

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 582 475**

51 Int. Cl.:

**G10L 21/038** (2013.01)

**G10L 19/26** (2013.01)

**G10L 19/12** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.09.2012 E 12845743 (9)**

97 Fecha y número de publicación de la concesión europea: **08.06.2016 EP 2791937**

54 Título: **Generación de una extensión de banda ancha de una señal de audio de ancho de banda extendido**

30 Prioridad:

**02.11.2011 US 201161554573 P**

**23.01.2012 US 201261589618 P**

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

**13.09.2016**

73 Titular/es:

**TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)  
(100.0%)  
164 83 Stockholm, SE**

72 Inventor/es:

**NORVELL, ERIK;  
GRANCHAROV, VOLODYA y  
TOFTGÅRD, TOMAS JANSSON**

74 Agente/Representante:

**DE ELZABURU MÁRQUEZ, Alberto**

**ES 2 582 475 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

Generación de una extensión de banda ancha de una señal de audio de ancho de banda extendido

### Sector técnico

5 La tecnología propuesta se refiere a la generación de una extensión de banda ancha de una señal de audio de ancho de banda extendido.

### Antecedentes

10 La mayoría de los sistemas de telecomunicación existentes operan en un ancho de banda de audio limitado. Como resultado de las limitaciones de los sistemas terrestres de telefonía, la mayoría de los servicios de voz están limitados a transmitir solo el extremo inferior del espectro. Aunque el ancho de banda de audio es suficiente para la mayoría de las conversaciones, se desea incrementar el ancho de banda para mejorar la inteligibilidad y el sentido de presencia. Aunque la capacidad en las redes de telecomunicación está aumentando continuamente, resulta todavía de gran interés limitar el ancho de banda necesario para el canal de comunicación. En las redes móviles menores anchos de banda de transmisión para cada llamada conducen a un menor consumo de energía, tanto en el dispositivo móvil como en la estación de base. Esto se traduce en un ahorro de energía y de coste para el operador de la red móvil, mientras que el usuario final experimentará una mayor vida de la batería y un mayor tiempo para hablar. Además, con un menor ancho de banda consumido por usuario, la red de telefonía móvil puede dar servicio, en paralelo, a un mayor número de usuarios.

20 Una propiedad del sistema auditivo humano es que la percepción depende de la frecuencia. En particular, nuestro oído es menos preciso para frecuencias más altas. Esto ha inspirado las llamadas técnicas de extensión del ancho de banda (BWE – BandWidth Extension, en inglés), en las que una banda de alta frecuencia es reconstruida a partir de una banda de baja frecuencia utilizando recursos limitados.

25 La BWE convencional utiliza una representación de la envolvente espectral de la señal de banda ancha extendida, y reproduce la estructura fina espectral de la señal, utilizando una versión modificada de la señal de banda baja. Si la envolvente de banda alta se representa mediante un filtro, la señal de estructura fina a menudo se denomina señal de excitación. Una representación precisa de la envolvente de banda alta es perceptualmente más importante que la estructura fina. En consecuencia, es habitual que los recursos disponibles en términos de bits se consuman en la representación de la envolvente, mientras que la estructura fina se reconstruye a partir de la señal codificada de banda baja, sin información lateral adicional. El concepto básico de la BWE se ilustra en la figura 1.

30 La tecnología de BWE ha sido aplicada en una variedad de sistemas de codificación de audio. Por ejemplo, el 3GPP AMR-WB+, [1], utiliza una BWE en el dominio del tiempo basada en un codificador de banda baja que conmuta entre la codificación de conversación de Predictor lineal excitado mediante código (CELP - Code Excited Lineal Predictor, en inglés) y la codificación de residuo codificado mediante transformada (TCX – Transform Coded Residual, en inglés). Otro ejemplo es el códec de audio basado en transformada eAAC del 3GPP, que efectúa una variante de la BWE en el dominio de la transformada, denominada replicación de banda espectral (SBR – Spectral Band Replication, en inglés), [2]. Aquí, la excitación se crea utilizando una mezcla de componentes tonales generados a partir de la excitación de banda baja y de una fuente de ruido con el fin de hacer coincidir la relación de tonal a ruido de la señal de entrada. En general, el nivel de ruido de la señal se puede describir como una medida de cuán plano es un espectro, por ejemplo, utilizando una medida de planicidad espectral. El nivel de ruido se puede describir asimismo como no tonalidad, aleatoriedad o ausencia de estructura de la excitación. Aumentar el nivel de ruido de una señal es hacerla más parecida a un ruido, por ejemplo, mezclando la señal con una señal de ruido desde, por ejemplo, un generador de números aleatorios o cualquier otra fuente de ruido. Esto se puede efectuar asimismo modificando el espectro de la señal para hacerla más plana.

45 La estructura espectral fina de la banda baja puede ser muy diferente de la estructura fina encontrada en la banda alta. En particular, la combinación de una excitación generada a partir de la señal de banda baja junto con la envolvente de banda alta puede producir aberraciones no deseadas, dado que el carácter armónico residente o la forma de la excitación pueden ser resaltadas mediante la conformación de la envolvente de una manera incontrolada. Como medida de seguridad, es habitual aplanar la envolvente de banda alta con el fin de limitar una interacción no deseada entre la excitación y la envolvente. Aunque esta solución puede proporcionar un compromiso razonable, la envolvente más plana se puede percibir como más ruidosa, y la envolvente de banda alta será menos precisa. Gustaffson et al: "Speech Band width Extension" describe la extensión del ancho de banda con el control del nivel de ruido de la excitación de banda alta y la aplicación de un post-filtro.

### Compendio

Un objetivo de la tecnología propuesta es un mayor control de la generación de la extensión de banda alta de una señal de audio de ancho de banda extendido.

55 Este objetivo se consigue de acuerdo con las reivindicaciones adjuntas.

Un primer aspecto de la tecnología propuesta implica un método de generar una extensión de banda alta de una señal de audio a partir de una envolvente y de una excitación. El método incluye la etapa de controlar conjuntamente la forma de la envolvente y el nivel de ruido de la excitación mediante un parámetro de control común.

5 Un segundo aspecto de la tecnología propuesta implica un decodificador de audio configurado para generar una extensión de banda alta de una señal de audio a partir de una envolvente y de una excitación. El decodificador de audio incluye una disposición de control configurada para controlar conjuntamente la forma de la envolvente y el nivel de ruido de la excitación, mediante un parámetro de control común.

Un tercer aspecto de la tecnología propuesta implica un equipo de usuario (UE – User Equipment, en inglés) que incluye un decodificador de audio de acuerdo con el segundo aspecto.

10 Un cuarto aspecto de la tecnología propuesta implica un codificador de audio que incluye un estimador de planicidad espectral configurado para determinar, para su transmisión a un decodificador, una medida de la planicidad espectral de una señal de banda alta.

15 La tecnología propuesta permite una estructura más pronunciada de la envolvente que enmascara aberraciones perceptuales creadas mediante excitaciones de banda alta generadas artificialmente. Al mismo tiempo, un control conjunto de la estructura de la envolvente y del nivel de ruido de la excitación mejora la naturalidad de la señal de audio reconstruida.

### Breve descripción de los dibujos

La tecnología propuesta, junto con otros objetivos y ventajas de la misma, se puede comprender mejor haciendo referencia a la siguiente descripción, tomada junto con los dibujos que se acompañan.

20 La figura 1 ilustra el concepto básico de la técnica de BWE en forma de un espectro de frecuencia. La señal codificada de banda baja es extendida con una banda alta mediante la utilización de una envolvente de banda alta y de una señal de excitación que se genera a partir de la señal de banda baja.

25 La figura 2 ilustra un sistema de BWE a modo de ejemplo con un códec de CELP para la banda baja, y en el que la banda superior está reconstruida utilizando una banda baja de predictor lineal (LP – Linear Predictor, en inglés) y una señal de excitación, que se genera a partir de parámetros de salida modificados del decodificador de CELP.

La figura 3 ilustra un decodificador de BWE a modo de ejemplo que tiene un codificador correspondiente, como se muestra en la figura 2. La excitación modulada se mezcla con una señal de ruido de un generador de ruido.

La figura 4 ilustra una realización a modo de ejemplo de la tecnología propuesta en un sistema de decodificador de CELP con una disposición de control conjunta para la mezcla de la excitación y la forma espectral.

30 La figura 5 ilustra un ejemplo de un espectro de LP de entrada y un espectro de LP que ha sido resaltado con un post-filtro.

La figura 6 ilustra una realización a modo de ejemplo de un codificador que utiliza un análisis de planicidad espectral basado en coeficientes de codificación predictiva lineal (LPC – Lineal Predictive Coding, en inglés).

35 La figura 7 ilustra una realización a modo de ejemplo de un decodificador correspondiente al codificador de la figura 6, que utiliza el parámetro de planicidad transmitido para un control conjunto de la envolvente espectral y de la estructura de la excitación.

La figura 8 ilustra un ejemplo de un códec de audio basado en transformada que tiene una codificación conjunta de la envolvente para todo el espectro, y que emplea técnicas de BWE para obtener la estructura fina espectral de la banda alta.

40 La figura 9 ilustra un ejemplo de un decodificador de BWE que pertenece a un codificador correspondiente, tal como se muestra en la figura 8. La excitación modulada se modifica utilizando un compresor para obtener una estructura fina más plana en la excitación de banda alta.

45 La figura 10 ilustra una realización a modo de ejemplo de la tecnología propuesta en un sistema de decodificador basado en transformada con un controlador conjunto para la compresión de la excitación y la expansión de la envolvente.

La figura 11 ilustra una realización a modo de ejemplo de un codificador, que tiene una unidad de decodificación local y un estimador de error de banda baja.

50 La figura 12 ilustra una realización a modo de ejemplo de la tecnología propuesta en un sistema de decodificador basado en transformada con una disposición de control conjunto para la compresión de la excitación y la expansión de la envolvente, en la que el control conjunto se adapta utilizando la estimación del error de banda baja a partir del codificador.

La figura 13 ilustra una realización a modo de ejemplo de una disposición de control.

La figura 14 ilustra un equipo de usuario (UE) que incluye un decodificador provisto de una disposición de control.

La figura 15 es un diagrama de flujo que ilustra una realización a modo de ejemplo de la tecnología propuesta.

La figura 16 es un diagrama de flujo que ilustra una realización a modo de ejemplo de la tecnología propuesta.

5 La figura 17 es un diagrama de flujo que ilustra una realización a modo de ejemplo de la tecnología propuesta.

La figura 18 es un diagrama de flujo que ilustra una realización a modo de ejemplo de la tecnología propuesta.

La figura 19 es un diagrama de flujo que ilustra una realización a modo de ejemplo de la tecnología propuesta.

### Descripción detallada

10 En la descripción detallada siguiente se han proporcionado bloques que realizan funciones iguales o similares con las mismas designaciones de referencia.

La tecnología propuesta puede ser utilizada tanto en la BWE en el dominio del tiempo como en la BWE en el dominio de la frecuencia. A continuación, se presentarán realizaciones a modo de ejemplo para ambas.

#### BWE en el dominio del tiempo

15 En la figura 2, se muestra una realización a modo de ejemplo de la BWE de una técnica anterior, prevista principalmente para aplicaciones de conversación. Este ejemplo utiliza un algoritmo de codificación de conversación de CELP de la señal de entrada, para la banda baja. La envolvente de banda alta se representa con un filtro de LP. La síntesis de la banda alta se crea utilizando una versión modificada de la señal de excitación de banda baja extraída a partir de la síntesis de CELP.

20 Cada trama  $y$  de la señal de entrada se divide en una señal de banda de baja frecuencia  $y_L$ , y una señal de banda de alta frecuencia  $y_H$ , mediante la utilización de un banco de filtros de análisis 10. Se puede utilizar cualquier banco de filtros adecuado, pero constaría esencialmente de un filtro de paso bajo y un filtro de paso alto, por ejemplo, un banco de filtros de filtro de espejo en cuadratura (QMF – Quadrature Mirror Filter, en inglés). La señal de banda baja se introduce en un algoritmo de codificación de CELP ejecutado en un codificador de CELP 12. El análisis de LP se lleva a cabo en la señal de banda alta, en un bloque de análisis de LP 14, para obtener una representación  $A$  de la envolvente de banda alta. Los coeficientes de LP que definen  $A$  están codificados con un cuantificador de LP o un codificador de LP 16, y los índices de cuantificación  $I_{LP}$  son multiplexados en un mux (multiplexador) de secuencia de bits 18, junto con los índices del codificador de CELP  $I_{CELP}$  para ser almacenados o transmitidos a un decodificador. El decodificador, a su vez, demultiplexa los índices  $I_{LP}$  e  $I_{CELP}$  en un demux (demultiplexador) de secuencia de bits 20, y los envía al decodificador de LP 22 y al decodificador de CELP 24, respectivamente. En la decodificación de CELP la señal de excitación de CELP  $X_L$  se extrae y procesa de tal manera que el espectro de frecuencia es modulado para generar la señal de excitación de banda alta  $X_H$ .

35 Existe una variedad de esquemas de modulación para crear una excitación de banda alta  $X_H$  a partir de una señal de excitación de banda baja  $X_L$  en un procesador de excitación 26. Por ejemplo, invertir el espectro garantiza que las propiedades de la señal sean similares en la región de cruce entre la banda baja y la banda alta, pero el extremo superior de la señal de banda alta puede presentar propiedades no deseadas. Otras maneras de generar una excitación de banda alta es realizar otros tipos de modulación, que pueden o no preservar la estructura armónica de una serie de armónicos. La señal de excitación se puede tomar solo de una parte de la banda baja, o incluso adaptativamente buscando en la banda baja partes adecuadas para su utilización para formar la señal de excitación de banda alta. Este último planteamiento puede requerir, asimismo, la codificación de los parámetros, de tal manera que el decodificador pueda identificar las regiones utilizadas en la excitación de banda alta.

40 La excitación modulada  $X_H$  se filtra utilizando el filtro de LP de banda alta  $1/\hat{A}$  para formar la síntesis de banda alta  $\hat{y}_H$ . Esto se realiza en un bloque de síntesis de LP 28. La  $\hat{y}_L$  de salida del decodificador de CELP se combina con la síntesis de banda alta  $\hat{y}_H$  en el banco de filtros de síntesis 30, para formar la señal de salida  $\hat{y}$ .

45 En la figura 2 y en las siguientes figuras, las líneas hacia y desde el mux de la secuencia de bits 18 y el demux de la secuencia de bits 20, respectivamente, se han marcado a trazos para indicar que transfieren índices que representan cantidades cuantificadas en lugar de los valores reales de las cantidades cuantificadas.

50 La excitación de la banda baja puede tener propiedades que no son adecuadas para su utilización como excitación de banda alta. Por ejemplo, la señal de banda baja a menudo contiene una estructura importante de armónicos, lo que provoca aberraciones molestas cuando se transfiere a la banda alta. Una solución de la técnica anterior para controlar la estructura de la excitación es mezclar la señal de excitación de banda baja con ruido. Un decodificador de ejemplo es un sistema tal como el mostrado en la figura 3. A continuación, se decodifican los coeficientes del filtro de LP de banda alta  $\hat{A}$ , y se ejecuta el decodificador de CELP 24, mientras se extrae la señal de excitación tal como se ha descrito en la figura 2. No obstante, la excitación modulada  $X_H$  se mezcla también, tal como se ilustra

mediante los multiplicadores 32, 34 y un sumador 36, con una señal de ruido gaussiano  $n$  de un generador de ruido 38 utilizando factores de mezcla respectivos  $g_x(i)$  y  $g_n(i)$  para cada subtrama  $i$ , es decir:

$$\tilde{x}_i = g_x(i) x_{H,i} + g_n(i) n_i \quad (1)$$

5 Aquí,  $X_{H,i}$  representa las muestras  $X_H$  de la subtrama  $i$ , de tal manera que  $X_H = [X_{H,1} X_{H,2} \dots X_{H,N_{sub}}]$ , en la que  $N_{sub}$  es el número de subtramas. En este ejemplo,  $N_{sub} = 4$ . Puede resultar además beneficioso adaptar la forma temporal de la señal de ruido  $n$ , de tal manera que coincida con la forma temporal de  $X_H$ .

En este ejemplo, los factores de mezcla se determinan en un controlador de mezcla 40, y se basan en un parámetro de sonorización  $v(i)$  de cada subtrama  $i$  del códec de CELP:

$$\begin{cases} g_x(i) = \sqrt{v(i)} \\ g_n(i) = \sqrt{E_1(1-v(i))/E_2} \end{cases} \quad (2)$$

10 en la que  $E_1$  y  $E_2$  son las energías de trama de  $X_H$  y  $n$ , respectivamente, es decir:

$$\begin{cases} E_1 = \sum_{k=0}^{L-1} x_H^2(k) \\ E_2 = \sum_{k=0}^{L-1} n^2(k) \end{cases} \quad (3)$$

15 en la que la trama actual está representada con las muestras  $k = 0, 1, 2, \dots, L-1$ . El parámetro de sonorización  $v(i)$  influye en el equilibrio entre la componente de ruido  $n$  y la excitación modulada  $X_H$ , y puede, por ejemplo, estar en el intervalo  $v(i) \in [0, 1]$ . El parámetro de sonorización expresa la periodicidad de la señal (o tonalidad o armonicidad), y se calcula a partir de la energía  $E_{ACB}$  del libro de códigos algebraico y de la energía  $E_{FCB}$  del libro de códigos fijo del códec de CELP, por ejemplo, de acuerdo con:

$$v(i) = 0,5(1 - r_v(i)) \quad (4)$$

en la que

$$r_v(i) = \frac{E_v(i) - E_c(i)}{E_v(i) + E_c(i)} \quad (5)$$

20 en la que  $E_v(i)$  y  $E_c(i)$  son las energías del vector escalado del código del paso (pitch, en inglés) y el vector escalado del código algebraico para la subtrama  $i$ .

La excitación mezclada  $\tilde{x}_H$  se filtra en el bloque de síntesis de LP 28 utilizando el filtro de LP de banda alta  $1/\hat{A}$  para formar la síntesis de banda alta  $\hat{y}_H$ . La  $\hat{y}_L$  del decodificador de CELP se combina con la síntesis de banda alta  $\hat{y}_H$  en el banco de filtros de síntesis 30 para formar la señal de salida  $\hat{y}$ .

25 Una realización a modo de ejemplo de una BWE en el dominio del tiempo basada en la tecnología propuesta en esta memoria se centra en un sistema de codificador y de decodificador de audio previsto principalmente para aplicaciones de conversación. Esta realización reside en el decodificador de un sistema de codificación y de decodificación tal como se muestra en la figura 2, y dispone de un sistema de mezcla de ruido de excitación tal como el descrito en la figura 3. La aportación a los sistemas de la técnica anterior es un control adicional, tanto de la  
 30 envolvente espectral como de la mezcla de excitación, mediante el control conjunto la forma de la envolvente y del nivel de ruido de la excitación, con un parámetro de control  $f$  común (o compartido), tal como se muestra a modo de ejemplo en el decodificador 200 de la figura 4. El parámetro de control  $f$  es "común" en el sentido de que el mismo parámetro de control  $f$  se utiliza para controlar tanto la forma de la envolvente como el nivel de ruido de la excitación. En este ejemplo se utiliza un solo parámetro de control  $f \in [0, 1]$ . Debe, no obstante, observarse, que se puede  
 35 utilizar cualquier intervalo del parámetro de control, por ejemplo  $[-A, A]$ ,  $[0, A]$ ,  $[A, 0]$ , o  $[A, B]$  para cualquier  $A$  y  $B$  adecuados. No obstante, existe el beneficio de tener un intervalo unidad simple, con el propósito de controlar conjuntamente dos o más procesos.

El control de la envolvente espectral puede, por ejemplo, realizarse mediante la utilización de un post-filtro conformador  $H(z)$  (ilustrado en 42 en la figura 4) de la forma:

$$H(z) = \frac{\hat{A}(z/\gamma_1)}{\hat{A}(z/\gamma_2)} \quad (6)$$

en la que

$\hat{A}$  es un filtro predictor lineal que representa la envolvente, y

$\gamma_1, \gamma_2$  son funciones del parámetro de control  $f$ .

- 5 Este post-filtro 42 se utiliza típicamente para limpiar los valles espectrales en un decodificador de CELP, y se controla mediante un controlador conjunto del post-filtro y de la excitación 44. Un ejemplo del resaltado de la envolvente del espectro obtenida con tal post-filtro, se puede ver en la figura 5. En esta realización a modo de ejemplo, el filtro 42 se ha realizado adaptativo, modificando  $\gamma_1, \gamma_2$  mediante la utilización del parámetro de control  $f$  de acuerdo con:

$$\begin{cases} \gamma_1 = \gamma_0 + f \cdot \Delta\gamma \\ \gamma_2 = \gamma_0 - f \cdot \Delta\gamma \end{cases} \quad (7)$$

- 10 en la que  $\gamma_0, \Delta\gamma$  son constantes predeterminadas. Valores adecuados para  $\gamma_0$  pueden ser  $\gamma_0 = 0,75$  o en el rango de  $\gamma_0 \in [0,5, 0,9]$ , y valores adecuados para  $\Delta\gamma$  pueden ser  $\Delta\gamma = 0,15$  o en el rango de  $\Delta\gamma \in [0,1, 0,3]$ . Debe observarse que  $\gamma_0$  y  $\Delta\gamma$  se deben elegir de tal manera que  $\gamma_1 \in [0, 1]$  y  $\gamma_2 \in [0, 1]$ . Con esta disposición, el valor de control  $f = 1$  proporcionará la modificación más importante del post-filtro, mientras que  $f = 0$  deshabilitará el post-filtro ajustando  $\gamma_1 = \gamma_2$ , lo que conduce a  $H(z) = 1$ .

- 15 En otra variante del post-filtro 42 el estado inactivo del filtro para  $f = 0$  se modifica, para proporcionar un efecto de aplanado sobre el espectro. Esto puede resultar útil para situaciones en las que el espectro inicial tiene demasiada estructura, de tal manera que una deshabilitación del post-filtro no es suficiente para conseguir la cantidad deseada de eliminación del resaltado del valle espectral. En ese caso, la expresión en la ecuación (7) se puede modificar como:

$$\begin{cases} \gamma_1 = \gamma_0 - \gamma_{\text{exp}} + f \cdot \Delta\gamma \\ \gamma_2 = \gamma_0 + \gamma_{\text{exp}} - f \cdot \Delta\gamma \end{cases} \quad (8)$$

o bien

$$\begin{cases} \gamma_1 = \gamma_0 - \gamma_{\text{exp}} + f \cdot (\Delta\gamma + \gamma_{\text{exp}}) \\ \gamma_2 = \gamma_0 + \gamma_{\text{exp}} - f \cdot (\Delta\gamma + \gamma_{\text{exp}}) \end{cases} \quad (9)$$

- 25 en las que la ecuación (9) tiene en cuenta implícitamente el desfase del filtro de aplanado. Debe observarse que  $f = 0$  en este caso genera  $\gamma_1 < \gamma_2$ , lo que significa que el post-filtro 42 tiene un efecto de aplanado en lugar de resaltar el efecto de la forma de la envolvente.

- 30 El efecto aplanador se puede conseguir asimismo ampliando el rango del parámetro de control  $f \in [-1, 1]$  o  $f \in [-A, A]$  o  $f \in [-A, B]$  para valores adecuados de  $A$  y  $B$ . En este caso, el post-filtro 42 se puede expresar como en la ecuación (7), de tal manera que un  $f$  negativo proporciona un efecto de aplanado a la envolvente espectral, mientras que un  $f$  positivo mejora la estructura de la envolvente espectral. Puede resultar asimismo deseable utilizar diferentes potencias del post-filtro para el resaltado de la estructura espectral y el aplanado espectral, respectivamente. Uno de tales métodos sería utilizar un  $\Delta\gamma$  diferente dependiendo del signo del parámetro de control  $f$ .

$$\begin{cases} \gamma_1 = \gamma_0 + f \cdot \Delta\gamma_{\text{sharp}}, f \geq 0 \\ \gamma_2 = \gamma_0 - f \cdot \Delta\gamma_{\text{sharp}} \\ \gamma_1 = \gamma_0 + f \cdot \Delta\gamma_{\text{flat}}, f < 0 \\ \gamma_2 = \gamma_0 - f \cdot \Delta\gamma_{\text{flat}} \end{cases} \quad (10)$$

- 35 en la que  $\Delta\gamma_{\text{flat}}$  y  $\Delta\gamma_{\text{sharp}}$  son constantes predeterminadas que controlan la potencia del aplanado y la potencia de la mejora espectral, respectivamente. Valores adecuados pueden ser  $\Delta\gamma_{\text{flat}} = 0,12$  o bien en el rango de  $\Delta\gamma_{\text{flat}} \in [0,01, 0,20]$  y  $\Delta\gamma_{\text{sharp}} = 0,08$  o bien en el rango de  $\Delta\gamma_{\text{sharp}} \in [0,01, 0,20]$ .

La mezcla de la excitación se controla a su vez mediante un controlador de mezcla 41, configurado para controlar el nivel de ruido mezclando la excitación de banda alta  $x_{H,i}$  de la subtrama  $i$  con el ruido  $n_i$  de acuerdo con (1), en la que los factores de mezclado  $g_x(i)$  y  $g_n(i)$  están definidos por:

$$\begin{cases} g_x(i) = \sqrt{v(i)(1-\alpha f)} \\ g_n(i) = \sqrt{E_1(1-v(i)(1-\alpha f))/E_2} \end{cases} \quad (11)$$

en la que

$v(i)$  es un parámetro de sonorización que controla parcialmente el nivel de ruido de la excitación,

$\alpha$  es una constante de sintonización predeterminada,

5  $E_1$  es la energía de trama de las excitaciones de banda alta  $x_{H,i}$  para todas las subtramas  $i$ , y

$E_2$  es la energía de trama del ruido  $n_i$  para todas las subtramas  $i$ .

La constante de sintonización  $\alpha$  decide la modificación máxima en comparación con la ecuación (2). Un valor adecuado para  $\alpha$  puede ser  $\alpha = 0,3$  o bien en el rango de  $\alpha \in [0, 1]$ . Cuando el parámetro de control  $f$  es cercano a 1 los factores de mezclado serán equilibrados para proporcionar más ruido, mientras que un  $f$  cercano a 0 proporcionará la proporción de ruido no modificada en la mezcla.

10 Si se permiten valores negativos del parámetro de control  $f$ , una expresión alternativa para los factores de mezcla de ruido generados por el controlador de mezcla 41 es

$$\begin{cases} g_x(i) = \sqrt{v(i)(1-\max(0,\alpha f))} \\ g_n(i) = \sqrt{E_1(1-v(i)(1-\max(0,\alpha f)))/E_2} \end{cases} \quad (12)$$

en la que

15  $v(i)$  es un parámetro de sonorización que controla parcialmente el nivel de ruido de la excitación,

$\alpha$  es la constante de sintonización predeterminada,

$E_1$  es la energía de trama de las excitaciones de banda alta  $x_{H,i}$  para todas las subtramas  $i$ , y

$E_2$  es la energía de trama del ruido  $n_i$  para todas las subtramas  $i$ .

20 Aquí, la función  $\max(a,b)$  devuelve el valor máximo de  $a$  y  $b$  definido en la ecuación (14) que se muestra a continuación. En la expresión anterior esto asegura que una  $f$  negativa no influya sobre los valores de mezclado del ruido.

En una realización, el parámetro de control  $f$  puede ser adaptado mediante la utilización de parámetros ya existentes en el decodificador 200. Un ejemplo es utilizar la inclinación espectral de la señal de banda alta, dado que el post-filtro 42 puede ser dañino en combinación con una inclinación espectral importante. De este modo, el controlador conjunto del post-filtro y de la excitación 44 pueden ser configurados para adaptar el parámetro de control  $f$  a una inclinación espectral de banda alta  $t_m$  de la trama  $m$ . La inclinación espectral de banda alta se puede aproximar utilizando el segundo coeficiente  $a_{1,m}$  del filtro de LP decodificado  $\hat{A} = \{1, a_{1,m}, a_{2,m}, \dots, a_{P,m}\}$  de la trama  $m$  actual, en el que  $P$  es el orden del filtro.

25

30 Generalmente resulta beneficioso suavizar la adaptación para evitar la aparición de cambios bruscos en la envolvente espectral, por ejemplo, de acuerdo con:

$$t_m = \beta \cdot a_{1,m} + (1 - \beta) \max(0, t_{m-1}) \quad (13)$$

en la que  $t_m$  es el valor de inclinación espectral de la trama  $n$ ,  $t_{m-1}$  es el valor de la inclinación espectral de la trama previa  $m-1$  y  $\beta = 0,1$  en el rango de  $\beta = [0, 0,5]$ . La función  $\max$  se puede definir como:

$$\max(a,b) = \begin{cases} a, & a \geq b \\ b, & a < b \end{cases} \quad (14)$$

35 Aquí, la función  $\max$  asegura que el valor de la inclinación espectral utilizado de la trama anterior no es negativo. Otros ejemplos para suavizar la inclinación espectral son:

$$t_m = \beta \cdot \max(0, a_{1,m}) + (1 - \beta)t_{m-1} \quad (15)$$

y

$$t_m = \beta \cdot a_{1,m} + (1 - \beta)t_{m-1} \quad (16)$$

Puede resultar asimismo deseable considerar inclinaciones espectrales tanto negativas como positivas. En este caso, se puede utilizar el valor absoluto de la aproximación de la inclinación espectral, es decir:

$$t_m = \beta \cdot |a_{1,m}| + (1 - \beta)t_{m-1} \quad (17)$$

- 5 El valor de la inclinación espectral suavizada puede ser mapeado al parámetro de control f mediante una función lineal definida a trozos:

$$f(t_m) = \begin{cases} 0, & t_m \geq C_{\max} \\ 1 - (t_m - C_{\min}) / (C_{\max} - C_{\min}), & C_{\min} \leq t_m < C_{\max} \\ 1, & t_m < C_{\min} \end{cases} \quad (18)$$

en la que  $C_{\min}$  y  $C_{\max}$  son constantes predeterminadas. En este ejemplo, los valores de la constante son ajustados a  $C_{\max} = 0,8$  y  $C_{\min} = 0,4$ , pero los valores de otras variables pueden ser elegidos de  $C_{\max} \in [0,5, 2,0]$  y  $C_{\min} \in [0, C_{\max}]$ .

- 10 Volviendo a la figura 4, utilizando  $g_x$  y el  $g_n$  modificados se obtiene una señal de excitación nueva  $\tilde{X}_H$ . Esta señal es filtrada utilizando el filtro de LP de banda alta  $1/\hat{A}$  (en 28), para formar una primera etapa de síntesis de banda alta  $y'_H$ . Esta señal es introducida en el post-filtro adaptativo  $H(z)$  (en 42) para obtener la síntesis de banda alta  $y_H$ . La salida  $\hat{y}_L$  del decodificador de CELP 24 se combina con la síntesis de banda alta  $\tilde{y}_H$  en el banco de filtros de síntesis 30 para formar la señal de salida  $\hat{y}$ .
- 15 Existen otras alternativas a la adaptación basada en la inclinación descrita anteriormente. Por ejemplo, se puede utilizar una medida de la planicidad espectral de la banda alta. La planicidad espectral  $\varphi$  se mide en alguna representación del espectro de banda alta. Se puede obtener, por ejemplo, a partir de los coeficientes A de LPC de banda alta, utilizando la expresión bien conocida:

$$\varphi = \frac{e^{\frac{1}{N} \sum_{i=0}^{N-1} \log(X_i)}}{\frac{1}{N} \sum_{i=0}^{N-1} X_i} \quad (19)$$

- 20 en la que

$$X_i = \frac{1}{|DFT(A, M)|^2}, i = 0, 1, 2, \dots, N - 1 \quad (20)$$

en la que  $DFT(A, M)$  indica la transformada de Fourier discontinua de la longitud M de los coeficientes A de LPC. La expresión  $|\cdot|$  indica la magnitud de los valores de la transformada compleja (el punto representa una expresión matemática) y, debido a la simetría de la transformada, solo se consideran los primeros  $N = M/2$  valores. Esta transformada se implementa preferiblemente mediante una FFT (Transformada rápida de Fourier – Fast Fourier Transform, en inglés), y la M sería la mayor potencia más cercana de 2 a la longitud del filtro P+1, es decir,  $M = 2^{\lceil \log_2(P+1) \rceil}$ .

- Si  $P+1 > M$ , el filtro de entrada A se rellena con ceros antes de que se ejecute la FFT. La planicidad espectral  $\varphi$  se puede calcular asimismo utilizando los coeficientes de LPC cuantificados  $\hat{A}$ . Si esto se lleva a cabo, la medida de la planicidad espectral se puede calcular en el decodificador sin señalización adicional. En este caso, el sistema se puede describir mediante la figura 4, siempre que A sea sustituida por  $\hat{A}$  en la ecuación (20).

- Puede resultar deseable determinar la medida de la planicidad espectral del lado del decodificador, para reducir la complejidad global, cuando se consideran tanto el codificador como el decodificador. En tal realización, el codificador incluye un estimador de planicidad espectral configurado para determinar, para su transmisión a un decodificador, una medida de la planicidad espectral de la señal de banda alta. En la figura 6, se representa un codificador que utiliza un estimador de planicidad espectral 46 basado en los coeficientes de LPC. En este caso, la medida de la planicidad se debe señalar en la secuencia de bits. La señalización puede consistir en una decisión binaria  $\hat{\varphi} \in \{0, 1\}$ , de si la planicidad espectral se considera alta o baja dependiendo de un valor de umbral  $\varphi_{thr}$ .

$$\begin{cases} \hat{\varphi} = 0, & \varphi \geq \varphi_{thr} \\ \hat{\varphi} = 1, & \varphi < \varphi_{thr} \end{cases} \quad (21)$$



El parámetro de control  $f$  correspondiente se puede obtener, por ejemplo, utilizando la decisión binaria  $\hat{\Phi}$ , es decir,  $f = 1 - 2\hat{\Phi}$ .

Con las definiciones anteriores, el parámetro de control  $f$  será de 1 para valores de planicidad por encima del umbral, y de -1 para valores de planicidad por debajo del umbral. Para limitar la influencia de la conmutación brusca entre estos valores, el parámetro de control se puede además suavizar utilizando, por ejemplo, un factor de olvido  $\beta$  de una manera similar a la del filtrado de inclinación:

$$f'_m = \beta \cdot f_m + (1 - \beta) \cdot f'_{m-1} \quad (22)$$

En la figura 7, se muestra un decodificador 200 correspondiente al codificador de la figura 6. Es similar al decodificador de la figura 4. No obstante, en la figura 7, el controlador conjunto del post-filtro y de la excitación 44 determinan el parámetro de control  $f$  sobre la base de la decisión binaria  $\hat{\Phi}$  recibida en lugar del filtro predictor lineal  $\hat{A}$  que representa a la envolvente. De manera general, el parámetro de control  $f$  está adaptado a una medida de la planicidad espectral ( $\phi$ ) de la banda alta.

Debe observarse que son posibles otras etapas del procesamiento antes del filtro de síntesis  $1/\hat{A}$ , o antes del post-filtro  $H(z)$ . Una de tales etapas del procesamiento podría ser un procedimiento de conformado temporal dirigido a reconstruir la estructura temporal de la señal de banda alta original. Tal conformado temporal puede ser codificado utilizando la cuantificación de un vector de forma de ganancia que representa los factores de corrección de la ganancia a nivel de subtrama. Parte de la conformación temporal será heredada asimismo de la señal de excitación de banda baja, que es parcialmente utilizada como base para la señal de excitación de banda alta.

El post-filtro y la mezcla de la excitación pueden afectar también a la energía de las señales. Mantener estable la energía resulta deseable, y existen muchos métodos para gestionar este aspecto. Una posible solución es medir la energía antes y después de la modificación, y la restauración de la energía al valor anterior a la mezcla de la excitación y del post-filtrado. La medición de la energía puede estar asimismo limitada a una cierta banda o a las regiones de mayor energía del espectro, permitiendo la pérdida de energía en los valles del espectro. En esta realización a modo de ejemplo, se puede utilizar la compensación de la energía como parte integral de las funciones de mezcla y de post-filtro.

BWE en el dominio de la frecuencia

A menudo, se utilizan codificadores de audio basados en la transformada de la frecuencia, para señales de audio generales, tales como música o conversación con ruidos de fondo y reverberación. A tasas de bits bajas, generalmente muestran un mal comportamiento. Una solución común de la técnica anterior es disminuir el ancho de banda para obtener una calidad aceptable para una banda más estrecha, y aplicar BWE para las frecuencias mayores. Una visión global de tal sistema se muestra en la figura 8.

El audio de entrada se divide en primer lugar en segmentos de tiempo o tramas, como etapa de preparación para la transformación de la frecuencia. Cada trama  $Y$  se transforma en el dominio de la frecuencia para formar un espectro en el dominio de la frecuencia  $Y$ . Esto se puede efectuar mediante la utilización de cualquier transformada adecuada, tal como la transformada discontinua del coseno modificada (MDCT – Modified Discrete Cosine Transform, en inglés), la transformada discontinua del coseno (DCT – Discrete Cosine Transform, en inglés) o la transformada de Fourier discontinua (DFT – Discrete Fourier discontinua (DFT – Discrete Fourier Transform, en inglés). El espectro de frecuencia se divide en vectores de fila más cortos, indicados como  $Y(b)$ . Estas funciones se realizan mediante un transformador de frecuencia 50. Cada vector representa ahora los coeficientes de una banda de frecuencia  $b$  del número total de bandas  $N_b$ . Desde una perspectiva perceptual, resulta beneficioso dividir el espectro utilizando una estructura de banda no uniforme de acuerdo con la resolución de frecuencia del sistema auditivo humano. Esto, en general, significa que se utilizan anchos de banda estrechos para frecuencias bajas, mientras que se utilizan anchos de banda más grandes para frecuencias altas.

A continuación, se calcula la norma de cada banda en un analizador de envolvente 52, para formar una secuencia de valores de ganancia  $E(b)$  que forman la envolvente espectral. Estos valores son a continuación cuantificados utilizando un codificador de envolvente 54 para formar la envolvente cuantificada  $\hat{E}(b)$ . La cuantificación de la envolvente se puede realizar utilizando cualquier técnica de cuantificación, por ejemplo, cuantificación escalar diferencial, o cualquier esquema de cuantificación de vectores. Los coeficientes de la envolvente cuantificada  $\hat{E}(b)$  se utilizan para normalizar los vectores de banda  $Y(b)$  en un normalizador de la envolvente 56 para formar vectores de forma normalizados  $X(b)$  correspondientes:

$$X(b) = \frac{1}{\hat{E}(b)} Y(b) \quad (23)$$

La secuencia de vectores de forma normalizados  $X(b)$  constituye la estructura fina del espectro. La importancia perceptual de la estructura fina espectral varía con la frecuencia, pero puede depender también de otras

propiedades de la señal, tal como la señal de envolvente espectral. Los codificadores de transformación a menudo emplean un modelo auditivo para determinar las partes importantes de la estructura fina, y asignan los recursos disponibles a las partes más importantes. La envolvente espectral se utiliza a menudo como entrada para este modelo auditivo, y la salida es típicamente una asignación de bits para cada una de las bandas correspondientes a los coeficientes de la envolvente. En esta memoria, un algoritmo de asignación de bits en un asignador de bits 58 utiliza la envolvente cuantificada  $\hat{E}(b)$  en combinación con un modelo auditivo interno para asignar un número de bits  $R(b)$  que, a su vez, son utilizados por un codificador de estructura fina 60. Cuando el codificador de transformación es operado a tasas de bits bajas, a algunas de las bandas se les asignarán cero bits, y los vectores de forma correspondientes no serán cuantificados. Los índices  $I_E$  e  $I_X$  de la cuantificación de la envolvente y los vectores codificados de la estructura fina, respectivamente, son multiplexados en un mux (multiplexador) de secuencia de bits 62 para ser almacenados o transmitidos a un decodificador.

El decodificador desmultiplexa los índices del canal de comunicación o los medios almacenados en un demux (desmultiplexador) de secuencia de bits 70 y transmite los índices  $I_X$  a un decodificador de estructura fina 72 e  $I_E$  a un decodificador de envolvente 74. La envolvente cuantificada  $\hat{E}(b)$  se obtiene e introduce en el algoritmo de asignación de bits en un asignador de bits 76 en el decodificador, que genera la asignación de bits  $R(b)$ . Utilizando  $R(b)$ , se encuentra la banda con el valor distinto de cero más alto en la asignación de bits. Esta banda se indica  $b_{max}$ .

El decodificador de estructura fina 72 utiliza los índices de estructura fina  $I_X$  y la asignación de bits  $R(b)$  para producir los vectores de estructura fina cuantificados  $\hat{X}_L(b)$  que están definidos para  $b = 1, 2, \dots, b_{max}$ .

En esta realización a modo de ejemplo, la frecuencia de cruce es adaptativa, dependiendo de la asignación de bits, y empieza en la banda  $b_{max} + 1$ , dada la restricción de que  $b_{max} + 1 \leq N_b$ .

Pueden existir bandas  $b < b_{max}$  que tienen cero bits asignados. En particular, para tasas de bits bajas, es habitual que aparezcan tales bandas de cero bits y, debido a las variaciones en el espectro, las posiciones de las bandas de cero bits habitualmente varían de trama en trama. Tales variaciones provocan efectos de modulación en la síntesis. Típicamente, las bandas de cero bits son manejadas con técnicas de llenado espectral, en las que las señales son inyectadas en las bandas de cero bits. La señal de llenado puede ser una señal de ruido pseudoaleatorio o una versión modificada de las bandas codificadas. La técnica de llenado no es una parte esencial de esta tecnología, y se asume que un llenado espectral adecuado forma parte del decodificador de estructura fina 72. Tras la realización del llenado espectral, la estructura fina de banda baja  $\hat{X}_L(b)$  es introducida en un conformador de envolvente 78 de baja frecuencia, que restaura el espectro de banda baja  $\hat{Y}_L(b)$  sintetizado de acuerdo con:

$$\hat{Y}_L(b) = \hat{X}_L(b) \cdot \hat{E}(b), \quad b = 1, 2, \dots, b_{max} \quad (24)$$

La estructura fina de banda baja  $\hat{X}_L(b)$  se introduce asimismo en un modificador o procesador de estructura fina 80, que identifica la longitud de la estructura de banda baja a partir del parámetro  $b_{max}$ , y crea una señal de excitación de banda alta  $\hat{X}_H(b)$  definida para  $b_{max} + 1, b_{max} + 2, \dots, N_b$ . Existen muchas técnicas para crear una excitación de banda alta a partir de la excitación de banda baja. En esta realización a modo de ejemplo, la mitad superior de la señal de excitación de banda baja se abate y duplica para llenar la excitación de banda alta. Así, se asume que  $\hat{X}_{LH}$  representa la mitad superior de la señal de excitación de banda baja y que la función  $rev(\cdot)$  invierte los elementos de un vector. A continuación, la secuencia  $[rev(\hat{X}_{LH}) \hat{X}_{LH} rev(\hat{X}_{LH}) \hat{X}_{LH} \dots]$  se repite tantas veces como sea necesario para llenar el espectro de excitación de banda alta  $\hat{X}_H(b), b_{max} + 1, b_{max} + 2, \dots, N_b$ . La señal de excitación de banda alta se introduce a continuación en un formador de envolvente de alta frecuencia 82 para formar el espectro de banda alta sintetizado  $\hat{Y}_H(b)$  de acuerdo con:

$$\hat{Y}_H(b) = \hat{X}_H(b) \cdot \hat{E}(b), \quad b = b_{max} + 1, b_{max} + 2, \dots, N_b \quad (25)$$

El espectro de banda baja sintetizado  $\hat{Y}_L(b)$  y el espectro de banda alta sintetizado  $\hat{Y}_H(b)$  se combinan en un combinador de espectro 84 para formar el espectro de síntesis  $\hat{Y}(b)$  o  $\hat{Y}$  con el índice de banda omitido. El espectro de síntesis se introduce en el transformador de frecuencia inverso 86 para formar la señal de salida  $\hat{y}$ . En este proceso, se llevan a cabo también las operaciones de creación de ventana y de adición de superposición que están ligadas a la transformación de la frecuencia.

Tal como era el caso de la BWE en el dominio del tiempo, la excitación de la banda baja puede tener propiedades que no son adecuadas para su utilización como excitación de banda alta. En particular, puede resultar deseable aplanar una parte de la estructura fina en la excitación de banda baja. En la figura 9, se muestra un decodificador de tal sistema a modo de ejemplo. Este sistema de la técnica anterior asume un codificador tal como el representado en la figura 8. La adición al esquema descrito es que existe un compresor H (en 88), que opera en la señal de

excitación de banda alta  $\hat{X}_H(b)$  para producir la señal de excitación de banda alta  $\hat{X}_H(b)$  comprimida. Una función de compresor a modo de ejemplo es:

$$H = \left( \frac{\max(|\hat{X}_H|)}{|\hat{X}_H|} \right)^\eta \quad (26)$$

5 que significa que H es un vector con la misma longitud que  $\hat{X}_H$ . Aquí se ha omitido el índice de banda b, y los vectores representan a todos los elementos para las bandas definidas, es decir:

$$\hat{X}_H = [\hat{X}_H(b_{\max} + 1) \hat{X}_H(b_{\max} + 2) \dots \hat{X}_H(N_b)] \quad (27)$$

El factor de compresión  $\eta$  es menor de 1, y un valor adecuado puede ser  $\eta = 0,5$  o bien en el rango de  $\eta \in [0,01, 0,99]$ , en el que valores cercanos a 0 no proporcionan ningún efecto, y valores cercanos a 1 proporcionan una compresión máxima. La síntesis de banda alta comprimida se obtiene mediante la multiplicación por elementos de H

10 y  $\hat{X}_H$ . Se puede expresar como matriz de multiplicación:

$$\tilde{X}_H = H \text{diag}(\hat{X}_H) \quad (28)$$

en la que  $\text{diag}(\hat{X}_H)$  produce una matriz cuadrada con  $\hat{X}_H$  en la diagonal. La excitación de banda alta comprimida  $\tilde{X}_H(b)$  se introduce en el formador de envolvente de alta frecuencia 82 para formar el espectro de banda alta  $\tilde{Y}_H(b)$  de acuerdo con:

$$15 \quad \hat{Y}_H(b) = \tilde{X}_H(b) \cdot \hat{E}(b), \quad b = b_{\max} + 1, b_{\max} + 2, \dots, N_b \quad (29)$$

Tal como se ilustra en la figura 9. el espectro de banda baja  $\hat{Y}_L(b)$  y el espectro de banda alta  $\hat{Y}_H(b)$  se combinan en el combinador de espectro 84 para formar el espectro de síntesis  $\hat{Y}$ , que es introducido en el transformador de frecuencia inverso 86 para formar la señal de salida  $\hat{y}$ .

20 Una realización a modo de ejemplo de una BWE en el dominio de la frecuencia basada en la tecnología propuesta, se centra en un sistema codificador y decodificador de audio previsto principalmente para señales de audio generales. La nueva tecnología reside principalmente en el decodificador de un sistema de codificación y de decodificación tal como el representado en la figura 8 con un sistema de compresión de excitación tal como se ilustra en la figura 9. En la figura 10, se ilustra una realización a modo de ejemplo de tal decodificador 200.

25 Adicionalmente a la técnica anterior se proporciona un control combinado de una compresión de la excitación de banda alta, que se controla conjuntamente con un amplificador de envolvente espectral 90, como se muestra en la figura 10. Como en el dominio del tiempo, se utiliza un parámetro de control  $f \in [0, 1]$  para dirigir tanto al compresor 88 como al amplificador 90. Esto se lleva a cabo mediante un controlador conjunto del amplificador y del compresor 92.

30 La potencia del compresor de excitación de banda alta 88 se adapta utilizando el parámetro de control f de acuerdo con:

$$H = \left( \frac{\max(|\hat{X}_H|)}{|\hat{X}_H|} \right)^{\eta + \Delta\eta f} \quad (30)$$

en la que  $\Delta\eta$  proporciona el exponente máximo del factor de compresión  $\eta + \Delta\eta$  cuando  $f = 1$ . Si  $\eta = 0,5$ , entonces un valor adecuado para  $\Delta\eta$  puede ser  $\Delta\eta = 0,3$ , o bien en el rango de  $\Delta\eta \in [0,01, 1-\eta]$ . Debe observarse que  $\eta + \Delta\eta \leq 1$ . La excitación de banda alta comprimida se obtiene mediante la multiplicación elemento a elemento de H y  $\hat{X}_H$ , es decir:

$$35 \quad \tilde{X}_H = H \text{diag}(\hat{X}_H) \quad (31)$$

El amplificador 90 utilizado en la envolvente de banda alta tiene una estructura similar al compresor de la excitación de banda alta:

$$G = \left( \frac{\max(\hat{E}(b))}{\hat{E}(b)} \right)^{-(\phi + \Delta\phi f)}, \quad b = b_{\max} + 1, b_{\max} + 2, \dots, N_b \quad (32)$$

5 Aquí, el valor absoluto  $|\cdot|$  se puede omitir, dado que los coeficientes de la envolvente  $\hat{E}(b) \geq 0$ . Para  $f = 0$ , el amplificador tendrá un efecto mínimo con el coeficiente de expansión  $\varphi$ . Un valor adecuado para  $\varphi$  puede ser  $\varphi = 0$ , dado que esto haría que la envolvente no resultase afectada para  $f = 0$ . Si un efecto de expansión pequeño es siempre deseable, se pueden elegir valores adecuados, por ejemplo, del rango  $\varphi \in [0, 0,5]$ . La expansión máxima se obtiene para  $f = 1$ , que proporciona el exponente del factor de expansión  $-(\varphi + \Delta\varphi)$ . El valor para  $\Delta\varphi$  se puede ajustar a  $\Delta\varphi = 1$ , pero el valor adecuado dependería en gran medida de la estructura de banda, y se puede elegir de un amplio rango, por ejemplo,  $\Delta\varphi \in [0,5, 10]$ . La envolvente ampliada  $E(b)$  se obtiene mediante multiplicación elemento a elemento de la envolvente con la función de expansión  $G$ , es decir:

$$\tilde{E}_H = G \text{diag}(\hat{E}_H) \quad (33)$$

10 en la que  $\hat{E}_H$  representa los elementos de la envolvente de la banda alta  $\hat{E}_H = [\hat{E}(b_{\max}+1) \hat{E}(b_{\max}+2) \dots \hat{E}(N_b)]$ . La envolvente expandida se aplica a la estructura fina de banda alta comprimida para formar el espectro de banda alta  $\hat{Y}_H(b)$  de acuerdo con:

$$\hat{Y}_H(b) = \tilde{X}_H(b) \cdot \tilde{E}(b), \quad b = b_{\max} + 1, b_{\max} + 2, \dots, N_b \quad (34)$$

15 El espectro de banda baja sintetizado  $\hat{Y}_L(b)$  y el espectro de banda alta sintetizado  $\hat{Y}_H(b)$  se combinan en el combinador de espectro 84 para formar el espectro de síntesis  $\hat{Y}$ , que se introduce en el transformador de frecuencia inversa 86 para formar la señal de salida  $\hat{y}$ .

20 El parámetro de control conjunto  $f$  se puede obtener a partir de los parámetros ya disponibles en el decodificador 200, o puede estar basado en un análisis proporcionado en el codificador y transmitido al decodificador. En esta memoria, como para el caso de BWE en el dominio del tiempo, nos basamos en una estimación de la inclinación espectral de banda alta. Tal estimación se puede obtener a partir de los parámetros de la envolvente, midiendo el cociente  $q_m$  de las sumas de los coeficientes de la envolvente en cada mitad de la señal de banda alta, es decir:

$$q_m = \frac{\sum_{b=b_{\max}+1}^{b_{\max}} \hat{E}(b)}{\sum_{b=t_{\max}+1}^{N_b} \hat{E}(b)} \quad (35)$$

en la que

$$b_{\max} = \lfloor (N_b - b_{\max}) / 2 \rfloor + b_{\max} + 1 \quad (36)$$

25 El suavizado de la inclinación espectral  $t_m$  para la trama  $m$  se puede efectuar de la misma manera que la realización en el dominio del tiempo, por ejemplo, utilizando:

$$t_m = \beta \cdot q_m + (1 - \beta)t_{m-1} \quad (37)$$

El mapeo de la inclinación espectral al parámetro de control  $f$  se puede efectuar asimismo utilizando la misma función lineal definida a trozos que en la realización en el dominio del tiempo, es decir:

$$f(t_m) = \begin{cases} 0, & t_m \geq C_{\max} \\ 1 - (t_m - C_{\min}) / (C_{\max} - C_{\min}), & C_{\min} \leq t_m < C_{\max} \\ 1, & t_m < C_{\min} \end{cases} \quad (38)$$

30 No obstante, dado que la definición de la inclinación espectral es diferente, las constantes  $C_{\max}$  y  $C_{\min}$  de la función de mapeo serán diferentes. Estas dependerán por ejemplo de la estructura de banda.

35 En una alternativa de la realización en el dominio de la frecuencia descrita anteriormente, el control conjunto de la envolvente y de la excitación está adaptado a la señal de error de banda baja, que se estima en el codificador, que es similar al codificador en el sistema representado en la figura 8, pero además tiene una decodificación local y una unidad de medición del error. Un ejemplo de tal sistema se muestra en la figura 11, en la que la decodificación local y una unidad de medición del error incluyen un decodificador local 96, un extractor de espectro de baja frecuencia 98, un sumador 100 y un codificador de error de baja frecuencia 102. En esta realización, una síntesis de banda baja local se obtiene utilizando la envolvente cuantificada  $\hat{E}(b)$  y una estructura fina de banda baja decodificada  $\hat{X}_L(b)$ , que es extraída del codificador de la estructura fina. Es asimismo posible ejecutar el decodificador de estructura fina

completa para extraer  $\hat{X}_L(b)$  de los índices  $l_x$ , pero se puede, en general, extraer una síntesis local del codificador con menos complejidad de cálculo. Un espectro de banda baja sintetizado localmente  $\hat{Y}_L(b)$  se genera mediante el conformado de la estructura de banda baja decodificada con la envolvente cuantificada:

$$\hat{Y}_L(b) = \hat{X}_L(b) \cdot \hat{E}(b), \quad b = 1, 2, \dots, b_{\max} \quad (39)$$

- 5 El espectro de banda baja de la señal de entrada  $Y_L(b)$  se extrae del espectro completo encontrando la última banda cuantificada utilizando la asignación de bit  $R(b)$ . Una señal de error de banda baja se forma como la relación de registro de la energía de la señal de entrada y la distancia euclidiana entre el espectro de banda baja sintetizado del espectro de banda baja de entrada, es decir, una medida  $D_L$  de la relación de señal a ruido (SNR – Signal to Noise Ratio, en inglés) en la síntesis de banda baja definida como:

$$D_L = 10 \log_{10} \left( \frac{Y_L Y_L^T}{(Y_L - \hat{Y}_L)(Y_L - \hat{Y}_L)^T} \right) \quad (40)$$

10 La SNR de banda baja es cuantificada y los índices de cuantificación  $l_{ERR}$  son multiplexados entre sí con los índices de la envolvente  $l_E$  y los índices de la estructura fina  $l_x$  para ser almacenados o transmitidos a un decodificador. La codificación de SNR baja se puede llevar a cabo, por ejemplo, utilizando un cuantificador escalar uniforme.

- 15 El decodificador 200 es similar al decodificador representado en la figura 9, pero además tiene un control combinado de la compresión de excitación de banda alta que es controlada a la vez que un amplificador de envolvente espectral, tal como se muestra en la figura 10. Como en las realizaciones del dominio del tiempo, se utiliza un parámetro de control  $f \in [0, 1]$  para dirigir tanto al compresor como al amplificador.

Utilizando el parámetro de control  $f$ , la potencia del compresor de extensión de banda alta se adapta de acuerdo con:

$$H = \left( \frac{\max(|\hat{X}_H|)}{|\hat{X}_H|} \right)^{\eta + \Delta\eta f} \quad (41)$$

- 20 en la que  $\Delta\eta$  proporciona el factor de compresión máximo  $\eta + \Delta\eta$  cuando  $f = 1$ . Si  $\eta = 0,5$ , un valor adecuado para  $\Delta\eta$  puede ser  $\Delta\eta = 0,3$ , o bien en el rango de  $\Delta\eta \in [0,01, 1-\eta]$ . Debe observarse que  $\eta + \Delta\eta \leq 1$ . La excitación de banda alta comprimida se obtiene mediante la multiplicación elemento a elemento de  $H$  y  $\hat{X}_H$  de acuerdo con:

$$\tilde{X}_H = H \text{diag}(\hat{X}_H) \quad (42)$$

- 25 El amplificador utilizado en la envolvente de banda alta tiene una estructura similar a la del compresor de la envolvente de banda alta:

$$G = \left( \frac{\max(\hat{E}(b))}{\hat{E}(b)} \right)^{-(\Phi + \Delta\Phi f)}, \quad b = b_{\max} + 1, b_{\max} + 2, \dots, N_b \quad (43)$$

- 30 Aquí, el valor absoluto  $|\cdot|$  se puede omitir, dado que los coeficientes de la envolvente  $\hat{E}(b) \geq 0$ . Para  $f = 0$ , el amplificador tendrá un mínimo efecto con el coeficiente de expansión  $\Phi$ . Un valor adecuado para  $\Phi$  puede ser  $\Phi = 0$ , dado que esto proporcionaría una envolvente no afectada para  $f = 0$ . Si un efecto de expansión pequeño resulta siempre deseable, se pueden elegir valores adecuados, por ejemplo, del rango de  $\Phi \in [0, 0,5]$ . La expansión máxima se obtiene para  $f = 1$ , lo que proporciona el exponente del factor de expansión  $-(\Phi + \Delta\Phi)$ . El valor para  $\Delta\Phi$  se puede ajustar a  $\Delta\Phi = 1$ , pero el valor adecuado dependería en gran medida de la estructura de la banda, y se puede elegir de un amplio rango, por ejemplo,  $\Phi\Delta \in [0,5, 10]$ . La envolvente ampliada  $\tilde{E}(b)$  se obtiene mediante la multiplicación elemento a elemento de la envolvente con la función de expansión  $G$ , es decir:

$$35 \quad \tilde{E}_H = G \text{diag}(\hat{E}_H) \quad (44)$$

en la que  $\hat{E}_H$  representa los elementos de la envolvente de banda alta  $\hat{E}_H = [\hat{E}(b_{\max}+1) \hat{E}(b_{\max}+2) \dots \hat{E}(N_b)]$ . La envolvente ampliada se aplica a la estructura fina de banda alta comprimida  $\tilde{X}_H(b)$  para formar el espectro de banda alta  $\hat{Y}_H(b)$  de acuerdo con:

$$\hat{Y}_H(b) = \tilde{X}_H(b) \cdot \tilde{E}(b), \quad b = b_{\max} + 1, b_{\max} + 2, \dots, N_b \quad (45)$$

El espectro de banda baja sintetizado  $\hat{Y}_L(b)$  y el espectro de banda alta sintetizado  $\hat{Y}_H(b)$  se combinan en el combinador de espectro para formar el espectro de síntesis  $\hat{Y}$ , que se introduce en el transformador de frecuencia inversa para formar la señal de salida  $\hat{y}$ .

- 5 En esta realización, el parámetro de control  $f$  está basado en la SNR de banda baja del análisis del codificador. En primer lugar, se obtiene una SNR de banda baja reconstruida  $\hat{D}_L$  del índice de error de banda baja  $I_{ERR}$ . La SNR de banda baja reconstruida se mapea a un parámetro de control  $f$  utilizando una función lineal definida a trozos:

$$f = \begin{cases} 0, & \hat{D}_L < D_{\min} \\ (\hat{D}_L - D_{\min}) / (D_{\max} - D_{\min}), & D_{\min} \leq \hat{D}_L \leq D_{\max} \\ 1, & \hat{D}_L > D_{\max} \end{cases} \quad (46)$$

- 10 en la que las constantes  $D_{\min}$  y  $D_{\max}$  dependen de los valores de distorsión de la banda baja típicos para este sistema. Un valor adecuado para  $D_{\min}$  puede ser  $D_{\min} = 10$ , o cualquier valor en el rango  $D_{\min} \in [5, 20]$ , mientras que valores adecuados para  $D_{\max}$  pueden ser  $D_{\max} = 20$  o en el rango  $D_{\max} \in [10, 50]$ . Esta relación proporcionará una modificación más importante para los valores altos de la SNR, correspondientes a la distorsión baja en la banda baja. Puede resultar asimismo deseable tener la relación opuesta, de tal manera que se utilizaría una modificación importante para las SNR bajas (valores de distorsión altos). Tal relación se puede obtener invirtiendo la relación descrita anteriormente, es decir:

$$f = \begin{cases} 1, & \hat{D}_L < D_{\min} \\ (D_{\max} - \hat{D}_L) / (D_{\max} - D_{\min}), & D_{\min} \leq \hat{D}_L \leq D_{\max} \\ 0, & \hat{D}_L > D_{\max} \end{cases} \quad (47)$$

- 20 Debe observarse que la función de compresor y de amplificador puede cambiar la energía global de los vectores. Preferiblemente, la energía debe mantenerse estable y existen muchos métodos para manejar esto. Una solución posible es medir la energía antes y después de la modificación, y restaurar la energía al valor antes de la compresión o la expansión. La medición de la energía se puede limitar asimismo a una cierta banda o a las regiones superiores de la energía del espectro, permitiendo una pérdida de energía en los valles del espectro. En esta realización a modo de ejemplo se asume que se utiliza una cierta compensación de energía y que forma una parte integral de las funciones del compresor y del amplificador.

- 25 Las etapas, funciones, procedimientos y/o bloques descritos en esta memoria se pueden implementar en hardware utilizando cualquier tecnología convencional, tal como la tecnología de circuitos separados o de circuitos integrados, que incluye tanto circuitos electrónicos de propósito general como circuitos específicos para una aplicación.

- 30 De manera alternativa, al menos algunas de las etapas, funciones, procedimientos y/o bloques descritos en esta memoria, pueden ser implementados en software para su ejecución mediante el equipo de procesamiento adecuado. Este equipo puede incluir, por ejemplo, uno o varios micro procesadores, uno o varios procesadores de señal digital (DSP – Digital Signal Processor, en inglés), uno o varios circuitos integrados específicos para una aplicación (ASIC – Application Specific Integrated Circuits, en inglés), hardware de video acelerado o uno o varios dispositivos lógicos programables, tales como las matrices de puertas programables en campo (FPGA – Field Programmable Gate Array, en inglés). También son factibles las combinaciones de tales elementos de procesamiento.

- 35 Debe comprenderse asimismo que es posible reutilizar las capacidades generales de procesamiento ya existentes en el codificador / decodificador. Esto se puede realizar, por ejemplo, mediante reprogramación del software existente, o mediante la adición de nuevos componentes de software.

- 40 La figura 13 ilustra una realización a modo de ejemplo de una disposición de control. Esta disposición se basa en un procesador 210, por ejemplo, un microprocesador, que ejecuta software 220 para controlar a la vez la forma de la envolvente y el nivel de ruido de la excitación con un parámetro de control común. El software está almacenado en la memoria 230. El procesador 210 se comunica con la memoria sobre un bus del sistema. Las señales de entrada son recibidas por un controlador de entrada / salida (I/O – Input / Output, en inglés) 240 que controla un bus de I/O, al cual están conectados el procesador 210 y la memoria 230. Las señales de salida obtenidas del software 220 son emitidas desde la memoria 230 mediante el controlador de I/O 240 sobre el bus I/O. Las señales de entrada y salida entre paréntesis corresponden a la BWE en el dominio del tiempo, y las señales de entrada y salida sin paréntesis corresponden a la BWE en el dominio de la frecuencia.

Una realización basada en una medida  $\phi$  de la planicidad espectral puede estar configurada estructuralmente como en la figura 13, con un procesador, memoria, bus de sistema, bus de I/O y controlador de I/O.

5 La tecnología descrita anteriormente está prevista para utilizar en un codificador / decodificador de audio, que se puede utilizar en un dispositivo móvil (por ejemplo, teléfono móvil, ordenador portátil) o en un dispositivo estacionario, tal como un ordenador personal. En esta memoria el término equipo de usuario (UE) se utilizará como nombre genérico para tales dispositivos. La figura 14 ilustra un UE que incluye un decodificador provisto de una disposición de control. Una señal de audio recibida por una unidad de radio 300 se convierte a banda base, se decodifica en el canal y se transmite a un decodificador de audio 200. El decodificador de audio está provisto de una disposición de control 310 que opera en el dominio del tiempo o de la frecuencia, tal como se ha descrito anteriormente. Las muestras de audio de ancho de banda extendido son transmitidas a una unidad de conversión y amplificación D/A 320, que transmite la señal final de audio a un altavoz 330.

La figura 15 es un diagrama de flujo que ilustra la tecnología propuesta. La etapa S1 controla conjuntamente la forma de la envolvente y el nivel de ruido con un parámetro de control  $f$  común.

15 La figura 16 es un diagrama de flujo que ilustra una realización a modo de ejemplo de la tecnología propuesta. En esta realización, la etapa S1 incluye una etapa S1A que controla la forma de la envolvente mediante la utilización de un post-filtro de formato  $H(z)$ , por ejemplo, que tiene la forma definida por las ecuaciones (6). Las constantes predeterminadas  $\gamma_1, \gamma_2$  se pueden determinar, por ejemplo, de acuerdo con una de las ecuaciones (7) – (10).

20 La figura 17 es un diagrama de flujo que ilustra una realización de la tecnología propuesta. En esta realización, la etapa S1 incluye una etapa S1B que controla el nivel de ruido de la excitación mediante la mezcla de la excitación de banda alta  $N_{H,i}$  de una subtrama  $i$  con ruido  $\eta_i$  de acuerdo con la ecuación (1), en la que los factores de mezcla  $g_x(i)$  y  $g_n(i)$  están definidos, por ejemplo, por la ecuación (11) o (12), dependiendo de la elección de las constantes predeterminadas  $\gamma_1, \gamma_2$ .

25 La figura 18 es un diagrama de flujo que ilustra una realización de la tecnología propuesta. En esta realización, la etapa S1 incluye una etapa S1C que adapta el parámetro de control  $f$  a una inclinación espectral de la banda alta  $t_m$  de la trama  $m$ , por ejemplo, de acuerdo con la ecuación (18). En una realización la inclinación espectral de banda alta  $t_m$  se puede aproximar utilizando el segundo coeficiente  $a_{1,m}$  del filtro predictor lineal decodificado  $\hat{A}_m = \{1, a_{1,m}, a_{2,m}, \dots, a_{P,m}\}$  de la trama  $m$ , en el que  $P$  es el orden del filtro. En general resulta asimismo beneficioso suavizar la inclinación espectral de banda alta  $t_m$ , por ejemplo, de acuerdo con una de las ecuaciones (13), (15) – (17). Una realización basada en una medida  $\phi$  de la planicidad espectral puede llevar a cabo la etapa S1C utilizando el planteamiento descrito con referencia a las ecuaciones (19) – (22).

30 La figura 19 es un diagrama de flujo que ilustra una realización de la tecnología propuesta. Esta realización combina las etapas S1A, S1B, S1C. Típicamente el parámetro de control  $f$  se determina en primer lugar. A continuación, se utiliza para realizar las etapas S1A y S1B. Asimismo, son posibles otras combinaciones, incluidas S1A + S1C o S1B + S1C.

35 Los expertos en la materia comprenderán que se pueden realizar varias modificaciones y cambios a la tecnología propuesta sin separarse del alcance de la misma, que está definida por las reivindicaciones adjuntas.

**Abreviaturas**

ASIC	Circuito integrado específico para una aplicación	Application Specific Integrated Circuit, en inglés
BWE	Extensión del ancho de banda	Extension Bandwidth, en inglés
CELP	Predictor lineal excitado mediante código	Code Excited Linear Predictor, en inglés
DCT	Transformada discontinua del coseno	Discrete Cosine Transform, en inglés
DFT	Transformada discontinua de Fourier	Discrete Fourier Transform, en inglés
DSP	Procesador de señal digital	Digital Signal Processor, en inglés
FFT	Transformada rápida de Fourier	Fast-Fourier Transform, en inglés
FPGA	Matrices de puertas programables en campo	Field Programmable Gate Array, en inglés
HF	Alta frecuencia	High Frequency, en inglés
LF	Baja frecuencia	Low Frequency, en inglés
LP	Predictor lineal	Linear Predictor, en inglés

LPC	Codificación lineal predictiva	Linear Predictive Coding, en inglés
MDCT	Transformada discontinua del coseno modificada	Modified Discrete Cosine Transform, en inglés
QMF	Filtro de espejo en cuadratura	Quadrature Mirror Filter, en inglés
SBR	Replicación de la banda espectral	Spectral Band Replication, en inglés
SNR	Relación de señal a ruido	Signal-to-Noise Ratio, en inglés
TCX	Residuo codificado de la transformada	Transform Coded Residual, en inglés
UE	Equipo de usuario	User Equipment, en inglés

**Referencias**

[1] "AMR-WB+: A new audio coding standard for 3rd generation mobile audio services", J. Mäkinen, B. Bessette, S. Bruhn, P. Ojala, R. Salami, A. Taleb, ICASSP 2005

[2] "Enhanced aacPlus encoder Spectral Band Replication (SBR) part", 3GPP TS 26.404 V10.0.0 (2011-03), sections 5.6.1 - 5.6.3, pp. 22-25.



**REIVINDICACIONES**

1. Método de generación de una extensión de la banda alta de una señal de audio a partir de una envolvente y de una excitación, en el que el método incluye la etapa (S1) de control simultáneo de la forma de la envolvente y del nivel de ruido de excitación con un parámetro de control común  $f$ , siendo la citada forma de envolvente controlada (S1A) mediante la utilización de un post-filtro formante  $H(z)$  de la forma:

$$H(z) = \frac{\hat{A}(z / \gamma_1)}{\hat{A}(z / \gamma_2)}$$

en la que

$\hat{A}$  es un predictor de filtro lineal que representa la envolvente, y

$\gamma_1, \gamma_2$  son funciones del parámetro de control  $f$ .

2. El método de la reivindicación 1, en el que

$$\begin{cases} \gamma_1 = \gamma_0 + f \cdot \Delta\gamma \\ \gamma_2 = \gamma_0 - f \cdot \Delta\gamma \end{cases}$$

en la que  $\gamma_0, \Delta\gamma$  son constantes predeterminadas.

3. El método de la reivindicación 1 o 2, que incluye la etapa de controlar (S1B) el nivel de ruido de la excitación mezclando una excitación de banda alta  $X_{H,i}$  de una subtrama  $i$  con ruido  $n_i$  de acuerdo con:

$$\tilde{x}_i = g_x(i) x_{H,i} + g_n(i) n_i$$

en la que los factores de mezclado  $g_x(i)$  y  $g_n(i)$  están definidos por:

$$\begin{cases} g_x(i) = \sqrt{v(i)(1-\alpha f)} \\ g_n(i) = \sqrt{E_1(1-v(i)(1-\alpha f))/E_2} \end{cases}$$

en la que

$v(i)$  es un parámetro de sonorización que controla parcialmente el nivel de ruido de la excitación,

- $\alpha$  es una constante de sintonización predeterminada

$E_1$  es la energía de trama de las excitaciones de banda alta  $X_{N,i}$  para todas las subtramas  $i$ , y

$E_2$  es la energía de trama del ruido  $n_i$  para todas las subtramas  $i$ .

4. El método de la reivindicación 1, en el que

$$\begin{cases} \gamma_1 = \gamma_0 + f \cdot \Delta\gamma_{sharp} \\ \gamma_2 = \gamma_0 - f \cdot \Delta\gamma_{sharp} \end{cases}, f \geq 0$$

$$\begin{cases} \gamma_1 = \gamma_0 + f \cdot \Delta\gamma_{flat} \\ \gamma_2 = \gamma_0 - f \cdot \Delta\gamma_{flat} \end{cases}, f < 0$$

- en la que  $\gamma_0, \Delta\gamma_{flat}$  y  $\Delta\gamma_{sharp}$  son constantes predeterminadas.

5. El método de la reivindicación 4, que incluye la etapa de controlar (S1B) el nivel de ruido de la excitación mezclando una excitación de banda alta  $X_{H,i}$  de una subtrama  $i$  con ruido  $n_i$  de acuerdo con:

$$\tilde{x}_i = g_x(i) x_{H,i} + g_n(i) n_i$$

en la que los factores de mezclado  $g_x(i)$  y  $g_n(i)$  se definen mediante:

$$\begin{cases} g_x(i) = \sqrt{v(i)(1 - \max(0, \alpha f))} \\ g_s(i) = \sqrt{E_1(1 - v(i)(1 - \max(0, \alpha f)))}/E_2 \end{cases}$$

en la que

$v(i)$  es un parámetro de sonorización que controla parcialmente el nivel de ruido de la excitación,

$\alpha$  es una constante de sintonización predeterminada

5  $E_1$  es la energía de trama de las excitaciones de banda alta  $X_{N,i}$  para todas las subtramas  $i$ , y

$E_2$  es la energía de trama del ruido  $n_i$  para todas las subtramas  $i$ .

6. El método de cualquiera de las reivindicaciones anteriores, que incluye la etapa de adaptar (S1C) el parámetro de control  $f$  a una inclinación espectral de la banda alta  $t_m$  de la trama  $m$ , y en la que el parámetro de control  $f$  depende de la inclinación espectral de la banda alta  $t_m$  de acuerdo con:

$$f(t_m) = \begin{cases} 0, & t_m \geq C_{\max} \\ 1 - (t_m - C_{\min}) / (C_{\max} - C_{\min}), & C_{\min} \leq t_m < C_{\max} \\ 1, & t_m < C_{\min} \end{cases}$$

en la que  $C_{\min}$  y  $C_{\max}$  son constantes predeterminadas.

7. El método de la reivindicación 6, en el que la inclinación espectral de banda alta  $t_m$  se aproxima mediante la utilización del segundo coeficiente  $a_{1,m}$  del filtro del predictor lineal decodificado  $\hat{A}_m = \{1, a_{1,m}, a_{2,m}, \dots, a_{P,m}\}$  de la trama  $m$ , en el que  $P$  es el orden del filtro, y en el que

$$t_m = \beta \cdot \max(0, a_{1,m}) + (1 - \beta)t_{m-1}$$

en el que

$t_m$  es el valor de la inclinación espectral de la trama  $m$ ,

$t_{m-1}$  es el valor de la inclinación espectral de la trama  $m-1$  anterior, y

$\beta$  es una constante en el rango de  $\beta = [0, 0.5]$ .

8. Decodificador de audio (200) configurado para generar una extensión de banda alta de una señal de audio a partir de una envolvente y de una excitación, que incluye una disposición de control (41, 42, 44; 88, 90, 92; 310) configurada para controlar a la vez la forma de la envolvente y el nivel de ruido de la excitación con un parámetro de control  $f$  común, incluyendo la citada disposición de control (41, 42, 44) un controlador conjunto del post-filtro y de la excitación (44) configurado para controlar la forma de la envolvente mediante la utilización de un post-filtro formante (42) ( $H(z)$ ) de la forma:

$$H(z) = \frac{\hat{A}(z/\gamma_1)}{\hat{A}(z/\gamma_2)}$$

en la que

$\hat{A}$  es un filtro predictor lineal que representa la envolvente, y

$\gamma_1, \gamma_2$  son funciones del parámetro de control  $f$ .

9. El decodificador de la reivindicación 8, en el que

$$\begin{cases} \gamma_1 = \gamma_0 + f \cdot \Delta\gamma \\ \gamma_2 = \gamma_0 - f \cdot \Delta\gamma \end{cases}$$

en la que  $\gamma_0, \Delta\gamma$  son constantes predeterminadas.

10. El decodificador de las reivindicaciones 8 o 9, que incluye un controlador de mezcla (41) configurado para controlar el nivel de ruido de la excitación mezclando una excitación de banda alta  $X_{H,i}$  de una subtrama  $i$  con ruido  $n_i$  de acuerdo con:

$$\tilde{x}_i = g_x(i) x_{H,i} + g_n(i) n_i$$

en la que los factores de mezclado  $g_x(i)$  y  $g_n(i)$  están definidos por:

$$\begin{cases} g_x(i) = \sqrt{v(i)(1-\alpha f)} \\ g_n(i) = \sqrt{E_1(1-v(i)(1-\alpha f))/E_2} \end{cases}$$

en la que

5  $v(i)$  es un parámetro de sonorización que controla parcialmente el nivel de ruido de la excitación,

$\alpha$  es una constante de sintonización predeterminada

$E_1$  es la energía de trama de las excitaciones de banda alta  $X_{N,i}$  para todas las subtramas  $i$ , y

$E_2$  es la energía de trama del ruido  $n_i$  para todas las subtramas  $i$ .

11. El decodificador de la reivindicación 8, en el que

$$\begin{cases} \gamma_1 = \gamma_0 + f \cdot \Delta\gamma_{sharp} \\ \gamma_2 = \gamma_0 - f \cdot \Delta\gamma_{sharp} \end{cases}, f \geq 0$$

$$\begin{cases} \gamma_1 = \gamma_0 + f \cdot \Delta\gamma_{flat} \\ \gamma_2 = \gamma_0 - f \cdot \Delta\gamma_{flat} \end{cases}, f < 0$$

10

en la que  $\gamma_0$ ,  $\Delta\gamma_{flat}$  y  $\Delta\gamma_{sharp}$  son constantes predeterminadas.

12. El decodificador de la reivindicación 11, que incluye un controlador de mezcla (41) configurado para controlar el nivel de ruido de la excitación mezclando una excitación de banda alta  $X_{H,i}$  de una subtrama  $i$  con el ruido  $n_i$  de acuerdo con:

$$\tilde{x}_i = g_x(i) x_{H,i} + g_n(i) n_i$$

15

en la que los factores de mezcla  $g_x(i)$  y  $g_n(i)$  están definidos por:

$$\begin{cases} g_x(i) = \sqrt{v(i)(1-\max(0,\alpha f))} \\ g_n(i) = \sqrt{E_1(1-v(i)(1-\max(0,\alpha f)))/E_2} \end{cases}$$

en la que

$v(i)$  es un parámetro de sonorización que controla parcialmente el nivel de ruido de la excitación,

20  $\alpha$  es una constante de sintonización predeterminada

$E_1$  es la energía de trama de las excitaciones de banda alta  $X_{N,i}$  para todas las subtramas  $i$ , y

$E_2$  es la energía de trama del ruido  $n_i$  para todas las subtramas  $i$ .

13. El decodificador de cualquiera de las reivindicaciones anteriores 8 – 12, en el que el controlador conjunto del post-filtro y de la excitación (44) están configurados para adaptar el parámetro de control  $f$  a una inclinación espectral de banda alta  $t_m$  de la trama  $m$ , y en el que el parámetro de control  $f$  depende de la inclinación espectral de la banda alta  $t_m$  de acuerdo con:

25

$$f(t_m) = \begin{cases} 0, & t_m \geq C_{max} \\ 1 - (t_m - C_{min}) / (C_{max} - C_{min}), & C_{min} \leq t_m < C_{max} \\ 1, & t_m < C_{min} \end{cases}$$

en el que  $C_{min}$  y  $C_{max}$  son constantes predeterminadas.

14. El decodificador de la reivindicación 13, en el que el post-filtro y el controlador conjunto del post-filtro y de la excitación (44) están configurados para aproximar la inclinación espectral de la banda alta  $t_m$  utilizando el segundo coeficiente  $a_{1,m}$  del filtro del predictor lineal decodificado  $\hat{A}_m = \{1, a_{1,m}, a_{2,m}, \dots, a_{P,m}\}$  de la trama  $m$ , en la que  $P$  es el orden del filtro, y en el que

$$5 \quad t_m = \beta \cdot \max(0, a_{1,m}) + (1 - \beta)t_{m-1}$$

en la que

$t_m$  es el valor de la inclinación espectral de la trama  $m$ ,

$t_{m-1}$  es el valor de la inclinación espectral de la trama  $m-1$  anterior, y

$\beta$  es una constante en el rango de  $\beta = [0, 0,5]$ .

10 15. Equipo de usuario (UE) que incluye un decodificador de audio de acuerdo con cualquiera de las reivindicaciones anteriores 8 – 14.

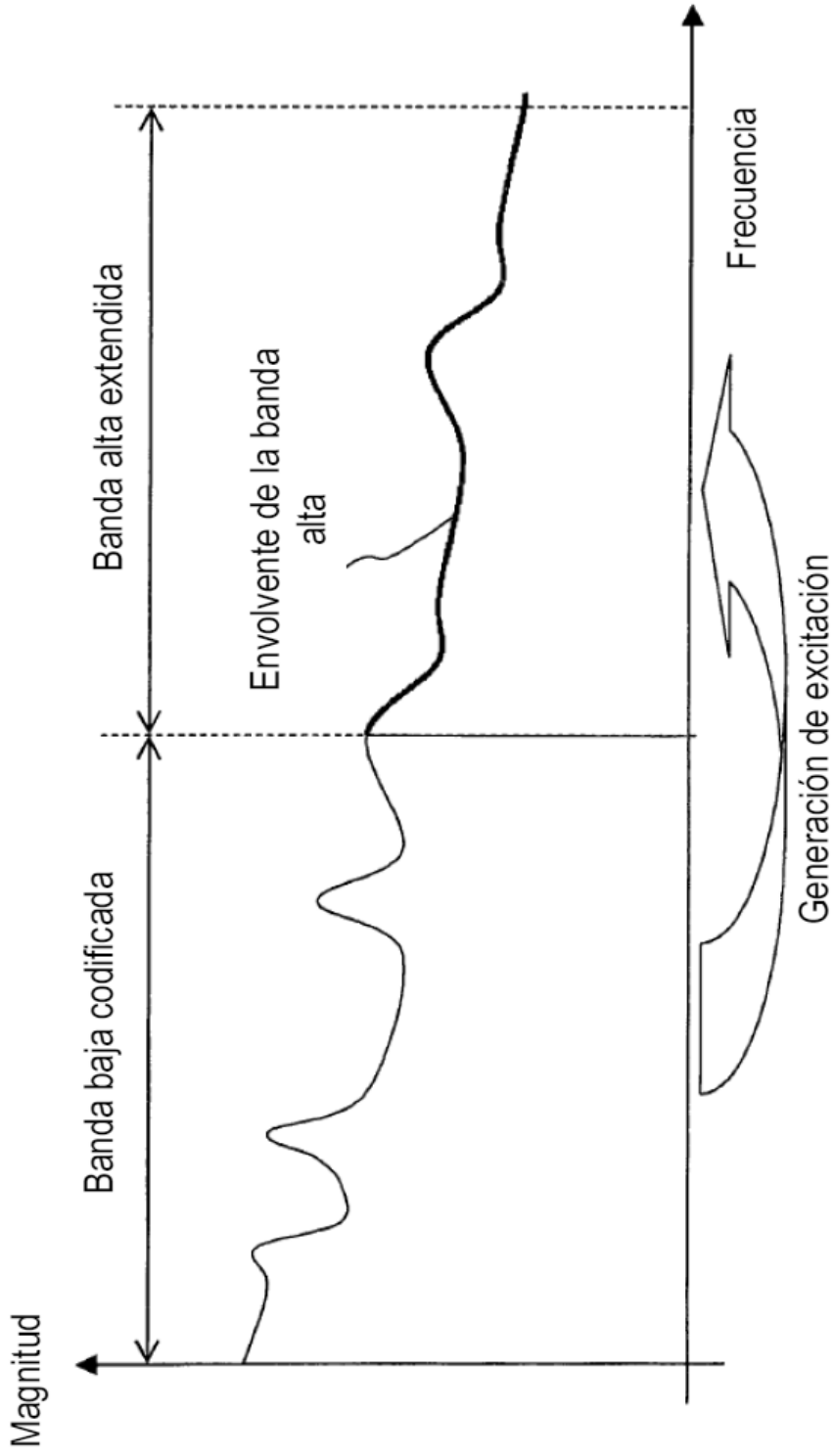
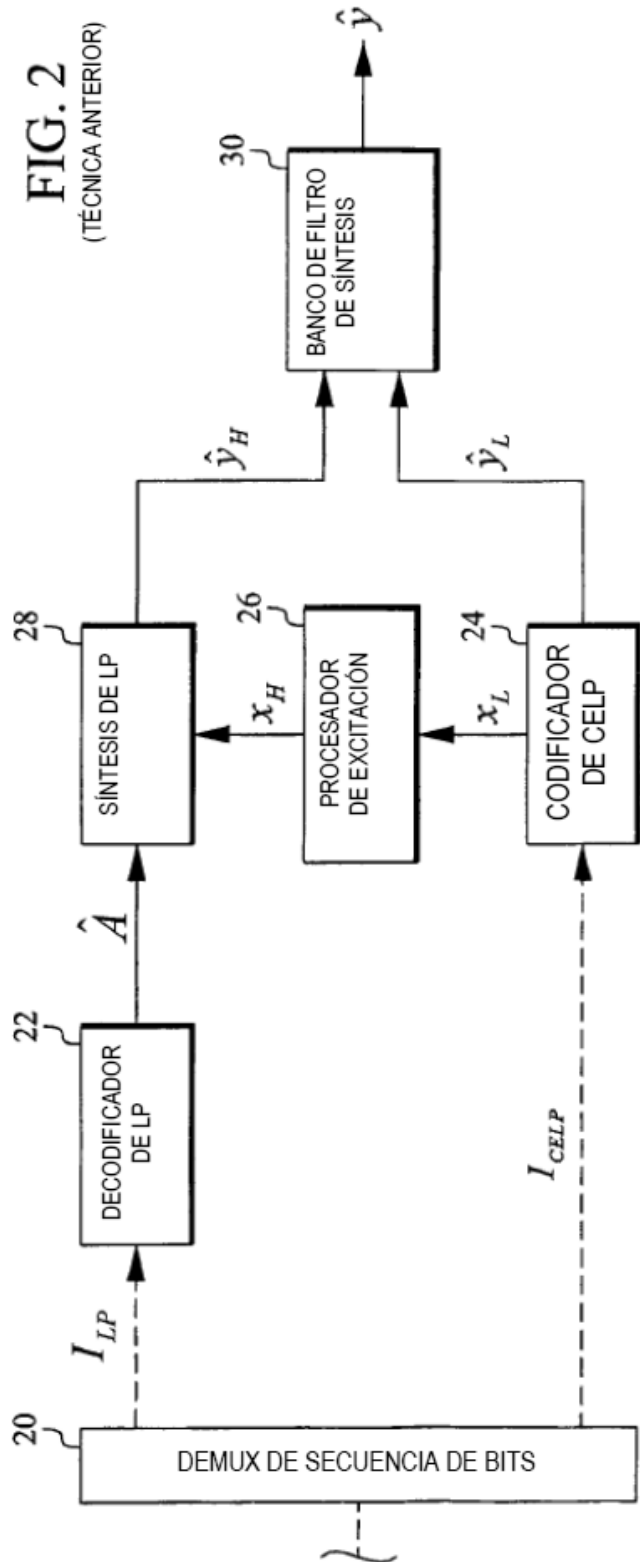
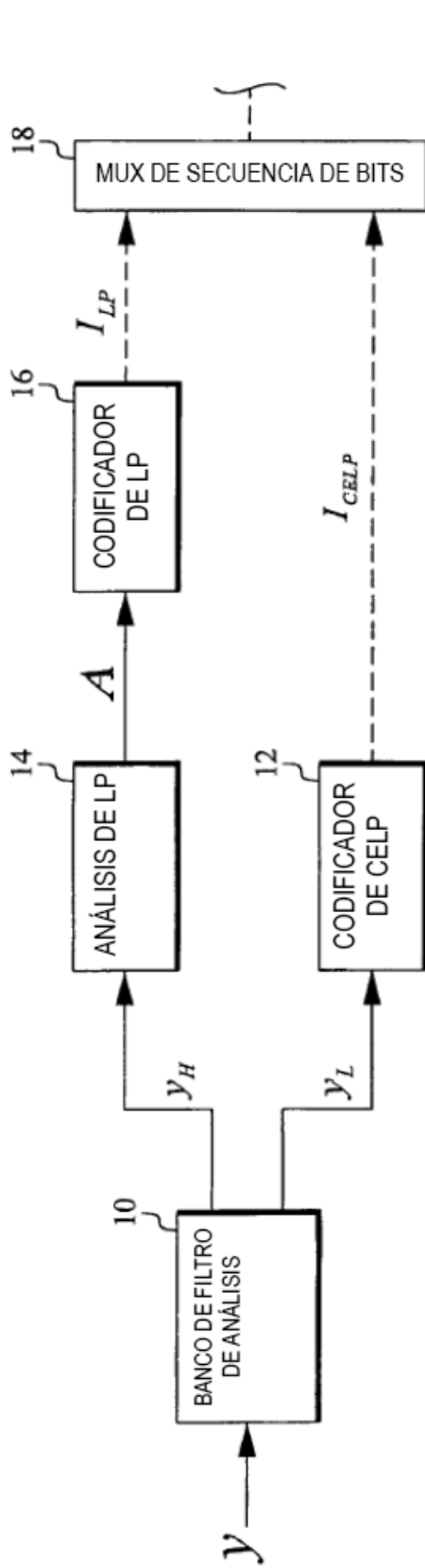


FIG. 1



**FIG. 2**  
(TÉCNICA ANTERIOR)

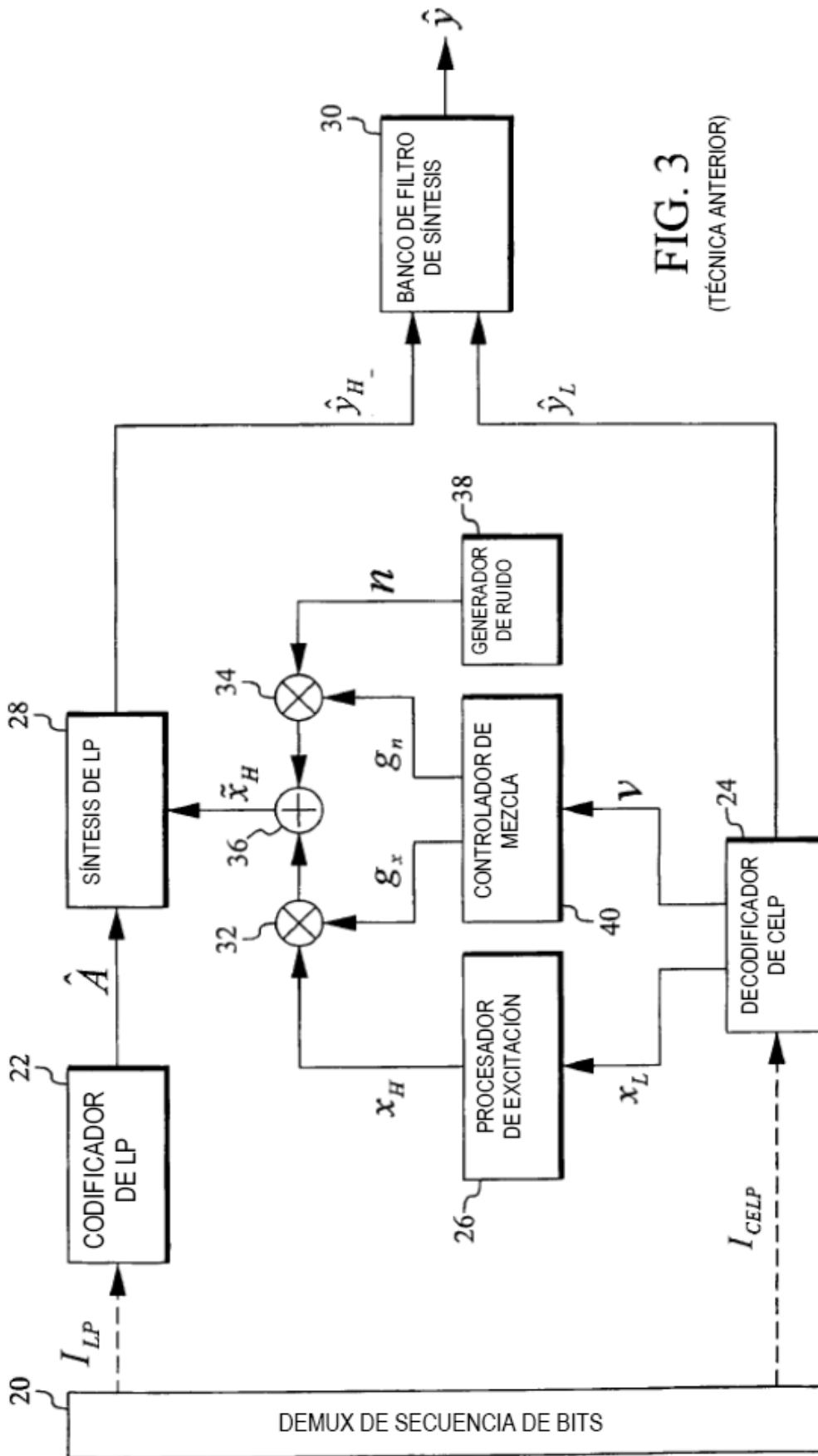


FIG. 3

(TÉCNICA ANTERIOR)

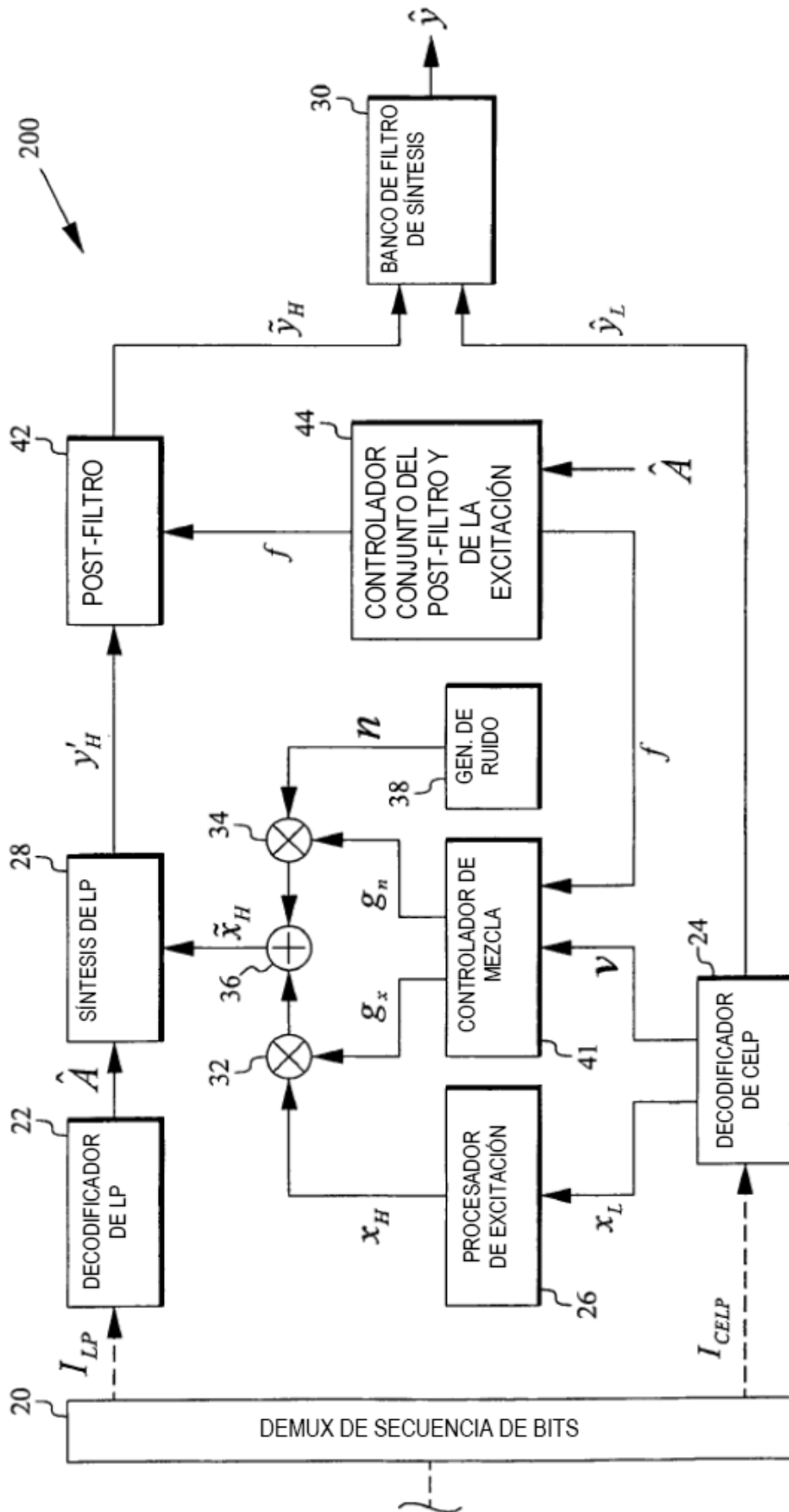


FIG. 4



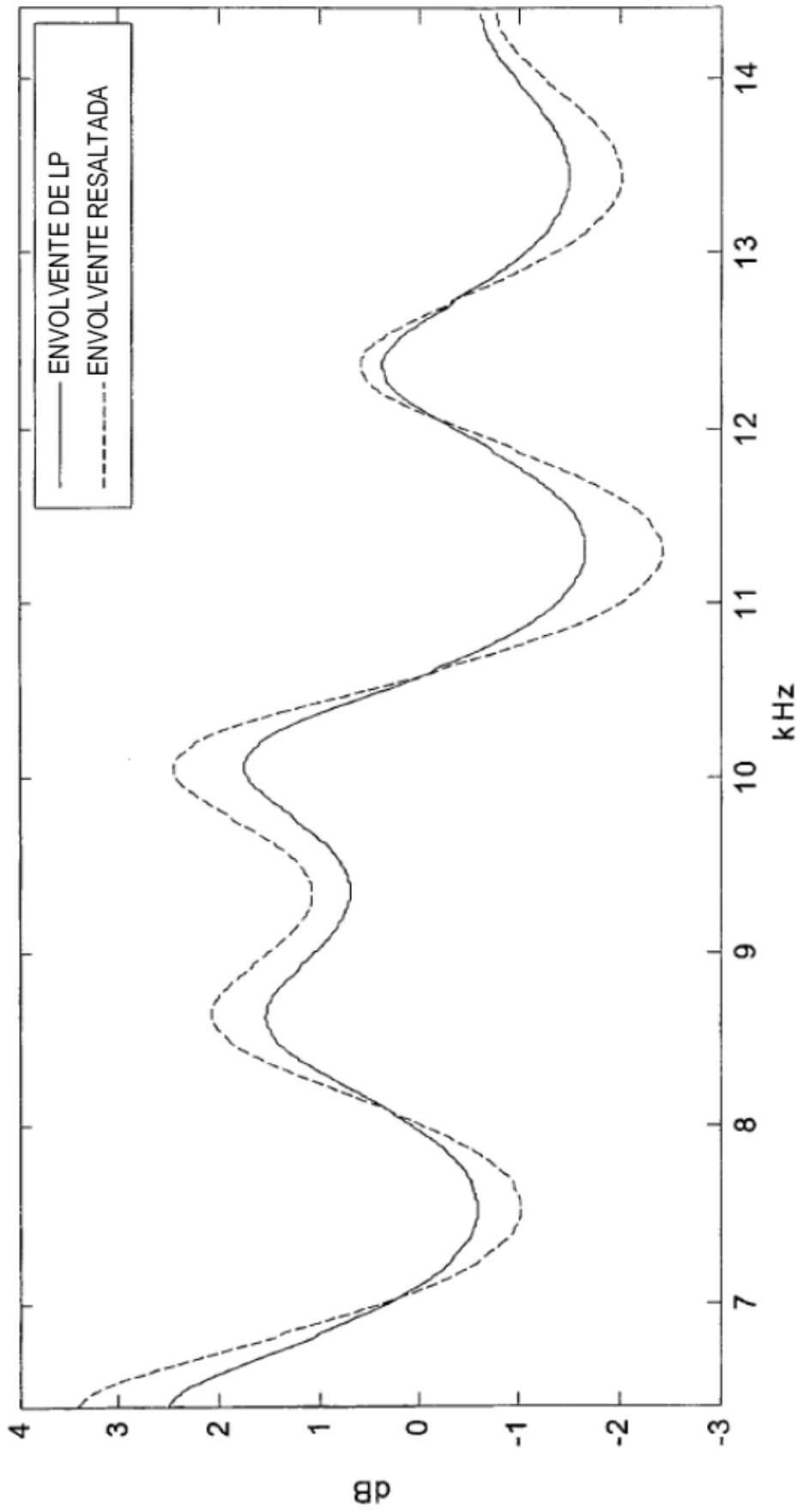


FIG. 5

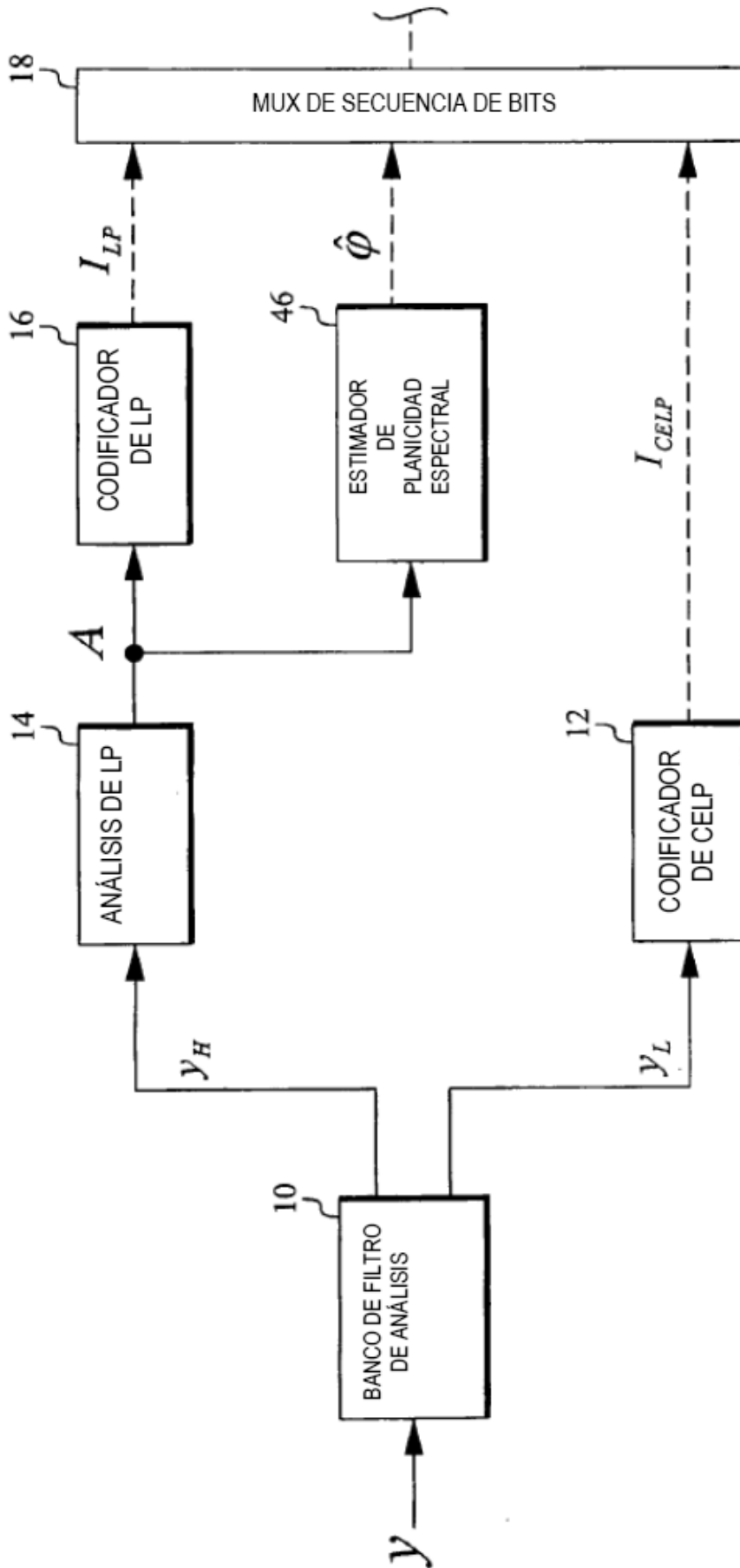


FIG. 6

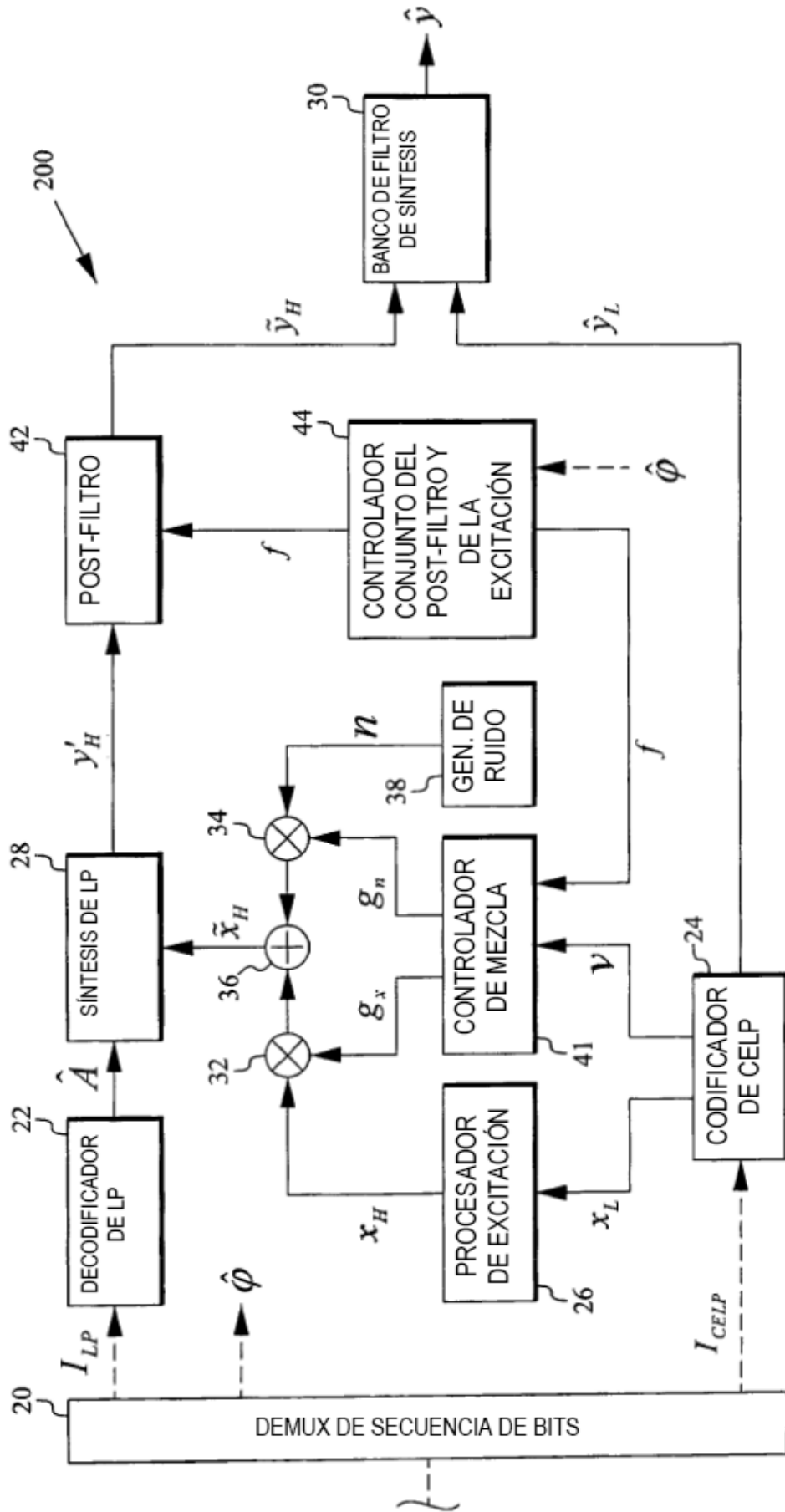


FIG. 7

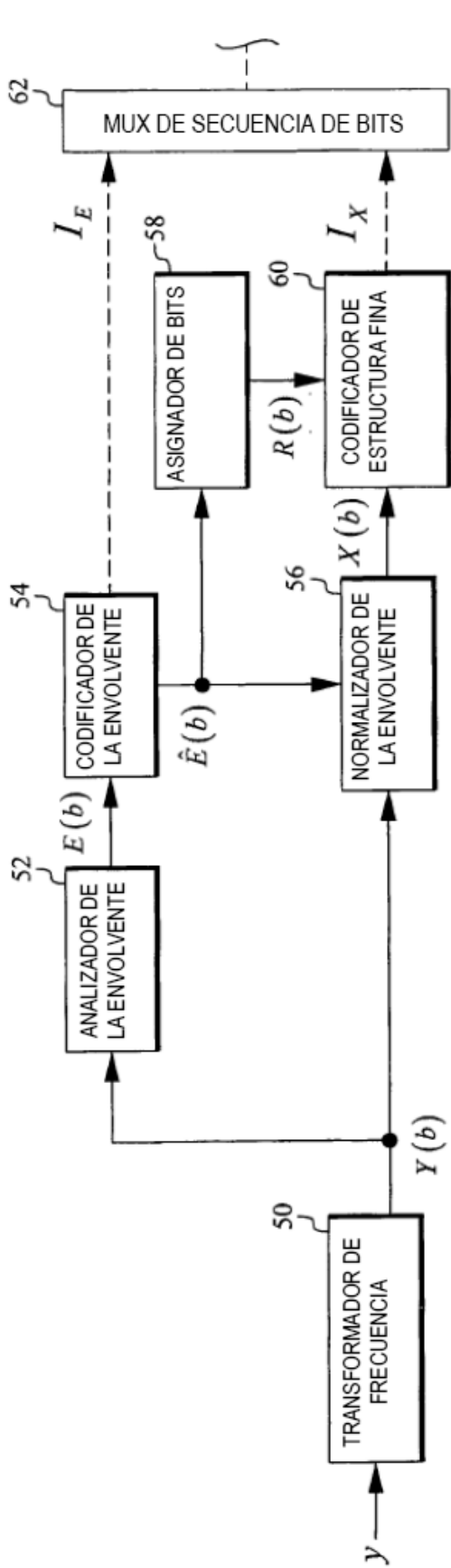


FIG. 8  
(TÉCNICA ANTERIOR)

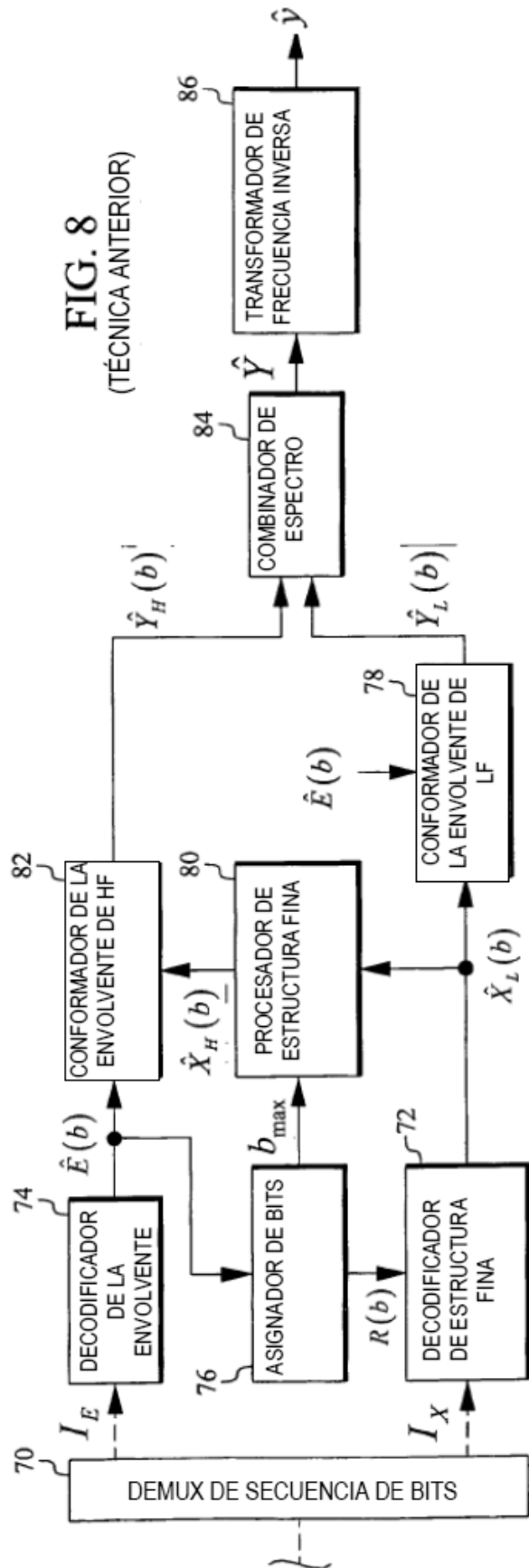
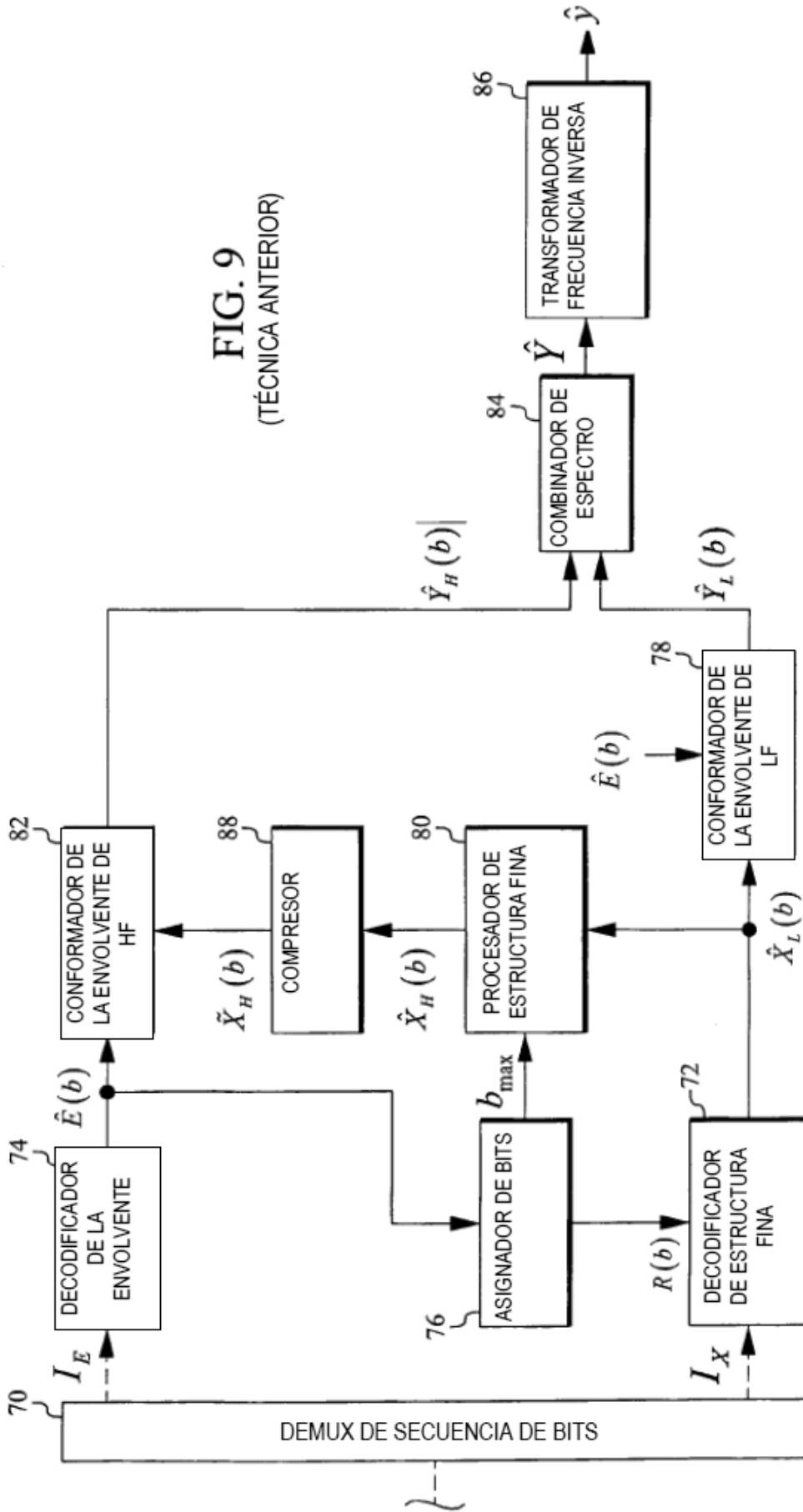
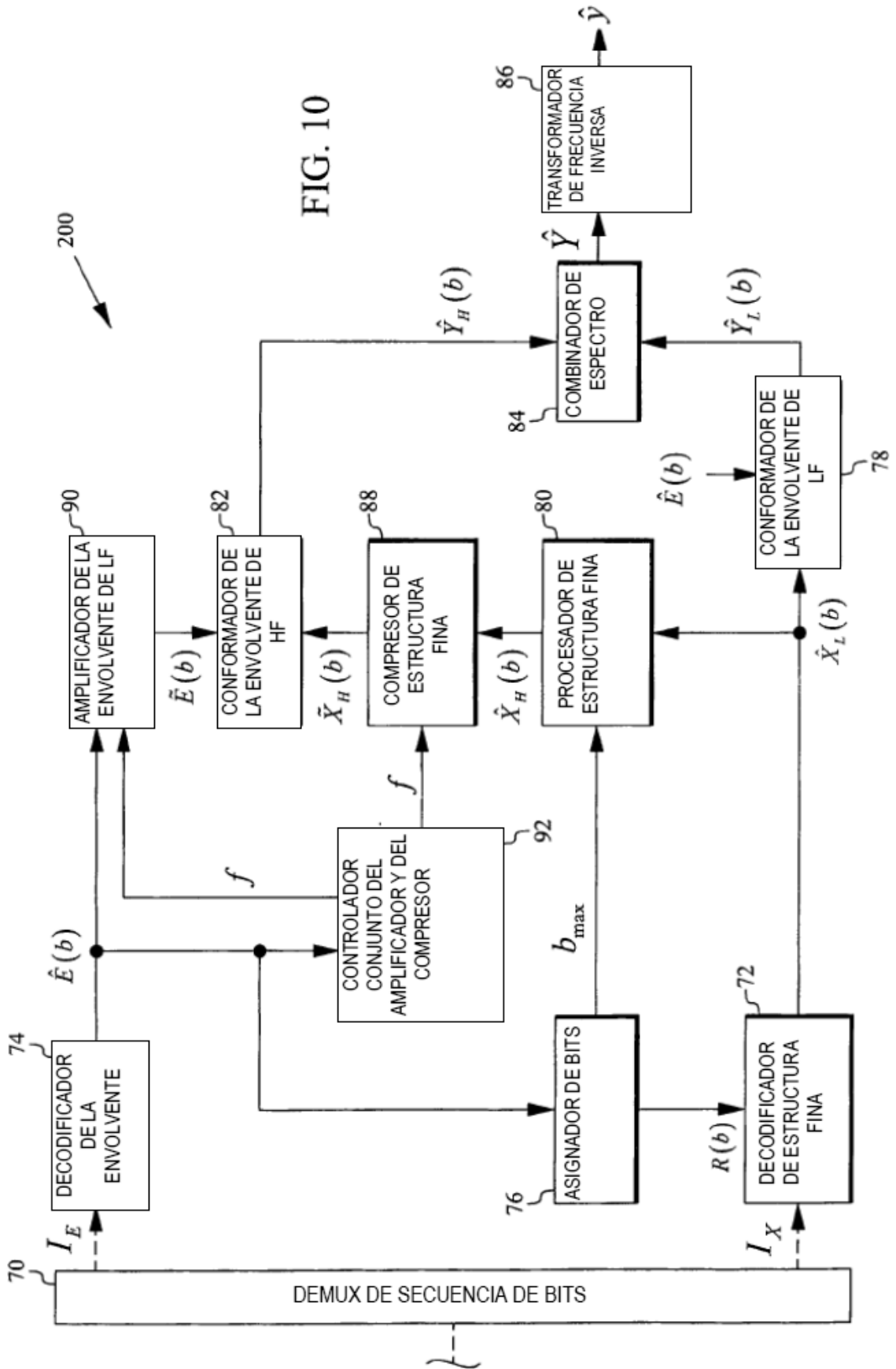


FIG. 9  
(TÉCNICA ANTERIOR)





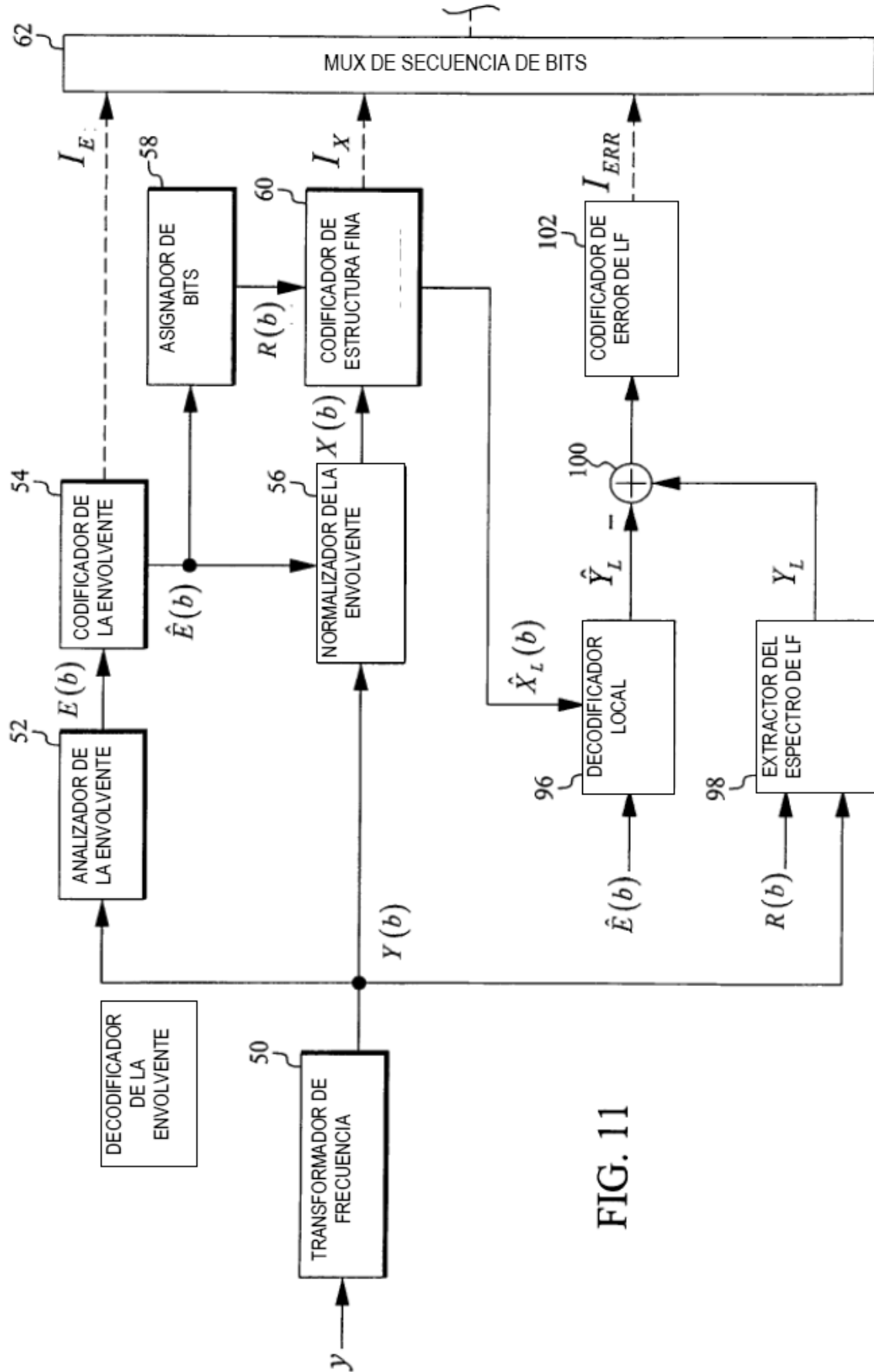
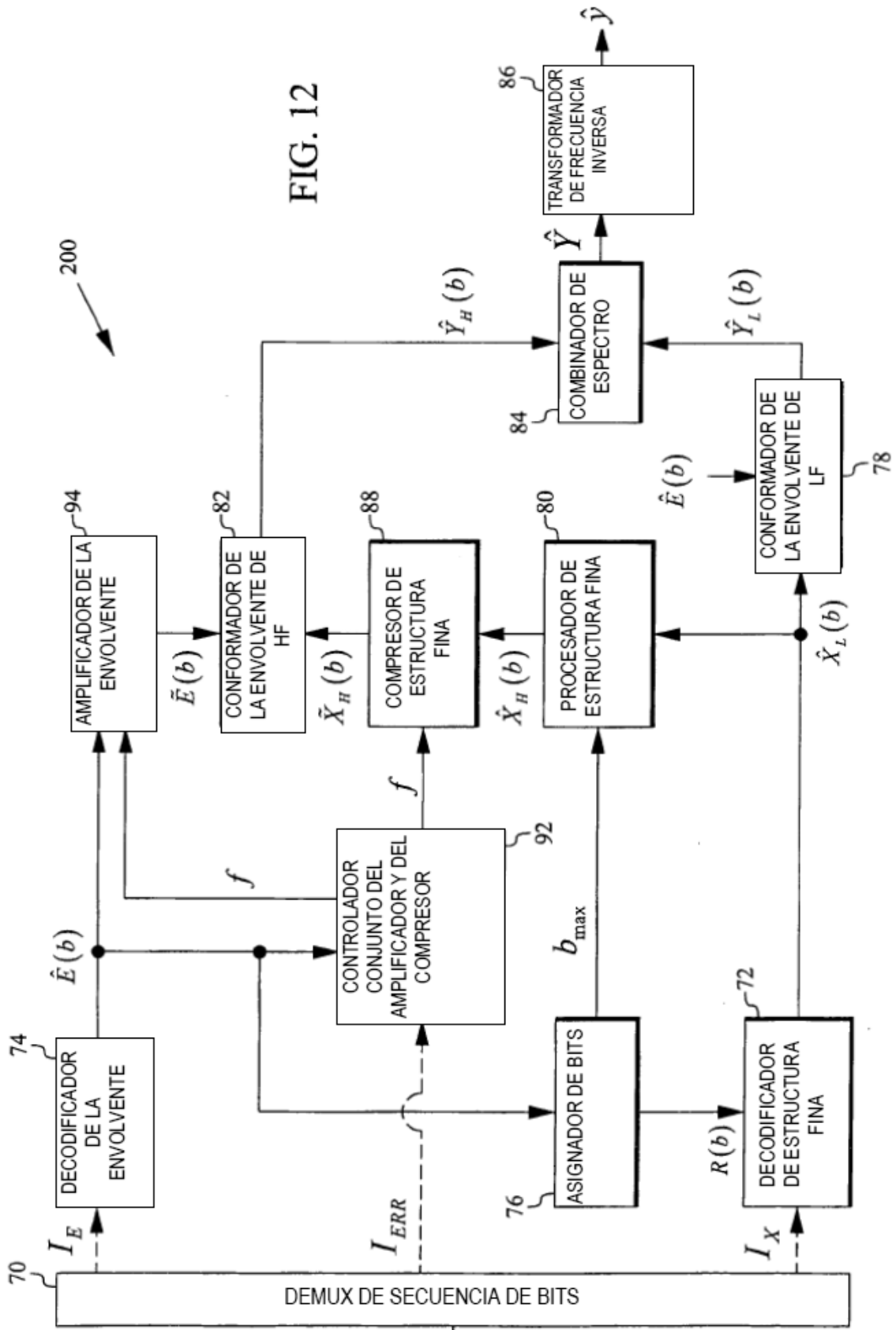


FIG. 11





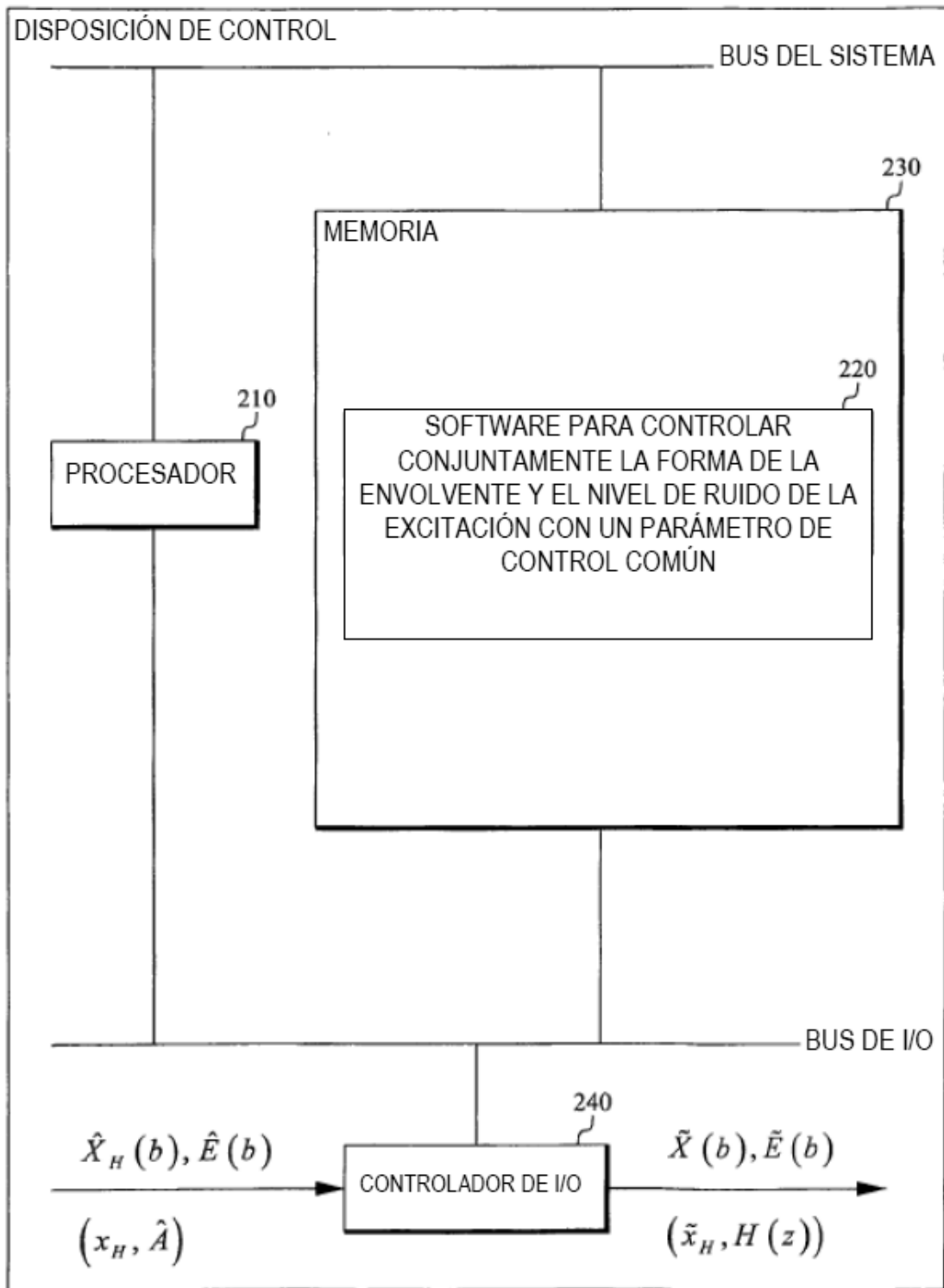


FIG. 13

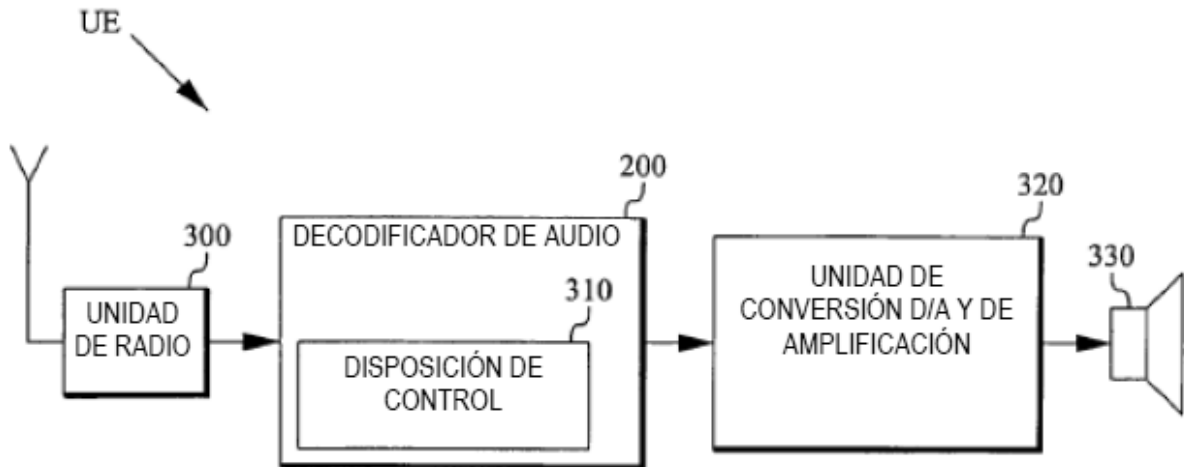


FIG. 14



FIG. 15

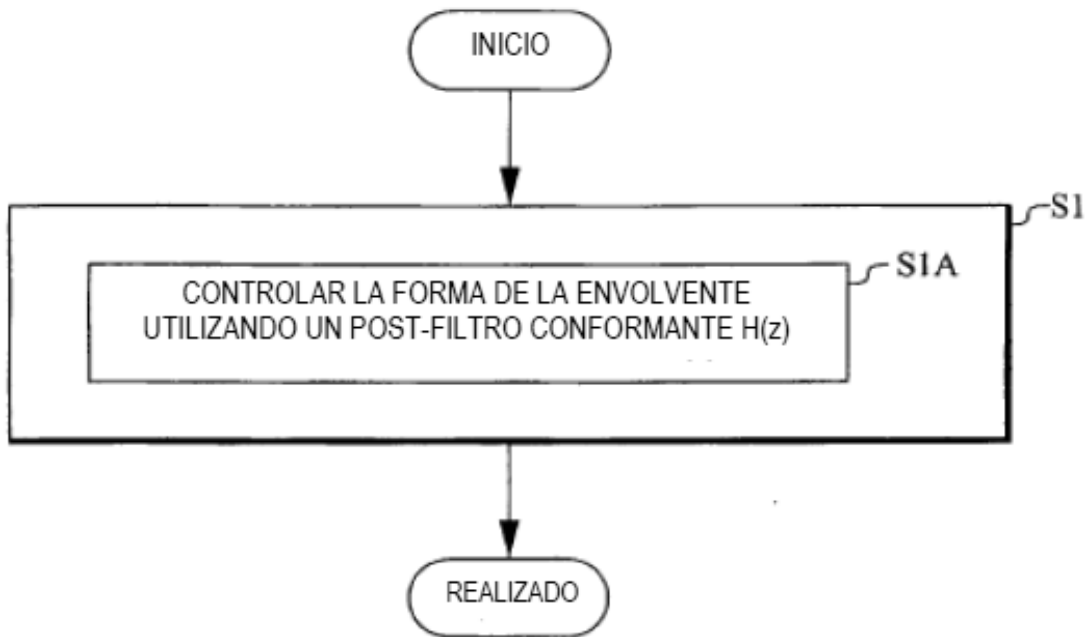


FIG. 16

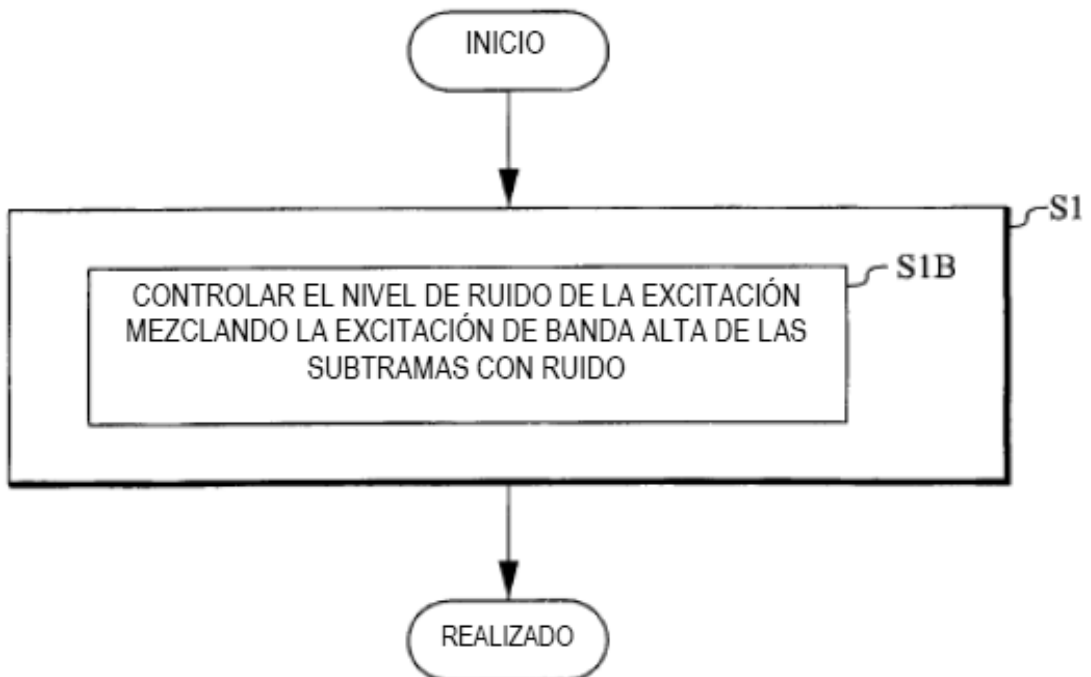


FIG. 17

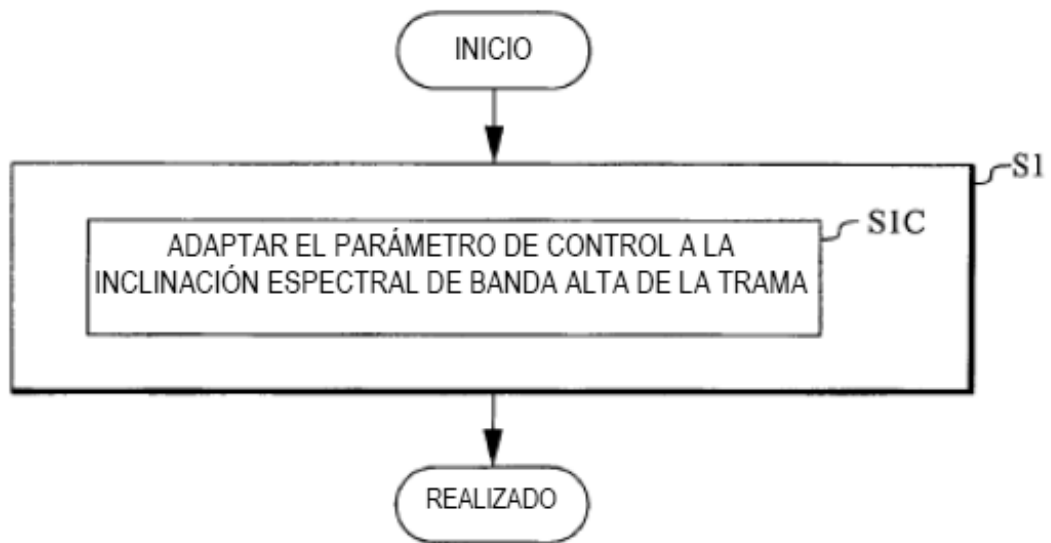


FIG. 18

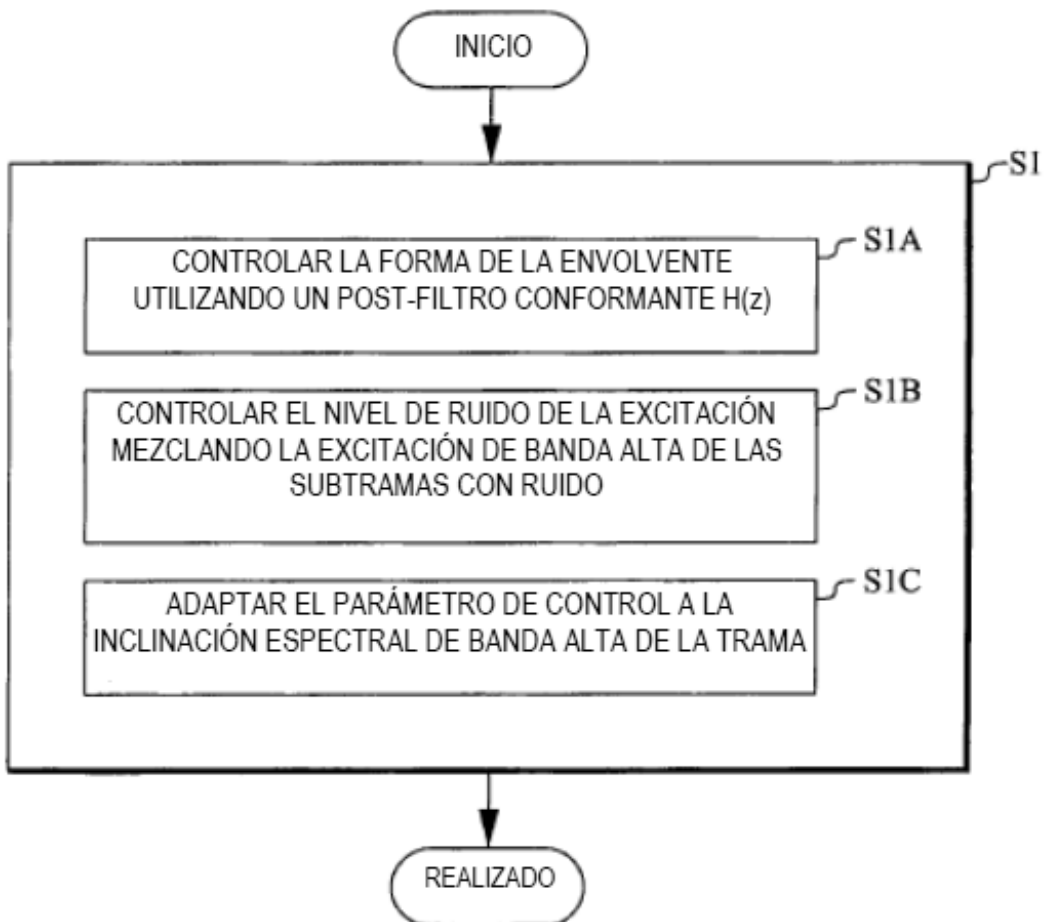


FIG. 19