

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 610 363**

51 Int. Cl.:

G10L 19/06 (2013.01)

G10L 21/04 (2013.01)

G10L 19/24 (2013.01)

G10L 21/038 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **02.04.2010 E 12171603 (9)**

97 Fecha y número de publicación de la concesión europea: **19.10.2016 EP 2509072**

54 Título: **Dispositivo de decodificación del habla, procedimiento de decodificación del habla y programa de decodificación del 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:
27.04.2017

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:

MARTÍN BADAJOZ, Irene

ES 2 610 363 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 del habla, procedimiento de decodificación del habla y programa de decodificación del habla

5

Campo técnico

La presente invención se refiere a un dispositivo de decodificación del habla, a un programa de decodificación del habla y a un procedimiento de decodificación del habla.

10

Antecedentes del estado de la técnica

Las técnicas de codificación del habla y del audio para comprimir la cantidad de datos de señales en unas cuantas decenas, extrayendo información no necesaria para la percepción humana, utilizando psico-acústica, son extremadamente importantes para transmitir y almacenar señales. Los ejemplos de técnicas perceptuales de codificación de audio ampliamente utilizadas incluyen "MPEG4 AAC", regularizada por las normas "ISO/IEC MPEG" o "MPEG-4 HE-AAC v2", regularizada por el Instituto Europeo de Normas de Telecomunicaciones (ETSI), presentada en un artículo en la EBU Technical Review de S. Meltzer y G. Moser, "MPEG-4 HE-AAC v2 - audio coding for today's digital media world" ["MPEG-4 HE-AAC v2 – codificación de audio para el mundo de los medios digitales de hoy"], 31/01/2006.

15

20

Una técnica de extensión de ancho de banda para generar componentes de alta frecuencia utilizando componentes del habla de baja frecuencia ha sido utilizada ampliamente en los últimos años como un procedimiento para mejorar el rendimiento de la codificación del habla y obtener una alta calidad del habla con una baja velocidad de bits. Entre los ejemplos típicos de la técnica de extensión del ancho de banda se encuentra la técnica de SBR (Replicación de Banda Espectral) que se utiliza en la "MPEG4 AAC". En la 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 la envolvente y la tonalidad espectrales de los coeficientes copiados. Dado que un procedimiento de codificación del habla que utiliza la técnica de extensión del 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 la velocidad de bits de la codificación del habla.

25

30

En la técnica de extensión del ancho de banda en el dominio de frecuencias representado por la SBR, la envolvente y la tonalidad espectrales de los coeficientes espectrales representados en el dominio de la frecuencia se ajustan ajustando una ganancia 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 la envolvente 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 la envolvente temporal del componente de alta frecuencia se transforma durante el proceso de ajuste y, en muchos casos, la envolvente temporal es más llana después del proceso de ajuste que antes del proceso de ajuste. La envolvente temporal del componente de alta frecuencia después del proceso de ajuste no coincide con la envolvente 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.

35

40

45

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 Estéreo Paramétrico. Un decodificador utilizado en la codificación de audio multicanal incluye medios para realizar la de-correlación en una señal decodificada utilizando un filtro de reverberación. Sin embargo, la envolvente temporal de la señal se transforma durante la de-correlación, provocando de esta manera una degradación de una señal de reproducción similar a la del pre-eco y el post-eco. Las soluciones para este problema incluyen una técnica de TES ("Temporal Envelope Shaping", formación de envolvente temporal) (Bibliografía de Patentes 1). En la técnica de TES, se realiza un análisis de predicción lineal en una dirección de frecuencia en una señal representada en un dominio de QMF, sobre el cual la de-correlación 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 de-correlación. Este proceso permite a la técnica de TES extraer la envolvente temporal de una señal en la cual no se ha realizado todavía la de-correlación y, de acuerdo a la envolvente temporal extraída, ajustar la envolvente temporal de la señal en la cual se ha realizado la de-correlación. Dado que la señal en la cual todavía no se ha realizado la de-correlación tiene una envolvente temporal menos distorsionada, la envolvente temporal de la señal en la cual se ha realizado la de-correlación 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.

50

55

60

Lista de citas

65

Bibliografía de patentes

Bibliografía de Patentes: 1. Publicación de Solicitud de Patente Estadounidense N° 2006/0239473

Sumario de la invención

5

Problema técnico

La técnica de TES descrita anteriormente es una técnica que utiliza el hecho de que una señal en la cual todavía no se ha realizado la de-correlación tiene una envolvente temporal menos distorsionada. Sin embargo, en un decodificador de 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 una envolvente temporal menos distorsionada con respecto al componente de alta frecuencia. Una de las soluciones a este problema es un procedimiento que analiza el componente de alta frecuencia de una señal de entrada en un codificador de SBR, cuantizando los coeficientes de predicción lineal obtenidos como resultado del análisis, y multiplexándolos en un flujo de bits que va a ser transmitido. Este procedimiento permite que el decodificador de SBR obtenga unos coeficientes de predicción lineal que incluyen información con una envolvente temporal menos distorsionada del componente de alta frecuencia. Sin embargo, en este caso, se necesita una gran cantidad de información para transmitir los coeficientes de predicción lineal cuantizados, incrementando por ello significativamente la velocidad de bits de todo el flujo 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 la velocidad de bits en la técnica de extensión de ancho de banda en el dominio de frecuencia representado por la SBR.

10

15

20

Solución al problema

La presente invención proporciona un dispositivo de decodificación del habla de acuerdo a la reivindicación 1, un dispositivo de decodificación del habla de acuerdo a la reivindicación 2, un procedimiento de decodificación del habla de acuerdo a la reivindicación 3, un procedimiento de decodificación del habla de acuerdo a la reivindicación 4, un programa de decodificación del habla de acuerdo a la reivindicación 5 y un programa de decodificación del habla de acuerdo a la reivindicación 6.

30

Efectos ventajosos de la invención

De acuerdo a 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 la velocidad de bits en la técnica de extensión de ancho de banda en el dominio de frecuencia representado por la SBR.

35

Breve descripción de los dibujos

La figura 1 es un diagrama que ilustra un dispositivo de codificación del habla de acuerdo a un primer ejemplo;

la figura 2 es un diagrama de flujo para describir un funcionamiento del dispositivo de codificación del habla de acuerdo al primer ejemplo;

la figura 3 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo al primer ejemplo;

la figura 4 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo al primer ejemplo;

la figura 5 es un diagrama que ilustra un dispositivo de codificación del habla de acuerdo a una primera modificación del primer ejemplo;

la figura 6 es un diagrama que ilustra un dispositivo de codificación del habla de acuerdo a una segunda realización;

la figura 7 es un diagrama de flujo para describir un funcionamiento del dispositivo de codificación del habla de acuerdo al segundo ejemplo;

la figura 8 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo al segundo ejemplo;

la figura 9 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo al segundo ejemplo;

la figura 10 es un diagrama que ilustra un dispositivo de codificación del habla de acuerdo a un tercer ejemplo;

la figura 11 es un diagrama de flujo para describir un funcionamiento del dispositivo de codificación del habla de acuerdo al tercer ejemplo;

40

45

50

55

60

65

- la figura 12 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo al tercer ejemplo;
- 5 la figura 13 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo al tercer ejemplo;
- la figura 14 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a una realización;
- 10 la figura 15 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a una modificación de la realización;
- la figura 16 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación de la realización;
- 15 la figura 17 es un diagrama de flujo que describe un funcionamiento del dispositivo de codificación del habla de acuerdo a la otra modificación de la realización;
- la figura 18 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación del primer ejemplo;
- 20 la figura 19 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo a la otra modificación del primer ejemplo;
- la figura 20 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación del primer ejemplo;
- 25 la figura 21 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo a la otra modificación del primer ejemplo;
- la figura 22 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a una modificación del segundo ejemplo;
- 30 la figura 23 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo a la modificación del segundo ejemplo;
- 35 la figura 24 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación del segundo ejemplo;
- la figura 25 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo a la otra modificación del segundo ejemplo;
- 40 la figura 26 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación de la realización;
- la figura 27 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo a la otra modificación de la realización;
- 45 la figura 28 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación de la realización;
- 50 la figura 29 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo a la otra modificación de la realización;
- la figura 30 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación de la realización;
- 55 la figura 31 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación de la realización;
- la figura 32 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo a la otra modificación de la realización;
- 60 la figura 33 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación de la realización;
- 65 la figura 34 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo a la otra modificación de la realización;

la figura 35 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación de la realización;

5 la figura 36 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo a la otra modificación de la realización;

la figura 37 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación de la realización;

10 la figura 38 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación de la realización;

15 la figura 39 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo a la otra modificación de la realización;

la figura 40 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a la otra modificación de la realización;

20 la figura 41 es un diagrama de flujo para describir un funcionamiento del dispositivo de decodificación del habla de acuerdo a otra modificación de la realización;

la figura 42 es un diagrama que ilustra un dispositivo de decodificación del habla de acuerdo a otra modificación de la realización;

25 la figura 43 es un diagrama de flujo para describir un funcionamiento del dispositivo de codificación de acuerdo a la otra modificación de la realización;

30 la figura 44 es un diagrama que ilustra un dispositivo de codificación del habla de acuerdo a otra modificación del primer ejemplo;

la figura 45 es un diagrama que ilustra un dispositivo de codificación del habla de acuerdo a otra modificación más del primer ejemplo;

35 la figura 46 es un diagrama que ilustra un dispositivo de codificación del habla de acuerdo a una modificación del segundo ejemplo;

la figura 47 es un diagrama que ilustra un dispositivo de codificación del habla de acuerdo a otra modificación del segundo ejemplo;

40 la figura 48 es un diagrama que ilustra un dispositivo de codificación del habla de acuerdo a la realización;

la figura 49 es un diagrama que ilustra un dispositivo de codificación del habla de acuerdo a otra modificación de la realización; y

45 la figura 50 es un diagrama que ilustra un dispositivo de codificación del habla de acuerdo a otra modificación de la realización.

Descripción de las realizaciones

50 A continuación se describen en detalle realizaciones preferibles 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, si corresponde.

55 (Primer ejemplo)

60 La figura 1 es un diagrama que ilustra un dispositivo de codificación del habla 11 de acuerdo a un primer ejemplo. El dispositivo de codificación del 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 del habla 11 cargando y ejecutando un programa informático predeterminado (como, por ejemplo, un programa informático para realizar los procesos ilustrados en el diagrama de flujo de la figura 2), almacenado en una memoria integrada del dispositivo de codificación del habla 11, como, por ejemplo, la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 11 recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 11, y emite un flujo de bits multiplexados y codificados al exterior del

dispositivo de codificación del habla 11.

El dispositivo de codificación del 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 códec 1c (medio central de codificación), una unidad de codificación de SBR 1d, una unidad de análisis de predicción lineal 1e (medio de cálculo de información suplementaria de envolvente temporal), una unidad de cálculo de parámetros de intensidad de filtro 1f (medio de cálculo de información suplementaria de envolvente temporal) y una unidad de multiplexado de flujo de bits 1g (medio de multiplexado de flujo de bits). Desde la unidad de transformación de frecuencia 1a hasta la unidad de multiplexado de flujo de bits 1g, del dispositivo de codificación del habla 11 ilustrado en la figura 1, son funciones realizadas cuando la CPU del dispositivo de codificación del habla 11 ejecuta el programa informático almacenado en la memoria integrada del dispositivo de codificación del habla 11. La CPU del dispositivo de codificación del habla 11 ejecuta procesos de forma secuencial (procesos desde la etapa Sa1 hasta la etapa Sa7), ilustrados en el diagrama de flujo de la figura 2, ejecutando el programa informático (o utilizando desde la unidad de transformación de frecuencia 1a hasta la unidad de multiplexado de flujo de bits 1g ilustrada en la figura 1). Diferentes tipos de datos necesarios para ejecutar el programa informático, y diferentes tipos de datos generados por la ejecución del programa informático, se almacenan, todos, en la memoria integrada, como por ejemplo la ROM y la RAM del dispositivo de codificación del habla 11.

La unidad de transformación de frecuencia 1a analiza una señal de entrada recibida desde el exterior del dispositivo de codificación del habla 11 mediante el dispositivo de comunicación del dispositivo de codificación del habla 11, utilizando un banco de filtros QMF de multi-división para obtener una señal $q(k, r)$ en un dominio de QMF (proceso en la etapa 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 sub-muestreada del dominio del tiempo que incluye solamente componentes de baja frecuencia de la señal de entrada (proceso en la etapa Sa2). La unidad central de codificación de códecs 1c codifica la señal sub-muestreada del dominio del tiempo para obtener un flujo de bits codificados (proceso en la etapa Sa3). La codificación realizada por la unidad central de codificación de códecs 1c puede basarse en un procedimiento de codificación del habla representado por un procedimiento de CELP, o puede estar basada en un procedimiento de codificación de audio, tal como una codificación de transformación representada por la AAC o un procedimiento de TCX (Transform Coded Excitation), excitación codificada por transformación.

La unidad de codificación de SBR 1d recibe la señal en el dominio de QMF desde la unidad de transformación de frecuencia 1a, y realiza la codificación de 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 la etapa Sa4). El procedimiento de análisis de QMF en la unidad de transformación de frecuencia 1a y el procedimiento de codificación de SBR en la unidad de codificación de SBR 1d se describen en detalle, por ejemplo, en la referencia de la Bibliografía "3GPP TS 26.404: Enhanced aacPlus encoder SBR part" ["3GPP TS 26.404: Codificador mejorado aacPlus – parte de SBR"].

La unidad de análisis de predicción lineal 1e recibe la señal en el dominio de 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 la etapa 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 de QMF. Puede utilizarse un procedimiento de covarianza o un procedimiento de auto-correlación 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 de códecs 1c. La unidad de análisis de predicción lineal 1e puede también realizar un análisis de predicción lineal en componentes de baja frecuencia distintos a los analizados cuando se obtienen los $a_H(n, r)$ para obtener coeficientes de predicción lineal de baja frecuencia $a_L(n, r)$ distintos a los $a_H(n, r)$ (los coeficientes de predicción lineal de acuerdo a dichos componentes de baja frecuencia corresponden a información de la envolvente temporal, y es lo mismo que 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 $0 \leq k < k_x$. 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 de la envolvente temporal y es el mismo en el primer ejemplo que en lo siguiente) (proceso en la etapa Sa6). Una ganancia de predicción $G_H(r)$ es calculada en primer lugar a partir de $a_H(n, r)$. El procedimiento para calcular la ganancia de predicción se describe, por ejemplo, en "Speech Coding, Takehiro Moriya, The Institute of Electronics, Information and Communication Engineers" ["Codificación del habla, Takehiro Moriya. El Instituto de Ingenieros de Electrónica, Información y Comunicación"]. Si se ha calculado $a_L(n, r)$, se calcula de forma similar una ganancia de predicción $G_L(r)$. 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 a 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 a 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 la envolvente temporal de los componentes de alta frecuencia durante la decodificación de la SBR. Un valor de la ganancia de predicción con respecto a los coeficientes de predicción lineal en la dirección de la frecuencia aumenta a medida que la variación de la envolvente temporal de una señal en el intervalo de análisis se hace más aguda. $K(r)$ es un parámetro para dar instrucciones a un decodificador de reforzar el proceso para agudizar la variación de la envolvente 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 (tal como, por ejemplo, un dispositivo de decodificación del habla 21) para debilitar el proceso para agudizar la variación de la envolvente 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 de la envolvente temporal. En lugar de transmitir $K(r)$ a cada ranura de tiempo, se puede transmitir un $K(r)$ 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 sobre los límites de la envolvente de SBR (límite temporal de envolvente de SBR) incluidos en la información suplementaria de SBR.

$K(r)$ se transmite a la unidad de multiplexado del flujo de bits 1g después de ser cuantizado. Es preferible calcular el $K(r)$ que representa la pluralidad de ranuras de tiempo, por ejemplo, calculando un promedio de $K(r)$ de una pluralidad de ranuras de tiempo r antes de realizar la cuantización. Para transmitir el $K(r)$ que representa la pluralidad de ranuras de tiempo, el $K(r)$ también puede obtenerse a partir del resultado del análisis de todo el segmento formado a partir 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 a la siguiente expresión (3). En este caso, $\text{mean}(\cdot)$ 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 modalidad de filtro inverso, incluida en la información suplementaria de SBR descrita en el documento "ISO/IEC 14496-3 subpart 4 General Audio Coding" ["ISO/IEC 14496-3 - sub-parte 4 - Codificación general de audio"]. En otras palabras, $K(r)$ no se transmite para las ranuras de tiempo para las cuales se transmite la información de modalidad de filtro inverso en la información suplementaria de SBR, y la información de modalidad de filtro inverso (modalidad_bs_invf en el documento "ISO/IEC 14496-3 - sub-parte 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 modalidad de filtro inverso incluida en la información suplementaria de SBR. $K(r)$ y la información de modalidad de filtro inverso incluida en la información suplementaria de SBR pueden combinarse para ser gestionadas como información vectorial, y realizar una codificación por entropía en el vector. En este caso, se puede restringir la combinación de $K(r)$ y el valor de la información de modalidad de filtro inverso incluida en la información suplementaria de SBR.

La unidad de multiplexado de flujo de bits 1g multiplexa el flujo de bits codificados, calculado 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 $K(r)$ calculado por la unidad de cálculo de parámetros de intensidad de filtro 1f, y emite un flujo de bits multiplexado (flujo de bits multiplexado codificado) a través del dispositivo de comunicación del dispositivo de codificación del habla 11 (proceso en la etapa Sa7).

La figura 3 es un diagrama que ilustra un dispositivo de decodificación del habla 21 de acuerdo al primer ejemplo. El dispositivo de decodificación del 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 del habla 21, cargando y ejecutando un programa informático predeterminado (tal como un programa informático para realizar procesos ilustrado en el diagrama de flujo de la figura 4) almacenado en una memoria integrada del

dispositivo de decodificación del habla 21, como, por ejemplo, la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 21 recibe la salida del flujo de bits multiplexados codificados desde el dispositivo de codificación del habla 11, un dispositivo de codificación del habla 11a de una modificación 1, que será descrito más adelante, o un dispositivo de codificación del habla de una modificación 2, que será descrito más adelante, y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 21. El dispositivo de decodificación del habla 21, tal como se ilustra en la figura 3, incluye de forma funcional una unidad de separación de flujos de bits 2a (medio de separación de flujos de bits), una unidad de decodificación central de códecs 2b (medio de decodificación central), 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 envolvente temporal de baja frecuencia), una unidad detectora de cambio de señal 2e, una unidad de ajuste de intensidad de filtro 2f (medio de ajuste de envolvente 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 de envolvente temporal), una unidad de adición de coeficiente 2m y una unidad de transformación inversa de frecuencia 2n. Tal como se menciona en el párrafo [0113], el dispositivo de decodificación del habla 24 de acuerdo a la realización de la presente invención incluye funcionalmente la estructura del dispositivo de decodificación del habla 21. La unidad de decodificación central de códecs 2b es una realización del medio de decodificación central de la presente invención. La unidad de transformación de frecuencia 2c es una realización del medio de transformación de frecuencia de la presente invención. La unidad de generación de alta frecuencia 2g es una realización del medio de generación de alta frecuencia de la presente invención. La unidad de ajuste de alta frecuencia 2j es una realización del medio de ajuste de alta frecuencia de la presente invención. Desde la unidad de separación de flujos de bits 2a hasta una unidad de cálculo de parámetros de forma de envolvente 1n, del dispositivo de decodificación del habla 21 ilustrado en la figura 3, son funciones que se realizan cuando la CPU del dispositivo de decodificación del habla 21 ejecuta el programa informático almacenado en la memoria integrada del dispositivo de decodificación del habla 21. La CPU del dispositivo de decodificación del habla 21 ejecuta secuencialmente procesos (procesos desde la etapa Sb1 hasta la etapa Sb11) ilustrados en el diagrama de flujo de la figura 4, ejecutando el programa informático (o utilizando desde la unidad de separación de flujos de bits 2a hasta la unidad de cálculo de parámetros de la forma de la envolvente 1n ilustrada en la figura 3). Diferentes tipos de datos necesarios para ejecutar el programa informático y diferentes tipos de datos generados por la ejecución del programa informático se almacenan todos en la memoria integrada, tal como, por ejemplo, la ROM y la RAM del dispositivo de decodificación del habla 21.

La unidad de separación de flujos de bits 2a separa el flujo de bits multiplexado proporcionado a través del dispositivo de comunicación del dispositivo de decodificación del habla 21 en un parámetro de intensidad de filtro, información suplementaria de SBR y el flujo de bits codificados. La unidad de decodificación central de códecs 2b decodifica el flujo de bits codificados, recibido desde la unidad de separación de flujos de bits 2a, para obtener una señal decodificada que incluye solamente los componentes de baja frecuencia (proceso en la etapa Sb1). En este momento, el procedimiento de decodificación puede basarse en el procedimiento de codificación del habla representado por el procedimiento de CELP, o puede basarse en una codificación de audio como la AAC o el procedimiento de TCX (Transform Coded Excitation), excitación codificada por transformación.

La unidad de transformación de frecuencia 2c analiza la señal decodificada recibida desde la unidad de decodificación central de códecs 2b utilizando el banco de filtros de QMF multi-división para obtener una señal $q_{dec}(k, r)$ en el dominio de QMF (proceso en la etapa 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 sub-muestra de la señal en el dominio de 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 sobre la $q_{dec}(k, r)$ 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 la etapa Sb3). El análisis de predicción lineal se realiza para una gama de $0 \leq k < k_x$ correspondiente a un ancho de banda de señal de la señal decodificada obtenida desde la unidad de decodificación central 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 de QMF, recibida desde la unidad de transformación de frecuencia 2c, y la emite como un resultado de detección $T(r)$. El cambio de señal puede ser detectado, por ejemplo, utilizando el procedimiento 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 a la siguiente expresión (4).

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

2. Una envolvente $p_{env}(r)$, obtenido allanando $p(r)$, se obtiene de acuerdo a 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 a la siguiente expresión (6), utilizando $p(r)$ y $p_{env}(r)$, donde β es una constante.

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

Los procedimientos que se han descrito anteriormente son simples ejemplos para detectar el cambio de señal en base al cambio de potencia, y el cambio de señal puede detectarse utilizando otros procedimientos 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 la etapa Sb4). La intensidad de filtro se ajusta, por ejemplo, de acuerdo a la expresión (7) siguiente, utilizando un parámetro de intensidad de filtro K recibido a través de la unidad de separación de flujos 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 $T(r)$ desde 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 de 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 de QMF de los componentes de alta frecuencia (proceso en la etapa Sb5). Los componentes de alta frecuencia se generan de acuerdo al procedimiento de generación de HF en la SBR, en el documento "MPEG4 AAC". ("ISO/IEC 14496-3 – sub-parte 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 las $q_{exp}(k, r)$ 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 la etapa Sb6). El análisis de predicción lineal se realiza para una gama 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 de 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 la etapa 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 a partir de un coeficiente a una frecuencia más baja, hacia un coeficiente de frecuencia más alta, o puede realizarse en sentido inverso. El filtrado inverso de predicción lineal es un proceso para allanar temporalmente la envolvente temporal de los componentes de alta frecuencia, antes de que se realice la formación de la envolvente temporal en la etapa 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 las salidas desde la unidad de ajuste de alta frecuencia 2j, que se describirá 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 de alta frecuencia de las salidas desde 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 los $a_{exp}(n, r)$. El ajuste de intensidad se realiza de acuerdo a 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 una salida desde la unidad de filtro inverso de predicción lineal 2i (proceso en la etapa Sb8). El ajuste se realiza de acuerdo a la información suplementaria de SBR recibida desde la unidad de separación de flujos de bits 2a. El proceso por parte de la unidad de ajuste de alta frecuencia 2j se realiza de acuerdo a la etapa de "ajuste de HF" en la SBR en el documento "MPEG4 AAC", y se ajusta realizando un filtrado inverso de predicción lineal en la dirección temporal, el ajuste de ganancia y la adición de ruido en la señal en el dominio de QMF de la banda de alta frecuencia. Los detalles de los procesos en las etapas descritas anteriormente se describen en el documento "ISO/IEC 14496-3 – sub-parte 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 alta frecuencia 2j funcionan, todas ellas, de acuerdo al decodificador de SBR en la "MPEG4 AAC" definida en el documento "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 los componentes de alta frecuencia $q_{adj}(n, r)$ de una señal en el dominio de QMF, salida desde la unidad de ajuste de alta frecuencia 2j, utilizando los $a_{adj}(n, r)$ obtenidos de la unidad de ajuste de intensidad de filtro 2f (proceso en la etapa 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 forma la envolvente temporal de los componentes de alta frecuencia generados en base a la SBR.

La unidad de adición de coeficientes 2m añade una señal en el dominio de QMF que incluye los componentes de baja frecuencia emitidos desde la unidad de transformación de frecuencia 2c y una señal en el dominio de QMF que incluye los componentes de alta frecuencia emitidos desde la unidad de filtro de predicción lineal 2k, y emite una señal en el dominio de QMF que incluye tanto los componentes de baja frecuencia como los componentes de alta frecuencia (proceso en la etapa Sb10).

La unidad de transformación inversa de frecuencia 2n procesa la señal en el dominio de QMF, obtenida desde la unidad de adición de coeficientes 2m, utilizando un banco de filtros de síntesis de QMF. En consecuencia, se obtiene una señal de habla decodificada del dominio del tiempo, que incluye tanto los componentes de baja frecuencia obtenidos por la decodificación central de códecs como los componentes de alta frecuencia generados por la SBR y cuyo envolvente temporal está formado por el filtro de predicción lineal, y se emite la señal de habla obtenida hacia fuera del dispositivo de decodificación del habla 21, a través del dispositivo de comunicación integrado (proceso en la etapa Sb11). Si $K(r)$ y la información de modalidad de filtro inverso de la información suplementaria de SBR, descrita en el documento "ISO/IEC 14496-3 – sub-parte 4 - Codificación General de Audio", se transmiten exclusivamente, la unidad de transformación inversa de frecuencia 2n puede generar información de modalidad de filtro inverso, de la información suplementaria de SBR, para una ranura de tiempo a la cual se transmite $K(r)$, pero la información de modalidad de filtro inverso de la información suplementaria de SBR no se transmite, utilizando información de modalidad de filtro inverso de la información suplementaria de SBR con respecto a al 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 modalidad de filtro inverso de la información suplementaria de SBR de la ranura de tiempo en una modalidad predeterminada 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 inverso de la información suplementaria de SBR, pero no se transmite $K(r)$, utilizando $K(r)$ para al menos una ranura de tiempo de las ranuras de tiempo antes y después de la ranura de tiempo. También es posible establecer el $K(r)$ 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 de la modalidad de filtro inverso de la información suplementaria de SBR, en base a la información que indica si se transmite $K(r)$ o la información de modalidad de filtro inverso de la información suplementaria de SBR.

(Modificación 1 del primer ejemplo)

La figura 5 es un diagrama que ilustra una modificación (dispositivo de codificación del habla 11a) del dispositivo de codificación del habla de acuerdo al primer ejemplo. El dispositivo de codificación del 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 del habla 11a, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de codificación del habla 11a, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 11a recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 11a, y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla 11a.

El dispositivo de codificación del habla 11a, tal como se ilustra en la figura 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 envolvente temporal), una unidad de cálculo de parámetros de intensidad de filtro 1f1 (medio de cálculo de información suplementaria de envolvente temporal) y una unidad de multiplexado de flujos de bits 1g1 (medio de multiplexado de flujos 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 multiplexado de flujos de bits 1g del dispositivo de codificación del habla 11. La unidad de multiplexado de flujos de bits 1g1 tiene la misma función que la de 1G. Desde 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 parámetros de intensidad de filtro 1f1 y la unidad de multiplexado de flujos de bits 1g1 del dispositivo de codificación del habla 11a ilustrado en la figura 5 son funciones realizadas cuando la CPU del dispositivo de codificación del habla 11a ejecuta el programa informático almacenado en la memoria integrada del dispositivo de codificación del habla 11a. Diferentes tipos de datos necesarios para ejecutar el programa informático y diferentes tipos de datos generados al ejecutar el programa informático se almacenan, todos, en la memoria integrada, tal como la ROM y la RAM del dispositivo de codificación del habla 11a.

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 de QMF para obtener una señal del dominio del 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 intensidad y calcula p(r). Como procedimiento alternativo, la potencia a corto plazo puede calcularse también de acuerdo a la siguiente expresión (12) utilizando la señal en el dominio de 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, también puede calcularse por el mismo procedimiento que para calcular T(r) por parte de la unidad de detección de cambio de señal 2e del dispositivo de decodificación del habla 21. El cambio de señal también puede detectarse utilizando otros procedimientos 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 los 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 procedimiento que para calcular T(r) por parte de la unidad de detección de cambio de señal 2e del dispositivo de decodificación del habla 21, y determinar el valor de K(r) utilizándolos. En este caso, por ejemplo, K(r) puede obtenerse de acuerdo a la siguiente expresión (13), en donde ε es una constante, tal como 3,0.

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

(Modificación 2 del primer ejemplo)

Un dispositivo de codificación del 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 del habla de la modificación 2, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de codificación del habla de la modificación 2, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla de la modificación 2 recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla, y emite un flujo de bits multiplexados codificados al exterior del dispositivo de

codificación del habla.

El dispositivo de codificación del habla de la modificación 2 incluye funcionalmente una unidad de codificación diferencial de coeficientes de predicción lineal (medio de cálculo de información suplementaria de envolvente temporal), y una unidad de multiplexado de flujos de bits (medio de multiplexado de flujos de bits) que recibe una salida desde la unidad de codificación diferencial de coeficientes de predicción lineal, que no está ilustrada, en lugar de la unidad de cálculo de parámetros de intensidad de filtro 1f y la unidad de multiplexado de flujos de bits 1g del dispositivo de codificación del habla 11. Desde la unidad de transformación de frecuencia 1a hasta la unidad de análisis de predicción lineal 1e, la unidad de codificación diferencial de coeficientes de predicción lineal y la unidad de multiplexado de flujos de bits del dispositivo de codificación del habla de la modificación 2, son funciones realizadas cuando la CPU del dispositivo de codificación del habla de la modificación 2 ejecuta el programa informático almacenado en la memoria integrada del dispositivo de codificación del habla de la modificación 2. Diferentes tipos de datos necesarios para ejecutar el programa informático y diferentes tipos de datos generados al ejecutar el programa informático se almacenan, todos, en la memoria integrada, tal como la ROM y la RAM del dispositivo de codificación del habla de la modificación 2.

La unidad de decodificación diferencial de coeficientes de predicción lineal calcula valores diferenciales $a_D(n, r)$ del coeficiente de predicción lineal de acuerdo a 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 diferencial de coeficientes de predicción lineal cuantiza los $a_D(n, r)$, y los transmite a la unidad de multiplexado de flujos de bits (estructura correspondiente a la unidad de multiplexado de flujos de bits 1g). La unidad de multiplexado de flujos de bits multiplexa los $a_D(n, r)$ en el flujo de bits en lugar de $K(r)$, y emite el flujo de bits multiplexado al exterior del dispositivo de codificación del habla, a través del dispositivo de comunicación integrado.

Un dispositivo de decodificación del 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 del habla de la modificación 2, cargando y ejecutando un programa informático predeterminado, almacenado en una memoria integrada del dispositivo de decodificación del habla de la modificación 2, como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla de la modificación 2 recibe el flujo de bits multiplexados codificados, emitido desde el dispositivo de codificación del habla 11, el dispositivo de codificación del habla 11a de acuerdo a la modificación 1 o el dispositivo de codificación del habla de acuerdo a la modificación 2, y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla.

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

La unidad de decodificación diferencial de coeficientes de predicción lineal obtiene los $a_{adj}(n, r)$ decodificados de forma diferencial, de acuerdo a la siguiente expresión (15), utilizando los $a_L(n, r)$ obtenidos desde la unidad de análisis de predicción lineal de baja frecuencia 2d y los $a_D(n, r)$ recibidos desde la unidad de separación de flujos 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 diferencial de coeficientes de predicción lineal transmite los $a_{adj}(n, r)$ decodificados diferencialmente de esta forma a la unidad de filtro de predicción lineal 2k. Un $a_D(n, r)$ 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, tal como LSP (Par de Espectro Lineal), ISP (Par de Espectro de Impedancia/Admitancia), LSF (Frecuencia de Espectro Lineal), ISF (Frecuencia de Espectro de Impedancia/Admitancia) y el coeficiente PARCOR (Correlación Parcial), un $a_D(n, r)$ puede ser un valor que muestra

una diferencia de los mismos. En este caso, la decodificación diferencial también tiene la misma forma de expresión.

(Segundo ejemplo)

5 La figura 6 es un diagrama que ilustra un dispositivo de codificación del habla 12 de acuerdo a un segundo ejemplo. El dispositivo de codificación del 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 del habla 12, cargando y ejecutando un programa informático predeterminado (como, por ejemplo, un programa informático para realizar procesos, ilustrado en el diagrama de flujo de la figura 7), almacenado en una memoria integrada del dispositivo de codificación del habla 12, como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 12 recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 12, y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla 12.

15 El dispositivo de codificación del habla 12 incluye funcionalmente una unidad de diezmado de coeficientes de predicción lineal 1j (medio de diezmado de coeficientes de predicción), una unidad de cuantización de coeficientes de predicción lineal 1k (medio de cuantización de coeficientes de predicción) y una unidad de multiplexado de flujos de bits 1g2 (medio de multiplexado de flujos de bits), en lugar de la unidad de cálculo de parámetros de intensidad de filtro 1f y la unidad de multiplexado de flujos de bits 1g del dispositivo de codificación del habla 11. Desde 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 diezmado de coeficientes de predicción lineal 1j, la unidad de cuantización de coeficientes de predicción lineal 1k y la unidad de multiplexado de flujos de bits 1g2 del dispositivo de codificación del habla 12, ilustradas en la figura 6, son funciones realizadas cuando la CPU del dispositivo de codificación del habla 12 ejecuta el programa informático almacenado en la memoria integrada en el dispositivo de codificación del habla 12. La CPU del dispositivo de codificación del habla 12 ejecuta procesos secuencialmente (procesos desde la etapa Sa1 hasta la etapa Sa5, y procesos desde la etapa Sc1 hasta la etapa Sc3), ilustrados en el diagrama de flujo de la figura 7, ejecutando el programa informático (o utilizando desde la unidad de transformación de frecuencia 1a hasta la unidad de análisis de predicción lineal 1e, la unidad de diezmado de coeficientes de predicción lineal 1j, la unidad de cuantización de coeficientes de predicción lineal 1k y la unidad de multiplexado de flujos de bits 1g2 del dispositivo de codificación del habla 12 ilustrado en la figura 6). Diferentes tipos de datos necesarios para ejecutar el programa informático y diferentes tipos de datos generados al ejecutar el programa informático se almacenan, todos, en la memoria integrada, como la ROM y la RAM del dispositivo de codificación del habla 12.

35 La unidad de diezmado del coeficiente de predicción lineal 1j diezma los $a_H(n, r)$ obtenidos desde 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 cuantización de coeficientes de predicción lineal 1k (proceso en la etapa Sc1). Se observa que $0 \leq i < N_{ts}$, y N_{ts} es el número de ranuras de tiempo en una trama para la cual se transmite $a_H(n, r)$. El diezmado de los coeficientes de predicción lineal puede ser realizado en un intervalo de tiempo predeterminado, o puede realizarse en un intervalo de tiempo no uniforme, en base a las características de $a_H(n, r)$. Por ejemplo, es posible un procedimiento que compare $G_H(r)$ de $a_H(n, r)$ en una trama con una cierta longitud, y haga que $a_H(n, r)$, al cual $G_H(r)$ supera en un cierto valor, sea un objeto de cuantización. Si el intervalo de diezmado 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.

45 La unidad de cuantización de coeficientes de predicción lineal 1k cuantiza los coeficientes de predicción lineal de alta frecuencia diezmados $a_H(n, r_i)$, recibidos desde la unidad de diezmado de coeficientes de predicción lineal 1j, y los índices r_i de las ranuras de tiempo correspondientes, y los transmite a la unidad de multiplexado de flujos de bits 1g2 (proceso en la etapa Sc2). Como estructura alternativa, en lugar de cuantizar los $a_H(n, r_i)$, pueden cuantizarse valores diferenciales $a_D(n, r_i)$ de los coeficientes de predicción lineal, como el dispositivo de codificación del habla de acuerdo a la modificación 2 del primer ejemplo.

55 La unidad de multiplexado de flujos de bits 1g2 multiplexa el flujo de bits codificados, calculado por la unidad central de codificación de códecs 1c, la información suplementaria de SBR calculada por la unidad de codificación de SBR 1d y los índices $\{r_i\}$ de las ranuras de tiempo correspondientes a los $a_H(n, r_i)$ que se cuantizan y reciben desde la unidad de cuantización de coeficientes de predicción lineal 1k, en un flujo de bits y emite el flujo de bits multiplexado a través del dispositivo de comunicación del dispositivo de codificación del habla 12 (proceso en la etapa Sc3).

60 La figura 8 es un diagrama que ilustra un dispositivo de decodificación del habla 22 de acuerdo al segundo ejemplo. El dispositivo de decodificación del 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 del habla 22, cargando y ejecutando un programa informático predeterminado (como, por ejemplo, un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 9), almacenado en una memoria integrada del dispositivo de decodificación del habla 22, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 22 recibe el flujo de bits multiplexados codificados, emitido desde el dispositivo de codificación del habla 12, y emite una señal de habla decodificada al exterior del dispositivo de

codificación del habla 12.

El dispositivo de decodificación del habla 22 incluye funcionalmente una unidad de separación de flujos de bits 2a1 (medio de separación de flujos de bits), una unidad de interpolación/extrapolación de coeficientes de predicción lineal 2p (medio de interpolación/extrapolación de coeficientes de predicción lineal) y una unidad de filtro de predicción lineal 2k1 (medio de formación de la envolvente temporal), en lugar de la unidad de separación de flujos 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 del habla 21. La unidad de separación de flujos de bits 2a1, la unidad de decodificación central de códecs 2b, la unidad de transformación de frecuencia 2c, desde 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 coeficientes 2m, la unidad de transformación inversa de frecuencia 2n y la unidad de interpolación/extrapolación de coeficientes de predicción lineal 2p del dispositivo de decodificación del habla 22, ilustrado en la figura 8, son funciones realizadas cuando la CPU del dispositivo de codificación del habla 12 ejecuta el programa informático almacenado en la memoria integrada del dispositivo de codificación del habla 12. La CPU del dispositivo de decodificación del habla 22 ejecuta secuencialmente los procesos (procesos desde la etapa Sb1 hasta la etapa Sd2, etapa Sd1, desde la etapa Sb5 hasta la etapa Sb8, etapa Sd2, y desde la etapa Sb10 hasta la etapa Sb11) ilustrados en el diagrama de flujo de la figura 9, ejecutando el programa informático (o utilizando la unidad de separación de flujos de bits 2a1, la unidad de decodificación central de códecs 2b, la unidad de transformación de frecuencia 2c, desde 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 coeficientes 2m, la unidad de transformación inversa de frecuencia 2n, y la unidad de interpolación/extrapolación de coeficientes de predicción lineal 2p, ilustradas en la figura 8). Diferentes tipos de datos necesarios para ejecutar el programa informático y diferentes tipos de datos generados al ejecutar el programa informático se almacenan, todos, en la memoria integrada, como la ROM y la RAM del dispositivo de decodificación del habla 22.

El dispositivo de decodificación del habla 22 incluye la unidad de separación de flujos de bits 2a1, la unidad de interpolación/extrapolación de coeficientes de predicción lineal 2p y la unidad de filtro de predicción lineal 2k1, en lugar de la unidad de separación de flujos 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 del habla 22.

La unidad de separación de flujos de bits 2a1 separa el flujo de bits multiplexado, suministrado a través del dispositivo de comunicación del dispositivo de decodificación del habla 22, en los índices r_i de las ranuras de tiempo correspondientes a los $a_H(n, r_i)$ que se cuantizan, la información suplementaria de SBR y el flujo de bits codificados.

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

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

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

La unidad de interpolación/extrapolación de coeficientes de predicción lineal 2p puede interpolar los coeficientes de predicción lineal, por ejemplo de acuerdo a la siguiente expresión (17), donde 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 de coeficientes de predicción lineal 2p puede convertir los coeficientes de predicción lineal en otras formas de expresión, tales como LSP (Par de Espectro Lineal), ISP (Par de Espectro de Impedancia/Admitancia), LSF (Frecuencia de Espectro Lineal), ISF (Frecuencia de Espectro de Impedancia/Admitancia) y coeficientes PARCOR (Correlación Parcial), 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 de predicción lineal en la unidad de filtro inverso de predicción lineal 2i. Si $a_D(n, r_i)$ es multiplexado en un flujo de bits en lugar de $a_H(n, r)$, la unidad de interpolación/extrapolación de coeficientes de predicción lineal 2p realiza la decodificación diferencial de manera parecida a la del dispositivo de decodificación del 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 los $q_{adj}(n, r)$ emitidos desde la unidad de ajuste de alta frecuencia 2j, utilizando los $a_H(n, r)$ que se interpolan o extrapolan, obtenidos desde la unidad de interpolación/extrapolación de coeficientes de predicción lineal 2p (proceso en la etapa 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 forma la envolvente temporal de los componentes de alta frecuencia generados por la SBR, realizando un filtrado de síntesis de predicción lineal, como la unidad de filtro de predicción lineal 2k del dispositivo de decodificación del habla 21.

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

(Tercer ejemplo)

La figura 10 es un diagrama que ilustra un dispositivo de codificación del habla 13 de acuerdo a un tercer ejemplo. El dispositivo de codificación del 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 del habla 13, cargando y ejecutando un programa informático predeterminado (como, por ejemplo, un programa informático para realizar los procesos ilustrados en el diagrama de flujo de la figura 11) almacenado en una memoria integrada del dispositivo de codificación del habla 13, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 13 recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 13, y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla 13.

El dispositivo de codificación del habla 13 incluye funcionalmente una unidad de cálculo de envolvente temporal 1m (medio de cálculo de información suplementaria de envolvente temporal), una unidad de cálculo de parámetros de forma de envolvente 1n (medio de cálculo de información suplementaria de envolvente temporal) y una unidad de multiplexado de flujos de bits 1g3 (medio de multiplexado de flujos 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 multiplexado de flujos de bits 1g del dispositivo de codificación del habla 11. Desde la unidad de transformación de frecuencia 1a hasta la unidad de codificación de SBR 1d, la unidad de cálculo de la envolvente temporal 1m, la unidad de cálculo de parámetros de forma de la envolvente 1n y la unidad de multiplexado de flujos de bits 1g3 del dispositivo de codificación del habla 13, ilustrado en la figura 10, son funciones realizadas cuando la CPU del dispositivo de codificación del habla 12 ejecuta el programa informático almacenado en la memoria integrada del dispositivo de codificación del habla 13. La CPU del dispositivo de codificación del habla 13 ejecuta procesos secuencialmente (procesos desde la etapa Sa1 hasta la etapa Sa4, y desde la etapa Se1 hasta la etapa Se3), ilustrados en el diagrama de flujo de la figura 11, ejecutando el programa informático (o utilizando desde la unidad de transformación de frecuencia 1a hasta la unidad de análisis de codificación de SBR 1d, la unidad de cálculo de envolvente temporal 1m, la unidad de cálculo de parámetros de forma de envolvente 1n y la unidad de multiplexado de flujos de bits 1g3 del dispositivo de codificación del habla 13 ilustrado en la figura 10). Diferentes tipos de datos necesarios para ejecutar el programa informático y diferentes tipos de datos generados al ejecutar el programa informático se almacenan, todos, en la memoria integrada, tal como la ROM y la RAM del dispositivo de codificación del habla 13.

La unidad de cálculo de envolvente temporal 1m recibe $q(k, r)$ y, por ejemplo, obtiene información de la envolvente 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 la etapa Se1). En este caso, $e(r)$ se obtiene de acuerdo a la siguiente expresión (19).

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

La unidad de cálculo de parámetros de forma de envolvente 1n recibe $e(r)$ desde la unidad de cálculo de la

envolvente temporal 1m y recibe límites de tiempo de envolvente de SBR {b_i} desde la unidad de codificación de SBR 1d. Debe observarse que 0 ≤ i ≤ Ne, y Ne es el número de envolventes de SBR en la trama codificada. La unidad de cálculo de parámetros de forma de envolvente 1n obtiene un parámetro de forma de envolvente s(i) (0 ≤ i < Ne) de cada una de las envolventes de SBR en la trama codificada de acuerdo a la siguiente expresión (20) (proceso en la etapa Se2). El parámetro de forma de envolvente s(i) corresponde a la información suplementaria de envolvente temporal, y es similar en el tercer ejemplo.

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

10 Debe observarse que:

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

15 donde s(i) en la expresión anterior es un parámetro que indica la magnitud de la variación de e(r) en la envolvente de la i-ésima SBR que cumple b_i ≤ r < b_{i+1}, y e(r) tiene un número mayor a medida que se incrementa la variación de la envolvente temporal. Las expresiones (20) y (21) descritas anteriormente son ejemplos del procedimiento para calcular s(i) y, por ejemplo, s(i) puede obtenerse también utilizando, por ejemplo, la SMF (Medida de Uniformidad Espectral) de e(r), una razón entre el valor máximo y el valor mínimo, y similares. A continuación, se cuantiza s(i) y se transmite a la unidad de multiplexado de flujos de bits 1g3.

20 La unidad de multiplexado de flujos de bits 1g3 multiplexa el flujo de bits codificados, calculado por la unidad central de codificación de códecs 1c, la información suplementaria de SBR calculada por la unidad de codificación de SBR 1d y s(i) en un flujo de bits, y emite el flujo de bits multiplexado a través del dispositivo de comunicación del dispositivo de codificación del habla 13 (proceso en la etapa Se3).

25 La figura 12 es un diagrama que ilustra un dispositivo de decodificación del habla 23 de acuerdo al tercer ejemplo. El dispositivo de decodificación del 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 del habla 23, cargando y ejecutando un programa informático predeterminado (tal como, por ejemplo, un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 13) almacenado en una memoria integrada del dispositivo de decodificación del habla 23, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 23 recibe el flujo de bits multiplexados codificados, emitido desde el dispositivo de codificación del habla 13, y produce una señal de habla decodificada al exterior del dispositivo de decodificación del habla 13.

35 El dispositivo de decodificación del habla 23 incluye funcionalmente una unidad de separación de flujos de bits 2a2 (medio de separación de flujos de bits), una unidad de cálculo de envolvente temporal de baja frecuencia 2r (medio de análisis de envolvente temporal de baja frecuencia), una unidad de ajuste de forma de envolvente 2s (medio de ajuste de envolvente temporal), una unidad de cálculo de envolvente temporal de alta frecuencia 2t, una unidad de allanamiento de envolvente temporal 2u y una unidad de formación de envolvente temporal 2v (medio de formación de la envolvente temporal), en lugar de la unidad de separación de flujos 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 del habla 21. Tal como se menciona en el párrafo [0113], el dispositivo de decodificación del habla 24 de acuerdo a la realización de la presente invención incluye funcionalmente la estructura del dispositivo de decodificación del habla 23. La unidad de cálculo de envolvente temporal de baja frecuencia 2r es una realización del medio de análisis de envolvente temporal de baja frecuencia de la presente invención. La unidad de ajuste de forma de envolvente 2s es una realización del medio de ajuste de envolvente temporal de la presente invención. La unidad de formación de la envolvente temporal 2v es una realización del medio de formación de la envolvente temporal de la presente invención. La unidad de separación de flujos de bits 2a2, desde la unidad de decodificación central 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 coeficientes 2m, la unidad de transformación inversa de frecuencia 2n, y desde la unidad de cálculo de envolvente temporal de baja frecuencia 2r hasta la unidad de formación de la envolvente temporal 2v del dispositivo de decodificación del habla 23, ilustrado en la figura 12, son

funciones realizadas cuando la CPU del dispositivo de codificación del habla 12 ejecuta el programa informático almacenado en la memoria integrada del dispositivo de codificación del habla 12. La CPU del dispositivo de decodificación del habla 23 ejecuta procesos secuencialmente (procesos desde la etapa Sb1 hasta la etapa Sb2, desde la etapa Sf1 a la etapa Sf2, etapa Sb5, desde la etapa Sf3 a la etapa Sf4, etapa Sb8, etapa Sf5 y desde la etapa Sb10 hasta la etapa Sb11), ilustrados en el diagrama de flujo de la figura 13, ejecutando el programa informático (o utilizando la unidad de separación de flujos de bits 2a2, desde la unidad de decodificación central 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 coeficientes 2m, la unidad de transformación inversa de frecuencia 2n, y desde la unidad de cálculo de envolvente temporal de baja frecuencia 2r hasta la unidad de formación de la envolvente temporal 2v del dispositivo de decodificación del habla 23, ilustrado en la figura 12). Diferentes tipos de datos necesarios para ejecutar el programa informático y diferentes tipos de datos generados al ejecutar el programa informático se almacenan, todos, en la memoria integrada, tal como la ROM y la RAM del dispositivo de codificación del habla 23.

La unidad de separación de flujos de bits 2a2 separa el flujo de bits multiplexados, suministrado a través del dispositivo de comunicación del dispositivo de decodificación del habla 23, en $s(i)$, la información suplementaria de SBR y el flujo de bits codificados. La unidad de cálculo de envolvente temporal de baja frecuencia 2r recibe los $q_{dec}(k, r)$, incluyendo los componentes de baja frecuencia desde la unidad de transformación de frecuencia 2c, y obtiene $e(r)$ de acuerdo a la siguiente expresión (22) (proceso en la etapa Sf1).

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

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

$$\begin{aligned} e_{adj}(r) &= \overline{e(i)} + \sqrt{s(i) - v(i)} \cdot (e(r) - \overline{e(i)}) \quad (s(i) > v(i)) \\ e_{adj}(r) &= e(r) \quad (\text{de lo contrario}) \end{aligned} \quad \text{---(23)}$$

Debe hacerse notar que:

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

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

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

La unidad de cálculo de envolvente temporal de alta frecuencia 2t calcula una envolvente temporal $e_{exp}(r)$ utilizando los $q_{exp}(k, r)$ obtenidos desde la unidad de generación de alta frecuencia 2g, de acuerdo a la siguiente expresión (26) (proceso en la etapa Sf3).

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

La unidad de allanamiento de envolvente temporal 2u allana la envolvente temporal de los $q_{\text{exp}}(k, r)$ obtenidos desde la unidad de generación de alta frecuencia 2g de acuerdo a la siguiente expresión (27), y transmite la señal obtenida $q_{\text{flat}}(k, r)$ en el dominio de QMF a la unidad de ajuste de alta frecuencia 2j (proceso en la etapa Sf4).

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

El allanamiento de la envolvente temporal por parte de la unidad de allanamiento de envolvente temporal 2u también puede omitirse. En lugar de calcular la envolvente temporal de los componentes de alta frecuencia de la salida desde la unidad de generación de alta frecuencia 2g y allanar la envolvente temporal de los mismos, puede calcularse la envolvente temporal de los componentes de alta frecuencia de una salida desde la unidad de ajuste de alta frecuencia 2j, y la envolvente temporal de la misma puede ser allanada. La envolvente temporal utilizada en la unidad de allanamiento de envolvente temporal 2u puede ser también $e_{\text{adj}}(r)$, obtenida desde la unidad de ajuste de forma de envolvente 2s, en lugar de la $e_{\text{exp}}(r)$ obtenida desde la unidad de cálculo de envolvente temporal de alta frecuencia 2t.

La unidad de formación de la envolvente temporal 2v forma los $q_{\text{adj}}(k, r)$ obtenidos desde la unidad de ajuste de alta frecuencia 2j utilizando los $e_{\text{adj}}(r)$ obtenidos desde la unidad de formación de la envolvente temporal 2v, y obtiene una señal $q_{\text{envadj}}(k, r)$ en el dominio de QMF en el cual se forma la envolvente temporal (proceso en la etapa Sf5). Tal como se ha mencionado anteriormente, la unidad de formación de envolvente temporal 2v es una realización del medio de formación de envolvente temporal de la presente invención. La formación se realiza de acuerdo a la siguiente expresión (28). La $q_{\text{envadj}}(k, r)$ se transmite a la unidad de adición de coeficientes 2m como una señal en el dominio de QMF correspondiente a los componentes de alta frecuencia.

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

(Realización)

La figura 14 es un diagrama que ilustra un dispositivo de decodificación del habla 24 de acuerdo a una realización de la presente invención. El dispositivo de decodificación del 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 del habla 24, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de decodificación del habla 24, tal como, por ejemplo, la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24 recibe la salida del flujo de bits multiplexados codificados desde el dispositivo de codificación del habla 11 o desde el dispositivo de codificación del habla 13, y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24.

El dispositivo de decodificación del habla 24 según la realización de la presente invención incluye funcionalmente la estructura del dispositivo de decodificación del habla 21 según el primer ejemplo (la unidad de decodificación central 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 coeficientes 2m y la unidad de transformación inversa de frecuencia 2n) y la estructura del dispositivo de decodificación del habla 23 según el tercer ejemplo (la unidad de cálculo de envolvente temporal de baja frecuencia 2r, la unidad de ajuste de forma de envolvente 2s y la unidad de formación de envolvente temporal 2v). El dispositivo de decodificación del habla 24 también incluye una unidad de separación de flujos de bits 2a3 (medio de separación de flujos de bits) y una unidad de conversión de información suplementaria 2w. La unidad de conversión de información suplementaria 2w es una realización del medio de conversión de información suplementaria de la presente invención. El orden de la unidad de filtro de predicción lineal 2k y de la unidad de formación de envolvente temporal 2v puede ser opuesto al ilustrado en la figura 14. El dispositivo de decodificación del habla 24 recibe, preferiblemente, el flujo de bits codificados por el dispositivo de codificación del habla 11 o el dispositivo de codificación del habla 13. La estructura del dispositivo de decodificación del habla 24 ilustrado en la figura 14 es una función realizada cuando la CPU del dispositivo de decodificación del habla 24 ejecuta el programa informático almacenado en la memoria integrada del dispositivo de decodificación del habla 24. Diferentes tipos de

datos necesarios para ejecutar el programa informático y diferentes tipos de datos generados al ejecutar el programa informático se almacenan, todos, en la memoria integrada, tal como la ROM y la RAM del dispositivo de decodificación del habla 24.

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

10 La unidad de conversión de información suplementaria 2w convierte la información suplementaria de envolvente temporal suministrada para obtener $K(r)$ y $s(i)$. Si la información suplementaria de la envolvente 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 $b_i \leq r < b_{i+1}$

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

20 y convertir el valor medio representado en la expresión (29) en $s(i)$, utilizando una tabla predeterminada. Si la información suplementaria de la envolvente 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 a fin de cumplir la relación de $b_i \leq r < b_{i+1}$.

25 Si la información suplementaria de la envolvente 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 envolvente de SBR. Las tablas para convertir $X(r)$ en $K(r)$ y $s(i)$ pueden ser distintas entre sí.

30 (Modificación 3 del primer ejemplo)

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

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

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

$$P_0(r) = \sum_{n=k_x}^{63} |q_{adj}(n, r)|^2 \quad \text{---(31)}$$

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

50 Al llevar a cabo el proceso de control de ganancia automático, la potencia de los componentes de alta frecuencia de la señal emitida desde 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 el cual se forma la envolvente temporal de los componentes de alta frecuencia generados en base a la SBR. El proceso de control de ganancia automático también puede realizarse de forma individual en un espectro de frecuencia específico de la señal en el dominio de 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-ésima 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 de QMF). F_i indica el límite del espectro de frecuencia, y es preferible que F_i sea una tabla de límites de frecuencia de un factor de escala de envolvente definido en la SBR en la "MPEG4 AAC". La tabla de límites de frecuencia está definida por la unidad de generación de alta frecuencia 2g en base a la definición de SBR en la "MPEG4 AAC". Al realizar el proceso de control de ganancia 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 filtrado de predicción lineal. 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 desde la unidad de filtro de predicción lineal 2k en la cual se forma la envolvente temporal de los componentes de alta frecuencia generados en base a la 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 envolvente 1n en el dispositivo de codificación del habla 13 del tercer ejemplo puede también realizarse mediante el proceso siguiente. La unidad de cálculo de parámetros de forma de envolvente 1n obtiene un parámetro de forma de envolvente $s(i)$ ($0 \leq i < N_e$) de acuerdo a la siguiente expresión (33) para cada envolvente de SBR en la trama 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 la envolvente de SBR, y el procedimiento de cálculo está basado en la expresión (21). Debe observarse que la envolvente de SBR indica el segmento de tiempo que cumple $b_i \leq r < b_{i+1}$. Los $\{b_i\}$ son los límites de tiempo de las envolventes 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 envolvente 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 de la gama de $b_i \leq r < b_{i+1}$. Por consiguiente, en este caso, el parámetro de forma de envolvente $s(i)$ es un parámetro para indicar una razón entre el valor mínimo y el valor medio de la información de la envolvente temporal ajustada en la envolvente de SBR. La unidad de ajuste de forma de envolvente 2s en el dispositivo de decodificación del habla 23 del tercer ejemplo también puede realizarse por medio del proceso siguiente. La unidad de ajuste de forma de envolvente 2s ajusta $e(r)$ utilizando $s(i)$ para obtener la información de la envolvente temporal ajustada $e_{adj}(r)$. El procedimiento de ajuste se basa en las siguientes expresión (35) o expresión (36).

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

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

La expresión 35 ajusta la forma de la envolvente para que la razón entre el valor mínimo y el valor medio de la información de la envolvente temporal ajustada $e_{adj}(r)$ en la envolvente de SBR se tome equivalente al valor del parámetro de forma de envolvente $s(i)$. Los cambios realizados en la modificación 1 del tercer ejemplo descrito anteriormente también pueden realizarse en la realización.

[Modificación de la realización]

La unidad de formación de la envolvente 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 la ganancia de la información de la envolvente temporal ajustada $e_{adj}(r)$, de manera que la potencia de $q_{envadj}(k, r)$ mantenga la de $q_{adj}(k, r)$ dentro de la envolvente de SBR. Tal como se indica en la expresión (38), en la presente modificación de la realización, $q_{envadj}(k, r)$ se obtiene multiplicando la señal $q_{adj}(k, r)$, en el dominio de QMF, por $e_{adj, scaled}(r)$ en lugar de $e_{adj}(r)$. Por consiguiente, la unidad de formación de la envolvente temporal $2v$ puede formar la envolvente temporal de la señal $q_{adj}(k, r)$ en el dominio de QMF, de manera que la potencia de señal dentro de la envolvente de SBR se torne equivalente antes y después de la formación de la envolvente temporal. Debe observarse que la envolvente de SBR indica el segmento de tiempo que cumple $b_i \leq r < b_{i+1}$. Los $\{b_i\}$ son los límites de tiempo de las envolventes de SBR incluidas en la información suplementaria de SBR como información, y son los límites del segmento de tiempo para el cual se dan el factor de escala de envolvente de SBR que representa la energía de señal media de un segmento de tiempo determinado y un espectro de frecuencia determinado. La terminología "envolvente de SBR" en las realizaciones de la presente invención corresponde a la terminología "segmento de tiempo de envolvente de SBR" en la "MPEG4 AAC", definida en el documento "ISO/IEC 14496-3", y la "envolvente de SBR" tiene el mismo contenido que el "segmento de tiempo de envolvente de SBR" en toda la extensión de las realizaciones.

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

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

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

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

Los cambios realizados en la presente modificación de la realización descrita anteriormente también pueden hacerse en el tercer ejemplo.

(Modificación 2 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).

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

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

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

5

Cuando se utilizan la expresión (39) y la expresión (40), la información de la envolvente temporal $e(r)$ es información en la cual la potencia de cada muestra de sub-banda de QMF es normalizada por la potencia media en la envolvente de SBR, y se extrae la raíz cuadrada. Sin embargo, la muestra de sub-banda de QMF es un vector de señal correspondiente al índice de tiempo "r" en la señal del dominio de QMF, y es una sub-muestra en el dominio de 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 sub-banda de QMF". En este caso, la información de la envolvente temporal $e(r)$ es un coeficiente de ganancia que debería ser multiplicado por cada muestra de sub-banda de QMF, y lo mismo vale para la información de la envolvente temporal ajustada $e_{adj}(r)$.

10

15

(Modificación 2 de la realización)

Un dispositivo de decodificación del habla 24a (que no se ilustra) 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 del habla 24a, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de decodificación del habla 24a, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24a recibe la salida de flujo de bits multiplexados codificados desde el dispositivo de codificación del habla 11 o del dispositivo de codificación del habla 13, y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24a. El dispositivo de decodificación del habla 24a incluye funcionalmente una unidad de separación de flujos de bits 2a4 (no ilustrada) en lugar de la unidad de separación de flujos de bits 2a3 del dispositivo de decodificación del habla 24, e incluye también una unidad de generación de información suplementaria de envolvente temporal 2y (no ilustrada) en lugar de la unidad de conversión de información suplementaria 2w. La unidad de separación de flujos de bits 2a4 separa el flujo de bits multiplexado en la información de SBR y el flujo de bits codificados. La unidad de generación de información suplementaria de envolvente temporal 2y genera información suplementaria de envolvente temporal en base a la información incluida en el flujo de bits codificados y la información suplementaria de SBR.

20

25

30

Para generar la información suplementaria de envolvente temporal en una envolvente de SBR determinada, por ejemplo, el ancho del tiempo $(b_{i+1}-b_i)$ de la envolvente de SBR, puede utilizarse una clase de trama, un parámetro de intensidad del filtro inverso, un suelo de ruido, la amplitud de la potencia de alta frecuencia, una razón entre la potencia de alta frecuencia y la potencia de baja frecuencia, un coeficiente de auto-correlación o una ganancia de predicción de un resultado de la realización de un análisis de predicción lineal en la dirección de frecuencia en una señal de baja frecuencia representada en el dominio de QMF, y similares. La información suplementaria de envolvente temporal puede ser generada determinando $K(r)$ o $s(i)$ en base a uno o a una pluralidad de valores de los parámetros. Por ejemplo, la información suplementaria de envolvente temporal puede ser generada determinando $K(r)$ o $s(i)$ en base a $(b_{i+1}-b_i)$ de manera que $K(r)$ o $s(i)$ sea reducido a medida que se incrementa el ancho del tiempo $(b_{i+1}-b_i)$ de la envolvente de SBR, o $K(r)$ o $s(i)$ se incrementa a medida que el ancho del tiempo $(b_{i+1}-b_i)$ de la envolvente de SBR se incrementa. Los cambios similares pueden hacerse también para el primer ejemplo y el tercer ejemplo.

35

40

45

(Modificación 3 de la realización)

Un dispositivo de decodificación del habla 24b (véase la figura 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 del habla 24b, cargando y ejecutando un programa informático predeterminado, almacenado en una memoria integrada del dispositivo de decodificación del habla 24b, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24b recibe el flujo de bits multiplexados codificados, enviado desde el dispositivo de codificación del habla 11 o el dispositivo de codificación del habla 13, y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24b. El dispositivo de decodificación del habla 24b, tal como se ilustra en la figura 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 de QMF de la banda de alta frecuencia, realizando un filtrado inverso de predicción lineal en la dirección temporal, el ajuste de ganancia y la adición de ruido, descrito en la etapa de "generación de HF" y la etapa de "ajuste de HF" en la SBR en la "MPEG4 AAC". En este momento, la señal de salida de la unidad de ajuste de alta frecuencia primaria 2j1 corresponde a una señal W_2 en la descripción en la "herramienta de SBR" en el documento "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 de envolvente temporal 2v forman la envolvente 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 etapa de "ajuste de HF" en la SBR en la "MPEG4 AAC". El proceso de la unidad de ajuste de alta frecuencia secundaria corresponde a un proceso de generación de una señal Y a partir de la señal W_2 en la descripción en la "herramienta de SBR" en el documento "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 de envolvente temporal 2v.

En la descripción anterior, solamente el proceso para añadir sinusoides es realizado por la unidad de ajuste de alta frecuencia secundaria 2j2. Sin embargo, cualquiera de los procesos de la etapa 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 el primer ejemplo y el segundo ejemplo, pero la unidad de formación de envolvente temporal no está incluida. En consecuencia, una señal de salida desde la unidad primaria de ajuste de alta frecuencia 2j1 es procesada 25 por la unidad de filtro de predicción lineal, y a continuación una señal de salida desde 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 de envolvente temporal 2v, pero la unidad de filtro de predicción lineal 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 formación de envolvente temporal 2v y, a continuación, una señal de salida desde la unidad de formación de envolvente temporal 2v es procesada por la unidad secundaria de ajuste de alta frecuencia.

En el dispositivo de decodificación del habla (dispositivo de decodificación del habla 24, 24a o 24b) de la realización, el orden de procesamiento de la unidad de filtro de predicción lineal 2k y la unidad de formación de la envolvente temporal 2v puede invertirse. En otras palabras, una señal de salida desde 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 de la envolvente temporal 2v y, a continuación, una señal de salida desde la unidad de formación de envolvente temporal 2v puede ser procesada por la unidad de filtro de predicción lineal 2k.

Asimismo, solamente si la información suplementaria de envolvente temporal incluye información de control binaria para indicar si el proceso es realizado por la unidad de filtro de predicción lineal 2k o la unidad de formación de envolvente temporal 2v, y la información de control indica que el proceso sea realizado por la unidad de filtro de predicción lineal 2k o la unidad de formación de la envolvente temporal 2v, la información suplementaria de envolvente temporal puede utilizar una forma que incluye además al menos un elemento entre el parámetro de intensidad de filtro $K(r)$, el parámetro de forma de envolvente $s(i)$ o $X(r)$, que es un parámetro para determinar, tanto $K(r)$ como $s(i)$, como información.

(Modificación 4 de la realización)

Un dispositivo de decodificación del habla 24c (véase la figura 16) de una 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 del habla 24c, cargando y ejecutando un programa informático predeterminado (tal como, por ejemplo, un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 17), almacenado en una memoria integrada del dispositivo de decodificación del habla 24c, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24c recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de

5 decodificación del habla 24c. Tal como se ilustra en la figura 16, el dispositivo de decodificación del 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 de la envolvente temporal 2v (las unidades de ajuste de componentes de señal individual corresponden al medio de formación de envolvente temporal).

10 La unidad primaria de ajuste de alta frecuencia 2j3 emite una señal en el dominio de QMF de la banda de alta frecuencia como un componente de señal de copia. La unidad primaria de ajuste de alta frecuencia 2j3 puede emitir una señal sobre la cual se realiza al menos el filtrado inverso de predicción lineal en la dirección temporal y el ajuste de ganancia (ajuste de características de frecuencia), sobre la señal en el dominio de QMF de la banda de alta frecuencia, utilizando la información suplementaria de SBR recibida desde la unidad de separación de flujos 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 senoide, utilizando la información suplementaria de SBR proporcionada desde la unidad de separación de flujos de bits 2a3, y emite el componente de señal de copia, el componente de señal de ruido y el componente de señal senoide por separado (proceso en la etapa Sg1). El componente de señal de ruido y el componente de señal senoide pueden no ser generados, dependiendo del contenido de la información suplementaria de SBR.

15 20 Las unidades de ajuste de componentes de señal individual 2z1, 2z2 y 2z3 realizan el procesamiento en toda la pluralidad de componentes de señal incluidos en la salida desde el medio primario de ajuste de alta frecuencia (proceso en la etapa 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, de manera similar a la de 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 multiplicación de cada muestra de sub-banda de QMF por un coeficiente de ganancia, utilizando la envolvente temporal obtenida desde la unidad de ajuste de forma de envolvente 2s, de manera similar a la de la unidad de formación de envolvente 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 realización de un filtrado de síntesis de predicción lineal en la dirección de frecuencia sobre la señal de entrada, utilizando los coeficientes de predicción lineal obtenidos desde la unidad de ajuste de intensidad de filtro 2f, de manera similar a la de la unidad de filtro de predicción lineal 2k y, a continuación, multiplicando cada muestra de sub-banda de QMF por un coeficiente de ganancia, utilizando la envolvente temporal obtenida desde la unidad de ajuste de forma de envolvente 2s, de manera similar a la de la unidad de formación de envolvente 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 multiplicación de cada muestra de sub-banda de QMF con respecto a la señal de entrada, por un coeficiente de ganancia, utilizando la envolvente temporal obtenida desde la unidad de ajuste de forma de envolvente 2s, de manera similar a la de la unidad de formación de la envolvente 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 desde la unidad de ajuste de intensidad de filtro 2f, de manera 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 de envolvente 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 formar la envolvente temporal de la señal de entrada, utilizando un procedimiento distinto a 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).

50 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 formar la envolvente temporal de cada uno de los componentes de señal incluidos en la salida del medio primario de ajuste de alta frecuencia, mediante distintos procedimientos. Por ejemplo, pueden realizarse distintos procesos sobre la señal de copia, la señal de ruido y la señal senoide, de manera que la unidad de ajuste de componentes de señal individuales 2z1 realice el proceso 2 sobre la señal de copia proporcionada, la unidad de ajuste de componentes de señal individuales 2z2 realice el proceso 3 sobre el componente de señal de ruido proporcionado, y la unidad de ajuste de componentes de señal individuales 2z3 realice el proceso 5 sobre la señal senoide proporcionada. En este caso, la unidad de ajuste de intensidad de filtro 2f y la unidad de ajuste de forma de envolvente 2s pueden transmitir los mismos coeficientes de predicción lineal y las envolventes temporales a las unidades de ajuste de componentes de señal individual 2z1, 2z2 y 2z3, pero también pueden transferir diferentes coeficientes de predicción lineal y envolventes temporales. También resulta posible transmitir los mismos coeficientes de predicción lineal y envolventes temporales a al menos dos de las unidades de ajuste de componentes de señal individual 2z1, 2z2 y 2z3. Dado que al menos una de las unidades de ajuste de componentes de señal individual 2z1, 2z2 y 2z3 puede no realizar el proceso de formación de la envolvente temporal, sino emitir la señal de entrada tal como está (proceso 5), las unidades de ajuste de componentes de señal individual 2z1, 2z2 y 2z3 realizan el proceso de envolvente temporal en al menos uno entre la pluralidad de componentes de señal emitidos desde la unidad primaria de ajuste de alta frecuencia 2j3 como un todo (si todas las unidades de ajuste de componentes de señal individual 2z1, 2z2 y 2z3

2z3 realizan el proceso 5, el proceso de formación de envolvente temporal no se realiza en ninguno de los componentes de señal, y los efectos de la presente invención no se exhiben).

Los procesos realizados por cada una de las unidades de ajuste de componentes de señal individual 2z1, 2z2 y 2z3 pueden estar fijados para uno de los procesos 1 a 7, pero pueden estar determinados dinámicamente para realizar uno de los procesos 1 a 7, en base a la información de control recibida desde el exterior del dispositivo de decodificación del habla 24c. En este momento, resulta preferible que la información de control esté incluida en el flujo de bits multiplexado. La información de control puede ser una instrucción para realizar uno cualquiera de los procesos 1 a 7 en un segmento específico de tiempo de envolvente de SBR, en la trama codificada o en el otro segmento de tiempo, o puede ser una instrucción para realizar uno 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 los componentes de señal procesados, emitidos desde las unidades de ajuste de componentes de señal individual 2z1, 2z2 y 2z3, y emite el resultado a la unidad de adición de coeficientes (proceso en la etapa Sg3). La unidad secundaria de ajuste de alta frecuencia 2j4 puede realizar al menos uno entre el filtrado inverso de predicción lineal en la dirección temporal y el ajuste de ganancia (ajuste de características de frecuencia) en el componente de señal de copia, utilizando la información suplementaria de SBR recibida desde la unidad de separación de flujos 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 etapa intermedia añadiendo al menos dos componentes de señal en los cuales se realiza uno cualquiera de los procesos 1 a 7, y realizando a continuación uno 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 etapa intermedia y un componente de señal que todavía no ha sido añadido a la señal de salida en la etapa intermedia, y emite el resultado a la unidad de adición de coeficientes. Más específicamente, resulta preferible generar una señal de salida en la etapa 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 sinusoide a la señal de salida en la etapa intermedia, y emite el resultado a la unidad de adición de coeficientes.

La unidad primaria de ajuste de alta frecuencia 2j3 puede producir uno cualquiera entre una pluralidad de componentes de señal, en una forma por separado 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 sinusoide. En este caso, el componente de señal puede obtenerse añadiendo al menos dos entre el componente de señal de copia, el componente de señal de ruido y el componente de señal sinusoide. El componente de señal puede ser también una señal obtenida dividiendo la banda de uno entre el componente de señal de copia, el componente de señal de ruido y la señal sinusoide. 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 a tres.

La señal de alta frecuencia generada por la 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 sinusoide. Dado que la señal de copia, la señal de ruido y la señal sinusoide tienen envolventes temporales distintas entre sí, si la envolvente temporal de cada uno de los componente de señal se forma utilizando procedimientos 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 una envolvente temporal llana, y la señal de copia tiene una envolvente temporal próxima a la de la señal en la banda de baja frecuencia, las envolventes temporales de la señal de copia y de la señal de ruido pueden ser controladas de forma independiente, gestionándolas por separado y aplicando diferentes procesos a las mismas. Por consiguiente, resulta efectivo para mejorar la calidad subjetiva de la señal decodificada. Más específicamente, resulta preferible realizar un proceso de formación de la envolvente temporal en la señal de ruido (proceso 3 o proceso 4), realizar un proceso distinto al de la señal de ruido en la señal de copia (proceso 1 o proceso 2), y realizar el proceso 5 en la señal sinusoide (en otras palabras, no se realiza el proceso de formación de envolvente temporal). También resulta preferible realizar un proceso de formación (proceso 3 o proceso 4) de la envolvente temporal en la señal de ruido, y realizar el proceso 5 en la señal de copia y la señal sinusoide (en otras palabras, no se realiza el proceso de formación de envolvente temporal).

(Modificación 4 del primer ejemplo)

Un dispositivo de codificación del habla 11b (figura 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 del habla 11b, cargando y ejecutando un programa informático predeterminado, almacenado en una memoria integrada del dispositivo de codificación del habla 11b, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 11b recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 11b y emite un flujo de bits

multiplexados codificados al exterior del dispositivo de codificación del habla 11b. El dispositivo de codificación del 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 del habla 11b, y también incluye una unidad de selección de ranura de tiempo 1p.

5 La unidad de selección de ranura de tiempo 1p recibe una señal en el dominio de QMF desde 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 de QMF en la ranura de tiempo seleccionada como la
10 unidad de análisis de predicción lineal 1e, en base al resultado de selección transmitido desde la unidad de selección de ranura de tiempo 1p, para obtener al menos uno entre los coeficientes de predicción lineal de alta frecuencia y los coeficientes 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
15 ranura de tiempo seleccionada por la unidad de selección de ranura de tiempo 1p, obtenidos 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 procedimiento de selección que utilice la potencia de señal de la señal del dominio de QMF de los componentes de alta frecuencia, de manera similar a la de una unidad de selección de ranura de tiempo 3a 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 de QMF de los componentes
20 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 de QMF recibidas desde la unidad de transformación de frecuencia 1a. El procedimiento de selección de la ranura de tiempo puede ser al menos uno de los procedimientos descritos anteriormente, puede incluir al menos un procedimiento distinto a los que se han descrito anteriormente, o puede ser la combinación de los mismos.

25 Un dispositivo de decodificación del habla 21a (véase la figura 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 del habla 21a, cargando y ejecutando un programa informático predeterminado (tal como, por ejemplo, un programa informático para realizar procesos ilustrados en el
30 diagrama de flujo de la figura 19) almacenado en una memoria integrada del dispositivo de decodificación del habla 21a, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 21a recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 21a. Tal como se ilustra en la figura 18, el dispositivo de decodificación del habla 21a incluye una unidad de análisis de predicción lineal de baja frecuencia 2d1, una unidad de detección de cambio de
35 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 del habla 21, y también incluye la unidad de selección de ranura de tiempo 3a.

40 La unidad de selección de ranura de tiempo 3a determina si va a realizarse o no el filtrado de síntesis de predicción lineal en la unidad de filtro de predicción lineal 2k, sobre la señal $q_{exp}(k, r)$ en el dominio de 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 la
45 etapa Sh1). La unidad de selección de ranura de tiempo 3a notifica, sobre el resultado de la selección de la ranura de tiempo, a 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 del dominio de QMF en la ranura de tiempo seleccionada $r1$, de la misma manera que la unidad de análisis de predicción lineal de baja frecuencia 2d, en base al resultado de la
50 selección transmitido desde la unidad de selección de ranura de tiempo 3a, para obtener coeficientes de predicción lineal de baja frecuencia (proceso en la etapa Sh2). La unidad de detección de cambio de señal 2e1 detecta la variación temporal en la señal del dominio de QMF en la ranura de tiempo seleccionada, como la unidad de detección de cambio de señal 2e, en base al resultado de selección transmitido desde la unidad de selección de
55 ranura de tiempo 3a, y emite 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, obtenidos por la unidad de análisis de predicción lineal de baja frecuencia 2d1, para obtener unos coeficientes de
60 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 del dominio de QMF de los componentes de alta frecuencia generados por la unidad de generación de alta frecuencia 2g para la ranura de tiempo $r1$ seleccionada, en base al 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 2h, para obtener unos coeficientes de predicción
65 lineal de alta frecuencia $a_{exp}(n, r1)$ (proceso en la etapa 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 de QMF de los componentes de alta frecuencia de la ranura de tiempo r_1 seleccionada, como la unidad de filtro inverso de predicción lineal $2i$, en base al resultado de la selección transmitido desde la unidad de selección de ranura de tiempo $3a$ (proceso en la etapa $Sh4$).

5 La unidad de filtro de predicción lineal $2k3$ realiza un filtrado de síntesis de predicción lineal en la dirección de frecuencia en una señal $q_{adj}(k, r_1)$ en el dominio de QMF de los componentes de alta frecuencia emitidos desde la unidad de ajuste de alta frecuencia $2j$ en la ranura de tiempo r_1 seleccionada, utilizando los $a_{adj}(n, r_1)$ obtenidos desde la unidad de ajuste de intensidad de filtro $2f$, como la unidad de filtro de predicción lineal $2k$, en base al resultado de la selección transmitido desde la unidad de selección de ranura de tiempo $3a$ (proceso en la etapa $Sh5$). 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 al menos una ranura de tiempo r en la cual la potencia de señal de la señal del dominio de 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 a la siguiente expresión.

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

20 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 del tiempo predeterminado que incluye la ranura de tiempo r . El ancho del tiempo predeterminado también puede ser la envolvente de SBR.

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

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

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

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

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

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

40 por ejemplo, puede calcularse mediante la expresión siguiente

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

45 donde c es un valor predeterminado para definir una gama para calcular el valor medio. La potencia de señal máxima puede calcularse mediante el procedimiento descrito anteriormente, o puede calcularse con un procedimiento diferente.

50 Al menos una ranura de tiempo puede seleccionarse a partir de ranuras de tiempo incluidas en un ancho del tiempo t durante el cual la señal del dominio de 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} . Al menos una ranura de tiempo puede ser también seleccionada a partir de ranuras de tiempo incluidas en un ancho del tiempo t durante el cual la potencia de señal de la señal del dominio de QMF de los componentes de alta frecuencia cambia desde un estado transitorio con una gran variación a un estado fijo con una pequeña variación, y que son mayores que el 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 procedimiento descrito anteriormente, o pueden definirse utilizando diferentes procedimientos. El procedimiento de selección de ranura de tiempo puede ser al menos uno de los procedimientos descritos anteriormente, puede incluir al menos un procedimiento distinto a los que se han descrito anteriormente o puede ser una combinación de los mismos.

(Modificación 5 del primer ejemplo)

Un dispositivo de codificación del habla 11c (figura 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 del habla 11c, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de codificación del habla 11c, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 11c recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 11c y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla 11c. El dispositivo de codificación del habla 11c incluye una unidad de selección de ranura de tiempo 1p1 y una unidad de multiplexado de flujos de bits 1g4, en lugar de la unidad de selección de ranura de tiempo 1p y la unidad de multiplexado de flujos de bits 1g del dispositivo de codificación del 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 la información de selección de ranura de tiempo a la unidad de multiplexado de flujos de bits 1g4. La unidad de multiplexado de flujos de bits 1g4 multiplexa el flujo de bits codificados, calculado 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 multiplexado de flujos de bits 1g, también multiplexa la información de selección de ranura de tiempo recibida desde la unidad de selección de ranura de tiempo 1p1 y emite el flujo de bits multiplexado a través del dispositivo de comunicación del dispositivo de codificación del 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 del 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 procedimiento de selección de ranura de tiempo de la unidad de selección de ranura de tiempo 3a1. El dispositivo de decodificación del habla 21b (véase la figura 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 del habla 21b, cargando y ejecutando un programa informático predeterminado (tal como, por ejemplo, un programa informático para realizar procesos, ilustrados en el diagrama de flujo de la figura 21) almacenado en una memoria integrada del dispositivo de decodificación del habla 21b, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 21b recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 21b.

Tal como se ilustra en la figura 20, el dispositivo de decodificación del habla 21b incluye una unidad de separación de flujos de bits 2a5 y la unidad de selección de ranura de tiempo 3a1, en lugar de la unidad de separación de flujos de bits 2a y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación del 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 flujos de bits 2a5 separa el flujo de bits multiplexado en el parámetro de intensidad de filtro, la información suplementaria de SBR y el flujo de bits codificados como la unidad de separación de flujo de bits 2a, y además 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 flujos de bits 2a5 (proceso en la etapa S11). La información de selección 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 procedimiento de selección de ranura de tiempo descrito en la modificación 4. En este caso, aunque no se ilustra, la señal del dominio de QMF de los componentes de alta frecuencia generados por la unidad de generación de señales 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 (tal como, por ejemplo, $P_{exp,Th}$ y t_{Th}) utilizado para seleccionar

la ranura de tiempo.

(Modificación 6 del primer ejemplo)

5 Un dispositivo de codificación del 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 del habla 11d, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de codificación del habla 11d, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 11d recibe una
10 señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 11d y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla 11d. El dispositivo de codificación del 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 del habla 11a de la modificación 1, y también incluye una unidad de selección de ranura de tiempo 1p2.

15 La unidad de selección de ranura de tiempo 1p2 recibe una señal en el dominio de 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 del habla 11a de la modificación 1.

(Modificación 7 del primer ejemplo)

25 Un dispositivo de codificación del 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 del habla 11e, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de codificación del habla 11e, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 11e recibe una
30 señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 11e y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla 11e. El dispositivo de codificación del 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 del habla 11d de la modificación 6. El dispositivo de codificación del habla 11e también incluye una unidad de multiplexado de flujos de bits que además
35 recibe una salida desde la unidad de selección de ranura de tiempo 1p3, en lugar de la unidad de multiplexado de flujos 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 multiplexado de flujos de bits.

40 (Modificación 8 del primer ejemplo)

45 Un dispositivo de codificación del 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 del habla de la modificación 8, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de codificación del habla de la modificación 8, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla de la modificación 8 recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla. El dispositivo de codificación del habla de la modificación 8 incluye además la unidad de selección de ranura de tiempo 1p, además de las del dispositivo de codificación del habla descrito en la modificación 2.

55 Un dispositivo de decodificación del 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 del habla de la modificación 8, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de decodificación del habla de la modificación 8, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla de la modificación 8 recibe el flujo de bits multiplexados codificados y emite una señal de
60 habla decodificada al exterior del dispositivo de decodificación del habla. El dispositivo de decodificación del habla de la modificación 8 incluye además 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 del habla descrito en la modificación 2, y además incluye la
65

unidad de selección de ranura de tiempo 3a.

(Modificación 9 del primer ejemplo)

5 Un dispositivo de codificación del 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 del habla de la modificación 9, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de codificación del habla de la modificación 9, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de
10 codificación del habla de la modificación 9 recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla. El dispositivo de codificación del 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 del habla descrito en la modificación 8. El dispositivo de codificación del habla de la modificación 9 incluye además una
15 unidad de multiplexado de flujos de bits que recibe una salida desde la unidad de selección de ranura de tiempo 1p1 además de la entrada suministrada a la unidad de multiplexado de flujos de bits descrita en la modificación 8, en lugar de la unidad de multiplexado de flujos de bits descrita en la modificación 8.

20 Un dispositivo de decodificación del 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 del habla de la modificación 9, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de decodificación del habla de la modificación 9, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de
25 decodificación del habla de la modificación 9 recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla. El dispositivo de decodificación del 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 del habla descrito en la modificación 8. El dispositivo de decodificación del habla de la modificación 9 también incluye una unidad de separación de flujo de bits que separa los $a_D(n, r)$ descritos en la modificación 2 en lugar del parámetro de intensidad de filtro de la unidad de separación
30 de flujos de bits 2a5, en lugar de la unidad de separación de flujos de bits 2a.

(Modificación 1 del Segundo ejemplo)

35 Un dispositivo de codificación del habla 12a (figura 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 del habla 12a, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de codificación del habla 12a, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 12a recibe una
40 señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 12a y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla 12a. El dispositivo de codificación del 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 del habla 12, y además incluye la unidad de selección de ranura de tiempo 1p.

45 Un dispositivo de decodificación del habla 22a (véase la figura 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 del habla 22a, cargando y ejecutando un programa informático predeterminado (tal como, por ejemplo, un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 23) almacenado en una memoria integrada del dispositivo de decodificación del habla
50 22a, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 22a recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 22a. El dispositivo de decodificación del habla 22a, tal como se ilustra en la figura 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 del habla 22 del segundo ejemplo, y además incluye la unidad de selección de
55 ranura de tiempo 3a.

60 La unidad de selección de ranura de tiempo 3a notifica el 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 coeficientes de predicción lineal 2p1. La unidad de interpolación/extrapolación de coeficientes de predicción lineal 2p1 obtiene los $a_H(n, r)$ correspondientes a la ranura de tiempo r1 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
65

coeficientes de predicción lineal 2p, en base al resultado de la selección transmitido desde la unidad de selección de ranura de tiempo 3a (proceso en la etapa Sj1). La unidad de filtro de predicción lineal 2k2 realiza un filtrado de síntesis de predicción lineal en la dirección de frecuencia sobre los $q_{adj}(n, r1)$ emitidos desde la unidad de ajuste de alta frecuencia 2j para la ranura de tiempo r1 seleccionada, utilizando los $a_H(n, r1)$ que se interpolan o extrapolan, y que se obtienen desde la unidad de interpolación/extrapolación de coeficientes de predicción lineal 2p1, como la unidad de filtro de predicción lineal 2k1 (proceso en la etapa Sj2), 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 del habla 12b (figura 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 del habla 11b, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de codificación del habla 12b, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 12b recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 12b y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla 12b. El dispositivo de codificación del habla 12b incluye la unidad de selección de ranura de tiempo 1p1, y una unidad de multiplexado de flujos de bits 1g5, en lugar de la unidad de selección de ranura de tiempo 1p y la unidad de multiplexado de flujos de bits 1g2 del dispositivo de codificación del habla 12a de la modificación 1. La unidad de multiplexado de flujos de bits 1g5 multiplexa el flujo de bits codificados, calculado por la unidad central de codificación de códecs 1c, la información suplementaria de SBR calculada por la unidad de codificación de SBR 1d y los índices de las ranuras de tiempo correspondientes a los coeficientes de predicción lineal cuantizados, recibidos desde la unidad de cuantización de coeficientes de predicción lineal 1k como la unidad de multiplexado de flujos de bits 1g2, multiplexa además la información de selección de ranura de tiempo, recibida desde la unidad de selección de ranura de tiempo 1p1, y emite el flujo de bits multiplexados a través del dispositivo de comunicación del dispositivo de codificación del habla 12b.

Un dispositivo de decodificación del habla 22b (véase la figura 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 del habla 22b, cargando y ejecutando un programa informático predeterminado (tal como, por ejemplo, un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 25) almacenado en una memoria integrada del dispositivo de decodificación del habla 22b, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 22b recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 22b. El dispositivo de decodificación del habla 22b, tal como se ilustra en la figura 24, incluye una unidad de separación de flujos de bits 2a6 y la unidad de selección de ranura de tiempo 3a1, en lugar de la unidad de separación de flujos de bits 2a1 y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación del 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 flujos de bits 2a6 separa el flujo de bits multiplexados en los $a_H(n, r_i)$ que están siendo cuantizados, el índice r_i de la ranura de tiempo correspondiente, la información suplementaria de SBR y el flujo de bits codificados, como la unidad de separación de flujos de bits 2a1, y además separa la información de selección de ranura de tiempo.

(Modificación 4 del tercer ejemplo)

El valor

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

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

(Modificación 4 del tercer ejemplo)

Tal como se describe en la modificación 2 del tercer ejemplo, es preferible que la unidad de ajuste de forma de envolvente 2s controle los $e_{adj}(r)$ utilizando un valor predeterminado $e_{adj,Th}(r)$, considerando que la envolvente temporal ajustada $e_{adj}(r)$ es un coeficiente de ganancia multiplicado por la muestra de sub-banda de QMF, por ejemplo, como la expresión (28) y las expresiones (37) y (38).

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

(Cuarto ejemplo)

Un dispositivo de codificación del habla 14 (figura 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 del habla 14, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de codificación del habla 14, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 14 recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 14 y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla 14. El dispositivo de codificación del habla 14 incluye una unidad de multiplexado de flujos de bits 1g7, en lugar de la unidad de multiplexado de flujos de bits 1g del dispositivo de codificación del habla 11b de la modificación 4 del primer ejemplo, y además incluye la unidad de cálculo de envolvente temporal 1m y la unidad de cálculo de parámetros de envolvente 1n del dispositivo de codificación del habla 13.

La unidad de multiplexado de flujos de bits 1g7 multiplexa el flujo de bits codificados, calculado 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 multiplexado de flujos 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 envolvente calculado por la unidad de cálculo de parámetros de forma de envolvente 1n en la información suplementaria de envolvente temporal, los multiplexa y emite el flujo de bits multiplexados (flujo de bits multiplexados codificados) a través del dispositivo de comunicación del dispositivo de codificación del habla 14.

(Modificación 4 del cuarto ejemplo)

Un dispositivo de codificación del habla 14a (figura 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 del habla 14a, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de codificación del habla 14a, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 14a recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 14a y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla 14a. El dispositivo de codificación del 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 del habla 14 del cuarto ejemplo, y además incluye la unidad de selección de ranura de tiempo 1p.

Un dispositivo de decodificación del habla 24d (véase la figura 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 del habla 24d, cargando y ejecutando un programa informático predeterminado (tal como un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 27) almacenado en una memoria integrada del dispositivo de decodificación del habla 24d, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24d recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24d. El dispositivo de decodificación del habla 24d, tal como se ilustra en la figura 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 del habla 24, y además incluye la unidad de selección de ranura de tiempo 3a. La unidad de formación de envolvente temporal 2v forma la señal en el dominio de QMF obtenido desde la unidad de filtro de predicción lineal 2k3 utilizando la información de envolvente temporal obtenida desde la unidad de ajuste de forma de envolvente 2s, como la unidad de formación de envolvente temporal 2v de la tercera realización, de la cuarta realización y de sus modificaciones (proceso en la etapa Sk1).

(Modificación 6 de la realización)

Un dispositivo de decodificación del habla 24e (véase la figura 28) 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 del habla 24e, cargando y ejecutando un programa informático predeterminado (tal como un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 29) almacenado en una memoria integrada del dispositivo de decodificación del habla 24e, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24e recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24e. En la modificación 6, el dispositivo de decodificación del habla 24e, tal como se ilustra en la figura 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 del habla 24d descrito en la modificación 4, que puede omitirse en toda la

extensión de la realización, como el primer ejemplo, e incluye una unidad de selección de ranura de tiempo 3a2 y una unidad de formación de envolvente temporal 2v1, en lugar de la unidad de selección de ranura de tiempo 3a y la unidad de formación de envolvente temporal 2v del dispositivo de decodificación del habla 24d. El dispositivo de decodificación del 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 de envolvente temporal realizado por la unidad de formación de envolvente temporal 2v1, cuyo orden de procesamiento es intercambiable en toda la extensión de la realización.

La unidad de formación de envolvente temporal 2v1 forma los $q_{adj}(k, r)$ obtenidos desde la unidad de ajuste de alta frecuencia 2j, utilizando los $e_{adj}(r)$ obtenidos desde la unidad de ajuste de forma de envolvente 2s, como la unidad de formación de envolvente temporal 2v, y obtiene una señal $q_{envadj}(k, r)$ en el dominio de QMF en la cual se forma la envolvente temporal. La unidad de formación de envolvente temporal 2v1 también notifica a la unidad de selección de ranura de tiempo 3a2 sobre los parámetros obtenidos cuando se forma la envolvente temporal, o sobre los parámetros calculados al menos utilizando los parámetros obtenidos cuando se está dando forma a la envolvente temporal como la información de selección de ranura 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 envolventes de SBR)

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

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

Los valores

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

también pueden 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 también utilizarse una pluralidad de segmentos de ranura de tiempo (como las envolventes de SBR)

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

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:

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

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

5 La información de selección de ranura de tiempo puede ser también $e_{adj}(r)$ de la expresión (23), la expresión (35) o la expresión (36), o puede ser $\left| e_{adj}(r) \right|^2$, 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 (tales como envolventes de SBR)

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

10 así como el valor medio de los mismos

$$\bar{e}_{adj}(i), \quad \left| \bar{e}_{adj}(i) \right|^2 \quad \text{---(57)}$$

15 como la información de selección de ranura de tiempo. Debe reseñarse que:

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

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

20 La información de selección de ranura de tiempo puede ser también $e_{adj, scaled}(r)$ 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 (tales como envolventes de SBR)

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

25 así como el valor medio de los mismos

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

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

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

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

5 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 del dominio de QMF, correspondiente a los componentes de alta frecuencia en los cuales se forma la envolvente temporal o un valor de amplitud de señal de los mismos, a los cuales se aplica la operación de raíz cuadrada.

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

10

En una pluralidad de segmentos de ranura de tiempo (tales como envolventes de SBR)

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

15

así como el valor medio de los mismos

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

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

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

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

25

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 espectro de frecuencia de los componentes de alta frecuencia generados por la unidad de generación de alta frecuencia $2g$ también puede estar representado como $k_x \leq k < k_x + M$.

30

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 o no el filtrado de síntesis de predicción lineal en la señal $q_{envadj}(k, r)$ en el dominio de QMF de los componentes de alta frecuencia de la ranura de tiempo r en la que se forma la envolvente temporal por parte de la unidad de formación de envolvente temporal $2v1$, en base a la información de selección de ranura de tiempo transmitida desde la unidad de

35

formación de envolvente temporal 2v1 (proceso en la etapa 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 al 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 de envolvente temporal 2v1, es mayor que un valor predeterminado u_{Th} , o al menos una ranura de tiempo r en la cual $u(r)$ es igual o mayor que un valor predeterminado u_{Th} . $u(r)$ puede incluir al 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 $P_{envadj}(r)$, descritos anteriormente, y;

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

y u_{Th} puede incluir al menos uno entre;

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

u_{Th} puede ser también un valor medio de $u(r)$ de un ancho del tiempo predeterminado (tal como una envolvente de SBR) que incluya la ranura de tiempo r . La selección puede hacerse también para estén incluidas las ranuras de tiempo en las que $u(r)$ alcanza sus máximos. Los máximos de $u(r)$ pueden calcularse tal como se calculan los máximos de la potencia de señal en la señal del dominio de 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 similar a los de la modificación 4 del primer ejemplo, utilizando $u(r)$, y las ranuras de tiempo pueden seleccionarse en base a esto. El procedimiento de selección de ranura de tiempo puede ser al menos uno de los procedimientos descritos anteriormente, puede incluir al menos un procedimiento distinto a los descritos anteriormente, o puede ser una combinación de los mismos.

(Modificación 7 de la realización)

Un dispositivo de decodificación del habla 24f (véase la figura 30) 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 del habla 24f, cargando y ejecutando un programa informático predeterminado (tal como, por ejemplo, un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 29) almacenado en una memoria integrada del dispositivo de decodificación del habla 24e, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24f recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24f. En la modificación 6, el dispositivo de decodificación del habla 24f, tal como se ilustra en la figura 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 del habla 24d descrito en la modificación 4, que puede omitirse en toda la extensión de la realización, como el primer ejemplo, e incluye la unidad de selección de ranura de tiempo 3a2 y la unidad de formación de envolvente temporal 2v1, en lugar de la unidad de selección de ranura de tiempo 3a y la unidad de formación de envolvente temporal 2v del dispositivo de decodificación del habla 24d. El dispositivo de decodificación del 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 de envolvente temporal realizado por la unidad de formación de envolvente temporal 2v1, cuyo orden de procesamiento es intercambiable en toda la extensión de la realización.

La unidad de selección de ranura de tiempo 3a2 determina si se realiza o no el filtrado de síntesis de predicción lineal por parte de la unidad de filtro de predicción lineal 2k3 en la señal $q_{envadj}(k, r)$ en el dominio de QMF de los componentes de alta frecuencia de las ranuras de tiempo r en las cuales la envolvente temporal es formada por la

unidad de formación de envolvente temporal 2v1, en base a la información de selección de ranura de tiempo transmitida desde la unidad de formación de envolvente 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 a la unidad de filtro de predicción lineal 2k3.

5

(Modificación 8 de la realización)

Un dispositivo de codificación del habla 14b (figura 50) 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 codificación del habla 14b, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de codificación del habla 14b, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de codificación del habla 14b recibe una señal de habla que va a ser codificada desde el exterior del dispositivo de codificación del habla 14b y emite un flujo de bits multiplexados codificados al exterior del dispositivo de codificación del habla 14b. El dispositivo de codificación del habla 14b incluye una unidad de multiplexado de flujos de bits 1g6 y la unidad de selección de ranura de tiempo 1p1, en lugar de la unidad de multiplexado de flujos de bits 1g7 y la unidad de selección de ranura de tiempo 1p del dispositivo de codificación del habla 14a de la modificación 4.

10

15

La unidad de multiplexado de flujos de bits 1g6 multiplexa el flujo de bits codificados, calculado 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 envolvente 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 envolvente calculado por la unidad de cálculo de parámetro de forma de envolvente 1n, también multiplexa la información de selección de ranura de tiempo recibida desde la unidad de selección de ranura de tiempo 1p1 y emite el flujo de bits multiplexados (flujo de bits multiplexados codificados) a través del dispositivo de comunicación del dispositivo de codificación del habla 14b.

20

25

Un dispositivo de decodificación del habla 24g (véase la figura 31) de la 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 del habla 24g, cargando y ejecutando un programa informático predeterminado (tal como un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 32) almacenado en una memoria integrada del dispositivo de decodificación del habla 24g, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24g recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24g. El dispositivo de decodificación del habla 24g incluye una unidad de separación de flujos de bits 2a7 y la unidad de selección de ranura de tiempo 3a1, en lugar de la unidad de separación de flujos de bits 2a3 y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación del habla 2d descrito en la modificación 4.

30

35

La unidad de separación de bits 2a7 separa el flujo de bits multiplexados, proporcionado a través del dispositivo de comunicación del dispositivo de decodificación del habla 24g, en la información suplementaria de envolvente temporal, la información suplementaria de SBR y el flujo de bits codificados, como la unidad de separación de flujos de bits 2a3, y además separa la información de selección de ranura de tiempo.

40

45

(Modificación 9 de la realización)

Un dispositivo de decodificación del habla 24h (véase la figura 33) de una 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 del habla 24h, cargando y ejecutando un programa informático predeterminado (tal como, por ejemplo, un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 34) almacenado en una memoria integrada del dispositivo de decodificación del habla 24h, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24h recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24h. El dispositivo de decodificación del habla 24h, tal como se ilustra en la figura 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 del habla 24b de la modificación 2, y además 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 etapa de "Ajuste de HF" en la SBR en la "MPEG-4 AAC", como la unidad primaria de ajuste de alta frecuencia 2j1 de la modificación 2 de la realización (proceso en la etapa Sm1). La unidad secundaria de ajuste de alta frecuencia 2j2 realiza al menos uno de los procesos en la etapa de "Ajuste de HF" en la SBR en la "MPEG-4 AAC", como la unidad secundaria de ajuste de alta frecuencia 2j2 de la modificación 2 de la realización (proceso en la etapa 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 etapa de "Ajuste de HF" en

50

55

60

65

la SBR en la "MPEG-4 AAC".

(Modificación 10 de la realización)

5 Un dispositivo de decodificación del habla 24i (véase la figura 35) de la 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 del habla 24i, cargando y ejecutando un programa informático predeterminado (tal como un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 36) almacenado en una memoria integrada del dispositivo de decodificación del habla 24i, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24i recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24i. El dispositivo de decodificación del habla 24i, tal como se ilustra en la figura 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 del habla 24h de la modificación 8, que puede omitirse en toda la extensión de la realización, como el primer ejemplo, e incluye la unidad de formación de envolvente temporal 2v1 y la unidad de selección de ranura de tiempo 3a2, en lugar de la unidad de formación de envolvente temporal 2v y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación del habla 24h de la modificación 8. El dispositivo de decodificación del 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 de envolvente temporal, realizado por la unidad de formación de envolvente temporal 2v1, cuyo orden de procesamiento es intercambiable en toda la extensión de la realización.

(Modificación 11 de la realización)

25 Un dispositivo de decodificación del habla 24j (véase la figura 37) 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 del habla 24j, cargando y ejecutando un programa informático predeterminado (tal como un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 36) almacenado en una memoria integrada del dispositivo de decodificación del habla 24j, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24j recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24j. El dispositivo de decodificación del habla 24j, tal como se ilustra en la figura 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 del habla 24h de la modificación 8, que puede omitirse en toda la extensión de la realización, como el primer ejemplo, e incluye la unidad de formación de envolvente temporal 2v1 y la unidad de selección de ranura de tiempo 3a2, en lugar de la unidad de formación de envolvente temporal 2v y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación del 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 de envolvente temporal realizado por la unidad de formación de envolvente temporal 2v1, cuyo orden de proceso es intercambiable en toda la extensión de la realización.

40

(Modificación 12 de la realización)

45 Un dispositivo de decodificación del habla 24k (véase la figura 38) 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 del habla 24k, cargando y ejecutando un programa informático predeterminado (tal como un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 39) almacenado en una memoria integrada del dispositivo de decodificación del habla 24k, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24k recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24k. El dispositivo de decodificación del habla 24k, tal como se ilustra en la figura 38, incluye la unidad de separación de flujos de bits 2a7 y la unidad de selección de ranura de tiempo 3a1, en lugar de la unidad de separación de flujos de bits 2a3 y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación del habla 24h de la modificación 8.

55

(Modificación 13 de la realización)

60 Un dispositivo de decodificación del habla 24q (véase la figura 40) 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 del habla 24q, cargando y ejecutando un programa informático predeterminado (tal como un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 41) almacenado en una memoria integrada del dispositivo de decodificación del habla 24q, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24q recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24q. El dispositivo de decodificación del habla 24q, tal como se ilustra en la figura 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

65

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 de envolvente 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 del habla 24c de la modificación 3, y además incluye la unidad de selección de ranura de tiempo 3a.

Al menos una de las unidades de ajuste de componentes de señal individual 2z4, 2z5 y 2z6 realiza el procesamiento en la señal del dominio de QMF de la ranura de tiempo seleccionada, para el componente de señal incluido en la salida 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 la etapa Sn1). Resulta preferible que el proceso que utiliza la información de selección de ranura de tiempo incluya al 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.

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 formar la envolvente temporal de cada uno entre la pluralidad de componentes de señal incluidos en la salida del medio primario de ajuste de alta frecuencia, por distintos procedimientos (si todas las unidades de ajuste de componentes de señal individual 2z4, 2z5 y 2z6 no realizan procesamiento en base al resultado de selección transmitido desde la unidad de selección de ranura de tiempo 3a, es lo mismo que la modificación 3 de la realización de la presente invención).

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.

En la figura 40, el resultado de la selección de ranura de tiempo se transmite a las unidades de ajuste de componentes de señal individual 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, sobre 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, con respecto 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 sub-banda de QMF por el coeficiente de ganancia se realiza sobre la señal de entrada utilizando la envolvente temporal obtenida desde la unidad de ajuste de forma de envolvente 2s, como la unidad de formación de envolvente temporal 2v, y a continuación el filtrado de síntesis de predicción lineal en la dirección de frecuencia también se realiza sobre la señal de salida, utilizando los coeficientes de predicción lineal recibidos desde la unidad de ajuste de intensidad de filtro 2f, como la unidad de filtro de predicción lineal 2k), descrita en la modificación 3 de la realización puede seleccionar la ranura de tiempo utilizando la información de selección de ranura de tiempo proporcionada desde la unidad de formación de envolvente temporal.

(Modificación 14 de la realización)

Un dispositivo de decodificación del habla 24m (véase la figura 42) 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 del habla 24m, cargando y ejecutando un programa informático predeterminado (tal como un programa informático para realizar procesos ilustrados en el diagrama de flujo de la figura 43) almacenado en una memoria integrada del dispositivo de decodificación del habla 24m, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24m recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24m. El dispositivo de decodificación del habla 24m, tal como se ilustra en la figura 42, incluye la unidad de separación de flujos de bits 2a7 y la unidad de selección de ranura de tiempo 3a1, en lugar de la unidad de separación de flujos de bits 2a3 y la unidad de selección de ranura de tiempo 3a del dispositivo de decodificación del habla 24q de la modificación 12.

(Modificación 15 de la realización)

Un dispositivo de decodificación del habla 24n (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 del habla 24n, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de decodificación del habla 24n, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24n recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de

5 decodificación del habla 24n. El dispositivo de decodificación del 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 del habla 24a de la modificación 1, y además incluye la unidad de selección de ranura de tiempo 3a.

10 (Modificación 16 de la realización)

10 Un dispositivo de decodificación del habla 24p (que no se ilustra) de una modificación 16 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 del habla 24p, cargando y ejecutando un programa informático predeterminado almacenado en una memoria integrada del dispositivo de decodificación del habla 24p, tal como la ROM en la RAM. El dispositivo de comunicación del dispositivo de decodificación del habla 24p recibe el flujo de bits multiplexados codificados y emite una señal de habla decodificada al exterior del dispositivo de decodificación del habla 24p. El dispositivo de decodificación del 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 del habla 24n de la modificación 14. El dispositivo de decodificación del habla 24p también incluye una unidad de separación de flujos de bits 2a8 (que no se ilustra), en lugar de la unidad de separación de flujos de bits 2a4.

25 La unidad de separación de flujos de bits 2a8 separa el flujo de bits multiplexados en la información suplementaria de SBR y en el flujo de bits codificados como la unidad de separación de flujos de bits 2a4, así como en la información de selección de ranura de tiempo.

Aplicabilidad industrial

30 La presente invención proporciona una técnica aplicable a la técnica de extensión de ancho de banda en el dominio de frecuencia representado por la 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 la velocidad de bits.

Lista de signos de referencia

35 11, 11a, 11b, 11c, 12, 12a, 12b, 13, 14, 14a, 14b: dispositivo de codificación del habla

1a: unidad de transformación de frecuencia

1b: unidad de transformación inversa de frecuencia

40 1c: unidad de codificación central de códecs

1d: unidad de codificación de SBR

45 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

50 1g, 1g1, 1g2, 1g3, 1g4, 1g5, 1g6, 1g7: unidad de multiplexado de flujos de bits

1h: unidad de transformación inversa de alta frecuencia

55 1i: unidad de cálculo de potencia a corto plazo

1j: unidad de diezmo de coeficiente de predicción lineal

1k: unidad de cuantización de coeficiente de predicción lineal

60 1m: unidad de cálculo de envolvente temporal

1n: unidad de cálculo de parámetros de forma de envolvente

65 1p, 1p1: unidad de selección de ranura de tiempo

ES 2 610 363 T3

- 21, 22, 23, 24, 24b, 24c: dispositivo de decodificación del habla
- 2a, 2a1, 2a2, 2a3, 2a5, 2a5, 2a6, 2a7: unidad de separación de flujos de bits
- 5 2b: unidad de decodificación central de códecs
- 2c: unidad de transformación de frecuencia
- 2d, 2d1: unidad de análisis de predicción lineal de baja frecuencia
- 10 2e, 2e1: unidad de detección de cambio de señal
- 2f: unidad de ajuste de intensidad de filtro
- 15 2g: unidad de generación de alta frecuencia
- 2h, 2h1: unidad de análisis de predicción lineal de alta frecuencia
- 2i, 2i1: unidad de filtro inverso de predicción lineal
- 20 2j, 2j1, 2j2, 2j3, 2j4: unidad de ajuste de alta frecuencia
- 2k, 2k1, 2k2, 2k3: unidad de filtro de predicción lineal
- 25 2m: unidad de adición de coeficientes
- 2n: unidad de transformación inversa de frecuencia
- 2p, 2p1: unidad de interpolación/extrapolación de coeficientes de predicción lineal
- 30 2r: unidad de cálculo de envolvente temporal de baja frecuencia
- 2s: unidad de ajuste de forma de envolvente
- 35 2t: unidad de cálculo de envolvente temporal de alta frecuencia
- 2u: unidad de allanamiento de envolvente temporal
- 2v, 2v1: unidad de formación de envolvente temporal
- 40 2w: unidad de conversión de información suplementaria
- 2z1, 2z2, 2z3, 2z4, 2z5, 2z6: unidad de ajuste de componentes de señal individual
- 45 3a, 3a1, 3a2: unidad de selección de ranura de tiempo

REIVINDICACIONES

1. Un dispositivo de decodificación del habla, adaptado para decodificar una señal de habla codificada, comprendiendo el dispositivo de decodificación del habla:
- 5 medio de separación de flujos de bits, adaptado para separar un flujo de bits que incluye la señal de habla codificada en un flujo de bits codificados e información suplementaria de envolvente temporal, recibándose el flujo de bits desde el exterior del dispositivo de decodificación del habla;
- 10 medio de decodificación central, adaptado para decodificar el flujo de bits codificados, separado por el medio de separación de flujos de bits, para obtener un componente de baja frecuencia;
- 15 transformación de frecuencia, adaptada para transformar el componente de baja frecuencia, obtenido por el medio de decodificación central, en un dominio de frecuencia;
- 20 medio de generación de alta frecuencia, adaptado 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, adaptado para ajustar el componente de alta frecuencia, generado por el medio de generación de alta frecuencia, para generar un componente de alta frecuencia ajustado;
- medio de análisis de envolvente temporal de baja frecuencia, adaptado para analizar el componente de baja frecuencia, transformado en el dominio de frecuencia por el medio de transformación de frecuencia, para obtener información de la envolvente temporal;
- 30 medio de conversión de información suplementaria, adaptado para convertir la información suplementaria de envolvente temporal en un parámetro para ajustar la información de la envolvente temporal;
- 35 medio de ajuste de envolvente temporal, adaptado para ajustar la información de envolvente temporal obtenida por el medio de análisis de envolvente temporal de baja frecuencia, para generar información de la envolvente temporal ajustada, usando el medio de ajuste de envolvente temporal el parámetro en dicho ajuste de la información de envolvente temporal; y
- 40 medio de formación de la envolvente temporal, adaptado para formar una envolvente temporal del componente de alta frecuencia ajustado, multiplicando el componente de alta frecuencia ajustado por la información de envolvente temporal ajustada.
2. Un dispositivo de decodificación del habla, adaptado para decodificar una señal de habla codificada, comprendiendo el dispositivo de decodificación del habla:
- 45 medio de decodificación central, adaptado para decodificar un flujo de bits que incluye la señal de habla codificada, para obtener un componente de baja frecuencia, recibándose el flujo de bits desde el exterior del dispositivo de decodificación del habla;
- 50 medio de transformación de frecuencia, adaptado para transformar el componente de baja frecuencia obtenido por el medio de decodificación central en un dominio de frecuencia;
- 55 medio de generación de alta frecuencia, adaptado 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;
- 60 medio de ajuste de alta frecuencia, adaptado para ajustar el componente de alta frecuencia generado por el medio de generación de alta frecuencia, para generar un componente de alta frecuencia ajustado;
- 65 medio de análisis de envolvente temporal de baja frecuencia, adaptado para analizar el componente de baja frecuencia transformado en el dominio de frecuencia por el medio de transformación de frecuencia, para obtener información de la envolvente temporal;
- medio de generación de información suplementaria de envolvente temporal, adaptado para analizar el flujo de bits para generar un parámetro para ajustar la información de la envolvente temporal;
- medio de ajuste de envolvente temporal, adaptado para ajustar la información de envolvente temporal obtenida por el medio de análisis de envolvente temporal de baja frecuencia, para generar información de la envolvente temporal ajustada, usando el medio de ajuste de envolvente temporal el parámetro en dicho ajuste de la información de envolvente temporal; y

medio de formación de la envolvente temporal, adaptado para formar una envolvente temporal del componente de alta frecuencia ajustado, multiplicando el componente de alta frecuencia ajustado por la información de envolvente temporal ajustada.

5
3. Un procedimiento de decodificación del habla que utiliza un dispositivo de decodificación del habla para decodificar una señal de habla codificada, comprendiendo el procedimiento de decodificación del habla:

10 una etapa de separación de flujos de bits, en que el dispositivo de decodificación del habla separa un flujo de bits que incluye la señal de habla codificada en un flujo de bits codificados e información suplementaria de la envolvente temporal, recibándose el flujo de bits desde el exterior del dispositivo de decodificación del habla;

15 una etapa de decodificación central en la que el dispositivo de decodificación del habla obtiene un componente de baja frecuencia decodificando el flujo de bits codificados, separado en la etapa de separación de flujos de bits;

20 una etapa de transformación de frecuencia en la que el dispositivo de decodificación del habla transforma el componente de baja frecuencia, obtenido en la etapa de decodificación central, en un dominio de frecuencia;

25 una etapa de generación de alta frecuencia en la que el dispositivo de decodificación del habla genera un componente de alta frecuencia copiando el componente de baja frecuencia transformado en el dominio de frecuencia en la etapa de transformación de frecuencia, desde una banda de baja frecuencia a una banda de alta frecuencia;

30 una etapa de ajuste de alta frecuencia en la que el dispositivo de decodificación del habla ajusta el componente de alta frecuencia generado en la etapa de generación de alta frecuencia, para generar un componente de alta frecuencia ajustado;

35 una etapa de análisis de envolvente temporal de baja frecuencia en la que el dispositivo de decodificación del habla obtiene información de la envolvente temporal analizando el componente de baja frecuencia transformado en el dominio de frecuencia en la etapa de transformación de frecuencia;

40 una etapa de conversión de información suplementaria en la que el dispositivo de decodificación del habla convierte la información suplementaria de envolvente temporal en un parámetro para ajustar la información de la envolvente temporal;

45 una etapa de ajuste de envolvente temporal en la que el dispositivo de decodificación del habla ajusta la información de la envolvente temporal, obtenida en la etapa de análisis de envolvente temporal de baja frecuencia, para generar información de envolvente temporal ajustada, en la que se usa el parámetro en dicho ajuste de la información de envolvente temporal; y

50 una etapa de formación de la envolvente temporal en la que el dispositivo de decodificación del habla forma una envolvente temporal del componente de alta frecuencia ajustado, multiplicando el componente de alta frecuencia ajustado por la información de envolvente temporal ajustada.

55 4. Un procedimiento de decodificación del habla que utiliza un dispositivo de decodificación del habla para decodificar una señal de habla codificada, comprendiendo el procedimiento de decodificación del habla:

60 una etapa de decodificación central, en la que el dispositivo de decodificación del habla decodifica un flujo de bits que incluye la señal de habla codificada, para obtener un componente de baja frecuencia, recibándose el flujo de bits desde el exterior del dispositivo de decodificación del habla;

65 una etapa de transformación de frecuencia en la que el dispositivo de decodificación del habla transforma el componente de baja frecuencia obtenido en la etapa de decodificación central en un dominio de frecuencia;

una etapa de generación de alta frecuencia en la que el dispositivo de decodificación del habla genera un componente de alta frecuencia copiando el componente de baja frecuencia transformado en el dominio de frecuencia en la etapa de transformación de frecuencia, desde una banda de baja frecuencia a una banda de alta frecuencia;

una etapa de ajuste de alta frecuencia en la que el dispositivo de decodificación del habla ajusta el componente de alta frecuencia generado en la etapa de generación de alta frecuencia, para generar un componente de alta frecuencia ajustado;

una etapa de análisis de envolvente temporal de baja frecuencia en la que el dispositivo de decodificación del habla obtiene información de la envolvente temporal analizando el componente de baja frecuencia transformado en el dominio de frecuencia en la etapa de transformación de frecuencia;

5 una etapa de generación de información suplementaria de envolvente temporal en la que el dispositivo de decodificación del habla analiza el flujo de bits para generar un parámetro para ajustar la información de la envolvente temporal;

10 una etapa de ajuste de envolvente temporal en la que el dispositivo de decodificación del habla ajusta la información de la envolvente temporal obtenida en la etapa de análisis de envolvente temporal de baja frecuencia, para generar información de envolvente temporal ajustada, en la que se usa el parámetro en dicho ajuste de la información de envolvente temporal; y

15 una etapa de formación de la envolvente temporal en la que el dispositivo de decodificación del habla forma una envolvente temporal del componente de alta frecuencia ajustado, multiplicando el componente de alta frecuencia ajustado por la información de envolvente temporal ajustada.

5. Un programa de decodificación del habla para decodificar una señal de habla codificada que hace que un dispositivo informático funcione como:

20 medio de separación de flujos de bits, para separar un flujo de bits que incluye la señal de habla codificada en un flujo de bits codificados e información suplementaria de envolvente temporal, recibiendo el flujo de bits desde el exterior del dispositivo de decodificación del habla;

25 medio de decodificación central para decodificar el flujo de bits codificados, separado por el medio de separación de flujo de bits, para obtener un componente de baja frecuencia;

30 medio de transformación de frecuencia para transformar el componente de baja frecuencia obtenido por el medio de decodificación central 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, desde una banda de baja frecuencia a una banda de alta frecuencia;

35 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;

40 medio de análisis de envolvente 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 de la envolvente temporal;

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

45 medio de ajuste de envolvente temporal para ajustar la información de envolvente temporal obtenida por el medio de análisis de envolvente temporal de baja frecuencia, para generar información de la envolvente temporal ajustada, usando el medio de ajuste de envolvente temporal el parámetro en dicho ajuste de la información de envolvente temporal; y

50 medio de formación de la envolvente temporal para formar una envolvente temporal del componente de alta frecuencia ajustado, multiplicando el componente de alta frecuencia ajustado por la información de envolvente temporal ajustada.

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

55 medio de decodificación central para decodificar un flujo de bits que incluye la señal de habla codificada, para obtener un componente de baja frecuencia, recibiendo el flujo de bits desde el exterior del dispositivo de decodificación del habla;

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

65 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, desde 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 envolvente 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 de la envolvente temporal;

10 medio de generación de información suplementaria de envolvente temporal para analizar el flujo de bits para generar un parámetro para ajustar la información de la envolvente temporal;

15 medio de ajuste de envolvente temporal para ajustar la información de envolvente temporal obtenida por el medio de análisis de envolvente temporal de baja frecuencia, para generar información de la envolvente temporal ajustada, usando el medio de ajuste de envolvente temporal el parámetro en dicho ajuste de la información de envolvente temporal; y

20 medio de formación de la envolvente temporal para formar una envolvente temporal del componente de alta frecuencia ajustado, multiplicando el componente de alta frecuencia ajustado por la información de envolvente temporal ajustada.

Fig. 1

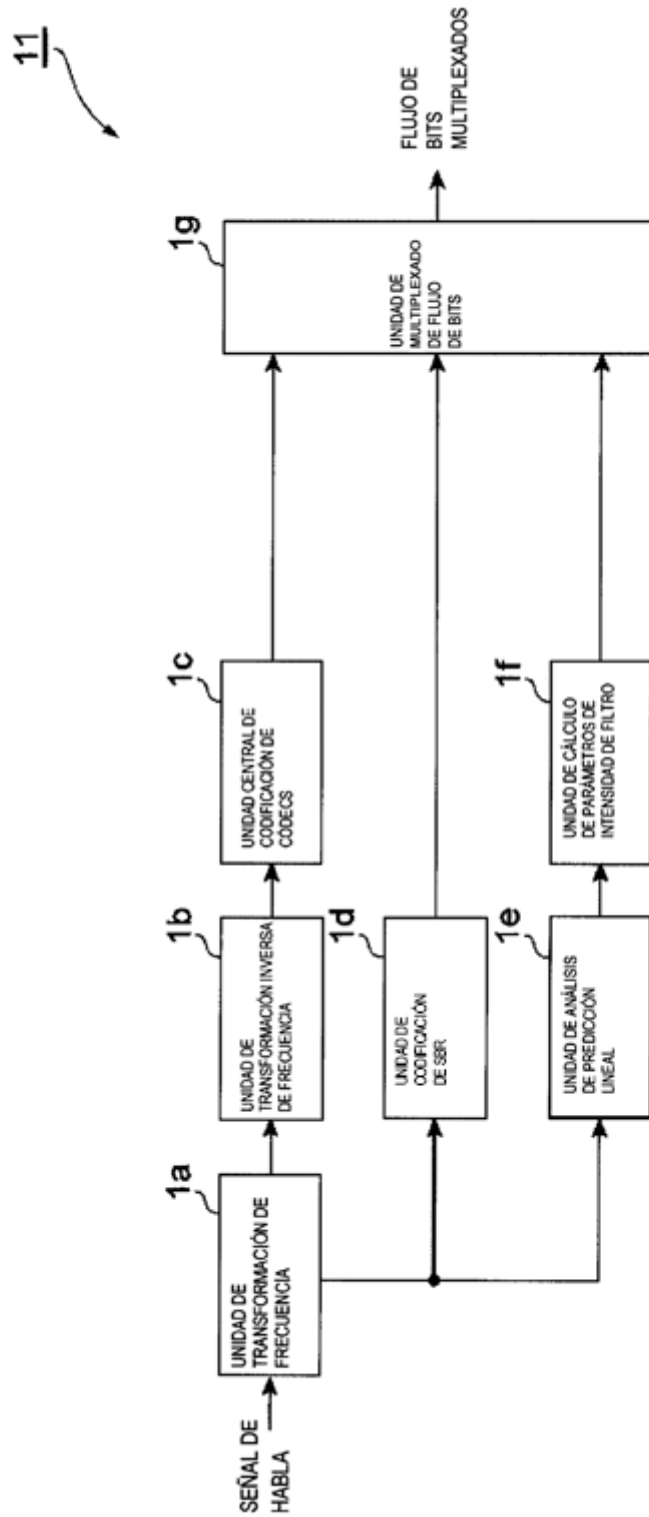
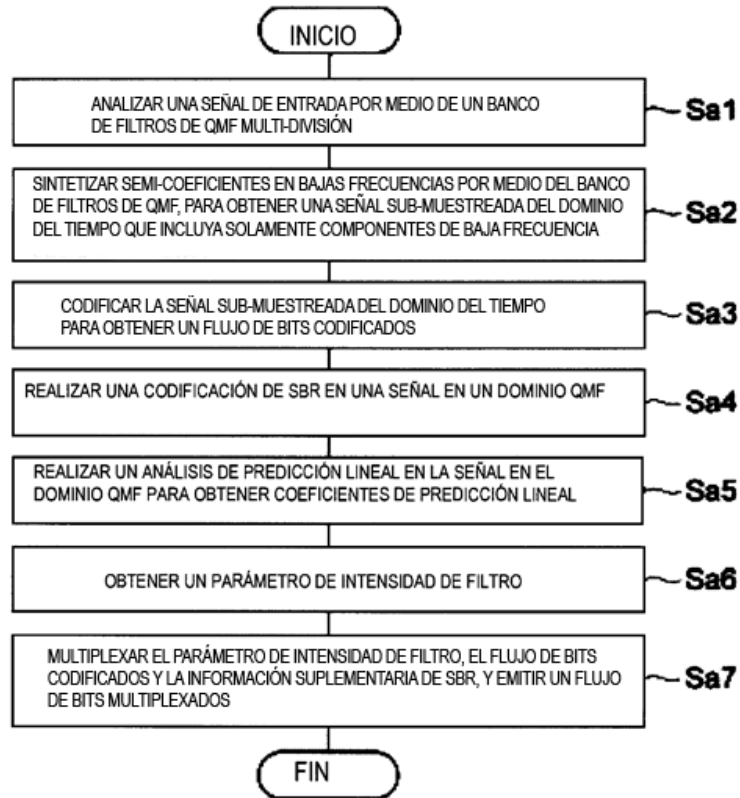


Fig. 2



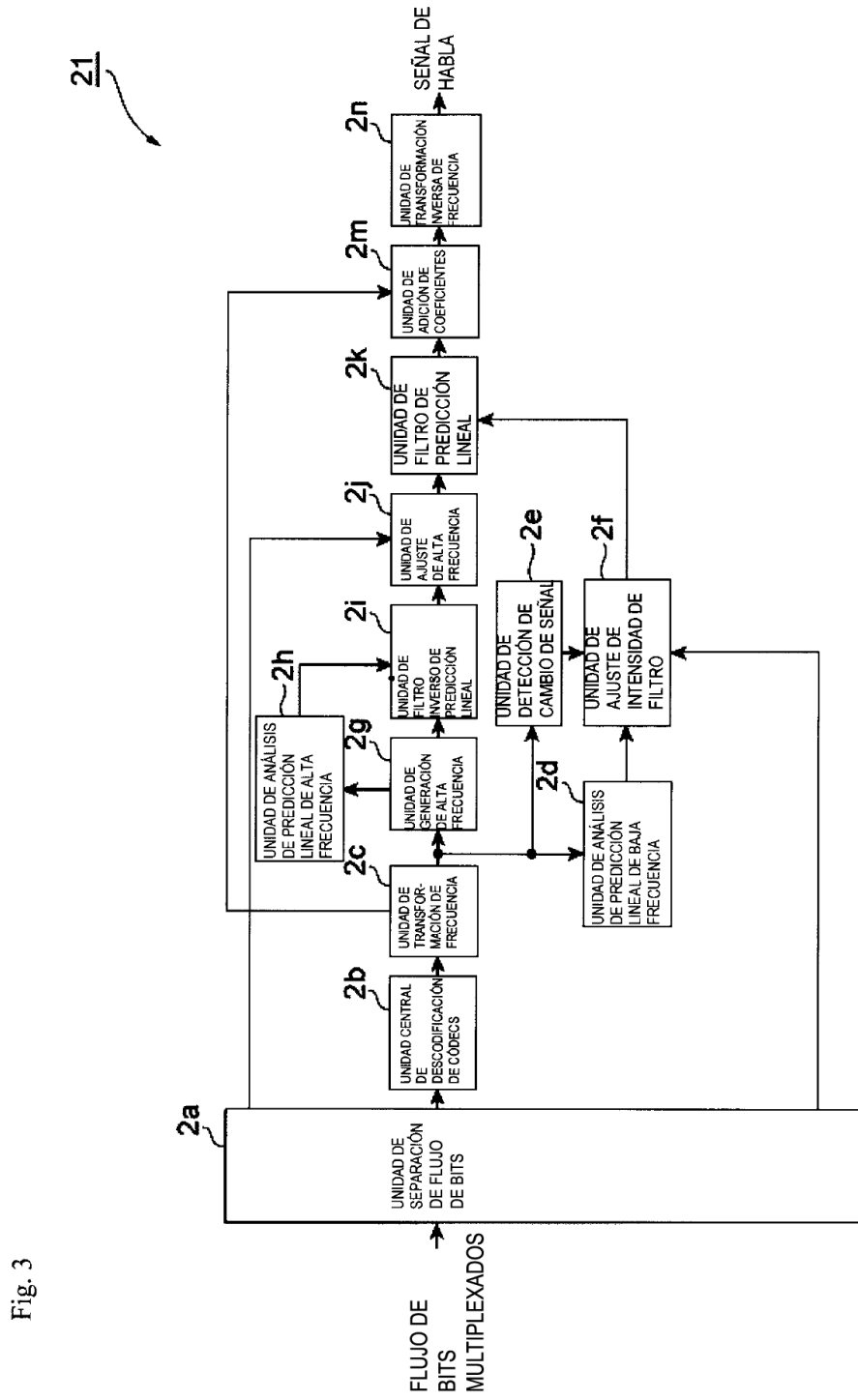


Fig. 3

21

Fig. 4

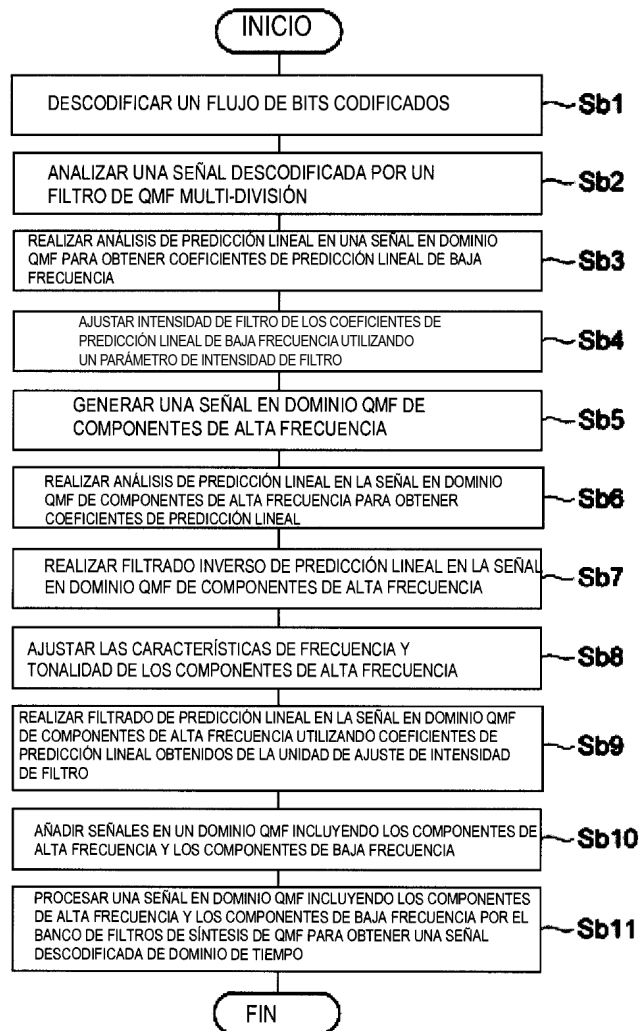
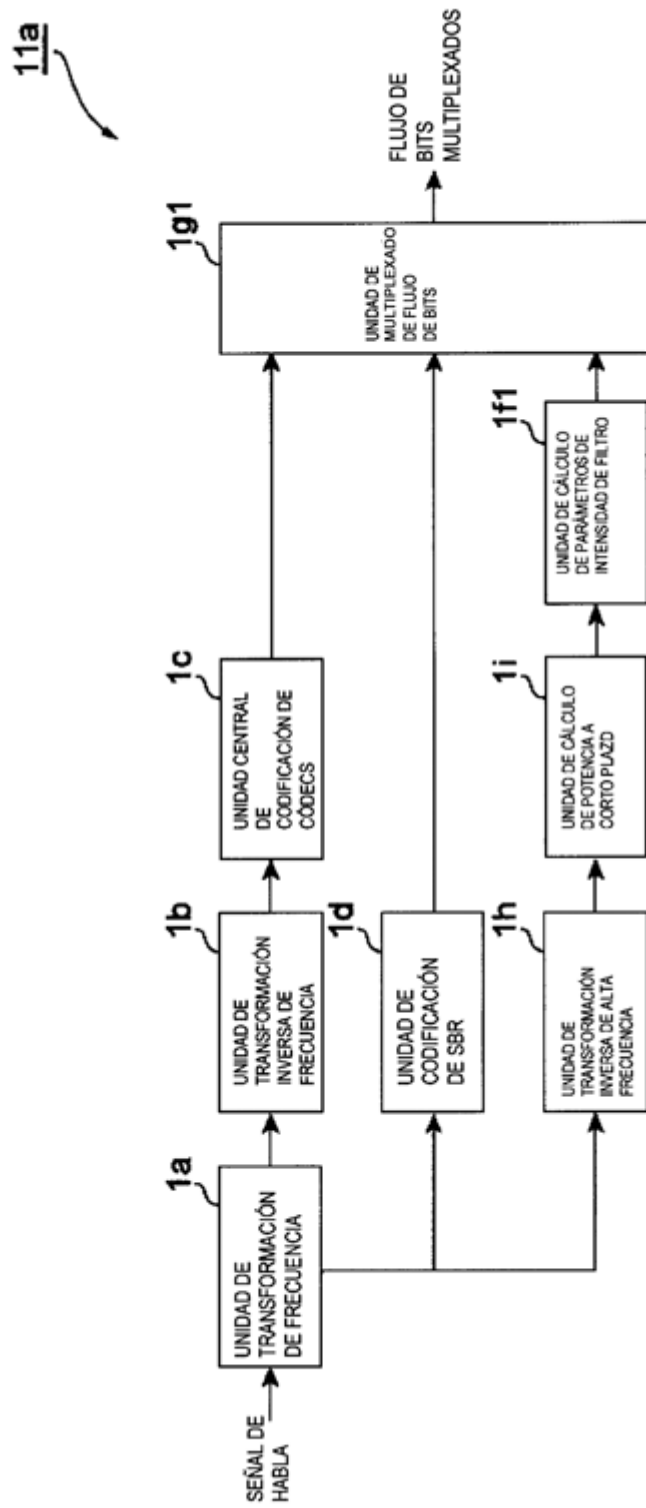


Fig. 5



12

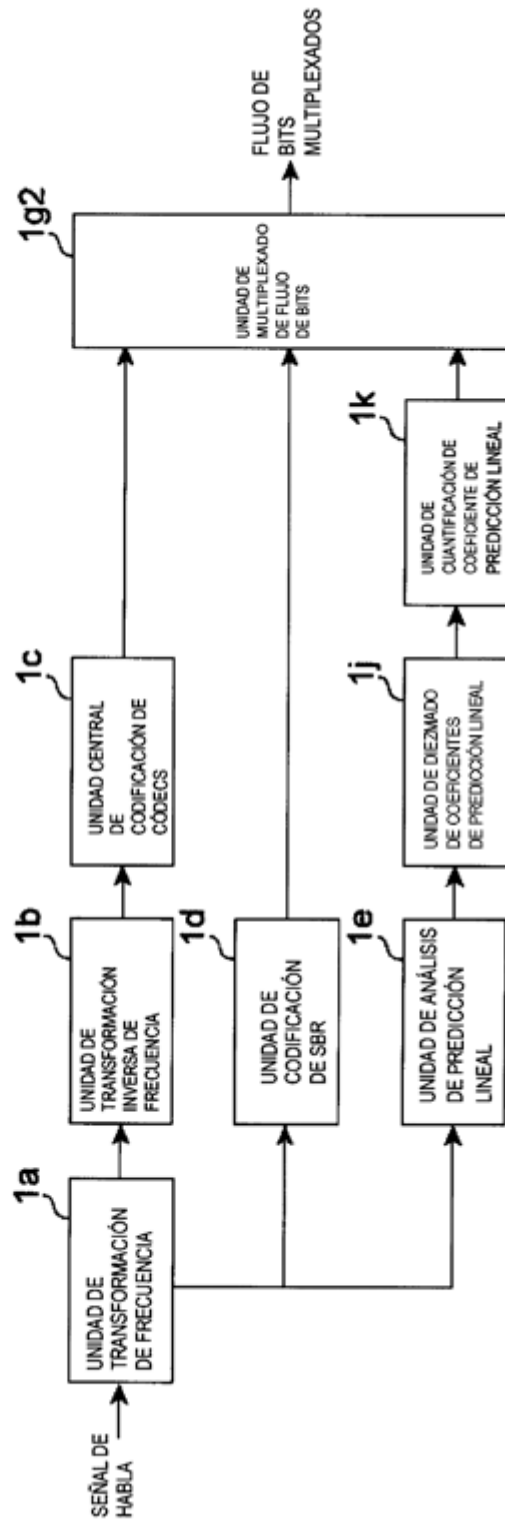


Fig. 6

Fig. 7

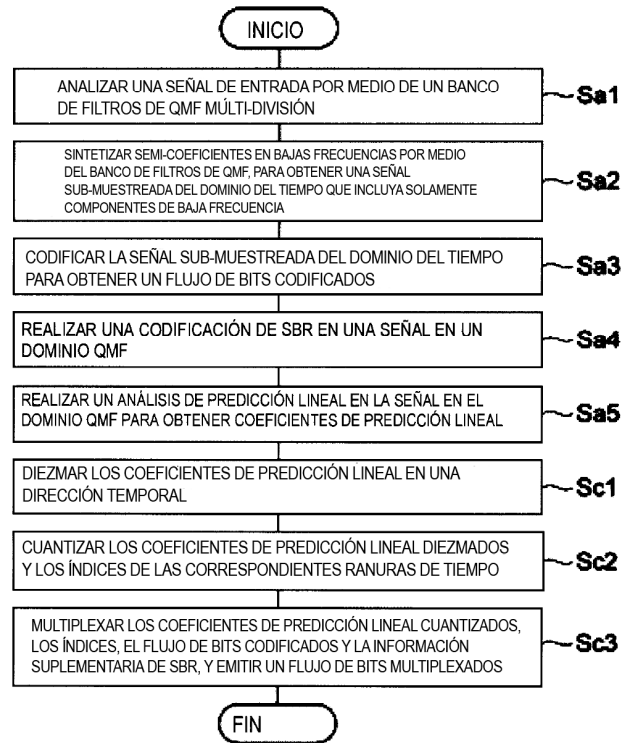


Fig. 8

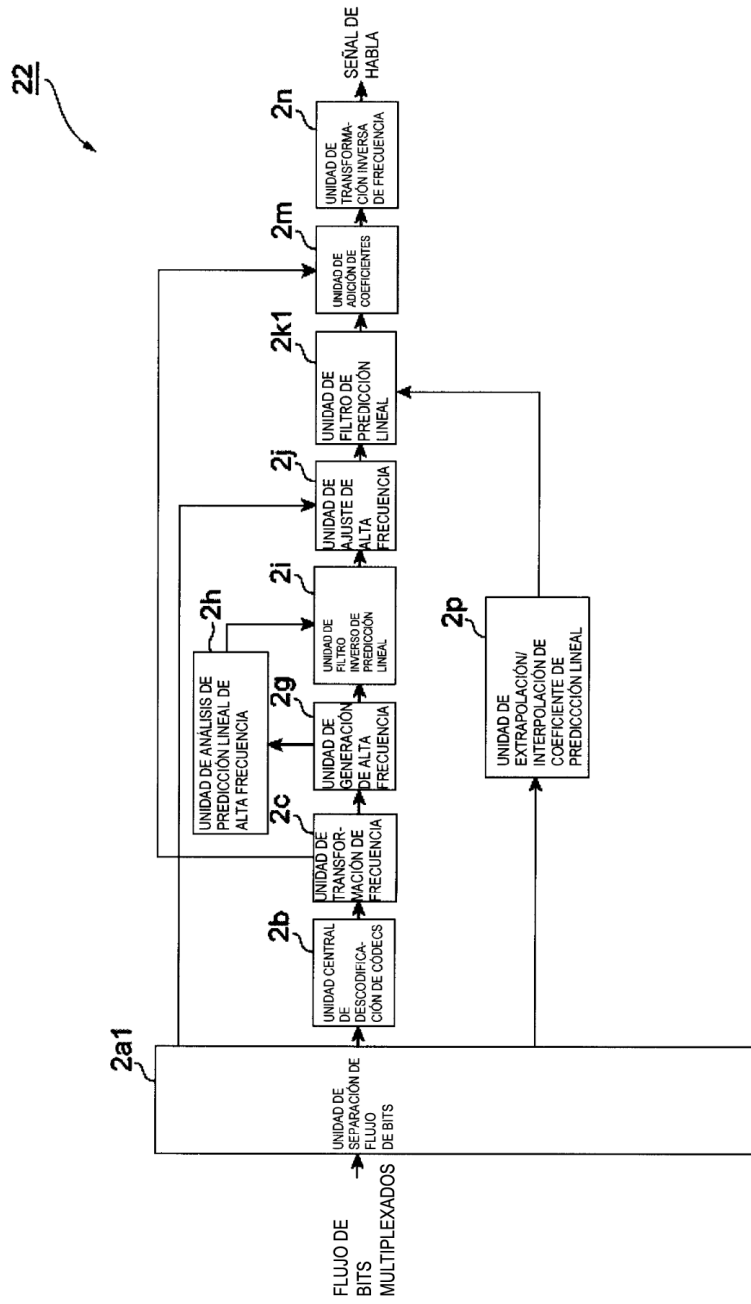


Fig. 9



13

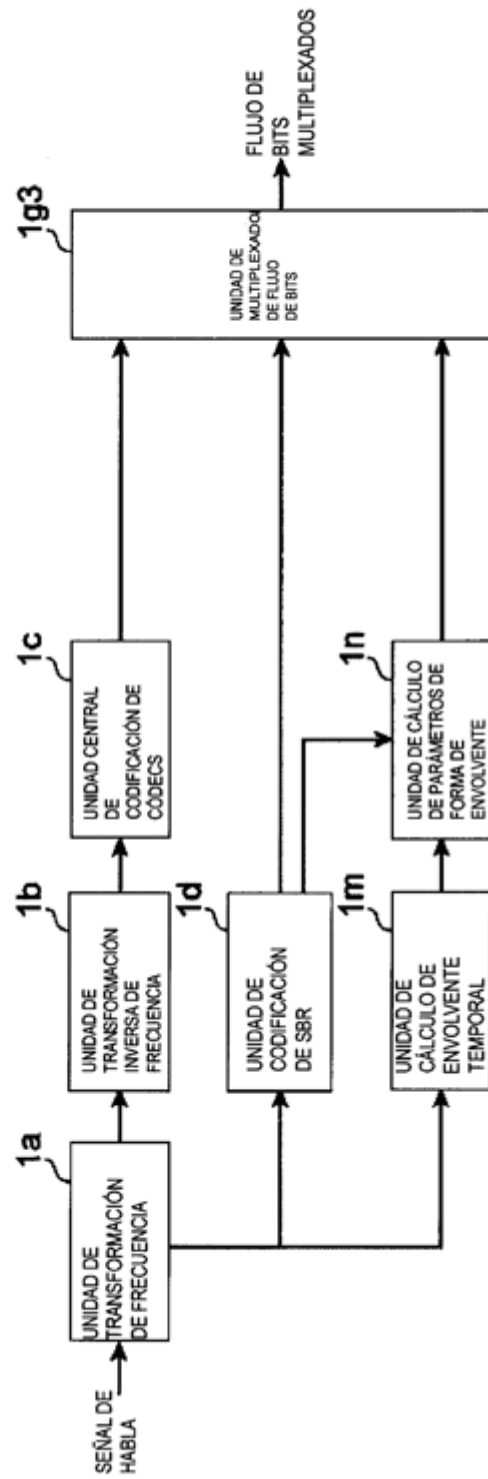


Fig. 10

Fig. 11



23

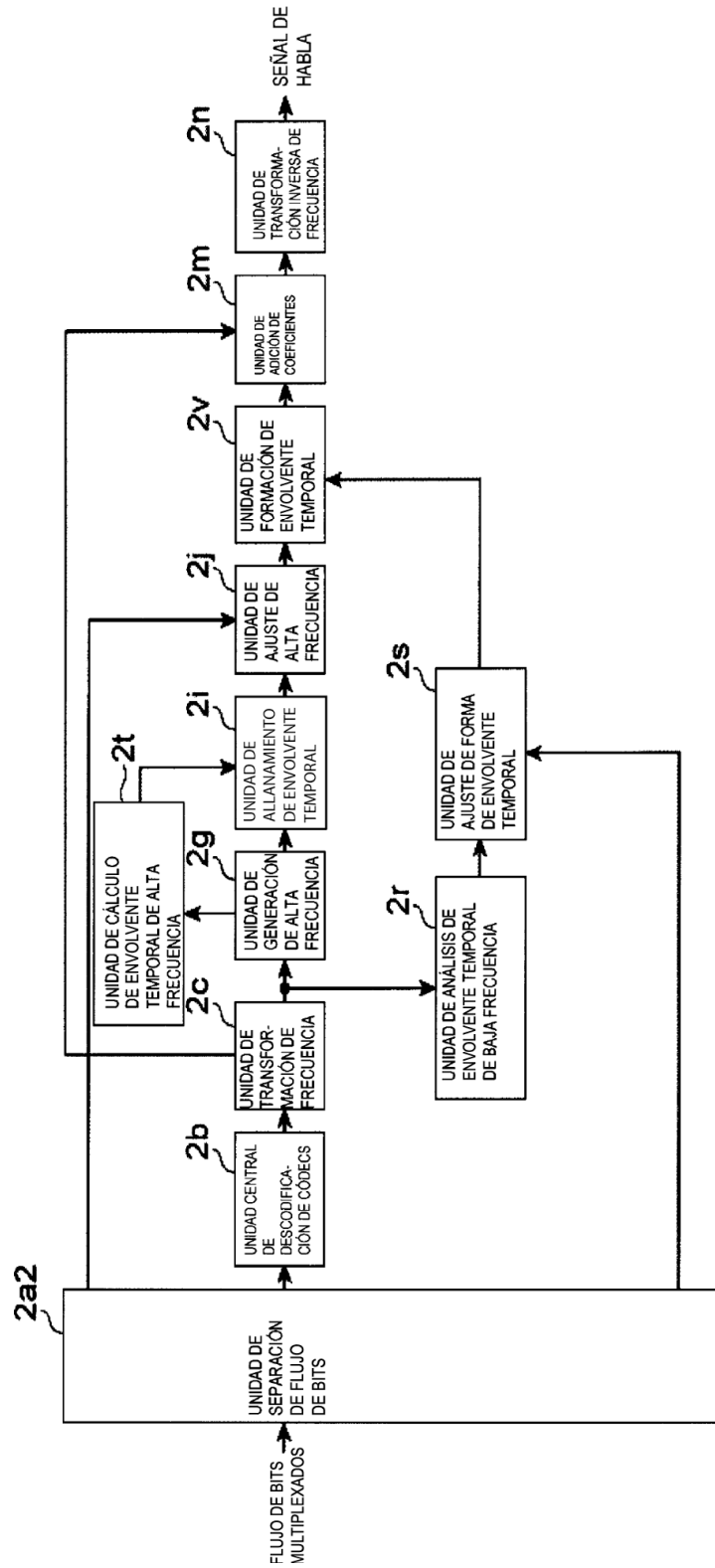


Fig. 12

Fig. 13



24

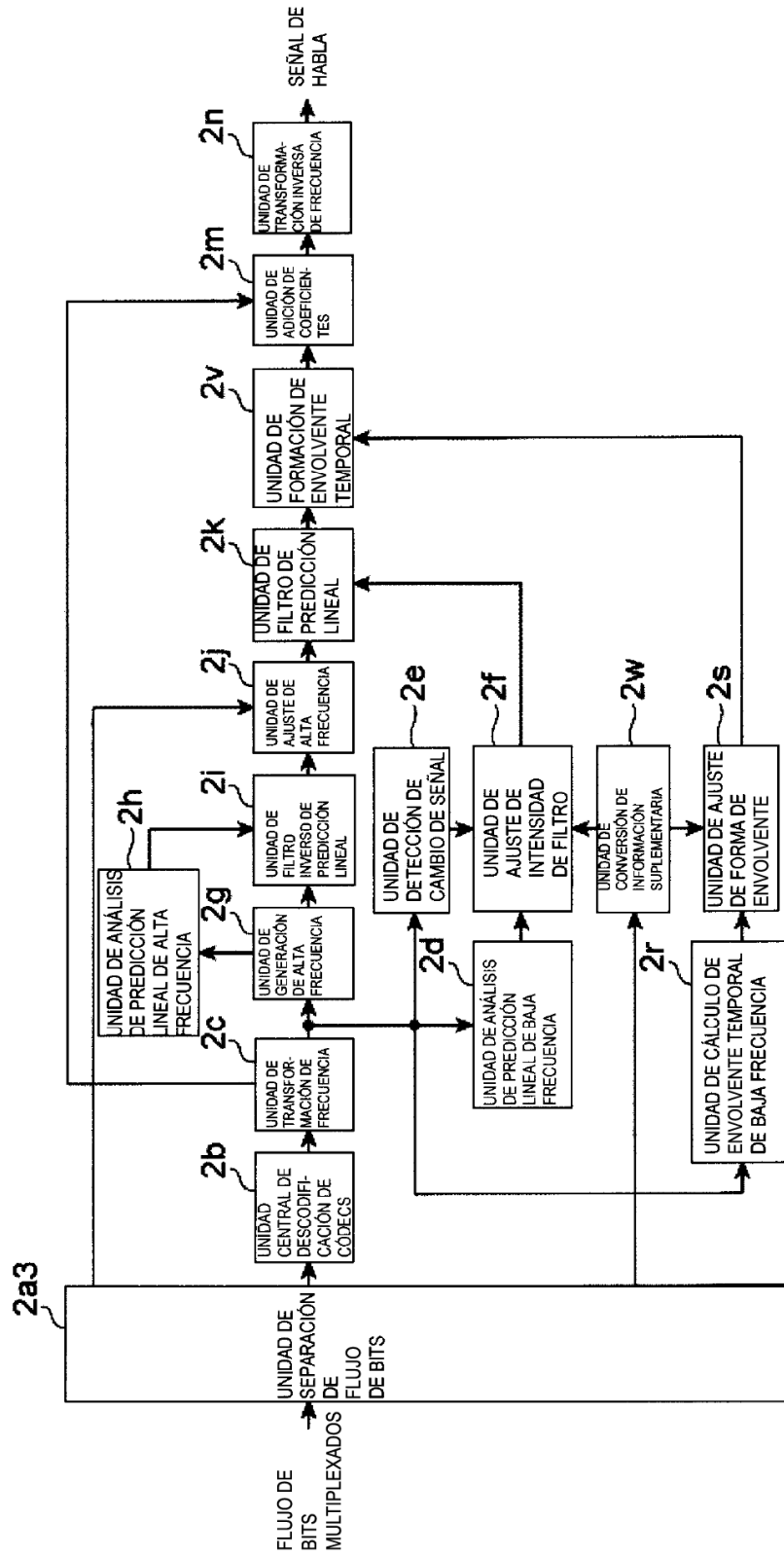
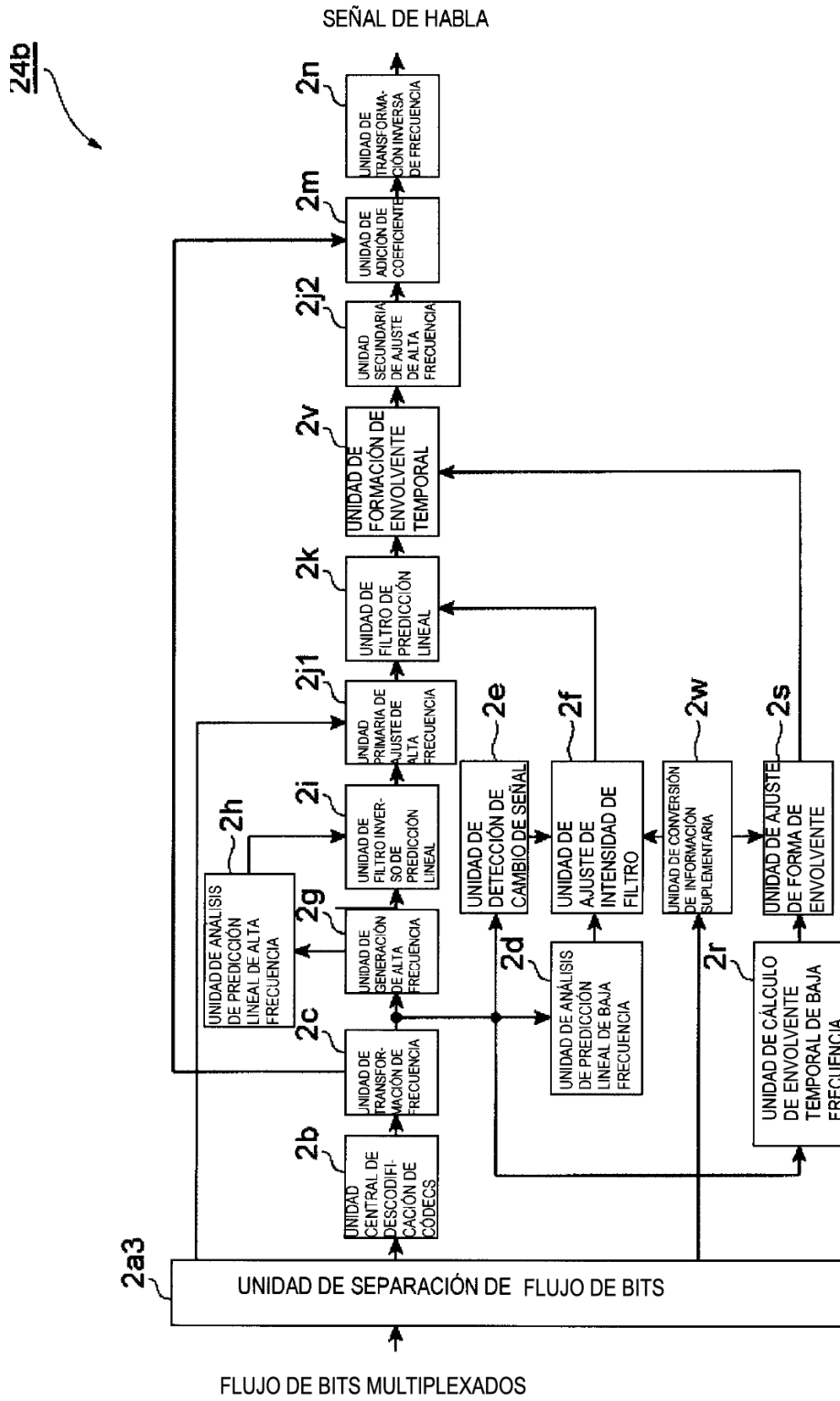


Fig. 14



24b

Fig. 15

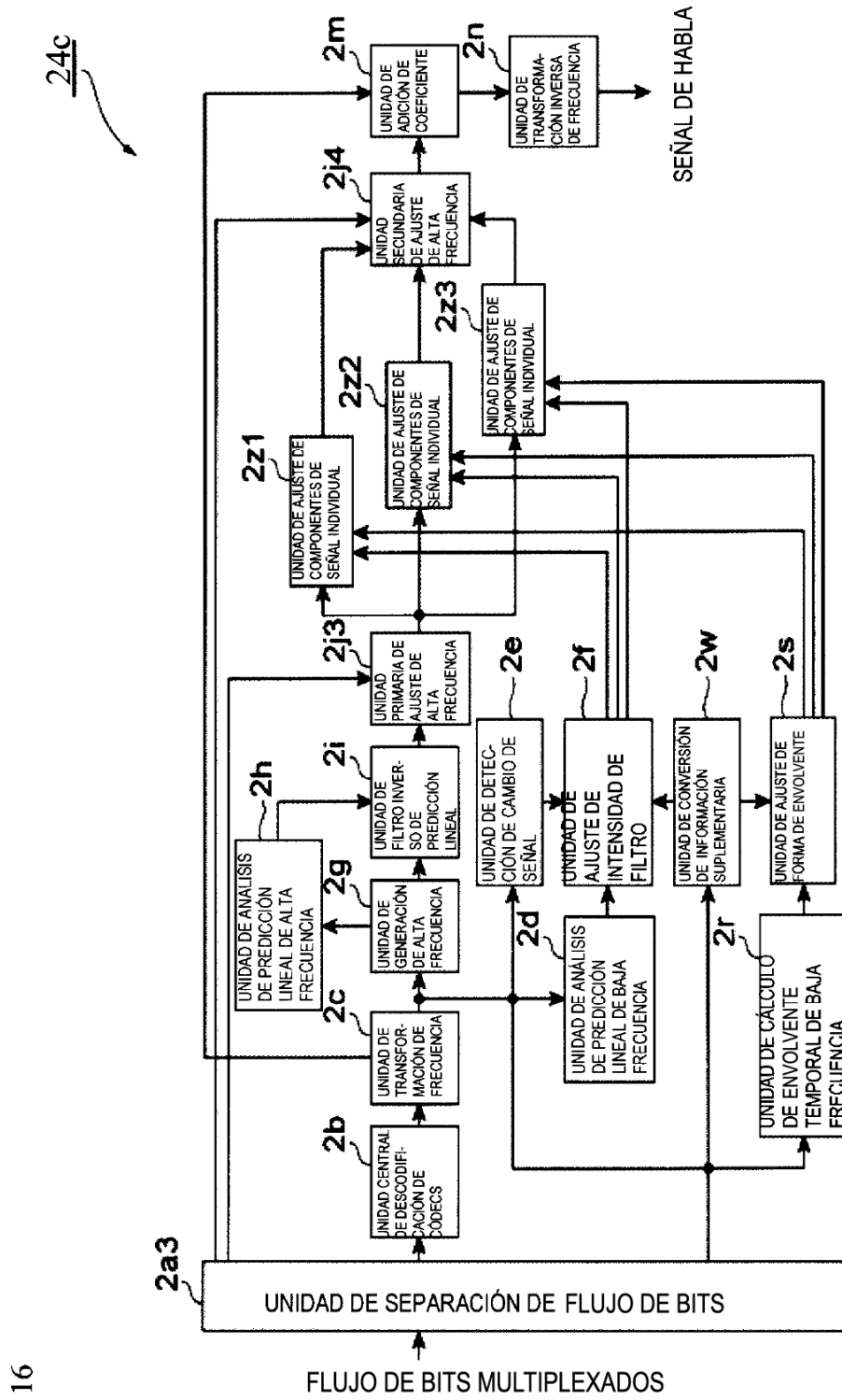


Fig. 16

Fig. 17



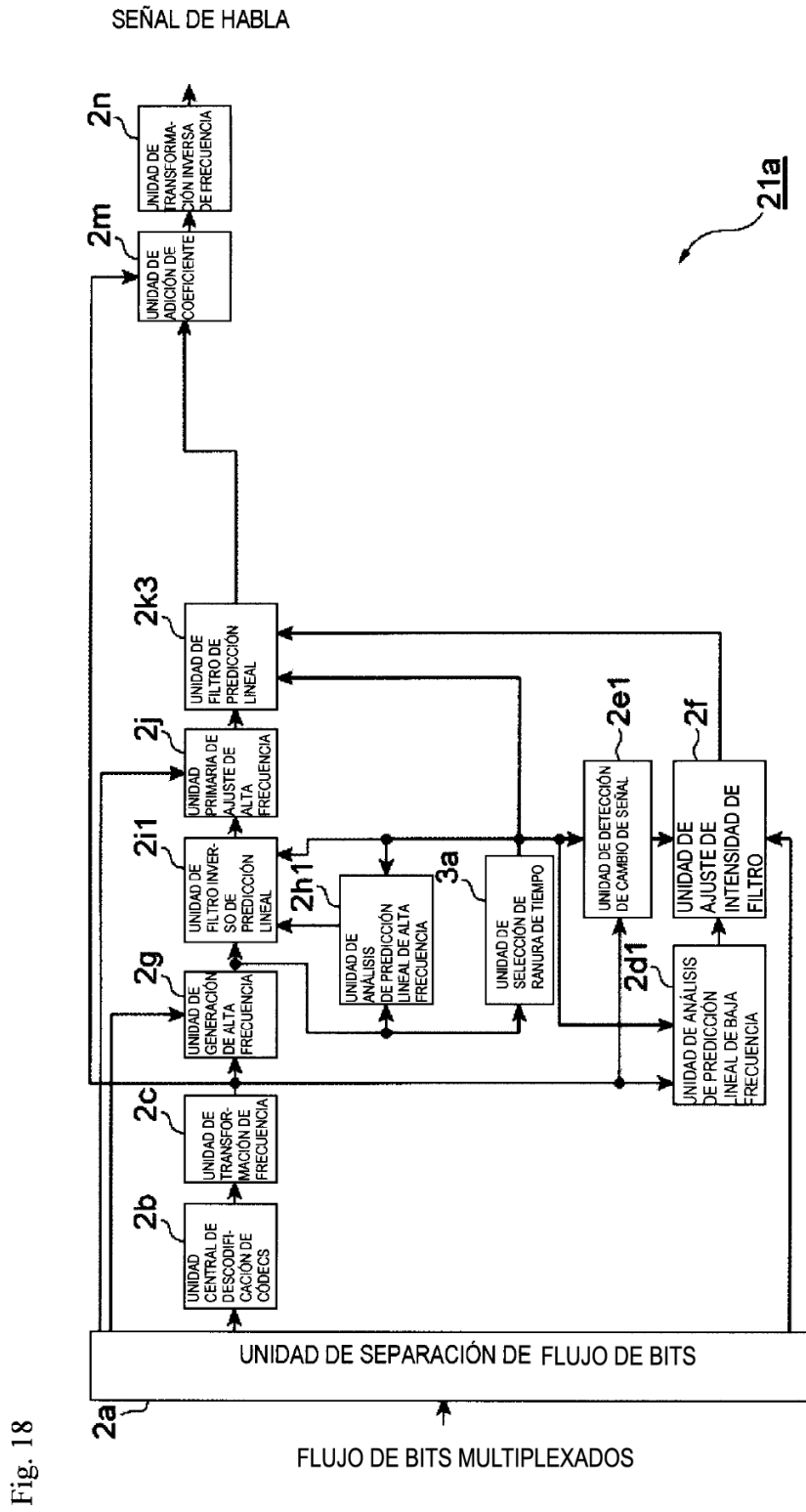


Fig. 18

Fig. 19



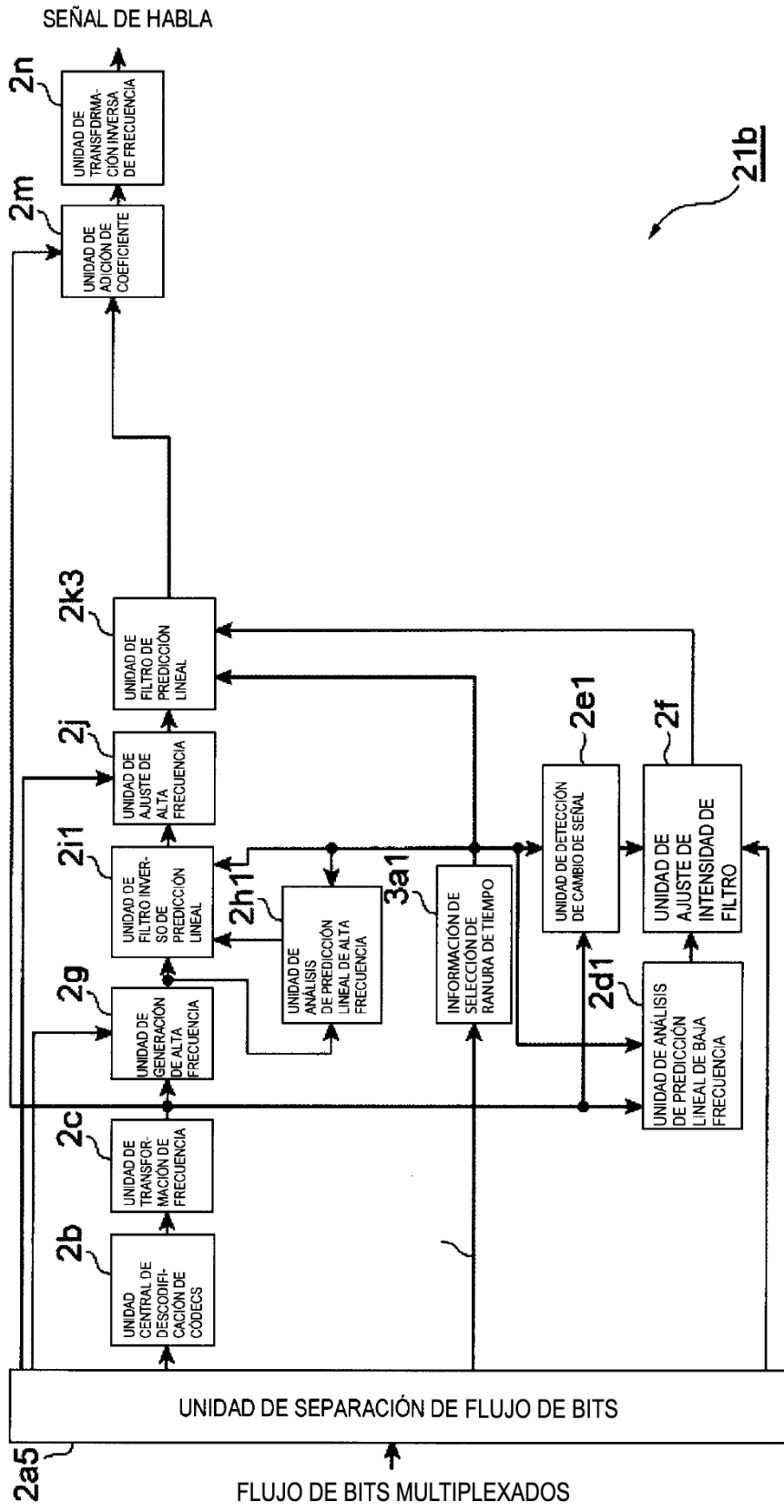


Fig. 20

Fig. 21



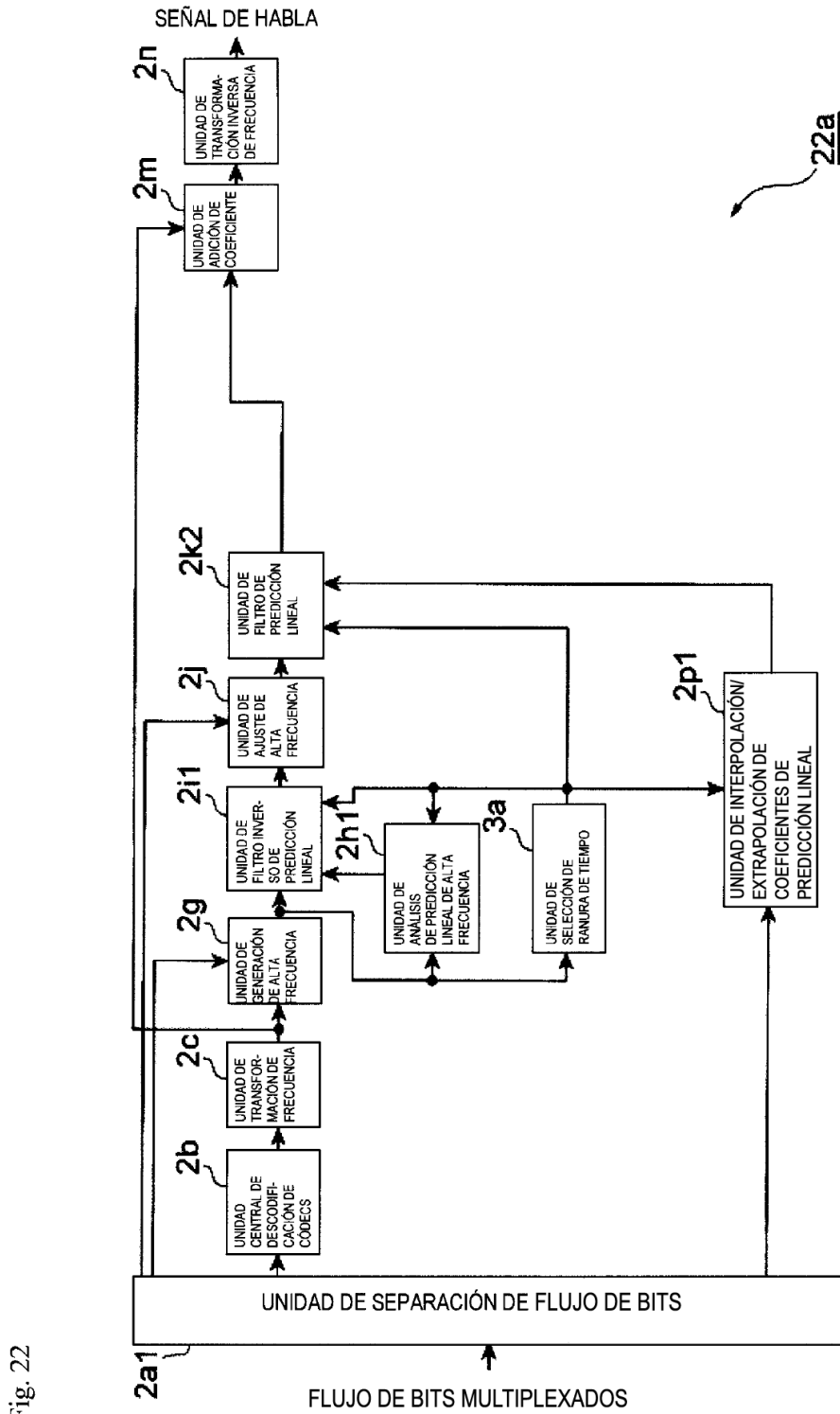


Fig. 22

Fig. 23



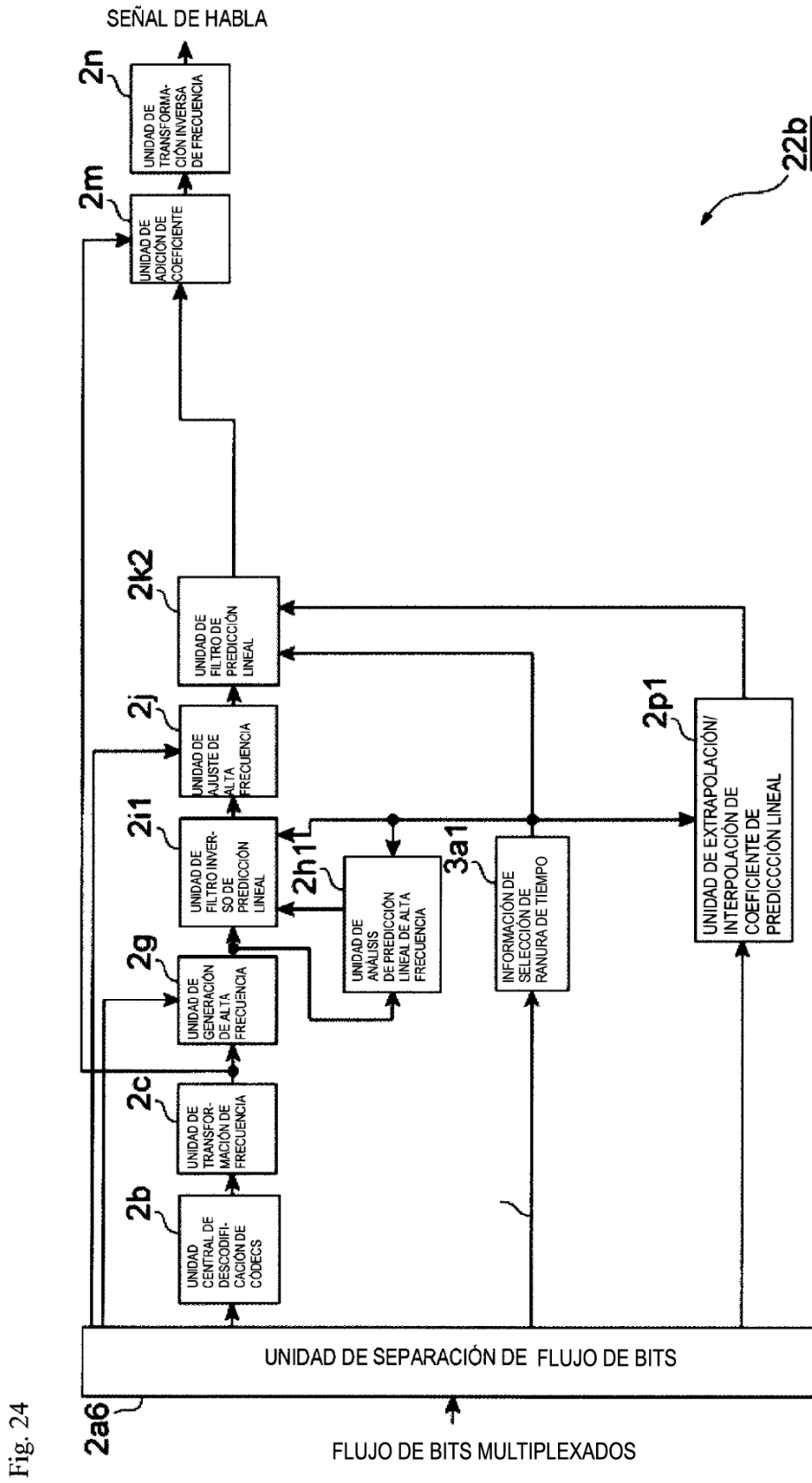


Fig. 24

Fig. 25



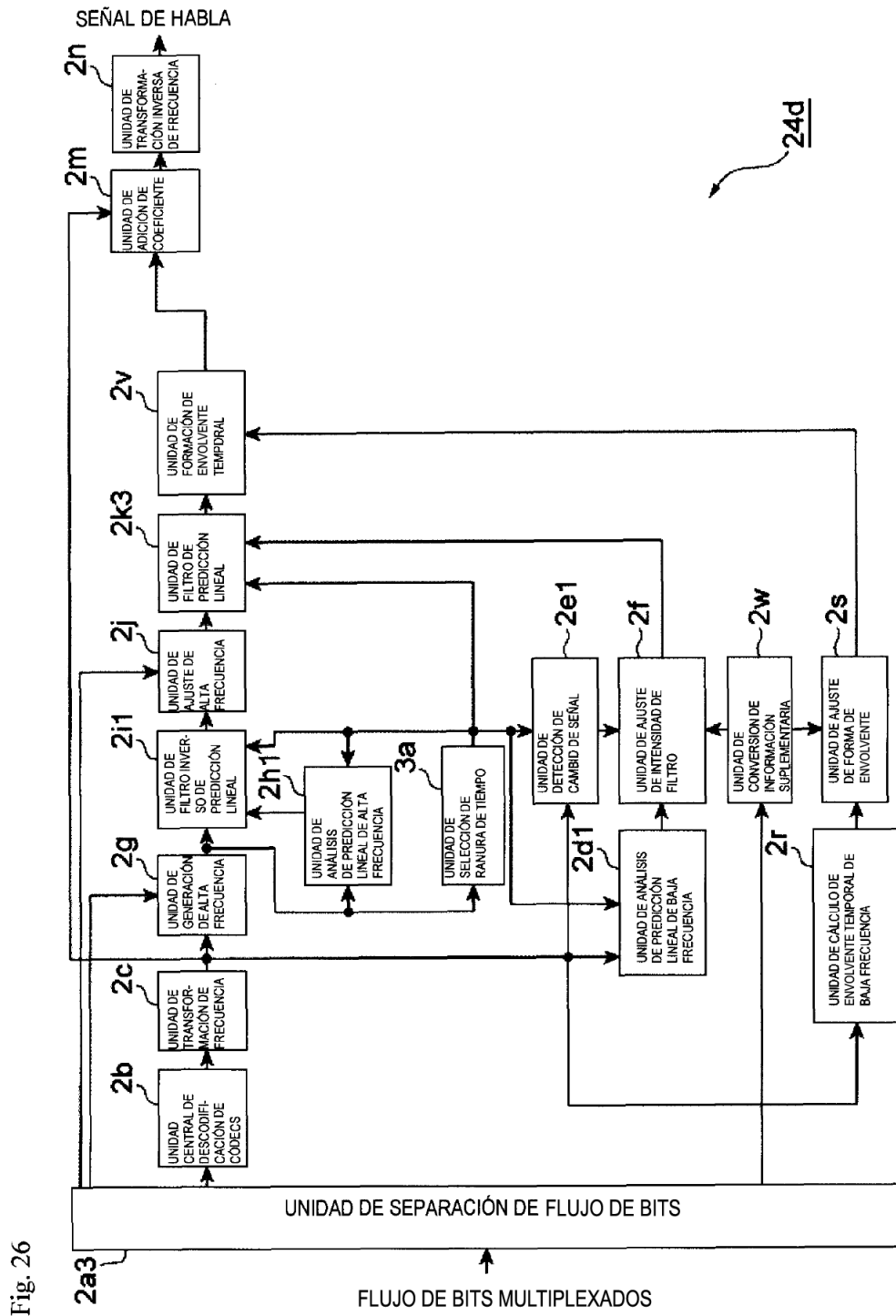
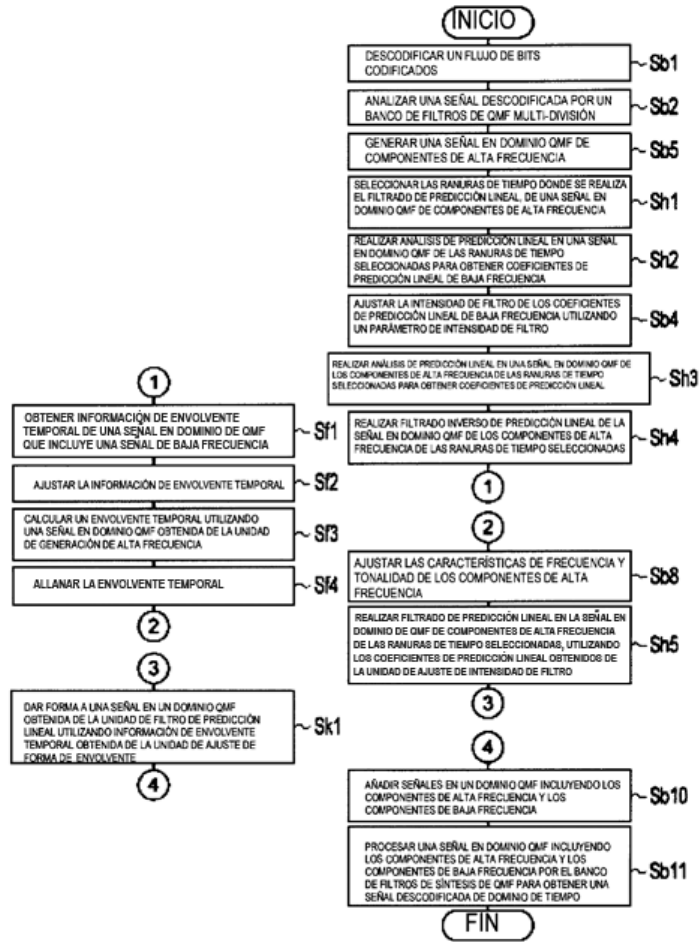


Fig. 26

Fig. 27



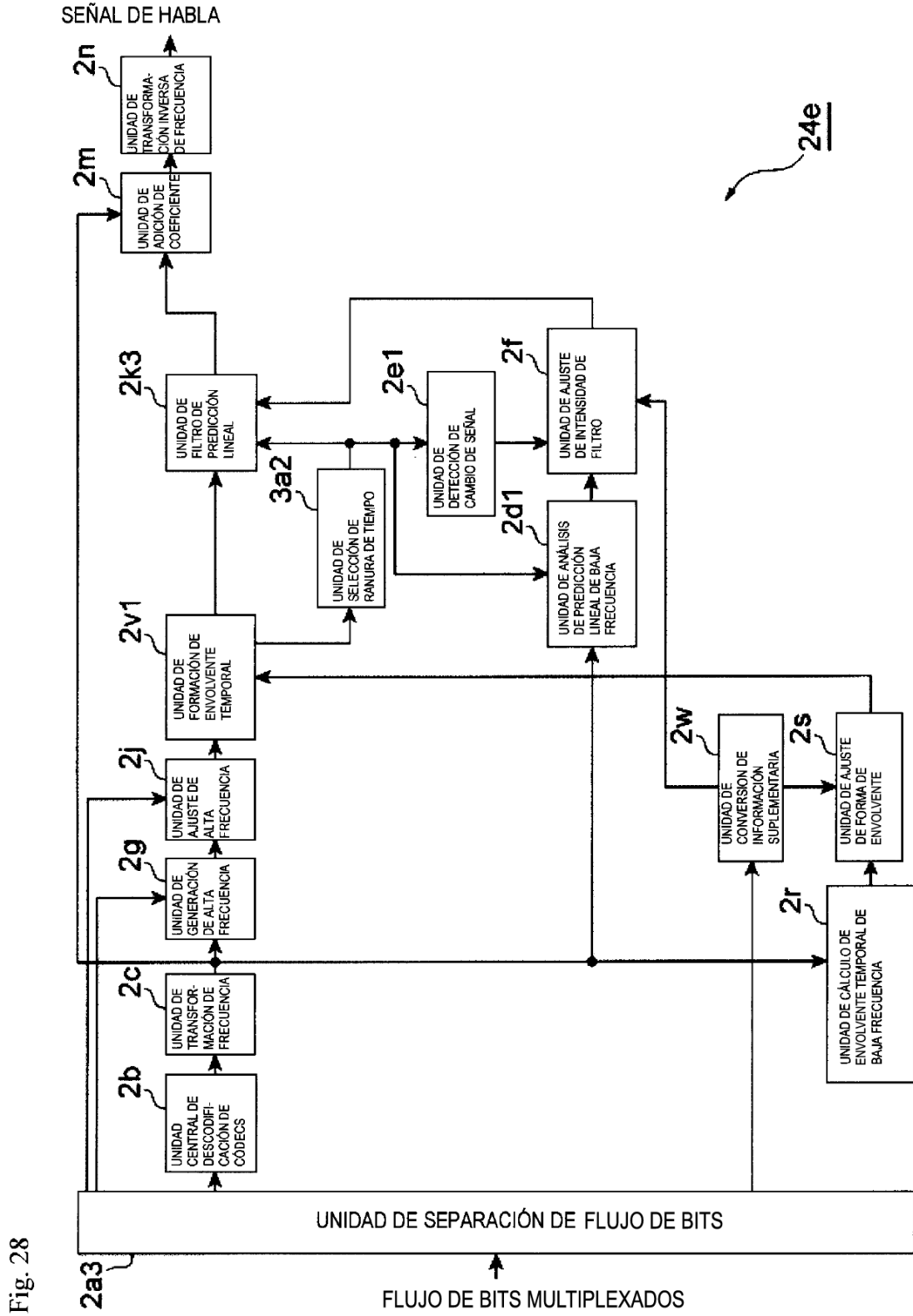
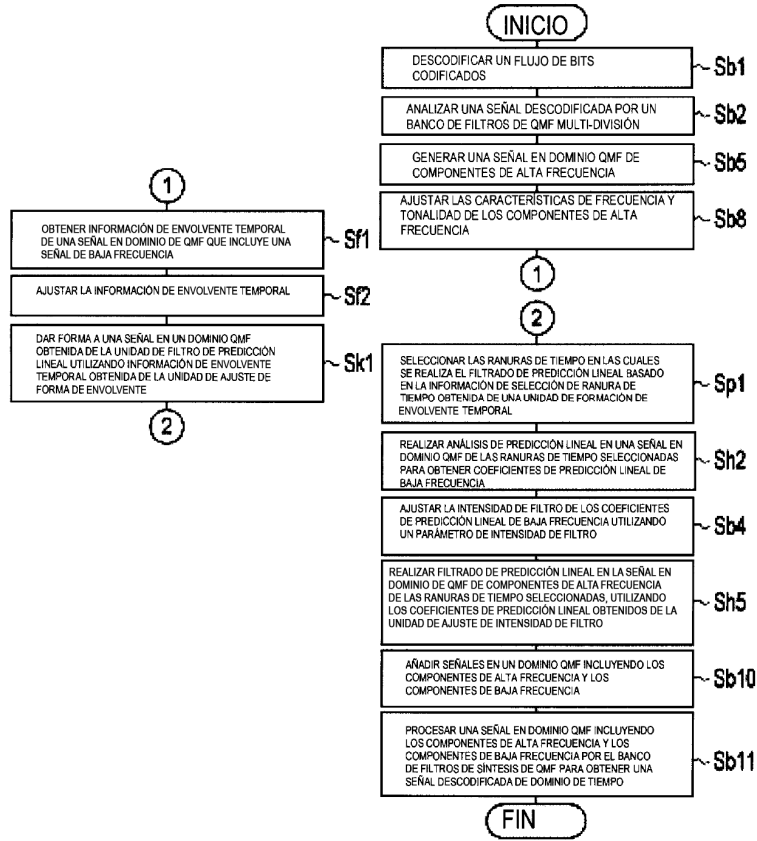


Fig. 28

Fig. 29



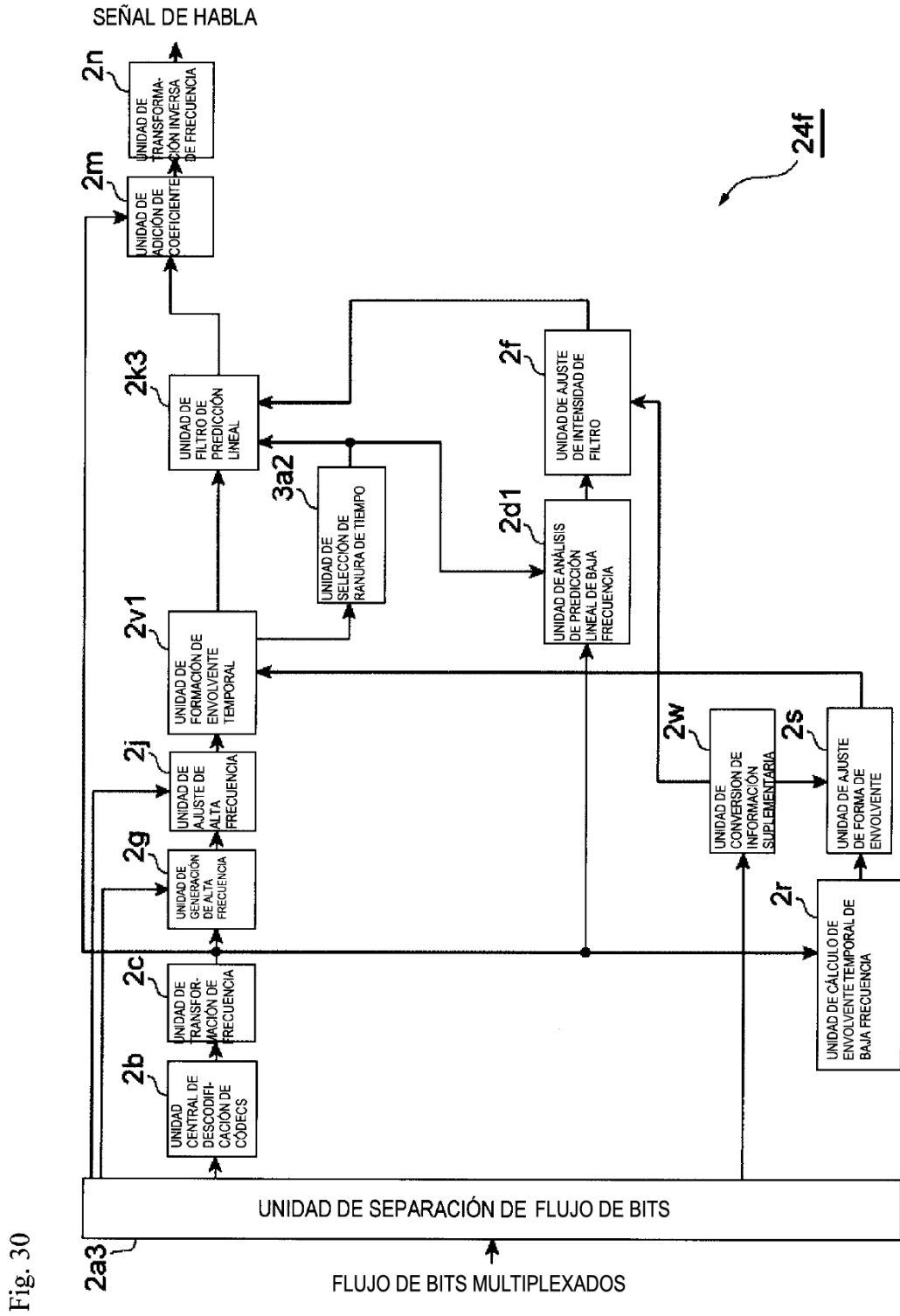


Fig. 30

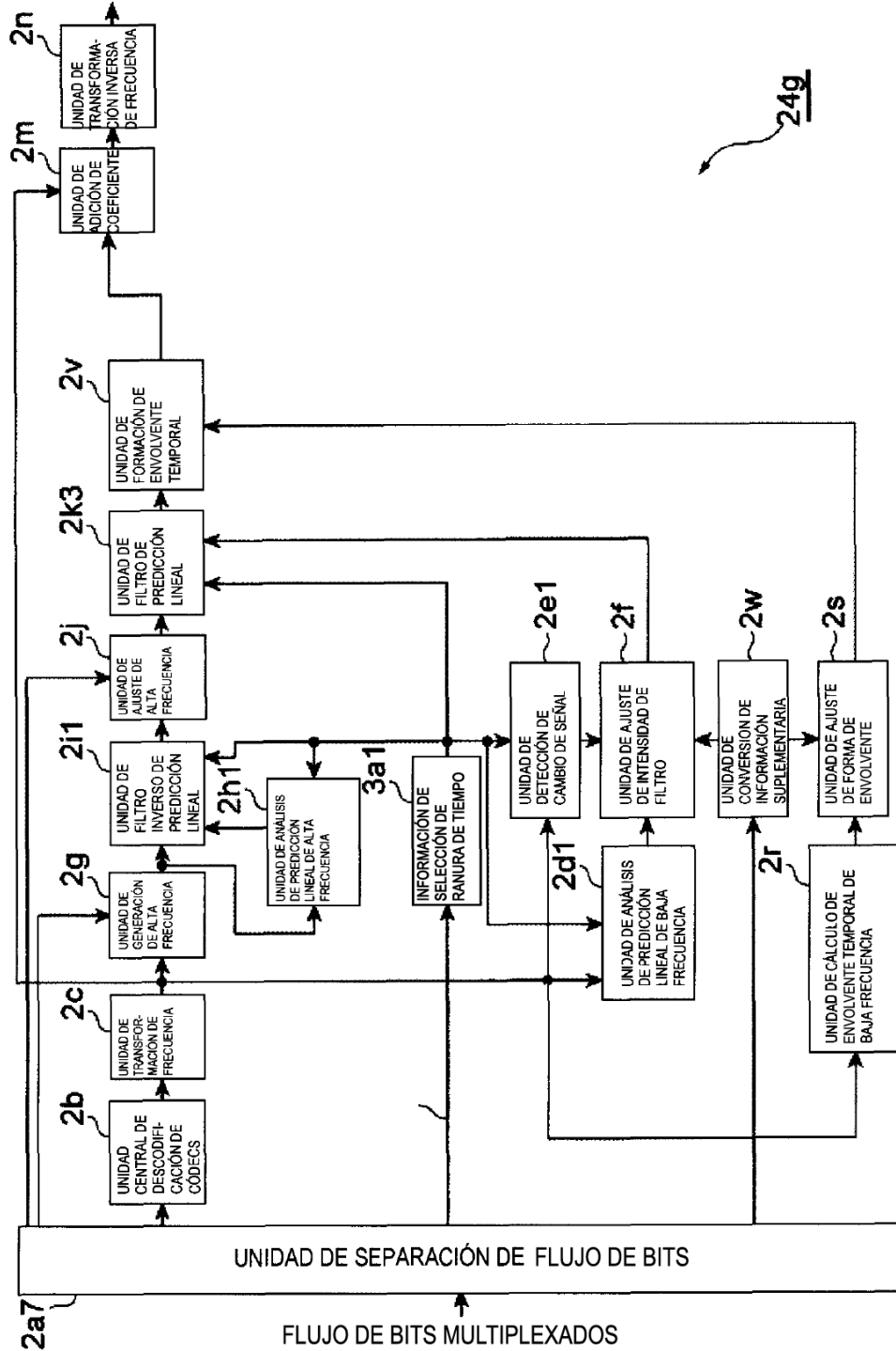
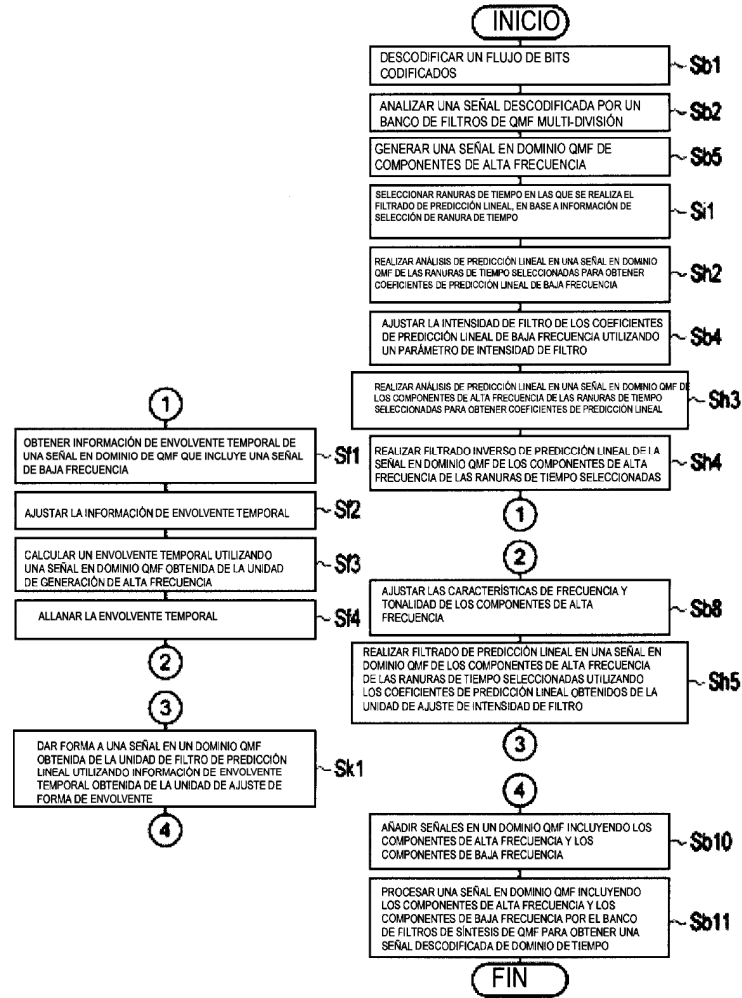


Fig. 31

Fig. 32



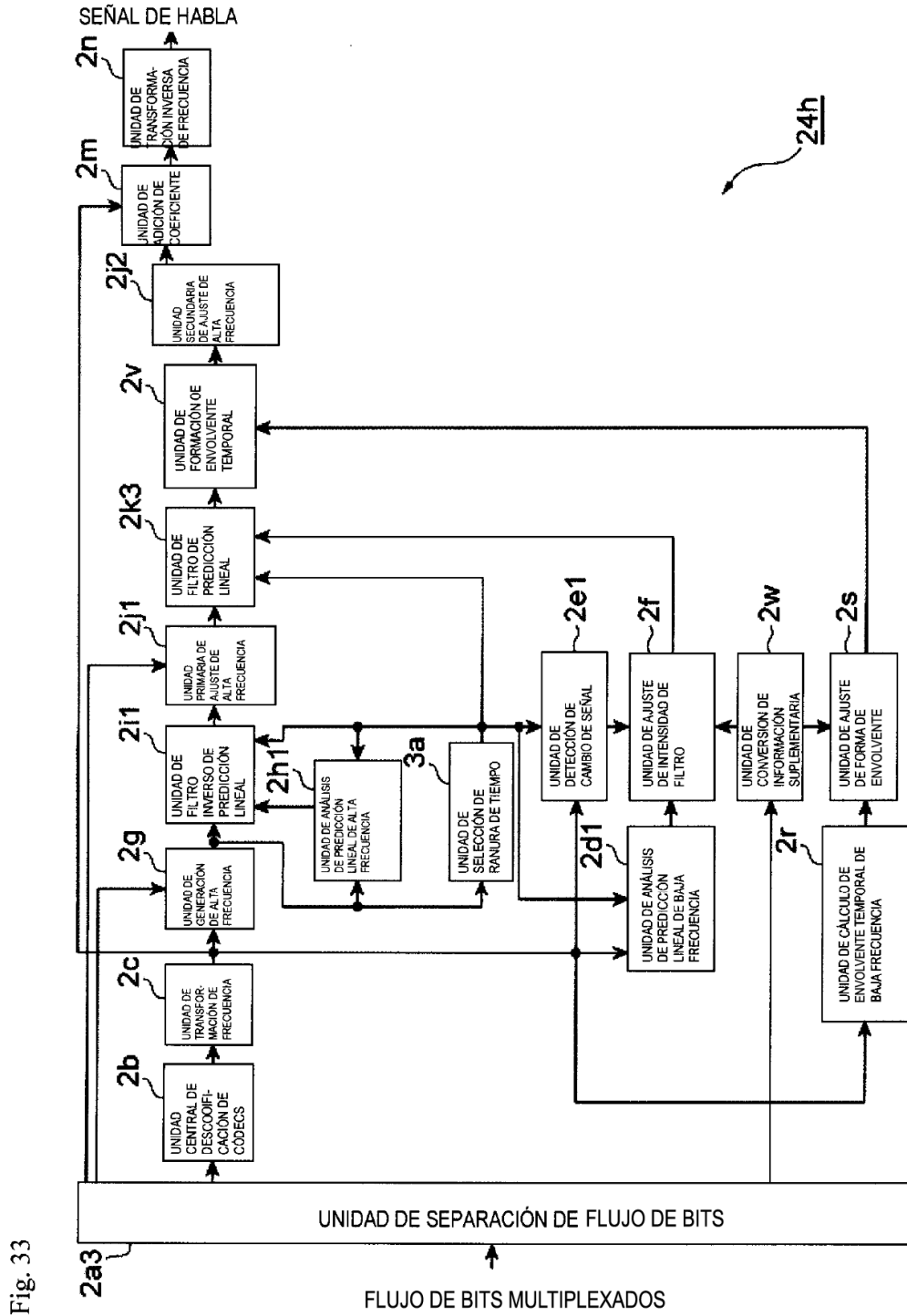
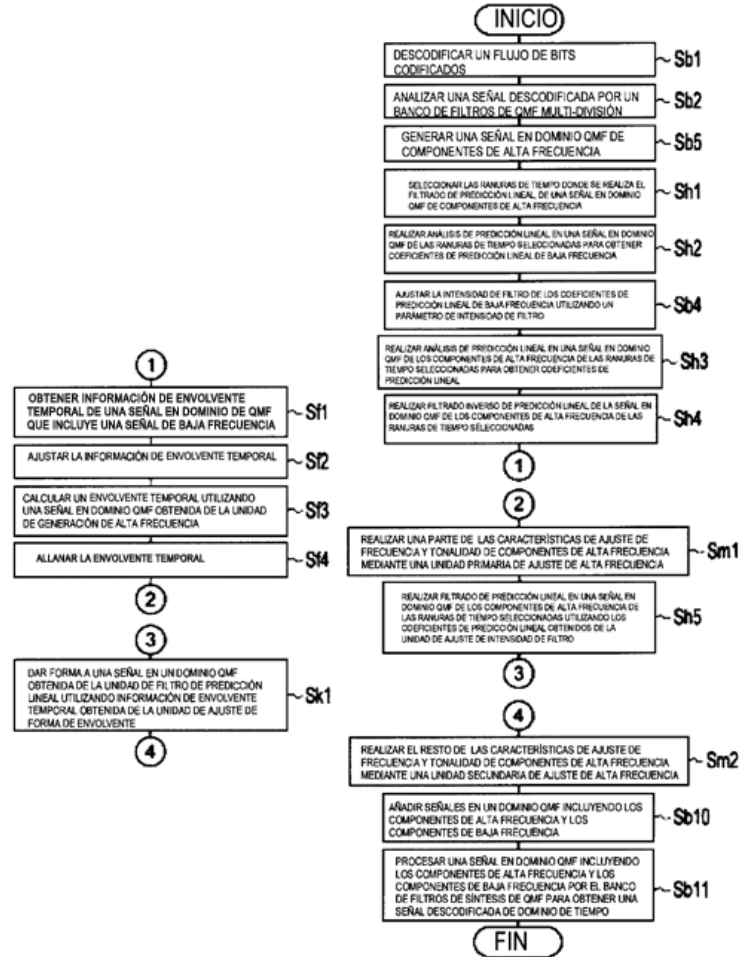


Fig. 33

Fig. 34



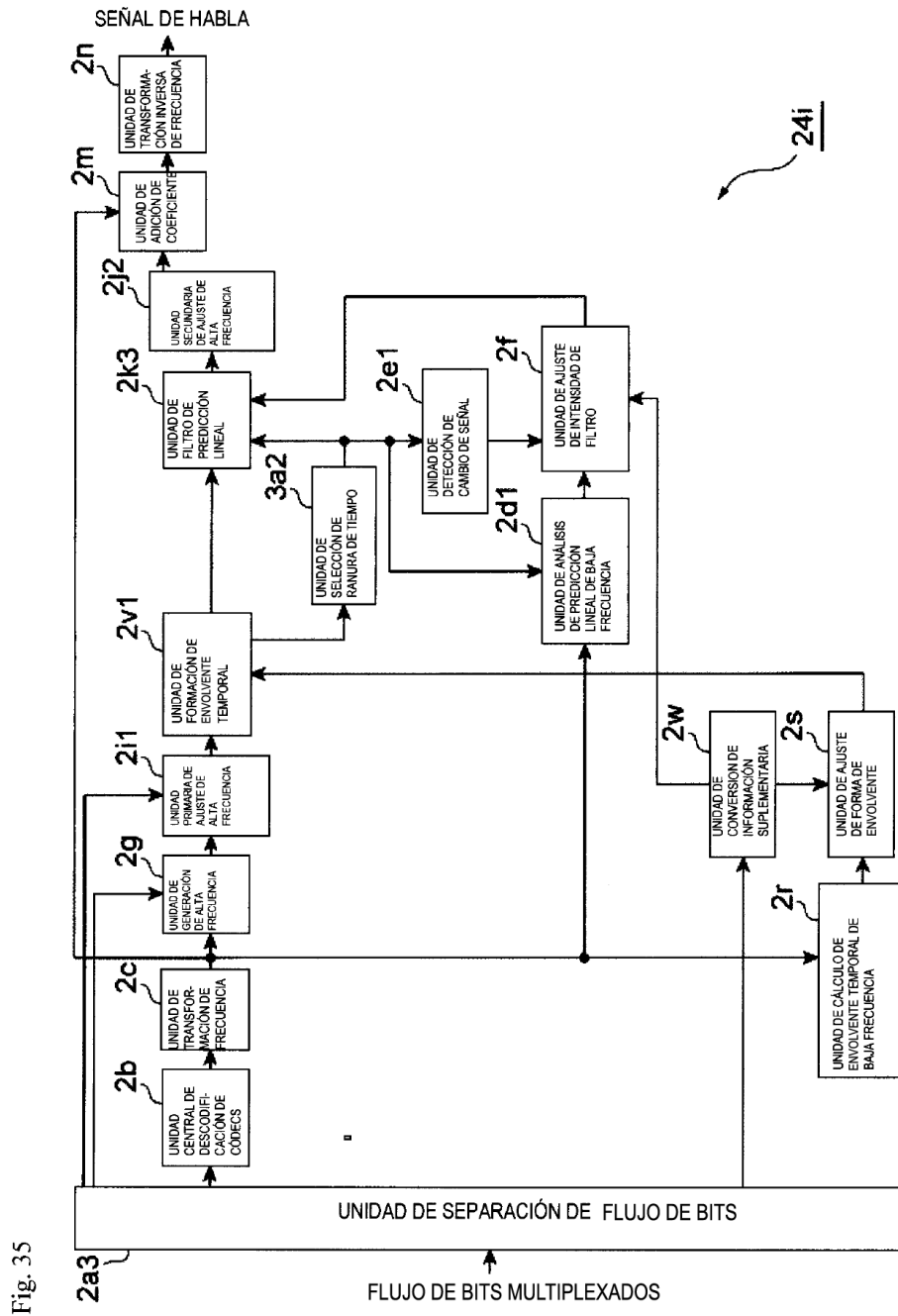
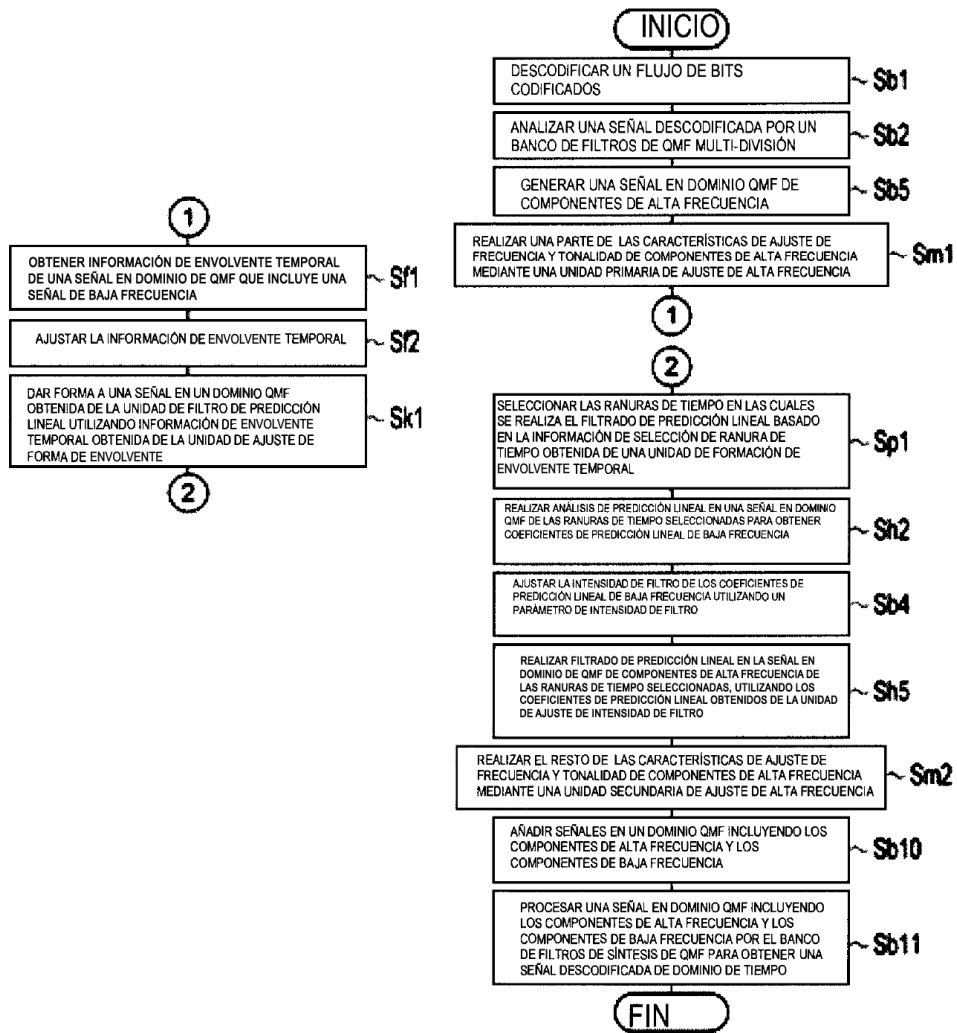


Fig. 35

Fig. 36



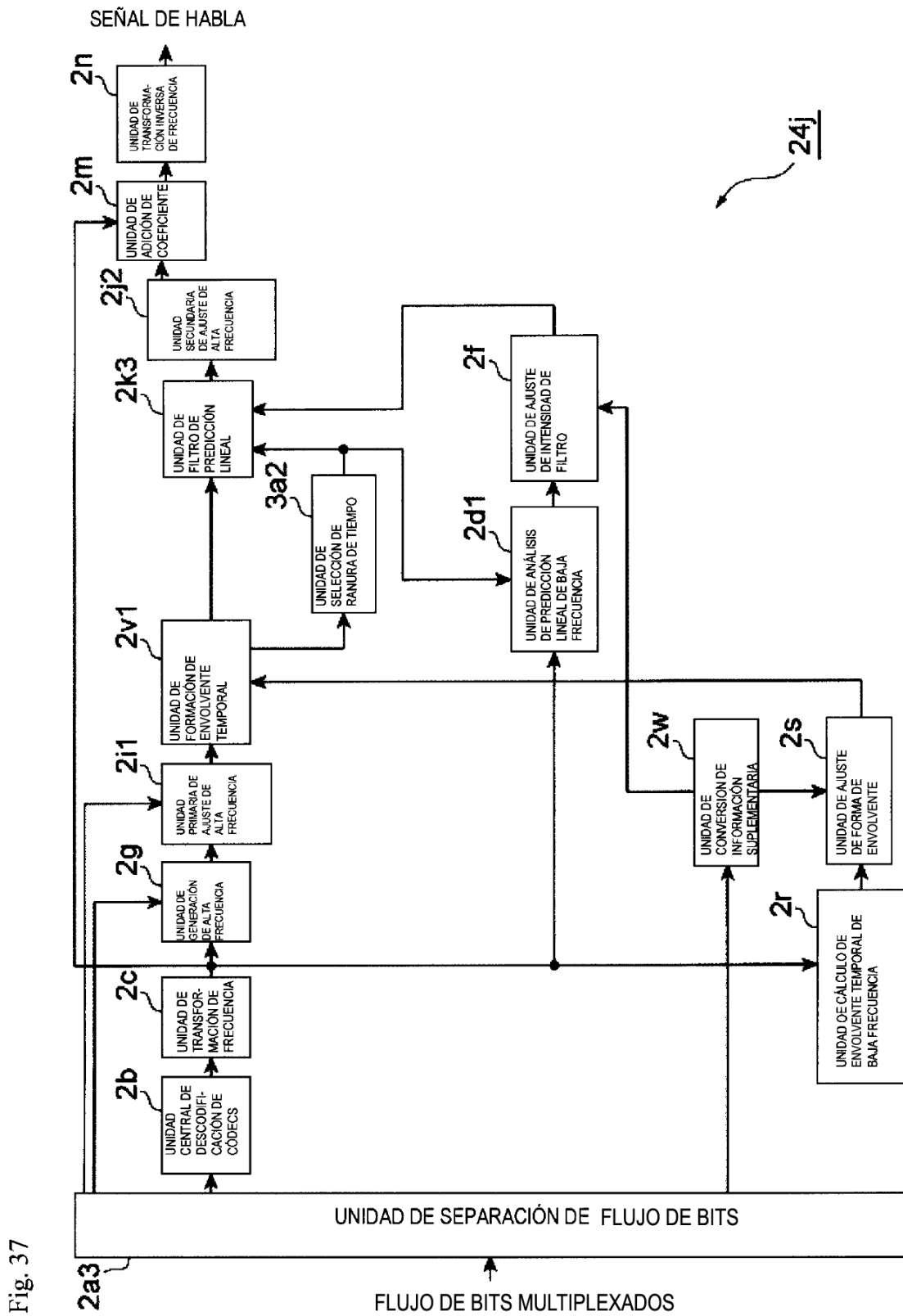


Fig. 37

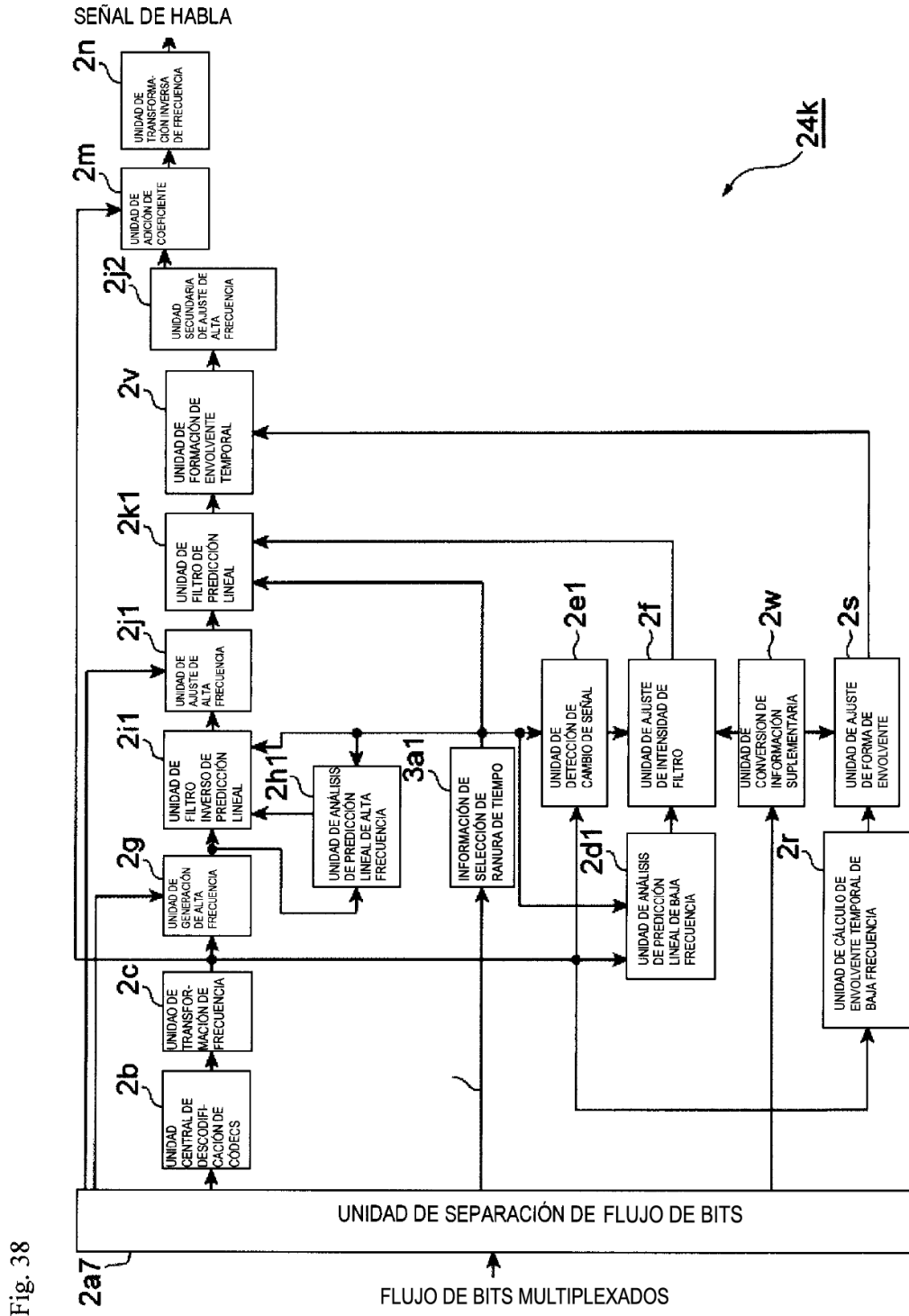
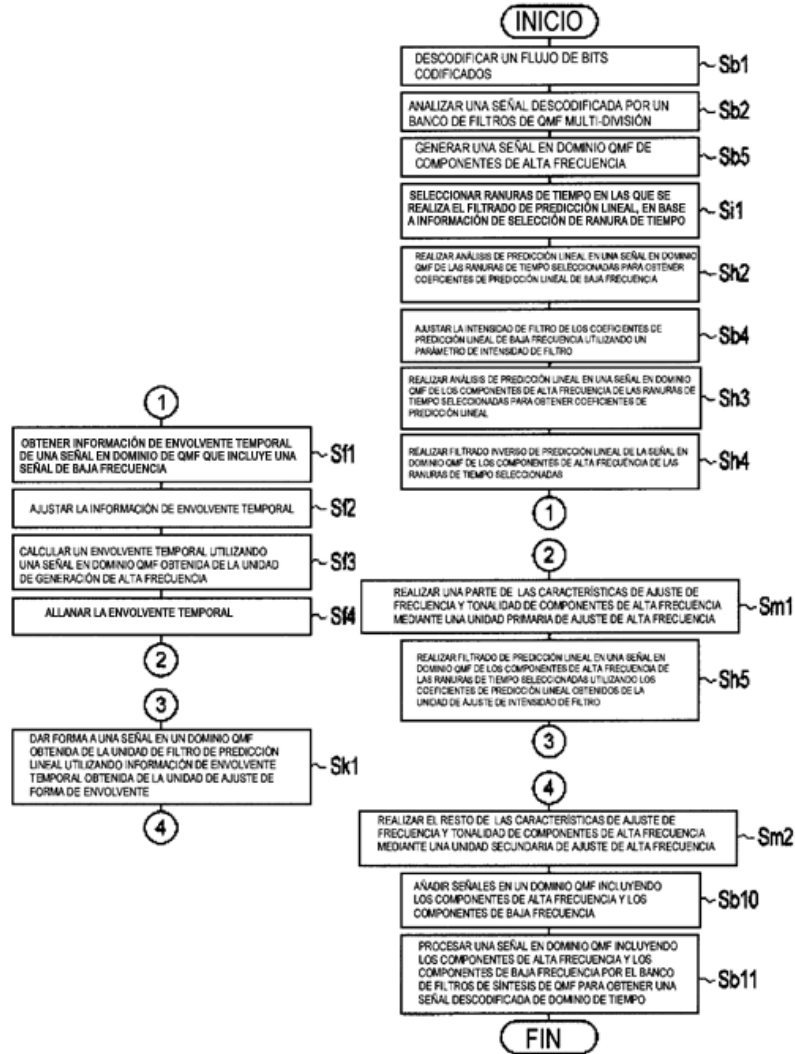


Fig. 38

Fig. 39



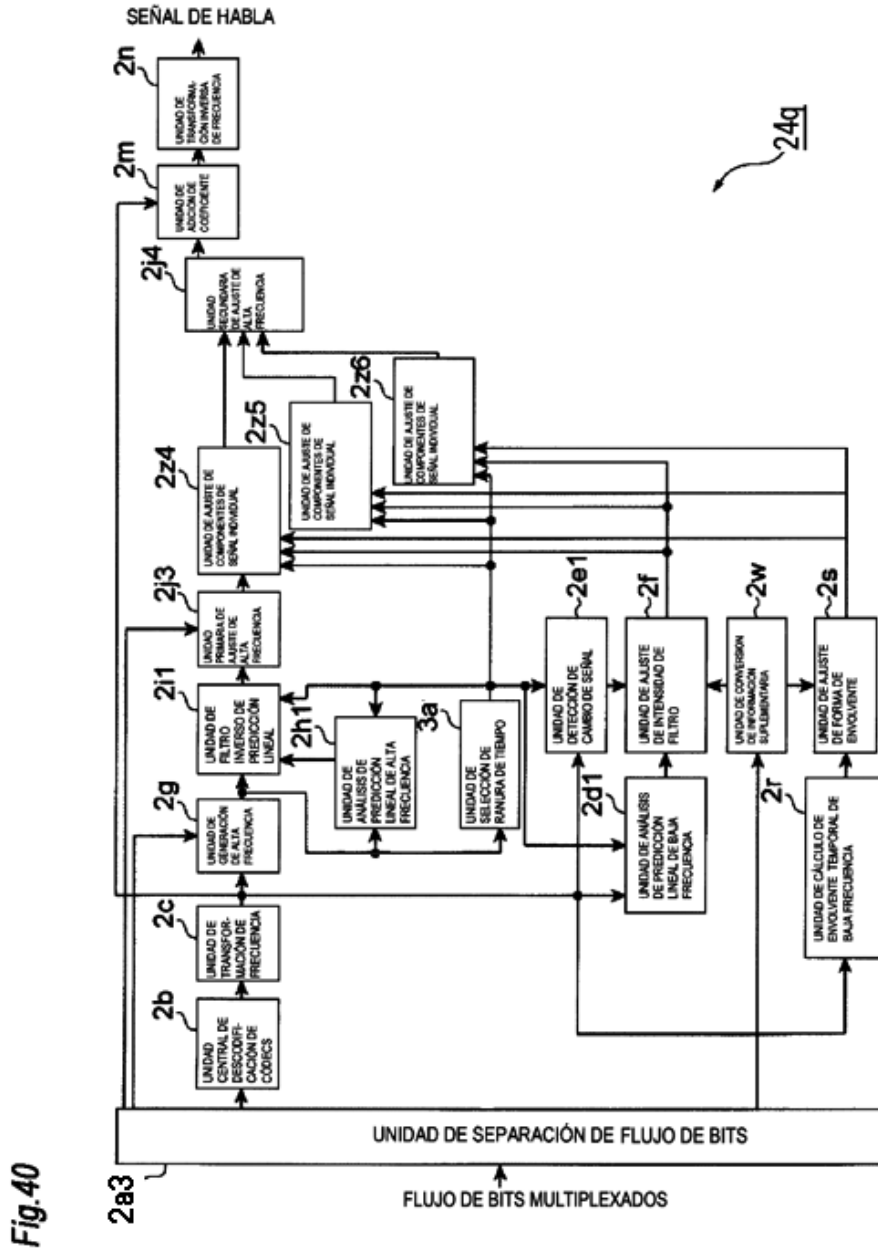
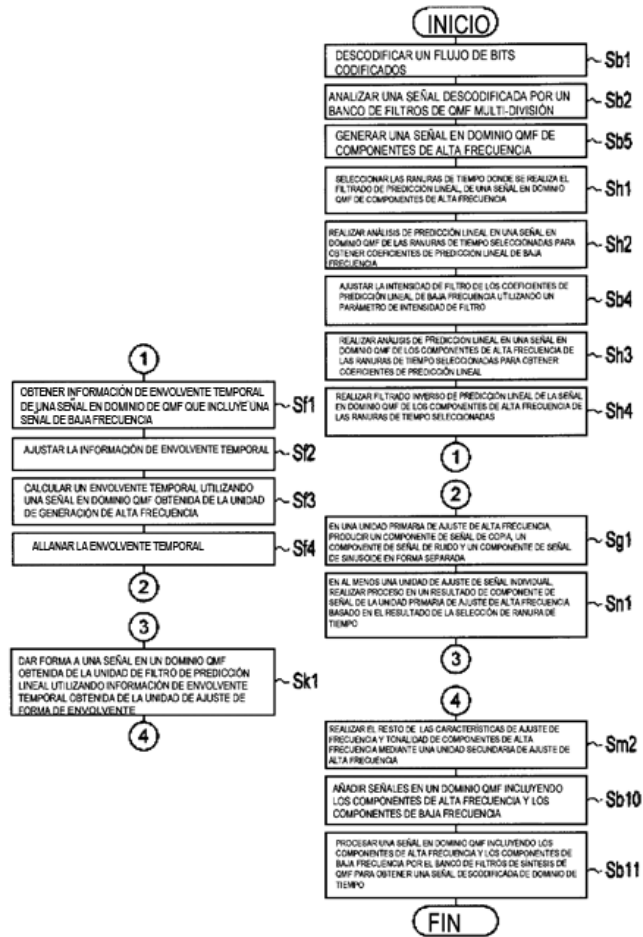


Fig.40

Fig. 41



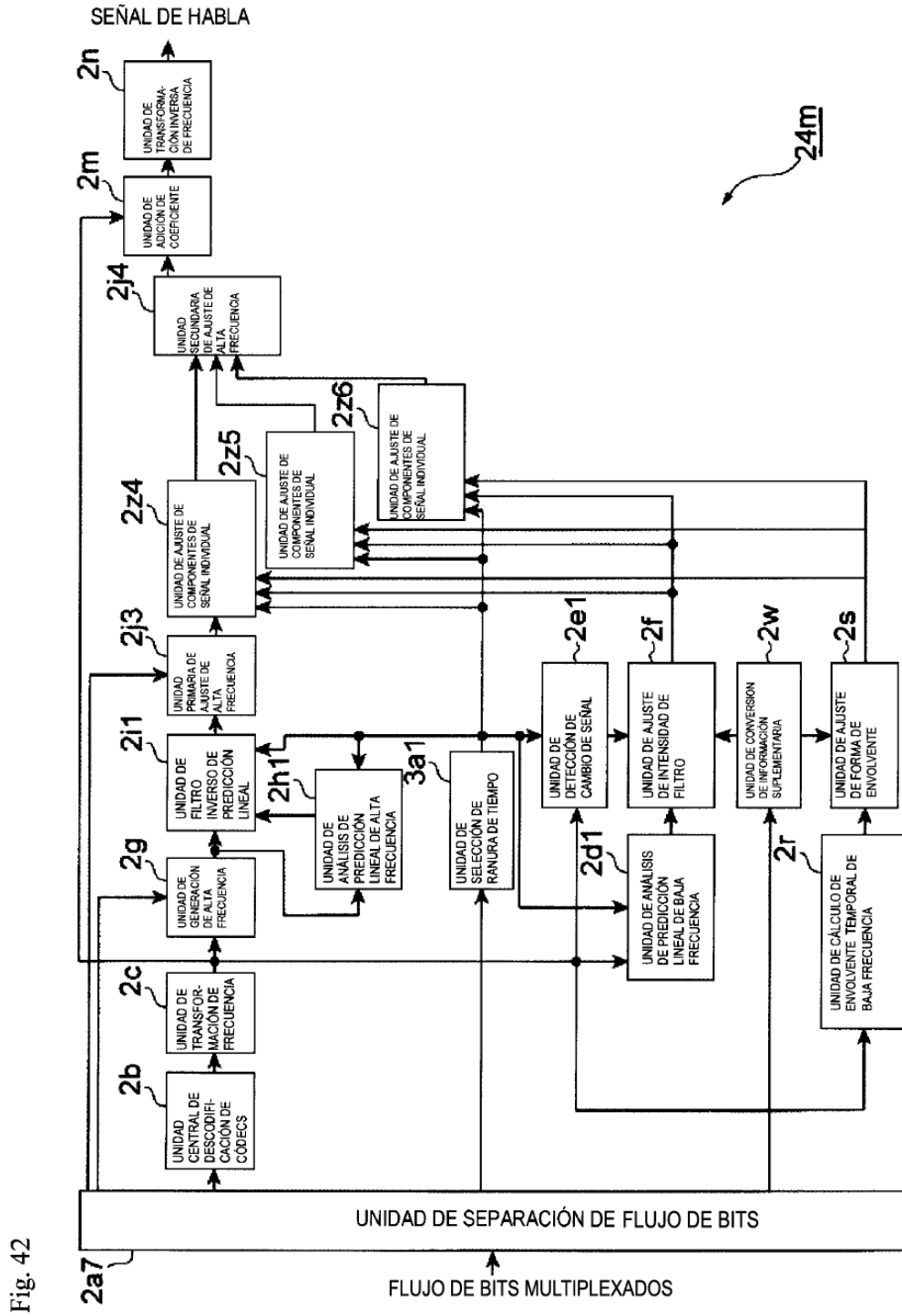
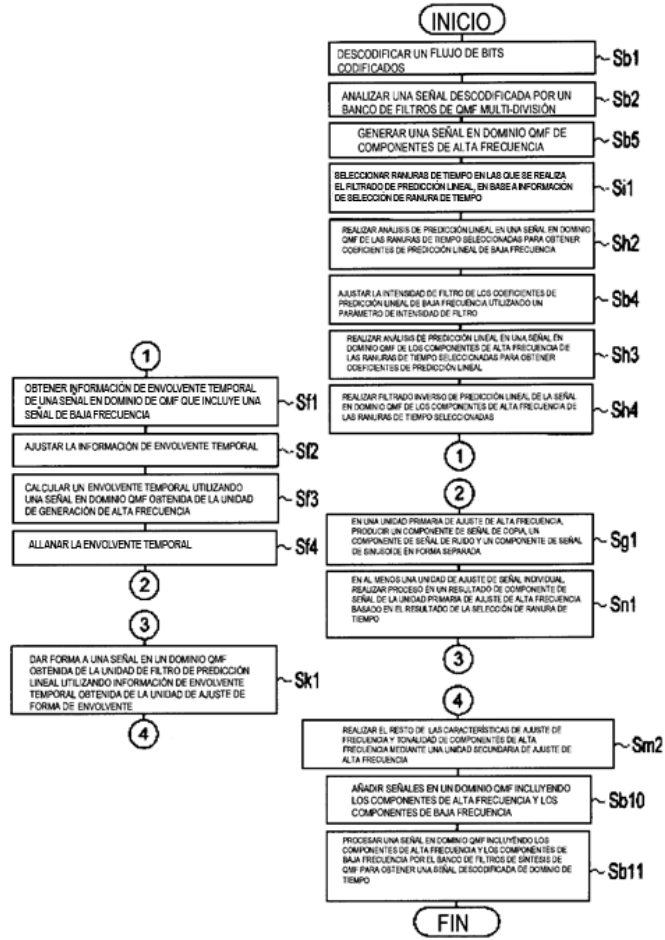


Fig. 42

Fig. 43



11b

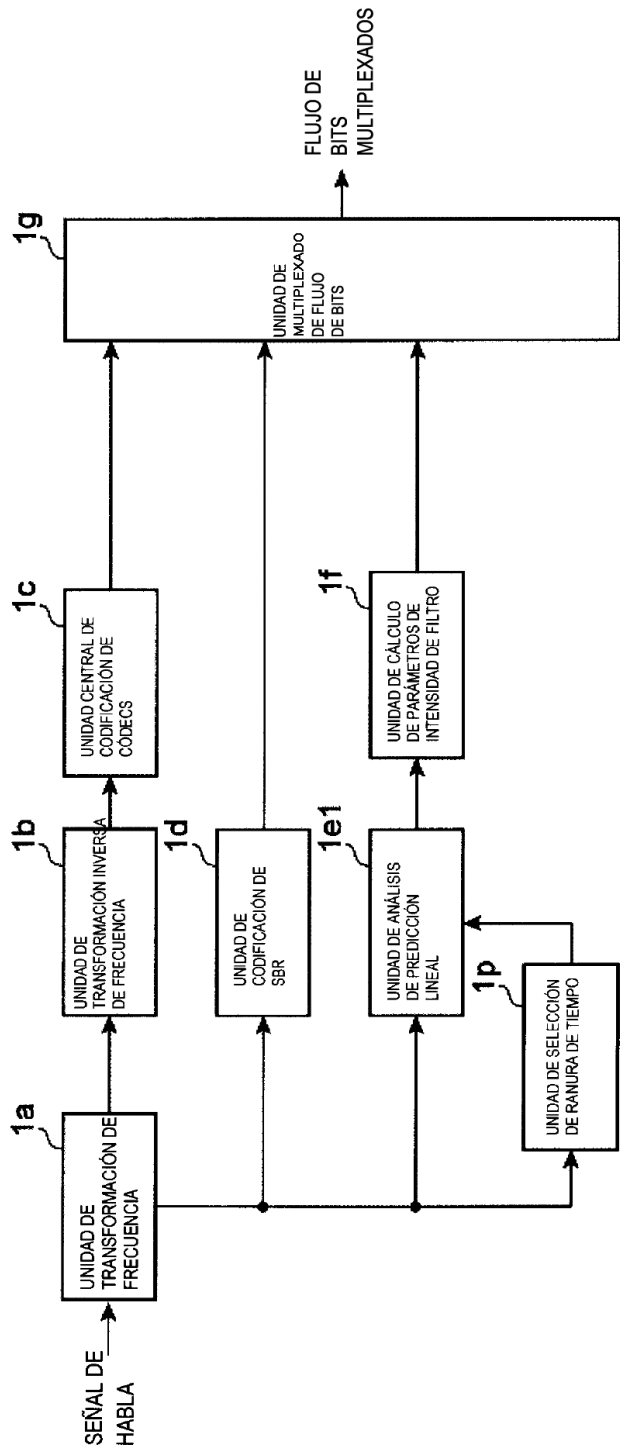


Fig. 44

Fig. 45

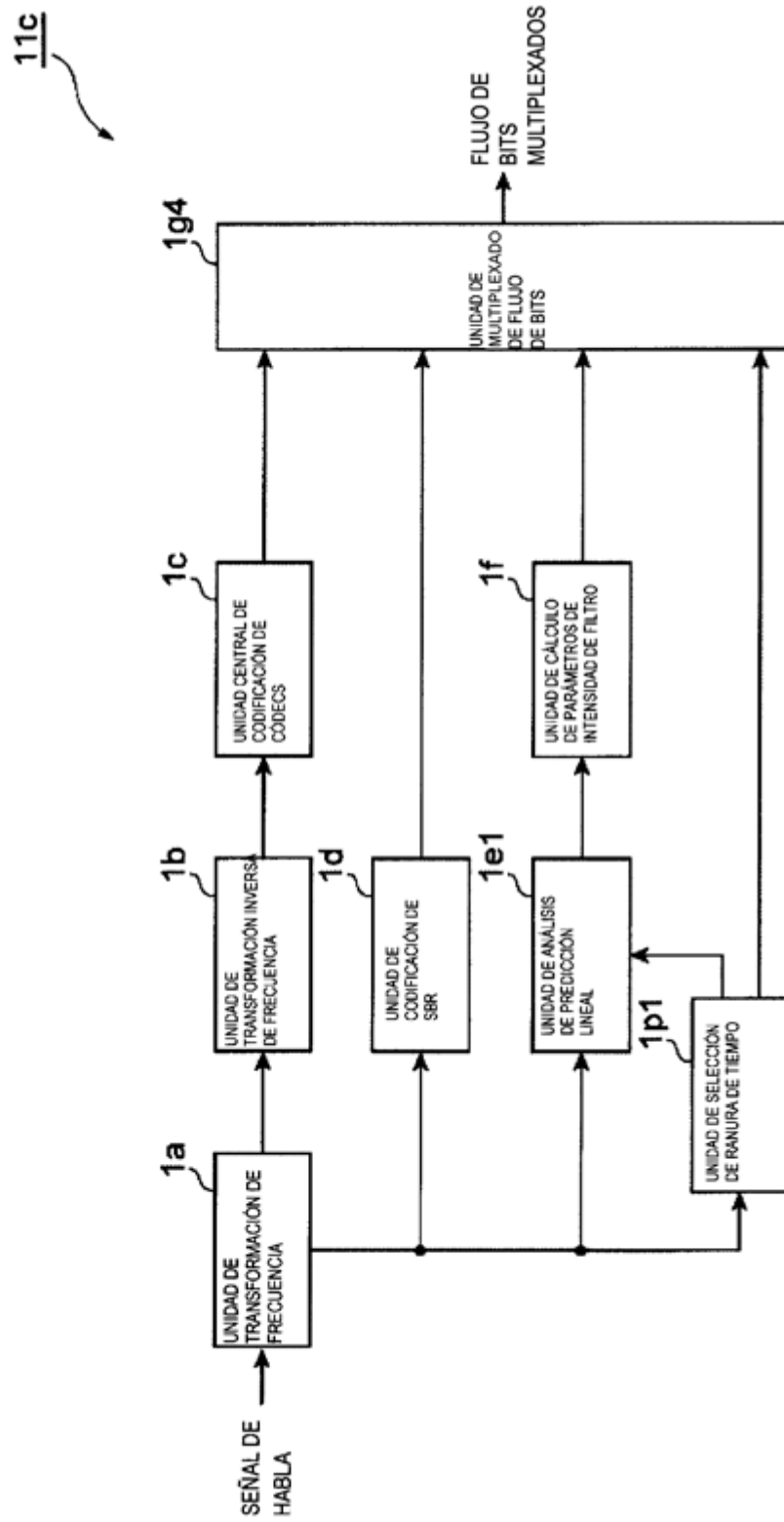
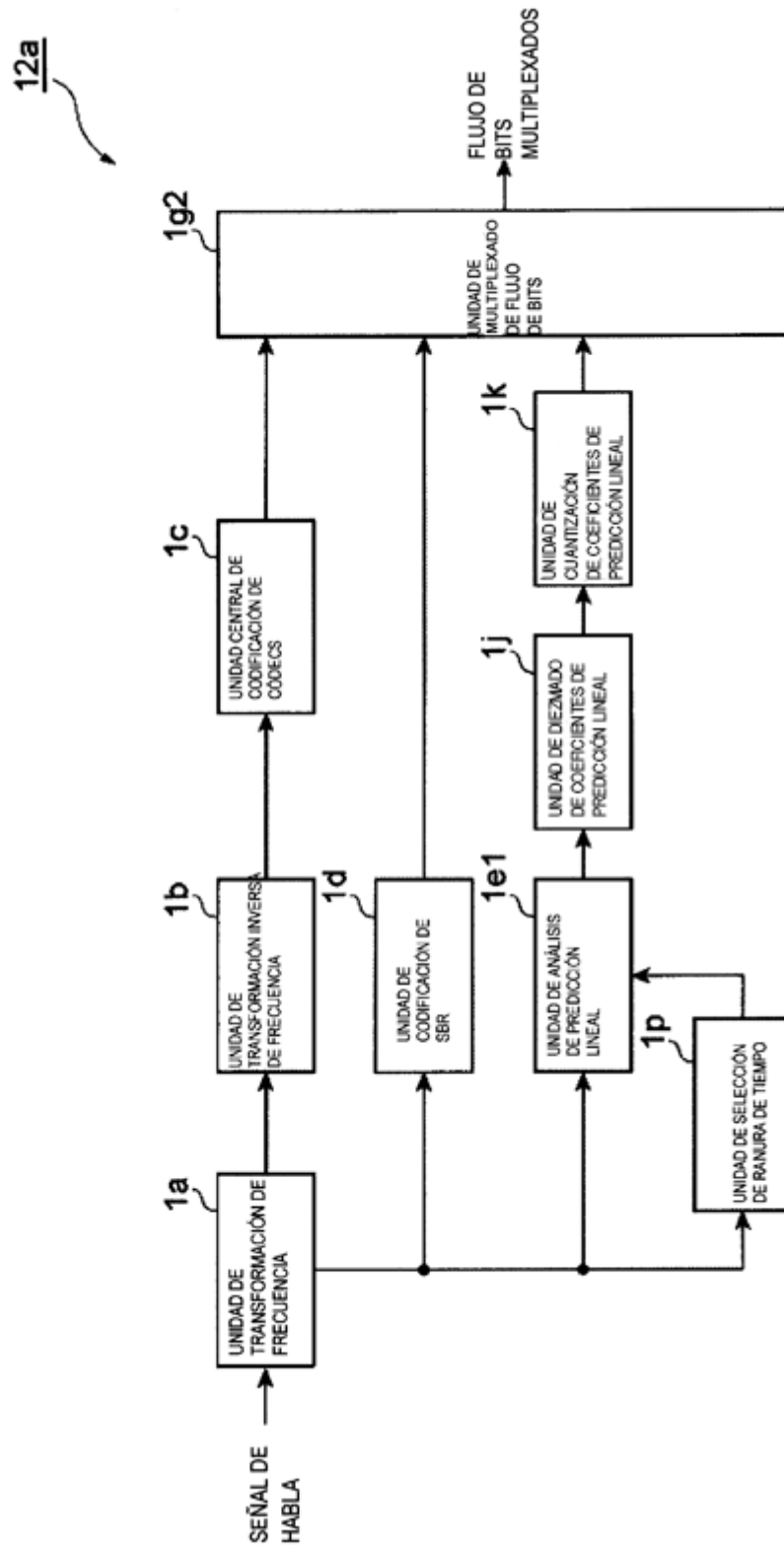


Fig. 46



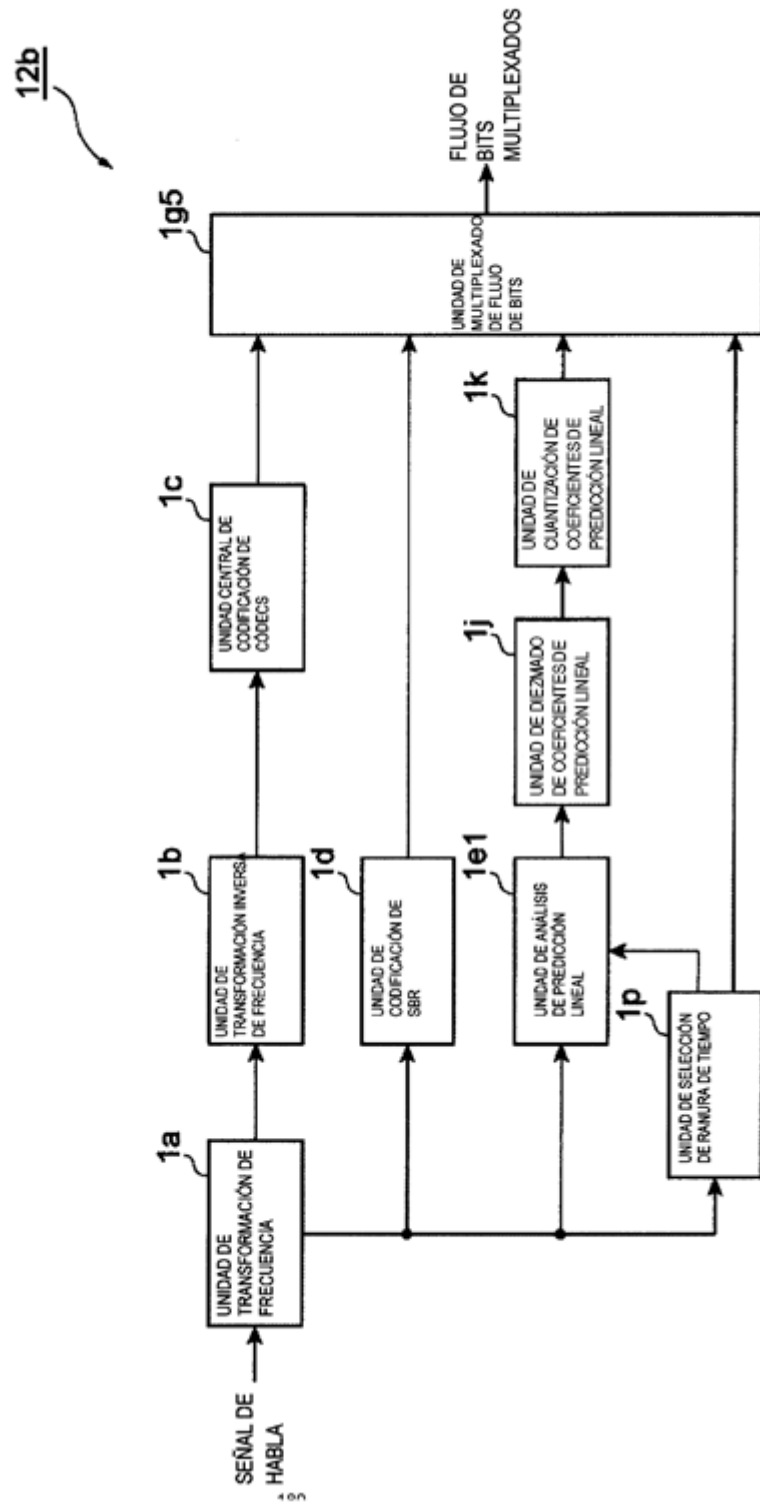


Fig. 47

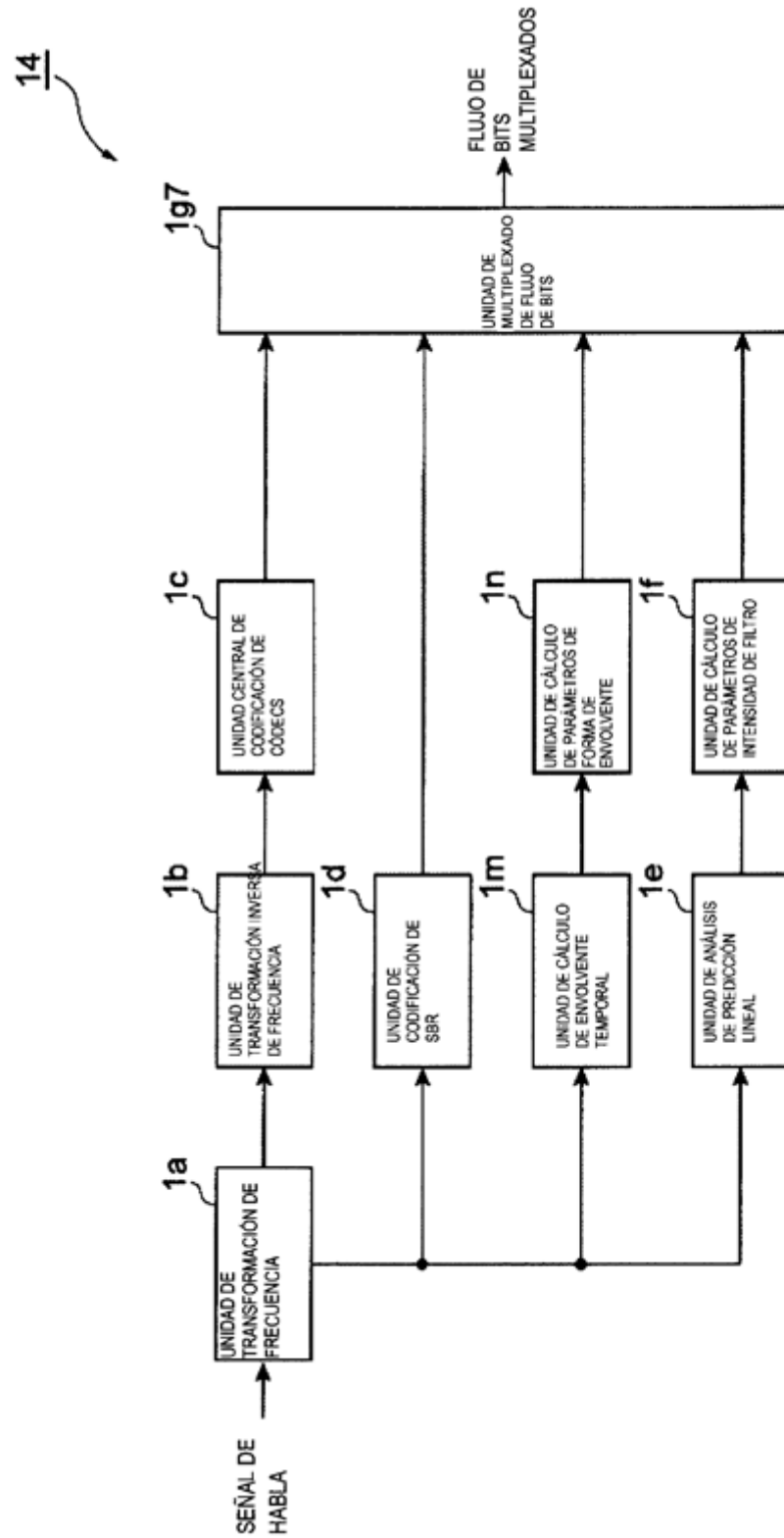


Fig. 48

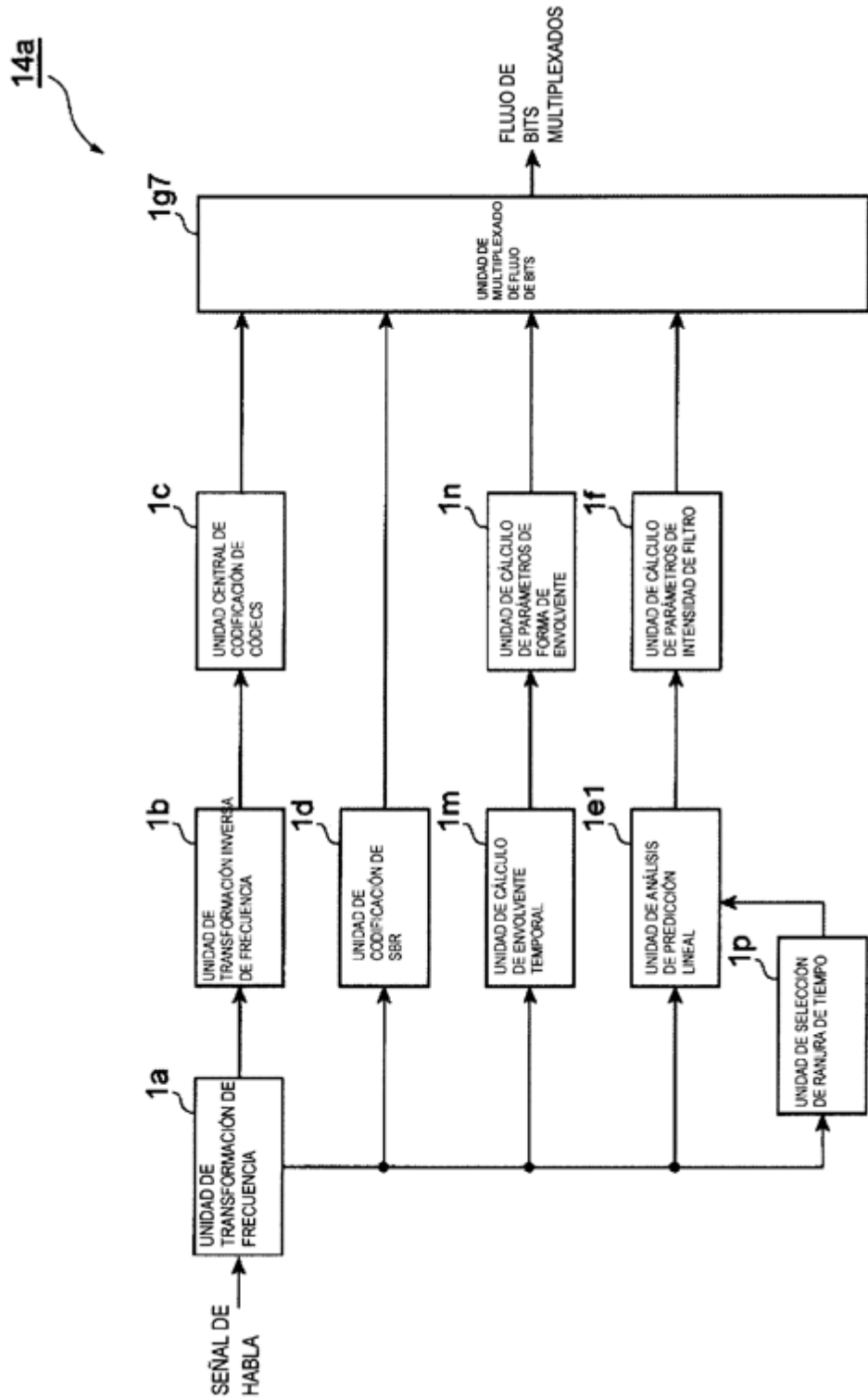


Fig. 49

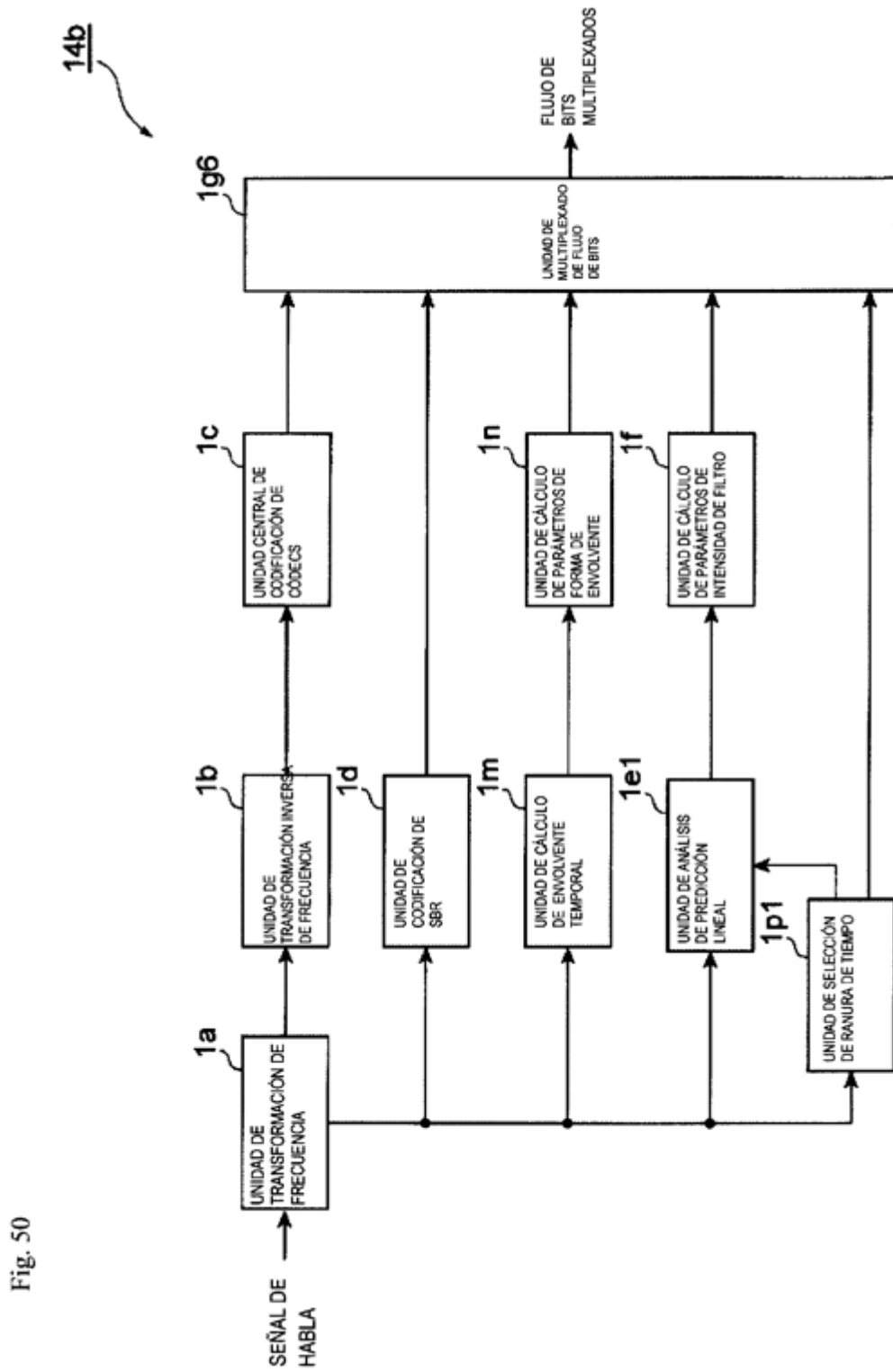


Fig. 50