

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 749 967**

51 Int. Cl.:

G10L 19/06 (2013.01)

G10L 19/032 (2013.01)

G10L 21/038 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.05.2012 E 17190535 (9)**

97 Fecha y número de publicación de la concesión europea: **10.07.2019 EP 3279895**

54 Título: **Codificación de audio en base a una representación eficiente de coeficientes autorregresivos**

30 Prioridad:

02.11.2011 US 201161554647 P

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

24.03.2020

73 Titular/es:

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

72 Inventor/es:

**GRANCHAROV, VOLODYA y
SVERRISSON, SIGURDUR**

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 749 967 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

Codificación de audio en base a una representación eficiente de coeficientes autorregresivos

5 Campo técnico

La tecnología propuesta se refiere a codificación de audio en base a una representación eficiente de coeficientes autorregresivos (AR).

10 Antecedentes

El análisis AR se utiliza comúnmente en la codificación de audio tanto en el dominio tiempo [1] como el dominio transformada [2]. Un enfoque de codificación se divulga en el documento EP1818913. Diferentes aplicaciones utilizan vectores de diferente longitud (el orden de modelo es dependiente principalmente del ancho de banda de la señal codificada; desde 10 coeficientes para señales con un ancho de banda de 4 kHz, hasta 24 coeficientes para señales con un ancho de banda de 16 kHz). Estos coeficientes AR están cuantificados con cuantificación vectorial (VQ) multietapa y dividida, que garantiza una reconstrucción casi transparente. Sin embargo, los esquemas de cuantificación convencionales no están diseñados para el caso en el que los coeficientes AR modelan altas frecuencias de audio (por ejemplo, por encima de 6 kHz), y operan a presupuestos de bit limitados (que no permiten codificación transparente de los coeficientes). Esto introduce grandes errores perceptuales en la señal reconstruida cuando estos esquemas de cuantificación convencionales se utilizan a rangos de frecuencia no óptimos y tasas de bit no óptimas.

Sumario

Un objeto de la tecnología propuesta es un esquema de cuantificación más eficiente para los coeficientes autorregresivos.

Este objeto se consigue en acuerdo con las reivindicaciones adjuntas.

Un primer aspecto de la tecnología propuesta involucra un aparato para codificar una representación espectral paramétrica de coeficientes autorregresivos que parcialmente representan una señal de audio. El aparato comprende:

- medios para codificar una parte de baja frecuencia de la representación espectral paramétrica cuantificando coeficientes de la representación espectral paramétrica que corresponde a una parte de baja frecuencia de la señal de audio;

- medios para codificar una parte de alta frecuencia de la representación espectral paramétrica por promediado ponderado en base a los coeficientes cuantificados volteados alrededor de una frecuencia de traslado especular cuantificada, que separa la parte de baja frecuencia de la parte de alta frecuencia, y una cuadrícula de frecuencia determinada a partir de un libro de códigos de cuadrícula de frecuencia en un procedimiento de búsqueda de ciclo cerrado;

- medios para dejar salir al menos un índice de cuantificación representando los coeficientes cuantificados, un índice de cuantificación que representa la frecuencia de traslado especular cuantificada y un índice de cuantificación que representa la cuadrícula de frecuencia para la transmisión a un descodificador.

Un segundo aspecto de la tecnología propuesta involucra un método de codificar una representación espectral paramétrica de coeficientes autorregresivos que representan parcialmente una señal de audio. El método comprende:

- codifica una parte de baja frecuencia de la representación espectral paramétrica cuantificando coeficientes de la representación espectral paramétrica que corresponden a una parte de baja frecuencia de la señal de audio;

- codifica una parte de alta frecuencia de la representación espectral paramétrica por promediado ponderado en base a los coeficientes cuantificados volteados alrededor de una frecuencia de traslado especular cuantificada, que separa la parte de baja frecuencia de la parte de alta frecuencia, y una cuadrícula de frecuencia determinada a partir de un libro de códigos de cuadrícula de frecuencia en un procedimiento de búsqueda de ciclo cerrado;

- deja salir al menos un índice de cuantificación que representa los coeficientes cuantificados, un índice de cuantificación que representa la frecuencia de traslado especular cuantificada y un índice de cuantificación que representa la cuadrícula de frecuencia para la transmisión a un codificador.

La tecnología propuesta proporciona un esquema de baja tasa de bit para la compresión o codificación de coeficientes autorregresivos. En adición a mejoras perceptuales, la tecnología propuesta también tiene la ventaja de

reducir la complejidad computacional en comparación con métodos de cuantificación de espectro entero.

Breve descripción de los dibujos

- 5 La tecnología propuesta, junto con más objetos y ventajas de la misma, puede entenderse mejor haciendo referencia a la siguiente descripción junto con los dibujos que se acompañan, en los que:
la figura 1 es un diagrama de flujo del método de codificación de acuerdo con la tecnología propuesta;
- 10 la figura 2 ilustra una realización del método de lado de codificador de la tecnología propuesta;
la figura 3 ilustra el volteo de elementos LSF de baja frecuencia cuantificados (representados por puntos negros) a alta frecuencia trasladándolos especularmente al espacio previamente ocupado por la mitad superior del vector LSF;
- 15 la figura 4 ilustra el efecto de alisado de cuadrícula en un espectro de señal;
la figura 5 es un diagrama de bloques de una realización del codificador de acuerdo con la tecnología propuesta;
la figura 6 es un diagrama de bloques de una realización del codificador de acuerdo con la tecnología propuesta;
- 20 la figura 7 es un diagrama de flujo del método de descodificación de acuerdo con la tecnología propuesta;
la figura 8 ilustra un ejemplo del método de lado de descodificador de la tecnología propuesta;
- 25 la figura 9 es un diagrama de bloques de un ejemplo del descodificador de acuerdo con la tecnología propuesta;
la figura 10 es un diagrama de bloques de un ejemplo del descodificador de acuerdo con la tecnología propuesta;
la figura 11 es un diagrama de bloques de una realización del codificador de acuerdo con la tecnología propuesta;
- 30 la figura 12 es un diagrama de bloques de un ejemplo del descodificador de acuerdo con la tecnología propuesta;
la figura 13 ilustra una realización de un equipo de usuario que incluye un codificador de acuerdo con la tecnología propuesta; y
- 35 la figura 14 ilustra un ejemplo de un equipo de usuario que incluye un descodificador de acuerdo con la tecnología propuesta.

Descripción detallada

- 40 La tecnología propuesta requiere como entrada un vector **a** de coeficientes AR (otro nombre utilizado comúnmente es coeficientes de predicción lineal (LP)). Se obtienen típicamente por computar primero las autocorrelaciones $r(j)$ del segmento de audio en ventana $s(n)$, $n=1, \dots, N$, es decir:

45
$$r(j) = \sum_{n=j}^N s(n)s(n-j), \quad j = 0, \dots, M \quad (1)$$

donde M es el orden de modelo predefinido. Entonces los coeficientes AR **a** se obtienen desde la secuencia de autocorrelación $r(j)$ a través del algoritmo Levinson-Durbin [3].

- 50 En un sistema de comunicación de audio AR los coeficientes tienen que transmitirse eficientemente desde la parte de codificador a la de decodificación del sistema. En la tecnología propuesta esto se consigue cuantificando solo ciertos coeficientes, y representando los coeficientes restantes con solo un número pequeño de bits.

Codificador

- 55 La figura 1 es un diagrama de flujo del método de codificación de acuerdo con la tecnología propuesta. El paso S1 codifica una parte de baja frecuencia de la representación espectral paramétrica cuantificando elementos de la representación espectral paramétrica que corresponden a una parte de baja frecuencia de la señal de audio. El paso S2 codifica una parte de alta frecuencia de la representación espectral paramétrica por promediado ponderado en base a los elementos cuantificados volteados alrededor una frecuencia de traslado especular cuantificada, que separa la parte de baja frecuencia de la parte de alta frecuencia, y una cuadrícula de frecuencia determinada a partir de un libro de códigos de cuadrícula de frecuencia en un procedimiento de búsqueda de ciclo cerrado.
- 60

La figura 2 ilustra pasos realizados en el lado de codificador de una realización de la tecnología propuesta. Primero

los coeficientes AR se convierten en una representación de frecuencias de espectro lineal (LSF) en el paso S3, por ejemplo por el algoritmo descrito en [4]. Entonces el vector LSF f se divide en dos partes, indicadas como partes de frecuencia baja (L) y alta (H) en el paso S4. Por ejemplo, en un vector 10 dimensional LSF, los primeros 5 coeficientes pueden asignarse al sub-vector f^l y los coeficientes restantes al sub-vector f^h .

5 Aunque la tecnología propuesta será descrita con referencia a una representación LSF, los conceptos generales pueden también aplicarse a una implementación alternativa en el que el vector AR se convierte en otra representación espectral paramétrica, tal como un par espectral lineal (LSP) o pares espectrales de inmitancia (ISO) en vez de LSF.

10 Solo el sub-vector f^l de baja frecuencia LSF se cuantifica en el paso S5, y sus índices de cuantificación l_n son transmitidos al decodificador. Las LSF de alta frecuencia del sub-vector f^h no son cuantificadas, sino que solamente se usan en la cuantificación de una frecuencia de traslado especular f_m (a \hat{f}_m) y la búsqueda de ciclo cerrado de una cuadrícula de frecuencia óptima g^{opt} desde un conjunto de cuadrículas de frecuencia g^i que forman un libro de
15 códigos de cuadrícula de frecuencia, como se describe con referencia a las ecuaciones (2)-(13) posteriores. Los índices de cuantificación l_m e l_g para la frecuencia de traslado especular y la cuadrícula de frecuencia óptima, respectivamente, representan el vector f^h de alta frecuencia LSF codificado y se transmiten al decodificador. La codificación del sub-vector de alta frecuencia f^h ocasionalmente será denominado "extrapolación" en la siguiente descripción.

20 En la realización propuesta la cuantificación está basada en un conjunto de cuantificadores escalares (SQs) individualmente optimizados en las propiedades estadísticas de los parámetros anteriores. En una implementación alternativa los elementos LSF podrían enviarse a un cuantificador vectorial (VQ) o se puede aún entrenar un VQ para el conjunto combinado de parámetros (los LSF, la frecuencia de traslado especular, y la cuadrícula óptima).

25 Los LSF de baja frecuencia del sub-vector f^l son volteados en el paso S6 adentro del espacio en el que se extienden los LSF de alta frecuencia del sub-vector f^h . Esta operación se ilustra en la figura 3. Primero la frecuencia de traslado especular cuantificada \hat{f}_m se calcula de acuerdo con:

$$30 \quad \hat{f}_m = Q\left(f(M/2) - \hat{f}(M/2 - 1)\right) + \hat{f}(M/2 - 1) \quad (2)$$

donde f denota todo el vector LSF y $Q(\cdot)$ es la cuantificación de la diferencia entre el primer elemento en f^h (concretamente $f(M/2)$) y el último elemento cuantificado en f^l (concretamente $\hat{f}(M/2-1)$), y donde M denota el número total de elementos en la representación espectral paramétrica.

35 Después los LSF volteados $f_{flip}(k)$ son calculados de acuerdo con:

$$f_{flip}(k) = 2\hat{f}_m - \hat{f}(M/2 - 1 - k), \quad 0 \leq k \leq M/2 - 1 \quad (3)$$

40 Entonces, los LSF volteados se re-escalan de manera que estarán limitados dentro del rango $[0 \dots 0.5]$ (como una alternativa el rango puede representarse en radianes como $[0 \dots \pi]$) de acuerdo con:

$$\tilde{f}_{flip}(k) = \begin{cases} (f_{flip}(k) - f_{flip}(0)) \cdot (f_{max} - \hat{f}_m) / \hat{f}_m + f_{flip}(0), & \hat{f}_m > 0.25 \\ f_{flip}(k), & \text{De otra manera} \end{cases} \quad (4)$$

45 Las cuadrículas de frecuencia g^i se re-escalan para caber en el intervalo entre el último elemento LSF cuantificado $\hat{f}(M/2-1)$ y un valor de punto de cuadrícula máximo g_{max} , es decir:

$$\tilde{g}^i(k) = g^i(k) \cdot (g_{max} - \hat{f}(M/2 - 1)) + \hat{f}(M/2 - 1) \quad (5)$$

50 Estos coeficientes volteados y re-escalados $\tilde{f}_{flip}(k)$ (colectivamente denotados \tilde{f}^h en la figura 2) son procesados adicionalmente en el paso S7 alisando con las cuadrículas de frecuencia re-escaladas $\tilde{g}^i(k)$. El alisado tiene la forma de una suma ponderada entre los LSF volteados y re-escalados $\tilde{f}_{flip}(k)$ y las cuadrículas de frecuencia re-escaladas $\tilde{g}^i(k)$, de acuerdo con:

$$55 \quad f_{smooth}(k) = [1 - \lambda(k)] \tilde{f}_{flip}(k) + \lambda(k) \tilde{g}^i(k) \quad (6)$$

donde $\lambda(k)$ y $[1 - \lambda(k)]$ son ponderaciones predefinidos.

Puesto que la ecuación (6) incluye un índice libre i , esto significa que un vector $f_{smooth}(k)$ se generará para cada $\tilde{g}^i(k)$.

Así, la ecuación (6) puede expresarse como:

$$f_{smooth}^i(k) = [1 - \lambda(k)] \tilde{f}_{flip}(k) + \lambda(k) \tilde{g}^i(k) \quad (7)$$

- 5 El alisado se realiza en el paso S7 en una búsqueda de ciclo cerrado sobre todas las cuadrículas de frecuencia g^i , para encontrar la que minimiza un criterio predefinido (descrito después de la ecuación (12) posterior).

Para $M/2 = 5$ las ponderaciones $\lambda(k)$ en la ecuación (7) pueden elegirse como:

10 $\lambda = \{ 0.2, 0.35, 0.5, 0.75, 0.8 \}$ (8)

- En una realización estas constantes están optimizadas perceptualmente (se sugieren diferentes conjuntos de valores, y finalmente se selecciona el conjunto que maximiza la calidad, como se declara por un panel de oyentes). Generalmente los valores de elementos en λ aumentan según el índice k aumenta. Puesto que un índice más alto corresponde a una frecuencia más alta, las frecuencias más altas del espectro resultante están más influenciadas por $\tilde{g}^i(k)$ que por \tilde{f}_{flip} (véase la ecuación (7)). Este resultado de este promediado alisado o ponderado es un espectro más plano hacia las frecuencias altas (la estructura de espectro introducida potencialmente por \tilde{f}_{flip} se retira progresivamente hacia frecuencias más altas).

- 20 Aquí g_{max} se selecciona cerca de, pero menos que 0,5. En este ejemplo, g_{max} se selecciona igual a 0,49.

El método en este ejemplo utiliza 4 cuadrículas entrenadas g^i (menos o más cuadrículas son posibles). Vectores de cuadrícula de plantilla en un rango [0... 1], pre-almacenados en la memoria, son de la forma:

25
$$\begin{cases} g^1 = \{ 0.17274857, 0.35811835, 0.52369229, 0.71552804, 0.85539771 \} \\ g^2 = \{ 0.16313042, 0.30782962, 0.43109281, 0.59395830, 0.81291897 \} \\ g^3 = \{ 0.17172427, 0.33157177, 0.48528862, 0.66492442, 0.82952486 \} \\ g^4 = \{ 0.16666667, 0.33333333, 0.50000000, 0.66666667, 0.83333333 \} \end{cases} \quad (9)$$

Si se asume que la posición del último coeficiente LSF cuantificado $\hat{f}(M/2-1)$ es 0,25, los vectores de cuadrícula re-escalados cogen la forma:

30
$$\begin{cases} \tilde{g}^1 = \{ 0.2915, 0.3359, 0.3757, 0.4217, 0.4553 \} \\ \tilde{g}^2 = \{ 0.2892, 0.3239, 0.3535, 0.3925, 0.4451 \} \\ \tilde{g}^3 = \{ 0.2912, 0.3296, 0.3665, 0.4096, 0.4491 \} \\ \tilde{g}^4 = \{ 0.2900, 0.3300, 0.3700, 0.4100, 0.4500 \} \end{cases} \quad (10)$$

- Un ejemplo del efecto de alisar los coeficientes LSF volteados y re-escalados a los puntos de cuadrícula se ilustra en la figura 4. Con un número creciente de vectores de cuadrícula utilizados en el procedimiento de ciclo cerrado, el espectro resultante se acerca mucho al espectro objetivo.

- 35 Si $g_{max} = 0,5$ en vez de 0,49, el libro de códigos de cuadrícula de frecuencia puede en cambio estar formado por:

$$\begin{cases} g^1 = \{ 0.15998503, 0.31215086, 0.47349756, 0.66540429, 0.84043882 \} \\ g^2 = \{ 0.15614473, 0.30697672, 0.45619822, 0.62493785, 0.77798001 \} \\ g^3 = \{ 0.14185823, 0.26648724, 0.39740108, 0.55685745, 0.74688616 \} \\ g^4 = \{ 0.15416561, 0.27238427, 0.39376780, 0.59287916, 0.86613986 \} \end{cases} \quad (11)$$

- 40 Si se asume de nuevo que la posición del último coeficiente LSF cuantificado $\hat{f}(M/2-1)$ es 0,25, los vectores de cuadrícula re-escalados toman la forma:

$$\begin{cases} \tilde{g}^1 = \{ 0.28999626, 0.32803772, 0.36837439, 0.41635107, 0.46010970 \} \\ \tilde{g}^2 = \{ 0.28903618, 0.32674418, 0.36404956, 0.40623446, 0.44449500 \} \\ \tilde{g}^3 = \{ 0.28546456, 0.31662181, 0.34935027, 0.38921436, 0.43672154 \} \\ \tilde{g}^4 = \{ 0.28854140, 0.31809607, 0.34844195, 0.39821979, 0.46653496 \} \end{cases} \quad (12)$$

Se hace notar que las cuadrículas re-escaladas \tilde{g}^i pueden ser diferentes de trama a trama, puesto que $\hat{f}(M/2-1)$ en la ecuación de re-escalado (5) puede no ser constante, sino variar con el tiempo. Sin embargo, el libro de códigos formado por las cuadrículas de plantilla g^i es constante. En este sentido las cuadrículas re-escaladas \tilde{g}^i pueden ser consideradas como un libro de códigos adaptativo formado a partir de un libro de códigos fijo de cuadrículas de plantilla g^i .

Los vectores LSF f_{smooth}^i creados por la suma ponderada en (7) se comparan con el vector LSF objetivo f^H y la cuadrícula óptima g^i se selecciona como la que minimiza el error cuadrático medio (MSE) entre estos dos vectores. El índice opt de esta cuadrícula óptima puede expresarse matemáticamente como:

$$opt = \arg \min_i \left(\sum_{k=0}^{M/2-1} \left(f_{smooth}^i(k) - f^H(k) \right)^2 \right) \quad (13)$$

donde $f^H(k)$ es un vector objetivo formado por los elementos de la parte de alta frecuencia de la representación espectral paramétrica.

En una implementación alternativa se puede utilizar medidas de error más avanzadas que imitan distorsión espectral (SD), por ejemplo media inversa armónica u otra ponderación en el dominio LSF.

En una realización el libro de códigos de cuadrícula de frecuencia se obtiene con un algoritmo de agrupación de medias de K en un gran conjunto de vectores LSF, que ha sido extraído de una base de datos de habla. Los vectores de cuadrícula en las ecuaciones (9) y (11) se seleccionan como los que, después de un re-escalado de acuerdo con la ecuación (5) y un promediado ponderado con \tilde{f}_{nip} de acuerdo con la ecuación (7), minimizan la distancia al cuadrado a f^H . En otras palabras, estos vectores de cuadrícula, cuando se utilizan en la ecuación (7), dan la mejor representación de los coeficientes LSF de alta frecuencia.

La figura 5 es un diagrama de bloques de una realización del descodificador de acuerdo con la tecnología propuesta. El descodificador 40 incluye un descodificador de baja frecuencia 10 configurado para descodificar una parte de baja frecuencia de la representación espectral paramétrica f cuantificando elementos de la representación espectral paramétrica que corresponden a una parte de baja frecuencia de la señal de audio. El descodificador 40 también incluye un descodificador de alta frecuencia 12 configurado para descodificar una parte de alta frecuencia f^H de la representación espectral paramétrica por promediado ponderado en base a los elementos cuantificados \hat{f}^L volteados alrededor de una frecuencia de traslado especular cuantificada que separa la parte de baja frecuencia de la parte de alta frecuencia, y una cuadrícula de frecuencia de un libro de códigos de cuadrícula de frecuencia 24 en un procedimiento de búsqueda de ciclo cerrado. Las entidades cuantificadas \hat{f}^L , \hat{f}_m , g^{opt} se representan por los índices correspondientes de cuantificación l_n , l_m , l_g , que se transmiten al descodificador.

La figura 6 es un diagrama de bloques de una realización del descodificador de acuerdo con la tecnología propuesta. El descodificador de baja frecuencia 10 recibe todo el vector LSF f , que está dividido en una parte de baja frecuencia o sub-vector f^L y una parte de alta frecuencia o sub-vector f^H por un divisor de vector 14. La parte de baja frecuencia se avanza a un cuantificador 16, que está configurado para codificar la parte de baja frecuencia f^L cuantificando sus elementos, por cuantificación bien escalar o bien vectorial, en una parte de baja frecuencia cuantificada o sub-vector \hat{f}^L . Al menos un índice de cuantificación l_n (dependiendo del método de cuantificación utilizado) se deja salir para la transmisión al descodificador.

El sub-vector \hat{f}^L de baja frecuencia cuantificado y el sub-vector f^H de alta frecuencia aún no codificado se avanzan al descodificador de alta frecuencia 12. Una calculadora de frecuencia de traslado especular 18 se configura para calcular la frecuencia de traslado especular cuantificada \hat{f}_m de acuerdo con la ecuación (2). Las líneas discontinuas indican que sólo se necesitan para ello el último elemento cuantificado $\hat{f}(M/2-1)$ en \hat{f}^L y el primer elemento $f(M/2)$ en f^H . El índice de cuantificación l_m que representa la frecuencia de traslado especular cuantificada \hat{f}_m se deja salir para la transmisión al descodificador.

La frecuencia de traslado especular cuantificada \hat{f}_m se avanza a una unidad de volteo de sub-vector de baja frecuencia cuantificada 20 configurada para voltear los elementos del sub-vector \hat{f}^L de baja frecuencia cuantificada alrededor de la frecuencia de traslado especular cuantificada \hat{f}_m de acuerdo con la ecuación (3). Los elementos volteados $\tilde{f}_{nip}(k)$ y la frecuencia de traslado especular cuantificada \hat{f}_m se avanzan a un re-escalador de elemento volteado 22 configurado para re-escalar los elementos volteados de acuerdo con la ecuación (4).

Las cuadrículas de frecuencia $g^i(k)$ se avanzan desde el libro de códigos de cuadrícula de frecuencia 24 a un re-escalador de cuadrícula de frecuencia 26, que también recibe el último elemento cuantificado $\hat{f}(M/2-1)$ en \hat{f}^L . El re-escalador 26 está configurado para re-escalar de acuerdo con la ecuación (5).

Los LSF volteados y re-escalados $\tilde{f}_{nip}(k)$ del re-escalador de elemento volteado 22 y las cuadrículas de frecuencia re-escaladas $\tilde{g}^i(k)$ desde el re-escalador de cuadrícula de frecuencia 26 se avanzan a una unidad de ponderación 28,

que está configurada para realizar un promediado ponderado de acuerdo con la ecuación (7). Los elementos alisados resultantes $f_{smooth}^i(k)$ y el vector objetivo de alta frecuencia \hat{f}^H se avanzan a una unidad de búsqueda de cuadrícula de frecuencia 30 configurada para seleccionar una cuadrícula de frecuencia g^{opt} de acuerdo con la ecuación (13). El índice l_g correspondiente se transmite al descodificador.

5

Descodificador

La figura 7 es un diagrama de flujo del método de descodificación de acuerdo con la tecnología propuesta. El paso S11 reconstruye elementos de una parte de baja frecuencia de la representación espectral paramétrica que corresponde a una parte de baja frecuencia de la señal de audio de al menos un índice de cuantificación que codifica esa parte de la representación espectral paramétrica. El paso S12 reconstruye elementos de una parte de alta frecuencia de la representación espectral paramétrica por promediado ponderado en base a los elementos decodificados volteados alrededor de una frecuencia de traslado especular descodificada, que separa la parte de baja frecuencia de la parte de alta frecuencia, y una cuadrícula de frecuencia descodificada.

10

15

Los pasos del método realizados en el descodificador se ilustran por el ejemplo en la figura 8. Primero se reciben los índices de cuantificación l_L , l_m , l_g para los LSF de baja frecuencia, la frecuencia óptima de traslado especular y la cuadrícula óptima, respectivamente.

20

En el paso S13 la parte de baja frecuencia cuantificada \hat{f}^L es reconstruida desde un libro de códigos de baja frecuencia utilizando el índice recibido l_L .

Los pasos del método realizados en el descodificador para reconstruir la parte de alta frecuencia \hat{f}^H son muy similares a pasos de procesamiento de codificador ya descritos en ecuaciones (3)-(7).

25

Los pasos de volteo y re-escalado realizados en el descodificador (en S14) son idénticos a las operaciones de codificador, y por lo tanto descritos exactamente en las ecuaciones (3)-(4).

30

Los pasos (en S15) de re-escalado de la cuadrícula (ecuación (5)), y alisado con ello (ecuación (6)), requieren solo ligera modificación en el descodificador, porque no se realiza la búsqueda de ciclo cerrado (búsqueda sobre i). Esto es porque el descodificador recibe el índice óptimo opt desde la corriente de bits. Estas ecuaciones en cambio cogen la siguiente forma:

$$\tilde{g}^{opt}(k) = g^{opt}(k) \cdot (g_{max} - \hat{f}(M/2 - 1)) + \hat{f}(M/2 - 1) \quad (14)$$

35

y

$$f_{smooth}(k) = [1 - \lambda(k)] \tilde{f}_{flip}(k) + \lambda(k) \tilde{g}^{opt}(k) \quad (15)$$

40

respectivamente. El vector f_{smooth} representa la parte de alta frecuencia \hat{f}^H de la señal descodificada.

Finalmente, las partes de baja y alta frecuencia \hat{f}^L , \hat{f}^H del vector LSF se combinan en el paso S16, y el vector resultante \hat{f} se transforma en coeficientes AR \hat{a} en el paso S17.

45

La figura 9 es un diagrama de bloques de un ejemplo del descodificador 50 de acuerdo con la tecnología propuesta. Un descodificador de baja frecuencia 60 se configura para reconstruir elementos \hat{f}^L de una parte de baja frecuencia f^L de la representación espectral paramétrica f que corresponde a una parte de baja frecuencia de la señal de audio de al menos un índice de cuantificación l_L que codifica esa parte de la representación espectral paramétrica. Un descodificador de alta frecuencia 62 está configurado para reconstruir elementos \hat{f}^H de una parte de alta frecuencia f^H de la representación espectral paramétrica por promediado ponderado en base a los elementos descodificados \hat{f}^L volteados alrededor de una frecuencia de traslado especular descodificada \hat{f}_m , que separa la parte de baja frecuencia de la parte de alta frecuencia, y una cuadrícula de frecuencia descodificada g^{opt} . La cuadrícula de frecuencia g^{opt} se obtiene recuperando la cuadrícula de frecuencia que corresponde a un índice recibido l_g de un libro de códigos de cuadrícula de frecuencia 24 (este es el mismo libro de códigos que en el codificador).

55

La figura 10 es un diagrama de bloques de un ejemplo del descodificador de acuerdo con la tecnología propuesta. El descodificador de baja frecuencia recibe al menos un índice de cuantificación l_L , dependiendo de si se utiliza cuantificación escalar o vectorial, y lo avanza a un descodificador de índice de cuantificación 66, que reconstruye elementos \hat{f}^L de la parte de baja frecuencia de la representación espectral paramétrica. El descodificador de alta frecuencia 62 recibe un índice de cuantificación de frecuencia de traslado especular l_m , que se avanza a un descodificador de frecuencia de traslado especular 66 para descodificar la frecuencia de traslado especular \hat{f}_m . Los bloques restantes 20, 22, 24, 26 y 28 realizan las mismas funciones que los bloques numerados correspondientemente en el codificador ilustrado en la figura 6. Las diferencias esenciales entre el codificador y el descodificador son que la frecuencia de traslado especular se descodifica desde el índice l_m en vez de calcularse

60

desde la ecuación (2), y que la unidad de búsqueda de cuadrícula de frecuencia 30 en el codificador no se requiere, ya que la cuadrícula de frecuencia óptima se obtiene directamente desde el libro de códigos de cuadrícula de frecuencia 24 mirando la cuadrícula de frecuencia g^{opt} que corresponde al índice recibido l_g .

5 Los pasos, funciones, procedimientos y/o bloques descritos aquí pueden implementarse en equipo físico informático utilizando cualquier tecnología convencional, tal como tecnología de circuito discreto o de circuito integrado, incluyendo tanto circuitos electrónicos de propósito general como circuitos específicos de aplicación.

10 Alternativamente, al menos algunos de los pasos, funciones, procedimientos y/o bloques descritos aquí pueden implementarse en equipo lógico informático para la ejecución por equipo de procesamiento adecuado. Este equipo puede incluir, por ejemplo, uno o varios microprocesadores, uno o varios procesadores de señal digital (DSP), uno o varios circuitos integrados de aplicación específica (ASIC), equipo físico informático acelerado de video o uno o varios dispositivos lógicos programables adecuados, tales como matrices de puertas programables de campo (FPGA). También son factibles combinaciones de tales elementos de procesamiento.

15 También debería entenderse que puede ser posible reutilizar las capacidades de procesamiento generales ya presentes en un UE. Esto se puede hacer, por ejemplo, reprogramando el equipo lógico informático existente o añadiendo nuevos componentes de equipo lógico informático.

20 La figura 11 es un diagrama de bloques de una realización del codificador 40 de acuerdo con la tecnología propuesta. Esta realización está basada en un procesador 110, por ejemplo un microprocesador, que ejecuta equipo lógico informático 120 para cuantificar la parte de baja frecuencia f^l de la representación espectral paramétrica, y equipo lógico informático 130 para la búsqueda de una extrapolación óptima representada por la frecuencia de traslado especular f_m y el vector de cuadrícula de frecuencia óptima g^{opt} . El equipo lógico informático se almacena en la memoria 140. El procesador 110 se comunica con la memoria sobre un sistema de bus. La representación espectral paramétrica entrante f se recibe por un controlador de entrada/salida (I/O) 150 que controla un bus de I/O, al que el procesador 110 y la memoria 140 están conectados. El equipo lógico informático 120 puede implementar la funcionalidad del codificador de baja frecuencia 10. El equipo lógico informático 130 puede implementar la funcionalidad del codificador de alta frecuencia 12. Los parámetros cuantificados f^l , f_m , g^{opt} (o preferiblemente los índices correspondientes l_{fL} , l_m , l_g) obtenidos del equipo lógico informático 120 y 130 se dejan salir desde la memoria 140 por el controlador I/O sobre el bus I/O.

35 La figura 12 es un diagrama de bloques de un ejemplo del descodificador 50 de acuerdo con la tecnología propuesta. Este ejemplo está basado en un procesador 210, por ejemplo un microprocesador, que ejecuta un equipo lógico informático 220 para descodificar la parte de baja frecuencia f^l de la representación espectral paramétrica, y equipo lógico informático 230 para descodificar la parte de baja frecuencia f^l de la representación espectral paramétrica por extrapolación. El equipo lógico informático se almacena en la memoria 240. El procesador 210 se comunica con la memoria sobre un bus de sistema. Los parámetros codificados entrantes f^l , f_m , g^{opt} (representados por l_{fL} , l_m , l_g) se reciben por un controlador de entrada/salida (I/O) 250 que controla un bus I/O, al que están conectados el procesador 210 y la memoria 240. El equipo lógico informático 220 puede implementar la funcionalidad del descodificador de baja frecuencia 60. El equipo lógico informático 230 puede implementar la funcionalidad del descodificador de alta frecuencia 62. La representación paramétrica descodificada f (f^l combinada con f^h) obtenida desde el equipo lógico informático 220 y 230 se deja salir desde la memoria 240 por el controlador I/O 250 sobre el bus I/O.

45 La figura 13 ilustra una realización de un equipo de usuario UE que incluye un descodificador de acuerdo con la tecnología propuesta. Un micrófono 70 avanza una señal de audio a un convertidor A/D 72. La señal de audio digitalizada es codificada por un codificador de audio 74. Solo los componentes relevantes para ilustrar la tecnología propuesta se ilustran en el codificador de audio 74. El codificador de audio 74 incluye un estimador de coeficiente AR 76, un convertidor de representación espectral paramétrica a AR 78 y un codificador 40 de la representación espectral paramétrica. La representación espectral paramétrica codificada (junto con otros parámetros de audio codificados que no se necesitan para ilustrar la presente tecnología) se avanza a una unidad de radio 80 para codificación de canal y conversión de frecuencia de radio y transmisión a un descodificador sobre una antena.

55 La figura 14 ilustra un ejemplo de un equipo de usuario UE que incluye un descodificador de acuerdo con la tecnología propuesta. Una antena recibe una señal que incluye la representación espectral paramétrica codificada y la avanza a una unidad de radio 82 para conversión desde la frecuencia de radio y descodificación de canal. La señal digital resultante se avanza a un descodificador de audio 84. Solo los componentes relevantes para ilustrar la tecnología propuesta se ilustran en el descodificador de audio 84. El descodificador de audio 84 incluye un descodificador 50 de una representación espectral paramétrica y un convertidor de representación espectral paramétrica a AR 86. Los coeficientes AR se utilizan (junto con otros parámetros de audio descodificados que no se necesitan para ilustrar la presente tecnología) para descodificar la señal de audio, y las muestras de audio resultantes se avanzan a una unidad de conversión D/A y amplificación 88, que deja salir la señal de audio a un altavoz 90.

65 En una aplicación de ejemplo el esquema de extrapolación-cuantificación AR propuesto se utiliza en un contexto

BWE. En este caso el análisis AR se realiza a una cierta banda de alta frecuencia, y solo se utilizan coeficientes AR para el filtro de síntesis. En vez de obtenerse con el filtro de análisis correspondiente, la señal de excitación para esta banda alta se extrapola desde una excitación de banda baja codificada independientemente.

5 En otra aplicación de ejemplo el esquema de extrapolación de cuantificación AR propuesto se utiliza en un esquema de codificación tipo ACELP. Los codificadores ACELP modelan un tracto vocal de un hablante con un modelo AR. Se genera una señal de excitación $e(n)$ pasando una forma de onda $s(n)$ a través de un filtro blanqueador $e(n) = A(z)s(n)$, donde $A(z) = 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_Mz^{-M}$, es el modelo AR de orden M. De una manera trama a trama un conjunto de coeficientes AR $a = [a_1 \ a_2 \ \dots \ a_M]^T$, y señal de excitación se cuantifican, y se transmiten índices de
10 cuantificación sobre la red. En el descodificador, se genera habla sintetizada de una manera trama a trama mandando la señal de excitación reconstruida a través del filtro de síntesis reconstruido $A(z)^{-1}$.

En una aplicación más de ejemplo el esquema de extrapolación de cuantificación AR propuesto se utiliza como una manera eficiente para parametrizar una envolvente de espectro de un código de audio de transformada. De una
15 manera a corto plazo la forma de onda se transforma a dominio frecuencia, y la respuesta de frecuencia de los coeficientes AR se utiliza para aproximar la envolvente de espectro y normalizar el vector transformado (para crear un vector residual). Después los coeficientes AR y el vector residual son codificados y transmitidos al descodificador.

Se entenderá por los expertos en la técnica que pueden hacerse varias modificaciones y cambios a la tecnología
20 propuesta sin salir del alcance de la misma, que está definido por las reivindicaciones adjuntas.

Abreviaturas

25 ACELP Predicción lineal excitada de código algebraico

ASIC Circuitos integrados de aplicación específica

AR Auto regresión

30 BWE Extensión de ancho de banda

DSP Procesador de señal digital

35 Campo FPGA Matriz de puerta programable

ISP Pares espectrales de inmitancia

LP Predicción lineal

40 LSF Frecuencias espectrales lineales

MSE Error cuadrático medio

45 SD Distorsión espectral

SQ Cuantificador escalar

UE Equipo de usuario

50 VQ Cuantificación vectorial

Referencias

55 [1] 3GPP TS 26.090, "Adaptive Multi-Rate (AMR) speech codec; Trans-coding functions", p.13, 2007

[2] N. Iwakami, et al., High-quality audio-coding at less than 64 kbit/s by using transform-domain weighted interleaved vector quantization (TWINVQ), IEEE ICASSP, vol. 5, pp. 3095-3098, 1995

60 [3] J. Makhoul, "Linear prediction: A tutorial review", Proc. IEEE, vol. 63, p. 566, 1975

[4] P. Kabal and R.P. Ramachandran, "The computation of line spectral frequencies using Chebyshev polynomials", IEEE Trans. on ASSP, vol. 34, nº 6, pp. 1419-1426, 1986

REIVINDICACIONES

1. Un codificador (40) para codificar una representación espectral paramétrica (f) de coeficientes autorregresivos (a) que representan parcialmente una señal de audio, incluyendo dicho codificador:

5 - medios (10) para codificar una parte de baja frecuencia (f^l) de la representación espectral paramétrica (f) por cuantificación de coeficientes de la representación espectral paramétrica que corresponden a una parte de baja frecuencia de la señal de audio;

10 - medios (12) para codificar una parte de alta frecuencia (f^h) de la representación espectral paramétrica (f) por promediado ponderado en base a los coeficientes cuantificados (\hat{f}^l) volteados alrededor una frecuencia de traslado especular cuantificada (\hat{f}_m), que separa la parte de baja frecuencia de la parte de alta frecuencia, y una cuadrícula de frecuencia (g^{opt}) determinada a partir de un libro de códigos de cuadrícula de frecuencia (24) en un procedimiento de búsqueda de ciclo cerrado; y

15 - medios para dejar salir al menos un índice de cuantificación (l_{fl}) que representa los coeficientes cuantificados (\hat{f}^l), un índice de cuantificación (l_m) que representa la frecuencia de traslado especular cuantificada (\hat{f}_m) y un índice de cuantificación (l_g) que representa la cuadrícula de frecuencia (g^{opt}) para la transmisión a un descodificador.

20 2. El codificador de la reivindicación 1, en el que los medios (12) para codificar una parte de alta frecuencia (f^h) de la representación espectral paramétrica (f) comprenden medios (18) para calcular la frecuencia de traslado especular cuantificada \hat{f}_m de acuerdo con:

$$\hat{f}_m = Q\left(f(M/2) - \hat{f}(M/2 - 1)\right) + \hat{f}(M/2 - 1),$$

25 donde:

Q denota cuantificación de la expresión en el paréntesis adyacente,

30 M denota el número total de coeficientes en la representación espectral paramétrica,

$f(M/2)$ denota el primer coeficiente en la parte de alta frecuencia, y

$\hat{f}(M/2 - 1)$ denota el último coeficiente cuantificado en la parte de baja frecuencia.

35 3. El codificador de la reivindicación 2, en el que los medios (12) para codificar una parte de alta frecuencia (f^h) de la representación espectral paramétrica (f) comprenden medios (20) para voltear los coeficientes cuantificados de la parte de baja frecuencia (f^l) de la representación espectral paramétrica (f) alrededor de la frecuencia de traslado especular cuantificada \hat{f}_m de acuerdo con:

$$40 \quad f_{flip}(k) = 2\hat{f}_m - \hat{f}(M/2 - 1 - k), \quad 0 \leq k \leq M/2 - 1$$

donde $\hat{f}(M/2 - 1 - k)$ denota coeficiente cuantificado $M/2 - 1 - k$.

45 4. El codificador de la reivindicación 3, en el que los medios (12) para codificar una parte de alta frecuencia (f^h) de la representación espectral paramétrica (f) comprenden medios (22) para re-escalar los coeficientes $f_{flip}(k)$ de acuerdo con:

$$\tilde{f}_{flip}(k) = \begin{cases} (f_{flip}(k) - f_{flip}(0)) \cdot (f_{max} - \hat{f}_m) / \hat{f}_m + f_{flip}(0), & \hat{f}_m > 0.25 \\ f_{flip}(k), & \text{De otra manera} \end{cases}$$

50 5. El codificador de la reivindicación 4, en el que los medios (12) para codificar una parte de alta frecuencia (f^h) de la representación espectral paramétrica (f) comprenden medios (26) para re-escalar las cuadrículas de frecuencia g^i del libro de códigos de cuadrícula de frecuencia (24) para caber en el intervalo entre el último coeficiente cuantificado $\hat{f}(M/2 - 1)$ en la parte de baja frecuencia y un máximo valor de punto de cuadrícula g_{max} de acuerdo con:

$$55 \quad \tilde{g}^i(k) = g^i(k) \cdot (g_{max} - \hat{f}(M/2 - 1)) + \hat{f}(M/2 - 1).$$

6. El codificador de la reivindicación 5, en el que los medios (12) para codificar una parte de alta frecuencia (f^H) de la representación espectral paramétrica (f) comprenden medios (28) para realizar promediado ponderado de los coeficientes volteados y re-escalados $\tilde{f}_{flip}(k)$ y las cuadrículas de frecuencia re-escaladas $\tilde{g}^i(k)$ de acuerdo con:

$$5 \quad f_{smooth}^i(k) = [1 - \lambda(k)] \tilde{f}_{flip}(k) + \lambda(k) \tilde{g}^i(k)$$

donde $\lambda(k)$ y $[1-\lambda(k)]$ son ponderaciones predefinidas.

10 7. El codificador de la reivindicación 6, en el que los medios (12) para codificar una parte de alta frecuencia (f^H) de la representación espectral paramétrica (f) comprenden medios (30) para seleccionar una cuadrícula de frecuencia g^{opt} , donde el índice opt satisface el criterio:

$$opt = \arg \min_i \left(\sum_{k=0}^{M/2-1} (f_{smooth}^i(k) - f^H(k))^2 \right)$$

15 donde $f^H(k)$ es un vector objetivo formado por los coeficientes de la parte de alta frecuencia de la representación espectral paramétrica.

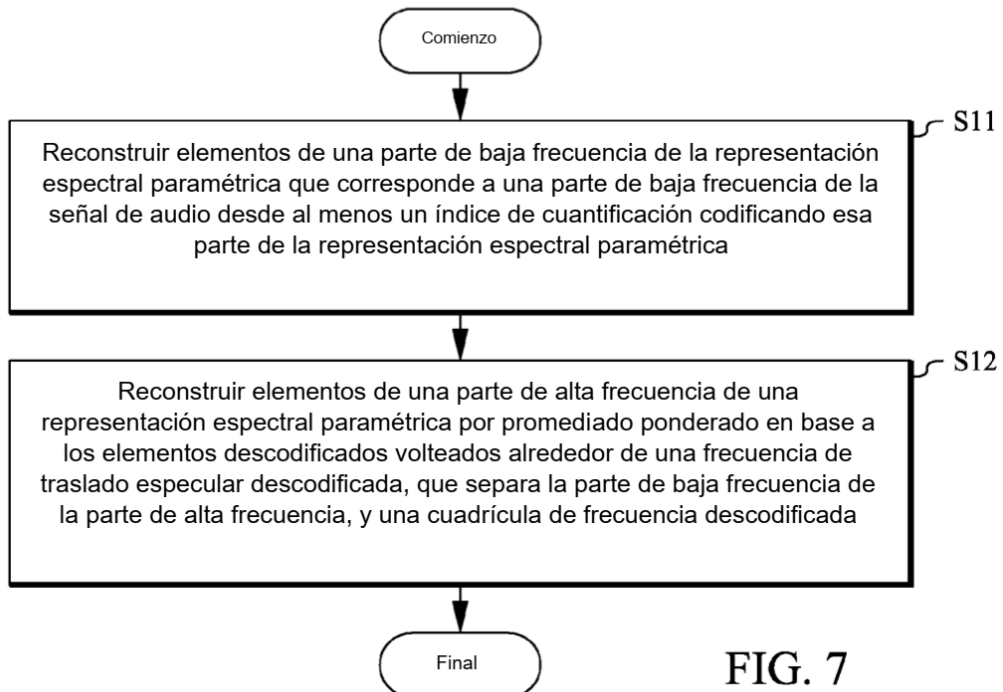
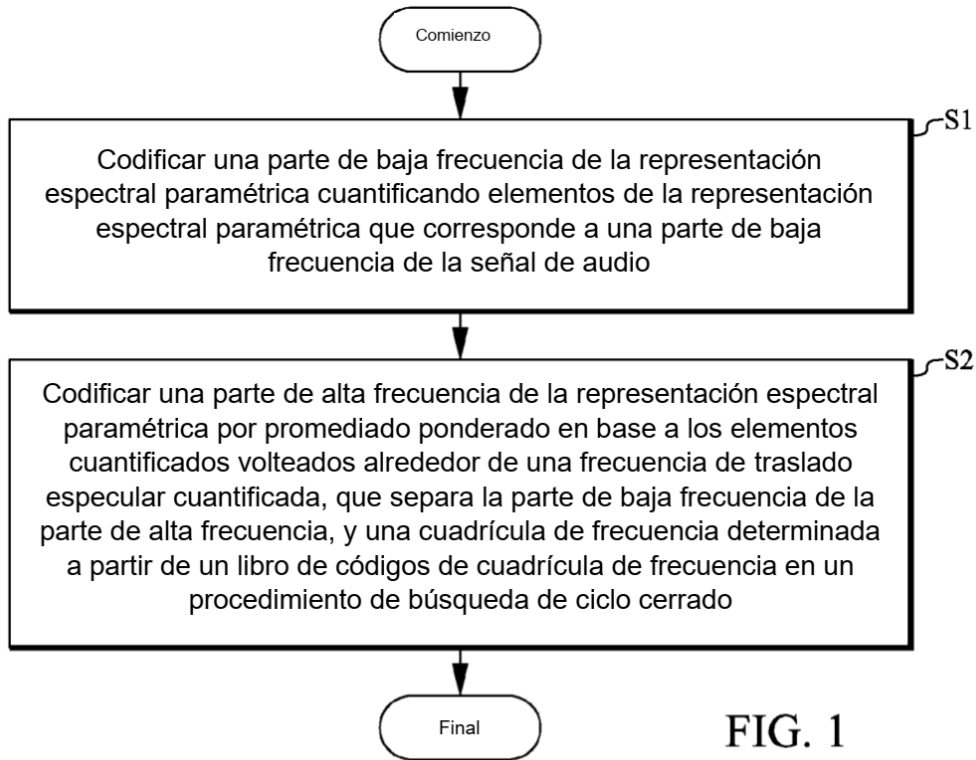
8. Un equipo de usuario que comprende un codificador (40) de acuerdo con una cualquiera de las reivindicaciones 1-7.

20 9. Un método de codificar una representación espectral paramétrica (f) de coeficientes autorregresivos (a) que parcialmente representan una señal de audio, comprendiendo dicho método:

25 - codificar una parte de baja frecuencia (f^L) de la representación espectral paramétrica (f) por cuantificación de coeficientes de la representación espectral paramétrica que corresponden a una parte de baja frecuencia de la señal de audio;

30 - codificar una parte de alta frecuencia (f^H) de la representación espectral paramétrica (f) por promediado ponderado en base a los coeficientes cuantificados (\hat{f}^L) volteados alrededor de una frecuencia de traslado especular cuantificada (\hat{f}_m), que separa la parte de baja frecuencia de la parte de alta frecuencia, y una cuadrícula de frecuencia (g^{opt}) determinada a partir de un libro de códigos de cuadrícula de frecuencia (24) en un procedimiento de búsqueda de cı́bulo cerrado; y

35 - dejar salir al menos un índice de cuantificación (I_{FL}) que representa los coeficientes cuantificados (\hat{f}^L), un índice de cuantificación (I_m) que representa la frecuencia de traslado especular cuantificada (\hat{f}_m) y un índice de cuantificación (I_g) que representa la cuadrícula de frecuencia (g^{opt}) para la transmisión a un descodificador.



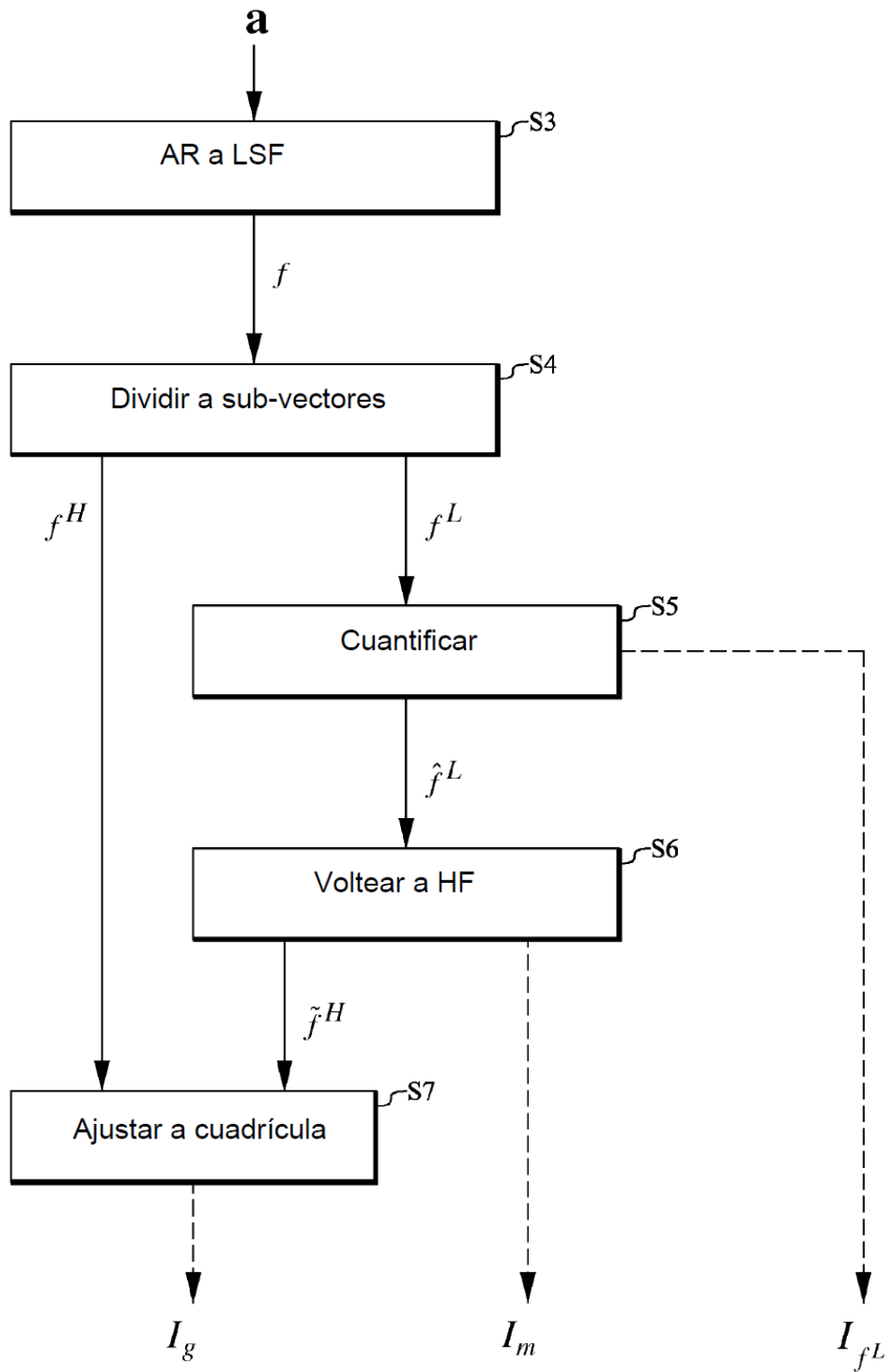


FIG. 2

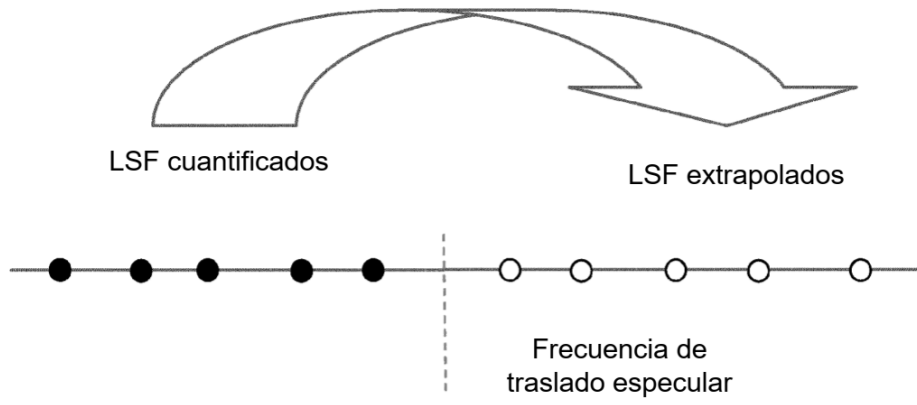


FIG. 3

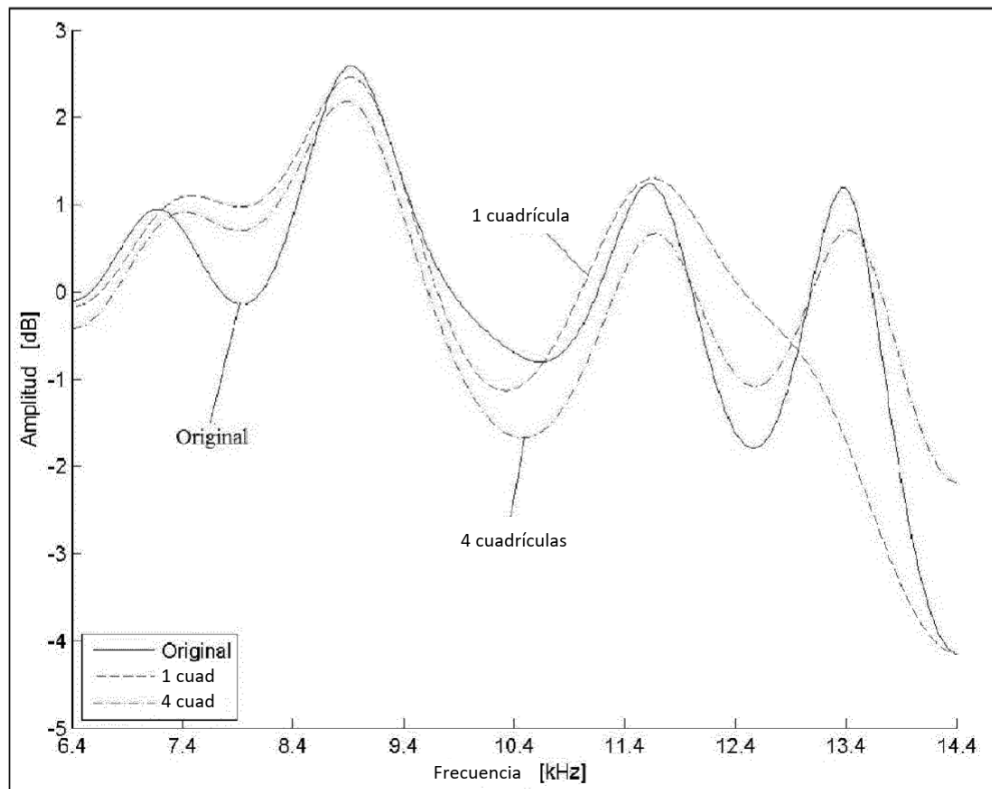


FIG. 4

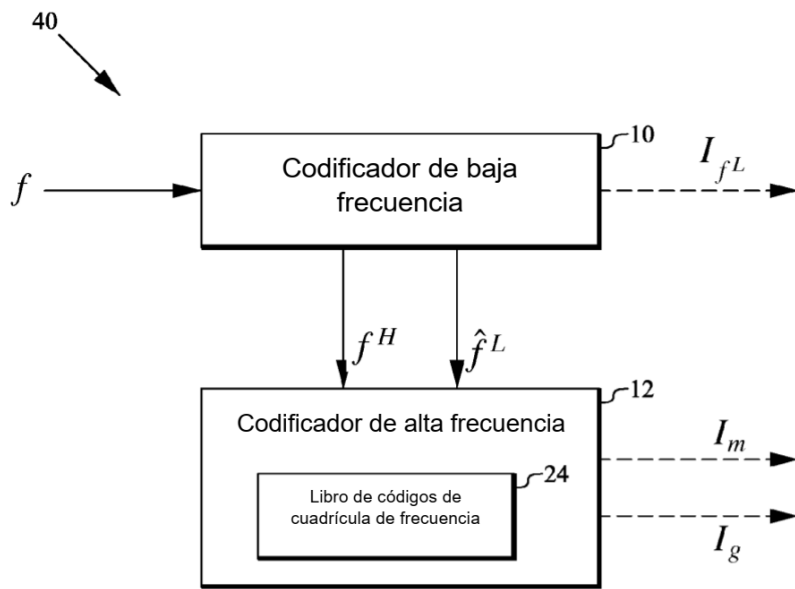


FIG. 5

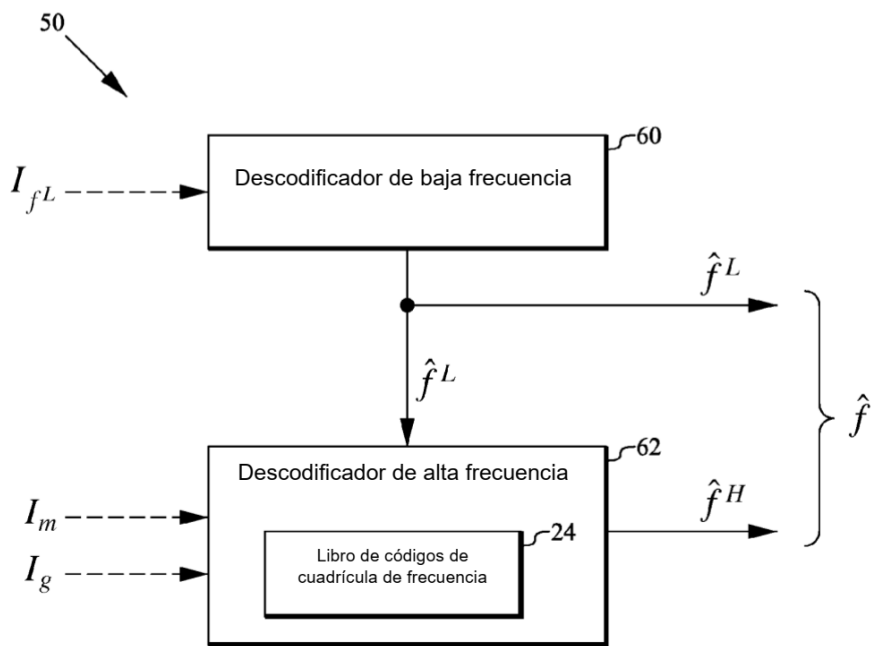


FIG. 9

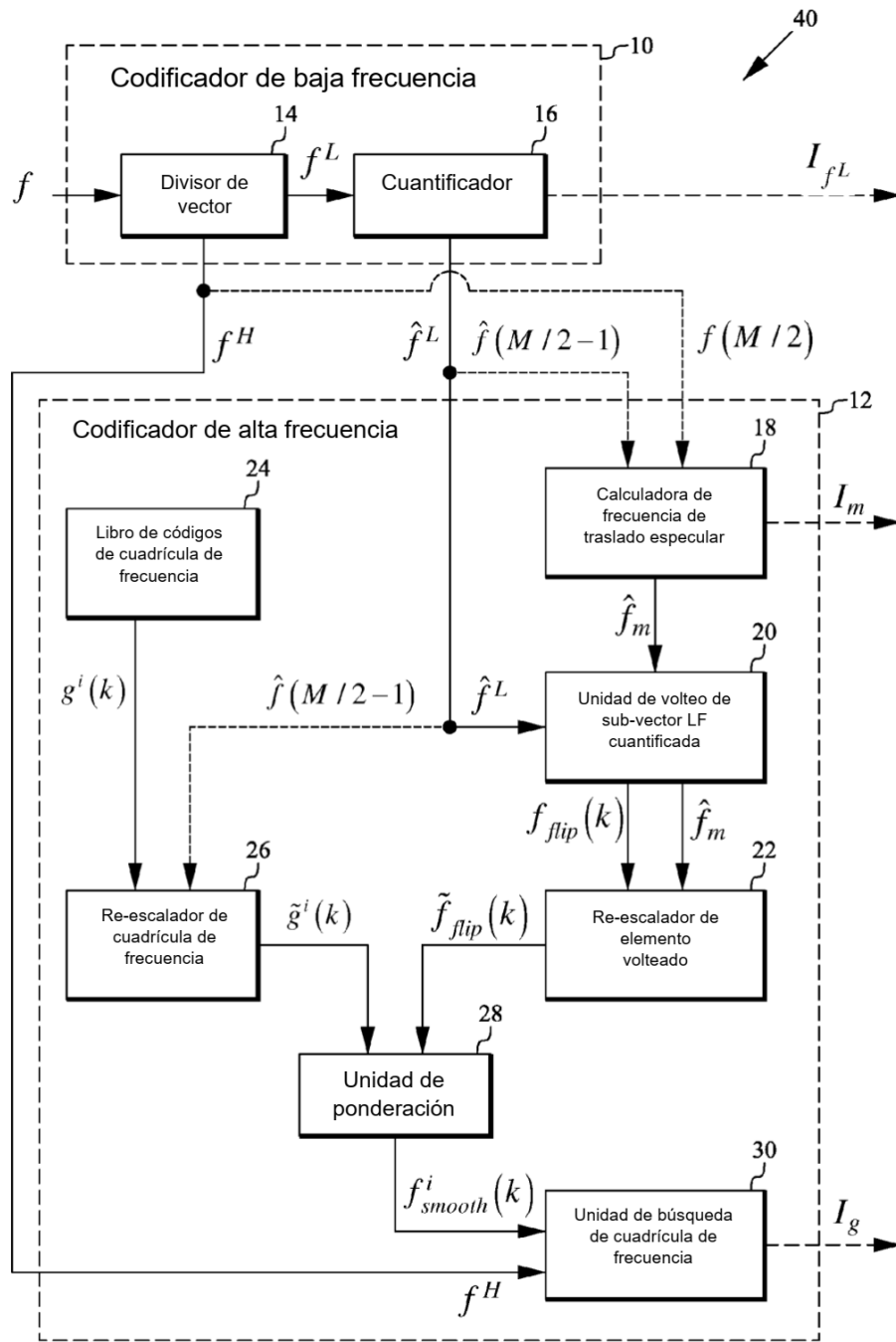


FIG. 6

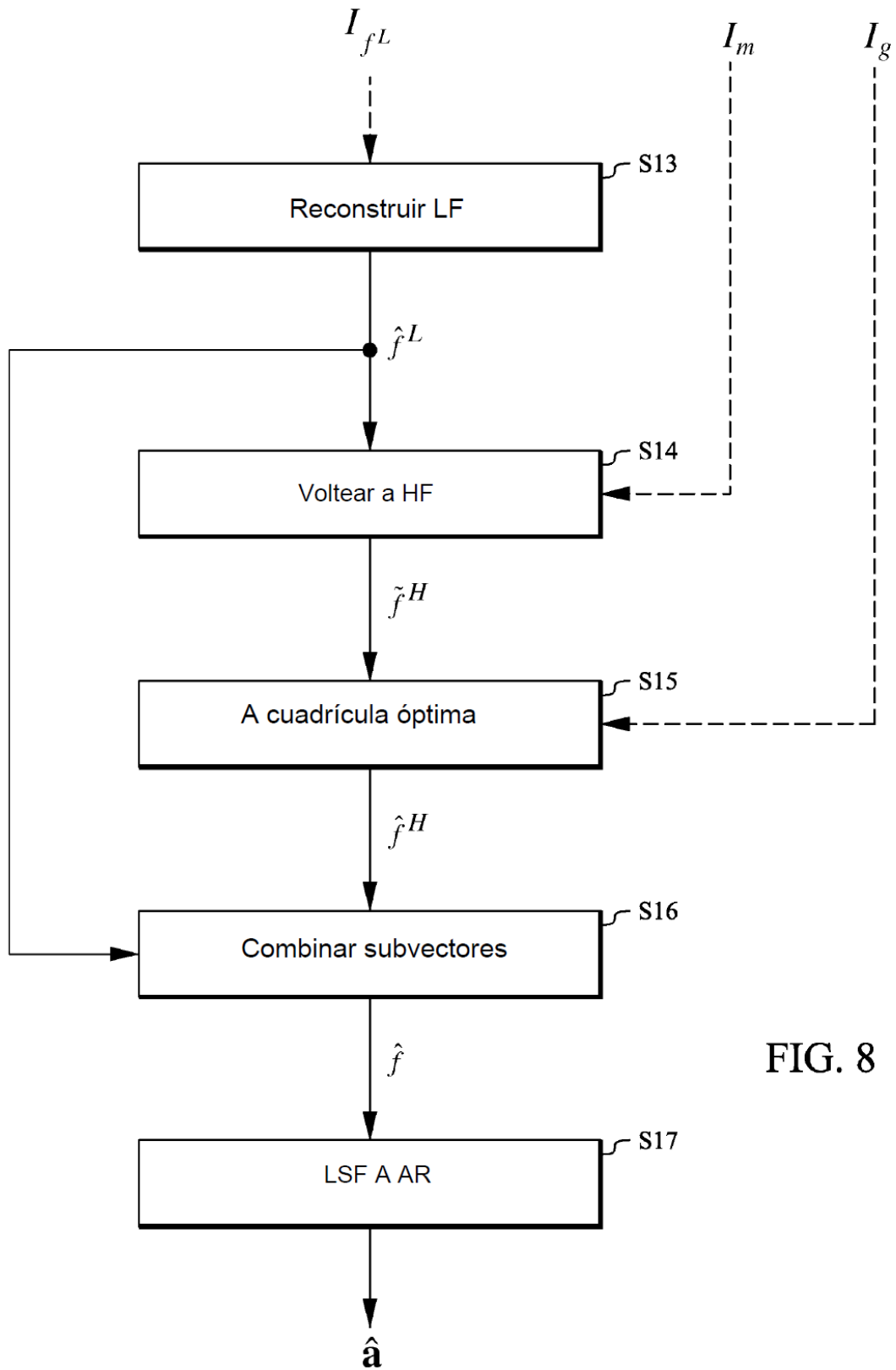


FIG. 8

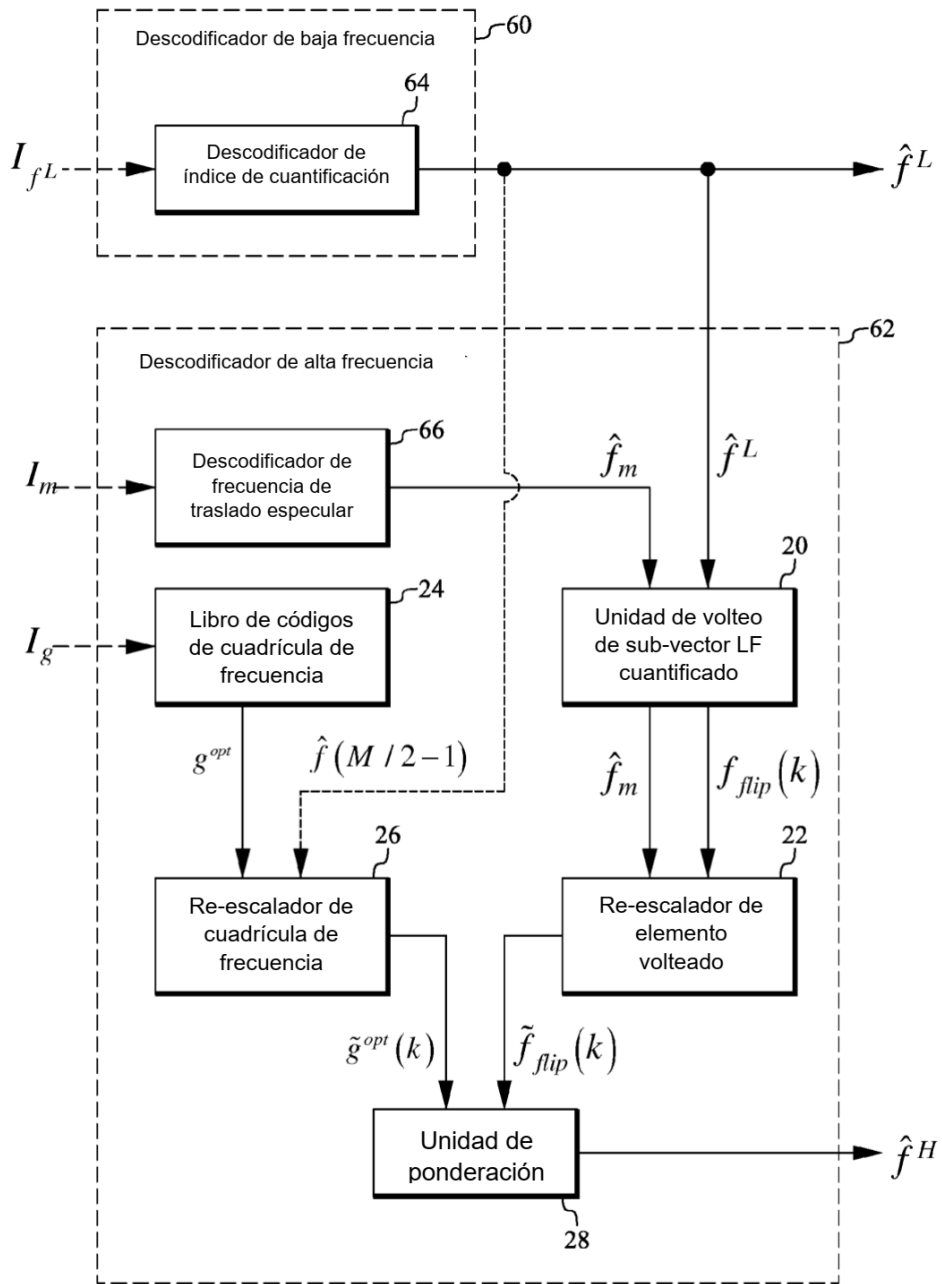


FIG. 10

50

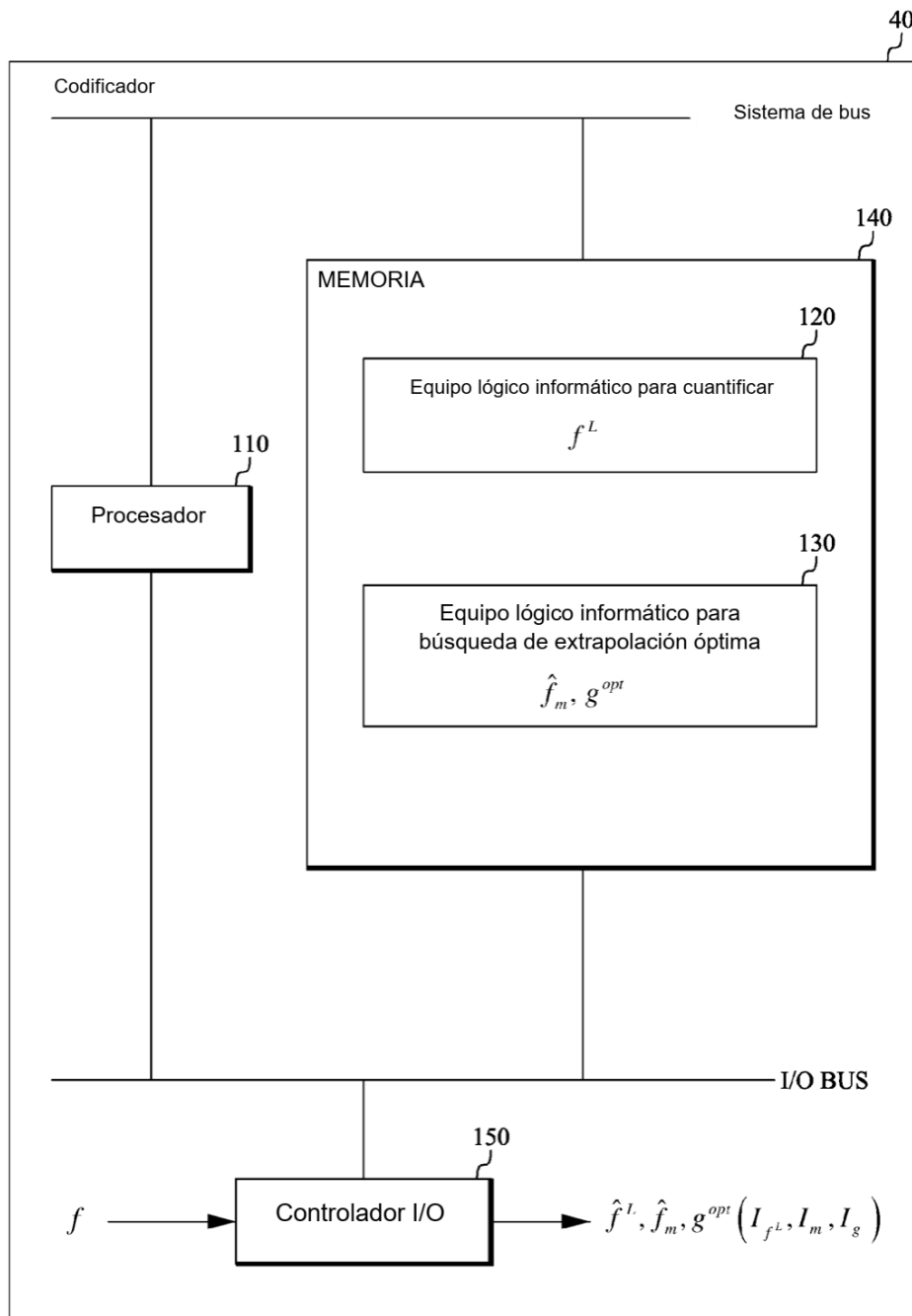


FIG. 11

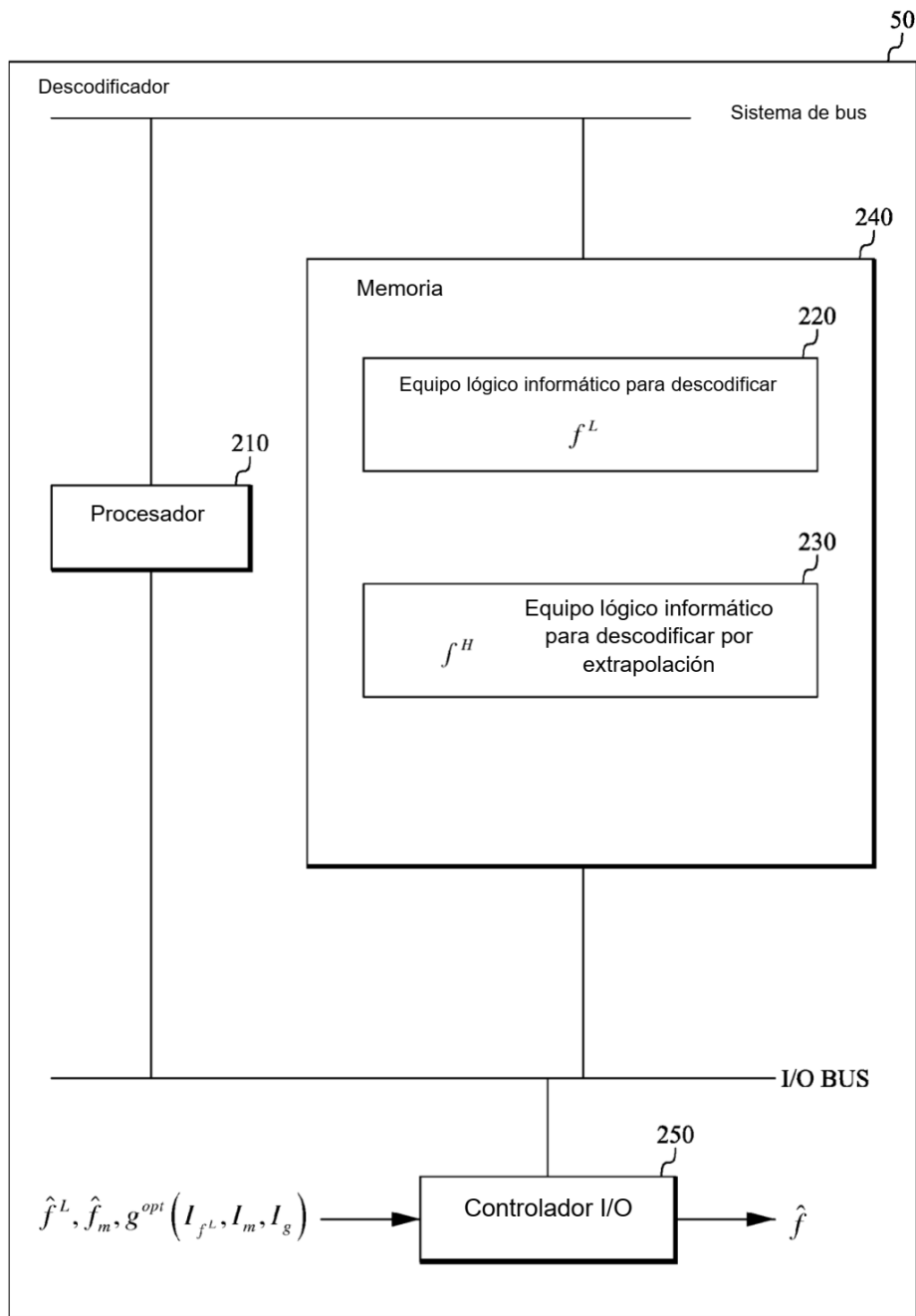


FIG. 12

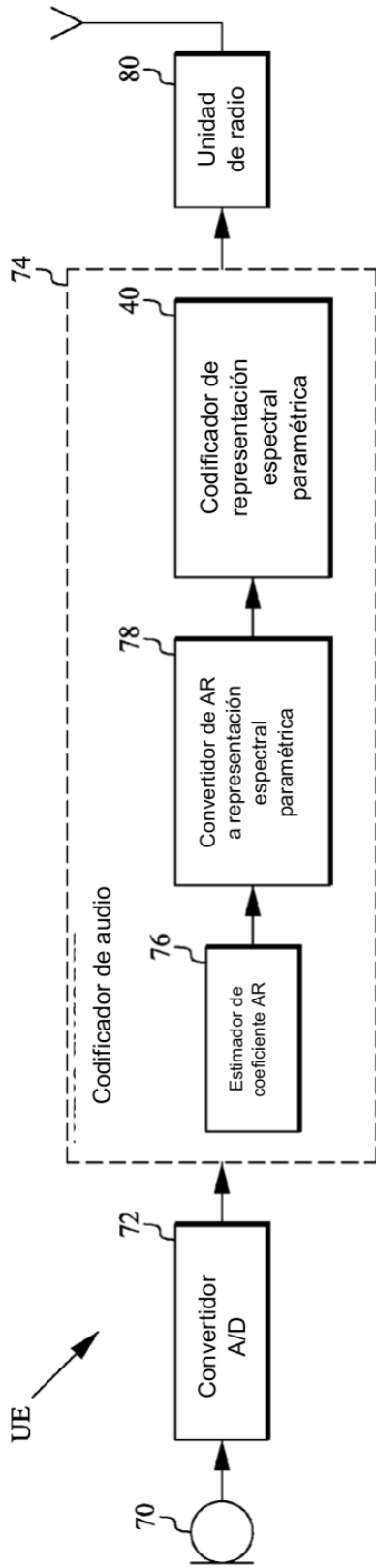


FIG. 13

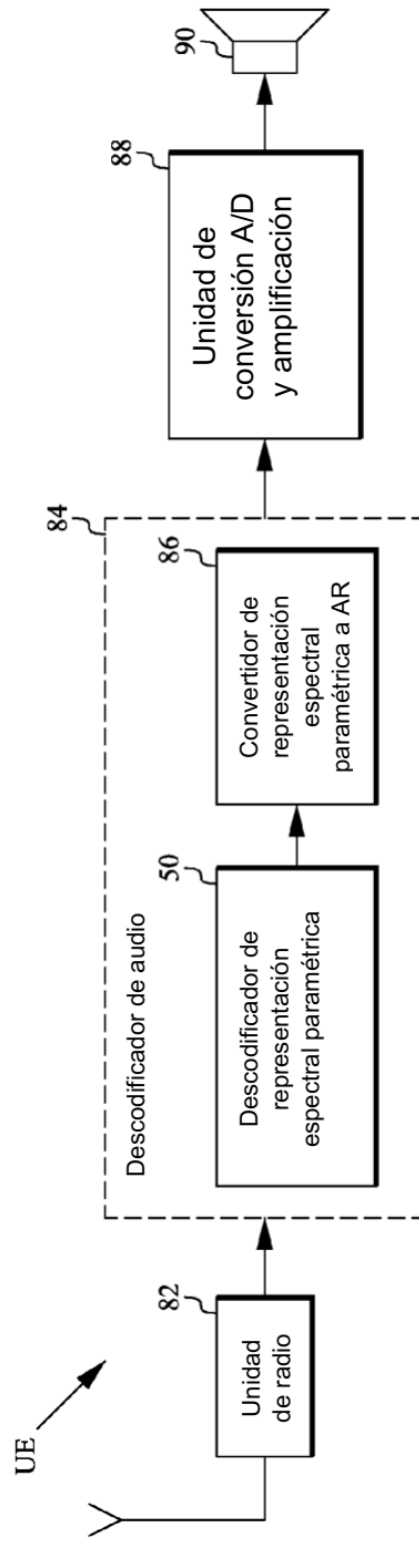


FIG. 14