

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 474 915**

51 Int. Cl.:

G10L 19/038 (2013.01)

G10L 19/083 (2013.01)

G10L 19/02 (2013.01)

G10L 19/24 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **12.12.2007** **E 07850512 (0)**

97 Fecha y número de publicación de la concesión europea: **04.06.2014** **EP 2101318**

54 Título: **Dispositivo de codificación, dispositivo de decodificación y métodos correspondientes**

30 Prioridad:

13.12.2006 JP 2006336270

02.03.2007 JP 2007053499

17.05.2007 JP 2007132078

13.07.2007 JP 2007185078

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

09.07.2014

73 Titular/es:

**PANASONIC INTELLECTUAL PROPERTY
CORPORATION OF AMERICA (100.0%)
20000 Mariner Avenue, Suite 200
Torrance CA 90503, US**

72 Inventor/es:

**YAMANASHI, TOMOFUMI y
OSHIKIRI, MASAHIRO**

74 Agente/Representante:

UNGRÍA LÓPEZ, Javier

ES 2 474 915 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

Dispositivo de codificación, dispositivo de decodificación y métodos correspondientes

5 **Campo técnico**

La presente invención se refiere a un aparato de codificación/aparato de decodificación y a un método de codificación/método de decodificación usado en un sistema de comunicación en el que se codifica y transmite una señal, y se recibe y decodifica.

10 **Técnica anterior**

Cuando se transmite una señal de habla/audio en un sistema de comunicación móvil o un sistema de comunicación de paquetes caracterizado por comunicación de internet, se usa a menudo tecnología de compresión/codificación para aumentar la eficacia de transmisión de señal de habla/audio. También, en los últimos años, se ha desarrollado un método de codificación/decodificación escalable que posibilita obtener una señal decodificada de buena calidad a partir de parte de información codificada incluso si ocurre un error de transmisión durante la transmisión.

Una tecnología de compresión/codificación anteriormente descrita es una tecnología de codificación predictiva de dominio de tiempo que aumenta la eficacia de compresión usando la correlación temporal de una señal de habla y/o señal de audio (en lo sucesivo denominada como "señal de habla/audio"). Por ejemplo, en el Documento de Patente 1, se predice una señal de trama actual a partir de una señal de trama pasada, y el método de codificación predictiva cambia de acuerdo con el error de predicción. También, en el Documento Distinto de Patente 1, se describe una tecnología en la cual un método de codificación predictiva cambia de acuerdo con el grado de cambio en el dominio de tiempo de un parámetro de habla tal como LSF (Frecuencia Espectral de Línea) y el estado de aparición de error de trama.

Documento de Patente 1: Solicitud de Patente Japonesa Abierta a Inspección Pública Nº HEI 8-211900

Documento distinto de Patente 1: Thomas Eriksson, Jan Linden y Jan Skoglund, "Exploiting Inter-frame Correlation In Spectral Quantization," "Acoustics, Speech, and Signal Processing," 1996. ICASSP-96. Conference Proceedings, 7-10 de mayo de 1996 página(s): 765 - 768 vol. 2

El documento US 2003/093271 A1 tiene por objeto proporcionar un dispositivo de codificación que es capaz de codificar una señal de audio con una alta tasa de compresión, al proporcionar un dispositivo de decodificación que es capaz de decodificar la señal de audio codificada, y al reproducir datos espectrales de frecuencia de banda ancha y una señal de audio de banda ancha. Para conseguir este objetivo, el sistema del documento se centra en una unidad de extensión de banda y una unidad de codificación. Un espectro de frecuencia, transformado desde una señal de entrada en un dominio de tiempo, está dividido en un espectro de frecuencia inferior y unos datos de extensión que especifican un espectro de frecuencia superior. La unidad de codificación se hace funcionar para codificar el espectro de frecuencia inferior y los datos de extensión y para emitir el espectro de frecuencia inferior codificado y los datos de extensión.

Adicionalmente, se copia un espectro parcial como el espectro de frecuencia superior de entre una pluralidad de los espectros parciales que forman el espectro de frecuencia inferior. En relación con el espectro de frecuencia superior, no el propio espectro ha de codificarse, sino principalmente los datos para copiar el espectro de frecuencia inferior, que sustituyen el espectro de frecuencia superior. Esto da como resultado que puede reducirse el efecto de que la cantidad de datos que se consume mediante la corriente de datos codificada que representa los componentes de frecuencia superior.

El artículo "APVQ Encoder Applied to Wideband Speech Coding" por Salavedra J. M. et al., 3 de octubre de 1996, se refiere a una técnica de codificación de habla de banda ancha en la que un codificador APVQ combina técnicas de codificación de subbanda, cuantificación de vector y predicción adaptativa. Se enseña que una señal de habla se divide en 16 subbandas por medio de un banco de filtro QMF, donde cada subbanda tiene un ancho de banda de frecuencia de 500 Hz. De acuerdo con el esquema de codificación desvelado, se forma un vector de señal con una muestra de la señal de error de predicción normalizada que procede de diferentes subbandas y a continuación se cuantifica el vector. Una señal de error de predicción se normaliza mediante su ganancia y la señal de error de predicción normalizada se introduce a la cuantificación del vector, y por lo tanto se considera una cuantificación de vector de ganancia-forma adaptativa.

60 **Divulgación de la invención****Problemas a resolver mediante la invención**

Sin embargo, con cualquiera de las anteriores tecnologías, se realiza codificación predictiva basándose en un parámetro de dominio de tiempo en una base trama a trama, y no se menciona codificación predictiva basándose en un parámetro de dominio no de tiempo tal como un parámetro de dominio de frecuencia. Si un método de

codificación predictiva basándose en un parámetro de dominio de tiempo, tal como se ha descrito anteriormente, se aplica simplemente a la codificación de parámetro de dominio de frecuencia, no hay problema si una banda objetivo de cuantificación es la misma en una trama pasada y trama actual, pero si la banda objetivo de cuantificación es diferente en una trama pasada y trama actual, el error de codificación y la degradación de calidad de audio de señal decodificada aumenta en gran medida, y puede no ser capaz de decodificarse una señal de habla/audio.

Es un objeto de la presente invención proporcionar un aparato de codificación y así sucesivamente capaz de reducir la cantidad de información codificada de una señal de habla/audio, y capaz también de reducir el error de codificación de señal de habla/audio y la degradación de calidad de audio de señal decodificada, cuando un componente de frecuencia de una banda diferente se hace un objetivo de cuantificación en cada trama.

Medios para resolver los problemas

Se define un aparato de codificación de acuerdo con la presente invención en la reivindicación 1.

Se define un aparato de decodificación de acuerdo con la presente invención en la reivindicación 16.

Se define un método de codificación de acuerdo con la presente invención en la reivindicación 17.

Se define un método de decodificación de acuerdo con la presente invención en la reivindicación 18.

Efecto ventajoso de la invención

La presente invención reduce la cantidad de información codificada de una señal de habla/audio o similar, y puede evitar también la degradación de calidad de nitidez de una señal decodificada, habla decodificada y así sucesivamente, y puede reducir el error de codificación de una señal de habla/audio o similar y la degradación de calidad de señal decodificada.

Breve descripción de los dibujos

La Figura 1 es un diagrama de bloques que muestra la configuración principal de un aparato de codificación de habla de acuerdo con la realización 1 de la presente invención;

La Figura 2 es un dibujo que muestra un ejemplo de la configuración de regiones obtenidas mediante una sección de selección de banda de acuerdo con la realización 1 de la presente invención;

La Figura 3 es un diagrama de bloques que muestra la configuración principal de un aparato de decodificación de habla de acuerdo con la realización 1 de la presente invención;

La Figura 4 es un diagrama de bloques que muestra la configuración principal de una variación de un aparato de codificación de habla de acuerdo con la realización no relacionada con la presente invención;

La Figura 5 es un diagrama de bloques que muestra la configuración principal de una variación de un aparato de decodificación de habla de acuerdo con la realización no relacionada con la presente invención;

La Figura 6 es un diagrama de bloques que muestra la configuración principal de un aparato de codificación de habla de acuerdo con la realización 2 de la presente invención;

La Figura 7 es un diagrama de bloques que muestra la configuración principal del interior de una sección de codificación de segunda capa de acuerdo con la realización 2 de la presente invención;

La Figura 8 es un diagrama de bloques que muestra la configuración principal de un aparato de decodificación de habla de acuerdo con la realización 2 de la presente invención;

La Figura 9 es un diagrama de bloques que muestra la configuración principal del interior de una sección de decodificación de segunda capa de acuerdo con la realización 2 de la presente invención;

La Figura 10 es un diagrama de bloques que muestra la configuración principal de un aparato de codificación de habla de acuerdo con la realización 3 de la presente invención;

La Figura 11 es un diagrama de bloques que muestra la configuración principal de un aparato de decodificación de habla de acuerdo con la realización 3 de la presente invención;

La Figura 12 es un diagrama de bloques que muestra la configuración principal de un aparato de codificación de habla de acuerdo con la realización 4 de la presente invención;

La Figura 13 es un diagrama de bloques que muestra la configuración principal de un aparato de decodificación de habla de acuerdo con la realización 4 de la presente invención;

La Figura 14 es un diagrama de bloques que muestra la configuración principal de un aparato de codificación de habla de acuerdo con la realización 5 de la presente invención;

La Figura 15 es un diagrama de bloques que muestra la configuración principal del interior de una sección de codificación de mejora de banda de acuerdo con la realización 5 de la presente invención;

La Figura 16 es un diagrama de bloques que muestra la configuración principal del interior de una sección de codificación de factor de escala correctiva de acuerdo con la realización 5 de la presente invención;

La Figura 17 es un diagrama de bloques que muestra la configuración principal del interior de una sección de codificación de segunda capa de acuerdo con la realización 5 de la presente invención;

La Figura 18 es un diagrama de bloques que muestra la configuración principal de un aparato de decodificación de habla de acuerdo con la realización 5 de la presente invención;

La Figura 19 es un diagrama de bloques que muestra la configuración principal del interior de una sección de decodificación de mejora de banda de acuerdo con la realización 5 de la presente invención;

La Figura 20 es un diagrama de bloques que muestra la configuración principal del interior de una sección de decodificación de segunda capa de acuerdo con la realización 5 de la presente invención;

La Figura 21 es un diagrama de bloques que muestra la configuración principal de un aparato de codificación de habla de acuerdo con la realización 6 de la presente invención;

La Figura 22 es un diagrama de bloques que muestra la configuración principal del interior de una sección de codificación de segunda capa de acuerdo con la realización 6 de la presente invención;

La Figura 23 es un dibujo que muestra un ejemplo de la configuración de regiones obtenidas mediante una sección de selección de banda de acuerdo con la realización 6 de la presente invención;

La Figura 24 es un diagrama de bloques que muestra la configuración principal de un aparato de decodificación de habla de acuerdo con la realización 6 de la presente invención;

La Figura 25 es un diagrama de bloques que muestra la configuración principal del interior de una sección de decodificación de segunda capa de acuerdo con la realización 6 de la presente invención.

La Figura 26 es un diagrama de bloques que muestra la configuración principal de un aparato de codificación de habla de acuerdo con la realización 7 de la presente invención;

La Figura 27 es un diagrama de bloques que muestra la configuración principal del interior de una sección de codificación de segunda capa de acuerdo con la realización 7 de la presente invención;

La Figura 28 es un diagrama de bloques que muestra la configuración principal de un aparato de decodificación de habla de acuerdo con la realización 7 de la presente invención; y

La Figura 29 es un diagrama de bloques que muestra la configuración principal del interior de una sección de decodificación de segunda capa de acuerdo con la realización 7 de la presente invención.

Mejor modo para llevar a cabo la invención

Como una vista general de un ejemplo de la presente invención, en la cuantificación de un componente de frecuencia de una banda diferente en cada trama, si el número de subbandas comunes a una banda objetivo de cuantificación de trama pasada y banda objetivo de cuantificación de trama actual se determina que es mayor que o igual a un valor predeterminado, se realiza codificación predictiva en un parámetro de dominio de frecuencia, y si se determina que el número de subbandas comunes es menor que el valor predeterminado, se codifica directamente un parámetro de dominio de frecuencia. Por este medio, se reduce la cantidad de información codificada de una señal de habla/audio o similar, y también puede evitarse la degradación de calidad de nitidez de una señal decodificada, habla decodificada y así sucesivamente, y puede reducirse el error de codificación de una señal de habla/audio o similar y la degradación de calidad de señal decodificada - y en particular, la degradación de calidad de audio de habla decodificada.

Se describirán ahora realizaciones de la presente invención en detalle con referencia a los dibujos adjuntos. En las siguientes descripciones, un aparato de codificación de habla y aparato de decodificación de habla se usan como ejemplos de un aparato de codificación y un aparato de decodificación de la presente invención.

(Realización 1)

La Figura 1 es un diagrama de bloques que muestra la configuración principal del aparato 100 de codificación de habla de acuerdo con la realización 1 de la presente invención.

En esta figura, el aparato 100 de codificación de habla está equipado con la sección 101 de transformación de dominio de frecuencia, la sección 102 de selección de banda, la sección 103 de cuantificación de forma, la sección 104 de decisión de ejecución/no ejecución de codificación predictiva, la sección 105 de cuantificación de ganancia y la sección 106 de multiplexación.

La sección 101 de transformación de dominio de frecuencia realiza una Transformada de Coseno Discreta Modificada (MDCT) usando una señal de entrada, para calcular un coeficiente de MDCT, que es un parámetro de dominio de frecuencia, y emite este a la sección 102 de selección de banda.

La sección 102 de selección de banda divide la entrada del coeficiente de MDCT de la sección 101 de transformación de dominio de frecuencia en una pluralidad de subbandas, selecciona una banda como una banda objetivo de cuantificación de la pluralidad de subbandas, y emite información de banda que indica la banda seleccionada a la sección 103 de cuantificación de forma, a la sección 104 de decisión de ejecución/no ejecución de codificación predictiva y a la sección 106 de multiplexación. Además, la sección 102 de selección de banda emite el coeficiente de MDCT a la sección 103 de cuantificación de forma. La entrada del coeficiente de MDCT a la sección 103 de cuantificación de forma puede realizarse también directamente desde la sección 101 de transformación de dominio de frecuencia por separado desde la entrada de la sección 101 de transformación de dominio de frecuencia a la sección 102 de selección de banda.

La sección 103 de cuantificación de forma realiza cuantificación de forma usando un coeficiente de MDCT que corresponde a una banda indicada mediante la entrada de información de banda de la sección 102 de selección de

banda de entre la entrada de los coeficientes de MDCT de la sección 102 de selección de banda, y emite la información codificada de forma obtenida a la sección 106 de multiplexación. Además, la sección 103 de cuantificación de forma encuentra un valor de ganancia ideal de cuantificación de forma, y emite el valor de ganancia ideal obtenido a la sección 105 de cuantificación de ganancia.

5 La sección 104 de decisión de ejecución/no ejecución de codificación predictiva encuentra un número de subbandas comunes a una banda objetivo de cuantificación de trama actual y una banda objetivo de cuantificación de trama pasada usando la entrada de información de banda de la sección 102 de selección de banda. A continuación, la sección 104 de decisión de ejecución/no ejecución de codificación predictiva determina que la codificación predictiva se ha de realizar en el coeficiente de MDCT de la banda objetivo de cuantificación indicada mediante la información de banda si el número de subbandas comunes es mayor que o igual a un valor predeterminado, o determina que la codificación predictiva no se ha de realizar en el coeficiente de MDCT de la banda objetivo de cuantificación indicada mediante la información de banda si el número de subbandas comunes es menor que el valor predeterminado. La sección 104 de decisión de ejecución/no ejecución de codificación predictiva emite el resultado de esta determinación a la sección 105 de cuantificación de ganancia.

Si la entrada del resultado de la determinación de la sección 104 de decisión de ejecución/no ejecución de codificación predictiva indica que la codificación predictiva se ha de realizar, la sección 105 de cuantificación de ganancia realiza codificación predictiva de la ganancia de banda objetivo de cuantificación de trama actual usando un valor de ganancia de cuantificación de trama pasada almacenado en una memoria intermedia interna y un libro de códigos de ganancia interno, para obtener información codificada de ganancia. Por otro lado, si la entrada del resultado de la determinación de la sección 104 de decisión de ejecución/no ejecución de codificación predictiva indica que la codificación predictiva no se ha de realizar, la sección 105 de cuantificación de ganancia obtiene información codificada de ganancia cuantificando directamente la entrada del valor de ganancia ideal de la sección 103 de cuantificación de forma. La sección 105 de cuantificación de ganancia emite la información codificada de ganancia obtenida a la sección 106 de multiplexación.

La sección 106 de multiplexación multiplexa la entrada de información de banda de la sección 102 de selección de banda, la entrada de información codificada de forma de la sección 103 de cuantificación de forma y la entrada de información codificada de ganancia de la sección 105 de cuantificación de ganancia y transmite la corriente de bits obtenida a un aparato de decodificación de habla.

El aparato 100 de codificación de habla que tiene una configuración tal como la anteriormente descrita separa una señal de entrada en secciones de N muestras (donde N es un número natural), y realiza codificación en una base trama a trama con N muestras como una trama. La operación de cada sección del aparato 100 de codificación de habla se describe en detalle a continuación.

En la siguiente descripción, una señal de entrada de una trama que es un objetivo de codificación se representa mediante x_n donde $n = 0, 1, \dots, N-1$. En este punto, n indica el índice de cada muestra en una trama que es un objetivo de codificación.

La sección 101 de transformación de dominio de frecuencia tiene N memorias intermedias internas, y en primer lugar inicializa cada memoria intermedia usando un valor de 0 de acuerdo con la Ecuación (1) a continuación.

$$buf_n = 0 \quad (n = 0, 1, \dots, N-1) \quad \dots \quad \text{(Ecuación 1)}$$

En esta ecuación, buf_n ($n = 0, \dots, N-1$) indica la (n+1)ésima de las N memorias intermedias en la sección 101 de transformación de dominio de frecuencia.

A continuación, la sección 101 de transformación de dominio de frecuencia encuentra el coeficiente de MDCT X_k realizando una transformada de coseno discreta modificada (MDCT) de la señal de entrada x_n de acuerdo con la Ecuación (2) a continuación

$$X_k = \frac{2}{N} \sum_{n=0}^{2N-1} x'_n \cos \left[\frac{(2n+1+N)(2k+1)\pi}{4N} \right] \quad (k = 0, \dots, N-1) \quad \dots \quad \text{(Ecuación 2)}$$

En esta ecuación, k indica el índice de cada muestra en una trama, y x'_n es un vector que enlaza la señal de entrada x_n y buf_n de acuerdo con la Ecuación (3) a continuación.

$$x'_n = \begin{cases} buf_n & (n = 0, \dots, N-1) \\ x_{n-N} & (n = N, \dots, 2N-1) \end{cases} \quad \dots \quad \text{(Ecuación 3)}$$

A continuación, la sección 101 de transformación de dominio de frecuencia actualiza buf_n ($n = 0, \dots, N-1$) como se muestra en la Ecuación (4) a continuación.

$$\text{buf}_n = x_n \quad (n = 0, \dots, N-1) \quad \dots \text{(Ecuación 4)}$$

A continuación la sección 101 de transformación de dominio de frecuencia emite el coeficiente de MDCT encontrado X_k a la sección 102 de selección de banda.

La sección 102 de selección de banda en primer lugar divide el coeficiente de MDCT X_k en una pluralidad de subbandas. En este punto, se proporcionará una descripción tomando un caso en el que el coeficiente de MDCT X_k se divide igualmente en J subbandas (donde J es un número natural) como un ejemplo. A continuación la sección 102 de selección de banda selecciona L subbandas consecutivas (donde L es un número natural) de entre las J subbandas, y obtiene M tipos de grupos de subbandas (donde M es un número natural). A continuación, estos M tipos de grupos de subbandas se denominan regiones.

La Figura 2 es un dibujo que muestra un ejemplo de la configuración de regiones obtenidas mediante la sección 102 de selección de banda.

En esta figura, el número de subbandas es 17 ($J=17$), el número de tipos de regiones es ocho ($M=8$), y cada región está compuesta de cinco subbandas consecutivas ($L=5$). De estas, por ejemplo, la región 4 está compuesta de las subbandas 6 a 10.

A continuación, la sección 102 de selección de banda calcula la energía media $E(m)$ de cada uno de los M tipos de regiones de acuerdo con la Ecuación (5) a continuación.

$$E(m) = \frac{\sum_{j=S(m)}^{S(m)+L-1} \sum_{k=B(j)}^{B(j)+W(j)} (X_k)^2}{L} \quad (m = 0, \dots, M-1) \quad \dots \text{(Ecuación 5)}$$

En esta ecuación, j indica el índice de cada una de las J subbandas, m indica el índice de cada uno de los M tipos de regiones, $S(m)$ indica el mínimo valor entre los índices de L subbandas que componen la región m , $B(j)$ indica el mínimo valor entre los índices de una pluralidad de coeficientes de MDCT que componen la subbanda j y $W(j)$ indica el ancho de banda de la subbanda j . En la siguiente descripción, se describirá como un ejemplo un caso en el que los anchos de banda de las J subbandas son todos iguales - es decir, un caso en el que $W(j)$ es una constante.

A continuación, la sección 102 de selección de banda selecciona una región - por ejemplo, una banda compuesta de las subbandas j'' a $j''+L-1$ - para la que la energía media $E(m)$ es un máximo como una banda que es un objetivo de cuantificación (una banda objetivo de cuantificación), y emite el índice $m_{\text{máx}}$ que indica esta región como información de banda a la sección 103 de cuantificación de forma, a la sección 104 de decisión de ejecución/no ejecución de codificación predictiva y a la sección 106 de multiplexación. La sección 102 de selección de banda emite también el coeficiente de MDCT X_k a la sección 103 de cuantificación de forma. En la siguiente descripción, los índices de banda que indican una banda objetivo de cuantificación seleccionada mediante la sección 102 de selección de banda se supone que son j'' a $j''+L-1$.

La sección 103 de cuantificación de forma realiza cuantificación de forma en una base subbanda a subbanda en un coeficiente de MDCT que corresponde a la banda indicada mediante la entrada $m_{\text{máx}}$ de información de banda de la sección 102 de selección de banda. Específicamente, la sección 103 de cuantificación de forma busca un libro de códigos de forma interno compuesto de cantidad SQ de vectores de código de forma para cada una de las L subbandas, y encuentra el índice de un vector de código de forma para el que el resultado de la Ecuación (6) a continuación es un máximo.

$$\text{Forma_}q(i) = \frac{\left\{ \sum_{k=0}^{W(j)} (X_{k+B(j)} \cdot SC_k^i) \right\}^2}{\sum_{k=0}^{W(j)} SC_k^i \cdot SC_k^i} \quad (j = j'', \dots, j''+L-1, i = 0, \dots, SQ-1)$$

... (Ecuación 6)

En esta ecuación, SC_k^i indica un vector de código de forma que compone un libro de códigos de forma, i indica un índice de vector de código de forma y k indica el índice de un elemento de vector de código de forma.

- 5 La sección 103 de cuantificación de forma emite el índice de vector de código de forma $S_máx$ para el que el resultado de la Ecuación (6) anterior es un máximo a la sección 106 de multiplexación como información codificada de forma. La sección 103 de cuantificación de forma calcula también el valor de ganancia ideal $Ganancia_i(j)$ de acuerdo con la Ecuación (7) a continuación, y emite este a la sección 105 de cuantificación de ganancia.

$$Ganancia_i(j) = \frac{\sum_{k=0}^{W(j)} (X_{k+B(j)} \cdot SC_k^{S_máx})}{\sum_{k=0}^{W(j)} SC_{k+B(j)}^{S_máx} \cdot SC_{k+B(j)}^{S_máx}} \quad (j = j'', \dots, j''+L-1)$$

... (Ecuación 7)

- 10 La sección 104 de decisión de ejecución/no ejecución de codificación predictiva tiene una memoria intermedia interna que almacena la entrada de información de banda $m_máx$ de la sección 102 de selección de banda en una trama pasada. En este punto, se describirá un caso a modo de ejemplo en el que la sección 104 de decisión de ejecución/no ejecución de codificación predictiva tiene una memoria intermedia interna que almacena información de
15 banda $m_máx$ para las tres tramas pasadas. La sección 104 de decisión de ejecución/no ejecución de codificación predictiva en primer lugar encuentra un número de subbandas comunes a una banda objetivo de cuantificación de trama pasada y banda objetivo de cuantificación de trama actual usando la entrada de información de banda $m_máx$ de la sección 102 de selección de banda en una trama pasada y la entrada de información de banda $m_máx$ de la sección 102 de selección de banda en la trama actual. A continuación la sección 104 de decisión de ejecución/no
20 ejecución de codificación predictiva determina que la codificación predictiva se ha de realizar si el número de subbandas comunes es mayor que o igual a un valor predeterminado, o determina que la codificación predictiva no se ha de realizar si el número de subbandas comunes es menor que el valor predeterminado. Específicamente, L subbandas indicadas mediante la entrada de información de banda $m_máx$ de la sección 102 de selección de banda una trama atrás en el tiempo se comparan con L subbandas indicadas mediante la entrada de información de banda
25 $m_máx$ de la sección 102 de selección de banda en la trama actual, y se determina que la codificación predictiva se ha de realizar si el número de subbandas comunes es P o más, o se determina que la codificación predictiva no se ha de realizar si el número de subbandas comunes es menor que P . La sección 104 de decisión de ejecución/no ejecución de codificación predictiva emite el resultado de esta determinación a la sección 105 de cuantificación de ganancia. A continuación la sección 104 de decisión de ejecución/no ejecución de codificación predictiva actualiza la
30 información de banda de almacenamiento de memoria intermedia interna usando la entrada de información de banda $m_máx$ de la sección 102 de selección de banda en la trama actual.

- La sección 105 de cuantificación de ganancia tiene una memoria intermedia interna que almacena un valor de ganancia de cuantificación obtenido en una trama pasada. Si una entrada del resultado de la determinación de la
35 sección 104 de decisión de ejecución/no ejecución de codificación predictiva indica que la codificación predictiva se ha de realizar, la sección 105 de cuantificación de ganancia realiza cuantificación prediciendo un valor de ganancia de trama actual usando el valor de ganancia de cuantificación de trama pasada C_j^t almacenado en la memoria intermedia interna. Específicamente, la sección 105 de cuantificación de ganancia busca un libro de códigos de ganancia interno compuesto de la cantidad de GQ de vectores de código de ganancia para cada una de las L
40 subbandas, y encuentra un índice de un vector de código de ganancia para el que el resultado de la Ecuación (8) a continuación es un mínimo.

$$Ganancia_q(i) = \sum_{j=0}^{L-1} \left\{ Ganancia_i(j+j'') - \sum_{t=1}^3 (\alpha_t \cdot C_{j,t}^{j''}) - \alpha_0 \cdot GC_j^{j''} \right\}^2 \quad (i = 0, \dots, GQ-1)$$

... (Ecuación 8)

- 45 En esta ecuación, GC_j^i indica un vector de código de ganancia que compone un libro de códigos de ganancia, i indica un índice de vector de código de ganancia y j indica un índice de un elemento de vector de código de ganancia. Por ejemplo, si el número de subbandas que componen una región es cinco ($L=5$), j tiene un valor de 0 a 4. En este punto, C_j^t indica un valor de ganancia de t tramas anteriores en tiempo, de modo que cuando $t=1$, por ejemplo, C_j^1 indica un valor de ganancia de una trama anterior en tiempo. También, α es un coeficiente de predicción lineal de 4º
50 orden almacenado en la sección 105 de cuantificación de ganancia. La sección 105 de cuantificación de ganancia trata L subbandas en una región como un vector L -dimensional y realiza cuantificación de vector.

La sección 105 de cuantificación de ganancia emite el índice de vector de código de ganancia $G_mín$ para el que el resultado de la Ecuación (8) anterior es un mínimo a la sección 106 de multiplexación como información codificada de ganancia. Si no existe valor de ganancia de una subbanda que corresponde a una trama pasada en la memoria intermedia interna, la sección 105 de cuantificación de ganancia sustituye el valor de ganancia de la subbanda más próxima en frecuencia en la memoria intermedia interna en la Ecuación (8) anterior.

Por otro lado, si la entrada del resultado de la determinación de la sección 104 de decisión de ejecución/no ejecución de codificación predictiva indica que la codificación predictiva no se ha de realizar, la sección 105 de cuantificación de ganancia cuantifica directamente la entrada del valor de ganancia ideal $Ganancia_i(j)$ de la sección 103 de cuantificación de forma de acuerdo con la Ecuación (9) anterior. En este punto, la sección 105 de cuantificación de ganancia trata un valor de ganancia ideal como un vector L -dimensional y realiza cuantificación de vector.

$$Ganancia_q(i) = \sum_{j=0}^{L-1} \{Ganancia_i(j + j'') - GC'_j\}^2 \quad (i = 0, \dots, GQ - 1) \dots \text{(Ecuación 9)}$$

En este punto, un índice de libro de códigos que hace la Ecuación (9) anterior un mínimo se indica mediante $G_mín$.

La sección 105 de cuantificación de ganancia emite $G_mín$ a la sección 106 de multiplexación como información codificada de ganancia. La sección 105 de cuantificación de ganancia actualiza también la memoria intermedia interna de acuerdo con la Ecuación (10) anterior usando la información codificada de ganancia $G_mín$ y el valor de ganancia de cuantificación C_j^1 obtenido en la trama actual.

$$\begin{cases} C_{j+j''}^3 = C_{j+j''}^2 \\ C_{j+j''}^2 = C_{j+j''}^1 \\ C_{j+j''}^1 = GC_j^{G_mín} \end{cases} \quad (j = 0, \dots, L - 1) \dots \text{(Ecuación 10)}$$

La sección 106 de multiplexación multiplexa la entrada de información de banda $m_máx$ de la sección 102 de selección de banda, la entrada de información codificada de forma $S_máx$ de la sección 103 de cuantificación de forma y la entrada de información codificada de ganancia $G_mín$ de la sección 105 de cuantificación de ganancia y transmite la corriente de bits obtenida a un aparato de decodificación de habla.

La Figura 3 es un diagrama de bloques que muestra la configuración principal del aparato 200 de decodificación de habla de acuerdo con esta realización.

En esta figura, el aparato 200 de decodificación de habla está equipado con la sección 201 de demultiplexación, la sección 202 de descuantificación de forma, la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva, la sección 204 de descuantificación de ganancia y la sección 205 de transformación de dominio de tiempo.

La sección 201 de demultiplexación demultiplexa información de banda, información codificada de forma e información codificada de ganancia de una corriente de bits transmitida desde un aparato 100 de codificación de habla, emite la información de banda obtenida a una sección 202 de descuantificación de forma y sección 203 de decisión de ejecución/no ejecución de decodificación predictiva, emite la información codificada de forma obtenida a la sección 202 de descuantificación de forma y emite la información codificada de ganancia obtenida a la sección 204 de descuantificación de ganancia.

La sección 202 de descuantificación de forma encuentra el valor de forma de un coeficiente de MDCT que corresponde a una banda objetivo de cuantificación indicada mediante la entrada de información de banda de la sección 201 de demultiplexación realizando descuantificación de entrada de información codificada de forma de la sección 201 de demultiplexación y emite el valor de forma encontrado a la sección 204 de descuantificación de ganancia.

La sección 203 de decisión de ejecución/no ejecución de decodificación predictiva encuentra un número de subbandas comunes a una banda objetivo de cuantificación de trama actual y una banda objetivo de cuantificación de trama pasada usando la entrada de información de banda de la sección 201 de demultiplexación. A continuación la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva determina que la decodificación predictiva se ha de realizar en el coeficiente de MDCT de la banda objetivo de cuantificación indicada mediante la información de banda si el número de subbandas comunes es mayor que o igual a un valor predeterminado, o determina que la decodificación predictiva no se ha de realizar en el coeficiente de MDCT de la banda objetivo de cuantificación indicada mediante la información de banda si el número de subbandas comunes es menor que el valor predeterminado. La sección 203 de decisión de ejecución/no ejecución de decodificación predictiva emite el

resultado de esta determinación a la sección 204 de descuantificación de ganancia.

Si la entrada del resultado de la determinación de la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva indica que la decodificación predictiva se ha de realizar, la sección 204 de descuantificación de ganancia realiza decodificación predictiva en la entrada de información codificada de ganancia de la sección 201 de multiplexación usando un valor de ganancia de trama pasada almacenado en una memoria intermedia interna y en un libro de códigos de ganancia interno para obtener un valor de ganancia. Por otro lado, si la entrada del resultado de la determinación de la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva indica que la decodificación predictiva no se ha de realizar, la sección 204 de descuantificación de ganancia obtiene un valor de ganancia realizando directamente descuantificación de la entrada de información codificada de ganancia de la sección 201 de demultiplexación usando el libro de códigos de ganancia interno. La sección 204 de descuantificación de ganancia emite el valor de ganancia obtenido a la sección 205 de transformación de dominio de tiempo. La sección 204 de descuantificación de ganancia encuentra también un coeficiente de MDCT de la banda objetivo de cuantificación usando el valor de ganancia obtenido y una entrada de valor de forma de la sección 202 de descuantificación de forma, y emite este a la sección 205 de transformación de dominio de tiempo como un coeficiente de MDCT decodificado.

La sección 205 de transformación de dominio de tiempo realiza una Transformada de Coseno Discreta Modificada Inversa (IMDCT) en la entrada del coeficiente de MDCT decodificado de la sección 204 de descuantificación de ganancia para generar una señal de dominio de tiempo, y emite esta como una señal decodificada.

El aparato 200 de decodificación de habla que tiene una configuración tal como la anteriormente descrita realiza las siguientes operaciones.

La sección 201 de demultiplexación demultiplexa información de banda $m_{\text{máx}}$, información codificada de forma $S_{\text{máx}}$ e información codificada de ganancia $G_{\text{mín}}$ de una corriente de bits transmitida del aparato 100 de codificación de habla, emite la información de banda obtenida $m_{\text{máx}}$ a la sección 202 de descuantificación de forma y sección 203 de decisión de ejecución/no ejecución de decodificación predictiva, emite la información codificada de forma obtenida $S_{\text{máx}}$ a la sección 202 de descuantificación de forma y emite la información codificada de ganancia obtenida $G_{\text{mín}}$ a la sección 204 de descuantificación de ganancia.

La sección 202 de descuantificación de forma tiene un libro de códigos de forma interno similar al libro de códigos de forma que se proporciona con la sección 103 de cuantificación de forma del aparato 100 de codificación de habla, y busca un vector de código de forma para el que la entrada de información codificada de forma $S_{\text{máx}}$ de la sección 201 de demultiplexación es un índice. La sección 202 de descuantificación de forma emite un vector de código buscado a la sección 204 de descuantificación de ganancia como el valor de forma de un coeficiente de MDCT de una banda objetivo de cuantificación indicado mediante la entrada de información de banda $m_{\text{máx}}$ de la sección 201 de demultiplexación. En este punto, un vector de código de forma buscado como un valor de forma se indica mediante $\text{Forma}_q(k)$ ($k = B(j'')$, ..., $B(j''+L)-1$).

La sección 203 de decisión de ejecución/no ejecución de decodificación predictiva tiene una memoria intermedia interna que almacena la entrada de información de banda $m_{\text{máx}}$ de la sección 201 de demultiplexación en una trama pasada. En este punto, se describirá un caso a modo de ejemplo en el que la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva tiene una memoria intermedia interna que almacena la información de banda $m_{\text{máx}}$ para las tres tramas pasadas. La sección 203 de decisión de ejecución/no ejecución de decodificación predictiva en primer lugar encuentra un número de subbandas comunes a una banda objetivo de cuantificación de trama pasada y banda objetivo de cuantificación de banda actual usando la entrada de información de banda $m_{\text{máx}}$ de la sección 201 de demultiplexación en una trama pasada y la entrada de información de banda $m_{\text{máx}}$ de la sección 201 de demultiplexación en la trama actual. A continuación la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva determina que la decodificación predictiva se ha de realizar si el número de subbandas comunes es mayor que o igual a un valor predeterminado, o determina que la decodificación predictiva no se ha de realizar si el número de subbandas comunes es menor que el valor predeterminado. Específicamente, la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva compara L subbandas indicadas mediante la entrada de información de banda $m_{\text{máx}}$ de la sección 201 de demultiplexación una trama atrás en tiempo con L subbandas indicadas mediante la entrada de información de banda $m_{\text{máx}}$ de la sección 201 de demultiplexación en la trama actual, y determina que la decodificación predictiva se ha de realizar si el número de subbandas comunes es P o más, o determina que la decodificación predictiva no se ha de realizar si el número de subbandas comunes es menor que P . La sección 203 de decisión de ejecución/no ejecución de decodificación predictiva emite el resultado de esta determinación a la sección 204 de descuantificación de ganancia. A continuación la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva actualiza la memoria intermedia interna que almacena la información de banda usando la entrada de información de banda $m_{\text{máx}}$ de la sección 201 de demultiplexación en la trama actual.

La sección 204 de descuantificación de ganancia tiene una memoria intermedia interna que almacena un valor de ganancia obtenido en una trama pasada. Si una entrada del resultado de la determinación de la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva indica que la decodificación predictiva se ha de

realizar, la sección 204 de descuantificación de ganancia realiza descuantificación prediciendo un valor de ganancia de trama actual usando un valor de ganancia de trama pasada almacenado en la memoria intermedia interna. Específicamente, la sección 204 de descuantificación de ganancia tiene el mismo tipo de libro de códigos de ganancia interno que la sección 105 de cuantificación de ganancia del aparato 100 de codificación de habla, y
 5 obtiene el valor de ganancia $Ganancia_q'$ realizando descuantificación de ganancia de acuerdo con la Ecuación (11) a continuación. En este punto, $C_j^{m^t}$ indica un valor de ganancia de t tramas anteriores en tiempo, de modo que cuando $t=1$, por ejemplo, $C_j^{m^1}$ indica un valor de ganancia de una trama anterior en tiempo. También, α es un coeficiente de predicción lineal de 4º orden almacenado en la sección 204 de descuantificación de ganancia. La sección 204 de descuantificación de ganancia trata L subbandas en una región como un vector L-dimensional, y
 10 realiza descuantificación de vector.

$$Ganancia_q'(j+j'') = \sum_{i=1}^3 (\alpha_i \cdot C_{j+j''}^{m^i}) + \alpha_0 \cdot GC_j^{G_min} \quad (j = 0, \dots, L-1)$$

... (Ecuación 11)

Si no existe valor de ganancia de una subbanda que corresponde a una trama pasada en la memoria intermedia interna, la sección 204 de descuantificación de ganancia sustituye el valor de ganancia de la subbanda más próxima en frecuencia en la memoria intermedia interna en la Ecuación (11) anterior.

Por otro lado, si la entrada del resultado de la determinación de la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva indica que la decodificación predictiva no se ha de realizar, la sección 204 de descuantificación de ganancia realiza descuantificación de un valor de ganancia de acuerdo con la Ecuación (12) a continuación usando el libro de códigos de ganancia anteriormente descrito. En este punto, un valor de ganancia se trata como un vector L-dimensional, y se realiza descuantificación de vector. Es decir, cuando la decodificación predictiva no se realiza, el vector de código de ganancia $GC_j^{G_min}$ que corresponde a la información codificada de ganancia G_min se toma directamente como un valor de ganancia.

$$Ganancia_q'(j+j'') = GC_j^{G_min} \quad (j = 0, \dots, L-1) \quad \dots \text{ (Ecuación 12)}$$

A continuación, la sección 204 de descuantificación de ganancia calcula un coeficiente de MDCT decodificado de acuerdo con la Ecuación (13) a continuación usando un valor de ganancia obtenido mediante la descuantificación de trama actual y una entrada de valor de forma de la sección 202 de descuantificación de forma, y actualiza la memoria intermedia interna de acuerdo con la Ecuación (14) a continuación. En este punto, un coeficiente de MDCT decodificado calculado se indica mediante X''_k . También, en la descuantificación de coeficiente de MDCT, si k está presente en $B(j'')$ a $B(j''+1)-1$, el valor de ganancia $Ganancia_q'(j)$ toma el valor de $Ganancia_q'(j'')$.

$$X''_k = Ganancia_q'(j) \cdot Forma_q'(k) \quad \left(\begin{matrix} k = B(j''), \dots, B(j''+L)-1 \\ j = j'', \dots, j''+L-1 \end{matrix} \right) \quad \dots \text{ (Ecuación 13)}$$

$$\begin{cases} C_j^{m^3} = C_j^{m^2} \\ C_j^{m^2} = C_j^{m^1} \\ C_j^{m^1} = Ganancia_q'(j) \end{cases} \quad (j = j'', \dots, j''+L-1) \quad \dots \text{ (Ecuación 14)}$$

La sección 204 de descuantificación de ganancia emite el coeficiente de MDCT decodificado X''_k calculado de acuerdo con la Ecuación (13) anterior a la sección 205 de transformación de dominio de tiempo.

La sección 205 de transformación de dominio de tiempo en primer lugar inicializa la memoria intermedia interna buf'_k a un valor de cero de acuerdo con la Ecuación (15) a continuación.

$$buf'_k = 0 \quad (k = 0, \dots, N-1) \quad \dots \text{ (Ecuación 15)}$$

A continuación la sección 205 de transformación de dominio de tiempo encuentra la señal decodificada Y_n de acuerdo con la Ecuación (16) a continuación usando la entrada del coeficiente de MDCT decodificado X''_k de la sección 204 de descuantificación de ganancia.

$$Y_n = \frac{2}{N} \sum_{k=0}^{2N-1} X''_k \cos \left[\frac{(2n+1+N)(2k+1)\pi}{4N} \right] \quad (n=0, \dots, N-1)$$

... (Ecuación 16)

En esta ecuación, X''_k es un vector que enlaza el coeficiente de MDCT decodificado X''_k y la memoria intermedia buf_k .

$$X''_k = \begin{cases} buf'_k & (k=0, \dots, N-1) \\ X''_k & (k=N, \dots, 2N-1) \end{cases} \quad \dots \text{ (Ecuación 17)}$$

A continuación, la sección 205 de transformación de dominio de tiempo actualiza la memoria intermedia buf'_k de acuerdo con la Ecuación (18) a continuación.

$$buf'_k = X''_k \quad (k=0, \dots, N-1) \quad \dots \text{ (Ecuación 18)}$$

La sección 205 de transformación de dominio de tiempo emite la señal decodificada obtenida Y_n como una señal de salida.

Por lo tanto, de acuerdo con esta realización, se selecciona una banda de alta energía en cada trama como una banda objetivo de cuantificación y se cuantifica un parámetro de dominio de frecuencia, posibilitando que se cree polarización en la distribución de valor de ganancia cuantificada, y que se mejore el rendimiento de cuantificación de vector.

También, de acuerdo con esta realización, en la cuantificación del parámetro de dominio de frecuencia de una banda objetivo de cuantificación diferente de cada trama, se realiza codificación predictiva en un parámetro de dominio de frecuencia si se determina que el número de subbandas comunes a una banda objetivo de cuantificación de trama pasada y la banda objetivo de cuantificación de trama actual es mayor que o igual a un valor predeterminado, y se codifica directamente un parámetro de dominio de frecuencia si se determina que el número de subbandas comunes es menor que el valor predeterminado. En consecuencia, se reduce la cantidad de información codificada en la codificación de habla, y también puede evitarse la degradación de calidad de habla nítida, y puede reducirse el error de codificación de señal de habla/audio y la degradación de calidad de audio de señal decodificada.

Adicionalmente, de acuerdo con esta realización, en el lado de codificación puede decidirse una banda objetivo de cuantificación, y realizarse cuantificación del parámetro de dominio de frecuencia, en unidades de región cada una compuestas de una pluralidad de subbandas, y la información en cuanto a un parámetro de dominio de frecuencia del que la región se ha hecho un objetivo de cuantificación puede transmitirse al lado de decodificación. En consecuencia, puede mejorarse la eficacia de cuantificación y la cantidad de información codificada transmitida al lado de decodificación puede reducirse adicionalmente como se compara con decidir si se ha de usar o no la codificación predictiva en una base subbanda a subbanda y transmitir información en cuanto a qué subbanda se ha hecho un objetivo de cuantificación para el lado de decodificación.

En esta realización, se ha descrito un caso a modo de ejemplo en el que se realiza cuantificación de ganancia en unidades de región cada una compuestas de una pluralidad de subbandas. Un objetivo de cuantificación puede seleccionarse también en una base subbanda a subbanda - es decir, la determinación de si se tiene que llevar a cabo o no cuantificación predictiva puede realizarse también en una base subbanda a subbanda.

En esta realización, se ha descrito un caso a modo de ejemplo en el que el método de cuantificación predictiva de ganancia es realizar predicción lineal en el dominio de tiempo para ganancia de la misma banda de frecuencia, pero la presente invención no está limitada a esto, y puede realizarse también predicción lineal en el dominio de tiempo para ganancia de diferentes bandas de frecuencia.

En esta realización, se ha descrito un caso en el que una señal de habla/audio normal se toma como un ejemplo de una señal que se hace un objetivo de cuantificación, pero la presente invención no está limitada a esto, y una señal de excitación obtenida mediante procesamiento de una señal de habla/audio por medio de un filtro inverso de LPC (Coeficiente de Predicción Lineal) puede usarse también como un objetivo de cuantificación.

En esta realización, se ha descrito un caso a modo de ejemplo en el que una región para la que la magnitud de energía de región individual - es decir, significancia perceptual - es la mayor se selecciona como una referencia para seleccionar una banda objetivo de cuantificación, pero la presente invención no está limitada a esto, y además de significancia perceptual, la correlación de frecuencia con una banda seleccionada en una trama pasada puede tenerse en consideración también al mismo tiempo. Es decir, si existen bandas candidatas para las que el número de subbandas comunes a una banda objetivo de cuantificación seleccionada en el pasado es mayor que o igual a un valor predeterminado y la energía es mayor que o igual a un valor predeterminado, la banda con la energía más alta entre las bandas candidatas anteriores puede seleccionarse como la banda objetivo de cuantificación, y si no existen tales bandas candidatas, la banda con la energía más alta entre todas las bandas de frecuencia puede seleccionarse como la banda objetivo de cuantificación. Por ejemplo, si una subbanda común a la región de energía más alta y una banda seleccionada en una trama pasada no existen, el número de subbandas comunes a la segunda región de energía más alta y una banda seleccionada en una trama pasada son mayores que o iguales a un valor umbral predeterminado, y la energía de la segunda región de energía más alta es mayor que o igual a un valor umbral predeterminado, la segunda región de energía más alta se selecciona en lugar de la región de energía más alta. También, una sección de selección de banda de acuerdo con esta realización selecciona una región más cercana a una banda objetivo de cuantificación seleccionada en el pasado de entre las regiones cuya energía es mayor que o igual a un valor predeterminado como una banda objetivo de cuantificación.

En esta realización, la cuantificación de coeficiente de MDCT puede realizarse después de que se realice interpolación usando una trama pasada. Por ejemplo, se describirá un caso con referencia a la Figura 2 en el que una banda objetivo de cuantificación de trama pasada es la región 3 (es decir, subbandas 5 a 9), una banda objetivo de cuantificación de trama actual es la región 4 (es decir, subbandas 6 a 10), y se realiza codificación predictiva de trama actual usando un resultado de cuantificación de trama pasada. En un caso de este tipo, se realiza codificación predictiva en las subbandas 6 a 9 de trama actual usando las subbandas 6 a 9 de trama pasada, y para la subbanda 10 de trama actual, se interpola la subbanda 10 de trama pasada usando las subbandas 6 a 9 de trama pasada, y a continuación se realiza codificación predictiva usando la subbanda 10 de trama pasada obtenida mediante interpolación.

En esta realización, se ha descrito un caso a modo de ejemplo en el que se realiza cuantificación usando el mismo libro de códigos independientemente de si se realiza o no codificación predictiva, pero la presente invención no está limitada a esto, y pueden usarse también diferentes libros de códigos de acuerdo con si se realiza codificación predictiva o no se realiza en la cuantificación de ganancia y en la cuantificación de forma.

En esta realización, se ha descrito un caso a modo de ejemplo en el que todas las anchuras de subbanda son la misma, pero la presente invención no está limitada a esto, y las anchuras de subbanda individuales pueden variar también.

En esta realización, se ha descrito un caso a modo de ejemplo en el que el mismo libro de códigos se usa para todas las subbandas en la cuantificación de ganancia y en la cuantificación de forma, pero la presente invención no está limitada a esto, y pueden usarse también diferentes libros de códigos en una base subbanda a subbanda en la cuantificación de ganancia y en la cuantificación de forma.

En esta realización, se ha descrito un caso a modo de ejemplo en el que se seleccionan subbandas consecutivas como una banda objetivo de cuantificación. Una pluralidad de subbandas no consecutivas puede seleccionarse también como una banda objetivo de cuantificación. En un caso de este tipo, la eficacia de codificación de habla puede mejorarse adicionalmente interpolando un valor de subbanda no seleccionado usando valores de subbanda adyacentes.

En esta realización, se ha descrito un caso a modo de ejemplo en el que el aparato 100 de codificación de habla está equipado con la sección 104 de decisión de ejecución/no ejecución de codificación predictiva. Un aparato de codificación de habla no está limitado a esto y puede tener también una configuración en la que no se proporcione la sección 104 de decisión de ejecución/no ejecución de codificación predictiva y la cuantificación predictiva no siempre se realice mediante la sección 105 de cuantificación de ganancia, como se ilustra mediante el aparato 100a de codificación de habla mostrado en la Figura 4. En este caso, como se muestra en la Figura 4, el aparato 100a de codificación de habla está equipado con la sección 101 de transformación de dominio de frecuencia, la sección 102 de selección de banda, la sección 103 de cuantificación de forma, la sección 105 de cuantificación de ganancia y la sección 106 de multiplexación. La Figura 5 es un diagrama de bloques que muestra la configuración del aparato 200a de decodificación de habla que corresponde al aparato 100a de codificación de habla, estando equipado el aparato 200a de decodificación de habla con la sección 201 de demultiplexación, la sección 202 de descuantificación de forma, la sección 204 de descuantificación de ganancia y la sección 205 de transformación de dominio de tiempo. En un caso de este tipo, el aparato 100a de codificación de habla realiza selección parcial de una banda a cuantificar de entre todas las bandas, divide adicionalmente la banda seleccionada en una pluralidad de subbandas, y cuantifica la ganancia de cada subbanda. Por este medio, puede realizarse la cuantificación a una tasa de bits inferior que con un método en el cual los componentes de todas las bandas están cuantificados, y la eficacia de la codificación puede mejorarse. También, la eficacia de codificación puede mejorarse adicionalmente cuantificando un vector de ganancia usando correlación de ganancia en el dominio de frecuencia.

Un aparato de codificación de habla puede tener también una configuración en la que no se proporciona la sección 104 de decisión de ejecución/no ejecución de codificación predictiva y la cuantificación predictiva siempre se realiza mediante la sección 105 de cuantificación de ganancia, como se ilustra mediante el aparato 100a de codificación de habla mostrado en la Figura 4. La configuración del aparato 200a de decodificación de habla que corresponde a este tipo de aparato 100a de codificación de habla es como se muestra en la Figura 5. En un caso de este tipo, el aparato 100a de codificación de habla realiza selección parcial de una banda a cuantificar de entre todas las bandas, divide adicionalmente la banda seleccionada en una pluralidad de subbandas, y realiza cuantificación de ganancia para cada subbanda. Por este medio, puede realizarse la cuantificación a una tasa de bits inferior que con un método en el cual los componentes de todas las bandas están cuantificados, y la eficacia de codificación puede mejorarse. También, la eficacia de codificación puede mejorarse adicionalmente mediante la cuantificación predictiva de un vector de ganancia usando correlación de ganancia en el dominio de tiempo.

En esta realización, se ha descrito un caso a modo de ejemplo en el que el método de seleccionar una banda objetivo de cuantificación en una sección de selección de banda es seleccionar la región con la energía más alta en todas las bandas, pero la presente invención no está limitada a esto, y la selección puede realizarse también usando información de una banda seleccionada en una trama temporalmente anterior además del criterio anterior. Por ejemplo, un posible método es seleccionar una región a cuantificar después de realizar multiplicación por un factor de ponderación de manera que una región que incluye una banda en la proximidad de una banda seleccionada en una trama temporalmente anterior se hace más propensa a selección. También, si existe una pluralidad de capas en las que se selecciona una banda a cuantificar, una banda cuantificada en una capa superior puede seleccionarse usando información de una banda seleccionada en una capa inferior. Por ejemplo, un posible método es seleccionar una región a cuantificar después de realizar multiplicación por un factor de ponderación de manera que una región que incluye una banda en la proximidad de una banda seleccionada en una capa inferior se hace más propensa a selección.

En esta realización, se ha descrito un caso a modo de ejemplo en el que el método de seleccionar una banda objetivo de cuantificación es seleccionar la región con la energía más alta en todas las bandas, pero la presente invención no está limitada a esto, y una cierta banda también puede seleccionarse preliminarmente de antemano, después de lo cual se selecciona finalmente una banda objetivo de cuantificación en la banda preliminarmente seleccionada. En un caso de este tipo, una banda seleccionada preliminarmente puede decidirse de acuerdo con la tasa de muestro de señal de entrada, tasa de bits de codificación o similar. Por ejemplo, un método es seleccionar una banda baja preliminarmente cuando la tasa de bits o tasa de muestreo es baja.

Por ejemplo, es posible para un método a emplear en la sección 102 de selección de banda en la cual una región a cuantificar se decide calculando la energía de la región después de limitar regiones seleccionables a regiones de banda baja de entre todas las candidatas de región seleccionables. Como un ejemplo de esto, un posible método es realizar limitación a cinco candidatas del lado de banda baja de entre el total de ocho regiones candidatas mostradas en la Figura 2, y seleccionar la región con la energía más alta entre estas. Como alternativa, la sección 102 de selección de banda puede comparar energías después de multiplicar la energía por un factor de ponderación de modo que una región de área inferior se hace proporcionalmente más propensa a selección. Otra posibilidad es para la sección 102 de selección de banda seleccionar una subbanda del lado de banda baja fija. Una característica de una señal de habla es que la estructura de armónicos se hace proporcionalmente más intensa hacia el lado de la banda baja, como resultado de que un pico intenso está presente en el lado de banda baja. Ya que este pico intenso es difícil de enmascarar, es propenso a percibirse como ruido. En este punto, aumentando la probabilidad de selección hacia el lado de banda baja en lugar de seleccionar simplemente una región basándose en la magnitud de energía, la posibilidad de una región que incluya un pico intenso estando seleccionado aumenta, y una sensación de ruido se reduce como resultado. Por lo tanto, la calidad de una señal decodificada puede mejorarse limitando regiones seleccionadas al lado de banda baja, o realizando multiplicación por un factor de ponderación de manera que, de esta manera, la probabilidad de selección aumenta hacia el lado de banda baja.

Un aparato de codificación de habla de acuerdo con la presente invención se ha descrito en términos de una configuración en la cual la cuantificación de forma (información de forma) se realiza en primer lugar en un componente de una banda a cuantificar, seguido por cuantificación de ganancia (información de ganancia), pero la presente invención no está limitada a esto, y puede usarse también una configuración en la cual se realiza en primer lugar cuantificación de ganancia seguido por cuantificación de forma.

(Realización 2)

La Figura 6 es un diagrama de bloques que muestra la configuración principal del aparato 300 de codificación de habla de acuerdo con la realización 2 de la presente invención.

En esta figura, el aparato 300 de codificación de habla está equipado con la sección 301 de submuestreo, la sección 302 de codificación de primera capa, la sección 303 de decodificación de primera capa, la sección 304 de sobremuestreo, la primera sección 305 de transformación de dominio de frecuencia, la sección 306 de retardo, la segunda sección 307 de transformación de dominio de frecuencia, la sección 308 de codificación de segunda capa y la sección 309 de multiplexación, y tiene una configuración escalable que comprende dos capas. En la primera capa,

se aplica un método de codificación de habla de CELP (Predicción Lineal con Excitación por Código), y en la segunda capa se aplica el método de codificación de habla descrito en la realización 1 de la presente invención.

La sección 301 de submuestreo realiza procesamiento de submuestreo en una señal de habla/audio de entrada, para convertir la tasa de muestreo de la señal de habla/audio de la Tasa 1 a la Tasa 2 (donde Tasa 1 > Tasa 2), y emite esta señal a la sección 302 de codificación de primera capa.

La sección 302 de codificación de primera capa realiza codificación de habla de CELP en la entrada de señal de habla/audio posterior a submuestreo de la sección 301 de submuestreo, y emite la información codificada de primera capa a la sección 303 de decodificación de primera capa y a la sección 309 de multiplexación. Específicamente, la sección 302 de codificación de primera capa codifica una señal de habla que comprende información de tracto vocal e información de excitación encontrando un parámetro de LPC para la información de tracto vocal y para la información de excitación, realiza codificación encontrando un índice que identifica qué modelo de habla almacenado previamente se ha de usar, es decir, un índice que identifica qué vector de excitación de un libro de códigos adaptativo y libro de códigos fijo se ha de generar.

La sección 303 de decodificación de primera capa realiza decodificación de habla de CELP en la entrada de información codificada de primera capa de la sección 302 de codificación de primera capa, y emite una señal decodificada de primera capa obtenida a la sección 304 de sobremuestreo.

La sección 304 de sobremuestreo realiza procesamiento de sobremuestreo en la entrada de la señal decodificada de primera capa de la sección 303 de decodificación de primera capa, para convertir la tasa de muestreo de señal decodificada de primera capa de la Tasa 2 a la Tasa 2, y emite esta señal a la primera sección 305 de transformación de dominio de frecuencia.

La primera sección 305 de transformación de dominio de frecuencia realiza una MDCT en la entrada de señal decodificada de primera capa posterior a sobremuestreo de la sección 304 de sobremuestreo, y emite un coeficiente de MDCT de primera capa obtenido como un parámetro de dominio de frecuencia a la sección 308 de codificación de segunda capa. El método de transformación real usado en la primera sección 305 de transformación de dominio de frecuencia es similar al método de transformación usado en la sección 101 de transformación de dominio de frecuencia del aparato 100 de codificación de habla de acuerdo con la realización 1 de la presente invención, y por lo tanto se omite en este punto una descripción del mismo.

La sección 306 de retardo emite una señal de habla/audio retardada a la segunda sección 307 de transformación de dominio de frecuencia emitiendo una señal de habla/audio de entrada después de almacenar esa señal de entrada en una memoria intermedia interna durante un tiempo predeterminado. El tiempo de retardo predeterminado en este punto es un tiempo que tiene en cuenta el retardo del algoritmo que aparece en la sección 301 de submuestreo, la sección 302 de codificación de primera capa, la sección 303 de decodificación de primera capa, la sección 304 de sobremuestreo, la primera sección 305 de transformación de dominio de frecuencia y la segunda sección 307 de transformación de dominio de frecuencia.

La segunda sección 307 de transformación de dominio de frecuencia realiza una MDCT en la entrada de la señal de habla/audio retardada de la sección 306 de retardo, y emite un coeficiente de MDCT de segunda capa obtenido como un parámetro de dominio de frecuencia a la sección 308 de codificación de segunda capa. El método de transformación real usado en la segunda sección 307 de transformación de dominio de frecuencia es similar al método de transformación usado en la sección 101 de transformación de dominio de frecuencia del aparato 100 de codificación de acuerdo con la realización 1 de la presente invención, y por lo tanto se omite en este punto una descripción del mismo.

La sección 308 de codificación de segunda capa realiza codificación de segunda capa usando la entrada del coeficiente de MDCT de primera capa de la primera sección 305 de transformación de dominio de frecuencia y la entrada del coeficiente de MDCT de segunda capa de la segunda sección 307 de transformación de dominio de frecuencia, y emite información codificada de segunda capa obtenida a la sección 309 de multiplexación. La configuración interna principal y la operación real de la sección 308 de codificación de segunda capa se describirá más adelante en el presente documento.

La sección 309 de multiplexación multiplexa la entrada de información codificada de primera capa de la sección 302 de codificación de primera capa y la entrada de información codificada de segunda capa de la sección 308 de codificación de segunda capa, y transmite la corriente de bits obtenida a un aparato de decodificación de habla.

La Figura 7 es un diagrama de bloques que muestra la configuración principal del interior de la sección 308 de codificación de segunda capa. La sección 308 de codificación de segunda capa tiene una configuración básica similar a la del aparato 100 de codificación de habla de acuerdo con la realización 1 (véase la Figura 1), y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

La sección 308 de codificación de segunda capa se diferencia del aparato 100 de codificación de habla en que está equipada con la sección 381 de cálculo de coeficiente de MDCT residual en lugar de la sección 101 de transformación de dominio de frecuencia. El procesamiento mediante la sección 106 de multiplexación es similar al procesamiento mediante la sección 106 de multiplexación del aparato 100 de codificación de habla, y en beneficio de la descripción, el nombre de una salida de señal de la sección 106 de multiplexación de acuerdo con esta realización se proporciona como "información codificada de segunda capa".

La información de banda, información codificada de forma e información codificada de ganancia pueden introducirse también directamente a la sección 309 de multiplexación y multiplexarse con la información codificada de primera capa sin pasar a través de la sección 106 de multiplexación.

La sección 381 de cálculo de coeficiente de MDCT residual encuentra un residuo de la entrada del coeficiente de MDCT de primera capa de la primera sección 305 de transformación de dominio de frecuencia y de la entrada del coeficiente de MDCT de segunda capa de la segunda sección 307 de transformación de dominio de frecuencia, y emite este a la sección 102 de selección de banda como un coeficiente de MDCT residual.

La Figura 8 es un diagrama de bloques que muestra la configuración principal del aparato 400 de decodificación de habla de acuerdo con la realización 2 de la presente invención.

En esta figura, el aparato 400 de decodificación de habla está equipado con la sección 401 de control, la sección 402 de decodificación de primera capa, la sección 403 de sobremuestreo, la sección 404 de transformación de dominio de frecuencia, la sección 405 de decodificación de segunda capa, la sección 406 de transformación de dominio de tiempo y el conmutador 407.

La sección 401 de control analiza elementos de configuración de una corriente de bits transmitida desde el aparato 300 de codificación de habla, y de acuerdo con estos elementos de configuración de corriente de bits, emite adaptativamente información codificada apropiada a la sección 402 de decodificación de primera capa y a la sección 405 de decodificación de segunda capa, y emite también información de control al conmutador 407. Específicamente, si la corriente de bits comprende información codificada de primera capa e información codificada de segunda capa, la sección 401 de control emite la información codificada de primera capa a la sección 402 de decodificación de primera capa y emite la información codificada de segunda capa a la sección 405 de decodificación de segunda capa, mientras que si la corriente de bits comprende únicamente información codificada de primera capa, la sección 401 de control emite esta información codificada de primera capa a la sección 402 de decodificación de primera capa.

La sección 402 de decodificación de primera capa realiza decodificación de CELP en la entrada de la información codificada de primera capa de la sección 401 de control, y emite la señal decodificada de primera capa obtenida a la sección 403 de sobremuestreo y al conmutador 407.

La sección 403 de sobremuestreo realiza procesamiento de sobremuestreo en la entrada de señal decodificada de primera capa de la sección 402 de decodificación de primera capa, para convertir la tasa de muestreo de la señal decodificada de primera capa de la Tasa 2 a la Tasa 1, y emite esta señal a la sección 404 de transformación de dominio de frecuencia.

La sección 404 de transformación de dominio de frecuencia realiza una MDCT en la entrada de la señal decodificada de primera capa posterior a sobremuestreo de la sección 403 de sobremuestreo, y emite un coeficiente de MDCT decodificado de primera capa obtenido como un parámetro de dominio de frecuencia a la sección 405 de decodificación de segunda capa. El método de transformación real usado en la sección 404 de transformación de dominio de frecuencia es similar al método de transformación usado en la sección 101 de transformación de dominio de frecuencia del aparato 100 de codificación de habla de acuerdo con la realización 1, y por lo tanto se omite en este punto una descripción del mismo.

La sección 405 de decodificación de segunda capa realiza descuantificación de ganancia y descuantificación de forma usando la entrada de información codificada de segunda capa de la sección 401 de control y la entrada del coeficiente de MDCT decodificado de primera capa de la sección 404 de transformación de dominio de frecuencia, para obtener un coeficiente de MDCT decodificado de segunda capa. La sección 405 de decodificación de segunda capa añade juntos el coeficiente de MDCT decodificado de segunda capa obtenido y el coeficiente de MDCT decodificado de primera capa, y emite el resultado de adición obtenido a la sección 406 de transformación de dominio de tiempo como un coeficiente de MDCT de adición. La configuración interna principal y operación real de la sección 405 de decodificación de segunda capa se describirá más adelante en el presente documento.

La sección 406 de transformación de dominio de tiempo realiza una IMDCT en la entrada del coeficiente de MDCT de adición de la sección 405 de decodificación de segunda capa, y emite una señal decodificada de segunda capa obtenida como un componente de dominio de tiempo al conmutador 407.

Basándose en la entrada de información de control de la sección 401 de control, si la entrada de la corriente de bits al aparato 400 de decodificación de habla comprende información codificada de primera capa e información codificada de segunda capa, el conmutador 407 emite la entrada de la señal decodificada de segunda capa de la sección 406 de transformación de dominio de tiempo como una señal de salida, mientras que si la corriente de bits comprende únicamente información codificada de primera capa, el conmutador 407 emite la entrada de la señal decodificada de primera capa de la sección 402 de decodificación de primera capa como una señal de salida.

La Figura 9 es un diagrama de bloques que muestra la configuración principal del interior de la sección 405 de decodificación de segunda capa. La sección 405 de decodificación de segunda capa tiene una configuración básica similar a la del aparato 200 de decodificación de habla de acuerdo con la realización 1 (véase la Figura 3), y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

La sección 405 de decodificación de segunda capa se diferencia del aparato 200 de decodificación de habla en que está equipada adicionalmente con la sección 452 de cálculo de coeficiente de MDCT de adición. También, el procesamiento se diferencia en parte entre la sección 451 de demultiplexación de la sección 405 de decodificación de segunda capa y la sección 201 de demultiplexación del aparato 200 de decodificación de habla, y se asigna un código de referencia diferente para indicar esto.

La sección 451 de demultiplexación demultiplexa información de banda, información codificada de forma e información codificada de ganancia de la entrada de información codificada de segunda capa de la sección 401 de control, y emite la información de banda obtenida a la sección 202 de descuantificación de forma y a la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva, la información codificada de forma obtenida a la sección 202 de descuantificación de forma, y la información codificada de ganancia obtenida a la sección 204 de descuantificación de ganancia.

La sección 452 de cálculo de coeficiente de MDCT de adición añade juntos la entrada del coeficiente de MDCT decodificado de primera capa de la sección 404 de transformación de dominio de frecuencia y la entrada del coeficiente de MDCT decodificado de segunda capa de la sección 204 de descuantificación de ganancia, y emite el resultado de adición obtenido a la sección 406 de transformación de dominio de tiempo como un coeficiente de MDCT de adición.

Por lo tanto, de acuerdo con esta realización, cuando un componente de frecuencia de una banda diferente se hace un objetivo de cuantificación en cada trama, se realiza codificación predictiva de parámetro no temporal adaptativamente además de aplicar codificación escalable, posibilitando de esta manera que se reduzca la cantidad de información codificada en la codificación de habla, y que se reduzca el error de codificación de señal de habla/audio y la degradación de calidad de audio de señal decodificada.

En esta realización, se ha descrito un caso a modo de ejemplo en el que la sección 308 de codificación de segunda capa toma un componente de diferencia de un coeficiente de MDCT de primera capa y el coeficiente de MDCT de segunda capa como un objetivo de codificación, pero la presente invención no está limitada a esto, y la sección 308 de codificación de segunda capa puede tomar también un componente de diferencia de un coeficiente de MDCT de primera capa y el coeficiente de MDCT de segunda capa como un objetivo de codificación para una banda de una frecuencia predeterminada o por debajo, o puede tomar un propio coeficiente de MDCT de señal de entrada como un objetivo de codificación para una banda superior a una frecuencia predeterminada. Es decir, la conmutación puede realizarse entre el uso o no uso de un componente de diferencia de acuerdo con la banda.

En esta realización, se ha descrito un caso a modo de ejemplo en el que el método de seleccionar una banda objetivo de cuantificación de codificación de segunda capa es seleccionar la región para la cual la energía de un componente residual de un coeficiente de MDCT de primera capa y el coeficiente de MDCT de segunda capa es la más alta, pero la presente invención no está limitada a esto, y la región para la que la energía de coeficiente de MDCT de primera capa es la más alta puede seleccionarse también. Por ejemplo, la energía de cada subbanda de coeficiente de MDCT de primera capa puede calcularse, después de lo cual las energías de cada subbanda se añaden juntas en una base región a región, y la región para la cual la energía es la más alta se selecciona como una banda objetivo de cuantificación de codificación de segunda capa. En el lado del aparato de decodificación, la región para la que la energía es la más alta entre las regiones del coeficiente de MDCT decodificado de primera capa obtenido mediante la decodificación de primera capa se selecciona como una banda objetivo de descuantificación de decodificación de segunda capa. Por este medio puede reducirse la tasa de bits de codificación, puesto que la información de banda relacionada con una banda de cuantificación de codificación de segunda capa no se transmite del lado del aparato de codificación.

En esta realización, se ha descrito un caso a modo de ejemplo en el que la sección 308 de codificación de segunda capa selecciona y realiza cuantificación en una banda objetivo de cuantificación para un componente residual de un coeficiente de MDCT de primera capa y el coeficiente de MDCT de segunda capa, pero la presente invención no está limitada a esto, y la sección 308 de codificación de segunda capa puede predecir también un coeficiente de MDCT de segunda capa desde un coeficiente de MDCT de primera capa, y seleccionar y realizar cuantificación en

una banda objetivo de cuantificación para un componente residual de ese coeficiente de MDCT predicho y un coeficiente de MDCT de segunda capa real. Esto posibilita que se mejore adicionalmente la eficacia de codificación utilizando una correlación entre un coeficiente de MDCT de primera capa y el coeficiente de MDCT de segunda capa.

5
(Realización 3)

10 La Figura 10 es un diagrama de bloques que muestra la configuración principal del aparato 500 de codificación de habla de acuerdo con la realización 3 de la presente invención. El aparato 500 de codificación de habla tiene una configuración básica similar a la del aparato 100 de codificación de habla mostrado en la Figura 1, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

15 El aparato 500 de codificación de habla se diferencia del aparato 100 de codificación de habla en que está equipado adicionalmente con la sección 504 de cálculo de valor de interpolación. También, el procesamiento se diferencia en parte entre la sección 505 de cuantificación de ganancia del aparato 500 de codificación de habla y la sección 105 de cuantificación de ganancia del aparato 100 de codificación de habla, y se asigna un código de referencia diferente para indicar esto.

20 La sección 504 de cálculo de valor de interpolación tiene una memoria intermedia interna que almacena información de banda que indica una banda objetivo de cuantificación de una trama pasada. Usando un valor de ganancia de cuantificación de una banda objetivo de cuantificación de una trama pasada leída de la sección 505 de cuantificación de ganancia, la sección 504 de cálculo de valor de interpolación interpola un valor de ganancia de una banda que no se cuantificó en una trama pasada entre bandas objetivo de cuantificación de trama actual indicadas mediante la
25 entrada de información de banda de la sección 102 de selección de banda. La sección 504 de cálculo de valor de interpolación emite un valor de interpolación de ganancia obtenido a la sección 505 de cuantificación de ganancia.

30 La sección 505 de cuantificación de ganancia se diferencia de la sección 105 de cuantificación de ganancia del aparato 100 de codificación de habla en que usa una entrada de valor de interpolación de ganancia de la sección 504 de cálculo de valor de interpolación además de un valor de ganancia de cuantificación de trama pasada almacenado en una memoria intermedia interna y un libro de códigos de ganancia interno cuando realiza codificación predictiva.

35 El método de interpolación de valor de ganancia usado mediante la sección 504 de cálculo de valor de interpolación se describirá ahora en detalle.

40 La sección 504 de cálculo de valor de interpolación tiene una memoria intermedia interna que almacena la entrada de información de banda $m_{\text{máx}}$ de la sección 102 de selección de banda en una trama pasada. En este punto, se describirá un caso a modo de ejemplo en el que se proporciona una memoria intermedia interna que almacena información de banda $m_{\text{máx}}$ para las tres tramas pasadas.

45 La sección 504 de cálculo de valor de interpolación en primer lugar calcula un valor de ganancia distinto de una banda indicada mediante la información de banda $m_{\text{máx}}$ para las tres tramas pasadas realizando interpolación lineal. Un valor de interpolación se calcula de acuerdo con la Ecuación (19) para un valor de ganancia de banda inferior a la banda indicada mediante la información de banda $m_{\text{máx}}$, y un valor de interpolación se calcula de acuerdo con la Ecuación (20) para un valor de ganancia de una banda superior a la banda indicada mediante la información de banda $m_{\text{máx}}$.

$$\beta_0 \cdot q_0 + \beta_1 \cdot q_1 + \beta_2 \cdot q_2 + \beta_3 \cdot g = 0 \quad \dots \text{(Ecuación 19)}$$

50

$$\beta_0' \cdot q_0 + \beta_1' \cdot q_1 + \beta_2' \cdot q_2 + \beta_3' \cdot g = 0 \quad \dots \text{(Ecuación 20)}$$

55 En la Ecuación (19) y Ecuación (20) β_i indica un coeficiente de interpolación, q_i indica un valor de ganancia de una banda objetivo de cuantificación indicada mediante la información de banda $m_{\text{máx}}$ de una trama pasada, y g indica un valor de interpolación de ganancia de una banda adyacente no cuantificada a una banda objetivo de cuantificación indicada mediante la información de banda $m_{\text{máx}}$ de una trama pasada. En este punto, un valor inferior de i indica una banda de frecuencia inferior proporcionalmente, y en la Ecuación (19) g indica un valor de interpolación de ganancia de una banda adyacente en el lado de banda alta de una banda objetivo de cuantificación indicada mediante la información de banda $m_{\text{máx}}$ de una trama pasada, mientras que en la Ecuación (20) g indica
60 un valor de interpolación de ganancia de una banda adyacente en el lado de banda baja de una banda objetivo de cuantificación indicada mediante la información de banda $m_{\text{máx}}$ de una trama pasada. Para el coeficiente de interpolación β_i , se supone que se usa un valor que se ha encontrado de antemano estadísticamente para satisfacer la Ecuación (19) y Ecuación (20). En este punto, se describe un caso en el que se usan diferentes coeficientes de

interpolación β_i en la Ecuación (19) y Ecuación (20), pero un conjunto similar de coeficientes de predicción α_i puede usarse también en la Ecuación (19) y Ecuación (20).

Como se muestra en la Ecuación (19) y Ecuación (20), es posible interpolar un valor de ganancia de una banda en el lado de banda alta o en el lado de banda baja adyacente a una banda objetivo de cuantificación indicada mediante la información de banda de trama pasada $m_máx$ de una trama pasada en la sección 504 de cálculo de valor de interpolación. La sección 504 de cálculo de valor de interpolación interpola sucesivamente valores de ganancia de bandas no cuantificadas adyacentes repitiendo las operaciones en la Ecuación (19) y Ecuación (20) usando los resultados obtenidos de la Ecuación (19) y Ecuación (20).

De esta manera, la sección 504 de cálculo de valor de interpolación interpola valores de ganancia de bandas distintas de una banda indicada mediante la información de banda $m_máx$ de las tres tramas pasadas entre las bandas objetivo de cuantificación de trama actual indicadas mediante la entrada de la información de banda de la sección 102 de selección de banda, usando valores de ganancia cuantificados de las tres tramas pasadas leídas de la sección 505 de cuantificación de ganancia.

A continuación, se describirá una operación de codificación predictiva en la sección 505 de cuantificación de ganancia.

La sección 505 de cuantificación de ganancia realiza cuantificación prediciendo un valor de ganancia de trama actual usando un valor de ganancia de cuantificación de trama pasada almacenado, una entrada de valor de interpolación de ganancia de la sección 504 de cálculo de valor de interpolación y un libro de códigos de ganancia interno. Específicamente, la sección 505 de cuantificación de ganancia busca un libro de códigos de ganancia interno compuesto de la cantidad de GQ de vectores de código de ganancia para cada una de las L subbandas, y encuentra un índice de un vector de código de ganancia para el que el resultado de la Ecuación (21) a continuación es un mínimo

$$Ganancia_q(i) = \sum_{j=0}^{L-1} \left\{ Ganancia_i(j + j'') - \sum_{t=1}^3 (\alpha_t \cdot C_{j+j'}^t) - \alpha_0 \cdot GC_j \right\}^2 \quad (i = 0, \dots, GQ-1)$$

... (Ecuación 21)

En la Ecuación (21), GC_j^i indica un vector de código de ganancia que compone un libro de códigos de ganancia, i indica un índice de vector de código de ganancia y j indica un índice de un elemento de vector de código de ganancia. En este punto, C_j^t indica un valor de ganancia de cuantificación de t tramas anteriores en tiempo, de modo que cuando $t=1$, por ejemplo, C_j^1 indica un valor de ganancia de cuantificación de una trama anterior en tiempo. También, α es un coeficiente de predicción lineal de 4º orden almacenado en la sección 505 de cuantificación de ganancia. Un valor de interpolación de ganancia calculado de acuerdo con la Ecuación (19) y Ecuación (20) mediante la sección 504 de cálculo de valor de interpolación se usa como un valor de ganancia de una banda no seleccionada como una banda objetivo de cuantificación en las tres tramas pasadas. La sección 505 de cuantificación de ganancia trata L subbandas en una región como un vector L-dimensional y realiza cuantificación de vector.

La sección 505 de cuantificación de ganancia emite el índice de vector de código de ganancia $G_mín$ para el que el resultado de la Ecuación (21) anterior es un mínimo a la sección 106 de multiplexación como información codificada de ganancia. La sección 505 de cuantificación de ganancia actualiza también la memoria intermedia interna de acuerdo con la Ecuación (22) a continuación usando la información codificada de ganancia $G_mín$ y el valor de ganancia de cuantificación C_j^1 obtenido en la trama actual.

$$\begin{cases} C_{j+j''}^3 = C_{j+j''}^2 \\ C_{j+j''}^2 = C_{j+j''}^1 \\ C_{j+j''}^1 = GC_j^{G_mín} \end{cases} \quad (j = 0, \dots, L-1) \quad \dots \text{ (Ecuación 22)}$$

La Figura 11 es un diagrama de bloques que muestra la configuración principal del aparato 600 de decodificación de habla de acuerdo con la realización 3 de la presente invención. El aparato 600 de decodificación de habla tiene una configuración básica similar a la del aparato 200 de decodificación de habla mostrado en la Figura 3, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las

descripciones de los mismos en este punto.

El aparato 600 de decodificación de habla se diferencia del aparato 200 de decodificación de habla en que está equipado adicionalmente con la sección 603 de cálculo de valor de interpolación. También, el procesamiento se diferencia en parte entre la sección 604 de descuantificación de ganancia del aparato 600 de decodificación de habla y la sección 204 de descuantificación de ganancia del aparato 200 de decodificación de habla, y se asigna un código de referencia diferente para indicar esto.

La sección 603 de cálculo de valor de interpolación tiene una memoria intermedia interna que almacena información de banda que indica información de banda descuantificada en una trama pasada. Usando un valor de ganancia de una banda descuantificada en una trama pasada leída de la sección 604 de descuantificación de ganancia, la sección 603 de cálculo de valor de interpolación interpola un valor de ganancia de una banda que no se descuantificó en una trama pasada entre las bandas objetivo de cuantificación de trama actual indicadas mediante la entrada de información de banda de la sección 201 de demultiplexación. La sección 603 de cálculo de valor de interpolación emite un valor de interpolación de ganancia obtenido a la sección 604 de descuantificación de ganancia.

La sección 604 de descuantificación de ganancia se diferencia de la sección 204 de descuantificación de ganancia del aparato 200 de decodificación de habla en que usa una entrada de valor de interpolación de ganancia de la sección 603 de cálculo de valor de interpolación además de un valor de ganancia descuantificada de trama pasada almacenado y un libro de códigos de ganancia interno cuando realiza codificación predictiva.

El método de interpolación de valor de ganancia usado mediante la sección 603 de cálculo de valor de interpolación es similar al método de interpolación de valor de ganancia usado mediante la sección 504 de cálculo de valor de interpolación, y por lo tanto se omite en este punto una descripción detallada del mismo.

A continuación, se describirá una operación de decodificación predictiva en la sección 604 de descuantificación de ganancia.

La sección 604 de descuantificación de ganancia realiza descuantificación prediciendo un valor de ganancia de trama actual usando un valor de ganancia descuantificada almacenado en una trama pasada, una entrada de valor de ganancia de interpolación de la sección 603 de cálculo de valor de interpolación y un libro de códigos de ganancia interno. Específicamente, la sección 604 de descuantificación de ganancia obtiene el valor de ganancia Ganancia_q' realizando descuantificación de ganancia de acuerdo con la Ecuación (23) a continuación.

$$Ganancia_q'(j+j'') = \sum_{t=1}^3 (\alpha_t \cdot C''_{t+j''}) + \alpha_0 \cdot GC_j^{G-\min} \quad (j=0, \dots, L-1)$$

... (Ecuación 23)

En la Ecuación (23), C''_t indica un valor de ganancia de t tramas anteriores en tiempo, de modo que cuando t=1, por ejemplo, C''_1 indica un valor de ganancia de una trama anterior. También, α es un coeficiente de predicción lineal de 4° orden almacenado en la sección 604 de descuantificación de ganancia. Un valor de interpolación de ganancia calculado mediante la sección 603 de cálculo de valor de interpolación se usa como un valor de ganancia de una banda no seleccionada como un objetivo de cuantificación en las tres tramas pasadas. La sección 604 de descuantificación de ganancia trata L subbandas en una región como un vector L-dimensional, y realiza descuantificación de vector.

A continuación, la sección 604 de descuantificación de ganancia calcula un coeficiente de MDCT decodificado de acuerdo con la Ecuación (24) a continuación usando un valor de ganancia obtenido mediante la descuantificación de la trama actual y una entrada de valor de forma de la sección 202 de descuantificación de forma, y actualiza la memoria intermedia interna de acuerdo con la Ecuación (25) a continuación. En este punto, un coeficiente de MDCT decodificado calculado se indica mediante X''_k . También, en la descuantificación del coeficiente de MDCT, si k está presente en B(j'') a B(j''+1)-1, el valor de ganancia Ganancia_q'(j) toma el valor de Ganancia_q'(j'').

$$X''_k = Ganancia_q'(j) \cdot Forma_q'(k) \quad \left(\begin{array}{l} k = B(j''), \dots, B(j''+1)-1 \\ j = j'', \dots, j''+L-1 \end{array} \right)$$

... (Ecuación 24)

$$\begin{cases} C_{j'}^{m3} = C_{j'}^{m2} \\ C_{j'}^{m2} = C_{j'}^{m1} \\ C_{j'}^{m1} = \text{Ganancia_}q''(j) \end{cases} \quad (j = j'', \dots, j'' + L - 1) \quad \dots \text{ (Ecuación 25)}$$

Por lo tanto, de acuerdo con esta realización, cuando se realiza una cuantificación del parámetro de dominio de frecuencia de una banda objetivo de cuantificación diferente de cada trama, los valores de bandas no cuantificadas adyacentes se interpolan sucesivamente a partir de un valor cuantificado en una trama pasada, y se realiza cuantificación predictiva usando un valor de interpolación. En consecuencia, la precisión de codificación de la codificación de habla puede mejorarse adicionalmente.

En esta realización, se ha descrito un caso a modo de ejemplo en el que un coeficiente de interpolación fijo β encontrado de antemano se usa cuando se calcula un valor de interpolación de ganancia, pero la presente invención no está limitada a esto, y puede realizarse también interpolación después de ajustar el coeficiente de interpolación previamente encontrado β . Por ejemplo, un coeficiente de predicción puede ajustarse de acuerdo con la distribución de ganancia de una banda cuantificada en cada trama. Específicamente, es posible mejorar la precisión de codificación de la codificación de habla realizando ajuste de modo que un coeficiente de predicción se debilite y el factor de ponderación de la ganancia de trama actual se aumente cuando la variación en la ganancia cuantificada en cada trama sea grande.

En esta realización, se ha descrito un caso a modo de ejemplo en el que una pluralidad de bandas consecutivas (una región) que comprende una banda cuantificada en cada trama se hace un objetivo, pero la presente invención no está limitada a esto, y una pluralidad de regiones puede hacerse también un objetivo de cuantificación. En un caso de este tipo, es posible mejorar la precisión de codificación de la codificación de habla empleando un método en el cual se realiza predicción lineal de valores finales de las respectivas regiones para una banda entre regiones seleccionadas además del método de interpolación de acuerdo con la Ecuación (19) y Ecuación (20).

(Realización 4)

La Figura 12 es un diagrama de bloques que muestra la configuración principal del aparato 700 de codificación de habla de acuerdo con la realización 4 de la presente invención. El aparato 700 de codificación de habla tiene una configuración básica similar a la del aparato 100 de codificación de habla mostrado en la Figura 1, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

El aparato 700 de codificación de habla se diferencia del aparato 100 de codificación de habla en que está equipado adicionalmente con la sección 704 de decisión de coeficiente de predicción. También, el procesamiento se diferencia en parte entre la sección 705 de cuantificación de ganancia del aparato 700 de codificación de habla y la sección 105 de cuantificación de ganancia del aparato 100 de codificación de habla, y se asigna un código de referencia diferente para indicar esto.

La sección 704 de decisión de coeficiente de predicción tiene una memoria intermedia interna que almacena información de banda que indica una banda objetivo de cuantificación de una trama pasada, decide un coeficiente de predicción a usarse en la cuantificación de la sección 705 de cuantificación de ganancia basándose en la información de banda de trama pasada, y emite un coeficiente de predicción decidido a la sección 705 de cuantificación de ganancia.

La sección 705 de cuantificación de ganancia se diferencia de la sección 105 de cuantificación de ganancia del aparato 100 de codificación de habla en que usa una entrada de coeficiente de predicción de la sección 704 de decisión de coeficiente de predicción en lugar de un coeficiente de predicción decidido de antemano cuando realiza codificación predictiva.

Se describirá ahora una operación de decisión de coeficiente de predicción en la sección 704 de decisión de coeficiente de predicción.

La sección 704 de decisión de coeficiente de predicción tiene una memoria intermedia interna que almacena la entrada de información de banda $m_{\text{máx}}$ de la sección 102 de selección de banda en una trama pasada. En este punto, se describirá un caso a modo de ejemplo en el que se proporciona una memoria intermedia interna que almacena información de banda $m_{\text{máx}}$ para las tres tramas pasadas.

Usando la información de banda $m_{\text{máx}}$ almacenada en la memoria intermedia interna y la entrada de información de banda $m_{\text{máx}}$ de la sección 102 de selección de banda en la trama actual, la sección 704 de decisión de

coeficiente de predicción encuentra un número de subbandas comunes a una banda objetivo de cuantificación de trama actual y banda objetivo de cuantificación de trama pasada. La sección 704 de decisión de coeficiente de predicción decide los coeficientes de predicción para ser el conjunto A y emite este a la sección 705 de cuantificación de ganancia si el número de subbandas comunes es mayor que o igual a un valor predeterminado, o decide los coeficientes de predicción para ser el conjunto B y emite este a la sección 705 de cuantificación de ganancia si el número de subbandas comunes es menor que el valor predeterminado. En este punto, el conjunto de coeficiente de predicción A es un conjunto de parámetro que enfatiza un valor más de trama pasada, y hace el factor de ponderación de un valor de ganancia de trama pasada más grande, que en el caso del conjunto de coeficiente de predicción B. Por ejemplo, en el caso de los coeficientes de predicción de 4º orden, es posible para el conjunto A decidirse como ($\alpha a_0=0,60$, $\alpha a_1=0,25$, $\alpha a_2=0,10$, $\alpha a_3=0,05$), y para el conjunto B decidirse como ($\alpha b_0=0,80$, $\alpha b_1=0,10$, $\alpha b_2=0,05$, $\alpha b_3=0,05$),

A continuación la sección 704 de decisión de coeficiente de predicción actualiza la memoria intermedia interna usando la entrada de información de banda $m_{\text{máx}}$ de la sección 102 de selección de banda en la trama actual.

A continuación, se describirá una operación de codificación predictiva en la sección 705 de cuantificación de ganancia.

La sección 705 de cuantificación de ganancia tiene una memoria intermedia interna que almacena un valor de ganancia de cuantificación obtenido en una trama pasada. La sección 705 de cuantificación de ganancia realiza cuantificación prediciendo un valor de ganancia de trama actual usando una entrada de coeficiente de predicción de la sección 704 de decisión de coeficiente de predicción y el valor de ganancia de cuantificación de trama pasada C_j^t almacenado en la memoria intermedia interna. Específicamente, la sección 705 de cuantificación de ganancia busca un libro de códigos de ganancia interno compuesto de la cantidad de GQ de vectores de código de ganancia para cada una de las L subbandas, y encuentra un índice de un vector de código de ganancia para el que el resultado de la Ecuación (26) a continuación es un mínimo si un coeficiente de predicción es el conjunto A, o encuentra un índice de un vector de código de ganancia para el que el resultado de la Ecuación (27) a continuación es un mínimo si un coeficiente de predicción es el conjunto B.

$$Ganancia_q(i) = \sum_{j=0}^{L-1} \left\{ Ganancia_i(j + j'') - \sum_{r=1}^3 (\alpha a_r \cdot C_{j,r}^t) - \alpha a_0 \cdot GC_j \right\}^2 \quad (i=0, \dots, GQ-1)$$

... (Ecuación 26)

$$Ganancia_q(i) = \sum_{j=0}^{L-1} \left\{ Ganancia_i(j + j'') - \sum_{r=1}^3 (\alpha b_r \cdot C_{j,r}^t) - \alpha b_0 \cdot GC_j \right\}^2 \quad (i=0, \dots, GQ-1)$$

... (Ecuación 27)

En la Ecuación (26) y Ecuación (27), GC_j^t indica un vector de código de ganancia que compone un libro de códigos de ganancia, i indica un índice de vector de código de ganancia, y j indica un índice de un elemento de vector de código de ganancia. En este punto, C_j^t indica un valor de ganancia de t tramas anteriores en tiempo, de modo que cuando $t=1$, por ejemplo, C_j^1 indica un valor de ganancia de una trama anterior en tiempo. También, α es un coeficiente de predicción lineal de 4º orden almacenado en la sección 705 de cuantificación de ganancia. La sección 705 de cuantificación de ganancia trata L subbandas en una región como un vector L-dimensional, y realiza cuantificación de vector. Si no existe valor de ganancia de una subbanda que corresponde a una trama pasada en la memoria intermedia interna, la sección 705 de cuantificación de ganancia sustituye el valor de ganancia de la subbanda más próxima en frecuencia en la memoria intermedia interna en la Ecuación (26) o Ecuación (27) anteriores.

La Figura 13 es un diagrama de bloques que muestra la configuración principal del aparato 800 de decodificación de habla de acuerdo con la realización 4 de la presente invención. El aparato 800 de decodificación de habla tiene una configuración básica similar a la del aparato 200 de decodificación de habla mostrado en la Figura 3, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

El aparato 800 de decodificación de habla se diferencia del aparato 200 de decodificación de habla en que está equipado adicionalmente con la sección 803 de decisión de coeficiente de predicción. También, el procesamiento se diferencia en parte entre la sección 804 de descuantificación de ganancia del aparato 800 de decodificación de habla y la sección 204 de descuantificación de ganancia del aparato 200 de decodificación de habla, y se asigna un código de referencia diferente para indicar esto.

La sección 803 de decisión de coeficiente de predicción tiene una memoria intermedia interna que almacena la entrada de información de banda de la sección 201 de demultiplexación en una trama pasada, decide un coeficiente de predicción a usarse en la cuantificación de la sección 804 de descuantificación de ganancia basándose en la información de banda de trama pasada, y emite un coeficiente de predicción decidido a la sección 804 de descuantificación de ganancia.

La sección 804 de descuantificación de ganancia se diferencia de la sección 204 de descuantificación de ganancia del aparato 200 de decodificación de habla en que usa una entrada de coeficiente de predicción de la sección 803 de decisión de coeficiente de predicción en lugar de un coeficiente de predicción decidido de antemano cuando realiza decodificación predictiva.

El método de decisión de coeficiente de predicción usado mediante la sección 803 de decisión de coeficiente de predicción es similar al método de decisión de coeficiente de predicción usado mediante la sección 704 de decisión de coeficiente de predicción del aparato 700 de codificación de habla, y por lo tanto una descripción detallada de la operación de la sección 803 de decisión de coeficiente de predicción se omite en este punto.

A continuación, se describirá una operación de decodificación predictiva en la sección 804 de descuantificación de ganancia.

La sección 804 de descuantificación de ganancia tiene una memoria intermedia interna que almacena un valor de ganancia obtenido en una trama pasada. La sección 804 de descuantificación de ganancia realiza descuantificación prediciendo un valor de ganancia de trama actual usando una entrada de coeficiente de predicción de la sección 803 de decisión de coeficiente de predicción y un valor de ganancia de trama pasada almacenado en la memoria intermedia interna. Específicamente, la sección 804 de descuantificación de ganancia tiene el mismo tipo de libro de códigos de ganancia interno que la sección 705 de cuantificación de ganancia del aparato 700 de codificación de habla, y obtiene el valor de ganancia $Ganancia_q'$ realizando descuantificación de ganancia de acuerdo con la Ecuación (28) a continuación si una entrada del coeficiente de predicción de la sección 803 de decisión de coeficiente de predicción es el conjunto A, o de acuerdo con la Ecuación (29) a continuación si el coeficiente de predicción es el conjunto B.

$$Ganancia_q'(j + j'') = \sum_{t=1}^3 (\alpha a_t \cdot C'''_{j+j''}) + \alpha a_0 \cdot GC_j^G - \min \quad (j = 0, \dots, L-1)$$

... (Ecuación 28)

$$Ganancia_q'(j + j'') = \sum_{t=1}^3 (\alpha b_t \cdot C'''_{j+j''}) + \alpha b_0 \cdot GC_j^G - \min \quad (j = 0, \dots, L-1)$$

... (Ecuación 29)

En la Ecuación (28) y Ecuación (29), C'''_j indica un valor de ganancia de t tramas anteriores en tiempo, de modo que cuando $t=1$, por ejemplo, C'''_j indica un valor de ganancia de una trama anterior. También, αa_i y αb_i indican la entrada del conjunto de coeficiente de predicción A y del conjunto B de la sección 803 de decisión de coeficiente de predicción. La sección 804 de descuantificación de ganancia trata L subbandas en una región como un vector L-dimensional, y realiza descuantificación de vector.

Por lo tanto, de acuerdo con esta realización, cuando se realiza cuantificación de parámetro de dominio de frecuencia de una banda objetivo de cuantificación diferente de cada trama, se realiza codificación predictiva seleccionado, a partir de una pluralidad de conjuntos de coeficientes de predicción, un conjunto de coeficiente de predicción que hace el factor de ponderación de un valor de ganancia de trama pasada proporcionalmente más grande que el mayor del número de subbandas comunes a una banda objetivo de cuantificación de trama pasada y banda objetivo de cuantificación de trama actual. En consecuencia, la precisión de codificación de la codificación de habla puede mejorarse adicionalmente.

En esta realización, se ha descrito un caso a modo de ejemplo en el que se proporcionan dos tipos de conjuntos de coeficientes de predicción de antemano, y un coeficiente de predicción usado en codificación predictiva se cambia de acuerdo con el número de subbandas comunes a una banda objetivo de cuantificación de trama pasada y banda objetivo de cuantificación de trama actual, pero la presente invención no está limitada a esto, y tres o más tipos de conjuntos de coeficientes de predicción pueden proporcionarse también de antemano.

En esta realización, se ha descrito un caso a modo de ejemplo en el que, si una banda objetivo de cuantificación en la trama actual no se ha cuantificado en una trama pasada, el valor de la banda más cercana en una trama pasada se sustituye, pero la presente invención no está limitada a esto, y si un valor de banda objetivo de cuantificación en la trama actual no se ha cuantificado en una trama pasada, la codificación predictiva puede realizarse también tomando el coeficiente de predicción de trama pasada relevante como cero, añadiendo un coeficiente de predicción de esa trama a un coeficiente de predicción de trama actual, calculando un nuevo conjunto de coeficiente de predicción y usando estos coeficientes de predicción. Por este medio, el efecto de la codificación predictiva puede cambiarse más flexiblemente y la precisión de codificación de la codificación de habla puede mejorarse adicionalmente.

(Realización 5)

La Figura 14 es un diagrama de bloques que muestra la configuración principal del aparato 1000 de codificación de habla de acuerdo con la realización 5 de la presente invención. El aparato 1000 de codificación de habla tiene una configuración básica similar a la del aparato 300 de codificación de habla mostrado en la Figura 6, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

El aparato 1000 de codificación de habla se diferencia del aparato 300 de codificación de habla en que está equipado adicionalmente con la sección 1007 de codificación de mejora de banda. También, el procesamiento se diferencia en parte entre la sección 1008 de codificación de segunda capa y la sección 1009 de multiplexación del aparato 1000 de codificación de habla y la sección 308 de codificación de segunda capa y la sección 309 de multiplexación del aparato 300 de codificación de habla, y se asignan diferentes códigos de referencia para indicar esto.

La sección 1007 de codificación de mejora de banda realiza codificación de mejora de banda usando una entrada de coeficiente de MDCT de primera capa de la primera sección 305 de transformación de dominio de frecuencia y una entrada de coeficiente de MDCT de entrada de la segunda sección 307 de transformación de dominio de frecuencia, y emite la información codificada de mejora de banda obtenida a la sección 1009 de multiplexación.

La sección 1009 de multiplexación se diferencia de la sección 309 de multiplexación únicamente en que multiplexa también información codificada de mejora de banda además de información codificada de primera capa e información codificada de segunda capa.

La Figura 15 es un diagrama de bloques que muestra la configuración principal del interior de la sección 1007 de codificación de mejora de banda.

En la Figura 15, la sección 1007 de codificación de mejora de banda está equipada con la sección 1071 de estimación de espectro de banda alta y la sección 1072 de codificación de factor de escala correctiva.

La sección 1071 de estimación de espectro de banda alta estima un espectro de banda alta de las bandas de señal FL a FH usando un espectro de banda baja de las bandas de señal 0 a FL de una entrada de coeficiente de MDCT de entrada de la segunda sección 307 de transformación de dominio de frecuencia, para obtener un espectro estimado. El método de derivación de espectro estimado es encontrar un espectro estimado de manera que el grado de similitud con el espectro de banda alta se hace un máximo transformando el espectro de banda baja basándose en este espectro de banda baja. La sección 1071 de estimación de espectro de banda alta codifica información relacionada con este espectro estimado (información de estimación), emite un parámetro de codificación obtenido, y proporciona también el propio espectro estimado a la sección 1072 de codificación de factor de escala correctiva.

En la siguiente descripción, una salida de espectro estimado de la sección 1071 de estimación de espectro de banda alta se denomina un primer espectro, y una salida de coeficiente de MDCT de primera capa (espectro de banda alta) de la primera sección 305 de transformación de dominio de frecuencia se denomina un segundo espectro.

Los tipos de espectros anteriormente descritos y las bandas de señal correspondientes pueden resumirse como sigue.

Espectro de banda estrecha (espectro de banda baja) ... 0 a FL espectro de banda ancha ... 0 a FH primer espectro (espectro estimado) ... FL a FH segundo espectro (espectro de banda alta) ... FL a FH

La sección 1072 de codificación de factor de escala correctiva corrige un factor de escala de primer espectro de modo que el factor de escala de primer espectro se acerca a un factor de escala de segundo espectro, y codifica y emite información relacionada con este factor de escala correctiva.

La salida de información codificada de mejora de banda de la sección 1007 de codificación de mejora de banda a la sección 1009 de multiplexación incluye una salida del parámetro de codificación de información de estimación de la sección 1071 de estimación de espectro de banda alta y una salida del parámetro de codificación de factor de escala

correctiva de la sección 1072 de codificación de factor de escala correctiva.

La Figura 16 es un diagrama de bloques que muestra la configuración principal del interior de la sección 1072 de codificación de factor de escala correctiva.

La sección 1072 de codificación de factor de escala correctiva está equipada con las secciones 1721 y 1722 de cálculo de factor de escala, el libro de códigos 1723 de factor de escala correctiva, el multiplicador 1724, el restador 1725, la sección 1726 de determinación, la sección 1727 de cálculo de error de ponderación y la sección 1728 de búsqueda. Estas secciones realizan las siguientes operaciones.

La sección 1721 de cálculo de factor de escala divide las bandas de señal del segundo espectro de entrada FL a FH en una pluralidad de subbandas, encuentra el tamaño de un espectro incluido en cada subbanda, y emite este al restador 1725. Específicamente, la división en subbandas se realiza asociada con una banda crítica, y se realiza la división en intervalos iguales en la escala de Bark. También, la sección 1721 de cálculo de factor de escala encuentra la amplitud media de los espectros incluidos en las subbandas, y toma este como el segundo factor de escala $SF2(k)$ $\{0 \leq k < NB\}$, donde NB representa el número de subbandas. Un valor de amplitud máxima o similar puede usarse en lugar de una amplitud media.

La sección 1722 de cálculo de factor de escala divide las bandas de señal del primer espectro de entrada FL a FH en una pluralidad de subbandas, calcula el primer factor de escala $SF1(k)$ $\{0 \leq k < NB\}$ de las subbandas, y emite este al multiplicador 1724. Como con la sección 1721 de cálculo de factor de escala, un valor de amplitud máxima o similar puede usarse en lugar de una amplitud media.

En el posterior procesamiento, los parámetros en la pluralidad de subbandas se integran en un valor de vector. Por ejemplo, la cantidad de NB de los factores de escala se representa como un vector. Se proporcionará una descripción tomando un caso en el que cada operación de procesamiento se realiza para cada uno de estos vectores como un ejemplo - es decir, un caso en el que se realiza cuantificación de vector.

El libro de códigos 1723 de factor de escala correctiva almacena una pluralidad de candidatos de factor de escala correctiva, y emite secuencialmente uno de los candidatos de factor de escala correctiva almacenados al multiplicador 1724 de acuerdo con una directiva de la sección 1728 de búsqueda. La pluralidad de candidatos de factor de escala correctiva almacenados en el libro de códigos 1723 de factor de escala correctiva se representa mediante un vector.

El multiplicador 1724 multiplica una primera salida de factor de escala de la sección 1722 de cálculo de factor de escala mediante una salida de candidato de factor de escala correctiva del libro de códigos 1723 de factor de escala correctiva, y proporciona el resultado de la multiplicación al restador 1725.

El restador 1725 resta la salida del multiplicador 1724 - es decir, el producto del primer factor de escala y el factor de escala correctiva - de la segunda salida del factor de escala de la sección 1721 de cálculo de factor de escala, y proporciona una señal de error obtenida de esta manera a la sección 1727 de cálculo de error de ponderación y a la sección 1726 de determinación.

La sección 1726 de determinación decide un vector de ponderación para proporcionarse a la sección 1727 de cálculo de error de ponderación basándose en el signo de la señal de error proporcionada del restador 1725. Específicamente, la señal de error $d(k)$ proporcionada del restador 1725 se representa mediante la Ecuación (30) a continuación.

$$d(k) = SF2(k) - v_i(k) \cdot SF1(k) \quad (0 \leq k < NB) \quad \dots \text{ (Ecuación 30)}$$

En este punto, $v_i(k)$ representa un iésimo candidato de factor de escala correctiva. La sección 1726 de determinación comprueba el signo de $d(k)$, selecciona w_{pos} como un factor de ponderación si $d(k)$ es positivo, o selecciona w_{neg} como un factor de ponderación si $d(k)$ es negativo, y emite el vector de ponderación $w(k)$ compuesto de estos a la sección 1727 de cálculo de error de ponderación. Estos factores de ponderación tienen la relación de tamaño relativo mostrada en la Ecuación (31) a continuación.

$$0 < w_{pos} < w_{neg} \quad \dots \text{ (Ecuación 31)}$$

Por ejemplo, si el número de subbandas $NB=4$, y los signos de $d(k)$ son $\{+, -, -, +\}$, la salida del vector de ponderación $w(k)$ a la sección 1727 de cálculo de error de ponderación se representa mediante $w(k) = \{w_{pos}, w_{neg}, w_{neg}, w_{pos}\}$.

La sección 1727 de cálculo de error de ponderación en primer lugar calcula el cuadrado de la señal de error proporcionada del restador 1725, y a continuación multiplica el vector de ponderación $w(k)$ proporcionado de la sección 1726 de determinación por el cuadrado de la señal de error para calcular el error cuadrático ponderado E , y proporciona el resultado de este cálculo a la sección 1728 de búsqueda. En este punto, el error cuadrático ponderado E se representa como se muestra en la Ecuación (32) a continuación.

$$E = \sum_{k=0}^{NB-1} w(k) \cdot d(k)^2 \quad \dots \text{ (Ecuación 32)}$$

La sección 1728 de búsqueda controla el libro de códigos 1723 de factor de escala correctiva y emite secuencialmente candidatos de factor de escala correctiva almacenados, y por medio de procesamiento de bucle cerrado encuentra un candidato de factor de escala correctiva para el que la salida del error cuadrático ponderado E de la sección 1727 de cálculo de error de ponderación es un mínimo. La sección 1728 de búsqueda emite el índice i_{opt} del candidato de factor de escala correctiva encontrado como un parámetro de codificación.

Cuando se establece un factor de ponderación usado cuando se calcula el error cuadrático ponderado E de acuerdo con el signo de una señal de error y el tipo de relación mostrada en la Ecuación (30) aplica a ese factor de ponderación, como se ha descrito anteriormente, se obtiene el siguiente tipo de efecto. En concreto, un caso en el que la señal de error $d(k)$ es positiva es un caso en el que un valor decodificado generado en el lado de decodificación (en términos del lado de codificación, un valor obtenido multiplicando un primer factor de escala por un factor de escala correctiva) es más pequeño que un segundo factor de escala, que es el valor objetivo. También, un caso en el que la señal de error $d(k)$ es negativa es un caso en el que un valor decodificado generado en el lado de decodificación es mayor que un segundo factor de escala, que es el valor objetivo. Por lo tanto, estableciendo un factor de ponderación cuando la señal de error $d(k)$ es positiva para que sea más pequeña que un factor de ponderación cuando la señal de error $d(k)$ es negativa, cuando los valores de error cuadrático son del mismo orden un candidato de factor de escala correctiva que genera un valor decodificado más pequeño que un segundo factor de escala se hace propenso a seleccionarse.

El siguiente tipo de efecto de mejora se obtiene mediante el procesamiento de la sección 1007 de codificación de mejora de banda. Por ejemplo, cuando se estima un espectro de banda alta usando un espectro de banda baja, como en esta realización, puede conseguirse generalmente una tasa de bits inferior. Sin embargo, mientras que puede conseguirse una tasa de bits inferior, la precisión de un espectro estimado - es decir, la similitud entre un espectro estimado y un espectro de banda alta - no puede decirse que es suficientemente alta, como se ha descrito anteriormente. En un caso de este tipo, si un valor decodificado de factor de escala se hace mayor que un valor objetivo, y un factor de escala posterior a cuantificación opera en la dirección de intensificación de un espectro estimado, la baja precisión del espectro estimado tiende a ser perceptible para el oído humano como degradación de calidad. A la inversa, cuando un valor decodificado de factor de escala se hace más pequeño que un valor objetivo, y un factor de escala posterior a cuantificación opera en la dirección de atenuación de este espectro estimado, la baja precisión del espectro estimado deja de ser perceptible, y se obtiene un efecto de mejorar la calidad de audio de la señal decodificada. Esta tendencia se ha confirmado también en una simulación por ordenador.

La Figura 17 es un diagrama de bloques que muestra la configuración principal del interior de la sección 1008 de codificación de segunda capa. La sección 1008 de codificación de segunda capa tiene una configuración básica similar a la de la sección 308 de codificación de segunda capa mostrada en la Figura 7, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto. El procesamiento se diferencia en parte entre la sección 1081 de cálculo de coeficiente de MDCT residual de la sección 1008 de codificación de segunda capa y la sección 381 de cálculo de coeficiente de MDCT residual de la sección 308 de codificación de segunda capa, y se asigna un código de referencia diferente para indicar esto.

La sección 1081 de cálculo de coeficiente de MDCT residual calcula un MDCT residual que es para ser un objetivo de cuantificación en la sección de codificación de segunda capa de un coeficiente de MDCT de entrada y coeficiente de MDCT de mejora de primera capa. La sección 1081 de cálculo de coeficiente de MDCT residual se diferencia de la sección 381 de cálculo de coeficiente de MDCT residual de acuerdo con la realización 2 en que toma un residuo del coeficiente de MDCT de entrada y del coeficiente de MDCT de mejora de primera capa como un coeficiente de MDCT residual para una banda no mejorada mediante la sección 1007 de codificación de mejora de banda y tomando un propio coeficiente de MDCT de entrada, en lugar de un residuo, como un coeficiente de MDCT residual para una banda mejorada mediante la sección 1007 de codificación de mejora de banda.

La Figura 18 es un diagrama de bloques que muestra la configuración principal del aparato 1010 de decodificación de habla de acuerdo con la realización 5 de la presente invención. El aparato 1010 de decodificación de habla tiene una configuración básica similar a la del aparato 400 de decodificación de habla mostrado en la Figura 8, y por lo

tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

El aparato 1010 de decodificación de habla se diferencia del aparato 400 de decodificación de habla en que está equipado adicionalmente con la sección 1012 de decodificación de mejora de banda y la sección 1013 de transformación de dominio de tiempo. También, el procesamiento se diferencia en parte entre la sección 1011 de control, la sección 1015 de decodificación de segunda capa y el conmutador 1017 del aparato 1010 de decodificación de habla y la sección 401 de control, la sección 405 de decodificación de segunda capa y el conmutador 407 del aparato 400 de decodificación de habla, y se asignan diferentes códigos de referencia para indicar esto.

La sección 1011 de control analiza elementos de configuración de una corriente de bits transmitida desde el aparato 1000 de codificación de habla, y de acuerdo con estos elementos de configuración de corriente de bits, emite adaptativamente información codificada apropiada a la sección 402 de decodificación de primera capa, a la sección 1012 de decodificación de mejora de banda y a la sección 1015 de decodificación de segunda capa, y emite también información de control al conmutador 1017. Específicamente, si la corriente de bits comprende información codificada de primera capa, información codificada de mejora de banda e información codificada de segunda capa, la sección 1011 de control emite la información codificada de primera capa a la sección 402 de decodificación de primera capa, emite la información codificada de mejora de banda a la sección 1012 de decodificación de mejora de banda, y emite la información codificada de segunda capa a la sección 1015 de decodificación de segunda capa. Si la corriente de bits comprende únicamente información codificada de primera capa y la información codificada de mejora de banda, la sección 1011 de control emite la información codificada de primera capa a la sección 402 de decodificación de primera capa, y emite la información codificada de mejora de banda a la sección 1012 de decodificación de mejora de banda. Si la corriente de bits comprende únicamente información codificada de primera capa, la sección 1011 de control emite esta información codificada de primera capa a la sección 402 de decodificación de primera capa. También, la sección 1011 de control emite información de control que controla el conmutador 1017 al conmutador 1017.

La sección 1012 de decodificación de mejora de banda realiza procesamiento de mejora de banda usando la entrada de información codificada de mejora de banda de la sección 1011 de control y una entrada de coeficiente de MDCT decodificado de primera capa de la sección 404 de transformación de dominio de frecuencia, para obtener un coeficiente de MDCT de mejora de primera capa. A continuación la sección 1012 de decodificación de mejora de banda emite el coeficiente de MDCT de mejora de primera capa obtenido a la sección 1013 de transformación de dominio de tiempo y a la sección 1015 de decodificación de segunda capa. La configuración interna principal y operación real de la sección 1012 de decodificación de mejora de banda se describirá más adelante en el presente documento.

La sección 1013 de transformación de dominio de tiempo realiza una IMDCT en la entrada del coeficiente de MDCT de mejora de primera capa de la sección 1012 de decodificación de mejora de banda, y emite una señal decodificada de mejora de primera capa obtenida como un componente de dominio de tiempo al conmutador 1017.

La sección 1015 de decodificación de segunda capa realiza descuantificación de ganancia y descuantificación de forma usando la entrada de información codificada de segunda capa de la sección 1011 de control y la entrada de coeficiente de MDCT de mejora de primera capa de la sección 1012 de decodificación de mejora de banda, para obtener un coeficiente de MDCT decodificado de segunda capa. La sección 1015 de decodificación de segunda capa añade juntos el coeficiente de MDCT decodificado de segunda capa obtenido y el coeficiente de MDCT decodificado de primera capa, y emite el resultado de la adición obtenido a la sección 406 de transformación de dominio de tiempo como un coeficiente de MDCT de adición. La configuración interna principal y operación real de la sección 1015 de decodificación de segunda capa se describirá más adelante en el presente documento.

Basándose en la entrada de información de control de la sección 1011 de control, si la entrada de corriente de bits al aparato 1010 de decodificación de habla comprende información codificada de primera capa, información codificada de mejora de banda e información codificada de segunda capa, el conmutador 1017 emite la entrada de señal decodificada de segunda capa de la sección 406 de transformación de dominio de tiempo como una señal de salida. Si la corriente de bits comprende únicamente información codificada de primera capa e información codificada de mejora de banda, el conmutador 1017 emite la entrada de señal decodificada de mejora de primera capa de la sección 1013 de transformación de dominio de tiempo como una señal de salida. Si la corriente de bits comprende únicamente información codificada de primera capa, el conmutador 1017 emite la entrada de señal decodificada de primera capa de la sección 402 de decodificación de primera capa como una señal de salida.

La Figura 19 es un diagrama de bloques que muestra la configuración principal del interior de la sección 1012 de decodificación de mejora de banda. La sección 1012 de decodificación de mejora de banda comprende la sección 1121 de decodificación de espectro de banda alta, la sección 1122 de decodificación de factor de escala correctiva, el multiplicador 1123 y la sección 1124 de enlace.

La sección 1121 de decodificación de espectro de banda alta decodifica un espectro estimado (espectro fino) de las bandas FL a FH usando un parámetro de codificación de información de estimación y el primer espectro incluido en la entrada de información codificada de mejora de banda de la sección 1011 de control. El espectro estimado obtenido se proporciona al multiplicador 1123.

La sección 1122 de decodificación de factor de escala correctiva decodifica un factor de escala correctiva usando un parámetro de codificación de factor de escala correctiva incluido en la entrada de información codificada de mejora de banda de la sección 1011 de control. Específicamente, la sección 1122 de decodificación de factor de escala correctiva hace referencia a un libro de códigos (no mostrado) de factor de escala correctiva interno y emite un factor de escala correctiva correspondiente al multiplicador 1123.

El multiplicador 1123 multiplica la salida del espectro estimado de la sección 1121 de decodificación de espectro de banda alta por la salida del factor de escala correctiva de la sección 1122 de decodificación de factor de escala correctiva y emite el resultado de multiplicación a la sección 1124 de enlace.

La sección 1124 de enlace enlaza el primer espectro y la salida del espectro estimado del multiplicador 1123 en el dominio de frecuencia, para generar un espectro decodificado de banda ancha de las bandas de señal 0 a FH, y emite este a la sección 1013 de transformación de dominio de tiempo como un coeficiente de MDCT de mejora de primera capa.

Por medio de la sección 1012 de decodificación de mejora de banda, cuando una señal de entrada se transforma a un coeficiente de dominio de frecuencia y un factor de escala se cuantifica en la codificación de dominio de frecuencia de la capa superior, se realiza cuantificación de factor de escala usando una escala de distorsión ponderada de manera que un candidato de cuantificación para el que el factor de escala se hace pequeño se hace propenso a seleccionarse. Es decir, un candidato de cuantificación en el cual el factor de escala después de cuantificación es más pequeño que el factor de escala antes de cuantificación, son más probables a seleccionarse. Por lo tanto, la degradación de calidad subjetiva perceptual puede suprimirse incluso cuando el número de bits asignados a la cuantificación de factor de escala es insuficiente.

La Figura 20 es un diagrama de bloques que muestra la configuración principal del interior de la sección 1015 de decodificación de segunda capa. La sección 1015 de decodificación de segunda capa tiene una configuración básica similar a la de la sección 405 de decodificación de segunda capa mostrada en la Figura 9, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

El procesamiento se diferencia en parte entre la sección 1151 de cálculo de coeficiente de MDCT de adición de la sección 1015 de decodificación de segunda capa y la sección 452 de cálculo de coeficiente de MDCT de adición de la sección 405 de decodificación de segunda capa, y se asignan un código de referencia diferente para indicar esto.

La sección 1151 de cálculo de coeficiente de MDCT de adición tiene un coeficiente de MDCT de mejora de primera capa como entrada de la sección 1012 de decodificación de mejora de banda, y un coeficiente de MDCT decodificado de segunda capa como entrada de la sección 204 de descuantificación de ganancia. La sección 1151 de cálculo de coeficiente de MDCT de adición añade juntos el coeficiente de MDCT decodificado de primera capa y el coeficiente de MDCT decodificado de segunda capa, y emite un coeficiente de MDCT de adición. Para una banda de banda mejorada, el valor del coeficiente de MDCT de mejora de primera capa se añade como cero en la sección 1151 de cálculo de coeficiente de MDCT de adición. Es decir, para una banda de mejora de banda, el valor de coeficiente de MDCT decodificado de segunda capa se toma como el valor de coeficiente de MDCT de adición.

Por lo tanto, de acuerdo con esta realización, cuando un componente de frecuencia de una banda diferente se hace un objetivo de cuantificación en cada trama, se realiza codificación predictiva de parámetro no temporal adaptativamente además de aplicar codificación escalable usando tecnología de mejora de banda. En consecuencia, la cantidad de información codificada en la codificación de habla puede reducirse, y el error de codificación de la señal de habla/audio y la degradación de la calidad de audio de la señal decodificada puede reducirse adicionalmente.

También, puesto que un residuo no se calcula para un componente de una banda mejorada mediante un método de codificación de mejora de banda, la energía de un componente objetivo de cuantificación no aumenta en una capa superior, y la eficacia de la cuantificación puede mejorarse.

En esta realización, se ha descrito un caso a modo de ejemplo en el que se aplica un método en el cual la información codificada de mejora de banda se calcula en un aparato de codificación usando la correlación entre un componente de banda baja decodificado mediante una sección de decodificación de primera capa y un componente de banda alta de una señal de entrada, pero la presente invención no está limitada a esto, y puede aplicarse también de manera similar a una configuración que emplee un método en el cual la información codificada de mejora de banda no se calcule, y se realice pseudo-generación de una banda alta por medio de un componente de ruido, como con AMR-WB (MultiTasa Adaptativa - Banda Ancha). Como alternativa, un método de selección de banda de la

presente invención puede aplicarse de manera similar al método de codificación de mejora de banda descrito en este ejemplo, o un método de codificación/decodificación escalable que no emplee un método de generación de componente de banda alta también usado en AMR-WB.

5 (Realización 6)

La Figura 21 es un diagrama de bloques que muestra la configuración principal del aparato 1100 de codificación de habla de acuerdo con la realización 6 de la presente invención.

10 En esta figura, el aparato 1100 de codificación de habla está equipado con la sección 301 de submuestreo, la sección 302 de codificación de primera capa, la sección 303 de decodificación de primera capa, la sección 304 de sobremuestreo, la primera sección 305 de transformación de dominio de frecuencia, la sección 306 de retardo, la segunda sección 307 de transformación de dominio de frecuencia, la sección 1108 de codificación de segunda capa y la sección 309 de multiplexación, y tiene una configuración escalable que comprende dos capas. En la primera
15 capa, se aplica un método de codificación de habla de CELP, y en la segunda capa, se aplica el método de codificación de habla descrito en la realización 1 de la presente invención.

Con la excepción de la sección 1108 de codificación de segunda capa, los elementos de configuración en el aparato 1100 de codificación de habla mostrado en la Figura 21 son idénticos a los elementos de configuración del aparato
20 300 de codificación de habla mostrado en la Figura 6, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

La Figura 22 es un diagrama de bloques que muestra la configuración principal del interior de la sección 1108 de codificación de segunda capa. La sección 1108 de codificación de segunda capa comprende principalmente la
25 sección 381 de cálculo de coeficiente de MDCT residual, la sección 1802 de selección de banda, la sección 103 de cuantificación de forma, la sección 104 de decisión de ejecución/no ejecución de codificación predictiva, la sección 1805 de cuantificación de ganancia y la sección 106 de multiplexación. Con la excepción de la sección 1802 de selección de banda y la sección 1805 de cuantificación de ganancia, los elementos de configuración en la sección 1108 de codificación de segunda capa son idénticos a los elementos de configuración de la sección 308 de
30 codificación de segunda capa mostrados en la Figura 7, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

La sección 1802 de selección de banda en primer lugar divide el coeficiente de MDCT X_k en una pluralidad de subbandas. En este punto, se proporcionará una descripción tomando un caso en el que el coeficiente de MDCT X_k
35 se divide igualmente en J subbandas (donde J es un número natural) como un ejemplo. A continuación la sección 1802 de selección de banda selecciona L subbandas (donde L es un número natural) de entre las J subbandas, y obtiene M tipos de regiones (donde M es un número natural).

La Figura 23 es un dibujo que muestra un ejemplo de la configuración de las regiones obtenidas mediante la sección
40 1802 de selección de banda.

En esta figura, el número de subbandas es 17 ($J=17$), el número de tipos de regiones es ocho ($M=8$), y cada región está compuesta de dos grupos de subbandas (siendo el número de bandas que componen estos dos grupos de subbanda tres y dos respectivamente). De estos dos grupos de subbanda, el grupo de subbanda que comprende
45 dos bandas localizadas en el lado de banda alta es fijo a lo largo de todas las tramas, siendo los índices de subbanda, por ejemplo, 15 y 16. Por ejemplo, la región 4 está compuesta de las subbandas 6 a 8, 15 y 16.

A continuación, la sección 1802 de selección de banda calcula la energía media $E(m)$ de cada uno de los M tipos de regiones de acuerdo con la Ecuación (33) a continuación.

50

$$E(m) = \frac{\sum_{j' \in \text{Región}(m)} \sum_{k=B(j')}^{B(j')+W(j')} (X_k)^2}{L} \quad (m = 0, \dots, M-1) \quad \dots \text{ (Ecuación 33)}$$

En esta ecuación, j' indica el índice de cada una de las J subbandas, y m indica el índice de cada uno de los M tipos de regiones. Región (m) significa una colección de índices de L subbandas que componen la región m, y $B(j')$
55 indica el mínimo valor entre los índices de una pluralidad de coeficientes de MDCT que componen la subbanda j' . $W(j')$ indica el ancho de banda de la subbanda j' , y en la siguiente descripción, se describirá como ejemplo un caso en el que los anchos de banda de las J subbandas son todos iguales - es decir, un caso en el que $W(j')$ es una constante.

60 A continuación, cuando una región para la que la energía media $E(m)$ es un máximo - por ejemplo, se selecciona la región $m_{\text{máx}}$, la sección 1802 de selección de banda selecciona una banda compuesta de las $j' \in \text{Región}(m_{\text{máx}})$

subbandas como una banda objetivo de cuantificación, y emite el índice $m_máx$ que indica a esta región como información de banda a la sección 103 de cuantificación de forma, a la sección 104 de decisión de ejecución/no ejecución de codificación predictiva y a la sección 106 de multiplexación. La sección 1802 de selección de banda emite también el coeficiente de MDCT residual X_k a la sección 103 de cuantificación de forma.

La sección 1805 de cuantificación de ganancia tiene una memoria intermedia interna que almacena un valor de ganancia de cuantificación obtenido en una trama pasada. Si una entrada del resultado de la determinación de la sección 104 de decisión de ejecución/no ejecución de codificación predictiva indica que la codificación predictiva se ha de realizar, la sección 1805 de cuantificación de ganancia realiza cuantificación prediciendo un valor de ganancia de trama actual usando el valor de ganancia de cuantificación de trama pasada C'_i , almacenado en la memoria intermedia interna. Específicamente, la sección 1805 de cuantificación de ganancia busca un libro de códigos de ganancia interno compuesto de la cantidad de GQ de vectores de código de ganancia para cada una de las L subbandas y encuentra un índice de un vector de código de ganancia para el que el resultado de la Ecuación (34) a continuación es un mínimo.

$$Ganancia_q(i) = \sum_{j' \in Regi\acute{o}n(m_m\acute{a}x)} \left\{ Ganancia_i(j') - \sum_{k=1}^3 (\alpha_k \cdot C'_i) - \alpha_0 \cdot GC_k \right\}^2 \begin{matrix} (i = 0, \dots, GQ-1) \\ (k = 0, \dots, L-1) \end{matrix}$$

... (Ecuación 34)

En esta ecuación, GC_k indica un vector de código de ganancia que compone un libro de códigos de ganancia, i indica un índice de vector de código de ganancia, y k indica un índice de un elemento de vector de código de ganancia. Por ejemplo, si el número de subbandas que compone una región es cinco ($L=5$), k tiene un valor de 0 a 4. En este punto, las ganancias de las subbandas de una región seleccionada se enlazan de modo que los índices de subbanda están en orden ascendente, las ganancias consecutivas se tratan como un vector de código de ganancia L-dimensional, y se realiza cuantificación de vector. Por lo tanto, para proporcionar una descripción usando la Figura 23, en el caso de la región 4, los valores de ganancia de los índices de subbanda 6, 7, 8, 15 y 16 se enlazan y tratan como un vector de código de ganancia 5-dimensional. También, C'_i indica un valor de ganancia de t tramas anteriores en tiempo, de modo que cuando $t=1$, por ejemplo, C'_i indica un valor de ganancia de una trama anterior en tiempo, y α es un coeficiente de predicción lineal de 4º orden almacenado en la sección 1805 de cuantificación de ganancia.

La sección 1805 de cuantificación de ganancia emite el índice de vector de código de ganancia $G_mín$ para el que el resultado de la Ecuación (34) anterior es un mínimo a la sección 106 de multiplexación como información codificada de ganancia. Si no existe valor de ganancia de una subbanda que corresponde a una trama pasada en la memoria intermedia interna, la sección 1805 de cuantificación de ganancia sustituye el valor de ganancia de la subbanda más próxima en frecuencia en la memoria intermedia interna en la Ecuación (34) anterior.

Por otro lado, si la entrada del resultado de la determinación de la sección 104 de decisión de ejecución/no ejecución de codificación predictiva indica que la codificación predictiva no se ha de realizar, la sección 1805 de cuantificación de ganancia cuantifica directamente la entrada del valor de ganancia ideal $Ganancia_i(j')$ de la sección 103 de cuantificación de forma de acuerdo con la Ecuación (35) a continuación. En este punto, la sección 1805 de cuantificación de ganancia trata un valor de ganancia ideal como un vector L-dimensional, y realiza cuantificación de vector.

$$Ganancia_q(i) = \sum_{j' \in Regi\acute{o}n(m_m\acute{a}x)} \left\{ Ganancia_i(j') - GC_k \right\}^2 \begin{matrix} (i = 0, \dots, GQ-1) \\ (k = 0, \dots, L-1) \end{matrix}$$

... (Ecuación 35)

En este punto, un índice de libro de códigos que hace a la Ecuación (35) anterior un mínimo se indica mediante $G_mín$.

La sección 1805 de cuantificación de ganancia emite $G_mín$ a la sección 106 de multiplexación como información codificada de ganancia. La sección 1805 de cuantificación de ganancia actualiza también la memoria intermedia interna de acuerdo con la Ecuación (36) a continuación usando la información codificada de ganancia $G_mín$ y el valor de ganancia de cuantificación c'_i , obtenido en la trama actual. Es decir, en la Ecuación (36), un valor C'_i , se actualiza con el índice del elemento de vector de código de ganancia $GC^{G_mín}_{j'}$ y j' satisfaciendo $j' \in Regi\acute{o}n(m_m\acute{a}x)$ asociados respectivamente en orden ascendente.

$$\begin{cases} C_{j'}^3 = C_{j'}^2, \\ C_{j'}^2 = C_{j'}^1, \\ C_{j'}^1 = GC_{j'}^{G_{\min}} \end{cases} \quad \left(\begin{array}{l} j' \in \text{Región}(m_{\max}) \\ j = 0, \dots, L-1 \end{array} \right) \quad \dots \text{ (Ecuación 36)}$$

La Figura 24 es un diagrama de bloques que muestra la configuración principal del aparato 1200 de decodificación de habla de acuerdo con esta realización.

En esta figura, el aparato 1200 de decodificación de habla está equipado con la sección 401 de control, la sección 402 de decodificación de primera capa, la sección 403 de sobremuestreo, la sección 404 de transformación de dominio de frecuencia, la sección 1205 de decodificación de segunda capa, la sección 406 de transformación de dominio de tiempo y el conmutador 407.

Con la excepción de la sección 1205 de decodificación de segunda capa, los elementos de configuración en el aparato 1200 de decodificación de habla mostrado en la Figura 24 son idénticos a los elementos de configuración del aparato 400 de decodificación de habla mostrado en la Figura 8, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

La Figura 25 es un diagrama de bloques que muestra la configuración principal del interior de la sección 1205 de decodificación de segunda capa. La sección 1205 de decodificación de segunda capa comprende principalmente la sección 451 de demultiplexación, la sección 202 de descuantificación de forma, la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva, la sección 2504 de descuantificación de ganancia y la sección 452 de cálculo de coeficiente de MDCT de adición. Con la excepción de la sección 2504 de descuantificación de ganancia, los elementos de configuración en la sección 1205 de decodificación de segunda capa son idénticos a los elementos de configuración de la sección 405 de decodificación de segunda capa mostrados en la Figura 9, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

La sección 2504 de descuantificación de ganancia tiene una memoria intermedia interna que almacena un valor de ganancia obtenido en una trama pasada. Si una entrada del resultado de la determinación de la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva indica que la codificación predictiva se ha de realizar, la sección 2504 de descuantificación de ganancia realiza descuantificación prediciendo un valor de ganancia de trama actual usando un valor de ganancia de trama pasada almacenado en la memoria intermedia interna. Específicamente, la sección 2504 de descuantificación de ganancia tiene el mismo tipo de libro de códigos de ganancia interno ($GC_k^{G_{\min}}$, donde k indica un índice de elemento) como la sección 105 de cuantificación de ganancia del aparato 100 de codificación de habla, y obtiene el valor de ganancia $Ganancia_q'$ realizando descuantificación de ganancia de acuerdo con la Ecuación (37) a continuación. En este punto, $C_{j'}^t$, indica un valor de ganancia de t tramas anteriores en tiempo, de modo que cuando $t=1$, por ejemplo, $C_{j'}^1$, indica un valor de ganancia de una trama anterior en tiempo. También, α es un coeficiente de predicción lineal de 4º orden almacenado en la sección 2504 de descuantificación de ganancia. La sección 2504 de descuantificación de ganancia trata L subbandas en una región como un vector L -dimensional, y realiza descuantificación de vector. Es decir, en la Ecuación (37), un valor de $Ganancia_q'$ (j') se calcula con el índice de elemento del vector de código de ganancia $GC_k^{G_{\min}}$ y j' satisfaciendo $k' \in \text{Región}(m_{\max})$ asociados respectivamente en orden ascendente.

$$Ganancia_q'(j') = \sum_{i=1}^3 (\alpha_i \cdot C_{j'}^{m_i}) + \alpha_0 \cdot GC_k^{G_{\min}} \quad \left(\begin{array}{l} j' \in \text{Región}(m_{\max}) \\ k = 0, \dots, L-1 \end{array} \right) \quad \dots \text{ (Ecuación 37)}$$

Si no existe valor de ganancia de una subbanda que corresponde a una trama pasada en la memoria intermedia interna, la sección 2504 de descuantificación de ganancia sustituye el valor de ganancia de la subbanda más próxima en frecuencia en la memoria intermedia interna en la Ecuación (37) anterior.

Por otro lado, si la entrada del resultado de la determinación de la sección 203 de decisión de ejecución/no ejecución de decodificación predictiva indica que la decodificación predictiva no se ha de realizar, la sección 2504 de descuantificación de ganancia realiza descuantificación de un valor de ganancia de acuerdo con la Ecuación (38) a continuación usando el libro de códigos de ganancia anteriormente descrito. En este punto, un valor de ganancia se trata como un vector L -dimensional, y se realiza descuantificación de vector. Es decir, cuando la decodificación predictiva no se realiza, la sección 2504 de descuantificación de ganancia toma el vector de código de ganancia $GC_k^{G_{\min}}$ que corresponde a la información codifica de ganancia $G_{\min}$ directamente como un valor de ganancia. En la Ecuación (38), k y j' están asociados respectivamente en orden ascendente de la misma manera que en la

Ecuación (37).

$$Ganancia_q'(j') = GC_k^{G-\min} \left(\begin{array}{l} j' \in \text{Región}(m_máx) \\ k = 0, \dots, L-1 \end{array} \right) \quad \dots \text{ (Ecuación 38)}$$

- 5 A continuación, la sección 2504 de descuantificación de ganancia calcula un coeficiente de MDCT decodificado de acuerdo con la Ecuación (39) a continuación usando un valor de ganancia obtenido mediante descuantificación de la trama actual y una entrada de valor de forma de la sección 202 de descuantificación de forma, y actualiza la memoria intermedia interna de acuerdo con la Ecuación (40) a continuación. En la Ecuación (40), un valor $C''_{j'}$, se actualiza con j de valor de ganancia descuantificada $Ganancia_q'(j)$ y j' satisfaciendo $j' \in \text{Región}(m_máx)$ asociado
10 respectivamente en orden ascendente. En este punto, un coeficiente de MDCT decodificado calculado se indica mediante X''_k . También, en la descuantificación del coeficiente de MDCT, si k está presente en $B(j')$ a $B(j'+1)-1$, el valor de ganancia toma el valor de $Ganancia_q'(j')$.

$$X''_k = Ganancia_q'(j') \cdot Forma_q'(k) \left(\begin{array}{l} k = B(j'), \dots, B(j'+1)-1 \\ j' \in \text{Región}(m_máx) \end{array} \right) \quad \dots \text{ (Ecuación 39)}$$

15

$$\left\{ \begin{array}{l} C''_{j'} = C''_{j'} \\ C''_{j'} = C''_{j'} \\ C''_{j'} = Ganancia_q'(j) \end{array} \right. \left(\begin{array}{l} j' \in \text{Región}(m_máx) \\ j = 0, \dots, L-1 \end{array} \right) \quad \dots \text{ (Ecuación 40)}$$

La sección 2504 de descuantificación de ganancia emite el coeficiente de MDCT decodificado X''_k calculado de acuerdo con la Ecuación (39) anterior a la sección 452 de cálculo de coeficiente de MDCT de adición.

20

Por lo tanto, de acuerdo con esta realización, como se compara con seleccionar una región compuesta de subbandas adyacentes de entre todas las bandas como una banda objetivo de cuantificación, una pluralidad de bandas para las que se desea mejorar la calidad de audio se establecen de antemano a través de un amplio intervalo, y una pluralidad de bandas no consecutivas que abarcan un amplio intervalo se seleccionan como bandas
25 objetivo de cuantificación. En consecuencia, tanto la calidad de banda baja como de banda alta puede mejorarse al mismo tiempo.

En esta realización, la razón para siempre fijar subbandas incluidas en una banda objetivo de cuantificación en el lado de la banda alta, como se muestra en la Figura 23, es que la distorsión de codificación es aún grande para una
30 banda alta en la primera capa de un códec escalable. Por lo tanto, la calidad de audio se mejora seleccionando de manera fija también una banda alta que no se haya codificado con muy alta precisión mediante la primera capa como un objetivo de cuantificación además de seleccionar una banda baja o media que tenga significancia perceptual para la selección como un objetivo de cuantificación en la segunda capa.

- 35 En esta realización, se ha descrito un caso a modo de ejemplo en el que una banda que se hace un objetivo de cuantificación de banda alta se fija incluyendo las mismas subbandas de banda alta (específicamente, los índices de subbanda 15 y 16) a lo largo de todas las tramas, pero la presente invención no está limitada a esto, y una banda que se hace un objetivo de cuantificación de banda alta puede seleccionarse también de entre una pluralidad de candidatas de banda objetivo de cuantificación para una subbanda de banda alta de la misma manera que para una
40 subbanda de banda baja. En un caso de este tipo, la selección puede realizarse después de multiplicar por un factor de ponderación más grande cuanto más alta es el área de subbanda. Es posible también para las bandas que se hacen candidatas que se cambien adaptativamente de acuerdo con la tasa de muestreo de la señal de entrada, tasa de bit de codificación, y las características espectrales de señal decodificada de primera capa, o las características espectrales de una señal diferencial para una señal de entrada y señal decodificada de primera capa o similar. Por
45 ejemplo, un posible método es dar prioridad como una candidata de banda objetivo de cuantificación a una parte donde la distribución de energía del espectro (coeficiente de MDCT residual) de una señal diferencial para la señal de entrada y la señal decodificada de primera capa es alta.

- 50 En esta realización, se ha descrito un caso a modo de ejemplo en el que un grupo de subbanda del lado de banda alta que compone una región es fijo, y si la codificación predictiva se ha de aplicar o no a una sección de cuantificación de ganancia se determina de acuerdo con el número de subbandas comunes a una banda objetivo de

cuantificación seleccionada en la trama actual y una banda objetivo de cuantificación seleccionada en una trama pasada, pero la presente invención no está limitada a esto, y la codificación predictiva puede también aplicarse siempre a la ganancia de un grupo de subbanda del lado de banda alta que compone una región, con la determinación de si la codificación predictiva se ha de realizar o no estando realizada únicamente para un grupo de subbanda del lado de banda baja. En este caso, el número de subbandas comunes a una banda objetivo de cuantificación seleccionada en la trama actual y una banda objetivo de cuantificación seleccionada en una trama pasada se tiene en consideración únicamente para un grupo de subbanda del lado de banda baja. Es decir, en este caso, un vector de cuantificación se cuantifica después de división en una parte para la que la codificación predictiva se realiza y una parte para la que la codificación predictiva no se realiza. De esta manera, puesto que no se realiza la determinación de si la codificación predictiva es necesaria o no para un grupo de subbanda fijo del lado de banda alta que compone una región, y la codificación predictiva siempre se realiza, la ganancia puede cuantificarse más eficazmente.

En esta realización, se ha descrito un caso a modo de ejemplo en el que la conmutación se realiza entre la aplicación y no aplicación de codificación predictiva en una sección de cuantificación de ganancia de acuerdo con el número de subbandas comunes a una banda objetivo de cuantificación seleccionada en la trama actual y una banda objetivo de cuantificación seleccionada una trama antes en tiempo, pero la presente invención no está limitada a esto, y un número de subbandas comunes a una banda objetivo de cuantificación seleccionadas en la trama actual y una banda objetivo de cuantificación seleccionada dos o más tramas antes en tiempo pueden usarse también. En este caso, incluso si el número de subbandas comunes a una banda objetivo de cuantificación seleccionada en la trama actual y una banda objetivo de cuantificación seleccionada una trama antes en tiempo es menor que o igual a un valor predeterminado, la codificación predictiva puede aplicarse en una sección de cuantificación de ganancia de acuerdo con el número de subbandas comunes a una banda objetivo de cuantificación seleccionada en la trama actual y una banda objetivo de cuantificación seleccionada dos o más tramas antes en tiempo.

En esta realización, se ha descrito un caso a modo de ejemplo en el que una región está compuesta de un grupo de subbanda del lado de banda baja y un grupo de subbanda del lado de banda alta, pero la presente invención no está limitada a esto, y, por ejemplo, un grupo de subbanda puede establecerse también en una banda media, y una región puede componerse de tres o más grupos de subbanda, el número de grupos de subbanda que compone una región cambiarse también adaptativamente de acuerdo con la tasa de muestreo de la señal de entrada, tasa de bits de codificación y características espectrales de señal decodificada de primera capa o las características espectrales de una señal diferencial para una señal de entrada y señal decodificada de primera capa o similar.

En esta realización, se ha descrito un caso a modo de ejemplo en el que un grupo de subbanda del lado de banda alta que compone una región está fijo a lo largo de todas las tramas, pero la presente invención no está limitada a esto, y un grupo de subbanda del lado de banda baja que compone una región puede fijarse también a lo largo de todas las tramas. También, tanto los grupos de subbanda del lado de banda alta como del lado de banda baja que componen una región pueden fijarse también a lo largo de todas las tramas, o tanto los grupos de subbanda del lado de banda alta como del lado de banda baja pueden buscarse y seleccionarse en una base trama a trama. Además, los diversos métodos anteriormente descritos pueden aplicarse a tres o más grupos de subbanda entre los grupos de subbanda que componen una región.

En esta realización, se ha descrito un caso a modo de ejemplo en el que, de las subbandas que componen una región, el número de subbandas que componen un grupo de subbanda del lado de banda alta es más pequeño que el número de subbandas que componen un grupo de subbanda del lado de banda baja (siendo el número de subbandas del grupo de subbanda del lado de banda alta dos, y siendo el número de subbandas del grupo de subbanda del lado de banda baja tres), pero la presente invención no está limitada a esto, y el número de subbandas que componen un grupo de subbanda del lado de banda alta puede también ser igual a, o mayor que, el número de subbandas que componen un grupo de subbanda del lado de banda baja. El número de subbandas que componen cada grupo de subbanda puede cambiarse también adaptativamente de acuerdo con la tasa de muestreo de la señal de entrada, tasa de bits de codificación, características espectrales de la señal decodificada de primera capa, características espectrales de una señal diferencial para una señal de entrada y señal decodificada de primera capa o similar.

En esta realización, se ha descrito un caso a modo de ejemplo en el que codificar usando un método de codificación de CELP se realiza mediante la sección 302 de codificación de primera capa, pero la presente invención no está limitada a esto, y la codificación usando un método de codificación distinto de CELP (tal como, por ejemplo, codificación de transformación) puede realizarse también.

(Realización 7)

La Figura 26 es un diagrama de bloques que muestra la configuración principal del aparato 1300 de codificación de habla de acuerdo con la realización 7 de la presente invención.

En esta figura, el aparato 1300 de codificación de habla está equipado con la sección 301 de submuestreo, la sección 302 de codificación de primera capa, la sección 303 de decodificación de primera capa, la sección 304 de

sobremuestreo, la sección 305 de transformación de dominio de frecuencia, la sección 306 de retardo, la segunda sección 307 de transformación de dominio de frecuencia, la sección 1308 de codificación de segunda capa y la sección 309 de multiplexación, y tiene una configuración escalable que comprende dos capas. En la primera capa, se aplica un método de codificación de habla de CELP, y en la segunda capa, se aplica el método de codificación de habla descrito en la realización 1 de la presente invención.

Con la excepción de la sección 1308 de codificación de segunda capa, los elementos de configuración en el aparato 1300 de codificación de habla mostrado en la Figura 26 son idénticos a los elementos de configuración del aparato 300 de codificación de habla mostrado en la Figura 6, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

La Figura 27 es un diagrama de bloques que muestra la configuración principal del interior de la sección 1308 de codificación de segunda capa. La sección 1308 de codificación de segunda capa comprende principalmente la sección 381 de cálculo de coeficiente de MDCT residual, la sección 102 de selección de banda, la sección 103 de cuantificación de forma, la sección 3804 de decisión de ejecución/no ejecución de codificación predictiva, la sección 3805 de cuantificación de ganancia y la sección 106 de multiplexación. Con la excepción de la sección 3804 de decisión de ejecución/no ejecución de codificación predictiva y la sección 3805 de cuantificación de ganancia, los elementos de configuración en la sección 1308 de codificación de segunda capa son idénticos a los elementos de configuración de la sección 308 de codificación de segunda capa mostrados en la Figura 7, y por lo tanto se asignan los mismos códigos de referencia a los elementos de configuración idénticos y se omiten las descripciones de los mismos en este punto.

La sección 3804 de decisión de ejecución/no ejecución de codificación predictiva tiene una memoria intermedia interna que almacena la entrada de la información de banda $m_{\text{máx}}$ de la sección 102 de selección de banda en una trama pasada. En este punto, se describirá un caso a modo de ejemplo en el que la sección 3804 de decisión de ejecución/no ejecución de codificación predictiva tiene una memoria intermedia interna que almacena información de banda $m_{\text{máx}}$ para las tres tramas pasadas. La sección 3804 de decisión de ejecución/no ejecución de codificación predictiva en primer lugar detecta una subbanda común a una banda objetivo de cuantificación de trama pasada y banda objetivo de cuantificación de trama actual usando la entrada de información de banda $m_{\text{máx}}$ de la sección 102 de selección de banda en una trama pasada y la entrada de información de banda $m_{\text{máx}}$ de la sección 102 de selección de banda en la trama actual. De las L subbandas indicadas mediante la entrada de información de banda $m_{\text{máx}}$ de la sección 102 de selección de banda, la sección 3804 de decisión de ejecución/no ejecución de codificación predictiva determina que la codificación predictiva se ha de aplicar, y establece $\text{Bandera_Pred}(j)=\text{ACTIVADO}$, para una subbanda seleccionada como un objetivo de cuantificación una trama atrás en tiempo. Por otro lado, de L subbandas indicadas mediante la entrada de información de banda $m_{\text{máx}}$ de la sección 102 de selección de banda, la sección 3804 de decisión de ejecución/no ejecución de codificación predictiva determina que la codificación predictiva no se ha de aplicar, y establece $\text{Bandera_Pred}(j)=\text{DESACTIVADO}$, para una subbanda no seleccionada como un objetivo de cuantificación una trama atrás en tiempo. En este punto, Bandera_Pred es una bandera que indica un resultado de la determinación de aplicación/no aplicación de codificación predictiva para cada subbanda, significando un valor ACTIVADO que la codificación predictiva se ha de aplicar a un valor de ganancia de subbanda, y significando un valor DESACTIVADO que la codificación predictiva no se ha de aplicar a un valor de ganancia de subbanda. La sección 3804 de decisión de ejecución/no ejecución de codificación predictiva emite un resultado de determinación para cada subbanda a la sección 3805 de cuantificación de ganancia. A continuación la sección 3804 de decisión de ejecución/no ejecución de codificación predictiva actualiza la memoria intermedia interna que almacena la información de banda usando la entrada de información de banda $m_{\text{máx}}$ de la sección 102 de selección de banda en la trama actual.

La sección 3805 de cuantificación de ganancia tiene una memoria intermedia interna que almacena un valor de ganancia de cuantificación obtenido en una trama pasada. La sección 3805 de cuantificación de ganancia cambia entre la ejecución/no ejecución de aplicación de codificación predictiva en el valor de ganancia de trama actual cuantificando de acuerdo con una entrada de resultado de la determinación de la sección 3804 de decisión de ejecución/no ejecución de codificación predictiva. Por ejemplo, si la codificación predictiva se ha de realizar, la sección 3805 de cuantificación de ganancia busca un libro de códigos de ganancia interno compuesto de la cantidad de GQ de vectores de código de ganancia para cada una de las L subbandas, realiza un cálculo de distancia que corresponde a la entrada del resultado de la determinación de la sección 3804 de decisión de ejecución/no ejecución de codificación predictiva, y encuentra un índice de vector de código de ganancia para el que el resultado de la Ecuación (41) anterior es un mínimo. En la Ecuación (41), uno u otro cálculo de distancia se realiza de acuerdo con $\text{Bandera_Pred}(j)$ que satisfacen $j \in \text{Región}(m_{\text{máx}})$, y se encuentra un índice de vector de ganancia para el que el valor total del error es un mínimo.

$$Ganancia_q(i) = \begin{cases} \sum_{j \in Regi\acute{o}n(m_m\acute{a}x)} \left\{ Ganancia_i(j) - \sum_{t=1}^L (\alpha_t \cdot C_j^t) - \alpha_0 \cdot GC_i \right\}^2 & (si(Bandera_Pred(j) == ACTIVADO)) \\ \sum_{j \in Regi\acute{o}n(m_m\acute{a}x)} \left\{ Ganancia_i(j) - GC_i \right\}^2 & (si(Bandera_Pred(j) == DESACTIVADO)) \end{cases} \quad \begin{matrix} (i = 0, \dots, GQ-1) \\ (k = 0, \dots, L-1) \end{matrix}$$

... (Ecuaci3n 41)

En esta ecuaci3n, GC_k^i indica un vector de c3digo de ganancia que compone un libro de c3digos de ganancia, i indica un 3ndice de vector de c3digo de ganancia y k indica un 3ndice de un elemento de vector de c3digo de ganancia. Por ejemplo, si el n3mero de subbandas que componen una regi3n es cinco ($L=5$), k tiene un valor de 0 a 4. En este punto, C_j^t indica un valor de ganancia de t tramas anteriores en tiempo, de modo que cuando $t=1$, por ejemplo, C_j^1 indica un valor de ganancia de una trama anterior en tiempo. Tambi3n, α es un coeficiente de predicci3n lineal de 4º orden almacenado en la secci3n 3805 de cuantificaci3n de ganancia. La secci3n 3805 de cuantificaci3n de ganancia trata L subbandas en una regi3n como un vector L -dimensional, y realiza cuantificaci3n de vector.

La secci3n 3805 de cuantificaci3n de ganancia emite el 3ndice de vector de c3digo de ganancia $G_m\acute{i}n$ para el que el resultado de la Ecuaci3n (41) anterior es un m3nimo a la secci3n 106 de multiplexaci3n como informaci3n codificada de ganancia.

La secci3n 3805 de cuantificaci3n de ganancia emite $G_m\acute{i}n$ a la secci3n 106 de multiplexaci3n como informaci3n codificada de ganancia. La secci3n 3805 de cuantificaci3n de ganancia actualiza tambi3n la memoria intermedia interna de acuerdo con la Ecuaci3n (42) a continuaci3n usando informaci3n codificada de ganancia $G_m\acute{i}n$ y el valor de ganancia de cuantificaci3n C_j^t obtenido en la trama actual. En la Ecuaci3n (42), un valor C_j^t , se actualiza con el 3ndice de elemento j del vector de c3digo de ganancia $GC_{j'}^{G_m\acute{i}n}$, j y j' satisfaciendo $j' \in Regi3n(m_m\acute{a}x)$ asociados respectivamente en orden ascendente.

$$\begin{cases} C_{j'}^3 = C_{j'}^2 \\ C_{j'}^2 = C_{j'}^1 \\ C_{j'}^1 = GC_{j'}^{G_m\acute{i}n} \end{cases} \quad \begin{matrix} (j' \in Regi3n(m_m\acute{a}x)) \\ (j = 0, \dots, L-1) \end{matrix} \quad \dots \text{ (Ecuaci3n 42)}$$

La Figura 28 es un diagrama de bloques que muestra la configuraci3n principal del aparato 1400 de decodificaci3n de habla de acuerdo con esta realizaci3n.

En esta figura, el aparato 1400 de decodificaci3n de habla est3 equipado con la secci3n 401 de control, la secci3n 402 de decodificaci3n de primera capa, la secci3n 403 de sobremuestreo, la secci3n 404 de transformaci3n de dominio de frecuencia, la secci3n 1405 de decodificaci3n de segunda capa, la secci3n 406 de transformaci3n de dominio de tiempo y el conmutador 407.

Con la excepci3n de la secci3n 1405 de decodificaci3n de segunda capa, los elementos de configuraci3n en el aparato 1400 de decodificaci3n de habla mostrado en la Figura 28 son id3nticos a los elementos de configuraci3n del aparato 400 de decodificaci3n de habla mostrado en la Figura 8, y por lo tanto se asignan los mismos c3digos de referencia a los elementos de configuraci3n id3nticos y se omiten las descripciones de los mismos en este punto.

La Figura 29 es un diagrama de bloques que muestra la configuraci3n principal del interior de la secci3n 1405 de decodificaci3n de segunda capa. La secci3n 1405 de decodificaci3n de segunda capa comprende principalmente la secci3n 451 de demultiplexaci3n, la secci3n 202 de descuantificaci3n de forma, la secci3n 4503 de decisi3n de ejecuci3n/no ejecuci3n de decodificaci3n predictiva, la secci3n 4504 de descuantificaci3n de ganancia, y la secci3n 452 de c3lculo de coeficiente de MDCT de adici3n. Con la excepci3n de la secci3n 4503 de decisi3n de ejecuci3n/no ejecuci3n de decodificaci3n predictiva y la secci3n 4504 de descuantificaci3n de ganancia, los elementos de configuraci3n en la secci3n 1405 de decodificaci3n de segunda capa mostrados en la Figura 29 son id3nticos a los elementos de configuraci3n de la secci3n 405 de decodificaci3n de segunda capa mostrados en la Figura 9, y por lo tanto se asignan los mismos c3digos de referencia a los elementos de configuraci3n id3nticos y se omiten las descripciones de los mismos en este punto.

La secci3n 4503 de decisi3n de ejecuci3n/no ejecuci3n de decodificaci3n predictiva tiene una memoria intermedia interna que almacena la entrada de informaci3n de banda $m_m\acute{a}x$ de la secci3n 451 de demultiplexaci3n en una trama pasada. En este punto, se describir3 un caso a modo de ejemplo en el que la secci3n 4503 de decisi3n de ejecuci3n/no ejecuci3n de decodificaci3n predictiva tiene una memoria intermedia interna que almacena la informaci3n de banda $m_m\acute{a}x$ para las tres tramas pasadas. La secci3n 4503 de decisi3n de ejecuci3n/no ejecuci3n de decodificaci3n predictiva en primer lugar detecta una subbanda com3n a una banda objetivo de cuantificaci3n de

trama pasada y banda objetivo de cuantificación de trama actual usando la entrada de información de banda $m_máx$ de la sección 451 de demultiplexación en una trama pasada y la entrada de información de banda $m_máx$ de la sección 451 de demultiplexación en la trama actual. De L subbandas indicadas mediante la entrada de información de banda $m_máx$ de la sección 451 de demultiplexación, la sección 4503 de decisión de ejecución/no ejecución de decodificación predictiva determina que la decodificación predictiva se ha de aplicar, y establece $Bandera_Pred(j)=ACTIVADO$, para una subbanda seleccionada como un objetivo de cuantificación una trama atrás en tiempo. Por otro lado, de L subbandas indicadas mediante la entrada de información de banda $m_máx$ de la sección 451 de demultiplexación, la sección 4503 de decisión de ejecución/no ejecución de decodificación predictiva determina que la decodificación predictiva no se ha de aplicar, y establece $Bandera_Pred(j)=DESACTIVADO$, para una subbanda no seleccionada como un objetivo de cuantificación una trama atrás en tiempo. En este punto, $Bandera_Pred$ es una bandera que indica un resultado de determinación de aplicación/no aplicación de decodificación predictiva para cada subbanda, significando un valor ACTIVADO que la decodificación predictiva se ha de aplicar a un valor de ganancia de subbanda, y significando un valor DESACTIVADO que la decodificación predictiva no se ha de aplicar a un valor de subbanda. A continuación, la sección 4503 de decisión de ejecución/no ejecución de decodificación predictiva emite un resultado de determinación para cada subbanda a la sección 4504 de descuantificación de ganancia. A continuación la sección 4503 de decisión de ejecución/no ejecución de decodificación predictiva actualiza la memoria intermedia interna que almacena la información de banda usando la entrada de información de banda $m_máx$ de la sección 451 de demultiplexación en la trama actual.

La sección 4504 de descuantificación de ganancia tiene una memoria intermedia interna que almacena un valor de ganancia obtenido en una trama pasada, y cambia entre la ejecución/no ejecución de la aplicación de decodificación predictiva en la decodificación del valor de ganancia de trama actual de acuerdo con una entrada del resultado de la determinación de la sección 4503 de decisión de ejecución/no ejecución de decodificación predictiva. La sección 4504 de descuantificación de ganancia tiene el mismo tipo de libro de códigos de ganancia interno que la sección 105 de cuantificación de ganancia del aparato 100 de codificación de habla, y cuando realiza decodificación predictiva, por ejemplo, obtiene el valor de $Ganancia_q'$ realizando descuantificación de ganancia de acuerdo con la Ecuación (43) a continuación. En este punto, C''^t_j indica un valor de ganancia de t tramas anteriores en tiempo, de modo que cuando $t=1$, por ejemplo, C''^t_j indica un valor de ganancia de una trama anterior. También, α es un coeficiente de predicción lineal de 4º orden almacenado en la sección 4504 de descuantificación de ganancia. La sección 4504 de descuantificación de ganancia trata L subbandas en una región como un vector L-dimensional, y realiza descuantificación de vector. En la Ecuación (43), un valor $Ganancia_q'$ (j') se calcula con el índice del elemento k y j' del vector de código de ganancia $GC^{G_mín}_k$ satisfaciendo $j' \in Región(m_máx)$ asociados respectivamente en orden ascendente.

$$Ganancia_q'(j') = \begin{cases} (si(Bandera_Pred(j') == ACTIVADO)) \\ \sum_{i=1}^3 (\alpha_i \cdot C''^i_j) + \alpha_0 \cdot GC^{G_mín}_k & \left(\begin{matrix} j' \in Región(m_máx) \\ k = 0, \dots, L-1 \end{matrix} \right) \\ (si(Bandera_Pred(j') == DESACTIVADO)) \\ GC^{G_mín}_k \end{cases}$$

... (Ecuación 43)

A continuación, la sección 4504 de descuantificación de ganancia calcula un coeficiente de MDCT decodificado de acuerdo con la Ecuación (44) a continuación usando un valor de ganancia obtenido mediante la descuantificación de trama actual y una entrada de valor de forma de la sección 202 de descuantificación de forma, y actualiza la memoria intermedia interna de acuerdo con la Ecuación (45) a continuación. En la Ecuación (45), un valor C''^t_j , se actualiza con j de valor de ganancia descuantificada $Ganancia_q'$ (j) y satisfaciendo j y $j' \in Región(m_máx)$ asociados respectivamente en orden ascendente. En este punto, un coeficiente de MDCT decodificado calculado se indica mediante X_k . También, en la descuantificación del coeficiente de MDCT, si k está presente en B (j') a B($j'+1$)-1, el valor de ganancia toma el valor de $Ganancia_q'(j')$.

$$X''_k = Ganancia_q'(j') \cdot Forma_q'(k) \left(\begin{matrix} k = B(j'), \dots, B(j'+L)-1 \\ j' \in Región(m_máx) \end{matrix} \right)$$

... (Ecuación 44)

$$\begin{cases} C_{j'}^{m3} = C_{j'}^{m2} \\ C_{j'}^{m2} = C_{j'}^{m1} \\ C_{j'}^{m1} = \text{Ganancia_} q'(j) \end{cases} \begin{pmatrix} j' \in \text{Región}(m_máx) \\ j = 0, \dots, L-1 \end{pmatrix} \quad \dots \text{ (Ecuación 45)}$$

La sección 4504 de descuantificación de ganancia emite el coeficiente de MDCT decodificado X''_k calculado de acuerdo con la Ecuación (44) anterior a la sección 452 de cálculo de coeficiente de MDCT de adición.

- 5 Por lo tanto, de acuerdo con esta realización, en el momento de la cuantificación de ganancia de una banda objetivo de cuantificación seleccionada en cada trama, se detecta si cada subbanda incluida en una banda objetivo de cuantificación se cuantificó o no en una trama pasada. A continuación se realiza cuantificación de vector, estando aplicada la codificación predictiva a una subbanda cuantificada en una trama pasada, y no estando aplicada la codificación predictiva a una subbanda no cuantificada en una trama pasada. Por este medio, la codificación del
- 10 parámetro de dominio de frecuencia puede llevarse a cabo más eficazmente que con un método en el cual la conmutación de la aplicación/no aplicación de codificación predictiva se realiza para un vector completo.

- En esta realización, se ha descrito un método en el cual se realiza conmutación entre aplicación y no aplicación de codificación predictiva en una sección de cuantificación de ganancia de acuerdo con el número de subbandas
- 15 comunes a una banda objetivo de cuantificación seleccionada en la trama actual y una banda objetivo de cuantificación seleccionada una trama atrás en tiempo, pero la presente invención no está limitada a esto, y un número de subbandas comunes a una banda objetivo de cuantificación seleccionada en la trama actual y una banda de objetivo de cuantificación seleccionada dos o más tramas atrás en tiempo pueden usarse también. En este caso, incluso si el número de subbandas comunes a una banda objetivo de cuantificación seleccionada en la trama actual
- 20 y una banda objetivo de cuantificación seleccionada una trama atrás en tiempo es menor que o igual a un valor predeterminado, la codificación predictiva puede aplicarse en una sección de cuantificación de ganancia de acuerdo con el número de subbandas comunes a una banda objetivo de cuantificación seleccionada en la trama actual y una banda objetivo de cuantificación seleccionada dos o más tramas atrás en tiempo.

- 25 Es posible también para el método de cuantificación descrito en esta realización combinarse con el método de selección de banda objetivo de cuantificación descrito en la realización 6. Se describirá un caso en el que, por ejemplo, una región que es una banda objetivo de cuantificación está compuesta de un grupo de subbanda del lado de banda baja y un grupo de subbanda del lado de banda alta, el grupo de subbanda del lado de banda alta está fijo a lo largo de todas las tramas, y se cuantifica un vector en el que la ganancia del grupo de subbanda del lado de
- 30 banda baja y el grupo de subbanda del lado de banda alta se hacen consecutivos. En este caso, en un vector de ganancia de banda objetivo de cuantificación, la cuantificación de vector se realiza con codificación predictiva estando siempre aplicada para un elemento que indica la ganancia del grupo de subbanda del lado de banda alta, y no estando aplicada la codificación predictiva para un elemento que indica ganancia de grupo de subbanda del lado de banda baja. Por este medio, la cuantificación de vector de ganancia puede llevarse a cabo más eficazmente que
- 35 cuando se realiza conmutación de aplicación/no aplicación de codificación predictiva para un vector completo. En este momento, en el grupo de subbanda del lado de banda baja, es también eficaz un método en el cual se realiza cuantificación de vector estando aplicada codificación predictiva a una subbanda cuantificada en una trama pasada, y no estando aplicada codificación predictiva a una subbanda no cuantificada en una trama pasada. También, para un elemento que indica ganancia de grupo de subbanda del lado de banda baja, se realiza cuantificación
- 40 conmutando entre la aplicación y no aplicación de codificación predictiva usando subbandas que componen una banda objetivo de cuantificación seleccionada en una trama pasada en tiempo, como se describe en la realización 1. Por este medio, la cuantificación de vector de ganancia puede realizarse aún más eficazmente. Es posible también para la presente invención aplicarse a una configuración que combine las configuraciones anteriormente descritas.

- 45 Esto concluye una descripción de realizaciones de la presente invención.

- En las anteriores realizaciones, se han descrito casos a modo de ejemplo en los que el método de seleccionar una banda objetivo de cuantificación es seleccionar la región con la energía más alta en todas las bandas, pero la
- 50 presente invención no está limitada a esto, y una cierta banda puede seleccionarse también preliminarmente de antemano, después de lo cual una banda objetivo de cuantificación se selecciona finalmente en la banda seleccionada preliminarmente. En un caso de este tipo, una banda seleccionada preliminarmente puede decidirse de acuerdo con la tasa de muestro de señal de entrada, tasa de bits de codificación, o similar. Por ejemplo, un método es seleccionar una banda baja preliminarmente cuando la tasa de muestreo es baja.

- 55 En las anteriores realizaciones, se usa la MDCT como un método de codificación de transformación, y por lo tanto "coeficiente de MDCT" usado en las anteriores realizaciones significa esencialmente "espectro". Por lo tanto, la expresión "coeficiente de MDCT" puede sustituirse por "espectro".

- En las anteriores realizaciones, se han mostrado ejemplos en los que los aparatos 200, 200a, 400, 600, 800, 1010,
- 60 1200 y 1400 de decodificación de habla reciben como entrada y procesan datos codificados transmitidos de los

aparatos 100, 100a, 300, 500, 700, 1000, 1100 y 1300 de codificación de habla, respectivamente, pero la salida de datos codificados mediante un aparato de codificación de una configuración diferente capaz de generar datos codificados que tengan una configuración similar puede introducirse y procesarse también.

5 Un aparato de codificación, aparato de decodificación y método de los mismos de acuerdo con la presente invención no están limitados a las realizaciones anteriormente descritas, y diversas variaciones y modificaciones pueden ser posibles sin alejarse del alcance de la presente invención. Por ejemplo, es posible para las realizaciones implementarse estando combinadas apropiadamente.

10 Es posible para un aparato de codificación y aparato de decodificación de acuerdo con la presente invención instalarse en un aparato de terminal de comunicación y aparato de estación base en un sistema de comunicación móvil, posibilitando de esta manera que se proporcione un aparato de terminal de comunicación, aparato de estación base y sistema de comunicación móvil que tienen el mismo tipo de efectos operacionales que los anteriormente descritos.

15 Se ha descrito en este punto un caso a modo de ejemplo en el que la presente invención está configurada como hardware, pero es también posible para la presente invención implementarse mediante software. Por ejemplo, el mismo tipo de funciones que aquellas de un aparato de codificación y aparato de decodificación de acuerdo con la presente invención pueden realizarse escribiendo un algoritmo de un método de codificación y método de decodificación de acuerdo con la presente invención en un lenguaje de programación, almacenar este programa en memoria y tenerlo ejecutado mediante un medio de procesamiento de información.

20 Los bloques de función usados en las descripciones de las anteriores realizaciones se implementan típicamente como LSI, que son circuitos integrados. Estos pueden implementarse individualmente como chips únicos, o un único chip puede incorporar algunos o todos ellos.

En este punto, se ha usado el término LSI, pero los términos IC, sistema LSI, súper LSI, ultra LSI y así sucesivamente pueden usarse también de acuerdo con diferencias en el grado de integración.

30 El método de implementación de circuitería integrada no está limitado a LSI, y la implementación por medio de circuitería especializada o un procesador de fin general puede usarse también. Un FPGA (Campo de Matriz de Puertas Programables) para el que es posible programación después de fabricación LSI, o un procesador reconfigurable que permite reconfiguración de conexiones de celdas de circuitos y ajustes en un LSI, puede usarse también.

35 En el caso de la introducción de una tecnología de implementación de circuito integrado en la cual se sustituya LSI por una tecnología diferente como un avance en, o derivación de, tecnología de semiconductores, la integración de los bloques de función puede, por supuesto, realizarse usando esa tecnología. La aplicación de biotecnología o similar es también una posibilidad.

40 **Aplicabilidad industrial**

Un aparato de codificación y así sucesivamente de acuerdo con la presente invención es adecuado para uso en un aparato de terminal de comunicación, aparato de estación base o similar, en un sistema de comunicación móvil.

45

REIVINDICACIONES

1. Un aparato (100, 100a, 500, 700) de codificación de habla adaptado para realizar codificación en una base trama a trama, comprendiendo el aparato de codificación de habla:

5 una sección (101) de transformación adaptada para transformar una señal de audio/habla de entrada a un dominio de frecuencia para obtener un parámetro de dominio de frecuencia;
 una sección (102) de selección adaptada para dividir el parámetro de dominio de frecuencia en una pluralidad de subbandas, y adaptada adicionalmente para seleccionar una banda objetivo de cuantificación compuesta de un
 10 número natural de subbandas consecutivas de la pluralidad de subbandas, y adaptada adicionalmente para generar una información de banda que indica la banda objetivo de cuantificación;
 una sección (103) de cuantificación de forma adaptada para cuantificar una forma del parámetro de dominio de frecuencia en la banda objetivo de cuantificación;
 una sección (105, 505, 705) de cuantificación de ganancia adaptada para codificar una ganancia del parámetro
 15 de dominio de frecuencia en la banda objetivo de cuantificación para obtener una información codificada de ganancia; y
 una sección (104) de determinación adaptada:

20 para determinar que la codificación predictiva se ha de realizar cuando el número de subbandas comunes a la banda objetivo de cuantificación de trama actual y una banda objetivo de cuantificación seleccionada en el pasado es igual a o mayor que un valor predeterminado, y
 para determinar que la codificación predictiva no se ha de realizar cuando el número de subbandas comunes es menor que el valor predeterminado,

25 donde la sección de cuantificación de ganancia está adaptada:

para obtener la información codificada de ganancia realizando codificación predictiva en la ganancia del parámetro de dominio de frecuencia en la banda objetivo de cuantificación usando información codificada de ganancia pasada cuando la sección de determinación ha determinado que la codificación predictiva se ha de
 30 realizar, y
 para obtener la información codificada de ganancia cuantificando directamente la ganancia del parámetro de dominio de frecuencia en la banda objetivo de cuantificación cuando la sección de determinación ha determinado que la codificación predictiva no se ha de realizar.

35 2. El aparato de codificación de habla de acuerdo con la reivindicación 1, donde la sección de cuantificación de ganancia está adaptada adicionalmente para obtener la información codificada de ganancia realizando cuantificación de vector de la ganancia del parámetro de dominio de frecuencia.

40 3. El aparato de codificación de habla de acuerdo con la reivindicación 1, donde la sección de cuantificación de ganancia está adaptada adicionalmente para obtener la información codificada de ganancia realizando cuantificación predictiva de la ganancia usando una ganancia de un parámetro de dominio de frecuencia en una trama pasada.

45 4. El aparato de codificación de habla de acuerdo con la reivindicación 1, donde la sección de selección está adaptada adicionalmente para seleccionar una región para la que la energía es la más alta entre las regiones compuestas de una pluralidad de subbandas como la banda objetivo de cuantificación.

5. El aparato de codificación de habla de acuerdo con la reivindicación 1, donde la sección de selección está adaptada adicionalmente para seleccionar una banda para la que la energía es la más alta entre las bandas candidatas como la banda objetivo de cuantificación, cuando las bandas candidatas existen para las que un número
 50 de subbandas comunes a la banda objetivo de cuantificación seleccionada en el pasado es igual a o mayor que un valor predeterminado y la energía es igual a o mayor que un valor predeterminado, y donde la sección de selección está adaptada adicionalmente para seleccionar una banda para la que la energía es la más alta en todas las bandas del dominio de frecuencia como la banda objetivo de cuantificación, cuando las bandas candidatas no existen.

55 6. El aparato de codificación de habla de acuerdo con la reivindicación 1, donde la sección de selección está adaptada adicionalmente para seleccionar una banda más cercana a la banda objetivo de cuantificación seleccionada en el pasado entre las bandas para las que la energía es igual a o mayor que un valor predeterminado que la banda objetivo de cuantificación.

60 7. El aparato de codificación de habla de acuerdo con la reivindicación 1, donde la sección de selección está adaptada adicionalmente para seleccionar la banda objetivo de cuantificación después de multiplicación por un factor de ponderación que es más grande cuanto más está una subbanda hacia un lado de banda baja.

65 8. El aparato de codificación de habla de acuerdo con la reivindicación 1, donde la sección de selección está adaptada adicionalmente para seleccionar una subbanda fijada al lado de banda baja como la banda objetivo de cuantificación.

9. El aparato de codificación de habla de acuerdo con la reivindicación 1, donde la sección de selección está adaptada adicionalmente para seleccionar la banda objetivo de cuantificación después de multiplicación por un factor de ponderación que es más grande cuanto más alta es la frecuencia de selección en el pasado de una subbanda.

5 10. El aparato de codificación de habla de acuerdo con la reivindicación 1, que comprende adicionalmente una sección (504) de interpolación adaptada para realizar interpolación en una ganancia de un parámetro de dominio de frecuencia en una subbanda no cuantificada en el pasado entre las subbandas indicadas mediante la información de banda usando información codificada de ganancia pasada, para obtener un valor de interpolación, donde la sección de cuantificación de ganancia está adaptada adicionalmente para usar también el valor de interpolación cuando realiza la codificación predictiva.

15 11. El aparato de codificación de habla de acuerdo con la reivindicación 1, que comprende adicionalmente una sección (704) de decisión adaptada para decidir un coeficiente de predicción de manera que un factor de ponderación de un valor de ganancia de una trama pasada es más grande cuanto más grande es una subbanda común a una banda objetivo de cuantificación de una trama pasada y una banda objetivo de cuantificación de una trama actual, donde la sección de cuantificación de ganancia está adaptada adicionalmente para usar el coeficiente de predicción cuando realiza la codificación predictiva.

20 12. El aparato de codificación de habla de acuerdo con la reivindicación 1, donde la sección de selección está adaptada adicionalmente para seleccionar de manera fija una subbanda predeterminada como parte de la banda objetivo de cuantificación.

25 13. El aparato de codificación de habla de acuerdo con la reivindicación 1, donde la sección de selección está adaptada adicionalmente para seleccionar la banda objetivo de cuantificación después de la multiplicación por un factor de ponderación que es más grande cuanto más está una subbanda hacia un lado de banda alta en parte de la banda objetivo de cuantificación.

30 14. El aparato de codificación de habla de acuerdo con la reivindicación 1, donde la sección de cuantificación de ganancia está adaptada adicionalmente para realizar codificación predictiva en una ganancia de un parámetro de dominio de frecuencia en parte de la banda objetivo de cuantificación, y para realizar codificación no predictiva en una ganancia de un parámetro de dominio de frecuencia en una parte restante.

35 15. El aparato de codificación de habla de acuerdo con la reivindicación 1, donde la sección de cuantificación de ganancia está adaptada adicionalmente para realizar cuantificación de vector de la ganancia de una pluralidad no consecutiva de subbandas.

16. Un aparato (200, 200a, 600, 800) de decodificación de habla que comprende:

40 una sección (201) de recepción adaptada para recibir información de banda que indica una banda objetivo de cuantificación, que está compuesta de un número natural de subbandas consecutivas de una pluralidad de subbandas, donde la pluralidad de subbandas es un parámetro de dominio de frecuencia subdividido de una señal de audio/habla de entrada;

45 una sección (202) de descuantificación de forma adaptada para decodificar una información codificada de forma en la que una forma de un parámetro de dominio de frecuencia en la banda objetivo de cuantificación se cuantifica, para generar una forma decodificada;

50 una sección (204, 604, 804) de descuantificación de ganancia adaptada para decodificar una información codificada de ganancia, en la que una ganancia del parámetro de dominio de frecuencia en la banda objetivo de cuantificación se cuantifica, para generar una ganancia decodificada, y para decodificar un parámetro de frecuencia usando la forma decodificada y la ganancia decodificada para generar un parámetro de dominio de frecuencia decodificado;

55 una sección (205) de transformación de dominio de tiempo adaptada para transformar el parámetro de dominio de frecuencia decodificado a un dominio de tiempo para obtener una señal decodificada de dominio de tiempo; y una sección (203) de determinación adaptada:

60 para encontrar un número de subbandas comunes a una banda objetivo de cuantificación de trama actual y una banda objetivo de cuantificación de trama pasada usando la información de banda recibida para determinar que la codificación predictiva se ha de realizar cuando el número de subbandas comunes a la banda objetivo de cuantificación de trama actual y una banda objetivo de cuantificación de trama pasada es igual a o mayor que un valor predeterminado, y para determinar que la decodificación predictiva no se ha de realizar cuando el número de subbandas comunes es menor que el valor predeterminado,

donde la sección de descuantificación de ganancia está adaptada:

65 para realizar decodificación predictiva de la ganancia del parámetro de dominio de frecuencia en la banda objetivo de cuantificación de trama actual usando la ganancia obtenida en la decodificación de ganancia

pasada, cuando la sección de determinación ha determinado que la decodificación predictiva se ha de realizar; y

para realizar directamente descuantificación de la información codificada de ganancia, en la que la ganancia de un parámetro de dominio de frecuencia se cuantifica en la banda objetivo de cuantificación de trama actual cuando la sección de determinación ha determinado que la decodificación predictiva no se ha de realizar.

17. Un método de codificación de habla para realizar codificación en una base trama a trama, comprendiendo el método las etapas de:

transformar una señal de audio/habla de entrada a un dominio de frecuencia para obtener un parámetro de dominio de frecuencia;
dividir el parámetro de dominio de frecuencia en una pluralidad de subbandas;
seleccionar una banda objetivo de cuantificación compuesta de un número natural de subbandas consecutivas de la pluralidad de subbandas;
generar una información de banda que indica la banda objetivo de cuantificación;
cuantificar una forma del parámetro de dominio de frecuencia en la banda objetivo de cuantificación para obtener una información codificada de forma;
codificar una ganancia del parámetro de dominio de frecuencia en la banda objetivo de cuantificación para obtener una información codificada de ganancia;
determinar que la codificación predictiva se ha de realizar cuando el número de subbandas comunes a la banda objetivo de cuantificación de trama actual y una banda objetivo de cuantificación seleccionada en el pasado es igual a o mayor que un valor predeterminado;
determinar que la codificación predictiva no se ha de realizar cuando el número de subbandas comunes es menor que el valor predeterminado;
obtener la información codificada de ganancia realizando codificación predictiva en la ganancia del parámetro de dominio de frecuencia en la banda objetivo de cuantificación usando información codificada de ganancia pasada, cuando se determina que la codificación predictiva se ha de realizar; y
obtener la información codificada de ganancia cuantificando directamente la ganancia del parámetro de dominio de frecuencia en la banda objetivo de cuantificación, cuando se determina que la codificación predictiva no se ha de realizar.

18. Un método de decodificación de habla que comprende las etapas de:

recibir información de banda que indica una banda objetivo de cuantificación, que está compuesta de un número natural de subbandas consecutivas de una pluralidad de subbandas, donde la pluralidad de subbandas es un parámetro de dominio de frecuencia subdividido de una señal de audio/habla de entrada;
decodificar una información codificada de forma en la que una forma de un parámetro de dominio de frecuencia en la banda objetivo de cuantificación se cuantifica, para generar una forma decodificada;
decodificar una información codificada de ganancia, en la que una ganancia de un parámetro de dominio de frecuencia en la banda objetivo de cuantificación se cuantifica, para generar una ganancia decodificada;
decodificar el parámetro de dominio de frecuencia usando la forma decodificada y la ganancia decodificada para generar un parámetro de dominio de frecuencia decodificada;
transformar el parámetro de dominio de frecuencia decodificada a un dominio de tiempo para obtener una señal decodificada de dominio de tiempo;
encontrar un número de subbandas comunes a una banda objetivo de cuantificación de trama actual y una banda objetivo de cuantificación de trama pasada usando la información de banda recibida;
determinar que la decodificación predictiva se ha de realizar cuando el número de subbandas comunes a la banda objetivo de cuantificación de trama actual y una banda objetivo de cuantificación de trama pasada es igual a o mayor que un valor predeterminado;
determinar que la decodificación predictiva no se ha de realizar cuando el número de subbandas comunes es menor que el valor predeterminado;
realizar decodificación predictiva de la ganancia del parámetro de dominio de frecuencia en la banda objetivo de cuantificación de trama actual usando la ganancia obtenida en la decodificación de ganancia pasada, cuando se determina que la decodificación predictiva se ha de realizar; y
realizar directamente descuantificación de la información codificada de ganancia, en la que la ganancia de un parámetro de dominio de frecuencia se cuantifica en la banda objetivo de cuantificación de trama actual, cuando se determina que la decodificación predictiva no se ha de realizar.

100

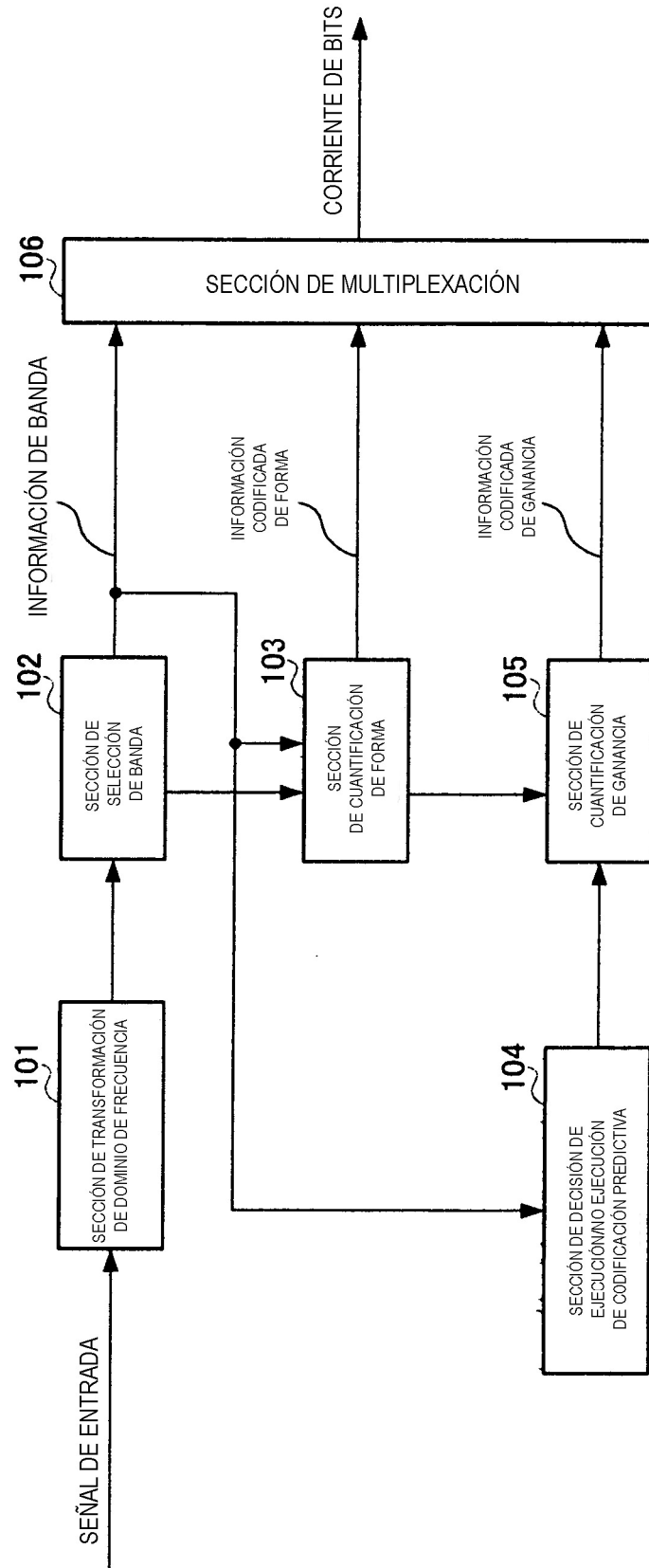


FIG.1

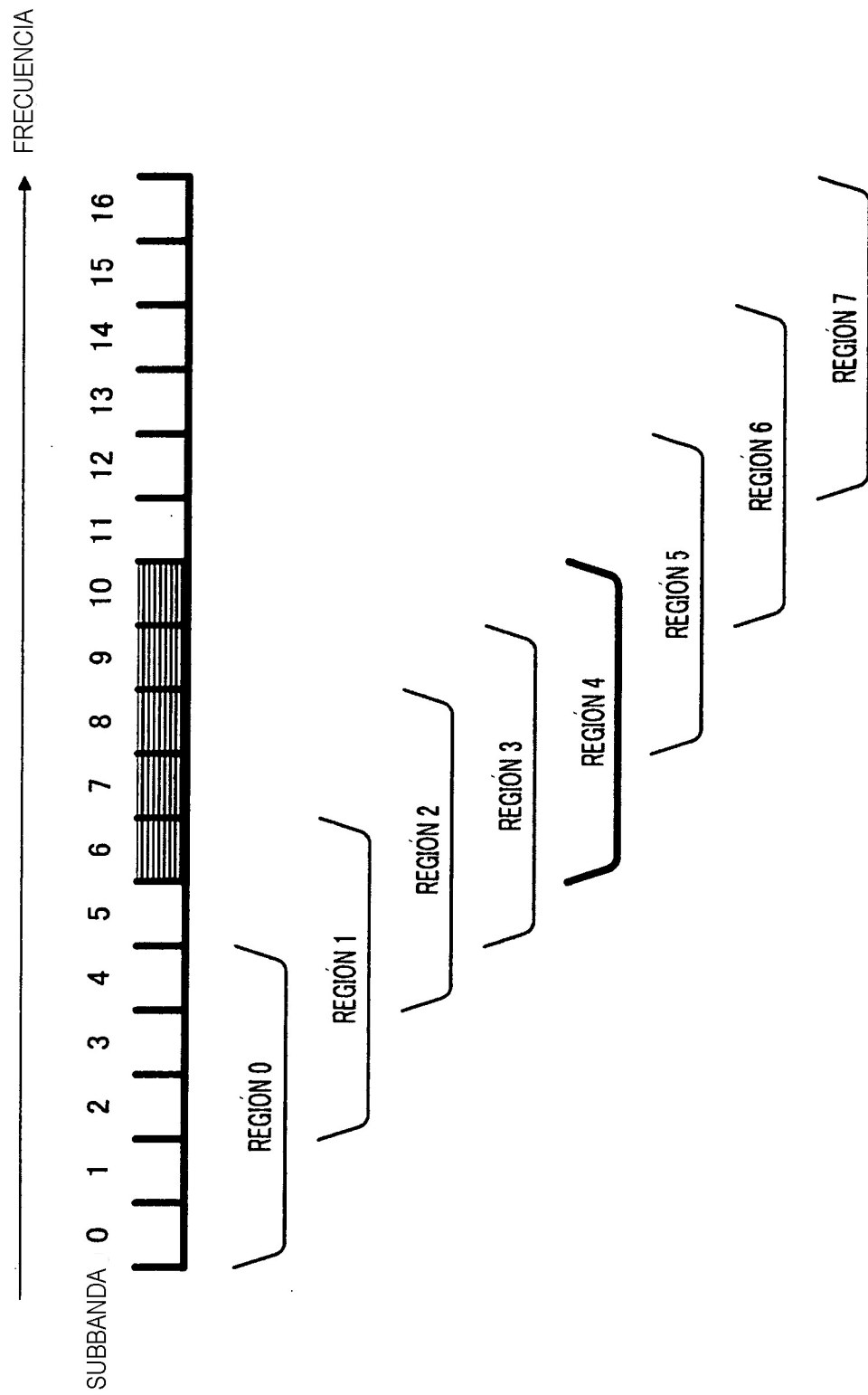


FIG.2

200

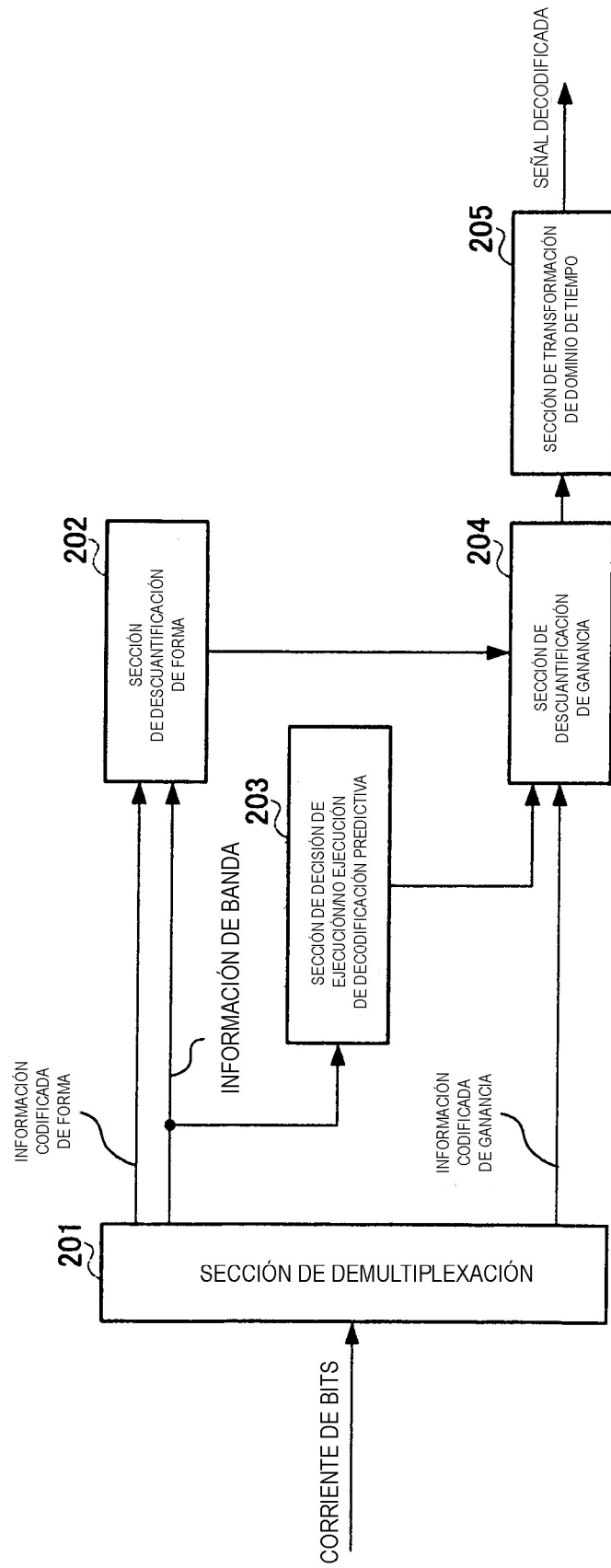


FIG.3

100a

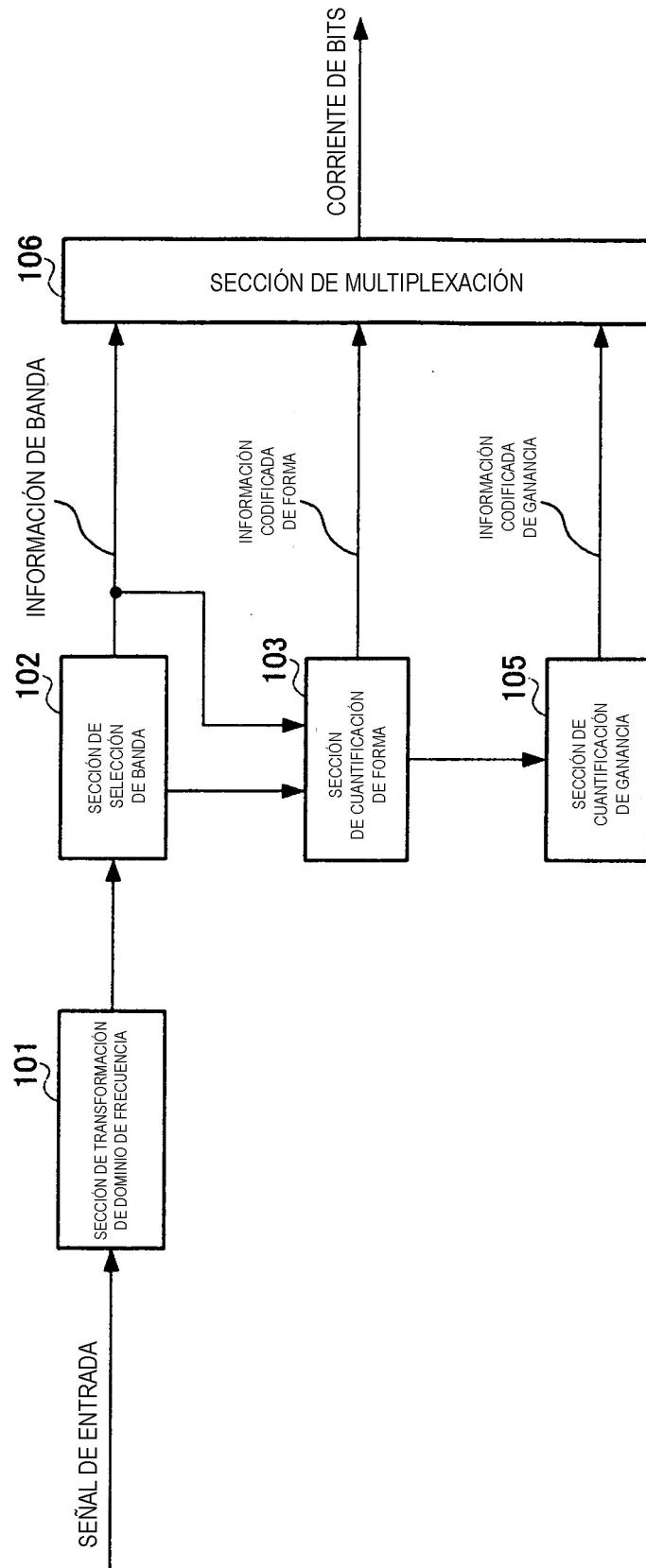


FIG.4

200a

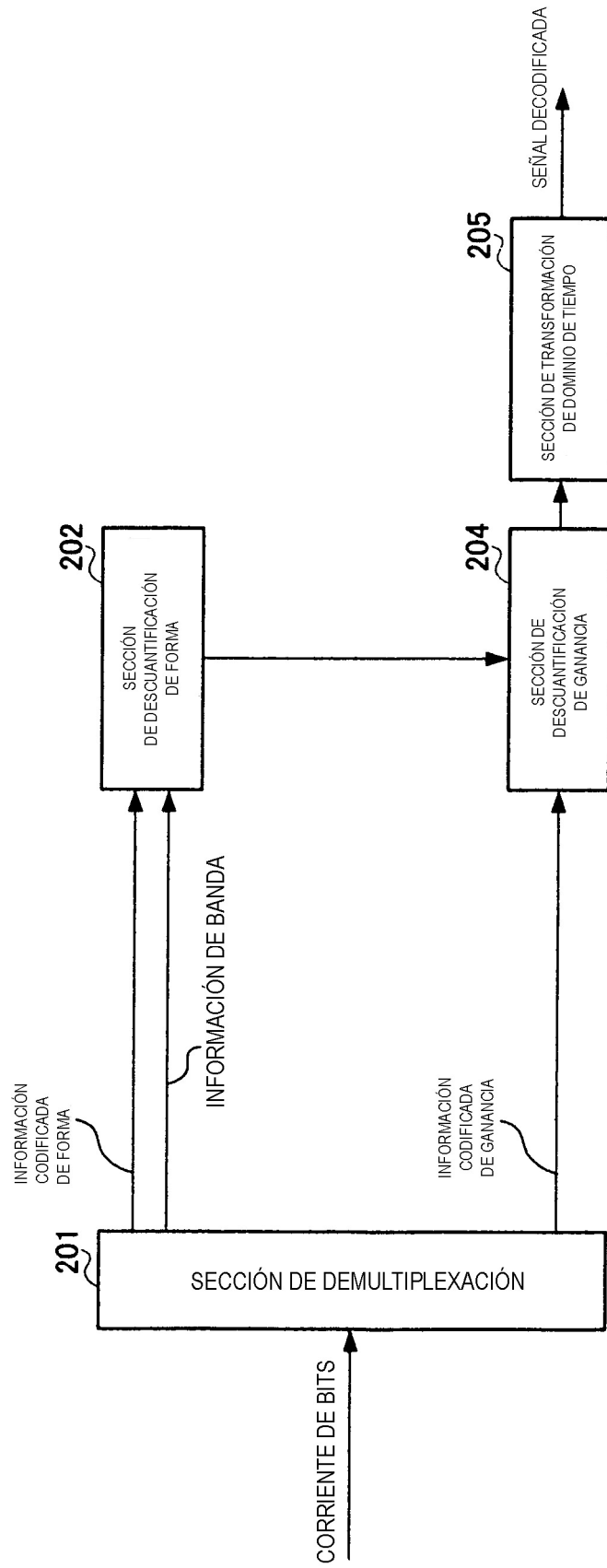


FIG.5

300

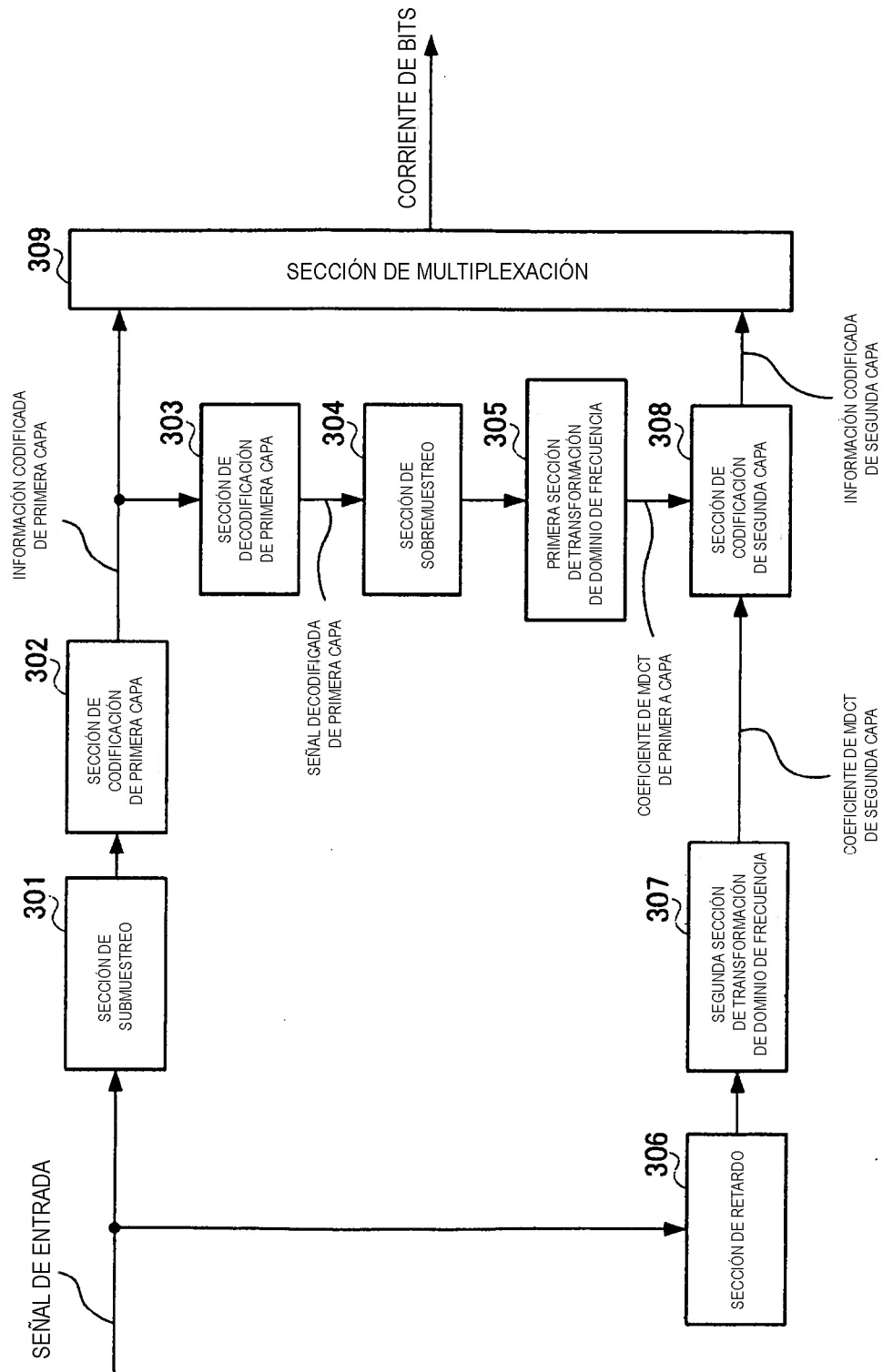


FIG.6

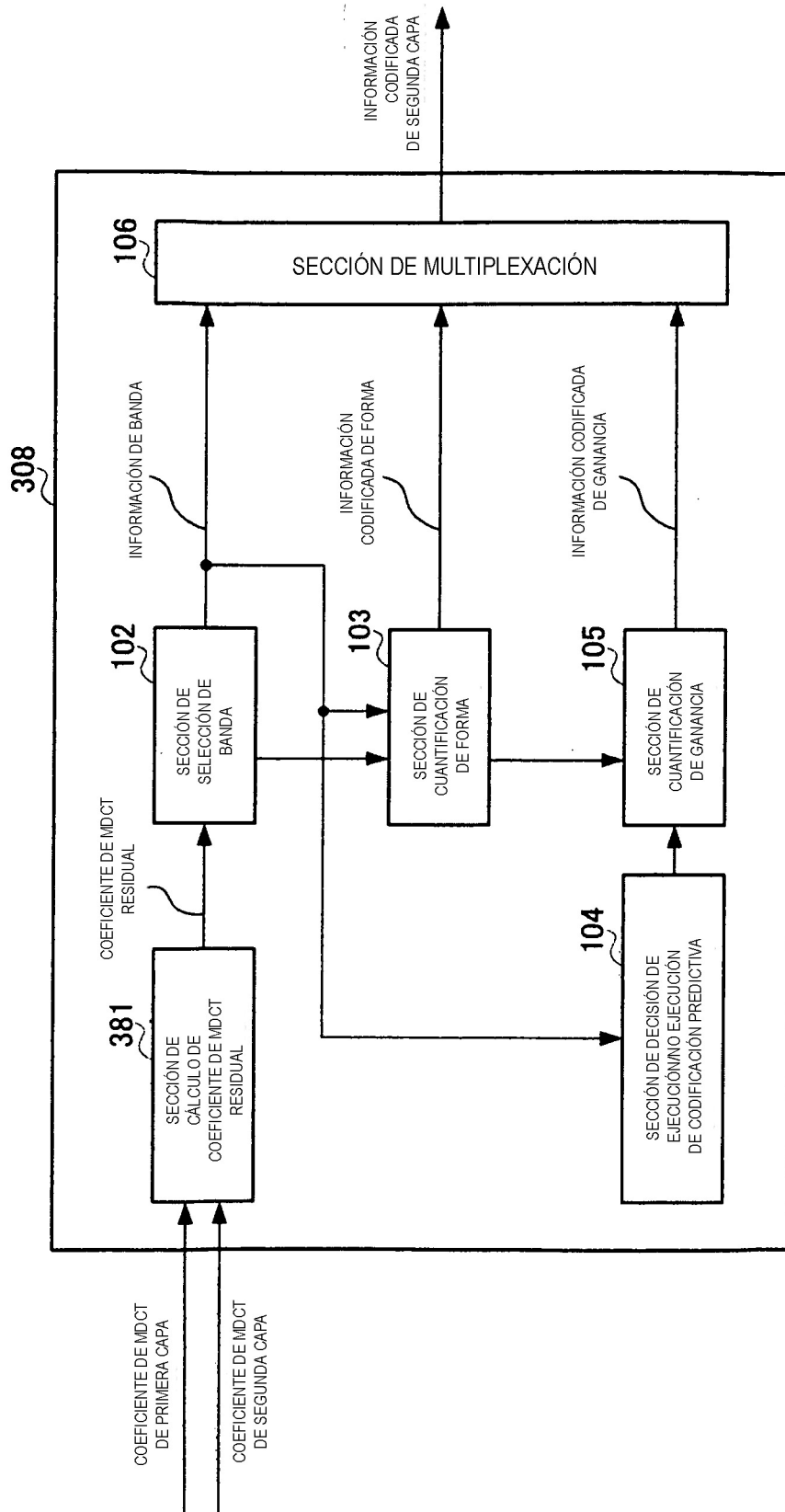


FIG.7

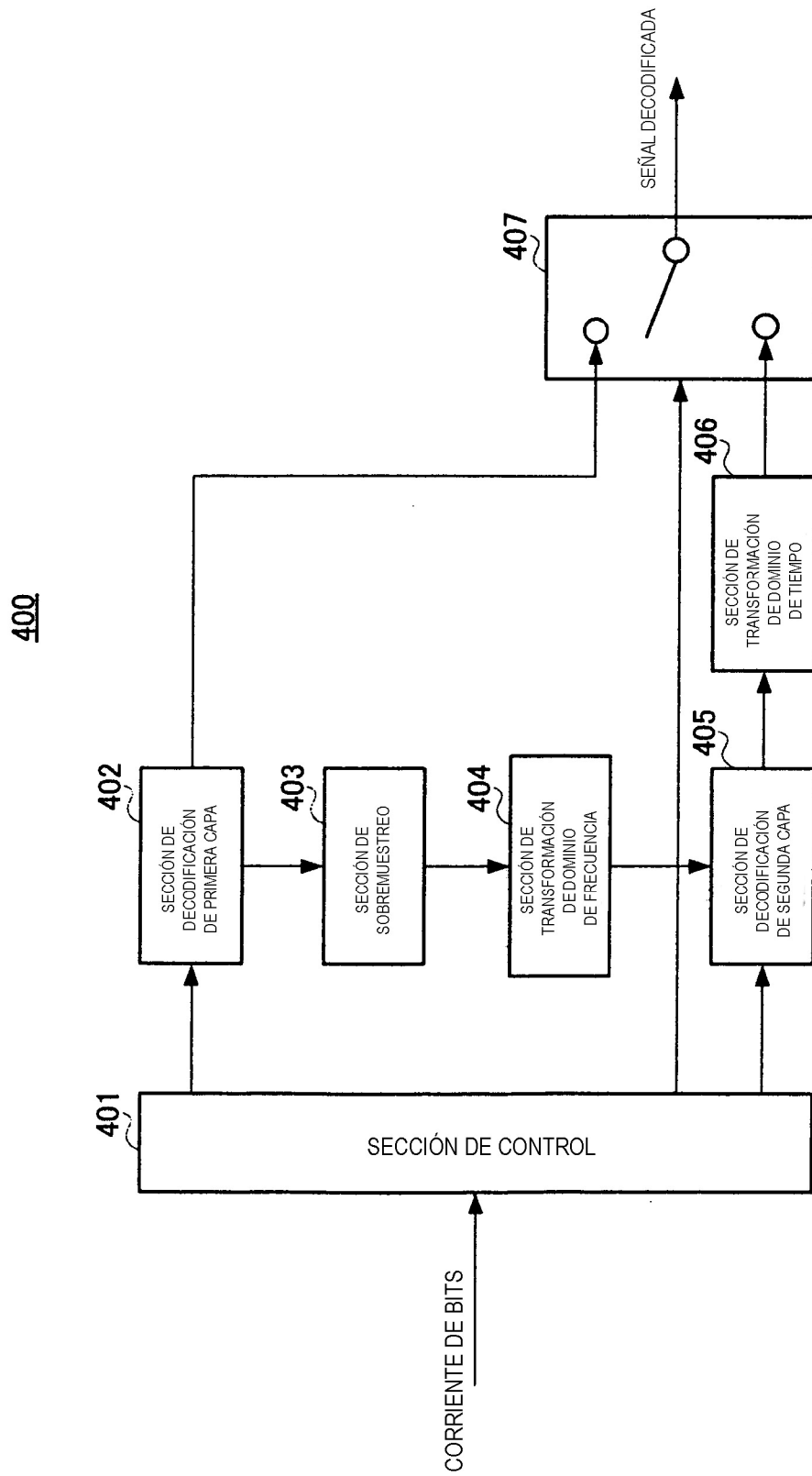


FIG.8

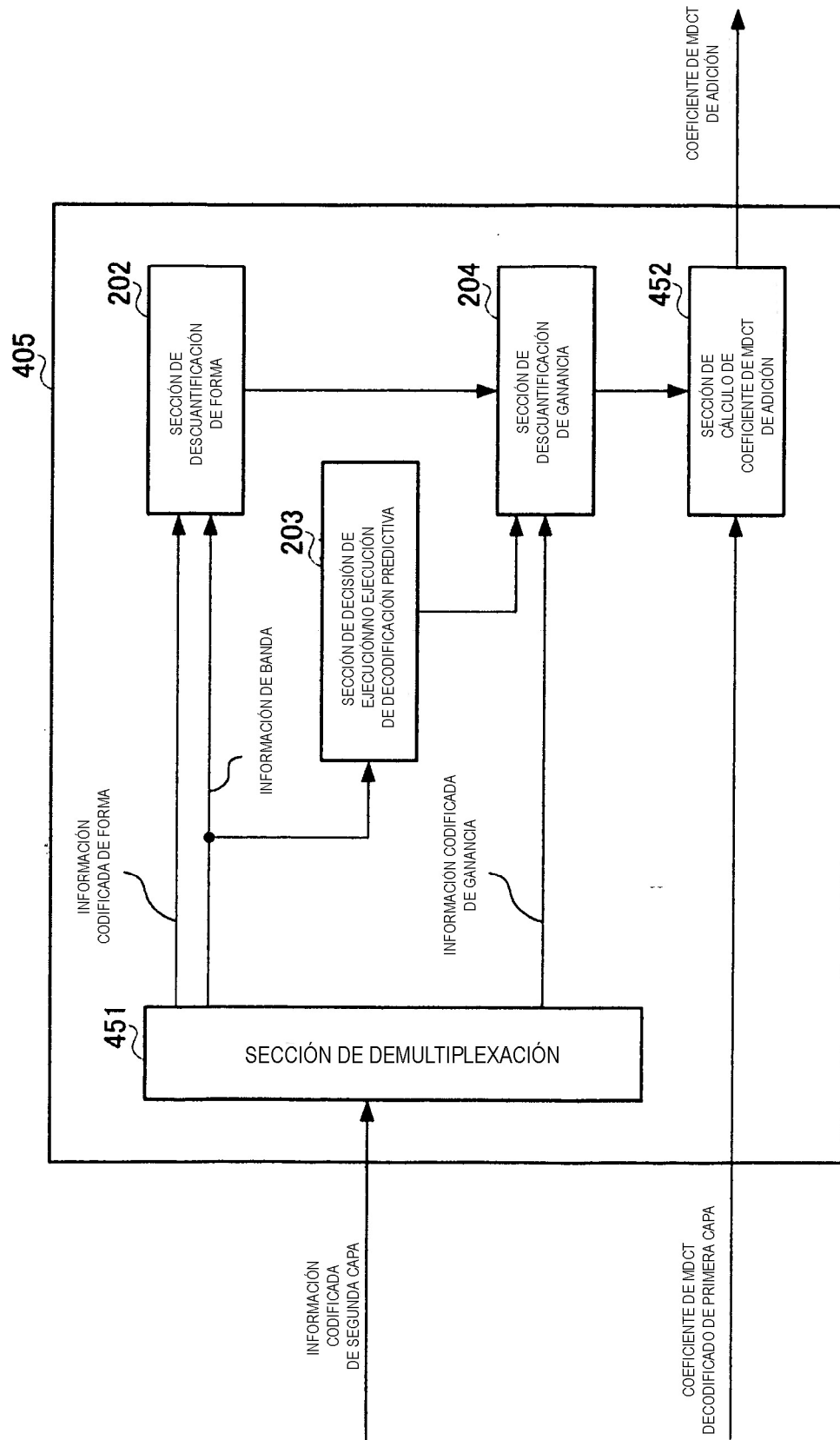


FIG.9

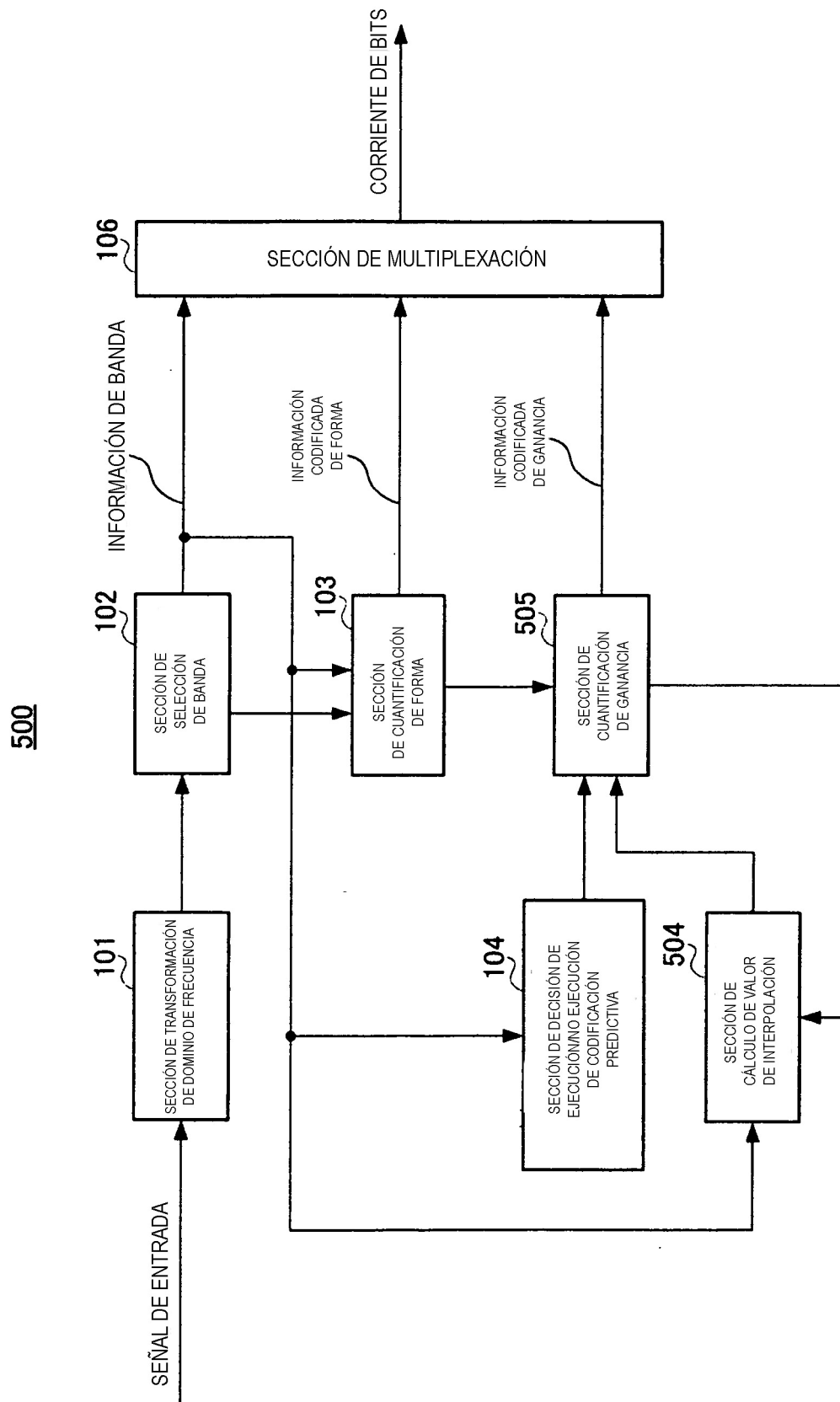


FIG.10

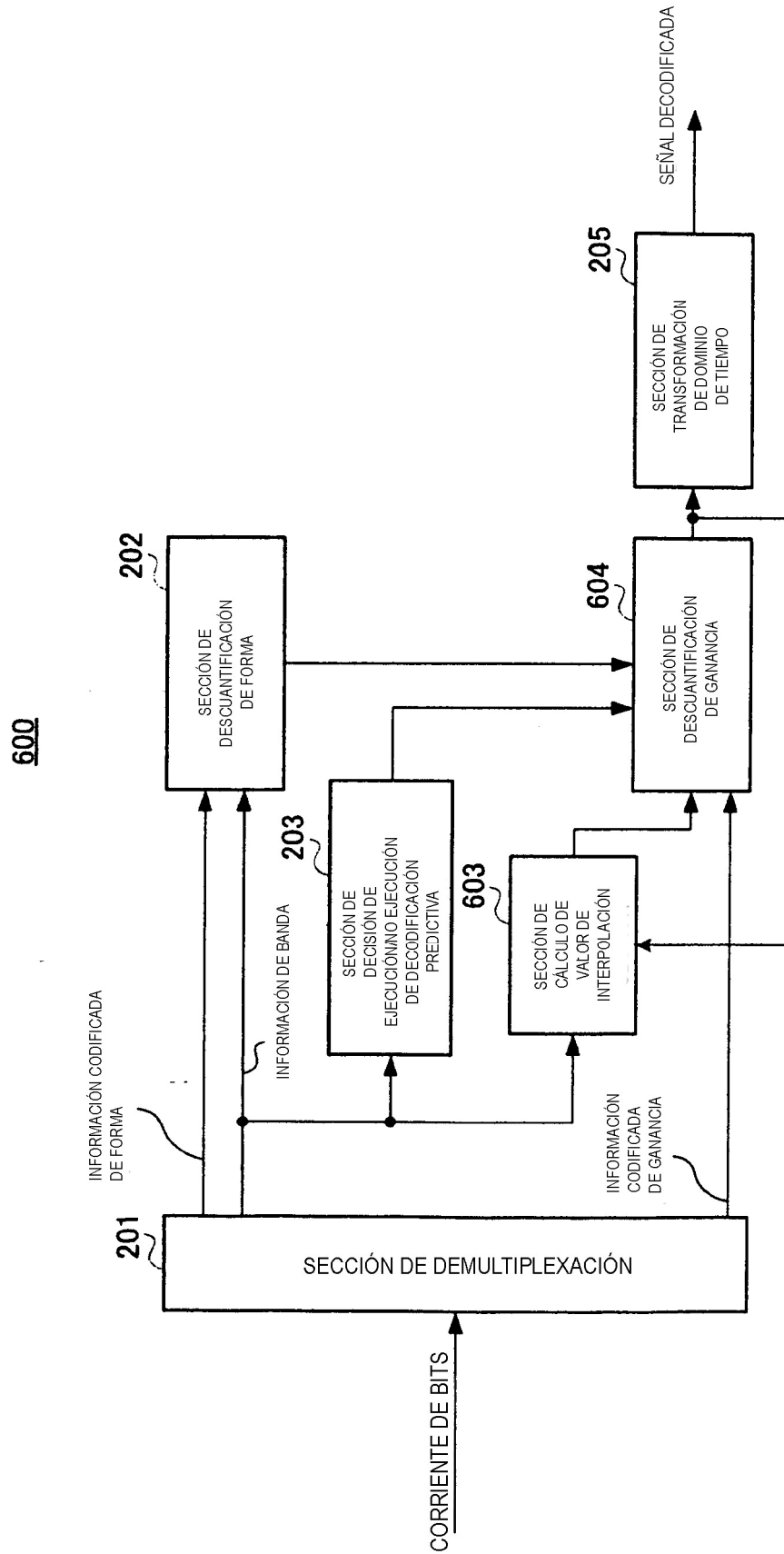


FIG.11

700

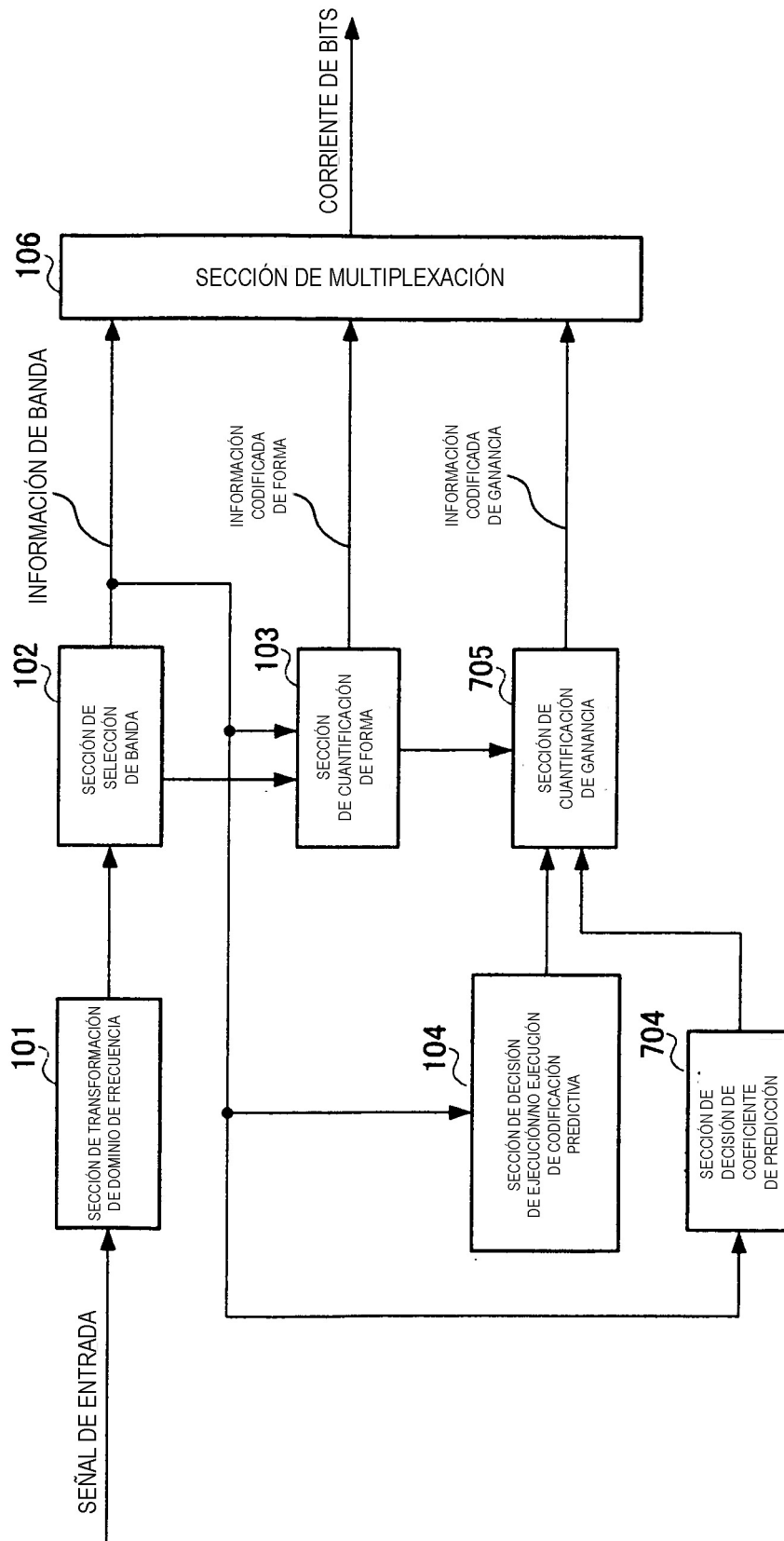


FIG.12

800

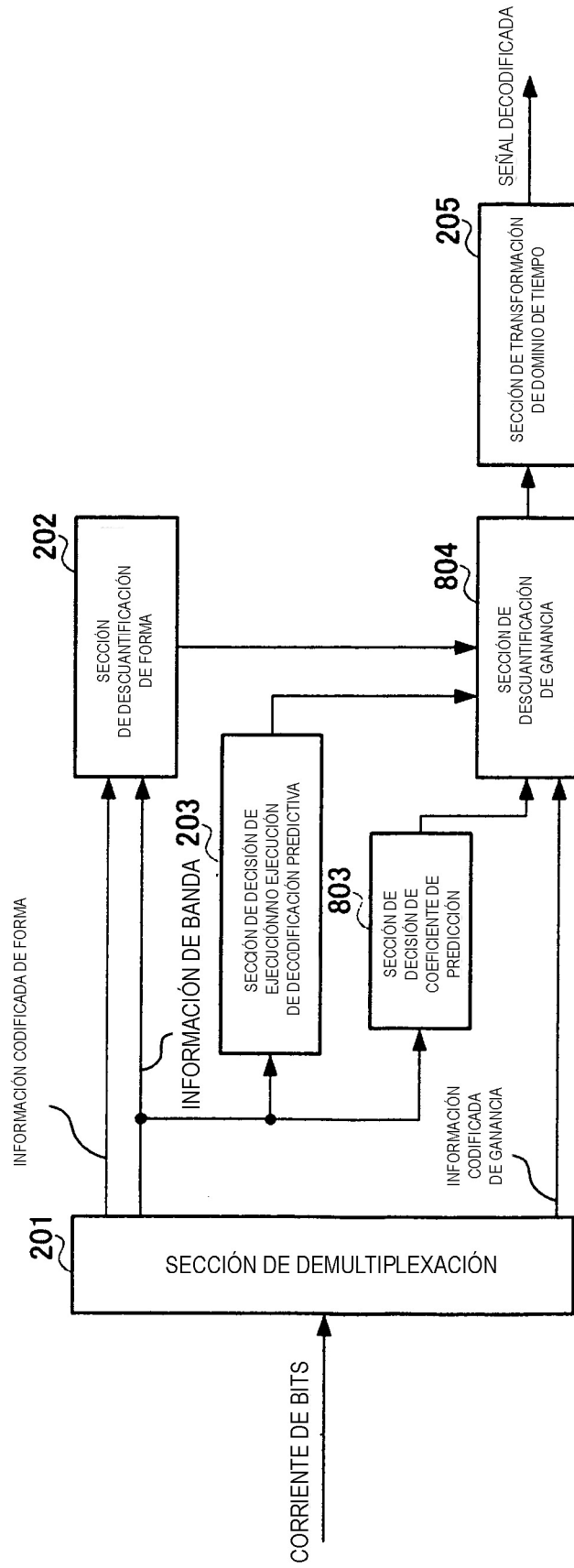


FIG.13

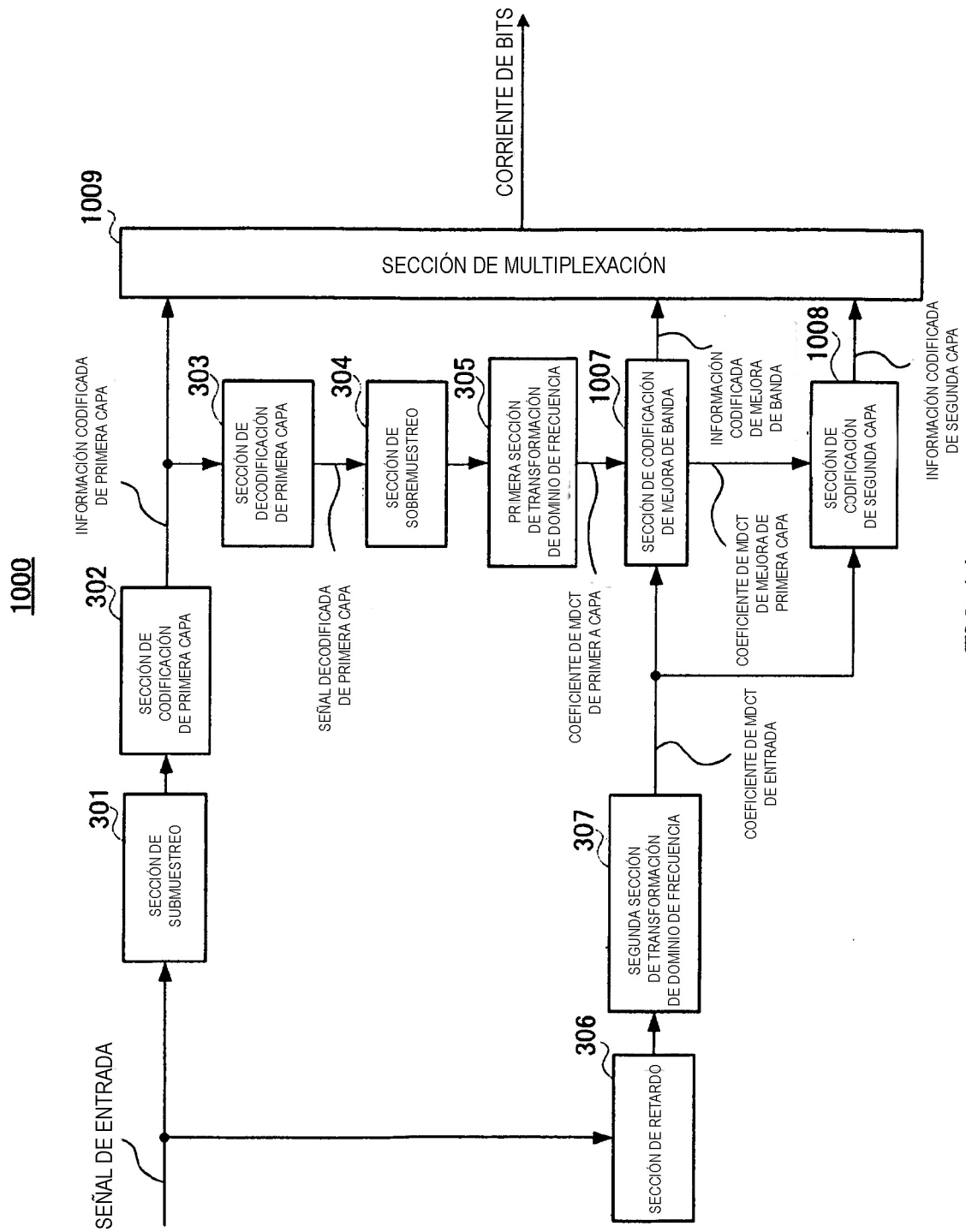


FIG.14

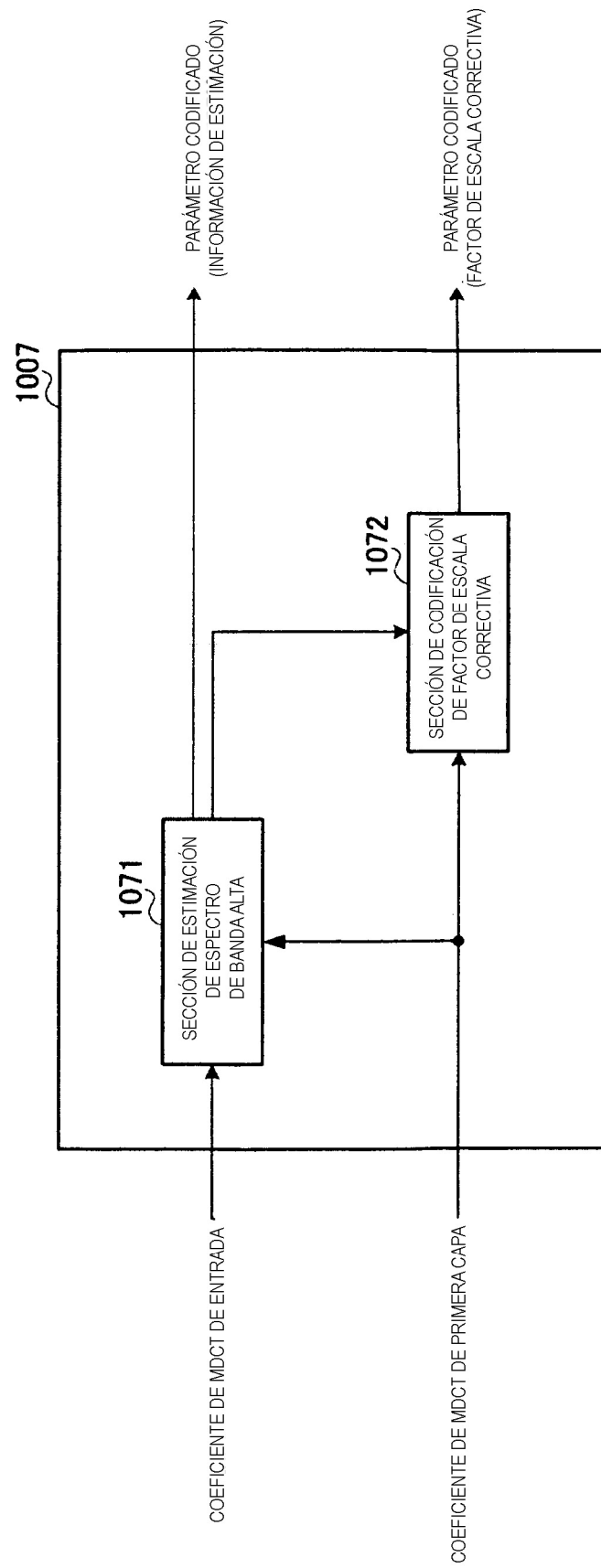


FIG.15

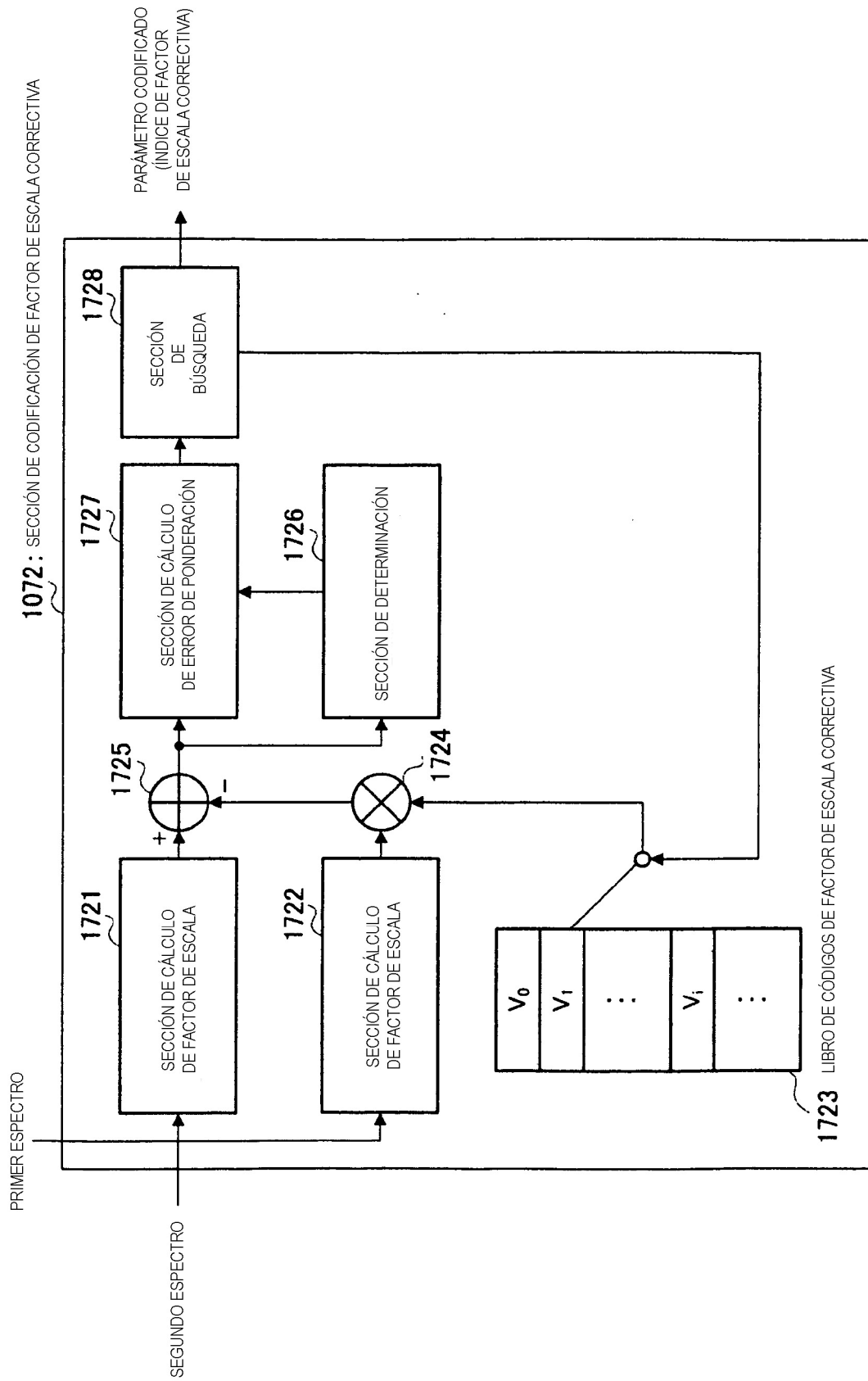


FIG.16

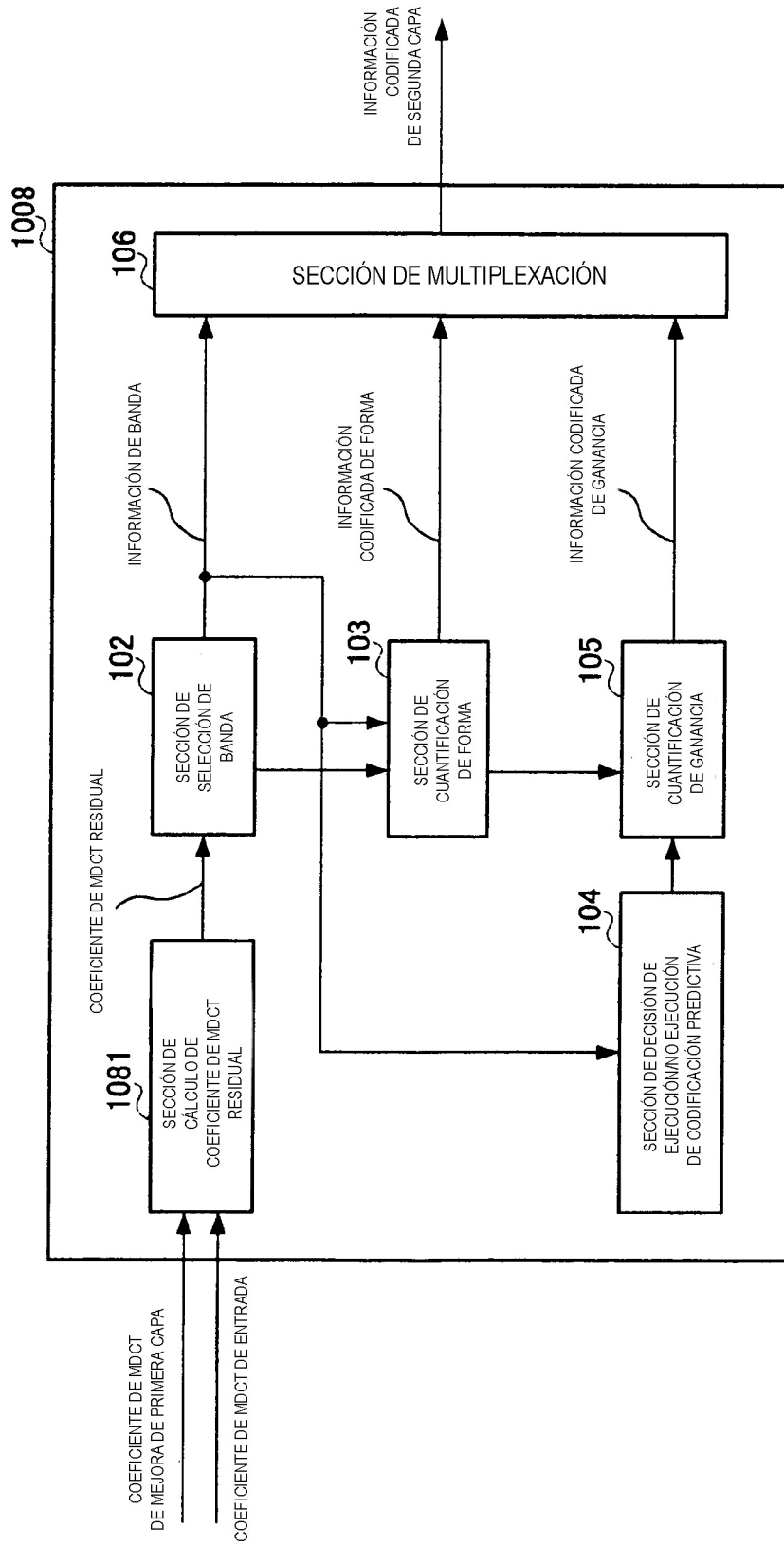


FIG.17

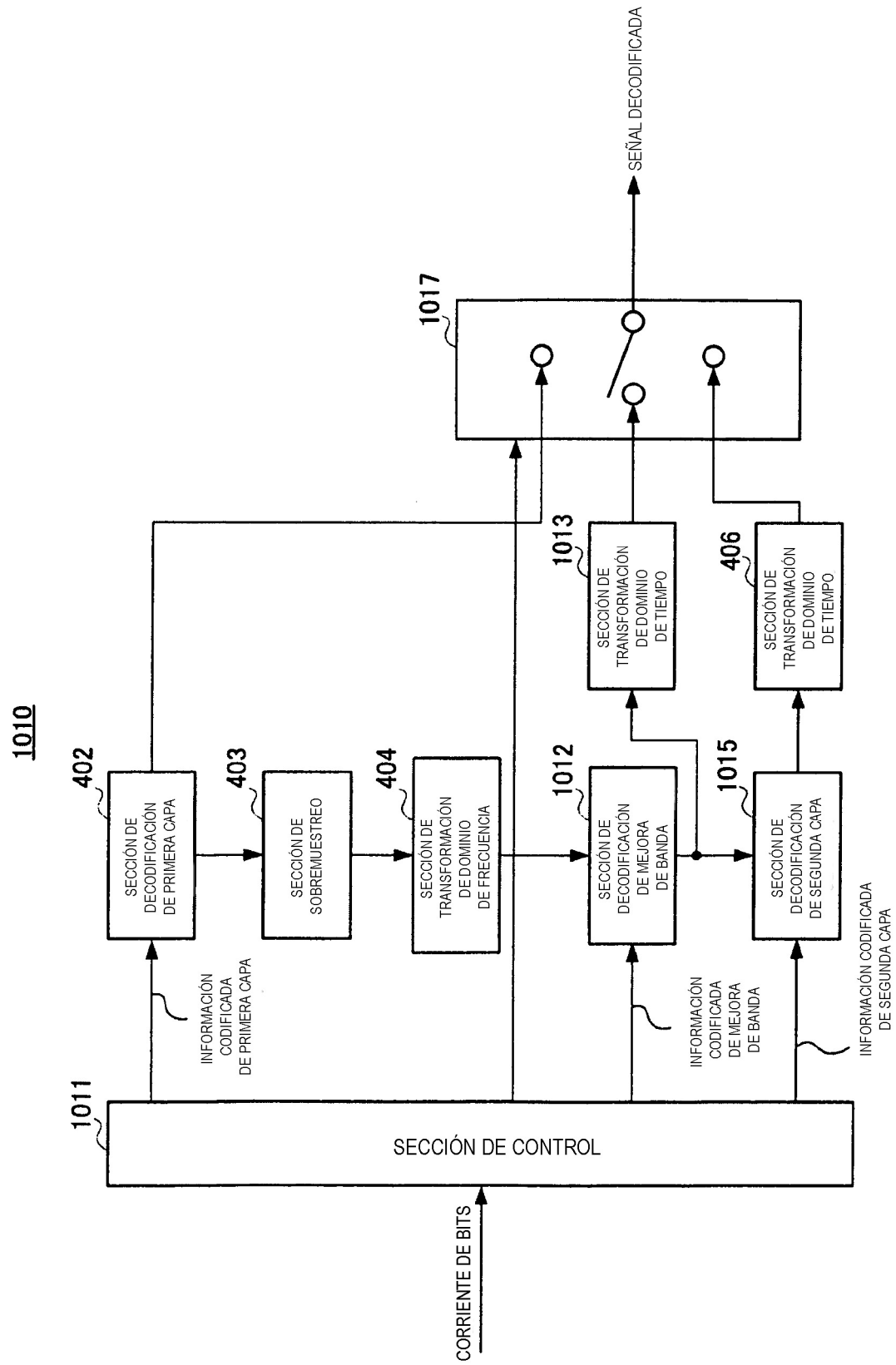


FIG.18

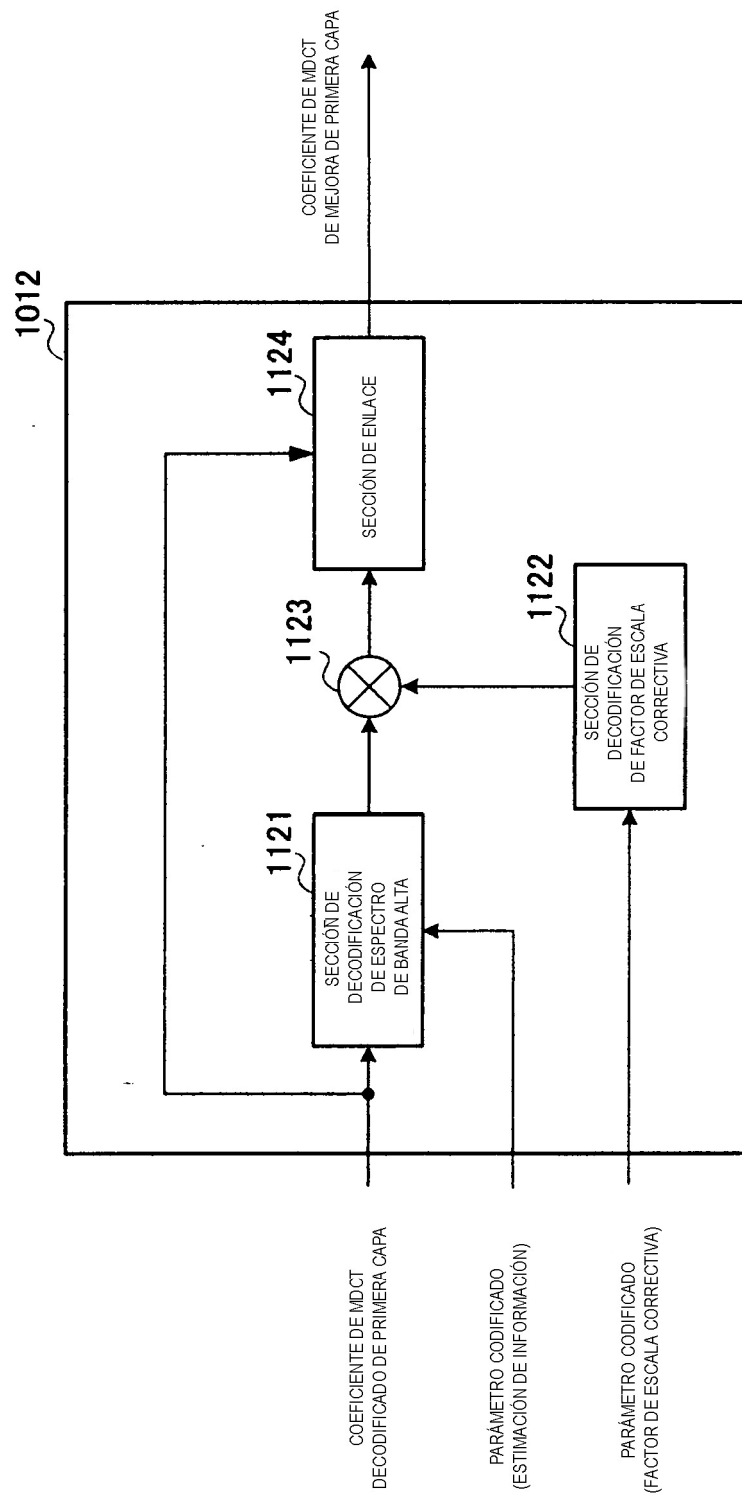


FIG.19

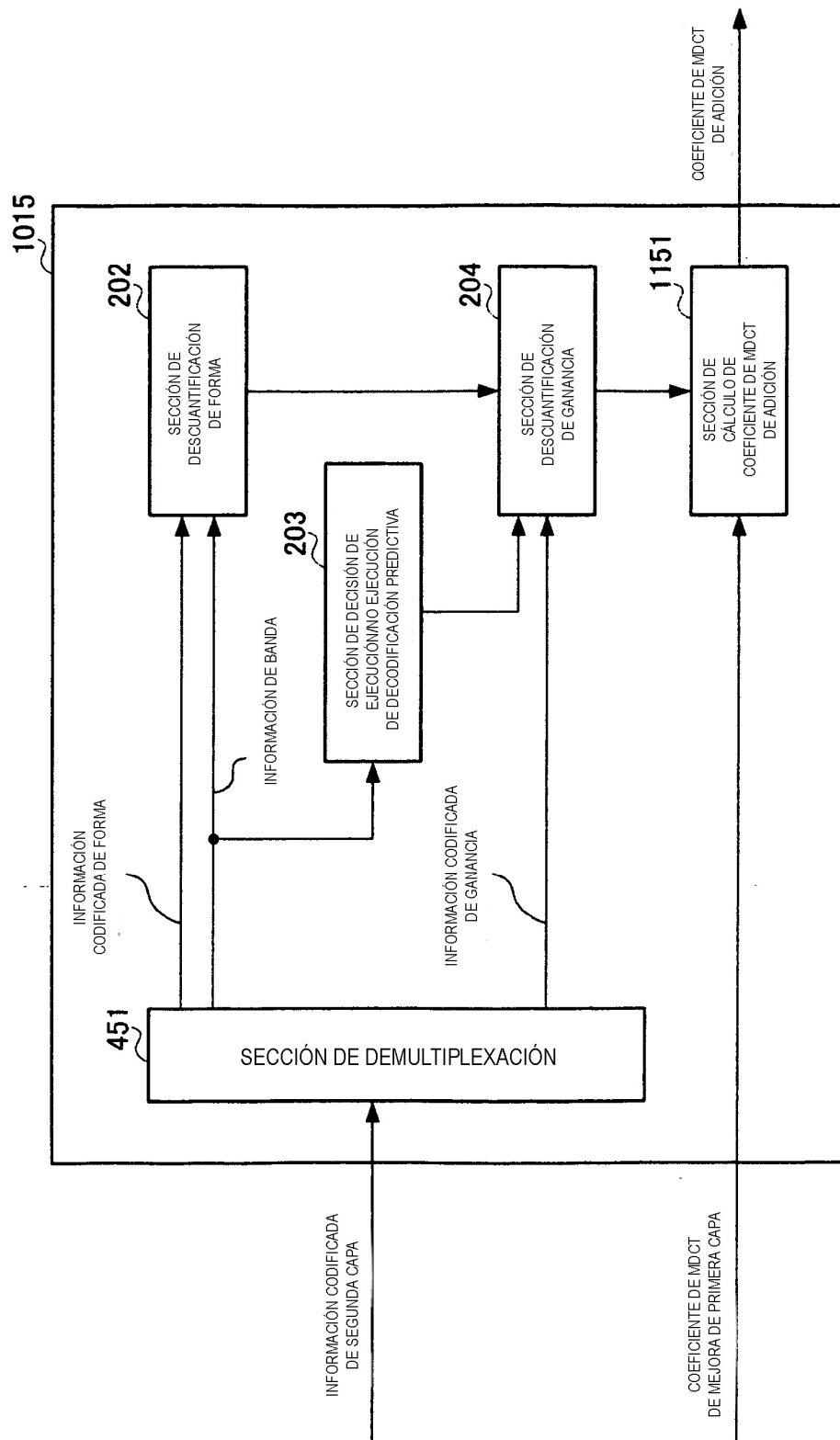


FIG.20

1100

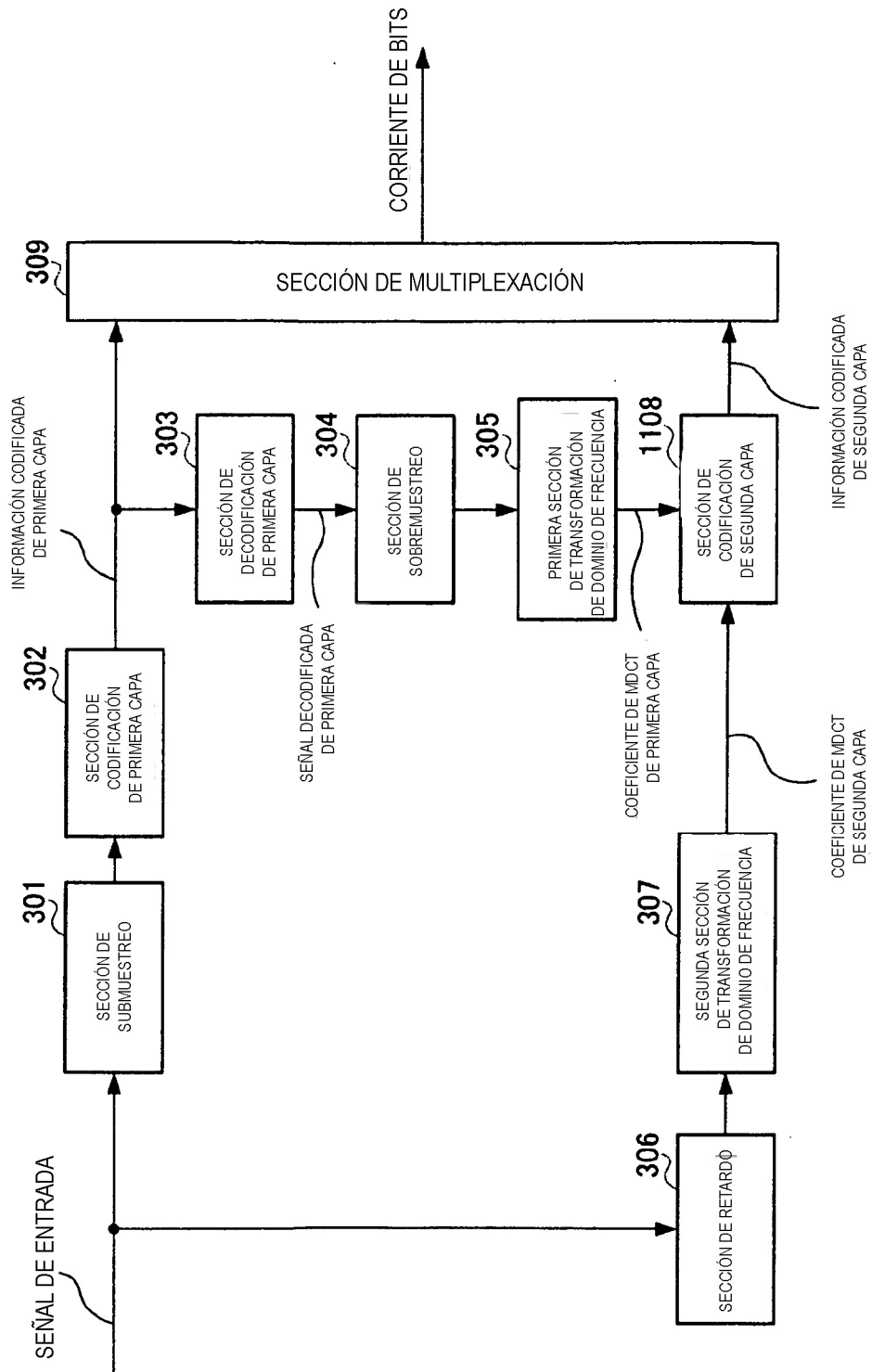


FIG.21

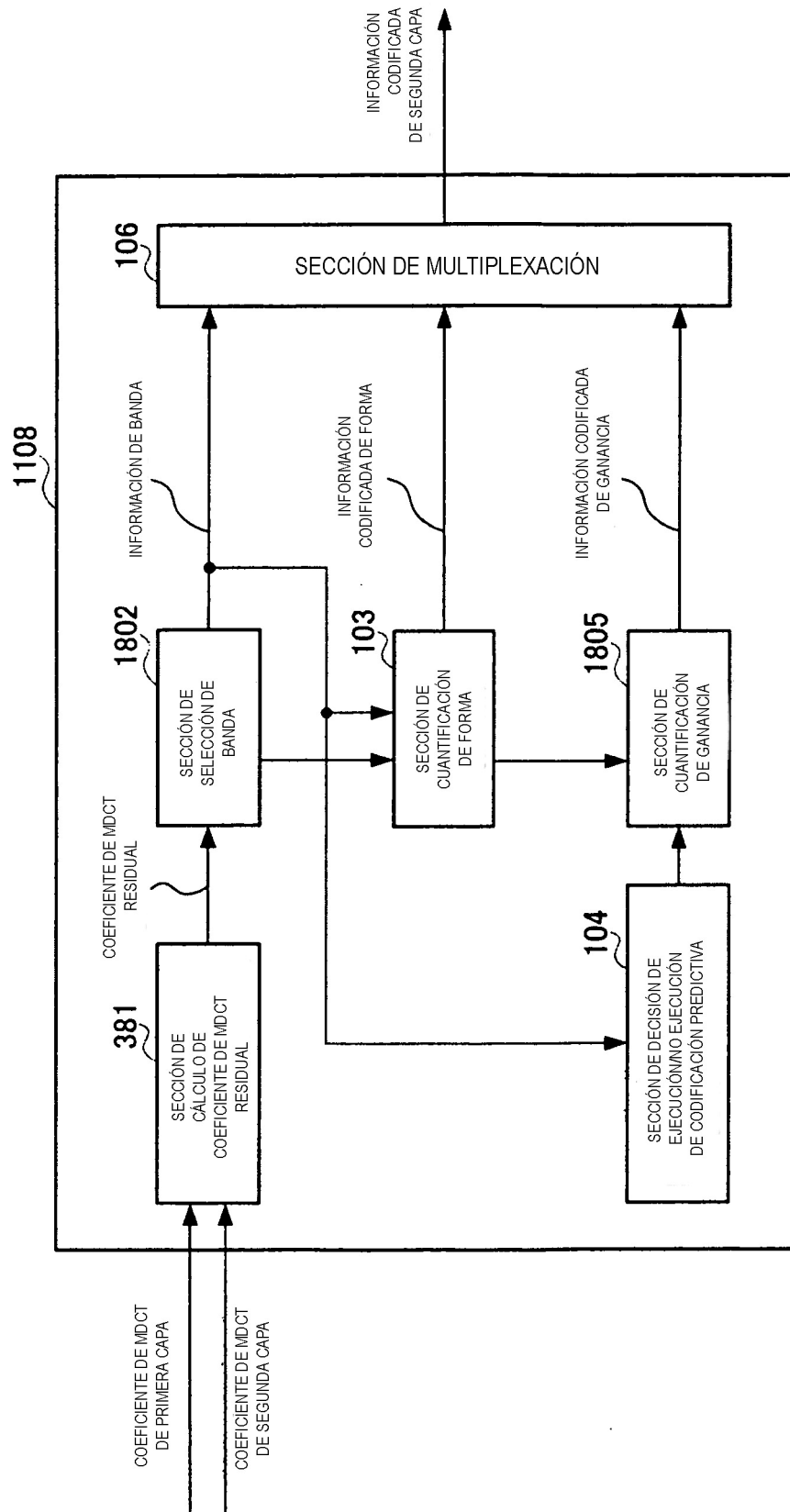


FIG.22

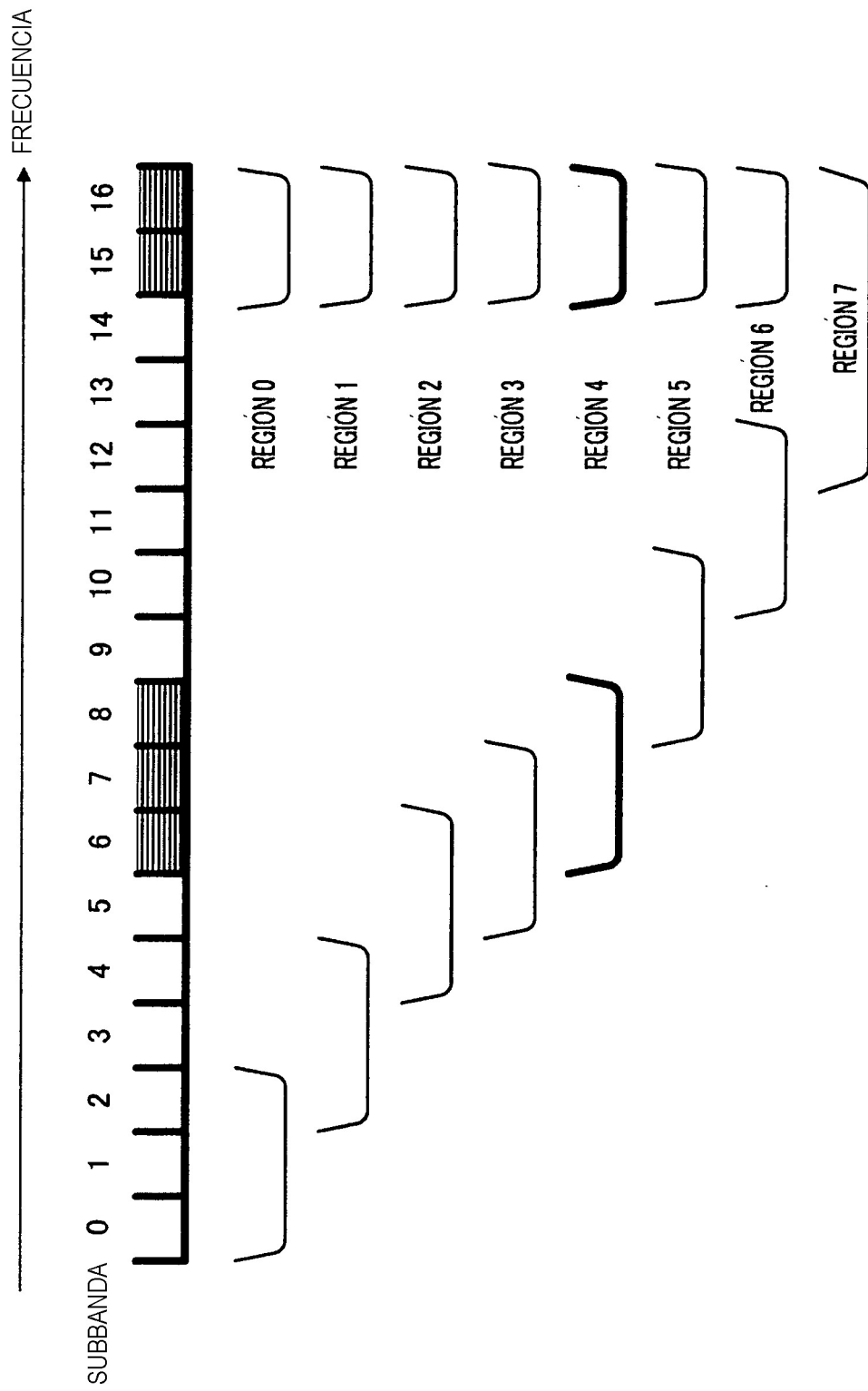


FIG.23

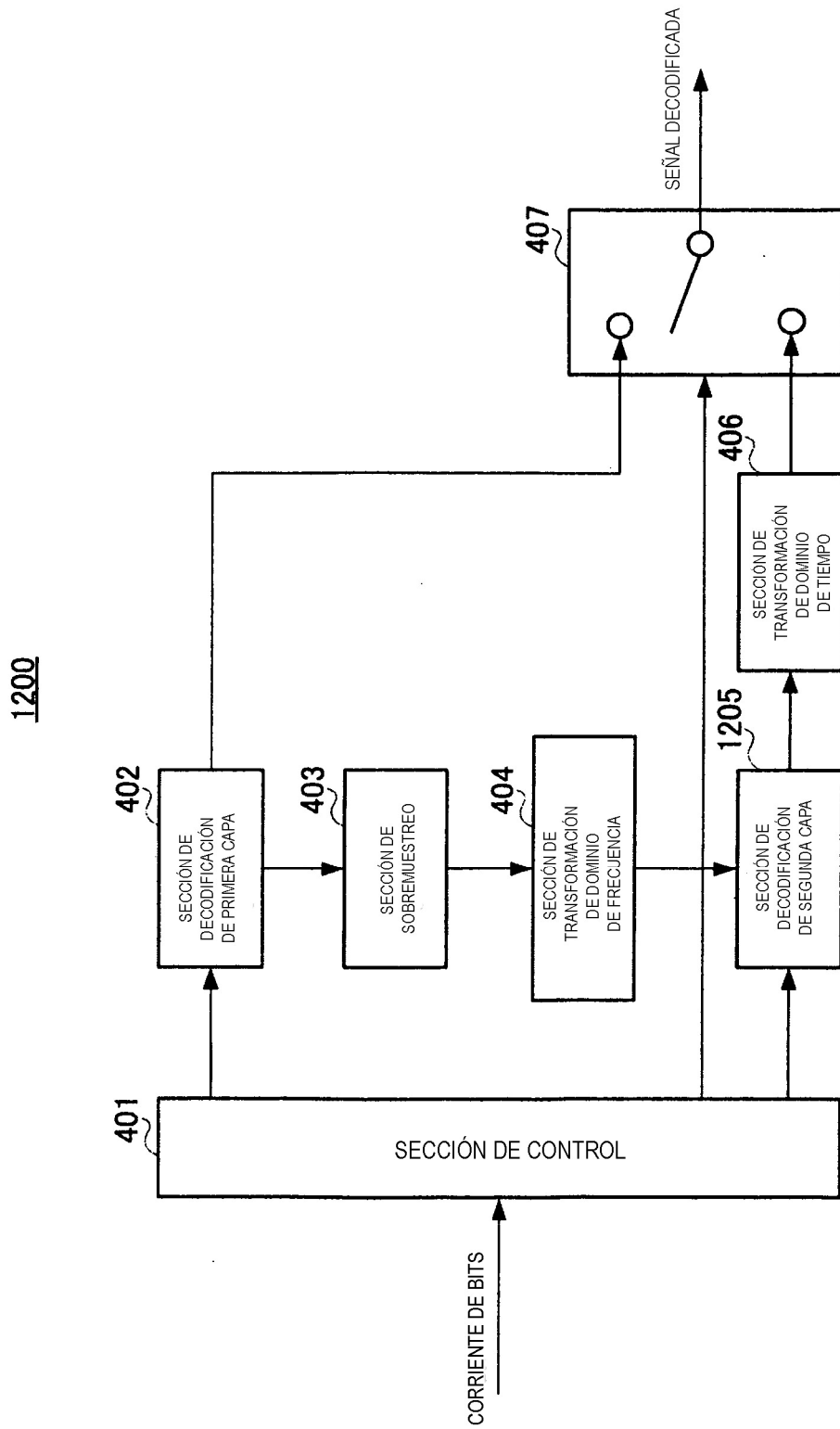


FIG.24

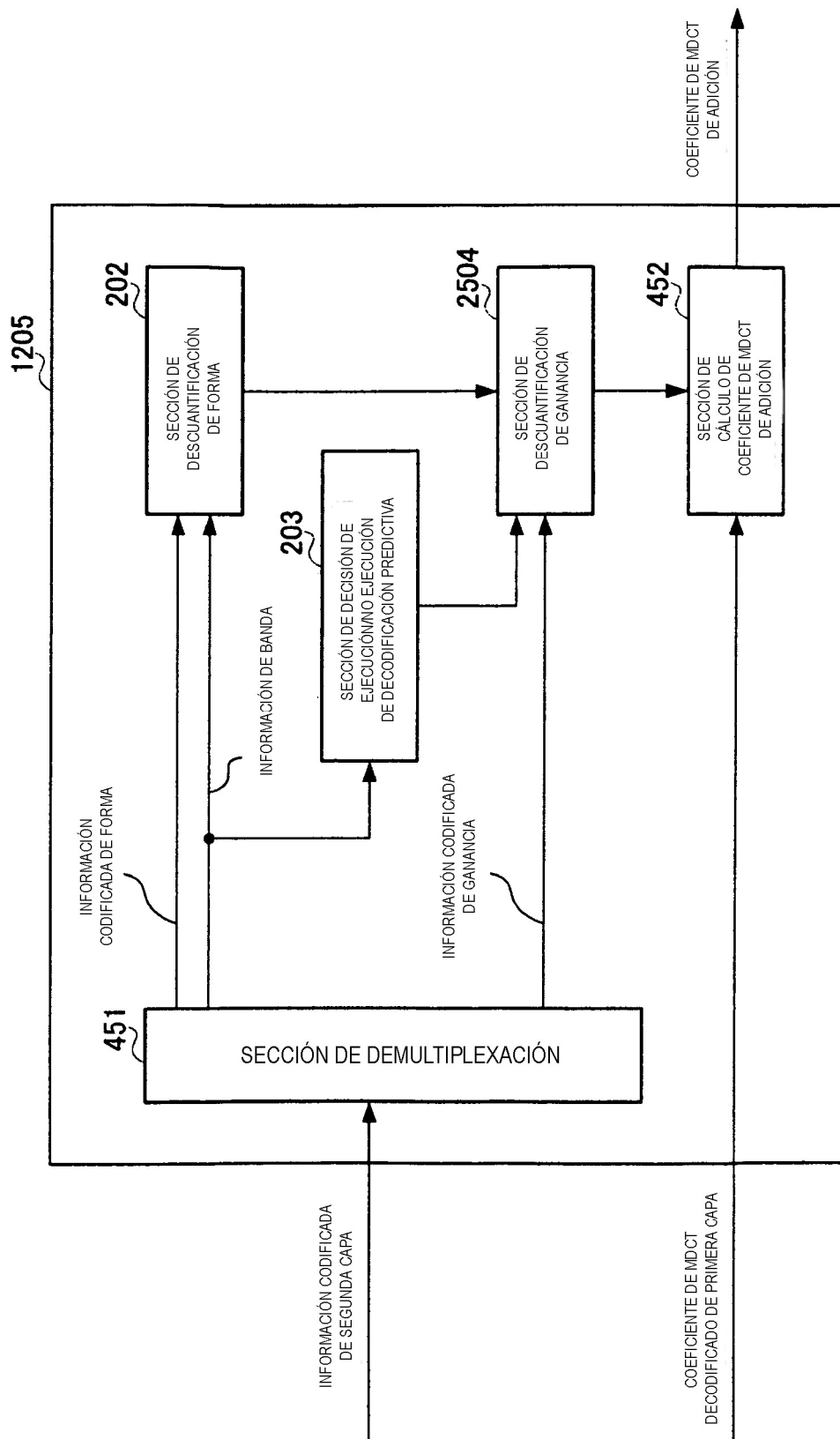


FIG.25

1300

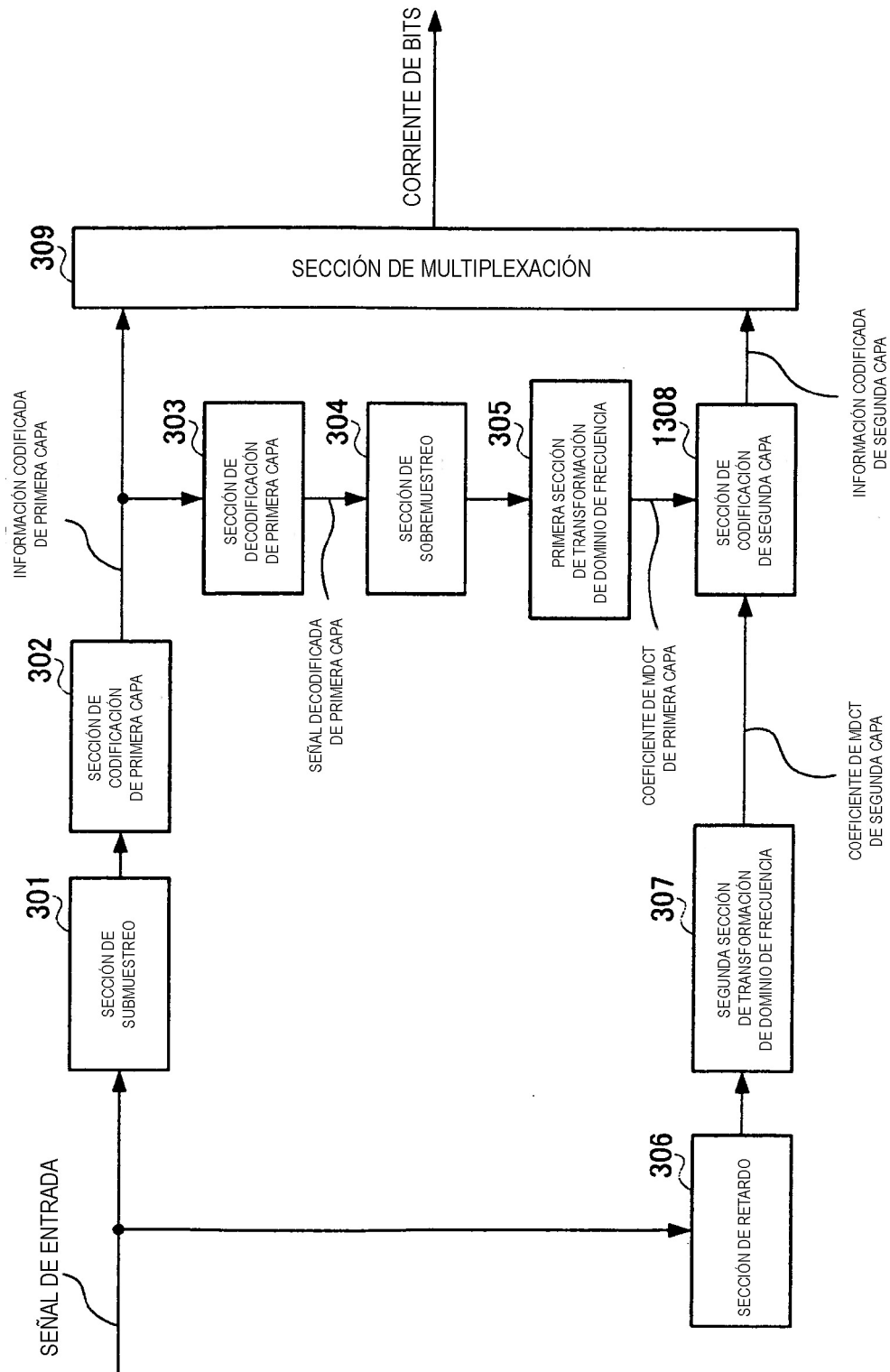


FIG.26

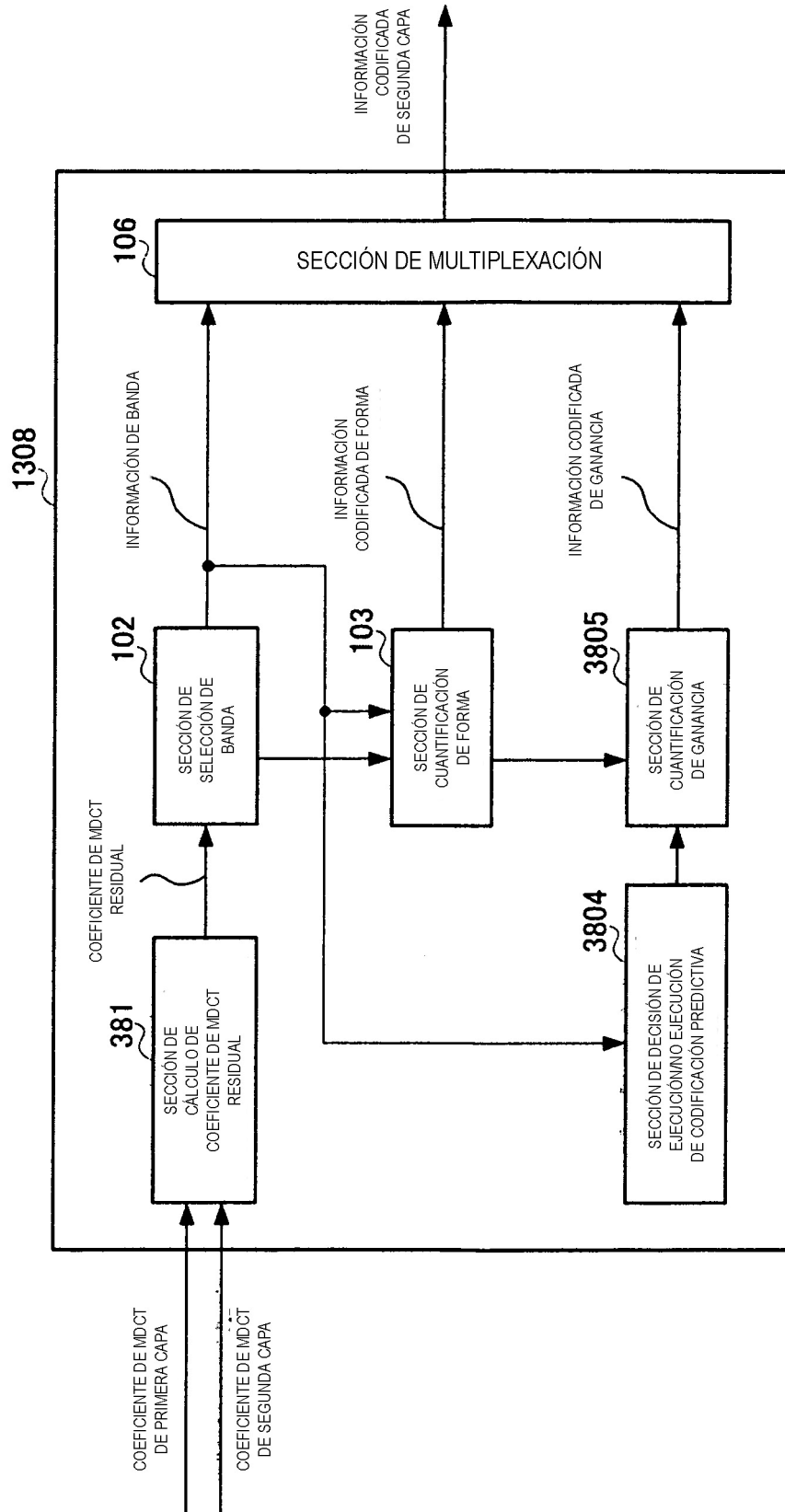


FIG.27

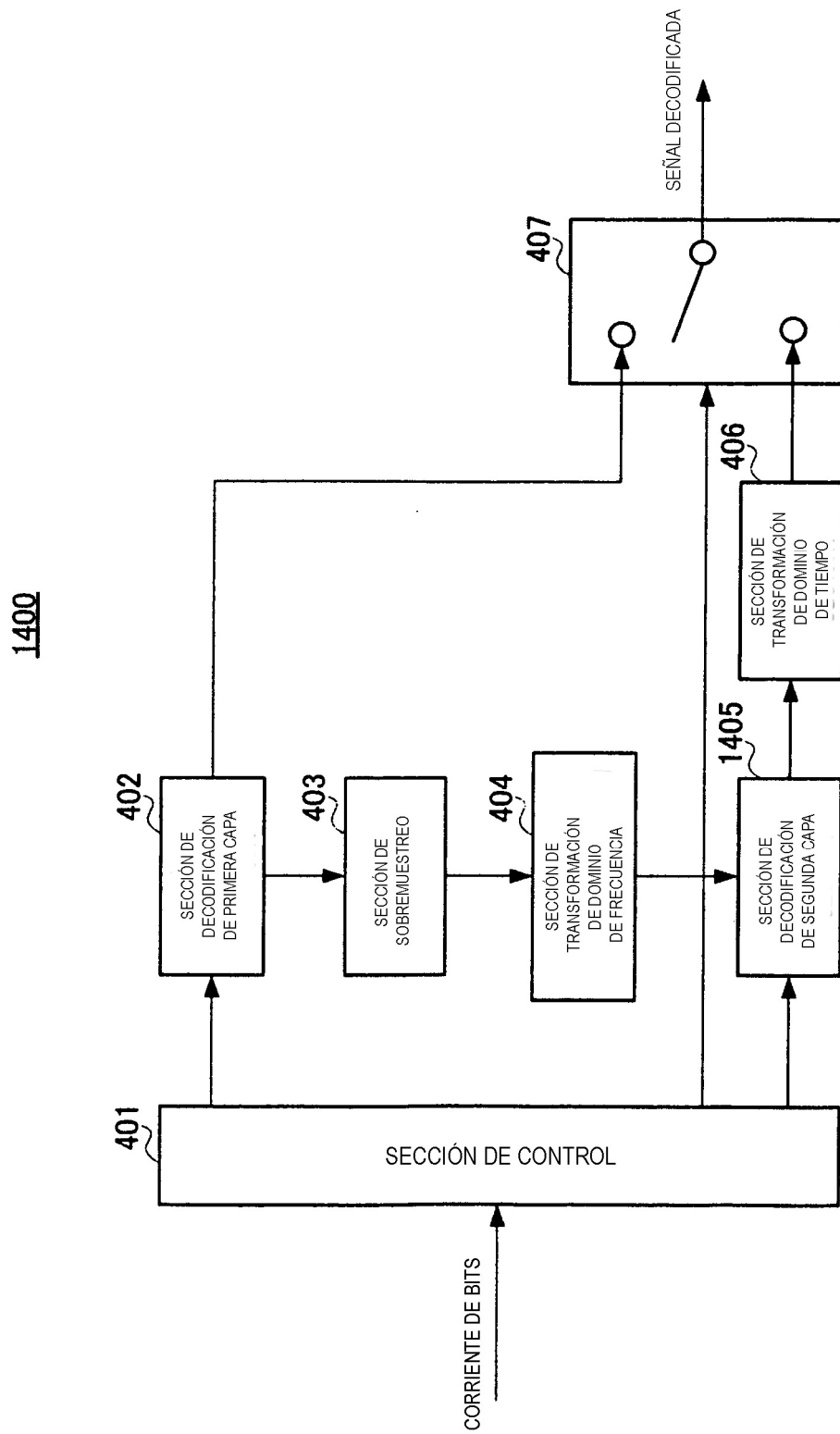


FIG.28

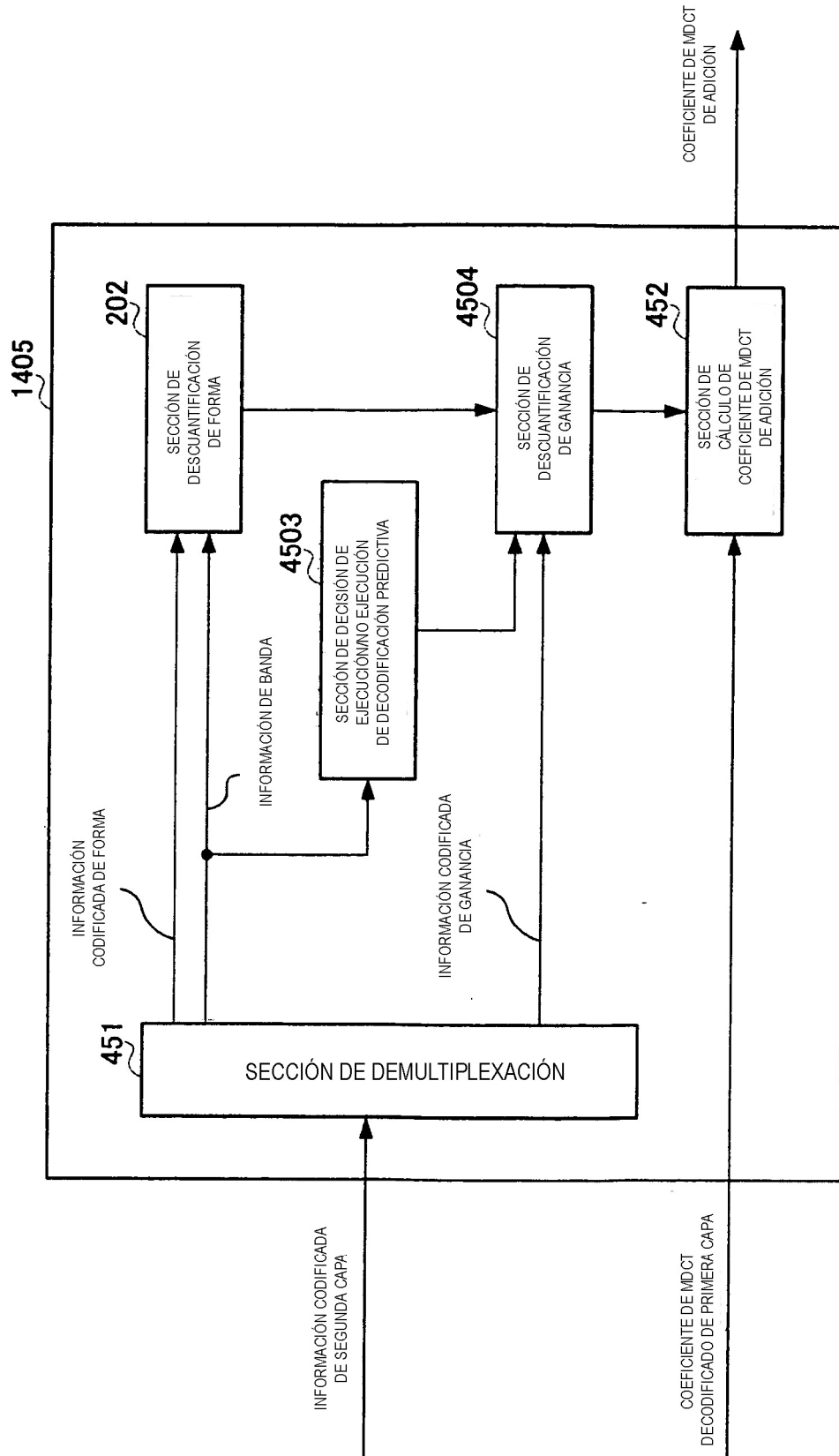


FIG.29