

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 716 652**

51 Int. Cl.:

G10L 19/06 (2013.01)

G10L 19/16 (2013.01)

G10L 19/038 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **06.11.2014 PCT/EP2014/073960**

87 Fecha y número de publicación internacional: **21.05.2015 WO15071173**

96 Fecha de presentación y número de la solicitud europea: **06.11.2014 E 14799376 (0)**

97 Fecha y número de publicación de la concesión europea: **19.12.2018 EP 3069338**

54 Título: **Codificador para la codificación de una señal de audio, sistema de transmisión de audio y procedimiento para la determinación de valores de corrección**

30 Prioridad:

13.11.2013 EP 13192735

28.07.2014 EP 14178815

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

13.06.2019

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)**

**Hansastraße 27c
80686 München, DE**

72 Inventor/es:

**SCHMIDT, KONSTANTIN;
FUCHS, GUILLAUME;
NEUSINGER, MATTHIAS y
DIETZ, MARTIN**

74 Agente/Representante:

SALVÀ FERRER, Joan

ES 2 716 652 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

Codificador para la codificación de una señal de audio, sistema de transmisión de audio y procedimiento para la determinación de valores de corrección

5 **[0001]** La presente invención se refiere a un codificador para la codificación de una señal de audio, un procedimiento de codificación de audio, un procedimiento para la determinación de valores de corrección y un programa informático. La invención se refiere además a la frecuencia espectral de inmitancia/ponderación de frecuencia espectral de línea.

10 **[0002]** En los códecs de voz y audio de hoy en día, lo más reciente es extraer la envolvente espectral de la señal de voz o audio mediante Predicción lineal y cuantificar y codificar además una transformación de los coeficientes de predicción lineal (LPC). Tales transformaciones son, por ejemplo, las frecuencias espectrales de línea (LSF) o las frecuencias espectrales de inmitancia (ISF).

15 **[0003]** La cuantificación vectorial (VQ) se prefiere generalmente a la cuantificación escalar para la cuantificación de LPC debido al aumento del rendimiento. Sin embargo, se ha observado que una codificación de LPC óptima muestra una sensibilidad escalar diferente para cada frecuencia del vector de LSF o ISF. Como consecuencia directa, el uso de una distancia euclidiana clásica como métrica en la etapa de cuantificación
 20 conducirá a un sistema subóptimo. Se puede explicar por el hecho de que el rendimiento de una cuantificación de LPC se mide generalmente por la distancia como la Distancia espectral logarítmica (LSD) o la Distancia espectral logarítmica ponderada (WLSD) que no tienen una relación proporcional directa con la distancia euclidiana.

[0004] LSD se define como el logaritmo de la distancia euclidiana de las envolturas espectrales de los
 25 coeficientes LPC originales y la versión cuantificada de ellos. WLSD es una versión ponderada que tiene en cuenta que las frecuencias bajas son perceptualmente más relevantes que las frecuencias altas.

[0005] Tanto la LSD como la WLSD son demasiado complejas para ser computadas dentro de un esquema de cuantificación de LPC. Por lo tanto, la mayoría de los esquemas de codificación de LPC utilizan tanto la distancia
 30 euclidiana simple como una versión ponderada de la misma (WED) definida como:

$$WED = \sum_i w_i * (lsf_i - qlsf_i)^2,$$

35 en la que lsf_i es el parámetro que se va a cuantificar y $qlsf_i$ es el parámetro cuantificado, w son ponderaciones que dan más distorsión a ciertos coeficientes y menos a otros.

[0006] Laroia et al. [1] presentó una estrategia heurística conocida como media armónica inversa para computar las ponderaciones que dan más importancia a las LSF cerradas a las regiones formantes. Si dos parámetros LSF están juntos, se espera que el espectro de la señal comprenda un pico cerca de esa frecuencia. Por
 40 lo tanto, un LSF que está cerca de uno de sus vecinos tiene una alta sensibilidad escalar y se le debe dar un peso mayor:

$$w_i = \frac{1}{(lsf_i - lsf_{i-1})} + \frac{1}{(lsf_{i+1} - lsf_i)}$$

45 **[0007]** El primer y el último coeficiente de ponderación se calculan con estos pseudo LSF:

$lsf_0 = 0$ y $lsf_{p+1} = \pi$, donde p es el orden del modelo LP. El orden suele ser 10 para la señal de voz muestreada a 8 kHz y 16 para la señal de voz muestreada a 16 kHz.

50 **[0008]** Gardner y Rao [2] derivaron la sensibilidad escalar individual para LSF a partir de una aproximación de alta velocidad (por ejemplo, cuando se usa una VQ con 30 o más bits). En tal caso, los pesos derivados son óptimos y minimizan la LSD. Los pesos escalares forman la diagonal de una matriz de sensibilidad llamada dada por:

$$D_\omega(\omega) = 4\beta J_\omega^T(\omega) R_A J_\omega(\omega)$$

[0009] Donde R_A es la matriz de autocorrelación de la respuesta de impulso del filtro de síntesis $1/A(z)$ derivada de los coeficientes predictivos originales del análisis LPC. $J_\omega(\omega)$ es una matriz jacobiana que transforma LSF en coeficientes de LPC.

5 **[0010]** El principal inconveniente de esta solución es la complejidad computacional para computar la matriz de sensibilidad.

[0011] La recomendación ITU G.718 [3] amplía la estrategia de Gardner al añadir algunas consideraciones psico-acústicas, en lugar de considerar la matriz R_A , considera la respuesta al impulso de un filtro de síntesis ponderado perceptual $W(z)$:
10

$$W(z) = W_B(z)/(A(z))$$

[0012] Donde $W_B(z)$ es un filtro IIR que se aproxima al filtro de ponderación Bark (corteza), se les da más importancia a las frecuencias bajas. La matriz de sensibilidad se computa a continuación reemplazando $1/A(z)$ por $W(z)$.
15

[0013] Aunque la ponderación utilizada en G.718 es teóricamente una estrategia casi óptima, hereda de la estrategia de Gardner una complejidad muy alta. Los códecs de audio de hoy están estandarizados con una limitación en complejidad y, por lo tanto, la compensación de complejidad y la ganancia en calidad perceptual no satisfacen esta estrategia.
20

[0014] La estrategia presentada por Laroia et al. puede producir pesos sub-óptimos pero es de baja complejidad. Los pesos generados con esta estrategia tratan todo el rango de frecuencia de igual modo, aunque la sensibilidad auditiva del ser humano es altamente no lineal. La distorsión en las frecuencias más bajas es mucho más audible que la distorsión en las frecuencias más altas.
25

[0015] En "Cuantificación de vectores codificados enrejados optimizada de parámetros LSF, aplicación al codificador de voz FS1016 de 4,8 kbps" (Bouzid M. et al, Signal Processing (procesamiento de señal), Elsevier Science Publishers BV Ámsterdam, NL, vol. 85, n.º 9, 1 de septiembre de 2005, páginas 1675-1694, ISSN: 0165-1684) se describen los codificadores de voz que funcionan a velocidades de bits bajas. Los coeficientes de LPC se derivan de la señal de entrada a través del análisis de predicción lineal.
30

[0016] En "Sobre el uso de la propiedad de entrelazado intermodal LSF para la cuantificación espectral" (Mi Suk Lee et al, Speech Coding Proceedings (procedimientos de codificación de voz), 1999 IEEE Taller en Porvoo, Finlandia 20-23 de junio de 1999, Piscataway, NJ, EE.UU., IEEE, US, 20 de junio de 1999, páginas 43-45, ISBN: 978-0-7803-5651-1) se describe un procedimiento en el que las frecuencias espectrales de línea extraídas de análisis sucesivos se entrelazan entre sí.
35

40 **[0017]** El documento WO 2012/053798 A2 describe un procedimiento y un aparato para la determinación de una función de ponderación para cuantificar un coeficiente de codificación predictiva lineal (LPC). El aparato de determinación de la función de ponderación puede convertir un coeficiente de LPC de una sub-trama media de una señal de entrada a uno de un coeficiente de frecuencia espectral de inmitancia (ISF) y un coeficiente de frecuencia espectral de línea (LSF), y puede determinar una función de ponderación asociada con una importancia del coeficiente de ISF o el coeficiente de LSF basado en el coeficiente de ISF convertido o el coeficiente de LSF.
45

[0018] Por lo tanto, existe una necesidad de mejorar los esquemas de codificación.

[0019] Un objeto de la presente invención es proporcionar esquemas de codificación que permitan la complejidad computacional de los algoritmos y/o una precisión incrementada de los mismos mientras se mantiene una buena calidad de audio al decodificar la señal de audio codificada.
50

[0020] Este objeto se logra mediante un codificador según la reivindicación 1, un procedimiento para determinar los valores de corrección según la reivindicación 10, un procedimiento según la reivindicación 11 y un programa informático según la reivindicación 12.
55

[0021] Los inventores han descubierto que determinando los factores de ponderación espectral mediante el uso de un procedimiento que comprende una complejidad computacional baja y corrigiendo al menos parcialmente los factores de ponderación espectral obtenidos mediante el uso de información de corrección precalculada, los factores de ponderación espectral corregidos obtenidos pueden permitir una codificación y decodificación de la señal de audio con un bajo esfuerzo computacional mientras se mantiene la precisión de la codificación y/o reducir las distancias espectrales de línea (LSD) reducidas.
60

[0022] Según una forma de realización de la presente invención, un codificador para la codificación de una

señal de audio comprende un analizador para el análisis de la señal de audio y para la determinación de los coeficientes de predicción de análisis a partir de la señal de audio. El codificador comprende además un convertidor configurado para derivar los coeficientes de predicción convertidos a partir de los coeficientes de predicción de análisis y una memoria configurada para almacenar una multitud de valores de corrección. El codificador comprende
 5 además una calculadora y un formador de flujo de bits. La calculadora comprende un procesador, un combinador y un cuantificador, en la que el procesador está configurado para procesar la conversión predicha para obtener factores de ponderación espectral. El combinador está configurado para combinar los factores de ponderación espectral y la multitud de valores de corrección para obtener factores de ponderación corregidos. El cuantificador está configurado para cuantificar los coeficientes de predicción convertidos mediante el uso de los factores de
 10 ponderación corregidos para obtener una representación cuantificada de los coeficientes de predicción convertidos, por ejemplo, un valor relacionado con una introducción de coeficientes de predicción en una base de datos. El formador de flujo de bits está configurado para formar una señal de salida basada en una información relacionada con la representación cuantificada de los coeficientes de predicción convertidos y basada en la señal de audio. Una ventaja de esta forma de realización es que el procesador puede obtener los factores de ponderación espectral
 15 mediante el uso de procedimientos y/o conceptos que comprenden una baja complejidad computacional. Un error posiblemente obtenido con respecto a otros conceptos o procedimientos se pueden corregir al menos parcialmente mediante la aplicación de la multitud de valores de corrección. Esto permite una complejidad computacional reducida de la derivación de peso cuando se compara con una regla de determinación basada en [3] y LSD reducida en comparación con una regla de determinación según [1].

20 **[0023]** Formas de realización adicionales proporcionan un codificador, en el que el combinador está configurado para combinar los factores de ponderación espectral, la multitud de valores de corrección y una información adicional relacionada con la señal de entrada para obtener los factores de ponderación corregidos. Al utilizar la información adicional relacionada con la señal de entrada, se puede lograr una mejora adicional de los
 25 factores de ponderación corregidos obtenidos a la vez que se mantiene una baja complejidad computacional, en particular cuando la información adicional relacionada con la señal de entrada se obtiene al menos parcialmente durante otras etapas de codificación, de tal modo que la información adicional pueda ser reciclada.

30 **[0024]** Formas de realización adicionales proporcionan un codificador, en el que el combinador está configurado cíclicamente, en cada ciclo, para la obtención de los factores ponderados corregidos. La calculadora comprende una configuración más fluida para combinar de manera ponderada los primeros factores de ponderación cuantificada obtenidos para un ciclo anterior y los segundos factores de ponderación cuantificada obtenidos para un ciclo que sigue al ciclo anterior para obtener factores de ponderación corregidos fluidos que comprenden un valor entre los valores del primer y el segundo factor de ponderación cuantificada. Esto permite una reducción o una
 35 prevención de distorsiones de transición, especialmente en un caso en el que los factores de ponderación corregidos de dos ciclos consecutivos se determinan de tal modo que comprenden una amplia diferencia cuando se comparan con cada uno.

40 **[0025]** Formas de realización adicionales proporcionan un sistema de transmisión de audio que comprende un codificador y un decodificador configurado para recibir la señal de salida del codificador o una señal derivada del mismo y para decodificar la señal recibida para proporcionar una señal de audio sintetizada, en la que la señal de salida del codificador se transmite a través de un medio de transmisión, tal como un medio por cable o un medio inalámbrico. Una ventaja del sistema de transmisión de audio es que el decodificador puede decodificar la señal de salida, la señal de audio, respectivamente, en base a procedimientos sin cambios.

45 **[0026]** Formas de realización adicionales proporcionan un procedimiento para la determinación de los valores de corrección para una primera multitud de primeros factores de ponderación. Cada factor de ponderación está adaptado para ponderar una porción de una señal de audio, por ejemplo, representada como una frecuencia espectral de línea o una frecuencia espectral de inmitancia. La primera multitud de primeros factores de ponderación
 50 se determina en función de una primera regla de determinación para cada señal de audio. Se calcula una segunda multitud de segundos factores de ponderación para cada señal de audio del conjunto de señales de audio basándose en una segunda regla de determinación. Cada una de la segunda multitud de los factores de ponderación se relaciona con un primer factor de ponderación, es decir, se puede determinar un factor de ponderación para una porción de la señal de audio en base a la primera regla de determinación y en base a la
 55 segunda regla de determinación para obtener dos resultados que pueden ser diferentes. Se calcula una tercera multitud de valores de distancia, teniendo los valores de distancia un valor relacionado con una distancia entre un primer factor de ponderación y un segundo factor de ponderación, ambos relacionados con la porción de la señal de audio. Se calcula una cuarta multitud de valores de corrección adaptados para reducir los valores de distancia cuando se combinan con los primeros factores de ponderación, de tal modo que cuando los primeros factores de
 60 ponderación se combinan con la cuarta multitud de valores de corrección, se reduce una distancia entre los primeros factores de ponderación corregidos cuando se compara con los segundos factores de ponderación. Esto permite computar los factores de ponderación basados en un conjunto de datos de entrenamiento una vez en base a la segunda regla de determinación que comprende una alta complejidad computacional y/o una alta precisión y otro tiempo basado en la primera regla de determinación que puede comprender una complejidad computacional más
 65 baja y puede ser una precisión más baja, en donde la precisión más baja y/o compensada o reducida al menos

parcialmente por corrección.

[0027] Formas de realización adicionales proporcionan un procedimiento en el que la distancia se reduce adaptando un polinomio, en el que los coeficientes polinomiales se relacionan con los valores de corrección. Formas de realización adicionales proporcionan un programa informático.

[0028] Las formas de realización preferidas de la presente invención se describirán en detalle haciendo referencia a las figuras adjuntas en las que:

- 10 La figura 1 muestra un diagrama de bloques esquemático de un codificador para la codificación de una señal de audio según una forma de realización;
 La figura 2 muestra un diagrama de bloques esquemático de una calculadora según una forma de realización en la que la calculadora se modifica cuando se compara con una calculadora mostrada en la figura 1;
 La figura 3 muestra un diagrama de bloques esquemático de un codificador que comprende adicionalmente un analizador espectral y un procesador espectral según una forma de realización;
 15 La Fig. 4a ilustra un vector que comprende 16 valores de frecuencias espectrales de línea que se obtienen mediante un convertidor basado en los coeficientes de predicción determinados según una forma de realización;
 La figura 4b ilustra una regla de determinación ejecutada por un combinador según una forma de realización;
 La figura 4c muestra una regla de determinación ejemplar para ilustrar la etapa de obtención de factores de ponderación corregidos según una forma de realización;
 20 La figura 5a representa un esquema de determinación ejemplar que puede implementarse mediante un cuantificador para determinar una representación cuantificada de los coeficientes de predicción convertidos según una forma de realización;
 La Fig. 5b muestra un vector ejemplar de valores de cuantificación que pueden combinarse en conjuntos de los mismos según una forma de realización;
 25 La figura 6 muestra un diagrama de bloques esquemático de un sistema de transmisión de audio según una forma de realización;
 La figura 7 ilustra una forma de realización de la derivación de los valores de corrección; y
 La figura 8 muestra un diagrama de flujo esquemático de un procedimiento para la codificación de una señal de audio según una forma de realización.

[0029] Elementos iguales o equivalentes o elementos con funcionalidad igual o equivalente se indican en la siguiente descripción con números de referencia iguales o equivalentes, incluso si aparecen en diferentes figuras.

- 35 **[0030]** En la siguiente descripción, se establece una pluralidad de detalles para proporcionar una explicación más completa de las formas de realización de la presente invención. Sin embargo, será evidente para los expertos en la técnica que las formas de realización de la presente invención pueden ponerse en práctica sin estos detalles específicos. En otros casos, se muestran estructuras y dispositivos bien conocidos en forma de diagrama de bloques en lugar de en detalle para evitar oscurecer las formas de realización de la presente invención. Además, las características de las diferentes formas de realización descritas en lo sucesivo se pueden combinar entre sí, a menos que se indique específicamente lo contrario.

[0031] La figura 1 muestra un diagrama de bloques esquemático de un codificador 100 para codificar una señal de audio. La señal de audio puede ser obtenida por el codificador 100 como una secuencia de tramas 102 de la señal de audio. El codificador 100 comprende un analizador para analizar la trama 102 y para determinar los coeficientes de predicción de análisis 112 a partir de la señal de audio 102. Los coeficientes de predicción de análisis (coeficientes de predicción) 112 se pueden obtener, por ejemplo, como coeficientes de predicción lineal (LPC). Alternativamente, también se pueden obtener coeficientes de predicción no lineal, en los que los coeficientes de predicción lineales se pueden obtener utilizando una menor potencia computacional y, por lo tanto, se pueden obtener más rápido.

[0032] El codificador 100 comprende un convertidor 120 configurado para derivar los coeficientes de predicción convertidos 122 de los coeficientes de predicción 112. El convertidor 120 puede estar configurado para determinar los coeficientes de predicción convertidos 122 para obtener, por ejemplo, frecuencias espectrales de línea (LSF) y/o frecuencias espectrales de inmitancia (ISF). Los coeficientes de predicción convertidos 122 pueden comprender una mayor robustez con respecto a los errores de cuantificación en una cuantificación posterior cuando se comparan con los coeficientes de predicción 112. Como la cuantificación generalmente se realiza de forma no lineal, la cuantificación de los coeficientes de predicción lineal puede conducir a distorsiones de una señal de audio descodificada.

[0033] El codificador 100 comprende una calculadora 130. La calculadora 130 comprende un procesador 140 que está configurado para procesar los coeficientes de predicción convertidos 122 para obtener factores de ponderación espectral 142. El procesador puede estar configurado para calcular y/o determinar los factores de ponderación 142 en función de una o más de una pluralidad de reglas de determinación conocidas tal como una media armónica inversa (IHM) como se conoce por [1] o según una estrategia más compleja como se describe en

[2]. La Norma G.718 de la Unión Internacional de Telecomunicaciones (ITU) describe una estrategia adicional para determinar los factores de ponderación al ampliar la estrategia de [2] como se describe en [3]. Preferiblemente, el procesador 140 está configurado para determinar los factores de ponderación 142 en base a una regla de determinación que comprende una baja complejidad computacional. Esto puede permitir un alto rendimiento de las 5 señales de audio codificadas y/o una simple realización del codificador 100 debido al hardware que puede consumir menos energía en función de menos esfuerzos computacionales.

[0034] La calculadora 130 comprende un combinador 150 configurado para combinar los factores de ponderación espectral 142 y una multitud de valores de corrección 162 para obtener factores de ponderación 10 corregidos 152. La multitud de valores de corrección se proporciona desde una memoria 160 en la que se almacenan los valores de corrección 162. Los valores de corrección 162 pueden ser estáticos o dinámicos, es decir, los valores de corrección 162 pueden actualizarse durante el funcionamiento del codificador 100 o pueden permanecer sin cambios durante el funcionamiento y/o pueden actualizarse solo durante un procedimiento de calibración para calibrar el codificador 100. Preferiblemente, la memoria 160 comprende valores de corrección 15 estática 162. Los valores de corrección 162 se pueden obtener, por ejemplo, mediante un procedimiento de precálculo como se describe más adelante. Alternativamente, la memoria 160 puede estar comprendida alternativamente por la calculadora 130 como se indica por las líneas de puntos.

[0035] La calculadora 130 comprende un cuantificador 170 configurado para cuantificar los coeficientes de predicción convertidos 122 mediante el uso de los factores de ponderación corregidos 152. El cuantificador 170 está 20 configurado para generar una representación cuantificada 172 de los coeficientes de predicción convertidos 122. El cuantificador 170 puede ser un cuantificador lineal, un cuantificador no lineal tal como un cuantificador logarítmico o un cuantificador similar a un vector, un cuantificador de vector respectivamente. Un cuantificador similar a un vector puede configurarse para cuantificar una pluralidad de porciones pf de los factores de ponderación corregidos 152 en 25 una pluralidad de valores cuantificados (porciones). El cuantificador 170 puede estar configurado para ponderar los coeficientes de predicción convertidos 122 con los factores de ponderación corregidos 152. El cuantificador puede estar configurado además para determinar una distancia de los coeficientes de predicción convertidos ponderados 122 a las entradas de una base de datos del cuantificador 170 y para seleccionar una palabra de código (representación) que se relaciona con una entrada en la base de datos en la que la entrada puede comprender una 30 distancia más baja a los coeficientes de predicción convertidos ponderados 122. Tal procedimiento se describe de manera ejemplar más adelante. El cuantificador 170 puede ser un vector cuantificador estocástico (VQ). Alternativamente, el cuantificador 170 también puede estar configurado para aplicar otros cuantificadores vectoriales como Lattice VQ o cualquier cuantificador de escalador. Alternativamente, el cuantificador 170 también puede estar 35 configurado para aplicar una cuantificación lineal o logarítmica.

[0036] La representación cuantificada 172 de los coeficientes de predicción convertidos 122, es decir, la palabra de código, se proporciona a un formador de flujo de bits 180 del codificador 100. El codificador 100 puede comprender una unidad de procesamiento de audio 190 configurada para procesar parte o toda la información de 40 audio de la señal de audio 102 y/o información adicional. La unidad de procesamiento de audio 190 está configurada para proporcionar datos de audio 192 tal como información de señal sonora o información de señal no sonora al formador de flujo de bits 180. El formador de flujo de bits 180 está configurado para formar una señal de salida (flujo de bits) 182 en función de la representación cuantificada 172 de los coeficientes de predicción convertidos 122 y en base a la información de audio 192, que se basa en la señal de audio 102.

[0037] Una ventaja del codificador 100 es que el procesador 140 puede estar configurado para obtener, es decir, calcular, los factores de ponderación 142 mediante el uso de una regla de determinación que comprende una complejidad computacional baja. Los valores de corrección 162 se pueden obtener, cuando se expresan de manera 45 simplificada, comparando un conjunto de factores de ponderación obtenidos por una regla de determinación (de referencia) con una alta complejidad computacional pero que comprenden, por lo tanto, una alta precisión y/o una buena calidad de audio y/o una LSD baja con factores de ponderación obtenidos por la regla de determinación ejecutada por el procesador 140. Esto se puede hacer para una multitud de señales de audio, en la que para cada una de las señales de audio se obtiene un número de factores de ponderación basado en ambas reglas de determinación. Para cada señal de audio, los resultados obtenidos pueden compararse para obtener una información relacionada con una falta de coincidencia o un error. La información relacionada con la discordancia o el 55 error se puede resumir y/o promediar con respecto a la multitud de señales de audio para obtener una información relacionada con un error promedio que realiza el procesador 140 con respecto a la regla de determinación de referencia cuando se ejecuta la regla de determinación con la menor complejidad computacional. La información obtenida relacionada con el error promedio y/o el desajuste puede representarse en los valores de corrección 162, de manera que los factores de ponderación 142 pueden combinarse con los valores de corrección 162 mediante el 60 combinador para reducir o compensar el error promedio. Esto permite reducir o casi compensar el error de los factores de ponderación 142 cuando se compara con la regla de determinación de referencia utilizada fuera de línea, al tiempo que permite una determinación menos compleja de los factores de ponderación 142.

[0038] La figura 2 muestra un diagrama de bloques esquemático de una calculadora modificada 130'. La 65 calculadora 130' comprende un procesador 140' configurado para calcular los pesos de la media armónica inversa

(IHM) a partir del LSF 122', que representan los coeficientes de predicción convertidos. La calculadora 130' comprende un combinador 150' que, cuando se compara con el combinador 150, está configurado para combinar los pesos IHM 142' del procesador 140', los valores de corrección 162 y una información adicional 114 de la señal de audio 102 indicada como "coeficientes de reflexión", en los que la información adicional 114 no se limita a los mismos. La información adicional puede ser un resultado provisional de otras etapas de codificación, por ejemplo, los coeficientes de reflexión 114 pueden ser obtenidos por el analizador 110 durante la determinación de los coeficientes de predicción 112 como se describe en la Fig. 1. Los coeficientes de predicción lineal pueden ser determinados por el analizador 110 cuando ejecuta una regla de determinación según el algoritmo de Levinson-Durbin en el que se determinan los algoritmos de reflexión. Una información relacionada con el espectro de potencia también se puede obtener durante el cálculo de los coeficientes de predicción 112. Una posible implementación del combinador 150' se describe más adelante. Alternativamente, o además, la información adicional 114 se puede combinar con los pesos 142 o 142' y los parámetros de corrección 162, por ejemplo, información relacionada con un espectro de potencia de la señal de audio 102. La información adicional 114 permite reducir además una diferencia entre los pesos 142 o 142' determinada por la calculadora 130 o 130' y los pesos de referencia. Un aumento de la complejidad computacional solo puede tener efectos menores, ya que la información adicional 114 ya puede estar determinada por otros componentes tales como el analizador 110 durante otras etapas de la codificación de audio.

[0039] La calculadora 130' comprende además un suavizado 155 configurado para recibir los factores de ponderación corregidos 152' del combinador 150' y una información opcional 157 (bandera de control) que permite el control del funcionamiento (estado ENCENDIDO/APAGADO) del suavizado 155. La bandera de control 157 se puede obtener, por ejemplo, a partir del analizador, lo que indica que se debe realizar un suavizado con el fin de reducir las transiciones difíciles. El suavizado 155 está configurado para combinar los factores de ponderación corregidos 152' y los factores de ponderación corregidos 152'', que son una representación retardada de los factores de ponderación corregidos determinados para una trama o sub-trama anterior de la señal de audio, es decir, los factores de ponderación corregidos determinados en un ciclo previo en estado ENCENDIDO. El suavizado 155 puede implementarse como un filtro de respuesta de impulso infinito (IIR). Por lo tanto, la calculadora 130' comprende un bloque de retraso 159 configurado para recibir y retrasar los factores de ponderación corregidos 152'' proporcionados por el suavizado 155 en un primer ciclo y para proporcionar esos pesos como los factores de ponderación corregidos 152'' en un ciclo siguiente.

[0040] El bloque de retraso 159 puede implementarse, por ejemplo, como un filtro de retraso o como una memoria configurada para almacenar los factores de ponderación corregidos recibidos 152''. El suavizado 155 está configurado para combinar de manera ponderada los factores de ponderación corregidos recibidos 152' y los factores de ponderación corregidos recibidos 152'' del pasado. Por ejemplo, los factores de ponderación corregidos (presentes) 152' pueden comprender un porcentaje del 25%, 50%, 75% o cualquier otro valor en los factores de ponderación corregidos suavizados 152'', en el que los factores de ponderación (pasados) 152'' pueden comprender una parte de (1 parte de los factores de ponderación corregidos 152'). Esto permite evitar transiciones bruscas entre las tramas de audio subsiguientes cuando la señal de audio, es decir, dos tramas subsiguientes de la misma, resultan en diferentes factores de ponderación corregidos que conducirían a distorsiones en una señal de audio descodificada. En el estado de apagado, el suavizado 155 está configurado para enviar los factores de ponderación corregidos 152'. Alternativamente, o además, el suavizado puede permitir una calidad de audio incrementada para las señales de audio que comprenden un alto nivel de periodicidad.

[0041] Alternativamente, el suavizado 155 puede estar configurado para combinar adicionalmente factores ponderados corregidos de más ciclos anteriores. Alternativamente, o además, los coeficientes de predicción convertidos 122' también pueden ser las frecuencias espectrales de inmitancia.

[0042] Se puede obtener un factor de ponderación w_i , por ejemplo, basado en la media armónica inversa (IHM). Una regla de determinación puede estar basada en una forma:

$$w_i = \frac{1}{(lsf_i - lsf_{i-1})} + \frac{1}{(lsf_{i+1} - lsf_i)}$$

en la que w_i indica un peso determinado 142' con índice i , LSF_i indica una frecuencia espectral de línea con índice i . El índice i corresponde a un número de factores de ponderación espectral obtenidos y puede ser igual a un número de coeficientes de predicción determinados por el analizador. El número de coeficientes de predicción y , por lo tanto, el número de coeficientes convertidos puede ser, por ejemplo, 16. Alternativamente, el número también puede ser 8 o 32. Alternativamente, el número de coeficientes convertidos también puede ser menor que el número de coeficientes de predicción, por ejemplo, si los coeficientes convertidos 122 se determinan como frecuencias espectrales de inmitancia que pueden comprender un número menor en comparación con el número de coeficientes de predicción.

[0043] En otras palabras, la Fig. 2 detalla el procesamiento realizado en la etapa de derivación del peso

ejecutado por el convertidor 120. Primero, los pesos IHM se computan a partir de las LSF. Según una forma de realización, se utiliza un orden de LPC de 16 para una señal muestreada a 16 kHz. Eso significa que las LSF están limitadas entre 0 y 8 kHz. Según una forma de realización adicional, el LPC es del orden 16 y la señal se muestrea a 12.8 kHz. En ese caso, las LSF están limitadas entre 0 y 6,4 kHz. Según una forma de realización adicional, la señal se muestrea a 8 kHz, lo que puede denominarse un muestreo de banda estrecha. Los pesos de IHM se pueden combinar entonces con información adicional, por ejemplo, en relación con algunos de los coeficientes de reflexión, dentro de un polinomio para el cual los coeficientes se optimizan fuera de línea durante una fase de entrenamiento. Finalmente, los pesos obtenidos pueden ser suavizados por el conjunto anterior de pesos en ciertos casos, por ejemplo, para señales estacionarias. Según una forma de realización, el suavizado nunca se realiza. Según otras formas de realización, se realiza solo cuando la trama de entrada se clasifica como sonora, es decir, la señal detectada es altamente periódica.

[0044] A continuación, se hará referencia a los detalles de la corrección de los factores de ponderación derivados. Por ejemplo, el analizador está configurado para determinar los coeficientes de predicción lineal (LPC) del orden 10 o 16, es decir, un número de 10 o 16 LPC. Aunque el analizador también puede configurarse para determinar cualquier otro número de coeficientes de predicción lineal o un tipo diferente de coeficiente, la siguiente descripción se hace con referencia a 16 coeficientes, ya que este número de coeficientes se utiliza en la comunicación móvil.

[0045] La figura 3 muestra un diagrama de bloques esquemático de un codificador 300 que comprende adicionalmente un analizador espectral 115 y un procesador espectral 145 que comprende cuando se compara con el codificador 100. El analizador espectral 115 está configurado para derivar los parámetros espectrales 116 de la señal de audio 102. Los parámetros espectrales pueden ser, por ejemplo, una curva de envolvente de un espectro de la señal de audio o de una trama de la misma y/o parámetros que caracterizan la curva de envolvente. Alternativamente, se pueden obtener coeficientes relacionados con el espectro de potencia.

[0046] El procesador espectral 145 comprende una calculadora de energía 145a que está configurada para computar una cantidad o una medida 146 para una energía de bins de frecuencia del espectro de la señal de audio 102 basada en los parámetros espectrales 116. El procesador espectral comprende además un normalizador 145b para normalizar los coeficientes de predicción convertidos 122' (LSF) para obtener coeficientes de predicción normalizados 147. Los coeficientes de predicción convertidos pueden normalizarse, por ejemplo, relativamente, con respecto a un valor máximo de una pluralidad de LSF y/o absolutamente, es decir, con respecto a un valor predeterminado tal como un valor máximo esperado o representable por las variables de computación utilizadas.

[0047] El procesador espectral 145 comprende además un primer determinador 145c configurado para determinar una energía de bin para cada parámetro de predicción normalizado, es decir, para relacionar cada parámetro de predicción normalizado 147 obtenido del normalizador 145b con una medida computada 146 para obtener un vector W1 que contiene la energía de bin para cada LSF. El procesador espectral 145 comprende además un segundo determinador 145d configurado para encontrar (determinar) una ponderación de frecuencia para cada LSF normalizada para obtener un vector W2 que comprende las ponderaciones de frecuencia. La información adicional 114 comprende los vectores W1 y W2, es decir, los vectores W1 y W2 son la característica que representa la información adicional 114.

[0048] El procesador 142' está configurado para determinar la IHM en función de los parámetros de predicción convertidos 122' y una potencia de IHM, por ejemplo, la segunda potencia, en la que alternativamente o además también se puede computar una potencia mayor, en la que la IHM y la potencia(s) de los mismos forman los factores de ponderación 142'.

[0049] Un combinador 150' está configurado para determinar los factores de ponderación corregidos (ponderaciones de LSF corregidas) 152' en función de la información adicional 114 y los factores de ponderación 142'.

[0050] Alternativamente, el procesador 140', el procesador espectral 145 y/o el combinador pueden implementarse como una unidad de procesamiento única tal como una unidad de procesamiento central, un (micro-) controlador, una matriz de puertas programables o similares.

[0051] En otras palabras, una primera y una segunda entrada al combinador son IHM e IHM², es decir, los factores de ponderación 142'. Una tercera entrada es para cada elemento de vector LSF i:

$$wfft_i = (\sqrt{wfft_i - \min} + 2) * FreqWTable[normLsf_i]$$

en la que *wfft* es la combinación de W1 y W2 y en la que *min* es el mínimo de *wfft*.

[0052] $i = 0.. M$ donde *M* puede ser 16 cuando 16 coeficientes de predicción se derivan de la señal de audio y

$$wfft_i = 10 * \log_{10}(\max(\text{binEner} \left[\left[\frac{lsfi}{50} + 0.5 \right] - 1 \right], \text{binEner} \left[\left[\frac{lsfi}{50} + 0.5 \right], \text{binEner} \left[\left[\frac{lsfi}{50} + 0.5 \right] + 1 \right] \right))$$

en la que *binEner* contiene la energía de cada bin del espectro, es decir, *binEner* corresponde a la medida 146.

5

[0053] El mapeo *binEner* $[\lfloor lsfi/50 + 0.5 \rfloor]$ es una aproximación aproximada de la energía de a. que forma en la envoltura espectral. *FreqWTable* es un vector que contiene pesos adicionales que se seleccionan dependiendo de la señal de entrada que es sonora o no.

10 **[0054]** Wfft es una aproximación de la energía espectral cercana a un coeficiente de predicción como un coeficiente de LSF. En términos simples, si un coeficiente de predicción (LSF) comprende un valor X, esto significa que el espectro de la señal de audio (trama) comprende un máximo de energía (formante) en la Frecuencia X o debajo de la misma. La wfft es una expresión logarítmica de la energía en la frecuencia X, es decir, corresponde a la energía logarítmica en esta ubicación. Cuando se compara con formas de realización descritas anteriormente que
 15 utilizan coeficientes de reflexión como información adicional, alternativamente, o además, se puede usar una combinación de wfft (W1) y *FreqWTable* (W2) para obtener la información adicional 114. *FreqWTable* describe una de una pluralidad de tablas posibles que se pueden utilizar. Basándose en un "modo de codificación" del codificador 300, por ejemplo, sonoro, fricativo o similar, se puede seleccionar al menos una de la pluralidad de tablas. Una o más de la pluralidad de tablas pueden entrenarse (programarse y adaptarse) durante el funcionamiento del
 20 codificador 300.

[0055] Un hallazgo del uso de wfft es mejorar la codificación de los coeficientes de predicción convertidos que representan un formante. En contraste con la formación de ruido clásica en la que el ruido está en frecuencias que comprenden grandes cantidades de energía (señal), la estrategia descrita se relaciona con la cuantificación de la
 25 curva de envolvente espectral. Cuando el espectro de potencia comprende una gran cantidad de energía (una gran medida) en frecuencias que comprenden o están dispuestas adyacentes a una frecuencia de un coeficiente de predicción convertido, este coeficiente de predicción convertido (LSF) puede cuantificarse mejor, es decir, con errores más bajos logrados por mayores ponderaciones, que otros coeficientes que comprenden una menor medida de energía.

30

[0056] La Fig. 4a ilustra un vector LSF que comprende 16 valores de entradas de las frecuencias espectrales de línea determinadas que se obtienen mediante el convertidor en función de los coeficientes de predicción determinados. El procesador está configurado para obtener también 16 pesos, por ejemplo, las medias armónicas inversas IHM representadas en un vector IHM. Los valores de corrección 162 se agrupan, por ejemplo, en un vector
 35 a, un vector b y un vector c. Cada uno de los vectores a, b y c comprende 16 valores a_{1-16} , b_{1-16} y c_{1-16} , en los que índices iguales indican que el valor de corrección respectivo está relacionado con un coeficiente de predicción, una representación convertida de los mismos y un factor de ponderación que comprende el mismo índice La Fig. 4b ilustra una regla de determinación ejecutada por el combinador 150 o 150' según una forma de realización. El combinador está configurado para computar o determinar un resultado para una función polinomial basada en una
 40 forma $y = a + bx + cx^2$, es decir, diferentes valores de corrección a, b, c se combinan (multiplican) con diferentes potencias de los factores de ponderación (ilustrados como x). y denota un vector de factores de ponderación corregidos obtenidos.

[0057] Alternativamente, o además, el combinador también puede configurarse para añadir valores de
 45 corrección adicionales (d, e, f, ...) y potencias adicionales de los factores de ponderación o de la información adicional. Por ejemplo, el polinomio representado en la Fig. 4b puede extenderse por un vector d que comprende 16 valores que se multiplican con una tercera potencia de la información adicional 114, un vector respectivo que comprende también 16 valores. Este puede ser, por ejemplo, un vector basado en IHM³ cuando el procesador 140' como se describe en la Fig. 3 está configurado para determinar potencias adicionales de IHM. Alternativamente, solo
 50 se pueden computar al menos el vector b y opcionalmente uno o más de los vectores de orden superior c, d, ... Simplificado el orden del polinomio aumenta con cada término, en el que cada tipo puede formarse basándose en el factor de ponderación y/u opcionalmente basándose en la información adicional, en la que el polinomio se basa en la forma $y = a + bx + cx^2$ también cuando comprende un término de orden superior. Los valores de corrección a, b, c y opcionalmente d, e, ... pueden comprender valores reales y/o imaginarios y también pueden comprender un valor de

cero.

[0058] La Fig. 4c representa una regla de determinación ejemplar para ilustrar la etapa de obtención de los factores de ponderación corregidos 152 o 152'. Los factores de ponderación corregidos se representan en un vector \underline{w} que comprende 16 valores, un factor de ponderación para cada uno de los coeficientes de predicción convertidos representados en la Fig. 4a. Cada uno de los factores de ponderación corregidos w_{1-16} se computa según la regla de determinación que se muestra en la Fig. 4b. Las descripciones anteriores solo deben ilustrar un principio de determinación de los factores de ponderación corregidos y no deben limitarse a las reglas de determinación descritas anteriormente. Las reglas de determinación descritas anteriormente también pueden ser variadas, escaladas, modificadas o similares. En general, los factores de ponderación corregidos se obtienen realizando una combinación de los valores de corrección con los factores de ponderación determinados.

[0059] La Fig. 5a representa un esquema de determinación ejemplar que puede ser implementado por un cuantificador tal como el cuantificador 170 para determinar la representación cuantificada de los coeficientes de predicción convertidos. El cuantificador puede resumir un error, por ejemplo, una diferencia o una potencia de la misma entre un coeficiente convertido determinado mostrado como LSF_i y un coeficiente de referencia indicado como LSF_i^r , en donde los coeficientes de referencia pueden almacenarse en una base de datos del cuantificador. La distancia determinada puede ser cuadrada de modo que solo se obtengan valores positivos. Cada una de las distancias (errores) está ponderada por un factor de ponderación respectivo w_i . Esto permite dar rangos de frecuencia o coeficientes de predicción convertidos con una mayor importancia para la calidad de audio, un mayor peso y rangos de frecuencia con una menor importancia para la calidad de audio y un menor peso. Los errores se resumen sobre algunos o todos los índices 1-16 para obtener un valor de error total. Esto se puede hacer para una pluralidad de combinaciones predefinidas (entradas de la base de datos) de coeficientes que se pueden combinar para los conjuntos Qu^1, Qu^2, \dots, Qu^n como se indica en la Fig. 5b. El cuantificador puede configurarse para seleccionar una palabra de código relacionada con un conjunto de coeficientes predefinidos que comprenden un error mínimo con respecto a los factores ponderados corregidos determinados y los coeficientes de predicción convertidos. La palabra de código puede ser, por ejemplo, un índice de una tabla de tal forma que un decodificador pueda restaurar el conjunto predefinido Qu^1, Qu^2, \dots basado en el índice recibido, la palabra de código recibida, respectivamente.

[0060] Para obtener los valores de corrección durante una fase de entrenamiento, se selecciona una regla de determinación de referencia según la cual se determinan los pesos de referencia. Dado que el codificador está configurado para corregir determinados factores de ponderación con respecto a los pesos de referencia y la determinación de los pesos de referencia se puede hacer fuera de línea, es decir, durante una etapa de calibración o similar, una regla de determinación que comprende una alta precisión (por ejemplo, LSD bajo) puede ser seleccionada mientras se descuida el esfuerzo computacional resultante. Preferiblemente, un procedimiento que comprende una alta precisión y tal vez una alta complejidad de cómputo puede seleccionarse para obtener factores de ponderación de referencia de tamaño pre-dimensionado. Por ejemplo, se puede usar un procedimiento para determinar los factores de ponderación según el Estándar G.718 [3].

[0061] También se ejecuta una regla de determinación según la cual el codificador determinará los factores de ponderación. Este puede ser un procedimiento que comprende una complejidad computacional baja mientras que acepta una menor precisión de los resultados determinados. Los pesos se computan según ambas reglas de determinación al usar un conjunto de material de audio que comprende, por ejemplo, voz y/o música. El material de audio se puede representar en varios vectores de entrenamiento M , en los que M puede comprender un valor de más de 100, más de 1.000 o más de 5.000. Ambos conjuntos de factores de ponderación obtenidos se almacenan en una matriz, comprendiendo cada matriz vectores que están relacionados cada uno con uno de los vectores de entrenamiento M .

[0062] Para cada uno de los vectores de entrenamiento M , se determina una distancia entre un vector que comprende los factores de ponderación determinados en función de la primera regla de determinación (de referencia) y un vector que comprende los vectores de ponderación determinados en función de la regla de determinación del codificador. Las distancias se suman para obtener una distancia total (error), en la que el error total se puede promediar para obtener un valor de error promedio.

[0063] Durante la determinación de los valores de corrección, un objetivo puede ser reducir el error total y/o el error promedio. Por lo tanto, un ajuste polinomial puede ejecutarse en base a la regla de determinación mostrada en la Fig. 4b, en la que los vectores $\underline{a}, \underline{b}, \underline{c}$ y/o vectores adicionales están adaptados al polinomio de tal manera que el error total y/o promedio se reduzca o minimice. El polinomio se ajusta a los factores de ponderación determinados en función de la regla de determinación, que se ejecutará en el decodificador. El polinomio puede ajustarse de tal manera que el error total o el error promedio esté por debajo de un valor de umbral, por ejemplo, 0,01, 0,1 o 0,2, en el que 1 indica una falta de coincidencia total. Alternativamente, o además, el polinomio puede ajustarse de tal manera que el error total se minimice utilizando un algoritmo de minimización de errores. Un valor de 0,01 puede indicar un error relativo que puede expresarse como una diferencia (distancia) y/o como un cociente de distancias. Alternativamente, el ajuste polinomial se puede hacer determinando los valores de corrección de tal manera que el

error total resultante o el error promedio comprenda un valor que esté cerca de un mínimo matemático. Esto se puede hacer, por ejemplo, mediante la derivación de las funciones utilizadas y una optimización basada en establecer la derivación obtenida en cero.

5 **[0064]** Se puede lograr una reducción adicional de la distancia (error), por ejemplo, la distancia euclidiana, al añadir la información adicional, como se muestra para 114 en el lado del codificador. Esta información adicional también se puede utilizar durante el cálculo de los parámetros de corrección. La información se puede usar combinando la misma con el polinomio para determinar el valor de corrección.

10 **[0065]** En otras palabras, primero los pesos IHM y los pesos G.718 se pueden extraer de una base de datos que contiene más de 5.000 segundos (o vectores de entrenamiento M) de material de voz y música. Los pesos IHM pueden almacenarse en la matriz I y los pesos G.718 pueden almacenarse en la matriz G . Sean I_i y G_i vectores que contienen todos los pesos IHM y G.718 W_i del i -th ISF o el coeficiente LSF de toda la base de datos de formación. La distancia euclidiana promedio entre estos dos vectores se puede determinar basada en:

15

$$d_i = \frac{1}{M} \sum^M (I_i - G_i)^2$$

[0066] Para minimizar la distancia entre estos dos vectores, se puede ajustar un polinomio de segundo orden:

20

$$d_i = \frac{1}{M} \sum^M (p_{0,i} + p_{1,i} I_i + p_{2,i} I_i^2 - G_i)^2$$

$$EI_i = \begin{bmatrix} 1 & I_{1,i} & I_{1,i}^2 \\ 1 & I_{2,i} & I_{2,i}^2 \\ \vdots & \vdots & \vdots \end{bmatrix}$$

[0067] Puede ser introducida una matriz EI_i y un vector $P_i = [p_{0,i} \ p_{1,i} \ p_{2,i}]^T$ con el fin de reescribir:

25

$$p_{0,i} + p_{1,i} I_i + p_{2,i} I_i^2 = EI_i P_i$$

y:

30

$$d_i = \frac{1}{M} \sum^M (EI_i P_i - G_i)^2$$

[0068] Para obtener el vector P_i que tiene la distancia euclidiana promedio más baja, la derivación $\frac{\partial d_i}{\partial P_i}$ se puede establecer en cero:

$$\frac{\partial d_i}{\partial P_i} = 2 EI_i^T (G_i - EI_i P_i) = 0$$

35

para obtener:

$$P_i = (EI_i^H EI_i)^{-1} EI_i^H G_i$$

40 **[0069]** Para reducir aún más la diferencia (distancia euclidiana) entre los pesos propuestos y los coeficientes de reflexión de los pesos G.718 de otra información, se puede añadir a la matriz EI_i . Debido a que, por ejemplo, los coeficientes de reflexión contienen cierta información sobre el modelo LPC que no es directamente observable en el

dominio LSF o ISF, ayudan a reducir la distancia euclidiana d . En la práctica, es probable que no todos los coeficientes de reflexión conduzcan a una reducción significativa de la distancia euclidiana. Los inventores descubrieron que puede ser suficiente utilizar el primer y el 14º coeficiente de reflexión. Añadiendo los coeficientes de reflexión, la matriz EI_i se verá como:

5

$$EI_i = \begin{bmatrix} 1 & I_{1,i} & I_{1,i}^2 & r_{1,1} & r_{1,2} & \dots \\ 1 & I_{2,i} & I_{2,i}^2 & r_{2,1} & r_{2,2} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

donde $r_{x,y}$ es el coeficiente de reflexión y -th (u otra información) de la instancia x -th en el conjunto de datos de entrenamiento. En consecuencia, la dimensión del vector P_i comprenderá dimensiones modificadas según el número de columnas en la matriz EI_i . El cálculo del vector óptimo P_i se mantiene igual que el anterior.

10

[0070] Al añadir información adicional, la regla de determinación representada en la Fig. 4b puede cambiarse (extenderse) según $y = a + b x + c x^2 + d x^3 + \dots$

15 **[0071]**

La figura 6 muestra un diagrama de bloques esquemático de un sistema de transmisión de audio 600 según una forma de realización. El sistema de transmisión de audio 600 comprende el codificador 100 y un decodificador 602 configurado para recibir la señal de salida 182 como un flujo de bits que comprende la LSF cuantificada, o una información relacionada con el mismo, respectivamente. El flujo de bits se envía a través de un medio de transmisión 604, tal como una conexión por cable (cable) o el aire.

20

[0072] En otras palabras, la Fig. 6 muestra una vista general del esquema de codificación de LPC en el lado del codificador. Vale la pena mencionar que la ponderación es utilizada solo por el codificador y no es necesaria por el decodificador. Primero se realiza un análisis de LPC en la señal de entrada. Produce coeficientes de LPC y coeficientes de reflexión (RC). Después del análisis de LPC, los coeficientes predictivos de LPC se convierten a LSF.

25

Estas LSF se cuantifican por vector utilizando un esquema como una cuantificación vectorial en varias etapas y, a continuación, se transmiten al decodificador. La palabra de código se selecciona según una distancia de error cuadrada ponderada llamada WED como se introdujo en la sección anterior. Para este fin, los pesos asociados deben ser computados de antemano. La derivación de pesos va en función de las LSF originales y los coeficientes de reflexión. Los coeficientes de reflexión están disponibles directamente durante el análisis de LPC como variables internas que necesita el algoritmo de Levinson-Durbin.

30

[0073] La figura 7 ilustra una forma de realización de derivación de los valores de corrección tal como se ha descrito anteriormente. Los coeficientes de predicción convertidos 122' (LSF) u otros coeficientes se utilizan para determinar los pesos según el codificador en un bloque A y para computar los pesos correspondientes en un bloque B. Los pesos obtenidos 142 se combinan directamente con los pesos de referencia obtenidos 142" en un bloque C para ajustar el modelado, es decir, para computar el vector P_i como lo indica la línea discontinua del bloque A al bloque C. Opcionalmente, si la información adicional 114 es tal como los coeficientes de reflexión o la información de potencia espectral se usa para determinar los valores de corrección 162, los pesos 142' se combinan con la información adicional 114 en un vector de regresión indicado como bloque D como lo ha descrito la EI_i extendida por los valores de reflexión. Los pesos 142''' obtenidos se combinan entonces con los factores de ponderación de referencia 142" en el bloque C.

35

40

[0074] En otras palabras, el modelo de ajuste del bloque C es el vector P que se describe anteriormente. A continuación, un pseudocódigo resume de manera ejemplar el procesamiento de la derivación de peso:

45

Input: lsf = original LSF vector
 order = order of LPC, length of lsf
 parcorr[0] = - 1st reflection coefficient
 parcorr[1] = - 14th reflection coefficient
 smooth_flag = flag for smoothing weights
 w_past = past weights

Output

weights = computed weights

```

/*Compute IHM weights*/
weights[0] = 1.f/( lsf[0] - 0 ) + 1.f/( lsf[1] - lsf[0] );
for(i=1; i<order-1; i++)
  weights[i] = 1.f/( lsf[i] - lsf[i-1] ) + 1.f/( lsf[i+1] - lsf[i] );

  weights[order-1] = 1.f/( lsf[order-1] - lsf[order-2] ) + 1.f/( 8000 - lsf[order-1] );

/* Fitting model*/
for(i=0; i<order; i++)
{
  weights[i] *= (8000/ PI);
  weights[i] = ((float)(lsf_fit_model[0][i])/(1<<12))
    + weights[i]*((float)(lsf_fit_model[1][i])/(1<<14))
    + weights[i]*weights[i]*((float)(lsf_fit_model[2][i])/(1<<19))
    + parcorr[0]* ((float)(lsf_fit_model[3][i])/(1<<13))
    + parcorr[1] * ((float)(lsf_fit_model[4][i])/(1<<10));

  /* avoid too low weights and negative weights*/
  if(weights[i] < 1.f/(i+1))
    weights[i] = 1.f/(i+1);
}

wherein "parcorr" indicates the extension of the matrix E1

if(smooth_flag){
  for(i=0; i<order; i++) {
    tmp = 0.75f*weights[i] * 0.25f*w_past[i];
    w_past[i]=weights[i];
    weights[i]=tmp;
  }
}

```

5 el cual indica el suavizado descrito anteriormente en el que los pesos actuales se ponderan con un factor de 0,75 y

los pesos anteriores se ponderan con un factor de 0,25.

[0075] Los coeficientes obtenidos para el vector P pueden comprender valores escalares como se indica a continuación como ejemplo para una señal muestreada a 16 kHz y con una orden de LPC de 16:

5
`lsf_fit_model[5][16] = {
 {679 , 10921 , 10643 , 4998 , 11223 , 6847 , 6637 , 5200 , 3347 , 3423 , 3208 , 3329 , 2785 , 2295 , 2287 ,
 1743},
 {23735 , 14092 , 9659 , 7977 , 4125 , 3600 , 3099 , 2572 , 2695 , 2208 , 1759 , 1474 , 1262 , 1219 , 931 ,
 1139},
 {-6548 , -2496 , -2002 , -1675 , -565 , -529 , -469 , -395 , -477 , -423 , -297 , -248 , -209 , -160 , -125 , -217},
 {-10830 , 10563 , 17248 , 19032 , 11645 , 9608 , 7454 , 5045 , 5270 , 3712 , 3567 , 2433 , 2380 , 1895 , 1962
 , 1801},
 {-17553 , 12265 , -758 , -1524 , 3435 , -2644 , 2013 , -616 , -25 , 651 , -826 , 973 , -379 , 301 , 281 , -165}};`

[0076] Como se ha indicado anteriormente, en lugar de la LSF, el convertidor también puede proporcionar la ISF como coeficientes convertidos 122. Una derivación de peso puede ser muy similar a lo indicado por el siguiente pseudocódigo. Las ISF de orden N son equivalentes a las LSF de orden N-1 para los primeros coeficientes N-1 a los que adjuntamos los coeficientes de reflexión Nth. Por lo tanto, la derivación de pesos está muy próxima a la derivación de pesos LSF. Está dado por el siguiente pseudocódigo:

Input: `lsf` = original ISF vector
 `order` = order of LPC, length of `lsf`
 `parcorr[0]` = - 1st reflection coefficient
 `parcorr[1]` = - 14th reflection coefficient
 `smooth_flag` = flag for smoothing weights
 `w_past` = past weights

Output
 `weights` = computed weights

`/*Compute IHM weights*/
 weights[0] = 1.f/(lsf[0] - 0) + 1.f/(lsf[1] - lsf[0]);`

15

```

for(i=1; i<order-2; i++)
    weights[i] = 1.f/( lsf[i] - lsf[i-1] ) + 1.f/( lsf[i+1] - lsf[i] );

    weights[order-2] = 1.f/( lsf[order-2] - lsf[order-3] ) + 1.f/( 6400 - lsf[order-2] );

/* Fitting model*/
for(i=0; i<order-1; i++)
{
    weights[i] *= (6400/PI);
    weights[i] = ((float)(isf_fit_model[0][i])/(1<<12))
        + weights[i]*((float)(isf_fit_model[1][i])/(1<<14))
        + weights[i]*weights[i]*((float)(isf_fit_model[2][i])/(1<<19))
        + parcorr[0]* ((float)(isf_fit_model[3][i])/(1<<13))
        + parcorr[1] * ((float)(isf_fit_model[4][i])/(1<<10));
    /* avoid too low weights and negative weights*/
    if(weights[i] < 1.f/(i+1))
        weights[i] = 1.f/(i+1);
}

if(smooth_flag){
    for(i=0; i<order-1; i++) {
        tmp = 0.75f*weights[i] * 0.25f*w_past[i];
        w_past[i]=weights[i];
        weights[i]=tmp;
    }
}
weights[order-1]=1;

```

donde se ajustan los coeficientes del modelo para la señal de entrada con componentes de frecuencia de hasta 6,4 kHz:

```

5 isf_fit_model[5][15] = {
{8112, 7326, 12119, 6264, 6398, 7690, 5676, 4712, 4776, 3789, 3059, 2908, 2862, 3266, 2740},
{16517, 13269, 7121, 7291, 4981, 3107, 3031, 2493, 2000, 1815, 1747, 1477, 1152, 761, 728},
{-4481, -2819, -1509, -1578, -1065, -378, -519, -416, -300, -288, -323, -242, -187, -7, -45},
10 {-7787, 5365, 12879, 14908, 12116, 8166, 7215, 6354, 4981, 5116, 4734, 4435, 4901, 4433, 5088},
{-11794, 9971, -3548, 1408, 1108, -2119, 2616, -1814, 1607, -714, 655279, 52, 972, -416});

```

donde se ajustan los coeficientes del modelo para la señal de entrada con componentes de frecuencia de hasta 4 kHz y con energía cero para el componente de frecuencia de 4 a 6,4 kHz:

```

15 isf_fit_model [5][15] = {
{21229, -746, 11940, 205, 3352, 5645, 3765, 3275, 3513, 2982, 4812, 4410, 1036, -6623, 6103},
{15704, 12323, 7411, 7416, 5391, 3658, 3578, 3027, 2624, 2086, 1686, 1501, 2294, 9648, -6401},
{-4198, -2228, -1598, -1481, -917, -538, -659, -529, -486, -295, -221, -174, -84, -11874, 27397},
{-29198, 25427, 13679, 26389, 16548, 9738, 8116, 6058, 3812, 4181, 2296, 2357, 4220, 2977, -
71},
{-16320, 15452, -5600, 3390, 589, -2398, 2453, -1999, 1351, -1853, 1628, -1404, 113, -765, -359}};

```

[0077] Básicamente, las órdenes de la ISF se modifican, lo que puede verse al comparar el bloque /* computar los pesos IHN */ de ambos pseudo-códigos.

20

[0078] La figura 8 muestra un diagrama de flujo esquemático de un procedimiento 800 para codificar una señal de audio. El procedimiento 800 comprende una etapa 802 en la que se analiza la señal de audio en la que los coeficientes de predicción de análisis se determinan a partir de la señal de audio. El procedimiento 800 comprende además una etapa 804 en el que los coeficientes de predicción convertidos se derivan de los coeficientes de

predicción de análisis. En una etapa 806 se almacena una multitud de valores de corrección, por ejemplo, en una memoria tal como la memoria 160. En una etapa 808, los coeficientes de predicción convertidos y la multitud de valores de corrección se combinan para obtener factores de ponderación corregidos. En una etapa 812, los coeficientes de predicción convertidos se cuantifican mediante el uso de los factores de ponderación corregidos para obtener una representación cuantificada de los coeficientes de predicción convertidos. En una etapa 814, se forma una señal de salida en función de la representación de los coeficientes de predicción convertidos y en función de la señal de audio.

[0079] En otras palabras, la presente invención propone una nueva forma eficaz de derivar los pesos óptimos w mediante el uso de un algoritmo heurístico de baja complejidad. Se presenta una optimización sobre la ponderación de IHM que resulta en una menor distorsión en las frecuencias más bajas al mismo tiempo que proporciona más distorsión a las frecuencias más altas y produce una distorsión general menos audible. Tal optimización se logra computando primero los pesos como se propone en [1] y, a continuación, modificándolos de manera que se aproximen mucho a los pesos que se hubieran obtenido mediante el uso de la estrategia del G.718 [3]. La segunda etapa consiste en un modelo polinomial de segundo orden simple durante una fase de entrenamiento, minimizando la distancia euclidiana promedio entre los pesos de IHM modificados y los pesos del G.718. Simplificada, la relación entre los pesos de IHM y G.718 se modela mediante una función polinomial (probablemente simple).

[0080] Aunque algunos aspectos se han descrito en el contexto de un aparato, está claro que estos aspectos también representan una descripción del procedimiento correspondiente, donde un bloque o dispositivo corresponde a una etapa de procedimiento o una característica de una etapa de procedimiento. Análogamente, los aspectos descritos en el contexto de una etapa de procedimiento también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente.

[0081] La señal de audio codificada de la invención se puede almacenar en un medio de almacenamiento digital o se puede transmitir en un medio de transmisión, tal como un medio de transmisión inalámbrico o un medio de transmisión por cable, tal como Internet.

[0082] Dependiendo de ciertos requisitos de implementación, las formas de realización de la invención pueden implementarse en hardware o en software. La implementación se puede llevar a cabo utilizando un medio de almacenamiento digital, por ejemplo, un disquete, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tienen señales de control legibles electrónicamente almacenadas en el mismo, que cooperan (o son capaces de cooperar) con un sistema informático programable de tal forma que se lleve a cabo el procedimiento respectivo.

[0083] Algunas formas de realización según la invención comprenden un soporte de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable, de tal forma que se lleve a cabo uno de los procedimientos descritos en esta invención.

[0084] En general, las formas de realización de la presente invención pueden implementarse como un producto de programa informático con un código de programa, siendo operativo el código de programa para llevar a cabo uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede, por ejemplo, almacenarse en un soporte legible por máquina.

[0085] Otras formas de realización comprenden el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por máquina.

[0086] En otras palabras, una forma de realización del procedimiento de la invención es, por lo tanto, un programa informático que tiene un código de programa para llevar a cabo uno de los procedimientos descritos en esta invención, cuando el programa informático se ejecuta en un ordenador.

[0087] Una forma de realización adicional de los procedimientos de la invención es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención.

[0088] Una forma de realización adicional del procedimiento de la invención es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención. El flujo de datos o la secuencia de señales pueden configurarse, por ejemplo, para ser transferidos a través de una conexión de comunicación de datos, por ejemplo, a través de Internet.

[0089] Una forma de realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los procedimientos descritos en esta invención.

65

[0090] Una forma de realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención.

[0091] En algunas formas de realización, se puede usar un dispositivo lógico programable (por ejemplo, una matriz de compuerta programable de campo) para llevar a cabo algunas o todas las funcionalidades de los procedimientos descritos en esta invención. En algunas formas de realización, una matriz de puerta programable de campo puede cooperar con un microprocesador con el fin de llevar a cabo uno de los procedimientos descritos en esta invención. En general, los procedimientos se llevan a cabo preferiblemente por cualquier aparato de hardware.

10 **[0092]** Las formas de realización descritas anteriormente son meramente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en esta invención serán evidentes para otros expertos en la técnica. Es la intención, por lo tanto, limitarse únicamente por el alcance de las reivindicaciones de patente inminentes y no por los detalles específicos presentados a modo de descripción y explicación de las formas de realización en esta invención.

15

Bibliografía

[0093]

20 [1] Laroia, R.; Phamdo, N.; Farvardin, N., "Robust and efficient quantization of speech LSP parameters using structured vector quantizers," Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on, vol., no., pp.641,644 vol. 1, 14-17 abril de 1991

[2] Gardner, William R.; Rao, B.D., "Theoretical analysis of the high-rate vector quantization of LPC parameters,"
25 Speech and Audio Processing, IEEE Transactions on, vol.3, no.5, pp.367,381, septiembre de 1995

[3] ITU-T G.718 "Frame error robust narrow-band and wideband embedded variable bitrate coding of speech and audio from 8-32 kbit/s", 06/2008, sección 6.8.2.4 "ISF weighting function for frame-end ISF quantization

REIVINDICACIONES

1. Codificador (100) para la codificación de una señal de audio (102), comprendiendo el codificador (100):

- 5 un analizador (100) configurado para analizar la señal de audio (102) y para determinar los coeficientes de predicción de análisis (112) a partir de la señal de audio (102);
- un convertidor (120) configurado para derivar los coeficientes de predicción convertidos (122; 122') de los coeficientes de predicción de análisis (112);
- 10 una memoria (160) configurada para almacenar una multitud de valores de corrección (162);
- una calculadora (130; 130') que comprende:

un procesador (140; 140') configurado para procesar los coeficientes de predicción convertidos (122; 122') para obtener factores de ponderación espectral (142; 142');

- 15 un combinador (150; 150') configurado para combinar los factores de ponderación espectral (142; 142') y la multitud de valores de corrección (162; a, b, c) para obtener factores de ponderación corregidos (152; 152'); y
- un cuantificador (170) configurado para cuantificar los coeficientes de predicción convertidos (122; 122') mediante el uso de los factores de ponderación corregidos (152; 152') para obtener una representación cuantificada (172) de los coeficientes de predicción convertidos (122; 122'); y

- 20 un formador de flujo de bits (180) configurado para formar una señal de salida (182) basada en la representación cuantificada (172) de los coeficientes de predicción convertidos (122) y basada en la señal de audio (102);
- en el que el combinador (150; 150') está configurado para aplicar un polinomio basado en una forma

25
$$w = a + bx + cx^2$$

en la que w denota un factor de ponderación corregido obtenido, x denota el factor de ponderación espectral y en la que a, b y c denotan valores de corrección.

- 30 2. Codificador según la reivindicación 1, en el que el combinador (150') está configurado para combinar los factores de ponderación espectral (142; 142'), la multitud de valores de corrección (162; a, b, c) y una información adicional (114) relacionada con la señal de entrada (102) para obtener los factores de ponderación corregidos (152').

- 35 3. Codificador según la reivindicación 2, en el que la información adicional (114) relacionada con la señal de entrada (102) comprende coeficientes de reflexión obtenidos por el analizador (110) o comprende una información relacionada con un espectro de potencia de la señal de audio (102).

- 4. Codificador según una de las reivindicaciones anteriores, en el que el analizador (110) está
- 40 configurado para determinar los coeficientes de predicción lineal (LPC) y en el que el convertidor (120) está configurado para derivar frecuencias espectrales de línea (LSF; 122') o frecuencias espectrales de inmitancia (ISF) a partir de los coeficientes de predicción lineal (LPC).

- 5. Codificador según una de las reivindicaciones anteriores, en el que el combinador (150; 150') está
- 45 configurado para cíclico, en cada ciclo, obteniendo los factores de ponderación corregidos (152; 152'); en el que la calculadora (130') comprende además un suavizado (155) configurado para combinar de manera ponderada los primeros factores de ponderación cuantificados (152''') obtenidos para un ciclo anterior y los segundos factores de ponderación cuantificados (152'') obtenidos para un ciclo que sigue al ciclo anterior para obtener factores de ponderación corregidos suavizados (152'') que comprenden un valor entre los valores del primer (152''') y el segundo
- 50 (152') factores de ponderación cuantificados.

- 6. Codificador según una de las reivindicaciones anteriores, en el que la multitud de valores de corrección (162; a, b, c) se deriva de pesos precalculados (LSF; 142''), en el que una complejidad computacional para determinar los pesos precalculados (LSF; 142'') es mayor cuando se compara con la complejidad
- 55 computacional de determinación de los factores de ponderación espectral (142; 142').

- 7. Codificador según una de las reivindicaciones anteriores, en el que el procesador (140; 140') se configura mediante la obtención de los factores de ponderación espectral (142; 142') por una media armónica inversa.

- 60 8. Codificador según una de las reivindicaciones anteriores, en el que el procesador (140; 140') se configura mediante la obtención de los factores de ponderación espectral (142; 142') basándose en una forma:

$$w_i = \frac{1}{(lsf_i - lsf_{i-1})} + \frac{1}{(lsf_{i+1} - lsf_i)}$$

en la que w_i denota un peso determinado con índice i , lsf_i denota una frecuencia espectral de línea con índice i , en la que el índice i corresponde a un número de factores de ponderación espectral (142; 142') obtenidos.

5

9. Sistema de transmisiones de audio (600) que comprende:

un codificador (100) según una de las reivindicaciones anteriores; y

un decodificador (602) configurado para recibir la señal de salida (182) del codificador o una señal derivada del mismo y para la decodificación de la señal recibida (182) para proporcionar una señal de audio sintetizada (102'); en el que el codificador (100) está configurado para acceder a un medio de transmisión (604) y para transmitir la señal de salida (182) a través de los medios de transmisión (604).

10. Procedimiento para la determinación de los valores de corrección (162; a, b, c) para una primera multitud (IHM) de los primeros factores de ponderación (142; 142'), cada factor de ponderación adaptado para ponderar una porción (LSF; ISF) de una señal de audio (102), comprendiendo el procedimiento (700):

el cálculo de la primera multitud (IHM) de los primeros factores de ponderación (142; 142') para cada señal de audio de un conjunto de señales de audio y en base a una primera regla de determinación;

el cálculo de una segunda multitud de segundos factores de ponderación (142'') para cada señal de audio del conjunto de señales de audio en base a una segunda regla de determinación, estando cada una de la segunda multitud de factores de ponderación (142'') relacionada con un primer factor de ponderación (142; 142');

el cálculo de una tercera multitud de valores de distancia d_i , teniendo cada valor de distancia d_i un valor relacionado con una distancia entre un primer factor de ponderación (142; 142') y un segundo factor de ponderación (142'') relacionado con una porción de la señal de audio (102) y

el cálculo de una cuarta multitud de valores de corrección adaptados para reducir los valores de distancia d_i cuando se combinan con los primeros factores de ponderación (142; 142');

en el que la cuarta multitud de valores de corrección se determina basándose en un ajuste polinomial que comprende:

30

la aplicación de los valores de los primeros factores de ponderación (142; 142') a un polinomio $y = p_{0,i} + p_{1,i}x + p_{2,i}x^2$ en el que al menos un coeficiente del polinomio es para la adaptación;

la adaptación de al menos un coeficiente del polinomio de tal forma que la tercera multitud de valores de distancia d_i satisfaga

35

$$\frac{\partial d_i}{\partial P_i} = 2 EI_i^T (G - EI_i P_i) = 0$$

que resulta en

40

$$P_i = (EI_i^H EI_i)^{-1} EI_i^H G_i$$

en donde P_i denota un vector en una forma $P_i = [p_{0,i} p_{1,i} p_{2,i}]^T$, y en donde EI_i denota una matriz basada en:

$$EI_i = \begin{bmatrix} 1 & I_{1,i} & I_{1,i}^2 \\ 1 & I_{2,i} & I_{2,i}^2 \\ \vdots & \vdots & \vdots \end{bmatrix}$$

45

en la que $I_{x,i}$ denota el factor de ponderación i -th (142; 142') determinado en base a la primera regla de determinación (IHM) para la porción x -th de la señal de audio (102) y G_i denota un factor de ponderación i -th determinado en base a la segunda regla de determinación; o

en la que la tercera multitud de valores de distancia (d_i) se calcula en base a una información adicional (114) que

comprende coeficientes de reflexión o una información relacionada con un espectro de potencia de al menos una del conjunto de señales de audio (102) basada en:

$$EI_i = \begin{bmatrix} 1 & I_{1,i} & I_{1,i}^2 & r_{1,1} & r_{1,2} & \dots \\ 1 & I_{2,i} & I_{2,i}^2 & r_{2,1} & r_{2,2} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

5

en donde $I_{x,i}$ denota el factor de ponderación i-th (142; 142') determinado en base a la primera regla de determinación (IHM) para la porción x-th de la señal de audio (102) y $r_{a,b}$ denota la información adicional (114) basada en el factor de ponderación b-th (142; 142') y la porción a-th de la señal de audio (102).

10 11. Procedimiento (800) para la codificación de una señal de audio, comprendiendo el procedimiento:

el análisis (802) de la señal de audio (102) y para determinar los coeficientes de predicción de análisis (112) a partir de la señal de audio (102);

la derivación (804) de los coeficientes de predicción convertidos (122; 122') a partir de los coeficientes de predicción de análisis (112);

15

el almacenamiento (806) de una multitud de valores de corrección (162; a-d);

el procesamiento de los coeficientes de predicción convertidos (122; 122') para obtener factores de ponderación espectral (142; 142') combinando (808) los factores de ponderación espectral (142; 142') y la multitud de valores de corrección (162; ad) para obtener los factores de ponderación corregidos (152; 152') que comprenden la aplicación

20

de un polinomio basado en una forma

$$w = a + bx + cx^2$$

en la que w denota un factor de ponderación corregido obtenido, x denota el factor de ponderación espectral y en la que a, b y c denotan valores de corrección

25

la cuantificación (812) de los coeficientes de predicción convertidos (122; 122') mediante el uso de los factores de ponderación corregidos (152; 152') para obtener una representación cuantificada (172) de los coeficientes de predicción convertidos (122; 122'); y

la formación (814) de una señal de salida (182) basada en la representación (172) de los coeficientes de predicción convertidos (122) y basada en la señal de audio (102).

30

12. Programa informático que tiene un código de programa para llevar a cabo, cuando se ejecuta en un ordenador, un procedimiento según la reivindicación 10 u 11.

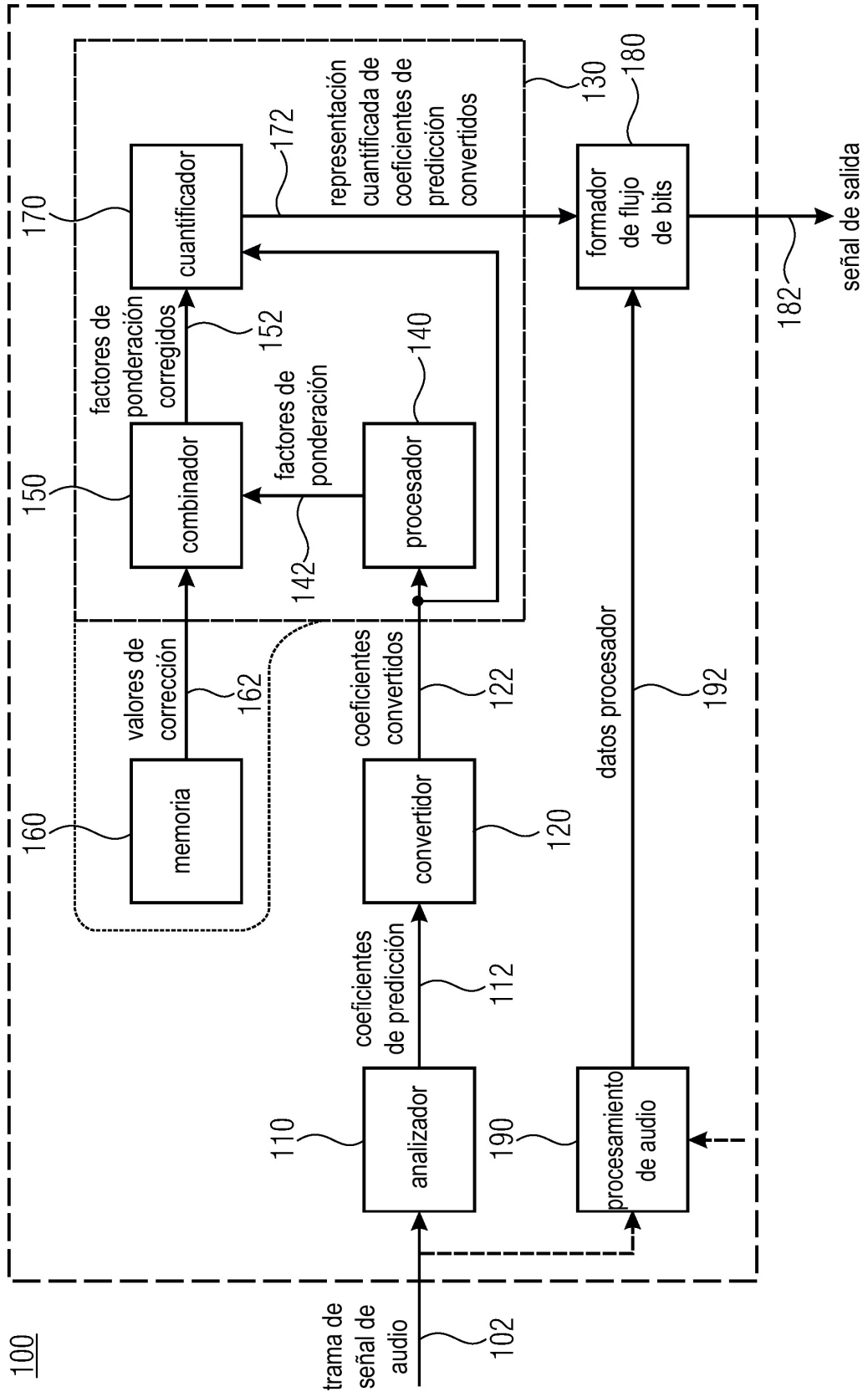


FIG 1

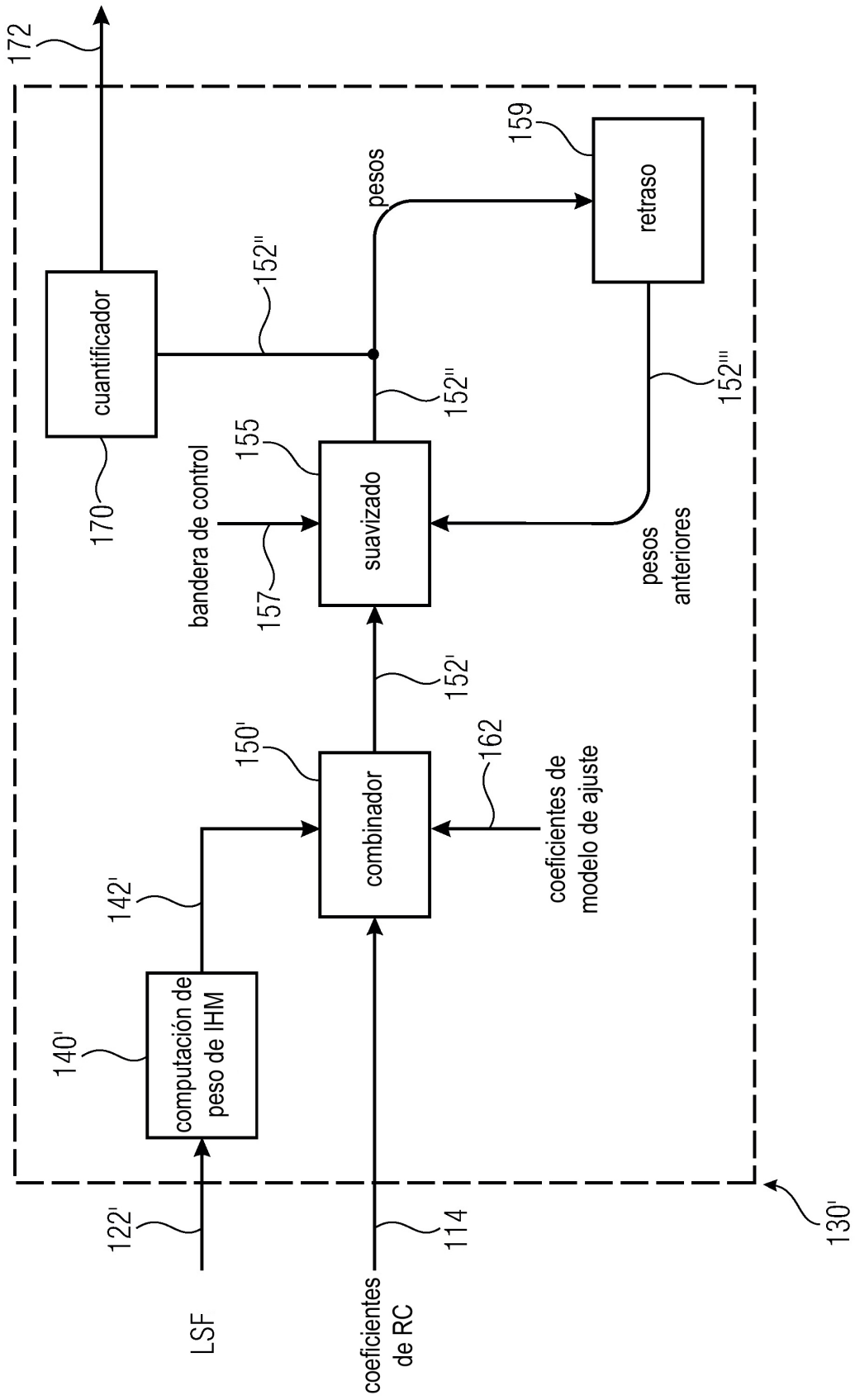


FIG 2

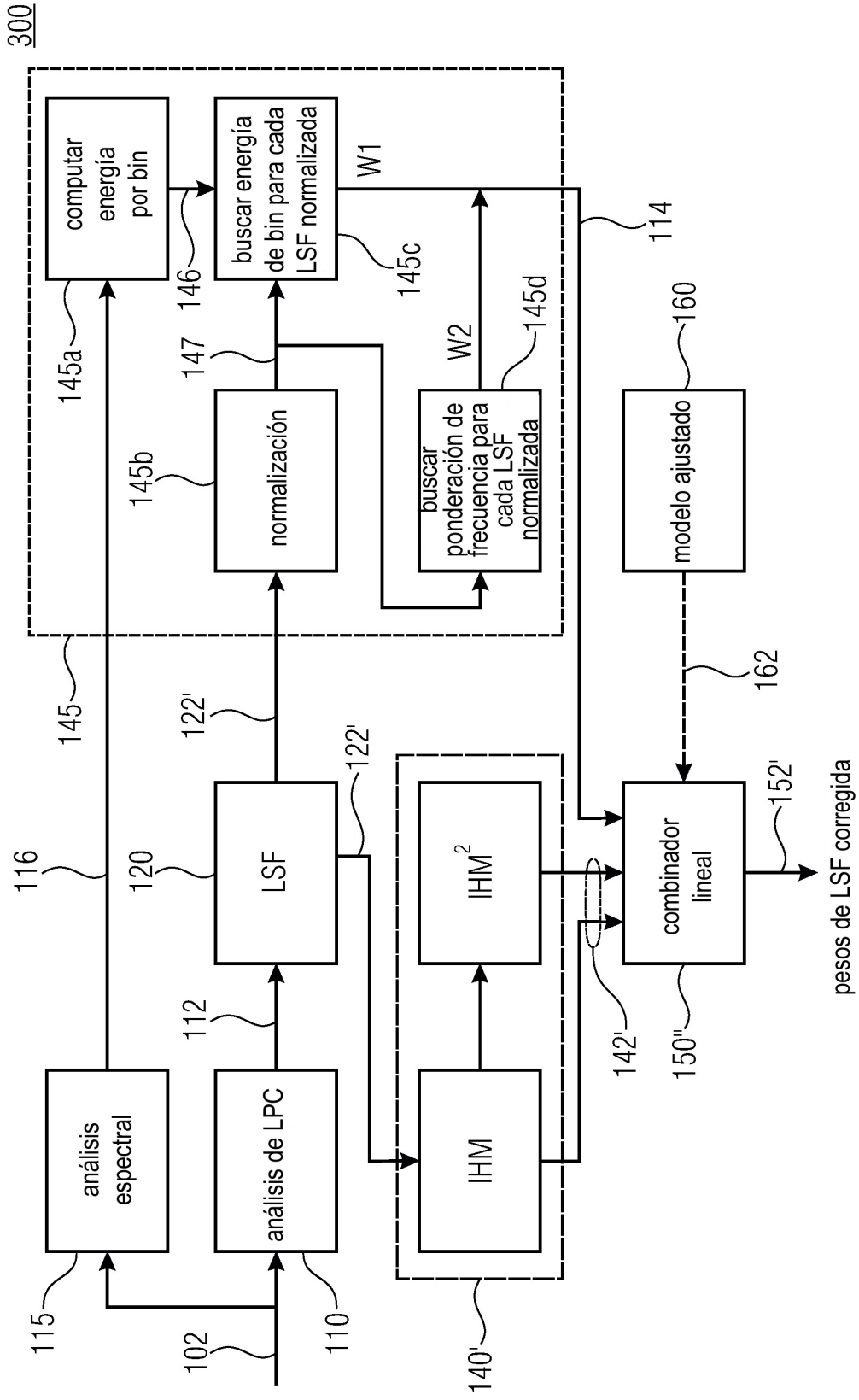


FIG 3

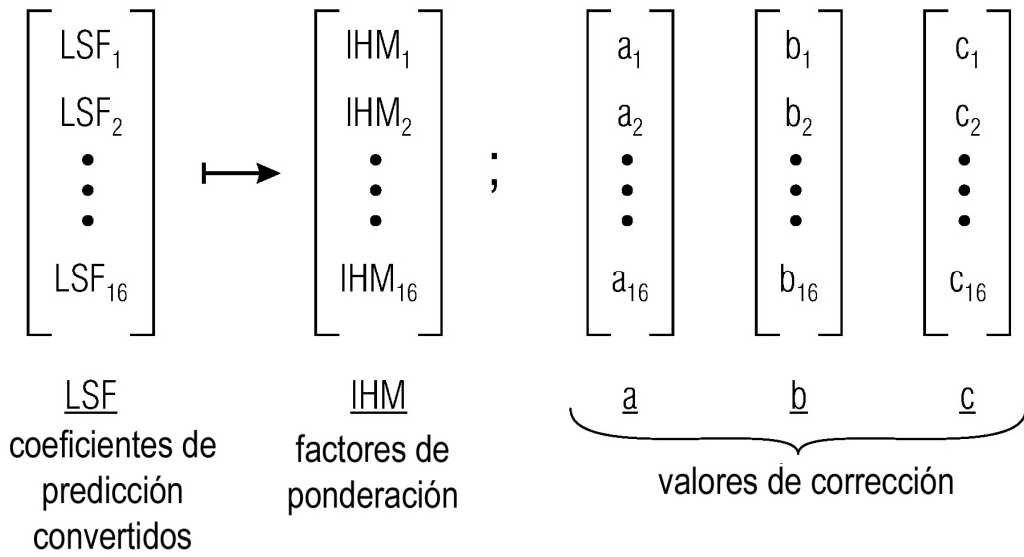


FIG 4A

$$\underline{y} = \underline{a} + \underline{b}x + \underline{c}x^2$$

FIG 4B

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{16} \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \text{IHM}_1 + c_1 \text{IHM}_1^2 \\ a_2 + b_2 \text{IHM}_2 + c_2 \text{IHM}_2^2 \\ \vdots \\ a_{16} + b_{16} \text{IHM}_{16} + c_{16} \text{IHM}_{16}^2 \end{bmatrix}$$

$$\underline{w}$$

 factores de
 ponderación
 corregidos

FIG 4C

$$w_1(\text{LSF}'_1 - \text{LSF}_1)^2 + w_2(\text{LSF}'_2 - \text{LSF}_2)^2 + \dots + w_{16}(\text{LSF}'_{16} - \text{LSF}_{16})^2$$

$$\cong \sum_{i=1}^{16} w_i(\text{LSF}'_i - \text{LSF}_i)^2$$

FIG 5A

$$\begin{array}{l} \text{LSF}'_1, \text{LSF}'_2 \dots \text{LSF}'_{16} = \text{Qu}' \\ \text{LSF}''_1, \text{LSF}''_2 \dots \text{LSF}''_{16} = \text{Qu}'' \end{array} ; \quad \text{Qu} = \begin{bmatrix} \text{Qu}' \\ \text{Qu}'' \\ \text{Qu}''' \\ \vdots \\ \text{Qu}^n \end{bmatrix}$$

FIG 5B

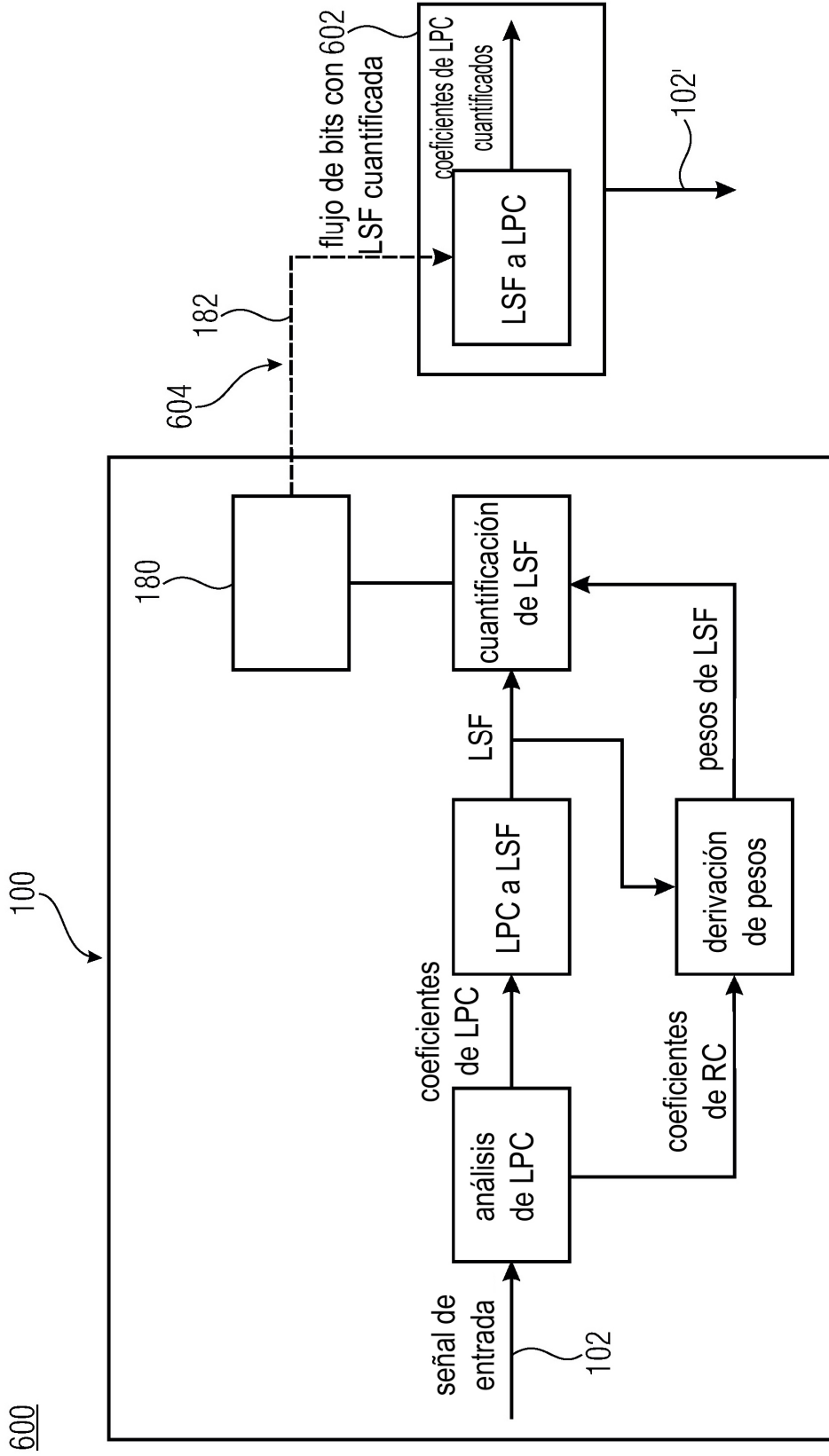


FIG 6

600

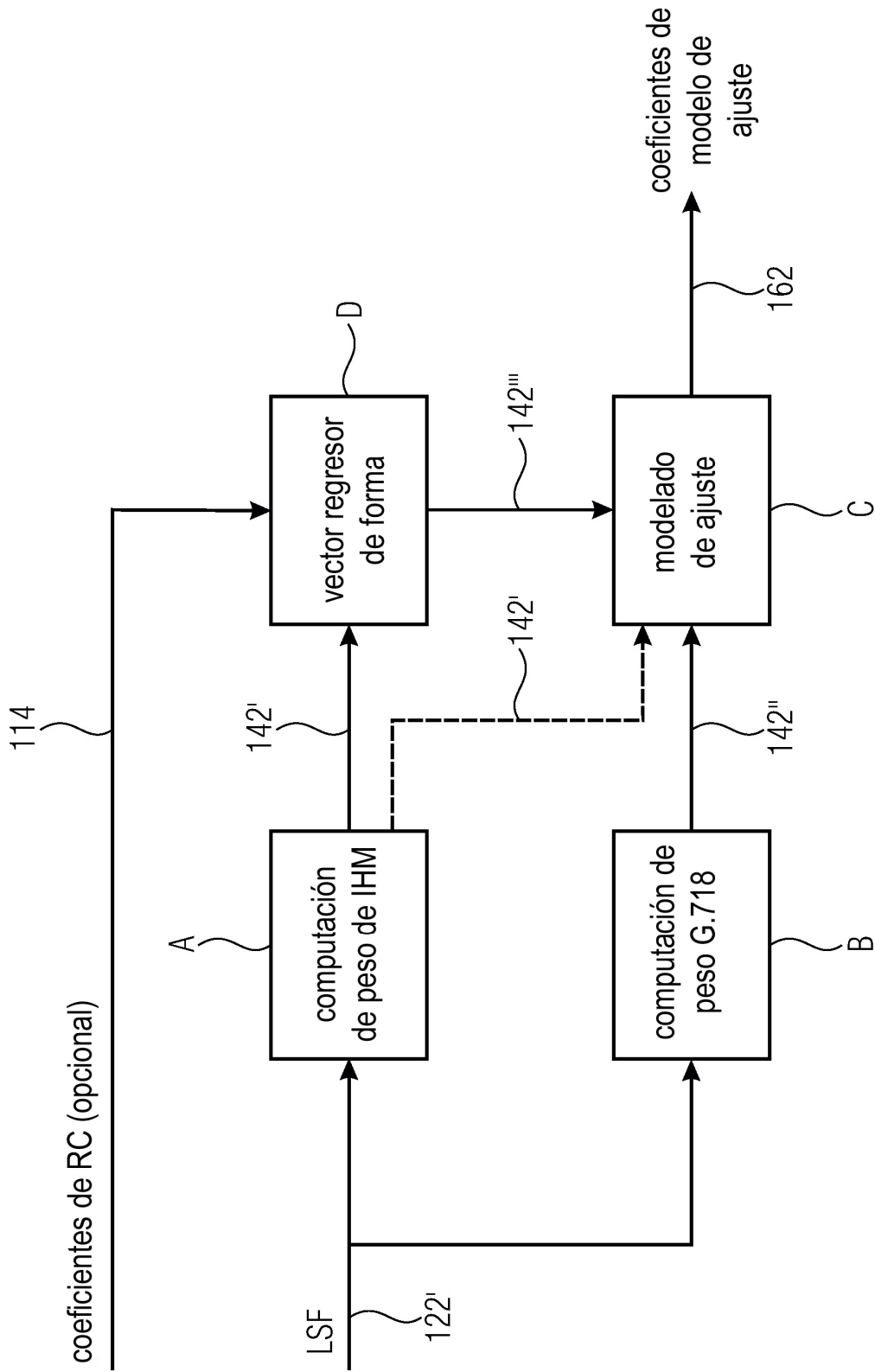


FIG 7

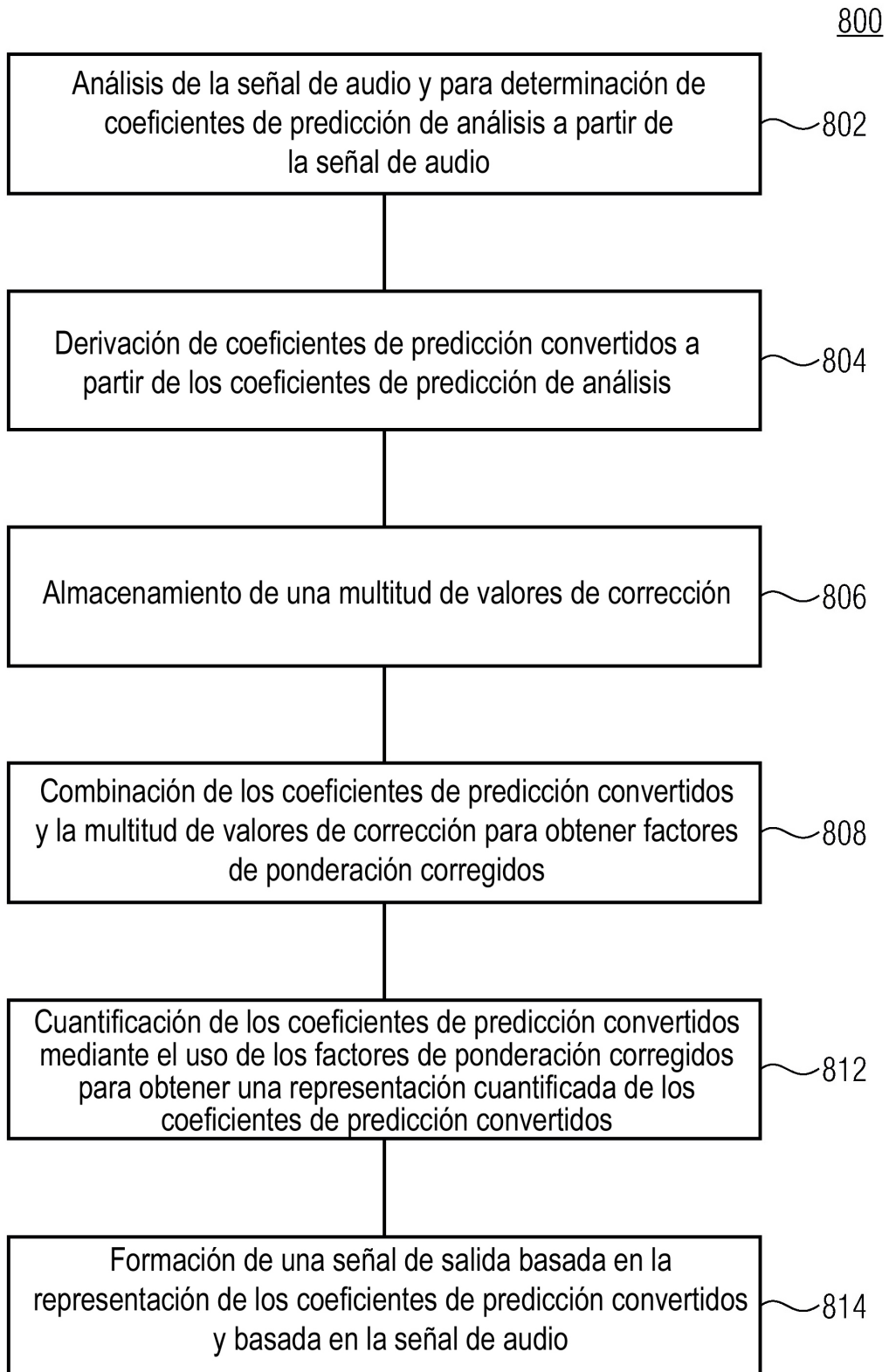


FIG 8