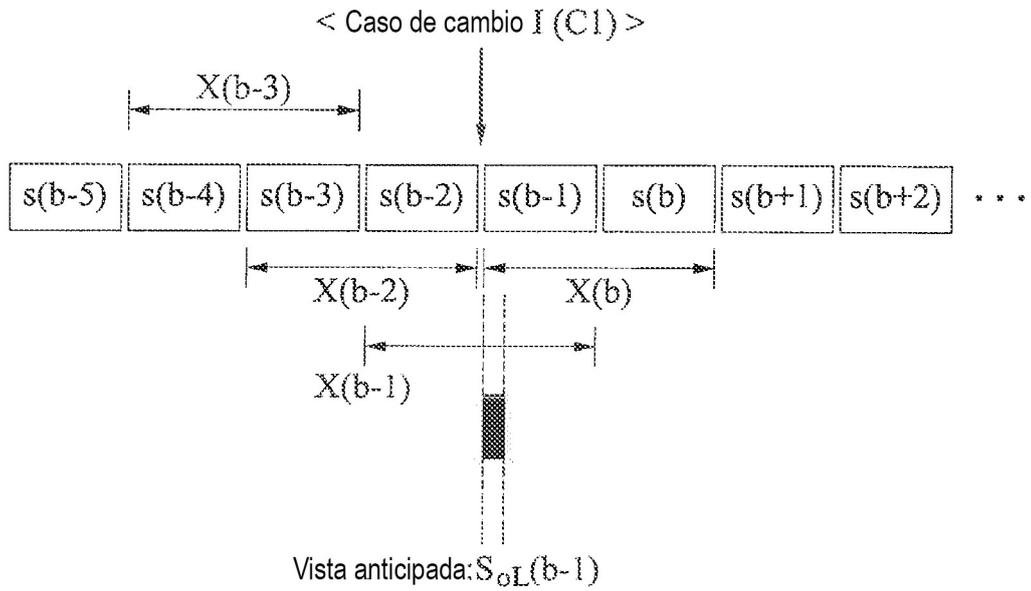


FIG. 7

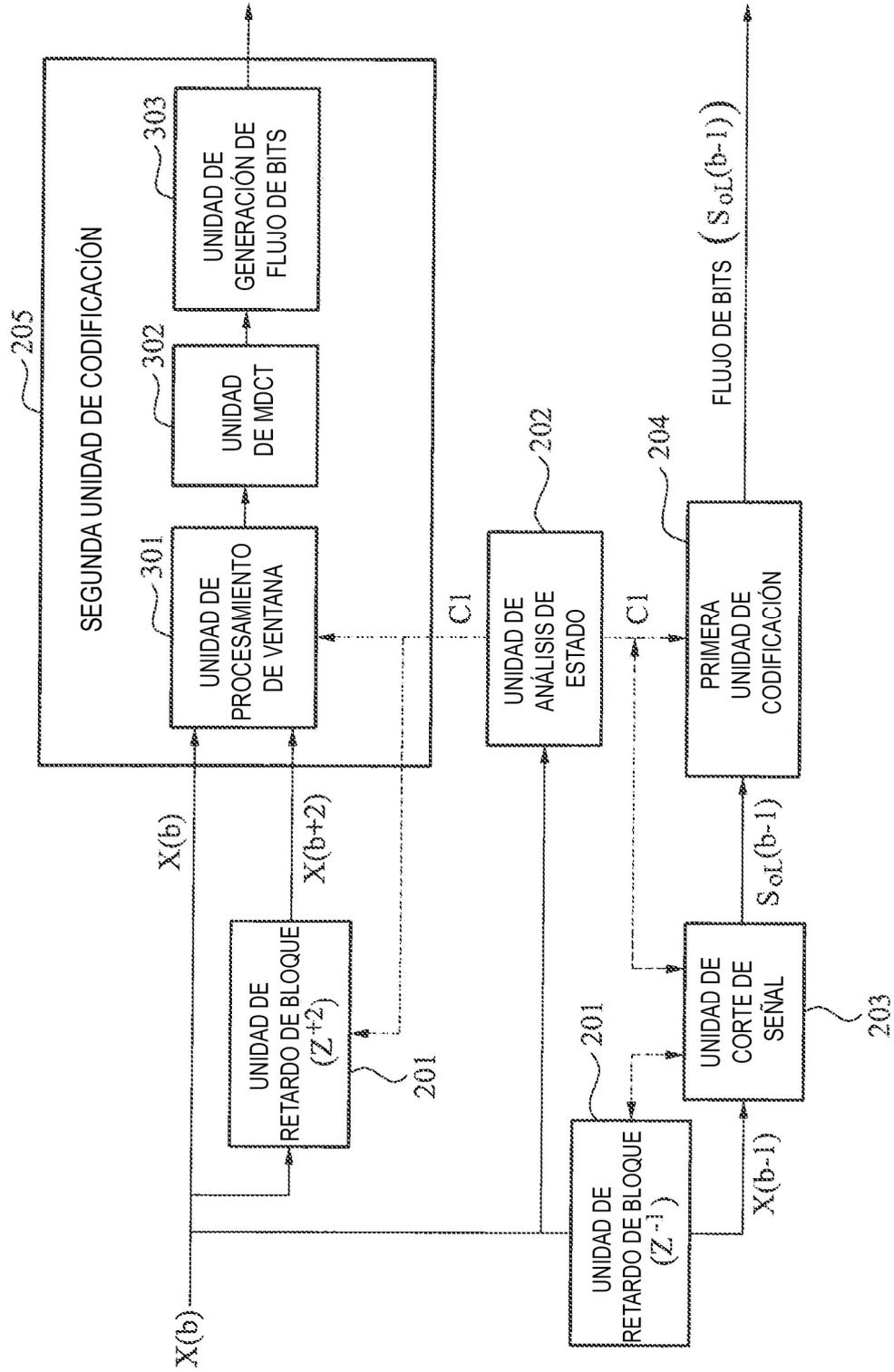


FIG. 8

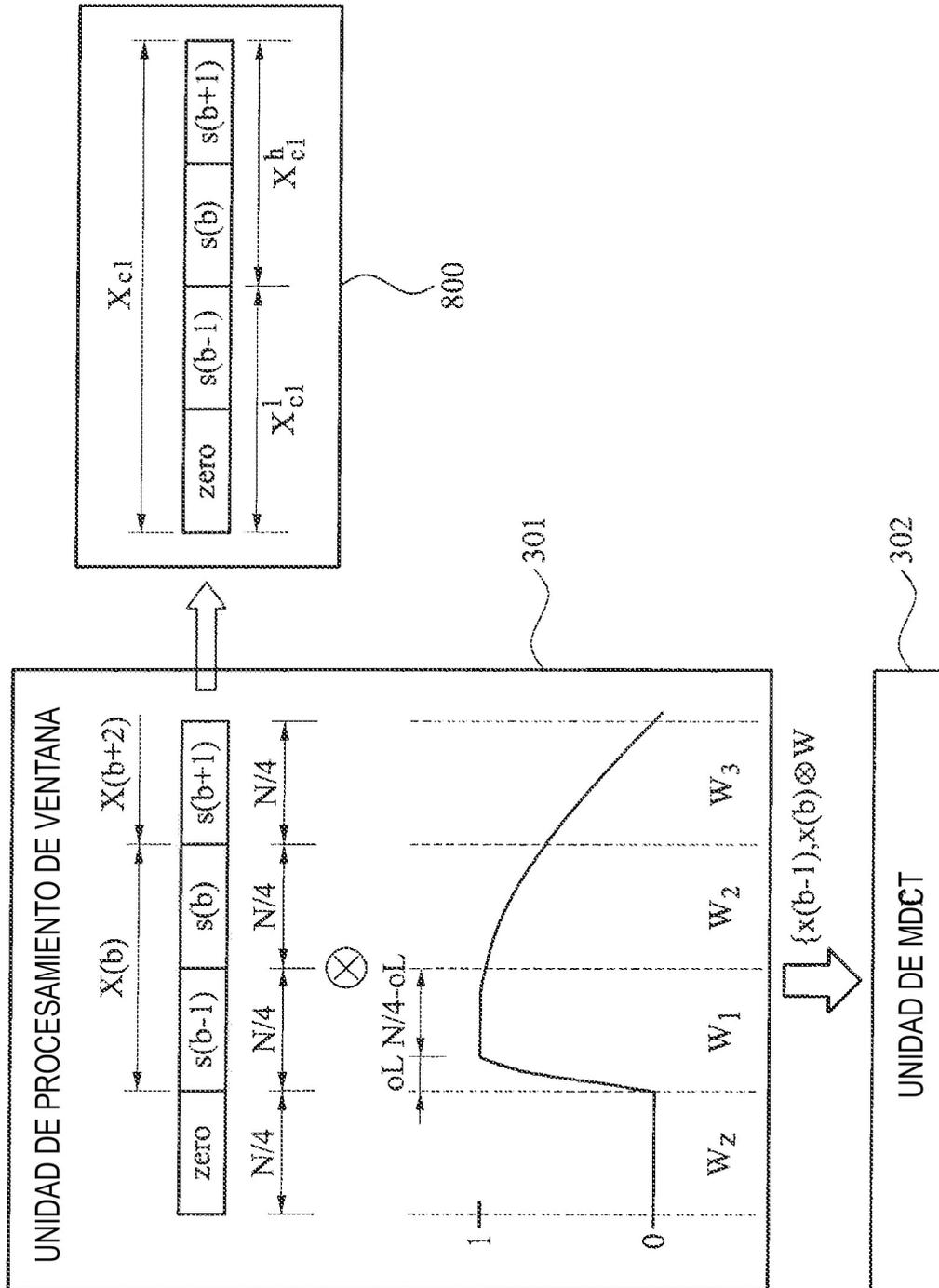


FIG. 9

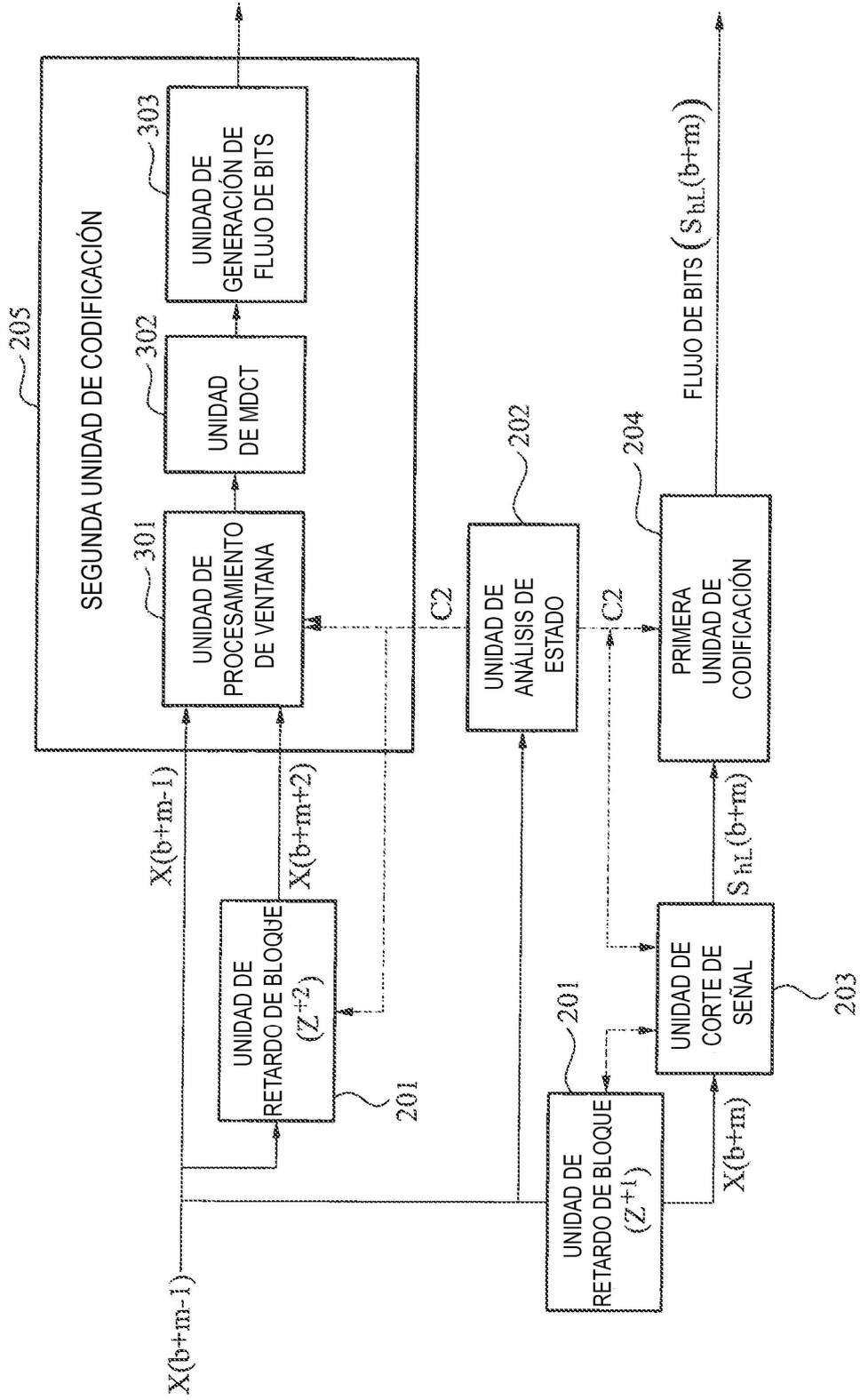


FIG. 10

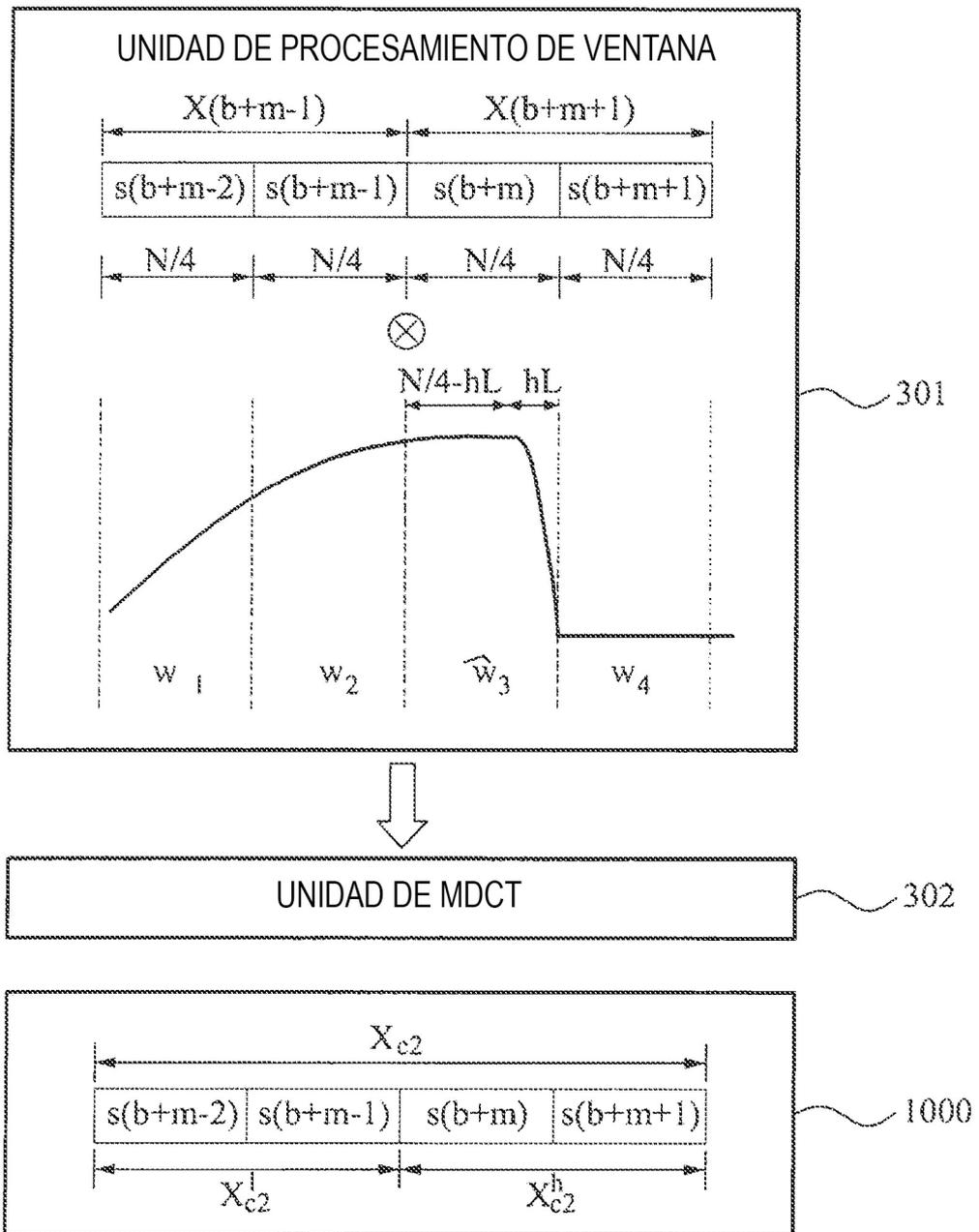


FIG. 11

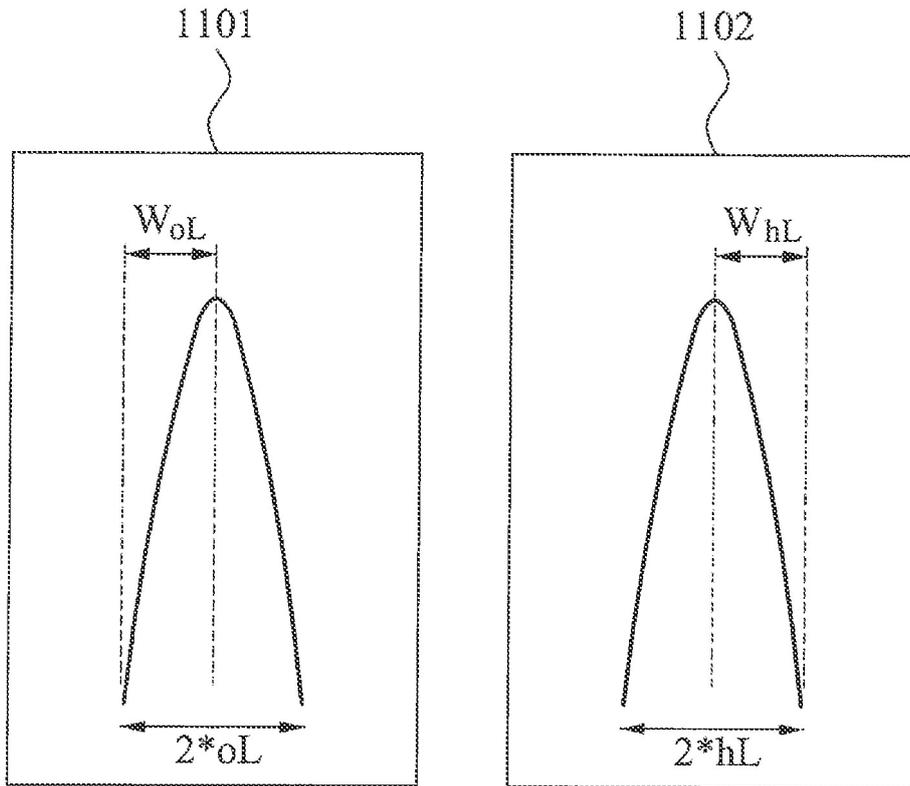


FIG. 12

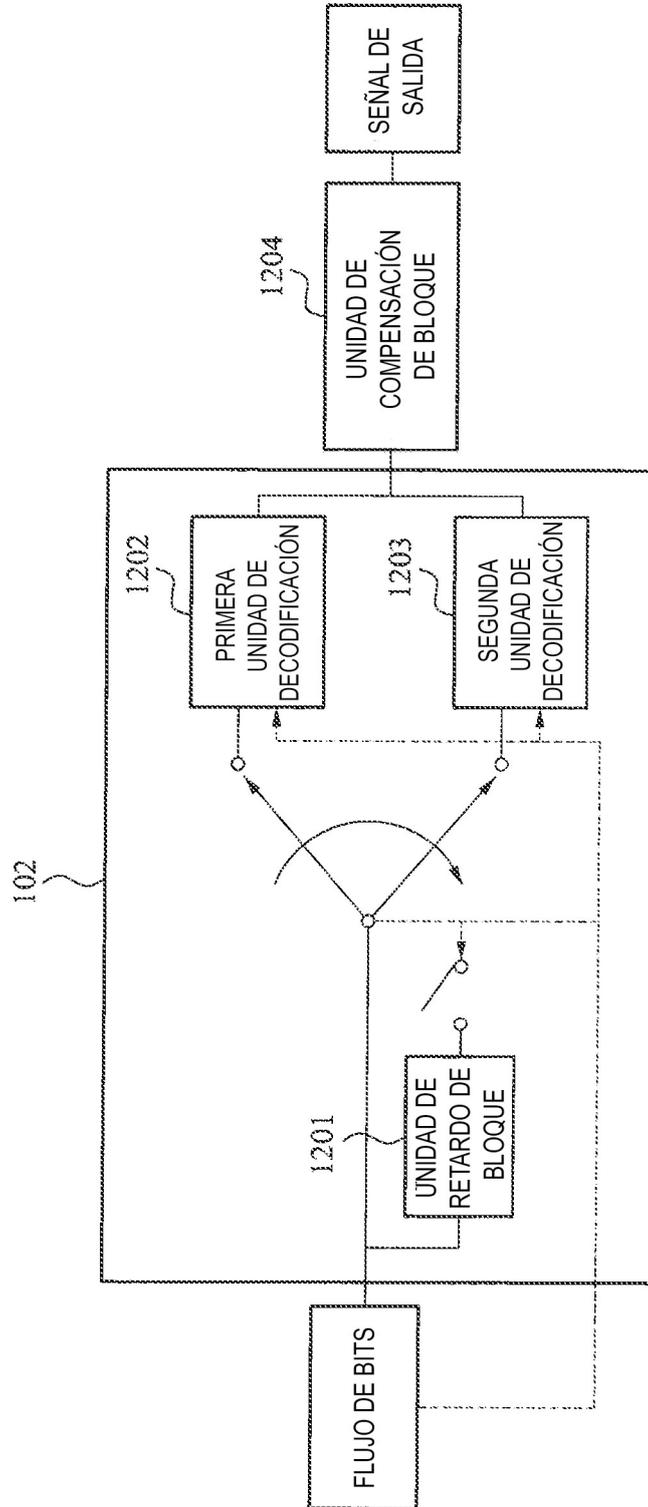


FIG. 13

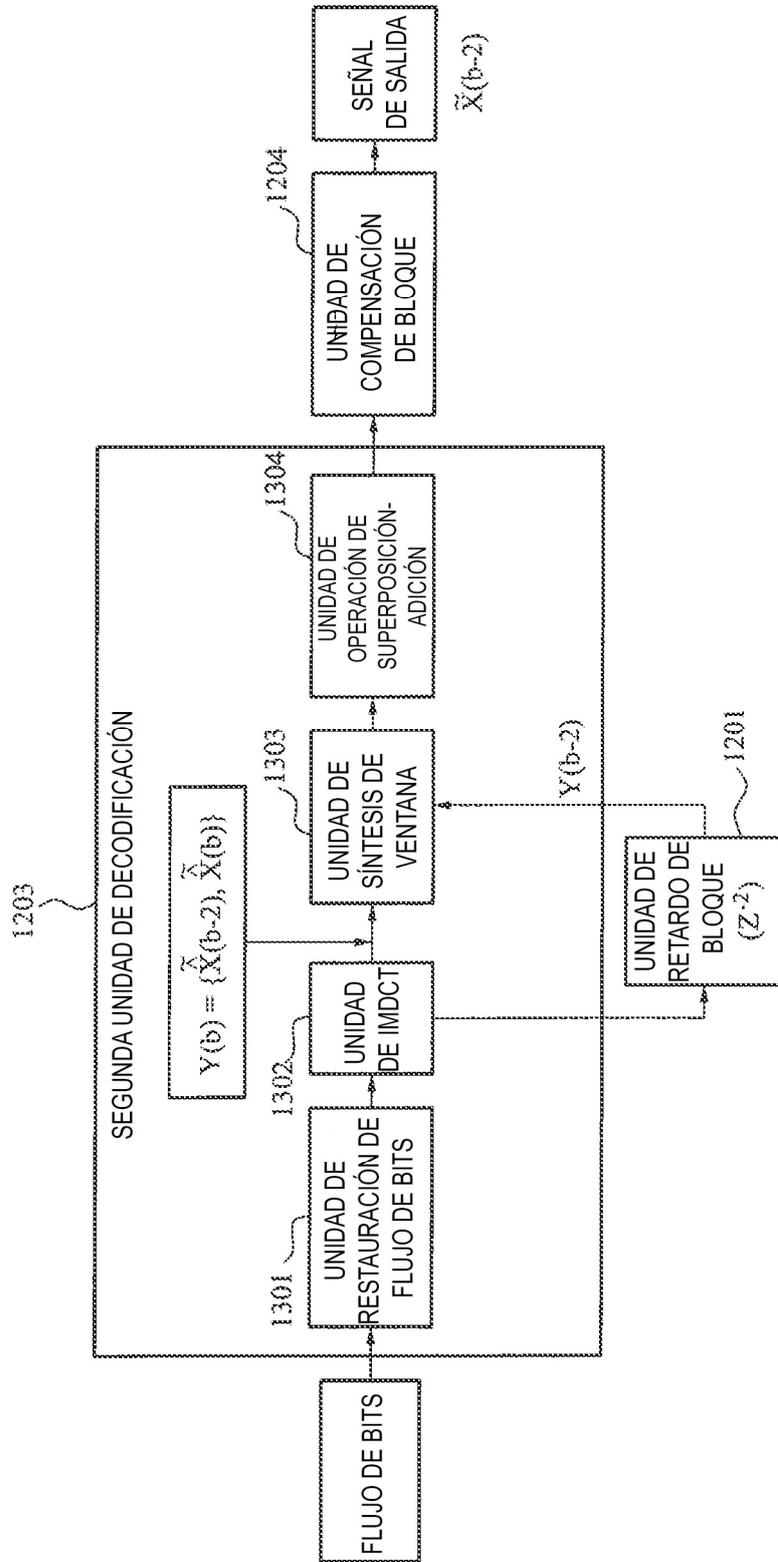


FIG. 14

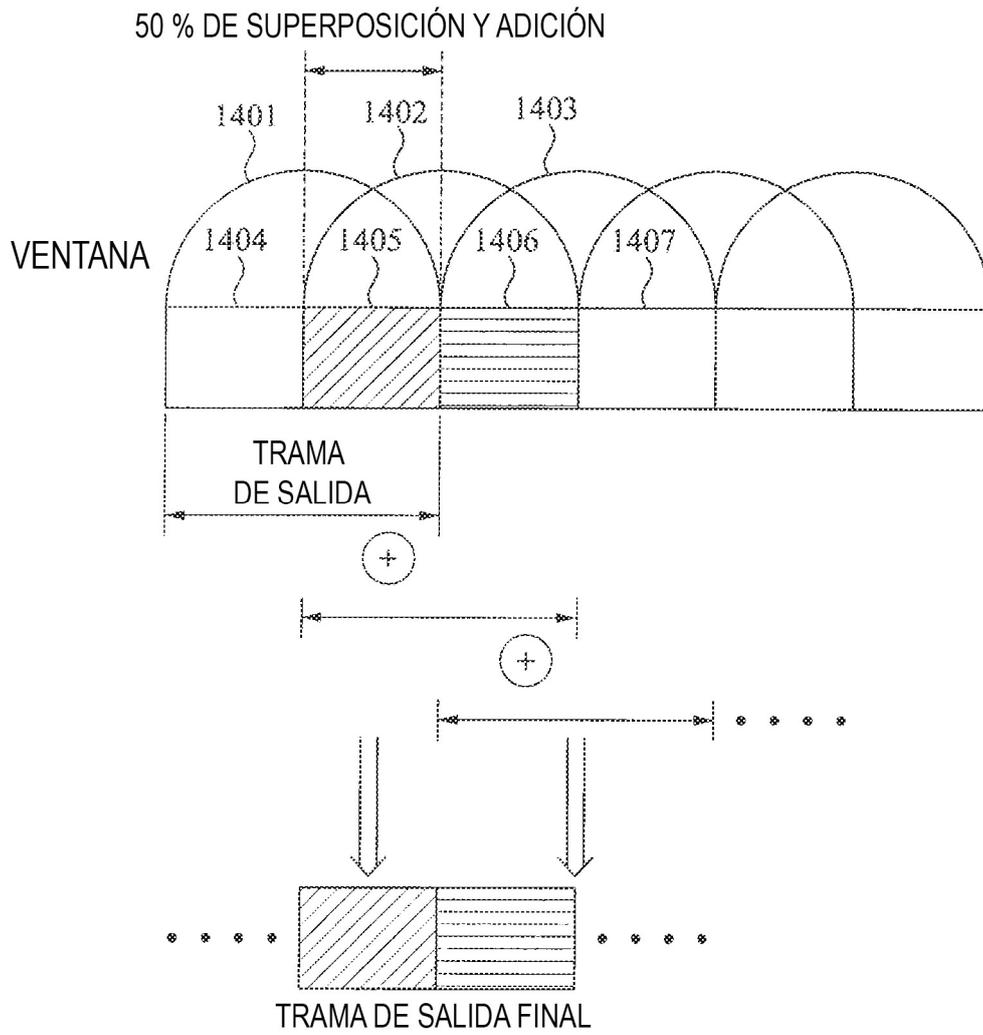


FIG. 15

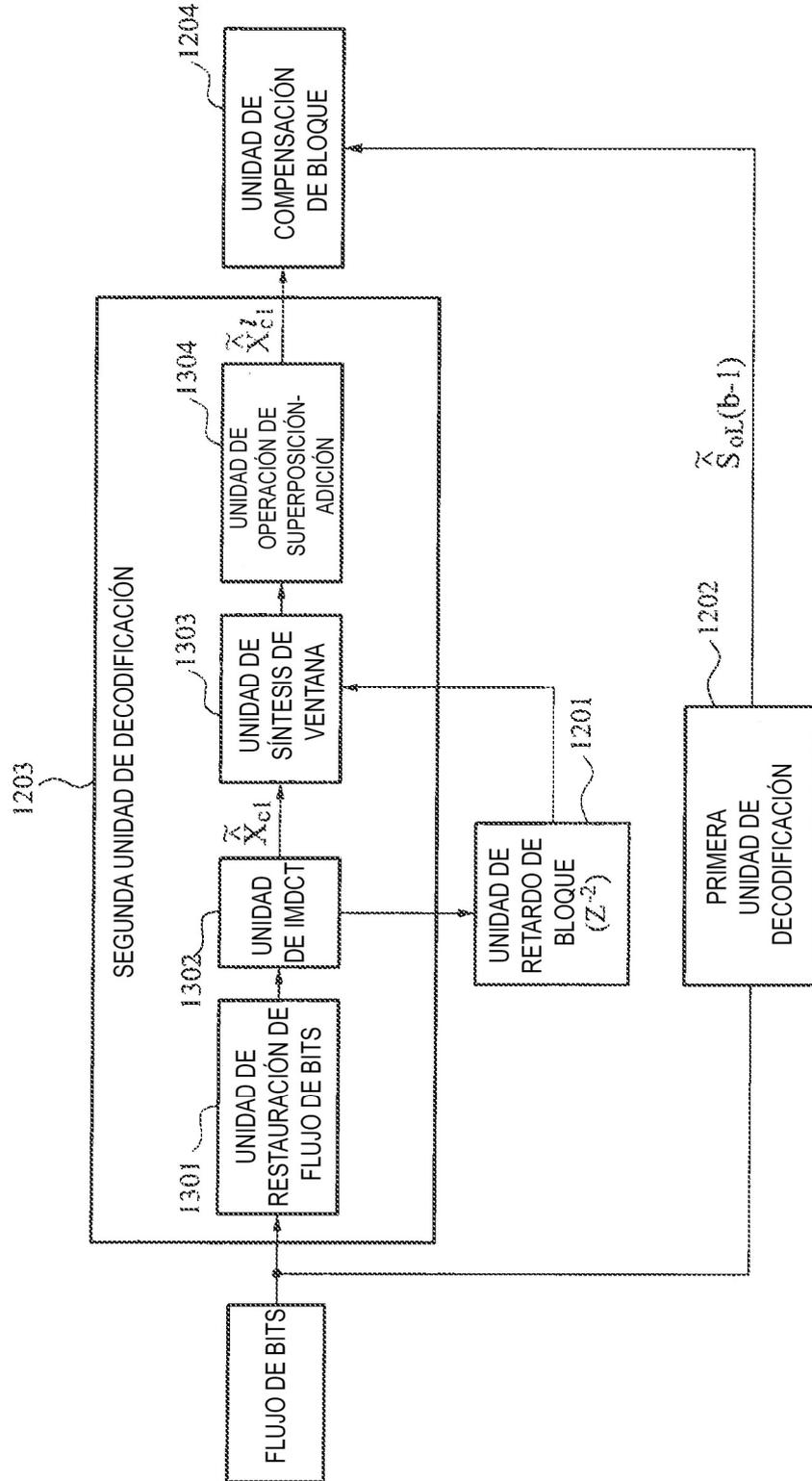


FIG. 17

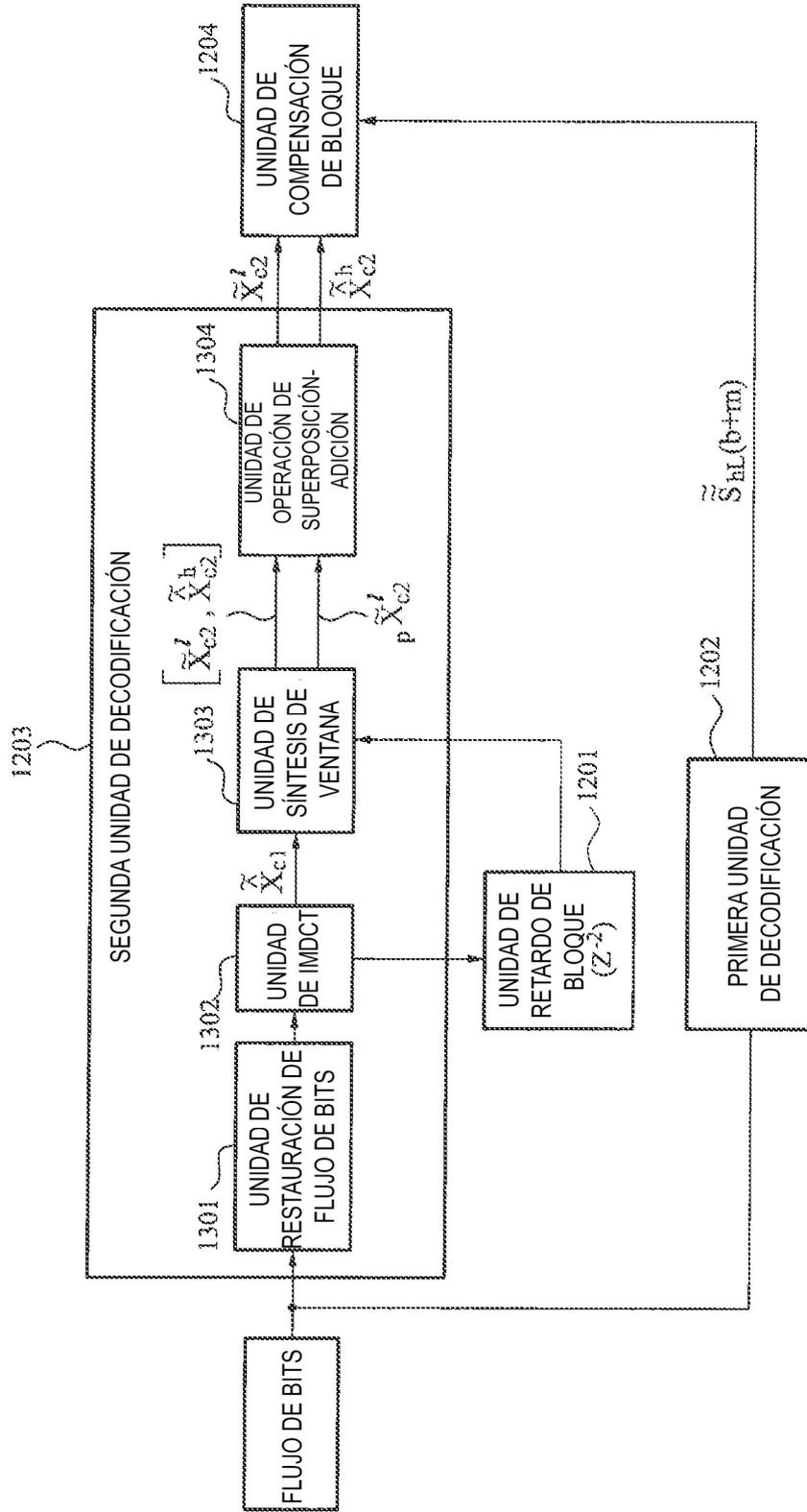


FIG. 18

