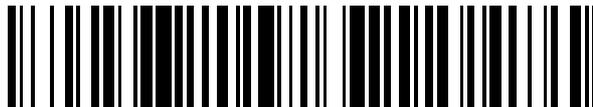


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 586 766**

51 Int. Cl.:

G10L 19/24 (2013.01)

G10L 21/04 (2013.01)

G10L 21/038 (2013.01)

G10L 19/06 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **02.04.2010 E 12171597 (3)**

97 Fecha y número de publicación de la concesión europea: **11.05.2016 EP 2503546**

54 Título: **Dispositivo de decodificación de habla, método de decodificación de habla y programa de decodificación de habla**

30 Prioridad:

03.04.2009 JP 2009091396

19.06.2009 JP 2009146831

08.07.2009 JP 2009162238

12.01.2010 JP 2010004419

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
18.10.2016

73 Titular/es:

NTT DOCOMO, INC. (100.0%)
11-1, Nagatacho 2-chome, Chiyoda-ku
Tokyo 100-6150, JP

72 Inventor/es:

TSUJINO, KOSUKE;
KIKUIRI, KEI y
NAKA, NOBUHIKO

74 Agente/Representante:

LLAGOSTERA SOTO, María Del Carmen

ES 2 586 766 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 decodificación de habla, método de decodificación de habla y programa de decodificación de habla

Ámbito Técnico

5 La presente invención se refiere a un dispositivo de decodificación de habla, a un método de decodificación de habla y a un programa de decodificación de habla

Antecedentes del Estado de la Técnica

10 Las técnicas de codificación de habla y audio para comprimir la cantidad de señales de datos unas cuantas decenas de veces extrayendo información no necesaria para la percepción humana utilizando psicoacústica son extremadamente importantes para transmitir y almacenar señales. Los ejemplos de técnicas perceptuales de codificación de audio ampliamente utilizadas incluyen "MPEG4 AAC" regularizado por "ISO/IEC MPEG", o el "MPEG HE-AAC v2 estandarizado por el Instituto Europeo de Estandarización de Telecomunicaciones (ETSI) presentado en un artículo en la EBU Technical Review de S. Meltzer & G. Moser, "MPEG-4 HE-AAC v2 – codificación de audio para el mundo de los medios digitales de hoy. 31.01.2006.

15 Una técnica de extensión de ancho de banda para generar componentes de alta frecuencia utilizando componentes de habla de baja frecuencia ha sido utilizada ampliamente en los últimos años como un método para mejorar el rendimiento de la codificación de habla y obtener una alta calidad de habla con un coeficiente de bits bajo. Entre los ejemplos típicos de la técnica de extensión de ancho de banda se encuentra la técnica de SBR (Replicación de Banda Espectral) que se utiliza en "MPEG4 AAC". En SBR, se genera un componente de alta frecuencia convirtiendo una señal en una región espectral utilizando un banco de filtros QMF (Filtro de Reflejo de Cuadratura) y copiando coeficientes espectrales de una banda de baja frecuencia a una banda de alta frecuencia con respecto a la señal transformada, y el componente de alta frecuencia se ajusta ajustando el envoltorio y la tonalidad espectrales de los coeficientes copiados. Dado que un método de codificación de habla que utiliza la técnica de extensión de ancho de banda puede reproducir los componentes de alta frecuencia de una señal utilizando únicamente una pequeña cantidad de información suplementaria, resulta efectivo a la hora de reducir el coeficiente de bits de la codificación de habla.

20 Una técnica de extensión de ancho de banda en el campo de frecuencias representado por SBR, el envoltorio y la tonalidad espectrales de los coeficientes espectrales representados en el campo de la frecuencia se ajustan ajustando un incremento en los coeficientes espectrales, realizando un filtrado inverso de predicción lineal en una dirección temporal, y superponiendo el ruido en el coeficiente espectral. Como resultado de este proceso de ajuste, cuando se codifica una señal con una gran variación en el envoltorio temporal, como por ejemplo una señal de habla, un aplauso o unas castañuelas, puede percibirse un sonido de reverberación llamado pre-eco o post eco en la señal decodificada. Este problema es debido a que el envoltorio temporal del componente de alta frecuencia se transforma durante el proceso de ajuste, y en muchos casos, el envoltorio temporal es más liso después del proceso de ajuste que antes del proceso de ajuste. El envoltorio temporal del componente de alta frecuencia después del proceso de ajuste no coincide con el envoltorio temporal del componente de alta frecuencia de una señal original antes de ser codificado, provocando de esta manera el pre-eco y el post-eco.

25 Un problema similar al del pre-eco y el post-eco se produce también en la codificación de audio multicanal que utiliza un proceso paramétrico representado por "MPEG Surround" y Stereo Paramétrico. Un decodificador utilizado en codificación de audio multicanal incluye medios para realizar la descorrelación en una señal decodificada utilizando un filtro de reverberación. Sin embargo, el envoltorio temporal de la señal se transforma durante la descorrelación, provocando de esta manera la degradación de la señal de reproducción similar a la del pre-eco y el post-eco. Las soluciones para este problema incluyen una técnica TES ("Temporal Envelope Shaping" formación del envoltorio Temporal) (Literatura de Patente 1). En la técnica TES, se realiza un análisis de predicción lineal en una dirección de frecuencia en una señal representada en un dominio QMF sobre el cual la descorrelación todavía no ha tenido lugar con el fin de obtener coeficientes de predicción lineal, y, utilizando los coeficientes de predicción lineal, se realiza el filtrado de síntesis de predicción lineal en la dirección de la frecuencia en la señal en la cual todavía no se ha realizado la descorrelación. Este proceso permite que la técnica TES extraiga el envoltorio temporal de una señal en la cual todavía no se ha realizado la descorrelación y, de acuerdo con el envoltorio temporal extraído, ajustar el envoltorio temporal de la señal en la cual se ha realizado la descorrelación. Dado que la señal en la cual todavía no se ha realizado la descorrelación tiene un envoltorio temporal menos distorsionado, el envoltorio temporal de la señal en la cual se ha realizado se ajusta a una forma menos distorsionada, obteniendo de esta forma una señal de reproducción en la cual se mejora el pre-eco y el post-eco.

Lista de citas

Literatura de patentes

Literatura de Patentes: 1. Publicación de Solicitud de Patente de Estados Unidos no. 2006/0239473

Resumen de la Invención

5 **Problema Técnico**

La técnica TES descrita más arriba es una técnica que utiliza el hecho de que una señal en la cual todavía no se ha realizado la descorrelación tiene un envoltorio temporal menos distorsionado. Sin embargo, en un decodificador SBR, el componente de alta frecuencia de una señal se copia a partir del componente de baja frecuencia de la señal. Por consiguiente, no es posible obtener un envoltorio temporal menos distorsionado con respecto al componente de alta frecuencia. Una de las soluciones a este problema es un método que analiza el componente de alta frecuencia de una señal de entrada en un codificador SBR, cuantificando los coeficientes de predicción lineal obtenidos como resultado del análisis, y multiplicándolos en una corriente de bits que va a ser transmitida. Este método permite que el decodificador SBR obtenga unos coeficientes de predicción lineal que incluyen información con un envoltorio temporal menos distorsionado del componente de alta frecuencia. Sin embargo, en este caso, se necesita una gran cantidad de información para transmitir los coeficientes de predicción lineal cuantificados, y de esta manera se incrementa de forma significativa el coeficiente de bits de toda la corriente de bits codificados. Así, la presente invención tiene la intención de reducir la aparición de pre-eco y de post-eco, a la vez que mejorar la calidad subjetiva de la señal decodificada, sin aumentar de forma significativa el coeficiente de bits en la técnica de extensión de ancho de banda en el dominio de frecuencia representado por SBR.

Solución al Problema

La presente invención proporciona un dispositivo de decodificación de habla de acuerdo con la reivindicación 1, un dispositivo de decodificación de habla de acuerdo con la reivindicación 2, un método de decodificación de habla de acuerdo con la reivindicación 4, un método de decodificación de habla de acuerdo con la reivindicación 5, un programa de decodificación de habla de acuerdo con la reivindicación 6 y un programa de decodificación de habla de acuerdo con la reivindicación 7.

Efectos Ventajosos de la Invención

De acuerdo con la presente invención, la aparición de pre-eco y post-eco puede reducirse, y la calidad subjetiva de una señal decodificada puede mejorarse sin aumentar de forma significativa el coeficiente de bits en la técnica de extensión de ancho de banda en el dominio de frecuencia representado por SBR.

Breve Descripción de los Dibujos

- La FIG. 1 es un diagrama que ilustra un dispositivo de codificación de habla de acuerdo con un primer ejemplo;
- 35 La FIG. 2 es un diagrama de flujo para describir un funcionamiento del dispositivo de codificación de habla de acuerdo con el primer ejemplo;
- La FIG. 3 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con el primer ejemplo;
- 40 La FIG. 4 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con el primer ejemplo;
- La FIG. 5 es un diagrama que ilustra un dispositivo de codificación de habla de acuerdo con una primera modificación del primer ejemplo;
- La FIG. 6 es un diagrama que ilustra un dispositivo de codificación de habla de acuerdo con un segundo ejemplo;
- 45 La FIG. 7 es un diagrama de flujo para describir un funcionamiento del dispositivo de codificación de habla de acuerdo con el segundo ejemplo;
- La FIG. 8 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con el segundo ejemplo;

ES 2 586 766 T3

La FIG. 9 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con el segundo ejemplo;

La FIG. 10 es un diagrama que ilustra un dispositivo de codificación de habla de acuerdo con un tercer ejemplo;

5 La FIG. 11 es un diagrama de flujo para describir un funcionamiento del dispositivo de codificación de habla de acuerdo con el tercer ejemplo;

La FIG. 12 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con el tercer ejemplo;

10 La FIG. 13 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con el tercer ejemplo;

La FIG. 14 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con una realización;

La FIG. 15 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con una modificación de la realización;

15 La FIG. 16 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación de la realización;

La FIG. 17 es un diagrama de flujo que describe el funcionamiento del dispositivo de codificación de habla de acuerdo con la otra modificación de la realización;

20 La FIG. 18 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación del primer ejemplo;

La FIG. 19 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con la otra modificación del primer ejemplo;

La FIG. 20 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación del primer ejemplo;

25 La FIG. 21 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con la otra modificación del primer ejemplo;

La FIG. 22 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con una modificación del segundo ejemplo;

30 La FIG. 23 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con la modificación del segundo ejemplo;

La FIG. 24 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación del segundo ejemplo;

La FIG. 25 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con la otra modificación del segundo ejemplo;

35 La FIG. 26 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación de la realización;

La FIG. 27 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con la otra modificación de la realización;

40 La FIG. 28 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación de la realización;

La FIG. 29 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con la otra modificación de la realización;

La FIG. 30 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación de la realización;

La FIG. 31 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación de la realización;

La FIG. 32 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con la otra modificación de la realización;

5 La FIG. 33 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación de la realización;

La FIG. 34 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con la otra modificación de la realización;

10 La FIG. 35 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación de la realización;

La FIG. 36 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con la otra modificación de la realización;

La FIG. 37 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación de la realización;

15 La FIG. 38 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación de la realización;

La FIG. 39 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con la otra modificación de la realización;

20 La FIG. 40 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación de la realización;

La FIG. 41 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación de habla de acuerdo con la otra modificación de la realización;

La FIG. 42 es un diagrama que ilustra un dispositivo de decodificación de habla de acuerdo con otra modificación de la realización;

25 La FIG. 43 es un diagrama de flujo para describir un funcionamiento del dispositivo de codificación de acuerdo con la otra modificación de la realización.

La FIG. 44 es un diagrama que ilustra un dispositivo de codificación de habla de acuerdo con otra modificación del primer ejemplo;

30 La FIG. 45 es un diagrama que ilustra un dispositivo de codificación de habla de acuerdo con todavía otra modificación del primer ejemplo;

La FIG. 46 es un diagrama que ilustra un dispositivo de codificación de habla de acuerdo con una modificación del segundo ejemplo;

La FIG. 47 es un diagrama que ilustra un dispositivo de codificación de habla de acuerdo con otra modificación del segundo ejemplo;

35 La FIG. 48 es un diagrama que ilustra un dispositivo de codificación de habla de acuerdo con el cuarto ejemplo;

La FIG. 49 es un diagrama que ilustra un dispositivo de codificación de habla de acuerdo con otra modificación del cuarto ejemplo; y

40 La FIG. 50 es un diagrama que ilustra un dispositivo de codificación de habla de acuerdo con otra modificación del cuarto ejemplo.

Descripción de las Realizaciones

A continuación se describen en detalle realizaciones preferentes de la presente invención y diferentes ejemplos que resultan útiles para comprender la presente invención en referencia a los dibujos adjuntos. En la descripción de los dibujos, los elementos que son el mismo elemento están etiquetados con los mismos símbolos de referencia, y la descripción duplicada de los mismos ha sido omitida, en su caso.

45

(Primer ejemplo)

La FIG. 1 es un diagrama que ilustra un dispositivo de codificación de habla 11 de acuerdo con un primer ejemplo. El dispositivo de codificación de habla 11 incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 11 cargando y ejecutando un programa de ordenador predeterminado (como por ejemplo un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 2) almacenado en una memoria integrada del dispositivo de codificación de habla 11 como por ejemplo la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 11 recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 11, y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 11.

El dispositivo de codificación de habla 11 incluye funcionalmente una unidad de transformación de frecuencia 1a (medio de transformación de frecuencia), una unidad de transformación de frecuencia inversa 1b, una unidad central de codificación de codec 1c (medio central de codificación), una unidad de codificación SBR 1d, una unidad de análisis de predicción lineal 1e (medio de cálculo de información suplementaria de envoltorio temporal), una unidad de cálculo de parámetros de intensidad de filtro 1f (medio de cálculo de información suplementaria de envoltorio temporal), y una unidad de multiplicación de corriente de bits 1g (medio de multiplicación de corriente de bits). La unidad de transformación de frecuencia 1a hasta la unidad de multiplicación de corriente de bits 1g del dispositivo de codificación de habla 11 ilustradas en la FIG. 1 son funciones realizadas cuando la CPU del dispositivo de codificación de habla 11 ejecuta el programa de ordenador almacenado en la memoria integrada del dispositivo de codificación de habla 11. La CPU del dispositivo de codificación de habla 11 ejecuta procesos de forma secuencial (procesos desde la Fase Sa1 a la Fase Sa7) ilustrados en el diagrama de flujo de la FIG. 2, ejecutando el programa de ordenador (o utilizando la unidad de transformación de frecuencia 1a hasta la unidad de multiplicación de corriente de bits 1g ilustradas en la FIG. 1). Diferentes tipos de datos necesarios para ejecutar el programa de ordenador y diferentes tipos de datos generados por la ejecución del programa de ordenador se almacenan en la memoria integrada, como por ejemplo la ROM y la RAM del dispositivo de codificación de habla 11.

La unidad de transformación de frecuencia 1a analiza una señal de entrada recibida desde fuera del dispositivo de codificación de habla 11 a través del dispositivo de comunicación del dispositivo de codificación de habla 11 utilizando un banco de filtros QMF multi-división para obtener una señal $q(k, r)$ en un dominio QMF (proceso en la Fase Sa1). Debe observarse que k ($0 \leq k \leq 63$) es un índice en una dirección de frecuencia, y r es un índice que indica una ranura de tiempo. La unidad de transformación inversa de frecuencia 1b sintetiza la mitad de los coeficientes en la parte de baja frecuencia en la señal del dominio de QMF obtenida por la unidad de transformación de frecuencia 1a utilizando el banco de filtros QMF para obtener una señal de dominio de tiempo de submuestreo que incluye solamente componentes de baja frecuencia de la señal de entrada (proceso en la Fase Sa2). La unidad central de codificación de códecs 1c codifica la señal de dominio de tiempo de submuestreo para obtener una corriente de bits codificados (proceso en la Fase Sa3). La codificación realizada por la unidad central de codificación de códecs 1c puede basarse en un método de codificación de habla representado por un método CELP, o puede estar basada en un método de codificación de audio como un código de transformación representado por un método AAC o TCX (Transform Coded Excitation) Transformar Excitación Codificada.

La unidad de codificación SBR 1d recibe la señal en el dominio QMF desde la unidad de transformación de frecuencia 1a, y realiza la codificación SBR basada en el análisis de la potencia, el cambio de señal, la tonalidad y similares de los componentes de alta frecuencia para obtener información suplementaria de SBR (proceso en Fase Sa4). El método de análisis de QMF en la unidad de transformación de frecuencia 1a y el método de codificación de SBR en la unidad de codificación de SBR 1d se describen en detalle en, por ejemplo, la referencia de Literatura "3GPP TS 26.404: Enhanced aacPlus encoder SBR part".

La unidad de análisis de predicción lineal 1e recibe la señal en el dominio QMF desde la unidad de transformación de frecuencia 1a, y realiza un análisis de predicción lineal en la dirección de frecuencia en los componentes de alta frecuencia de la señal para obtener coeficientes de predicción lineal de alta frecuencia $a_H(n, r)$ ($1 \leq n \leq N$) (proceso en Fase Sa5). Debe hacerse constar que N es un orden de predicción lineal. El índice r es un índice en una dirección temporal para una sub-muestra de las señales en el dominio QMF. Puede utilizarse un método de covariación o un método de autocorrelación para el análisis de predicción lineal de la señal. El análisis de predicción lineal para obtener $a_H(n, r)$ se realiza en los componentes de alta frecuencia que cumplen $k_x < k \leq 63$ en $q(k, r)$. Debe hacerse constar que k_x es un índice de frecuencia que corresponde a una frecuencia de límite superior de la banda de frecuencia codificada por la unidad central de codificación 1c. La unidad de análisis de predicción lineal 1e puede también realizar un análisis de predicción lineal en componentes de baja frecuencia distintos de los analizados cuando se obtiene $a_H(n, r)$ para obtener coeficientes de predicción lineal de baja frecuencia a_L

(n, r) distintos de $a_H(n, r)$ (los coeficientes de predicción lineal de acuerdo con dichos componentes de baja frecuencia corresponden a información del envoltorio temporal, y es la misma en el primer ejemplo más abajo). El análisis de predicción lineal para obtener $a_L(n, r)$ se realiza en componentes de baja frecuencia que cumplen $0 \leq k \leq k_x$. El análisis de predicción lineal puede realizarse también en una parte de la banda de frecuencia incluida en una sección de $0 \leq k < k_x$.

La unidad de cálculo de parámetros de intensidad de filtro 1f, por ejemplo, utiliza los coeficientes de predicción lineal obtenidos por la unidad de análisis de predicción lineal 1e para calcular un parámetro de intensidad de filtro (el parámetro de intensidad de filtro corresponde a la información suplementaria del envoltorio temporal y es la misma en el primer ejemplo que en las siguientes) (proceso en Fase Sa6). Un incremento de predicción $G_H(r)$ es calculado en primer lugar a partir de $a_H(n, r)$. El método para calcular el incremento de predicción se describe en detalle, por ejemplo, en "Speech Coding, Takehiro Moriya, The Institute of Electronics, Information and Communication Engineers". Si se ha calculado $a_L(n, r)$, se calcula de forma similar un incremento $G_L(r)$. El parámetro de intensidad de filtro $K(r)$ es un parámetro que aumenta a medida que aumenta $G_H(r)$, y por ejemplo, puede obtenerse de acuerdo con la expresión (1) siguiente. Aquí, $\max(a, b)$ indica el valor máximo de a y b, y $\min(a, b)$ indica el valor mínimo de a y b.

$$K(r) = \max(0, \min(1, G_H(r) - 1)) \quad \text{---(1)}$$

Si se ha calculado $G_L(r)$, puede obtenerse $K(r)$ como un parámetro que aumenta a medida que aumenta $G_H(r)$, y disminuye a medida que aumenta $G_L(r)$. En este caso, por ejemplo, puede obtenerse K de acuerdo con la expresión (2) siguiente.

$$K(r) = \max(0, \min(1, G_H(r)/G_L(r) - 1)) \quad \text{---(2)}$$

$K(r)$ es un parámetro que indica la intensidad para ajustar el envoltorio temporal de los componentes de alta frecuencia durante la decodificación de SBR. Un valor del incremento de predicción con respecto a los coeficientes de predicción lineal en la dirección de la frecuencia aumenta a medida que el envoltorio temporal de una señal en el intervalo de análisis se hace más agudo. $K(r)$ es un parámetro para dar instrucciones a un decodificador de reforzar el proceso para agudizar la variación del envoltorio temporal de los componentes de alta frecuencia generados por SBR, con el incremento de su valor. $K(r)$ puede ser también un parámetro para dar instrucciones a un decodificador (como un dispositivo de decodificación de habla 21) para debilitar el proceso para agudizar la variación del envoltorio temporal de los componentes de alta frecuencia generados por SBR, con el descenso de su valor, o puede incluir un valor para no ejecutar el proceso para agudizar la variación del envoltorio temporal. En lugar de transmitir $K(r)$ a cada ranura de tiempo, se puede transmitir $K(r)$ que representa una pluralidad de ranuras de tiempo. Para determinar el segmento de las ranuras de tiempo en las cuales se comparte el mismo valor de $K(r)$, es preferible utilizar información en los límites de tiempo del envoltorio SBR (límite de tiempo de envoltorio SBR) incluidos en la información suplementaria de SBR.

$K(r)$ se transmite a la unidad de multiplicación de la corriente de bits 1g después de ser cuantificado. Es preferible calcular $K(r)$ que representa la pluralidad de ranuras de tiempo, por ejemplo, calculando un $K(r)$ medio de una pluralidad de ranuras de tiempo r antes de realizar la cuantificación. Para transmitir $K(r)$ que representa la pluralidad de ranuras de tiempo, $K(r)$ también puede obtenerse a partir del resultado del análisis de todo el segmento formado de la pluralidad de ranuras de tiempo, en lugar de calcular independientemente $K(r)$ a partir del resultado de analizar cada ranura de tiempo como la expresión (2). En este caso, $K(r)$ puede calcularse, por ejemplo, de acuerdo con la siguiente expresión (3). En que "mean" (·) indica un valor medio en el segmento de las ranuras de tiempo representadas por $K(r)$.

$$K(r) = \max(0, \min(1, \text{mean}(G_H(r))/\text{mean}(G_L(r)) - 1)) \quad \text{---(3)}$$

$K(r)$ puede ser transmitido de forma exclusiva con información de modo de filtro inverso incluida en la información suplementaria de SBR descrita en "ISO/IEC 14496-3 subpart 4 General Audio Coding". En otras palabras, $K(r)$ no se transmite para las ranuras de tiempo para las cuales se transmite la información de modo de filtro inverso en la información suplementaria de SBR, y la información de modo de filtro inverso (bs_invf_mode en "ISO/IEC 14496-3 subparte 4 Codificación General de Audio") en la información suplementaria de SBR no necesita ser transmitida para la ranura de tiempo para la cual se transmite $K(r)$. También puede añadirse información indicando que se transmite $K(r)$ o la información de modo de filtro inverso incluida en la información suplementaria de SBR. $K(r)$ y la información de modo de filtro inverso

incluida en la información suplementaria de SBR pueden combinarse para ser gestionadas como información de vector, y realizar una codificación de entropía en el vector. En este caso, se puede restringir la combinación de $K(r)$ y el valor de la información de modo de filtro inverso incluida en la información suplementaria de SBR.

- 5 La unidad de multiplicación de corriente de bits 1g multiplica la corriente de bits codificados calculada por la unidad central de codificación de códec 1c, la información suplementaria de SBR calculada por la unidad de codificación de SBR 1d, y $K(r)$ calculado por la unidad de cálculo de parámetros de intensidad de filtro 1f, y envía una corriente de bits multiplicados (corriente de bits multiplicados codificados) a través del dispositivo de comunicación del dispositivo de codificación de habla 11 (proceso en Fase Sa7).
- 10 La FIG. 3 es un diagrama que ilustra un dispositivo de decodificación de habla 21 de acuerdo con el primer ejemplo. El dispositivo de decodificación de habla 21 incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 21 cargando y ejecutando un programa de ordenador predeterminado (como un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 4) almacenado en una memoria integrada del dispositivo de decodificación de habla 21, como por ejemplo la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 21 recibe la salida de corriente de bits multiplicados codificados desde el dispositivo de codificación de habla 11, un dispositivo de codificación de habla 11a de una modificación 1, que será descrito más adelante, o un dispositivo de codificación de habla de una modificación 2, que será descrito más adelante, y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 21. El dispositivo de decodificación de habla 21, tal como se ilustra en la FIG. 3, incluye de forma funcional una unidad de separación de corriente de bits 2a (medio de separación de corriente de bits), una unidad central de decodificación de códecs 2b (medio central de decodificación), una unidad de transformación de frecuencia 2c (medio de transformación de frecuencia), una unidad de análisis de predicción lineal de baja frecuencia 2d (medio de análisis de envoltorio temporal de baja frecuencia), una unidad de detector de cambio de señal 2e, una unidad de ajuste de intensidad de filtro 2f (medio de ajuste de envoltorio temporal), una unidad de generación de alta frecuencia 2g, (medio de generación de alta frecuencia), una unidad de análisis de predicción lineal de alta frecuencia 2h, una unidad de filtro inverso de predicción lineal 2i, una unidad de ajuste de alta frecuencia 2j (medio de ajuste de alta frecuencia), una unidad de filtro de predicción lineal 2k (medio de formación del envoltorio temporal), una unidad de adición de coeficiente 2m, y una unidad de transformación de frecuencia inversa 2n. La unidad de separación de corriente de bits 2a hasta una unidad de cálculo de parámetros de forma de envoltorio 1n del dispositivo de decodificación de habla 21 ilustrado en la FIG. 3 son funciones que se realizan cuando la CPU del dispositivo de decodificación de habla 21 ejecuta el programa de ordenador almacenado en la memoria integrada del dispositivo de decodificación de habla 21. La CPU del dispositivo de decodificación de habla 21 ejecuta secuencialmente procesos (procesos desde Fase Sb1 a Fase Sb11) ilustrados en el diagrama de flujo de la FIG. 4, ejecutando el programa de ordenador (o utilizando la unidad de separación de corriente de bits 2a hasta la unidad a de transformación de frecuencia inversa 2n ilustrada en la FIG. 3). Diferentes tipos de datos necesarios para ejecutar el programa de ordenador y diferentes tipos de datos generados por el programa de ordenador se almacenan en la memoria integrada, como por ejemplo la ROM y la RAM del dispositivo de decodificación de habla 21.

La unidad de separación de corriente de bits 2a separa la corriente de bits multiplicados proporcionados a través del dispositivo de comunicación del dispositivo de decodificación de habla 21 en un parámetro de intensidad de filtro, información suplementaria de SBR y la corriente de bits codificados. La unidad central de decodificación de códecs 2b decodifica la corriente de bits codificados recibida de la unidad de separación de corriente de bits 2a para obtener una señal decodificada que incluye solamente los componentes de baja frecuencia (proceso en la Fase Sb1). En este momento, el método de decodificación puede basarse en el método de codificación de habla representado por el método CELP, o puede basarse en una codificación de audio como el método AAC o TCX (Transform Coded Excitation) Transformar Excitación Códificada.

La unidad de transformación de frecuencia 2c analiza la señal decodificada recibida desde la unidad central de decodificación de códecs 2b utilizando el banco de filtros QMF multi-división para obtener una señal $q_{dec}(k, r)$ en el dominio QMF (proceso en Fase Sb2). Debe hacerse notar que k ($0 \leq k \leq 63$) es un índice en la dirección de frecuencia, y r es un índice que indica un índice para la submuestra de la señal en el dominio QMF en la dirección temporal.

La unidad de análisis de predicción lineal de baja frecuencia 2d realiza un análisis de predicción lineal en la dirección de frecuencia en $q_{dec}(k, r)$ de cada ranura de tiempo r , obtenida a partir de la unidad de transformación de frecuencia 2c, para obtener coeficientes de predicción lineal de baja frecuencia $a_{dec}(n, r)$ (proceso en Fase Sb3). El análisis de predicción lineal se realiza para un espectro de $0 \leq k < k_x$ correspondiente a un ancho de banda de una señal obtenida desde la unidad central de decodificación de

códecs 2b. El análisis de predicción lineal puede ser realizado en una parte de la banda de frecuencia incluida en la sección de $0 \leq k < k_x$.

La unidad de detección de cambio de señal 2e detecta la variación temporal de la señal en el dominio QMF recibida desde la unidad de transformación de frecuencia 2c, y la emite como resultado de detección $T(r)$. El cambio de señal puede ser detectado, por ejemplo, utilizando el método que se describe más abajo.

1. La potencia a corto plazo $p(r)$ de una señal en la ranura de tiempo r se obtiene de acuerdo con la siguiente expresión (4).

$$p(r) = \sum_{k=0}^{63} |q_{dec}(k, r)|^2 \quad \text{---(4)}$$

2. Un envoltorio $p_{env}(r)$ obtenido nivelando $p(r)$ se obtiene de acuerdo con la siguiente expresión (5). Debe hacerse notar que α es una constante que cumple $0 < \alpha < 1$.

$$p_{env}(r) = \alpha \cdot p_{env}(r-1) + (1-\alpha) \cdot p(r) \quad \text{---(5)}$$

3. $T(r)$ se obtiene de acuerdo con la siguiente expresión (6) utilizando $p(r)$ y $p_{env}(r)$, en que β es una constante.

$$T(r) = \max(1, p(r)/(\beta \cdot p_{env}(r))) \quad \text{---(6)}$$

Los métodos que se han descrito anteriormente son simples ejemplos para detectar el cambio de señal basado en el cambio de potencia, y el cambio de señal puede detectarse utilizando otros métodos más sofisticados. Asimismo, puede omitirse la unidad de detección de cambio de señal 2e.

La unidad de ajuste de intensidad de filtro 2f ajusta la intensidad del filtro con respecto a $a_{dec}(n, r)$ obtenida a partir de la unidad de análisis de predicción lineal de baja frecuencia 2d para obtener coeficientes de predicción lineal ajustados $a_{adj}(n, r)$, (proceso en Fase Sb4). La intensidad de filtro se ajusta, por ejemplo, de acuerdo con la expresión (7) siguiente, utilizando un parámetro de intensidad de filtro K recibido a través de la unidad de separación de corriente de bits 2a.

$$a_{adj}(n, r) = a_{dec}(n, r) \cdot K(r)^n \quad (1 \leq n \leq N) \quad \text{---(7)}$$

Si se obtiene un resultado $T(r)$ de la unidad de detección de cambio de señal 2e, la intensidad puede ajustarse según la siguiente expresión (8).

$$a_{adj}(n, r) = a_{dec}(n, r) \cdot (K(r) \cdot T(r))^n \quad (1 \leq n \leq N) \quad \text{---(8)}$$

La unidad de generación de alta frecuencia 2g copia la señal en el dominio QMF obtenida de la unidad de transformación de frecuencia 2c desde la banda de baja frecuencia a la banda de alta frecuencia para generar una señal $q_{exp}(k, r)$ en el dominio QMF de los componentes de alta frecuencia (proceso en Fase Sb5). Los componentes de alta frecuencia se generan de acuerdo con el método de generación HF en SBR en "MPEG4 AAC" . ("ISO/IEC 14496-3 subparte 4 Codificación General de Audio").

La unidad de análisis de predicción lineal de alta frecuencia 2h realiza un análisis de predicción lineal en la dirección de frecuencia en $q_{exp}(k, r)$ de cada una de las ranuras de tiempo r generadas por la unidad de generación de alta frecuencia 2g para obtener coeficientes de predicción lineal de alta frecuencia $a_{exp}(n, r)$ (proceso en Fase Sb6). El análisis de predicción lineal se realiza para un espectro de $k_x \leq k \leq 63$ correspondiente a los componentes de alta frecuencia generados por la unidad de generación de alta frecuencia 2g.

La unidad de filtro inverso de predicción lineal 2i realiza un filtrado inverso de predicción lineal en la dirección de frecuencia en una señal en el dominio QMF de la banda de alta frecuencia generado por la unidad de generación de alta frecuencia 2g utilizando $a_{exp}(n, r)$ como coeficientes (proceso en Fase Sb7). La función de transferencia del filtro inverso de predicción lineal puede expresarse mediante la siguiente expresión (9).

$$f(z) = 1 + \sum_{n=1}^N a_{exp}(n, r) z^{-n} \quad \text{---(9)}$$

El filtrado inverso de predicción lineal puede realizarse de un coeficiente a una baja frecuencia a un coeficiente de alta frecuencia, o puede realizarse en sentido inverso. El filtrado inverso de predicción lineal es un proceso para alisar temporalmente el envoltorio temporal de los componentes de alta frecuencia, antes que se realice la formación del envoltorio temporal en la fase siguiente, y la unidad de filtro inverso de predicción lineal 2i puede omitirse. También resulta posible realizar un análisis de predicción lineal y filtrado inverso a partir de resultados de la unidad de ajuste de alta frecuencia 2j, que se describe más adelante, por parte de la unidad de análisis de predicción lineal de alta frecuencia 2h y la unidad de filtro inverso de predicción lineal 2i, en lugar de realizar un análisis de predicción lineal y filtrado inverso en los componentes e alta frecuencia de los resultados de la unidad de generación de alta frecuencia 2g. Los coeficientes de predicción lineal utilizados para el filtrado inverso de predicción lineal pueden ser también $a_{dec}(n, r)$ o $a_{adj}(n, r)$, en lugar de $a_{exp}(n, r)$. Los coeficientes de predicción lineal utilizados para el filtrado inverso de predicción lineal pueden ser también coeficientes de predicción lineal $a_{exp,adj}(n, r)$ obtenidos realizando un ajuste de intensidad de filtro en $a_{exp}(n, r)$. El ajuste de intensidad se realiza de acuerdo con la siguiente expresión (10), similar a la que se utiliza cuando se obtiene $a_{adj}(n, r)$.

$$a_{exp,adj}(n, r) = a_{exp}(n, r) \cdot K(r)^n \quad (1 \leq n \leq N) \quad \text{---(10)}$$

La unidad de ajuste de alta frecuencia 2j ajusta las características y la tonalidad de frecuencia de los componentes de alta frecuencia de un resultado de la unidad de filtro inverso de predicción lineal 2i (proceso en Fase Sb8). El ajuste se realiza de acuerdo con la información suplementaria de SBR recibida de la unidad de separación de corriente de bits 2a. El proceso por parte de la unidad de ajuste de alta frecuencia 2j se realiza de acuerdo con la fase de "ajuste HF" en SBR en "MPEG4 AAC", y se ajusta realizando un filtrado inverso de predicción lineal en la dirección temporal, el ajuste de incremento y la adición de ruido en la señal en el dominio QMF de la banda de alta frecuencia. Los detalles de los procesos en las fases descritas anteriormente se describen en "ISO/IEC 14496-3 subparte 4 Codificación General de Audio". Tal como se ha descrito anteriormente, la unidad de transformación de frecuencia 2c, la unidad de generación de alta frecuencia 2g, y la unidad de ajuste de frecuencia 2j operan todas ellas de acuerdo con el decodificador en "MPEG4 AAC" definido en "ISO/IEC 14496-3".

La unidad de filtro de predicción lineal 2k realiza un filtrado de síntesis de predicción lineal en la dirección de frecuencia en un componente de alta frecuencia $q_{adj}(n, r)$ de una señal en el resultado de dominio QMF de la unidad de ajuste de alta frecuencia 2j, utilizando $a_{adj}(n, r)$ obtenido de la unidad de ajuste de intensidad de filtro 2f (proceso en Fase Sb9). La función de transferencia del filtrado de síntesis de predicción lineal puede expresarse como la siguiente expresión (11).

$$g(z) = \frac{1}{1 + \sum_{n=1}^N a_{adj}(n, r) z^{-n}} \quad \text{---(11)}$$

Al realizar el filtrado de síntesis de predicción lineal, la unidad de filtro de predicción lineal 2k da forma al envoltorio temporal de los componentes de alta frecuencia generados basados en SBR.

La unidad de adición de coeficiente 2m añade una señal en el dominio QMF que incluye los componentes de baja frecuencia enviados desde la unidad de transformación de frecuencia 2c y una señal en el dominio QMF que incluye los componentes de alta frecuencia enviados desde la unidad de filtro de

predicción lineal 2k, y envía una señal en el dominio QMF que incluye tanto los componentes de baja frecuencia como los componentes de alta frecuencia (proceso en Fase Sb10).

5 La unidad de transformación inversa de frecuencia 2n procesa la señal en el dominio QMF obtenida de la unidad de adición de coeficiente 2m utilizando un banco de filtros de síntesis QMF. En concordancia, se obtiene una señal de habla decodificada del dominio de tiempo que incluye tanto los componentes de baja frecuencia obtenidos por el decodificador central de códecs como los componentes de alta frecuencia generados por SBR y cuyo envoltorio temporal está formado por el filtro de predicción lineal, y la señal de habla obtenida se envía a fuera del dispositivo de decodificación de habla 21 a través del dispositivo de comunicación integrado (proceso en Fase Sb11). Si K(r) y la información de modo de filtro inverso de la información suplementaria de SBR descrita en "ISO/IEC 14496-3 subparte 4 Codificación General de Audio" se transmiten exclusivamente, la unidad de transformación inversa de frecuencia 2n puede generar información de modo de filtro inverso de la información suplementaria de SBR para una ranura de tiempo para la cual se transmite K(r), pero la información de modo de filtro inverso de la información suplementaria de SBR no se transmite, utilizando información de modo de filtro inverso de la información suplementaria de SBR con respecto a por lo menos una ranura de tiempo de las ranuras de tiempo antes y después de la ranura de tiempo. También es posible establecer la información de modo de filtro inverso de la información suplementaria de SBR de la ranura de tiempo en un modo predeterminado con antelación. La unidad de transformación inversa de frecuencia 2n puede generar K(r) para una ranura de tiempo a la cual se transmiten los datos de filtrado de la información suplementaria de SBR, pero no se transmite K(r), utilizando K(r) para por lo menos una ranura de tiempo de las ranuras de tiempo antes y después de la ranura de tiempo. También es posible establecer K(r) de la ranura de tiempo en un valor predeterminado con antelación. La unidad de transformación inversa de frecuencia 2n puede también determinar si la información transmitida es K(r) o la información del modo de filtro inverso de la información suplementaria de SBR, basada en la información que indica si se transmite K(r) o la información de modo de filtro inverso de la información suplementaria de SBR.

(Modificación 1 del Primer Ejemplo)

30 La FIG. 5 es un diagrama que ilustra una modificación (dispositivo de codificación de habla 11a) del dispositivo de codificación de habla de acuerdo con el primer ejemplo. El dispositivo de codificación de habla 11a incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 11a cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla 11a como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 11a recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 11a, y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 11a.

40 El dispositivo de codificación de habla 11a, tal como se ilustra en la FIG. 5, incluye funcionalmente una unidad de transformación inversa de alta frecuencia 1h, una unidad de cálculo de potencia a corto plazo 1i (medio de cálculo de información suplementaria de envoltorio temporal), una unidad de cálculo de parámetro de intensidad de filtro 1f1 (medio de cálculo de información suplementaria de envoltorio temporal), y una unidad de multiplicación de corriente de bits 1g1 (medio de multiplicación de corriente de bits), en lugar de la unidad de análisis de predicción lineal 1e, la unidad de cálculo de parámetro de intensidad de filtro 1f, y la unidad de multiplicación de corriente de bits 1g del dispositivo de codificación de habla 11. La unidad de multiplicación de corriente de bits 1g1 tiene la misma función que 1g. La unidad de transformación de frecuencia 1a hasta la unidad de codificación de SBR 1d, la unidad de transformación inversa de alta frecuencia 1h, la unidad de cálculo de potencia a corto plazo 1i, la unidad de cálculo de parámetro de intensidad de filtro 1f1, y la unidad de multiplicación de corriente de bits 1g1 del dispositivo de codificación de habla 11a ilustrado en la FIG. 5 son funciones realizadas cuando la CPU del dispositivo de codificación de habla 11a ejecuta el programa de ordenador almacenado en la memoria integrada del dispositivo de codificación de habla 11a. Diferentes tipos de datos necesarios para ejecutar el programa de ordenador y diferentes tipos de datos generados al ejecutar el programa de ordenador se almacenan en la memoria integrada, como la ROM y la RAM del dispositivo de codificación de habla 11a.

55 La unidad de transformación inversa de alta frecuencia 1h sustituye los coeficientes de la señal en el dominio de QMF obtenidos de la unidad de transformación de frecuencia 1a con "0", que corresponde a los componentes de baja frecuencia codificados por la unidad central de codificación de códecs 1c, y procesa los coeficientes utilizando el banco de filtros de síntesis QMF para obtener una señal de dominio de tiempo que incluye solamente los componentes de alta frecuencia. La unidad de cálculo de potencia a corto plazo 1i divide los componentes de alta frecuencia en el dominio de tiempo obtenidos de la unidad de transformación inversa de alta frecuencia 1h en segmentos cortos, calcula la potencia y calcula p(r). Como método alternativo, la potencia a corto plazo puede calcularse también según la siguiente expresión (12) utilizando la señal en el dominio QMF.

$$p(r) = \sum_{k=0}^{63} |q(k, r)|^2 \quad \text{---(12)}$$

La unidad de cálculo de parámetros de intensidad de filtro 1f1 detecta la parte cambiada de $p(r)$, y determina un valor de $K(r)$, para que $K(r)$ sea incrementado con el cambio grande. El valor de $K(r)$, por ejemplo, puede calcularse también por el mismo método que para calcular $T(r)$ por la unidad de detección de cambio de señal 2e del dispositivo de decodificación de habla 21. El cambio de señal también puede detectarse utilizando otros métodos más sofisticados. La unidad de cálculo de parámetros de intensidad de filtro 1f1 puede también obtener potencia a corto plazo de cada uno de los componentes de baja frecuencia y los componentes de alta frecuencia, obtener cambios de señal $Tr(r)$ y $Th(r)$ de cada uno de los componentes de baja frecuencia y de los componentes de alta frecuencia utilizando el mismo método que para calcular $T(r)$ por la unidad de detección de cambio de señal 2e del dispositivo de decodificación de habla 21 y determinar el valor de $K(r)$ utilizándolos. En este caso, por ejemplo, $K(r)$ puede obtenerse de acuerdo con la siguiente expresión (13), en que ε es una constante como por ejemplo 3.0.

$$K(r) = \max(0, \varepsilon \cdot (Th(r) - Tr(r))) \quad \text{---(13)}$$

15 (Modificación 2 del Primer ejemplo)

Un dispositivo de codificación de habla (que no se ilustra) de una modificación 2 del primer ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla de la modificación 2 cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla de la modificación 2 como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla de la modificación 2 recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla, y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla.

El dispositivo de codificación de habla de la modificación 2 incluye funcionalmente una unidad de codificación diferencial de coeficiente de predicción lineal (medio de cálculo de información suplementaria de envoltorio temporal), y una unidad de multiplicación de corriente de bits (medio de multiplicación de corriente de bits) que recibe una información de la unidad de codificación diferencial del coeficiente de predicción lineal, que no está ilustrado, en lugar de la unidad de cálculo de parámetros de intensidad de filtro 1f y la unidad de multiplicación de corriente de bits 1g del dispositivo de codificación de habla 11. La unidad de transformación de frecuencia 1a hasta la unidad de análisis de predicción lineal 1e, la unidad de codificación diferencial de coeficiente de predicción lineal, y la unidad de multiplicación de corriente de bits del dispositivo de codificación de habla de la modificación 2 son funciones realizadas cuando la CPU del dispositivo de codificación de habla de la modificación 2 ejecuta el programa de ordenador almacenado en la memoria integrada del dispositivo de codificación de habla de la modificación 2. Diferentes tipos de datos necesarios para ejecutar el programa de ordenador y diferentes tipos de datos generados al ejecutar el programa de ordenador son almacenados en la memoria integrada, como la ROM y la RAM del dispositivo de codificación de habla de la modificación 2.

La unidad de decodificación diferencial del coeficiente de predicción lineal calcula valores de diferencial $a_D(n, r)$ del coeficiente de predicción lineal de acuerdo con la expresión (14) siguiente, utilizando $a_H(n, r)$ de la señal de entrada y $a_L(n, r)$ de la señal de entrada.

$$a_D(n, r) = a_H(n, r) - a_L(n, r) \quad (1 \leq n \leq N) \quad \text{---(14)}$$

A continuación, la unidad de codificación del diferencial de coeficiente de predicción lineal cuantifica $a_D(n, r)$, y los transmite a la unidad de multiplicación de corriente de bits (estructura correspondiente a la unidad de multiplicación de corriente de bits 1g). La unidad de multiplicación de corriente de bits multiplica $a_D(n, r)$ en la corriente de bits en lugar de $K(r)$, y envía la corriente de bits multiplicados a fuera del dispositivo de codificación de habla a través del dispositivo de comunicación integrado.

Un dispositivo de decodificación de habla (que no se ilustra) de la modificación 2 del primer ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla de la modificación 2 cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de decodificación de habla de la modificación 2 como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla de la modificación 2 recibe la salida de la corriente de bits multiplicados codificados del dispositivo de codificación de habla 11, el dispositivo de codificación de habla 11a de acuerdo con la modificación 1, o el dispositivo de codificación de habla de acuerdo con la modificación 2, y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla.

El dispositivo de decodificación de habla de la modificación 2 incluye funcionalmente una unidad de decodificación de diferencial de coeficiente de predicción lineal, que no está ilustrada, en lugar de la unidad de ajuste de intensidad de filtro 2f del dispositivo de decodificación de habla 21. La unidad de separación de corriente de bits 2a hasta la unidad de detección de cambio de señal 2e, la unidad de decodificación de diferencial del coeficiente de predicción lineal y la unidad de generación de alta frecuencia 2g hasta la unidad de transformación inversa de frecuencia 2n del dispositivo de decodificación de habla de la modificación 2 son funciones realizadas cuando la CPU del dispositivo de decodificación de habla de la modificación 2 ejecuta el programa de ordenador almacenado en la memoria integrada del dispositivo de decodificación de habla de la modificación 2. Diferentes tipos de datos necesarios para ejecutar el programa de ordenador y diferentes tipos de datos generados al ejecutar el programa del ordenador están almacenados en la memoria integrada como la ROM y la RAM del dispositivo de decodificación de habla de la modificación 2.

La unidad de decodificación de diferencial del coeficiente de predicción lineal obtiene $a_{adj}(n, r)$ decodificado de forma diferencial de acuerdo con la siguiente expresión (15), utilizando $a_L(n, r)$ obtenido de la unidad de análisis de predicción lineal de baja frecuencia 2d y $a_D(n, r)$ recibido de la unidad de separación de corriente de bits 2a.

$$a_{adj}(n, r) = a_{dec}(n, r) + a_D(n, r), \quad 1 \leq n \leq N \quad \text{---(15)}$$

La unidad de decodificación de diferencial del coeficiente de predicción lineal transmite $a_{adj}(n, r)$ decodificado diferencialmente de esta forma a la unidad de filtro de predicción lineal 2k. $a_D(n, r)$ puede ser un valor diferencial en el dominio de coeficientes de predicción tal como se ilustra en la expresión (14). Pero, después de convertir los coeficientes de predicción a la otra forma de expresión, como coeficientes LSP (Linear Spectrum Pair) Par de Espectro Lineal, ISP (Immittance Spectrum Pair) Par de Espectro de Impedancia, LSF (Linear Spectrum Frequency) Frecuencia de Espectro Lineal, ISF (Immittance Spectrum Frequency) Frecuencia de Espectro de Impedancia, y PARCOR, $a_D(n, r)$ puede ser un valor que muestra una diferencia. En este caso, la decodificación de diferencial también tiene la misma forma de expresión.

(Segundo Ejemplo)

La FIG. 6 es un diagrama que ilustra un dispositivo de codificación de habla 12 de acuerdo con un segundo ejemplo. El dispositivo de codificación de habla 12 incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 12 cargando y ejecutando un programa de ordenador predeterminado (como por ejemplo un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 7) almacenado en una memoria integrada del dispositivo de codificación de habla 12 como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 12 recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 12, y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 12.

El dispositivo de codificación de habla 12 incluye funcionalmente una unidad de decimación del coeficiente de predicción lineal 1j (medio de decimación del coeficiente de predicción), una unidad de cuantificación del coeficiente de predicción lineal 1k (medio de cuantificación del coeficiente de predicción), y una unidad de multiplicación de corriente de bits 1g2 (medio de multiplicación de corriente de bits), en lugar de la unidad de cálculo de parámetros de intensidad de filtro 1f y la unidad de multiplicación de corriente de bits 1g del dispositivo de codificación de habla 11. La unidad de transformación de frecuencia 1a hasta la unidad de análisis de predicción lineal 1e (medio de análisis de predicción lineal), la unidad de decimación del coeficiente de predicción lineal 1j, la unidad de cuantificación de coeficiente de predicción lineal 1k y la unidad de multiplicación de corriente de bits 1g2 del dispositivo de codificación de habla 12 ilustrados en la FIG. 6 son funciones realizadas cuando la CPU del dispositivo de codificación de habla 12 ejecuta el programa de ordenador almacenado en la memoria integrada en el dispositivo de codificación de habla 12. La CPU del dispositivo de codificación de habla 12

ejecuta procesos secuencialmente (procesos desde la Fase Sa1 a Fase Sa5, y procesos desde la Fase Sc1 a Fase Sc3) ilustrados en el diagrama de flujo de la FIG. 7, ejecutando el programa de ordenador (o utilizando la unidad de transformación de frecuencia 1a hasta la unidad de análisis de predicción lineal 1e, la unidad de decimación del coeficiente de predicción lineal 1j, la unidad de cuantificación del coeficiente de predicción lineal 1k, y la unidad de multiplicación de corriente de bits 1g2 del dispositivo de codificación de habla 12 ilustrado en la FIG. 6). Diferentes tipos de datos necesarios para ejecutar el programa de ordenador y diferentes tipos de datos generados al ejecutar el programa del ordenador se almacenan en la memoria integrada, como la ROM y la RAM del dispositivo de codificación de habla 12.

La unidad de decimación del coeficiente de predicción lineal 1j decima $a_H(n, r)$ obtenido de la unidad de análisis de predicción lineal 1e en la dirección temporal, y transmite un valor de $a_H(n, r)$ para una parte de la ranura de tiempo r_i y un valor del r_i correspondiente, a la unidad de cuantificación del coeficiente de predicción lineal 1k (proceso en Fase Sc1). Se observa que $0 \leq i < N_{is}$, y N_{is} es el número de ranuras de tiempo en una secuencia para la cual se transmite $a_H(n, r)$. La decimación de los coeficientes de predicción lineal puede ser realizada en un intervalo de tiempo predeterminado, o puede realizarse en un intervalo de tiempo no uniforme basado en las características de $a_H(n, r)$. Por ejemplo, es posible un método que compare $G_H(r)$ de $a_H(n, r)$ en una secuencia con una cierta longitud, y haga que $a_H(n, r)$, del cual $G_H(r)$ excede un cierto valor, sea un objeto de cuantificación. Si el intervalo de decimación de los coeficientes de predicción lineal es un intervalo predeterminado en lugar de utilizar las características de $a_H(n, r)$, $a_H(n, r)$ no tiene que ser calculado para la ranura de tiempo en la cual no se realiza la transmisión.

La unidad de cuantificación del coeficiente de predicción lineal 1k cuantifica los coeficientes de predicción lineal de alta frecuencia decimados $a_H(n, r_i)$ recibidos de la unidad de decimación del coeficiente de predicción lineal 1j e indexa r_i de las ranuras de tiempo correspondientes, y las transmite a la unidad de multiplicación de la corriente de bits 1g2 (proceso en Fase Sc2). Como estructura alternativa, en lugar de cuantificar $a_H(n, r_i)$, pueden cuantificarse valores de diferencial $a_D(n, r_i)$ de los coeficientes de predicción lineal como el dispositivo de codificación de habla de acuerdo con la modificación 2 del primer ejemplo.

La unidad de multiplicación de corriente de bits 1g2 multiplica la corriente de bits codificados calculada por la unidad central de codificación de códecs 1c, la información suplementaria de SBR calculada por la unidad de codificación de SBR 1d, e indexa $\{r_i\}$ de las ranuras de tiempo correspondientes a $a_H(n, r_i)$ que se cuantifican y reciben desde la unidad de cuantificación del coeficiente de predicción lineal 1k en una corriente de bits, y envía la corriente de bits multiplicados a través del dispositivo de comunicación del dispositivo de codificación de habla 12 (proceso en Fase Sc3).

La FIG. 8 es un diagrama que ilustra un dispositivo de decodificación de habla 22 de acuerdo con el segundo ejemplo. El dispositivo de decodificación de habla 22 incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 22 cargando y ejecutando un programa de ordenador predeterminado (como por ejemplo un programa de ordenador para realizar procesos ilustrado en el diagrama de flujo de la FIG. 9) almacenado en una memoria integrada del dispositivo de decodificación de habla 22 como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 22 recibe la corriente de bits multiplicados codificados emitida por el dispositivo de codificación de habla 12, y envía una señal de habla decodificada hacia fuera del dispositivo de codificación de habla 12.

El dispositivo de decodificación de habla 22 incluye funcionalmente una unidad de separación de corriente de bits 2a1 (medio de separación de corriente de bits), una unidad de interpolación/extrapolación del coeficiente de predicción lineal 2p (medio de interpolación/extrapolación del coeficiente de predicción lineal) y una unidad de filtro de predicción lineal 2k1 (medio de formación del envoltorio temporal) en lugar de la unidad de separación de la corriente de bits 2a, la unidad de análisis de predicción lineal de baja frecuencia 2d, la unidad de detección de cambio de señal 2e, la unidad de ajuste de intensidad de filtro 2f, y la unidad de filtro de predicción lineal 2k del dispositivo de decodificación de habla 21. La unidad de separación de corriente de bits 2a1, la unidad central de decodificación de códecs 2b, la unidad de transformación de frecuencia 2c, la unidad de generación de alta frecuencia 2g hasta la unidad de ajuste de alta frecuencia 2j, la unidad de filtro de predicción lineal 2k1, la unidad de adición de coeficiente 2m, la unidad de transformación inversa de frecuencia 2n, y la unidad de interpolación/extrapolación del coeficiente de predicción lineal 2p del dispositivo de decodificación de habla 22 ilustrados en la FIG. 8 son funciones realizadas cuando la CPU del dispositivo de codificación de habla 12 ejecuta el programa de ordenador almacenado en la memoria integrada del dispositivo de codificación de habla 12. La CPU del dispositivo de decodificación de habla 22 ejecuta secuencialmente los procesos (procesos desde Fase Sb1 a Fase Sd2, Fase Sd1, desde Fase Sb5 a Fase Sb8, Fase Sd2, y desde Fase Sb10 a Fase Sb11) ilustrados en el diagrama de flujo de la FIG. 9, ejecutando el programa de ordenador (o utilizando la unidad de separación de corriente de bits 2a1, la unidad central de decodificación de códecs 2b, la unidad de transformación de frecuencia 2c, la unidad de generación de alta frecuencia 2g hasta la unidad de ajuste de alta frecuencia 2j, la unidad de filtro de predicción lineal 2k1, la unidad de adición de coeficiente

2m, la unidad de transformación inversa de frecuencia 2n, y la unidad de interpolación/extrapolación del coeficiente de predicción lineal 2p ilustradas en la FIG. 8). Diferentes tipos de datos necesarios para ejecutar el programa de ordenador y diferentes tipos de datos generados al ejecutar el programa del ordenador se almacenan en la memoria integrada, como la ROM y la RAM del dispositivo de decodificación de habla 22.

El dispositivo de decodificación de habla 22 incluye la unidad de separación de corriente de bits 2a1, la unidad de interpolación/extrapolación del coeficiente de predicción lineal 2p y la unidad de filtro de predicción lineal 2k1, en lugar de la unidad de separación de corriente de bits 2a, la unidad de análisis de predicción lineal de baja frecuencia 2d, la unidad de detección de cambio de señal 2e, la unidad de ajuste de intensidad de filtro 2f y la unidad de filtro de predicción lineal 2k del dispositivo de decodificación de habla 22.

La unidad de separación de corriente de bits 2a1 separa la corriente de bits multiplicados suministrada a través del dispositivo de comunicación del dispositivo de decodificación de habla 22 en los índices r_i de las ranuras de tiempo correspondientes a $a_H(n, r_i)$ que se cuantifican, la información suplementaria de SBR y la corriente de bits codificados.

La unidad de interpolación/extrapolación del coeficiente de predicción lineal 2p recibe los índices r_i de las ranuras de tiempo correspondientes a $a_H(n, r_i)$ que se cuantifican desde la unidad de separación de corriente de bits 2a1, y obtiene $a_H(n, r)$ correspondiente a las ranuras de tiempo de las cuales no se transmiten los coeficientes de predicción lineal, por interpolación o extrapolación (procesos en Fase Sd1). La unidad de interpolación/extrapolación del coeficiente de predicción lineal 2p puede extrapolar los coeficientes de predicción lineal, por ejemplo de acuerdo con la siguiente expresión (16).

$$a_H(n, r) = \delta^{|r-r_{i0}|} a_H(n, r_{i0}) \quad (1 \leq n \leq N) \quad \text{---(16)}$$

en que r_{i0} es el valor más próximo a r en las ranuras de tiempo $\{r_i\}$ de las cuales se transmiten los coeficientes de predicción lineal. δ es una constante que cumple $0 < \delta < 1$.

La unidad de interpolación/extrapolación del coeficiente de predicción lineal 2p puede interpolar los coeficientes de predicción lineal, por ejemplo de acuerdo con la siguiente expresión (17), en que se cumple $r_{i0} < r < r_{i0+1}$.

$$a_H(n, r) = \frac{r_{i0+1} - r}{r_{i0+1} - r_i} \cdot a_H(n, r_i) + \frac{r - r_{i0}}{r_{i0+1} - r_{i0}} \cdot a_H(n, r_{i0+1}) \quad (1 \leq n \leq N) \quad \text{---(17)}$$

La unidad de interpolación/extrapolación del coeficiente de predicción lineal 2p puede convertir los coeficientes de predicción lineal en otras formas de expresión, como coeficientes LSP (Linear Spectrum Pair) Par de Espectro Lineal, ISP (Immittance Spectrum Pair) Par de Espectro de Impedancia, LSF (Linear Spectrum Frequency) Frecuencia de Espectro Lineal, ISF (Immittance Spectrum Frequency) Frecuencia de Espectro de Impedancia, y PARCOR, interpolarlos o extrapolarlos, y convertir los valores obtenidos en coeficientes de predicción lineal para ser utilizados. Los $a_H(n, r)$ interpolados o extrapolados se transmiten a la unidad de filtro de predicción lineal 2k1 y se utilizan como coeficientes de predicción lineal para el filtrado de síntesis de predicción lineal, pero también pueden utilizarse como coeficientes e predicción lineal en la unidad de filtro inverso de predicción lineal 21. Si $a_D(n, r_i)$ es multiplicado en una corriente de bits en lugar de $a_H(n, r)$, la unidad de interpolación/extrapolación del coeficiente de predicción lineal 2p realiza la decodificación del diferencial de manera parecida a la del dispositivo de decodificación de habla según la modificación 2 del primer ejemplo, antes de realizar el proceso de interpolación o extrapolación descrito anteriormente.

La unidad de filtro de predicción lineal 2k1 realiza un filtrado de síntesis de predicción lineal en la dirección de frecuencia en la salida de $q_{adj}(n, r)$ desde la unidad de ajuste de alta frecuencia 2j, utilizando $a_H(n, r)$ que se interpola o extrapola obtenido de la unidad de interpolación/extrapolación de coeficiente de predicción lineal 2p (proceso en Fase Sd2). Una función de transferencia de la unidad de filtro de predicción lineal 2k1 puede expresarse como la siguiente expresión (18). La unidad de filtro de predicción lineal 2k1 da forma al envoltorio temporal de los componentes de alta frecuencia generados por el SBR

realizando un filtrado de síntesis de predicción lineal, como la unidad de filtro de predicción lineal 2k del dispositivo de decodificación de habla 21.

$$g(z) = \frac{1}{1 + \sum_{n=1}^N a_H(n, r) z^{-n}} \quad \text{---(18)}$$

5 (Tercer Ejemplo)

La FIG. 10 es un diagrama que ilustra un dispositivo de codificación de habla 13 de acuerdo con un tercer ejemplo. El dispositivo de codificación de habla 13 incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 13 cargando y ejecutando un programa de ordenador predeterminado (como por ejemplo un programa de ordenador para realizar procesos, ilustrados en el diagrama de flujo de la FIG. 11) almacenado en una memoria integrada del dispositivo de codificación de habla 13 como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 13 recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 13, y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 13.

15 El dispositivo de codificación de habla 13 incluye funcionalmente una unidad de cálculo de envoltorio temporal 1m (medio de cálculo de información suplementaria de envoltorio temporal), una unidad de cálculo de parámetros de forma de envoltorio 1n (medio de cálculo de información suplementaria de envoltorio temporal), y una unidad de multiplicación de corriente de bits 1g3 (medio de multiplicación de corriente de bits) en lugar de la unidad de análisis de predicción lineal 1e, la unidad de cálculo de parámetros de intensidad de filtro 1f, y la unidad de multiplicación de corriente de bits 1g del dispositivo de codificación de habla 11. La unidad de transformación de frecuencia 1a hasta la unidad de codificación de SBR 1d, la unidad de cálculo del envoltorio temporal 1m, la unidad de cálculo de parámetros de forma de envoltorio 1n y la unidad de multiplicación de corriente de bits 1g3 del dispositivo de codificación de habla 13 ilustrado en la FIG. 10 son funciones realizadas cuando la CPU del dispositivo de codificación de habla 12 ejecuta el programa de ordenador almacenado en la memoria integrada del dispositivo de codificación de habla 12. La CPU del dispositivo de codificación de habla 13 ejecuta procesos secuencialmente (procesos desde la Fase Sa1 a Fase Sa4, y procesos desde la Fase Se1 a Fase Se3) ilustrados en el diagrama de flujo de la FIG. 11, ejecutando el programa de ordenador (o utilizando la unidad de transformación de frecuencia 1a hasta la unidad de codificación de SBR 1d, la unidad de cálculo de envoltorio temporal 1m, la unidad de cálculo de parámetros de forma de envoltorio 1n y la unidad de multiplicación de corriente de bits 1g3 del dispositivo de codificación de habla 13 ilustrado en la FIG. 10). Diferentes tipos de datos requeridos para ejecutar el programa de ordenador y diferentes tipos de datos generados al ejecutar el programa de ordenador se almacenan en la memoria integrada, como la ROM y la RAM del dispositivo de codificación de habla 13.

35 La unidad de cálculo de envoltorio temporal 1m recibe q(k, r), y por ejemplo, obtiene información del envoltorio temporal e(r) de los componentes de alta frecuencia de una señal, obteniendo la potencia de cada ranura de tiempo de q(k, r) (proceso en Fase Se1). En este caso, e(r) se obtiene de acuerdo con la siguiente expresión (19).

$$e(r) = \sqrt{\sum_{k=k_x}^{63} |q(k, r)|^2} \quad \text{---(19)}$$

40 La unidad de cálculo de parámetros de forma de envoltorio 1n recibe e(r) de la unidad de cálculo del envoltorio temporal 1m y recibe límites de tiempo de envoltorio de SBR {bi} de la unidad de codificación de SBR 1d. Debe observarse que 0 ≤ i ≤ Ne, y Ne es el número de envoltorios de SBR en la secuencia codificada. La unidad de cálculo de parámetros de forma de envoltorio 1n obtiene un parámetro de forma de envoltorio s(i) (0 ≤ i < Ne) de cada uno de los envoltorios de SBR en la secuencia codificada de acuerdo

con la siguiente expresión (20) (proceso en Fase Se2). El parámetro de forma de envoltorio $s(i)$ corresponde a la información suplementaria de envoltorio temporal, y es similar en el tercer ejemplo.

$$s(i) = \frac{1}{b_{i+1} - b_i - 1} \sum_{r=b_i}^{b_{i+1}-1} (\overline{e(i)} - e(r))^2 \quad \text{---(20)}$$

5 Debe observarse que:

$$\overline{e(i)} = \frac{\sum_{r=b_i}^{b_{i+1}-1} e(r)}{b_{i+1} - b_i} \quad \text{---(21)}$$

10 en que $s(i)$ en la expresión anterior es un parámetro que indica la magnitud de la variación de $e(r)$ en el envoltorio de i -ésimo SBR donde se cumple $b_i \leq r < b_{i+1}$, y $e(r)$ tiene un número mayor a medida que se incrementa la variación del envoltorio temporal. Las expresiones (20) y (21) descritas anteriormente son ejemplos del método para calcular $s(i)$, y por ejemplo, $s(i)$ puede obtenerse también utilizando, por ejemplo, SMF (Spectral Flatness Measure) Medida de Plano Espectral, de $e(r)$, un coeficiente del valor máximo por el valor mínimo, y similares. A continuación, se cuantifica $s(i)$, y se transmite a la unidad de multiplicación de corriente de bits 1g3.

15 La unidad de multiplicación de corriente de bits 1g3 multiplica la corriente de bits codificados calculada por la unidad central de codificación 1c, la información suplementaria de SBR calculada por la unidad de codificación de SBR 1d, y $s(i)$ en una corriente de bits, y envía la corriente de bits multiplicados a través del dispositivo de comunicación del dispositivo de codificación de habla 13 (proceso en Fase Se3).

20 La FIG. 12 es un diagrama que ilustra un dispositivo de decodificación de habla 23 de acuerdo con el tercer ejemplo. El dispositivo de decodificación de habla 23 incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 23 cargando y ejecutando un programa de ordenador predeterminado (como por ejemplo un programa de ordenador para realizar procesos, ilustrados en el diagrama de flujo de la FIG. 13) almacenado en una memoria integrada del dispositivo de decodificación de habla 23 como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 23 recibe la corriente de bits multiplicados codificados enviados desde el dispositivo de codificación de habla 13, y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 13.

30 El dispositivo de decodificación de habla 23 incluye funcionalmente una unidad de separación de corriente de bits 2a2 (medio de separación de corriente de bits), una unidad de cálculo de envoltorio temporal de baja frecuencia 2r (medio de análisis de envoltorio temporal de baja frecuencia), una unidad de ajuste de forma de envoltorio 2s (medio de ajuste de envoltorio temporal), una unidad de cálculo de envoltorio temporal de alta frecuencia 2t, una unidad de alisamiento de envoltorio temporal 2u, y una unidad de formación del envoltorio temporal 2v (medio de formación del envoltorio temporal), en lugar de la unidad de separación de corriente de bits 2a, la unidad de análisis de predicción lineal de baja frecuencia 2d, la unidad de detección de cambio de señal 2e, la unidad de ajuste de intensidad de filtro 2f, la unidad de análisis de predicción lineal de alta frecuencia 2h, la unidad de filtro inverso de predicción lineal 2i, y la unidad de filtro de predicción lineal 2k del dispositivo de decodificación de habla 21. La unidad de separación de corriente de bits 2a2, la unidad central de decodificación de códecs 2b hasta la unidad de transformación de frecuencia 2c, la unidad de generación de alta frecuencia 2g, la unidad de ajuste de alta frecuencia 2j, la unidad de adición de coeficiente 2m, la unidad de transformación inversa de frecuencia 2n y la unidad de cálculo de envoltorio temporal de baja frecuencia 2r hasta la unidad de formación del envoltorio temporal 2v del dispositivo de decodificación de habla 23 ilustrado en la FIG. 12 son funciones realizadas cuando la CPU del dispositivo de codificación de habla 12 ejecuta el programa de ordenador almacenado en la memoria integrada del dispositivo de codificación de habla 12. La CPU del dispositivo de decodificación de habla 23 ejecuta procesos secuencialmente (procesos desde la Fase Sb1 a Fase Sb2, de Fase Sf1 a Fase Sf2, Fase Sb5, de Fase Sf3 a Fase Sf4, Fase Sb8, Fase Sf5 y

desde Fase Sb10 a Fase Sb11) ilustrados en el diagrama de flujo de la FIG. 13, ejecutando el programa de ordenador (o utilizando la unidad de separación de corriente de bits 2a2, la unidad central de decodificación de códecs 2b hasta la unidad de transformación de frecuencia 2c, la unidad de generación de alta frecuencia 2g, la unidad de ajuste de alta frecuencia 2j, la unidad de adición de coeficiente 2m, la
 5 unidad de transformación inversa de frecuencia 2n, y la unidad de cálculo de envoltorio temporal de baja frecuencia 2r hasta la unidad de formación del envoltorio temporal 2v del dispositivo de decodificación de habla 23 ilustrado en la FIG. 12). Diferentes tipos de datos necesarios para ejecutar el programa de ordenador y diferentes tipos de datos generados al ejecutar el programa del ordenador se almacenan en la memoria integrada, como la ROM y la RAM del dispositivo de codificación de habla 23.

10 La unidad de separación de corriente de bits 2a2 separa la corriente de bits multiplicados suministrada a través del dispositivo de comunicación del dispositivo de decodificación de habla 23 en $s(i)$, la información suplementaria de SBR y la corriente de bits codificados. La unidad de cálculo de envoltorio temporal de baja frecuencia 2r recibe $q_{dec}(k, r)$ incluyendo los componentes de baja frecuencia de la unidad de transformación de frecuencia 2c, y obtiene $e(r)$ de acuerdo con la siguiente expresión (22) (proceso en Fase Sf1).
 15

$$e(r) = \sqrt{\sum_{k=0}^{63} |q_{dec}(k, r)|^2} \quad \text{---(22)}$$

La unidad de ajuste de forma de envoltorio 2s ajusta $e(r)$ utilizando $s(i)$, y obtiene la información del envoltorio temporal ajustada $e_{adj}(r)$ (proceso en Fase Sf2). $e(r)$ puede ser ajustada, por ejemplo, de acuerdo con las siguientes expresiones (23) a (25).

$$e_{adj}(r) = \overline{e(i)} + \sqrt{s(i) - v(i)} \cdot (e(r) - \overline{e(i)}) \quad (s(i) > v(i)) \quad \text{---(23)}$$

20 $e_{adj}(r) = e(r)$ de lo contrario

Debe hacerse notar que:

$$\overline{e(i)} = \frac{\sum_{r=b_i}^{b_{i+1}-1} e(r)}{b_{i+1} - b_i} \quad \text{---(24)}$$

$$v(i) = \frac{1}{b_{i+1} - b_i - 1} \sum_{r=b_i}^{b_{i+1}-1} (\overline{e(i)} - e(r))^2 \quad \text{---(25)}$$

25 Las expresiones (23) a (25) descritas más arriba son ejemplos de método de ajuste, y también puede utilizarse el otro método de ajuste por el cual la forma de $e_{adj}(r)$ se convierte en similar a la forma ilustrada por $s(i)$.

La unidad de cálculo de envoltorio temporal de alta frecuencia 2t calcula un envoltorio temporal $e_{exp}(r)$ utilizando $q_{exp}(k, r)$ obtenido de la unidad de generación de alta frecuencia 2g, de acuerdo con la siguiente expresión (26) (proceso en Fase Sf3).

$$e_{exp}(r) = \sqrt{\sum_{k=k_x}^{63} |q_{exp}(k, r)|^2} \quad \text{---(26)}$$

- 5 La unidad de alisado de envoltorio temporal 2u alisa el envoltorio temporal de $q_{exp}(k, r)$ obtenido de la unidad de generación de alta frecuencia 2g de acuerdo con la siguiente expresión (27), y transmite la señal obtenida $q_{flat}(k, r)$ en el dominio QMF a la unidad de ajuste de alta frecuencia 2j (proceso en Fase Sf4).

$$q_{flat}(k, r) = \frac{q_{exp}(k, r)}{e_{exp}(r)} \quad (k_x \leq k \leq 63) \quad \text{---(27)}$$

- 10 El alisado del envoltorio temporal por parte de la unidad de alisado de envoltorio temporal 2u también puede omitirse. En lugar de calcular el envoltorio temporal de los componentes de alta frecuencia del resultado de la unidad de generación de alta frecuencia 2g y alisar el envoltorio temporal de la misma, puede calcularse el envoltorio temporal de los componentes de alta frecuencia de un resultado de la unidad de ajuste de alta frecuencia 2j, y el envoltorio temporal del mismo puede ser alisado. El envoltorio temporal utilizado en la unidad de alisado de envoltorio temporal 2u puede ser también $e_{adj}(r)$ obtenido de la unidad de ajuste de forma de envoltorio 2s en lugar de $e_{exp}(r)$ obtenido de la unidad de cálculo de envoltorio temporal de alta frecuencia 2t.

- 20 La unidad de formación del envoltorio temporal 2v da forma a $q_{adj}(k, r)$ obtenido de la unidad de ajuste de alta frecuencia 2j utilizando $e_{adj}(r)$ obtenido de la unidad de formación del envoltorio temporal 2v, y obtiene una señal $q_{envadj}(k, r)$ en el dominio QMF en el cual se forma el envoltorio temporal (proceso en Fase Sf5). La formación se realiza de acuerdo con la siguiente expresión (28). $q_{envadj}(k, r)$ se transmite a la unidad de adición de coeficiente 2m como una señal en el dominio QMF correspondiente a los componentes de alta frecuencia.

$$q_{envadj}(k, r) = q_{adj}(k, r) \cdot e_{adj}(r) \quad (k_x \leq k \leq 63) \quad \text{---(28)}$$

- 25 (Realización)

- 30 La FIG. 14 es un diagrama que ilustra un dispositivo de decodificación de habla 24 de acuerdo con una realización de la presente invención. El dispositivo de decodificación de habla 24 incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24 cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de decodificación de habla 24 como por ejemplo la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24 recibe la salida de corriente de bits multiplicados codificados del dispositivo de codificación de habla 11 o del dispositivo de codificación de habla 13, y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24.

- 40 El dispositivo de decodificación de habla 23 incluye funcionalmente la estructura del dispositivo de decodificación de habla 21 (la unidad central de decodificación de códecs 2b, la unidad de transformación de frecuencia 2c, la unidad de análisis de predicción lineal de baja frecuencia 2d, la unidad de detección de cambio de señal 2e, la unidad de ajuste de intensidad de filtro 2f, la unidad de generación de alta frecuencia 2g, la unidad de análisis de predicción lineal de alta frecuencia 2h, la unidad de filtro inverso de predicción lineal 2i, la unidad de ajuste de alta frecuencia 2j, la unidad de filtro de predicción lineal 2k, la unidad de adición de coeficiente 2m y la unidad de transformación inversa de frecuencia 2n) y la estructura del dispositivo de decodificación de habla 24 (la unidad de cálculo de envoltorio temporal de

baja frecuencia $2r$, la unidad de ajuste de forma de envoltorio $2s$, y la unidad de formación del envoltorio temporal $2v$). El dispositivo de decodificación de habla 24 también incluye una unidad de separación de corriente de bits $2a3$ (medio de separación de corriente de bits) y una unidad de conversión de información suplementaria $2w$. El orden de la unidad de filtro de predicción lineal $2k$ y la unidad de formación del envoltorio temporal $2v$ puede ser opuesto al ilustrado en la FIG. 14. El dispositivo de decodificación de habla 24 recibe preferiblemente la corriente de bits codificados por el dispositivo de codificación de habla 11 o el dispositivo de codificación de habla 13 . La estructura del dispositivo de decodificación de habla 24 ilustrado en la FIG. 14 es una función realizada cuando la CPU del dispositivo de decodificación de habla 24 ejecuta el programa de ordenador almacenado en la memoria integrada del dispositivo de decodificación de habla 24 . Diferentes tipos de datos necesarios para ejecutar el programa de ordenador y diferentes tipos de datos generados al ejecutar el programa del ordenador se almacenan en la memoria integrada, como la ROM y la RAM del dispositivo de decodificación de habla 24 .

La unidad de separación de corriente de bits $2a3$ separa la corriente de bits multiplicados a través del dispositivo de comunicación del dispositivo de decodificación de habla 24 en la información suplementaria del envoltorio temporal, la información suplementaria de SBR y la corriente de bits codificados. La información suplementaria del envoltorio temporal también puede ser $K(r)$ descrito en el primer ejemplo o $s(i)$ descrito en el tercer ejemplo. La información suplementaria del envoltorio temporal también puede ser otro parámetro $X(r)$ que no es ni $K(r)$ ni $s(i)$.

La unidad de conversión de información suplementaria $2w$ convierte la información suplementaria de envoltorio temporal suministrada para obtener $K(r)$ y $s(i)$. Si la información suplementaria del envoltorio temporal es $K(r)$, la unidad de conversión de información suplementaria $2w$ convierte $K(r)$ en $s(i)$. La unidad de conversión de información suplementaria $2w$ puede también obtener, por ejemplo, un valor medio de $K(r)$ en una sección de $b_i \leq r < b_i + 1$

$$\overline{K(i)} \quad \text{---(29)}$$

y convertir el valor medio representado en la expresión (29) en $s(i)$ utilizando una tabla predeterminada. Si la información suplementaria del envoltorio temporal es $s(i)$, la unidad de conversión de información suplementaria $2w$ convierte $s(i)$ en $K(r)$. La unidad de conversión de información suplementaria $2w$ puede también realizar la conversión convirtiendo $s(i)$ en $K(r)$, por ejemplo, utilizando una tabla predeterminada. Debe observarse que i y r se asocian entre sí con el fin de cumplir la relación de $b_i \leq r < b_i + 1$.

Si la información suplementaria del envoltorio temporal es un parámetro $X(r)$ que no es ni $s(i)$ ni $K(r)$, la unidad de conversión de información suplementaria $2w$ convierte $X(r)$ en $K(r)$ y $s(i)$. Es preferible que la unidad de conversión de información suplementaria $2w$ convierta $X(r)$ en $K(r)$ y $s(i)$, por ejemplo, utilizando una tabla predeterminada. También es preferible que la unidad de conversión de información suplementaria $2w$ transmita $X(r)$ como valor representativo cada envoltorio de SBR. Las tablas para convertir $X(r)$ en $K(r)$ y $s(i)$ pueden ser distintas entre sí.

(Modificación 3 del Primer Ejemplo)

En el dispositivo de decodificación de habla 21 del Primer ejemplo, la unidad de filtro de predicción lineal $2k$ del dispositivo de decodificación de habla 21 puede incluir un proceso de control de incremento automático. El proceso de control de incremento automático es un proceso para ajustar la potencia de la señal en el dominio de QMF enviada desde la unidad de filtro de predicción lineal $2k$ a la potencia de la señal en el dominio de QMF que se proporciona. En general, se realiza una señal $q_{syn,pow}(n, r)$ en el dominio QMF cuyo incremento se ha controlado mediante la siguiente expresión:

$$q_{syn,pow}(n, r) = q_{syn}(n, r) \cdot \sqrt{\frac{P_0(r)}{P_1(r)}} \quad \text{---(30)}$$

Aquí, $P_0(r)$ y $P_1(r)$ se expresan mediante las siguientes expresiones (31) y (32).

$$P_0(r) = \sum_{n=k_x}^{\infty} |q_{adj}(n, r)|^2 \quad \text{---(31)}$$

$$P_1(r) = \sum_{n=k_x}^{63} |q_{syn}(n, r)|^2 \quad \text{---(32)}$$

Al realizar el proceso de control de incremento automático, la potencia de los componentes de alta frecuencia de la salida de señal de la unidad de filtro de predicción lineal 2k se ajusta a un valor equivalente al de antes del filtrado de predicción lineal. Como resultado, el efecto de ajustar la intensidad de la señal de alta frecuencia realizado por la unidad de ajuste de alta frecuencia 2j puede mantenerse para la señal de salida de la unidad de filtro de predicción lineal 2k en la cual se ha formado el envoltorio temporal de los componentes de alta frecuencia generados basados en SBR. El proceso de control de incremento automático también puede realizarse de forma individual en un espectro de frecuencia específico de la señal en el dominio QMF. El proceso realizado en el espectro de frecuencia individual puede realizarse limitando n en la expresión (30), la expresión (31) y la expresión (32) dentro de un espectro de frecuencia específico. Por ejemplo, el espectro de frecuencia i-ésimo puede expresarse como $F_i \leq n < F_{i+1}$ (en este caso, i es un índice que indica el número de un espectro de frecuencia específico de la señal en el dominio QMF). F_i indica los límites del espectro de frecuencia, y es preferible que F_i sea una tabla de límite de frecuencia de un factor de escala de envoltorio definido en SBR en "MPEG4 AAC". La tabla de límite de frecuencia está definida por la unidad de generación de alta frecuencia 2g basada en la definición de SBR en "MPEG4 AAC". Al realizar el proceso de control de incremento automático, la intensidad de la señal de salida de la unidad de filtro de predicción lineal 2k en un espectro de frecuencias específico de los componentes de alta frecuencia se ajusta a un valor equivalente al de antes del filtro de predicción temporal. Como resultado, el efecto para ajustar la intensidad de la señal de alta frecuencia realizado por la unidad de ajuste de alta frecuencia 2j en la señal de salida de la unidad de filtro de predicción lineal 2k en la cual se forma el envoltorio temporal de los componentes de alta frecuencia generados basados en SBR, se mantiene por unidad de espectro de frecuencia. Los cambios realizados en la presente modificación 3 del primer ejemplo pueden también hacerse para la unidad de filtro de predicción lineal 2k de la realización.

([Modificación 1 del Tercer Ejemplo])

La unidad de cálculo de parámetros de forma de envoltorio 1n en el dispositivo de codificación de habla 13 del Tercer ejemplo puede también realizarse mediante el proceso siguiente. La unidad de cálculo de parámetros de forma de envoltorio 1n obtiene un parámetro de forma de envoltorio $s(i)$ ($0 \leq i < N_e$) de acuerdo con la siguiente expresión (33) para cada envoltorio de SBR en la secuencia codificada.

$$s(i) = 1 - \min\left(\frac{e(r)}{e(i)}\right) \quad \text{---(33)}$$

Debe observarse que:

$$\overline{e(i)} \quad \text{---(34)}$$

es un valor medio de $e(r)$ en el envoltorio de SBR, y el método de cálculo está basado en la expresión (21). Debe observarse que el envoltorio de SBR indica el segmento de tiempo que cumple $b_i \leq r < b_{i+1}$. $\{b_i\}$ son los límites de tiempo de los envoltorios de SBR incluidos en la información suplementaria de SBR como información, y son los límites del segmento de tiempo para el cual se da el factor de escala de envoltorio de SBR que representa la energía de señal media en un segmento de tiempo determinado y un espectro de frecuencia determinado. $\min(\cdot)$ representa el valor mínimo dentro del espectro de $b_i \leq r < b_{i+1}$. Por consiguiente, en este caso, el parámetro de forma de envoltorio $s(i)$ es un parámetro para indicar un coeficiente del valor mínimo para el valor medio de la información del envoltorio temporal ajustado en el

envoltorio de SBR. La unidad de ajuste de forma de envoltorio 2s en el dispositivo de decodificación de habla 23 del tercer ejemplo también puede realizarse por medio del proceso siguiente. La unidad de ajuste de forma de envoltorio 2s ajusta $e(r)$ utilizando $s(i)$ para obtener la información del envoltorio temporal ajustada $e_{adj}(r)$. El método de ajuste se basa en la siguiente expresión (35) o expresión (36).

$$e_{adj}(r) = \overline{e(i)} \left(1 + s(i) \frac{(e(r) - \overline{e(i)})}{e(i) - \min(e(r))} \right) \quad \text{---(35)}$$

$$e_{adj}(r) = \overline{e(i)} \left(1 + s(i) \frac{(e(r) - \overline{e(i)})}{e(i)} \right) \quad \text{---(36)}$$

5

La expresión 35 ajusta la forma del envoltorio para que el coeficiente del valor mínimo para el valor medio de la información del envoltorio temporal ajustada $e_{adj}(r)$ en el envoltorio de SBR se convierta en equivalente al valor del parámetro de forma de envoltorio $s(i)$. Los cambios realizados en la modificación 1 del Tercer ejemplo descrito más arriba también pueden realizarse en la realización.

10 [Modificación 2 del Tercer Ejemplo]

La unidad de formación del envoltorio temporal 2v puede utilizar también la siguiente expresión en lugar de la expresión (28). Tal como se indica en la expresión (37), $e_{adj, scaled}(r)$ se obtiene controlando el incremento de la información del envoltorio temporal ajustado $e_{adj}(r)$, de manera que la potencia de $q_{envadj}(k,r)$ mantenga la de $q_{adj}(k,r)$ dentro del envoltorio de SBR. Tal como se indica en la expresión (38), en la presente modificación 2 del Tercer ejemplo, $q_{envadj}(k,r)$ se obtiene multiplicando la señal $q_{adj}(k,r)$ en el dominio QMF mediante $e_{adj, scaled}(r)$ en lugar de $e_{adj}(r)$. Por consiguiente, la unidad de formación del envoltorio temporal 2v puede formar el envoltorio temporal de la señal $q_{adj}(k,r)$ en el dominio QMF, de manera que la potencia de señal dentro del envoltorio de SBR se convierta en equivalente antes y después de la formación del envoltorio temporal. Debe observarse que el envoltorio de SBR indica el segmento de tiempo que cumple $b_i \leq r < b_{i+1}$. $\{b_i\}$ son los límites de tiempo de los envoltorios de SBR incluidos en la información suplementaria de SBR como información, y son los límites del segmento de tiempo para el cual se dan el factor de escalado de envoltorio de SBR que representa la energía de señal media de un segmento de tiempo determinado y un espectro de frecuencia determinado. La terminología "envoltorio de SBR" en las realizaciones la presente invención corresponde a la terminología "segmento de tiempo de envoltorio de SBR" en "MPEG4 AAC" definido en "ISO/IEC 14496-3", y el "envoltorio de SBR" tiene el mismo contenido que el "segmento de tiempo de envoltorio de SBR" a lo largo de las realizaciones.

15

20

25

$$e_{adj, scaled}(r) = e_{adj}(r) \cdot \sqrt{\frac{\sum_{k=k_x}^{63} \sum_{r=b_i}^{b_{i+1}-1} |q_{adj}(k,r)|^2}{\sum_{k=k_x}^{63} \sum_{r=b_i}^{b_{i+1}-1} |q_{adj}(k,r) \cdot e_{adj}(r)|^2}} \quad \text{---(37)}$$

($k_x \leq k \leq 63, b_i \leq r < b_{i+1}$)

$$q_{envadj}(k, r) = q_{adj}(k, r) \cdot e_{adj, scaled}(r) \quad \text{---(38)}$$

($k_x \leq k \leq 63, b_i \leq r < b_{i+1}$)

Los cambios realizados en la presente modificación 2 del Tercer ejemplo descrita anteriormente también pueden hacerse en la cuarta realización.

5 (Modificación 3 del Tercer ejemplo)

La expresión (19) puede ser también la siguiente expresión (39).

$$e(r) = \sqrt{\frac{(b_{i+1} - b_i) \sum_{k=0}^{63} |q(k, r)|^2}{\sum_{r=b_i}^{b_{i+1}-1} \sum_{k=0}^{63} |q(k, r)|^2}} \quad \text{---(39)}$$

La expresión (22) puede ser también la siguiente expresión (40).

$$e(r) = \sqrt{\frac{(b_{i+1} - b_i) \sum_{k=0}^{63} |q_{dec}(k, r)|^2}{\sum_{r=b_i}^{b_{i+1}-1} \sum_{k=0}^{63} |q_{dec}(k, r)|^2}} \quad \text{---(40)}$$

10

La expresión (26) puede ser también la siguiente expresión (41).

$$e_{exp}(r) = \sqrt{\frac{(b_{i+1} - b_i) \sum_{k=k_x}^{63} |q_{exp}(k, r)|^2}{\sum_{r=b_i}^{b_{i+1}-1} \sum_{k=k_x}^{63} |q_{exp}(k, r)|^2}} \quad \text{---(41)}$$

15

Cuando se utilizan la expresión (39) y la expresión (40), la información del envoltorio temporal $e(r)$ es información en la cual la potencia de cada muestra de subbanda de QMF es normalizada por la potencia media en el envoltorio de SBR, y se extrae la raíz cuadrada. Sin embargo, la muestra de subbanda de QMF es un vector de señal correspondiente al índice de tiempo "r" en la señal del dominio de QMF, y es una submuestra en el dominio QMF. En todos los ejemplos y las realizaciones de la presente invención, la terminología "ranura de tiempo" tiene el mismo contenido que la "muestra de subbanda de QMF". En este

caso, la información del envoltorio temporal $e(r)$ es un coeficiente de incremento que debería ser multiplicado por cada muestra de subbanda de QMF, y lo mismo se aplica a la información del envoltorio temporal ajustada $e_{adj}(r)$.

(Modificación 1 de la Realización)

5 Un dispositivo de decodificación de habla 24a (que no se ilustra) de una modificación 1 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se
 10 ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24a cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de decodificación de habla 24a como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24a recibe la salida de corriente de bits multiplicados codificados del dispositivo de codificación de habla 11 o del dispositivo de codificación de habla 13, y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24a. El dispositivo de decodificación de habla 24a incluye funcionalmente una unidad de separación de corriente de bits 2a4 (no
 15 ilustrada) en lugar de la unidad de separación de corriente de bits 2a3 del dispositivo de decodificación de habla 24, e incluye también una unidad de generación de información suplementaria de envoltorio temporal 2y (no ilustrada) en lugar de la unidad de conversión de información suplementaria 2w. La unidad de separación de corriente de bits 2a4 separa la corriente de bits multiplicados en la información de SBR y la corriente de bits codificados. La unidad de generación de información suplementaria de envoltorio temporal 2y genera información suplementaria de envoltorio temporal basada en la información
 20 incluida en la corriente de bits codificados y la información suplementaria de SBR.

Para generar la información suplementaria de envoltorio temporal en un envoltorio SBR determinado, por ejemplo, el ancho de tiempo (b_i+1-b_i) del envoltorio de SBR, puede utilizarse una clase de secuencia, un parámetro de intensidad del filtro inverso, un suelo de ruido, la amplitud de la potencia de alta frecuencia, un coeficiente de la potencia de alta frecuencia en relación con la potencia de baja frecuencia, un
 25 coeficiente de autocorrelación o un incremento de predicción de un resultado de realizar un análisis de predicción lineal en la dirección de frecuencia en una señal de baja frecuencia representada en el dominio QMF, y similares. La información suplementaria de envoltorio temporal puede ser generada determinando $K(r)$ o $s(i)$ basados en uno o en una pluralidad de valores de los parámetros. Por ejemplo, la información suplementaria de envoltorio temporal puede ser generada determinando $K(r)$ o $s(i)$ basados e (b_i+1-b_i) de
 30 manera que $K(r)$ o $s(i)$ sea reducido a medida que se incrementa el ancho de tiempo (b_i+1-b_i) del envoltorio de SBR, o $K(r)$ o $s(i)$ se incrementa a medida que el ancho de tiempo (b_i+1-b_i) del envoltorio de SBR se incrementa. Los cambios similares pueden hacerse también para el primer ejemplo y el tercer ejemplo.

(Modificación 2 de la Realización)

35 Un dispositivo de decodificación de habla 24b (ver FIG. 15) de una modificación 2 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24b cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de decodificación de habla 24b como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24b recibe la corriente de bits multiplicados codificados enviados desde el
 40 dispositivo de codificación de habla 11 o del dispositivo de codificación de habla 13, y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24b. El dispositivo de decodificación de habla 24b, tal como se ilustra en la FIG. 15, incluye una unidad de ajuste de alta frecuencia primaria 2j1 y una unidad de ajuste de alta frecuencia secundaria 2j2, en lugar de la unidad de ajuste de alta frecuencia 2j.

En este punto, la unidad de ajuste de alta frecuencia primaria 2j1 ajusta una señal en el dominio QMF de la banda de alta frecuencia realizando un filtrado inverso de predicción lineal en la dirección temporal, el ajuste de incremento, y la adición de ruido, descrito en la fase de "generación de HF" y la fase de "ajuste de HF" en SBR en "MPEG4 AAC". En este momento, la señal de salida en la unidad de ajuste de alta
 50 frecuencia primaria 2j1 corresponde a una señal W_2 en la descripción en "herramienta de SBR" en "ISO/IEC 14496-3:2005", cláusulas 4.6.18.7.6 de "Montaje de señales de HF". La unidad de filtro de predicción lineal 2k (o la unidad de filtro de predicción lineal 2k1) y la unidad de formación del envoltorio temporal 2v dan forma al envoltorio temporal de la señal de salida desde la unidad de ajuste de alta frecuencia primaria. La unidad de ajuste de alta frecuencia secundaria 2j2 realiza un proceso de adición de sinusoides en la fase de "ajuste de HF" en SBR en "MPEG4 AAC". El proceso de la unidad de ajuste de alta frecuencia secundario corresponde a un proceso de generación de una señal Y a partir de la señal W_2 en la descripción en "herramienta de SBR" en "ISO/IEC 14496-3:2005", cláusulas 4.6.18.7.6 de "Montaje de señales de HF", en las cuales la señal W_2 es sustituida por una señal de salida de la unidad de formación del envoltorio temporal 2v.

En la descripción indicada más arriba, solamente el proceso para añadir sinusoides es realizado por la unidad de ajuste de alta frecuencia 2j2. Sin embargo, cualquiera de los procesos de la fase de "ajuste de HF" puede ser realizado por la unidad secundaria de ajuste de alta frecuencia 2j2. También pueden realizarse modificaciones similares en el primer ejemplo, el segundo ejemplo y el tercer ejemplo. En estos casos, la unidad de filtro de predicción lineal (unidades de filtro de predicción lineal (2k y 2k1) se incluye en la primera realización y la segunda realización, pero la unidad de formación del envoltorio temporal no está incluida. En consecuencia, una señal de salida desde la unidad primaria de ajuste de alta frecuencia 2j1 es procesada por la unidad de filtro de predicción lineal, y a continuación una señal de salida de la unidad de filtro de predicción lineal es procesada por la unidad secundaria de ajuste de alta frecuencia 2j2.

En el tercer ejemplo, se encuentra incluida la unidad de formación del envoltorio temporal 2v, pero la unidad de filtro de predicción lineal no está incluida. En consecuencia, una señal de salida de la unidad primaria de ajuste de alta frecuencia 2j1 es procesada por la unidad de formación del envoltorio temporal 2v, y a continuación una señal enviada desde la unidad de formación del envoltorio temporal 2v es procesada por la unidad secundaria de ajuste de alta frecuencia.

En el dispositivo de decodificación de habla (dispositivo de decodificación de habla 24, 24a, o 24b) de la realización, el orden de proceso de la unidad de filtro de predicción lineal 2k y la unidad de formación del envoltorio temporal 2v puede invertirse. En otras palabras, una señal de salida de la unidad de ajuste de alta frecuencia 2j o la unidad primaria de ajuste de alta frecuencia 2j1 puede ser procesada primero por la unidad de formación del envoltorio temporal 2v, y a continuación una señal de salida de la unidad de formación del envoltorio temporal 2v puede ser procesada por la unidad de filtro de predicción lineal 2k.

Asimismo, solamente si la información suplementaria de envoltorio temporal incluye información de control binario para indicar si el proceso es realizado por la unidad de filtro de predicción lineal 2k o la unidad de formación del envoltorio temporal 2v, y la información de control indica realizar el proceso por parte de la unidad de filtro de predicción lineal 2k o la unidad de formación del envoltorio temporal 2v, la información suplementaria del envoltorio temporal puede utilizar una forma que incluye también por lo menos un elemento entre el parámetro de intensidad de filtro $K(r)$, el parámetro de intensidad de filtro $s(i)$, o $X(r)$ que es un parámetro para determinar $K(r)$ y $s(i)$ como información.

(Modificación 3 de la Realización)

Un dispositivo de decodificación de habla 24c (ver FIG. 16) de una modificación 3 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24c cargando y ejecutando un programa de ordenador predeterminado (como por ejemplo un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 17) almacenado en una memoria integrada del dispositivo de decodificación de habla 24c como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24c recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24c. Tal como se ilustra en la FIG. 16, el dispositivo de decodificación de habla 24c incluye una unidad primaria de ajuste de alta frecuencia 2j3 y una unidad secundaria de ajuste de alta frecuencia 2j4 en lugar de la unidad de ajuste de alta frecuencia 2j, y también incluye unidades de ajuste de componentes de señal individual 2z1, 2z2 y 2z3 en lugar de la unidad de filtro de predicción lineal 2k y la unidad de formación del envoltorio temporal 2v (las unidades de ajuste de componentes de señal individual corresponden al medio de formación del envoltorio temporal).

La unidad primaria de ajuste de alta frecuencia 2j3 envía una señal de salida en el dominio QMF de la banda de alta frecuencia como componente de señal de copia. La unidad primaria de ajuste de alta frecuencia 2j3 puede proporcionar una señal sobre la cual se realiza por lo menos el filtrado inverso de predicción lineal en la dirección temporal y/o el ajuste de incremento (ajuste de características de frecuencia) sobre la señal en el dominio QMF de la banda de alta frecuencia, utilizando la información suplementaria de SBR recibida de la unidad de separación de corriente de bits 2a3, como componente de señal de copia. La unidad primaria de ajuste de alta frecuencia 2j3 también genera un componente de señal de ruido y un componente de señal de senoide utilizando la información suplementaria de SBR proporcionada por la unidad de separación de corriente de bits 2a3, y emite el componente de señal de copia, el componente de señal de ruido y el componente de señal de senoide separadamente (proceso en fase Sg1). El componente de señal de ruido y el componente de señal de senoide pueden no ser generados, dependiendo del contenido de la información suplementaria de SBR.

Las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 realizan el proceso en toda la pluralidad de componentes de señal incluidos en la señal de salida del medio primario de ajuste de alta frecuencia (proceso en Fase Sg2). El proceso con las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 puede ser un filtrado de síntesis de predicción lineal en la dirección de frecuencia obtenido de la unidad de ajuste de intensidad de filtro 2f utilizando los coeficientes de

predicción lineal, similar a la unidad de filtro de predicción lineal 2k (proceso 1). El proceso con las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 puede ser también un proceso de multiplicar cada muestra de subbanda de QMF por un coeficiente de incremento utilizando el envoltorio temporal obtenido de la unidad de ajuste de forma de envoltorio 2s, similar al de la unidad de formación del envoltorio temporal 2v (proceso 2). El proceso con las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 puede ser también un proceso de realizar un filtrado de síntesis de predicción lineal en la dirección de frecuencia sobre la señal de entrada utilizando los coeficientes de predicción lineal obtenidos de la unidad de ajuste de intensidad de filtro 2f similar a la de la unidad de filtro de predicción lineal 2k, y a continuación multiplicando cada muestra de subbanda de QMF por un coeficiente de incremento utilizando el envoltorio temporal obtenido de la unidad de ajuste de forma de envoltorio 2s, similar al de la unidad de formación del envoltorio temporal 2v (proceso 3). El proceso con las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 puede ser también un proceso de multiplicar cada muestra de subbanda de QMF con relación a la señal de entrada por un coeficiente de incremento utilizando el envoltorio temporal obtenido de la unidad de ajuste de forma de envoltorio 2s, similar a la de la unidad de formación del envoltorio temporal 2v, y seguidamente realizar un filtrado de síntesis de predicción lineal en la dirección de frecuencia sobre la señal de salida utilizando los coeficientes de predicción lineal obtenidos de la unidad de ajuste de intensidad de filtro 2f, similar a la de la unidad de filtro de predicción lineal 2k (proceso 4). Las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 pueden no realizar el proceso de formación del envoltorio temporal sobre la señal de entrada, pero pueden emitir la señal de entrada tal como está (proceso 5). El proceso con las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 puede incluir cualquier proceso para dar forma al envoltorio temporal de la señal de entrada utilizando un método distinto de los procesos 1 a 5 (proceso 6). El proceso con las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 puede ser también un proceso en el cual una pluralidad de procesos entre los procesos 1 a 6 se combinan en un orden arbitrario (proceso 7).

Los procesos con las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 pueden ser los mismos, pero las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 pueden dar forma al envoltorio temporal de cada uno de los componentes de señal incluidos en la salida del medio primario de ajuste de alta frecuencia mediante distintos métodos. Por ejemplo, pueden realizarse distintos procesos sobre la señal de copia, la señal de ruido y la señal de senoide, de manera que una unidad de ajuste de componentes de señal individuales 2z1 realice el proceso 2 en la señal de copia proporcionada, la unidad de ajuste de componentes de señal individuales 2z2 realice el proceso 3 en el componente de señal de ruido proporcionado, y la unidad de ajuste de componentes de señal individuales 2z3 realice el proceso 5 en la señal de senoide proporcionada. En este caso, la unidad de ajuste de intensidad de filtro 2f y la unidad de ajuste de forma de envoltorio 2s pueden transmitir los mismos coeficientes de predicción lineal, y los envoltorios temporales a las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3, pero también puede transferir diferentes coeficientes de predicción lineal y envoltorios temporales. También resulta posible transmitir los mismos coeficientes de predicción lineal y envoltorios temporales a por lo menos dos de las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3. Dado que por lo menos una de las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 puede no realizar el proceso de formación del envoltorio temporal sino emitir la señal de entrada tal como se encuentra (proceso 5), las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 realizan el proceso de envoltorio temporal en por lo menos uno de la pluralidad de componentes de señal emitidos desde la unidad primaria de ajuste de alta frecuencia 2j3 en conjunto (si todas las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 realizan el proceso 5, el proceso de formación del envoltorio temporal no se realiza en ninguno de los componentes de señal, y los efectos de la presente invención no se muestran).

Los procesos realizados por cada una de las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 pueden estar fijados a uno de los procesos 1 a 7, pero pueden estar determinados dinámicamente para realizar uno de los procesos 1 a 7, basados en la información de control recibida desde fuera del dispositivo de decodificación de habla 24c. En este momento, resulta preferible que la información de control esté incluida en la corriente de bits multiplicados. La información de control puede ser una instrucción para realizar cualquiera de los procesos 1 a 7 en un segmento de tiempo de envoltorio de SBR específico, en la secuencia codificada o en el otro segmento de tiempo, o puede ser una instrucción para realizar cualquiera de los procesos 1 a 7 sin especificar el segmento de tiempo de control.

La unidad secundaria de ajuste de alta frecuencia 2j4 añade la señal de salida de los componentes de señal procesados desde las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3, y envía el resultado a la unidad de adición de coeficiente (proceso en Fase Sg3). La unidad secundaria de ajuste de alta frecuencia 2j4 puede realizar por lo menos uno de entre el filtrado inverso de predicción lineal en la dirección temporal y el ajuste de incremento (ajuste de características de frecuencia) en el componente de señal de copia, utilizando la información suplementaria de SBR recibida de la unidad de separación de corriente de bits 2a3.

Las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 pueden funcionar en cooperación entre sí, y generar una señal de salida en una fase intermedia añadiendo por lo menos dos componentes de señal en los cuales se realiza cualquiera de los procesos 1 a 7, y a continuación realizando cualquiera de los procesos 1 a 7 en la señal añadida. En este momento, la unidad secundaria de ajuste de alta frecuencia 2j4 añade la señal de salida en la fase intermedia y un componente de señal que todavía no ha sido añadido a la señal de salida en la fase intermedia, y envía el resultado a la unidad de adición de coeficiente. Más específicamente, resulta preferible generar una señal de salida en la fase intermedia realizando el proceso 5 en el componente de señal de copia, aplicando el proceso 1 al componente de ruido, añadiendo los dos componentes de señal y posteriormente aplicando el proceso 2 a la señal añadida. En este momento, la unidad secundaria de ajuste de alta frecuencia 2j4 añade el componente de señal de senoide a la señal de salida en la fase intermedia, y envía el resultado a la unidad de adición de coeficiente.

La unidad primaria de ajuste de alta frecuencia 2j3 puede enviar cualquiera de entre una pluralidad de componentes de señal en una forma separada de las otras además de los tres componentes de señal del componente de señal de copia, el componente de señal de ruido y el componente de señal de senoide. En este caso, el componente de señal puede obtenerse añadiendo por lo menos dos de entre el componente de señal de copia, el componente de señal de ruido y el componente de señal de senoide. El componente de señal puede ser también una señal obtenida dividiendo la banda de uno de entre el componente de señal de copia, el componente de señal de ruido y la señal de senoide. El número de componentes de señal puede ser distinto de tres, y en ese caso, el número de unidades de ajuste de componentes de señal individual puede ser distinto de tres.

La señal de alta frecuencia generada por SBR consiste en tres elementos del componente de señal de copia obtenidos copiando desde la banda de baja frecuencia a la banda de alta frecuencia, la señal de ruido y la señal de senoide. Dado que la señal de copia, la señal de ruido y la señal de senoide tienen envoltorios temporales distintos entre sí, si el envoltorio temporal de cada componente de la señal se forma utilizando métodos diferentes tales como las unidades de ajuste de componentes de señal individual de la presente modificación, resulta posible mejorar todavía más la calidad subjetiva de la señal decodificada en comparación con los otros ejemplos. En particular, dado que la señal de ruido generalmente tiene un envoltorio temporal liso, y la señal de copia tiene un envoltorio temporal próximo al de la señal en la banda de baja frecuencia, los envoltorios temporales de la señal de copia y de la señal de ruido pueden ser controlados de forma independiente, manejándolos por separado y aplicando diferentes procesos. Por consiguiente, resulta efectivo para mejorar la calidad subjetiva de la señal decodificada. Más específicamente, resulta preferible realizar un proceso de formación del envoltorio temporal en la señal de ruido (proceso 3 o proceso 4), realizar un proceso distinto del de la señal de ruido en la señal de copia (proceso 1 o proceso 2), y realizar el proceso 5 en la señal de senoide (en otras palabras, no se realiza el proceso de formación del envoltorio temporal). También resulta preferible realizar un proceso de formación (proceso 3 o proceso 4) del envoltorio temporal en la señal de ruido, y realizar el proceso 5 en la señal de copia y la señal de senoide (en otras palabras, no se realiza el proceso de formación del envoltorio temporal).

(Modificación 4 del Primer Ejemplo)

Un dispositivo de codificación de habla 11b (FIG. 44) de una modificación 4 del Primer ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 11b cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla 11b como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 11b recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 11b y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 11b. El dispositivo de codificación de habla 11b incluye una unidad de análisis de predicción lineal 1e1 en lugar de la unidad de análisis de predicción lineal 1e del dispositivo de codificación de habla 11b, y también incluye una unidad de selección de ranura de tiempo 1p.

La unidad de selección de ranura de tiempo 1p recibe una señal en el dominio QMF de la unidad de transformación de frecuencia 1a y selecciona una ranura de tiempo en la cual se realiza el análisis de predicción lineal por parte de la unidad de análisis de predicción lineal 1e1. La unidad de análisis de predicción lineal 1e1 realiza un análisis de predicción lineal en la señal de dominio QMF en la ranura de tiempo seleccionada como la unidad de análisis de predicción lineal 1e, basada en el resultado de selección transmitido desde la unidad de selección de ranura de tiempo 1p, para obtener por lo menos uno de entre el coeficiente de predicción lineal de alta frecuencia y el coeficiente de predicción lineal de baja frecuencia. La unidad de cálculo de parámetros de intensidad de filtro 1f calcula un parámetro de intensidad de filtro utilizando coeficientes de predicción lineal de la ranura de tiempo seleccionada por la unidad de selección de ranura de tiempo 1p, obtenido por la unidad de análisis de predicción lineal 1e1. Para seleccionar una ranura de tiempo por parte de la unidad de selección de ranura de tiempo 1p, por

ejemplo puede utilizarse al menos un método de selección que utilice la potencia de señal de la señal de dominio QMF de los componentes de alta frecuencia, similar al de una unidad de selección de ranura de tiempo 3 en un dispositivo de decodificación 21a de la presente modificación, que se describirá más adelante. En este momento, resulta preferible que la señal de dominio QMF de los componentes de alta frecuencia en la unidad de selección de ranura de tiempo 1p sea un componente de frecuencia codificado por la unidad de codificación de SBR 1d, entre las señales en el dominio QMF recibidas de la unidad de transformación de frecuencia 1a. El método de selección de la ranura de tiempo puede ser por lo menos uno de los métodos descritos más arriba, puede incluir al menos un método distinto de los que se han descrito anteriormente, o puede ser la combinación de los mismos.

Un dispositivo de decodificación de habla 21a (ver FIG. 18) de la modificación 4 del primer ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 21a cargando y ejecutando un programa de ordenador predeterminado (como por ejemplo un programa de ordenador para realizar procesos, ilustrados en el diagrama de flujo de la FIG. 19) almacenado en una memoria integrada del dispositivo de decodificación de habla 21a como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 21a recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 21a. Tal como se ilustra en la FIG. 18, el dispositivo de decodificación de habla 21a incluye una unidad de análisis de predicción lineal de baja frecuencia 2d1, una unidad de detección de cambio de señal 2e1, una unidad de análisis de predicción lineal de alta frecuencia 2h1, una unidad de filtro inverso de predicción lineal 2i1, y una unidad de filtro de predicción lineal 2k3 en lugar de la unidad de análisis de predicción lineal de baja frecuencia 2d, la unidad de detección de cambio de señal 2e, la unidad de análisis de predicción lineal de alta frecuencia 2h, la unidad de filtro inverso de predicción lineal 2i, y la unidad de filtro de predicción lineal 2k del dispositivo de decodificación de habla 21, y también incluye la unidad de selección de ranura de tiempo 3a.

La unidad de selección de ranura de tiempo 3a determina si va a realizarse el filtrado de síntesis de predicción lineal en la unidad de filtro de predicción lineal 2k en la señal $q_{exp}(k, r)$ en el dominio QMF de los componentes de alta frecuencia de la ranura de tiempo r generados por la unidad de generación de alta frecuencia 2g, y selecciona una ranura de tiempo en la cual se realiza el filtrado de síntesis de predicción lineal (proceso en Fase Sh1). La unidad de selección de ranura de tiempo 3a notifica, del resultado de la selección de la ranura de tiempo, la unidad de análisis de predicción lineal de baja frecuencia 2d1, la unidad de detección de cambio de señal 2e1, la unidad de análisis de predicción lineal de alta frecuencia 2h1, la unidad de filtro inverso de predicción lineal 2i1 y la unidad de filtro de predicción lineal 2k3. La unidad de análisis de predicción lineal de baja frecuencia 2d1 realiza un análisis de predicción lineal en la señal de dominio QMF en la ranura de tiempo seleccionada r1, de la misma manera en que la unidad de análisis de predicción lineal de baja frecuencia 2d, basándose en el resultado de la selección transmitido desde la unidad de selección de ranura de tiempo 3a, para obtener coeficientes de predicción lineal de baja frecuencia (proceso en Fase Sh2). La unidad de detección de cambio de señal 2e1 detecta la variación temporal en la señal del dominio QMF en la ranura de tiempo seleccionada, como la unidad de detección de cambio de señal 2e, basándose en el resultado de selección transmitido desde la unidad de selección de ranura de tiempo 3a, y produce un resultado de detección T (r1).

La unidad de ajuste de intensidad de filtro 2f realiza un ajuste de intensidad de filtro en los coeficientes de predicción lineal de baja frecuencia de la ranura de tiempo seleccionada por la unidad de selección de ranura de tiempo 3a obtenida por la unidad de análisis de predicción lineal de baja frecuencia 2d1, para obtener unos coeficientes de predicción lineal $a_{dec}(n, r1)$ ajustados. La unidad de análisis de predicción lineal de alta frecuencia 2h1 realiza un análisis de predicción lineal en la dirección de frecuencia en la señal de dominio QMF de los componentes de alta frecuencia generados por la unidad de generación de alta frecuencia 2g para la ranura de tiempo r1 seleccionada, basándose en el resultado de la selección transmitido desde la unidad de selección de ranura de tiempo 3a, como la unidad de análisis de predicción lineal de alta frecuencia 2k, para obtener unos coeficientes de predicción lineal de alta frecuencia $a_{exp}(n, r1)$ (proceso en Fase Sh3). La unidad de filtro inverso de predicción lineal 2i1 realiza un filtrado inverso de predicción lineal, en el cual $a_{exp}(n, r1)$ son coeficientes, en la dirección de frecuencia en la señal $q_{exp}(k, r)$ en el dominio QMF de los componentes de alta frecuencia de la ranura de tiempo r1 seleccionada, como la unidad de filtro inverso de predicción lineal 2i, basado en los resultados de la selección transmitidos desde la unidad de selección de ranura de tiempo 3a (proceso en Fase Sh4).

La unidad de filtro de predicción lineal 2k3 realiza un filtrado de síntesis de predicción lineal en la dirección de frecuencia en una señal $q_{adj}(k, r1)$ en el dominio QMF de los componentes de alta frecuencia enviados desde la unidad de ajuste de intensidad de filtro 2f en la ranura de tiempo r1 seleccionada utilizando $a_{adj}(n, r1)$ obtenido de la unidad de ajuste de intensidad de filtro 2f, como la unidad de filtro de predicción lineal 2k, basado en el resultado de la selección transmitido desde la unidad de selección de ranura de tiempo 3a (proceso en Fase Sh5). Los cambios realizados en la unidad de filtro de predicción lineal 2k descritos en la modificación 3 también pueden realizarse en la unidad de filtro de predicción lineal 2k3. Para

seleccionar una ranura de tiempo en la cual se realiza el filtrado de síntesis de predicción lineal, por ejemplo, la unidad de selección de ranura de tiempo 3a puede seleccionar por lo menos una ranura de tiempo r en la cual la potencia de señal de la señal de dominio QMF $q_{exp}(k, r)$ de los componentes de alta frecuencia es mayor que un valor $P_{exp, Th}$ predeterminado. Es preferible calcular la potencia de señal de $q_{exp}(k, r)$ de acuerdo con la siguiente expresión.

$$P_{exp}(r) = \sum_{k=k_x}^{k_x+M-1} |q_{exp}(k, r)|^2 \quad \text{---(42)}$$

donde M es un valor que representa un espectro de frecuencia superior a una frecuencia de límite inferior k_x de los componentes de alta frecuencia generados por la unidad de generación de alta frecuencia 2g, y el espectro de frecuencia de los componentes de alta frecuencia generados por la unidad de generación de alta frecuencia 2g puede estar representado como $k_x \leq k < k_x + M$. El valor predeterminado $P_{exp, Th}$ puede ser también un valor medio de $P_{exp}(r)$ de un ancho de tiempo predeterminado que incluye la ranura de tiempo r. El ancho de tiempo predeterminado también puede ser el envoltorio de SBR.

La selección también puede hacerse para incluir una ranura de tiempo en la cual la potencia de señal de la señal de dominio QMF de los componentes de alta frecuencia alcanza su pico. La potencia de señal pico puede calcularse, por ejemplo, utilizando un valor medio móvil:

$$P_{exp, MA}(r) \quad \text{---(43)}$$

de la potencia de señal, y la potencia de señal pico puede ser la potencia de señal en el dominio QMF de los componentes de alta frecuencia de la ranura de tiempo r en la cual el resultado de

$$P_{exp, MA}(r+1) - P_{exp, MA}(r) \quad \text{---(44)}$$

cambia de valor positivo a valor negativo. El valor móvil medio de la potencia de señal,

$$P_{exp, MA}(r) \quad \text{---(45)}$$

por ejemplo, puede calcularse mediante la expresión siguiente

$$P_{exp, MA}(r) = \frac{1}{c} \sum_{r'=r-\frac{c}{2}}^{r+\frac{c}{2}-1} P_{exp}(r') \quad \text{---(46)}$$

donde c es un valor predeterminado para definir un espectro para calcular el valor medio. La potencia de señal pico puede calcularse mediante el método descrito más arriba, o puede calcularse con un método diferente.

Por lo menos una ranura de tiempo puede seleccionarse a partir de ranuras de tiempo incluidas en un ancho de tiempo t durante el cual la señal de dominio QMF de los componentes de alta frecuencia transita desde un estado fijo con una pequeña variación de su potencia de señal a un estado transitorio con una gran variación de su potencia de señal, y que es inferior a un valor predeterminado t_{th} . Por lo menos una ranura de tiempo puede ser también seleccionada a partir de ranuras de tiempo incluidas en un ancho de tiempo t durante el cual la potencia de señal de la señal de dominio QMF de los componentes de alta frecuencia se cambia de un estado transitorio con una gran variación a un estado fijo con una pequeña variación, y que son superiores al valor predeterminado t_{th} . La ranura de tiempo r en la cual $|P_{exp}(r+1)-P_{exp}(r)|$ es menor que un valor predeterminado (o igual o menor que un valor predeterminado) puede ser el estado fijo, y la ranura de tiempo r en la cual $|P_{exp}(r+1)-P_{exp}(r)|$ es igual o superior a un valor predeterminado (o mayor que un valor predeterminado) puede ser el estado transitorio. La ranura de tiempo r en la cual $|P_{exp,MA}(r+1)-P_{exp,MA}(r)|$ es menor que un valor predeterminado (o igual o menor que un valor predeterminado) puede ser el estado fijo, y la ranura de tiempo r en la cual $|P_{exp,MA}(r+1)-P_{exp,MA}(r)|$ es igual o superior a un valor predeterminado (o mayor que un valor predeterminado) puede ser el estado transitorio. El estado transitorio y el estado fijo pueden definirse utilizando el método descrito más arriba, o pueden definirse utilizando diferentes métodos. El método de selección de ranura de tiempo puede ser por lo menos alguno de los métodos descritos más arriba, puede incluir por lo menos un método distinto de los que se han descrito más arriba o puede ser una combinación de los mismos.

(Modificación 5 del Primer ejemplo)

Un dispositivo de codificación de habla 11c (FIG. 45) de una modificación 5 del Primer ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 11c cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla 11c como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 11c recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 11c y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 11c. El dispositivo de codificación de habla 11c incluye una unidad de selección de ranura de tiempo $1p1$, y una unidad de multiplicación de corriente de bits $1g4$ en lugar de la unidad de selección de ranura de tiempo $1p$ y la unidad de multiplicación de corriente de bits $1g$ del dispositivo de codificación de habla 11b de la modificación 4.

La unidad de selección de ranura de tiempo $1p1$ selecciona una ranura de tiempo como la unidad de selección de ranura de tiempo $1p$ descrita en la modificación 4 del Primer ejemplo, y transmite información de selección de ranura de tiempo a la unidad de multiplicación de corriente de bits $1g4$. La unidad de multiplicación de corriente de bits $1g4$ multiplica la corriente de bits codificados calculada por la unidad central de codificación de códecs $1c$, la información suplementaria de SBR calculada por la unidad de codificación de SBR $1d$ y el parámetro de intensidad de filtro calculado por la unidad de cálculo de parámetros de intensidad de filtro $1f$ como la unidad de multiplicación de corriente de bits $1g$, también multiplica la información de selección de ranura de tiempo recibida de la unidad de selección de ranura de tiempo $1p1$, y envía la corriente de bits multiplicados a través del dispositivo de comunicación del dispositivo de codificación de habla 11c. La información de selección de ranura de tiempo es información de selección de ranura de tiempo recibida por una unidad de selección de ranura de tiempo $3a1$ en un dispositivo de decodificación de habla 21b, que será descrito más adelante, y por ejemplo, puede incluirse un índice $r1$ de una ranura de tiempo a seleccionar. La información de selección de ranura de tiempo puede ser también un parámetro utilizado en el método de selección de ranura de tiempo de la unidad de selección de ranura de tiempo $3a1$. El dispositivo de decodificación de habla 21b (ver FIG. 20) de la modificación 5 del Primer ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 21b cargando y ejecutando un programa de ordenador predeterminado (como por ejemplo un programa de ordenador para realizar procesos, ilustrados en el diagrama de flujo de la FIG. 21) almacenado en una memoria integrada del dispositivo de decodificación de habla 21b como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 21b recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 21b.

Tal como se ilustra en la FIG. 20, el dispositivo de decodificación de habla 21b incluye una unidad de separación de corriente de bits $2a5$ y la unidad de selección de ranura de tiempo $3a1$ en lugar de la unidad de separación de corriente de bits $2a$ y la unidad de selección de ranura de tiempo $3a$ del dispositivo de decodificación de habla 21a de la modificación 4, y la información de selección de ranura de tiempo se suministra a la unidad de selección de ranura de tiempo $3a1$. La unidad de separación de corriente de bits $2a5$ separa la corriente de bits multiplicados en el parámetro de intensidad de filtro, la información suplementaria de SBR y la corriente de bits codificados como la unidad de separación de corriente de bits $2a$, y asimismo separa la información de selección de ranura de tiempo. La unidad de selección de ranura de tiempo $3a1$ selecciona una ranura de tiempo sobre la base de la información de

selección de ranura de tiempo transmitida desde la unidad de separación de corriente de bits 2a5 (proceso en Fase Si1). La información de ranura de tiempo es información utilizada para seleccionar una ranura de tiempo, y por ejemplo puede incluir el índice r1 de la ranura de tiempo que se va a seleccionar. La información de selección de ranura de tiempo puede ser también un parámetro, por ejemplo, utilizado en el método de selección de ranura de tiempo descrito en la modificación 4. En este caso, aunque no se ilustra, la señal de dominio QMF de los componentes de alta frecuencia generada por la unidad de generación de señal de alta frecuencia 2g puede ser suministrada a la unidad de selección de ranura de tiempo 3a1, además de la información de selección de ranura de tiempo. El parámetro puede ser también un valor predeterminado (como por ejemplo P_{exp} , T_h y tTh) utilizado para seleccionar la ranura de tiempo.

10 (Modificación 6 del Primer ejemplo)

Un dispositivo de codificación de habla 11d (que no se ilustra) de una modificación 6 del primer ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 11d cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla 11d como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 11d recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 11d y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 11d. El dispositivo de codificación de habla 11d incluye una unidad de cálculo de potencia a corto plazo 1i1, que no está ilustrada, en lugar de la unidad de cálculo de potencia a corto plazo 1i del dispositivo de codificación de habla 11a de la modificación 1, y también incluye una unidad de selección de ranura de tiempo 1p2.

La unidad de selección de ranura de tiempo 1p2 recibe una señal en el dominio QMF desde la unidad de transformación de frecuencia 1a, y selecciona una ranura de tiempo correspondiente al segmento de tiempo en el cual se realiza el proceso de cálculo de potencia a corto plazo por parte de la unidad de cálculo de potencia a corto plazo 1i. La unidad de cálculo de potencia a corto plazo 1i1 calcula la potencia a corto plazo de un segmento de tiempo correspondiente a la ranura de tiempo seleccionada sobre la base del resultado de la selección transmitido desde la unidad de selección de ranura de tiempo 1p2, como la unidad de cálculo de potencia a corto plazo 1i del dispositivo de codificación de habla 11a de la modificación 1.

30 (Modificación 7 del Primer ejemplo)

Un dispositivo de codificación de habla 11e (que no se ilustra) de una modificación 7 del Primer ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 11e cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla 11e como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 11e recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 11e y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 11e. El dispositivo de codificación de habla 11e incluye una unidad de selección de ranura de tiempo 1p3, que no está ilustrada, en lugar de la unidad de selección de ranura de tiempo 1p2 del dispositivo de codificación de habla 11d de la modificación 6. El dispositivo de codificación de habla 11e también incluye una unidad de multiplicación de corriente de bits que también recibe información de la unidad de selección de ranura de tiempo 1p3, en lugar de la unidad de multiplicación de corriente de bits 1g1. La unidad de selección de ranura de tiempo 1p3 selecciona una ranura de tiempo como la unidad de selección de ranura de tiempo 1p2 descrita en la modificación 6 del Primer ejemplo, y transmite la información de selección de ranura de tiempo a la unidad de multiplicación de corriente de bits.

(Modificación 8 del Primer ejemplo)

Un dispositivo de codificación de habla (que no se ilustra) de una modificación 8 del primer ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla de la modificación 8 cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla de la modificación 8 como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla de la modificación 8 recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla. El dispositivo de codificación de habla de la modificación 8 incluye también una unidad de selección de ranura de tiempo 1p además de las del dispositivo de codificación de habla descrito en la modificación 2.

Un dispositivo de decodificación de habla (que no se ilustra) de la modificación 8 del Primer ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla de la modificación 8 cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de decodificación de habla de la modificación 8 como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla de la modificación 8 recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla. El dispositivo de decodificación de habla de la modificación 8 incluye también la unidad de análisis de predicción lineal de baja frecuencia 2d1, la unidad de detección de cambio de señal 2e1, la unidad de análisis de predicción lineal de alta frecuencia 2h1, la unidad de filtro inverso de predicción lineal 2i1, y la unidad de filtro de predicción lineal 2k3 en lugar de la unidad de análisis de predicción lineal de baja frecuencia 2d, la unidad de detección de cambio de señal 2e, la unidad de análisis de predicción lineal de alta frecuencia 2h, la unidad de filtro inverso de predicción lineal 2i, y la unidad de filtro de predicción lineal 2k del dispositivo de decodificación de habla descrito en la modificación 2, y también incluye la unidad de selección de ranura de tiempo 3a.

(Modificación 9 del Primer ejemplo)

Un dispositivo de codificación de habla (que no se ilustra) de una modificación 9 del primer ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla de la modificación 9 cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla de la modificación 9 como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla de la modificación 9 recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla. El dispositivo de codificación de habla de la modificación 9 incluye la unidad de selección de ranura de tiempo 1p1 en lugar de la unidad de selección de ranura de tiempo 1p del dispositivo de codificación de habla descrito en la modificación 8. El dispositivo de codificación de habla de la modificación 9 incluye también una unidad de multiplicación de corriente de bits que recibe información de la unidad de selección de ranura de tiempo 1p1 además la información suministrada a la unidad de multiplicación de corriente de bits descrita en la modificación 8, en lugar de la unidad de multiplicación de corriente de bits descrita en la modificación 8.

Un dispositivo de decodificación de habla (que no se ilustra) de la modificación 9 del primer ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla de la modificación 9 cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de decodificación de habla de la modificación 9 como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla de la modificación 9 recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla. El dispositivo de decodificación de habla de la modificación 9 incluye la unidad de selección de ranura de tiempo 3a1 en lugar de la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación de habla descrito en la modificación 8. El dispositivo de decodificación de habla de la modificación 9 también incluye una unidad de separación de corriente de bits que separa $a_p(n, r)$ descrito en la modificación 2 en lugar del parámetro de intensidad de filtro de la unidad de separación de corriente de bits 2a5, en lugar de la unidad de separación de corriente de bits 2a.

(Modificación 1 del Segundo ejemplo)

Un dispositivo de codificación de habla 12a (FIG. 46) de una modificación 1 del segundo ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 12a cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla 12a como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 12a recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 12a y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 12a. El dispositivo de codificación de habla 12a incluye la unidad de análisis de predicción lineal 1e1 en lugar de la unidad de análisis de predicción lineal 1e del dispositivo de codificación de habla 12, y también incluye la unidad de selección de ranura de tiempo 1p.

Un dispositivo de decodificación de habla 22a (ver FIG. 22) de la modificación 1 del segundo ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 22a cargando y ejecutando un programa de ordenador predeterminado (como por ejemplo un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 23) almacenado en una memoria integrada del dispositivo de decodificación de habla 22a como la ROM en la RAM. El dispositivo de

comunicación del dispositivo de decodificación de habla 22a recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 22a. El dispositivo de decodificación de habla 22a, tal como se ilustra en la FIG. 22, incluye la unidad de análisis de predicción lineal de baja frecuencia 2d1, la unidad de detección de cambio de señal 2e1, la unidad de análisis de predicción lineal de alta frecuencia 2h1, la unidad de filtro inverso de predicción lineal 2i1, una unidad de filtro de predicción lineal 2k2, y una unidad de interpolación/extrapolación de predicción lineal 2p1, en lugar de la unidad de análisis de predicción lineal de alta frecuencia 2h, la unidad de filtro inverso de predicción lineal 2i, la unidad de filtro de predicción lineal 2k1, y la unidad de interpolación/extrapolación de predicción lineal 2p del dispositivo de decodificación de habla 22 del Segundo ejemplo, y también incluye la unidad de selección de ranura de tiempo 3a.

La unidad de selección de ranura de tiempo 3a notifica del resultado de la selección de la ranura de tiempo, a la unidad de análisis de predicción lineal de alta frecuencia 2h1, la unidad de filtro inverso de predicción lineal 2i1, la unidad de filtro de predicción lineal 2k2, y la unidad de interpolación/extrapolación de coeficiente de predicción lineal 2p1. La unidad de interpolación/extrapolación de coeficiente de predicción lineal 2p1 obtiene $a_H(n, r)$ correspondiente a la ranura de tiempo r_1 que es la ranura de tiempo seleccionada y cuyos coeficientes de predicción lineal no se transmiten por interpolación o extrapolación, como la unidad de interpolación/extrapolación de coeficiente de predicción lineal 2p, en base al resultado de la selección transmitido desde la unidad de selección de ranura de tiempo 3a (proceso en Fase S_{j1}). La unidad de filtro de predicción lineal 2k2 realiza un filtrado de síntesis de predicción lineal en la dirección de frecuencia en el resultado de $q_{adj}(n, r_1)$ desde la unidad de ajuste de alta frecuencia 2j para la ranura de tiempo r_1 seleccionada utilizando $a_H(n, r_1)$ que se interpola o extrapola y se obtiene desde la unidad de interpolación/extrapolación de coeficiente de predicción lineal 2p1, como la unidad de filtro de predicción lineal 2k1 (proceso en Fase S_{j2}), en base al resultado de selección transmitido desde la unidad de selección de ranura de tiempo 3a. Los cambios realizados en la unidad de filtro de predicción lineal 2k descritos en la modificación 3 del Primer ejemplo también pueden realizarse en la unidad de filtro de predicción lineal 2k2.

(Modificación 2 del Segundo ejemplo)

Un dispositivo de codificación de habla 12b (FIG. 47) de una modificación 2 del segundo ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 11b cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla 12b como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 12b recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 12b y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 12b. El dispositivo de codificación de habla 12b incluye la unidad de selección de ranura de tiempo 1p1, y una unidad de multiplicación de corriente de bits 1g5 en lugar de la unidad de selección de ranura de tiempo 1p y la unidad de multiplicación de corriente de bits 1g2 del dispositivo de codificación de habla 12a de la modificación 1. La unidad de multiplicación de corriente de bits 1g5 multiplica la corriente de bits codificados calculada por la unidad central de codificación de códecs 1c, la información suplementaria de SBR calculada por la unidad de codificación de SBR 1d, e indexa las ranuras de tiempo correspondientes a los coeficientes de predicción lineal cuantificados recibidos de la unidad de cuantificación de coeficientes de predicción lineal 1k como la unidad de multiplicación de corriente de bits 1g2, también multiplica la información de selección de ranura de tiempo recibida de la unidad de selección de ranura de tiempo 1p1, y envía la corriente de bits multiplicados a través del dispositivo de comunicación del dispositivo de codificación de habla 12b.

Un dispositivo de decodificación de habla 22b (ver FIG. 24) de la modificación 2 del Segundo ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 22b cargando y ejecutando un programa de ordenador predeterminado (como por ejemplo un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 25) almacenado en una memoria integrada del dispositivo de decodificación de habla 22b como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 22b recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 22b. El dispositivo de decodificación de habla 22b, tal como se ilustra en la FIG. 24, incluye una unidad de separación de corriente de bits 2a6, y la unidad de selección de ranura de tiempo 3a1 en lugar de la unidad de separación de corriente de bits 2a1 y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación de habla 22a descrito en la modificación 1, y la información de selección de ranura de tiempo es suministrada a la unidad de selección de ranura de tiempo 3a1. La unidad de separación de corriente de bits 2a6 separa la corriente de bits multiplicados en $a_H(n, r_i)$ que está siendo cuantificada, el índice r_i de la ranura de tiempo correspondiente, la información suplementaria de SBR y la corriente de bits codificados como la unidad de separación de corriente de bits 2a1, y también separa la información de selección de ranura de tiempo.

(Modificación 4 del Tercer ejemplo)

$$\overline{e(i)} \quad \text{---(47)}$$

descrito en la modificación 1 del tercer ejemplo puede ser un valor medio de $e(r)$ en el envoltorio de SBR, o puede ser un valor definido de alguna otra manera.

5 (Modificación 5 del Tercer ejemplo)

Tal como se describe en la modificación 3 del tercer ejemplo, es preferible que la unidad de ajuste de forma de envoltorio 2s controle $e_{adj}(r)$ utilizando un valor predeterminado $e_{adj,Th}(r)$, considerando que el envoltorio temporal ajustado $e_{adj}(r)$ es un coeficiente de incremento multiplicado por la muestra de subbanda de QMF, por ejemplo, como la expresión (28) y las expresiones (37) y (38).

$$e_{adj}(r) \geq e_{adj,Th} \quad \text{---(48)}$$

10

(Cuarto ejemplo)

Un dispositivo de codificación de habla 14 (FIG. 48) del cuarto ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 14 cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla 14 como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 14 recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 14 y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 14. El dispositivo de codificación de habla 14 incluye una unidad de multiplicación de corriente de bits 1g7 en lugar de la unidad de multiplicación de corriente de bits 1g del dispositivo de codificación de habla 11b de la modificación 4 del Primer ejemplo, y también incluye la unidad de cálculo de envoltorio temporal 1m y la unidad de cálculo de parámetros de envoltorio de la unidad 1n del dispositivo de codificación de habla 13.

15

20

25

30

La unidad de multiplicación de corriente de bits 1g7 multiplica la corriente de bits codificados calculada por la unidad central de codificación de códecs 1c y la información suplementaria de SBR calculada por la unidad de codificación de SBR 1d como la unidad de multiplicación de corriente de bits 1g, convierte el parámetro de intensidad de filtro calculado por la unidad de cálculo de parámetros de intensidad de filtro y el parámetro de forma de envoltorio calculado por la unidad de cálculo de parámetros de forma de envoltorio 1n en la información suplementaria de envoltorio temporal, los multiplica y envía la corriente de bits multiplicados (corriente de bits multiplicados codificados) a través del dispositivo de comunicación del dispositivo de codificación de habla 14.

(Modificación 4 del Cuarto ejemplo)

Un dispositivo de codificación de habla 14a (FIG. 49) de una modificación 4 del cuarto ejemplo incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 14a cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla 14a como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 14a recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 14a y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 14a. El dispositivo de codificación de habla 14a incluye la unidad de análisis de predicción lineal 1e1 en lugar de la unidad de análisis de predicción lineal 1e del dispositivo de codificación de habla 14 del cuarto ejemplo, y también incluye la unidad de selección de ranura de tiempo 1p.

35

40

45

Un dispositivo de decodificación de habla 24d (ver FIG. 26) de la modificación 4 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24d cargando y ejecutando un programa de ordenador predeterminado (como un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 27) almacenado en una memoria integrada del dispositivo de decodificación de habla 24d como la ROM en la RAM. El dispositivo de comunicación del dispositivo de

5 decodificación de habla 24d recibe la corriente de bits multiplicados codificados y envía una señal de
 habla decodificada hacia fuera del dispositivo de decodificación de habla 24d. El dispositivo de
 decodificación de habla 24d, tal como se ilustra en la FIG. 26, incluye la unidad de análisis de predicción
 lineal de baja frecuencia 2d1, la unidad de detección de cambio de señal 2e1, la unidad de análisis de
 predicción lineal de alta frecuencia 2h1, la unidad de filtro inverso de predicción lineal 2i1, y la unidad de
 filtro de predicción lineal 2k3 en lugar de la unidad de análisis de predicción lineal de baja frecuencia 2d,
 la unidad de detección de cambio de señal 2e, la unidad de análisis de predicción lineal de alta frecuencia
 2h, la unidad de filtro inverso de predicción lineal 2i, y la unidad de filtro de predicción lineal 2k del
 dispositivo de decodificación de habla 24, y también incluye la unidad de selección de ranura de tiempo
 3a. La unidad de formación del envoltorio temporal 2v da forma a la señal en el dominio QMF obtenido de
 la unidad de filtro de predicción lineal 2k3 utilizando la información del envoltorio temporal obtenida de la
 unidad de ajuste de forma de envoltorio 2s, como la unidad de formación del envoltorio temporal 2v del
 tercer ejemplo, la realización y sus modificaciones (proceso en Fase Sk1).

(Modificación 5 de la Realización)

15 Un dispositivo de decodificación de habla 24e (ver FIG. 28) de una modificación 5 de la realización incluye
 físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran,
 y la CPU controla de forma integral el dispositivo de decodificación de habla 24e cargando y ejecutando
 un programa de ordenador predeterminado (como un programa de ordenador para realizar procesos,
 20 ilustrado en el diagrama de flujo de la FIG. 29) almacenado en una memoria integrada del dispositivo de
 decodificación de habla 24e como la ROM en la RAM. El dispositivo de comunicación del dispositivo de
 decodificación de habla 24e recibe la corriente de bits multiplicados codificados y envía una señal de
 habla decodificada hacia fuera del dispositivo de decodificación de habla 24e. En la modificación 5, el
 dispositivo de decodificación de habla 24e, tal como se ilustra en la FIG. 28, omite la unidad de análisis de
 predicción lineal de alta frecuencia 2h1, y la unidad de filtro inverso de predicción lineal 2i1 del dispositivo
 de decodificación de habla 24d descrito en la modificación 4 que puede omitirse a lo largo de la
 25 realización como el primer ejemplo, e incluye una unidad de selección de ranura de tiempo 3a2, y una
 unidad de formación del envoltorio temporal 2v1 en lugar de la unidad de selección de ranura de tiempo
 3a y la unidad de formación del envoltorio temporal 2v del dispositivo de decodificación de habla 24d. El
 dispositivo de decodificación de habla 24e también cambia el orden del filtrado de síntesis de predicción
 lineal realizado por la unidad de filtro de predicción lineal 2k3 y el proceso de formación del envoltorio
 30 temporal realizado por la unidad de formación del envoltorio temporal 2v1 cuyo orden de proceso es
 intercambiable a lo largo de la realización.

La unidad de formación del envoltorio temporal 2v1 da forma a $q_{adj}(k, r)$ obtenido de la unidad de ajuste
 de alta frecuencia 2j utilizando $e_{adj}(r)$ obtenido de la unidad de ajuste de forma de envoltorio 2s, como la
 35 unidad de formación del envoltorio temporal 2v, y obtiene una señal $q_{envadj}(k, r)$ en el dominio QMF en la
 cual se da forma al envoltorio temporal. La unidad de formación del envoltorio temporal 2v1 también
 notifica a la unidad de selección de ranura de tiempo 3a2 sobre los parámetros obtenidos cuando se da
 forma al envoltorio temporal, o sobre los parámetros calculados por lo menos utilizando los parámetros
 obtenidos cuando se está dando forma al envoltorio temporal como la información de selección de ranura
 40 de tiempo. La información de selección de ranura de tiempo puede ser $e(r)$ de la expresión (22) o la
 expresión (40), o $|e(r)| < 2 >$ a la cual no se aplica la operación de raíz cuadrada durante el proceso de
 cálculo. También pueden utilizarse una pluralidad de secciones de ranura de tiempo (como envoltorios de
 SBR)

$$b_i \leq r < b_{i+1} \quad \text{---(49)}$$

45 así como la expresión (24) que es el valor medio de las mismas

$$\overline{e(i)}, |e(i)|^2 \quad \text{---(50)}$$

también puede utilizarse como la información de selección de ranura de tiempo. Debe reseñarse que:

$$|\overline{e(i)}|^2 = \frac{\sum_{r=b_i}^{b_{i+1}-1} |e(r)|^2}{b_{i+1} - b_i} \quad \text{---(51)}$$

La información de selección de ranura de tiempo puede ser también $e_{exp}(r)$ de la expresión (26) y la expresión (41), o $|e_{exp}(r)|^2$ a la cual no se aplica la operación de raíz cuadrada durante el proceso de cálculo. Puede utilizarse una pluralidad de segmentos de ranura de tiempo (como los envoltorios de SBR)

5 $b_i \leq r < b_{i+1} \quad \text{---(52)}$

Y el valor medio de los mismos

$$\overline{e_{exp}(i)}, |\overline{e_{exp}(i)}|^2 \quad \text{---(53)}$$

Como información de selección de ranura de tiempo. Debe indicarse que:

$$\overline{e_{exp}(i)} = \frac{\sum_{r=b_i}^{b_{i+1}-1} e_{exp}(r)}{b_{i+1} - b_i} \quad \text{---(54)}$$

$$|\overline{e_{exp}(i)}|^2 = \frac{\sum_{r=b_i}^{b_{i+1}-1} |e_{exp}(r)|^2}{b_{i+1} - b_i} \quad \text{---(55)}$$

10

La información de selección de ranura de tiempo puede ser también $e_{adj}(r)$ de la expresión (23) o la expresión (35), o la expresión (36), o puede ser $|e_{adj}(r)|^2$ a la cual no se aplica la operación de raíz cuadrada durante el proceso de cálculo. También pueden utilizarse una pluralidad de segmentos de ranura de tiempo (como envoltorios de SBR)

15 $b_i \leq r < b_{i+1} \quad \text{---(56)}$

así como el valor medio de las mismas

$$\overline{e_{adj}(i)}, |\overline{e_{adj}(i)}|^2 \quad \text{---(57)}$$

también puede utilizarse como la información de selección de ranura de tiempo. Debe reseñarse que:

$$\bar{e}_{adj}(i) = \frac{\sum_{r=b_i}^{b_{i+1}-1} e_{adj}(r)}{b_{i+1} - b_i} \quad \text{---(58)}$$

$$\left| \bar{e}_{adj}(i) \right|^2 = \frac{\sum_{r=b_i}^{b_{i+1}-1} \left| e_{adj}(r) \right|^2}{b_{i+1} - b_i} \quad \text{---(59)}$$

La información de selección de ranura de tiempo puede también ser $e_{adj, scaled}(r)$ de la expresión (37) o puede ser $\left| e_{adj, scaled}(r) \right|^2$ a la cual no se aplica la operación de raíz cuadrada durante el proceso de cálculo. En una pluralidad de segmentos de ranura de tiempo (como envoltorios de SBR)

5

$$b_i \leq r < b_{i+1} \quad \text{---(60)}$$

así como el valor medio de los mismos

$$\bar{e}_{adj, scaled}(i), \quad \left| \bar{e}_{adj, scaled}(i) \right|^2 \quad \text{---(61)}$$

también puede utilizarse como la información de selección de ranura de tiempo. Debe reseñarse que:

$$\bar{e}_{adj,scaled}(i) = \frac{\sum_{r=b_i}^{b_{i+1}-1} e_{adj,scaled}(r)}{b_{i+1} - b_i} \quad \text{---(62)}$$

$$|\bar{e}_{adj,scaled}(i)|^2 = \frac{\sum_{r=b_i}^{b_{i+1}-1} |e_{adj,scaled}(r)|^2}{b_{i+1} - b_i} \quad \text{---(63)}$$

La información de selección de ranura de tiempo puede ser también una potencia de señal $P_{envadj}(r)$ de la ranura de tiempo r de la señal de dominio de QMF correspondiente a los componentes de alta frecuencia en los cuales se da forma al envoltorio temporal o un valor de amplitud de señal de los mismos a los cuales se aplica la operación de raíz cuadrada.

5

$$\sqrt{P_{envadj}(r)} \quad \text{---(64)}$$

En una pluralidad de segmentos de ranura de tiempo (como envoltorios de SBR)

$$b_i \leq r < b_{i+1} \quad \text{---(65)}$$

así como el valor medio de los mismos

$$\bar{P}_{envadj}(i), \quad \sqrt{\bar{P}_{envadj}(i)} \quad \text{---(66)}$$

10

también puede utilizarse como la información de selección de ranura de tiempo. Debe reseñarse que:

$$P_{envadj}(r) = \sum_{k=k_x}^{k_x+M-1} |q_{envadj}(k, r)|^2 \quad \text{---(67)}$$

$$\bar{P}_{envadj}(i) = \frac{\sum_{r=b_i}^{b_{i+1}-1} P_{envadj}(r)}{b_{i+1} - b_i} \quad \text{---(68)}$$

M es un valor que representa un espectro de frecuencia superior al de la frecuencia de límite inferior k_x de los componentes de alta frecuencia generados por la unidad de generación de alta frecuencia 2g, y el intervalo de frecuencia de los componentes de alta frecuencia generados por la unidad de generación de alta frecuencia 2g también pueden estar representados como $k_x \leq k < k_x + M$.

La unidad de selección de ranura de tiempo 3a2 selecciona ranuras de tiempo en las cuales se realiza el filtrado de síntesis de predicción lineal por parte de la unidad de filtro de predicción lineal 2k, determinando si se realiza el filtrado de síntesis de predicción lineal en la señal $q_{envadj}(k, r)$ en el dominio QMF de los componentes de alta frecuencia de la ranura de tiempo r en que se da forma al envoltorio temporal por parte de la unidad de formación del envoltorio temporal 2v1, en base a la información de selección de ranura de tiempo transmitida desde la unidad de formación del envoltorio temporal 2v1 (proceso en Fase Sp1).

Para seleccionar las ranuras de tiempo en las cuales se realiza el filtrado de síntesis de predicción lineal por parte de la unidad de selección de ranura de tiempo 3a2 en la presente modificación, puede seleccionarse por lo menos una ranura de tiempo r en la cual un parámetro $u(r)$ incluido en la información de selección de ranura de tiempo transmitida desde la unidad de formación del envoltorio temporal 2v1 es mayor que un valor predeterminado u_{Th} , o en el que al menos se puede seleccionar una ranura de tiempo r en la cual $u(r)$ es igual o mayor que un valor predeterminado u_{Th} . $u(r)$ puede incluir por lo menos un elemento entre $e(r)$, $|e(r)|^2$, $e_{exp}(r)$, $|e_{exp}(r)|^2$, $e_{adj}(r)$, $|e_{adj}(r)|^2$, $e_{adj, scaled}(r)$, $|e_{adj, scaled}(r)|^2$, y $P_{envadj}(r)$, tal como se ha descrito anteriormente, y;

$$\sqrt{P_{envadj}(r)} \quad \text{---(69)}$$

Y u_{Th} puede incluir por lo menos uno entre;

$$\begin{aligned} & \overline{e(i)}, \left| \overline{e(i)} \right|^2, e_{\text{exp}}(i), \\ & \left| \overline{e_{\text{exp}}(i)} \right|^2, \overline{e_{\text{adj}}(i)}, \left| \overline{e_{\text{adj}}(i)} \right|^2 \\ & \overline{e_{\text{adj,scaled}}(i)}, \left| \overline{e_{\text{adj,scaled}}(i)} \right|^2, \\ & \overline{P_{\text{envadj}}(i)}, \sqrt{\overline{P_{\text{envadj}}(i)}}, \end{aligned} \quad \text{---(70)}$$

5 u_{Th} puede ser también un valor medio de $u(r)$ de un ancho de tiempo predeterminado (como un envoltorio de SBR) incluyendo la ranura de tiempo r . La selección puede hacerse también para incluir las ranuras de tiempo en las que $u(r)$ alcanza sus picos. Los picos de $u(r)$ pueden calcularse tal como se calculan los picos de la potencia de señal en la señal de dominio QMF de los componentes de alta frecuencia en la modificación 4 del Primer ejemplo. El estado fijo y el estado transitorio en la modificación 4 del Primer ejemplo pueden determinarse de forma parecida a los de la modificación 4 del Primer ejemplo utilizando $u(r)$, y las ranuras de tiempo pueden seleccionarse sobre esta base. El método de selección de ranura de tiempo puede ser por lo menos uno de los métodos descritos anteriormente, puede incluir por lo menos un método distinto de los descritos anteriormente, o puede ser una combinación de los mismos.

(Modificación 6 de la Realización)

15 Un dispositivo de decodificación de habla 24f (ver FIG. 30) de una modificación 6 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24f cargando y ejecutando un programa de ordenador predeterminado (como el programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 29) almacenado en una memoria integrada del dispositivo de decodificación de habla 24e como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24f recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24f. En la modificación 6, el dispositivo de decodificación de habla 24f, tal como se ilustra en la FIG. 30, omite la unidad de detección de cambio de señal 2e1, la unidad de análisis de predicción lineal de alta frecuencia 2h1, y la unidad de filtro inverso de predicción lineal 2i1 del dispositivo de decodificación de habla 24d descrito en la modificación 4 que puede omitirse a lo largo de la realización como el primer ejemplo, e incluye la unidad de selección de ranura de tiempo 3a2 y la unidad de formación del envoltorio temporal 2v1 en lugar de la unidad de selección de ranura de tiempo 3a y la unidad de formación del envoltorio temporal 2v del dispositivo de decodificación de habla 24d. El dispositivo de decodificación de habla 24f también cambia el orden del filtrado de síntesis de predicción lineal realizado por la unidad de filtro de predicción lineal 2k3 y el proceso de formación del envoltorio temporal realizado por la unidad de formación del envoltorio temporal 2v1 cuyo orden de proceso es intercambiable a lo largo de la realización.

30 La unidad de selección de ranura de tiempo 3a2 determina si se realiza el filtrado de síntesis de predicción lineal por parte de la unidad de filtro de predicción lineal 2k3 en la señal $q_{\text{envadj}}(k, r)$ en el dominio QMF de los componentes de alta frecuencia de las ranuras de tiempo r en las cuales el envoltorio temporal es formado por la unidad de formación del envoltorio temporal 2v1, sobre la base de la información de selección de ranura de tiempo transmitida desde la unidad de formación del envoltorio temporal 2v1, selecciona las ranuras de tiempo en que se realiza el filtrado de síntesis de predicción lineal, y notifica sobre las ranuras de tiempo seleccionadas a la unidad de análisis de predicción lineal de baja frecuencia 2d1 y la unidad de filtro de predicción lineal 2k3.

(Modificación 7 de la Realización)

Un dispositivo de codificación de habla 14b (FIG. 50) de una modificación 7 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de codificación de habla 14b cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de codificación de habla 14b como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación de habla 14b recibe una señal de habla que va a ser codificada desde fuera del dispositivo de codificación de habla 14b y envía una corriente de bits multiplicados codificados hacia fuera del dispositivo de codificación de habla 14b. El dispositivo de codificación de habla 14b incluye una unidad de multiplicación de corriente de bits 1g6 y la unidad de selección de ranura de tiempo 1p1 en lugar de la unidad de multiplicación de corriente de bits 1g7 y la unidad de selección de ranura de tiempo 1p del dispositivo de codificación de habla 14a de la modificación 4.

La unidad de multiplicación de corriente de bits 1g6 multiplica la corriente de bits codificados calculada por la unidad central de codificación de códecs 1c, la información suplementaria de SBR calculada por la unidad de codificación de SBR 1d y la información suplementaria de envoltorio temporal en la que se convierten el parámetro de intensidad de filtro calculado por la unidad de cálculo de parámetros de intensidad de filtro y el parámetro de forma de envoltorio calculado por la unidad de cálculo de parámetro de forma de envoltorio 1n, también multiplica la información de selección de ranura de tiempo recibida de la unidad de selección de ranura de tiempo 1p1, y envía la corriente de bits multiplicados (corriente de bits multiplicados codificados) a través del dispositivo de comunicación del dispositivo de codificación de habla 14b.

Un dispositivo de decodificación de habla 24g (ver FIG. 31) de una modificación 7 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24g cargando y ejecutando un programa de ordenador predeterminado (como el programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 32) almacenado en una memoria integrada del dispositivo de decodificación de habla 24g como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24g recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24g. El dispositivo de decodificación de habla 24g incluye una unidad de separación de corriente de bits 2a7 y la unidad de selección de ranura de tiempo 3a1 en lugar de la unidad de separación de corriente de bits 2a3 y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación de habla 2d descrito en la modificación 4.

La unidad de separación de bits 2a7 separa la corriente de bits multiplicados proporcionada a través del dispositivo de comunicación del dispositivo de decodificación de habla 24g en la información suplementaria de envoltorio temporal, la información suplementaria de SBR y la corriente de bits codificados, como la unidad de separación de corriente de bits 2a3, y también separa la información de selección de ranura de tiempo.

(Modificación 8 de la Realización)

Un dispositivo de decodificación de habla 24h (ver FIG. 33) de una modificación 8 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24h cargando y ejecutando un programa de ordenador predeterminado (como por ejemplo un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 34) almacenado en una memoria integrada del dispositivo de decodificación de habla 24h como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24h recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24h. El dispositivo de decodificación de habla 24h, tal como se ilustra en la FIG. 33, incluye la unidad de análisis de predicción lineal de baja frecuencia 2d1, la unidad de detección de cambio de señal 2e1, la unidad de análisis de predicción lineal de alta frecuencia 2h1, la unidad de filtro inverso de predicción lineal 2i1, y la unidad de filtro de predicción lineal 2k3 en lugar de la unidad de análisis de predicción lineal de baja frecuencia 2d, la unidad de detección de cambio de señal 2e, la unidad de análisis de predicción lineal de alta frecuencia 2h, la unidad de filtro inverso de predicción lineal 2i, y la unidad de filtro de predicción lineal 2k del dispositivo de decodificación de habla 24 b de la modificación 2, y también incluye la unidad de selección de ranura de tiempo 3a. La unidad primaria de ajuste de alta frecuencia 2j1 realiza al menos uno de los procesos en la fase "Ajuste de HF" en SBR en "MPEG-4 AAC", como la unidad primaria de ajuste de alta frecuencia 2j1 de la modificación 2 de la realización (proceso en Fase Sm1). La unidad secundaria de ajuste de alta frecuencia 2j2 realiza por lo menos uno de los procesos en la fase "Ajuste de HF" en SBR en "MPEG-4 AAC", como la unidad secundaria de ajuste de alta frecuencia 2j2 de la modificación 2 de la realización (proceso en Fase Sm2). Resulta preferible que el proceso realizado por la unidad secundaria

de ajuste de alta frecuencia 2j2 sea un proceso no realizado por la unidad primaria de ajuste de alta frecuencia 2j1, entre los procesos en la fase de "Ajuste de HF" en SBR en "MPEG-4 AAC".

(Modificación 9 de la Realización)

5 Un dispositivo de decodificación de habla 24i (ver FIG. 35) de la modificación 9 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24i cargando y ejecutando un programa de ordenador predeterminado (como un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 36) almacenado en una memoria integrada del dispositivo de decodificación de habla 24i como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24i recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24i. El dispositivo de decodificación de habla 24i, tal como se ilustra en la FIG. 35, omite la unidad de análisis de predicción lineal de alta frecuencia 2h1, y la unidad de filtro inverso de predicción lineal 2i1 del dispositivo de decodificación de habla 24h de la modificación 8 que puede omitirse a lo largo de la realización como el primer ejemplo, e incluye la unidad de formación del envoltorio temporal 2v1 y la unidad de selección de ranura de tiempo 3a2 en lugar de la unidad de formación del envoltorio temporal 2v y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación de habla 24h de la modificación 8. El dispositivo de decodificación de habla 24i también cambia el orden del filtrado de síntesis de predicción lineal realizado por la unidad de filtro de predicción lineal 2k3 y el proceso de formación del envoltorio temporal realizado por la unidad de formación del envoltorio temporal 2v1 cuyo orden de proceso es intercambiable a lo largo de la realización.

(Modificación 10 de la Realización)

25 Un dispositivo de decodificación de habla 24j (ver FIG. 37) de una modificación 10 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24j cargando y ejecutando un programa de ordenador predeterminado (como un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 36) almacenado en una memoria integrada del dispositivo de decodificación de habla 24j como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24j recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24j. El dispositivo de decodificación de habla 24j, tal como se ilustra en la FIG. 37, omite la unidad de detección de cambio de señal 2e1, la unidad de análisis de predicción lineal de alta frecuencia 2h1 y la unidad de filtro inverso de predicción lineal 2i1 del dispositivo de decodificación de habla 24h de la modificación 8 que puede omitirse a lo largo de la realización como el primer ejemplo, e incluye la unidad de formación del envoltorio temporal 2v1 y la unidad de selección de ranura de tiempo 3a2 en lugar de la unidad de formación del envoltorio temporal 2v y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación de habla 24h de la modificación 8. Se cambia el orden del filtrado de síntesis de predicción lineal realizado por la unidad de filtro de predicción lineal 2k3 y el proceso de formación del envoltorio temporal realizado por la unidad de formación del envoltorio temporal 2v1 cuyo orden de proceso es intercambiable a lo largo de la realización.

(Modificación 11 de la Realización)

45 Un dispositivo de decodificación de habla 24k (ver FIG. 38) de una modificación 11 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24k cargando y ejecutando un programa de ordenador predeterminado (como un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 39) almacenado en una memoria integrada del dispositivo de decodificación de habla 24k como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24k recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24k. El dispositivo de decodificación de habla 24k, tal como se ilustra en la FIG. 38, incluye la unidad de separación de corriente de bits 2a7 y la unidad de selección de ranura de tiempo 3a1 en lugar de la unidad de separación de corriente de bits 2a3 y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación de habla 24h de la modificación 8.

55 (Modificación 12 de la Realización)

Un dispositivo de decodificación de habla 24q (ver FIG. 40) de una modificación 12 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24q cargando y

5 ejecutando un programa de ordenador predeterminado (como un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 41) almacenado en una memoria integrada del dispositivo de decodificación de habla 24q como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24q recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24q. El dispositivo de decodificación de habla 24q, tal como se ilustra en la FIG. 40, incluye la unidad de análisis de predicción lineal de baja frecuencia 2d1, la unidad de detección de cambio de señal 2e1, la unidad de análisis de predicción lineal de alta frecuencia 2h1, la unidad de filtro inverso de predicción lineal 2i1, y las unidades de ajuste de componentes de señal individual 2z4, 2z5, y 2z6 (las unidades de ajuste de componentes de señal individual corresponden al medio de formación del envoltorio temporal) en lugar de la unidad de análisis de predicción lineal de baja frecuencia 2d, la unidad de detección de cambio de señal 2e, la unidad de análisis de predicción lineal de alta frecuencia 2h, la unidad de filtro inverso de predicción lineal 2i y las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 del dispositivo de decodificación de habla 24c de la modificación 3, y también incluye la unidad de selección de ranura de tiempo 3a.

10 Por lo menos una de las unidades de ajuste de componentes de señal individual 2z4, 2z5, y 2z6 realiza el proceso en la señal de dominio QMF de la ranura de tiempo seleccionada, para el componente de señal incluido en el resultado del medio primario de ajuste de alta frecuencia, como las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3, en base al resultado de selección transmitido desde la unidad de selección de ranura de tiempo 3a (proceso en Fase Sn1). Resulta preferible que el proceso que utiliza la información de selección de ranura de tiempo incluya por lo menos un proceso que incluya el filtrado de síntesis de predicción lineal en la dirección de frecuencia, entre los procesos de las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 descritas en la modificación 3 de la realización.

25 Los procesos realizados por las unidades de ajuste de componentes de señal individual 2z4, 2z5, y 2z6 pueden ser los mismos que los procesos realizados por las unidades de ajuste de componentes de señal individual 2z1, 2z2, y 2z3 descritos en la modificación 3 de la realización, pero las unidades de ajuste de componentes de señal individual 2z4, 2z5, y 2z6 pueden dar forma al envoltorio temporal de cada uno de los diversos componentes de señal incluidos en el resultado del medio primario de ajuste de alta frecuencia por distintos métodos (si todas las unidades de ajuste de componentes de señal individual 2z4, 2z5, y 2z6 no realizan un proceso en base al resultado de selección transmitido desde la unidad de selección de ranura de tiempo 3a, es el mismo de la modificación 3 de la realización de la presente invención).

30 No es necesario que todos los resultados de selección de la ranura de tiempo transmitidos a las unidades de ajuste de componentes de señal individual 2z4, 2z5, y 2z6 desde la unidad de selección de ranura de tiempo 3a sean el mismo, y pueden ser distintos, en parte o en su totalidad.

35 En la FIG. 40, el resultado de la selección de ranura de tiempo se transmite a las unidades individuales de ajuste de componentes 2z4, 2z5, y 2z6 desde una unidad de selección de ranura de tiempo 3a. Sin embargo, resulta posible incluir una pluralidad de unidades de selección de ranura de tiempo para notificar, de los diferentes resultados de la selección de ranura de tiempo, a todos o a una parte de las unidades de ajuste de componentes de señal individual 2z4, 2z5, y 2z6. En este momento, la unidad de selección de ranura de tiempo relativa a la unidad de ajuste de componentes de señal individual entre las unidades de ajuste de componentes de señal individual 2z4, 2z5, y 2z6 que realiza el proceso 4 (el proceso de multiplicar cada muestra de subbanda de QMF por el coeficiente de incremento se realiza en la señal de entrada utilizando el envoltorio temporal obtenido de la unidad de ajuste de forma de envoltorio 2s como la unidad de formación del envoltorio temporal 2v, y a continuación el filtrado de síntesis de predicción lineal en la dirección de frecuencia también se realiza en la señal de salida utilizando los coeficientes de predicción lineal recibidos de la unidad de ajuste de intensidad de filtro 2f como la unidad de filtro de predicción lineal 2k) descrito en la modificación 3 de la cuarta realización puede seleccionar la ranura de tiempo utilizando la información de selección de ranura de tiempo proporcionada por la unidad de formación del envoltorio temporal.

(Modificación 13 de la Realización)

40 Un dispositivo de decodificación de habla 24m (ver FIG. 42) de una modificación 13 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24m cargando y ejecutando un programa de ordenador predeterminado (como un programa de ordenador para realizar procesos, ilustrado en el diagrama de flujo de la FIG. 43) almacenado en una memoria integrada del dispositivo de decodificación de habla 24m como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24m recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24m. El dispositivo

de decodificación de habla 24m, tal como se ilustra en la FIG. 42, incluye la unidad de separación de corriente de bits 2a7 y la unidad de selección de ranura de tiempo 3a1 en lugar de la unidad de separación de corriente de bits 2a3 y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación de habla 24q de la modificación 12.

5 (Modificación 14 de la Realización)

Un dispositivo de decodificación de habla 24n (que no se ilustra) de una modificación 14 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24n cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de decodificación de habla 24n como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24n recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24n. El dispositivo de decodificación de habla 24n incluye funcionalmente la unidad de análisis de predicción lineal de baja frecuencia 2d1, la unidad de detección de cambio de señal 2e1, la unidad de análisis de predicción lineal de alta frecuencia 2h1, la unidad de filtro inverso de predicción lineal 2i1, y la unidad de filtro de predicción lineal 2k3 en lugar de la unidad de análisis de predicción lineal de baja frecuencia 2d, la unidad de detección de cambio de señal 2e, la unidad de análisis de predicción lineal de alta frecuencia 2h, la unidad de filtro inverso de predicción lineal 2i, y la unidad de filtro de predicción lineal 2k del dispositivo de decodificación de habla 24a de la modificación 1, y también incluye la unidad de selección de ranura de tiempo 3a.

(Modificación 15 de la Realización)

Un dispositivo de decodificación de habla 24p (que no se ilustra) de una modificación 15 de la realización incluye físicamente una CPU, una ROM, una RAM, un dispositivo de comunicación y similares, que no se ilustran, y la CPU controla de forma integral el dispositivo de decodificación de habla 24p cargando y ejecutando un programa de ordenador predeterminado almacenado en una memoria integrada del dispositivo de decodificación de habla 24p como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación de habla 24p recibe la corriente de bits multiplicados codificados y envía una señal de habla decodificada hacia fuera del dispositivo de decodificación de habla 24p. El dispositivo de decodificación de habla 24p incluye funcionalmente la unidad de selección de ranura de tiempo 3a1 en lugar de la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación de habla 24n de la modificación 14. El dispositivo de decodificación de habla 24p también incluye una unidad de separación de corriente de bits 2a8 (que no se ilustra), en lugar de la unidad de separación de corriente de bits 2a4.

La unidad de separación de corriente de bits 2a8 separa la corriente de bits multiplicados en la información suplementaria de SBR y la corriente de bits codificados como la unidad de separación de corriente de bits 2a4, así como en la información de selección de ranura de tiempo.

Aplicabilidad Industrial

La presente invención proporciona una técnica aplicable a la técnica de extensión de ancho de banda en el dominio de frecuencia representado por SBR, y para reducir la aparición de pre-eco y post-eco y mejorar la calidad subjetiva de la señal decodificada sin aumentar de forma significativa el coeficiente de bits.

Lista de Signos de Referencia

- 11, 11a, 11b, 11c, 12, 12a, 12b, 13, 14, 14a, 14b : dispositivo de codificación de habla
- 1a : unidad de transformación de frecuencia
- 1b : unidad de transformación inversa de frecuencia
- 1c : unidad central de codificación de códecs
- 1d : unidad de codificación de SBR
- 1e, 1e1 : unidad de análisis de predicción lineal
- 1f : unidad de cálculo de parámetros de intensidad de filtro
- 1f1 : unidad de cálculo de parámetros de intensidad de filtro
- 1g, 1g1, 1g2, 1g3, 1g4, 1g5, 1g6, 1g7 : unidad de multiplicación de corriente de bits
- 1h : unidad de transformación inversa de alta frecuencia
- 1i : unidad de cálculo de potencia a corto plazo
- 1j : unidad de decimación de coeficiente de predicción lineal
- 1k : unidad de cuantificación de coeficiente de predicción lineal
- 1m : unidad de cálculo de envoltorio temporal
- 1n : unidad de cálculo de parámetros de forma de envoltorio

ES 2 586 766 T3

- 1p, 1p1 : unidad de selección de ranura de tiempo
- 21, 22, 23, 24, 24b, 24c : dispositivo de decodificación de habla
- 2a, 2a1, 2a2, 2a3, 2a5, 2a6, 2a7 : unidad de separación de corriente de bits
- 2b : unidad central de decodificación de códecs
- 5 2c : unidad de transformación de frecuencia
- 2d, 2d1 : unidad de análisis de predicción lineal de baja frecuencia
- 2e, 2e1 : unidad de detección de cambio de señal
- 2f : unidad de ajuste de intensidad de filtro
- 2g : unidad de generación de alta frecuencia
- 10 2h, 2h1 : unidad de análisis de predicción lineal de alta frecuencia
- 2i, 2i1 : unidad de filtro inverso de predicción lineal
- 2j, 2j1, 2j2, 2j3, 2j4 : unidad de ajuste de alta frecuencia
- 2k, 2k1, 2k2, 2k3 : unidad de filtro de predicción lineal
- 2m : unidad de adición de coeficiente
- 15 2n : unidad de transformación inversa de frecuencia
- 2p, 2p1 : unidad de interpolación/extrapolación de coeficiente de predicción lineal
- 2r : unidad de cálculo de envoltorio temporal de baja frecuencia
- 2s : unidad de ajuste de forma de envoltorio
- 2t : unidad de cálculo de envoltorio temporal de alta frecuencia
- 20 2u : unidad de alisamiento de envoltorio temporal
- 2v, 2v1 : unidad de formación del envoltorio temporal
- 2w : unidad de conversión de información suplementaria
- 2z1, 2z2, 2z3, 2z4, 2z5, 2z6 : unidad de ajuste de componentes de señal individual
- 3a, 3a1, 3a2 : unidad de selección de ranura de tiempo

Reivindicaciones

1. Un dispositivo de decodificación de habla para decodificar una señal de habla codificada, en que el dispositivo de decodificación de habla comprende:

5 medio de separación de corriente de bits para separar una corriente de bits que incluye la señal de habla codificada en una corriente de bits codificados e información suplementaria de envoltorio temporal, en que la corriente de bits es recibida desde fuera del dispositivo de decodificación de habla;

10 medio de decodificación central para decodificar la corriente de bits codificados separada por el medio de separación de corriente de bits para obtener un componente de baja frecuencia;

15 medio de transformación de frecuencia para transformar el componente de baja frecuencia obtenido por el medio central de decodificación en un dominio de frecuencia;

20 medio de generación de alta frecuencia para generar un componente de alta frecuencia copiando el componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia de una banda de baja frecuencia a una banda de alta frecuencia;

25 medio de ajuste de alta frecuencia para ajustar el componente de alta frecuencia generado por el medio de generación de alta frecuencia para generar un componente de alta frecuencia ajustado;

30 medio de análisis de envoltorio temporal de baja frecuencia para analizar el componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia para obtener información del envoltorio temporal, en que el medio de análisis del envoltorio temporal de baja frecuencia obtiene la información de envoltorio temporal obteniendo potencia de cada muestra de subbanda de QMF del componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia;

medio de conversión de información suplementaria para convertir la información suplementaria de envoltorio temporal en un parámetro para ajustar la información del envoltorio temporal;

medio de ajuste de envoltorio temporal para ajustar la información de envoltorio temporal obtenida por el medio de análisis de envoltorio temporal de baja frecuencia para generar información del envoltorio temporal ajustada, en que el medio de ajuste de envoltorio temporal utiliza el parámetro en dicho ajuste de la información del envoltorio temporal; y

medio de formación del envoltorio temporal para formar un envoltorio temporal del componente de alta frecuencia ajustado, utilizando la información del envoltorio temporal ajustada.

2. Un dispositivo de decodificación de habla para decodificar una señal de habla codificada, en que el dispositivo de decodificación de habla comprende:

35 medio central de decodificación para decodificar una corriente de bits que incluye la señal de habla codificada para obtener un componente de baja frecuencia, en que la corriente de bits es recibida desde fuera del dispositivo de decodificación de habla;

40 medio de transformación de frecuencia para transformar el componente de baja frecuencia obtenido por el medio central de decodificación en un dominio de frecuencia;

45 medio de generación de alta frecuencia para generar un componente de alta frecuencia copiando el componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia de una banda de baja frecuencia a una banda de alta frecuencia;

50 medio de ajuste de alta frecuencia para ajustar el componente de alta frecuencia generado por el medio de generación de alta frecuencia para generar un componente de alta frecuencia ajustado;

55 medio de análisis de envoltorio temporal de baja frecuencia para analizar el componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia para obtener información del envoltorio temporal, en que el medio de análisis del envoltorio temporal de baja frecuencia obtiene la información de envoltorio temporal obteniendo potencia de cada muestra de subbanda de QMF del componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia;

medio de generación de información suplementaria de envoltorio temporal para analizar la corriente de bits para generar un parámetro para ajustar la información del envoltorio temporal;

medio de ajuste de envoltorio temporal para ajustar la información de envoltorio temporal obtenida por el medio de análisis de envoltorio temporal de baja frecuencia para generar información del envoltorio temporal ajustada, en que el medio de ajuste de envoltorio temporal utiliza el parámetro en dicho ajuste de la información del envoltorio temporal; y

medio de formación del envoltorio temporal para formar un envoltorio temporal del componente de alta frecuencia ajustado, utilizando la información del envoltorio temporal ajustada.

60 3. El dispositivo de decodificación de habla de acuerdo con la Reivindicación 1 o 2, en que el medio de análisis de envoltorio temporal de baja frecuencia obtiene la información del envoltorio temporal

normalizando la potencia de cada muestra de subbanda de QMF utilizando una potencia media en un segmento de tiempo de envoltorio de SBR.

4. Un método de decodificación de habla que utiliza un dispositivo de decodificación de habla para decodificar una señal de habla codificada, en que el método de decodificación de habla comprende:

- 5 una fase de separación de corriente de bits, en que el dispositivo de decodificación de habla separa una corriente de bits que incluye la señal de habla codificada en una corriente de bits codificados e información suplementaria del envoltorio temporal, en que la corriente de bits se recibe desde fuera del dispositivo de decodificación de habla;
- 10 una fase central de decodificación en que el dispositivo de decodificación de habla obtiene un componente de baja frecuencia decodificando la corriente de bits codificados separada en la fase de separación de corriente de bits;
- 15 una fase de transformación de frecuencia en que el dispositivo de decodificación de habla transforma el componente de baja frecuencia obtenido en la fase central de decodificación en un dominio de frecuencia;
- 20 una fase de generación de alta frecuencia en que el dispositivo de decodificación de habla genera un componente de alta frecuencia copiando el componente de baja frecuencia transformado en el dominio de frecuencia en la fase de transformación de frecuencia de una banda de baja frecuencia a una banda de alta frecuencia;
- 25 una fase de ajuste de alta frecuencia en que el dispositivo de decodificación de habla ajusta el componente de alta frecuencia generado en la fase de generación de alta frecuencia para generar un componente de alta frecuencia ajustado;
- 30 una fase de análisis de envoltorio temporal de baja frecuencia en que el dispositivo de decodificación de habla obtiene información del envoltorio temporal analizando el componente de baja frecuencia transformado en el dominio de frecuencia en la fase de transformación de frecuencia;
- 35 una fase de conversión de información suplementaria en que el dispositivo de decodificación de habla convierte la información suplementaria de envoltorio temporal en un parámetro para ajustar la información del envoltorio temporal;
- 40 una fase de ajuste de envoltorio temporal en que el dispositivo de decodificación de habla ajusta la información del envoltorio temporal obtenida en la fase de análisis de envoltorio temporal de baja frecuencia para generar información del envoltorio temporal ajustada, en que el parámetro se utiliza en dicho ajuste de la información del envoltorio temporal; y
- 45 una fase de formación del envoltorio temporal en que el dispositivo de decodificación de habla da forma a un envoltorio temporal del componente de alta frecuencia ajustado, utilizando la información del envoltorio temporal ajustada.

5. Un método de decodificación de habla que utiliza un dispositivo de decodificación de habla para decodificar una señal de habla codificada, en que el método de decodificación de habla comprende:

- 45 una fase central de decodificación, en que el dispositivo de decodificación de habla decodifica una corriente de bits que incluye la señal de habla codificada para obtener un componente de baja frecuencia, en que la corriente de bits es recibida desde fuera del dispositivo de decodificación de habla;
- 50 una fase de transformación de frecuencia en que el dispositivo de decodificación de habla transforma el componente de baja frecuencia obtenido en la fase central de decodificación en un dominio de frecuencia;
- 55 una fase de generación de alta frecuencia en que el dispositivo de decodificación de habla genera un componente de alta frecuencia copiando el componente de baja frecuencia transformado en el dominio de frecuencia en la fase de transformación de frecuencia de una banda de baja frecuencia a una banda de alta frecuencia;
- 60 una fase de ajuste de alta frecuencia en que el dispositivo de decodificación de habla ajusta el componente de alta frecuencia generado en la fase de generación de alta frecuencia para generar un componente de alta frecuencia ajustado;
- 65 una fase de análisis de envoltorio temporal de baja frecuencia en que el dispositivo de decodificación de habla obtiene información del envoltorio temporal analizando el componente de baja frecuencia transformado en el dominio de frecuencia en que la información de envoltorio temporal se obtiene obteniendo potencia de cada muestra de subbanda de QMF del componente de baja frecuencia transformado en el dominio de frecuencia en la fase de transformación de frecuencia;

una fase de generación de información suplementaria de envoltorio temporal en que el dispositivo de decodificación de habla analiza la corriente de bits para generar un parámetro para ajustar la información del envoltorio temporal;

5

una fase de ajuste de envoltorio temporal en que el dispositivo de decodificación de habla ajusta la información del envoltorio temporal obtenida en la fase de análisis de envoltorio temporal de baja frecuencia para generar información de envoltorio temporal ajustada, en que el parámetro se utiliza en dicho ajuste de la información del envoltorio temporal; y

10

una fase de formación del envoltorio temporal en que el dispositivo de decodificación de habla da forma a un envoltorio temporal del componente de alta frecuencia ajustado, utilizando la información del envoltorio temporal ajustada.

6. Un programa de decodificación de habla para decodificar una señal de habla codificada que hace que un dispositivo de ordenador funcione como:

15

medio de separación de corriente de bits para separar una corriente de bits que incluye la señal de habla codificada en una corriente de bits codificados e información suplementaria de envoltorio temporal, en que la corriente de bits es recibida desde fuera del dispositivo de decodificación de habla;

medio central de decodificación para decodificar la corriente de bits codificados separada por el medio de separación de corriente de bits para obtener un componente de baja frecuencia;

20

medio de transformación de frecuencia para transformar el componente de baja frecuencia obtenido por el medio central de decodificación en un dominio de frecuencia;

medio de generación de alta frecuencia para generar un componente de alta frecuencia copiando el componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia de una banda de baja frecuencia a una banda de alta frecuencia;

25

medio de ajuste de alta frecuencia para ajustar el componente de alta frecuencia generado por el medio de generación de alta frecuencia para generar un componente de alta frecuencia ajustado;

30

medio de análisis de envoltorio temporal de baja frecuencia para analizar el componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia para obtener información del envoltorio temporal, en que el medio de análisis del envoltorio temporal de baja frecuencia obtiene la información de envoltorio temporal obteniendo potencia de cada muestra de subbanda de QMF del componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia;

medio de conversión de información suplementaria para convertir la información suplementaria de envoltorio temporal en un parámetro para ajustar la información del envoltorio temporal;

35

medio de ajuste de envoltorio temporal para ajustar la información del envoltorio temporal obtenida por el medio de análisis de envoltorio temporal de baja frecuencia para generar información del envoltorio general ajustada, en que el medio de ajuste de envoltorio temporal utiliza el parámetro en dicho ajuste de información del envoltorio temporal; y

40

medio de formación del envoltorio temporal para dar forma a un envoltorio temporal del componente de alta frecuencia ajustado, utilizando la información del envoltorio temporal ajustada.

7. Un programa de decodificación de habla para decodificar una señal de habla codificada que hace que un dispositivo de ordenador funcione como:

45

medio central de decodificación para decodificar una corriente de bits que incluye la señal de habla codificada para obtener un componente de baja frecuencia, en que la corriente de bits es recibida desde fuera del dispositivo de decodificación de habla;

medio de transformación de frecuencia para transformar el componente de baja frecuencia obtenido por el medio central de decodificación en un dominio de frecuencia;

50

medio de generación de alta frecuencia para generar un componente de alta frecuencia copiando el componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia de una banda de baja frecuencia a una banda de alta frecuencia;

medio de ajuste de alta frecuencia para ajustar el componente de alta frecuencia generado por el medio de generación de alta frecuencia para generar un componente de alta frecuencia ajustado;

5 medio de análisis de envoltorio temporal de baja frecuencia para analizar el componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia para obtener información del envoltorio temporal, en que el medio de análisis del envoltorio temporal de baja frecuencia obtiene la información de envoltorio temporal obteniendo potencia de cada muestra de subbanda de QMF del componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia;

10 medio de generación de información suplementaria de envoltorio temporal para analizar la corriente de bits para generar un parámetro para ajustar la información del envoltorio temporal; medio de ajuste de envoltorio temporal para ajustar la información del envoltorio temporal obtenida por el medio de análisis de envoltorio temporal de baja frecuencia para generar información del envoltorio temporal ajustada, en que el medio de ajuste de envoltorio temporal utiliza el parámetro en dicho ajuste de la información del envoltorio temporal; y

15 medio de formación del envoltorio temporal para dar forma a un envoltorio temporal del componente de alta frecuencia ajustado, utilizando la información del envoltorio temporal ajustada.

Fig. 1

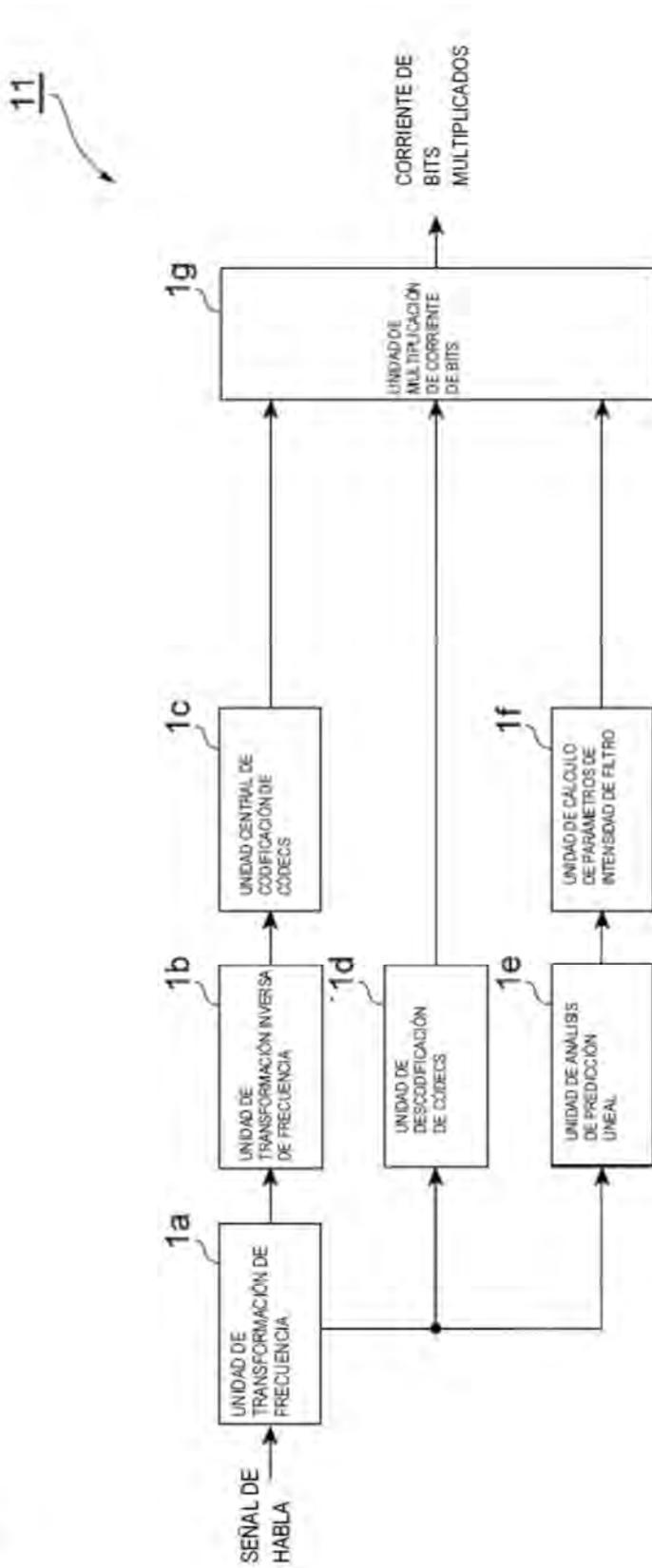


Fig. 2

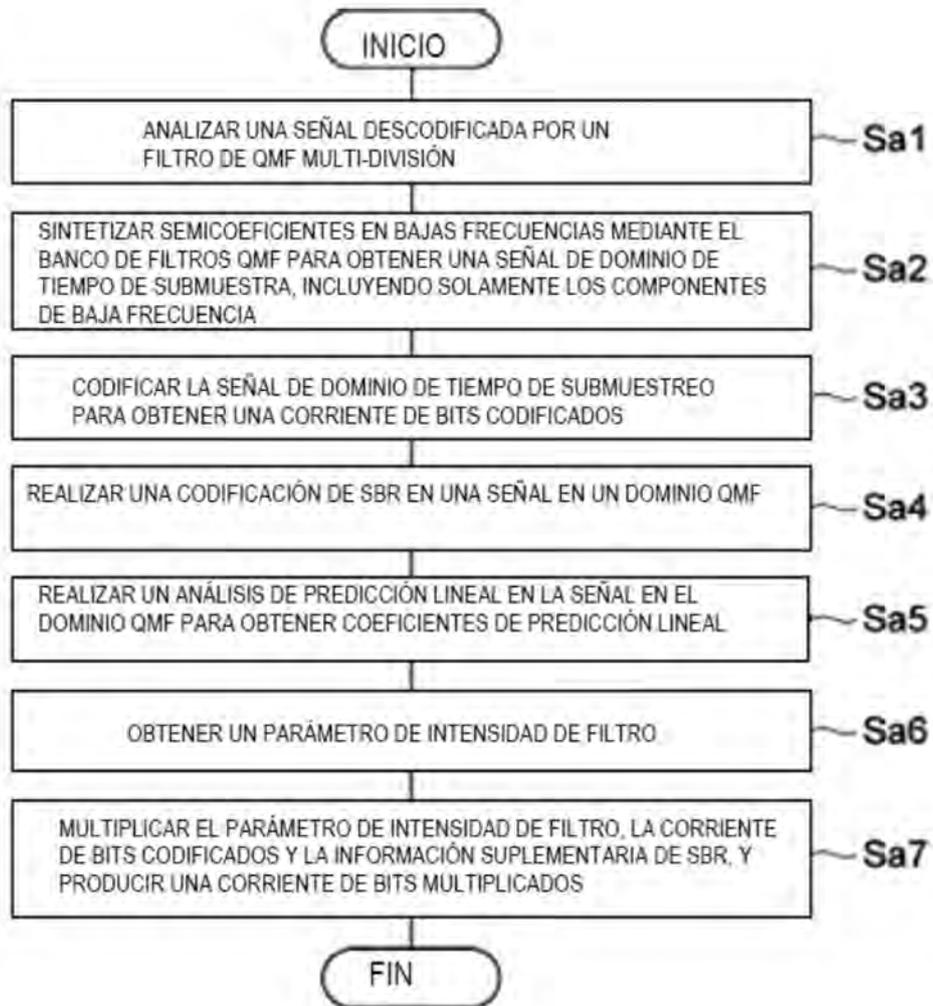
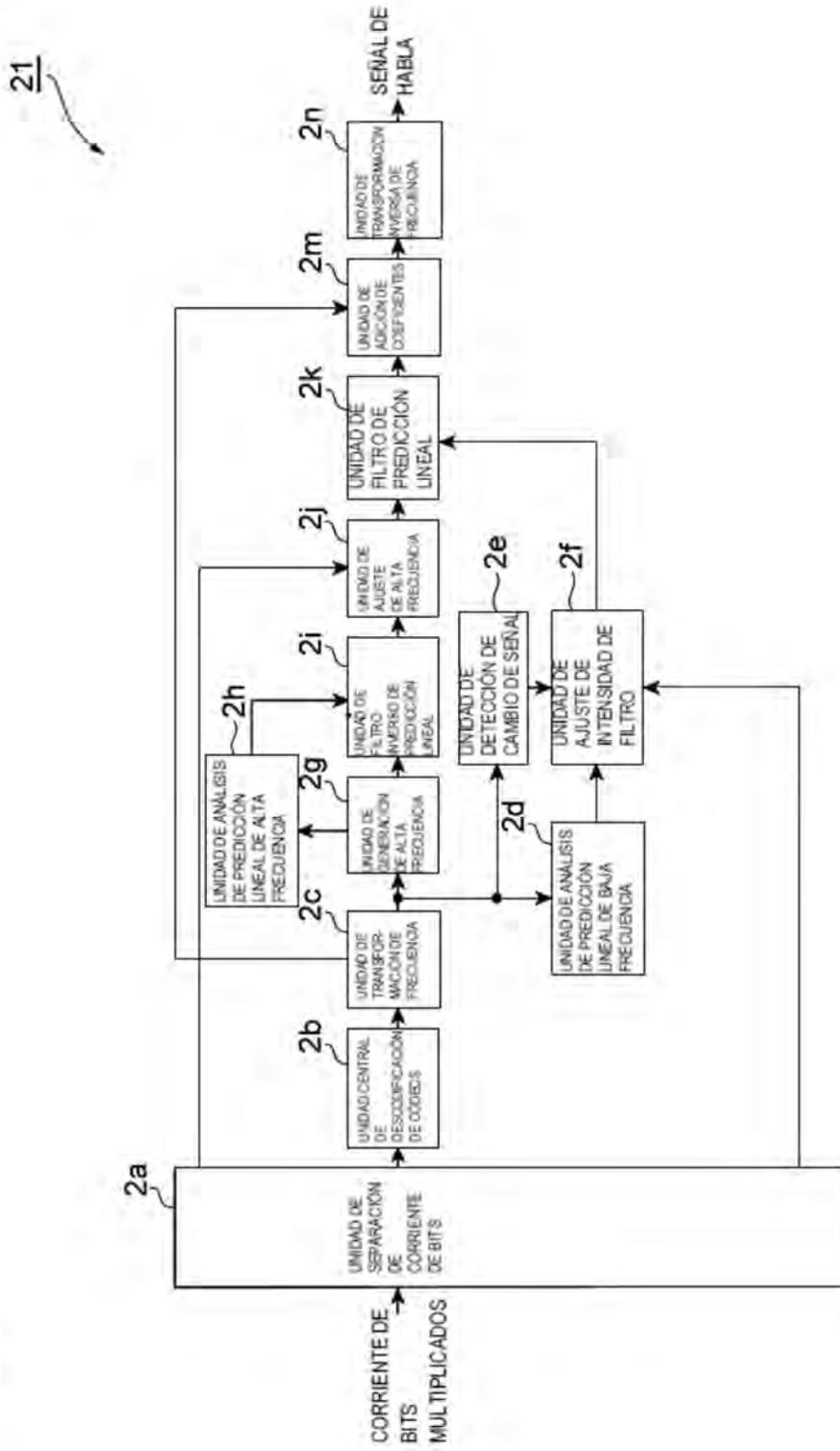


Fig. 3



21

Fig. 4



Fig. 5

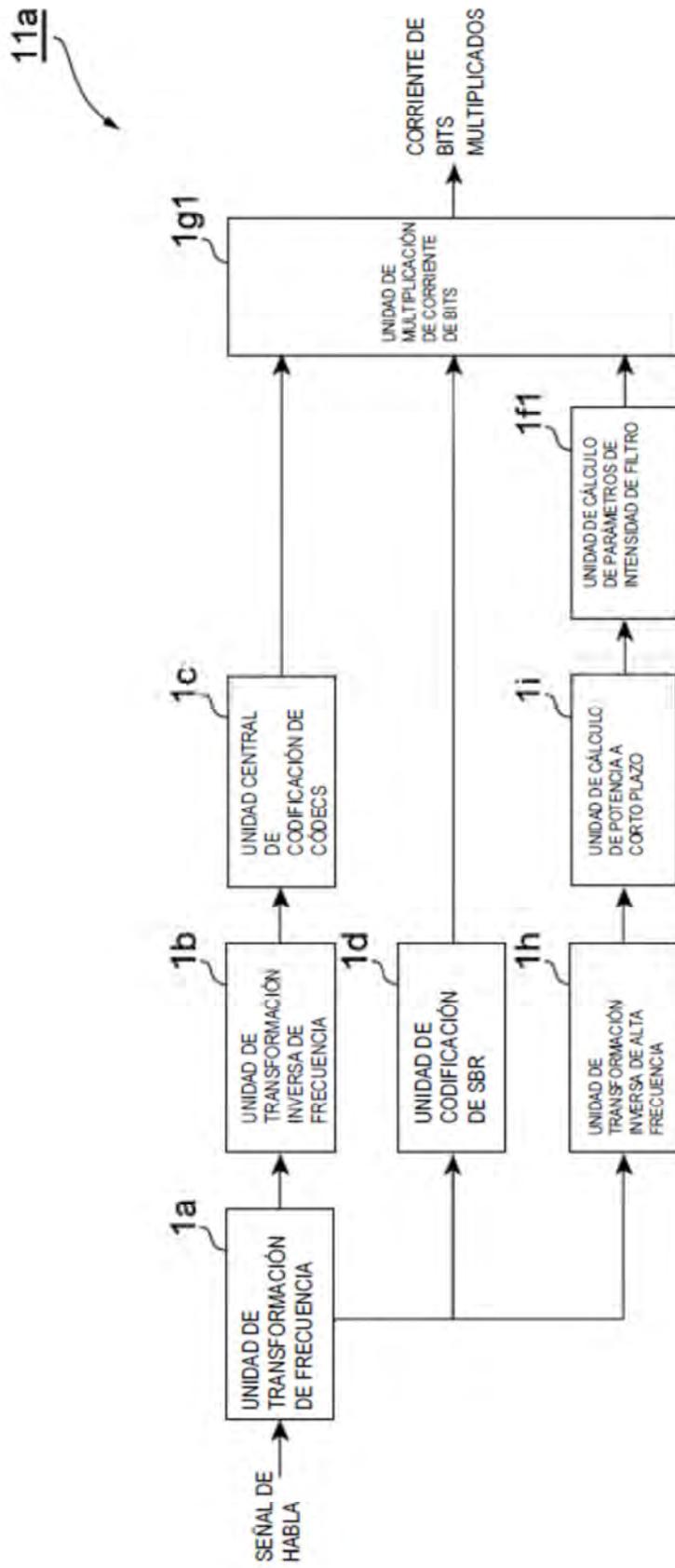


Fig. 6

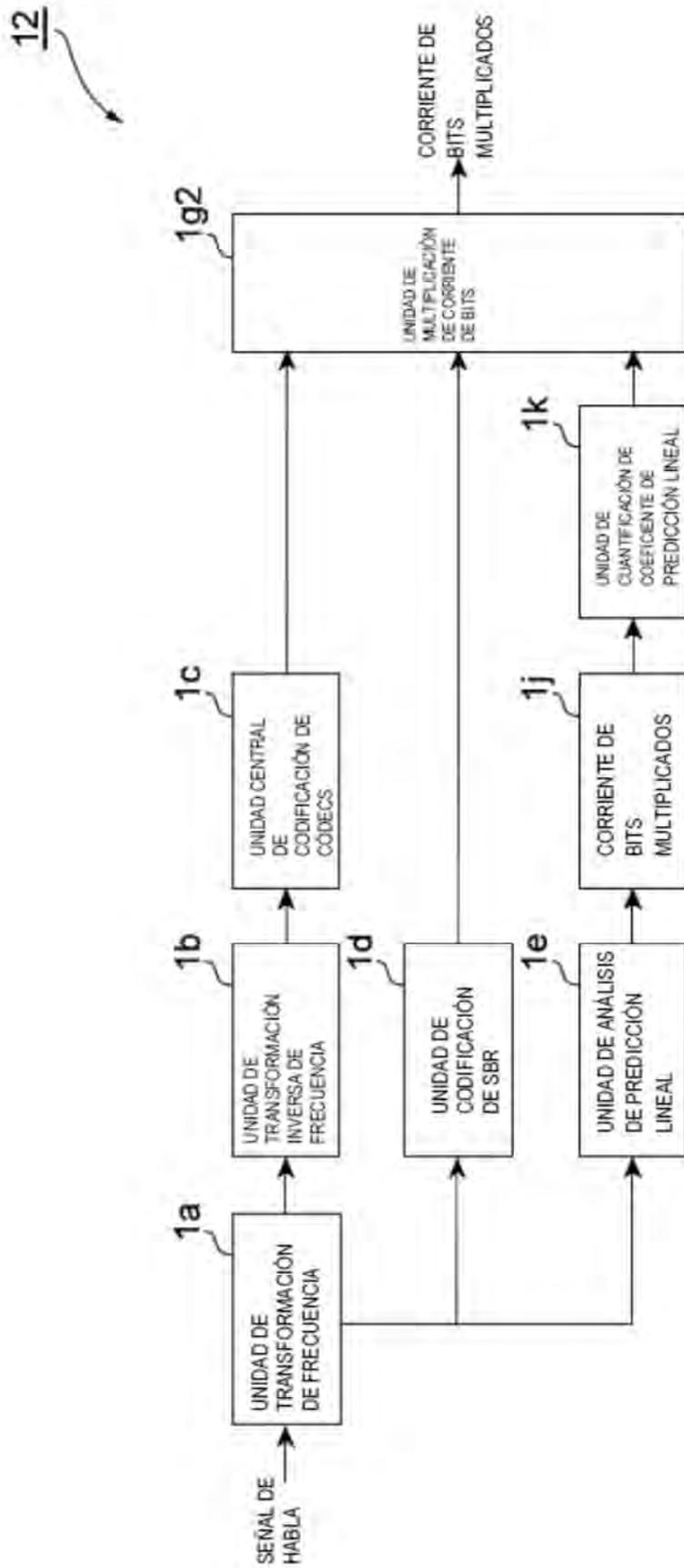


Fig. 7



Fig. 8

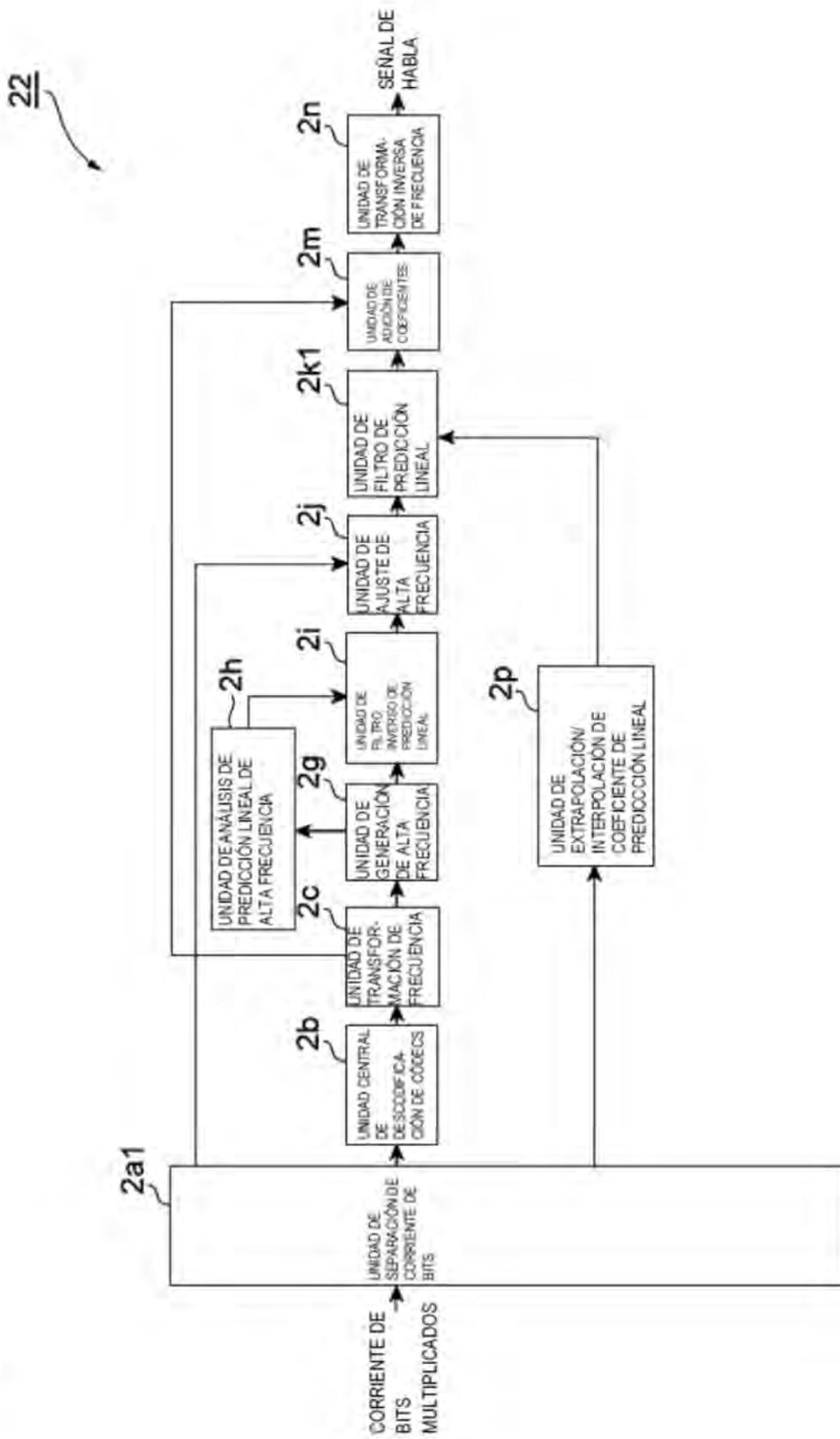


Fig 9



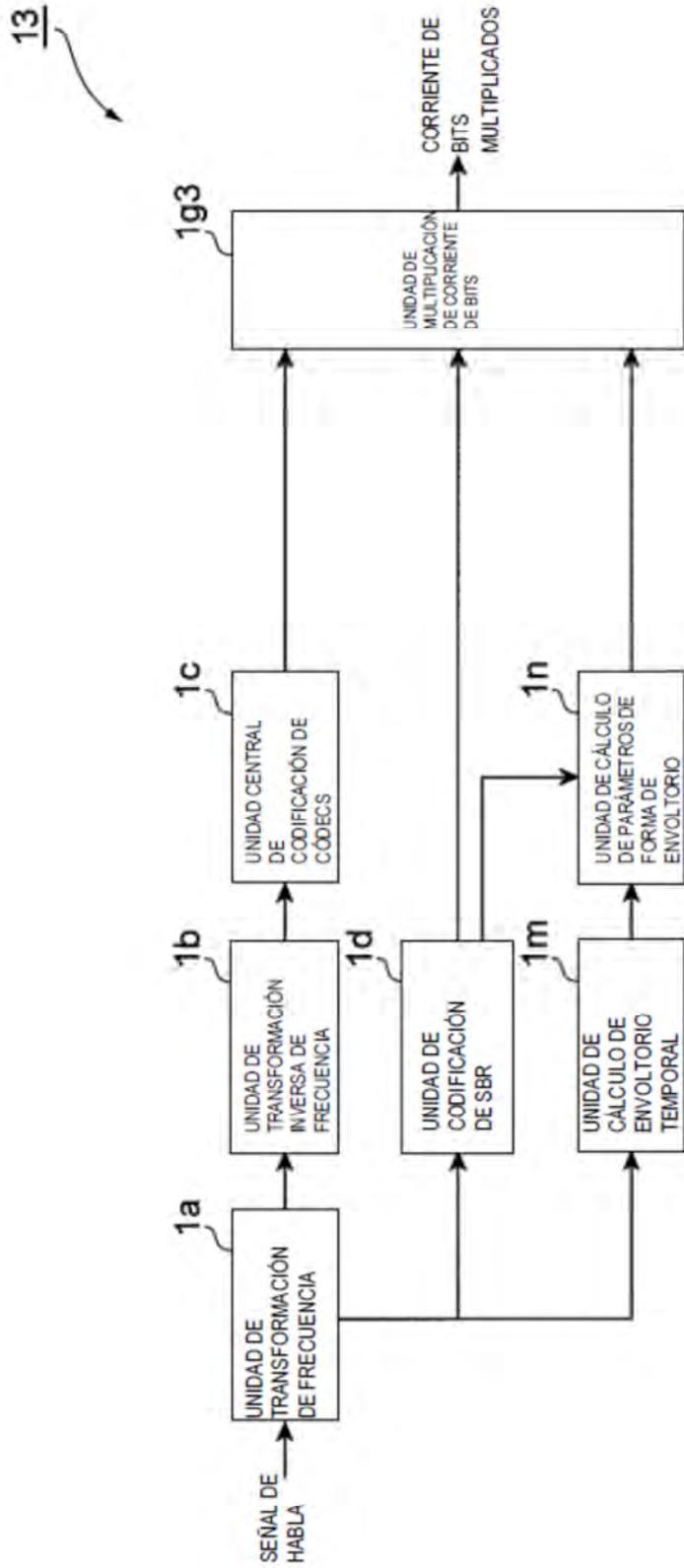


Fig. 10

Fig. 11



23

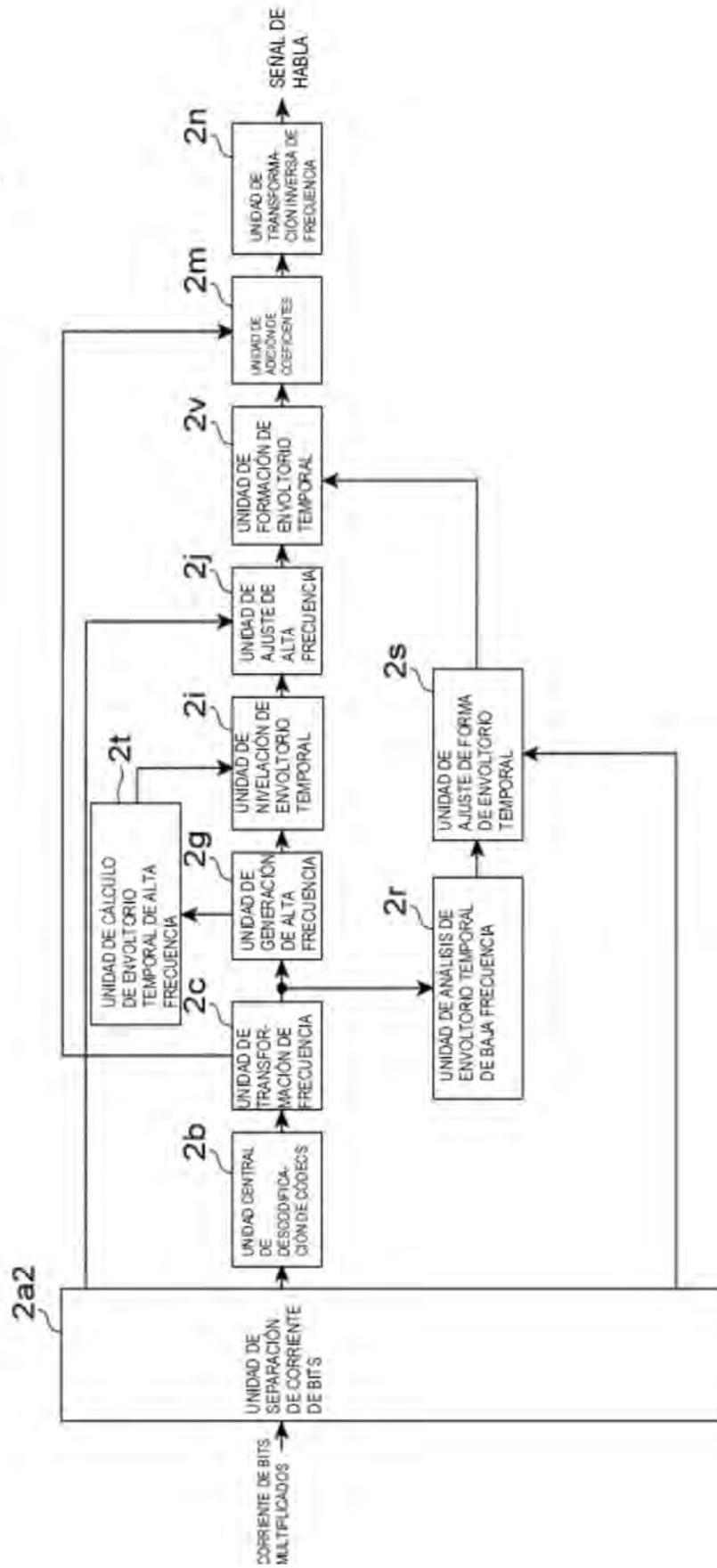


Fig. 12

Fig. 13



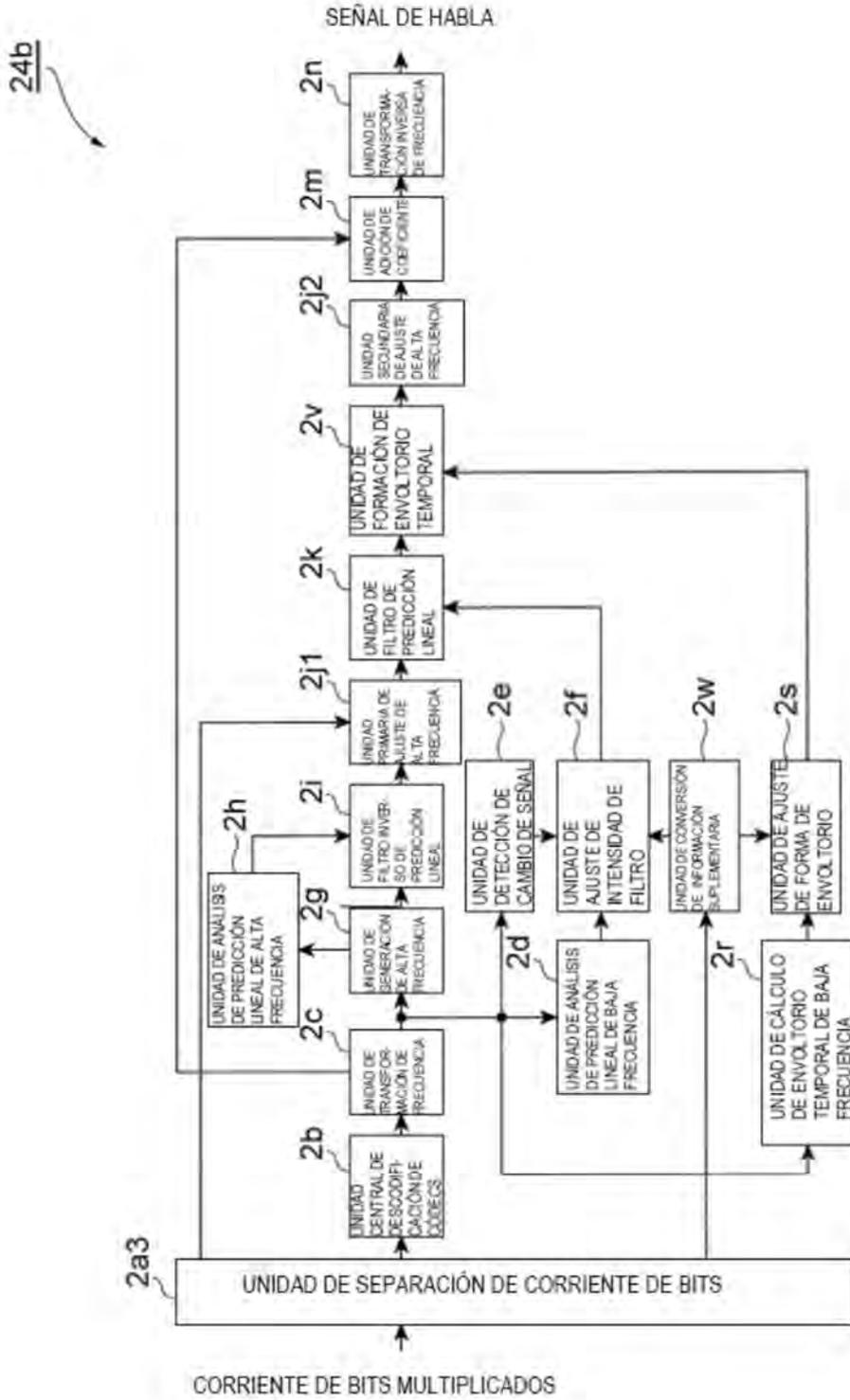


Fig. 15

Fig. 16

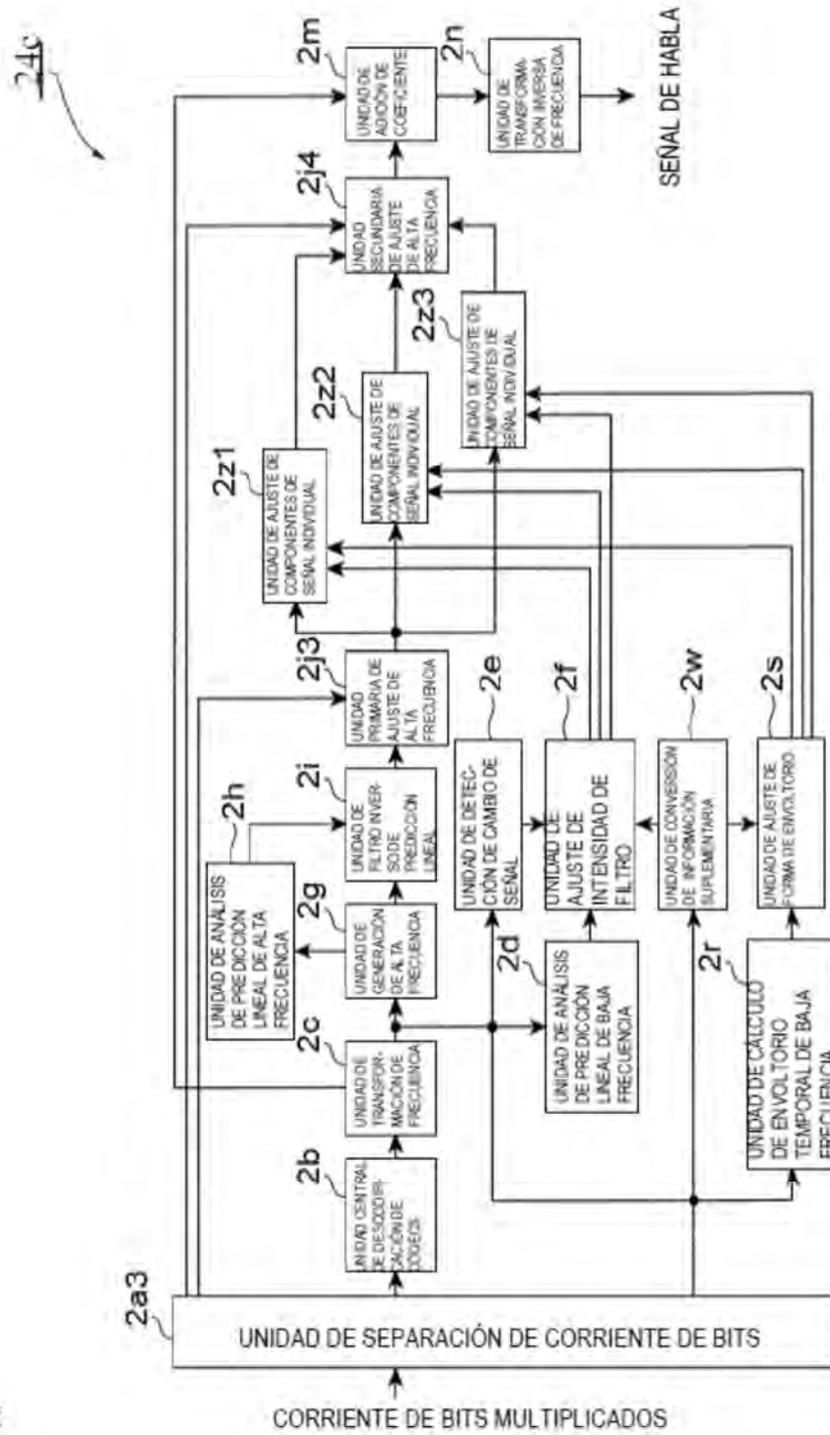


Fig. 17



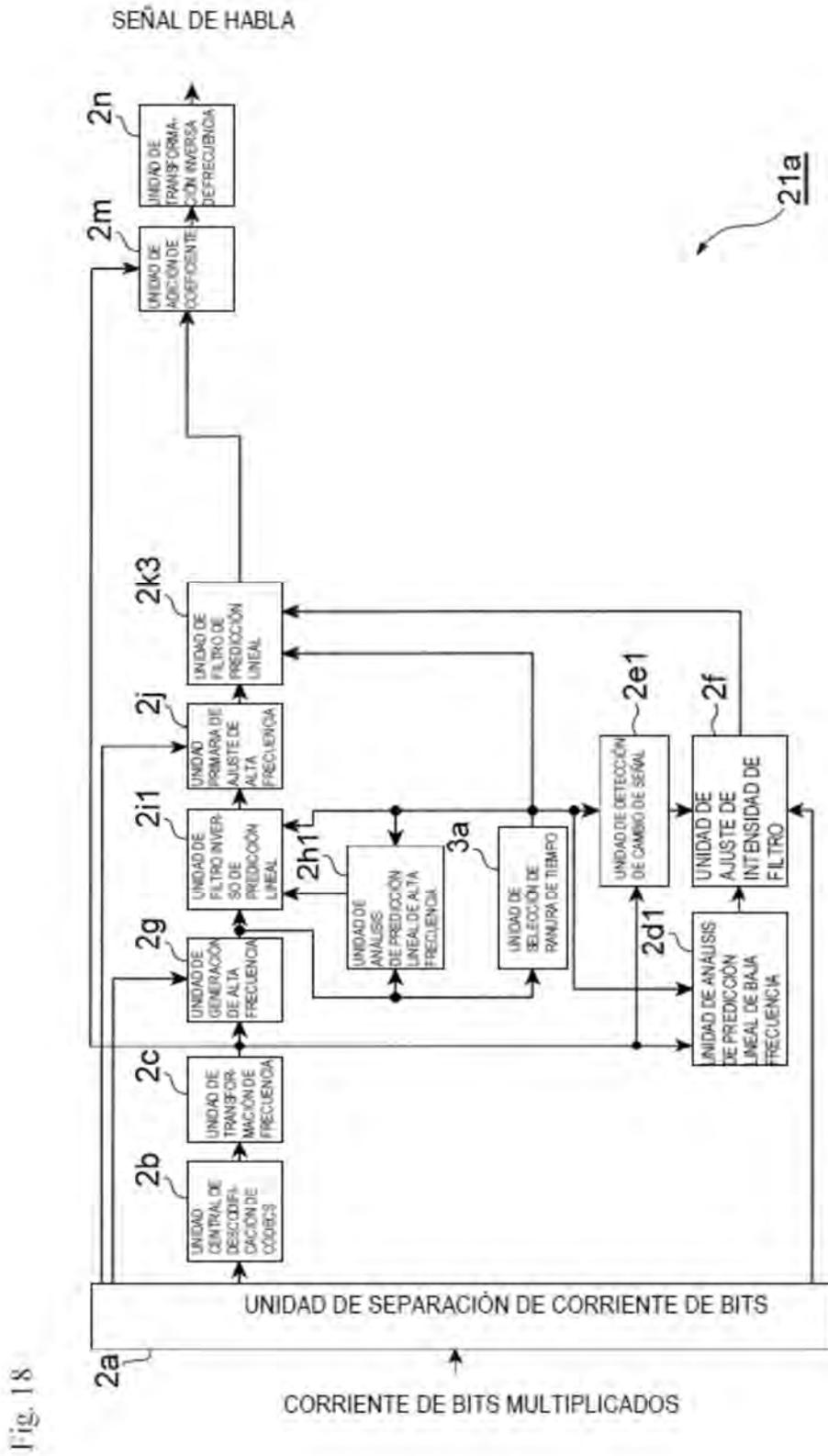


Fig. 19



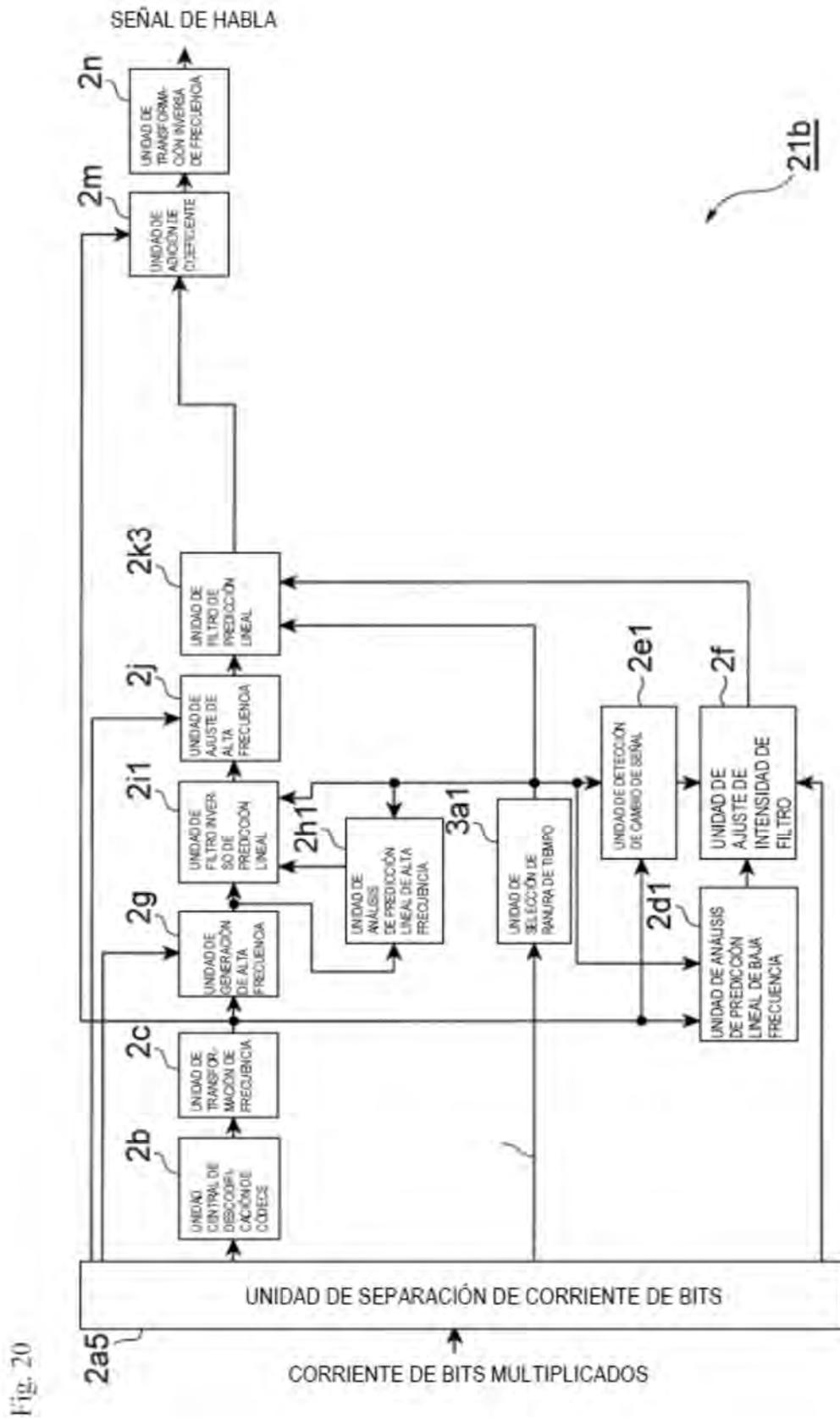


Fig. 21



Fig. 22

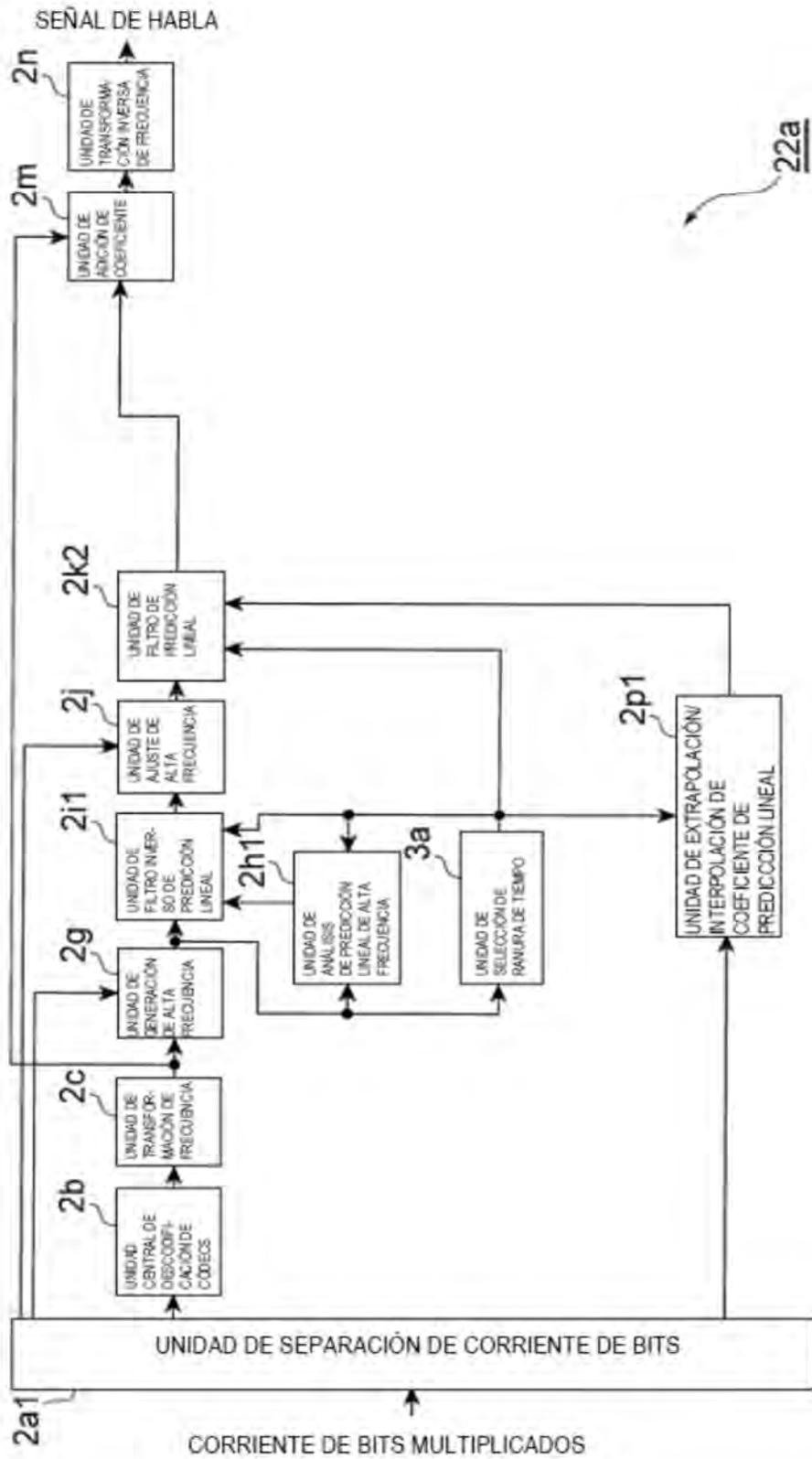


Fig. 23



Fig. 24

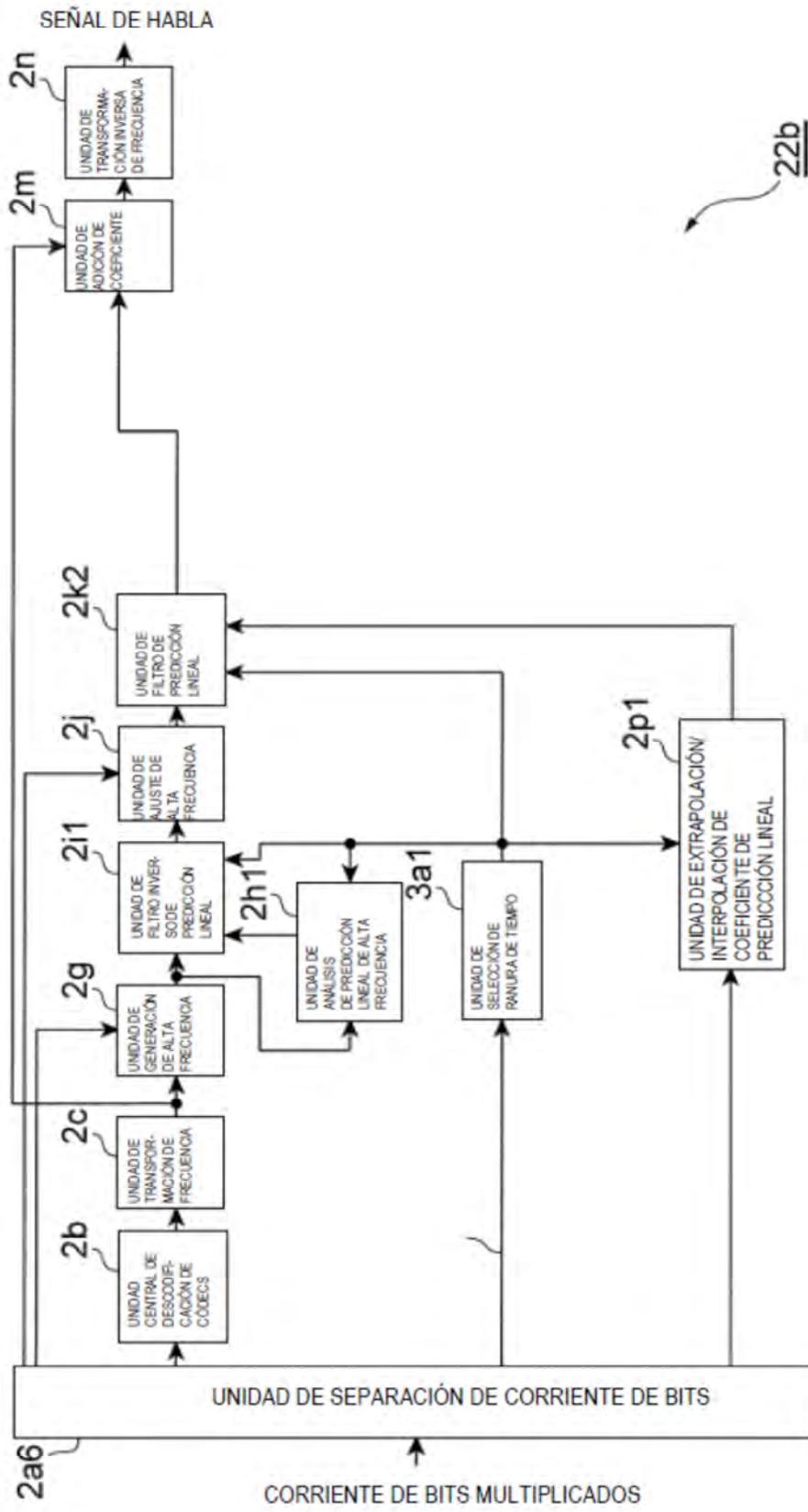


Fig. 25



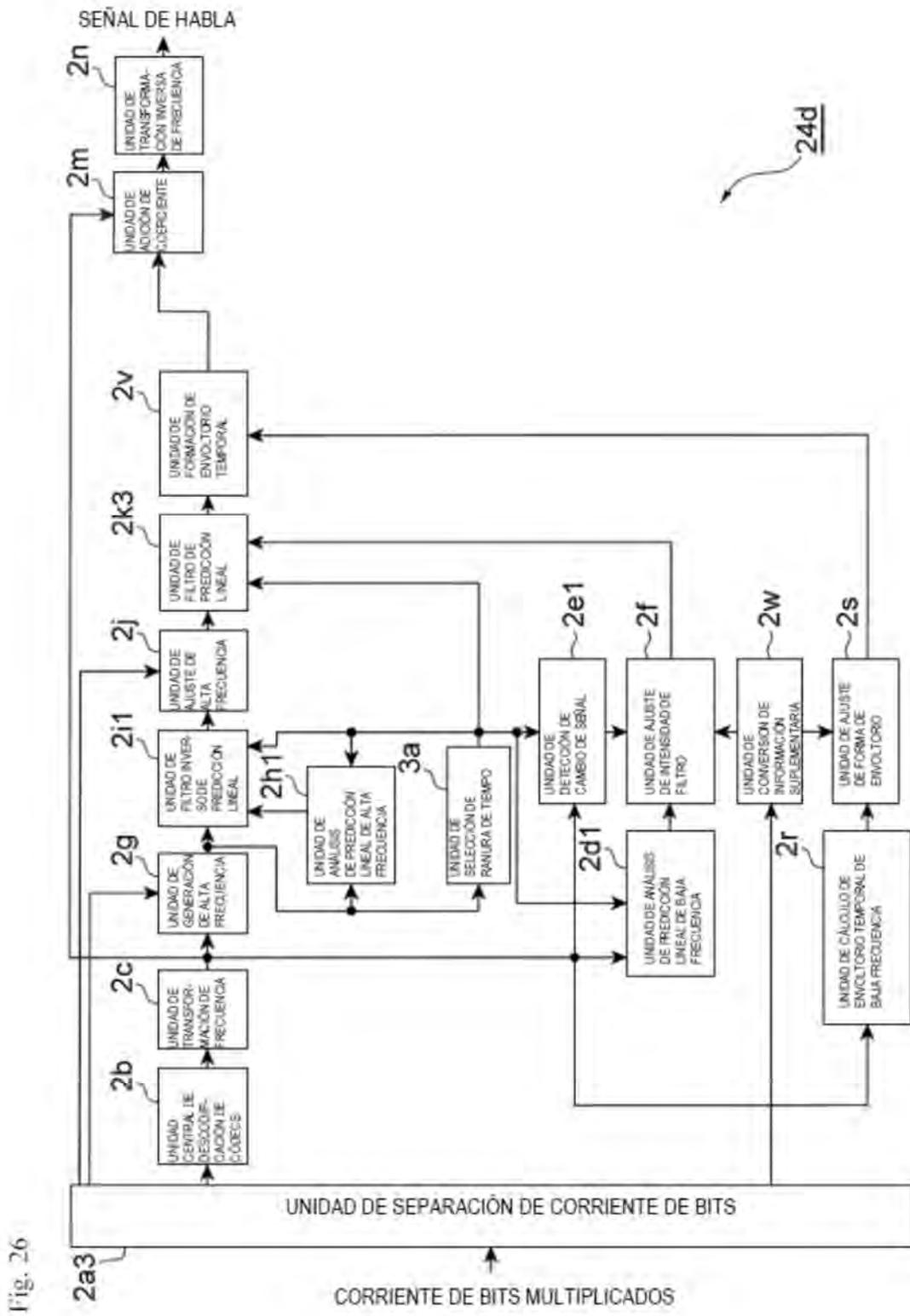
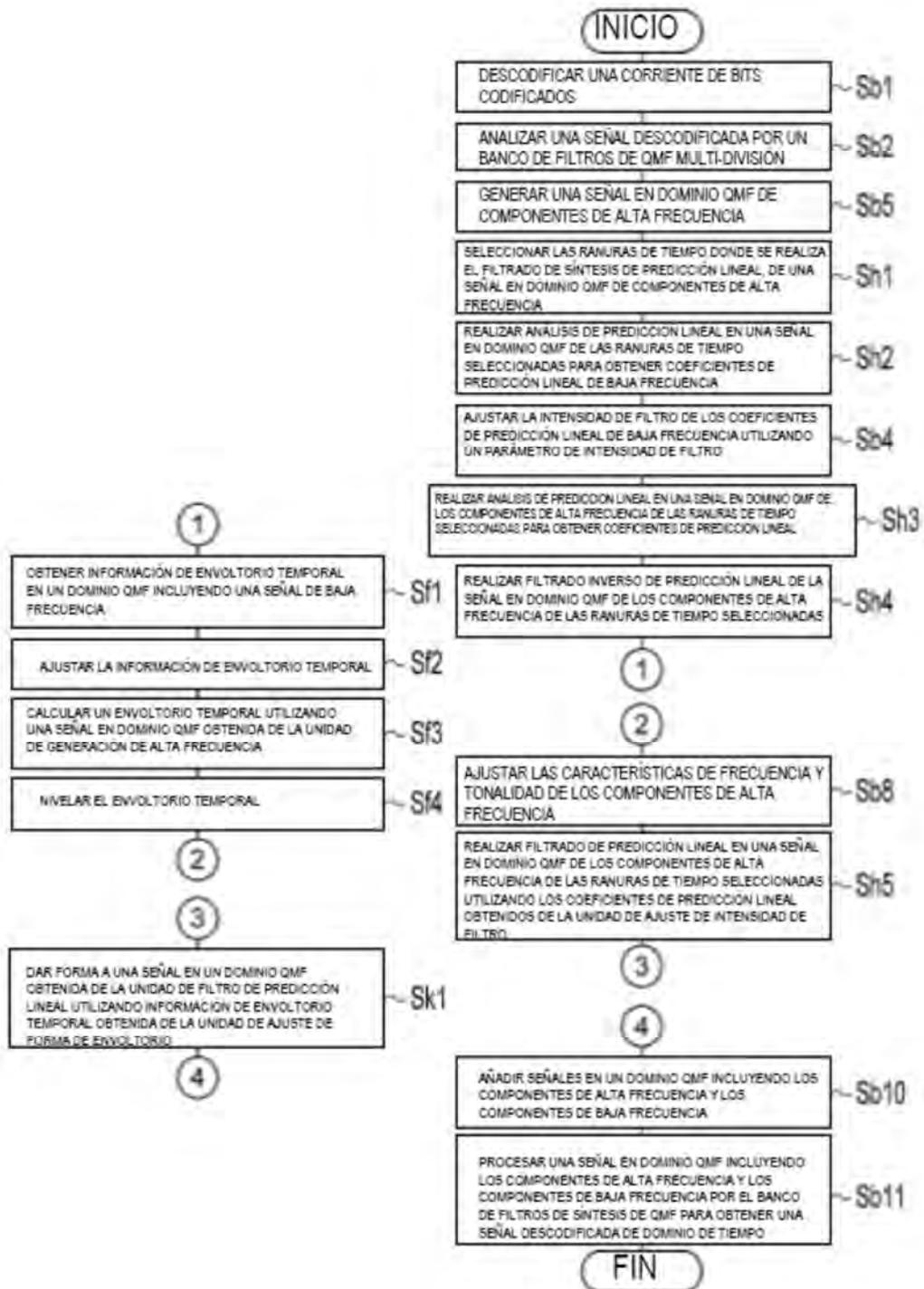


Fig. 26

Fig. 27



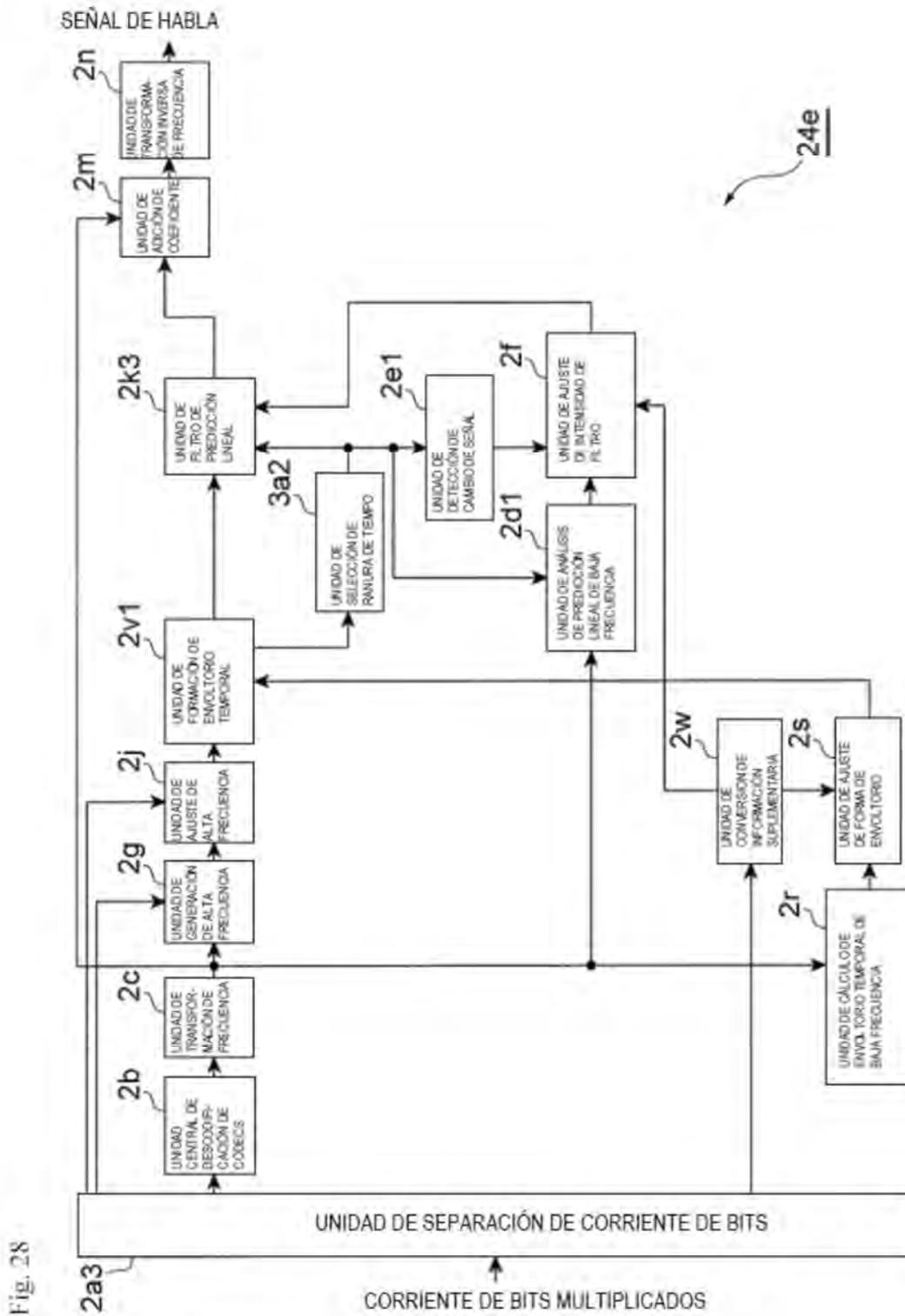
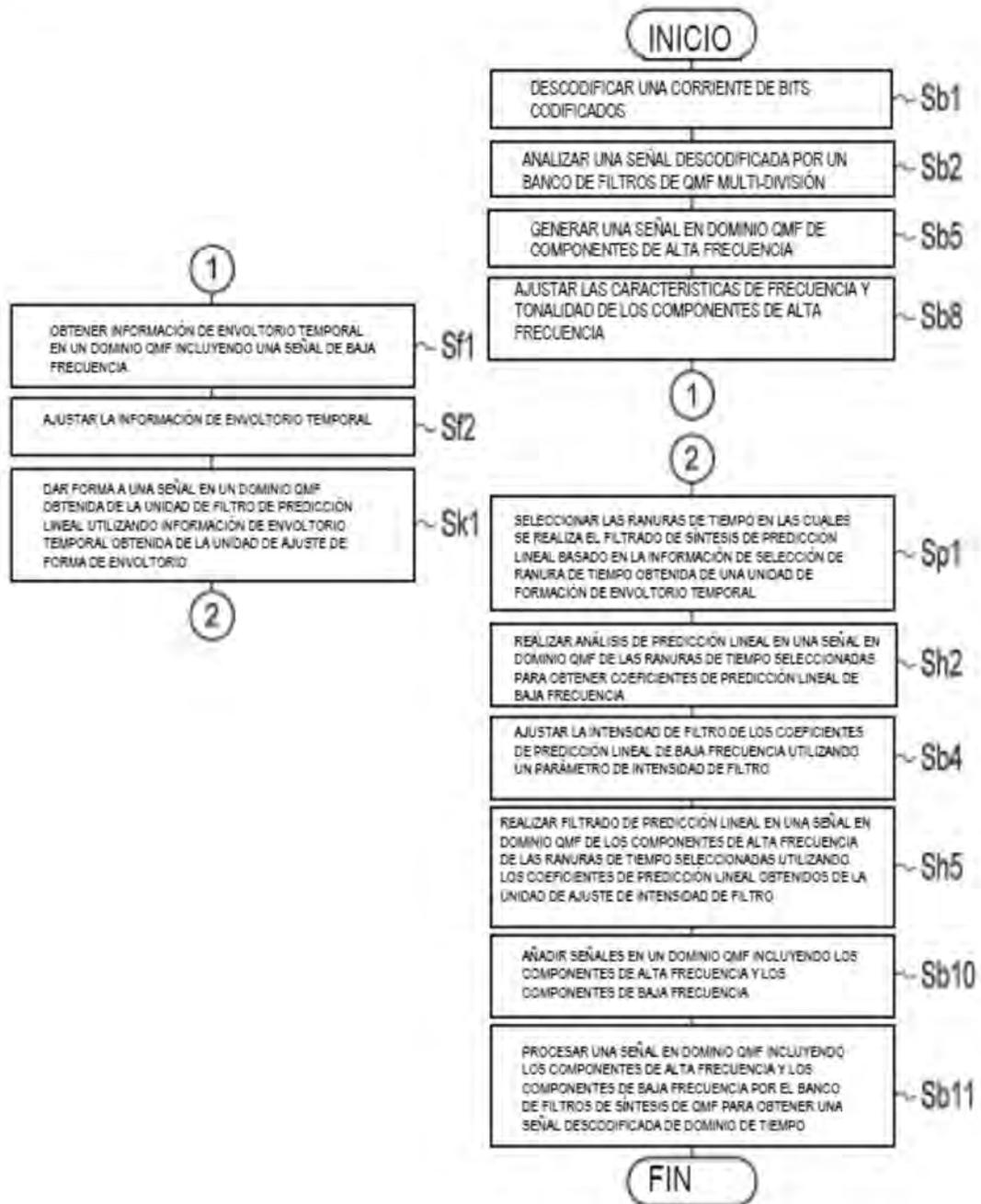


Fig. 29



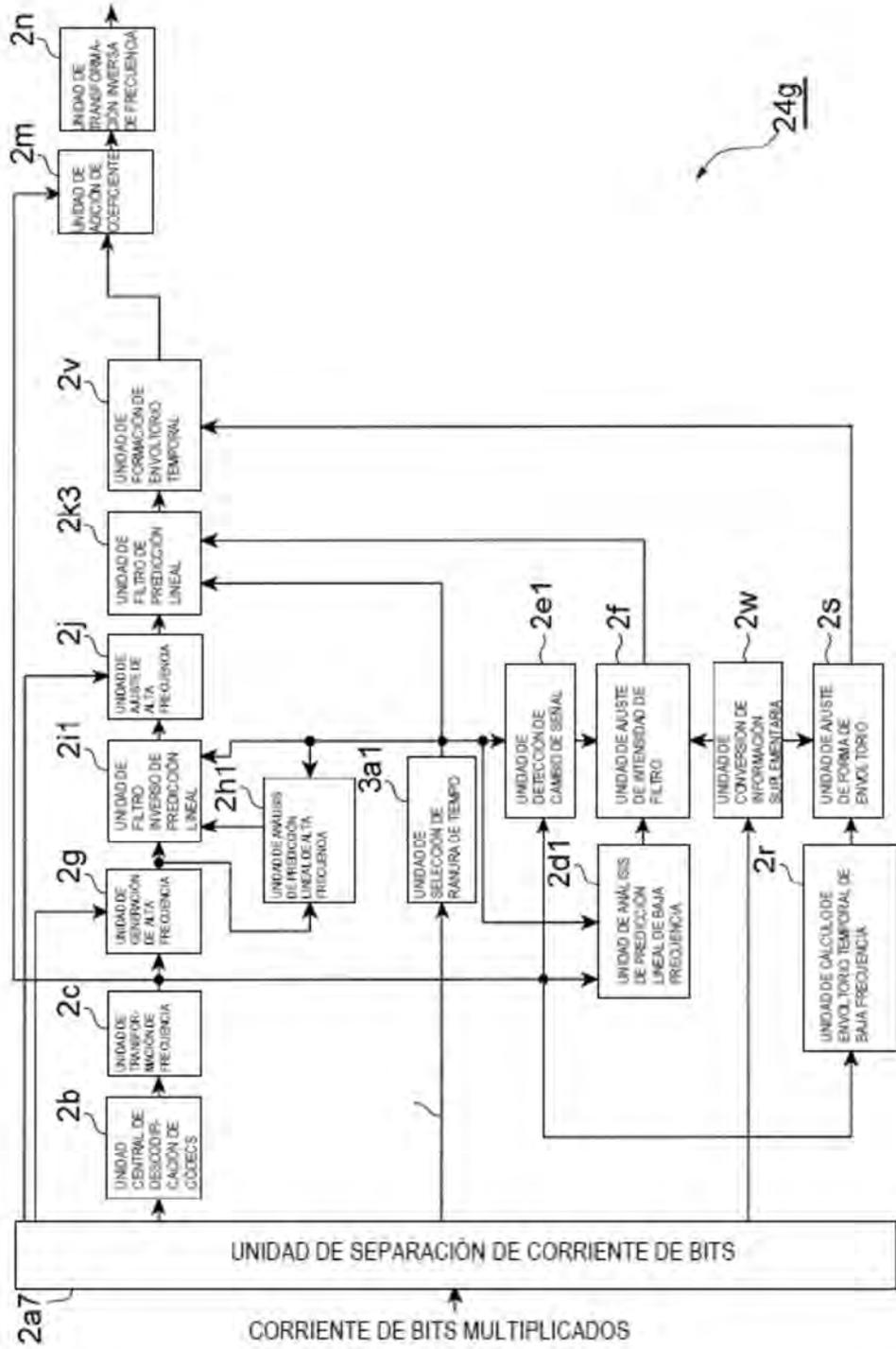
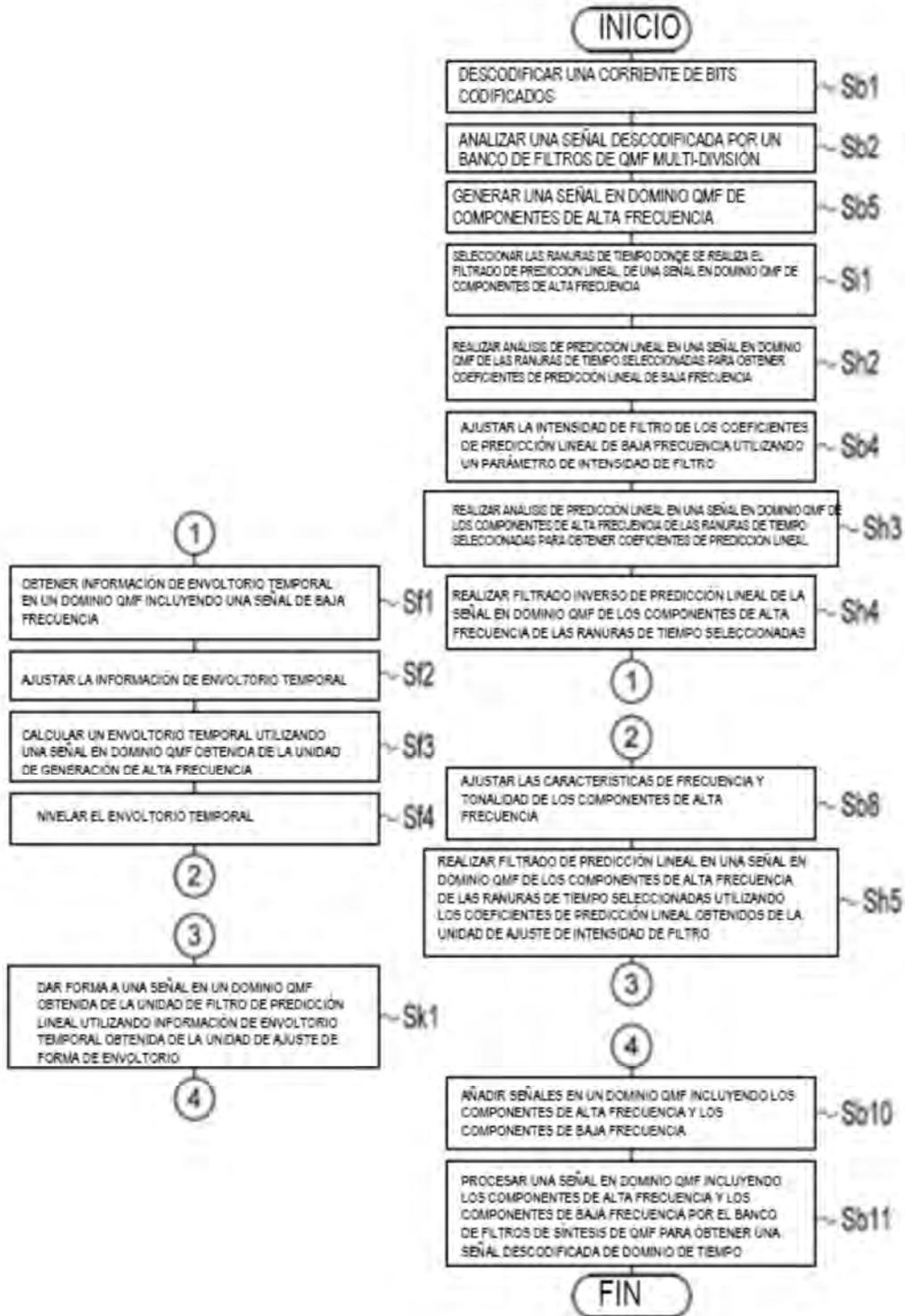


Fig. 31

Fig. 32



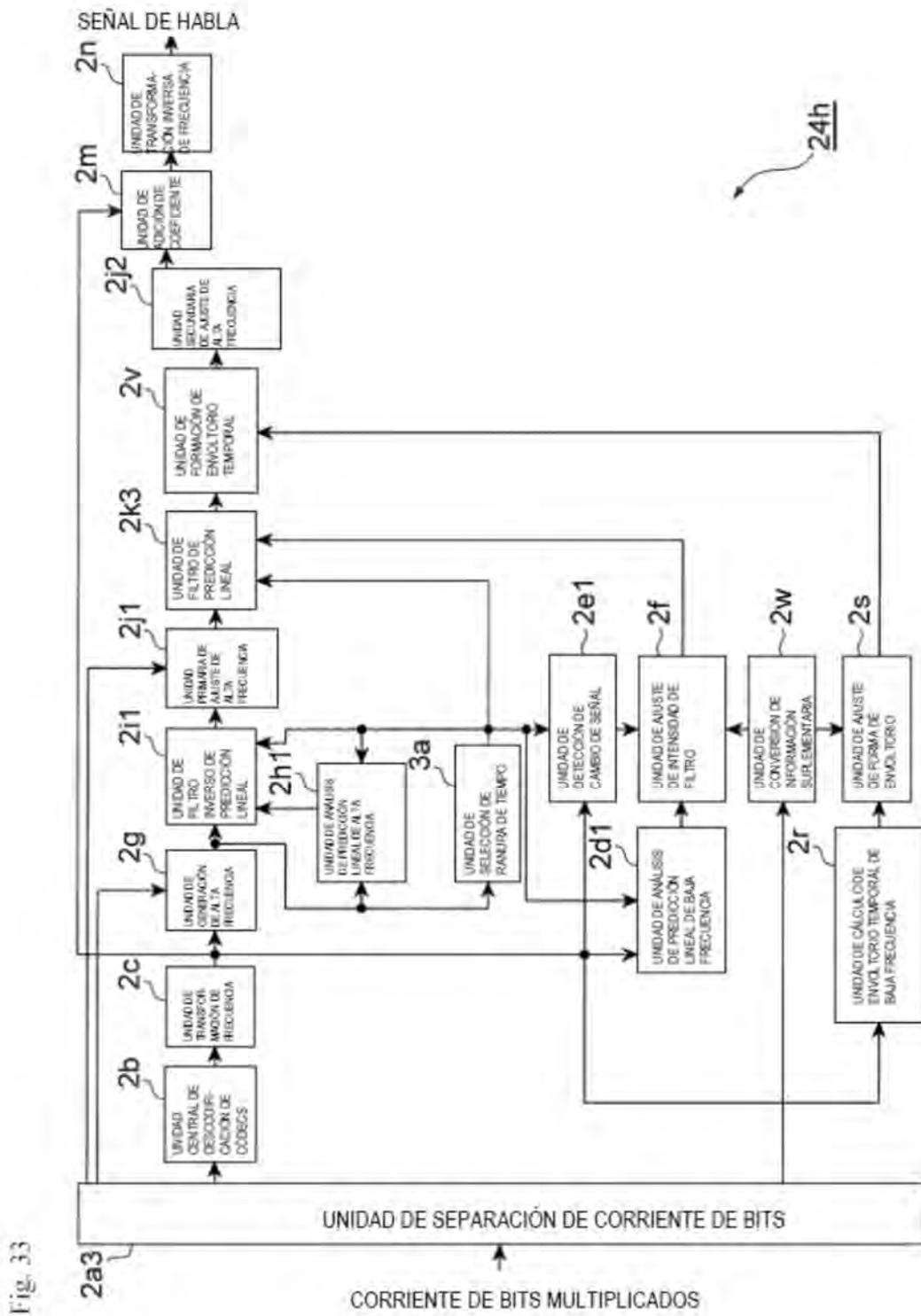
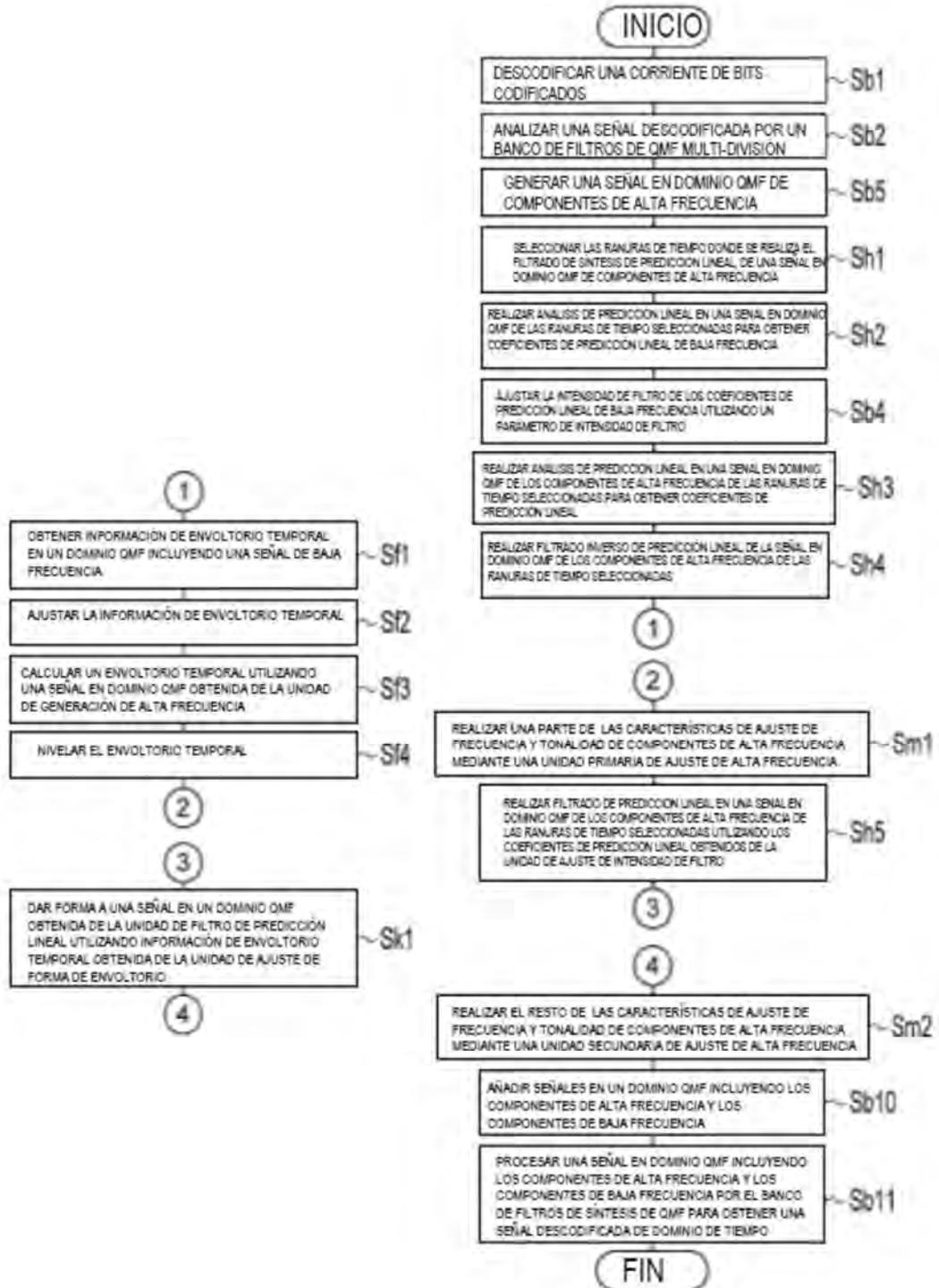


Fig. 33

Fig. 34



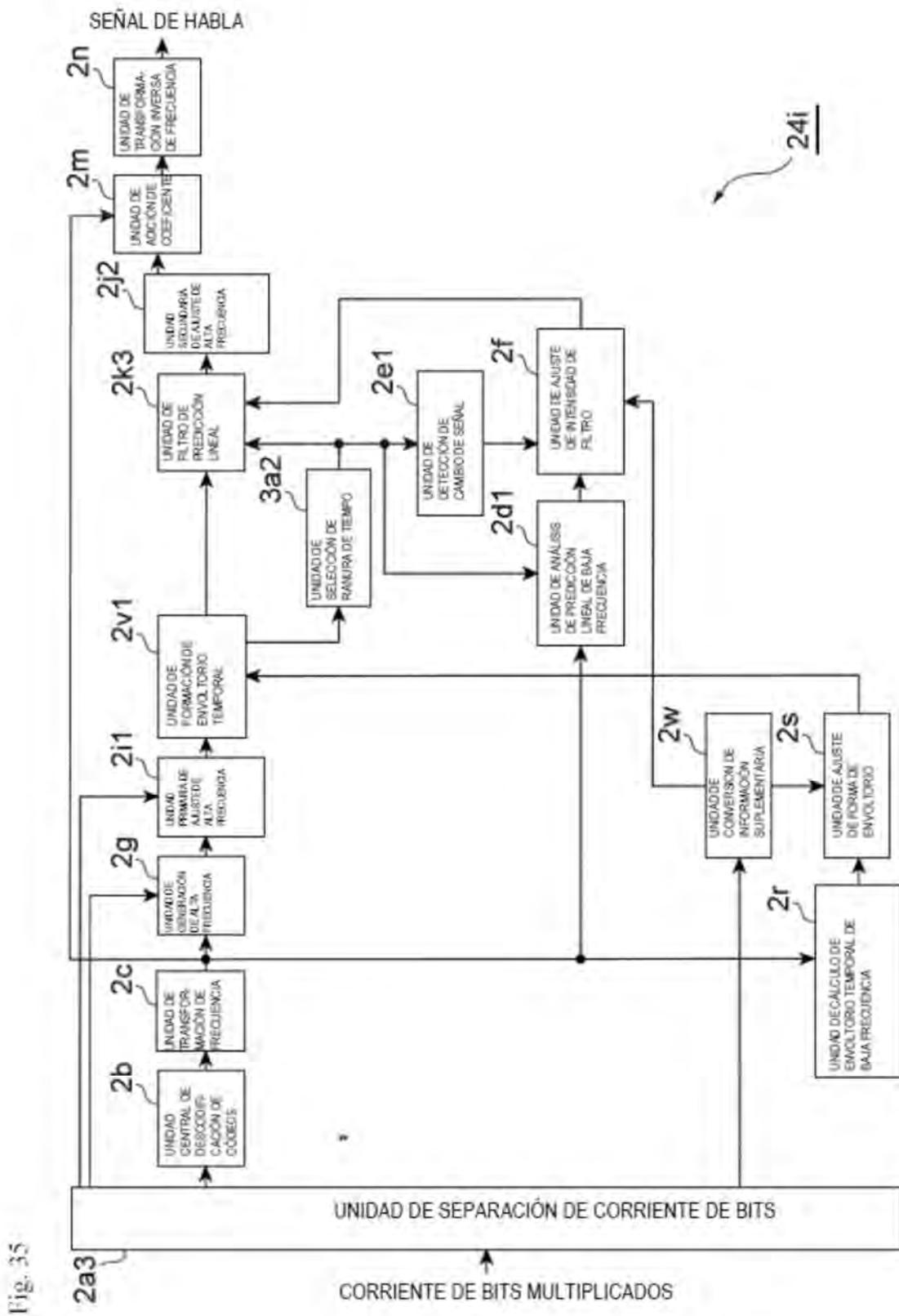
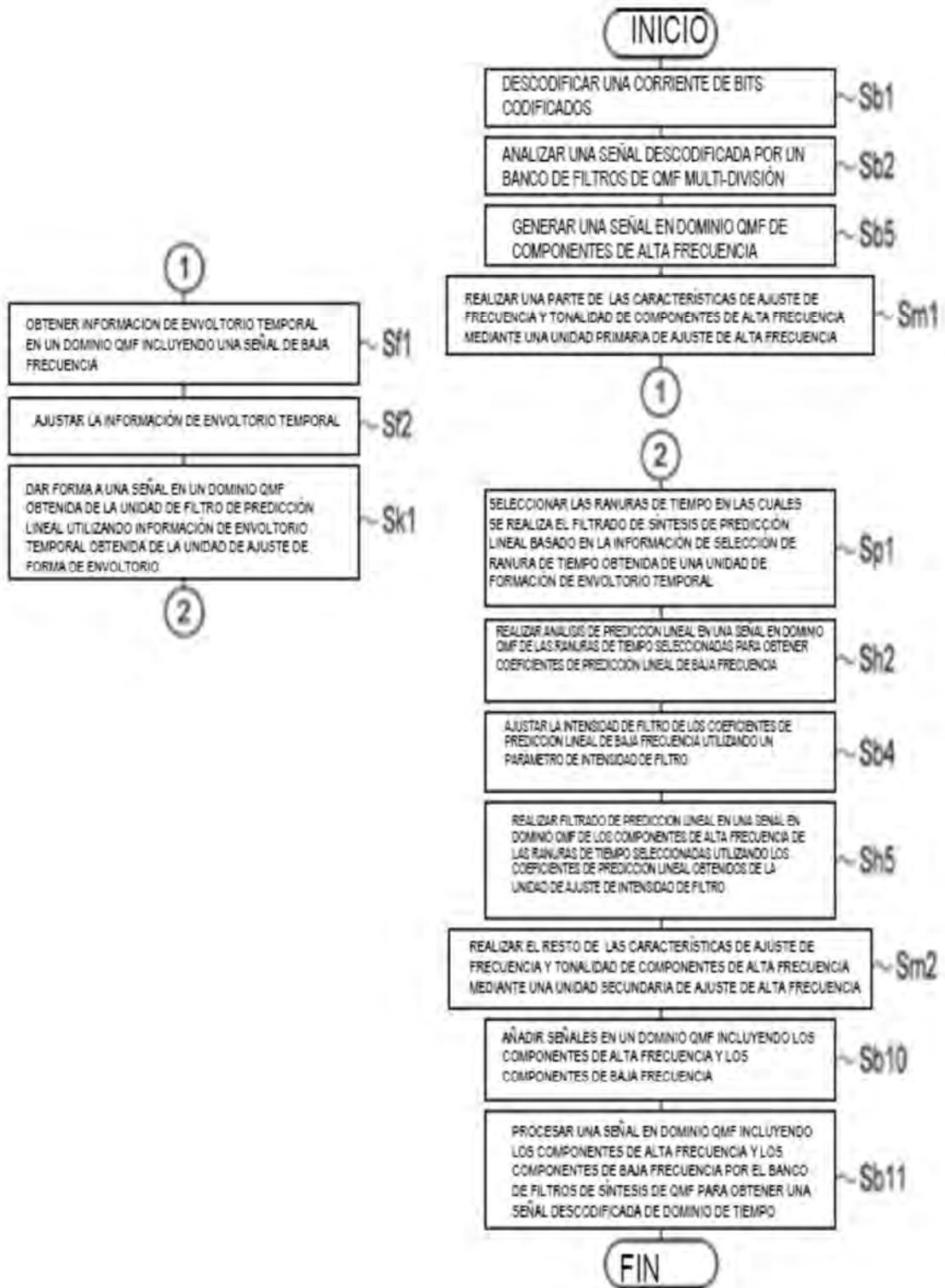
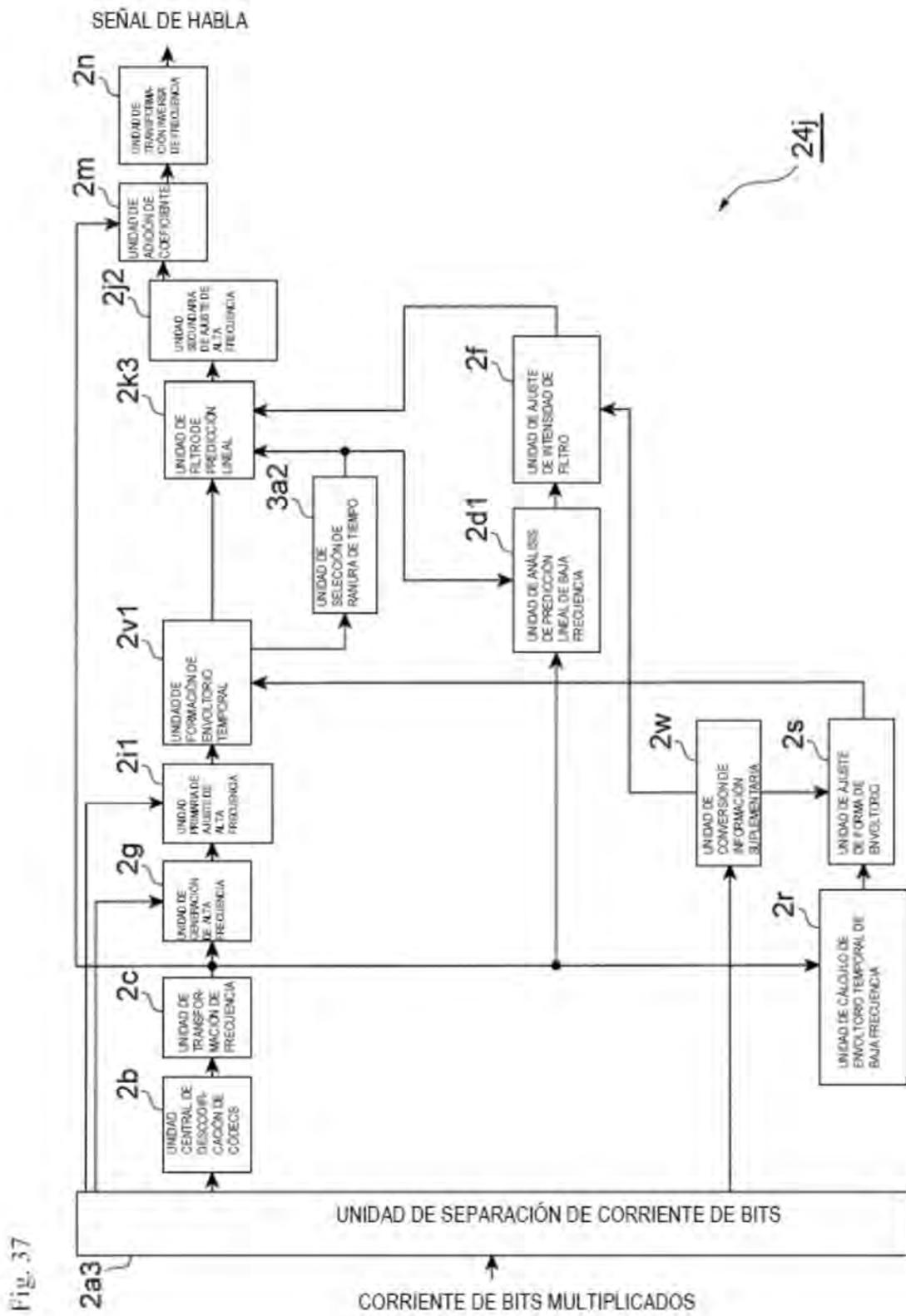


Fig. 35

Fig. 36





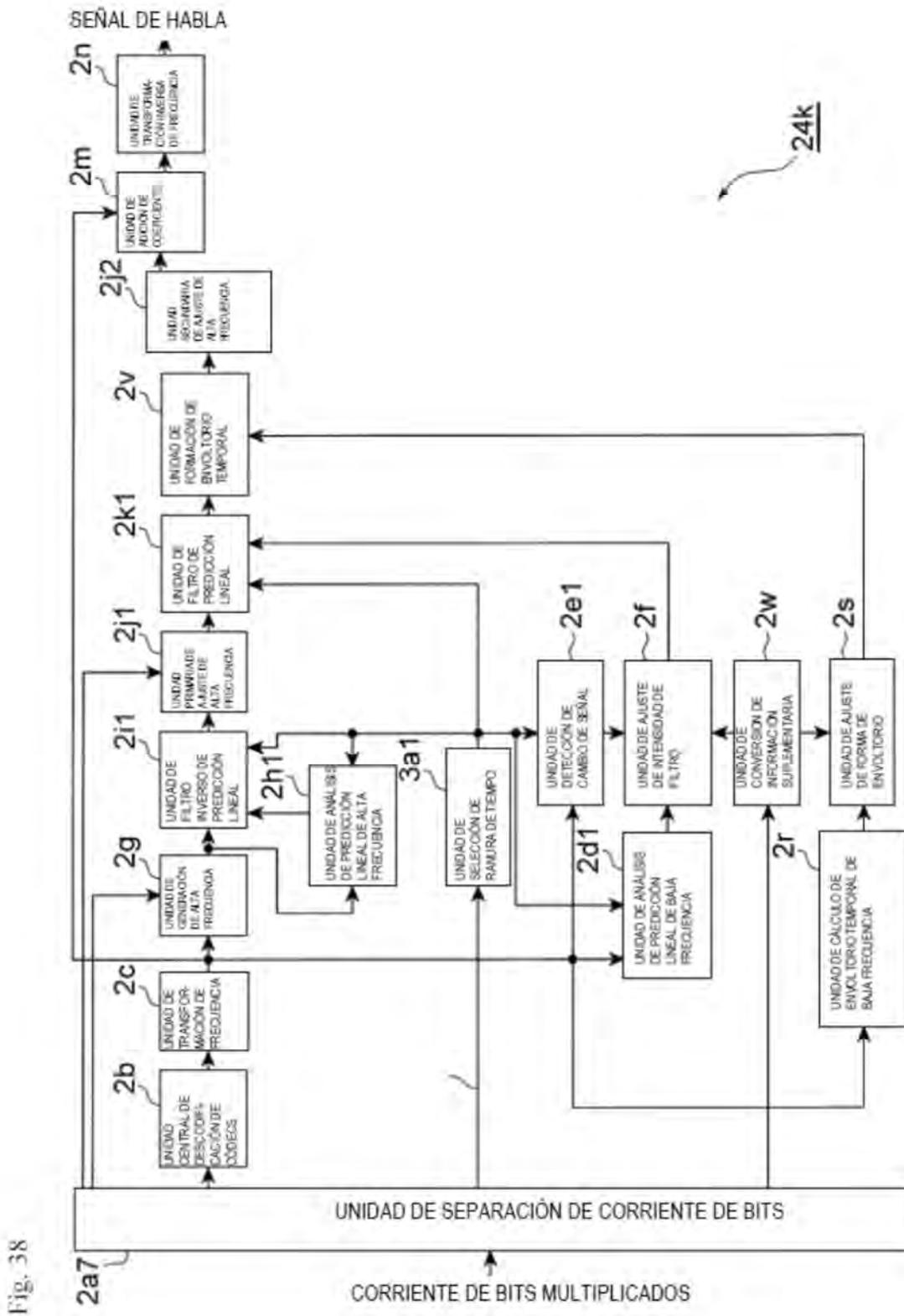
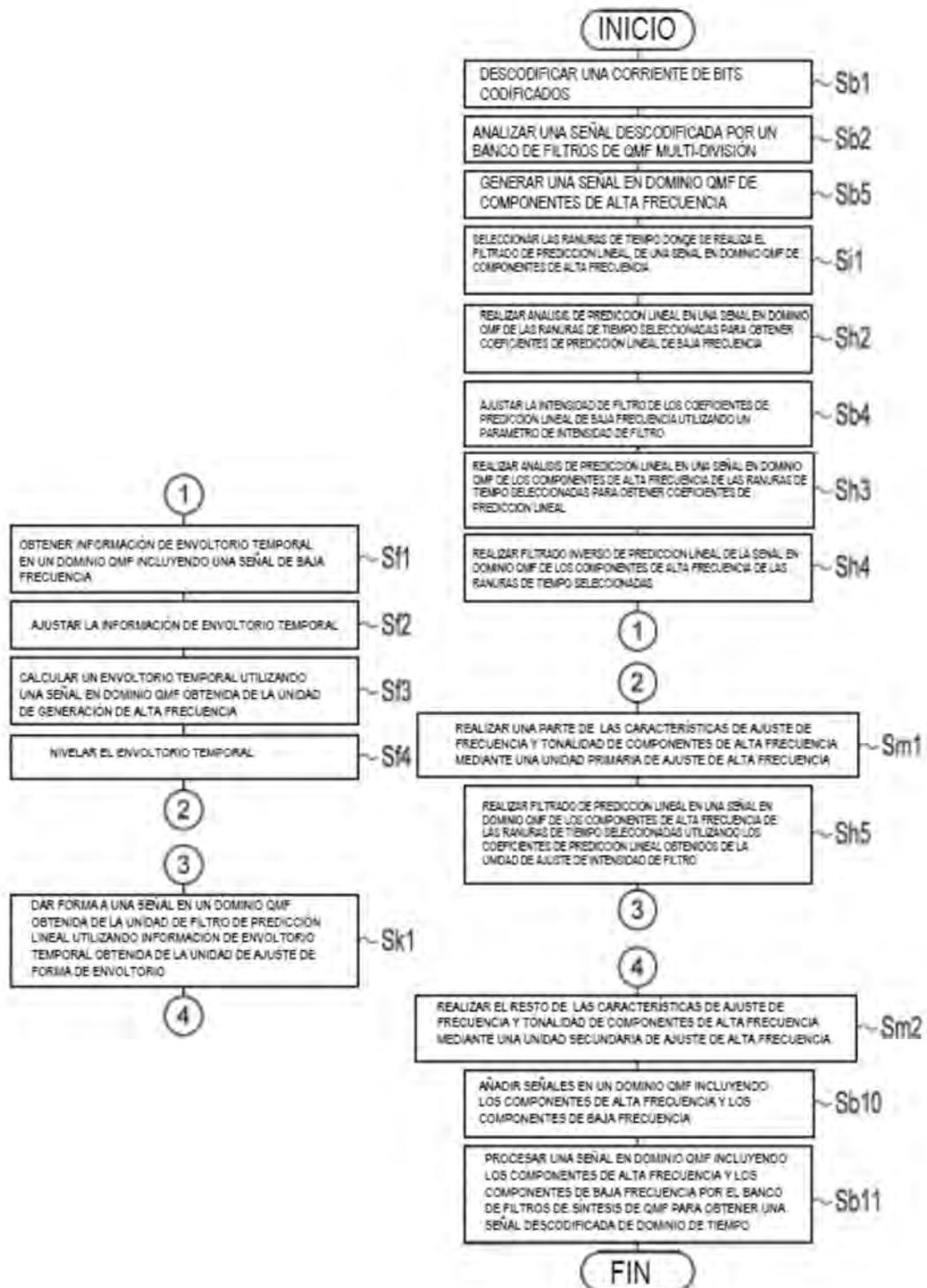


Fig. 38

Fig. 39



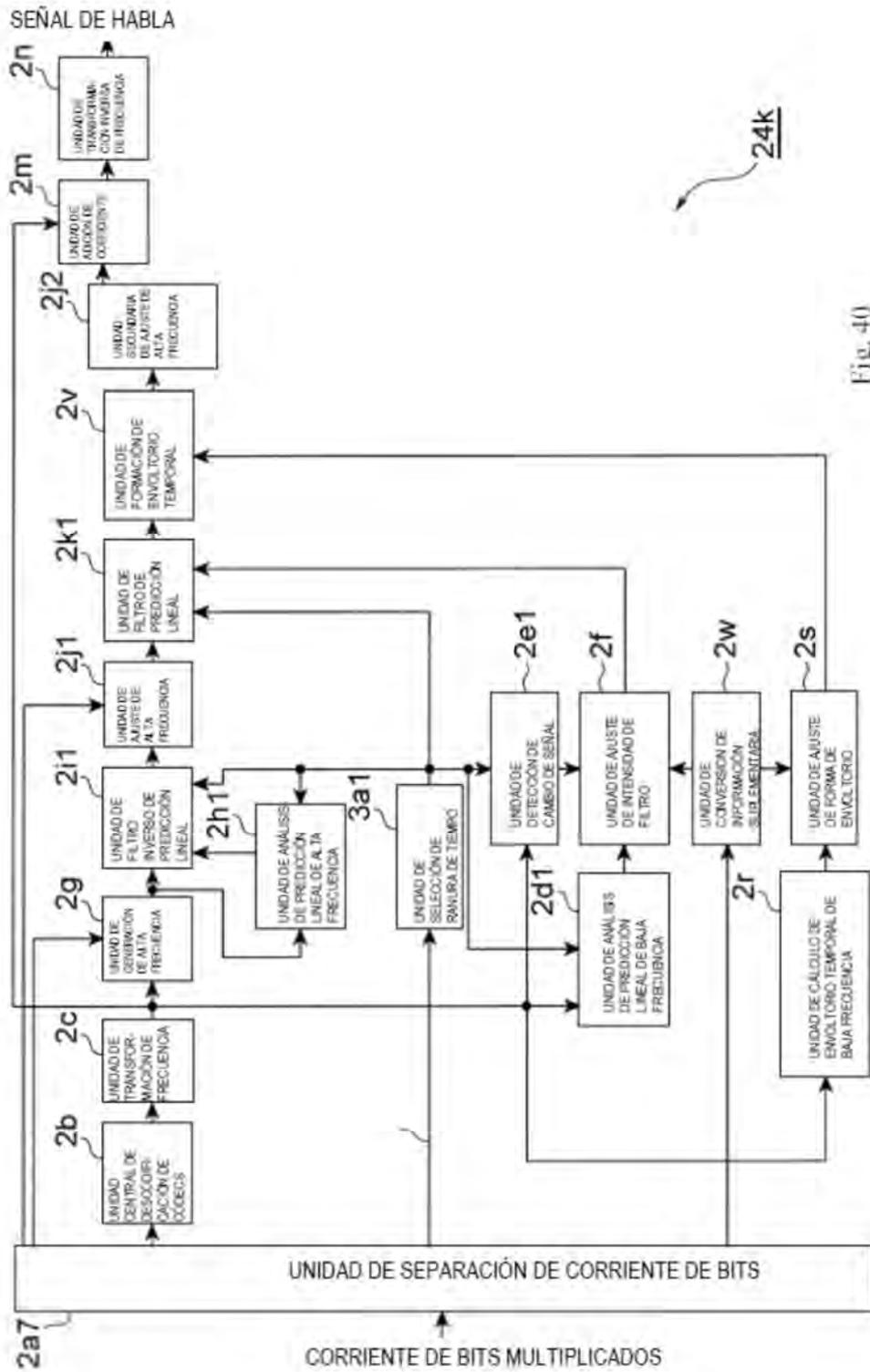


Fig. 40

Fig. 41

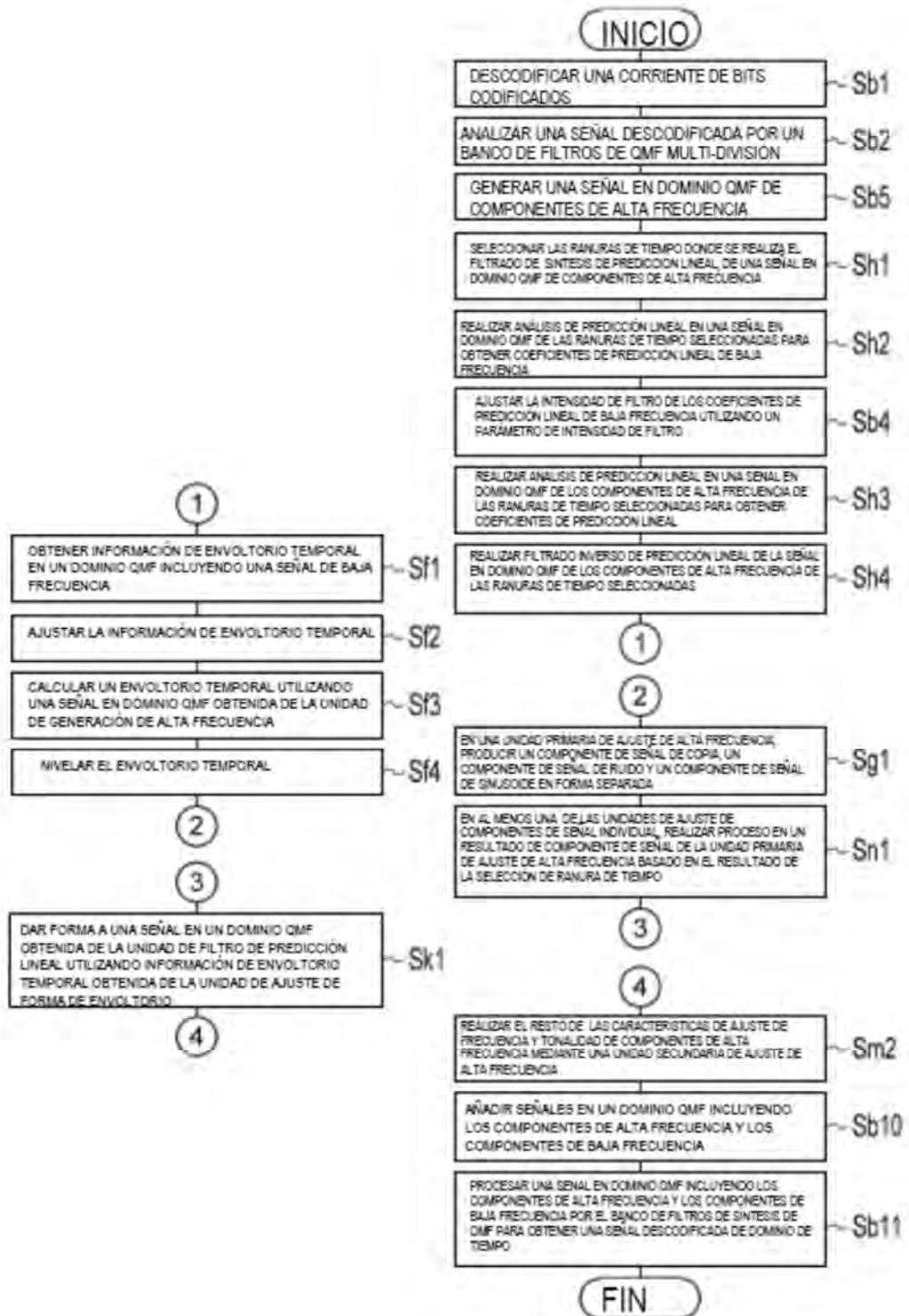


Fig. 42

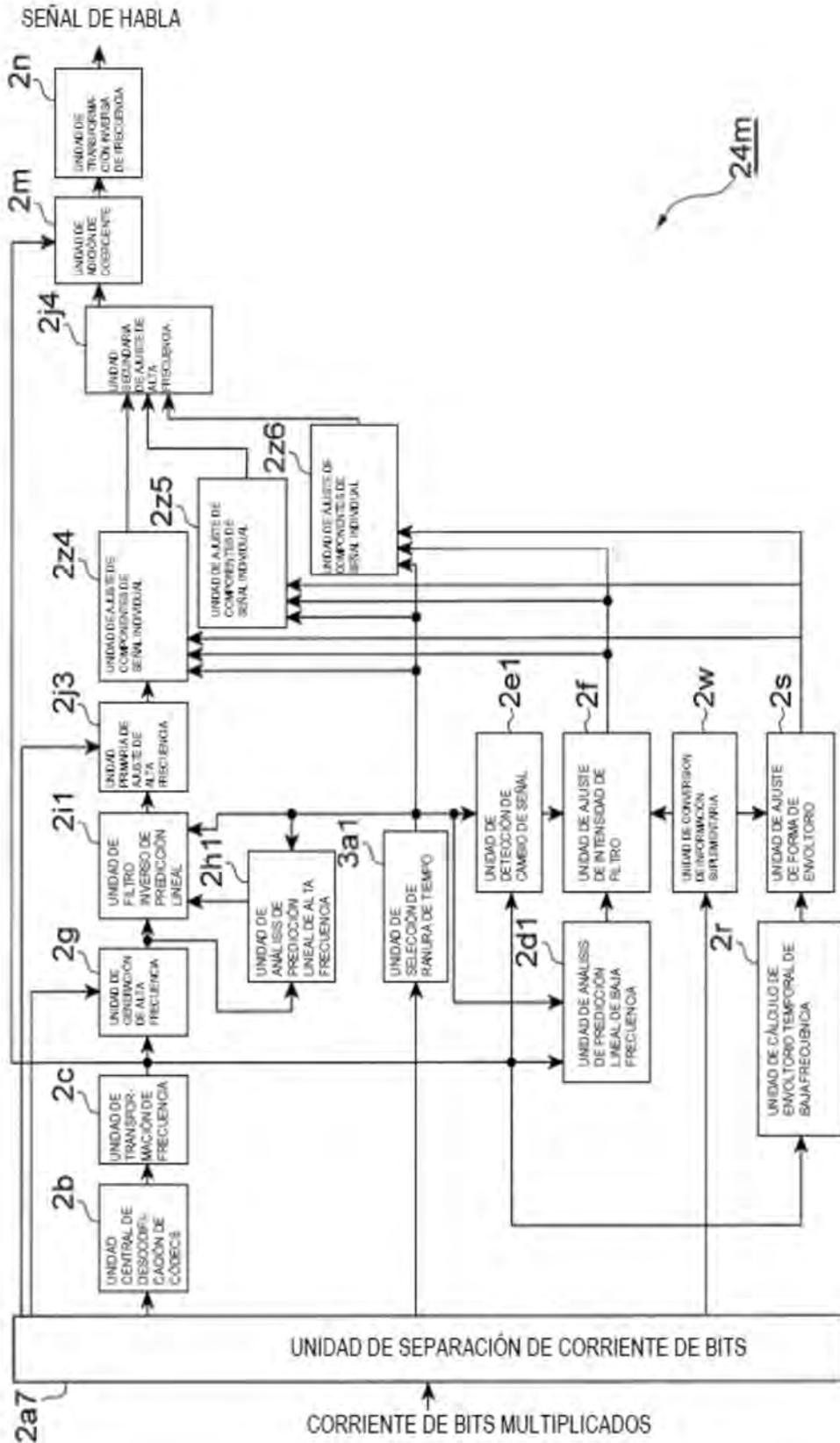


Fig. 43

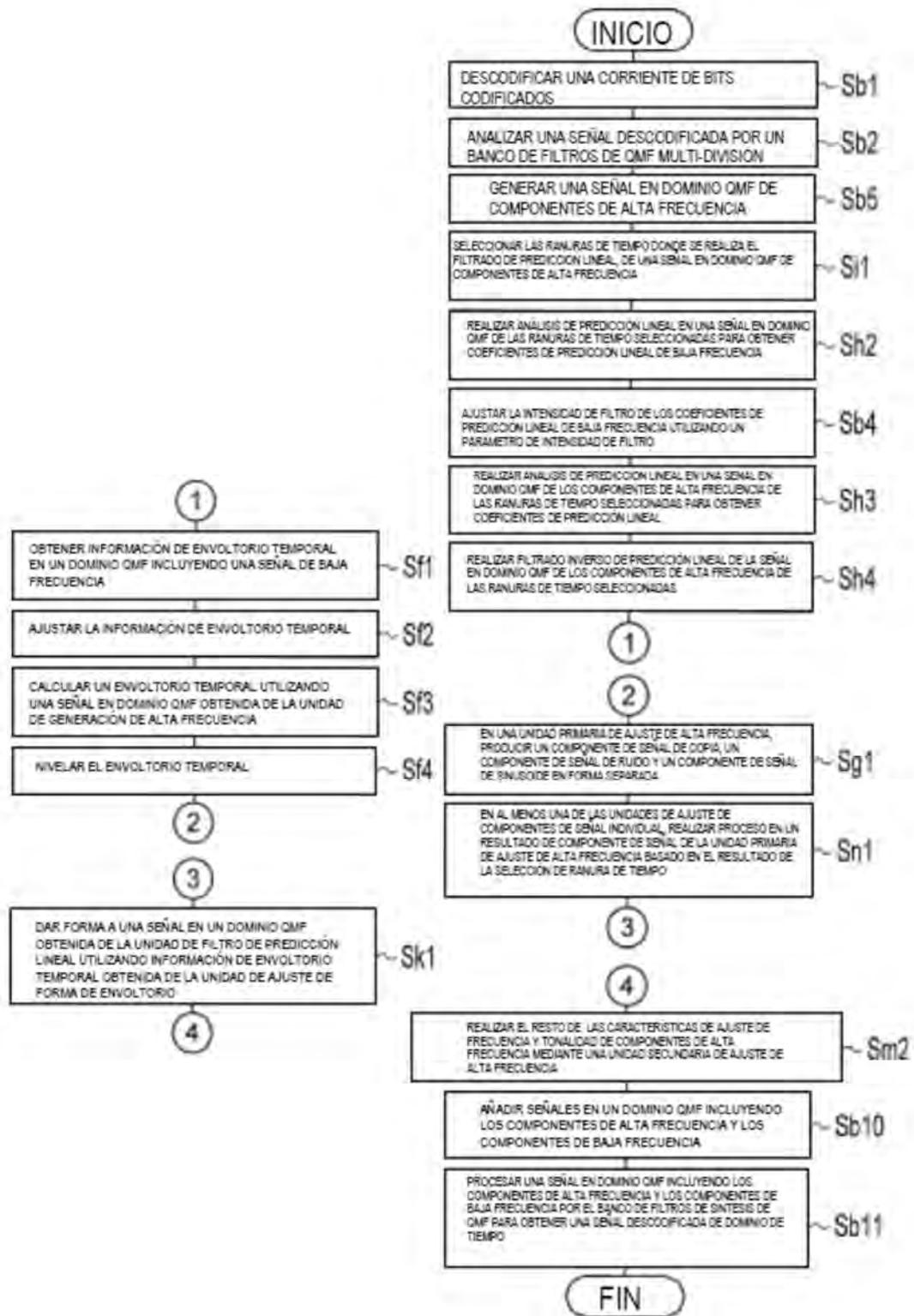


Fig. 44

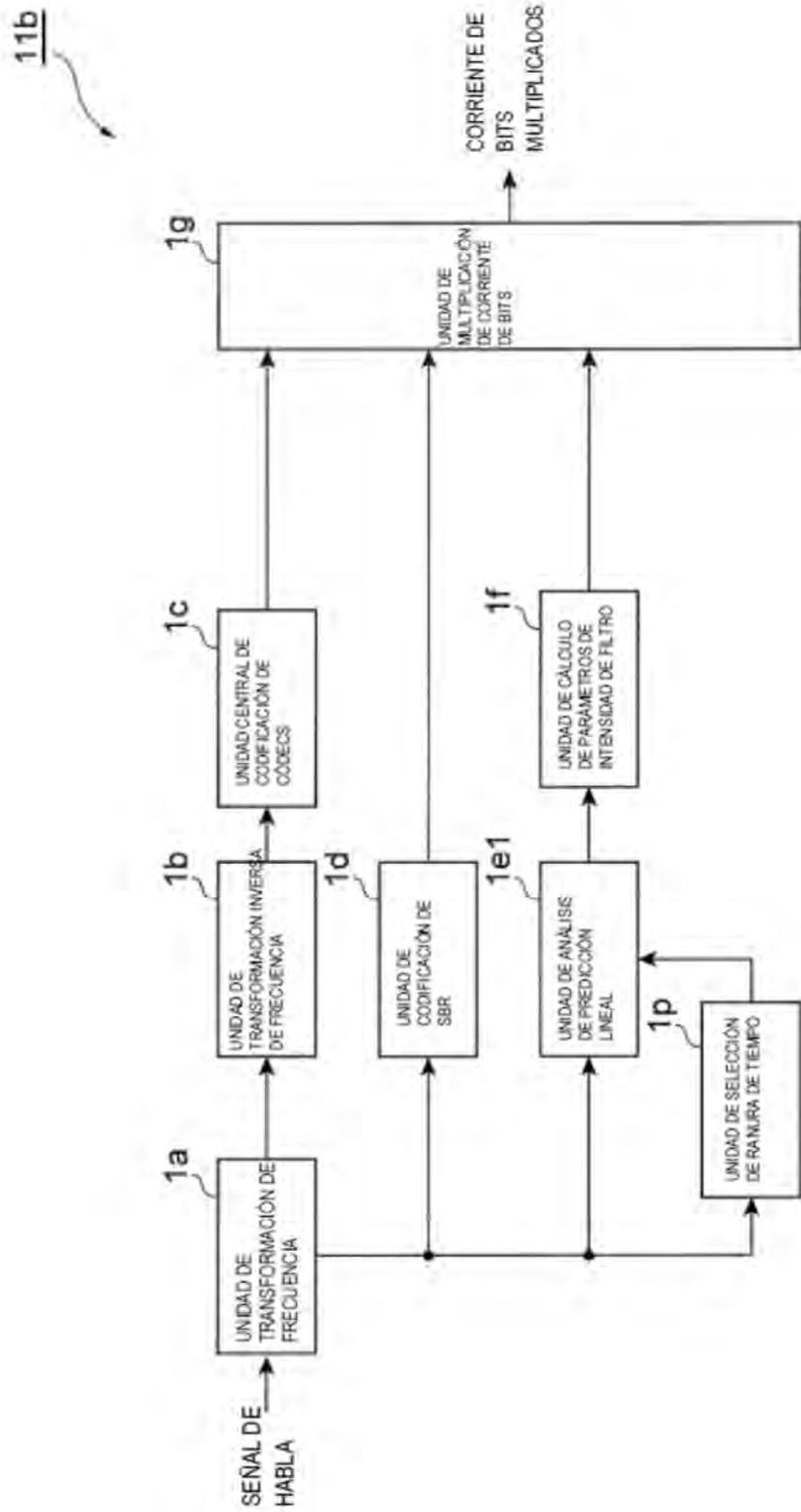


Fig. 45

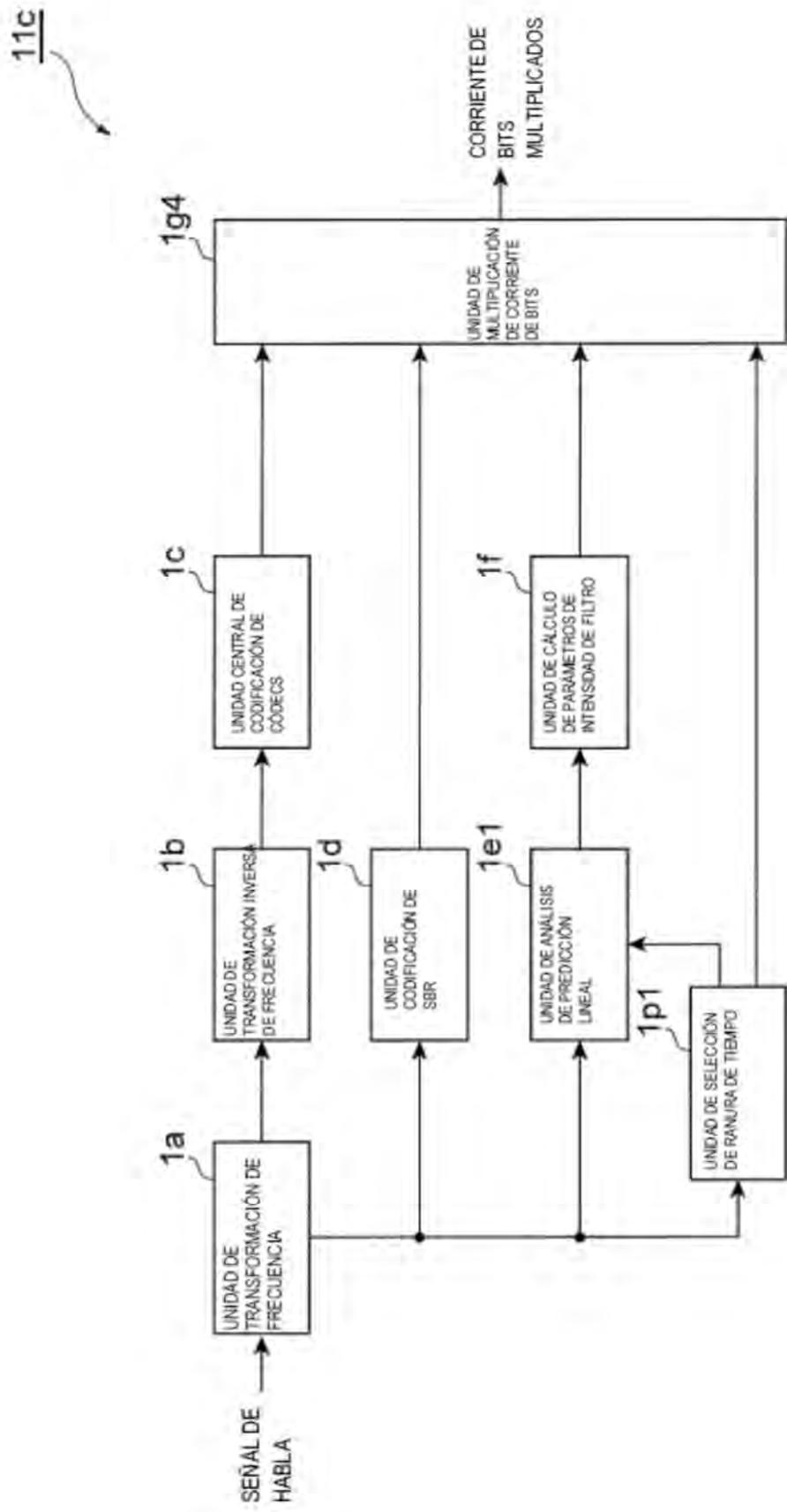
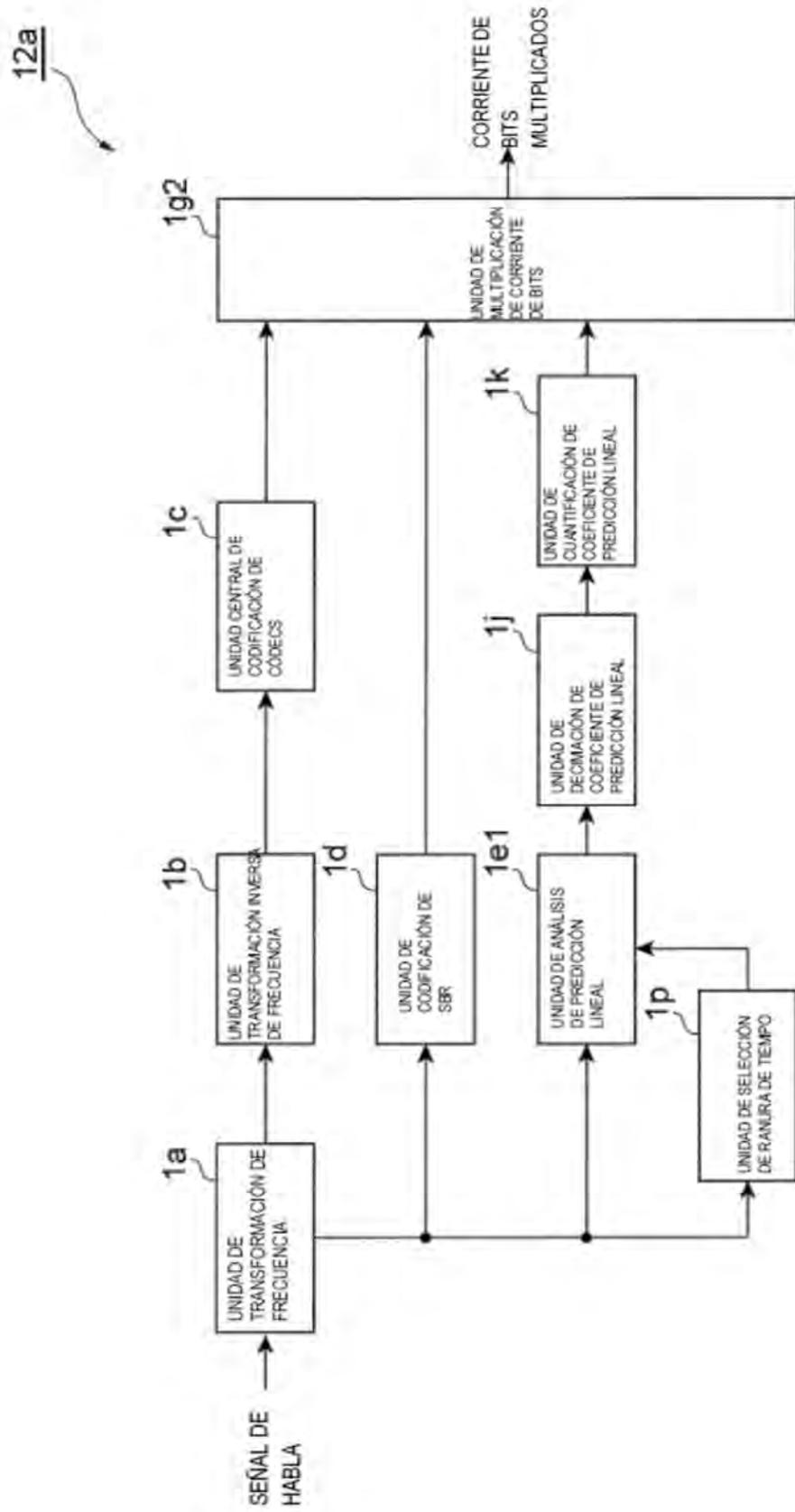


Fig. 46



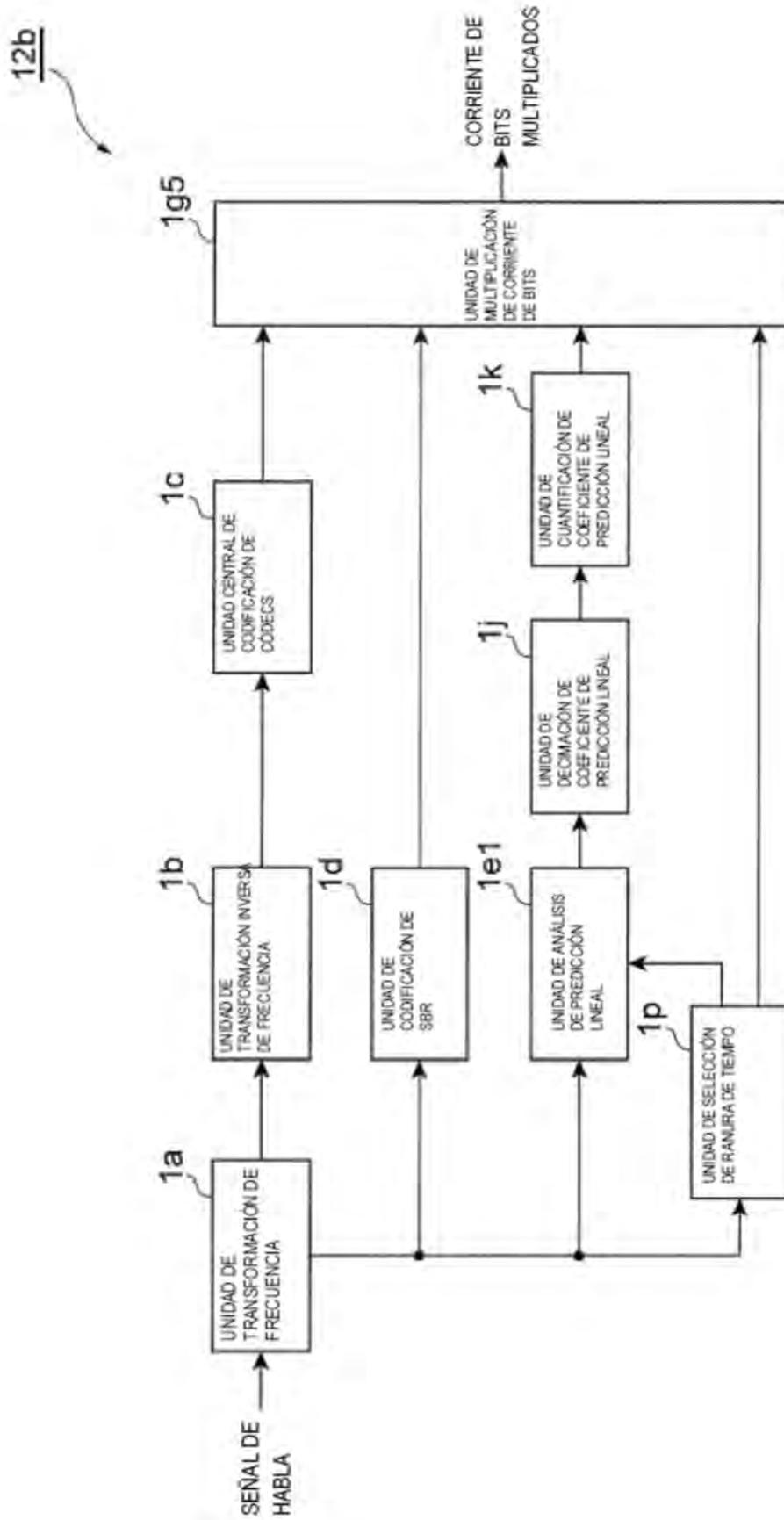


Fig. 47

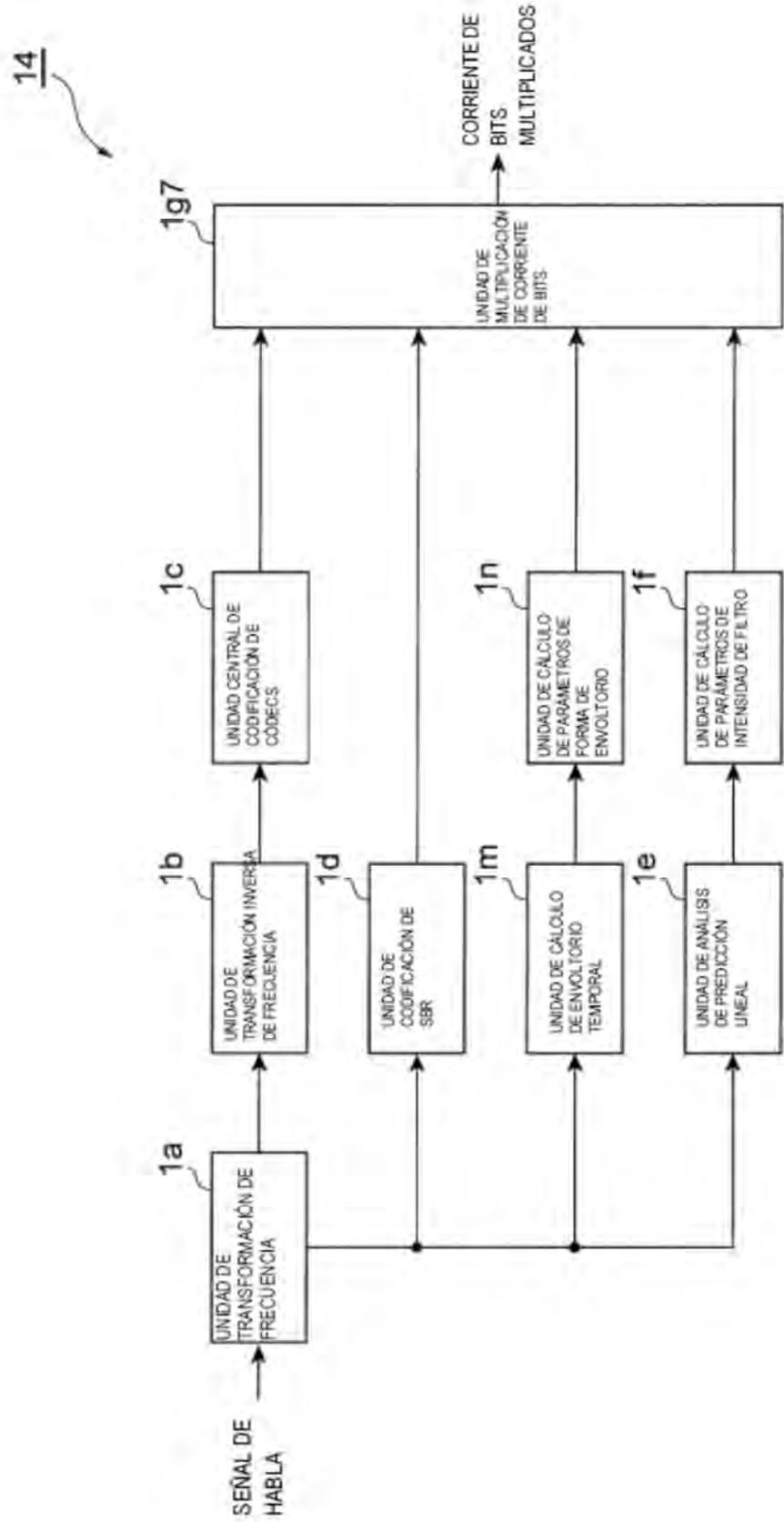


Fig. 48

Fig. 49

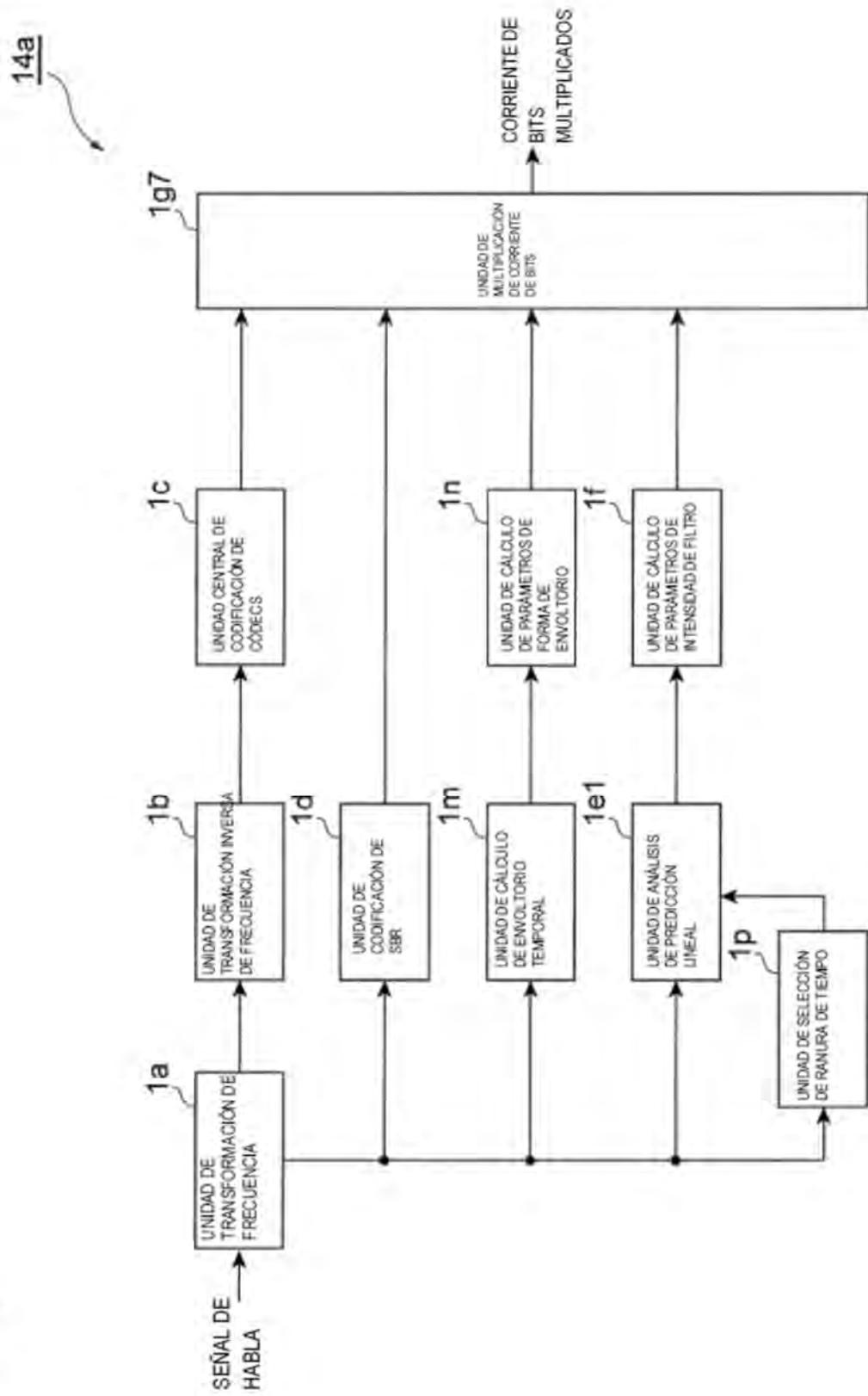


Fig. 50

