



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 358 125**

51 Int. Cl.:  
**G10L 21/02** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06740357 .6**

96 Fecha de presentación : **03.04.2006**

97 Número de publicación de la solicitud: **1866915**

97 Fecha de publicación de la solicitud: **19.12.2007**

54 Título: **Procedimiento y aparato para un filtrado de antidispersión de una señal ensanchada de excitación de predicción de velocidad de ancho de banda.**

30 Prioridad: **01.04.2005 US 667901 P**  
**22.04.2005 US 673965 P**

45 Fecha de publicación de la mención BOPI:  
**05.05.2011**

45 Fecha de la publicación del folleto de la patente:  
**05.05.2011**

73 Titular/es: **QUALCOMM Incorporated**  
**5775 Morehouse Drive**  
**San Diego, California 92121, US**

72 Inventor/es: **Vos, Koen Bernard y**  
**Kandhadai, Ananthapadmanabhan A.**

74 Agente: **Fàbrega Sabaté, Xavier**

ES 2 358 125 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

Procedimiento y aparato para un filtrado de antidispersión de una señal ensanchada de excitación de predicción de velocidad de ancho de banda.

## CAMPO DE LA INVENCIÓN

5 Esta invención se refiere al procesamiento de señales.

## ANTECEDENTES

10 Tradicionalmente, las comunicaciones de voz a través de la red telefónica pública conmutada (PSTN) han tenido limitado el ancho de banda al intervalo de frecuencias comprendido entre 300 y 3400 kHz. Las nuevas redes de comunicaciones de voz, tales como telefonía celular y voz sobre IP (Protocolo de Internet, VoIP), pueden tener otros límites de ancho de banda diferentes y puede ser deseable transmitir y recibir comunicaciones de voz que incluyan un intervalo de frecuencias de banda ancha en tales redes. Por ejemplo, puede ser deseable soportar un intervalo de frecuencias de audio con un valor inferior de 50 Hz y/o con un valor superior de 7 u 8 kHz. También puede ser deseable soportar otras aplicaciones, tales como conferencias de audio o de audio/vídeo de alta calidad, que puedan tener contenido de voz y audio en intervalos fuera de los límites PSTN tradicionales.

15 La ampliación del intervalo soportado por un codificador de voz en frecuencias superiores puede mejorar la inteligibilidad. Por ejemplo, la información que diferencia sonidos fricativos tales como 's' y 'f' está presente principalmente en las altas frecuencias. La ampliación de banda alta también puede mejorar otras características de la voz, tales como la presencia. Por ejemplo, incluso una vocal sonora puede presentar energía espectral muy por encima del límite PSTN.

20 Un enfoque a la codificación de voz de banda ancha requiere escalar una técnica de codificación de voz de banda estrecha (por ejemplo, una técnica configurada para codificar el intervalo comprendido entre 0 y 4 kHz) para cubrir el espectro de banda ancha. Por ejemplo, una señal de voz puede muestrearse a una tasa superior para incluir componentes a altas frecuencias, y una técnica de codificación de banda estrecha puede reconfigurarse para utilizar más coeficientes de filtro para representar esta señal de banda ancha. Sin embargo, las técnicas de codificación de banda estrecha tales como CELP (predicción lineal excitada por libro de códigos) requieren un gran esfuerzo computacional y un codificador CELP de banda ancha puede consumir demasiados ciclos de procesamiento para que resulte práctico en muchas aplicaciones móviles y en otras aplicaciones embebidas. La codificación de todo el espectro de una señal de banda ancha en una calidad deseada utilizando una técnica de este tipo también puede dar lugar a un incremento inaceptablemente grande del ancho de banda. Además, la transcodificación de una señal codificada de este tipo sería necesaria incluso antes de que su parte de banda estrecha pudiera transmitirse a y/o descodificarse por un sistema que solo soporte codificación de banda estrecha.

30 Otro enfoque a la codificación de voz de banda ancha requiere extrapolar la envolvente espectral de banda alta a partir de la envolvente espectral de banda estrecha codificada. Aunque un enfoque de este tipo puede implementarse sin aumentar el ancho de banda y sin necesidad de transcodificación, la envolvente espectral aproximada o la estructura de formantes de la parte de banda alta de una señal de voz no pueden predecirse generalmente de manera precisa a partir de la envolvente espectral de la parte de banda estrecha.

35 Puede ser deseable implementar la codificación de voz de banda ancha de manera que al menos la parte de banda estrecha de la señal codificada pueda enviarse a través de un canal de banda estrecha (tal como un canal PSTN) sin transcodificación u otra modificación significativa. La eficacia de la extensión de codificación de banda ancha también puede ser deseable, por ejemplo, para evitar una reducción significativa del número de usuarios que pueden recibir servicio en aplicaciones tales como telefonía celular inalámbrica y difusión a través de canales cableados y canales inalámbricos.

40 Un procedimiento y un aparato para una codificación de voz flexible con escalabilidades de ancho de banda se conocen a partir del documento "A Bitrate and Bandwidth scalable CELP Coder" de Toshiyuki et al., actas de la Conferencia Internacional de Seattle IEEE 1998, WA, USA, del 12 al 15 de mayo de 1998, Nueva York, NY, USA, IEEE, US, vol. 1, 12 de mayo de 1998 (12/05/1998), páginas 341 a 344, XP010279059 ISBN: 0-7803-4428-6. Según este procedimiento se utiliza una cuantificación vectorial de múltiples impulsos para la codificación, donde la herramienta de ampliación de ancho de banda codifica directamente una señal de voz de banda ancha sin utilizar una estructura de subbandas. Para la descodificación, la señal de voz de banda estrecha se extrae de una parte del flujo de bits codificado mediante un descodificador MP-CELP escalable de velocidad binaria, mientras que la señal de voz de banda ancha se obtiene utilizando todo el flujo de bits a través del descodificador MP-CELP escalable de velocidad binaria y de una herramienta de descodificación de ampliación de ancho de banda.

45 Otro procedimiento se describe en el documento EP 1 300 833 A2, en el que una representación de envolvente espectral de ancho de banda se genera a partir de una representación espectral de banda estrecha de entrada utilizando

coeficientes predictivos lineales (LPC). Este procedimiento también utiliza coeficientes de correlación y coeficientes de área asociados con los LPC de banda estrecha, donde esta representación está relacionada con un modelo de tubo acústico discretizado. Este procedimiento también aplica filtrado para reducir las deficiencias de una excitación de banda alta generada.

## 5 RESUMEN

La invención proporciona un procedimiento para generar una señal de excitación de banda alta según la reivindicación 1, un medio de almacenamiento de datos según la reivindicación 18, un aparato según la reivindicación 19 y un programa informático según la reivindicación 40.

### BREVE DESCRIPCIÓN DE LOS DIBUJOS

- 10 La figura 1a muestra un diagrama de bloques de un codificador A100 de voz de banda ancha según una realización.  
 La figura 1b muestra un diagrama de bloques de una implementación A102 del codificador A100 de voz de banda ancha.  
 La figura 2a muestra un diagrama de bloques de un decodificador B100 de voz de banda ancha según una realización.  
 La figura 2b muestra un diagrama de bloques de una implementación B102 del codificador B100 de voz de banda ancha.  
 La figura 3a muestra un diagrama de bloques de una implementación A112 de un banco A110 de filtros.
- 15 La figura 3b muestra un diagrama de bloques de una implementación B122 de un banco B120 de filtros.  
 La figura 4a muestra una cobertura de ancho de banda de una banda alta y de una banda baja de un ejemplo del banco A110 de filtros.  
 La figura 4b muestra una cobertura de ancho de banda de una banda alta y de una banda baja de otro ejemplo del banco A110 de filtros.
- 20 La figura 4c muestra un diagrama de bloques de una implementación A114 de un banco A112 de filtros.  
 La figura 4d muestra un diagrama de bloques de una implementación B124 de un banco B122 de filtros.  
 La figura 5a muestra un ejemplo de una representación gráfica de la amplitud logarítmica frente a la frecuencia de una señal de voz.  
 La figura 5b muestra un diagrama de bloques de un sistema de codificación básico de predicción lineal.
- 25 La figura 6 muestra un diagrama de bloques de una implementación A122 de un codificador A120 de banda estrecha.  
 La figura 7 muestra un diagrama de bloques de una implementación B112 de un decodificador B110 de banda estrecha.  
 La figura 8a muestra un ejemplo de una representación gráfica de la amplitud logarítmica frente a la frecuencia de una señal residual de voz sonora.  
 La figura 8b muestra un ejemplo de una representación gráfica de la amplitud logarítmica frente al tiempo de una señal residual de voz sonora.
- 30 La figura 9 muestra un diagrama de bloques de un sistema de codificación básico de predicción lineal que también lleva a cabo una predicción a largo plazo.  
 La figura 10 muestra un diagrama de bloques de una implementación A202 de un codificador A200 de banda alta.
- 35 La figura 11 muestra un diagrama de bloques de una implementación A302 de un generador A300 de excitación de banda alta.  
 La figura 12 muestra un diagrama de bloques de una implementación A402 de un ensanchador A400 de espectro.  
 La figura 12a muestra representaciones gráficas de espectros de señales en varios puntos en un ejemplo de una operación de ensanchamiento espectral.  
 La figura 12b muestra representaciones gráficas de espectros de señales en varios puntos en otro ejemplo de una operación de ensanchamiento espectral.
- 40 La figura 13 muestra un diagrama de bloques de una implementación A304 de un generador A302 de excitación de banda alta.

- La figura 14 muestra un diagrama de bloques de una implementación A306 del generador A302 de excitación de banda alta.
- La figura 15 muestra un diagrama de flujo de una tarea T100 de cálculo de envolvente.
- La figura 16 muestra un diagrama de bloques de una implementación 492 de un combinador 490.
- 5 La figura 17 muestra un enfoque para calcular una medida de periodicidad de una señal S30 de banda alta.
- La figura 18 muestra un diagrama de bloques de una implementación A312 del generador A302 de excitación de banda alta.
- La figura 19 muestra un diagrama de bloques de una implementación A314 del generador A302 de excitación de banda alta.
- 10 La figura 20 muestra un diagrama de bloques de una implementación A316 del generador A302 de excitación de banda alta.
- La figura 21 muestra un diagrama de flujo de una tarea T200 de cálculo de ganancia.
- La figura 22 muestra un diagrama de flujo de una implementación T210 de la tarea T200 de cálculo de ganancia.
- La figura 23a muestra un diagrama de una función de división en ventanas.
- 15 La figura 23b muestra una aplicación de una función de división en ventanas, como la mostrada en la figura 23a, en subtramas de una señal de voz.
- La figura 24 muestra un diagrama de bloques de una implementación B202 de un descodificador B200 de banda alta.
- La figura 25 muestra un diagrama de bloques de una implementación AD10 del codificador A100 de voz de banda ancha.
- La figura 26a muestra un diagrama esquemático de una implementación D122 de una línea D120 de retardo.
- 20 La figura 26b muestra un diagrama esquemático de una implementación D124 de la línea D120 de retardo.
- La figura 27 muestra un diagrama esquemático de una implementación D130 de la línea D120 de retardo.
- La figura 28 muestra un diagrama de bloques de una implementación AD12 del codificador AD10 de voz de banda ancha.
- La figura 29 muestra un diagrama de flujo de un procedimiento de procesamiento MD100 de señales según una realización.
- 25 La figura 30 muestra un diagrama de flujo de un procedimiento M100 según una realización.
- La figura 31a muestra un diagrama de flujo de un procedimiento M200 según una realización.
- La figura 31b muestra un diagrama de flujo de una implementación M210 del procedimiento M200.
- La figura 32 muestra un diagrama de flujo de un procedimiento M300 según una realización.
- 30 En las figuras y en la descripción adjunta, las mismas etiquetas de referencia se refieren a elementos o señales idénticos o análogos.

#### DESCRIPCIÓN DETALLADA

- Las realizaciones descritas en este documento incluyen sistemas, procedimientos y aparatos que pueden configurarse para ampliar un codificador de voz de banda estrecha para que soporte la transmisión y/o el almacenamiento de señales de voz de banda ancha con un incremento de ancho de banda comprendido solamente entre 800 y 1000 bps (bits por segundo) aproximadamente. Las ventajas potenciales de estas implementaciones incluyen codificación embebida para permitir la compatibilidad con sistemas de banda estrecha, una asignación y reasignación de bits relativamente sencillas entre los canales de codificación de banda estrecha y de banda alta, la eliminación de una operación de síntesis de banda ancha de un alto esfuerzo computacional y el mantenimiento de una baja tasa de muestreo para señales que van a procesarse mediante rutinas de codificación de forma de onda de un alto esfuerzo computacional.
- 35
- 40 A no ser que esté limitado expresamente por su contexto, el término "calcular" se utiliza en este documento para indicar cualquiera de sus significados habituales, tales como computar, generar y seleccionar a partir de una lista de valores. Cuando se utilice el término "comprender" en la presente descripción y en las reivindicaciones, no excluye otros elementos u operaciones. El término "A se basa en B" se utiliza para indicar cualquiera de sus significados habituales, incluyendo los

casos (i) "A es igual a B" y (ii) "A se basa al menos en B". El término "protocolo de Internet" incluye la versión 4, tal y como se describe en la RFC (solicitud de comentarios) 791 del IETF (grupo de trabajo de ingeniería de Internet) y versiones posteriores tales como la versión 6.

5 La figura 1a muestra un diagrama de bloques de un codificador A100 de voz de banda ancha según una realización. Un banco A110 de filtros está configurado para filtrar una señal S10 de voz de banda ancha para generar una señal S20 de banda estrecha y una señal S30 de banda alta. Un codificador A120 de banda estrecha está configurado para codificar la señal S20 de banda estrecha para generar parámetros S40 de filtro de banda estrecha (NB) y una señal S50 residual de banda estrecha. Tal y como se describe en mayor detalle en este documento, el codificador A120 de banda estrecha está configurado normalmente para generar parámetros S40 de filtro de banda estrecha y una señal S50 de excitación de banda estrecha codificada como índices de libro de código o en otra forma cuantificada. Un codificador A200 de banda alta está configurado para codificar la señal S30 de banda alta según información de la señal S50 de excitación de banda estrecha codificada para generar parámetros S60 de codificación de banda alta. Tal y como se describe en mayor detalle en este documento, el codificador A200 de banda alta está configurado normalmente para generar parámetros S60 de codificación de banda alta como índices de libro de código o en otra forma cuantificada. Un ejemplo particular del codificador A100 de voz de banda ancha está configurado para codificar la señal S10 de voz de banda ancha a una velocidad de 8,55 kbps aproximadamente (kilobits por segundo), utilizándose 7,55 kbps aproximadamente para los parámetros S40 de filtro de banda estrecha y para la señal S50 de excitación de banda estrecha codificada, y utilizándose 1 kbps aproximadamente para los parámetros S60 de codificación de banda alta.

10  
15  
20 Puede ser deseable combinar las señales de banda estrecha y de banda alta codificadas en un único flujo de bits. Por ejemplo, puede ser deseable multiplexar juntas las señales codificadas para su transmisión (por ejemplo, a través de un canal de transmisión cableado, óptico o inalámbrico), o para su almacenamiento, como una señal de voz de banda ancha codificada. La figura 1b muestra un diagrama de bloques de una implementación A102 del codificador A100 de voz de banda ancha que incluye un multiplexor A130 configurado para combinar los parámetros S40 de filtro de banda estrecha, la señal S50 de excitación de banda estrecha codificada y los parámetros S60 de filtro de banda alta en una señal S70 multiplexada.

25  
30 Un aparato que incluye al codificador A102 también puede incluir un sistema de circuitos configurado para transmitir la señal S70 multiplexada en un canal de transmisión tal como un canal cableado, óptico o inalámbrico. Un aparato de este tipo también puede configurarse para llevar a cabo una o más operaciones de codificación de canal en la señal, tal como codificación de corrección de errores (por ejemplo, codificación convolucional de velocidades compatibles) y/o codificación de detección de errores (por ejemplo, codificación de redundancia cíclica), y/o una o más capas de codificación de protocolo de red (por ejemplo, Ethernet, TCP/IP, cdma2000).

35  
40 Puede ser deseable que el multiplexor A130 esté configurado para encapsular la señal codificada de banda estrecha (incluyendo los parámetros S40 de filtro de banda estrecha y la señal S50 de excitación de banda estrecha) como un subflujo separable de la señal S70 multiplexada de manera que la señal codificada de banda estrecha pueda recuperarse y descodificarse independientemente de otra parte de la señal S70 multiplexada tal como una señal de banda alta y/o de banda baja. Por ejemplo, la señal S70 multiplexada puede estar dispuesta de manera que la señal codificada de banda estrecha pueda recuperarse extrayendo los parámetros S60 de filtro de banda alta. Una ventaja potencial de esta característica es evitar la necesidad de transcodificar la señal codificada de banda ancha antes de suministrarla a un sistema que soporte la descodificación de la señal de banda estrecha pero que no soporte la descodificación de la parte de banda alta.

45  
50 La figura 2a es un diagrama de bloques de un descodificador B100 de voz de banda ancha según una realización. Un descodificador B110 de banda estrecha está configurado para descodificar los parámetros S40 de filtro de banda estrecha y la señal S50 codificada de excitación de banda estrecha para generar una señal S90 de banda estrecha. Un descodificador B200 de banda alta está configurado para descodificar los parámetros S60 de codificación de banda alta según una señal S80 de excitación de banda estrecha, basada en la señal S50 codificada de excitación de banda estrecha, para generar una señal S100 de banda alta. En este ejemplo, el descodificador B110 de banda estrecha está configurado para proporcionar la señal S80 de excitación de banda estrecha al descodificador B200 de banda alta. Un banco B120 de filtros está configurado para combinar la señal S90 de banda estrecha y la señal S100 de banda alta para generar una señal S110 de voz de banda ancha.

55 La figura 2b es un diagrama de bloques de una implementación B102 del descodificador B100 de voz de banda ancha que incluye un demultiplexor B130 configurado para generar las señales S40, S50 y S60 codificadas de la señal S70 multiplexada. Un aparato que incluye al descodificador B102 puede incluir un sistema de circuitos configurado para recibir la señal S70 multiplexada desde un canal de transmisión tal como un canal cableado, óptico o inalámbrico. Un aparato de este tipo también puede configurarse para llevar a cabo una o más operaciones de descodificación de canal en la señal, tal como descodificación de corrección de errores (por ejemplo, descodificación convolucional de velocidades compatibles) y/o descodificación de detección de errores (por ejemplo, descodificación de redundancia cíclica), y/o una o más capas de descodificación de protocolo de red (por ejemplo, Ethernet, TCP/IP, cdma2000).

El banco A110 de filtros está configurado para filtrar una señal de entrada según un esquema de banda dividida para generar una subbanda de baja frecuencia y una subbanda de alta frecuencia. Dependiendo de los criterios de diseño de la aplicación particular, las subbandas de salida pueden tener anchos de banda iguales o diferentes y pueden estar solapadas o no solapadas. También es posible una configuración del banco A110 de filtros que genere más de dos subbandas. Por ejemplo, un banco de filtros de este tipo puede configurarse para generar una o más señales de banda baja que incluyan componentes en un intervalo de frecuencias inferior al de la señal S20 de banda estrecha (tal como el intervalo comprendido entre 50 y 300 Hz). También es posible que un banco de filtros de este tipo esté configurado para generar una o más señales de banda alta adicionales que incluyan componentes en un intervalo de frecuencias superior al de la señal S30 de banda alta (tal como un intervalo comprendido entre 14 y 20, 16 y 20 o 16 y 32 KHz). En ese caso, el codificador A100 de voz de banda ancha puede implementarse para codificar esta señal o señales por separado, y el multiplexor A130 puede configurarse para incluir la señal o señales codificada(s) adicional(es) en la señal S70 multiplexada (por ejemplo, como una parte separable).

La figura 3a muestra un diagrama de bloques de una implementación A112 del banco A110 de filtros que está configurada para generar dos señales de subbanda que presentan tasas de muestreo reducidas. El banco A110 de filtros está dispuesto para recibir una señal S10 de voz de banda ancha que presenta una parte de alta frecuencia (o de banda alta) y una parte de baja frecuencia (o de banda baja). El banco A112 de filtros incluye una trayectoria de procesamiento de banda baja configurada para recibir una señal S10 de voz de banda ancha y para generar una señal S20 de voz de banda estrecha, y una trayectoria de procesamiento de banda alta configurada para recibir la señal S10 de voz de banda ancha y para generar una señal S30 de voz de banda alta. Un filtro 110 paso bajo filtra la señal S10 de voz de banda ancha para hacer pasar una subbanda de baja frecuencia seleccionada, y un filtro 130 paso alto filtra la señal S10 de voz de banda ancha para hacer pasar una subbanda de alta frecuencia seleccionada. Puesto que ambas señales de subbanda presentan anchos de banda más estrechos que la señal S10 de voz de banda ancha, sus tasas de muestreo pueden reducirse hasta cierto punto sin pérdida de información. Un muestreador 120 descendente reduce la tasa de muestreo de la señal de paso bajo según un factor de diezmado deseado (por ejemplo, eliminando muestras de la señal y/o sustituyendo las muestras por valores promedio), y un muestreador 140 descendente reduce igualmente la tasa de muestreo de la señal de paso alto según otro factor de diezmado deseado.

La figura 3b muestra un diagrama de bloques de una implementación B122 correspondiente del banco B120 de filtros. Un muestreador 150 ascendente aumenta la tasa de muestreo de la señal S90 de banda estrecha (por ejemplo, rellenando con ceros y/o duplicando las muestras), y un filtro 160 paso bajo filtra la señal muestreada de manera ascendente para hacer pasar solamente una parte de banda baja (por ejemplo, para impedir la superposición). Asimismo, un muestreador 170 ascendente aumenta la tasa de muestreo de la señal S100 de banda alta y un filtro 180 paso alto filtra la señal muestreada de manera ascendente para hacer pasar solamente una parte de banda alta. Después, las dos señales de banda de paso se suman para formar una señal S110 de voz de banda ancha. En algunas implementaciones del decodificador B100, el banco B120 de filtros está configurado para generar una suma ponderada de las dos señales de banda de paso según una o más ponderaciones recibidas y/o calculadas por el decodificador B200 de banda alta. También se contempla una configuración del banco B120 de filtros que combine más de dos señales de banda de paso.

Cada uno de los filtros 110, 130, 160 y 180 pueden implementarse como un filtro de respuesta de impulsos finita (FIR) o como un filtro de respuesta de impulsos infinita (IIR). Las respuestas de frecuencia de los filtros 110 y 130 de codificador pueden presentar regiones de transición simétricas o de formas distintas entre la banda de detención y la banda de paso. Asimismo, las respuestas de frecuencia de los filtros 160 y 180 de decodificador pueden presentar regiones de transición simétricas o de formas distintas entre la banda de detención y la banda de paso. Puede ser deseable, pero no es estrictamente necesario, que el filtro 110 paso bajo tenga la misma respuesta que el filtro 160 paso bajo y que el filtro 130 paso alto tenga la misma respuesta que el filtro 180 paso alto. En un ejemplo, los dos pares 110, 130 y 160, 180 de filtros son bancos de filtros de espejo en cuadratura (QMF), presentando el par 110, 130 de filtros los mismos coeficientes que el par 160, 180 de filtros.

En un ejemplo típico, el filtro 110 paso bajo presenta una banda de paso que incluye el intervalo PSTN limitado de 300 a 3400 Hz (por ejemplo, la banda de 0 a 4 kHz). Las figuras 4a y 4b muestran anchos de banda relativos de la señal S10 de voz de banda ancha, de la señal S20 de banda estrecha y de la señal S30 de banda alta en dos ejemplos de implantación diferentes. En estos dos ejemplos particulares, la señal S10 de voz de banda ancha presenta una tasa de muestreo de 16 kHz (representando componentes de frecuencia dentro del intervalo de 0 a 8 kHz), y la señal S20 de banda estrecha presenta una tasa de muestreo de 8 kHz (representando componentes de frecuencia dentro del intervalo de 0 a 4 kHz).

En el ejemplo de la figura 4a no se produce un solapamiento significativo entre las dos subbandas. Una señal S30 de banda alta como la mostrada en este ejemplo puede obtenerse utilizando un filtro 130 paso alto con una banda de paso de 4 a 8 kHz. En este caso, puede ser deseable reducir la tasa de muestreo a 8kHz realizando un muestreo descendente de la señal filtrada mediante un factor de dos. Una operación de este tipo, mediante la cual se espera reducir significativamente la complejidad computacional de operaciones de procesamiento adicionales en la señal, reducirá la energía de banda de paso al intervalo de 0 a 4 kHz sin pérdida de información.

5 En el ejemplo alternativo de la figura 4b, la subbanda superior y la subbanda inferior presentan un solapamiento apreciable, de manera que la región de 3,5 a 4 kHz se describe mediante ambas señales de subbanda. Una señal S30 de banda alta como en este ejemplo puede obtenerse utilizando un filtro 130 paso alto con una banda de paso de 3,5 a 7 kHz. En este caso, puede ser deseable reducir la tasa de muestreo a 7kHz muestreando de manera descendente la señal filtrada mediante un factor de 16/7. Una operación de este tipo, mediante la cual se espera reducir significativamente la complejidad computacional de operaciones de procesamiento adicionales en la señal, reducirá la energía de banda de paso al intervalo de 0 a 3,5 kHz sin pérdida de información.

10 En un dispositivo de microteléfono típico para una comunicación telefónica, uno o más de los transductores (es decir, el micrófono y el auricular u altavoz) carece(n) de una respuesta apreciable en el intervalo de frecuencias de 7 a 8 kHz. En el ejemplo de la figura 4b, la parte de la señal S10 de voz de banda ancha entre 7 y 8 kHz no está incluida en la señal codificada. Otros ejemplos particulares del filtro 130 paso alto presentan bandas de paso de 3,5 a 7,5 kHz y de 3,5 a 8 kHz.

15 En algunas implementaciones, la provisión de un solapamiento entre subbandas como en el ejemplo de la figura 4b permite la utilización de un filtro paso bajo y/o de un filtro paso alto que presenten una suave caída en la región solapada. Tales filtros son normalmente más fáciles de diseñar, menos complejos computacionalmente y/o introducen menos retardo que los filtros con respuestas más angulosas o de "muro de ladrillos". Los filtros que presentan regiones de transición angulosas tienden a tener mayores lóbulos laterales (lo que puede provocar superposición) que los filtros de orden similar que presentan suaves caídas. Los filtros que presentan regiones de transición angulosas también pueden presentar largas respuestas de impulsos que pueden provocar artefactos oscilantes. En implementaciones de bancos de filtros que presenten uno o más filtros IIR, permitiendo una suave caída en la región solapada puede adoptarse la utilización de un filtro o filtros cuyos polos estén alejados del círculo unitario, lo que puede ser importante para garantizar una implementación de punto fijo estable.

20 El solapamiento de las subbandas permite una mezcla uniforme de banda baja y de banda alta que puede dar lugar a artefactos menos audibles, una menor superposición y/o una transición menos apreciable de una banda a otra. Además, la eficacia de codificación del codificador A120 de banda estrecha (por ejemplo, un codificador de forma de onda) puede disminuir con una frecuencia creciente. Por ejemplo, la calidad de codificación del codificador de banda estrecha puede reducirse a bajas velocidades binarias, especialmente en presencia de ruido de fondo. En tales casos, proporcionar un solapamiento de las subbandas puede aumentar la calidad de las componentes de frecuencia reproducidas en la región solapada.

25 Además, el solapamiento de las subbandas permite una mezcla uniforme de banda baja y de banda alta que puede dar lugar a artefactos menos audibles, una superposición reducida y/o una transición menos perceptible de una banda a otra. Esta característica puede ser especialmente deseable para una implementación en la que el codificador A120 de banda estrecha y el codificador A200 de banda alta funcionen según diferentes metodologías de codificación. Por ejemplo, diferentes técnicas de codificación pueden generar señales que suenen de manera muy diferente. Un codificador que codifique una envolvente espectral en forma de índices de libro de código puede generar una señal que tenga un sonido diferente al de un codificador que codifique en cambio el espectro de amplitud. Un codificador de dominio de tiempo (por ejemplo, un codificador PCM o de modulación por códigos de impulsos) puede generar una señal que tenga un sonido diferente al de un codificador de dominio de frecuencia. Un codificador que codifique una señal con una representación de la envolvente espectral y la señal residual correspondiente puede generar una señal que tenga un sonido diferente al de un codificador que codifique una señal con solamente una representación de la envolvente espectral. Un codificador que codifique una señal como una representación de su forma de onda puede generar una salida que tenga un sonido diferente al de un codificador sinusoidal. En tales casos, la utilización de filtros que presenten regiones de transición angulosas para definir subbandas no solapadas puede dar lugar a una transición abrupta y perceptualmente apreciable entre las subbandas de la señal de banda ancha sintetizada.

30 Aunque los bancos de filtros QMF que presentan respuestas de frecuencia solapadas complementarias se utilizan normalmente en técnicas de subbanda, tales filtros no son adecuados para al menos algunas de las implementaciones de codificación de banda ancha descritas en este documento. Un banco de filtros QMF en el codificador está configurado para crear un grado significativo de superposición que se suprime en el banco de filtros QMF correspondiente en el decodificador. Una disposición de este tipo puede no ser apropiada para una aplicación en la que la señal adquiere una cantidad importante de distorsión entre los bancos de filtros, ya que la distorsión puede reducir la eficacia de la propiedad de supresión de superposición. Por ejemplo, las aplicaciones descritas en este documento incluyen implementaciones de codificación configuradas para funcionar a velocidades binarias muy bajas. Como consecuencia de la reducida velocidad binaria, es probable que la señal descodificada aparezca significativamente distorsionada en comparación con la original, de manera que la utilización de bancos de filtros QMF puede dar lugar a una superposición no eliminada. Las aplicaciones que utilizan bancos de filtros QMF presentan normalmente velocidades binarias superiores (por ejemplo, por encima de los 12 kbps para AMR y de los 64 kbps para G.722).

35 Además, un codificador puede configurarse para generar una señal sintetizada que sea perceptualmente similar a la señal

original pero que realmente sea muy diferente de la señal original. Por ejemplo, un codificador que obtenga la excitación de banda alta a partir de la señal residual de banda estrecha tal y como se describe en este documento puede generar una señal de este tipo, ya que la señal residual de banda alta real puede estar totalmente ausente de la señal descodificada. La utilización de bancos de filtros QMF en tales aplicaciones puede dar lugar a un grado de distorsión significativo provocado por una superposición no eliminada.

La cantidad de distorsión provocada por la superposición QMF puede reducirse si la subbanda afectada es estrecha, ya que el efecto de la superposición está limitado a un ancho de banda igual al ancho de la subbanda. Sin embargo, para los ejemplos descritos en este documento en los que cada subbanda incluye casi la mitad del ancho de banda de banda ancha, la distorsión provocada por la superposición no eliminada podría afectar a una parte significativa de la señal. La calidad de la señal también puede verse afectada por la posición de la banda de frecuencia sobre la cual se produce la superposición no eliminada. Por ejemplo, la distorsión creada cerca del centro de una señal de voz de banda ancha (por ejemplo, entre 3 y 4 kHz) puede ser mucho más inaceptable que la distorsión que se produce cerca de un borde de la señal (por ejemplo, por encima de los 6 kHz).

Aunque las respuestas de los filtros de un banco de filtros QMF están estrictamente relacionadas entre sí, la trayectoria de banda baja y la trayectoria de banda alta de los bancos A110 y B120 de filtros pueden configurarse para presentar espectros que no estén en absoluto relacionados aparte del solapamiento de las dos subbandas. El solapamiento de las dos subbandas se define como la distancia desde el punto en el que la respuesta de frecuencia del filtro de banda alta desciende a -20 dB hasta el punto en que la respuesta de frecuencia del filtro de banda baja desciende a 20 dB. En varios ejemplos del banco A110 y/o B120 de filtros, este solapamiento oscila entre 200 Hz aproximadamente y 1 kHz aproximadamente. El intervalo entre 400 Hz aproximadamente y 600 Hz aproximadamente puede representar un equilibrio deseable entre la eficacia de codificación y una uniformidad perceptual. En un ejemplo particular como el mencionado anteriormente, el solapamiento es de 500 Hz aproximadamente.

Puede ser deseable implementar el banco A122 y/o B122 de filtros para llevar a cabo operaciones como las ilustradas en las figuras 4a y 4b en varias fases. Por ejemplo, la figura 4c muestra un diagrama de bloques de una implementación A114 del banco A112 de filtros que realiza un equivalente funcional de operaciones de muestreo descendente y de filtrado de paso alto utilizando una serie de operaciones de interpolación, remuestreo, diezmado, y otras aplicaciones. Una implementación de este tipo puede ser más fácil de diseñar y/o puede permitir la reutilización de bloques funcionales de lógica y/o código. Por ejemplo, puede utilizarse el mismo bloque funcional para realizar las operaciones de diezmado a 14 kHz y de diezmado a 7 kHz tal y como se muestra en la figura 4c. La operación de inversión espectral puede implementarse multiplicando la señal con la función  $e^{in\pi}$  o la secuencia  $(-1)^n$ , cuyos valores alternan entre +1 y -1. La operación de conformación espectral puede implementarse como un filtro paso bajo configurado para conformar la señal para obtener una respuesta de filtro global deseada.

Debe observarse que como consecuencia de la operación de inversión espectral, el espectro de la señal S30 de banda alta se invierte. Operaciones posteriores en el codificador y en el descodificador correspondiente pueden configurarse consecuentemente. Por ejemplo, el generador A300 de excitación de banda alta descrito en este documento puede configurarse para generar una señal S120 de excitación de banda alta que también presente una forma invertida espectralmente.

La figura 4d muestra un diagrama de bloques de una implementación B124 del banco B122 de filtros que realiza un equivalente funcional de operaciones de muestreo ascendente y de filtrado de paso alto utilizando una serie de operaciones de interpolación, remuestreo, etc. El banco B124 de filtros incluye una operación de inversión espectral en la banda alta que invierte una operación similar a la realizada, por ejemplo, en un banco de filtros del codificador tal como el banco A114 de filtros. En este ejemplo particular, el banco B124 de filtros también incluye filtros de muesca en la banda baja y en la banda alta que atenúan una componente de la señal a 7100 Hz, aunque tales filtros son opcionales y no es necesario que estén incluidos. La solicitud de patente "SYSTEMS, METHODS, AND APPARATUS FOR SPEECH SIGNAL FILTERING", presentada como US 2007/0088558A1, incluye una descripción y figuras adicionales relacionadas con las respuestas de elementos de implementaciones particulares de los bancos A110 y B120 de filtros.

El codificador A120 de banda estrecha está implementado según un modelo de filtro fuente que codifica la señal de voz de entrada como (A) un conjunto de parámetros que describe un filtro y (B) una señal de excitación que activa al filtro descrito para generar una reproducción sintetizada de la señal de voz de entrada. La figura 5a muestra un ejemplo de una envolvente espectral de una señal de voz. Los picos que caracterizan esta envolvente espectral representan resonancias del tracto vocal y se denominan formantes. La mayoría de codificadores de voz codifican al menos esta estructura espectral aproximada como un conjunto de parámetros tales como coeficientes de filtro.

La figura 5b muestra un ejemplo de una disposición básica de filtro fuente aplicada a la codificación de la envolvente espectral de la señal S20 de banda estrecha. Un módulo de análisis calcula un conjunto de parámetros que caracteriza a un filtro correspondiente al sonido de voz durante un periodo de tiempo (normalmente 20 ms). Un filtro blanqueador (también denominado como filtro de análisis o de error de predicción) configurado según esos parámetros de filtro elimina

la envolvente espectral para aplanar espectralmente la señal. La señal blanqueada resultante (también denominada como señal residual) tiene menos energía y, por lo tanto, menos varianza y es más fácil de codificar que la señal de voz original. Los errores generados por la codificación de la señal residual también pueden distribuirse de manera más homogénea a través del espectro. Los parámetros de filtro y la señal residual se cuantifican normalmente para su transmisión eficaz a través del canal. En el descodificador, un filtro de síntesis configurado según los parámetros de filtro se excita mediante una señal basada en la señal residual para generar una versión sintetizada del sonido de voz original. El filtro de síntesis está configurado normalmente para presentar una función de transferencia que sea la inversa de la función de transferencia del filtro blanqueador.

La figura 6 muestra un diagrama de bloques de una implementación A122 básica del codificador A120 de banda estrecha. En este ejemplo, un módulo 210 de análisis de codificación por predicción lineal (LPC) codifica la envolvente espectral de la señal S20 de banda estrecha como un conjunto de coeficientes de predicción lineal (LP) (por ejemplo, coeficientes de un filtro todo polos  $1/A(z)$ ). El módulo de análisis procesa normalmente la señal de entrada como una serie de tramas no solapadas, con un nuevo conjunto de coeficientes que se calculan para cada trama. El periodo de trama es generalmente un periodo sobre el que puede esperarse que la señal sea localmente estacionaria; un ejemplo común es 20 milisegundos (equivalente a 160 muestras a una tasa de muestreo de 8 kHz). En un ejemplo, el módulo 210 de análisis LPC está configurado para calcular un conjunto de diez coeficientes de filtro LP para caracterizar la estructura de formante de cada trama de 20 milisegundos. También es posible implementar el módulo de análisis para procesar la señal de entrada como una serie de tramas solapadas.

El módulo de análisis puede configurarse para analizar directamente las muestras de cada trama, o las muestras pueden ponderarse primero según una función de división en ventanas (por ejemplo, una ventana de Hamming). El análisis también puede realizarse sobre una ventana que sea más larga que la trama, tal como una ventana de 30 ms. Esta ventana puede ser simétrica (por ejemplo, 5-20-5, de manera que incluya los 5 milisegundos inmediatamente antes y después de la trama de 20 milisegundos) o asimétrica (por ejemplo, 10-20, de manera que incluya los últimos 10 milisegundos de la trama anterior). Un módulo de análisis LPC está configurado normalmente para calcular los coeficientes de filtro LP utilizando la recursividad de Levinson-Durbin o el algoritmo de Leroux-Gueguen. En otra implementación, el módulo de análisis puede configurarse para calcular un conjunto de coeficientes cepstrales para cada trama en lugar de un conjunto de coeficientes de filtro LP.

La velocidad de salida del codificador A120 puede reducirse significativamente, sin afectar demasiado a la calidad de reproducción, cuantificando los parámetros de filtro. Los coeficientes de filtro de predicción lineal son difíciles de cuantificar eficazmente y normalmente se correlacionan en otra representación, tal como pares espectrales lineales (LSP) o frecuencias espectrales lineales (LSF), para la cuantificación y/o la codificación por entropía. En el ejemplo de la figura 6, la transformada 220 de coeficientes de filtro LP a LSF transforma el conjunto de coeficientes de filtro LP en un conjunto correspondiente de LSF. Otras representaciones biunívocas de coeficientes de filtro LP incluyen coeficientes paracor, valores del logaritmo de relación de área, pares espectrales de inmitancia (ISP) y frecuencias espectrales de inmitancia (ISF), que se utilizan en el códec AMR-WB (banda ancha adaptable a múltiples velocidades) de GSM (sistema global para comunicaciones móviles). Normalmente, una transformada entre un conjunto de coeficientes de filtro LP y un conjunto correspondiente de LSF es reversible, pero las realizaciones también incluyen implementaciones del codificador A120 en las que la transformada no es reversible sin errores.

Un cuantificador 230 está configurado para cuantificar el conjunto de LSF de banda estrecha (u otra representación de coeficientes), y el codificador A122 de banda estrecha está configurado para transmitir el resultado de esta cuantificación como los parámetros S40 de filtro de banda estrecha. Un cuantificador de este tipo incluye normalmente un cuantificador vectorial que codifica el vector de entrada como un índice a una entrada vectorial correspondiente de una tabla o libro de códigos.

Tal y como se observa en la figura 6, el codificador A122 de banda estrecha también genera una señal residual haciendo pasar la señal S20 de banda estrecha a través de un filtro 260 blanqueador (también denominado como filtro de análisis o de error de predicción) que está configurado según el conjunto de coeficientes de filtro. En este ejemplo particular, el filtro 260 blanqueador está implementado como un filtro FIR, aunque también pueden utilizarse implementaciones IIR. Esta señal residual contendrá normalmente información perceptualmente importante de la trama de voz, tal como una estructura de larga duración relacionada con el tono, que no está representada en los parámetros S40 de filtro de banda estrecha. El cuantificador 270 está configurado para calcular una representación cuantificada de esta señal residual para transmitirse como una señal S50 de excitación de banda estrecha codificada. Un cuantificador de este tipo incluye normalmente un cuantificador vectorial que codifica el vector de entrada como un índice a una entrada de vector correspondiente de una tabla o libro de códigos. Como alternativa, un cuantificador de este tipo puede configurarse para enviar uno o más parámetros a partir de los cuales el vector puede generarse dinámicamente en el descodificador, en lugar de recuperarse de un medio de almacenamiento, como en un procedimiento de libro de códigos disperso. Un procedimiento de este tipo se utiliza en esquemas de codificación tales como CELP (predicción lineal de excitación por libro de códigos) algebraica y códecs tales como EVRC (códec de velocidad variable mejorado) del 3GPP2 (segundo proyecto de colaboración de tercera generación).

Es deseable que el codificador A120 de banda estrecha genere la señal codificada de excitación de banda estrecha según los mismos valores de parámetro de filtro que estarán disponibles para el descodificador de banda estrecha correspondiente. De esta manera, la señal codificada de excitación de banda estrecha resultante ya puede soportar, hasta cierto punto, imprecisiones en esos valores de parámetros, tales como errores de cuantificación. Por consiguiente, es deseable configurar el filtro blanqueador utilizando los mismos valores de coeficiente que estarán disponibles en el descodificador. En el ejemplo básico del codificador A122 mostrado en la figura 6, un cuantificador 240 inverso descuantifica los parámetros S40 de codificación de banda estrecha, una transformada 250 de LSF a coeficientes de filtro LP correlaciona los valores resultantes con un conjunto correspondiente de coeficientes de filtro LP, y este conjunto de coeficientes se utiliza para configurar un filtro 260 blanqueador para generar la señal residual que se cuantifica mediante un cuantificador 270.

Algunas implementaciones del codificador A120 de banda estrecha están configuradas para calcular la señal S50 codificada de excitación de banda estrecha identificando un vector de libro de códigos de un conjunto de vectores de libro de códigos que se ajuste mejor a la señal residual. Sin embargo, debe observarse que el codificador A120 de banda estrecha también puede implementarse para calcular una representación cuantificada de la señal residual sin generar realmente la señal residual. Por ejemplo, el codificador A120 de banda estrecha puede configurarse para utilizar una pluralidad de vectores de libro de códigos para generar señales sintetizadas correspondientes (por ejemplo, según un conjunto actual de parámetros de filtro), y para seleccionar el vector de libro de códigos asociado a la señal generada que se ajuste mejor a las señal S20 de banda estrecha original en un dominio ponderado de manera perceptual.

La figura 7 muestra un diagrama de bloques de una implementación B112 del descodificador B110 de banda estrecha. Un cuantificador 310 inverso descuantifica los parámetros S40 de filtro de banda estrecha (en este caso, a un conjunto de LSF), y una transformada 320 de LSF a coeficientes de filtro de LP transforma las LSF en un conjunto de coeficientes de filtro (por ejemplo, tal y como se ha descrito anteriormente con referencia al cuantificador 240 inverso y a la transformada 250 del codificador A122 de banda estrecha). El cuantificador 340 inverso descuantifica la señal S40 residual de banda estrecha para generar una señal S80 de excitación de banda estrecha. En función de los coeficientes de filtro y de la señal S80 de excitación de banda estrecha, un filtro 330 de síntesis de banda estrecha sintetiza la señal S90 de banda estrecha. Dicho de otro modo, el filtro 330 de síntesis de banda estrecha está configurado para conformar de manera espectral la señal S80 de excitación de banda estrecha según los coeficientes de filtro descuantificados para generar la señal S90 de banda estrecha. El descodificador B112 de banda estrecha también proporciona la señal S80 de excitación de banda estrecha al codificador A200 de banda alta, el cual la utiliza para obtener la señal S120 de excitación de banda alta como la descrita en este documento. En algunas implementaciones descritas posteriormente, el descodificador B110 de banda estrecha puede configurarse para proporcionar al descodificador B200 de banda alta información adicional relacionada con la señal de banda estrecha, tal como la inclinación espectral, retardo y ganancia de tono, y modo de voz.

El sistema del codificador A122 de banda estrecha y del descodificador B112 de banda estrecha es un ejemplo básico de un códec de voz de análisis por síntesis. La codificación de predicción lineal excitada por libro de códigos (CELP) es un tipo conocido de codificación de análisis por síntesis, y las implementaciones de tales codificadores pueden llevar a cabo la codificación de forma de onda de la señal residual, incluyendo operaciones tales como la selección de entradas de libros de códigos fijos y adaptativos, operaciones de minimización de errores y/u operaciones de ponderación perceptuales. Otras implementaciones de codificación de análisis por síntesis incluyen codificación de predicción lineal de excitación mixta (MELP), de CELP algebraica (ACELP), de CELP de relajación (RCELP), de excitación por impulsos regulares (RPE), de CELP de múltiples impulsos (MPE) y de predicción lineal excitada por suma vectorial (VSELP). Procedimientos de codificación relacionados incluyen codificación de excitación por múltiples bandas (MBE) y codificación de interpolación de forma de onda de prototipo (PWI). Ejemplos de códecs de voz normalizados de análisis por síntesis incluyen el códec de velocidad total GSM (GSM 06.10) del ETSI (Instituto Europeo de Normas de Telecomunicaciones), el cual utiliza predicción lineal excitada por señal residual (RELP); el códec de velocidad total mejorado GSM (ETSI-GSM 06.60); el codificador normalizado del anexo E del G.729 de 11,8 kb/s de la ITU (Unión Internacional de Telecomunicaciones); códecs IS-641 (norma provisional) para IS-136 (un esquema de acceso múltiple por división de tiempo); los códecs GSM adaptables a múltiples velocidades GSM (GSM-AMR); y el códec 4GV<sup>TM</sup> (vocoder<sup>TM</sup> de cuarta generación) (QUALCOMM Incorporated, San Diego, CA). El codificador A120 de banda estrecha y el descodificador B110 correspondiente pueden implementarse según cualquiera de estas tecnologías o según cualquier otra tecnología de codificación de voz (ya sea conocida o desarrollada en un futuro) que represente una señal de voz como (A) un conjunto de parámetros que describa un filtro y (B) una señal de excitación utilizada para activar el filtro descrito para generar la señal de voz.

Incluso después de que el filtro blanqueador haya eliminado la envolvente espectral aproximada de la señal S20 de banda estrecha, puede haber una considerable cantidad de una fina estructura de armónico, especialmente para voz sonora. La figura 8a muestra una representación gráfica espectral de un ejemplo de una señal residual, como la que puede generarse por un filtro blanqueador, para una señal sonora tal como una vocal. La estructura periódica visible en este ejemplo está relacionada con el tono, y diferentes sonidos sonoros producidos por el mismo hablante pueden tener diferentes estructuras de formante pero estructuras de tono similares. La figura 8b muestra una representación gráfica en el dominio de tiempo de un ejemplo de una señal residual de este tipo que muestra una secuencia de impulsos de tono en el tiempo.

La eficacia de codificación y/o la calidad de la voz pueden aumentarse utilizando uno o más valores de parámetro para codificar las características de la estructura de tono. Una característica importante de la estructura de tono es la frecuencia del primer armónico (también denominada como frecuencia fundamental), que está normalmente en el intervalo de 60 a 400 Hz. Esta característica se codifica normalmente como la inversa de la frecuencia fundamental, denominada también desfase de tono (*pitch lag*). El desfase de tono indica el número de muestras de un periodo de tono y puede codificarse como uno o más índices de libro de código. Las señales de voz de los hombres tienden a tener mayores desfases de tono que las señales de voz de las mujeres.

Otra característica de señal relacionada con la estructura de tono es la periodicidad, la cual indica la intensidad de la estructura de armónico o, dicho de otro modo, el grado en que la señal es armónica o no armónica. Dos indicadores típicos de la periodicidad son los cruces por cero y las funciones de autocorrelación normalizada (NACF). La periodicidad también puede indicarse mediante la ganancia de tono, la cual se codifica normalmente como una ganancia de libro de códigos (por ejemplo, una ganancia cuantificada de libro de códigos adaptativo).

El codificador A120 de banda estrecha puede incluir uno o más módulos configurados para codificar la estructura de armónico de larga duración de la señal S20 de banda estrecha. Tal y como se muestra en la figura 9, un paradigma CELP típico que puede utilizarse incluye un módulo de análisis LPC de bucle abierto, el cual codifica las características de corta duración de la envolvente espectral aproximada, seguido por una etapa de análisis de predicción a largo plazo de bucle cerrado que codifica la fina estructura de tono o de armónico. Las características de corta duración se codifican como coeficientes de filtro, y las características de larga duración se codifican como valores para parámetros tales como el desfase de tono y la ganancia de tono. Por ejemplo, el codificador A120 de banda estrecha puede configurarse para transmitir la señal S50 de excitación de banda estrecha codificada en un forma que incluya uno o más índices de libro de códigos (por ejemplo, un índice de libro de códigos fijo y un índice de libro de códigos adaptativo) y valores de ganancia correspondientes. El cálculo de esta representación cuantificada de la señal residual de banda estrecha (por ejemplo, mediante el cuantificador 270), puede incluir seleccionar tales índices y calcular tales valores. La codificación de la estructura de tono también puede incluir interpolación de una forma de onda de prototipo de tono, operación que puede incluir calcular una diferencia entre impulsos de tono sucesivos. La modelización de la estructura de larga duración puede inhabilitarse para tramas correspondientes a voz sorda, que normalmente es similar al ruido y no está estructurada.

Una implementación del descodificador B110 de banda estrecha según un paradigma como el mostrado en la figura 9 puede configurarse para transmitir la señal S80 de excitación de banda estrecha al descodificador B200 de banda alta después de que la estructura de larga duración (estructura de tono o de armónico) se haya restablecido. Por ejemplo, un descodificador de este tipo puede configurarse para transmitir la señal S80 de excitación de banda estrecha como una versión descuantificada de la señal S50 de excitación de banda estrecha codificada. Por supuesto, también es posible implementar el descodificador B110 de banda estrecha de manera que el descodificador B200 de banda alta realice la descuantificación de la señal S50 de excitación de banda estrecha codificada para obtener la señal S80 de excitación de banda estrecha.

En una implementación del codificador A100 de voz de banda ancha según un paradigma como el mostrado en la figura 9, el codificador A200 de banda alta puede configurarse para recibir la señal de excitación de banda estrecha generada por el análisis a corto plazo o por el filtro blanqueador. Dicho de otro modo, el codificador A120 de banda estrecha puede configurarse para transmitir la señal de excitación de banda estrecha al codificador A200 de banda alta antes de codificar la estructura de larga duración. Sin embargo, es deseable que el codificador A200 de banda alta reciba del canal de banda estrecha la misma información de codificación que se recibirá mediante el descodificador B200 de banda alta, de manera que los parámetros de codificación generados por el codificador A200 de banda alta ya puedan soportar hasta cierto punto imprecisiones en esa información. Por lo tanto, puede ser preferible que el codificador A200 de banda alta reconstruya la señal S80 de excitación de banda estrecha a partir de la misma señal S50 codificada de excitación de banda estrecha cuantificada y/o parametrizada que se transmitirá por el codificador A100 de voz de banda ancha. Una ventaja potencial de este enfoque es un cálculo más preciso de los factores S60b de ganancia de banda alta que se describirán posteriormente.

Además de los parámetros que caracterizan la estructura de corta duración y/o de larga duración de la señal S20 de banda estrecha, el codificador A120 de banda estrecha puede generar valores de parámetro que se refieren a otras características de la señal S20 de banda estrecha. Estos valores, que pueden cuantificarse de manera adecuada para su transmisión mediante el codificador A100 de voz de banda ancha, pueden incluirse entre los parámetros S40 de filtro de banda estrecha o transmitirse por separado. El codificador A200 de banda alta también puede estar configurado para calcular los parámetros S60 de codificación de banda alta según uno o más de estos parámetros adicionales (por ejemplo, después de la descuantificación). En el descodificador B100 de voz de banda ancha, el descodificador B200 de banda alta puede configurarse para recibir los valores de parámetro a través del descodificador B110 de banda estrecha (por ejemplo, después de la descuantificación). Como alternativa, el descodificador B200 de banda alta puede configurarse para recibir (y posiblemente para descuantificar) los valores de parámetro directamente.

En un ejemplo de parámetros de codificación de banda estrecha adicionales, el codificador A120 de banda estrecha

- 5 genera valores de inclinación espectral y de parámetros de modo de voz para cada trama. La inclinación espectral se refiere a la forma de la envolvente espectral sobre la banda de paso y se representa normalmente por el primer coeficiente de reflexión cuantificado. Para la mayoría de sonidos sonoros, la energía espectral disminuye con frecuencias crecientes, de manera que el primer coeficiente de reflexión es negativo y puede aproximarse a -1. La mayoría de sonidos sordos tienen un espectro que es o bien plano, de manera que el primer coeficiente de reflexión vale casi cero, o bien tiene más energía a altas frecuencias, de manera que el primer coeficiente de reflexión es positivo y puede aproximarse a +1.
- 10 El modo de voz (también denominado como modo de sonoridad) indica si la trama actual representa voz sonora o sorda. Este parámetro puede tener un valor binario en función de una o más medidas de periodicidad (por ejemplo, cruces por cero, NACF, ganancia de tono) y/o actividad de voz para la trama, tal como una relación entre una medida de este tipo y un valor de umbral. En otras implementaciones, el parámetro de modo de voz tiene uno o más estados diferentes para indicar modos tales como silencio o ruido de fondo, o una transición entre silencio y voz sonora.
- 15 El codificador A200 de banda alta está configurado para codificar la señal S30 de banda alta según un modelo de filtro fuente, basándose la excitación de este filtro en la señal codificada de excitación de banda estrecha. La figura 10 muestra un diagrama de bloques de una implementación A202 del codificador A200 de banda alta que está configurada para generar un flujo de parámetros S60 de codificación de banda alta que incluye parámetros S60a de filtro de banda alta y factores S60b de ganancia de banda alta. El generador A300 de excitación de banda alta obtiene una señal S120 de excitación de banda alta a partir de la señal S50 de excitación de banda estrecha codificada. Un módulo A210 de análisis genera un conjunto de valores de parámetro que caracteriza a la envolvente espectral de la señal S30 de banda alta. En este ejemplo particular, un módulo A210 de análisis está configurado para realizar un análisis LPC para generar un conjunto de coeficientes de filtro LP para cada trama de la señal S30 de banda alta. Una transformada 410 de coeficientes de filtro de predicción lineal a LSF transforma el conjunto de coeficientes de filtro LP en un conjunto correspondiente de LSF. Tal y como se ha indicado anteriormente con referencia al módulo 210 de análisis y a la transformada 220, el módulo A210 de análisis y/o la transformada 410 pueden configurarse para utilizar otros conjuntos de coeficientes (por ejemplo, coeficientes cepstrales) y/o representaciones de coeficientes (por ejemplo, ISP).
- 20
- 25 Un cuantificador 420 está configurado para cuantificar el conjunto de LSF de banda alta (u otra representación de coeficientes, tal como ISP), y un codificador A202 de banda alta está configurado para transmitir el resultado de esta cuantificación como los parámetros S60a de filtro de banda alta. Un cuantificador de este tipo incluye normalmente un cuantificador vectorial que codifica el vector de entrada como un índice a una entrada de vector correspondiente de una tabla o libro de códigos.
- 30 El codificador A202 de banda alta también incluye un filtro A220 de síntesis configurado para generar una señal S130 de banda alta sintetizada según la señal S120 de excitación de banda alta y la envolvente espectral codificada (por ejemplo, el conjunto de coeficientes de filtro LP) generada por el módulo A210 de análisis. El filtro A220 de síntesis está implementado normalmente como un filtro IIR, aunque también pueden usarse implementaciones FIR. En un ejemplo particular, el filtro A220 de síntesis está implementado como un filtro autorregresivo lineal de sexto orden.
- 35 Un calculador A230 de factor de ganancia de banda alta calcula una o más diferencias entre los niveles de la señal S30 de banda alta original y la señal S130 de banda alta sintetizada para especificar una envolvente de ganancia para la trama. El cuantificador 430, que puede implementarse como un cuantificador vectorial que codifica el vector de entrada como un índice a una entrada de vector correspondiente de una tabla o libro de códigos, cuantifica el valor o valores que especifican la envolvente de ganancia, y el codificador A202 de banda alta está configurado para transmitir el resultado de esta cuantificación como factores S60b de ganancia de banda alta.
- 40
- 45 En una implementación como la mostrada en la figura 10, el filtro A220 de síntesis está dispuesto para recibir los coeficientes de filtro del módulo A210 de análisis. Una implementación alternativa del codificador A202 de banda alta incluye un cuantificador inverso y una transformada inversa configurados para descodificar los coeficientes de filtro de los parámetros S60a de filtro de banda alta y, en este caso, el filtro A220 de síntesis está dispuesto en cambio para recibir los coeficientes de filtro descodificados. Una disposición alternativa de este tipo puede soportar un cálculo más preciso de la envolvente de ganancia mediante el calculador A230 de ganancia de banda alta.
- 50 En un ejemplo particular, el módulo A210 de análisis y el calculador A230 de ganancia de banda alta transmiten un conjunto de seis LSF y un conjunto de cinco valores de ganancia por trama, respectivamente, de manera que puede conseguirse un ensanchamiento de banda ancha de la señal S20 de banda estrecha solamente con once valores adicionales por trama. El oído tiende a ser menos sensible a los errores de frecuencia a altas frecuencias, de manera que la codificación de banda alta en un orden LPC bajo puede generar una señal que presente una calidad perceptual comparable a la codificación de banda estrecha en un orden LPC superior. Una implementación típica del codificador A200 de banda alta puede configurarse para transmitir entre 8 y 12 bits por trama para una reconstrucción de alta calidad de la envolvente espectral y entre 8 y 12 bits adicionales por trama para una reconstrucción de alta calidad de la envolvente temporal. En otro ejemplo particular, el módulo A210 de análisis transmite un conjunto de ocho LSF por trama.
- 55
- Algunas implementaciones del codificador A200 de banda alta están configuradas para generar la señal S120 de

excitación de banda alta generando una señal de ruido aleatorio que presenta componentes de frecuencia de banda alta y modulando en amplitud la señal de ruido según la envolvente de dominio de tiempo de la señal S20 de banda estrecha, de la señal S80 de excitación de banda estrecha o de la señal S30 de banda alta. Aunque un procedimiento de este tipo basado en ruido puede generar resultados adecuados para sonidos sordos, puede no ser deseable para sonidos sonoros, cuyas señales residuales son normalmente armónicas y, por consiguiente, presentan alguna estructura periódica.

El generador A300 de excitación de banda alta está configurado para generar la señal S120 de excitación de banda alta ensanchado el espectro de la señal S80 de excitación de banda estrecha en el intervalo de frecuencias de banda alta. La figura 11 muestra un diagrama de bloques de una implementación A302 del generador A300 de excitación de banda alta. Un cuantificador 450 inverso está configurado para descuantificar la señal S50 codificada de excitación de banda estrecha para generar la señal S80 de excitación de banda estrecha. Un ensanchador A400 de espectro está configurado para generar una señal S160 ensanchada de manera armónica en función de la señal S80 de excitación de banda estrecha. Un combinador 470 está configurado para combinar una señal de ruido aleatorio generada por un generador 480 de ruido y una envolvente de dominio de tiempo calculada por un calculador 460 de envolvente para generar una señal S170 de ruido modulada. Un combinador 490 está configurado para mezclar la señal S160 ensanchada de manera armónica y la señal S170 de ruido modulada para generar la señal S120 de excitación de banda alta.

En un ejemplo, el ensanchador A400 de espectro está configurado para realizar una operación de plegado espectral (también denominada como reflexión) en la señal S80 de excitación de banda estrecha para generar la señal S160 ensanchada de manera armónica. El plegado espectral puede llevarse a cabo rellenando con ceros la señal S80 de excitación y aplicando después un filtro paso alto para mantener la superposición. En otro ejemplo, el ensanchador A400 de espectro está configurado para generar la señal S160 ensanchada de manera armónica trasladando de manera espectral la señal S80 de excitación de banda estrecha hacia la banda alta (por ejemplo, mediante un muestreo ascendente seguido de una multiplicación por una señal de coseno de frecuencia constante).

Los procedimientos de traslación y de plegado espectral pueden generar señales ensanchadas de manera espectral cuya estructura armónica es discontinua con la estructura armónica original de la señal S80 de excitación de banda estrecha en fase y/o frecuencia. Por ejemplo, tales procedimientos pueden generar señales que presenten picos que no estén situados generalmente en múltiplos de la frecuencia fundamental, lo que puede provocar artefactos de sonido metálico en la señal de voz reconstruida. Estos procedimientos también tienden a generar armónicos de alta frecuencia que presentan características tonales extrañamente fuertes. Además, puesto que una señal PSTN puede muestrearse a 8 kHz pero limitarse en banda a no más de 3400 Hz, el espectro superior de la señal S80 de excitación de banda estrecha puede contener poca o ninguna energía, de manera que una señal ensanchada generada según una operación de plegado espectral o de traslación espectral puede presentar un agujero espectral por encima de 3400 Hz.

Otros procedimientos de generación de la señal S160 ensanchada de manera armónica incluyen identificar una o más frecuencias fundamentales de la señal S80 de excitación de banda estrecha y generar tonos de armónico según esa información. Por ejemplo, la estructura armónica de una señal de excitación puede caracterizarse por la frecuencia fundamental junto con información de amplitud y de fase. Otra implementación del generador A300 de excitación de banda alta genera una señal S160 ensanchada de manera armónica en función de la frecuencia fundamental y de la amplitud (según se indica, por ejemplo, por el desfase de tono y la ganancia de tono). Sin embargo, a no ser que la señal ensanchada de manera armónica sea coherente en fase con la señal S80 de excitación de banda estrecha, la calidad de la voz descodificada resultante puede no ser aceptable.

Puede utilizarse una función no lineal para crear una señal de excitación de banda alta que sea coherente en fase con la excitación de banda estrecha y que conserve la estructura armónica sin discontinuidad de fase. Una función no lineal también puede proporcionar un mayor nivel de ruido entre los armónicos de alta frecuencia, lo que tiende a sonar más natural que los armónicos de alta frecuencia tonales generados por procedimientos tales como el plegado espectral y la traslación espectral. Funciones típicas no lineales sin memoria que pueden aplicarse mediante varias implementaciones del ensanchador A400 de espectro incluyen la función de valor absoluto (también denominada como rectificación de onda completa), rectificación de media onda, elevación al cuadrado, elevación al cubo y truncamiento. Otras implementaciones del ensanchador A400 de espectro pueden configurarse para aplicar una función no lineal que tenga memoria.

La figura 12 es un diagrama de bloques de una implementación A402 del ensanchador A400 de espectro que está configurada para aplicar una función no lineal para ensanchar el espectro de la señal S80 de excitación de banda estrecha. Un muestreador 510 ascendente está configurado para muestrear de manera ascendente la señal S80 de excitación de banda estrecha. Puede ser deseable muestrear de manera ascendente la señal lo suficiente como para minimizar la superposición tras la aplicación de la función no lineal. En un ejemplo particular, el muestreador 510 ascendente muestrea de manera ascendente la señal por un factor de ocho. El muestreador 510 ascendente puede configurarse para realizar la operación de muestreo ascendente rellenando con ceros la señal de entrada y filtrando mediante paso bajo el resultado. Un calculador 520 de función no lineal está configurado para aplicar una función no lineal a la señal muestreada de manera ascendente. Una ventaja potencial de la función de valor absoluto sobre otras funciones no lineales de ensanchamiento espectral, tales como la elevación al cuadrado, es que la normalización de energía no es

necesaria. En algunas implementaciones, la función de valor absoluto puede aplicarse de manera eficaz suprimiendo o eliminando el bit de signo de cada muestra. El calculador 520 de función no lineal también puede configurarse para realizar una distorsión de amplitud de la señal muestreada de manera ascendente o ensanchada de manera espectral.

5 Un muestreador 530 descendente está configurado para muestrear de manera descendente el resultado, ensanchado de manera espectral, de aplicar la función no lineal. Puede ser deseable que el muestreador 530 descendente realice una operación de filtrado de paso de banda para seleccionar una banda de frecuencia deseada de la señal ensanchada de manera espectral antes de reducir la tasa de muestreo (por ejemplo, para reducir o evitar la superposición o la corrupción mediante una imagen no deseada). También puede ser deseable que el muestreador 530 descendente reduzca la tasa de muestreo en más de una etapa.

10 La figura 12a es un diagrama que muestra los espectros de señal en varios puntos en un ejemplo de una operación de ensanchamiento espectral, donde la escala de frecuencia es la misma en todas las representaciones gráficas. La representación gráfica (a) muestra el espectro de un ejemplo de la señal S80 de excitación de banda estrecha. La representación gráfica (b) muestra el espectro después de que la señal S80 se haya muestreado de manera ascendente por un factor de ocho. La representación gráfica (c) muestra un ejemplo del espectro ensanchado después de la aplicación de una función no lineal. La representación gráfica (d) muestra el espectro después de un filtrado de paso bajo. En este ejemplo, la banda de paso se ensancha hasta el límite de frecuencia superior de la señal S30 de banda alta (por ejemplo, 7 kHz u 8 kHz).

15 La representación gráfica (e) muestra el espectro después de una primera etapa de muestreo descendente, en la que la tasa de muestreo se reduce por un factor de cuatro para obtener una señal de banda ancha. La representación gráfica (f) muestra el espectro después de una operación de filtrado de paso alto para seleccionar la parte de banda alta de la señal ensanchada, y la representación gráfica (g) muestra el espectro después de una segunda etapa de muestreo descendente, en la que la tasa de muestreo se reduce por un factor de dos. En un ejemplo particular, el muestreador 530 descendente realiza el filtrado de paso alto y la segunda etapa de muestreo descendente haciendo pasar la señal de banda ancha a través del filtro 130 paso alto y del muestreador 140 descendente del banco A112 de filtros (o de otras estructuras o rutinas que tengan la misma respuesta) para generar una señal ensanchada de manera espectral que presente el intervalo de frecuencias y la tasa de muestreo de la señal S30 de banda alta.

20 Como puede observarse en la representación gráfica (g), el muestreo descendente de la señal de paso alto mostrada en la representación gráfica (f) provoca la inversión de su espectro. En este ejemplo, el muestreador 530 descendente también está configurado para realizar una operación de inversión espectral en la señal. La representación gráfica (h) muestra un resultado de la aplicación de la operación de inversión espectral, la cual puede realizarse multiplicando la señal por la función  $e^{-jn\pi}$  o por la secuencia  $(-1)^n$ , cuyos valores alternan entre +1 y -1. Una operación de este tipo es equivalente a desplazar el espectro digital de la señal en el dominio de frecuencia en un distancia de  $\pi$ . Debe observarse que también puede obtenerse el mismo resultado aplicando las operaciones de muestreo descendente y de inversión espectral. Las operaciones de muestreo ascendente y/o de muestreo descendente también pueden configurarse para incluir remuestreo para obtener una señal ensanchada de manera espectral que presente la tasa de muestreo de la señal S30 de banda alta (por ejemplo, 7 kHz).

25 Tal y como se ha indicado anteriormente, los bancos A110 y B120 de filtros pueden implementarse de manera que una o ambas señales S20, S30 de banda estrecha y de banda alta presente(n) una forma invertida de manera espectral en la salida del banco A110 de filtros, se codifique(n) y se descodifique(n) en la forma invertida de manera espectral y se invierta(n) de nuevo de manera espectral en el banco B120 de filtros antes de transmitirse en la señal S110 de voz de banda ancha. En ese caso, por supuesto, una operación de inversión espectral como la mostrada en la figura 12a no sería necesaria ya que sería deseable que la señal S120 de excitación de banda alta también presentase una forma invertida de manera espectral.

30 Las diversas tareas de muestreo ascendente y de muestreo descendente de una operación de ensanchamiento espectral como la realizada por el ensanchador A402 de espectro pueden configurarse y disponerse de muchas maneras diferentes. Por ejemplo, la figura 12b es un diagrama que muestra los espectros de señal en varios puntos en otro ejemplo de una operación de extensión espectral, donde la escala de frecuencia es la misma en todas las representaciones gráficas. La representación gráfica (a) muestra el espectro de un ejemplo de la señal S80 de excitación de banda estrecha. La representación gráfica (b) muestra el espectro después de que la señal S80 se haya muestreado de manera ascendente por un factor de dos. La representación gráfica (c) muestra un ejemplo del espectro ensanchado después de la aplicación de una función no lineal. En este caso se acepta la superposición que puede producirse en las frecuencias más altas.

35 La representación gráfica (d) muestra el espectro después de una operación de inversión espectral. La representación gráfica (e) muestra el espectro después de una única etapa de muestreo descendente, en la que la tasa de muestreo se reduce por un factor de dos para obtener la señal ensanchada de manera espectral deseada. En este ejemplo, la señal está en una forma invertida de manera espectral y puede utilizarse en una implementación del codificador A200 de banda alta que procese la señal S30 de banda alta en esta forma.

Es probable que la señal ensanchada de manera espectral generada por el calculador 520 de función no lineal presente una gran disminución de amplitud a medida que aumente la frecuencia. El ensanchador A402 espectral incluye un aplanador 540 espectral configurado para realizar una operación de blanqueo en la señal muestreada de manera descendente. El aplanador 540 espectral puede configurarse para realizar una operación de blanqueo fija o para realizar una operación de blanqueo adaptativa. En un ejemplo particular de blanqueo adaptativo, el aplanador 540 espectral incluye un módulo de análisis LPC configurado para calcular un conjunto de cuatro coeficientes de filtro a partir de la señal muestreada de manera descendente y un filtro de análisis de cuarto orden configurado para blanquear la señal según esos coeficientes. Otras implementaciones del ensanchador A400 de espectro incluyen configuraciones en las que el aplanador 540 espectral actúa en la señal ensanchada de manera espectral antes del muestreador 530 descendente.

El generador A300 de excitación de banda alta puede implementarse para transmitir la señal S160 ensanchada de manera armónica como la señal S120 de excitación de banda alta. Sin embargo, en algunos casos, utilizando solamente una señal ensanchada de manera armónica como la excitación de banda alta puede dar como resultado artefactos audibles. La estructura armónica de la voz es generalmente menos pronunciada en la banda alta que en la banda baja, y la utilización de una estructura demasiado armónica en la señal de excitación de banda alta puede dar como resultado un sonido con zumbidos. Este artefacto puede apreciarse especialmente en señales de voz de mujeres.

Las realizaciones incluyen implementaciones del generador A300 de excitación de banda alta que están configuradas para mezclar la señal S160 ensanchada de manera armónica con una señal de ruido. Tal y como se muestra en la figura 11, el generador A302 de excitación de banda alta incluye un generador 480 de ruido que está configurado para generar una señal de ruido aleatorio. En un ejemplo, el generador 480 de ruido está configurado para generar una señal de ruido blanco pseudoaleatorio de varianza unitaria aunque en otras implementaciones la señal de ruido no necesita ser blanca y puede presentar una densidad de potencia que varíe con la frecuencia. Puede ser deseable que los generadores 480 de ruido estén configurados para transmitir la señal de ruido como una función determinista de manera que su estado pueda duplicarse en el decodificador. Por ejemplo, el generador 480 de ruido puede configurarse para transmitir la señal de ruido como una función determinista de información codificada anteriormente dentro de la misma trama, tal como los parámetros S40 de filtro de banda estrecha y/o la señal S50 codificada de excitación de banda estrecha.

Antes de mezclarse con la señal S160 ensanchada de manera armónica, la señal de ruido aleatorio generada por el generador 480 de ruido puede modularse en amplitud para presentar una envolvente de dominio de tiempo que aproxime la distribución de energía en el tiempo de la señal S20 de banda estrecha, de la señal S30 de banda alta, de la señal S80 de excitación de banda estrecha o de la señal S160 ensanchada de manera armónica. Tal y como se muestra en la figura 11, el generador A302 de excitación de banda alta incluye un combinador 470 configurado para modular en amplitud la señal de ruido generada por el generador 480 de ruido según una envolvente de dominio de tiempo calculada por el calculador 460 de envolvente. Por ejemplo, el combinador 470 puede implementarse como un multiplicador dispuesto para escalar la salida del generador 480 de ruido según la envolvente de dominio de tiempo calculada por el calculador 460 de envolvente para generar la señal S170 de ruido modulada.

En una implementación A304 del generador A302 de excitación de banda alta, como la mostrada en el diagrama de bloques de la figura 13, el calculador 460 de envolvente está dispuesto para calcular la envolvente de la señal S160 ensanchada de manera armónica. En una implementación A306 del generador A302 de excitación de banda alta, como la mostrada en el diagrama de bloques de la figura 14, el calculador 460 de envolvente está dispuesto para calcular la envolvente de la señal S80 de excitación de banda estrecha. Implementaciones adicionales del generador A302 de excitación de banda alta pueden configurarse de otro modo para añadir ruido a la señal S160 ensanchada de manera armónica según las posiciones de los impulsos de tono de banda estrecha en el tiempo.

El calculador 460 de envolvente puede configurarse para realizar un cálculo de envolvente como una tarea que incluya una serie de subtareas. La figura 15 muestra un diagrama de flujo de un ejemplo T100 de una tarea de este tipo. La subtarea T110 calcula el cuadrado de cada muestra de la trama de la señal cuya envolvente va a modelarse (por ejemplo, la señal S80 de excitación de banda estrecha o la señal S160 ensanchada de manera armónica) para generar una secuencia de valores al cuadrado. La subtarea T120 realiza una operación de suavizado en la secuencia de valores al cuadrado. En un ejemplo, la subtarea T120 aplica un filtro paso bajo IIR de primer orden a la secuencia según la expresión

$$y(n) = ax(n) + (1 - a)y(n - 1), \quad (1)$$

donde  $x$  es la entrada del filtro,  $y$  es la salida del filtro,  $n$  es un índice de dominio de tiempo, y  $a$  es un coeficiente de suavizado que presenta un valor entre 0,5 y 1. El valor del coeficiente de suavizado  $a$  puede ser fijo o, en una implementación alternativa, puede ser adaptativo según una indicación de ruido en la señal de entrada, de manera que  $a$  está más cerca de 1 en ausencia de ruido y más cerca de 0,5 en presencia de ruido. La subtarea T130 aplica una función de raíz cuadrada a cada muestra de la secuencia suavizada para generar la envolvente de dominio de tiempo.

Una implementación de este tipo del calculador 460 de envolvente puede configurarse para realizar las diversas subtareas

de la tarea T100 en serie y/o en paralelo. En implementaciones adicionales de la tarea T100, la subtarea T110 puede estar precedida por una operación de paso banda configurada para seleccionar una parte de frecuencia deseada de la señal cuya envolvente va a modelarse, tal como el intervalo de 3 a 4 kHz.

5 El combinador 490 está configurado para mezclar la señal S160 ensanchada de manera armónica y la señal S170 de ruido modulada para generar la señal S120 de excitación de banda alta. Implementaciones del combinador 490 pueden configurarse, por ejemplo, para calcular la señal S120 de excitación de banda alta como una suma de la señal S160 ensanchada de manera armónica y la señal S170 de ruido modulada. Una implementación de este tipo del combinador 490 puede configurarse para calcular la señal S120 de excitación de banda alta como una suma ponderada aplicando un factor de ponderación a la señal S160 ensanchada de manera armónica y/o a la señal S170 de ruido modulada antes de la suma. Cada factor de ponderación puede calcularse según uno o más criterios y puede ser un valor fijo o, como alternativa, un valor adaptativo que se calcule para cada trama o para cada subtrama.

10 La figura 16 muestra un diagrama de bloques de una implementación 492 del combinador 490 que está configurada para calcular la señal S120 de excitación de banda alta como una suma ponderada de la señal S160 ensanchada de manera armónica y la señal S170 de ruido modulada. El combinador 492 está configurado para ponderar la señal S160 ensanchada de manera armónica según un factor S180 de ponderación armónico, para ponderar la señal S170 de ruido modulada según un factor S190 de ponderación de ruido y para transmitir la señal S120 de excitación de banda alta como una suma de las señales ponderadas. En este ejemplo, el combinador 492 incluye un calculador 550 de factor de ponderación que está configurado para calcular el factor S180 de ponderación de armónico y el factor S190 de ponderación de ruido.

15 El calculador 550 de factor de ponderación puede configurarse para calcular los factores S180 y S190 de ponderación según una relación deseada de contenido de armónico con respecto a contenido de ruido en la señal S120 de excitación de banda alta. Por ejemplo, puede ser deseable que el combinador 492 genere la señal S120 de excitación de banda alta para que presente una relación de energía de armónico con respecto a energía de ruido similar a la de la señal S30 de banda alta. En algunas implementaciones del calculador 550 de factor de ponderación, los factores S180, S190 de ponderación se calculan según uno o más parámetros relacionados con una periodicidad de la señal S20 de banda estrecha o de la señal residual de banda estrecha, tal como la ganancia de tono y/o el modo de voz. Una implementación de este tipo del calculador 550 de factor de ponderación puede configurarse para asignar un valor al factor S180 de ponderación de armónico que sea proporcional a la ganancia de tono, por ejemplo, y/o para asignar un valor más elevado al factor S190 de ponderación de ruido para señales de voz sordas en lugar de para señales de voz sonoras.

20 En otras implementaciones, el calculador 550 de factor de ponderación está configurado para calcular valores para el factor S180 de ponderación de armónico y/o para el factor S190 de ponderación de ruido según una medida de periodicidad de la señal S30 de banda alta. En un ejemplo de este tipo, el calculador 550 de factor de ponderación calcula el factor S180 de ponderación de armónico como el valor máximo del coeficiente de autocorrelación de la señal S30 de banda alta para la trama o subtramas actuales, donde la autocorrelación se lleva a cabo a través de un intervalo de búsqueda que incluye un retardo de un desfase de tono y que no incluye un retardo de cero muestras. La figura 17 muestra un ejemplo de tal intervalo de búsqueda de muestras de longitud  $n$  que está centrado en torno a un retardo de un desfase de tono y que presenta un ancho no superior a un desfase de tono.

25 La figura 17 también muestra un ejemplo de otro enfoque en el que el calculador 550 de factor de ponderación calcula una medida de periodicidad de la señal S30 de banda alta en varias etapas. En una primera etapa, la trama actual se divide en una pluralidad de subtramas, y el retardo para el que el coeficiente de autocorrelación es máximo se identifica por separado para cada subtrama. Tal y como se ha mencionado anteriormente, la autocorrelación se realiza en un intervalo de búsqueda que incluye un retardo de un desfase de tono y que no incluye un retardo de cero muestras.

30 En una segunda etapa, una trama retardada se construye aplicando el retardo identificado correspondiente a cada subtrama, concatenando las subtramas resultantes para construir una trama retardada de manera óptima, y calculando el factor S180 de ponderación de armónico como el coeficiente de correlación entre la trama original y la trama retardada de manera óptima. En una alternativa adicional, el calculador 550 de factor de ponderación calcula el factor S180 de ponderación de armónico como la media de los coeficientes de autocorrelación máximos obtenidos en la primera etapa para cada subtrama. Implementaciones del calculador 550 de factor de ponderación también pueden calcularse para escalar el coeficiente de correlación, y/o para combinarlo con otro valor, para calcular el valor para el factor S180 de ponderación de armónico.

35 Puede ser deseable que el calculador 550 de factor de ponderación calcule una medida de periodicidad de la señal S30 de banda alta solamente cuando se indique de otro modo una presencia de periodicidad en la trama. Por ejemplo, el calculador 550 de factor de ponderación puede configurarse para calcular una medida de periodicidad de la señal S30 de banda alta según una relación entre otro indicador de periodicidad de la trama actual, tal como la ganancia de tono, y un valor de umbral. En un ejemplo, el calculador 550 de factor de ponderación está configurado para realizar una operación de autocorrelación en la señal S30 de banda alta solamente si la ganancia de tono de la trama (por ejemplo, la ganancia

de libro de códigos adaptativo de la señal residual de banda estrecha) tiene un valor superior a 0,5 (como alternativa, de al menos 0,5). En otro ejemplo, el calculador 550 de factor de ponderación está configurado para realizar una operación de autocorrelación en la señal S30 de banda alta solamente para tramas que presenten estados particulares de modo de voz (por ejemplo, solamente para señales sonoras). En estos casos, el calculador 550 de factor de ponderación puede configurarse para asignar un factor de ponderación por defecto para tramas que presenten otros estados de modo de voz y/o valores inferiores de ganancia de tono.

Las realizaciones incluyen implementaciones adicionales del calculador 550 de factor de ponderación que están configuradas para calcular factores de ponderación según características distintas de o adicionales a la periodicidad. Por ejemplo, una implementación de este tipo puede configurarse para asignar un valor más elevado al factor S190 de ganancia de ruido para señales de voz que presenten un gran desfase de tono en lugar de para señales de voz que presenten un pequeño desfase de tono. Otra implementación de este tipo del calculador 550 de factor de ponderación está configurada para determinar una medida de armonicidad de la señal S10 de voz de banda ancha, o de la señal S30 de banda alta, según una medida de la energía de la señal en múltiplos de la frecuencia fundamental con respecto a la energía de la señal en otras componentes de frecuencia.

Algunas implementaciones del codificador A100 de voz de banda ancha están configuradas para transmitir una indicación de periodicidad o armonicidad (por ejemplo, una bandera de un bit que indique si la trama es armónica o no armónica) en función de la ganancia de tono y/u otra medida de periodicidad o armonicidad descrita en este documento. En un ejemplo, un descodificador B100 de voz de banda ancha correspondiente utiliza esta indicación para configurar una operación tal como el cálculo de los factores de ponderación. En otro ejemplo, una indicación de este tipo se utiliza en el codificador y/o en el descodificador para calcular un valor para un parámetro de modo de voz.

Puede ser deseable que el generador A302 de excitación de banda alta genere la señal S120 de excitación de banda alta de manera que la energía de la señal de excitación no quede afectada en gran medida por los valores particulares de los factores S180 y S190 de ponderación. En este caso, el calculador 550 de factor de ponderación puede configurarse para calcular un valor para el factor S180 de ponderación de armónico o para el factor S190 de ponderación de ruido (o para recibir un valor de este tipo de un medio de almacenamiento u otro elemento del codificador A200 de banda alta) y para obtener un valor para el otro factor de ponderación según una expresión como la siguiente

$$(W_{\text{armónico}})^2 + (W_{\text{ruido}})^2 = 1, \quad (2)$$

donde  $W_{\text{armónico}}$  denota el factor S180 de ponderación de armónico y  $W_{\text{ruido}}$  denota el factor S190 de ponderación de ruido. Como alternativa, el calculador 550 de factor de ponderación puede configurarse para seleccionar, según un valor de una medida de periodicidad para la trama o subtrama actual, un par correspondiente de entre una pluralidad de pares de factores S180, S190 de ponderación, donde los pares se precálculan para satisfacer una relación de energía constante tal como la expresión (2). Para una implementación del calculador 550 de factor de ponderación en la que se observa la expresión (2), los valores típicos para el factor S180 de ponderación de armónico oscilan entre 0,7 aproximadamente y 1,0 aproximadamente, y los valores típicos para el factor S190 de ponderación de ruido oscilan entre 0,1 aproximadamente y 0,7 aproximadamente. Otras implementaciones del calculador 550 de factor de ponderación pueden configurarse para funcionar según una versión de la expresión (2) que está modificada según una ponderación de línea de base deseada entre la señal S160 ensanchada de manera armónica y la señal S170 de ruido modulada.

Pueden producirse artefactos en una señal de voz sintetizada cuando un libro de códigos disperso (uno cuyas entradas son en gran parte valores de cero) se haya utilizado para calcular la representación cuantificada de la señal residual. La dispersión de libro de códigos se produce especialmente cuando la señal de banda estrecha se codifica a una baja velocidad binaria. Los artefactos producidos por la dispersión de libros de códigos son normalmente cuasiperiódicos en el tiempo y se producen principalmente por encima de 3 kHz. Puesto que el oído humano tiene una mejor resolución de tiempo a frecuencias más altas, estos artefactos pueden apreciarse más en la banda alta.

Las realizaciones incluyen implementaciones del generador A300 de excitación de banda alta que están configuradas para realizar un filtrado de antidispersión. La figura 18 muestra un diagrama de bloques de una implementación A312 del generador A302 de excitación de banda alta que incluye un filtro 600 de antidispersión dispuesto para filtrar la señal de excitación de banda estrecha descuantificada generada por el cuantificador 450 inverso. La figura 19 muestra un diagrama de bloques de una implementación A314 del generador A302 de excitación de banda alta que incluye un filtro 600 de antidispersión dispuesto para filtrar la señal ensanchada de manera espectral generada por el ensanchador A400 de espectro. La figura 20 muestra un diagrama de bloques de una implementación A316 del generador A302 de excitación de banda alta que incluye un filtro 600 de antidispersión dispuesto para filtrar la salida del combinador 490 para generar la señal S120 de excitación de banda alta. Por supuesto, se contemplan implementaciones del generador A300 de excitación de banda alta que combinen las características de cualquiera de las implementaciones A304 y A306 con las características de cualquiera de las implementaciones A312, A314 y A316 y, por lo tanto, se indican de manera expresa. El filtro 600 de antidispersión también puede estar dispuesto dentro del ensanchador A400 de espectro: por ejemplo,

después de cualquiera de los elementos 510, 520, 530 y 540 del ensanchador A402 de espectro. Debe observarse expresamente que el filtro 600 de antidispersión también puede usarse con implementaciones del ensanchador A400 de espectro que realicen plegado espectral, traslación espectral o ensanchamiento de armónico.

5 El filtro 600 de antidispersión puede configurarse para alterar la fase de su señal de entrada. Por ejemplo, puede ser deseable que el filtro 600 de antidispersión esté configurado y dispuesto de manera que la fase de la señal S120 de excitación de banda alta se aleatorice, o se distribuya de otro modo de manera más uniforme, en el tiempo. También puede ser deseable que la respuesta del filtro 600 de antidispersión sea plana de manera espectral, de manera que el espectro de magnitud de la señal filtrada no cambie de manera apreciable. En un ejemplo, el filtro 600 de antidispersión está implementado como un filtro de todo paso que presenta una función de transferencia según la siguiente expresión:

10

$$H(z) = \frac{-0.7 + z^{-4}}{1 - 0.7z^{-4}} \cdot \frac{0.6 + z^{-6}}{1 + 0.6z^{-6}} \quad (3).$$

Un efecto de un filtro de este tipo puede ser distribuir la energía de la señal de entrada de manera que no se concentre solamente en algunas muestras.

15

Los artefactos producidos por la dispersión de libro de códigos son normalmente más apreciables para señales similares al ruido, donde la señal residual incluye menos información de tono, y también para voz en ruido de fondo. La dispersión provoca normalmente menos artefactos cuando la excitación presenta una estructura de larga duración y, de hecho, la modificación de fase puede provocar ruido en señales sonoras. Por lo tanto, puede ser deseable configurar el filtro 600 de antidispersión para filtrar señales sordas y para hacer pasar al menos las mismas señales sonoras sin alteración. Las señales sordas están caracterizadas por una baja ganancia de tono (por ejemplo, una ganancia de libro de códigos adaptativo de banda estrecha cuantificada) y por una inclinación espectral (por ejemplo, primer coeficiente de reflexión cuantificado) cercana a cero o positiva, indicando una envolvente espectral que es plana o que está inclinada hacia arriba sin una frecuencia creciente. Las implementaciones típicas del filtro 600 de antidispersión están configuradas para filtrar sonidos sordos (por ejemplo, como los indicados por el valor de la inclinación espectral), para filtrar sonidos sonoros cuando la ganancia de tono está por debajo de un valor de umbral (como alternativa, no superior al valor de umbral) y además para hacer pasar la señal sin alteraciones.

20

25

Implementaciones adicionales del filtro 600 de antidispersión incluyen dos o más filtros que están configurados para presentar diferentes ángulos de modificación de fase máxima (por ejemplo, de hasta 180 grados). En ese caso, el filtro 600 de antidispersión puede estar configurado para realizar una selección entre esos filtros de componente según un valor de la ganancia de tono (por ejemplo, el libro de códigos adaptativo cuantificado o la ganancia LTP), de manera que se utiliza un mayor ángulo de modificación de fase máxima para tramas que tengan valores de ganancia de tono más bajos. Una implementación del filtro 600 de antidispersión también puede incluir diferentes filtros de componente que estén configurados para modificar la fase más o menos en el espectro de frecuencia, de manera que un filtro configurado para modificar la fase en un intervalo de frecuencias más amplio de la señal de entrada se utiliza para tramas que tengan valores de ganancia de tono más bajos.

30

35

Para una reproducción precisa de la señal de voz codificada puede ser deseable que la relación entre los niveles de la parte de banda alta y la parte de banda estrecha de la señal S100 de voz de banda ancha sintetizada sea similar a la de la señal S10 de voz de banda ancha original. Además de una envolvente espectral como la representada por los parámetros S60a de codificación de banda alta, el codificador A200 de banda alta puede configurarse para caracterizar la señal S30 de banda alta especificando una envolvente temporal o de ganancia. Tal y como se muestra en la figura 10, el codificador A202 de banda alta incluye un calculador A230 de factor de ganancia de banda alta que está configurado y dispuesto para calcular uno o más factores de ganancia según una relación entre la señal S30 de banda alta y la señal S130 de banda alta sintetizada, tal como una diferencia o relación entre las energías de las dos señales en la trama o en alguna parte de la misma. En otras implementaciones del codificador A202 de banda alta, el calculador A230 de ganancia de banda alta puede estar configurado de la misma manera pero dispuesto en cambio para calcular la envolvente de ganancia según una relación variable en el tiempo de este tipo entre la señal S30 de banda alta y la señal S80 de excitación de banda estrecha o la señal S120 de excitación de banda alta.

40

45

Es probable que las envolventes temporales de la señal S80 de excitación de banda estrecha y de la señal S30 de banda alta sean similares. Por lo tanto, codificar una envolvente de ganancia que esté basada en una relación entre la señal S30 de banda alta y la señal S80 de excitación de banda estrecha (o una señal que se obtenga de las mismas, tal como la señal S120 de excitación de banda alta o la señal S130 de banda alta sintetizada) será generalmente más eficaz que codificar una envolvente de ganancia basada solamente en la señal S30 de banda alta. En una implementación típica, el codificador A202 de banda alta está configurado para transmitir un índice cuantificado de ocho a doce bits que especifica cinco factores de ganancia para cada trama.

50

El calculador A230 de factor de ganancia de banda alta puede configurarse para realizar el cálculo del factor de ganancia como una tarea que incluya una o más series de subtareas. La figura 21 muestra un diagrama de flujo de un ejemplo T200 de una tarea de este tipo que calcula un valor de ganancia para una subtrama correspondiente según las energías relativas de la señal S30 de banda alta y de la señal S130 de banda alta sintetizada. Las tareas 220a y 220b calculan las energías de las subtramas correspondientes de las señales respectivas. Por ejemplo, las tareas 220a y 220b pueden configurarse para calcular la energía como una suma de los cuadrados de las muestras de la subtrama respectiva. La tarea T230 calcula un factor de ganancia para la subtrama como la raíz cuadrada de la relación de esas energías. En este ejemplo, la tarea T230 calcula el factor de ganancia como la raíz cuadrada de la relación de la energía de la señal S30 de banda alta con respecto a la energía de la señal S130 de banda alta sintetizada en la subtrama.

Puede ser deseable que el calculador A230 de factor de ganancia de banda alta esté configurado para calcular las energías de subtrama según una función de división en ventanas. La figura 22 muestra un diagrama de flujo de una implementación T210 de este tipo de la tarea T200 de cálculo de factor de ganancia. La tarea T215a aplica una función de división en ventanas a la señal S30 de banda alta, y la tarea T215b aplica la misma función de división en ventanas a la señal S130 de banda alta sintetizada. Las implementaciones 222a y 222b de las tareas 220a y 220b calculan las energías de las ventanas respectivas, y la tarea T230 calcula un factor de ganancia para la subtrama como la raíz cuadrada de la relación de las energías.

Puede ser deseable aplicar una función de división en ventanas que solape subtramas adyacentes. Por ejemplo, una función de división en ventanas que genere factores de ganancia que puedan aplicarse mediante solapamiento y suma puede ayudar a reducir o evitar discontinuidades entre subtramas. En un ejemplo, el calculador A230 de factor de ganancia de banda alta está configurado para aplicar una función trapezoidal de división en ventanas, tal como la mostrada en la figura 23a, en la que la ventana se solapa con cada una de las dos subtramas adyacentes en un milisegundo. La figura 23b muestra una aplicación de esta función de división en ventanas a cada una de las cinco subtramas de una trama de 20 milisegundos. Otras implementaciones del calculador A230 de factor de ganancia de banda alta pueden configurarse para aplicar funciones de división en ventanas que presenten diferentes periodos de solapamiento y/o diferentes formas de ventana (por ejemplo, rectangulares, de Hamming) que pueden ser simétricas o asimétricas. También es posible que una implementación del calculador A230 de factor de ganancia de banda alta esté configurada para aplicar diferentes funciones de división en ventanas a diferentes subtramas de una trama y/o que una trama incluya subtramas de diferentes longitudes.

De manera no limitativa, los siguientes valores se presentan como ejemplos de implementaciones particulares. Se considera una trama de 20 ms para estos casos, aunque puede utilizarse cualquier otra duración. Para una señal de banda alta muestreada a 7 kHz, cada trama tiene 140 muestras. Si una trama de este tipo se divide en cinco subtramas de igual longitud, cada subtrama tendrá 28 muestras, y la ventana mostrada en la figura 23a tendrá un ancho de 42 muestras. Para una señal de banda alta muestreada a 8 kHz, cada trama presenta 160 muestras. Si una trama de este tipo se divide en cinco subtramas de igual longitud, cada subtrama tendrá 32 muestras, y la ventana mostrada en la figura 23a tendrá un ancho de 48 muestras. En otras implementaciones pueden usarse subtramas de cualquier ancho e incluso es posible que una implementación del calculador A230 de ganancia de banda alta esté configurada para generar un factor de ganancia diferente para cada muestra de una trama.

La figura 24 muestra un diagrama de bloques de una implementación B202 del descodificador B200 de banda alta. El descodificador B202 de banda alta incluye un generador B300 de excitación de banda alta que está configurado para generar la señal S120 de excitación de banda alta en función de la señal S80 de excitación de banda estrecha. Dependiendo de las elecciones particulares de diseño del sistema, el generador B300 de excitación de banda alta puede implementarse según cualquiera de las implementaciones del generador A300 de excitación de banda alta descrito en este documento. Normalmente es deseable implementar el generador B300 de excitación de banda alta para que tenga la misma respuesta que el generador de excitación de banda alta del codificador de banda alta del sistema de codificación particular. Puesto que el descodificador B110 de banda estrecha realizará normalmente la descuantificación de la señal S50 de excitación de banda estrecha codificada, en la mayoría de los casos el generador B300 de excitación de banda alta puede implementarse para recibir la señal S80 de excitación de banda estrecha del descodificador B110 de banda estrecha y no necesita incluir un cuantificador inverso configurado para descuantificar la señal S50 codificada de excitación de banda estrecha. También es posible implementar el descodificador B110 de banda estrecha para que incluya una instancia del filtro 600 de antidispersión dispuesta para filtrar la señal de excitación de banda estrecha descuantificada antes de que se introduzca en un filtro de síntesis de banda estrecha tal como el filtro 330.

Un cuantificador 560 inverso está configurado para descuantificar los parámetros S60a de filtro de banda alta (en este ejemplo, a un conjunto de LSF), y una transformada 570 de LSF a coeficientes de filtro LP está configurada para transformar las LSF en un conjunto de coeficientes de filtro (por ejemplo, como el descrito anteriormente con referencia al cuantificador 240 inverso y a la transformada 250 del codificador A122 de banda estrecha). En otras implementaciones, como las mencionadas anteriormente, pueden utilizarse diferentes conjuntos de coeficientes (por ejemplo, coeficientes cepstrales) y/o diferentes representaciones de coeficientes (por ejemplo, ISP). Un filtro B200 de síntesis de banda alta está configurado para generar una señal de banda alta sintetizada según la señal S120 de excitación de banda alta y el

conjunto de coeficientes de filtro. Para un sistema en el que el codificador de banda alta incluye un filtro de síntesis (por ejemplo, como en el ejemplo del codificador A202 descrito anteriormente), puede ser deseable implementar el filtro B200 de síntesis de banda alta para que tenga la misma respuesta (por ejemplo, la misma función de transferencia) que el filtro de síntesis.

5 El descodificador B202 de banda alta incluye un cuantificador 580 inverso configurado para descuantificar los factores S60b de ganancia de banda alta y un elemento 590 de control de ganancia (por ejemplo, un multiplicador o amplificador) configurado y dispuesto para aplicar los factores de ganancia descuantificados a la señal de banda alta sintetizada para generar la señal S100 de banda alta. Para un caso en el que la envolvente de ganancia de una trama está especificada por más de un factor de ganancia, el elemento 590 de control de ganancia puede incluir lógica configurada para aplicar los factores de ganancia a las subtramas respectivas, posiblemente según una función de división en ventanas que puede ser la misma o una función de división en ventanas diferente como la aplicada por un calculador de ganancia (por ejemplo, el calculador A230 de ganancia de banda alta) del codificador de banda alta correspondiente. En otras implementaciones del descodificador B202 de banda alta, el elemento 590 de control de ganancia está configurado de manera similar pero está dispuesto en cambio para aplicar los factores de ganancia descuantificados a la señal S80 de excitación de banda estrecha o a la señal S120 de excitación de banda alta.

10 Tal y como se ha mencionado anteriormente, puede ser deseable obtener el mismo estado en el codificador de banda alta y en el descodificador de banda alta (por ejemplo, utilizando valores descuantificados durante la codificación). Por lo tanto, en un sistema de codificación según una implementación de este tipo puede ser deseable garantizar el mismo estado para generadores de ruido correspondientes en los generadores A300 y B300 de excitación de banda alta. Por ejemplo, los generadores A300 y B300 de excitación de banda alta de una implementación de este tipo pueden configurarse de manera que el estado del generador de ruido sea una función determinista de información ya codificada dentro de la misma trama (por ejemplo, los parámetros S40 de filtro de banda estrecha o una parte de los mismos y/o la señal S50 de excitación de banda estrecha codificada o una parte de la misma).

15 Uno o más de los cuantificadores de los elementos descritos en este documento (por ejemplo, los cuantificadores 230, 420 ó 430) pueden configurarse para realizar una cuantificación vectorial clasificada. Por ejemplo, un cuantificador de este tipo puede configurarse para seleccionar un libro de códigos de un conjunto de libros de códigos en función de información que ya se haya codificado dentro de la misma trama en el canal de banda estrecha y/o en el canal de banda alta. Una técnica de este tipo proporciona normalmente una mayor eficacia de codificación a expensas de un almacenamiento adicional de libros de códigos.

20 Tal y como se ha mencionado anteriormente con referencia a, por ejemplo, las figuras 8 y 9, una considerable cantidad de estructura periódica puede permanecer en la señal residual después de la eliminación de la envolvente espectral aproximada de la señal S20 de voz de banda estrecha. Por ejemplo, la señal residual puede contener una secuencia de valores de pico o de impulsos aproximadamente periódicos en el tiempo. Es especialmente probable que tal estructura, que está relacionada normalmente con el tono, se produzca en señales de voz sonoras. El cálculo de una representación cuantificada de la señal residual de banda estrecha puede incluir la codificación de esta estructura de tono según un modelo de periodicidad de larga duración como el representado, por ejemplo, por uno o más libros de códigos.

25 La estructura de tono de una señal residual real puede no coincidir exactamente con el modelo de periodicidad. Por ejemplo, la señal residual puede incluir pequeñas fluctuaciones en la regularidad de las posiciones de los impulsos de tono, de manera que las distancias entre impulsos de tono sucesivos en una trama no son exactamente iguales y la estructura es bastante irregular. Estas irregularidades tienden a reducir la eficacia de codificación.

30 Algunas implementaciones del codificador A120 de banda estrecha están configuradas para realizar una regularización de la estructura de tono aplicando una distorsión de tiempo adaptativa a la señal residual antes o durante la cuantificación o incluyendo de otro modo una distorsión de tiempo adaptativa en la señal codificada de excitación. Por ejemplo, un codificador de este tipo puede configurarse para seleccionar o calcular de otro modo un grado de distorsión en el tiempo (por ejemplo, según uno o más criterios de minimización de errores y/o de ponderación perceptual) de manera que la señal de excitación resultante se ajuste de manera óptima al modelo de periodicidad de larga duración. La regularización de la estructura de tono se realiza por un subconjunto de codificadores CELP denominados codificadores de predicción lineal excitada por código de relajación (RCELP).

35 Un codificador RCELP está configurado normalmente para realizar la distorsión de tiempo como un desplazamiento de tiempo adaptativo. Este desplazamiento de tiempo puede ser un retardo que oscile entre algunos milisegundos negativos y algunos milisegundos positivos, y normalmente se modifica de manera uniforme para evitar discontinuidades audibles. En algunas implementaciones, un codificador de este tipo está configurado para aplicar la regularización por tramos, donde cada trama o subtrama está distorsionada mediante un desplazamiento de tiempo fijo correspondiente. En otras implementaciones, el codificador está configurado para aplicar la regularización como una función de distorsión continua, de manera que una trama o subtrama se distorsiona según un contorno de tono (denominado también como trayectoria de tono). En algunos casos (por ejemplo, según se describe en la publicación de solicitud de patente estadounidense

2004/0098255), el codificador está configurado para incluir una distorsión de tiempo en la señal codificada de excitación aplicando el desplazamiento a una señal de entrada ponderada de manera perceptual que se utiliza para calcular la señal de excitación codificada.

5 El codificador calcula una señal codificada de excitación que está regularizada y cuantificada, y el decodificador descuantifica la señal codificada de excitación para obtener una señal de excitación que se utiliza para sintetizar la señal de voz descodificada. Por lo tanto, la señal de salida descodificada presenta el mismo retardo variable que el que estaba incluido en la señal codificada de excitación debido a la regularización. Normalmente, no se transmite al decodificador ninguna información que especifique las cantidades de regularización.

10 La regularización tiende a hacer la señal residual más fácil de codificar, lo que mejora la ganancia de codificación del elemento de predicción a largo plazo y, por lo tanto, aumenta la eficacia de codificación global, generalmente sin generar artefactos. Puede ser deseable realizar la regularización solamente en tramas que sean sonoras. Por ejemplo, el codificador A124 de banda estrecha puede configurarse para desplazar solamente aquellas tramas o subtramas que tengan una estructura de larga duración, tal como las señales sonoras. También puede ser deseable realizar la regularización solamente en subtramas que incluyan energía de impulso de tono. Varias implementaciones de la codificación RCELP se describen en las patentes estadounidenses números 5.704.003 (Kleijn et al.) y 6.879.955 (Rao) y en la publicación de solicitud de patente estadounidense 2004/0098255 (Kovesi et al.). Implementaciones existentes de codificadores RCELP incluyen el códec de velocidad variable mejorado (EVRC), como el descrito en la especificación IS-127 de la Asociación de la Industria de Telecomunicaciones (TIA), y el vocodificador de modo seleccionable (SMV) del segundo proyecto de colaboración de tercera generación (3GPP2).

20 Desafortunadamente, la regularización puede provocar problemas en un codificador de voz de banda ancha en el que la excitación de banda alta se obtenga de la señal codificada de excitación de banda estrecha (tal como un sistema que incluye el codificador A100 de voz de banda ancha y el decodificador B100 de voz de banda ancha). Debido a esta obtención a partir de una señal distorsionada en el tiempo, la señal de excitación de banda alta presentará generalmente un perfil de tiempo que será diferente del de la señal de voz de banda alta original. Dicho de otro modo, la señal de excitación de banda alta ya no será sincrónica con la señal de voz de banda alta original.

25 Una desalineación en el tiempo entre la señal de excitación de banda alta distorsionada y la señal de voz de banda alta original puede provocar varios problemas. Por ejemplo, la señal de excitación de banda alta distorsionada ya no puede proporcionar una excitación fuente adecuada para un filtro de síntesis que esté configurado según los parámetros de filtro extraídos de la señal de voz de banda alta original. Como resultado, la señal de banda alta sintetizada puede contener artefactos audibles que reduzcan la calidad percibida de la señal de voz de banda ancha descodificada.

30 Esta desalineación en el tiempo también puede provocar ineficiencias en la codificación de envolvente de ganancia. Tal como se ha mencionado anteriormente, es probable que exista una correlación entre las envolventes temporales de la señal S80 de excitación de banda estrecha y de la señal S30 de banda alta. Codificando la envolvente de ganancia de la señal de banda alta según una relación entre estas dos envolventes temporales puede obtenerse un aumento de la eficacia de codificación en comparación con la codificación directa de la envolvente de ganancia. Sin embargo, cuando la señal codificada de excitación de banda estrecha está regularizada, esta correlación puede debilitarse. La desalineación en el tiempo entre la señal S80 de excitación de banda estrecha y la señal S30 de banda alta puede provocar que aparezcan fluctuaciones en los factores S60b de ganancia de banda alta, y la eficacia de codificación puede disminuir.

40 Las realizaciones incluyen procedimientos de codificación de voz de banda ancha que llevan a cabo la distorsión de tiempo de una señal de voz de banda alta según una distorsión de tiempo incluida en una señal codificada de excitación de banda estrecha correspondiente. Las ventajas potenciales de tales procedimientos incluyen mejorar la calidad de una señal de voz de banda ancha descodificada y/o mejorar la eficacia de codificación de una envolvente de ganancia de banda alta.

45 La figura 25 muestra un diagrama de bloques de una implementación AD10 del codificador A100 de voz de banda ancha. El codificador AD10 incluye una implementación A124 del codificador A120 de banda estrecha que está configurada para realizar la regularización durante el cálculo de la señal S50 codificada de excitación de banda estrecha. Por ejemplo, el codificador A124 de banda estrecha puede configurarse según una o más de las implementaciones RCELP descritas anteriormente.

50 El codificador A124 de banda estrecha también está configurado para transmitir una señal SD10 de datos de regularización que especifica el grado de distorsión de tiempo aplicado. Para varios casos en los que el codificador A124 de banda estrecha esté configurado para aplicar un desplazamiento de tiempo fijo a cada trama o subtrama, la señal SD10 de datos de regularización puede incluir una serie de valores que indiquen cada cantidad de desplazamiento de tiempo como un valor entero o no entero en lo que se refiere a muestras, milisegundos o algún otro incremento de tiempo. Para un caso en el que el codificador A124 de banda estrecha esté configurado para modificar de otro modo la escala de tiempo de una trama u otra secuencia de muestras (por ejemplo, comprimiendo una parte y ampliando otra parte), la señal SD10 de información de regularización puede incluir una descripción correspondiente de la modificación, tal como un

conjunto de parámetros de función. En un ejemplo particular, el codificador A124 de banda estrecha está configurado para dividir una trama en tres subtramas y para calcular un desplazamiento de tiempo fijo para cada subtrama, de manera que la señal SD10 de datos de regularización indica tres cantidades de desplazamiento de tiempo para cada trama regularizada de la señal codificada de banda estrecha.

5 El codificador AD10 de voz de banda ancha incluye una línea D120 de retardo configurada para adelantar o retrasar partes de la señal S30 de voz de banda alta, según cantidades de retardo indicadas por una señal de entrada, para generar una señal S30a de voz de banda alta distorsionada en el tiempo. En el ejemplo mostrado en la figura 25, la línea D120 de retardo está configurada para distorsionar en el tiempo la señal S30 de voz de banda alta según la distorsión indicada por la señal SD10 de datos de regularización. De esta manera, la misma cantidad de distorsión de tiempo que estaba incluida en la señal S50 codificada de excitación de banda estrecha también se aplica a la parte correspondiente de la señal S30 de voz de banda alta antes del análisis. Aunque este ejemplo muestra la línea D120 de retardo como un elemento separado del codificador A200 de banda alta, en otras implementaciones la línea D120 de retardo está dispuesta como parte del codificador de banda alta.

10 Implementaciones adicionales del codificador A200 de banda alta pueden configurarse para realizar un análisis espectral (por ejemplo, un análisis LPC) de la señal S30 de voz de banda alta no distorsionada y para distorsionar en el tiempo la señal S30 de voz de banda alta antes del cálculo de los parámetros S60b de ganancia de banda alta. Un codificador de este tipo puede incluir, por ejemplo, una implementación de la línea D120 de retardo dispuesta para realizar la distorsión en el tiempo. Sin embargo, en tales casos, los parámetros S60a de filtro de banda alta basados en el análisis de la señal S30 no distorsionada pueden describir una envolvente espectral que esté desalineada en el tiempo con la señal S120 de excitación de banda alta.

15 La línea D120 de retardo puede configurarse según cualquier combinación de elementos lógicos y de elementos de almacenamiento adecuados para aplicar las operaciones de distorsión de tiempo deseadas a la señal S30 de voz de banda alta. Por ejemplo, la línea D120 de retardo puede configurarse para leer la señal S30 de voz de banda alta de un almacenamiento intermedio según los desplazamientos de tiempo deseados. La figura 26a muestra un diagrama esquemático de una implementación D122 de este tipo de la línea D120 de retardo que incluye un registro SR1 de desplazamiento. El registro SR1 de desplazamiento es un almacenamiento intermedio de longitud  $m$  que está configurado para recibir y almacenar las  $m$  muestras más recientes de la señal S30 de voz de banda alta. El valor  $m$  es igual a al menos la suma del desplazamiento de tiempo positivo máximo (o "adelanto") y del desplazamiento de tiempo negativo máximo (o "retraso") que van a soportarse. Puede ser conveniente que el valor  $m$  sea igual a la longitud de una trama o subtrama de la señal S30 de banda alta.

20 La línea D122 de retardo está configurada para transmitir la señal S30a de banda alta distorsionada en el tiempo desde una ubicación de desfase OL del registro SR1 de desplazamiento. La posición de la ubicación de desfase OL varía en torno a una posición de referencia (desplazamiento de tiempo cero) según el desplazamiento de tiempo actual indicado, por ejemplo, por la señal SD10 de datos de regularización. La línea D122 de retardo puede configurarse para soportar límites de adelanto o de retraso iguales o, como alternativa, un límite más grande que otro de manera que pueda llevarse a cabo un mayor desplazamiento en una dirección que en la otra. La figura 26a muestra un ejemplo particular que soporta un mayor desplazamiento de tiempo positivo que negativo. La línea D122 de retardo puede configurarse para transmitir una o más muestras a la vez (dependiendo del ancho de un bus de salida, por ejemplo).

25 Un desplazamiento de tiempo de regularización que presente una magnitud de más de algunos milisegundos puede provocar artefactos audibles en la señal descodificada. Normalmente, la magnitud de un desplazamiento de tiempo de regularización como el generado por un codificador A124 de banda estrecha no será mayor que algunos milisegundos, de manera que los desplazamientos de tiempo indicados por la señal SD10 de datos de regularización serán limitados. Sin embargo, en tales casos puede desearse que la línea D122 de retardo esté configurada para imponer un límite máximo en los desplazamientos de tiempo en la dirección positiva y/o en la dirección negativa (por ejemplo, para observar un límite más preciso que el impuesto por el codificador de banda estrecha).

30 La figura 26b muestra un diagrama esquemático de una implementación D124 de la línea D122 de retardo que incluye una ventana de desplazamiento SW. En este ejemplo, la posición de la ubicación de desfase OL está limitada por la ventana de desplazamiento SW. Aunque la figura 26b muestra un caso en que la longitud  $m$  del almacenamiento intermedio es mayor que el ancho de la ventana de desplazamiento SW, la línea D124 de retardo también puede implementarse de manera que el ancho de la ventana de desplazamiento SW sea igual a  $m$ .

35 En otras implementaciones, la línea D120 de retardo está configurada para escribir la señal S30 de voz de banda alta en un almacenamiento intermedio según los desplazamientos de tiempo deseados. La figura 27 muestra un diagrama esquemático de una implementación D130 de este tipo de la línea D120 de retardo que incluye dos registros SR2 y SR3 de desplazamiento configurados para recibir y almacenar la señal S30 de voz de banda alta. La línea D130 de retardo está configurada para escribir una trama o subtrama del registro SR2 de desplazamiento en el registro SR3 de desplazamiento según un desplazamiento de tiempo como el indicado, por ejemplo, por la señal SD10 de datos de

regularización. El registro SR3 de desplazamiento está configurado como un almacenamiento intermedio FIFO dispuesto para transmitir la señal S30 de banda alta distorsionada en el tiempo.

En el ejemplo particular mostrado en la figura 27, el registro SR2 de desplazamiento incluye una parte FB1 de almacenamiento intermedio de trama y una parte DB de almacenamiento intermedio de retardo, y el registro SR3 de desplazamiento incluye una parte FB2 de almacenamiento intermedio de trama, una parte AB de almacenamiento intermedio de adelanto y una parte RB de almacenamiento intermedio de retraso. Las longitudes del almacenamiento AB intermedio de adelanto y del almacenamiento RB intermedio de retraso pueden ser iguales, o una puede ser más grande que la otra, de manera que se soporta un mayor desplazamiento en una dirección que en la otra. El almacenamiento DB intermedio de retardo y la parte RB del almacenamiento intermedio de retraso pueden configurarse para tener la misma longitud. Como alternativa, el almacenamiento DB intermedio de retardo puede ser más corto que el almacenamiento RB intermedio de retraso para soportar un intervalo de tiempo requerido para transferir muestras desde el almacenamiento FB1 intermedio de trama al registro SR3 de desplazamiento, lo que puede incluir otras operaciones de procesamiento tales como la distorsión de las muestras antes de su almacenamiento en el registro SR3 de desplazamiento.

En el ejemplo de la figura 27, el almacenamiento FB1 intermedio de trama está configurado para tener la misma longitud que una trama de la señal S30 de banda alta. En otro ejemplo, el almacenamiento FB1 intermedio de trama está configurado para tener la misma longitud que una subtrama de la señal S30 de banda alta. En ese caso, la línea D130 de retardo puede configurarse para incluir lógica para aplicar el mismo retardo (por ejemplo, un promedio) a todas las subtramas de una trama que va a desplazarse. La línea D130 de retardo también puede incluir lógica para calcular la media de valores del almacenamiento FB1 intermedio de trama con valores que van a sobrescribirse en el almacenamiento RB intermedio de retraso o en el almacenamiento AB intermedio de adelanto. En un ejemplo adicional, el registro SR3 de desplazamiento puede configurarse para que reciba valores de la señal S30 de banda alta solamente a través del almacenamiento FB1 intermedio de trama y, en ese caso, la línea D130 de retardo puede incluir lógica para realizar una interpolación a través de los espacios entre tramas o subtramas sucesivas escritas en el registro SR3 de desplazamiento. En otras implementaciones, la línea D130 de retardo puede configurarse para realizar una operación de distorsión en muestras del almacenamiento FB1 intermedio de trama antes de que se escriban en el registro SR3 de desplazamiento (por ejemplo, según una función descrita por la señal SD10 de datos de regularización).

Puede ser deseable que la línea D120 de retardo aplique una distorsión de tiempo basada en, pero que no sea idéntica a, la distorsión especificada por la señal SD10 de datos de regularización. La figura 28 muestra un diagrama de bloques de una implementación AD12 del codificador AD10 de voz de banda ancha que incluye un elemento D110 de correlación de valores de retardo. El elemento D110 de correlación de valores de retardo está configurado para transformar la distorsión indicada por la señal SD10 de datos de regularización en valores SD10a de retardo correlacionados. La línea D120 de retardo está dispuesta para generar la señal S30a de voz de banda alta distorsionada en el tiempo según la distorsión indicada por los valores SD10a de retardo correlacionados.

Puede esperarse que el desplazamiento de tiempo aplicado por el codificador de banda estrecha se desarrolle de manera uniforme en el tiempo. Por lo tanto, normalmente es suficiente calcular el desplazamiento de tiempo medio de banda estrecha aplicado a las subtramas durante una trama de voz y desplazar una trama correspondiente de la señal S30 de voz de banda alta según este promedio. En un ejemplo de este tipo, el elemento D110 de correlación de valores de retardo está configurado para calcular un promedio de los valores de retardo de subtrama para cada trama, y la línea D120 de retardo está configurada para aplicar el promedio calculado a una trama correspondiente de la señal S30 de banda alta. En otros ejemplos, puede calcularse y aplicarse un promedio en un periodo más corto (tal como dos subtramas o media trama) o en un periodo más largo (tal como dos tramas). En un caso en el que el promedio no sea un valor entero de muestras, el elemento D110 de correlación de valores de retardo puede configurarse para redondear el valor a un número entero de muestras antes de transmitirse a la línea D120 de retardo.

El codificador A124 de banda estrecha puede configurarse para incluir un desplazamiento de tiempo de regularización de un número no entero de muestras en la señal codificada de excitación de banda estrecha. En tal caso, puede ser deseable que el elemento D110 de correlación de valores de retardo esté configurado para redondear el desplazamiento de tiempo de banda estrecha a un número entero de muestras y que la línea D120 de retardo aplique el desplazamiento de tiempo redondeado a la señal S30 de voz de banda alta.

En algunas implementaciones del codificador AD10 de voz de banda ancha, las tasas de muestreo de la señal S20 de voz de banda estrecha y de la señal S30 de voz de banda alta pueden ser diferentes. En tales casos, el elemento D110 de correlación de valores de retardo puede configurarse para ajustar cantidades de desplazamiento de tiempo indicadas en la señal SD10 de datos de regularización para soportar una diferencia entre las tasas de muestreo de la señal S20 de voz de banda estrecha (o la señal S80 de excitación de banda estrecha) y de la señal S30 de voz de banda alta. Por ejemplo, el elemento D110 de correlación de valores de retardo puede configurarse para escalar las cantidades de desplazamiento de tiempo según una relación de las tasas de muestreo. En un ejemplo particular como el mencionado anteriormente, la señal S20 de voz de banda estrecha se muestrea a 8 kHz y la señal S30 de voz de banda alta se muestrea a 7 kHz. En este caso, el elemento D110 de correlación de valores de retardo está configurado para multiplicar cada cantidad de

desplazamiento por 7/8. Implementaciones del elemento D110 de correlación de valores de retardo también pueden configurarse para realizar una operación de escalado de este tipo junto con una operación de redondeo a enteros y/o una operación de cálculo de la media de desplazamientos de tiempo como las descritas en este documento.

5 En implementaciones adicionales, la línea D120 de retardo está configurada para modificar de otro modo la escala de tiempo de una trama u otra secuencia de muestras (por ejemplo, comprimiendo una parte y ampliando otra parte). Por ejemplo, el codificador A124 de banda estrecha puede configurarse para realizar la regularización según una función tal como una trayectoria o contorno de tono. En ese caso, la señal SD10 de datos de regularización puede incluir una descripción correspondiente de la función, tal como un conjunto de parámetros, y la línea D120 de retardo puede incluir lógica configurada para distorsionar tramas o subtramas de la señal S30 de voz de banda alta según la función. En otras implementaciones, el elemento D110 de correlación de valores de retardo está configurado para calcular la media, escalar, y/o redondear la función antes de aplicarse a la señal S30 de voz de banda alta mediante la línea D120 de retardo. Por ejemplo, el elemento D110 de correlación de valores de retardo puede configurarse para calcular uno o más valores de retardo según la función, indicando cada valor de retardo una pluralidad de muestras, que se aplican posteriormente mediante la línea D120 de retardo para distorsionar en el tiempo una o más tramas o subtramas correspondientes de la señal S30 de voz de banda alta.

10 La figura 29 muestra un diagrama de flujo de un procedimiento MD100 para distorsionar en el tiempo una señal de voz de banda alta según una distorsión de tiempo incluida en una señal codificada de excitación de banda estrecha correspondiente. La tarea TD100 procesa una señal de voz de banda ancha para obtener una señal de voz de banda estrecha y una señal de voz de banda alta. Por ejemplo, la tarea TD100 puede configurarse para filtrar la señal de voz de banda ancha utilizando un banco de filtros que contenga filtros paso bajo y filtros paso alto, tal como una implementación del banco A110 de filtros. La tarea TD200 codifica la señal de voz de banda estrecha en al menos una señal codificada de excitación de banda estrecha y una pluralidad de parámetros de filtro de banda estrecha. La señal codificada de excitación de banda estrecha y/o los parámetros de filtro pueden cuantificarse, y la señal codificada de voz de banda estrecha también puede incluir otros parámetros tales como un parámetro de modo de voz. La tarea TD200 también incluye una distorsión de tiempo en la señal codificada de excitación de banda estrecha.

20 La tarea TD300 genera una señal de excitación de banda alta en función de una señal de excitación de banda estrecha. En este caso, la señal de excitación de banda estrecha está basada en la señal codificada de excitación de banda estrecha. Según al menos la señal de excitación de banda alta, la tarea TD400 codifica la señal de voz de banda alta en al menos una pluralidad de parámetros de filtro de banda alta. Por ejemplo, la tarea TD400 puede configurarse para codificar la señal de voz de banda alta en una pluralidad de LSF cuantificadas. La tarea TD500 aplica un desplazamiento de tiempo a la señal de voz de banda alta basada en información relacionada con una distorsión de tiempo incluida en la señal codificada de excitación de banda estrecha.

25 La tarea TD400 puede configurarse para realizar un análisis espectral (tal como un análisis LPC) en la señal de voz de banda alta y/o para calcular una envolvente de ganancia de la señal de voz de banda alta. En tales casos, la tarea TD500 puede configurarse para aplicar el desplazamiento de tiempo a la señal de voz de banda alta antes del análisis y/o del cálculo de envolvente de ganancia.

30 Otras implementaciones del codificador A100 de voz de banda ancha están configuradas para invertir una distorsión de tiempo de la señal S120 de excitación de banda alta provocada por una distorsión de tiempo incluida en la señal codificada de excitación de banda estrecha. Por ejemplo, el generador A300 de excitación de banda alta puede implementarse para incluir una implementación de la línea D120 de retardo que esté configurada para recibir la señal SD10 de datos de regularización o los valores SD10a de retardo correlacionados, y para aplicar un desplazamiento de tiempo inverso correspondiente a la señal S80 de excitación de banda estrecha y/o a una señal posterior basada en la misma tal como la señal S160 ensanchada de manera armónica o la señal S120 de excitación de banda alta.

35 Implementaciones adicionales del codificador de voz de banda ancha pueden configurarse para codificar la señal S20 de voz de banda estrecha y la señal S30 de voz de banda alta de manera independiente entre sí, de modo que la señal S30 de voz de banda alta se codifica como una representación de una envolvente espectral de banda alta y de una señal de excitación de banda alta. Una implementación de este tipo puede configurarse para realizar una distorsión en el tiempo de la señal residual de banda alta o para incluir de otro modo una distorsión de tiempo en una señal codificada de excitación de banda alta, según información relacionada con una distorsión de tiempo incluida en la señal codificada de excitación de banda estrecha. Por ejemplo, el codificador de banda alta puede incluir una implementación de la línea D120 de retardo y/o una implementación del elemento D110 de correlación de valores de retardo como las descritas en este documento que estén configuradas para aplicar una distorsión de tiempo a la señal residual de banda alta. Ventajas potenciales de una operación de este tipo incluyen una codificación más eficaz de la señal residual de banda alta y una mejor correspondencia entre las señales de voz de banda alta y de banda estrecha sintetizadas.

40 Tal y como se ha mencionado anteriormente, las realizaciones descritas en este documento incluyen implementaciones que pueden utilizarse para realizar codificación embebida, permitiendo la compatibilidad con sistemas de banda estrecha y

5 evitando la necesidad de transcodificación. Los medios que soportan la codificación de banda alta también pueden diferenciarse, en lo que a los costes se refiere, entre chips, conjuntos de chips, dispositivos y/o redes que soportan banda ancha con compatibilidad retroactiva y aquellos que soportan solamente banda estrecha. Los medios que soportan la codificación de banda alta como la descrita en este documento también pueden utilizarse junto con una técnica para soportar codificación de banda baja, y un sistema, procedimiento o aparato según una realización de este tipo pueden soportar la codificación de componentes de frecuencia desde, por ejemplo, 50 ó 100 Hz aproximadamente hasta 7 u 8 kHz aproximadamente.

10 Tal y como se ha mencionado anteriormente, añadir medios de soporte de banda alta en un codificador de voz puede mejorar la inteligibilidad, especialmente en lo que respecta a la diferenciación de sonidos fricativos. Aunque tal diferenciación puede obtenerse normalmente escuchando a una persona en un contexto particular, los medios que soportan banda alta pueden servir como una característica de habilitación en el reconocimiento de voz y en otras aplicaciones de interpretación mediante máquinas, tales como sistemas para la navegación por menús mediante voz automatizada y/o el procesamiento automático de llamadas.

15 Un aparato según una realización puede incorporarse en un dispositivo portátil de comunicaciones inalámbricas tal como un teléfono celular o un asistente personal digital (PDA). Como alternativa, un aparato de este tipo puede incluirse en otro dispositivo de comunicaciones tal como un microteléfono de VoIP, un ordenador personal configurado para soportar comunicaciones de VoIP, o un dispositivo de red configurado para encaminar comunicaciones telefónicas o de VoIP. Por ejemplo, un aparato según una realización puede implementarse en un chip o en un conjunto de chips de un dispositivo de comunicaciones. Dependiendo de la aplicación particular, un dispositivo de este tipo también puede incluir características tales como conversión de analógico a digital y/o de digital a analógico de una señal de voz, un sistema de circuitos para realizar amplificación y/u otras operaciones de procesamiento de señales en una señal de voz, y/o un sistema de circuitos de radiofrecuencia para la transmisión y/o recepción de la señal de voz codificada.

25 La presentación anterior de las realizaciones descritas se proporciona para que un experto en la técnica fabrique o utilice la presente invención. Son posibles varias modificaciones de estas realizaciones, y los principios genéricos presentados en este documento también pueden aplicarse a otras realizaciones. Por ejemplo, una realización puede implementarse en parte o en su totalidad como un circuito cableado, como una configuración de circuitos fabricada en un circuito integrado de aplicación específica o como un programa de firmware cargado en un medio de almacenamiento no volátil o un programa de software cargado desde o en un medio de almacenamiento de datos como un código legible por máquina, siendo tal código instrucciones que pueden ejecutarse por una disposición de elementos lógicos tal como un microprocesador u otra unidad de procesamiento de señales digitales. El medio de almacenamiento de datos puede ser una disposición de elementos de almacenamiento tal como una memoria de semiconductor (que puede incluir de manera no limitativa una RAM (memoria de acceso aleatorio) dinámica o estática, una ROM (memoria de solo lectura), y/o una RAM flash), o una memoria ferroeléctrica, magnetorresistiva, ovónica, polimérica o de cambio de fase; o un medio de disco tal como un disco magnético u óptico. Debe entenderse que el término "software" incluye código fuente, código de lenguaje ensamblador, código máquina, código binario, firmware, macrocódigo, microcódigo, uno o más conjuntos o secuencias de instrucciones ejecutables por una disposición de elementos lógicos, y cualquier combinación de tales ejemplos.

40 Los diversos elementos de implementaciones de los generadores A300 y B300 de excitación de banda alta, del codificador A100 de banda alta, del descodificador B200 de banda alta, del codificador A100 de voz de banda ancha y del descodificador B100 de voz de banda ancha pueden implementarse como dispositivos electrónicos y/u ópticos que residen, por ejemplo, en el mismo chip o entre dos o más chips de un conjunto de chips, aunque también se contemplan otras disposiciones sin tal limitación. Uno o más elementos de un aparato de este tipo pueden implementarse en su totalidad o en parte como uno o más conjuntos de instrucciones dispuestos para ejecutarse en una o más disposiciones fijas o programables de elementos lógicos (por ejemplo, transistores, puertas) tales como microprocesadores, procesadores incorporados, núcleos IP, procesadores de señales digitales, FPGA (matrices de puertas programables de campo), ASSP (productos estándar de aplicación específica) y ASIC (circuitos integrados de aplicación específica). También es posible que uno o más de tales elementos tengan una estructura en común (por ejemplo, un procesador utilizado para ejecutar partes de código correspondientes a diferentes elementos en momentos diferentes, un conjunto de instrucciones ejecutadas para llevar a cabo tareas correspondientes a diferentes elementos en momentos diferentes, o una disposición de dispositivos electrónicos y/u ópticos que realizan operaciones para diferentes elementos en momentos diferentes). Además, es posible que uno o más de estos elementos se utilicen para realizar tareas o para ejecutar otros conjuntos de instrucciones que no estén directamente relacionados con una operación del aparato, tal como una tarea relacionada con otra operación de un dispositivo o sistema en el que el aparato esté incorporado.

55 La figura 30 muestra un diagrama de flujo de un procedimiento M100, según una realización, para codificar una parte de banda alta de una señal de voz que presenta una parte de banda estrecha y la parte de banda alta. La tarea X100 calcula un conjunto de parámetros de filtro que caracterizan una envolvente espectral de la parte de banda alta. La tarea X200 calcula una señal ensanchada de manera espectral aplicando una función no lineal a una señal obtenida de la parte de banda estrecha. La tarea X300 genera una señal de banda alta sintetizada según (A) el conjunto de parámetros de filtro y

(B) una señal de excitación de banda alta basada en la señal ensanchada de manera espectral. La tarea X400 calcula una envolvente de ganancia en función de una relación entre (C) la energía de la parte de banda alta y (D) la energía de una señal obtenida de la parte de banda estrecha.

5 La figura 31a muestra un diagrama de flujo de un procedimiento M200 de generación de una señal de excitación de banda alta según una realización. La tarea Y100 calcula una señal ensanchada de manera armónica aplicando una función no lineal a una señal de excitación de banda estrecha obtenida de una parte de banda estrecha de una señal de voz. La tarea Y200 mezcla la señal ensanchada de manera armónica con una señal de ruido modulada para generar una señal de excitación de banda alta. La figura 31b muestra un diagrama de flujo de un procedimiento M210 de generación de una señal de excitación de banda alta según otra realización que incluye las tareas Y300 e Y400. La tarea Y300 calcula una envolvente de dominio de tiempo según la energía en el tiempo de una señal de entre la señal de excitación de banda estrecha y la señal ensanchada de manera armónica. La tarea Y400 modula una señal de ruido según la envolvente de dominio de tiempo para generar la señal de ruido modulada.

10 La figura 32 muestra un diagrama de flujo de un procedimiento M300, según una realización, para decodificar una parte de banda alta de una señal de voz que presenta una parte de banda estrecha y la parte de banda alta. La tarea Z100 recibe un conjunto de parámetros de filtro que caracterizan a una envolvente espectral de la parte de banda alta y un conjunto de factores de ganancia que caracterizan a una envolvente temporal de la parte de banda alta. La tarea Z200 calcula una señal ensanchada de manera espectral aplicando una función no lineal a una señal obtenida de la parte de banda estrecha. La tarea Z300 genera una señal de banda alta sintetizada según (A) el conjunto de parámetros de filtro y (B) una señal de excitación de banda alta basada en la señal ensanchada de manera espectral. La tarea Z400 modula una envolvente de ganancia de la señal de banda alta sintetizada en función del conjunto de factores de ganancia. Por ejemplo, la tarea Z400 puede configurarse para modular la envolvente de ganancia de la señal de banda alta sintetizada aplicando el conjunto de factores de ganancia a una señal de excitación obtenida de la parte de banda estrecha, a la señal ensanchada de manera espectral, a la señal de excitación de banda alta o a la señal de banda alta sintetizada.

## REIVINDICACIONES

1. Un procedimiento para generar una señal de excitación de banda alta, comprendiendo dicho procedimiento:  
 generar (A400) una señal (S160) ensanchada de manera espectral ensanchando el espectro de una señal (S80) que está basada en una señal codificada de excitación de banda baja;  
 5 realizar un filtrado (600) de antidispersión de una señal (S80) que está basada en la señal codificada de excitación de banda baja,  
 en el que la señal de excitación de banda alta está basada en la señal ensanchada de manera espectral, y  
 en el que la señal de excitación de banda alta está basada en un resultado de dicha acción de filtrado de antidispersión;  
 10 y en el que la acción de dicho filtrado de antidispersión está basada en un valor de al menos uno de entre un parámetro de inclinación espectral, un parámetro de ganancia de tono y un parámetro de modo de voz.
2. El procedimiento según la reivindicación 1, en el que dicha acción de filtrado de antidispersión incluye realizar un filtrado de antidispersión de la señal ensanchada de manera espectral.
3. El procedimiento según la reivindicación 1, en el que dicha acción de filtrado de antidispersión incluye realizar un filtrado de antidispersión de la señal de excitación de banda alta.
- 15 4. El procedimiento según la reivindicación 1, en el que dicha acción de filtrado de antidispersión de una señal incluye realizar una operación de filtrado en la señal según una función de transferencia de todo paso.
5. El procedimiento según la reivindicación 1, en el que dicha acción de filtrado de antidispersión de una señal incluye cambiar el espectro de fase de la señal sin modificar sustancialmente el espectro de magnitud de la señal.
- 20 6. El procedimiento según la reivindicación 1, en el que dicha generación de una señal ensanchada de manera espectral comprende ensanchar de manera armónica el espectro de una señal que está basada en la señal codificada de excitación de banda baja para obtener la señal ensanchada de manera espectral.
7. El procedimiento según la reivindicación 1, en el que dicha generación de una señal ensanchada de manera espectral comprende aplicar una función no lineal a una señal que está basada en la señal codificada de  
 25 excitación de banda baja para generar la señal ensanchada de manera espectral.
8. El procedimiento según la reivindicación 7, en el que la función no lineal comprende al menos una de entre la función de valor absoluto, la función de elevación al cuadrado y una función de truncamiento.
9. El procedimiento según la reivindicación 1, comprendiendo dicho procedimiento mezclar una señal que está basada en la señal ensanchada de manera espectral con una señal de ruido modulada, en el que la señal de  
 30 excitación de banda alta está basada en la señal mezclada.
10. El procedimiento según la reivindicación 9, en el que dicha mezcla incluye calcular una suma ponderada de la señal de ruido modulada y una señal que está basada en la señal ensanchada de manera espectral, en el que la señal de excitación de banda alta está basada en la suma ponderada.
11. El procedimiento según la reivindicación 9, en el que dicha señal de ruido modulada está basada en un resultado de modular una señal de ruido según una envolvente de dominio de tiempo de una señal basada en al menos una de entre la señal codificada de excitación de banda baja y la señal ensanchada de manera espectral.
- 35 12. El procedimiento según la reivindicación 11, comprendiendo dicho procedimiento generar la señal de ruido según una función determinista de información en una señal de voz codificada.
13. El procedimiento según la reivindicación 1, en el que dicha generación de una señal ensanchada de manera espectral incluye ensanchar de manera armónica el espectro de una señal muestreada de manera ascendente que está basada en la señal codificada de excitación de banda baja.
- 40 14. El procedimiento según la reivindicación 1, comprendiendo dicho procedimiento al menos uno de entre (A) aplanar de manera espectral la señal ensanchada de manera espectral y (B) aplanar de manera espectral la señal de excitación de banda alta.
- 45 15. El procedimiento según la reivindicación 14, en el que dicho aplanamiento espectral comprende:

calcular una pluralidad de coeficientes de filtro en función de una señal que va a aplanarse de manera espectral;  
y

filtrar la señal que va a aplanarse de manera espectral con un filtro blanqueador configurado según la pluralidad de coeficientes de filtro.

- 5 16. El procedimiento según la reivindicación 15, en el que dicho cálculo de una pluralidad de coeficientes de filtro incluye realizar un análisis de predicción lineal de la señal que va a aplanarse de manera espectral.
17. El procedimiento según la reivindicación 1, comprendiendo dicho procedimiento al menos uno de (i) codificar una señal de voz de banda alta según la señal de excitación de banda alta y (ii) descodificar una señal de voz de banda alta según la señal de excitación de banda alta.
- 10 18. El procedimiento según la reivindicación 1, que comprende además  
decidir si realizar dicho filtrado de antidispersión, donde dicha decisión está basada en un valor de al menos uno de entre un parámetro de inclinación espectral, un parámetro de ganancia de tono y un parámetro de modo de voz.
- 15 19. Un medio de almacenamiento de datos que presenta instrucciones ejecutables por máquina adaptadas para realizar el procedimiento de procesamiento de señales según la reivindicación 1, cuando dichas instrucciones se ejecutan en un ordenador.
- 20 20. Un aparato, que comprende:  
un ensanchador (A400) de espectro configurado para generar una señal (S160) ensanchada de manera espectral ensanchando el espectro de una señal (S80) que está basada en una señal codificada de excitación de banda baja;  
un filtro (600) de antidispersión configurado para filtrar una señal (S80) que está basada en la señal codificada de excitación de banda baja,  
en el que la señal de excitación de banda alta está basada en la señal ensanchada de manera espectral, y  
en el que la señal de excitación de banda alta está basada en una salida de dicho filtro de antidispersión, y
- 25 lógica de decisión configurada para decidir si filtrar una señal que está basada en la señal codificada de excitación de banda baja con dicho filtro de antidispersión, donde dicha lógica de decisión está configurada para decidir en función de un valor de al menos uno de entre un parámetro de inclinación espectral, un parámetro de ganancia de tono y un parámetro de modo de voz.
- 30 21. El aparato según la reivindicación 20, en el que dicho filtro de antidispersión está configurado para filtrar la señal ensanchada de manera espectral.
22. El aparato según la reivindicación 20, en el que dicho filtro de antidispersión está configurado para filtrar la señal de excitación de banda alta.
23. El aparato según la reivindicación 20, en el que dicho filtro de antidispersión está configurado para filtrar la señal según una función de transferencia de todo paso.
- 35 24. El aparato según la reivindicación 20, en el que dicho filtro de antidispersión está configurado para cambiar el espectro de fase de la señal sin modificar sustancialmente el espectro de magnitud de la señal.
25. El aparato según la reivindicación 20, en el que dicho ensanchador de espectro está configurado para ensanchar de manera armónica el espectro de una señal que está basada en la señal codificada de excitación de banda baja para obtener la señal ensanchada de manera espectral.
- 40 26. El aparato según la reivindicación 20, en el que dicho ensanchador de espectro está configurado para aplicar una función no lineal a una señal que está basada en la señal codificada de excitación de banda baja para generar la señal ensanchada de manera espectral.
27. El aparato según la reivindicación 26, en el que la función no lineal comprende al menos una de entre la función de valor absoluto, la función de elevación al cuadrado y una función de truncamiento.
- 45 28. El aparato según la reivindicación 20, comprendiendo dicho aparato un combinador configurado para mezclar una señal que está basada en la señal ensanchada de manera espectral con una señal de ruido modulada, en el

que la señal de excitación de banda alta está basada en una salida de dicho combinador.

29. El aparato según la reivindicación 28, en el que dicho mezclador está configurado para calcular una suma ponderada de la señal de ruido modulada y una señal que está basada en la señal ensanchada de manera espectral, en el que la señal de excitación de banda alta está basada en la suma ponderada.
- 5 30. El aparato según la reivindicación 28, incluyendo dicho aparato un segundo combinador configurado para modular una señal de ruido según una envolvente de dominio de tiempo de una señal basada en al menos una de entre la señal codificada de excitación de banda baja y la señal ensanchada de manera espectral, en el que la señal de ruido modulada está basada en una salida de dicho segundo combinador.
- 10 31. El aparato según la reivindicación 30, comprendiendo dicho aparato un generador de ruido configurado para generar la señal de ruido según una función determinista de información en una señal de voz codificada.
32. El aparato según la reivindicación 20, en el que dicho ensanchador de espectro está configurado para ensanchar de manera armónica el espectro de una señal muestreada de manera ascendente que está basada en la señal codificada de excitación de banda baja.
- 15 33. El aparato según la reivindicación 20, comprendiendo dicho aparato un aplanador espectral configurado para aplanar de manera espectral al menos una de entre la señal ensanchada de manera espectral y la señal de excitación de banda alta.
- 20 34. El aparato según la reivindicación 33, en el que dicho aplanador espectral está configurado para calcular una pluralidad de coeficientes de filtro en función de una señal que va a aplanarse de manera espectral y para filtrar la señal que va a aplanarse de manera espectral con un filtro blanqueador configurado según la pluralidad de coeficientes de filtro.
35. El aparato según la reivindicación 34, en el que dicho aplanador espectral está configurado para calcular la pluralidad de coeficientes de filtro en función de un análisis de predicción lineal de la señal que va a aplanarse de manera espectral.
- 25 36. El aparato según la reivindicación 20, comprendiendo dicho aparato al menos uno de (i) un codificador de voz de banda alta configurado para codificar una señal de voz de banda alta según la señal de excitación de banda alta y (ii) un decodificador de voz de banda alta configurado para decodificar una señal de voz de banda alta según la señal de excitación de banda alta.
37. El aparato según la reivindicación 20, comprendiendo dicho aparato un teléfono celular.
- 30 38. El aparato según la reivindicación 20, comprendiendo dicho aparato un dispositivo configurado para transmitir una pluralidad de paquetes compatibles con una versión del Protocolo de Internet, en el que la pluralidad de paquetes describe la señal de excitación de banda estrecha.
39. El aparato según la reivindicación 20, comprendiendo dicho aparato un dispositivo configurado para recibir una pluralidad de paquetes compatibles con una versión del Protocolo de Internet, en el que la pluralidad de paquetes describe la señal de excitación de banda estrecha.
- 35 40. El aparato según la reivindicación 20, comprendiendo dicho aparato un teléfono celular.
41. Programa informático que comprende instrucciones ejecutables por ordenador adaptadas para realizar las etapas del procedimiento según cualquiera de las reivindicaciones 1 a 18 cuando el programa informático se ejecuta en un ordenador.

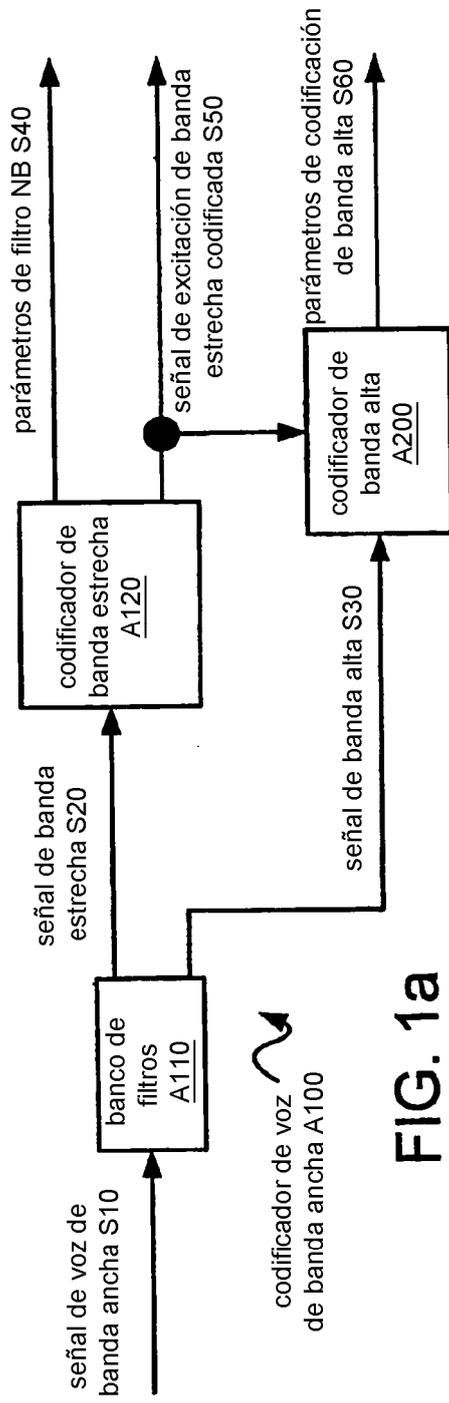


FIG. 1a

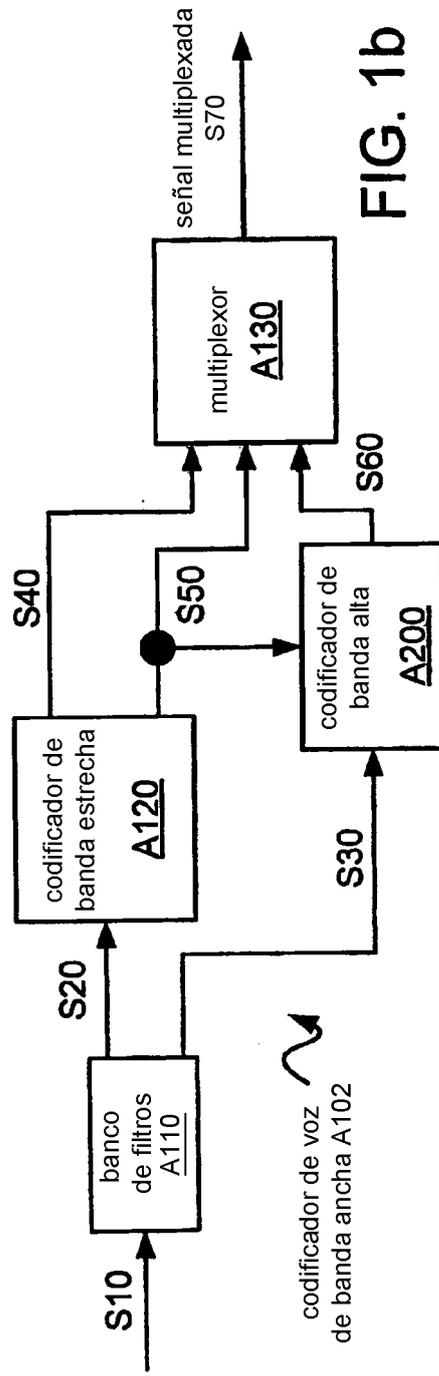


FIG. 1b

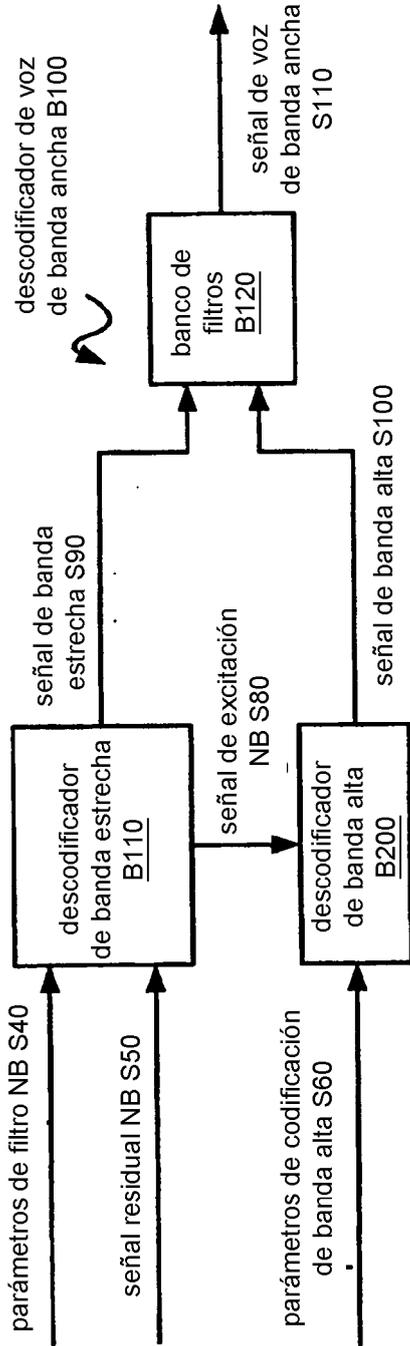


FIG. 2a

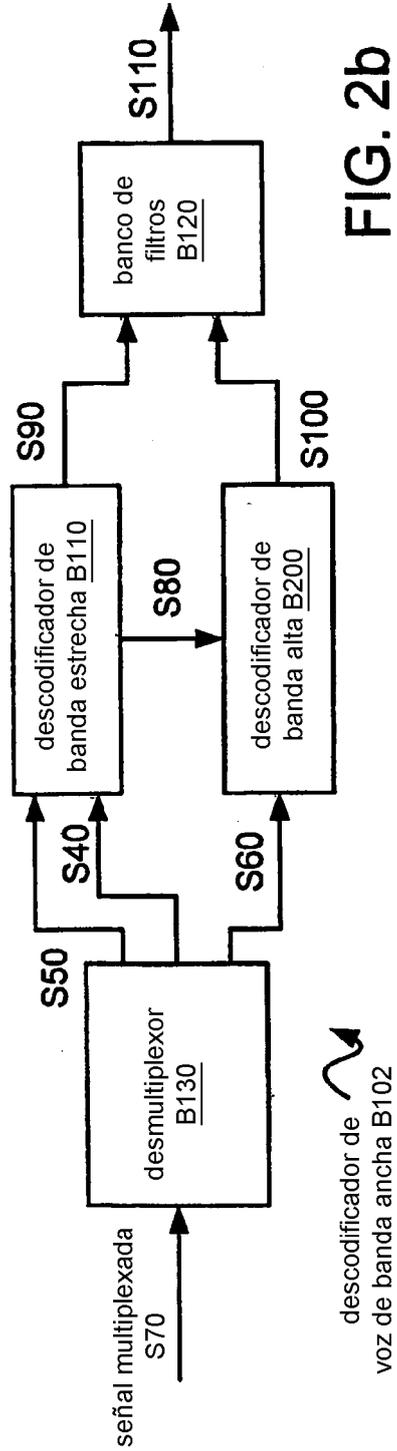
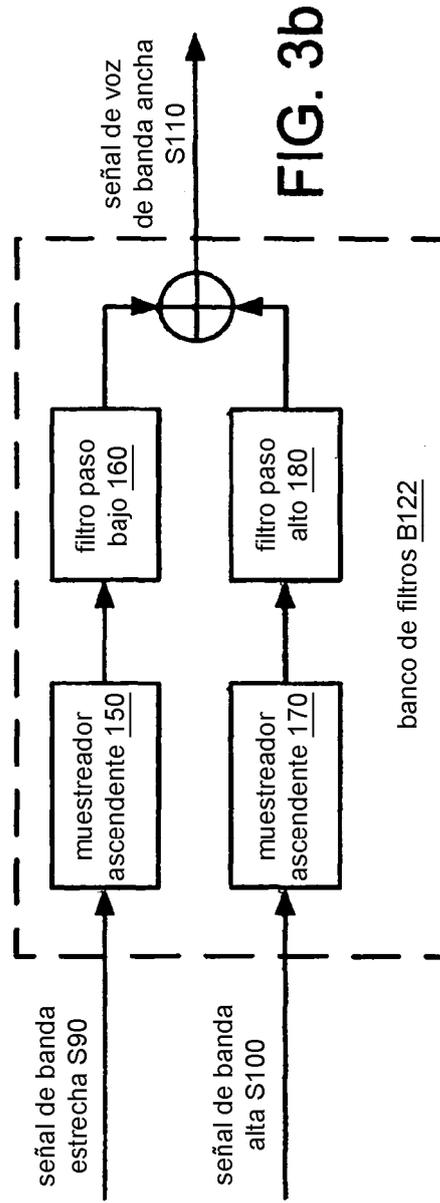
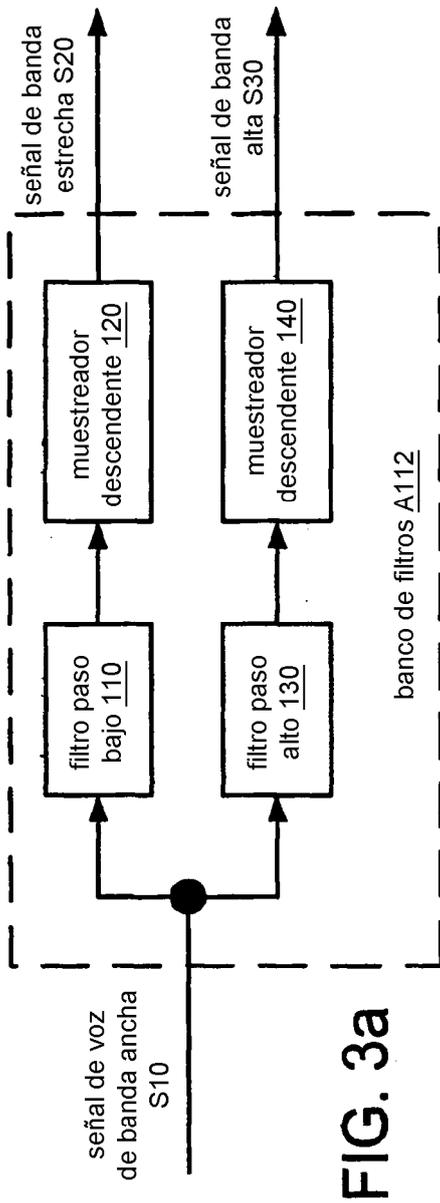


FIG. 2b



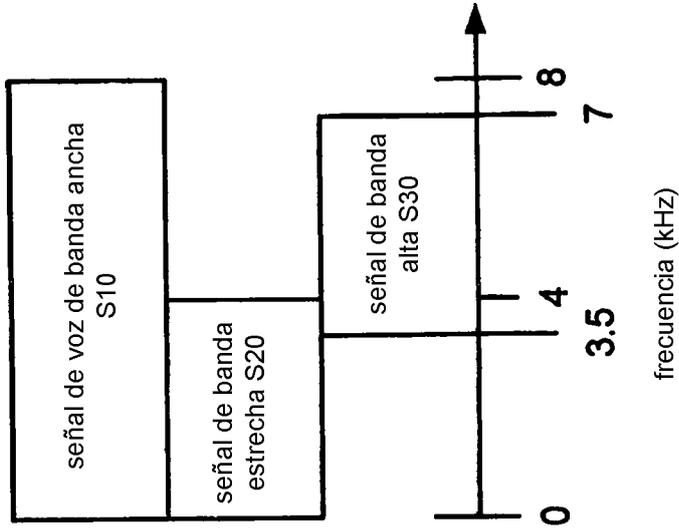


FIG. 4a

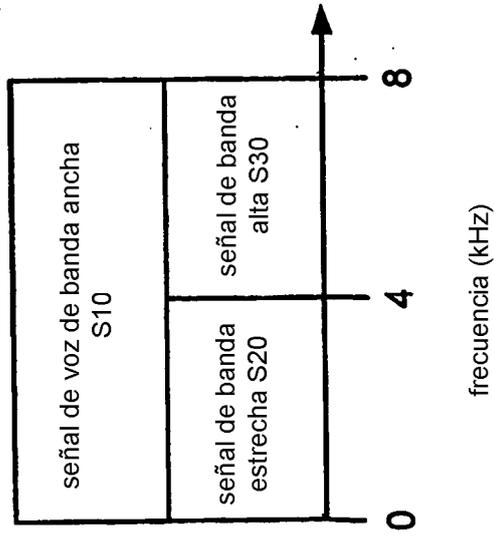


FIG. 4b

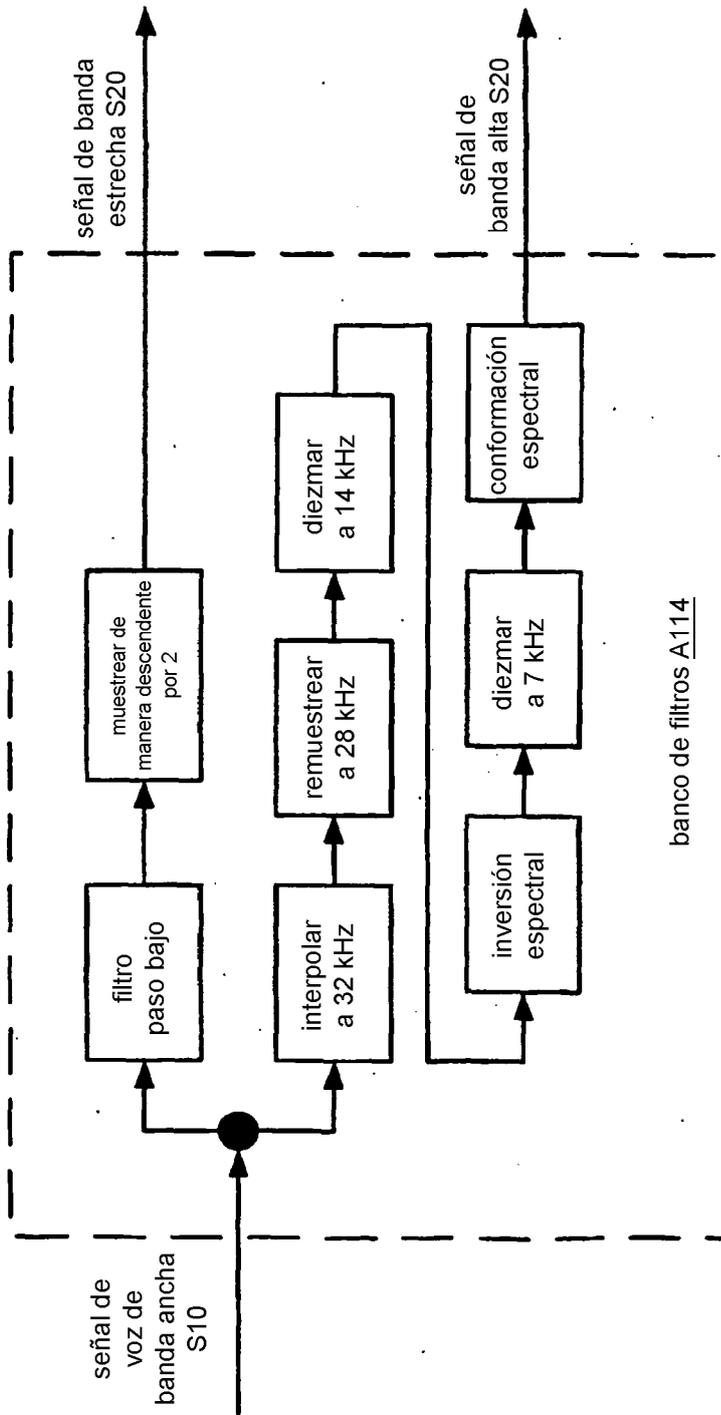


FIG. 4c

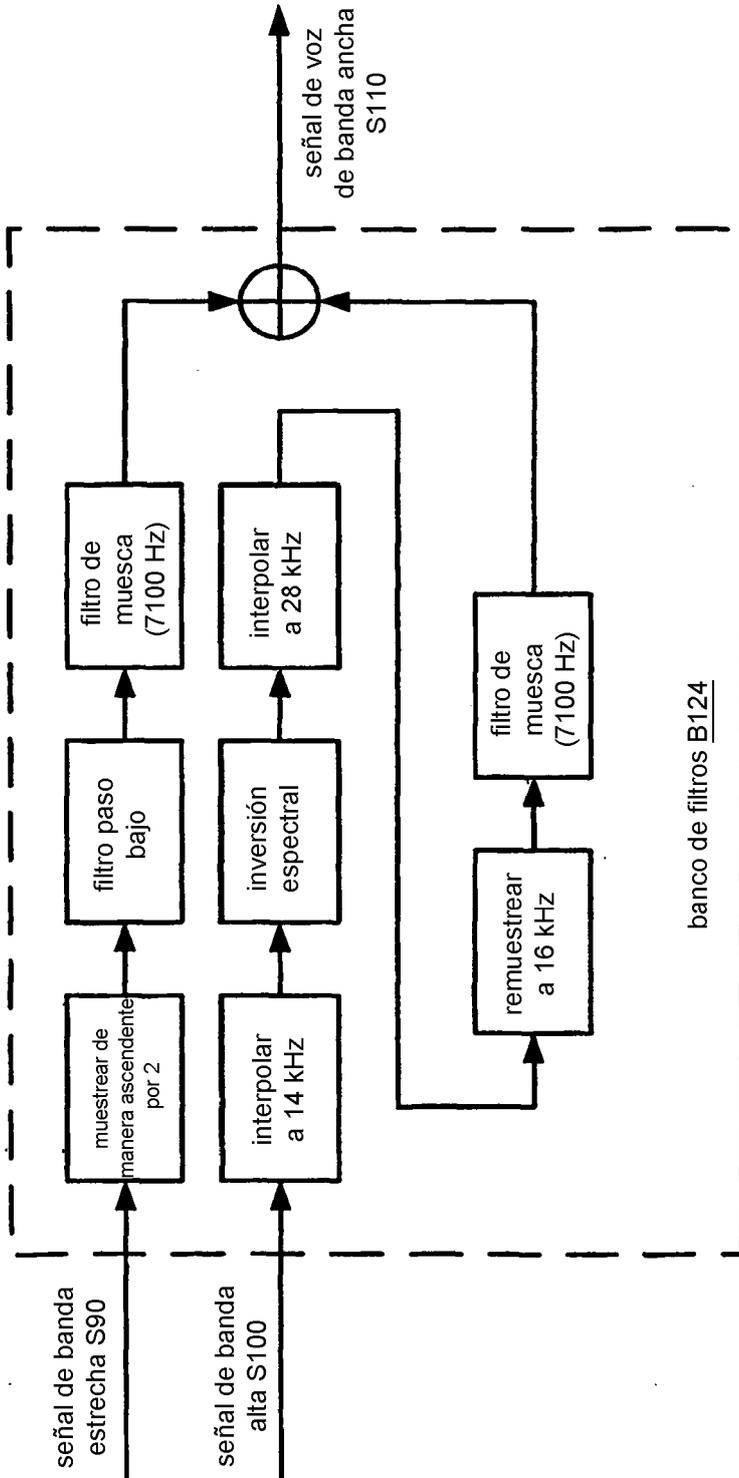


FIG. 4d

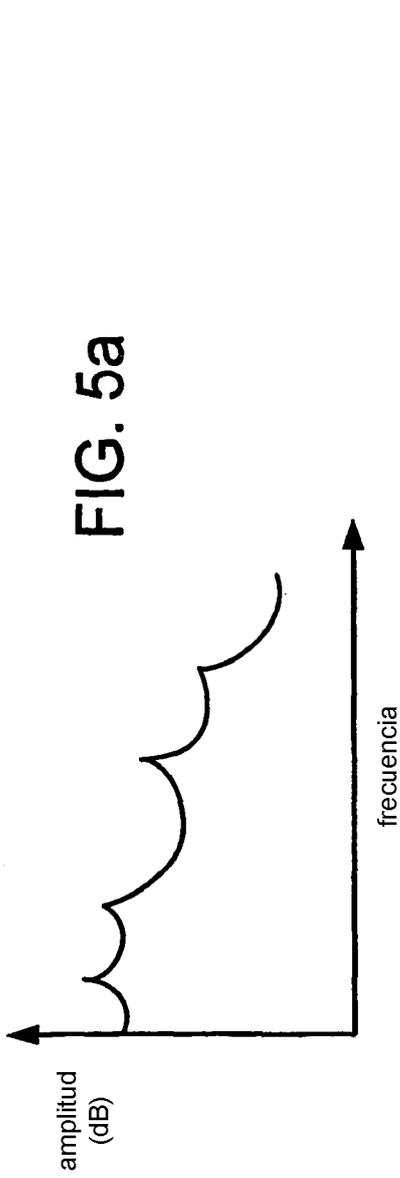


FIG. 5a

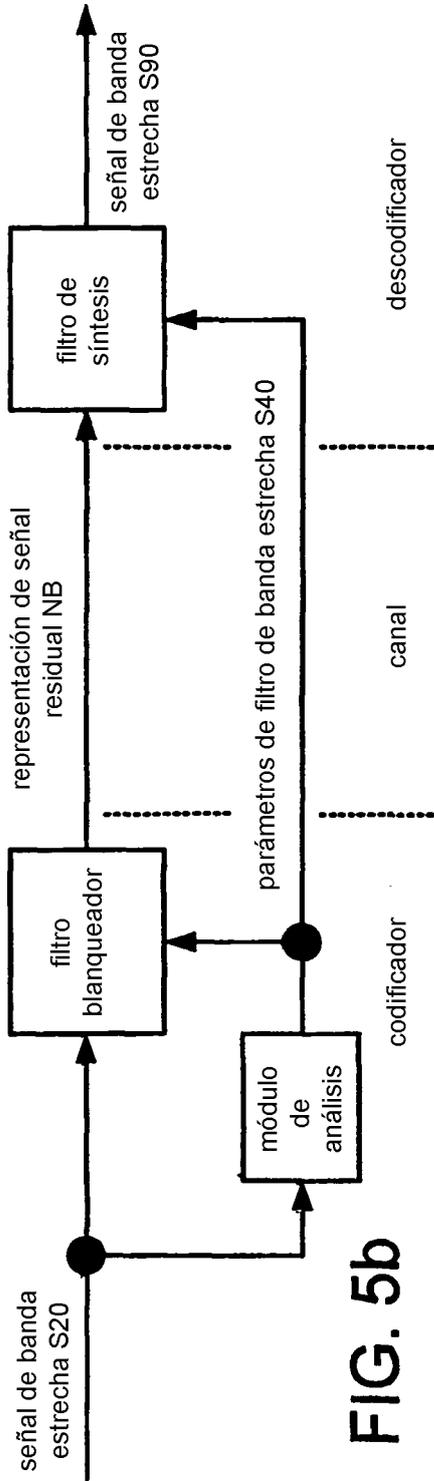


FIG. 5b

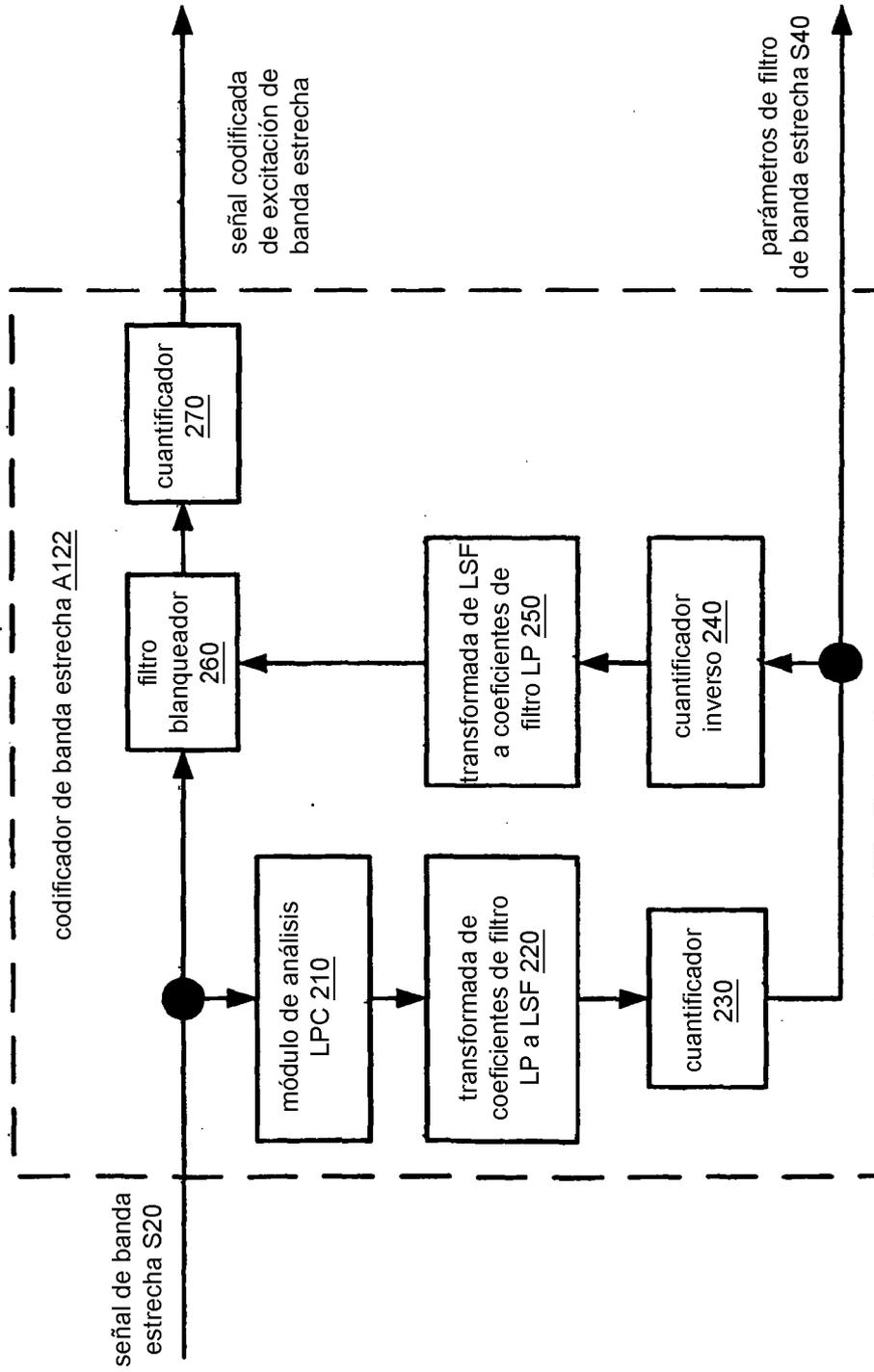


FIG. 6

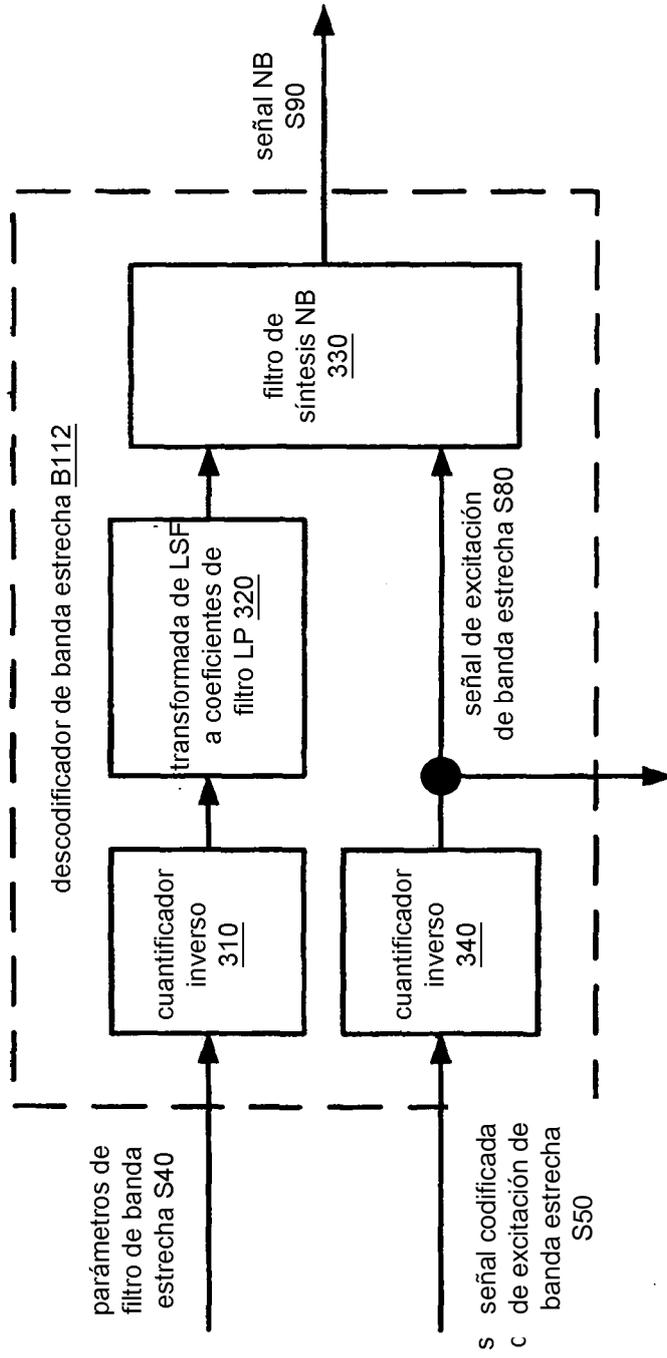


FIG. 7

FIG. 8a

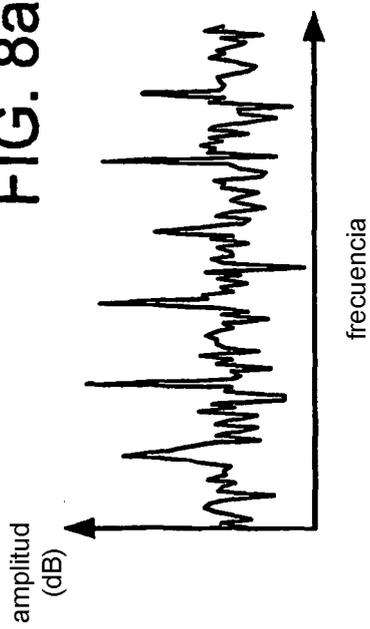


FIG. 8b

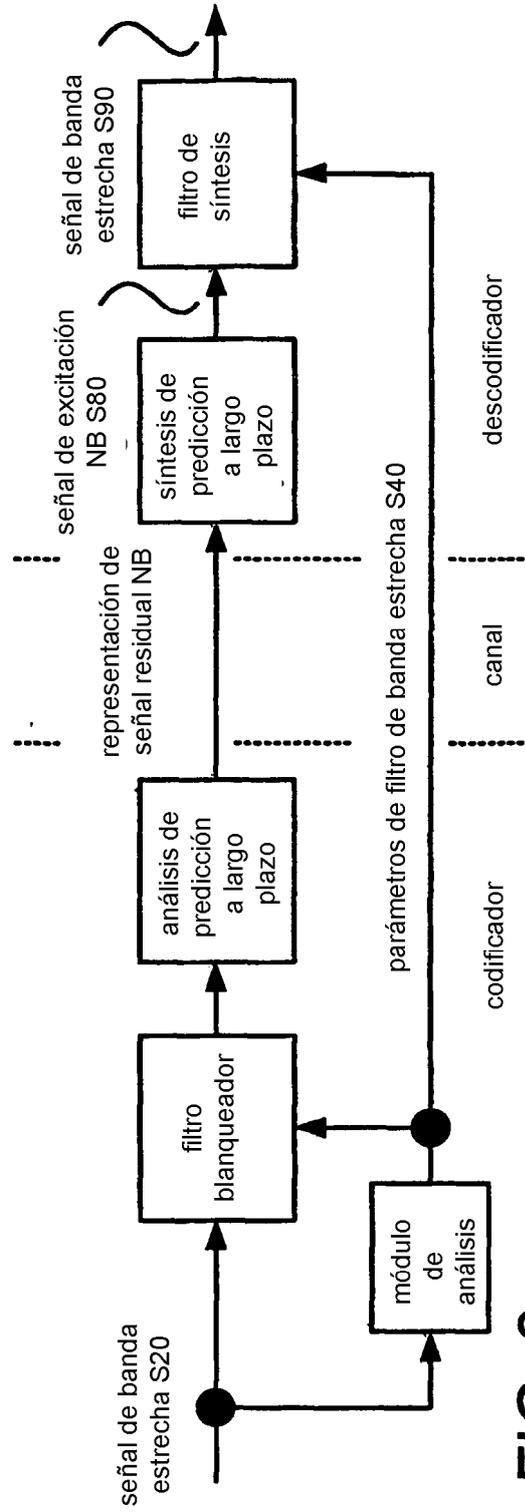
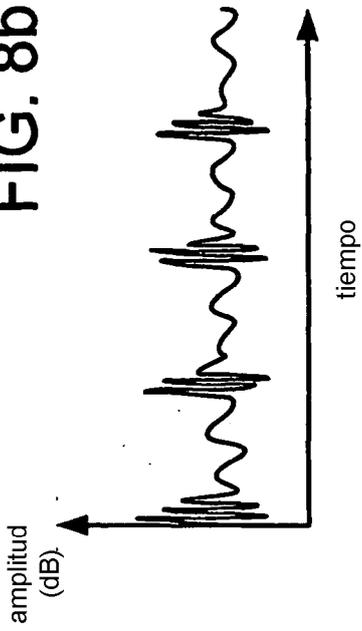


FIG. 9

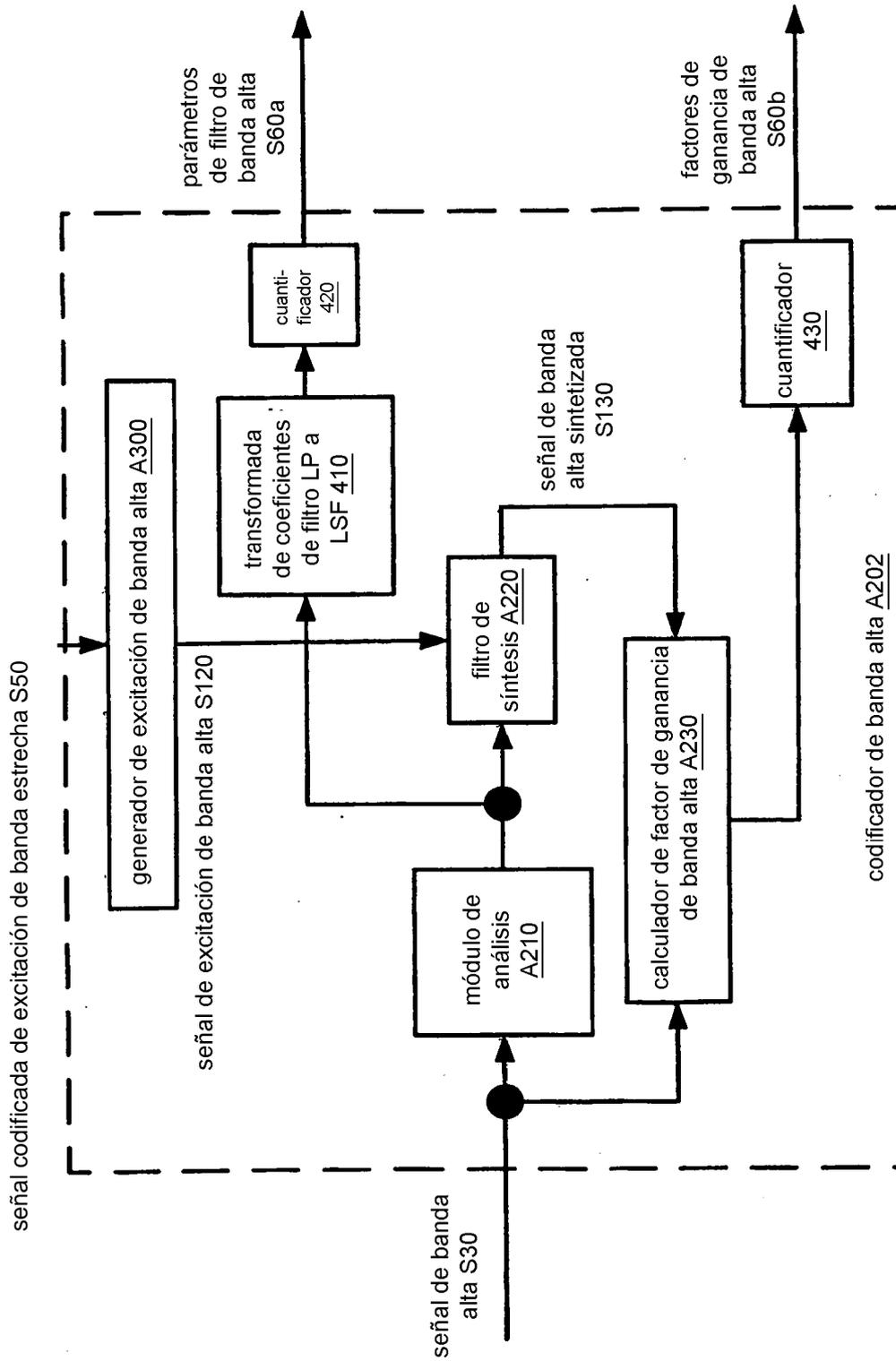


FIG. 10

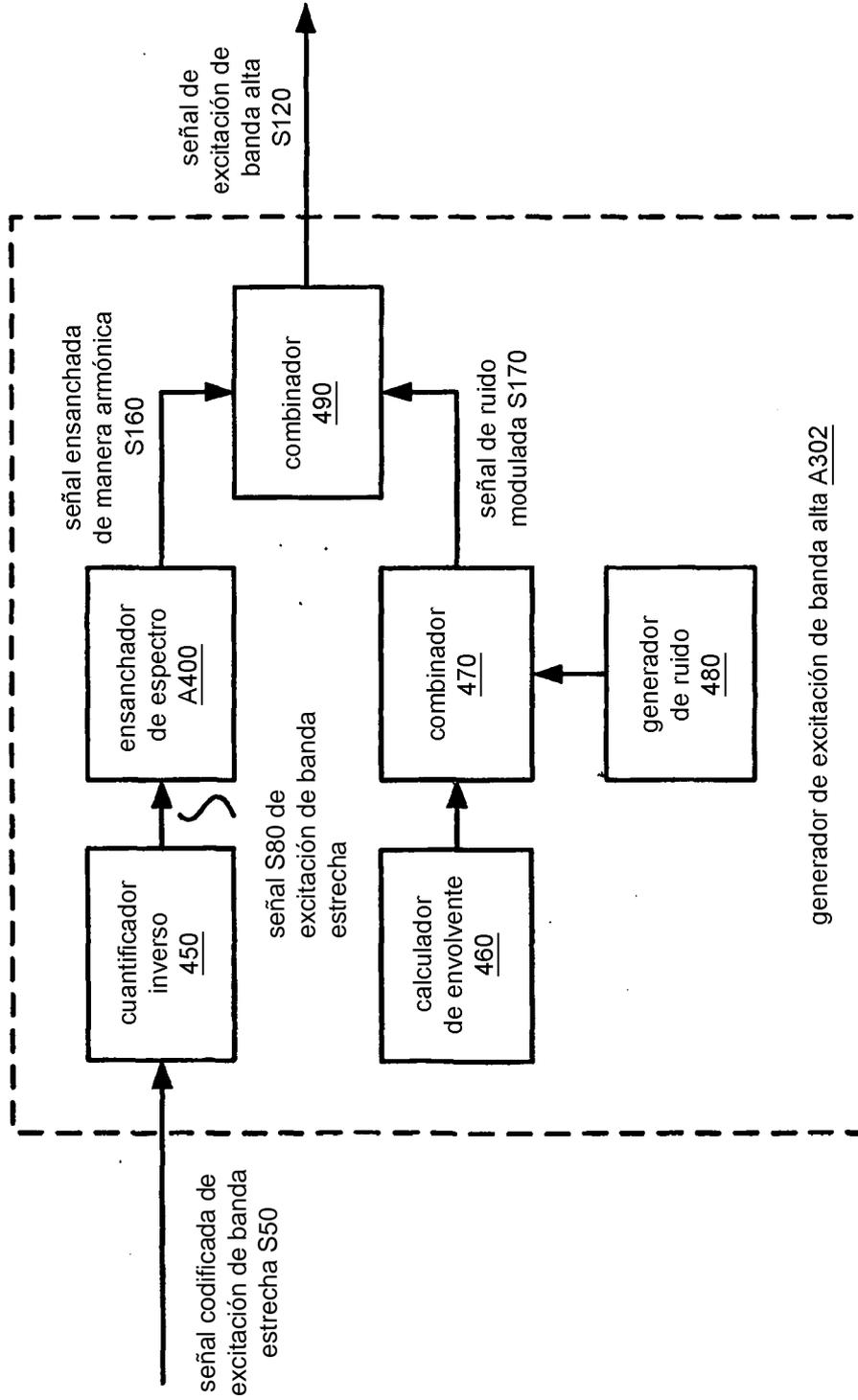


FIG. 11

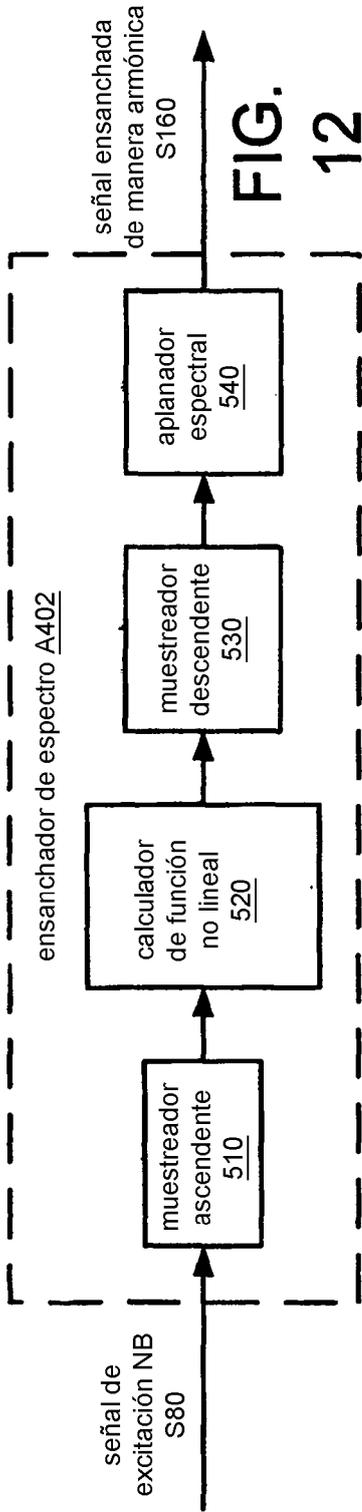


FIG. 12

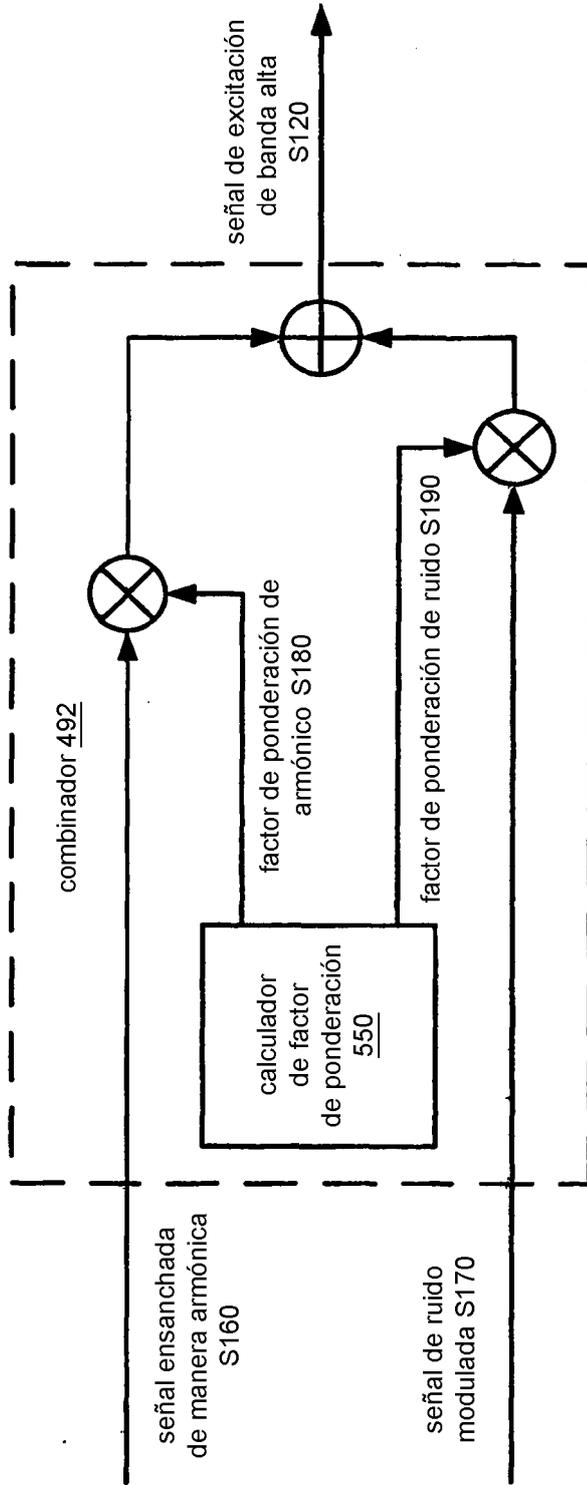


FIG. 16

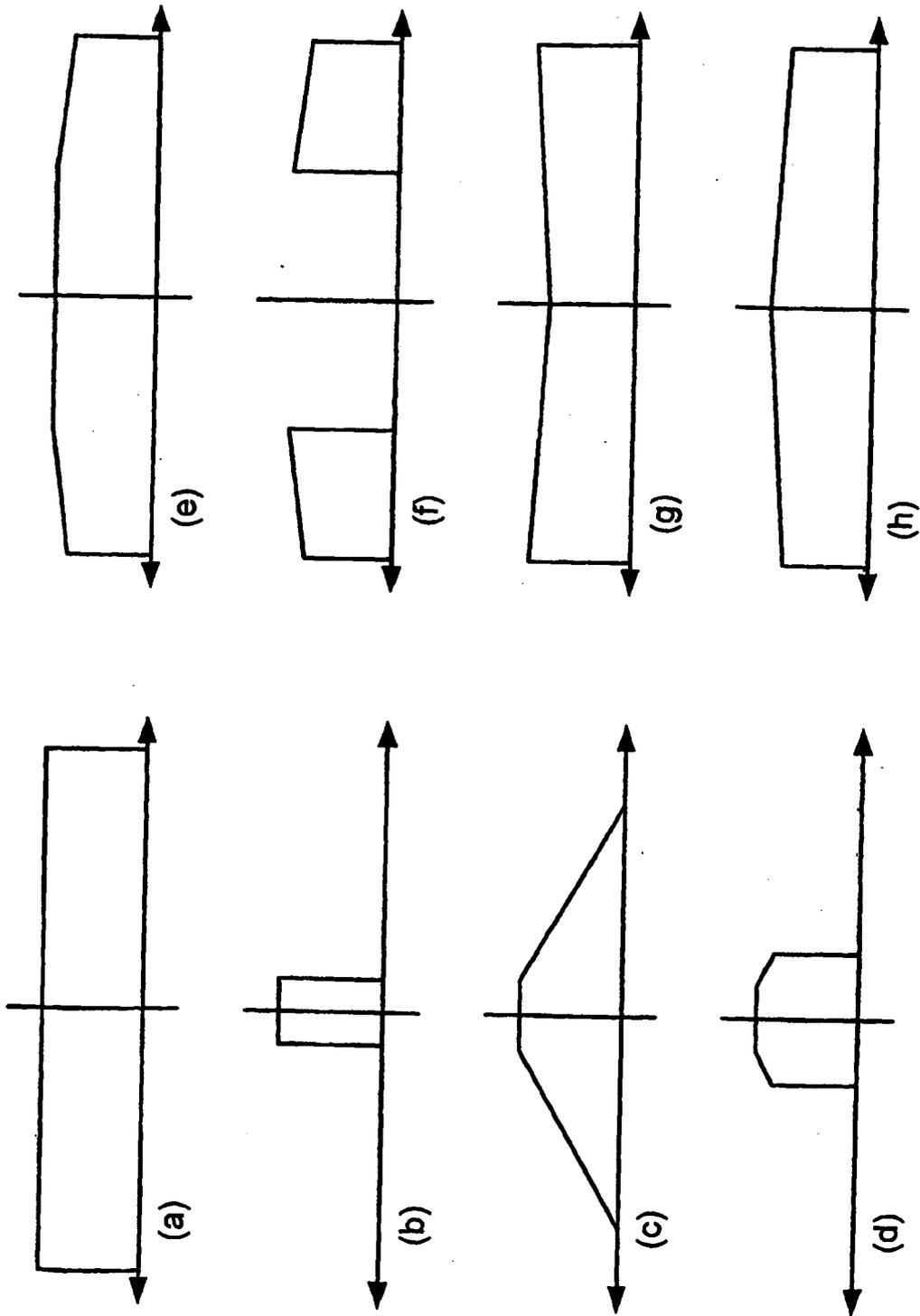


FIG. 12a

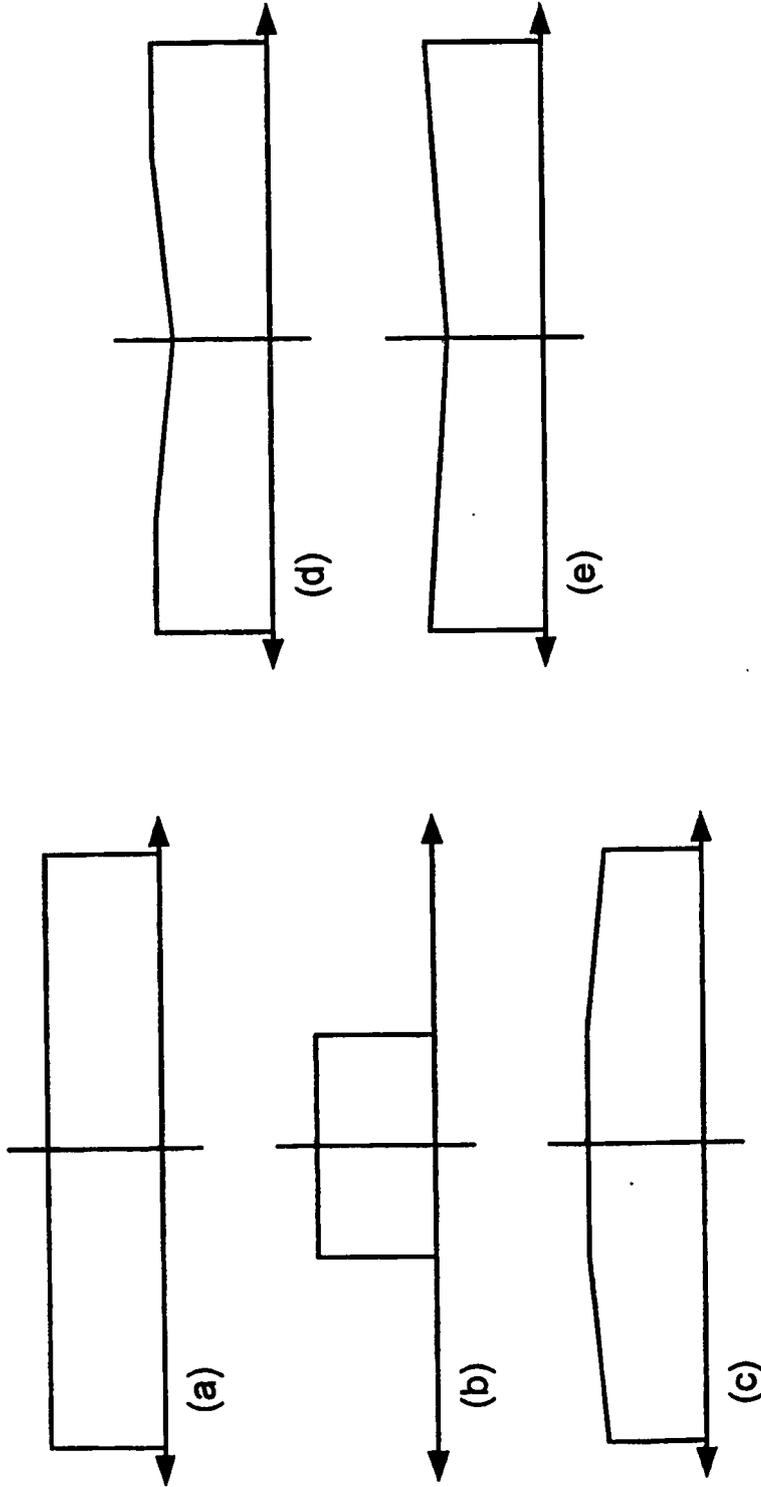


FIG. 12b

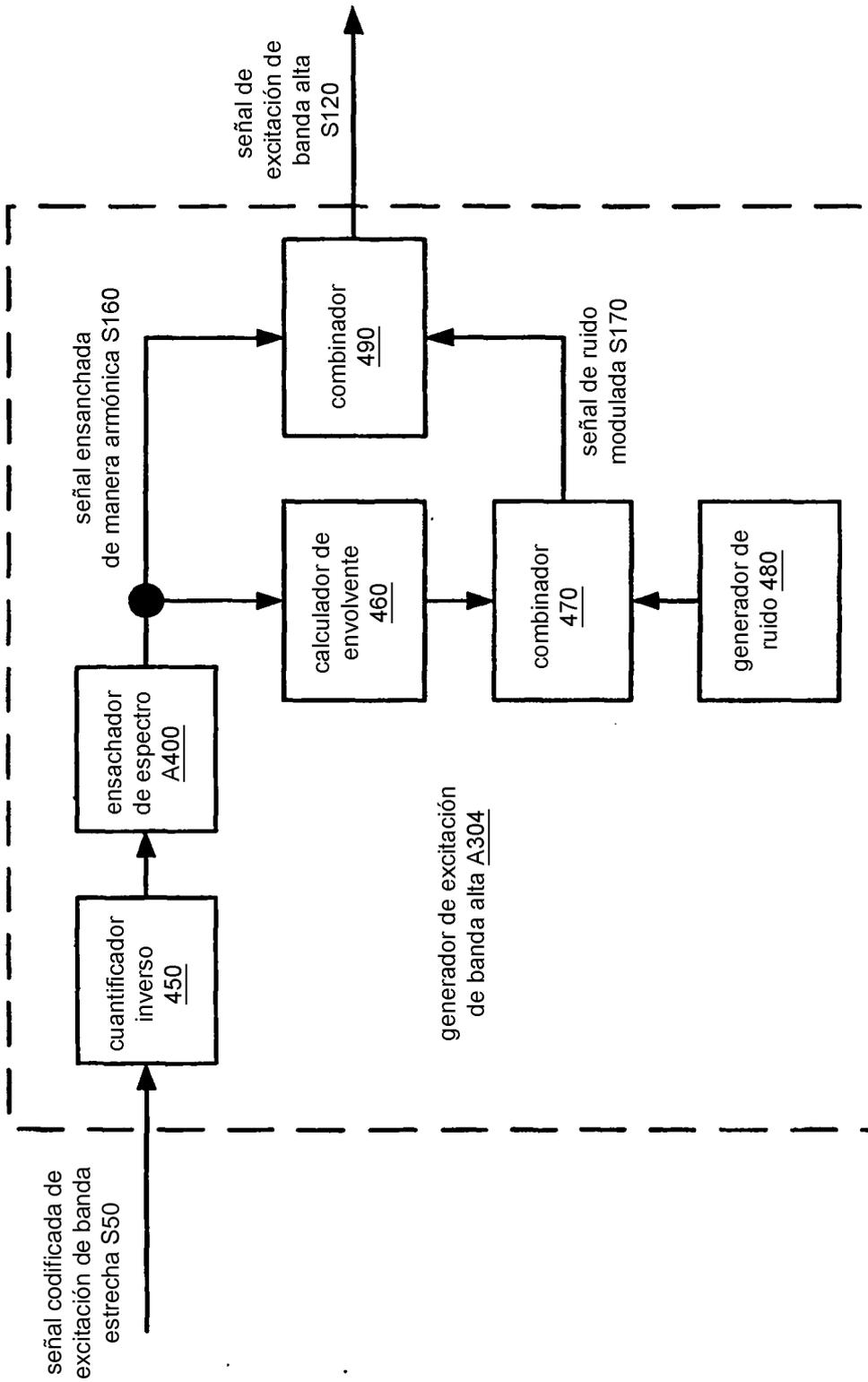


FIG. 13

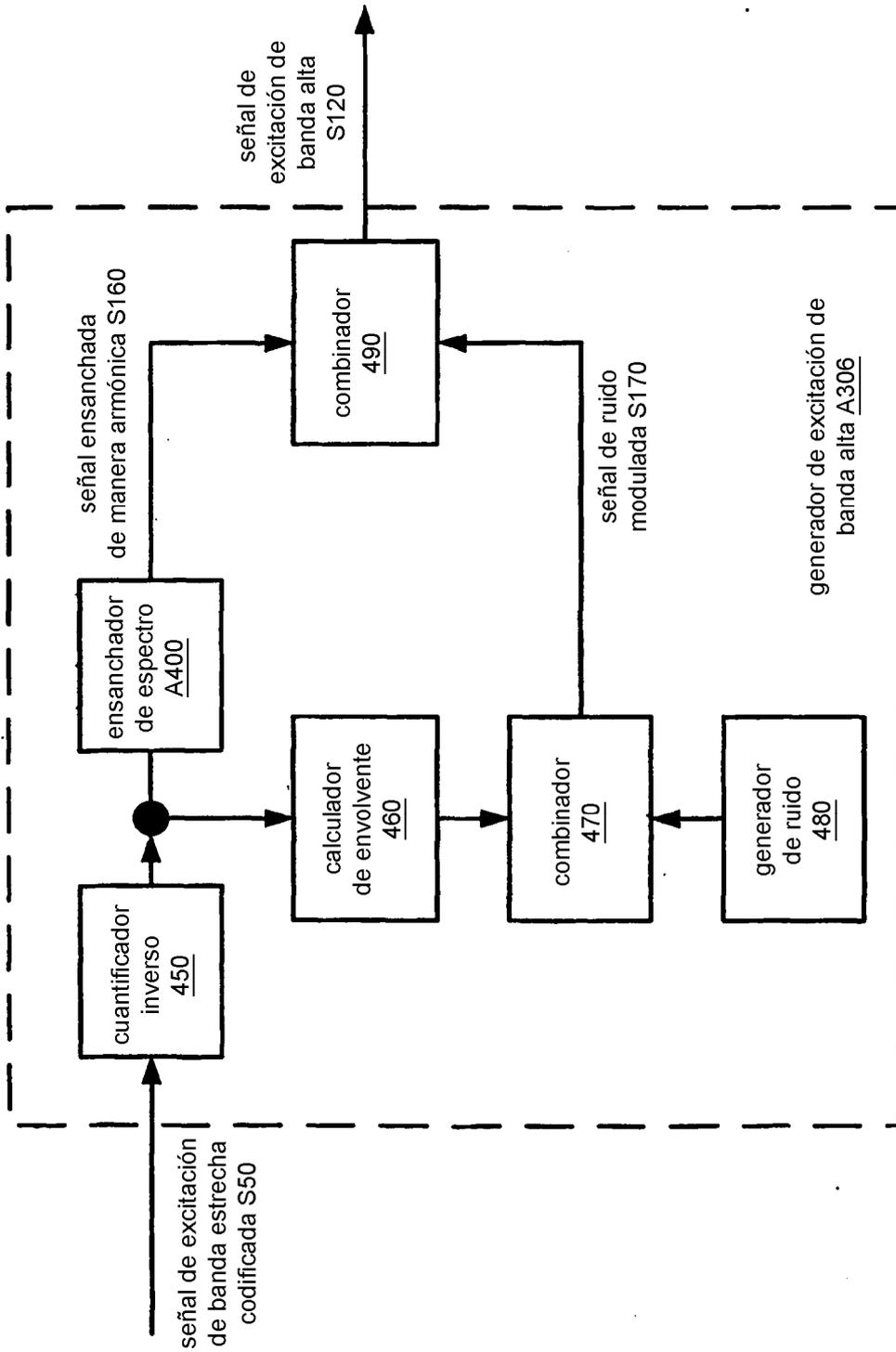


FIG. 14

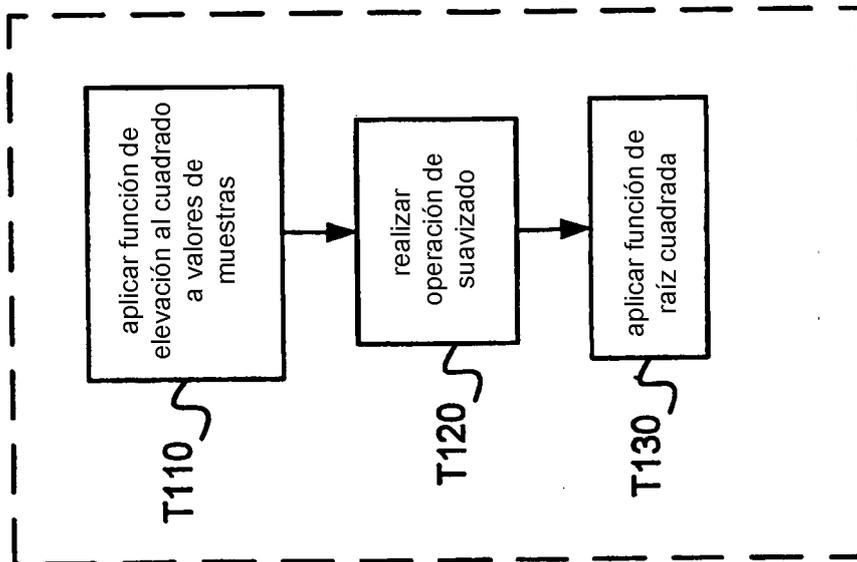


FIG. 15

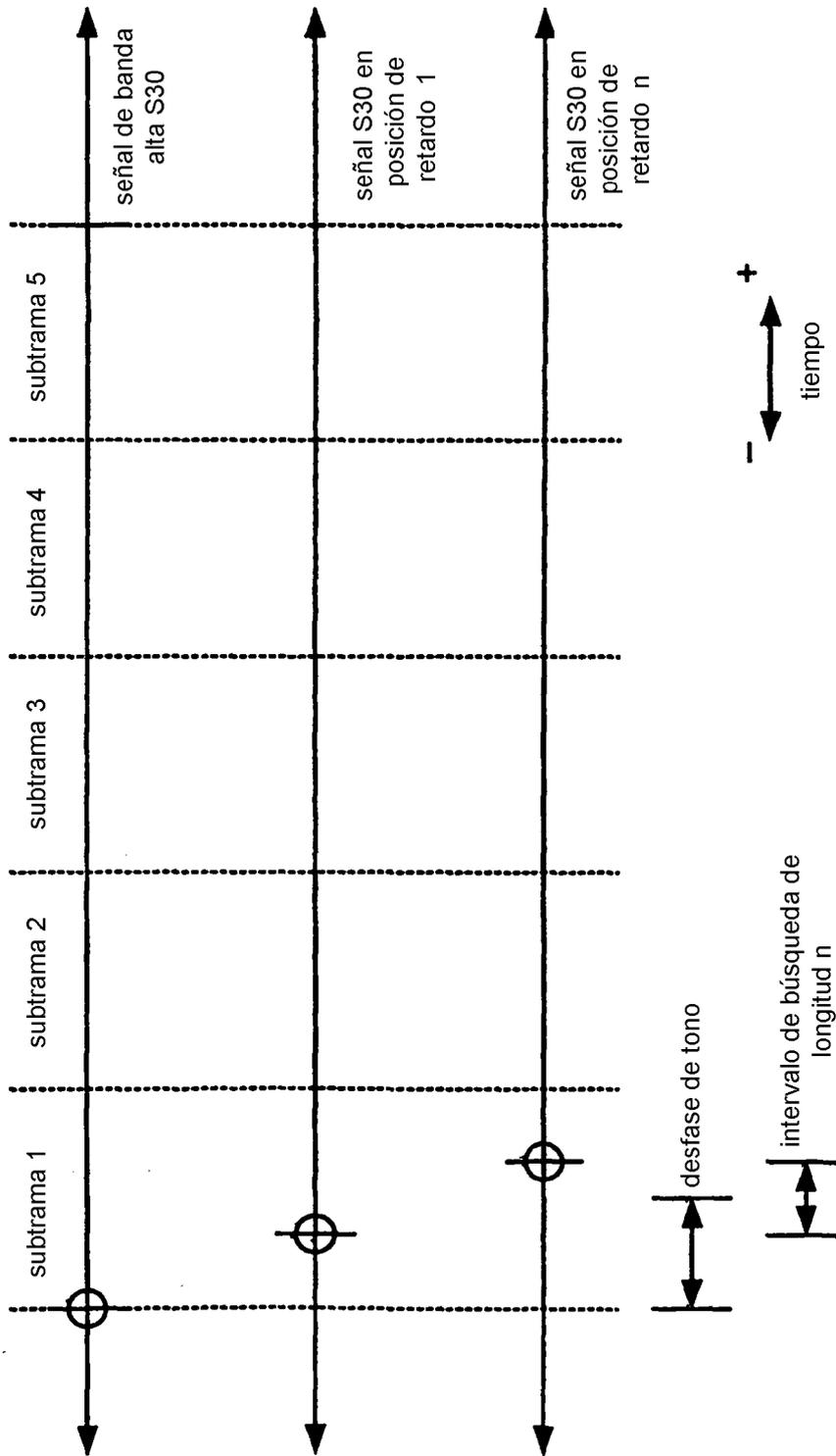


FIG. 17

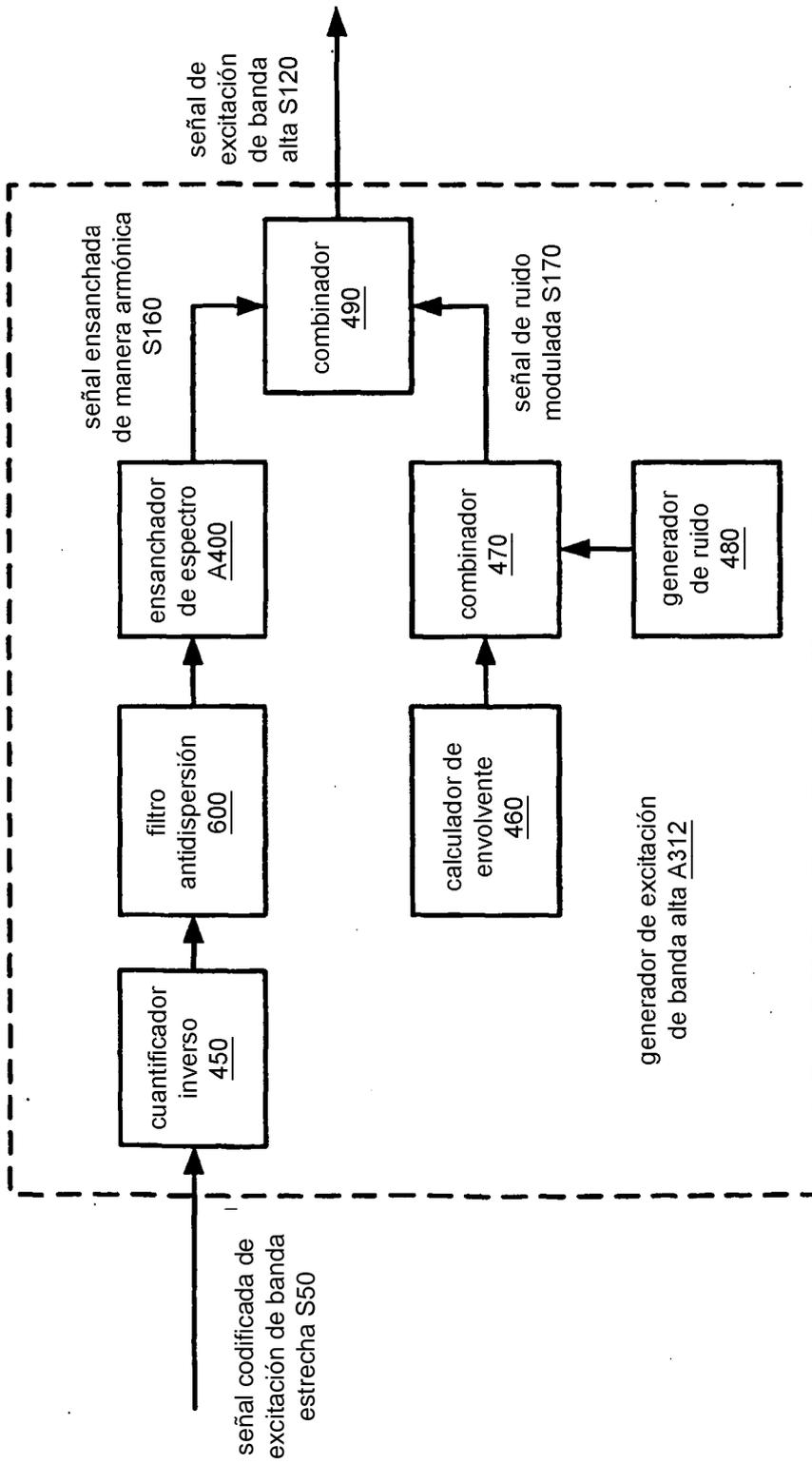


FIG. 18

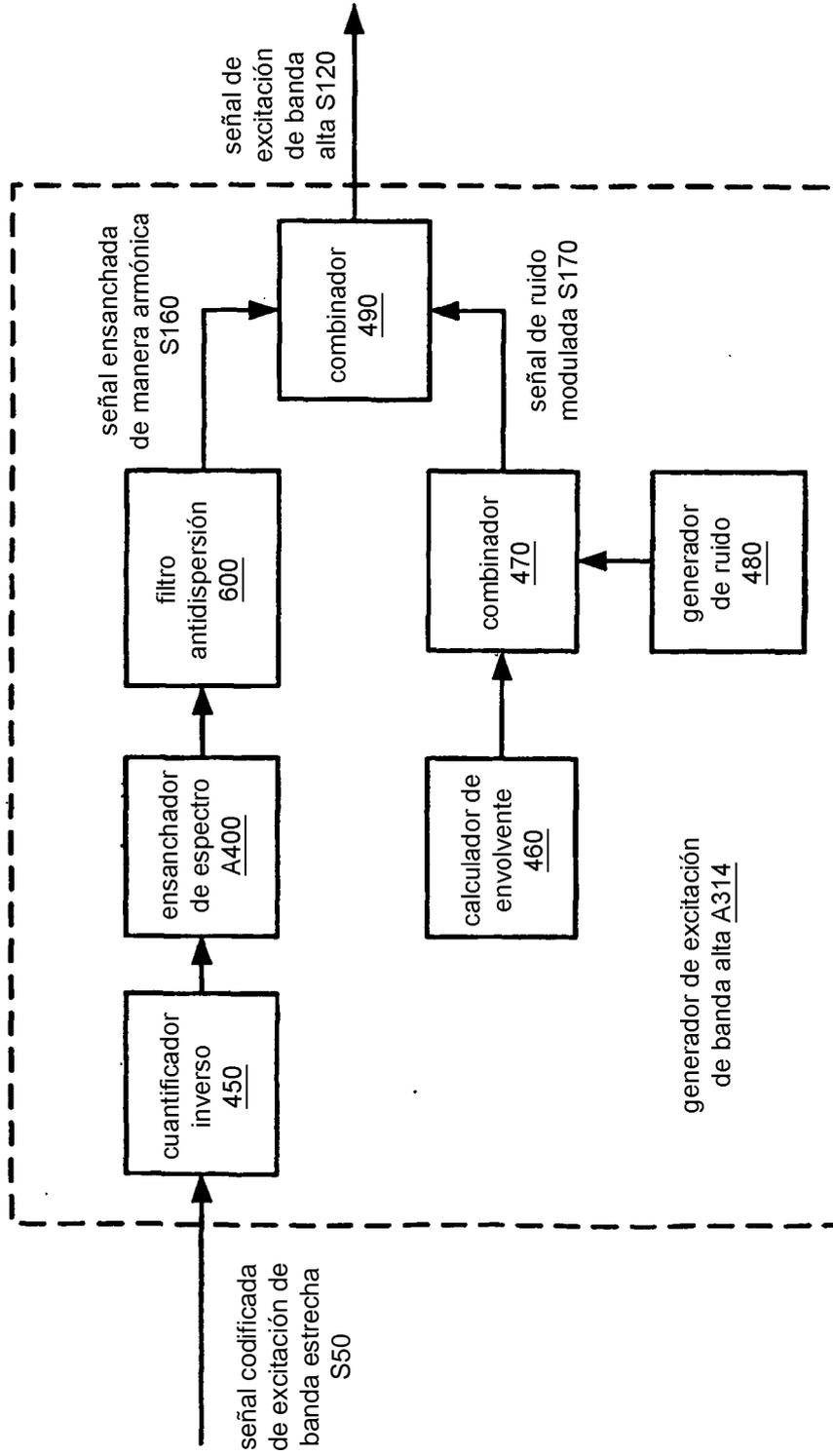


FIG. 19

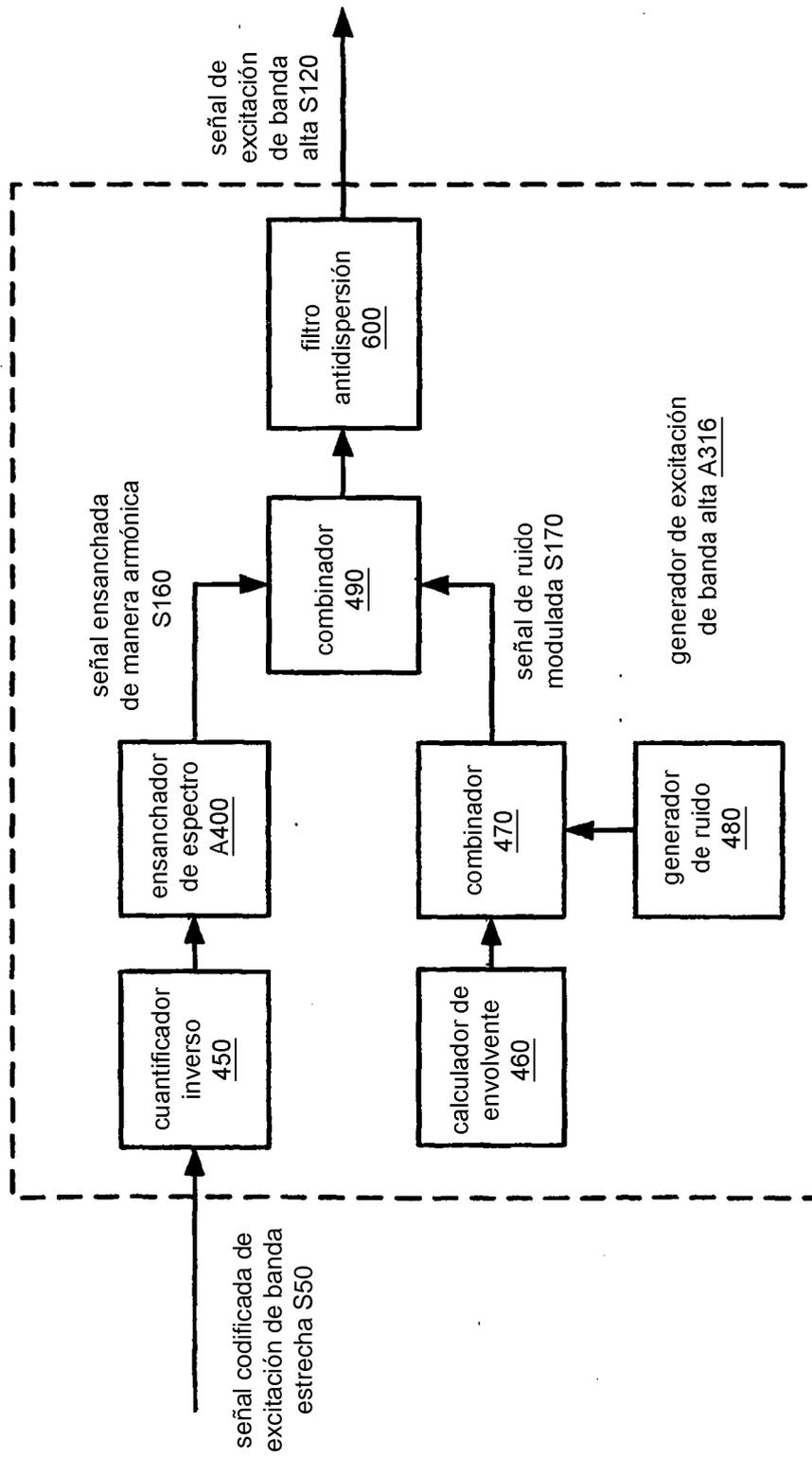


FIG. 20

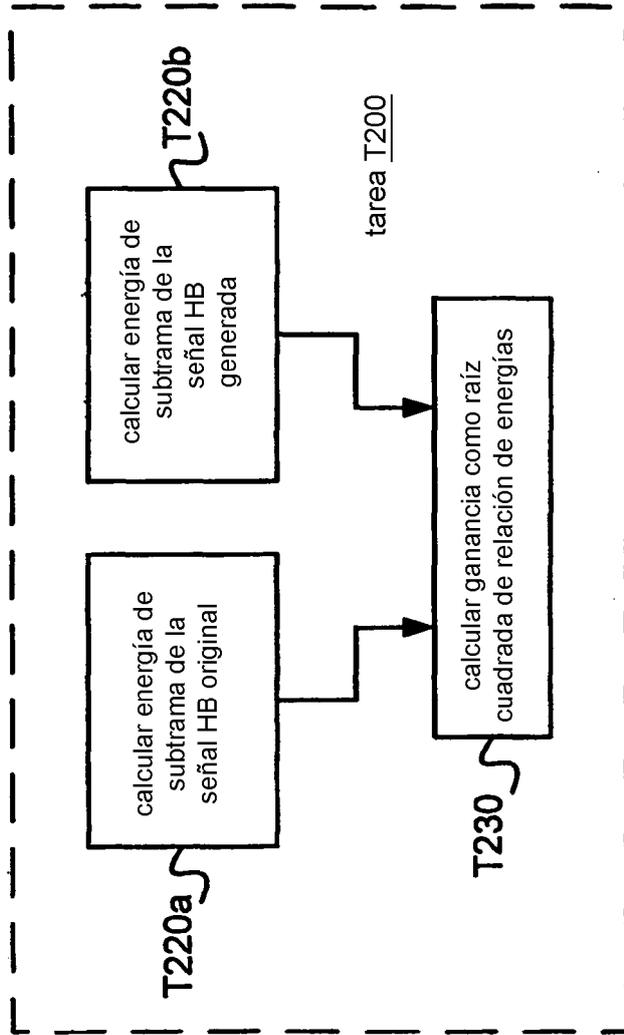


FIG. 21

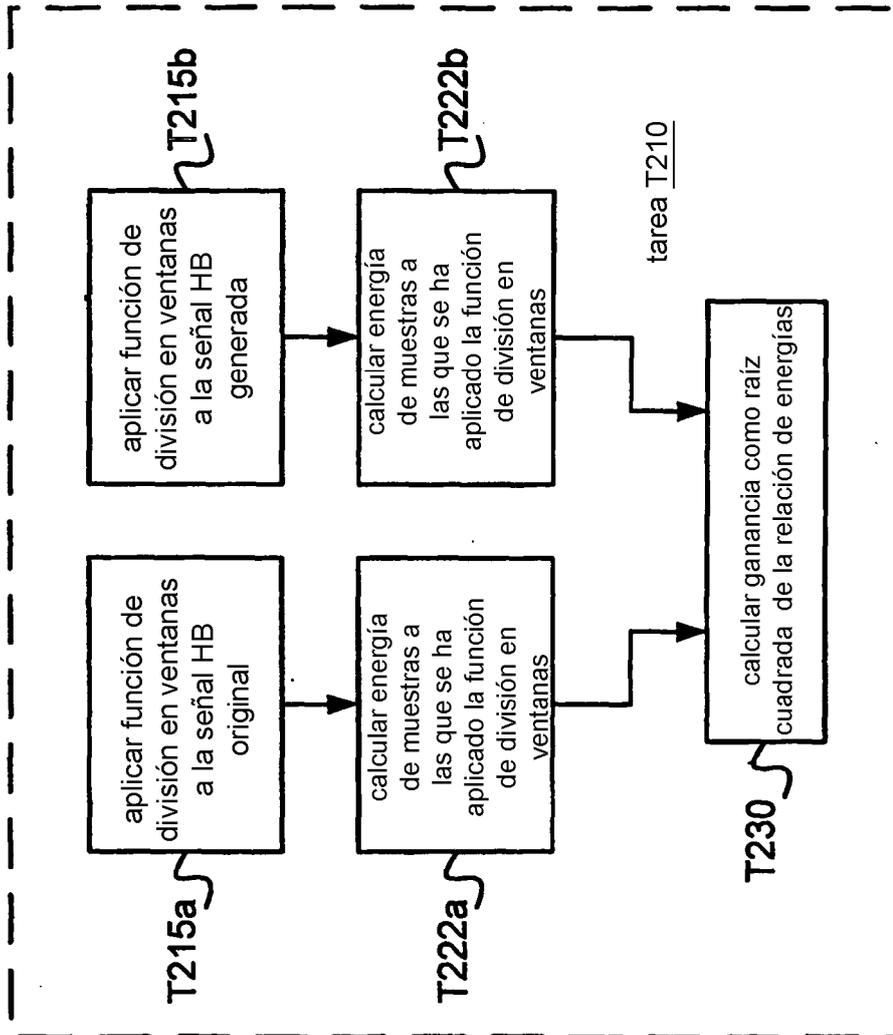


FIG. 22

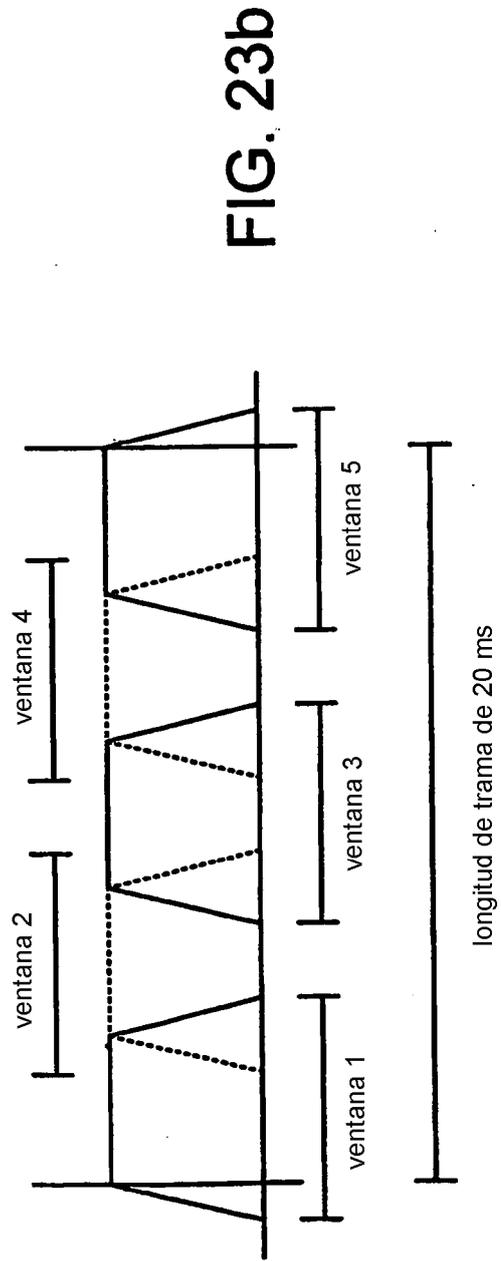
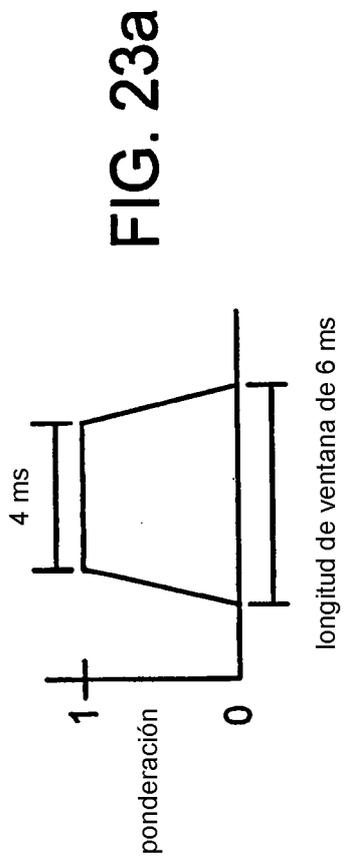
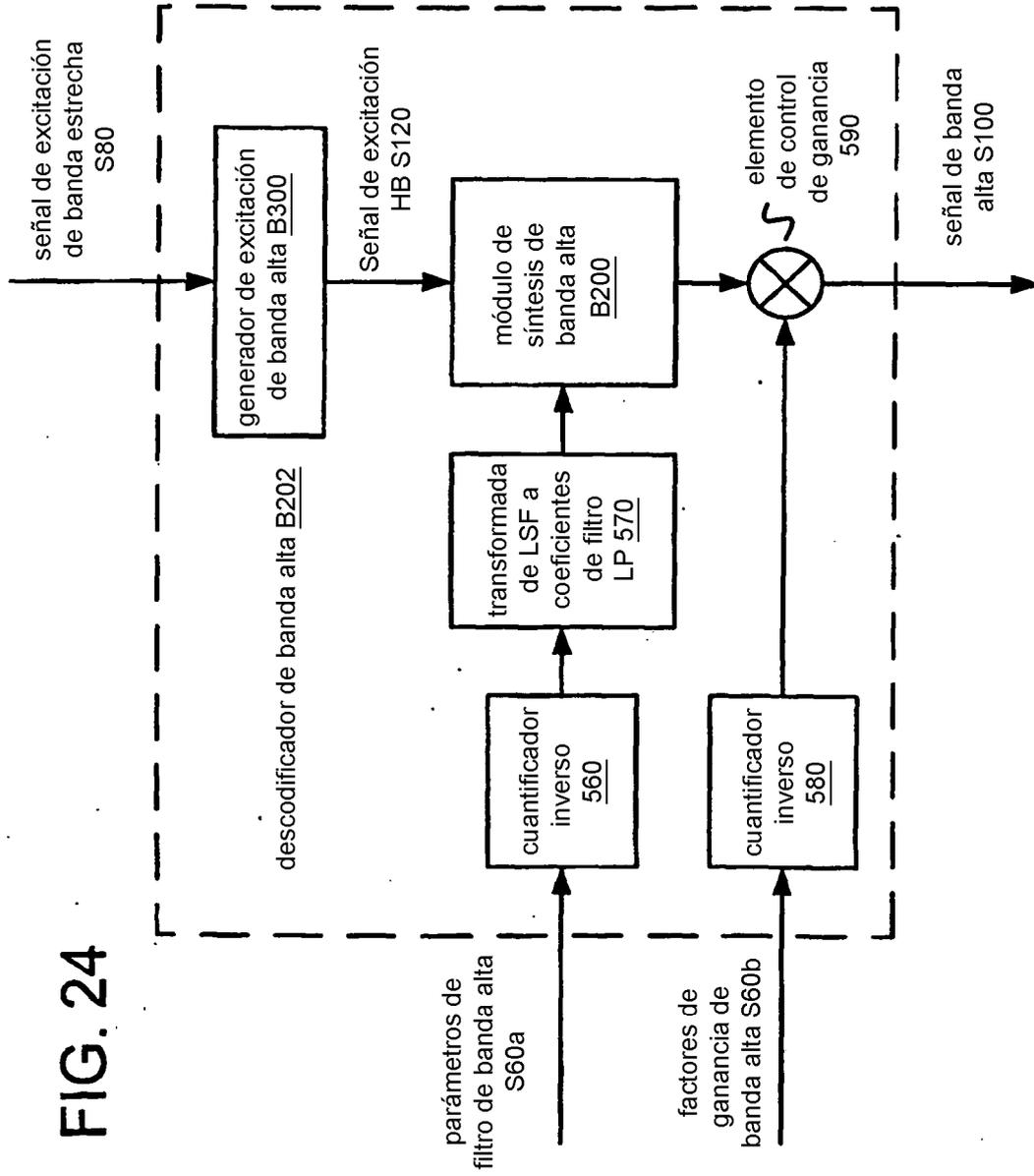


FIG. 24



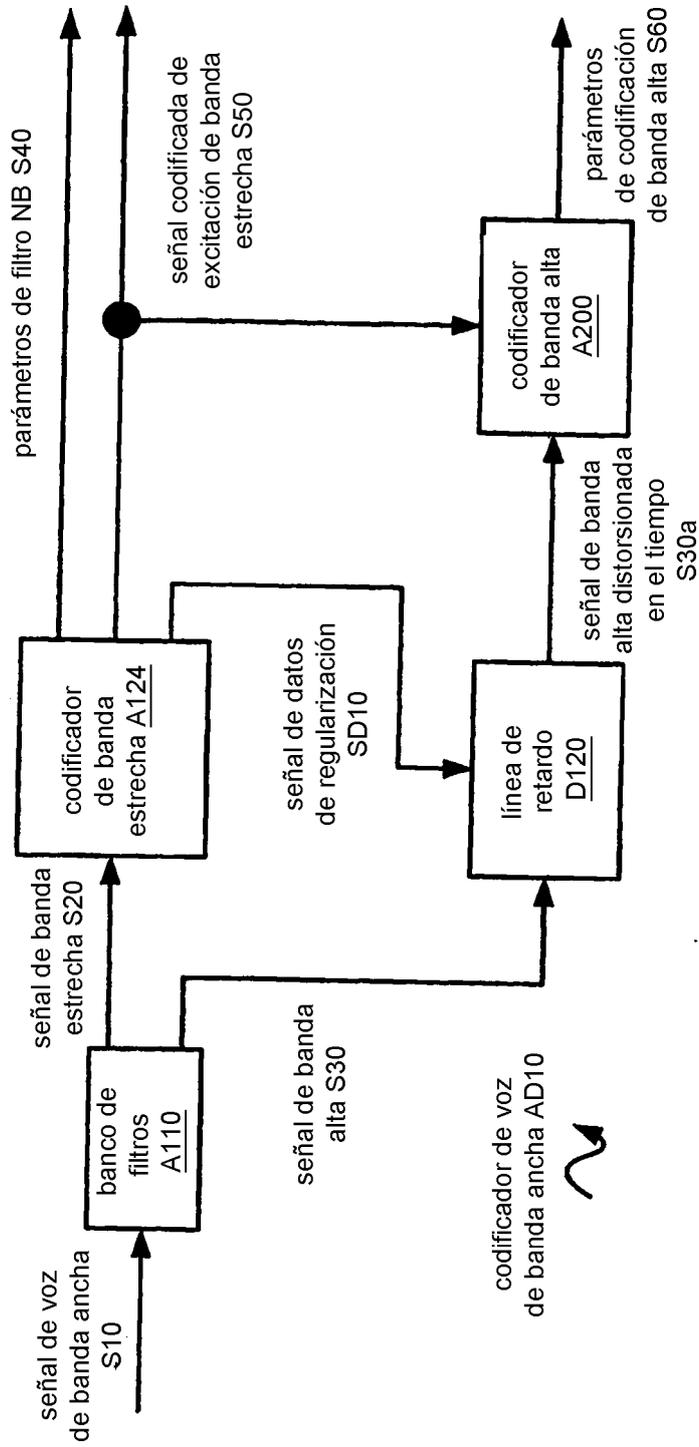
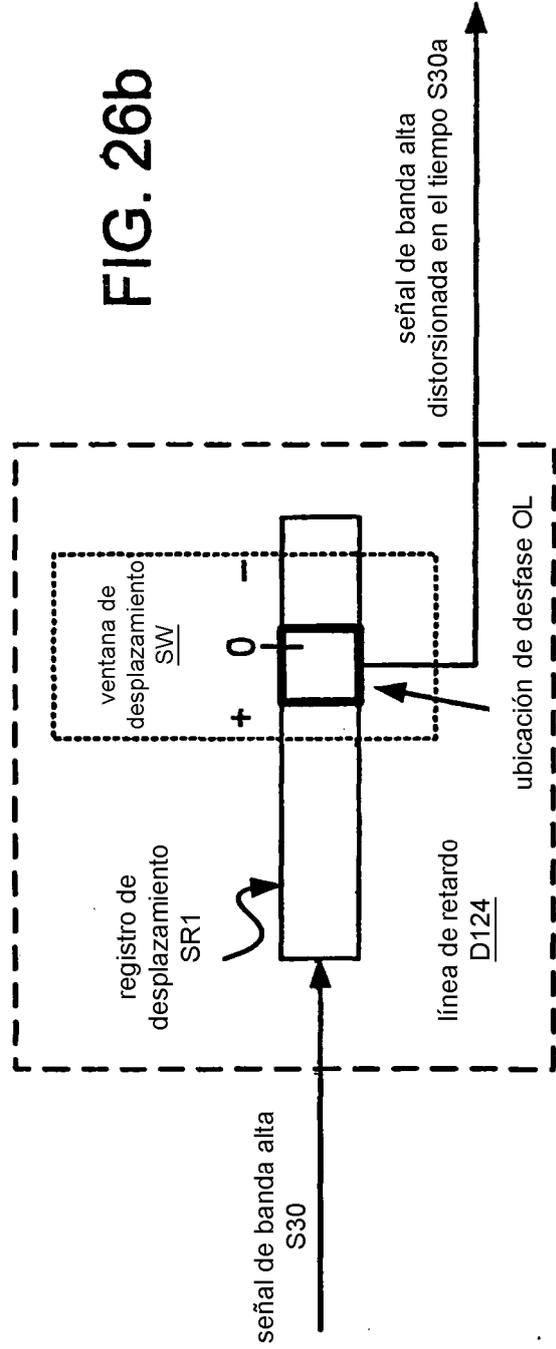
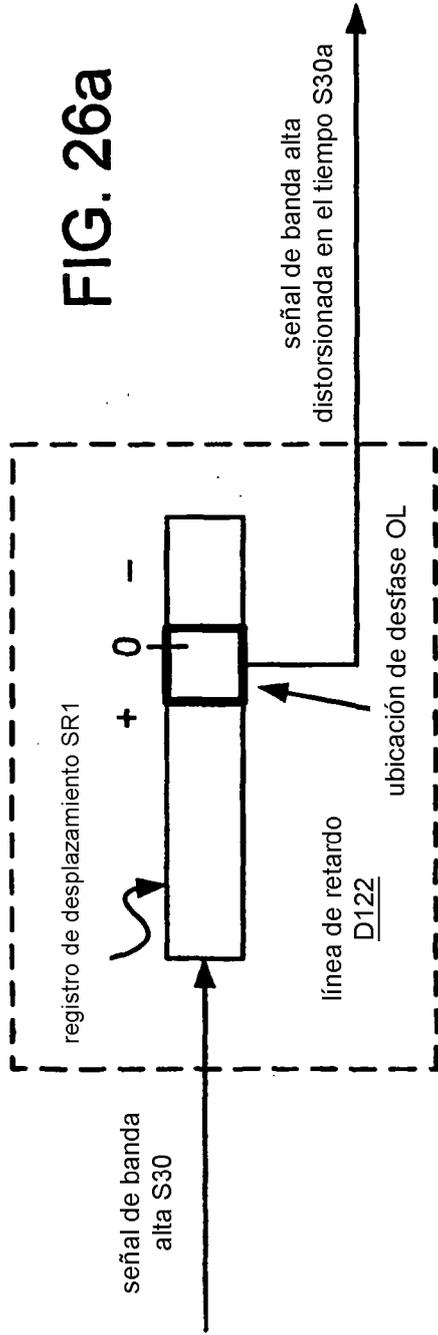


FIG. 25



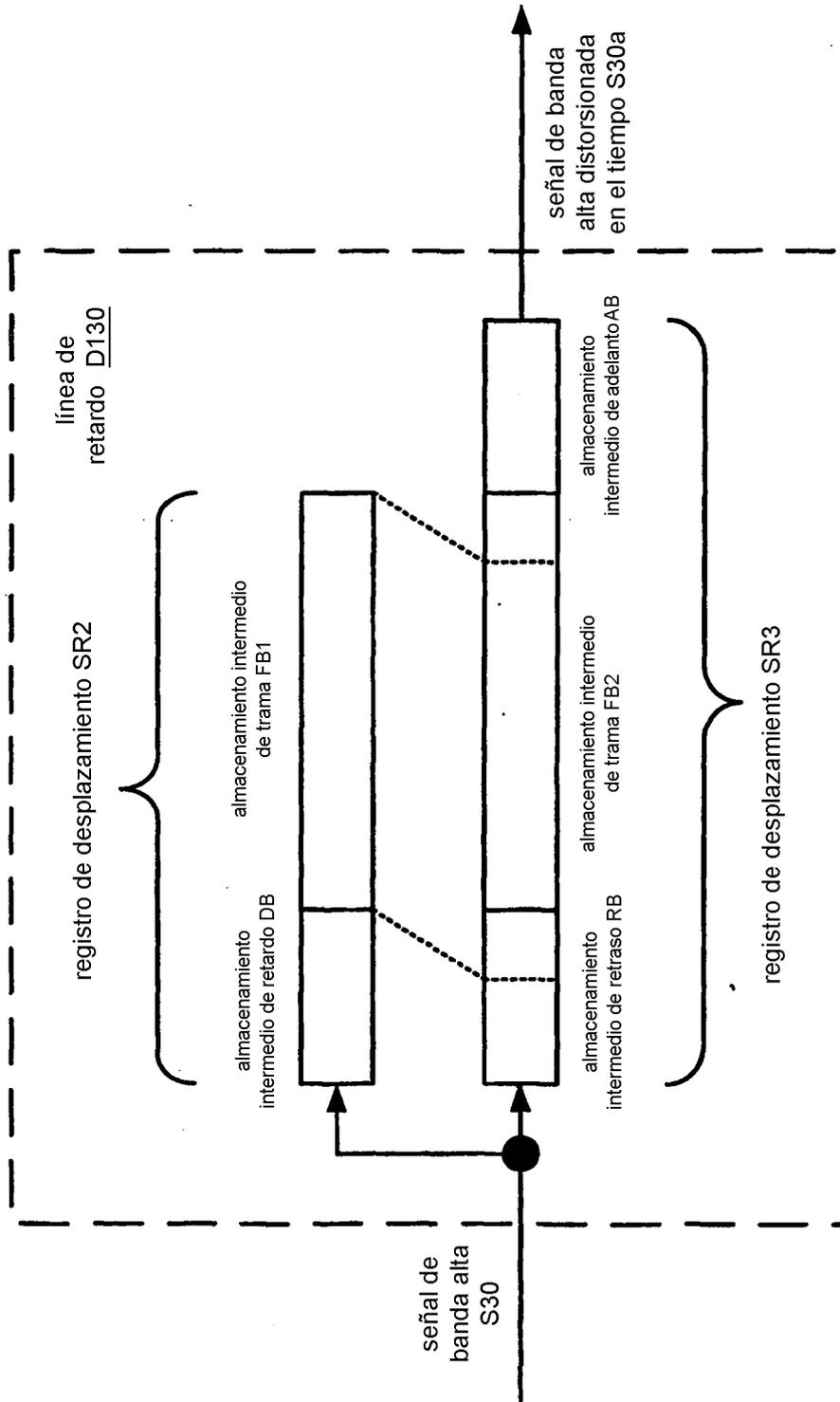


FIG. 27

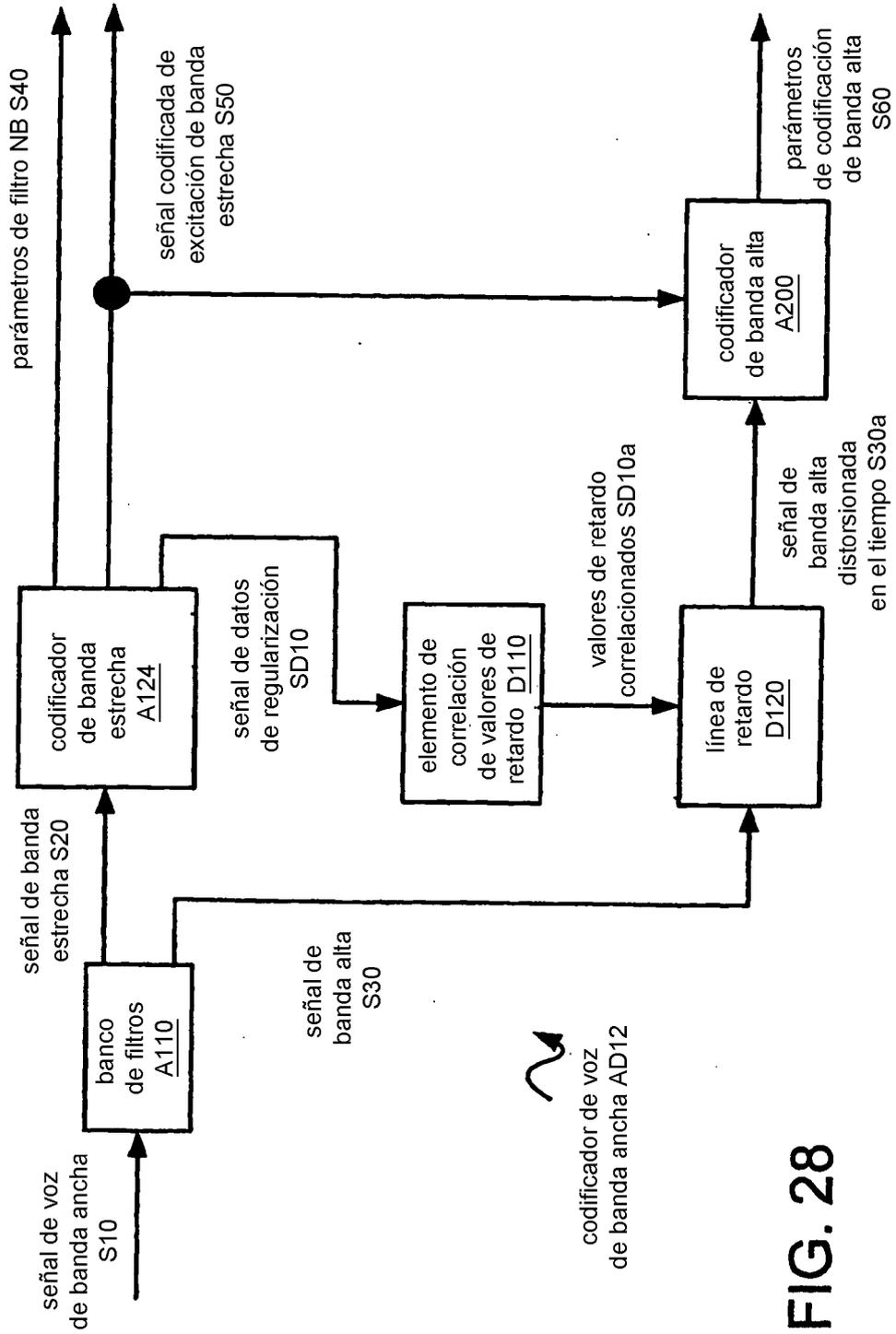


FIG. 28

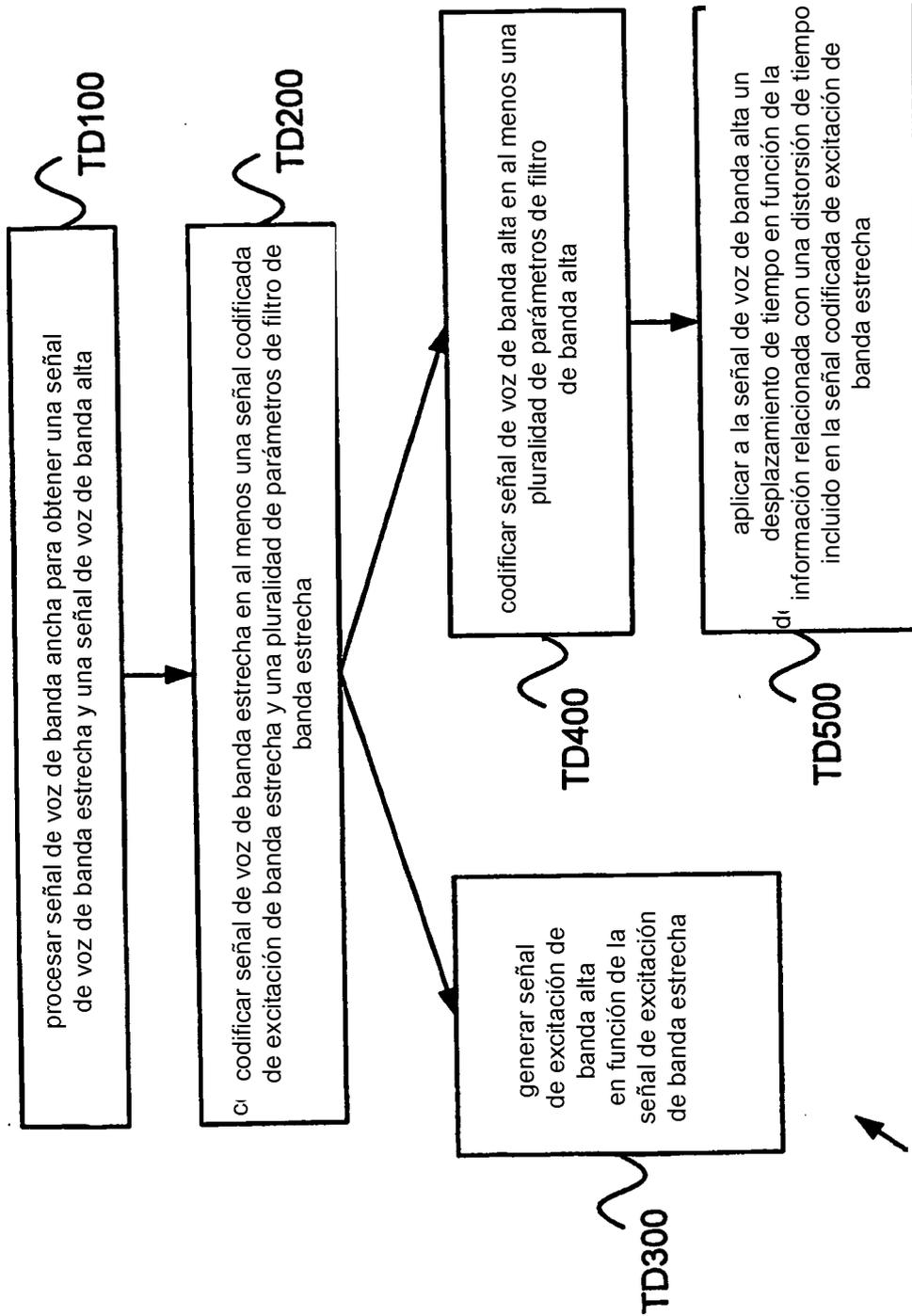


FIG. 29

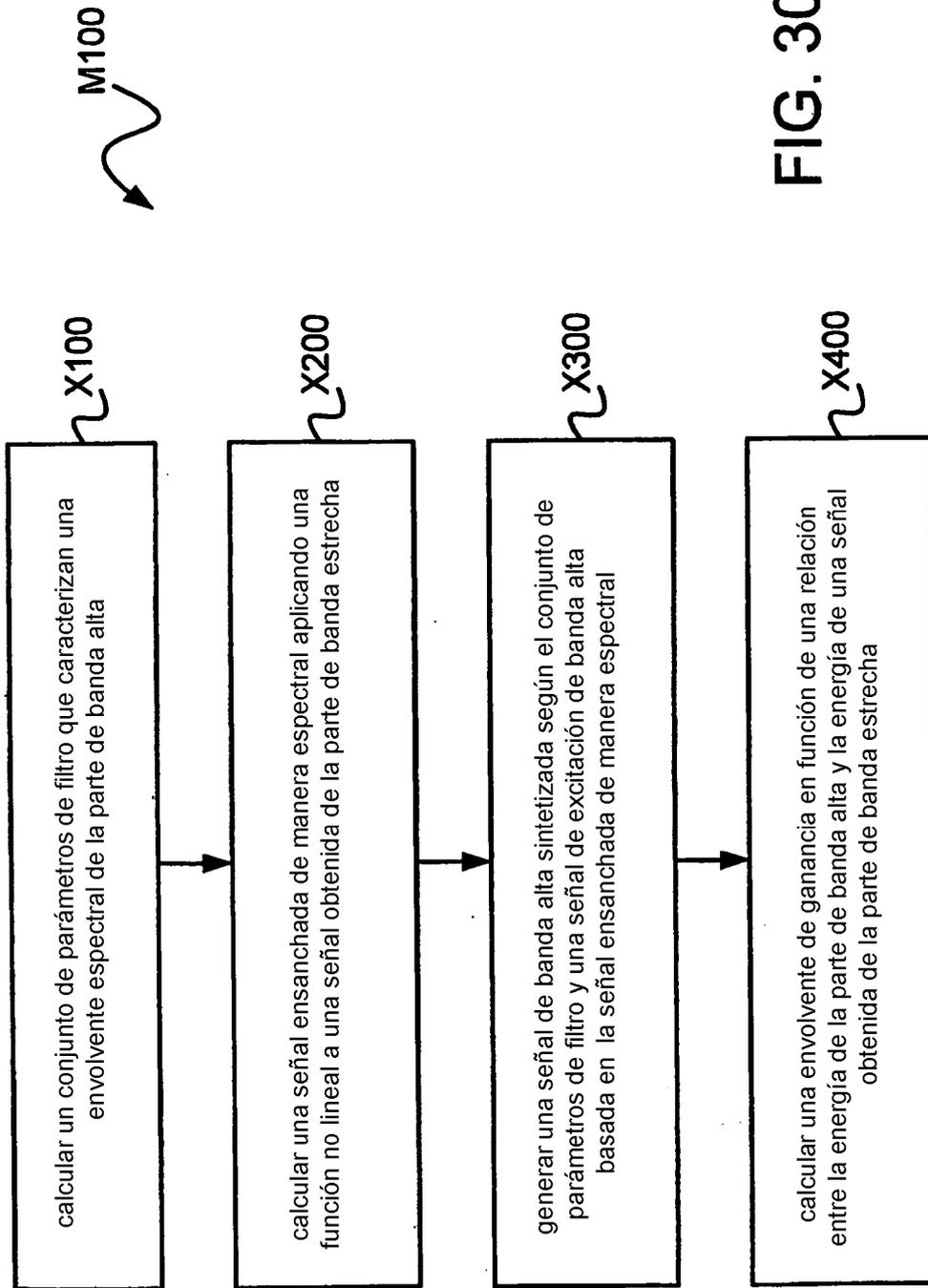


FIG. 30

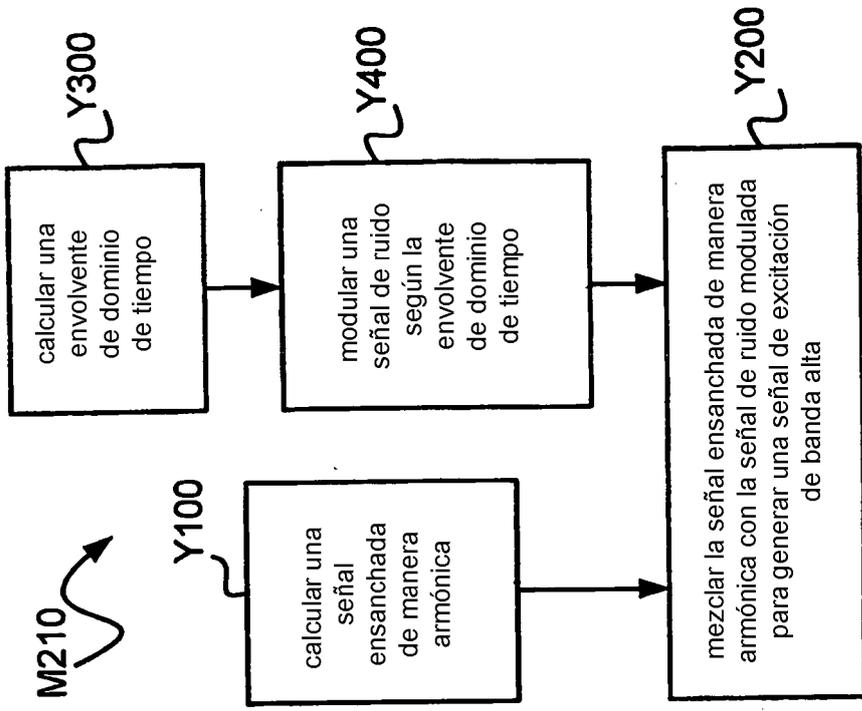


FIG. 31b

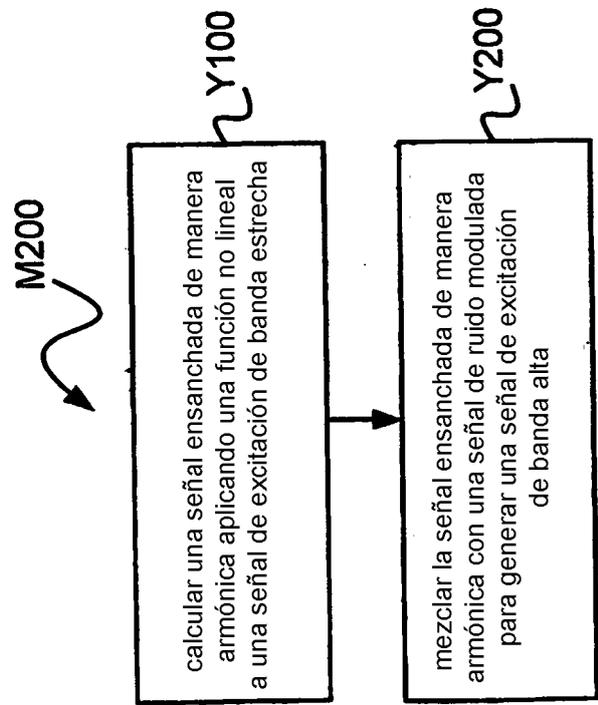


FIG. 31a

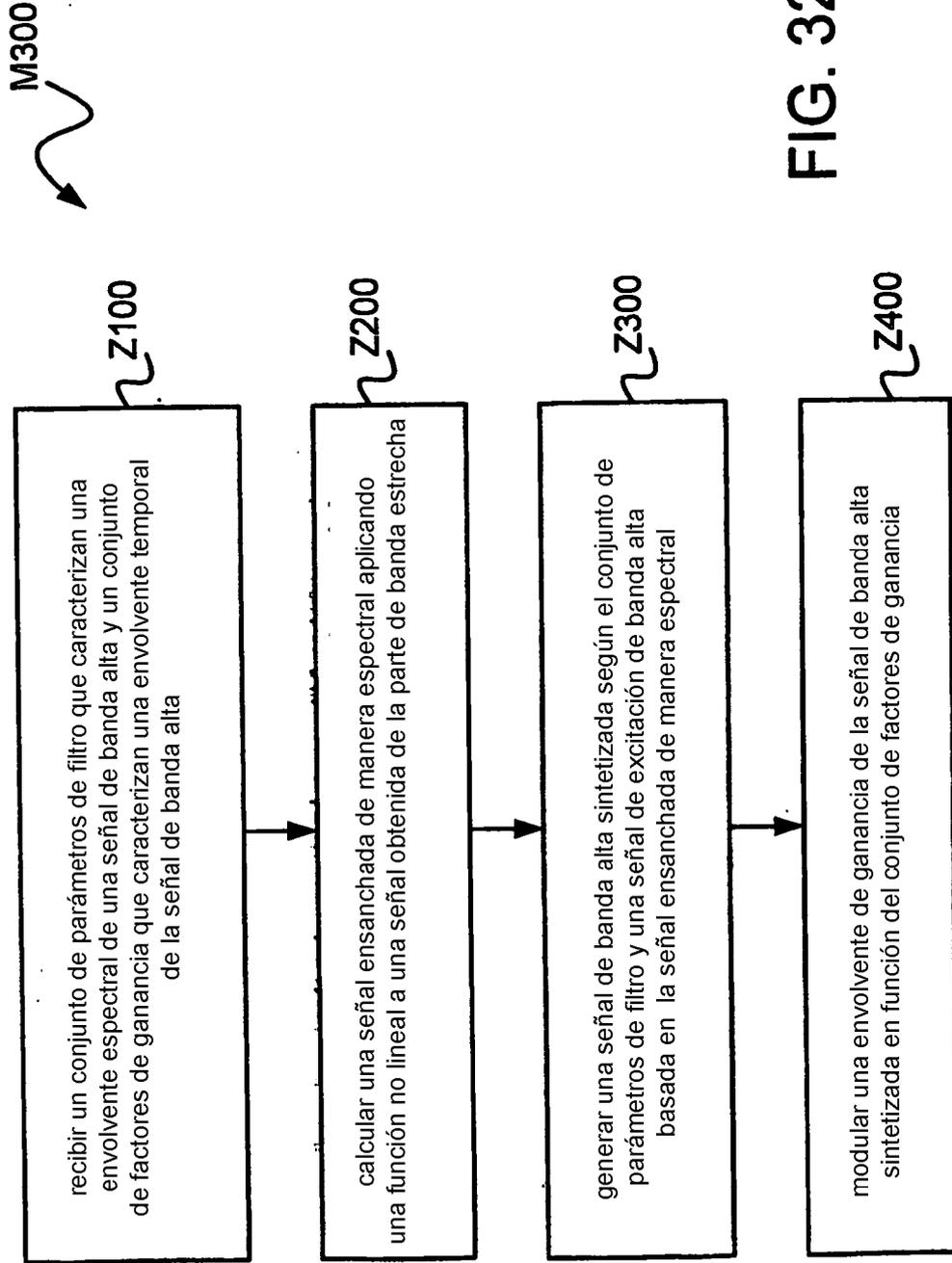


FIG. 32