

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 588 483**

51 Int. Cl.:

**G10L 19/03** (2013.01)

**G10L 19/012** (2013.01)

**G10L 19/22** (2013.01)

**G10L 21/0216** (2013.01)

**G10L 25/78** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **14.02.2012** **E 14196262 (1)**

97 Fecha y número de publicación de la concesión europea: **01.06.2016** **EP 2866228**

54 Título: **Decodificador de audio que comprende un estimador de ruido de fondo**

30 Prioridad:

**14.02.2011 US 201161442632 P**

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

**03.11.2016**

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR  
FÖRDERUNG DER ANGEWANDTEN  
FORSCHUNG E.V. (100.0%)  
Hansastraße 27c  
80686 München, DE**

72 Inventor/es:

**SETIAWAN, PANJI;  
SCHMIDT, KONSTANTIN y  
WILDE, STEPHAN**

74 Agente/Representante:

**PONTI SALES, Adelaida**

**ES 2 588 483 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**

Decodificador de audio que comprende un estimador de ruido de fondo

5

La presente invención concierne a un códec de audio que soporta síntesis de ruido durante fases inactivas.

10

En la técnica se conoce la posibilidad de reducir el ancho de banda de transmisión aprovechando los períodos inactivos de voz o de otras fuentes de ruido. Tales esquemas generalmente usan alguna forma de detección para distinguir entre fases inactivas (de silencio) y activas (de no silencio). Durante fases inactivas, se logra una menor frecuencia de bits parando la transmisión de la secuencia de datos ordinaria codificando precisamente la señal grabada, y enviando en cambio sólo actualizaciones de descripción de inserción de silencio (SID). Las actualizaciones de SID pueden ser transmitidas a intervalos regulares o cuando se detectan cambios en las características del ruido de fondo. Entonces se pueden usar los cuadros de SID en el lado de decodificación para generar un ruido de fondo con características similares al ruido de fondo durante las fases activas de modo que el freno de la transmisión de la secuencia de datos ordinaria que codifica la señal grabada no conduce a una transición desagradable desde la fase activa a la fase inactiva en el lado del receptor.

15

20

Sin embargo, todavía hay una necesidad de reducir más el ritmo de transmisión. Un creciente número de consumidores de cantidad de bits transmitidos (*bitrate*), tal como un creciente número de teléfonos móviles, y un creciente número de aplicaciones que hacen uso más o menos intensivo en cuando a cantidad de bits transmitidos, requieren una permanente reducción de la cantidad de bits transmitidos consumidos.

25

Por el otro lado, el ruido sintetizado debe emular estrechamente el ruido real de modo que la síntesis sea transparente para los usuarios.

30

La solicitud WO 02/101722 A1 describe, entre otras cosas, un sistema para generar ruido de confort coloreado. En particular, el ruido de confort se genera coloreando o conformando espectralmente una señal de ruido blanco. La información de filtro se estima basándose en pausas o intervalos durante una señal vocal.

35

Consiguientemente, es un objetivo de la presente invención proveer un esquema de códec de audio que soporta generación de ruido durante fases inactivas, el cual permite reducir la cantidad de bits transmitidos de transmisión mientras se mantiene la calidad de generación de ruido alcanzable.

40

La idea básica de la presente invención es que se puede ahorrar valiosa cantidad de bits transmitidos manteniendo la calidad de generación de ruido dentro de las fases inactivas, si se actualiza continuamente una estimación de ruido de fondo paramétrico durante una fase activa de modo que la generación de ruido puede ser iniciada inmediatamente luego de la entrada de la fase inactiva que sigue a la fase activa. Por ejemplo, la actualización continua puede ser realizada en el lado de decodificación, y no hay necesidad de proveer en forma preliminar el lado de decodificación, con una representación codificada del ruido de fondo durante una fase de calentamiento inmediatamente a continuación de la detección de la fase inactiva cuya provisión consumiría valiosa cantidad de bits transmitidos, ya que el lado de decodificación tiene la estimación de ruido de fondo paramétrica continuamente actualizada durante la fase activa y por ende, está preparada en todo momento para entrar inmediatamente a la fase activa con una generación de ruido apropiada. De manera similar se puede evitar una fase de calentamiento así si la estimación de ruido de fondo paramétrica se hace en el lado de codificación. En lugar de continuar en forma preliminar con proveer al lado de decodificación una representación convencionalmente codificada del ruido de fondo al detectar la entrada de la fase inactiva para aprender el ruido de fondo e informar consiguientemente al lado de decodificación después de la fase de aprendizaje, el codificador puede proveer al decodificador la necesaria estimación de ruido de fondo paramétrica inmediatamente al detectar la entrada de la fase inactiva recurriendo a la estimación de ruido ambiente paramétrica continuamente actualizada durante la fase activa parada, evitando de ese modo la prosecución adicional preliminar que consume cantidad de bits transmitidos de la muy trabajosa codificación del ruido de fondo.

55

60

De acuerdo con realizaciones específicas de la presente invención, se alcanza una generación de ruido más realista con moderado factor que rebaja el rendimiento en términos de, por ejemplo, cantidad de bits transmitidos y complejidad computacional. En particular, de acuerdo con estas realizaciones, se usa el dominio espectral para parametrizar el ruido de fondo produciendo de ese modo una síntesis de ruido de fondo que es más realista y por ende conduce a una más transparente conmutación de fase activa a inactiva. Asimismo, se ha descubierto que parametrizar el ruido de fondo en el dominio espectral permite separar ruido de la señal útil y por consiguiente, parametrizar el ruido de fondo en el dominio espectral tiene ventaja cuando se combina con la actualización continua mencionada antes de la estimación de ruido de fondo paramétrica durante las fases activas ya que se puede lograr una mejor separación entre ruido y señal útil en el dominio espectral de modo que no es necesaria transición adicional desde un dominio a otro cuando se combinan ambos aspectos ventajosos de la presente solicitud.

65

Otros detalles ventajosos de realizaciones de la presente invención son tema de las reivindicaciones dependientes.

Abajo se describen realizaciones preferidas de la presente solicitud con respecto a las Figuras entre las cuales:

- 5 la Figura 1 muestra un diagrama de bloques de un codificador de audio, de acuerdo con una realización de comparación;
- 10 la Figura 2 muestra una implementación posible del motor de codificación 14;
- la Figura 3 muestra un diagrama de bloques de un decodificador de audio de acuerdo con una realización;
- 15 la Figura 4 muestra una posible implementación del motor de decodificación de la Figura 3 de acuerdo con una realización;
- 20 la Figura 5 muestra un diagrama de bloques de un codificador de audio de acuerdo con otra realización de comparación más detallada;
- la Figura 6 muestra un diagrama de bloques de un decodificador que podría ser usado en conexión con el codificador de la Figura 5;
- 25 la Figura 7 muestra un diagrama de bloques de un decodificador de audio de acuerdo con otra realización más detallada;
- la Figura 8 muestra un diagrama de bloques de una parte de extensión de ancho de banda espectral de un codificador de audio;
- 30 la Figura 9 muestra una implementación del codificador de extensión de ancho de banda espectral CNG de la Figura 8;
- la Figura 10 muestra un diagrama de bloques de un decodificador de audio de acuerdo con un ejemplo de realización que usa extensión de ancho de banda espectral;
- 35 la Figura 11 muestra un diagrama de bloques de una posible descripción más detallada de una realización para un decodificador de audio que usa replicación de ancho de banda espectral;
- la Figura 12 muestra un diagrama de bloques de un codificador de audio que usa extensión de ancho de banda espectral; y
- 40 la Figura 13 muestra un diagrama de bloque de otra realización de un decodificador de audio.

La Figura 1 ilustra un codificador de audio 100 de acuerdo con una realización de comparación. El codificador de audio de la Figura 1 comprende un estimador de ruido de fondo 12, un motor de codificación 14, un detector 16, una entrada de señal de audio 18 y una salida de secuencia de datos 20. El proveedor 12, el motor de codificación 14 y el detector 16 tienen una entrada conectada a la entrada de señal de audio 18, respectivamente. Las salidas del estimador 12 y del motor de codificación 14 están respectivamente conectadas a la salida de secuencia de datos 20 a través de un conmutador 22. El conmutador 22, el estimador 12 y el motor de codificación 14 tienen una entrada de control conectada a una salida del detector 16, respectivamente.

50 El estimador de ruido de fondo 12 está configurado para actualizar continuamente una estimación de ruido de fondo paramétrica durante una fase activa 24 basado en una señal de audio de entrada que ingresa al codificador de audio 100 en la entrada 18. A pesar de que la Figura 1 sugiere que el estimador de ruido de fondo 12 puede establecer la actualización continua de la estimación de ruido de fondo paramétrica basado en la señal de audio como entrada en la entrada 18, no necesariamente es este el caso. Alternativamente o adicionalmente, el estimador de ruido de fondo 12 puede obtener una versión de la señal de audio del motor de codificación 14 como se ilustra mediante la línea de trazos 26. En ese caso, el estimador de ruido de fondo 12, alternativamente o adicionalmente, sería conectado a la entrada 18 indirectamente vía la línea de conexión 26 y el motor de codificación 14, respectivamente. En particular, existen diferentes posibilidades para que el estimador de ruido de fondo 12 actualice continuamente la estimación de ruido de fondo y abajo se describen algunas de esas posibilidades.

60 El motor de codificación 14 está configurado para codificar la señal de audio de entrada que llega a la entrada 18 en una secuencia de datos durante la fase activa 24. La fase activa abarcará todo tiempo en que una información útil está contenida dentro de la señal de audio como voz u otro sonido útil de una fuente de ruido. Por el otro lado, los sonidos con una característica casi invariable en el tiempo tales como un espectro con invariancia temporal como el causado por ejemplo por lluvia o tráfico en el fondo de un locutor, será clasificado como ruido de fondo y siempre

que esté presente meramente este ruido de fondo, el respectivo período de tiempo será clasificado como una fase inactiva 28. El detector 16 es responsable de detectar la entrada de una fase inactiva 28 a continuación de la fase activa 24 basado en la señal de audio de entrada en la entrada 18. En otras palabras, el detector 16 distingue entre dos fases, a saber la fase activa y la fase inactiva en donde el detector 16 decide en cuanto a qué fase está presente actualmente. El detector 16 informa al motor de codificación 14 acerca de la fase actualmente presente y como ya se mencionó, el motor de codificación 14 realiza la codificación de la señal de audio de entrada en la secuencia de datos durante las fases activas 24. El detector 16 controla el conmutador 22 consiguientemente de modo que la secuencia de datos entregada por el motor de codificación 14 sea entregada en la salida 20. Durante las fases inactivas, el motor de codificación 14 puede dejar de codificar la señal de audio de entrada. Por lo menos, la secuencia de datos entregada en la salida 20 ya no es alimentada por ninguna secuencia de datos posiblemente entregada por el motor de codificación 14. Además de eso, el motor de codificación 14 puede solo realizar mínimo procesamiento para dar soporte al estimador 12 con algunas actualizaciones de estado variables. Esta acción reducirá mucho la potencia computacional. Por ejemplo, se pone el conmutador 22 tal que la salida del estimador 12 esté conectada a la salida 20 en lugar de a la salida del motor de codificación. De esta manera se reduce valiosa cantidad de bits transmitidos para transmitir la serie de bits en el tiempo entregada en la salida 20.

El estimador de ruido de fondo 12 está configurado para actualizar continuamente una estimación de ruido de fondo paramétrica durante la fase activa 24 basado en la señal de audio de entrada 18 como ya mencionó arriba, y debido a esto, el estimador 12 puede insertar en la secuencia de datos 30 entregada en la salida 20, la estimación de ruido de fondo paramétrica como está actualizada continuamente durante la fase activa 24, inmediatamente a continuación de la transición desde la fase activa 24 a la fase inactiva 28, esto es, inmediatamente luego de la entrada en la fase inactiva 28. Por ejemplo, el estimador de ruido de fondo 12 puede insertar un cuadro descriptor de inserción de silencio 32 en la secuencia de datos 30 inmediatamente siguiente a la finalización de la fase activa 24 e inmediatamente siguiente al instante de tiempo 34 en el cual el detector 16 detectó la entrada de la fase inactiva 28. En otras palabras, no hay espacio de tiempo entre la detección de la entrada de la fase inactiva 28 en los detectores y la inserción del SID 32 necesaria debido a la actualización continua del estimador de ruido de fondo de la estimación de ruido de fondo paramétrica durante la fase activa 24.

Así, sintetizando la descripción de arriba, el codificador de audio 10 de la Figura 1 puede funcionar como sigue. Imagine, con propósitos de ilustración, que actualmente está presente una fase activa 24. En este caso, el motor de codificación 14 actualmente codifica la señal de audio de entrada en la entrada 18, formando la secuencia de datos 20. El conmutador 22 conecta la salida del motor de codifica 14 a la salida 20. El motor de codificación 14 puede usar codificación paramétrica y/codificación por transformada para codificar la señal de audio de entrada 18 formando la secuencia de datos. En particular, el motor de codificación 14 puede codificar la señal de audio de entrada en unidades de cuadros con cada cuadro que codifica uno de los intervalos de tiempo consecutivos — que se superponen parcialmente entre sí — de la señal de audio de entrada. El motor de codificación 14 puede tener adicionalmente la capacidad de conmutar entre diferentes modos de codificación entre cuadros consecutivos de la secuencia de datos. Por ejemplo, algunos cuadros pueden ser codificados usando codificación predictiva tal como codificación CELP, y algunos otros cuadros pueden ser codificados usando codificación por transformada tal como codificación TCX o AAC. Se hace referencia, por ejemplo, a USAC y sus modos de codificación según se describen en ISO/IEC CD 23003—3 fechado el 24 de septiembre de 2010.

El estimador de ruido de fondo 12 actualiza continuamente la estimación de ruido de fondo paramétrica durante la fase activa 24. Consiguientemente, el estimador de ruido de fondo 12 puede estar configurado para distinguir entre un componente de ruido y un componente de señal útil dentro de la señal de audio de entrada para determinar la estimación de ruido de fondo paramétrica meramente del componente de ruido. Tal y como se describe adicionalmente abajo, el estimador de ruido de fondo 12 puede realizar esta actualización en un dominio espectral tal como un dominio espectral también usado para codificación por transformada dentro del motor de codificación 14. Sin embargo, también hay disponibles otras alternativas, tal como el dominio del tiempo. Si es en el dominio espectral, el mismo puede ser un dominio de la transformada lapped tal como un dominio de MDCT (*Modified Discrete Cosine Transform*), o un dominio de banco de filtros tal como un dominio de banco de filtros de valor complejo tal como un dominio QMF (*Quadrature Mirror Filters*).

Asimismo, el estimador de ruido de fondo 12 puede realizar la actualización basado en una excitación o señal residual obtenida como un resultado intermedio dentro del motor de codificación 14, por ejemplo, durante la codificación predictivo y/o por transformada más que la señal de audio según entra a la entrada 18 o un codificada con pérdida de datos (*lossy*) formando la secuencia de datos. Haciéndolo esto, una gran cantidad del componente útil de señal dentro de la señal de audio ya habría sido quitado de modo que la detección del componente de ruido es más fácil para el estimador de ruido de fondo 12.

Durante la fase activa 24, el detector 16 también está funcionando continuamente para detectar una entrada de la fase inactiva 28. El detector 16 puede ser puesto en práctica como un detector de actividad de voz/sonido (VAD/SAD) o algún otro mecanismo que decida si hay un componente de señal útil presente actualmente dentro de la señal de audio o no. Un criterio de base para el detector 16 para decidir si continúa o no una fase activa 24, podría

ser verificar si una potencia filtrada por pasa bajos de la señal de audio, queda por debajo de un cierto umbral, suponiendo que se ingresa a una fase inactiva tan pronto como se excede el umbral.

Independientemente de la forma exacta en que el detector 16 realiza la detección de la entrada de la fase inactiva 28 a continuación de la fase activa 24, el detector 16 informa inmediatamente a las otras entidades 12, 14 y 22, de la entrada de la fase inactiva 28. Debido a la actualización continua del estimador de ruido de fondo de la estimación de ruido de fondo paramétrica durante la fase activa 24, se puede impedir inmediatamente que la secuencia de datos 30 entregada en la salida 20 siga siendo alimentada desde el motor de codificación 14. En cambio, el estimador de ruido de fondo 12 insertaría, inmediatamente luego de ser informado de la entrada de la fase inactiva 28, insertaría la información sobre la última actualización de la estimación de ruido de fondo paramétrica en la secuencia de datos 30, en la forma del cuadro SID 32. Eso es, el cuadro SID 32 podría seguir inmediatamente al último cuadro del motor de codificación que codifica el cuadro de la señal de audio concerniente al intervalo de tiempo dentro del cual el detector 16 detectó la entrada de fase inactiva.

Normalmente, el ruido de fondo no cambia con mucha frecuencia. En la mayoría de los casos, el ruido de fondo tiende a ser algo invariante con el tiempo. Por consiguiente, después que el estimador de ruido de fondo 12 insertó el cuadro de SID 32 inmediatamente después que el detector 16 que detecta el comienzo de la fase inactiva 28, toda transmisión de secuencia de datos puede ser interrumpida de modo que en esta fase de interrupción 34, la secuencia de datos 30 no consume cantidad de bits transmitidos o meramente un mínimo de cantidad de bits transmitidos requerida para algún propósito de transmisión. Para mantener una mínima cantidad de bits transmitidos, el estimador de ruido de fondo 12 puede repetir en forma intermitente la salida de SID 32.

Sin embargo, a pesar de la tendencia del ruido de fondo a no cambiar con el tiempo, de todos modos puede ocurrir que el ruido de fondo cambie. Por ejemplo, imagine un usuario de teléfono móvil saliendo del auto de modo que el ruido de fondo cambia del ruido de motor al ruido del tráfico afuera del auto durante la llamada del usuario. Para rastrear tales cambios del ruido de fondo, el estimador de ruido de fondo 12 puede ser configurado para escrutar continuamente el ruido de fondo durante la fase inactiva 28. Siempre que el estimador de ruido de fondo 12 determina que la estimación de ruido de fondo paramétrica cambia en una magnitud que excede algún umbral, el estimador de ruido de fondo 12 puede insertar una versión actualizada de la estimación de ruido de fondo paramétrica dentro de la secuencia de datos 20 vía otro SID 38, después de lo cual puede seguir otra fase de interrupción 40, por ejemplo, empieza otra fase activa 42 según lo detectado por el detector 16, y así de seguido. Naturalmente, los cuadros de SID que revelan la estimación de ruido de fondo paramétrica actualizada actualmente pueden, adicionalmente o alternativamente, pueden ser intercalados dentro de las fases inactivas de una manera intermedia dependiente de los cambios en la estimación de ruido de fondo paramétrica.

Obviamente, la secuencia de datos 44 entregada por el motor de codificación 14 e indicada en la Figura 1 mediante el uso de sombreado, consume más cantidad de bits transmitidos de transmisión que los fragmentos de secuencia de datos 32 y 38 a ser transmitidos durante las fases inactivas 28 y por consiguiente los ahorros de cantidad de bits transmitidos son considerables. Asimismo, como el estimador de ruido de fondo 12 puede empezar inmediatamente con proceder a alimentar adicionalmente la secuencia de datos 30, no es necesario continuar en forma preliminar transmitiendo la secuencia de datos 44 del motor de codifica 14 más allá del punto de tiempo de detección de fase inactiva 34, reduciendo de ese modo todavía más la cantidad de bits transmitidos consumidos en total.

Como se explicará en más detalle abajo, el motor de codificación 14 puede ser configurado para, al codificar la señal de audio de entrada, codificar predictivamente la señal de audio de entrada en coeficientes de predicción lineal y un señal de excitación con codificar por transformada la señal de excitación y codificar los coeficientes de predicción lineal en la secuencia de datos 30 y 44, respectivamente. En la Figura 2 se muestra una posible implementación. De acuerdo con la Figura 2, el motor de codificación 14 comprende un transformador 50, un corrector (*shaper*) de ruido del dominio de la frecuencia 52 y un cuantificador 54 los cuales están conectados en serie en el orden en que se mencionan entre una entrada de señal de audio 56 y una salida de secuencia de datos 58 del motor de codificación 14. Además, el motor de codificación 14 de la Figura 2 comprende un módulo de análisis de predicción lineal 60 el cual está configurado para determinar coeficientes de predicción lineal a partir de la señal de audio 56 mediante respectivo ventaneo de análisis de porciones de la señal de audio y aplicando una autocorrelación a las porciones ventaneadas, o determinar una autocorrelación sobre la base de las transformadas en el dominio de la transformada de la señal de audio de entrada como salida mediante el transformador 50 usando su espectro de potencia y aplicando una DFT inversa a ello de modo de determinar la autocorrelación, realizar subsiguientemente estimación LPC en base a la autocorrelación tal como usando un algoritmo de (Wiener—) Levinson—Durbi.

En base a los coeficientes de predicción lineal determinados por el módulo de análisis de predicción lineal 60, la secuencia de datos entregada en la salida 58 es alimentada con respectiva información sobre los LPCs, y el corrector de ruido del dominio de la frecuencia es controlado de modo que corregir espectralmente el espectrograma de la señal de audio en concordancia con una función transferencia correspondiente a la función transferencia de un filtro de análisis de predicción lineal determinado por los coeficientes de predicción lineal entregados por el módulo 60. Se puede realizar una cuantificación de los LPCs para transmitirlos en la secuencia de datos, en el dominio LSP/LSF usando interpolación de modo de reducir la tasa de trasmisión comparado con la tasa de análisis en el

analizador 60. Además, la conversión de LPC a ponderación espectral realizada en el FDNS puede involucrar aplicar una ODFT sobre los LPCs y aplicar los valores de ponderación resultantes a los espectros del transformador como divisor.

5 El cuantificador 54 entonces cuantifica los coeficientes de transformación del espectrograma espectralmente formado (aplanado). Por ejemplo, el transformador 50 usa una transformada lapped tal como una MDCT para transferir la señal de audio desde el dominio del tiempo al dominio espectral, obteniendo de ese modo transformadas consecutivas correspondientes a porciones ventaneadas que se superponen de la señal de audio, las cuales luego son formadas espectralmente mediante el corrector de ruido del dominio de la frecuencia 52 ponderando estas transformadas de acuerdo con la función de transferencia del filtro de análisis de LP.

10 El espectrograma corregido puede ser interpretado como una señal de excitación y como tal se ilustra mediante la flecha de trazos 62, el estimador de ruido de fondo 12 puede estar configurado para actualizar la estimación de ruido paramétrica usando esta señal de excitación. Alternativamente, como está indicado mediante la flecha de trazos 62, el estimador de ruido de fondo 12 puede usar la representación de transformada lapped como salida mediante el transformador 50 como base para la actualización directamente, esto es, sin la corrección de ruido del dominio de la frecuencia mediante el corrector de ruido 52.

15 Más detalles relacionados con posibles implementaciones de los elementos mostrados en las Figuras 1 a 2 se pueden establecer a partir de la descripción subsiguiente más detallada y se hace notar que todos estos detalles son individualmente transferibles a los elementos de las Figuras 1 y 2.

20 Sin embargo, antes de esta descripción más detallada, se hace referencia a la Figura 3, la cual muestra que, adicionalmente o alternativamente, la actualización de estimación de ruido de fondo paramétrica se puede hacer del lado del decodificador.

25 El decodificador de audio 80 de la Figura 3 está configurado para decodificar una secuencia de datos que entra en la entrada 82 del decodificador 80 de modo de reconstruir a partir de ella una señal de audio a ser entregada en una salida 84 del decodificador 80. La secuencia de datos comprende por lo menos una fase activa 86 seguida por una fase inactiva 88. Internamente, el decodificador de audio 80 comprende un estimador de ruido de fondo 90, un motor de decodificación 92, un generador aleatorio paramétrico 94 y un generador de ruido de fondo 96. El motor de decodificación 92 está conectado entre la entrada 82 y la salida 84 y de manera similar, la conexión en serie del proveedor 90, el generador de ruido de fondo 96 y el generador aleatorio paramétrico 94 están conectados entre la entrada 82 y la salida 84. El decodificador 92 está configurado para reconstruir la señal de audio a partir de la secuencia de datos durante la fase activa, de modo que la señal de audio 98 como se entrega en la salida 84 comprende ruido y sonido útil en una calidad apropiada. El estimador de ruido de fondo 90 está configurado para actualizar continuamente una estimación de ruido de fondo paramétrica a partir de la secuencia de datos durante la fase activa. Con este fin, el estimador de ruido de fondo 90 puede no estar conectado a la entrada 82 directamente sino vía el motor de decodificación 92 como se ilustra mediante la línea de trazos 100 de modo de obtener del motor de decodificación 92 alguna versión reconstruida de la señal de audio. En principio, el estimador de ruido de fondo 90 puede estar configurado para funcionar en modo muy similar al estimador de ruido de fondo 12, además del hecho de que el estimador de ruido de fondo 90 solo tiene acceso a la versión reconstruible de la señal de audio, esto es, que incluye la pérdida causada por cuantificación en el lado de codificación.

40 El generador aleatorio paramétrico 94 puede comprender uno o más generadores o pseudogeneradores de números aleatorios, la secuencia de valores entregada por el cual puede conformarse a una distribución estadística que puede ser fijada en forma paramétrica vía el generador de ruido de fondo 96.

45 El generador de ruido de fondo 96 está configurado para sintetizar la señal de audio 98 durante la fase inactiva 88 controlando el generador aleatorio paramétrico 94 durante la fase inactiva 88 dependiendo de la estimación de ruido de fondo paramétrica según se obtuvo del estimador de ruido de fondo 90. A pesar de que ambas entidades, 96 y 94, se muestran conectadas en serie, la conexión en serie no debe ser interpretada como limitación. Los generadores 96 y 94 podrían estar interconectados. De hecho, el generador 94 podría ser interpretado como siendo parte del generador 96.

50 Así, el modo de funcionamiento del decodificador de audio 80 de la Figura 3 puede ser como sigue. Durante una fase activa 86, la entrada 82 es provista continuamente con una porción de secuencia de datos 102 la cual ha de ser procesada por el motor de decodificación 92 durante la fase activa 86. La secuencia de datos 104 que entra en la entrada 82 entonces frena la transmisión de la porción de secuencia de datos 102 dedicada para el motor de decodificación 92 en algún instante de tiempo 106. Eso es, no hay otro cuadro de la porción de secuencia de datos en ese instante de tiempo 106 para decodificación por el motor 92. La señalización de la entrada de la fase inactiva 88 puede ser la perturbación de la transmisión de la porción de secuencia de datos 102, o bien puede ser señalizada mediante alguna información 108 dispuesta inmediatamente en el inicio de la fase inactiva 88.

En cualquier caso, la entrada de la fase inactiva 88 ocurre muy repentinamente, pero esto no es un problema ya que el estimador de ruido de fondo 90 tiene actualizada continuamente la estimación de ruido de fondo paramétrica durante la fase activa 86 sobre la base de la porción de secuencia de datos 102. Debido a esto, el estimador de ruido de fondo 90 puede proveer al generador de ruido de fondo 96 la versión más nueva de la estimación de ruido de fondo paramétrica tan pronto como empieza en 106, la fase inactiva 88. Consiguientemente, a partir del instante 106 en adelante, el motor de decodificación 92 deja de entregar reconstrucción de señal de audio ya que el motor de decodificación 92 ya no es alimentado con una porción de secuencia de datos 102, pero el generador aleatorio paramétrico 94 es controlado por el generador de ruido de fondo 96 de acuerdo con una estimación de ruido de fondo paramétrica tal que puede entregar una emulación del ruido de fondo en la salida 84 inmediatamente a continuación del instante de tiempo 106 de modo seguir sin interrupción la señal de audio reconstruida según lo entregado por el motor de decodificación 92 hasta el instante de tiempo 106. Se puede usar desvanecimiento cruzado para transitar desde el último cuadro reconstruido de la fase activa según lo entregado por el motor 92 al ruido de fondo según está determinado por la versión recientemente actualizada de la estimación de ruido de fondo paramétrica.

Como el estimador de ruido de fondo 90 está configurado para actualizar continuamente la estimación de ruido de fondo paramétrica a partir de la secuencia de datos 104 durante la fase activa 86, el mismo puede ser configurado para distinguir entre un componente de ruido y un componente de señal útil dentro de la versión de la señal de audio según lo reconstruido a partir de la secuencia de datos 104 en la fase activa 86 y para determinar la estimación de ruido de fondo paramétrica meramente a partir del componente de ruido en lugar del componente de señal útil. La manera en que el estimador de ruido de fondo 90 realizar esta distinción/separación corresponde a la manera delineada arriba con respecto al estimador de ruido de fondo 12. Por ejemplo, se puede usar la excitación o señal residual interna reconstruida a partir de la secuencia de datos 104 dentro del motor de decodificación 92.

Similar a la Figura 2, la Figura 4 muestra una posible implementación para el motor de decodificación 92. De acuerdo con la Figura 4, el motor de decodificación 92 comprende una entrada 110 para recibir la porción de secuencia de datos 102 y una salida 112 para entregar la señal reconstruida dentro de la fase activa 86. Conectados en serie entre ellos, el motor de decodificación 92 comprende un descuantificador 114, un corrector de ruido del dominio de la frecuencia 116 y un transformador inverso 118, los cuales están conectados entre la entrada 110 y la salida 112 en el orden en que se mencionan. La porción de secuencia de datos 102 que llega a la entrada 110 comprende una versión codificada de transformada de la señal de excitación, esto es, niveles de coeficientes de transformada que la representan, los cuales son alimentados a la entrada del descuantificador 114, así como también información sobre coeficientes de predicción lineal, cuya información es alimentada al corrector de ruido del dominio de la frecuencia 116. El descuantificador 114 descuantifica la representación espectral de la señal de excitación y la remite al corrector de ruido del dominio de la frecuencia 116 el cual, a su vez, forma espectralmente el espectrograma de la señal de excitación (junto con el ruido de cuantificación plano) de acuerdo con una función transferencia que corresponde a un filtro de síntesis de predicción lineal, formando de ese modo ruido de cuantificación. En principio, el FDNS 116 de la Figura 4 actúa similar al FDNS de la Figura 2: Se extraen los LPCs de la secuencia de datos y luego son sometidos a conversión de LPC a peso espectral, por ejemplo, aplicando una ODFT sobre los LPCs extraídos, aplicando luego las ponderaciones espectrales resultantes a los espectros descuantificados entrantes desde el descuantificador 114 como multiplicadores. El transformador 118 entonces transfiere la reconstrucción de señal de audio así obtenida desde el dominio espectral al dominio del tiempo y entre la señal de audio reconstruida así obtenida en la salida 112. El transformador inverso 118 puede usar una transformada lapped tal como una IMDCT. Como se ilustra mediante la flecha de trazos 120, el espectrograma de la señal de excitación puede ser usado por el estimador de ruido de fondo 90 para la actualización de ruido de fondo paramétrica. Alternativamente, pero no de acuerdo con la invención, el espectrograma de la señal de audio en sí mismo puede ser usado como se indica mediante la flecha de trazos 122.

Con respecto a la Figura 2 y a la Figura 4, se debe notar que las implementaciones de los motores de codificación/decodificación descritas con respecto a los mismos no han de ser interpretadas como restrictivas. También son factibles ejemplos alternativos. Asimismo, los motores de codificación/decodificación pueden ser un tipo códec multimodal donde las partes de las Figuras 2 y 4 meramente asumen responsabilidad para codificar/decodificar cuadros que tiene un modo de codificación de cuadro específico asociado con el mismo, mientras que otros cuadros son sometidos a otras partes de los motores de codificación/decodificación no mostrados en las Figuras 2 y 4. Tal otro modo de codificación de cuadro podría también ser un modo de codificación predictivo que usa codificación de predicción lineal por ejemplo, pero con codificación en el dominio del tiempo en lugar de usar codificación de transformada.

La Figura 5 muestra un ejemplo más detallado del codificador de la Figura 1. En particular, el estimador de ruido de fondo 12 se muestra en más detalle en la Figura 5.

De acuerdo con la Figura 5, el estimador de ruido de fondo 12 comprende un transformador 140, un FDNS 142, un módulo de análisis de LP 144, un estimador de ruido 146, un estimador de parámetros 148, un medidor de estacionariedad 150, y un cuantificador 152. Algunos de los componentes recién mencionados pueden estar total o parcialmente en el motor de codificación 14. Por ejemplo, el transformador 140 y el transformador 50 de la Figura 2

pueden ser iguales, los módulos de análisis de LP 60 y 144 pueden ser iguales, los FDNSs 52 y 142 pueden ser iguales y/o los cuantificadores 54 y 152 pueden ser implementados en un módulo.

La Figura 5 también muestra un empaquetador de serie de bits en el tiempo (*bitstream packager*) 154 el cual asume una responsabilidad pasiva para la operación del conmutador 22 en la Figura 1. En particular, el VAD como el detector 16 del codificador de la Figura 5 es llamado en forma ejemplar, simplemente decide en lo que hace a cuál camino se debe tomar, ya sea el camino de la codificación de audio 14 o el camino del estimador de ruido de fondo 12. Para ser más preciso, el motor de codificación 14 y el estimador de ruido de fondo 12 están ambos conectados en paralelo entre la entrada 18 y el empaquetador 154, en donde adentro del estimador de ruido de fondo 12, el transformador 140, el FDNS 142, el módulo de análisis de LP 144, el estimador de ruido 146, el estimador de parámetros 148, y el cuantificador 152, están conectados en serie entre la entrada 18 y el empaquetador 154 (en el orden en que se mencionan), mientras el módulo de análisis de LP 144 está conectado entre la entrada 18 y una entrada de LPC del módulo FDNS 142 y una entrada adicional del cuantificador 153, respectivamente, y un medidor de estacionariedad 150 está conectado adicionalmente entre el módulo de análisis de LP 144 y una entrada de control del cuantificador 152. El empaquetador de serie de bits en el tiempo 154 simplemente realiza el empaquetado si recibe una entrada proveniente de cualquiera de las entidades conectadas a sus entradas.

En el caso de transmitir cuadros cero, esto es, durante la fase de interrupción de la fase inactiva, el detector 16 informa al estimador de ruido de fondo 12, en particular al cuantificador 152, que pare de procesar y que no envíe nada al empaquetador de serie de bits en el tiempo 154.

De acuerdo con la Figura 5, el detector 16 puede funcionar en el dominio del tiempo y/o de la transformada/espectral de modo de detectar fases activas/inactivas.

El modo de funcionamiento del codificador de la Figura 5 es como sigue. Como se irá aclarando, el codificador de la Figura 5 puede mejorar la calidad de ruido de comodidad tal como ruido estacionario en general, tal como ruido de auto, ruido de charla con muchos participantes, algunos instrumentos musicales, y en particular aquellos que tienen muchos armónicos tales como gotas de lluvia.

En particular, el codificador de la Figura 5 es para controlar un generador aleatorio en el lado de decodificación de modo de excitar coeficientes de transformada tal que sea emulado el ruido detectado del lado de codificación. Consiguientemente, antes de discutir la funcionalidad del codificador de la Figura 5 adicionalmente, se hace una breve referencia a la Figura 6 que muestra un posible decodificador el cual podría emular el ruido de comodidad en el lado de decodificación según instrucción del codificador de la Figura 5. Más en general, la Figura 6 muestra una posible implementación de un decodificador que se ajusta al codificador de la Figura 1.

En particular, el decodificador de la Figura 6 comprende un motor de decodificación 160 como para decodificar la porción de secuencia de datos 44 durante las fases activas y una parte generadora de ruido de comodidad 162 para generar el ruido de comodidad en base a la información 32 y 38 provista en la secuencia de datos concerniente a las fases inactivas 28. La parte generadora de ruido de comodidad 162 comprende un generador aleatorio paramétrico 164, un FDNS 166 y un transformador (o sintetizador) inverso 168. Los módulos 164 y 168 están conectados en serie entre sí de modo que en la salida del sintetizador 168 resulta el ruido de comodidad, el cual llena el espacio entre la señal de audio reconstruida como es entregada por el motor de decodificación 160 durante las fases inactivas 28, como se discutió con respecto a la Figura 1. Los procesadores FDNS 166 y transformador inverso 168 pueden ser parte del motor de decodificación 160. En particular, pueden ser iguales al FDNS 116 y 118 de la Figura 4, por ejemplo.

El modo de funcionamiento y la funcionalidad de los módulos individuales de las Figuras 5 y 6 resultarán más claro a partir de la siguiente discusión.

En particular, el transformador 140 descompone espectralmente la señal de entrada en un espectrograma tal como usando una transformada lapped. Un estimador de ruido 146 está configurado para determinar parámetros de ruido a partir de ello. En forma concurrente, del detector de actividad de voz o sonido 16 evalúa los rasgos establecidos a partir de la señal de entrada de modo de detectar si tiene lugar una transición desde una fase activa a una fase inactiva o vice versa. Estos rasgos usados por el detector 16 pueden ser en forma de detector de componente transitorio/arranque, medición de tonalidad, y medición residual de LPC. El detector de componente transitorio/arranque puede ser usado para detectar ataque (aumento repentino de energía) o el comienzo de voz activa en una ambiente limpio o señal limpia de ruido; la medición de tonalidad puede ser usada para distinguir ruido de fondo útil tal como una sirena, teléfono sonando o música; residual LPC puede ser usado para obtener una indicación de la presencia de voz en la señal. Basado en estos rasgos, el detector 16 puede dar aproximadamente una información sobre si el cuadro actual puede ser clasificado, por ejemplo, como voz, silencio, música o ruido.

Mientras el estimador de ruido 146 puede ser responsable de distinguir el ruido de dentro del espectrograma del componente de señal útil allí dentro, tal como se propuso [R. Martin, Estimación de Densidad Espectral de Potencia de Ruido Basada en Alisamiento Óptimo y Estadística de Mínimos (*Noise Power Spectral Density Estimation Based*

on *Optimal Smoothing and Minimum Statistics*), 2001], el estimador de parámetros 148 puede ser responsable de analizar estadísticamente los componentes de ruido y determinar parámetros para cada componente espectral, por ejemplo, basado en el componente de ruido.

5 El estimador de ruido 146 puede estar configurado, por ejemplo, para buscar mínimos locales en el espectrograma y el estimador de parámetros 148 puede estar configurado para determinar la estadística de ruido en estas porciones suponiendo que los mínimos del espectrograma son principalmente un atributo de ruido de fondo más que sonido de primer plano.

10 Como una nota intermedia se pone énfasis en que también es posible realizar la estimación mediante el estimador de ruido sin el FDNS 142 ya que los mínimos sí ocurren en el espectro no corregido. La mayor parte de la descripción de la Figuras 5 quedaría igual.

15 El cuantificador de parámetros 152 a su vez puede ser configurado para parametrizar los parámetros estimados por el estimador de parámetros 148. Por ejemplo, los parámetros pueden describir una amplitud media y un momento de primero orden, u orden más alto, de una distribución de valores espectrales de dentro del espectrograma de la señal de entrada en tanto se refiere al componente de ruido. Para ahorrar cantidad de bits transmitidos, los parámetros pueden ser remitidos a la secuencia de datos para inserción dentro de la misma dentro de cuadros SID en una resolución espectral más baja que la resolución espectral provista por el transformador 140.

20 El medidor de estacionariedad 150 puede estar configurado para establecer una medida de estacionariedad para la señal de ruido. El estimador de parámetros 148 a su vez puede usar la medida de estacionariedad de modo de decidir si una actualización de parámetro debe ser iniciada o no enviando otro cuadro SID tal como el cuadro 38 de la Figura 1 o para influenciar la manera en que se estiman los parámetros.

25 El módulo 152 cuantifica los parámetros calculados por el estimador de parámetros 148 y el análisis de LP 144 y envía las señales al lado de decodificación. En particular, antes de cuantificar, los componentes espectrales pueden ser agrupados en grupos. Tal agrupamiento puede ser seleccionado de acuerdo con aspectos psicoacústicos tal como conformando la escala de Bark o algo similar. El detector 16 informa al cuantificador 152 si se necesita hacer la cuantificación o no. En el caso de no ser necesaria la cuantificación, deben seguir cuadros cero.

30 Cuando se transfiere la descripción a un escenario concreto de conmutación desde una fase activa a una fase inactiva, entonces los módulos de la Figura 5 actúan como sigue.

35 Durante una fase activa, el motor de codificación 14 sigue codificando la señal de audio vía el empaquetador en series de bits en el tiempo. La codificación puede realizarse a modo de cuadros. Cada cuadro de la secuencia de datos puede representar una porción/intervalo de tiempo de la señal de audio. El codificador de audio 14 puede estar configurado para codificar todos los cuadros que usan codificación LPC. El codificador de audio 14 puede estar configurado para codificar algunos cuadros como se describe con respecto a la Figura 2, llamado modo de codificación de cuadro TCX, por ejemplo. Los remanentes pueden ser codificados usando codificación de predicción lineal código—excitado (*code—excited linear prediction*) (CELP) tal como modo de codificación ACELP (*algebraic—code—excited linear prediction*), por ejemplo. Eso es, la porción 44 de la secuencia de datos puede comprender una actualización continua de coeficientes de LPC usando alguna tasa de transmisión de LPC que puede ser igual o mayor que la tasa de cuadro.

45 En paralelo, el estimador de ruido 146 inspecciona los espectros aplanados LPC (filtrado por análisis de LPC) de modo de identificar los mínimos  $k_{\min}$  de dentro del espectrograma TCX representado por la secuencia de estos espectros. Por supuesto, estos mínimos pueden variar con el tiempo  $t$ , esto es  $k_{\min}(t)$ . En todo caso, los mínimos pueden formar trazas en la salida del espectrograma mediante FDNS 142 y por ende, para cada espectro consecutivo  $i$  en el instante  $t_i$ , el mínimo puede ser asociable con los mínimos en espectro precedente y siguiente, respectivamente.

50 El estimador de parámetros entonces establece parámetros de estimación de ruido de fondo a partir de eso tal como, por ejemplo, una tendencia central (valor medio, mediana o similar)  $m$  y/o dispersión (desviación estándar, varianza o similar)  $d$  para diferentes componentes o bandas espectrales. La derivación puede involucrar análisis estadístico de los coeficientes espectrales consecutivos de los espectros del espectrograma en los mínimos, produciendo de ese modo  $m$  y  $d$  para cada mínimo en  $k_{\min}$ . La interpolación a lo largo de la dimensión espectral entre los mínimos de espectro antedichos puede realizarse de modo de obtener  $m$  y  $d$  para otros componentes o bandas espectrales predeterminados. La resolución espectral para la derivación y/o interpolación de la tendencia central (promedio medio) y la derivación de la dispersión (desviación estándar, varianza o similar) puede diferir.

60 Los parámetros recién mencionados son actualizados continuamente para cada salida de espectro por el FDNS 142, por ejemplo.

Tan pronto como el detector 16 detecta la entrada de una fase inactiva, el detector 16 puede informar al motor 14 consiguientemente para que no sean remitidos más cuadros activos al empaquetador 154. Sin embargo, el cuantificador 152 entrega los parámetros de ruido estadísticos recién mencionados en un primer cuadro SID dentro de la fase inactiva, en cambio. El primer cuadro SID puede comprender o no, una actualización de los LPCs. Si hay una actualización de LPC presente, la misma puede ser conducida adentro de la secuencia de datos en el cuadro SID 32 en el formato usado en la porción 44, esto es, durante la fase activa, tal como usando cuantificación en el dominio LSF/LSP, o en forma diferente, tal como usando ponderaciones espectrales correspondientes al análisis de LPC o a la función transferencia del filtro de síntesis de LPC tal como los que hubieran sido aplicados por el FDNS 142 dentro del marco de trabajo del motor de codificación 14 al proceder con una fase activa.

Durante la fase inactiva, el estimador de ruido 146, el estimador de parámetros 148 y el medidor de estacionariedad 150 sigue cooperando de modo de mantener actualizado el lado de decodificación sobre cambios en el ruido de fondo. En particular, el medidor 150 verifica la ponderación espectral definida por los LPCs, de modo de identificar cambios e informar al estimador 148 cuándo un cuadro SID debe ser enviado al decodificador. Por ejemplo, el medidor 150 podría activar el estimador consiguientemente siempre que la medida antedicha de estacionariedad indique un grado de fluctuación en los LPCs que excede una cierta magnitud. Adicionalmente o alternativamente, el estimador podría ser disparado para enviar los parámetros actualizados regularmente. Entre estos cuadros de actualización SID 40 no se enviaría nada en las secuencias de datos, esto es, "cuadros cero".

En el lado del decodificador, durante la fase activa, el motor de decodificación 160 asume la responsabilidad de reconstruir la señal de audio. Tan pronto como empieza la fase inactiva, el generador aleatorio de parámetros adaptivos 164 usa los parámetros del generador aleatorio descuantificados enviados durante la fase inactiva adentro de la secuencia de datos desde el cuantificador de parámetros 150, para generar componentes espectrales aleatorios, formando de ese modo un espectrograma aleatorio el cual es formado espectralmente adentro del procesador de energía espectral 166 con el sintetizador 168 realizando entonces una retransformación desde el dominio espectral al dominio del tiempo. Para la formación espectral adentro del FDNS 166, se puede usar, ya sea los coeficientes LPC más recientes provenientes de los cuadros activos más recientes, o bien la ponderación espectral a ser aplicada por el FDNS 166 puede ser derivada de allí mediante extrapolación, o bien el cuadro SID 32 en sí mismo puede conducir la información. Mediante esta medida, en el inicio de la fase inactiva, el FDNS 166 continúa ponderando espectralmente el espectro entrante de acuerdo con una función de transferencia de un filtro de síntesis de LPC, con el LPS que define el filtro de síntesis de LPC que se deriva de la porción de datos activos 44 o del cuadro SID 32. Sin embargo, con el inicio de la fase inactiva, el espectro a ser corregido por el FDNS 166 es el espectro generado aleatoriamente en lugar de una transformación codificada como en caso del modo de codificación de cuadro TCX. Asimismo, la corrección espectral aplicada en 166 es actualizada meramente en forma discontinua mediante el uso de los cuadros SID 38. Se podría realizar una interpolación o un debilitamiento para conmutar gradualmente desde una definición de corrección espectral a la siguiente durante las fases de interrupción 36.

Como se muestra en la Figura 6, el generador aleatorio paramétrico adaptivo 164 adicionalmente puede, opcionalmente, usar los coeficientes de transformada descuantificados como están contenidos dentro de las porciones más recientes de la última fase activa en la secuencia de datos, a saber, adentro de la porción de secuencia de datos 44 inmediatamente antes de la entrada de la fase inactiva. Por ejemplo, el significado puede ser entonces que se realiza una transición suave desde el espectrograma dentro de la fase activa al espectrograma aleatorio dentro de la fase inactiva.

Con referencia brevemente de nuevo a las Figuras 1 y 3, se deduce de las Figuras 5 y 6 (y la explicada subsiguientemente Figura 7) que la estimación de ruido de fondo paramétrica como es generada dentro del codificador y/o decodificador, puede comprender información estadística sobre una distribución de valores espectrales temporalmente consecutivos para distintas porciones espectrales tal como bandas de Bark o diferentes componentes espectrales. Para cada tal porción espectral, por ejemplo, la información estadística puede contener una medida de dispersión. La medida de dispersión, consiguientemente, sería definida en la información espectral de una manera resuelta espectralmente, a saber, muestreada en/para las porciones espectrales. La resolución espectral, esto es, el número de medidas para dispersión y tendencia central se desparramas a lo largo del eje espectral, pueden diferir entre, por ejemplo, medida de dispersión y la opcionalmente presente media o medida de tendencia central. La información estadística está contenida dentro de los cuadros SID. Se puede referir a un espectro corregido tal como el espectro filtrado de análisis LPC (esto es, LPC aplanado) tal como un espectro de MDCT corregido el cual permite síntesis sintetizando un espectro aleatorio de acuerdo con el espectro estadístico y des—corrigiendo el mismo de acuerdo con una función transferencia del filtro de síntesis de LPC. En ese caso, la información de corrección espectral puede estar presente dentro de los cuadros SID, a pesar de que puede no utilizarse en el primer cuadro SID 32, por ejemplo. Sin embargo, como se mostrará luego, esta información estadística, alternativamente, puede referir a un espectro no corregido. Asimismo, en lugar de usar una representación del espectro de valores reales tal como una MDCT, se puede usar un espectro de banco de filtros de valores complejos tal como espectro QMF de la señal de audio. Por ejemplo, se puede usar el espectro QMF de la señal de audio en forma no corregida y puede describirse en forma estadística mediante la información estadística en cuyo caso no hay corrección espectral más que la contenida dentro de la información estadística en sí misma.

En forma similar a la relación entre la Figura 3 con respecto a la Figura 1, la Figura 7 muestra una posible implementación del decodificador de la Figura 3. Como se muestra mediante el uso de los mismos signos de referencia que en la Figura 5, el decodificador de la Figura 7 puede comprender un estimador de ruido 146, un estimador de parámetros 148 y un medidor de estacionariedad 150, los cuales funcionan de manera similar a los mismos elementos que están en la Figura 5, con el estimador de ruido 146 de la Figura 7 que sin embargo, opera sobre el espectrograma transmitido y descuantificado tal como 120 o 122 de la Figura 4. El estimador de parámetros 146 entonces opera como el discutido en la Figura 5. Lo mismo vale con respecto al medidor de estacionariedad 148, el cual opera sobre la energía y los valores espectrales o datos de LPC revelando el desarrollo temporal del espectro del filtro de análisis de LPC (o del filtro de síntesis de LPC) según es transmitido y descuantificado vía/desde la secuencia de datos durante la fase activa.

Mientras los elementos 146, 148 y 150 actúan como el estimador de ruido de fondo 90 de la Figura 3, el decodificador de la Figura 7 también comprende un generador aleatorio paramétrico adaptivo 164 y un FDNS 166 así como también un transformador inverso 168 y ellos están conectados en serie entre sí como en la Figura 6, de modo de entregar el ruido de comodidad en la salida del sintetizador 168. Los módulos 164, 166, y 168 actúan como el generador de ruido de fondo 96 de la Figura 3 con el módulo 164 que asume la responsabilidad de la funcionalidad del generador aleatorio paramétrico 94. El generador aleatorio paramétrico adaptivo 94 o 164 entre componentes espectrales generadas aleatoriamente del espectrograma de acuerdo con los parámetros determinados por el estimador de parámetros 148 el cual a su vez es disparado usando la medida de estacionariedad entregada por el medidor de estacionariedad 150. El procesador 166 entonces corrige espectralmente el espectrograma así generado con el transformador inverso 168, realizando luego la transición desde el dominio espectral al dominio del tiempo. Notar que cuando durante la fase inactiva 88 el decodificador está recibiendo la información 108, el estimador de ruido de fondo 90 está realizando una actualización de las estimaciones de ruido seguida por algún medio de interpolación. Por lo demás, si se reciben cuadros cero, simplemente hace el procesamiento tal como interpolación y/o desvanecimiento.

Sintetizando las Figuras 5 a 7, éstas muestran que es técnicamente posible aplicar un generador aleatorio controlado 164 para excitar los coeficientes TCX, lo cuales pueden ser valores reales tal como en MDCT o valores complejos como en FFT. También podría ser ventajoso aplicar el generador aleatorio 164 sobre grupos de coeficientes usualmente conseguidos a través de bancos de filtros.

El generador aleatorio 164 preferiblemente es controlado tal que el mismo modela el tipo de ruido lo más fielmente posible. Esto se podría llevar a cabo si el ruido blando es conocido de antemano. Algunas aplicaciones pueden permitirlo. En muchas aplicaciones realistas donde un sujeto puede encontrar diferentes tipos de ruido, se requiere un método adaptivo como se muestra en las Figuras 5 a 7. Por consiguiente, se usa un generador aleatorio de parámetros adaptivo 164, el cual podría ser definido brevemente como  $g = f(\mathbf{x})$ , donde  $\mathbf{x} = (x_1, x_2, \dots)$  es un conjunto de parámetros de generador aleatorio provisto por los estimadores de parámetros 146 y 150, respectivamente.

Para hacer el generador aleatorio de parámetros adaptivo, el estimador de parámetros de generador aleatorio 146 controla adecuadamente el generador aleatorio. Se puede incluir compensación de deslizamiento para compensar los casos donde se consideran que los datos son estadísticamente insuficientes. Esto se hace para generar un modelo estadísticamente apareado del ruido basado en los cuadros pasados y siempre actualizará los parámetros estimados. Un ejemplo se da cuando se supone que el generador aleatorio 164 genera un ruido Gaussiano. En este caso, por ejemplo, sólo se pueden necesitar los parámetros media y varianza y se puede calcular un deslizamiento y aplicarlo a esos parámetros. Un método más avanzado puede manejar cualquier tipo de ruido o distribución y los parámetros no son necesariamente los momentos de una distribución.

Para el ruido no estacionario, se necesita tener una medida de estacionariedad y entonces se puede usar un generador aleatorio paramétrico menos adaptivo. La medida de estacionariedad determinada por el medidor 148 puede ser derivada de la forma espectral de la señal de entrada usando diversos métodos como por ejemplo, la medida de distancia de Itakura, la medida de distancia de Kullback—Leibler, etc.

Para manejar la naturaleza discontinua de actualizaciones de ruido enviadas a través de cuadros SID tal como se ilustra mediante 38 en la Figura 1, usualmente se envía información adicional tal como la energía y la forma espectral del ruido. Esta información es útil para generar el ruido en el decodificador teniendo una transición suave incluso durante un período de discontinuidad dentro de la fase inactiva. Finalmente, se pueden aplicar diversas técnicas de alisamiento o filtrado para ayudar a mejorar la calidad del emulador de ruido de comodidad.

Como ya se notó arriba, las Figuras 5 y 6 por un lado, y la Figura 7 por otro, pertenecen a diferentes escenarios. En un escenario que corresponde a las Figuras 5 y 6, la estimación de ruido de fondo paramétrica se hace en el codificador basado en la señal de entrada procesada y después los parámetros son transmitidos al decodificador. La Figura 7 corresponde al otro escenario donde el decodificador puede ocuparse de la estimación de ruido de fondo paramétrica basado en los cuadros pasados recibidos dentro de la fase activa. El uso de un detector de actividad de voz/señal o estimador de ruido puede ser beneficioso para ayudar a extraer componentes de ruido incluso durante voz activa, por ejemplo.

Entre los escenarios mostrados en las Figuras 5 a 7, el escenario de la Figura 7 puede ser preferido ya que este escenario resulta en que se transmite una menor cantidad de bits transmitidos. El escenario de las Figuras 5 y 6 sin embargo, tiene la ventaja de tener a disposición más precisa estimación de ruido.

5 Todas las realizaciones de arriba podrían ser combinadas con técnicas de extensión de ancho de banda tal como replicación de banda espectral (SBR), a pesar de que en general se puede usar extensión de ancho de banda.

10 Para ilustrar esto, ver la Figura 8. La Figura 8 muestra módulos mediante los cuales los codificadores de las Figuras 1 y 5 podrían ser extendidos para realizar codificación paramétrica con relación a una porción de más alta frecuencia de la señal de entrada. En particular, de acuerdo con la Figura 8, una señal de audio de entrada del dominio del tiempo es descompuesta espectralmente mediante un banco de filtros de análisis 200 tal como un banco de filtros de análisis QMF como se muestra en la Figura 8. Las técnicas de arriba de las Figuras 1 y 5 entonces sólo serían aplicadas a una porción de frecuencia más baja de la descomposición espectral generada por el banco de filtros 200. Para llevar información sobre la porción de frecuencia más alta al lado del decodificador, también se usa codificación paramétrica. Con este fin, un codificador de replicación de banda espectral regular 202 está configurado para parametrizar la porción de frecuencia más alta durante fases activas y alimentar información sobre ello en la forma de información de replicación de banda espectral dentro de la secuencia de datos al lado de decodificación. Un conmutador 204 puede estar provisto entre la salida del banco de filtros QMF 200 y la entrada del codificador de replicación de banda espectral 202 para conectar la salida del banco de filtros 200 con una entrada de un codificador de replicación de banda espectral 206 conectado en paralelo con el codificador 202 de modo de asumir la responsabilidad de la extensión de ancho de banda durante fases inactivas. Eso es, el conmutador 204 puede ser controlado como el conmutador 22 de la Figura 1. Como se describirá en más detalle abajo, el módulo de codificador de replicación de banda espectral 206 puede estar configurado para funcionar de manera similar al codificador de replicación de banda espectral 202: ambos pueden estar configurados para parametrizar la envolvente espectral de la señal de audio de entrada dentro de la porción de frecuencia más alta, esto es, la porción de frecuencia más alta remanente no sometida a codificación central por el motor de codificación, por ejemplo. Sin embargo, el módulo de codificador de replicación de banda espectral 206 puede usar una mínima resolución de tiempo/frecuencia a la cual la envolvente espectral es parametrizada y conducida adentro de la secuencia de datos, mientras el codificador de replicación de banda espectral 202 puede estar configurado para adaptar la resolución de tiempo/frecuencia a la señal de audio de entrada dependiendo de las ocurrencias de transitorios dentro de la señal de audio.

35 La Figura 9 muestra una implementación posible del módulo de codificación de extensión de ancho de banda 206. Un fijador de rejilla tiempo/frecuencia 208, un calculador de energía 210 y un codificador de energía 212 están conectados en serie entre sí entre una entrada y una salida del módulo de codificación 206. El fijador de rejilla de tiempo/frecuencia 208 puede estar configurado para fijar la resolución de tiempo/frecuencia en la cual se determina la envolvente de la porción de frecuencia más alta. Por ejemplo, una mínima resolución de tiempo/frecuencia permitida es usada continuamente por el módulo de codificación 206. El calculador de energía 210 entonces puede determinar la energía de la porción de frecuencia más alta del espectrograma entregado por el banco de filtros 200 dentro de la porción de frecuencia más alta en baldosas de tiempo/frecuencia correspondientes a la resolución de tiempo/frecuencia, y el codificador de energía 212 puede usar codificación de entropía, por ejemplo, para insertar las energías calculadas por el calculador 210 en la secuencia de datos 40 (ver Figura 1) durante las fases inactivas tal como dentro de cuadros SID, tal como el cuadro SID 38.

45 Se debe notar que la información de extensión de ancho de banda generada de acuerdo con las Figuras 8 y 9 también puede ser usada en conexión con usar un decodificador de acuerdo con cualquiera de los ejemplos de realización descritos arriba, tal como los de las Figuras 3, 4 y 7.

50 Así, las Figuras 8 y 9 aclaran que la generación de ruido de comodidad como se explicó con respecto a las Figuras 1 a 7, también puede ser usada en conexión con replicación de banda espectral. Por ejemplo, los codificadores y decodificadores de audio descritos arriba pueden operar en diferentes modos de operación, entre los cuales algunos comprenden replicación de banda espectral y algunos no. Los modos de operación de banda súper ancha por ejemplo, podrían involucrar replicación de banda espectral. En cualquier caso, los ejemplos de las Figuras 1 a 7 que muestran ejemplos para generar ruido de comodidad pueden ser combinadas con técnicas de extensión de ancho de banda de la manera descrita con respecto a las Figuras 8 y 9. El módulo de codificación de replicación de banda espectral 206 que es responsable de la extensión de ancho de banda durante las fases inactivas puede estar configurado para operar sobre una muy baja resolución de tiempo y frecuencia. Comparado con el procesamiento de replicación de banda espectral regular, el codificador 206 puede operar en una resolución de frecuencia diferente lo cual acarrea una tabla de banda de frecuencia adicional con resolución de muy baja frecuencia junto con filtros de alisamiento IIR en el decodificador para toda banda de factor de escala de generación de ruido de comodidad que interpola los factores de escala de energía aplicados en el ajustador de envolvente durante las fases inactivas. Como se mencionó recién, la rejilla de tiempo/frecuencia puede estar configurada para corresponder a una resolución temporal lo más baja posible.

Eso es, la codificación de extensión de ancho de banda puede ser realizada de modo diferente en QMF o dominio espectral dependiendo del silencio o fase activo que esté presente. En la fase activa, esto es, durante cuadros activos, se lleva a cabo codificación SBR regular por medio del codificador 202, dando por resultado una secuencia de datos SBR normal que acompaña a las secuencias de datos 44 102, respectivamente. En las fases inactivas o durante cuadros clasificados como cuadros SID, sólo se puede extraer información acerca de la envolvente espectral, representada como factores de escala de energía, mediante la aplicación de una rejilla de tiempo/frecuencia que exhibe una muy baja resolución de frecuencia, y por ejemplo, la resolución de tiempo más baja posible. Los factores de escala resultantes podrían ser codificados eficientemente mediante el codificador 212 y escritos a la secuencia de datos. En cuadros cero o durante fases de interrupción 36, no se puede escribir información lateral a la secuencia de datos mediante el módulo de codificación de replicación de banda espectral 206 y por lo tanto no se puede llevar a cabo ningún cálculo de energía mediante el calculador 210.

En conformidad con la Figura 8, la Figura 10 muestra una posible extensión de los ejemplos de realización del decodificador de las Figuras 3 y 7 a técnicas de codificación de extensión de ancho de banda. Para ser más precisos, la Figura 10 muestra un posible ejemplo de realización de un decodificador de audio de acuerdo con la presente invención. Un decodificador por núcleo 92 está conectado en paralelo con un generador de ruido de comodidad, estando indicado el generador de ruido de comodidad con el signo de referencia 220 y comprendiendo, por ejemplo, el módulo de generación de ruido 162 o los módulos 90, 94 y 96 de la Figura 3. Un conmutador 222 se muestra como distribuyendo los cuadros dentro de la secuencia de datos 104 y 30 respectivamente, sobre el decodificador por núcleo 92 o el generador de ruido de comodidad 220 dependiendo del tipo de cuadro, a saber, si el cuadro concierne o pertenece a una fase activa, o concierne o pertenece a una fase inactiva tal como cuadros SID o cuadros cero que conciernen a fases de interrupción. Las salidas del decodificador por núcleo 92 y del generador de ruido de comodidad 220 se conectan a una entrada de un decodificador de extensión de ancho de banda 224, la salida del cual revela la señal de audio reconstruida.

La Figura 11 muestra una realización más detallada de una posible implementación de un decodificador de extensión de ancho de banda 224.

Como se muestra en la Figura 11, el decodificador de extensión de ancho de banda 224 de acuerdo con la realización de la Figura 11 comprende un entrada 226 para recibir la reconstrucción del dominio del tiempo de la porción de baja frecuencia de la señal de audio completa a ser reconstruida. Es la entrada 226 que conecta el decodificador de extensión de ancho de banda 224 con las salidas del decodificador por núcleo 92 y el generador de ruido de comodidad 220 de modo que la entrada del dominio del tiempo en el entrada 226 pueda ser reconstruida la porción de baja frecuencia de una señal de audio que comprende tanto ruido como componente útil, o bien el ruido de comodidad generado para puentear el tiempo entre fases activas.

Como de acuerdo con la realización de la Figura 11, el decodificador de extensión de ancho de banda 224 está construido para realizar una replicación de ancho de banda espectral, el decodificador se llama decodificador SBR en lo que sigue. Con respecto a las Figuras 8 a 10, sin embargo, se enfatiza que estos ejemplos no están restringidos a replicación de ancho de banda espectral. Más bien, también se puede usar una manera alternativa, más general, de extensión de ancho de banda, con relación a estos ejemplos.

Además, el decodificador SBR 224 de la Figura 11 comprende una salida del dominio del tiempo 228 para entregar la señal de audio reconstruida, esto es, o bien en fases activas o bien en fases inactivas. Entre la entrada 226 y la salida 228 del decodificador SBR 224, están, conectados en serie en el orden en que se mencionan, un descomponedor espectral 230 el cual puede ser, como se muestra en la Figura 11, un banco de filtros de análisis tal como un banco de filtros de análisis QMF, un generador de HF 232, un ajustador de envolvente 234, y un convertidor del dominio espectral al tiempo 236 el cual puede estar, como se muestra en la Figura 11, realizado como un banco de filtros de síntesis tal como un banco de filtros de síntesis QMF.

Los módulos 230 a 236 operan como sigue. El descomponedor espectral 230 descompone espectralmente la señal de entrada del dominio del tiempo de modo de obtener una porción de baja frecuencia reconstruida. El generador de HF 232 genera una porción de réplica de alta frecuencia basado en porción de baja frecuencia reconstruida y el ajustador de envolvente 234 forma o corrige espectralmente la réplica de alta frecuencia usando una representación de una envolvente espectral de la porción de alta frecuencia como es llevada por medio de la secuencia de datos SBR y provista por módulos todavía no discutidos pero mostrados en la Figura 11 arriba del ajustador de envolvente 234. Así, el ajustador de envolvente 234 ajusta la envolvente de la porción de réplica de alta frecuencia de acuerdo con la representación de rejilla de tiempo/frecuencia de la envolvente de alta frecuencia transmitida, y remite la porción de alta frecuencia así obtenida al convertidor del dominio espectral a temporal 236 para una conversión del espectro de frecuencia completo, esto es, porción de alta frecuencia formada espectralmente junto con la porción de baja frecuencia reconstruida, a la señal del dominio del tiempo reconstruida en la salida 228.

Como ya se mencionó arriba con respecto a las Figuras 8 a 10, la envolvente espectral de porción de alta frecuencia puede ser llevada adentro de la secuencia de datos en la forma de factores de escala de energía y el decodificador SBR 224 comprende una entrada 238 para recibir esta información sobre la envolvente espectral de porciones de

alta frecuencia. Como se muestra en la Figura 11, en el caso de fases activas, esto es, cuadros activos presentes en la secuencia de datos durante fases activas, las entradas 238 pueden estar conectadas directamente a la entrada de envolvente espectral del ajustador de envolvente 234 vía un respectivo conmutador 240. Sin embargo, el decodificador SBR 224 comprende adicionalmente un combinador de factor de escala 242, un almacenamiento de datos de factor de escala 244, una unidad filtradora de interpolación 246, tal como una unidad filtradora IIR, y un ajustador de ganancia 248. Los módulos 242, 244, 246 y 248 están conectados en serie entre sí entre 238 y la entrada de envolvente espectral del ajustador de envolvente 234 con el conmutador 240 que están conectado entre el ajustador de ganancia 248 y el ajustador de envolvente 234 y un conmutador adicional 250 que está conectado entre el almacenamiento de datos de factor de escala 244 y la unidad filtradora 246. El conmutador 250 está configurado para conectar este almacenamiento de datos de factor de escala 244 con la entrada de la unidad filtradora 246, o bien un restaurador de datos de factor de escala 252. En el caso de cuadros SID durante fases inactivas — y opcionalmente en casos de cuadros activos para los cuales es aceptable una representación muy grosera de la envolvente espectral de porción de alta frecuencia — los conmutadores 250 y 240 conectan la secuencia de los módulos 242 a 248 entre la entrada 238 y el ajustador de envolvente 234. El combinador de factor de escala 242 adapta la resolución de frecuencia en la cual ha sido transmitida la envolvente espectral de porciones de alta frecuencia vía la secuencia de datos a la resolución, la cual espera recibir el ajustador de envolvente 234 y un almacenamiento de datos de factor de escala 244 almacena la envolvente espectral hasta una siguiente actualización. La unidad filtradora 246 filtra la envolvente espectral en dimensión temporal y/o espectral y el ajustador de ganancia 248 adapta la ganancia de la envolvente espectral de la porción de alta frecuencia. Con ese fin, el ajustador de ganancia puede combinar los datos de envolvente según fueron obtenidos por la unidad 246 con la envolvente real como se puede derivar de la salida del banco de filtros QMF. El restaurador de datos de factor de escala 252 reproduce los datos de factor de escala que representan la envolvente espectral dentro de fases de interrupción o cuadros cero según lo almacenado por el almacenamiento de factor de escala 244.

Así, en el lado del decodificador, se puede llevar a cabo el siguiente procesamiento. En los cuadros activos o durante fases activas, se puede aplicar procesamiento de replicación de banda espectral regular. Durante estos períodos, los factores de escala provenientes de la secuencia de datos, los cuales típicamente están disponibles para un mayor número de bandas de factor de escala comparado con procesamiento de generación de ruido de comodidad, son convertidos a la resolución de frecuencia de generación de ruido de comodidad por el combinador de factor de escala 242. El combinador de factor de escala combina los factores de escala para la resolución de frecuencia más alta para resultar en un número de factores de escala conforme con CNG aprovechando bordes de banda de frecuencia comunes de las diferentes tablas de banda de frecuencia. Los valores de factor de escala resultantes en la salida de la unidad combinadora de factor de escala 242 son almacenados para la reutilización en cuadros cero y posterior reproducción por el recuperador 252 y son usados subsiguientemente para actualizar la unidad filtradora 246 para el modo de operación CNG. En los cuadros SID se aplica un lector de secuencia de datos SBR modificado el cual extrae la información de factor de escala de la secuencia de datos. La configuración restante del procesamiento SBR es inicializada con valores predefinidos, se inicializa la rejilla de tiempo/frecuencia a la misma resolución de tiempo/frecuencia usada en el codificador. Los factores de escala extraídos son alimentados a la unidad filtradora 246 donde, por ejemplo, un filtro de alisamiento IIR interpola la progresión de la energía para una banda de factor de escala de baja resolución a lo largo del tiempo. En caso de caso cuadros cero, no hay carga útil para leer desde la serie de bits en el tiempo y la configuración SBR que incluye la rejilla de tiempo/frecuencia es la misma que la usada en los cuadros SID. En cuadros cero, los filtros de alisamiento de la unidad filtradora 246 son alimentados con un valor de factor de escala entregado desde la unidad combinadora de factores de escala 242 que han sido almacenados en el último cuadro que contiene información de factores de escala válida. En el caso que el cuadro actual es clasificado como un cuadro inactivo o cuadro SID, el ruido de comodidad es generado en el dominio TCX y se vuelve a transformar al dominio del tiempo. Subsiguientemente, la señal del dominio del tiempo que contiene el ruido de comodidad es alimentada al banco de filtros de análisis QMF 230 del módulo SBR 224. En el dominio de QMF, la extensión de ancho de banda del ruido de comodidad se realiza por medio de transposición de copy—up dentro del generador de HF 232 y finalmente la envolvente espectral de la parte de alta frecuencia creada artificialmente, es ajustada mediante aplicación de información de factores de escala de energía en el ajustador de envolvente 234. Estos factores de escala de energía se obtienen mediante la salida de la unidad filtradora 246 y son puestos en escala mediante la unidad de ajuste de ganancia 248 antes de la aplicación en el ajustador de envolvente 234. En esta unidad de ajuste de ganancia 248 se calcula un valor de ganancia para ajustar los factores de escala y se aplica para compensar grandes diferencias de energía en el borde entre la porción de baja frecuencia y el contenido de alta frecuencia de la señal.

Las realizaciones descritas arriba se usan en común en las Figuras 12 y 13. La Figura 12 muestra un codificador de audio de acuerdo con una realización de comparación, y la Figura 13 muestra una realización de un decodificador de audio. Los detalles revelados con respecto a estas figuras se aplican igual a los elementos mencionados previamente en forma individual.

El codificador de audio de la Figura 12 comprende un banco de filtros de análisis QMF 200 para descomponer espectralmente una señal de audio de entrada. Un detector 270 y un estimador de ruido 262 están conectados a una salida del banco de filtros de análisis QMF 200. El estimador de ruido 262 asume la responsabilidad de la funcionalidad del estimador de ruido de fondo 12. Durante las fases activas, los espectros QMF provenientes del

banco de filtros de análisis QMF son procesador por una conexión en paralelo de un estimador de parámetros de replicación de banda espectral 260 seguido por algún codificador SBR 264 por un lado, y una concatenación de un banco de filtros de síntesis QMF 272 seguido por un codificador por núcleo 14 por el otro lado. Ambos caminos paralelos se conectan a una respectiva salida del empaquetador de serie de bits en el tiempo 266. En el caso de cuadros SID salientes, el codificador de cuadro SID 274 recibe los datos desde el estimador de ruido 262 y entrega los cuadros SID al empaquetador de serie de bits en el tiempo 266.

Los datos de extensión de ancho de banda espectral entregados por el estimador 260 describen la envolvente espectral de la porción de alta frecuencia del espectrograma o espectro entregado por el banco de filtros de análisis QMF 200, los cuales luego son codificados, tal como mediante codificación por entropía, por el codificador SBR 264. El multiplexador de secuencia de datos 266 inserta los datos de extensión de ancho de banda espectral de fases activas en la secuencia de datos entregada en una salida 268 del multiplexador 266.

El detector 270 detecta si actualmente está activa una fase activa o inactiva. En base a esta detección, un cuadro activo, un cuadro SID o un cuadro cero, este cuadro inactivo, se ha de entregar actualmente. En otras palabras, el módulo 270 decide si una fase activa o una fase inactiva está activa y si la fase inactiva está activa, si se ha de entregar o no un cuadro SID. Las decisiones están indicadas en la Figura 12 usando I para cuadros cero, A para cuadros activos, y S para cuadros SID. Los cuadros que corresponden a intervalos de tiempo de la señal de entrada donde la fase activa está presente, también son remitidos a la concatenación del banco de filtros de síntesis QMF 272 y al codificador por núcleo 14. El banco de filtros de síntesis QMF 272 tiene una resolución de menor frecuencia u opera a un menor número de sub-bandas QMF cuando se compara con el banco de filtros de análisis QMF 200 de modo de lograr por medio del cociente de número de sub-bandas una reducción de la tasa de muestreo para transferir las porciones de cuadro activas de la señal de entrada al dominio del tiempo nuevamente. En particular, el banco de filtros de síntesis QMF 272 es aplicado a las porciones de menor frecuencia o sub-bandas de menor frecuencia del espectrograma de banco de filtros de análisis QMF dentro de los cuadros activos. El codificador por núcleo 14 recibe así una versión de tasa de muestreo reducida de la señal de entrada, la cual convierte así meramente una porción de menor frecuencia de la señal de entrada original ingresada al banco de filtros de análisis QMF 200. La restante porción de frecuencia más alta es codificada paramétricamente por los módulos 260 y 264.

Los cuadros SID (o, para ser más precisos, la información a ser llevada por los mismos) son remitidos al codificador SID 274 el cual asume la responsabilidad de las funcionalidades del módulo 152 de la Figura 5, por ejemplo. La única diferencia: el módulo 262 opera sobre el espectro de la señal de entrada directamente — sin corrección por LPC. Asimismo, como se usa el filtrado de análisis QMF, la operación del módulo 262 es independiente del modo de cuadro elegido por el codificador por núcleo o la opción de extensión de ancho de banda espectral que se aplique o no. Las funcionalidades del módulo 148 y 150 de la Figura 5 pueden ser implementadas dentro del módulo 274.

El multiplexador 266 multiplexa la respectiva información codificada formando la secuencia de datos en la salida 268.

El decodificador de audio de la Figura 13 puede operar sobre una secuencia de datos como la entregada por el codificador de la Figura 12. Eso es, un módulo 280 está configurado para recibir la secuencia de datos y para clasificar los cuadros dentro de la secuencia de datos en cuadros activos, cuadros SID y cuadros cero, esto es, una carencia de cuadro en la secuencia de datos, por ejemplo. Los cuadros activos son remitidos a una concatenación de un decodificador por núcleo 92, un banco de filtros de análisis QMF 282 y un módulo de extensión de ancho de banda espectral 284. Opcionalmente, un estimador de ruido 286 está conectado a la salida del banco de filtros de análisis QMF. El estimador de ruido 286 puede operar en forma similar, y puede asumir la responsabilidad de las funcionalidades de, el estimador de ruido de fondo 90 de la Figura 3, por ejemplo, con la excepción de que el estimador de ruido opera sobre el espectro no corregido en lugar de sobre los espectros de excitación. La concatenación de módulos 92, 282 y 284 se conecta a una entrada del banco de filtros de síntesis QMF 288. Los cuadros SID son remitidos a un decodificador de cuadros SID 290 el cual asume la responsabilidad de la funcionalidad del generador de ruido de fondo 96 de la Figura 3, por ejemplo. Un actualizador de parámetros de generación de ruido de comodidad 292 es alimentado por la información proveniente del decodificador 290 y el estimador de ruido 286 con este actualizador 292 que gobierna el generador aleatorio 294 el cual asume la responsabilidad de la funcionalidad de los generadores aleatorios paramétricos de la Figura 3. Como faltan cuadros inactivos o cero, no tiene que ser remitidos a ningún lado, pero ellos disparan otro ciclo de generación aleatoria del generador aleatorio 294. La salida del generador aleatorio 294 está conectada al banco de filtros de síntesis QMF 288, la salida del cual revela la señal de audio reconstruida en silencio y fases activas en el dominio del tiempo.

Así, durante las fases activas, el decodificador por núcleo 92 reconstruye la porción de baja frecuencia de la señal de audio incluyendo tanto ruido como componentes de señal útiles. El banco de filtros de análisis QMF 282 descompone espectralmente la señal reconstruida y el módulo de extensión de ancho de banda espectral 284 usa la información de extensión de ancho de banda espectral dentro de la secuencia de datos y cuadros activos, respectivamente, para agregar la porción de alta frecuencia. El estimador de ruido 286, si está presente, realiza la estimación de ruido basado en una porción de espectro según lo reconstruido por el decodificador por núcleo, esto es, la porción de baja frecuencia. En las fases inactivas, los cuadros SID llevan información que describe

paramétricamente la estimación de ruido de fondo derivada por la estimación de ruido 262 en el lado del codificador. El actualizador de parámetros 292 puede usar principalmente la información del codificador para actualizar su estimación de ruido de fondo paramétrica, usando la información provista por el estimador de ruido 286 principalmente como una posición de soporte de información para caso de pérdida de transmisión concerniente a cuadros SID. El banco de filtros de síntesis QMF 288 convierte la señal espectralmente descompuesta como salida mediante el módulo de replicación de banda espectral 284 en fases activas y el espectro de señal generada de ruido de comodidad en el dominio del tiempo. Así, las Figuras 12 y 13 ponen en claro que se puede usar un marco de trabajo de banco de filtros QMF como una base para generación de ruido de comodidad basado en QMF. El marco de trabajo QMF provee una manera conveniente para reducir la tasa de muestreo de la señal de entrada a la tasa de muestreo del codificador por núcleo, para aumentar la tasa de muestreo de la señal de salida del codificador por núcleo del decodificador por núcleo 92 en el lado del decodificador usando el banco de filtros de síntesis QMF 288. Al mismo tiempo, el marco de trabajo QMF también puede ser usado en combinación con extensión de ancho de banda para extraer y procesar los componentes de alta frecuencia de la señal los cuales son postergados por los módulos de codificador por núcleo y decodificador por núcleo 14 y 92. Consiguientemente, el banco de filtros QMF puede ofrecer un marco de trabajo común para diversas herramientas de procesamiento de señal. De acuerdo con las Figuras 12 y 13, la generación de ruido de comodidad se incluye exitosamente dentro de este marco de trabajo.

En particular, de acuerdo con las Figuras 12 y 13, se puede ver que es posible generar ruido de comodidad en el lado del decodificador después del análisis QMF, pero antes de la síntesis QMF aplicando un generador aleatorio 294 para excitar las partes real e imaginaria de cada coeficiente QMF del banco de filtros de síntesis QMF 288, por ejemplo. La amplitud de las secuencias aleatorias, por ejemplo, son computadas individualmente en cada banda QMF tal que el espectro del ruido de comodidad generado se parezca al espectro de la señal de ruido de fondo de entrada real. Esto se puede lograr en cada banda QMF usando un estimador de ruido después del análisis QMF en el lado de codificación. Estos parámetros entonces pueden ser transmitidos a través de los cuadros SID para actualizar la amplitud de las secuencias aleatorias aplicada en cada banda QMF en el lado del decodificador.

Idealmente, notar que la estimación de ruido 262 aplicada en el lado del codificador debe poder operar tanto durante períodos inactivos (esto es, sólo ruido) como activos (que típicamente contienen voz ruidos) de modo que los parámetros de ruido de comodidad puedan ser actualizados inmediatamente al final de cada período activo. Además, la estimación de ruido podría ser usada también del lado del decodificador. Como se descartan los cuadros de sólo ruido en un sistema de codificación/decodificación basado en DTX, la estimación de ruido en el lado del decodificador puede, favorablemente, operar sobre contenidos de voz ruidosos. La ventaja de realizar la estimación de ruido en el lado de decodificador, además de en el lado de codificador, es que la forma espectral del ruido de comodidad puede ser actualizada incluso cuando la transmisión de paquete desde el codificador al decodificador falla para el(los) primer(os) cuadro(s) SID siguientes a un período de actividad.

La estimación de ruido debe poder seguir, precisa y rápidamente, variaciones del contenido espectral del ruido de fondo e idealmente lo debe poder realizar durante ambos, cuadros, activo e inactivo, como se expresó arriba. Una manera de lograr estos objetivos es rastrear los mínimos tomados en cada banda por el espectro de potencia usando una ventana deslizante de longitud finita, como se propone en [R. Martin, Estimación de Densidad Espectral de Potencia de Ruido Basada en Alisamiento Óptimo y Estadística de Mínimos (*Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics*), 2001]. La idea detrás de esto es que la potencia de un espectro de voz ruidosa frecuentemente decae a la potencia del ruido de fondo, por ejemplo, entre palabras o sílabas. Rastrear el mínimo del espectro de potencia, por lo tanto, provee una estimación del piso de ruido en cada banda, incluso durante actividad de voz. Sin embargo, estos pisos de ruido son subestimados en general. Asimismo, ellos no permiten capturar rápidas fluctuaciones de las potencias espectrales, especialmente repentinos incrementos de energía.

De todos modos, el piso de ruido computado como se describe arriba en cada banda provee información lateral muy útil para aplicar una segunda etapa de estimación de ruido. De hecho, podemos esperar que la potencia de un espectro ruidoso esté cerca del piso de ruido estimado durante inactividad, mientras que la potencia espectral estará muy por arriba del piso de ruido durante la actividad. Los pisos de ruido computados separadamente en cada banda pueden ser usados entonces como detectores groseros de actividad para cada banda. En base a este conocimiento, se puede estimar fácilmente la potencia de ruido de fondo como una versión recursivamente alisada del espectro de potencia como sigue:

$$\sigma_N^2(m, k) = \beta(m, k) \cdot \sigma_N^2(m - 1, k) + (1 - \beta(m, k)) \cdot \sigma_X^2(m, k) ,$$

donde  $\sigma_X^2(m, k)$  denota la densidad espectral de potencia de la señal de entrada en el cuadro  $m$  y banda  $k$ ,  $\sigma_N^2(m, k)$  refiere a la estimación de potencia de ruido, y  $\beta(m, k)$  es un factor de olvido (necesariamente entre 0 y 1) que controla la magnitud del alisamiento para cada banda y cada cuadro separadamente. Usar la información de piso de ruido para reflejar el estado de actividad, debe tomar un pequeño valor durante períodos inactivos (esto es, cuando el espectro de potencia está cerca del piso de ruido), mientras se debe elegir un valor alto para aplicar más

alisamiento (idealmente manteniendo  $\sigma_N^2(m,k)$  constante) durante cuadros activos. Para lograr esto se puede tomar una decisión débil calculando los factores de olvido como sigue:

$$\beta(m,k) = 1 - e^{-\alpha \left( \frac{\sigma_N^2(m,k)}{\sigma_{NF}^2(m,k)} - 1 \right)},$$

5 donde  $\sigma_{NF}^2$  es la potencia de piso de ruido y  $\alpha$  es un parámetro de control. Un valor más alto para  $\alpha$  da por resultado factores de olvido más grandes y por ende más alisamiento global.

10 Así, se ha descrito un concepto de Generación de Ruido de Comodidad (CNG) donde se produce ruido artificial en el lado del decodificador en un dominio de transformada. Las realizaciones de arriba pueden ser aplicadas en combinación con virtualmente cualquier tipo de herramienta de análisis espectro—temporal (esto es, una transformación o un banco de filtros) que descompone una señal del dominio del tiempo en múltiples bandas espectrales.

15 Por ende, las realizaciones de arriba, inter alias, describieron un CNG basado en TCX donde un generador de ruido de comodidad emplea pulsos aleatorio para modelar lo residual.

20 A pesar de que se han descrito algunos aspectos en el contexto de un aparato, es claro que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a una etapa de método o a una característica de una etapa de método. Análogamente, los aspectos descritos en el contexto de una etapa de método también representan una descripción de un correspondiente bloque o componente o característica de un correspondiente aparato. Algunos o todas las etapas de método pueden ser ejecutadas por (o usando) un aparato de hardware, como por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico.

25 En algunas realizaciones, alguno o más de los pasos de método más importantes pueden ser ejecutados mediante un aparato tal.

30 Dependiendo de ciertos requerimientos de implementación, las realizaciones de la invención pueden ser implementadas en hardware o en software. La implementación puede ser realizada usando un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un CD, un Blu—Ray, una memoria de sólo lectura (ROM), una PROM, una EPROM, una EEPROM o una memoria FLASH, teniendo señales de control legibles electrónicamente almacenadas en las mismas, las cuales cooperan (o son capaces de cooperar) con un sistema de ordenador programable tal que se ejecute el respectivo método. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

35 Algunas realizaciones de acuerdo con la invención comprenden un portador de datos que tiene señales de control legibles electrónicamente, las cuales son capaces de cooperar con un sistema de ordenador programable, tal que uno de los métodos descritos en la presente sea ejecutado.

40 Generalmente, realizaciones de la presente invención pueden ser implementadas como un producto de programa de ordenador con un código de programa, siendo código de programa operativo para ejecutar uno de los métodos cuando el producto de programa de ordenador se ejecuta en un ordenador. El código de programa puede ser almacenado, por ejemplo, sobre un portador legible por una máquina.

45 Otras realizaciones comprenden el programa de ordenador para ejecutar uno de los métodos descritos en la presente, almacenado en un portador legible por una máquina.

50 En otras palabras, una realización del método inventivo es, por lo tanto, un programa de ordenador con un código de programa para ejecutar uno de los métodos descritos en la presente, cuando el programa de ordenador se ejecuta en un ordenador.

55 Una realización adicional de los métodos inventivos es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa de ordenador para ejecutar uno de los métodos descritos en la presente. El portador de datos, el medio de almacenamiento digital o el medio de registro típicamente son tangibles y/o no transitorios.

60 Una realización adicional del método inventivo es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa de ordenador para ejecutar uno de los métodos descritos en la presente. El flujo de datos o la secuencia de señales pueden ser configurados, por ejemplo, para ser transferidos vía una conexión de comunicación de datos, por ejemplo, vía Internet.

Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado para o adaptado para ejecutar uno de los métodos descritos en la presente.

5 Una realización adicional comprende un ordenador que tiene instalado en él el programa de ordenador para ejecutar uno de los métodos descritos aquí.

10 Otra realización de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa de ordenador para ejecutar uno de los métodos descritos aquí, a un receptor. El receptor puede ser, por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o sistema puede comprender, por ejemplo, un servidor de archivos para transferir el programa de ordenador al receptor.

15 En algunas realizaciones se puede usar un dispositivo de lógica programable (por ejemplo un arreglo de compuesta programable de campo) para realizar algunas o todas las funcionalidades de los métodos descritos en la presente. En algunas realizaciones, el arreglo de compuerta programable de campo puede cooperar con un microprocesador para realizar uno de los métodos descritos en la presente. Generalmente, los métodos preferiblemente son realizados mediante algún aparato de hardware.

20 Las realizaciones que se describieron más arriba son puramente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y variaciones posibles de las disposiciones y de los detalles descritos en la presente serán evidentes para los expertos en la materia. Por lo tanto, es la intención que la invención esté limitada sólo por el alcance de las siguientes reivindicaciones de patente y no por los detalles específicos presentados por la descripción y la explicación de las realizaciones en la presente.

**REIVINDICACIONES**

1. Un decodificador de audio para decodificar una secuencia de datos para reconstruir una señal de audio a partir de la misma, en el que la secuencia de datos comprende al menos una fase activa (86) seguida por una fase inactiva (88), donde el decodificador de audio comprende:
  - un estimador de ruido de fondo (90) configurado para actualizar continuamente una estimación de ruido de fondo paramétrica a partir de la secuencia de datos (104) durante la fase activa (86);
  - un decodificador (92) configurado para reconstruir la señal de audio a partir de la secuencia de datos durante la fase activa;
  - un generador aleatorio paramétrico (94);
  - un generador de ruido de fondo (96) configurado para sintetizar la señal de audio durante la fase inactiva (88) controlando el generador aleatorio paramétrico (94) durante la fase inactiva (88) dependiendo de la estimación de ruido de fondo paramétrica;
  - donde el decodificador (92) está configurado para, al reconstruir la señal de audio a partir de la secuencia de datos, conformar una señal de excitación codificada por transformada en la secuencia de datos, de acuerdo a coeficientes de predicción lineal también codificados en la secuencia de datos; y
  - donde el estimador de ruido de fondo (90) está configurado para actualizar la estimación de ruido de fondo paramétrica utilizando la señal de excitación.
2. Un decodificador de audio de acuerdo con la reivindicación 1, en donde el estimador de ruido de fondo está configurado para, al actualizar la estimación de ruido de fondo paramétrica, identificar mínimos locales en la señal de excitación y realizar un análisis estadístico de la señal de excitación en los mínimos locales para derivar la estimación de ruido de fondo paramétrica.
3. Un decodificador de audio de acuerdo con una cualquiera de las reivindicaciones anteriores, en donde el decodificador está configurado para, al reconstruir la señal de audio, utilizar decodificación predictiva y/o por transformada para reconstruir una porción de frecuencia más baja de la señal de audio a partir de la secuencia de datos, y para sintetizar una porción de frecuencia más alta de la señal de audio.
4. Un decodificador de audio de acuerdo con la reivindicación 3, en donde el decodificador está configurado para sintetizar la porción de frecuencia más alta de la señal de audio a partir de una envolvente espectral de la porción de frecuencia más alta de la señal de audio de entrada, codificada paraméricamente en la secuencia de datos, o para sintetizar la porción de frecuencia más alta de la señal de audio mediante extensión ciega de ancho de banda basada en la porción de frecuencia más baja.
5. Un decodificador de audio de acuerdo con la reivindicación 4, en donde el decodificador está configurado para interrumpir la decodificación predictiva y/o por transformada en fases inactivas y realizar la síntesis de la porción de frecuencia más alta de la señal de audio formando espectralmente una réplica de la porción de frecuencia más baja de la señal de audio de acuerdo con la envolvente espectral en la fase activa, y formando espectralmente una réplica de la señal de audio sintetizada de acuerdo a la envolvente espectral en la fase inactiva.
6. Un decodificador de audio de acuerdo con la reivindicación 4 ó 5, en donde el decodificador comprende un banco de filtros inverso para componer espectralmente la señal de audio de entrada a partir de un conjunto de sub-bandas de la porción de frecuencia más baja, y un conjunto de sub-bandas de la porción de frecuencia más alta.
7. Un decodificador de audio de acuerdo con una cualquiera de las reivindicaciones 1 a 6, en donde el decodificador de audio está configurado para detectar una entrada de la fase inactiva siempre que se interrumpa la secuencia de datos, y/o siempre que la secuencia de datos señale la entrada de la secuencia de datos.
8. Un decodificador de audio de acuerdo con una cualquiera de las reivindicaciones 1 a 7, en donde el generador de ruido de fondo (96) está configurado para sintetizar la señal de audio durante la fase inactiva (88) controlando el generador aleatorio paramétrico (94) durante la fase inactiva (88) dependiendo del ruido de fondo paramétrico tal y como es actualizado continuamente por el estimador de ruido de fondo solo en caso de ausencia de cualquier información de estimación de ruido de fondo paramétrica en la secuencia de datos inmediatamente después de una transición de una fase activa a una fase inactiva.
9. Un método de decodificación de audio para decodificar una secuencia de datos para reconstruir una señal de audio a partir de la misma, en el que la secuencia de datos comprende al menos una fase activa (86) seguida por una fase inactiva (88), donde el método comprende:
  - actualizar continuamente una estimación de ruido de fondo paramétrica a partir de la secuencia de datos (104) durante la fase activa (86);
  - reconstruir la señal de audio a partir de la secuencia de datos durante la fase activa;
  - sintetizar la señal de audio durante la fase inactiva (88) controlando un generador aleatorio paramétrico (94) durante la fase inactiva (88) dependiendo de la estimación de ruido de fondo paramétrica;

donde la reconstrucción de la señal de audio a partir de la secuencia de datos comprende conformar una señal de excitación codificada por transformada en la secuencia de datos, de acuerdo a coeficientes de predicción lineal también codificados en la secuencia de datos; y

5 donde la actualización continua de la estimación de ruido de fondo paramétrica se lleva a cabo utilizando la señal de excitación.

10. Un programa de ordenador que tiene un código de programa configurado para implementar, cuando se ejecuta en un ordenador, un método de acuerdo con la reivindicación 9.

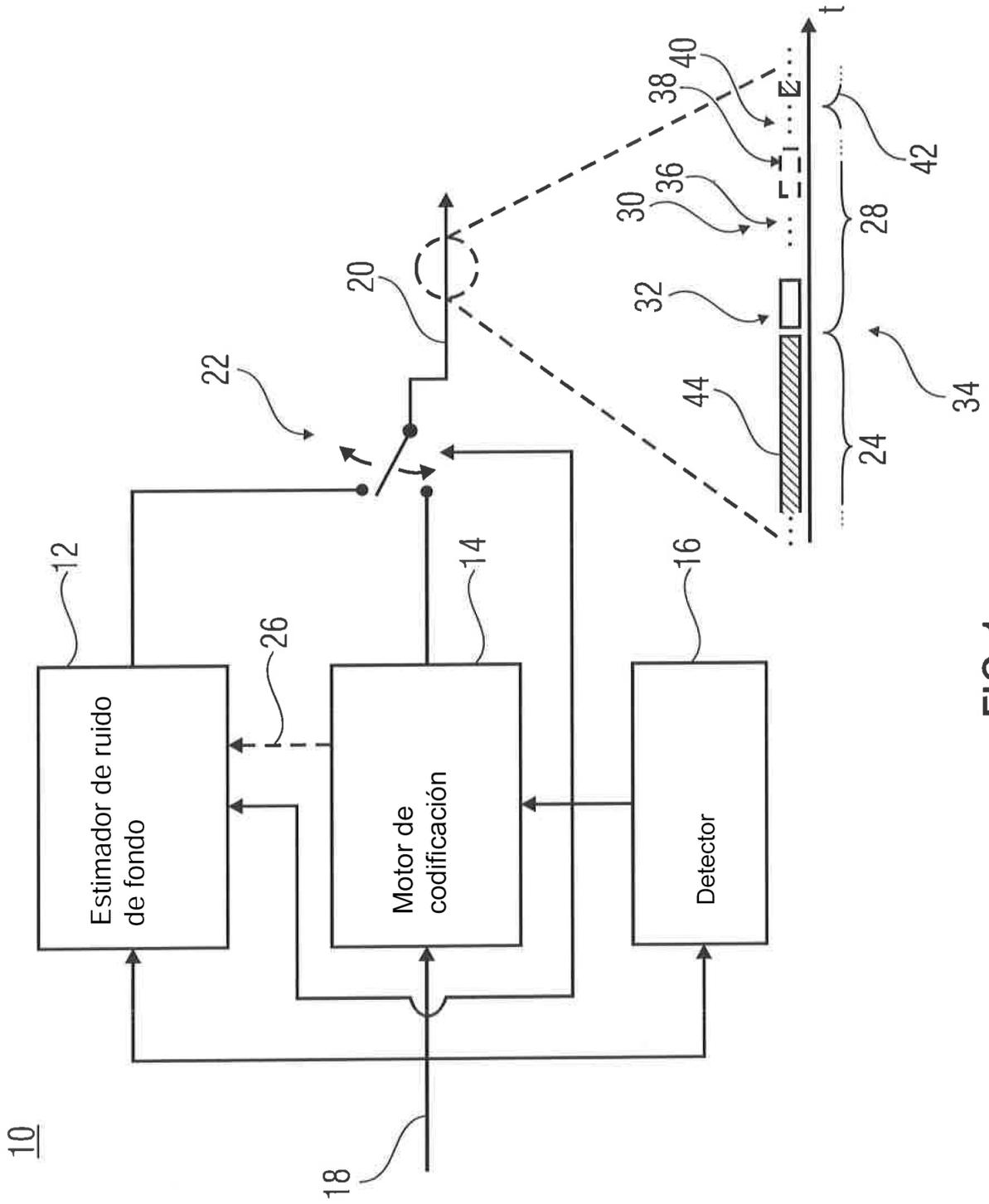


FIG 1

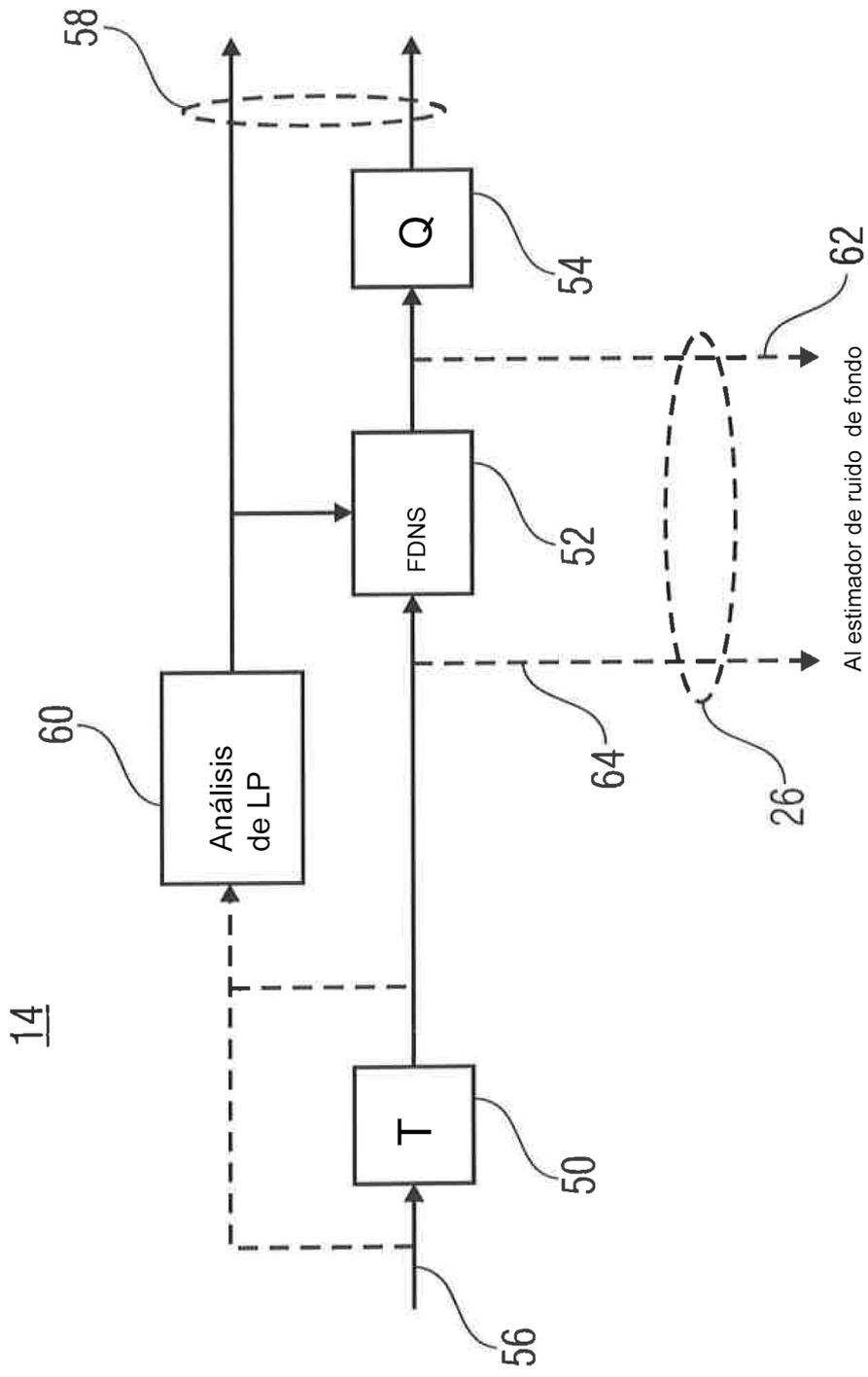


FIG 2

80

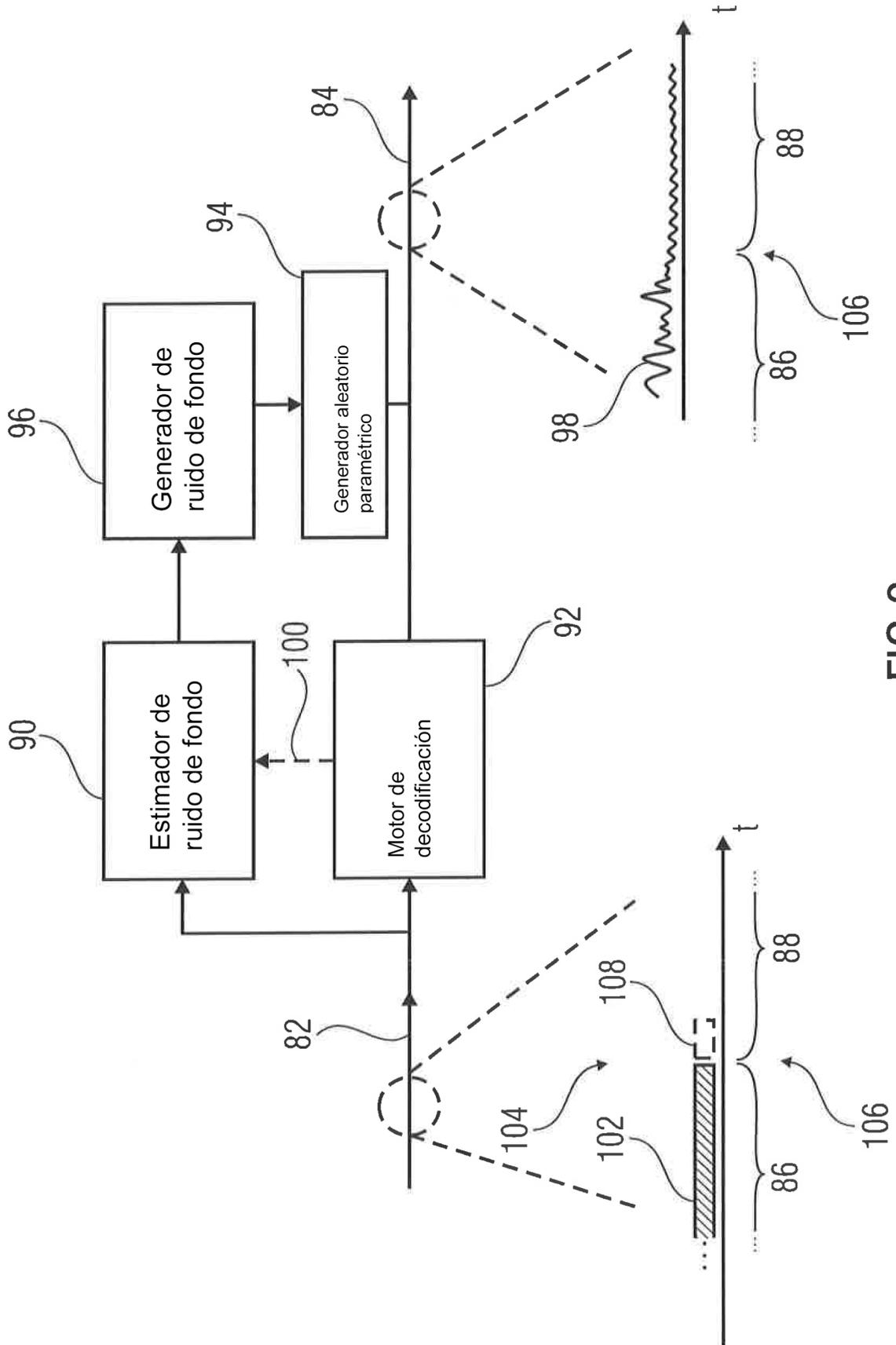


FIG 3

92

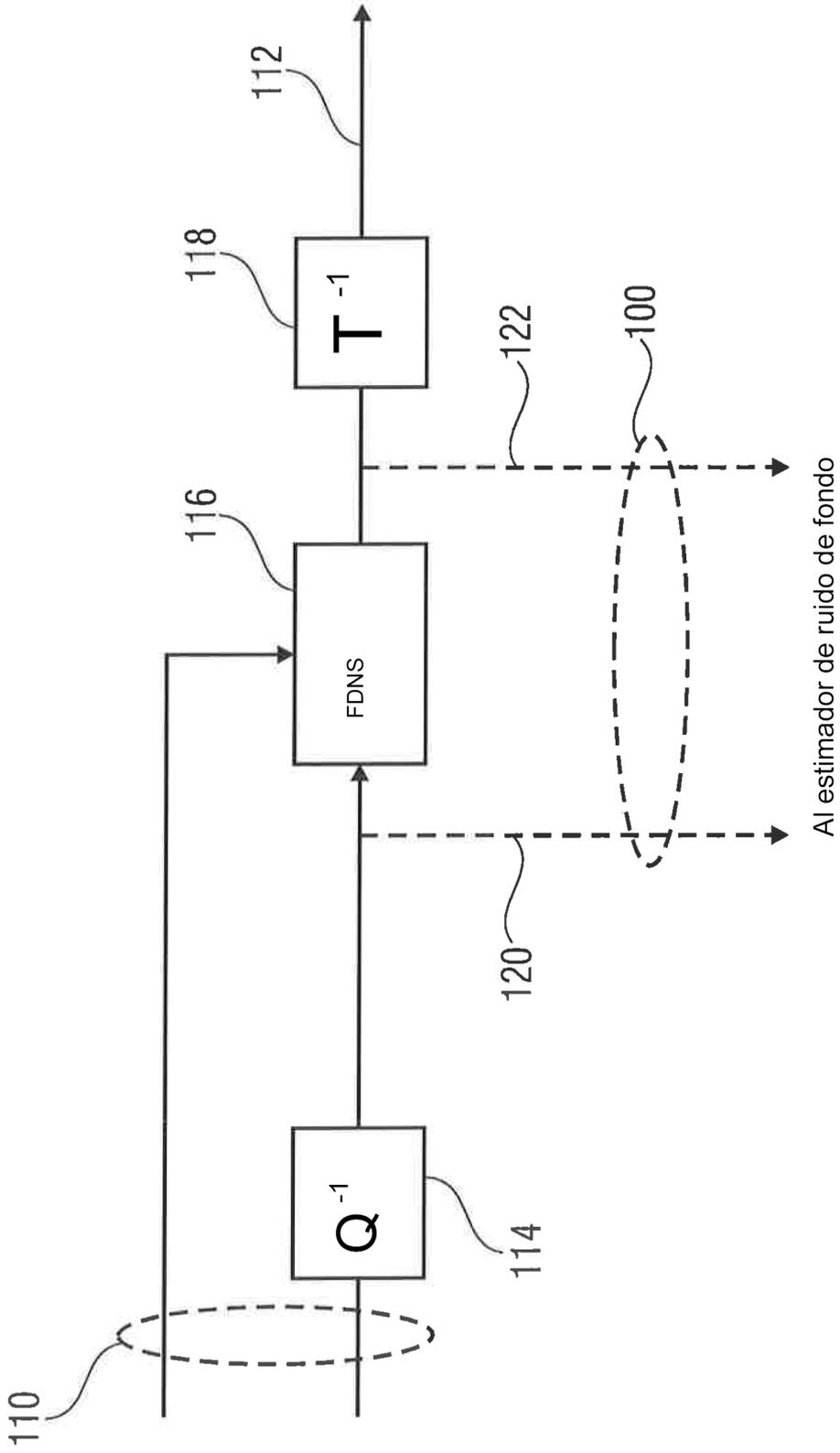


FIG 4

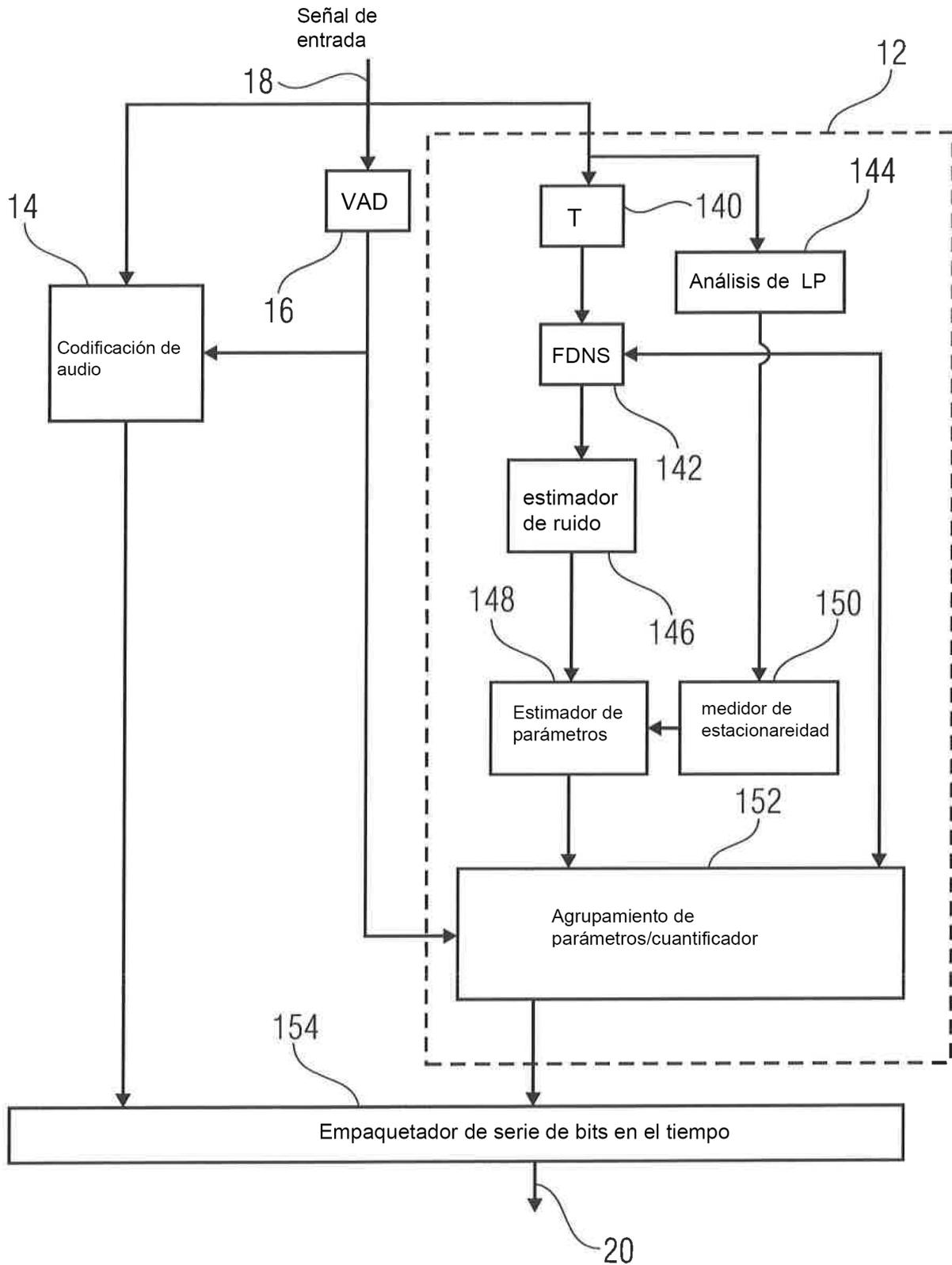


FIG 5

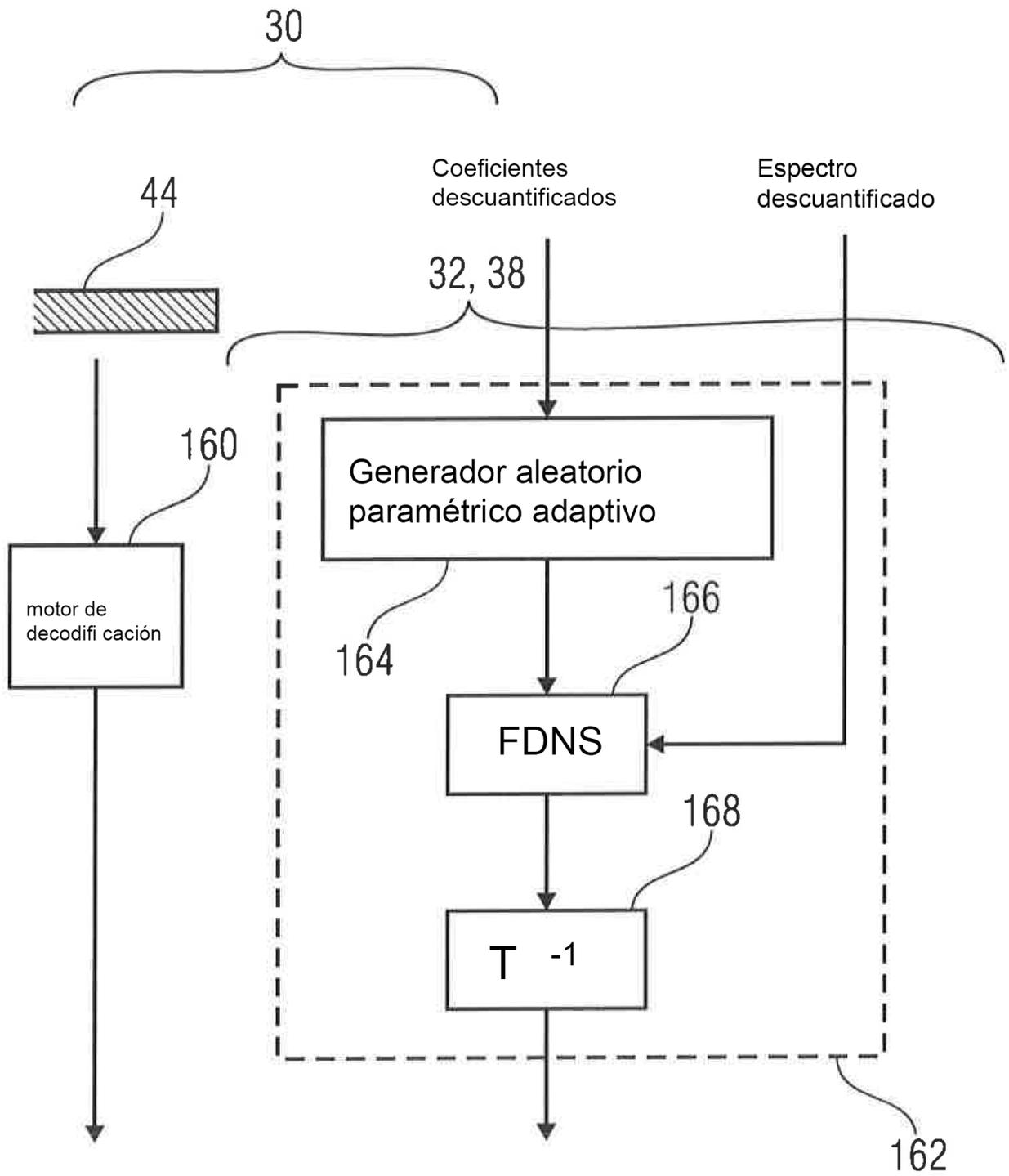


FIG 6

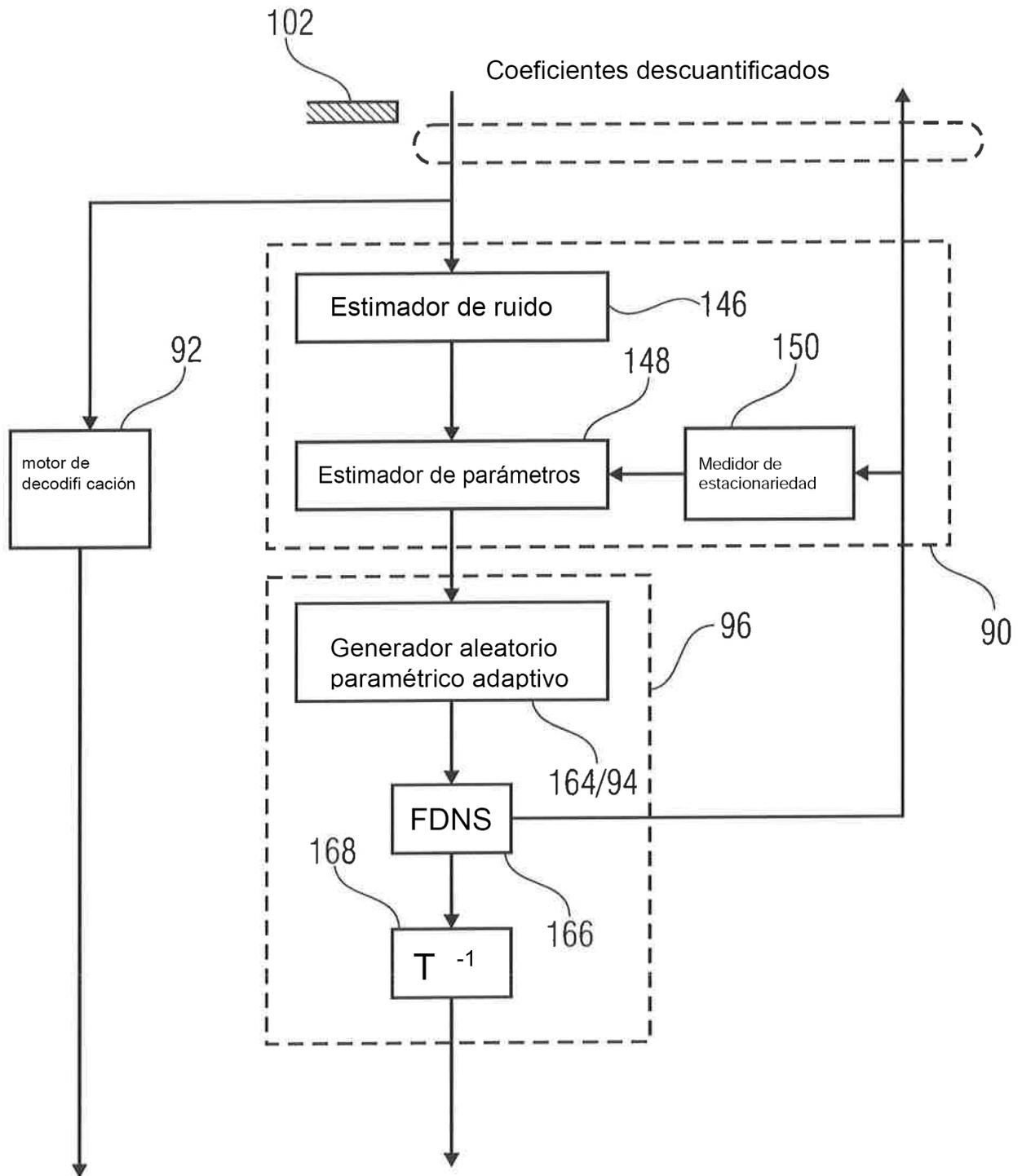


FIG 7

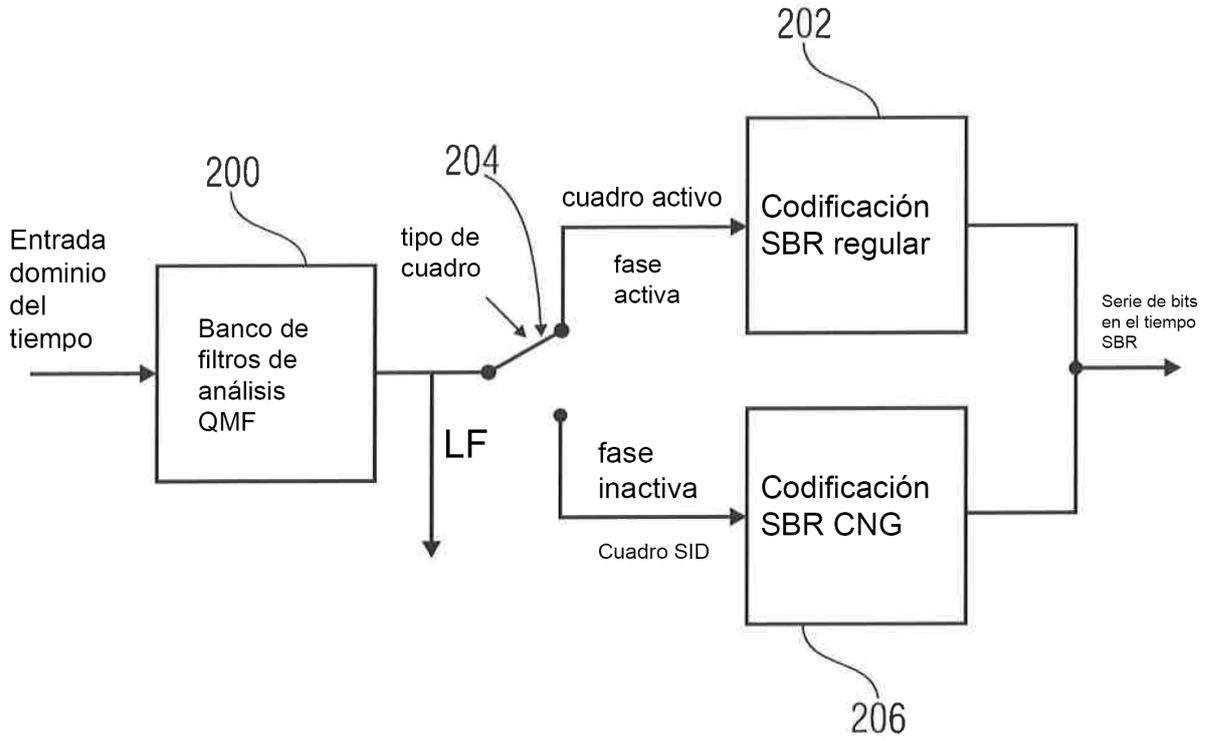


FIG 8

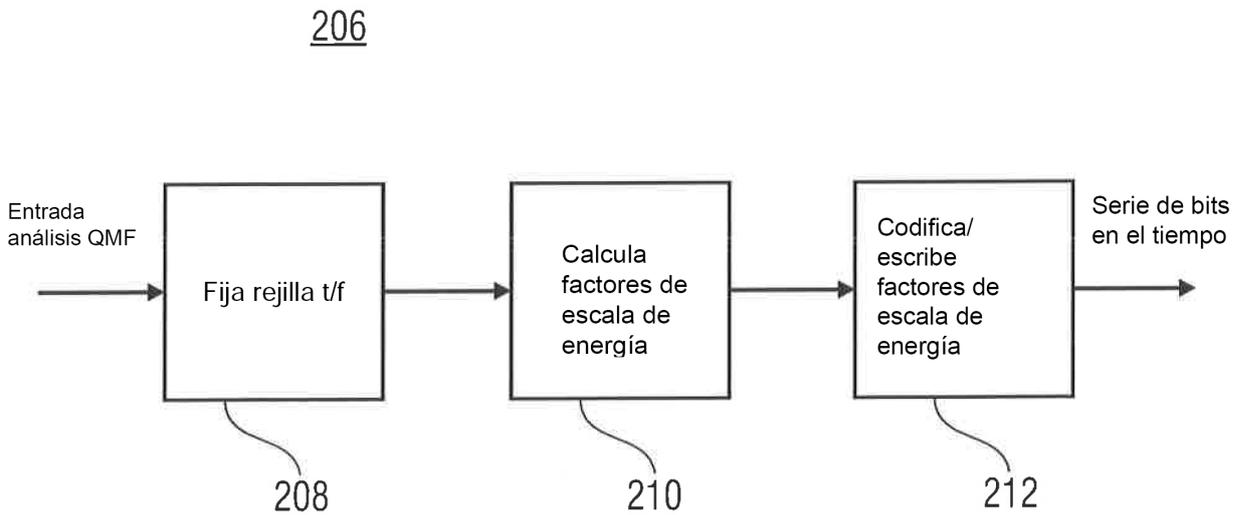


FIG 9

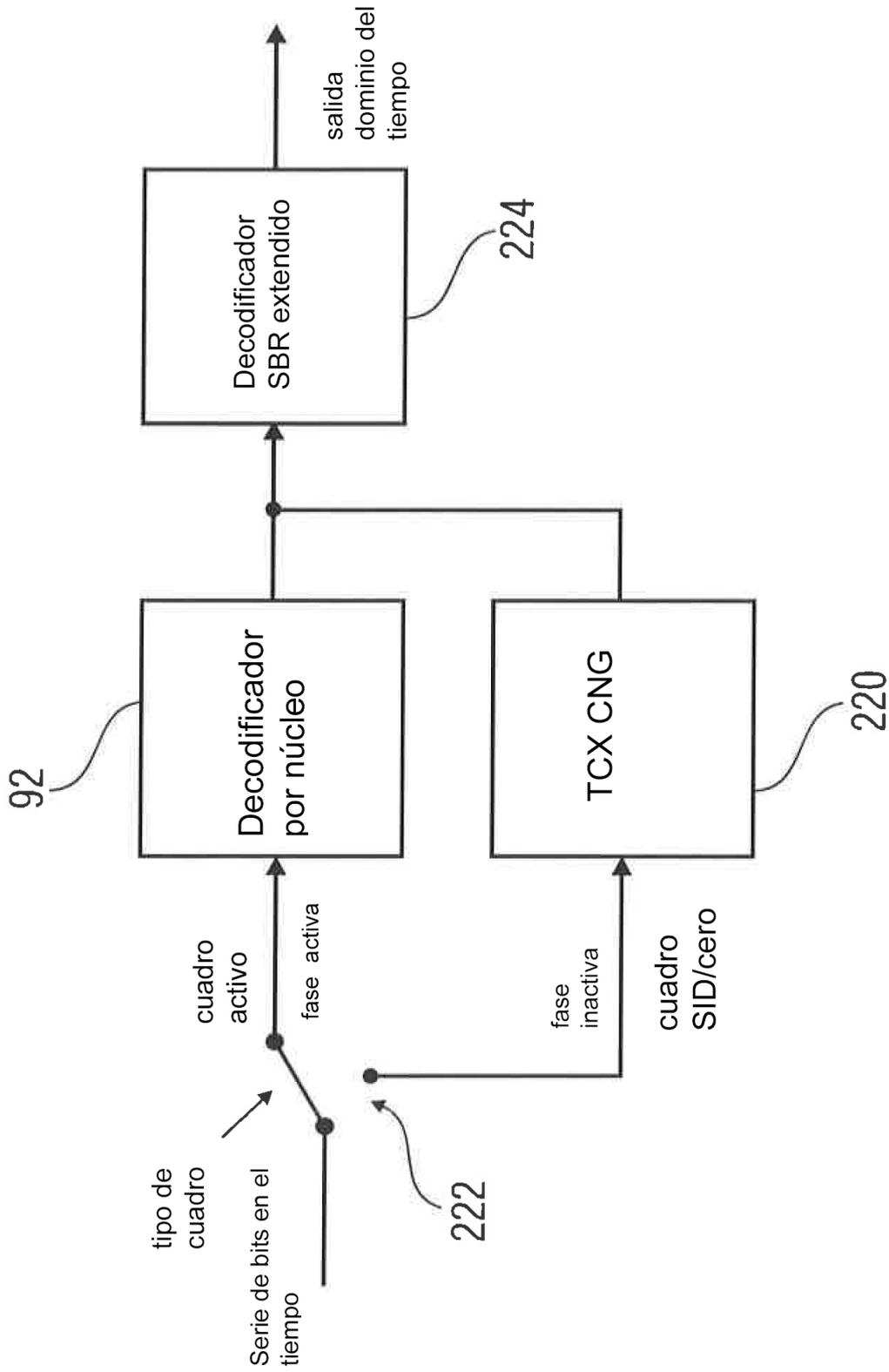


FIG 10

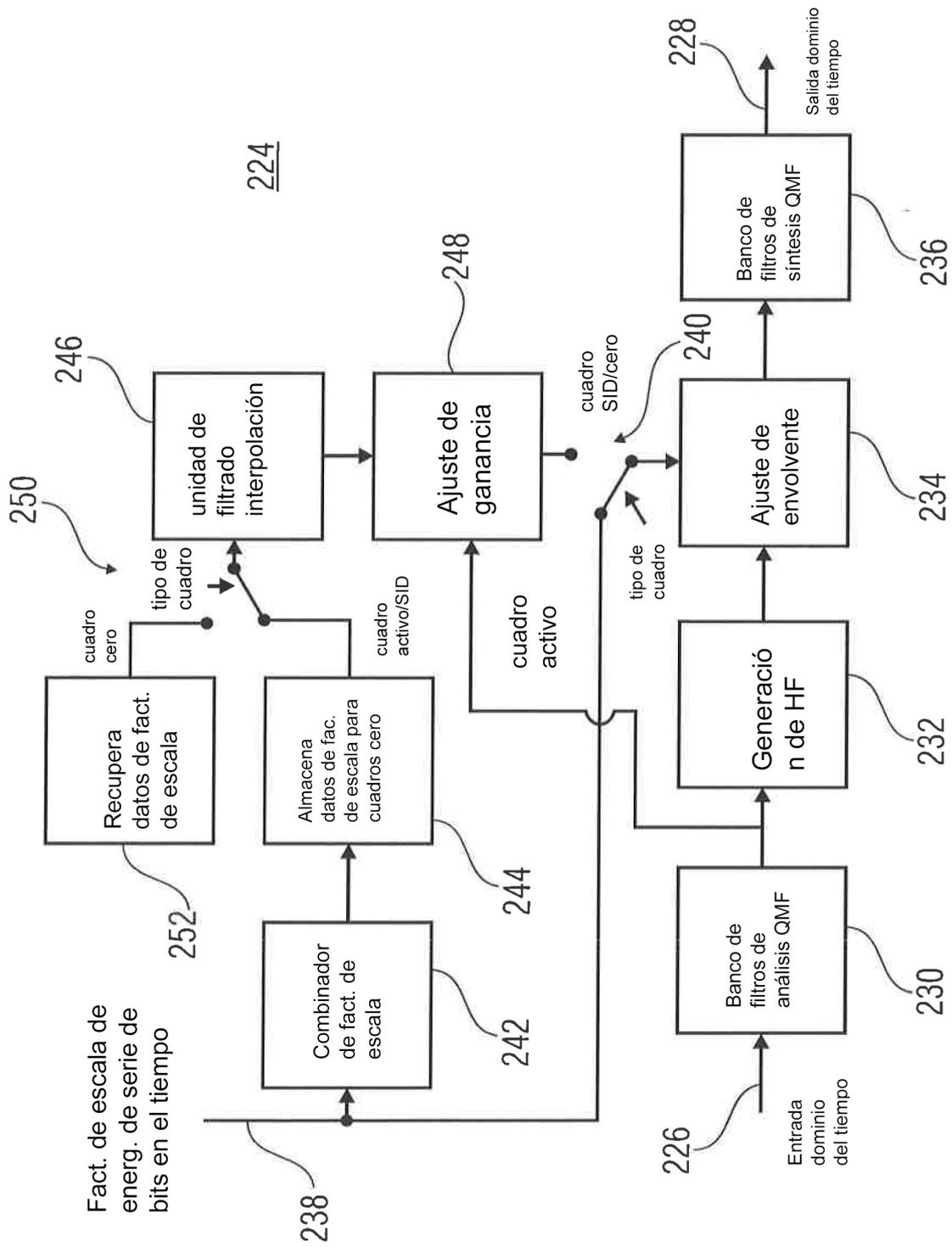


FIG 11

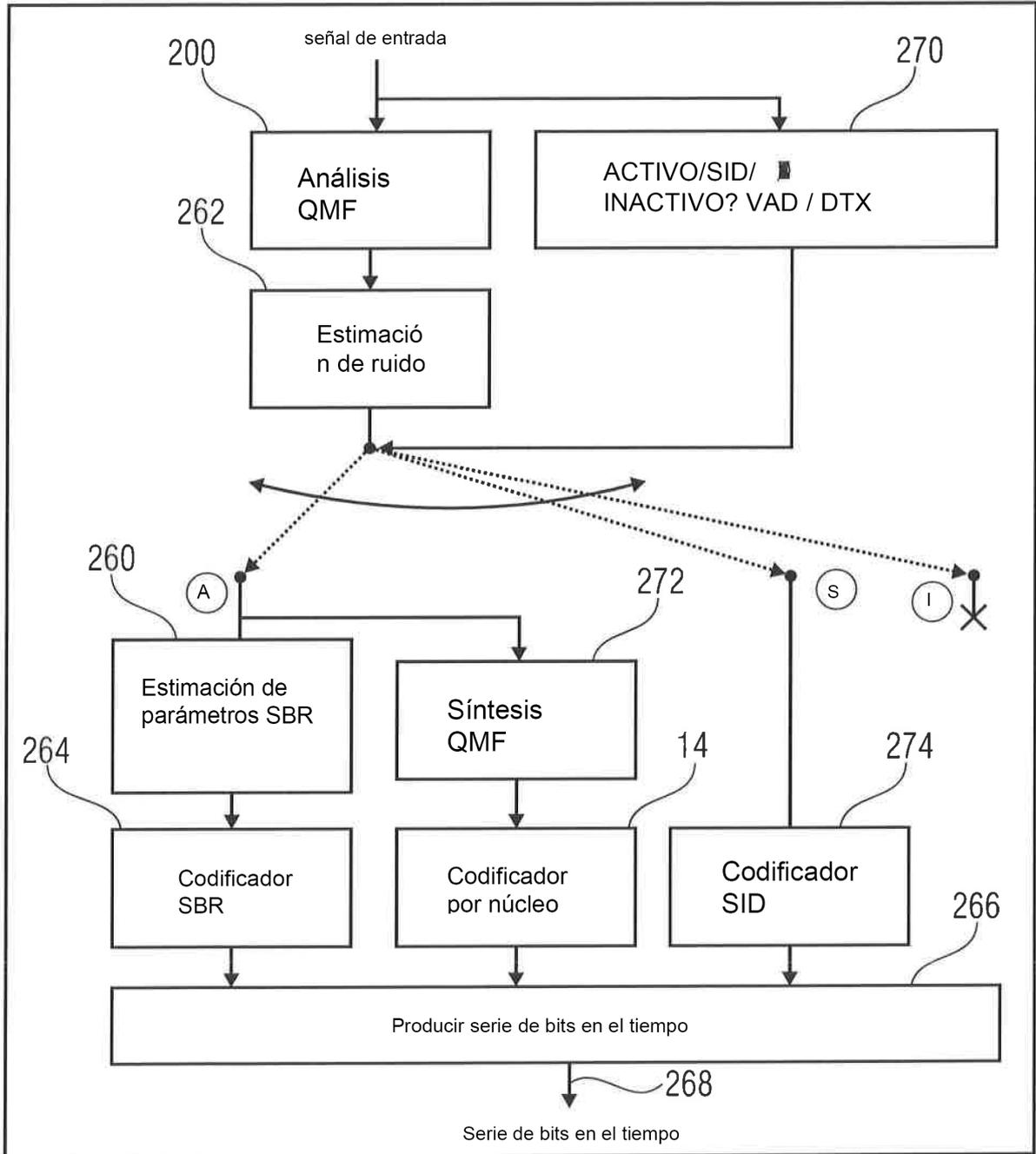


FIG 12

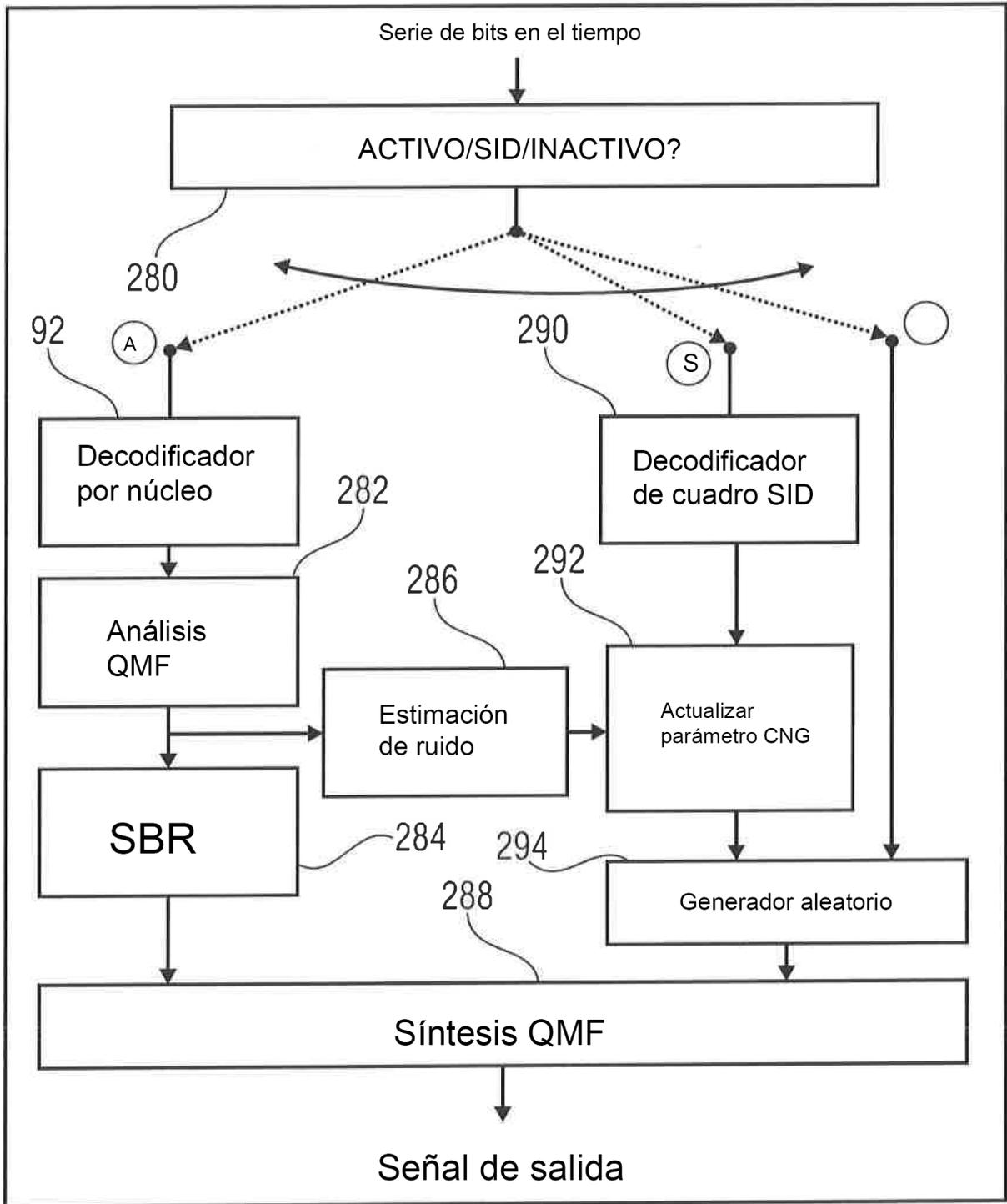


FIG 13