

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 569 912**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/18 (2013.01)

G10L 19/16 (2013.01)

G10L 19/02 (2013.01)

G10L 19/00 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.06.2009 E 09793874 (0)**

97 Fecha y número de publicación de la concesión europea: **23.03.2016 EP 2301023**

54 Título: **Esquema de codificación/decodificación de audio a baja tasa de bits y conmutadores en cascada**

30 Prioridad:

11.07.2008 US 79854

08.10.2008 EP 08017663

18.02.2009 EP 09002271

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

13.05.2016

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (50.0%)
Hansastrasse 27c
80686 München, DE y
VOICEAGE CORPORATION (50.0%)**

72 Inventor/es:

**GRILL, BERNHARD;
LEFEBVRE, ROCH;
BESSETTE, BRUNO;
LAPIERRE, JIMMY;
GOURNAY, PHILIPPE;
SALAMI, REDWAN;
BAYER, STEFAN;
FUCHS, GUILLAUME;
GEYERSBERGER, STEFAN;
GEIGER, RALF;
HILPERT, JOHANNES;
KRAEMER, ULRICH;
LECOMTE, JEREMIE;
MULTRUS, MARKUS;
NEUENDORF, MAX;
POPP, HARALD y
RETTELBACH, NIKOLAUS**

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 569 912 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).

Esquema de codificación/decodificación de audio a baja tasa de bits y conmutadores en cascada

DESCRIPCIÓN

5 Campo de la invención

La presente invención se refiere a la codificación de audio y, en particular, a esquemas de codificación de audio a baja tasa de bits.

10 Antecedentes de la invención y técnica anterior

En la técnica, se conocen los esquemas de codificación de dominio de frecuencia tales como MP3 o AAC (sigla en inglés correspondiente a la traducción codificación de audio avanzada). Estos codificadores de dominio de frecuencia se basan en una conversión de dominio del tiempo/dominio de frecuencia, una etapa de cuantificación posterior, en la cual el error de cuantificación se controla utilizando información desde un módulo psico-acústico, y una etapa de codificación, en la cual los coeficientes espectrales cuantificados y la información secundaria correspondiente se codifican por entropía utilizando tablas de códigos.

Por otro lado, existen codificadores que son muy apropiados para el procesamiento de voz tal como AMR-WB+ (sigla en inglés correspondiente a la traducción velocidad múltiple adaptativa - banda ancha) como se describe en 3GPP TS 26.290. Dichos esquemas de codificación de voz realizan un filtrado Predictivo Lineal de una señal de dominio del tiempo. Dicho filtrado PL se obtiene de un análisis de Predicción Lineal de la señal de entrada de dominio del tiempo. Los coeficientes del filtro Predictivo Lineal resultantes a continuación se cuantifican/codifican y se transmiten como información secundaria. El proceso se conoce como Codificación de Predicción Lineal (LPC, por su sigla en inglés). En la salida del filtro, la señal residual de predicción o señal de error de predicción que también se conoce como la señal de excitación se codifica utilizando las etapas de análisis-por-síntesis del codificador ACELP (sigla en inglés que corresponde a la traducción predicción lineal con excitación por código algebraico) o, como alternativa, se codifica utilizando un codificador de transformación, que utiliza una transformada de Fourier con un solapamiento. La decisión entre la codificación ACELP y la codificación de excitación Codificada de Transformación que también se denomina codificación TCX (por su sigla en inglés) se realiza utilizando un algoritmo de bucle cerrado o bucle abierto.

Los esquemas de codificación de audio en dominio de frecuencia tal como el esquema de codificación de alta eficiencia AAC (sigla en inglés que corresponde a la traducción codificador de audio avanzado), que combina un esquema de codificación AAC y una técnica de replicación de banda espectral pueden combinarse también con una herramienta de codificación de estero conjunto o multi-canal que se conoce bajo el término "MPEG surround" (sigla en inglés que corresponde a la traducción Grupo de Expertos en Imágenes en Movimiento).

Por otro lado, los codificadores de voz tales como AMR-WB+ (sigla en inglés que corresponde a la traducción múltiple velocidad adaptativa-banda ancha) también tienen una etapa de mejora de la frecuencia alta y una funcionalidad estéreo.

Los esquemas de codificación de dominio de frecuencia son ventajosos ya que muestran una alta calidad a bajas tasas de bits para señales musicales. Sin embargo, resulta problemático la calidad de señales del habla a bajas tasas de bits.

Los esquemas de codificación del habla muestran una alta calidad para señales del habla incluso a bajas tasas de bits, pero muestran una calidad pobre para señales musicales a bajas tasas de bits.

El documento WO 2008/071353 A2 y RAMPRASHAD, S. A. The Multimode Transform Predictive Coding Paradigm. IEEE Transactions on Speech and Audio Processing, Vol. 11, N.º 2, marzo de 2003, páginas 117-129, desvela ejemplos de esquemas híbridos para codificar y decodificar señales del audio.

Sumario de la Invención

Un objetivo de la presente invención es proporcionar una coordinación mejorada de codificación/decodificación.

Este objetivo se logra por medio de un codificador de audio como se expone la reivindicación 1, un método de codificación de audio como se expone en la reivindicación 15, un decodificador como se expone en la reivindicación 16, un método de decodificación como se expone en la reivindicación 23, una señal codificada como se expone en la reivindicación 24 o un programa informático como se expone en la reivindicación 25.

Se exponen realizaciones preferidas en las reivindicaciones dependientes.

Breve descripción de los dibujos

5 A continuación se describen realizaciones preferidas de la presente invención con respecto a los dibujos adjuntos, en los cuales:

- La Fig. 1a es un diagrama de bloques de un esquema de codificación de acuerdo con un primer aspecto de la presente invención;
- 10 La Fig. 1b es un diagrama de bloques de un esquema de decodificación de acuerdo con el primer aspecto de la presente invención;
- La Fig. 1c es un diagrama de bloques de un esquema de codificación de acuerdo con un aspecto adicional de la presente invención;
- 15 La Fig. 2a es un diagrama de bloques de un esquema de codificación de acuerdo con un segundo aspecto de la presente invención;
- La Fig. 2b es un diagrama esquemático de un esquema de decodificación de acuerdo con el segundo aspecto de la presente invención.
- 20 La Fig. 2c es un diagrama de bloques de un esquema de codificación de acuerdo con un aspecto adicional de la presente invención
- 25 La Fig. 3a ilustra un diagrama de bloques de un esquema de codificación de acuerdo con un aspecto adicional de la presente invención;
- La Fig. 3b ilustra un diagrama de bloques de un esquema de decodificación de acuerdo con el aspecto adicional de la presente invención;
- 30 La Fig. 3c ilustra una representación esquemática del aparato/método de codificación con conmutadores en cascada;
- La Fig. 3d ilustra un diagrama esquemático de un aparato o método de decodificación, en el cual se utilizan combinadores en cascada;
- 35 La Fig. 3e ilustra una ilustración de una señal de dominio del tiempo y una representación correspondiente de la señal codificada que ilustra regiones de transición gradual cortas que se incluyen en ambas señales codificadas;
- 40 La Fig. 4a ilustra un diagrama de bloques con un conmutador situado antes de las ramificaciones de codificación;
- La Fig. 4b ilustra un diagrama de bloques de un esquema de codificación con el conmutador situado después de la codificación de las ramificaciones;
- 45 La Fig. 4c ilustra un diagrama de bloques para una realización de combinador preferido;
- La Fig. 5a ilustra una forma de onda de un segmento del habla de dominio del tiempo como un segmento cuasi-periódico o señal del tipo impulso;
- 50 La Fig. 5b ilustra un espectro del segmento de la Fig. 5a;
- La Fig. 5c ilustra un segmento de voz de dominio del tiempo de un habla no vocalizada como ejemplo para un segmento del tipo ruido;
- 55 La Fig. 5d ilustra un espectro de la forma de onda de dominio del tiempo de la Fig. 5c;
- La Fig. 6 ilustra un diagrama de bloques de un análisis por medio de un codificador de análisis por síntesis CELP;
- 60 Las Figs. 7a a 7d ilustran señales de excitación vocalizadas/no vocalizadas como ejemplo para señales del tipo impulso;

- La Fig. 7e ilustra una etapa del lado del codificador LPC que proporciona información de predicción a corto plazo y la señal del error de predicción (excitación);
- 5 La Fig. 7f ilustra una realización adicional de un dispositivo LPC para generar una señal ponderada;
- La Fig. 7g ilustra una implementación para transformar una señal ponderada en una señal de excitación aplicando una operación de ponderación inversa y un análisis de excitación posterior como se requiere en el convertidor 537 de la Fig. 2b;
- 10 La Fig. 8 ilustra un diagrama de bloques de un algoritmo multi-canal conjunto de acuerdo con una realización de la presente invención;
- La Fig. 9 ilustra una realización preferida de un algoritmo de extensión de ancho de banda;
- 15 La Fig. 10a ilustra una descripción detallada del conmutador cuando se realiza una decisión de bucle abierto; y
- La Fig. 10b ilustra una ilustración del conmutador cuando se opera en un modo de decisión de bucle cerrado.
- 20

Descripción detallada de las realizaciones preferidas

25 La Fig. 1a ilustra una realización de la invención con dos conmutadores en cascada. Una señal mono, una señal estéreo o una señal multi-canal se introduce a un conmutador 200. El conmutador 200 se controla por una etapa de decisión 300. La etapa de decisión recibe, como entrada, una entrada de señal en el bloque 200. Como alternativa, la etapa de decisión 300 puede también recibir una información secundaria que se incluye en la señal mono, la señal estéreo o señal multi-canal o al menos está asociada con una señal de este tipo, donde existe la información, la cual se generó, por ejemplo, cuando se produce originalmente la señal mono, la señal estéreo o la señal multi-canal.

30

La etapa de decisión 300 activa el conmutador 200 con el fin de suministrar una señal en una porción de codificación de frecuencia 400 ilustrada en una ramificación superior de la Fig. 1a o una porción de codificación de dominio LPC-500 ilustrada en una ramificación inferior en la Fig. 1a. Un elemento clave de la ramificación de codificación de dominio de frecuencia es un bloque de conversión espectral 410 el cual es operativo para convertir una señal de salida de etapa de pre-procesamiento común (como se analiza más adelante) en un dominio espectral. El bloque de conversión espectral puede incluir un algoritmo MDCT, un algoritmo QMF, un análisis de Ondícula o un banco de filtro tal como un banco de filtro muestreado en forma crítica que tiene una cierta cantidad de canales de banco de filtro, donde las señales de sub-bandas en este banco de filtro pueden ser señales con valor real o señales con valor complejo. La salida del bloque de conversión espectral 410 se codifica utilizando un codificador de audio espectral 421, el cual puede incluir bloques de procesamiento como se conocen a partir del esquema de codificación AAC.

35

40

Por lo general, el procesamiento en la ramificación 400 es un procesamiento en un modelo basándose en la percepción o modelo de sumidero de información. De esta manera, estas ramificaciones modelan el sistema auditivo humano que recibe el sonido. Contrario a ello, el procesamiento en la ramificación 500 debe generar una señal en el dominio de excitación, residual o LPC. Por lo general, el procesamiento en la ramificación 500 es un procesamiento en un modelo del habla o modelo de generación de información. Para señales del habla, este modelo es un modelo del habla humana/sistema de generación de sonido. Sin embargo, si un sonido de una fuente diferente que requiera un modelo de generación de sonido diferente se ha de codificar, entonces el procesamiento en la ramificación 500 puede ser diferente.

45

50

En la ramificación de codificación inferior 500, un elemento clave consiste en un dispositivo LPC 510, que emite una información LPC que se usa para controlar las características de un filtro LPC. Esta información LPC se transmite a un decodificador. La señal de salida 510 de la etapa LPC es una señal de dominio LPC que consiste en una señal de excitación y/o una señal ponderada.

55

El dispositivo LPC generalmente emite una señal de dominio LPC, la cual puede ser cualquier señal en el dominio LPC tal como la señal de excitación en la Fig. 7e o una señal ponderada en la Fig. 7f u otra señal, que se ha generado aplicando coeficientes de filtro LPC a una señal de audio. Adicionalmente, un dispositivo LPC puede también determinar estos coeficientes y puede además cuantificar/codificar estos coeficientes.

60

La decisión en la etapa de decisión puede ser una señal adaptable para que la etapa de decisión realice una discriminación de música/ voz y controle el conmutador 200 de tal modo que las señales musicales se introduzcan

en la ramificación superior 400, y las señales del habla se introduzcan en la ramificación inferior 500. En una realización, la etapa de decisión alimenta su información de decisión en una corriente de bits de salida para que un decodificador pueda utilizar esta información de decisión con el fin de llevar a cabo las operaciones de decodificación correctas.

5 Un decodificador de este tipo se ilustra en la Fig. 1b. La salida de la señal por el codificador de audio espectral 421, después de la transmisión, se introduce en un decodificador de audio espectral 431. La salida de señal del decodificador de audio espectral 431 se introduce en un convertidor de dominio de tiempo 440. De forma análoga, la salida de la ramificación de codificación de dominio de LPC 500 de la Fig. 1a se recibe en el lado del decodificador y se procesa por los elementos 531, 533, 534, y 532 para obtener una señal de excitación LPC. La señal de excitación LPC se introduce en una etapa de síntesis LPC 540, que recibe, como una entrada adicional, la información LPC generada por la etapa de análisis LPC correspondiente 510. La salida del convertidor de tiempo-dominio 440 y/o la salida de la etapa de síntesis LPC 540 se introducen en un conmutador 600. El conmutador 600 se controla a través de una señal de control del conmutador que se generó, por ejemplo, por la etapa de decisión 300, o que se proporcionó en forma externa por medio de un creador de la señal mono, señal estéreo o señal multi-canal original. La salida del conmutador 600 es una señal mono, señal estéreo o señal multi-canal completa.

La señal de entrada en el conmutador 200 y la etapa de decisión 300 puede ser una señal mono, señal estéreo o señal multi-canal o en general una señal de audio. Dependiendo de la decisión que puede obtenerse de la señal de entrada del conmutador 200 o de cualquier fuente externa como un productor de la señal de audio original que subyace la señal introducida en la etapa 200, el conmutador conmuta entre la ramificación de codificación de frecuencia 400 y la ramificación de codificación de frecuencia LPC 500. La ramificación de codificación de frecuencia 400 comprende una etapa de conversión espectral 410 y una etapa de cuantificación/codificación posteriormente conectada 421. La etapa de cuantificación/codificación puede incluir cualquiera de las funcionalidades conocidas de los codificadores de dominio de frecuencia modernos como el codificador AAC. Además, la operación de cuantificación en la etapa de cuantificación/codificación 421 puede controlarse a través de un módulo psicoacústico que genera información psicoacústica tal como un valor umbral del enmascaramiento psicoacústico sobre la frecuencia, donde esta información se introduce en la etapa 421.

En la ramificación de codificación LPC, la señal de salida del conmutador se procesa a través de una etapa de análisis LPC 510 que genera información del lado LPC y una señal de dominio LPC. El codificador de excitación de manera inventiva comprende un conmutador adicional para conmutar el procesamiento adicional de la señal de dominio LPC entre una operación de cuantificación/codificación 522 en el dominio LPC o una etapa de cuantificación/codificación 524, que procesa valores en el dominio espectral LPC. Para este fin, un convertidor espectral 523 se proporciona en la entrada de la etapa de cuantificación/codificación 524. El conmutador 521 se controla en un modo de bucle abierto o un modo de bucle cerrado dependiendo de las configuraciones específicas como se describen, por ejemplo, en la especificación técnica AMR-WB+.

Para el modo de control de bucle cerrado, el codificador además incluye un cuantificador/codificador inverso 531 para la señal de dominio LPC, un cuantificador/codificador inverso 533 para la señal de dominio espectral LPC y un convertidor espectral inverso 534 para la salida del elemento 533. Ambas señales codificadas y nuevamente decodificadas en las ramificaciones de procesamiento de la segunda ramificación de codificación se introducen en el dispositivo de control del conmutador 525. En el dispositivo de control del conmutador 525, estas dos señales de salida se comparan entre sí y/o con una función objetivo o se calcula una función objetivo la cual puede basarse en una comparación de la distorsión en ambas señales para que la señal que tenga la menor distorsión se use para decidir, que posición debería tomar el conmutador 521. De manera alternativa, en caso de que ambas ramificaciones provean tasas de bits no-constantas, la ramificación que proporcione la tasa de bits inferior podría seleccionarse incluso cuando la relación señal-ruido de esta ramificación es menor que la relación señal-ruido de la otra ramificación. De manera alternativa, la función objetivo podría utilizar, como una entrada, la relación señal-ruido de cada señal y una tasa de bits de cada señal y/o criterios adicionales con el fin de encontrar la mejor decisión para un objetivo específico. Si, por ejemplo, el objetivo es tal que la tasa de bits debería ser lo más baja posible, entonces la función objetivo se basaría en gran medida en la tasa de bits de las dos señales emitidas por los elementos 531, 534. Sin embargo, cuando el objetivo principal es obtener la mejor calidad para una cierta tasa de bits, entonces el control del conmutador 525 podría descartar, por ejemplo, cada señal que se encuentre sobre la tasa de bits permitida y cuando ambas señales se encuentran por debajo de la tasa de bits permitida, el control del conmutador seleccionaría la señal que tenga la mejor relación señal-ruido, es decir, que tenga la menores distorsiones de cuantificación/codificación.

El esquema de codificación de acuerdo con la presente invención, como se ha establecido previamente, se ilustra en la Fig. 1b. Para cada tipo de las tres posibles señales de salida, existe una etapa específica de decodificación/re-cuantificación 431, 531 o 533. Mientras que la etapa 431 emite un espectro de tiempo que se convierte en dominio de tiempo utilizando el convertidor frecuencia/tiempo 440, la etapa 531 emite una señal de dominio LPC, y el elemento 533 emite un espectro LPC. Con el fin de asegurarse que las señales de entrada en el conmutador 532 se

encuentran ambas en el dominio LPC, se proporciona el espectro LPC/convertidor LPC 534. Los datos de salida del conmutador 532 se transforman nuevamente en el dominio de tiempo utilizando una etapa de síntesis LPC 540, que se controla a través de información LPC generada y transmitida mediante el lado del codificador. A continuación, después del bloque 540, ambas ramificaciones tienen información de dominio de tiempo que se conmuta de acuerdo con una señal de control de conmutación con el fin de obtener finalmente una señal de audio tal como una señal mono, señal estéreo, o señal multi-canal, que depende de la señal introducida en el esquema de codificación de la Fig. 1a.

La Fig. 1c ilustra una realización adicional con una disposición diferente del conmutador 521 similar al principio de la Fig. 4b.

La Fig. 2a ilustra un esquema de codificación preferido de acuerdo con un segundo aspecto de la invención. Un esquema de pre-procesamiento común conectado a la entrada del conmutador 200 puede comprender un bloque estéreo envolvente/conjunto 101 el cual genera, como una salida, parámetros de estéreo conjunto y una señal de salida mono, que se genera al mezclar en forma descendente la señal de entrada que es una señal que tiene dos o más canales. Generalmente, la señal en la salida del bloque 101 puede también ser una señal que tiene más canales, pero debido a la funcionalidad de mezcla descendente del bloque 101, la cantidad de canales en la salida del bloque 101 será menor a la cantidad de canales introducidos en el bloque 101.

El esquema de pre-procesamiento común puede comprender como alternativa al bloque 101 o además del bloque 101 una etapa de extensión de ancho de banda 102. En la realización de la Fig. 2a, la salida del bloque 101 se introduce al bloque de extensión de ancho de banda 102 que, en el codificador de la Fig. 2a, emite una señal de banda limitada como la señal de banda baja o señal de paso bajo en su salida. Preferentemente, esta señal se somete también a submuestreo (por ejemplo por un factor de dos). Adicionalmente, para la banda alta de la señal introducida en el bloque 102, los parámetros de extensión de ancho de banda tales como los parámetros de envolvente espectral, parámetros de filtrado inverso, parámetros de ruido de fondo etc., como se conocen del perfil HE-AAC de MPEG-4 se generan y reenvían a un multiplexor de secuencia de bits 800.

Preferentemente, la etapa de decisión 300 recibe la señal introducida en el bloque 101 o introducida en el bloque 102 con el fin de decidir entre, por ejemplo, un modo de música o modo del habla. En el modo de música, se selecciona la ramificación de codificación superior 400, mientras que en el modo del habla, se selecciona la ramificación de codificación más baja 500. Preferentemente, la etapa de decisión además controla el bloque de estéreo conjunto 101 y/o el bloque de extensión de ancho de banda 102 para adaptar la funcionalidad de estos bloques a la señal específica. Por lo tanto, cuando la etapa de decisión determina que cierta porción de tiempo de la señal de entrada es del primer modo tal como el modo de música, se pueden controlar las características específicas del bloque 101 y/o bloque 102 mediante la etapa de decisión 300. Como alternativa, cuando la etapa de decisión 300 determina que la señal se encuentra en un modo del habla o, generalmente, en un segundo modo de dominio LPC, se pueden controlar las características específicas de los bloques 101 y 102 de acuerdo con la salida de la etapa de decisión.

Preferentemente, la conversión espectral de la ramificación de codificación 400 se realiza utilizando una operación MDCT (Transformada de coseno discreta modificada) la cual, incluso más preferentemente, es la operación MDCT comprimida en el tiempo, donde la intensidad o, generalmente, la fuerza de compresión puede controlarse entre cero y una alta intensidad de compresión. En una intensidad de compresión cero, la operación MDCT en el bloque 411 es una operación MDCT sencilla conocida en la técnica. La intensidad de compresión en el tiempo junto con la información secundaria de compresión en el tiempo pueden transmitirse/introducirse en el multiplexor de secuencia de bits 800 como información secundaria.

En la ramificación de codificación LPC, el codificador de dominio LPC puede incluir un núcleo ACELP 526 que calcula una ganancia de tono, un retardo de tono y/o información del libro de códigos tal como un índice y ganancia del libro de códigos. El modo TCX como se conoce a partir de 3GPP TS 26.290 incurre en un procesamiento de una señal perceptualmente ponderada en el dominio de transformación. Una señal ponderada transformada de Fourier se cuantifica utilizando una cuantificación entramada multi-tasa dividida (algebraica VQ) con cuantificación de factor ruido. Una transformación se calcula en 1024, 512, o 256 ventanas de muestra. La señal de excitación se recupera por filtrado inverso de la señal ponderada cuantificada a través de un filtro de ponderación inverso. En la primera ramificación de codificación 400, un convertidor espectral preferentemente comprende una operación MDCT específicamente adaptada que tiene ciertas funciones de ventana seguidas por una etapa de codificación de cuantificación/entropía que puede consistir en una etapa de cuantificación de un solo vector, pero preferentemente es un codificador de cuantificación/entropía escalar combinado similar al cuantificador/codificador en la ramificación de codificación de dominio de frecuencia, es decir, en el elemento 421 de la Fig. 2a.

En la segunda ramificación de codificación, se encuentra el bloque LPC 510 seguido por un conmutador 521, nuevamente seguido de un bloque ACELP 526 o un bloque TCX 527. ACELP se describe en 3GPP TS 26,190 y TCX se describe en 3GPP TS 26,290. Generalmente, el bloque ACELP 526 recibe una señal de excitación LPC

como se calcula por el procedimiento como se describe en la Fig. 7e. El bloque TCX 527 recibe una señal ponderada como se genera mediante la Fig. 7f.

5 En TCX, la transformación se aplica a la señal ponderada calculada filtrando la señal de entrada a través de un filtro de ponderación basándose en LPC. El filtro de ponderación utilizado en las realizaciones preferidas de la invención se proporciona por $(1-A(z/\gamma))/(1-\mu z^{-1})$. En consecuencia, la señal ponderada es una señal de dominio LPC y su transformación es un dominio espectral LPC. La señal procesada por el bloque ACELP 526 es la señal de excitación y es diferente de la señal procesada por el bloque 527, pero ambas señales se encuentran en el dominio LPC.

10 En el lado del decodificador ilustrado en la Fig. 2b, después de la transformación espectral inversa en el bloque 537, se aplica la inversa del filtro de ponderación, es decir $(1-\mu z^{-1})/(1-A(z/\gamma))$. A continuación, la señal se filtra a través de $(1-A(z))$ para dirigirse al dominio de excitación LPC. De esta manera, la conversión al bloque de dominio LPC 534 y el bloque TCX¹ 537 incluye la transformación inversa y a continuación el filtrado a través de $\frac{(1-\mu z^{-1})}{(1-A(z/\gamma))} (1-A(z))$ para convertir desde dominio ponderado al dominio de excitación.

15 Aunque el elemento 510 en las Figs. 1a, 1c, 2a, 2c ilustra un solo bloque, el bloque 510 puede emitir diferentes señales siempre y cuando estas señales se encuentren en el dominio LPC. El modo real del bloque 510 tal como el modo de señal de excitación o el modo de señal ponderada puede depender del estado real del conmutador. Como alternativa, el bloque 510 puede tener dos dispositivos de procesamiento paralelos, donde un dispositivo se implementa de manera similar a la Fig. 7e y el otro dispositivo se implementa como la Fig. 7f. Por lo tanto, el dominio LPC en la salida de 510 puede representar la señal de excitación LPC o la señal ponderada LPC o cualquier otra señal de dominio LPC.

20 En la segunda ramificación de codificación (ACELP/TCX) de la Fig. 2a o 2c, la señal preferentemente se pre-enfatiza a través de un filtro $1-0,68z^{-1}$ antes de la codificación. En el decodificador ACELP/TCX en la Fig. 2b la señal sintetizada se desenfatiza con el filtro $1/(1-0,68z^{-1})$. La pre-énfasis puede ser parte del bloque LPC 510 donde la señal se pre-enfatiza antes del análisis y cuantificación LPC. De manera similar, la desenfatización puede ser parte del bloque de síntesis LPC-1 540.

30 La Fig. 2c ilustra una realización adicional para la implementación de la Fig. 2a, pero con una disposición diferente del conmutador 521 similar al principio de la Fig. 4b.

35 En una realización preferida, el primer conmutador 200 (véase Fig. 1a o 2a) se controla a través de una decisión de bucle abierto (como en la Fig. 4a) y el segundo conmutador se controla a través de una decisión de bucle cerrado (como en la figura 4b).

40 Por ejemplo, la Fig. 2c, tiene el segundo conmutador situado después de las ramificaciones ACELP y TCX como en la Fig. 4b. A continuación, en la primera ramificación de procesamiento, el primer dominio LPC representa la excitación LPC, y en la segunda ramificación de procesamiento, el segundo dominio LPC representa la señal ponderada LPC. Es decir, la primera señal de dominio LPC se obtiene por el filtrado a través de $(1-A(z))$ para convertir al dominio residual LPC, mientras que el segundo dominio LPC se obtiene a través del filtrado del filtro $(1-A(z/\gamma))/(1-\mu z^{-1})$ para convertir al dominio ponderado LPC.

45 La Fig. 2b ilustra un esquema de decodificación correspondiente al esquema de codificación de la Fig. 2a. La secuencia de bits generada por un multiplexor de secuencia de bits 800 de la Fig. 2a se introduce en un demultiplexor de secuencia de bits 900. Dependiendo de una información obtenida por ejemplo de una secuencia de bits a través de un bloque de detección de modo 601, un conmutador del lado del decodificador 600 se controla para reenviar señales desde la ramificación superior o señales desde la ramificación inferior al bloque de extensión de ancho de banda 701. El bloque de extensión de ancho de banda 701 recibe, desde el demultiplexor de secuencia de bits 900, información secundaria y, basándose en esta información secundaria y la salida de la decisión de modo 601, reconstruye la banda alta basándose en la salida de banda baja por el conmutador 600.

50 La señal de banda completa generada por el bloque 701 se introduce en la etapa de procesamiento conjunta estéreo/envolvente 702, la cual reconstruye dos canales estéreo o varios multi-canales. Generalmente, el bloque 702 emitirá más canales de los que se introdujeron en este bloque. Dependiendo de la aplicación, la entrada en el bloque 702 puede incluir dos canales tales como en el modo estéreo y puede incluir más canales siempre que la salida por este bloque tenga más canales que la entrada en este bloque.

60 Se ha observado que el conmutador 200 conmuta entre ambas ramificaciones para que sólo una ramificación reciba una señal para procesar y la otra ramificación no reciba una señal para procesar. En una realización alternativa, sin embargo, el conmutador puede también estar dispuesto en forma posterior a por ejemplo el codificador de audio 421

y al codificador de excitación 522, 523, 524, lo que significa que ambas ramificaciones 400, 500 procesan la misma señal en paralelo. Con el fin de no duplicar la tasa de bits, sin embargo, sólo la señal emitida por una de esas ramificaciones de codificación 400 o 500 se selecciona para ser escrita en la secuencia de bits de salida. La etapa de decisión operará para que la señal escrita en la secuencia de bits minimice una cierta función de coste, donde la función de coste puede ser la tasa de bits generada o la distorsión perceptual generada o una función de coste combinada de tasa/distorsión. En consecuencia, ya sea en este modo o en el modo ilustrado en las Figuras, la etapa de decisión puede también operar en un modo de bucle cerrado con el fin de asegurarse que, finalmente, sólo la salida de ramificación de codificación se escribe en la secuencia de bits la cual tiene una distorsión perceptual dada la tasa de bits más baja o, para una tasa de bits dada, tiene la distorsión perceptual más baja. En el modo de bucle cerrado, la entrada de realimentación puede obtenerse de las salidas de los tres bloques cuantificadores/escaladores 421, 522 y 424 en la Fig. 1a.

En la implementación que tiene dos conmutadores, es decir, el primer conmutador 200 y el segundo conmutador 521, se prefiere que la resolución de tiempo para el primer conmutador sea menor que la resolución de tiempo para el segundo conmutador. Expresado de manera diferente, los bloques de la señal de entrada en el primer conmutador, que pueden conmutarse a través de una operación del conmutador son mayores que los bloques conmutados por el segundo conmutador operando en el dominio LPC. De manera ejemplar, el conmutador del dominio de frecuencia/dominio LPC 200 puede conmutar bloques de una longitud de 1024 muestras, y el segundo conmutador 521 puede conmutar bloques que tienen 256 muestras cada uno.

Aunque algunas de las Figs. 1a a 10b se ilustran como diagramas de bloques de un aparato, estas figuras simultáneamente son una ilustración de un método, donde las funcionalidades del bloque corresponden a las etapas del método.

La Fig. 3a ilustra un codificador de audio para generar una señal de audio codificada que era una salida de la primera ramificación de codificación 400 y una segunda ramificación de codificación 500. Adicionalmente, la señal de audio codificada preferentemente incluye información secundaria tal como parámetros de pre-procesamiento desde la etapa de pre-procesamiento común o, como se analizó en relación con las Figs. previas, información de control del conmutador.

Preferentemente, la primera ramificación de codificación es operativa para codificar una señal intermedia de audio 195 de acuerdo con un primer algoritmo de codificación, donde el primer algoritmo de codificación posee un modelo de sumidero de información. La primera ramificación de codificación 400 genera la primera señal de salida del codificador que es una representación de información espectral codificada de la señal intermedia de audio 195.

Adicionalmente, la segunda ramificación de codificación 500 se adapta para codificar la señal intermedia de audio 195 de acuerdo con un segundo algoritmo de codificación, teniendo el segundo algoritmo de codificación un modelo de fuente de información y que genera, en una segunda señal de salida del codificador, parámetros codificados para el modelo de fuente de información que representa la señal intermedia de audio.

El codificador de audio además comprende la etapa de pre-procesamiento común para pre-procesar una señal de entrada de audio 99 para obtener la señal intermedia de audio 195. Específicamente, la etapa de pre-procesamiento común es operativa para procesar la señal de entrada de audio 99 para que la señal intermedia de audio 195, es decir, la salida del algoritmo de pre-procesamiento común sea una versión comprimida de la señal de entrada de audio.

Un método de codificación de audio preferido para generar una señal de audio codificada, comprende una etapa de codificar 400 una señal intermedia de audio 195 de acuerdo con un primer algoritmo de codificación, teniendo el primer algoritmo de codificación un modelo de sumidero de información y generar, en una primera señal de salida, información espectral codificada que representa la señal de audio; una etapa de codificar 500 una señal intermedia de audio 195 de acuerdo con un segundo algoritmo de codificación, teniendo el segundo algoritmo de codificación un modelo de fuente de información y generar, en una segunda señal de salida, parámetros codificados para el modelo de fuente de información que representa la señal intermedia 195, y una etapa de pre-procesar de manera común 100 una señal de entrada de audio 99 para obtener la señal intermedia de audio 195, donde, en la etapa de pre-procesar de manera común la señal de entrada de audio 99 se procesa para que la señal intermedia de audio 195 sea una versión comprimida de la señal de entrada de audio 99, donde la señal de audio codificada incluye, para una cierta porción de la señal de audio ya sea la primera señal de salida o la segunda señal de salida. El método preferentemente incluye la etapa adicional de codificar una cierta porción de la señal intermedia de audio ya sea utilizando el primer algoritmo de codificación o utilizando el segundo algoritmo de codificación o codificando la señal utilizando ambos algoritmos y emitiendo en una señal codificada ya sea el resultado del primer algoritmo de codificación o el resultado del segundo algoritmo de codificación.

Generalmente, el algoritmo de codificación de audio utilizado en la primera ramificación de codificación 400 refleja y

modela la situación en un sumidero de audio. El sumidero de una información de audio es normalmente el oído humano. El oído humano puede modelarse como un analizador de frecuencia. Por lo tanto, la primera ramificación de codificación emite información espectral codificada. Preferentemente, la primera ramificación de codificación además incluye un modelo psicoacústico para aplicar adicionalmente un umbral de enmascaramiento psicoacústico.

5 Este umbral de enmascaramiento psicoacústico se utiliza cuando se cuantifican valores espectrales de audio donde, preferentemente, la cuantificación se realiza de modo tal que un ruido de cuantificación se introduce cuantificando los valores de audio espectrales, que están ocultos debajo del umbral de enmascaramiento psicoacústico.

La segunda ramificación de codificación representa un modelo de fuente de información, que refleja la generación de sonido de audio. En consecuencia, el modelo de fuente de información puede incluir un modelo del habla reflejado por una etapa de análisis LPC, es decir, transformando una señal de dominio de tiempo en un dominio LPC y procesando posteriormente la señal residual LPC, es decir, la señal de excitación. Sin embargo, los modelos de fuente de sonido alternativos, son modelos de fuente de sonido para representar un cierto instrumento u otros generadores de sonido como una fuente de sonido específica que existe en el mundo real. Una selección entre diferentes modelos de fuente de sonido puede realizarse cuando varios modelos de fuente de sonido están disponibles, por ejemplo teniendo en cuenta un cálculo de SNR, es decir, teniendo en cuenta un cálculo, cuál de los modelos de fuente es el más adecuado para codificar una cierta porción de tiempo y/o porción de frecuencia de una señal de audio. Preferentemente, sin embargo, la conmutación entre las ramificaciones de codificación se realiza en el dominio de tiempo, es decir, una cierta porción de tiempo se codifica utilizando un modelo y una cierta porción de tiempo diferente de la señal intermedia se codifica utilizando la otra ramificación de codificación.

10
15
20

Los modelos de fuentes de información están representados por ciertos parámetros. Con respecto al modelo del habla, los parámetros son parámetros LPC y parámetros de excitación codificados, cuando se considera un codificador moderno del habla tal como AMR-WB+. El AMR-WB+ comprende un codificador ACELP y un codificador TCX. En este caso, los parámetros de excitación codificados pueden ser ganancia global, ruido de fondo, y códigos de longitud variable.

25

La Fig. 3b ilustra un decodificador que corresponde al codificador ilustrado en la Fig. 3a. Generalmente, la Fig. 3b ilustra un decodificador de audio para decodificar una señal de audio codificada y obtener una señal de audio decodificada 799. El decodificador incluye la primera ramificación de decodificación 450 para decodificar una señal codificada de acuerdo con un primer algoritmo de codificación que tiene un modelo de sumidero de información. El decodificador de audio además incluye una segunda ramificación de decodificación 550 para decodificar una señal de información codificada de acuerdo con un segundo algoritmo de codificación que tiene un modelo de fuente de información. El decodificador de audio además incluye un combinador para combinar señales de salida desde la primera ramificación de decodificación 450 y la segunda ramificación de decodificación 550 para obtener una señal combinada. La señal combinada ilustrada en la Fig. 3b como la señal intermedia de audio decodificada 699 se introduce en una etapa de post procesamiento común para post procesar la señal intermedia de audio decodificada 699, que es la señal combinada emitida por el combinador 600 para que una señal emitida de la etapa de pre-procesamiento común sea una versión expandida de la señal combinada. En consecuencia, la señal de audio decodificada 799 posee un contenido de información ampliado en comparación con la señal intermedia de audio decodificada 699. Esta expansión de información se proporciona por la etapa de post procesamiento común con la ayuda de parámetros de pre/post procesamiento que pueden transmitirse de un codificador a un decodificador, o que pueden obtenerse desde la propia señal intermedia de audio decodificada. Preferentemente, sin embargo, los parámetros de pre/post procesamiento se transmiten desde un codificador a un decodificador, ya que este procedimiento permite una mejor calidad de la señal de audio decodificada.

30
35
40
45

La Fig. 3c ilustra un codificador de audio para codificar una señal de entrada de audio 195, que puede ser igual a la señal de audio intermedia 195 de la Fig. 3a de acuerdo con la realización preferida de la presente invención. La señal de entrada de audio 195 se encuentra presente en un primer dominio el cual puede, por ejemplo, ser el dominio de tiempo pero que también puede ser cualquier otro dominio tal como un dominio de frecuencia, un dominio LPC, un dominio espectral LPC o cualquier otro dominio. Generalmente, la conversión de un dominio a otro dominio se realiza por medio de un algoritmo de conversión tal como cualquiera de los algoritmos de conversión de tiempo/frecuencia bien conocidos o algoritmos de conversión frecuencia/tiempo.

50

Una transformación alternativa del dominio de tiempo, por ejemplo en el dominio LPC es el resultado del filtrado LPC de una señal de dominio de tiempo que da como resultado una señal residual LPC o señal de excitación. Cualquier otra operación de filtrado que produzca una señal filtrada que tiene un impacto en una cantidad sustancial de muestras de señal antes de la transformación puede utilizarse como algoritmo de transformación, según el caso. Por lo tanto, ponderar una señal de audio utilizando un filtro de ponderación basándose en LPC es una transformación adicional, que genera una señal en el dominio LPC. En una transformación de tiempo/frecuencia, la modificación de un solo valor espectral tendrá un impacto sobre todos los valores de dominio de tiempo antes de la transformación. De manera análoga, una modificación de cualquier muestra de dominio de tiempo tendrá un impacto sobre cada muestra de dominio de frecuencia. De manera similar, una modificación de una muestra de la señal de excitación en

55
60

una situación de dominio LPC tendrá, debido a la longitud del filtro LPC, un impacto sobre una cantidad sustancial de muestras antes del filtrado de LPC. De manera similar, una modificación de una muestra antes de una transformación LPC tendrá un impacto sobre muchas muestras obtenidas por esta transformación LPC debido al efecto de memoria intrínseco del filtro LPC.

5 El codificador de audio de la Fig. 3c incluye una primera ramificación de codificación 400 que genera una primera señal codificada. La primera señal codificada puede estar en un cuarto dominio el cual, en la realización preferida, es el dominio de tiempo espectral, es decir, el dominio obtenido cuando una señal de dominio de tiempo se procesa a través de una conversión de tiempo/frecuencia.

10 Por lo tanto, la primera ramificación de codificación 400 para codificar una señal de audio utiliza un primer algoritmo de codificación para obtener una primera señal codificada, donde este primer algoritmo de codificación puede o no incluir un algoritmo de conversión de tiempo/frecuencia.

15 El codificador de audio además incluye una segunda ramificación de codificación 500 para codificar una señal de audio. La segunda ramificación de codificación 500 utiliza un segundo algoritmo de codificación para obtener una segunda señal codificada, que es diferente del primer algoritmo de codificación.

20 El codificador de audio además incluye un primer conmutador 200 para conmutar entre la primera ramificación de codificación 400 y la segunda ramificación de codificación 500 para que una porción de la señal de entrada de audio, ya sea la primera señal codificada en la salida del bloque 400 o la segunda señal codificada en la salida de la segunda ramificación de codificación se incluya en una señal de salida del codificador. De esta manera, cuando para una cierta porción de la señal de entrada de audio 195, la primera señal codificada en el cuarto dominio se incluye en la señal de salida del codificador, la segunda señal codificada ya sea la primera señal procesada en el segundo dominio o la segunda señal procesada en el tercer dominio no está incluida en la señal de salida del codificador. Se asegura así que este codificador tenga una tasa de bits eficiente. En las realizaciones, cualquier porción de tiempo de la señal de audio que se incluye en dos señales codificadas diferentes es pequeña en comparación con una longitud de trama de una trama como se analizará en relación con la Fig. 3e. Estas pequeñas porciones son útiles para una transición gradual de una señal codificada a la otra señal codificada en caso de un evento de conmutación con el fin de reducir artefactos que puedan ocurrir sin transición gradual. Por lo tanto, aparte de la región de transición gradual, cada bloque de dominio de tiempo se representa por una señal codificada de un solo dominio.

35 Como se ilustra en la Fig. 3c, la segunda ramificación de codificación 500 comprende un convertidor 510 para convertir la señal de audio en el primer dominio, es decir, la señal 195 en un segundo dominio. Además, la segunda ramificación de codificación 500 comprende una primera ramificación de procesamiento 522 para procesar una señal de audio en el segundo dominio para obtener una primera señal procesada que, preferentemente, se encuentra también en el segundo dominio para que la primera ramificación de procesamiento 522 no realice un cambio de dominio.

40 La segunda ramificación de codificación 500 además comprende una segunda ramificación de procesamiento 523, 524 que convierte la señal de audio en el segundo dominio en un tercer dominio, que es diferente al primer dominio y que es también diferente al segundo dominio y que procesa la señal de audio en el tercer dominio para obtener una segunda señal procesada en la salida de la segunda ramificación de procesamiento 523, 524.

45 Adicionalmente, la segunda ramificación de codificación comprende un segundo conmutador 521 para conmutar entre la primera ramificación de procesamiento 522 y la segunda ramificación de procesamiento 523, 524 para que, para una porción de la entrada de la señal de audio en la segunda ramificación de codificación, ya sea la primera señal procesada en el segundo dominio o la segunda señal procesada en el tercer dominio esté en la segunda señal codificada.

50 La Fig. 3d ilustra un decodificador correspondiente para decodificar una señal de audio codificada generada por el codificador de la Fig. 3c. Generalmente, cada bloque de la señal de audio del primer dominio está representado por una segunda señal de dominio, una tercera señal de dominio o una cuarta señal de dominio codificada aparte de una región de transición gradual opcional la cual, preferentemente, es corta en comparación con la longitud de una trama con el fin de obtener un sistema que se encuentra en el límite de muestreo crítico en la medida de lo posible. La señal de audio codificada incluye la primera señal codificada, una segunda señal codificada en un segundo dominio y una tercera señal codificada en un tercer dominio, donde la primera señal codificada, la segunda señal codificada y la tercera señal codificada se relacionan todas con diferentes porciones de tiempo de la señal de audio decodificada y donde el segundo dominio, el tercer dominio y el primer dominio para una señal de audio decodificada son diferentes entre sí.

60 El decodificador comprende una primera ramificación de decodificación para la decodificación basándose en el primer algoritmo de codificación. La primera ramificación de decodificación se ilustra en 431, 440 en la Fig. 3d y

preferentemente comprende un convertidor de frecuencia/tiempo. La primera señal codificada está preferentemente en un cuarto dominio y se convierte en el primer dominio que es el dominio para la señal de salida decodificada.

El decodificador de la Fig. 3d además comprende una segunda ramificación de decodificación la cual comprende varios elementos. Estos elementos incluyen una primera ramificación de procesamiento inversa 531 para el procesamiento inverso de la segunda señal codificada para obtener una primera señal procesada inversa en el segundo dominio en la salida del bloque 531. La segunda ramificación de decodificación además comprende una segunda ramificación de procesamiento inversa 533, 534 para el procesamiento inverso de una tercera señal codificada para obtener una segunda señal procesada inversa en el segundo dominio, donde la segunda ramificación de procesamiento inversa comprende un convertidor para la conversión desde el tercer dominio al segundo dominio.

La segunda ramificación de decodificación además comprende un primer combinador 532 para combinar la primera señal procesada inversa y la segunda señal procesada inversa para obtener una señal en el segundo dominio, donde esta señal combinada se ve únicamente influenciada, en el primer instante de tiempo, por la primera señal procesada inversa y se ve únicamente influenciada, en un instante posterior, por la segunda señal procesada inversa.

La segunda ramificación de decodificación además comprende un convertidor 540 para convertir la señal combinada al primer dominio.

Finalmente, el decodificador ilustrado en la Fig. 3d comprende un segundo combinador 600 para combinar la primera señal del bloque decodificada 431, 440 y la señal de salida del convertidor 540 para obtener una señal de salida decodificada en el primer dominio. Nuevamente, la señal de salida decodificada en el primer dominio se ve únicamente influenciada, en el primer instante de tiempo, por la salida de señal por el convertidor 540 y se ve únicamente influenciada, en un instante de tiempo posterior, por la primera salida de señal decodificada por el bloque 431, 440.

Esta situación se ilustra, desde una perspectiva del codificador, en la Fig. 3e. La porción superior en la Fig. 3e ilustra en la representación esquemática, una primera señal de audio de dominio como una señal de audio de dominio de tiempo, donde el índice de tiempo aumenta de izquierda a derecha y el elemento 3 podría considerarse como una secuencia de muestras de audio que representa la señal 195 en la Fig. 3c. La Fig. 3e ilustra las tramas 3a, 3b, 3c, 3d que pueden generarse conmutando entre la primera señal codificada y la primera señal procesada y la segunda señal procesada como se ilustra en el elemento 4 en la Fig. 3e. La primera señal codificada, la primera señal procesada y la segunda señal procesada están todas en diferentes dominios y con el fin de asegurarse que la conmutación entre los diferentes dominios no dé como resultado un artefacto en el lado del decodificador, las tramas 3a, 3b de la señal de dominio de tiempo poseen un rango de solapamiento que se indica como una región de transición gradual, y una región de transición gradual de este tipo se encuentra en la trama 3b y 3c. Sin embargo, no existe tal región de transición gradual entre la trama 3d, 3c lo que significa que la trama 3d está también representada por una segunda señal procesada, es decir, una señal en el tercer dominio, y no hay cambio de dominio entre la trama 3c y 3d. Por lo tanto, generalmente, se prefiere no proporcionar una región de transición gradual donde no exista cambio de dominio y proporcionar una región de transición gradual, es decir, una porción de la señal de audio que se codifica por dos señales codificadas/procesadas posteriormente cuando no existe un cambio de dominio, es decir, una acción de conmutación de cualquiera de los dos conmutadores. Preferentemente, las transiciones graduales se realizan para otros cambios de dominio.

En la realización, en la cual la primera señal codificada o la segunda señal procesada se han generado por un procesamiento MDCT que tiene por ejemplo un 50 por ciento de solapamiento, cada vez que se incluye una muestra de dominio en dos tramas posteriores. Debido a las características de MDCT, sin embargo, esto no da como resultado una tara, ya que el MDCT es un sistema críticamente muestreado. En este contexto, críticamente muestreado significa que la cantidad de valores espectrales es la misma que la cantidad de valores de dominio de tiempo. El MDCT es ventajoso ya que el efecto de transición gradual se proporciona sin una región de transición gradual específica para que se proporcione una transición gradual desde el bloque MDCT al próximo bloque MDCT sin ninguna tara que podría violar el requisito de muestreo crítico.

Preferentemente, el primer algoritmo de codificación en la primera ramificación de codificación se basa en un modelo de sumidero de información, y el segundo algoritmo de codificación en la segunda ramificación de codificación se basa en una fuente de información o un modelo SNR. Un modelo SNR es un modelo que no está específicamente relacionado con un mecanismo de generación de un sonido específico pero que puede seleccionarse un modo de codificación entre una pluralidad de modos de codificación basándose, por ejemplo, en una decisión de bucle cerrado. De esta manera, un modelo SNR es cualquier modelo de codificación disponible pero que no necesariamente tiene que estar relacionado con la constitución física del generador de sonido pero que es cualquier modelo de codificación parametrizado diferente del modelo de sumidero de información, que puede seleccionarse

por una decisión de bucle cerrado y, específicamente, comparando diferentes resultados de SNR de los diferentes modelos.

5 Como se ilustra en la Fig. 3c, se proporciona un controlador 300, 525. Este controlador puede incluir las funcionalidades de la etapa de decisión 300 de la Fig. 1a y, además, puede incluir la funcionalidad del dispositivo de control del conmutador 525 en la Fig. 1a. Generalmente, el controlador es para controlar el primer conmutador y el segundo conmutador en un modo adaptativo de señal. El controlador es operativo para analizar la entrada de la señal en el primer conmutador o la salida por medio de la primera o segunda ramificación de codificación o las señales obtenidas por codificación y decodificación desde la primera y la segunda ramificación de codificación con respecto a la función objetivo. Como alternativa, o adicionalmente, el controlador es operativo para analizar la señal introducida en el segundo conmutador o emitida por medio de la primera ramificación de procesamiento o la segunda ramificación de procesamiento u obtenida mediante el procesamiento y procesamiento inverso desde la primera ramificación de procesamiento y la segunda ramificación de procesamiento, nuevamente con respecto a una función objetivo.

15 En una realización, la primera ramificación de codificación o la segunda ramificación de codificación comprende un algoritmo de conversión de tiempo/frecuencia de introducción antisolapamiento tal como un algoritmo MDCT o MDST, que es diferente de una transformada FFT directa (Transformada rápida de Fourier), que no introduce un efecto de solapamiento. Adicionalmente, una o ambas ramificaciones comprenden un bloque codificador cuantificador/entropía. Específicamente sólo la segunda ramificación de procesamiento de la segunda ramificación de codificación incluye el convertidor de tiempo/frecuencia que introduce una operación de solapamiento y la primera ramificación de procesamiento de la segunda ramificación de codificación comprende un codificador cuantificador y/o de entropía y no introduce efectos de solapamiento. El convertidor de tiempo/frecuencia que introduce solapamiento preferentemente comprende un generador de partición en ventanas para aplicar una ventana de análisis y un algoritmo de transformación MDCT. Específicamente, el generador de partición en ventanas es operativo para aplicar la función de ventana a tramas posteriores en un modo de solapamiento para que una muestra de una señal partida en ventanas tenga lugar en al menos dos tramas partidas en ventanas posteriores.

30 En una realización, la primera ramificación de procesamiento comprende un codificador ACELP y una segunda ramificación de procesamiento que comprende un convertidor espectral MDCT y el cuantificador para cuantificar componentes espectrales para obtener componentes espectrales cuantificados, donde cada componente espectral cuantificado es cero o se define por un índice cuantificador de la pluralidad de diferentes posibles índices cuantificadores.

35 Adicionalmente, se prefiere que el primer conmutador 200 opere en una manera de bucle abierto y el segundo conmutador opere en una manera de bucle cerrado.

40 Como se estableció anteriormente, ambas ramificaciones de codificación son operativas para codificar la señal de audio en el sentido del bloque, en el cual el primer conmutador o el segundo conmutador conmuta en sentido del bloque para que tenga lugar una acción de conmutación, al mínimo, después de un bloque con una cantidad de muestras pre-definidas de una señal, formando la cantidad pre-definida una longitud de trama para el conmutador correspondiente. En consecuencia, el gránulo para conmutar con el primer conmutador puede ser, por ejemplo, un bloque de 2048 o 1028 muestras, y la longitud de la trama, basándose en cuál del primer conmutador 200 está conmutando puede ser variable pero está, preferentemente, fijado a un periodo suficientemente largo de este tipo.

45 Contrario a lo anterior, la longitud del bloque para el segundo conmutador 521, es decir, cuando el segundo conmutador 521 conmuta de un modo al otro, es sustancialmente menor a la longitud del bloque para el primer conmutador. Preferentemente, ambas longitudes de bloque para los conmutadores se seleccionan de modo que la longitud de bloque mayor es un múltiplo entero de la longitud de bloque menor. En la realización preferida, la longitud del bloque del primer conmutador es 2048 o 1024 y la longitud del bloque del segundo conmutador es 1024 o más preferentemente 512 e incluso más preferentemente 256 e incluso más preferentemente 128 muestras para que, al máximo, el segundo conmutador pueda conmutar 16 veces cuando el primer conmutador conmuta sólo una vez. Una relación preferida de longitud máxima de bloque es, sin embargo, 4:1.

55 En una realización adicional, el controlador 300, 525 es operativo para realizar una discriminación del habla y de música para el primer conmutador de tal manera que una decisión del habla se favorece con respecto a una decisión de música. En esta realización, se toma una decisión de voz incluso cuando una porción menor del 50 % de una trama para el primer conmutador es del habla y la porción de más del 50 % de la trama es de música.

60 Adicionalmente, el controlador es operativo para conmutar ya al modo del habla, cuando una pequeña porción de la primera trama es del habla y, específicamente, cuando una porción de la primera trama es del habla, que es el 50 % de la longitud de la segunda trama más pequeña. Por lo tanto, una decisión de conmutación preferida del habla/favorecimiento ya conmuta al habla incluso cuando, por ejemplo, sólo el 6 % o el 12 % de un bloque

correspondiente a la longitud de la trama del primer conmutador es habla.

Este procedimiento preferentemente se lleva cabo para explotar completamente la capacidad de ahorro de tasa de bits de la primera ramificación de procesamiento, la cual tiene un núcleo del habla vocalizada en una realización y para no perder calidad alguna incluso para el resto de la primera trama larga, al cual no es del habla debido a que la segunda ramificación de procesamiento incluye un convertidor y, por lo tanto, es útil también para las señales de audio que poseen señales que no son del habla. Preferentemente, esta segunda ramificación de procesamiento incluye un solapamiento MDCT, que está muestreado en forma crítica, y que incluso en tamaños pequeños de ventana proporciona una operación libre de solapamiento y altamente eficiente debido al procesamiento de cancelación de solapamiento de dominio de tiempo tal como el solapamiento y adición en el lado del decodificador. Adicionalmente, es útil una gran longitud de bloque para la primera ramificación de codificación que preferentemente es una ramificación de codificación MDCT tipo AAC, ya que las señales que no son del habla son normalmente bastante estacionarias y una larga ventana de transformación proporciona una resolución de alta frecuencia y por lo tanto, una alta calidad y, además, proporciona una eficiencia en tasa de bits debido a un módulo de cuantificación controlado en forma psicoacústica, que puede también aplicarse al modo de codificación basándose en la transformación en la segunda ramificación de procesamiento o de la segunda ramificación de codificación.

Con respecto a la ilustración del decodificador de la Fig. 3d, se prefiere que la señal transmitida incluya un indicador explícito como información secundaria 4a como se ilustra en la Fig. 3e. Esta información secundaria 4a se extrae por medio de un analizador de secuencia de bits que no se ilustra en la Fig. 3d con el fin de reenviar la primera señal codificada correspondiente, la primera señal procesada o segunda señal procesada al procesador correcto tal como la primera ramificación de decodificación, la primera ramificación de procesamiento inversa o la segunda ramificación de procesamiento inversa en la Fig. 3d. Por lo tanto, una señal codificada no sólo posee señales codificadas/procesadas sino que también incluye información secundaria relacionada con estas señales. En otras realizaciones, sin embargo, puede haber una señalización implícita que permite a un analizador de secuencia de bits del lado del decodificador distinguir entre las señales ciertas. Con referencia a la Fig. 3e, se describe que la primera señal procesada o la segunda señal procesada son la salida de la segunda ramificación de codificación y, en consecuencia, la segunda señal codificada.

Preferentemente, la primera ramificación de decodificación y/o la segunda ramificación de procesamiento inversa incluyen una transformación MDCT para convertir del dominio espectral al dominio de tiempo. Para tal fin, se proporciona un sumador de solapamiento para llevar a cabo una funcionalidad de cancelación de solapamiento en dominio de tiempo la cual, al mismo tiempo, proporciona un efecto de transición gradual con el fin de evitar artefactos de bloqueo. Generalmente, la primera ramificación de decodificación convierte una señal codificada en el cuarto dominio en el primer dominio, mientras que la segunda ramificación de procesamiento inversa realiza una conversión del tercer dominio al segundo dominio y el convertidor posteriormente conectado al primer combinador proporciona una conversión del segundo dominio al primer dominio para que, en la entrada del combinador 600, únicamente las señales del primer dominio se encuentran allí, las cuales representan, en la realización de la Fig. 3d, la señal de salida decodificada.

Las Fig. 4a y 4b ilustran dos realizaciones diferentes, que difieren en la posición del conmutador 200. En la Fig. 4a, el conmutador 200 se encuentra entre una salida de la etapa de pre-procesamiento común 100 y la entrada de las dos ramificaciones codificadas 400, 500. La realización de la Fig. 4a asegura que la señal de audio se introduce en una única ramificación de codificación, y la otra ramificación de codificación, no conectada a la salida de la etapa de pre-procesamiento común no opera y, por lo tanto, se encuentra apagada o en modo de reposo. Se prefiere esta realización porque la ramificación de codificación no activa no consume energía y los recursos computacionales que son útiles para aplicaciones móviles en particular, que se alimentan por baterías y, por lo tanto, poseen limitación general de consumo de energía.

Por otro lado, sin embargo, la realización de la Fig. 4b puede preferirse cuando el consumo de energía no es un problema. En esta realización, ambas ramificaciones de codificación 400, 500 se encuentran activas todo el tiempo, y sólo la salida de la ramificación de codificación seleccionada para una cierta porción de tiempo y/o una cierta porción de frecuencia se reenvía al formateador de secuencia de bits el cual puede implementarse como un multiplexor de secuencia de bits 800. Por lo tanto, en la realización de la Fig. 4b, ambas ramificaciones de codificación se encuentran activas todo el tiempo, y la salida de una ramificación de codificación que se selecciona por la etapa de decisión 300 se introduce en la secuencia de bits de salida, mientras que la salida de la otra ramificación de codificación no seleccionada 400 se descarta, es decir, no se introduce en la secuencia de bits de salida, es decir, la señal de audio codificada.

La Fig. 4c ilustra un aspecto adicional de una implementación de decodificador preferida. Con el fin de evitar artefactos audibles específicamente en la situación, en la cual el primer decodificador es un decodificador generador de solapamiento en tiempo o generalmente establecido como un decodificador de dominio de frecuencia y el segundo decodificador es un dispositivo de dominio, los límites entre los bloques o tramas emitidos por el primer

5 decodificador 450 y el segundo decodificador 550 no deberían ser totalmente continuos, específicamente en una
situación de conmutación. De esta manera, cuando se emite el primer bloque del primer decodificador 450 es
emitido y, cuando para la porción de tiempo posterior, se emite un bloque del segundo decodificador, se prefiere
realizar una operación de transición gradual como se ilustra por medio del bloque de transición gradual 607. Para tal
fin, el bloque de transición gradual 607 podría implementarse como se ilustra en la Fig. 4c en 607a, 607b y 607c.
Cada ramificación podría tener una ponderación con un factor de ponderación m_1 entre 0 y 1 en la escala
normalizada, donde el factor de ponderación puede variar como se indica en el diagrama 609, una regla de
transición gradual de este tipo asegura la continua y uniforme transición gradual, la cual, además, asegura al usuario
que no va a percibir variaciones de volumen. Se pueden aplicar reglas de transición gradual no lineales como una
10 regla de transición gradual \sin^2 en lugar de una regla de transición gradual lineal.

En ciertas instancias, el último bloque del primer decodificador se generó utilizando una ventana donde la ventana
en realidad desarrolló un desvanecimiento de este bloque. En ese caso, el factor de ponderación m_1 en el bloque
607a es igual a 1 y, en realidad, no se necesita ponderación alguna para esta ramificación.

15 Cuando tiene lugar una conmutación del segundo decodificador al primer decodificador, y cuando el segundo
decodificador incluye una ventana que en realidad desvanece la salida hasta el final del bloque, entonces la
ponderación indicada con " m_2 " no sería necesaria o el parámetro de ponderación puede establecerse a 1 para toda
la región de transición gradual.

20 Cuando el primer bloque después de que se ha generado una conmutación utilizando una operación de generación
de ventanas, y cuando esta ventana en realidad realizó una operación de desvanecimiento, entonces el factor de
ponderación correspondiente, puede establecerse también a 1 para que no sea realmente necesario un ponderador.
Por lo tanto, cuando el último bloque se parte en ventanas con el fin de desvanecerse por el decodificador y cuando
25 el primer bloque después que la conmutación se parte en ventanas utilizando el decodificador para proporcionar un
desvanecimiento, los ponderadores 607a, 607b no son del todo necesarios y una operación de suma por parte del
sumador 607c es suficiente.

30 En este caso, la porción de desvanecimiento de la última trama y la porción de aumento gradual de la próxima trama
definen la región de transición gradual indicada en el bloque 609. Adicionalmente, se prefiere en una situación de
este tipo que el último bloque de un decodificador posea un cierto solapamiento de tiempo con el primer bloque del
otro decodificador.

35 Si una operación de transición gradual no es necesaria o no es posible o no se desea, y si sólo se encuentra un
codificador definitivo desde un decodificador al otro decodificador, se prefiere realizar una conmutación de este tipo
en pasajes silenciosos de la señal de audio o al menos en pasajes de la señal de audio donde haya baja energía, es
decir, percibidos como silenciosos o casi silenciosos. Preferentemente, la etapa de decisión 300 asegura en una
realización de este tipo que el conmutador 200 sólo se activa cuando la porción de tiempo correspondiente que
sigue al evento de conmutación posee energía que, por ejemplo, es menor a la energía media de la señal de audio y
40 es, preferentemente, menor del 50 % de la energía media de la señal de audio relacionada con, por ejemplo, dos o
incluso más porciones/tramas de tiempo de la señal de audio.

45 Preferentemente, la segunda regla de codificación/regla de decodificación es un algoritmo de codificación basado en
LPC. En una codificación del habla basada en LPC, se establece una diferencia entre segmentos o porciones de
señal de excitación del tipo impulso cuasi-periódico, y segmentos o porciones de señal de excitación del tipo ruido.
Esto se lleva cabo para codificadores de voz LPC (2,4 kbps) con tasa de bits muy baja como en la Fig 7b. Sin
embargo, en los codificadores CELP con tasa media, la excitación se obtiene para la adición de vectores en escala
de un libro de códigos adaptativo y un libro de códigos fijo.

50 Los segmentos de señal de excitación del tipo impulso cuasi-periódico, es decir, segmentos de señal con tonos
específicos se codifican con diferentes mecanismos que las señales de excitación del tipo ruido. Mientras que las
señales de excitación del tipo impulso cuasi-periódico están conectadas el habla vocalizada, las señales del tipo
ruido están relacionadas con el habla no vocalizada.

55 Como ejemplo se hace referencia a las Figs. 5a a 5d. En este caso, los segmentos o porciones de señal de
excitación del tipo impulso cuasi-periódico y segmentos o porciones de señal de excitación del tipo ruido se analizan
a modo de ejemplo. Específicamente, un habla vocalizada como se ilustra en la Fig. 5a en el dominio de tiempo y en
la Fig. 5b en el dominio de frecuencia se analiza como ejemplo para una porción de señal de excitación del tipo
impulso cuasi-periódico, y un segmento del habla no vocalizada como ejemplo de una porción de señal del tipo ruido
60 en relación con las Figs. 5c y 5d. El habla generalmente puede clasificarse en vocalizada, no vocalizada o mixta. Las
representaciones de dominio de tiempo y frecuencia para segmentos de muestras vocalizadas y no vocalizadas se
muestran en la Fig. 5a a 5d. El habla vocalizada es cuasi periódica en el dominio de tiempo y estructurada
armónicamente en el dominio de frecuencia, mientras que el habla no vocalizada es del tipo aleatorio y de banda

ancha. El espectro de corto tiempo del habla vocalizada se caracteriza por su estructura de formante armónica fina. La estructura armónica fina es consecuencia de la cuasi-periodicidad del habla y puede atribuirse a las cuerdas vocales vibratorias. La estructura de formante (envolvente espectral) se debe a la interacción de la fuente y tractos vocales. Los tractos vocales comprenden la faringe y la cavidad bucal. La forma de la envolvente espectral que “se ajusta” al espectro de tiempo corto del habla vocalizada se asocia con las características de transferencia del tracto vocal y la inclinación espectral (6 dB /octava) debido al pulso glotal. La envolvente espectral se caracteriza por un grupo de picos llamados formantes. Los formantes son modos resonantes del tracto vocal. Para el tracto vocal promedio existen de tres a cinco formantes por debajo de 5 kHz. Las amplitudes y locaciones de los primeros tres formantes, normalmente tienen lugar por debajo de 3 kHz, son ambas muy importantes, en la síntesis y percepción del habla. Los formantes mayores son también importantes para representaciones de banda ancha y el habla no vocalizada. Las propiedades del habla están relacionadas con el sistema de producción del habla físico como sigue. El habla vocalizada se produce excitando el tracto vocal con pulsos de aire glotales cuasi-periódico generados por las cuerdas vocales vibratorias. La frecuencia de los pulsos periódicos se denomina como la frecuencia fundamental o tono. El habla no vocalizada se produce forzando el aire a través de una constricción del tracto vocal. Los sonidos nasales se deben al acoplamiento acústico del tracto nasal con el tracto vocal, y se producen sonidos plosivos liberando de manera abrupta la presión de aire que se creó detrás del cierre en el tracto.

En consecuencia, una porción del tipo ruido de la señal de audio no muestra una estructura de dominio de tiempo del tipo impulso ni una estructura de dominio de frecuencia armónica como se ilustra en la Fig. 5c y Fig. 5d, la cual es diferente de la porción tipo impulso cuasi-periódica como se ilustra por ejemplo en la Fig. 5a y Fig. 5b. Como se describe más adelante, sin embargo, la diferencia entre las porciones del tipo ruido y las porciones del tipo impulso cuasi-periódico pueden también observarse después de un LPC para la señal de excitación. El LPC es un método que modela el tracto vocal y extrae desde la señal la excitación de los tractos vocales.

Adicionalmente, las porciones del tipo impulso cuasi-periódico y las porciones del tipo ruido pueden tener lugar de manera oportuna, es decir, significa que una porción de la señal de audio en el tiempo es ruidosa y otra porción de la señal de audio en el tiempo es cuasi-periódica, es decir tonal. De manera alternativa, o además, las características de una señal pueden ser diferentes en diferentes bandas de frecuencia. De esta manera, la determinación, acerca de si la señal de audio es ruidosa o tonal, puede también realizarse seleccionando la frecuencia para que una cierta banda de frecuencia o varias bandas de frecuencia se consideren como ruidosas y otras bandas de frecuencia se consideren tonales. En este caso, una cierta porción de tiempo de la señal de audio puede incluir componentes tonales y componentes de ruido.

La Fig. 7a ilustra un modelo lineal de un sistema de producción del habla. Este sistema supone una excitación de dos etapas, es decir, un tren de impulsos para el habla vocalizada como se indica en la Fig. 7c, y un ruido aleatorio para el habla no vocalizada como se indica en la Fig. 7d. El tracto vocal se modela como un filtro omni-polar 70 el cual procesa pulsos de la Fig. 7c o Fig. 7d, generados por el modelo glotal 72. Por lo tanto, el sistema de la Fig. 7a puede reducirse a un modelo de filtro omni-polar de la Fig. 7b con una etapa de ganancia 77, una trayectoria de ida 78, una trayectoria de realimentación 79, y una etapa de adición 80. En la trayectoria de realimentación 79, existe un filtro de predicción 81, y todo el sistema de síntesis del modelo fuente ilustrado en la Fig. 7b puede estar representado utilizando funciones de dominio z de la siguiente manera:

$$S(z)=g/(1-A(z))\cdot X(z),$$

donde g representa la ganancia, A(z) es el filtro de predicción como lo determina un análisis LP, X(z) es la señal de excitación, y S(z) es la salida de la síntesis del habla.

Las Figs. 7c y 7d proporcionan una descripción de dominio de tiempo gráfica de síntesis del habla vocalizada y no vocalizada utilizando el modelo de sistema de fuente lineal. Este sistema y los parámetros de excitación en la ecuación anterior son desconocidos y deben determinarse desde un grupo finito de muestras del habla. Los coeficientes de A(z) se obtienen utilizando una predicción lineal de la señal de entrada y una cuantificación de los coeficientes del filtro. En un predictor lineal progresivo de orden p-ésimo, la muestra presente de la secuencia del habla se predice desde una combinación lineal de p muestras pasadas. Los coeficientes del predictor pueden determinarse por algoritmos bien conocidos tales como el algoritmo Levinson-Durbin, o generalmente un método de autocorrelación o método de reflexión.

La Fig. 7e ilustra una implementación más detallada del bloque de análisis de LPC. La señal de audio se introduce en un bloque de determinación de filtro que determina la información del filtro A(z). Esta información se emite como la información de predicción a corto plazo necesaria para un decodificador. La información de predicción a corto plazo se requiere por el filtro de predicción real 85. En un restador 86, se introduce una muestra actual de la señal de audio y se resta un valor predicho para la muestra actual de modo que para esta muestra, la señal de error de predicción se genera en la línea 84. Una secuencia de tales muestras de señal de error de predicción se ilustra muy esquemáticamente en las Fig. 7c o 7d. En consecuencia, las Fig. 7a, 7b pueden considerarse como un tipo de

una señal del tipo impulso rectificado.

Mientras que la Fig. 7e ilustra un modo preferido para calcular la señal de excitación, la Fig. 7f ilustra un modo preferido para calcular la señal ponderada. En contraposición a la Fig. 7e, el filtro 85 es diferente, cuando γ es diferente de 1. Un valor menor a 1 se prefiere para γ . Adicionalmente, el bloque 87 está presente, y se prefiere para μ un número menor a 1. Generalmente, los elementos en la Fig. 7e y 7f pueden implementarse como en 3GPP TS 26.190 o 3GPP TS 26.290.

La Fig. 7g ilustra un procesamiento inverso, que puede aplicarse del lado del decodificador como en el elemento 537 de la Fig. 2b. En particular, el bloque 88 genera una señal no ponderada desde la señal ponderada y el bloque 89 calcula una excitación desde la señal no ponderada. En general, todas las señales excepto la señal no ponderada en la Fig. 7g se encuentran en el dominio LPC, sin embargo, la señal de excitación y la señal ponderada son señales diferentes en el mismo dominio. El bloque 89 emite una señal de excitación que a continuación puede utilizarse junto con la salida del bloque 536. A continuación, la transformación LPC inversa común puede realizarse en el bloque 540 de la Fig. 2b.

Posteriormente, se analizará un codificador CELP de análisis por síntesis será analizado en relación con la Fig. 6 con el fin de ilustrar las modificaciones aplicadas a este algoritmo. Este codificador CELP se analiza en detalle en "Speech Coding: A Tutorial Review", Andreas Spanias, Procedimientos de IEEE, Vol. 82, N.º 10, octubre de 1994, páginas 1541-1582. El codificador CELP como se ilustra en la Fig. 6 incluye un componente de predicción a largo plazo 60 y un componente de predicción a corto plazo 62. Adicionalmente, se usa un libro de codificación como se indica en 64. Un filtro de ponderación perceptual $W(z)$ se implementa en 66, y un controlador de minimización de error se proporciona en 68. $s(n)$ es la señal de entrada de dominio de tiempo. Después de haber sido perceptualmente ponderada, la señal ponderada se introduce en un restador 69, que calcula el error entre la señal de síntesis ponderada en la salida del bloque 66 y la señal ponderada original $s_w(n)$. Generalmente, los coeficientes del filtro de predicción a corto plazo $A(z)$ se calculan mediante una etapa de análisis LP y sus coeficientes se cuantifican en $\hat{A}(z)$ como se indica en la Fig. 7e. La información de predicción a largo plazo $AL(z)$ incluyendo la ganancia de predicción a largo plazo g y el índice de cuantificación del vector, es decir, referencias del libro de codificación se calculan sobre la señal de error de predicción en la salida de la etapa de análisis LPC denominada como 10a en la Fig. 7e. Los parámetros LTP son el retardo de tono y ganancia. En CELP esto se implementa normalmente como un libro de codificación adaptativo que contiene la señal de excitación pasada (no la residual). El retardo CB adaptativo y al ganancia se encuentran al minimizar el cuadrado medio del ponderado (búsqueda de tono en bucle cerrado).

El algoritmo CELP codifica la señal residual señal obtenida después de las predicciones a largo plazo y a corto plazo utilizando un libro de codificación de por ejemplo secuencias Gaussianas. El algoritmo ACELP, donde la "A" significa "Algebraico" posee un libro de codificación específico algebraicamente diseñado.

Un libro de codificación puede contener más o menos vectores donde cada vector es de algunas muestras de largo. Un factor de ganancia g adapta al vector de codificación y la codificación obtenida se filtra por el filtro de síntesis de predicción a largo plazo y el filtro de síntesis de predicción a corto plazo. El vector de codificación "óptimo" se selecciona de manera que se minimiza el error cuadrado medio perceptualmente ponderado en la salida del restador 69. El proceso de búsqueda en CELP se realiza por medio de una optimización de análisis por síntesis ilustrada en la Fig. 6.

Para casos específicos, cuando una trama es una mezcla del habla vocalizada y no vocalizada o cuando tiene lugar el habla sobre música, una codificación TCX puede ser más apropiada para codificar la excitación en el dominio LPC. La codificación TCX procesa la señal a ponderada en el dominio de frecuencia sin presuponer producciones de excitación. TCX es más genérica que la codificación CELP y no está restringida a un modelo de fuente vocalizada o no vocalizada de la excitación. TCX es aún una codificación con modelo filtro-fuente que utiliza un filtro de predicción lineal para modelar los formantes de las señales del tipo voz.

En la codificación tipo AMR-WB+-, se lleva a cabo una selección entre los diferentes modos TCX y ACELP como se conoce de la descripción AMR-WB+. Los modos TCX son diferentes en que la longitud de la Transformada Discreta de Fourier en sentido del bloque es diferente para diferentes modos y el mejor modo puede seleccionarse mediante un enfoque de análisis por síntesis o modo de "prealimentación" directa.

Como se analizó en relación con las Fig. 2a y 2b, la etapa de pre-procesamiento común 100 preferentemente incluye un multi-canal conjunto (dispositivo envolvente/estéreo conjunto) 101 y, además, una etapa de extensión de ancho de banda 102. De manera correspondiente, el decodificador incluye una etapa de extensión de ancho de banda 701 y una etapa de multicanal conjunto posteriormente conectada 702. Preferentemente, la etapa de multicanal conjunto 101 está, con respecto al codificador, conectada antes de la etapa de extensión de ancho de banda 102, y, en el lado del decodificador, la etapa de extensión de ancho de banda 701 está conectada antes de la etapa de multicanal

conjunto 702 con respecto a la dirección de procesamiento de señal. Como alternativa, sin embargo, la etapa de pre-procesamiento común puede incluir una etapa de multicanal conjunto sin la extensión de ancho de banda posteriormente conectada o una etapa de extensión de ancho de banda sin una etapa de multicanal conjunto conectada.

5 Un ejemplo preferido para una etapa de multicanal conjunto en el lado del codificador 101a, 101b y en el lado del decodificador 702a y 702b se ilustra en el contexto de la Fig. 8. Un número de canales de entrada originales E se introduce en el mezclador descendente 101a para que el mezclador descendente genere un número de canales K transmitidos, donde el número K es mayor a o igual a uno y menor a o igual a E.

10 Preferentemente, los canales de entrada E se introducen en un analizador de parámetros de multicanal conjunto 101b que genera información paramétrica. Esta información paramétrica está preferentemente codificada por entropía tal como por una codificación de diferencia y posterior codificación Huffman o, como alternativa, codificación aritmética posterior. La salida de información paramétrica codificada por el bloque 101b se transmite a un decodificador de parámetro 702b que puede ser parte del elemento 702 en la Fig. 2b. El decodificador de parámetro 702b decodifica la información paramétrica transmitida y reenvía la información paramétrica decodificada al mezclador ascendente 702a. El mezclador ascendente 702a recibe los canales K transmitidos y genera un número de canales de salida L, donde el número de L es mayor a o igual a K y menor a o igual a E.

20 La información paramétrica puede incluir diferencias de nivel de inter canal, diferencias de tiempo de inter canal, diferencias de fase de inter canal y/o mediciones de coherencia de inter canal como se conoce a partir de la técnica de BCC o como se conoce y describe en detalle en la norma de sonido envolvente MPEG. La cantidad de canales transmitidos puede ser un solo canal mono para aplicaciones con tasa de bits ultra bajas o puede incluir una aplicación estéreo compatible o puede incluir una señal estéreo compatible, es decir, dos canales. Típicamente, el número de canales de entrada E puede ser cinco o incluso mayor. Como alternativa, el número de canales de entrada E puede ser también E objetos de audio como se conoce en el contexto de codificación de objeto de audio espacial (SAOC, por su sigla en inglés).

30 En una implementación, el mezclador descendente realiza una adición ponderada o no ponderada de los canales de entrada originales E de los objetos de audio de entrada E. En caso de objetos de audio como canales de entrada, el analizador de parámetros multicanal conjunto 101b calculará parámetros de objeto de audio como una matriz de correlación entre los objetos de audio preferentemente para cada porción de tiempo y aún más preferentemente para cada banda de frecuencia. Para tal fin, todo el rango de frecuencia puede dividirse en al menos 10 y preferentemente 32 o 64 bandas de frecuencia.

35 La Fig. 9 ilustra una realización preferida para la implementación de la etapa de extensión de ancho de banda 102 en la Fig. 2a y la etapa de extensión de ancho de banda correspondiente 701 en la Fig. 2b. En el lado del codificador, el bloque de extensión de ancho de banda 102 preferentemente incluye un bloque de filtro paso bajo 102b, un bloque de submuestreador, que sigue al paso bajo, o que es parte del QMF inverso, que actúa sobre sólo la mitad de las bandas QMF, y un analizador de banda alta 102a. La entrada de señal de audio original en el bloque de extensión de ancho de banda 102 se filtra con paso bajo para generar la señal de banda baja la cual se introduce en las ramificaciones de codificación y/o el conmutador. El filtro paso bajo posee una frecuencia recortada que puede estar en el rango de 3 kHz a 10 kHz. Adicionalmente, el bloque de extensión de ancho de banda 102 además incluye un analizador de banda alta para calcular los parámetros de extensión de ancho de banda como la información de parámetros de envolvente espectral, una información de parámetros de ruido de fondo, una información de parámetros de filtro inverso, información paramétrica adicional relacionada con ciertas líneas armónicas en la banda alta y parámetros adicionales como se analizan en detalle en la norma MPEG-4 en el capítulo relacionado con la replicación de banda espectral.

50 En el lado del decodificador, el bloque de extensión de ancho de banda 701 incluye un parcheador 701a, un ajustador 701b y un combinador 701c. El combinador 701c combina la señal de banda baja decodificada y la señal de banda alta ajustada emitida mediante el ajustador 701b. La entrada en el ajustador 701b se proporciona por un parcheador que se opera para obtener la señal de banda alta desde la señal de banda baja tal como por la replicación de banda espectral o, generalmente, por extensión de ancho de banda. El parcheo realizado mediante el parcheador 701a puede ser un parcheo realizado de una manera armónica o no armónica. La señal generada por el parcheador 701a se ajusta, posteriormente, por el ajustador 701b utilizando la información de extensión de ancho de banda paramétrica.

60 Como se indica en la Fig. 8 y Fig. 9, los bloques descritos pueden tener una entrada de control de modo en una realización preferida. Esta entrada de control de modo se obtiene desde la señal de salida de la etapa de decisión 300. En dicha realización preferida, una característica de un bloque correspondiente puede adaptarse a la salida de la etapa de decisión, es decir, si, en una realización preferida, se lleva a cabo una decisión del habla o una decisión de música para una cierta porción de tiempo de la señal de audio. Preferentemente, el control de modo sólo se

refiere a una o más de las funcionalidades de estos bloques pero a todas las funcionalidades de los bloques. Por ejemplo, la decisión puede influenciar sólo el parcheador 701a pero puede no influenciar los otros bloques en la Fig. 9, o puede, por ejemplo, influenciar sólo el analizador de parámetro multicanal conjunto 101b en la Fig. 8 pero no los demás bloques en la Fig. 8. Esta implementación se prefiere de manera que se obtiene una mayor flexibilidad y mayor calidad y menor señal de salida de tasa de bits proporcionando flexibilidad en la etapa de pre-procesamiento común. Por otro lado, sin embargo, el uso de algoritmos en la etapa de pre-procesamiento común para ambos tipos de señales permite la implementación de un esquema de codificación/decodificación eficiente.

La Fig. 10a y Fig. 10b ilustra dos implementaciones diferentes de la etapa de decisión 300. En la Fig. 10a, se indica una decisión de bucle abierto. En este caso, el analizador de señal 300a en la etapa de decisión posee ciertas reglas con el fin de decidir si la porción cierta de tiempo o una cierta porción de frecuencia de la señal de entrada posee una característica que requiere que esta porción de señal se codifique por la primera ramificación de codificación 400 o por la segunda ramificación de codificación 500. Para este fin, el analizador de señal 300a puede analizar la señal de entrada de audio en la etapa de pre-procesamiento común o puede analizar la salida de señal de audio por la etapa de pre-procesamiento común, es decir, la señal intermedia de audio o puede analizar una señal intermedia dentro de la etapa de pre-procesamiento común tal como la emisión de la señal de mezcla descendente que puede ser una señal mono o una señal con canales k indicados en la Fig. 8. En el lado de la salida, el analizador de señal 300a genera la decisión de conmutación para controlar el conmutador 200 en el lado del codificador y el conmutador correspondiente 600 o el combinador 600 en el lado del decodificador.

Aunque no se analiza en detalle para el segundo conmutador 521, debe destacarse que el segundo conmutador 521 puede colocarse de modo similar al primer conmutador 200 como se analiza en relación a las Fig. 4a y Fig. 4b. De esta manera, una posición alternativa del conmutador 521 en la Fig. 3c se encuentra en la salida de ambas ramificaciones de los procesamientos 522, 523, 524 de modo que, ambas ramificaciones de los procesamientos operen en paralelo y sólo la salida de una ramificación de procesamiento se escribe en una secuencia de bits a través de un formador de secuencia de bits que no se ilustra en la Fig. 3c.

Adicionalmente, el segundo combinador 600 puede tener una funcionalidad de transición gradual específica como se analiza en la Fig. 4c. Como alternativa o además, el primer combinador 532 podría tener la misma una funcionalidad de transición gradual. Adicionalmente, ambos combinadores pueden tener la misma una funcionalidad de transición gradual o pueden tener diferentes funcionalidades de transición gradual o pueden no tener funcionalidades de transición gradual para que ambos combinadores sean conmutadores sin una funcionalidad de transición gradual adicional.

Como se analizó anteriormente, ambos conmutadores pueden controlarse a través de una decisión de bucle abierto o de bucle cerrado como se analizó en relación a la Fig. 10a y Fig. 10b, donde el controlador 300, 525 de la Fig. 3c puede tener diferentes o iguales funcionalidades para ambos conmutadores.

Adicionalmente, una funcionalidad de compresión en el tiempo la cual es una señal adaptativa puede existir no sólo en la primera ramificación de codificación o primera ramificación de decodificación sino también puede existir en la segunda ramificación de procesamiento de la segunda ramificación de codificación en el lado del codificador así como en el lado del decodificador. Dependiendo de una señal procesada, ambas funcionalidades de compresión en el tiempo pueden tener la misma información de compresión en el tiempo para que se aplique la misma de compresión en el tiempo a señales en el primer dominio y en el segundo dominio. Esto ahorra la carga de procesamiento y podría ser útil en algunas instancias, en casos donde bloques posteriores poseen una característica de compresión en el tiempo similar. En realizaciones alternativas, sin embargo, se prefiere tener estimadores de compresión de tiempo independientes para la primera ramificación de codificación y la segunda ramificación de procesamiento en la segunda ramificación de codificación.

La señal de audio codificada inventiva puede almacenarse en un medio de almacenamiento digital o transmitirse en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable tal como internet.

En una realización diferente, el conmutador 200 de la Fig. 1a o 2a conmuta entre dos ramificaciones de codificación 400, 500. En otra realización, puede haber ramificaciones de codificación adicionales como una tercera ramificación de codificación o incluso una cuarta ramificación de codificación o incluso más ramificaciones de codificación. En el lado del decodificador, el conmutador 600 de la Fig. 1b o 2b conmuta entre las dos ramificaciones de decodificación 431, 440 y 531, 532, 533, 534, 540. En una realización adicional, puede haber ramificaciones de decodificación adicionales tal como una tercera ramificación de decodificación o incluso una cuarta ramificación de decodificación o incluso más ramificaciones de decodificación. De modo similar, los otros conmutadores 521 o 532 pueden conmutar entre más de dos algoritmos de codificación diferentes, cuando se proporcionan tales ramificaciones de codificación/decodificación.

Las realizaciones descritas anteriormente son meramente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en la presente resultarán evidentes para los expertos en la técnica. Por lo tanto, se intenta limitar sólo por el alcance de las reivindicaciones de la patente inminentes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones de la presente. Dependiendo de ciertos requisitos de implementación de los métodos de la invención, los métodos de la invención pueden implementarse en hardware o software. La implementación puede realizarse utilizando un medio de almacenamiento digital, en particular, un disco, DVD o CD con señales de control electrónicamente legibles almacenadas en los mismos, que co-operan con sistemas informáticos programables de modo que se realizan los métodos de la invención. Generalmente, la presente invención es por lo tanto un producto de programa informático con un código de programa almacenado en un portador legible por una máquina, el código de programa se opera para llevar a cabo los métodos de la invención cuando el producto de programa informático se ejecuta en un ordenador. En otras palabras, los métodos de la invención son, por lo tanto, un programa informático con un código de programa para llevar a cabo al menos uno de los métodos de la invención cuando el programa informático se ejecuta en un ordenador.

15

REIVINDICACIONES

1. Un codificador de audio para codificar una señal de entrada de audio (195), estando la señal de entrada de audio en un primer dominio, que comprende:

5 una primera ramificación de codificación (400) para codificar una señal de audio utilizando un primer algoritmo de codificación para obtener una primera señal codificada;

10 una segunda ramificación de codificación (500) para codificar una señal de audio utilizando un segundo algoritmo de codificación para obtener una segunda señal codificada, donde el primer algoritmo de codificación es diferente del segundo algoritmo de codificación; y

un primer conmutador (200) para conmutar entre la primera ramificación de codificación y la segunda ramificación de codificación de modo que, para una porción de la señal de entrada de audio, ya sea la primera señal codificada o la segunda señal codificada está comprendida en una señal de salida del codificador, en el que la segunda ramificación de codificación comprende:

15 un convertidor (510) para convertir la señal de audio en un segundo dominio diferente del primer dominio, una primera ramificación de procesamiento (522) para procesar una señal de audio en el segundo dominio para obtener una primer señal procesada;

20 una segunda ramificación de procesamiento (523, 524) para convertir una señal de audio en un tercer dominio diferente del primer dominio y del segundo dominio y para procesar la señal en el tercer dominio para obtener una segunda señal procesada; y

25 un segundo conmutador (521) para conmutar entre la primera ramificación de procesamiento (522) y la segunda ramificación de procesamiento (523, 524) de modo que, para una porción de la entrada de la señal de audio en la segunda ramificación de codificación, ya sea la primera señal procesada o la segunda señal procesada está comprendida en la segunda señal codificada.

2. Codificador de audio de acuerdo con la reivindicación 1, en el que el primer algoritmo de codificación en la primera ramificación de codificación (400) se basa en un modelo de sumidero de información, o en el que el segundo algoritmo de codificación en la segunda ramificación de codificación (500) se basa en una fuente de información o en un modelo de relación de señal a ruido.

3. Codificador de audio de acuerdo con la reivindicación 1 o 2, en el que la primera ramificación de codificación comprende un convertidor (410) para convertir la señal de entrada de audio en un cuarto dominio diferente del primer dominio, del segundo dominio, y del tercer dominio.

4. Codificador de audio de acuerdo con una de las reivindicaciones precedentes, en el que el primer dominio es el dominio de tiempo, el segundo dominio es un dominio LPC obtenido al filtrar de primera señal de dominio LPC, el tercer dominio es un dominio espectral LPC obtenido al convertir una señal LPC filtrada en un dominio espectral, y el cuarto dominio es un dominio espectral obtenido por el dominio de frecuencia al convertir la primera señal de dominio.

5. Codificador de audio de acuerdo una de las reivindicaciones precedentes, que además comprende un controlador (300, 525) para controlar el primer conmutador (200) o el segundo conmutador (521) en un modo adaptativo de señal,

45 en el que el controlador es operativo para analizar una señal introducida en el primer conmutador (200) o emitida por la primera ramificación de codificación o la segunda ramificación de codificación o una señal obtenida por la decodificación de una señal de salida de la primera ramificación de codificación o la segunda ramificación de codificación con respecto a una función objetivo, o

50 en el que el controlador (300, 525) es operativo para analizar una señal introducida en el segundo conmutador (521) o emitida por la primera ramificación de procesamiento o la segunda ramificación de procesamiento o señales obtenidas por señales de salida de procesamiento inverso desde la primera ramificación de procesamiento (522) y la segunda ramificación de procesamiento (523, 524) con respecto a una función objetivo.

6. Codificador de audio de acuerdo con una de las reivindicaciones precedentes, en el que, la primera ramificación de codificación (400) o la segunda ramificación de procesamiento (523, 524) de la segunda ramificación de codificación (500) comprende un convertidor de tiempo/frecuencia que introduce solapamiento y una etapa de cuantificador/codificador por entropía (421) y en el que la primera ramificación de procesamiento de la segunda ramificación de codificación comprende una etapa de cuantificador o de codificador por entropía (522) sin una conversión que introduce que introduce solapamiento.

7. Codificador de audio de acuerdo con la reivindicación 6, en el que el convertidor de tiempo/frecuencia que introduce solapamiento comprende un generador de partición en ventanas para aplicar una ventana de análisis y un algoritmo de transformación de coseno discreta modificada (MDCT), siendo operativo el generador de partición en

ventanas para aplicar la función de ventana a tramas posteriores de manera solapada de modo que una muestra de señal de entrada en el generador de partición en ventanas tenga lugar en al menos dos tramas posteriores.

5 8. Codificador de audio de acuerdo con una de las reivindicaciones precedentes, en el que la primera ramificación de procesamiento (522) comprende la codificación de excitación LPC de un codificador de predicción lineal con excitación por código algebraico y la segunda ramificación de procesamiento comprende un convertidor espectral MDCT y un cuantificador para cuantificar componentes espectrales para obtener componentes espectrales cuantificados, en el que cada componente espectral cuantificado es cero o se define por un índice de cuantificación de una pluralidad de índices de cuantificación.

10 9. Codificador de audio de acuerdo con la reivindicación 5, en el que el controlador es operativo para controlar el primer conmutador (200) en una manera de bucle abierto y para controlar el segundo conmutador (521) en una manera de bucle cerrado.

15 10. Codificador de audio de acuerdo con una de las reivindicaciones precedentes, en el que la primera ramificación de codificación y la segunda ramificación de codificación son operativas para codificar la señal de audio en una manera de sentido en bloque, en el que el primer conmutador o el segundo conmutador se conmutan en una manera de sentido en bloque de modo que tenga lugar una acción de conmutación, al mínimo, después de un bloque de una cantidad predefinida de muestras de una señal, formando la cantidad predefinida de muestras una longitud de trama para el conmutador correspondiente (521, 200).

20 11. Codificador de audio de acuerdo con la reivindicación 10, en el que la longitud de la trama para el primer conmutador es al menos el doble de tamaño de la longitud de la trama del segundo conmutador.

25 12. Codificador de audio de acuerdo con la reivindicación 5, en el que el controlador es operativo para realizar una discriminación del habla/música de tal manera que se favorece una decisión para el habla con respecto a una decisión a música de modo que se toma una decisión para el habla incluso cuando una porción menor del 50 % de una trama para el primer conmutador es del habla y una porción mayor del 50 % de la trama para el primer conmutador es de música.

30 13. Codificador de audio de acuerdo con la reivindicación 5 o 12, en el que una trama para el segundo conmutador es menor a la trama para el primer conmutador, y en el que el controlador (525, 300) es operativo para tomar una decisión para el habla cuando se encuentra que únicamente una porción de la primera trama que tiene una longitud que es mayor del 50 % de la longitud de la segunda trama incluye música.

35 14. Codificador de audio de acuerdo con una de las reivindicaciones precedentes, en el que, la primera ramificación de codificación (400) o la segunda ramificación de procesamiento de la segunda ramificación de codificación incluye una funcionalidad de compresión en el tiempo variable.

40 15. Método de codificación de una señal de entrada de audio (195), estando la señal de entrada de audio en un primer dominio, que comprende:

45 codificar (400) una señal de audio utilizando un primer algoritmo de codificación para obtener una primera señal codificada;

codificar (500) una señal de audio utilizando un segundo algoritmo de codificación para obtener una segunda señal codificada, en el que el primer algoritmo de codificación es diferente del segundo algoritmo de codificación;

50 y conmutar (200) entre la codificación que utiliza el primer algoritmo de codificación y la codificación que utiliza el segundo algoritmo de codificación de modo que, para una porción de la señal de entrada de audio, ya sea la primera señal codificada o la segunda señal codificada esté comprendida en una señal de salida codificada, en el que codificar (500) utilizando el segundo algoritmo de codificación comprende:

55 convertir (510) la señal de audio en un segundo dominio diferente del primer dominio, procesar (522) una señal de audio en el segundo dominio para obtener una primera señal procesada; convertir (523) una señal en un tercer dominio diferente del primer dominio y del segundo dominio y procesar (524) la señal en el tercer dominio para obtener una segunda señal procesada; y conmutar (521) entre procesar (522) la señal de audio y convertir (523) y procesar (524) de modo que, para una porción de la señal de audio codificada utilizando el segundo algoritmo de codificación, ya sea la primera señal procesada o la segunda señal procesada esté comprendida en la segunda señal codificada.

60 16. Decodificador para decodificar una señal de audio codificada, comprendiendo la señal de audio codificada una primera señal codificada, una primera señal procesada en un segundo dominio, y una segunda señal procesada en un tercer dominio, en el que la primera señal codificada, la primera señal procesada, y la segunda señal procesada

están relacionadas con diferentes porciones de tiempo de una señal de audio decodificada, y en el que un primer dominio, el segundo dominio y el tercer dominio son diferentes entre sí, que comprende:

- 5 una primera ramificación de decodificación (431, 440) para decodificar la primera señal codificada basándose en un primer algoritmo de codificación;
una segunda ramificación de decodificación para decodificar la primera señal procesada o la segunda señal procesada,
en el que la segunda ramificación de decodificación comprende
- 10 una primera ramificación de procesamiento inversa (531) para el procesamiento inverso de la primera señal procesada para obtener una primera señal procesada inversa en el segundo dominio;
una segunda ramificación de procesamiento inversa (533, 534) para el procesamiento inverso de la segunda señal procesada para obtener una segunda señal procesada inversa en el segundo dominio;
15 un primer combinador (532) para combinar la primera señal procesada inversa y la segunda señal procesada inversa para obtener una señal combinada en el segundo dominio; y
un convertidor (540) para convertir la señal combinada al primer dominio; y
- un segundo combinador (600) para combinar la señal convertida en el primer dominio y la primera señal decodificada emitida por la primera ramificación de decodificación para obtener la señal de audio decodificada en el primer dominio.
- 20 17. El decodificador de la reivindicación 16, en el que el primer combinador (532) o el segundo combinador (600) comprende un conmutador con funcionalidad de transición gradual.
- 25 18. El decodificador de la reivindicación 16 o 17, en el que el primer dominio es un dominio de tiempo, el segundo dominio es un dominio LPC, el tercer dominio es un dominio LPC dominio espectral, o la primera señal codificada se codifica en un cuarto dominio, que es un dominio de tiempo espectral obtenido por la conversión de una señal en tiempo/frecuencia en el primer dominio.
- 30 19. El decodificador de acuerdo con cualquiera de las reivindicaciones 16 a 18, en el que la primera ramificación de decodificación (431, 440) comprende un codificador inverso y un des-cuantificador y un convertidor de dominio de tiempo y dominio de frecuencia (440), o la segunda ramificación de decodificación comprende un codificador inverso y un des-cuantificador en la primera ramificación de procesamiento inverso o un codificador inverso y un des-cuantificador y un dominio espectral LPC a convertidor de dominio LPC (534) en la segunda ramificación de
35 procesamiento inversa.
20. El decodificador de la reivindicación 19, en el que la primera ramificación de decodificación o la segunda ramificación de procesamiento inversa comprende un sumador de solapamiento para realizar una funcionalidad de cancelación de solapamiento en dominio de tiempo.
- 40 21. El decodificador de acuerdo con una de las reivindicaciones 16 a 20, en el que la primera ramificación de decodificación o la segunda ramificación de procesamiento inversa comprende un descompresor controlado por una característica de compresión incluida en la señal de audio codificada.
- 45 22. El decodificador de acuerdo con una de las reivindicaciones 16 a 21, en el que la señal codificada comprende, como información secundaria (4a), una indicación de si una señal codificada debe ser codificada por una primera ramificación de codificación o una segunda ramificación de codificación o una primera ramificación de procesamiento de la segunda ramificación de codificación o una segunda ramificación de procesamiento de la segunda ramificación de codificación, y
50 que además comprende un analizador para analizar la señal codificada para determinar, basándose en la información secundaria (4a), si una señal codificada debe ser procesada por la primera ramificación de decodificación, o la segunda ramificación de decodificación, o la primera ramificación de procesamiento inversa de la segunda ramificación de decodificación o la segunda ramificación de procesamiento inversa de la segunda ramificación de decodificación.
- 55 23. Método de decodificación de una señal de audio codificada, comprendiendo la señal de audio codificada una primera señal codificada, una primera señal procesada en un segundo dominio, y una segunda señal procesada en un tercer dominio, en el que la primera señal codificada, la primera señal procesada, y la segunda señal procesada están relacionadas con diferentes porciones de tiempo de una señal de audio decodificada, y en el que un primer
60 dominio, el segundo dominio y el tercer dominio son diferentes entre sí, que comprende:
- decodificar (431, 440) la primera señal codificada basándose en un primer algoritmo de codificación;
decodificar la primera señal procesada o la segunda señal procesada,

en el que decodificar la primera señal procesada o la segunda señal procesada comprende:

- 5 el procesamiento inverso (531) de la primera señal procesada para obtener una primera señal procesada inversa en el segundo dominio;
- el procesamiento inverso (533, 534) de la segunda señal procesada para obtener una segunda señal procesada inversa en el segundo dominio;
- combinar (532) la primera señal procesada inversa y la segunda señal procesada inversa para obtener una señal combinada en el segundo dominio; y
- 10 convertir (540) de la señal combinada al primer dominio; y
- combinar (600) la señal convertida en el primer dominio y la primera señal decodificada para obtener la señal de audio decodificada en el primer dominio.

24. Señal de audio codificada que comprende:

- 15 una primera señal codificada o a decodificar utilizando un primer algoritmo de codificación o de decodificación, una primera señal procesada en un segundo dominio, y una segunda señal procesada en un tercer dominio, en el que la primera señal procesada y la segunda señal procesada se codifican utilizando un segundo algoritmo de codificación, en el que la primera señal codificada, la primera señal procesada, y la segunda señal procesada están relacionadas con diferentes porciones de tiempo de una señal de audio decodificada,
- 20 en el que un primer dominio, el segundo dominio y el tercer dominio son diferentes entre sí, e indicando la información secundaria (4a) si una porción de la señal de audio codificada es la primera señal codificada, la primera señal procesada o la segunda señal procesada.

- 25 Programa informático para realizar, cuando se ejecuta en el ordenador, el método de codificación de una señal de audio de acuerdo con la reivindicación 15 o el método de decodificación de una señal de audio codificada de acuerdo con la reivindicación 23.

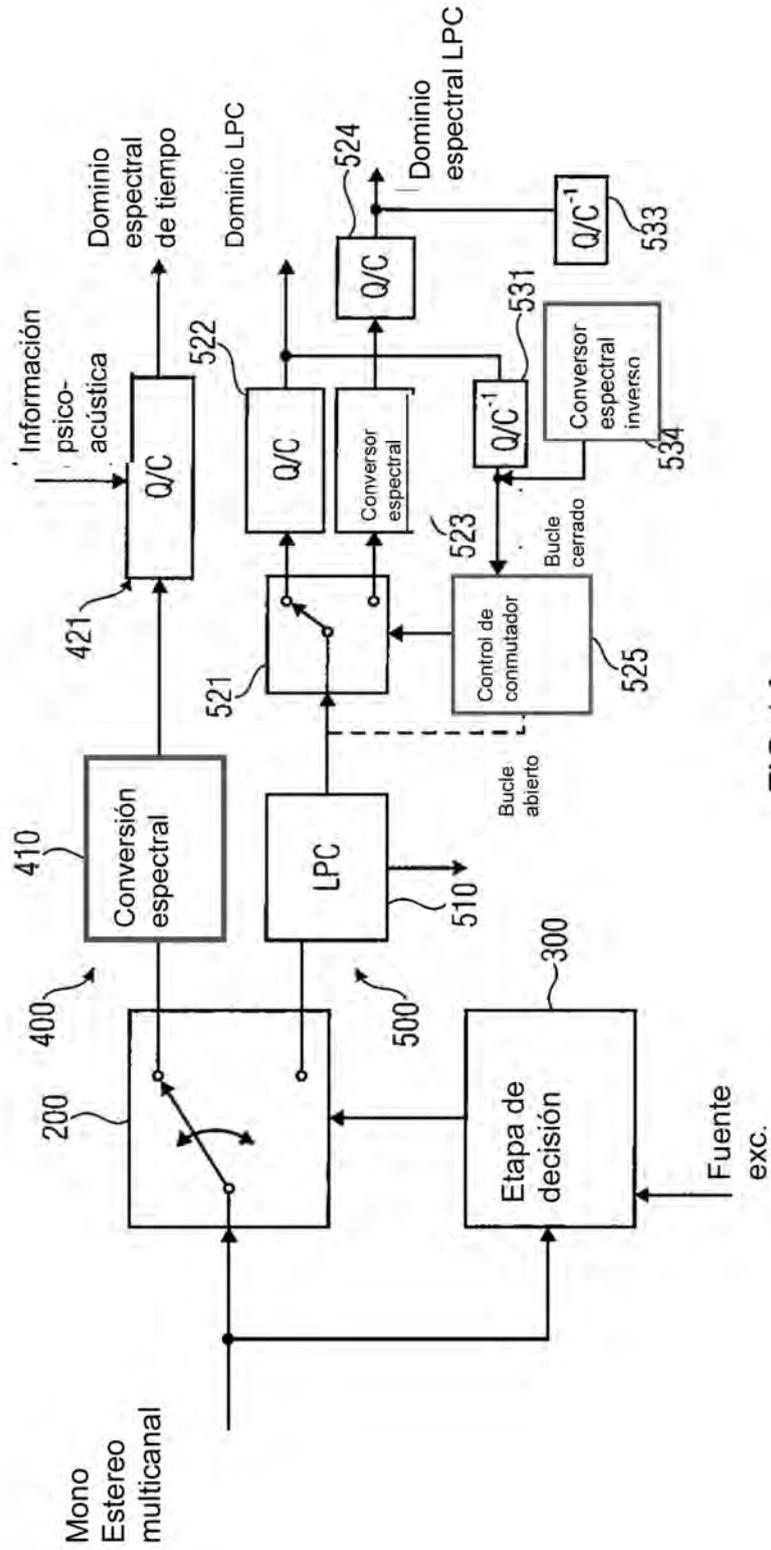


FIG 1A

(Codificador)

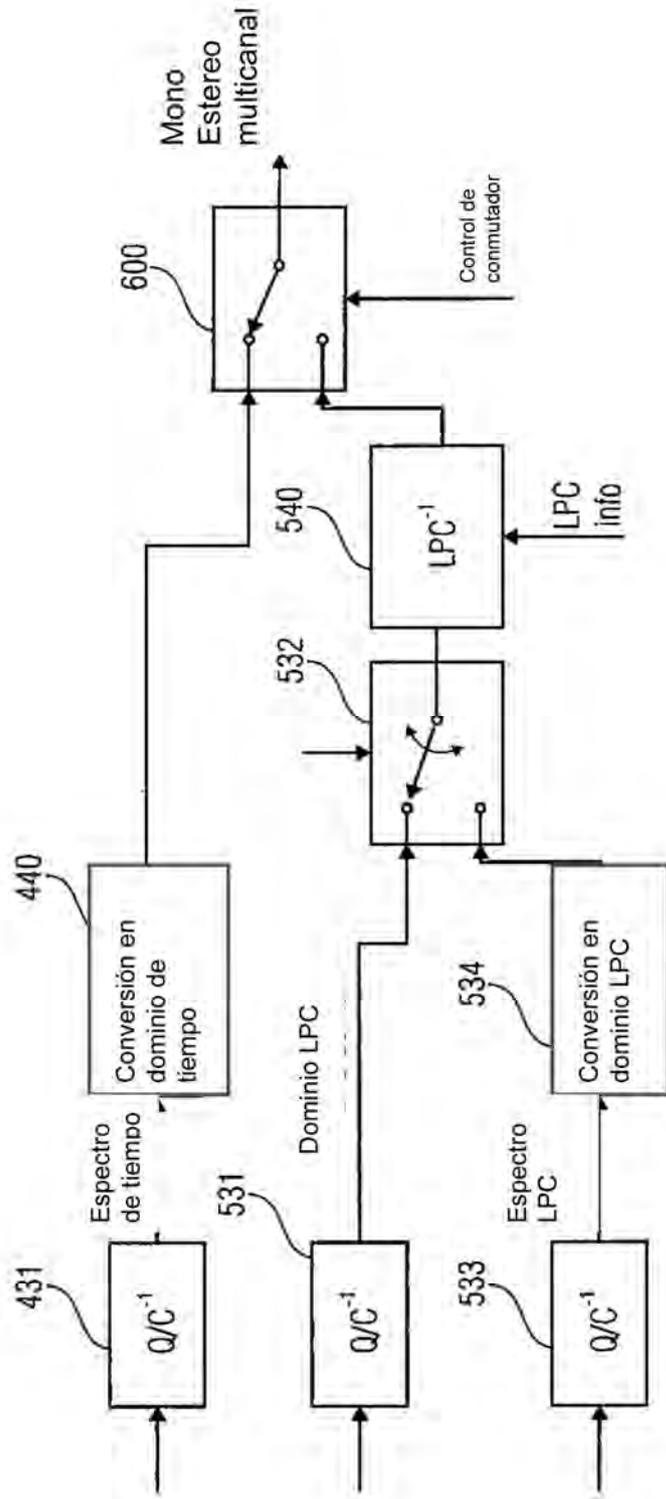


FIG 1B
(Decodificador)

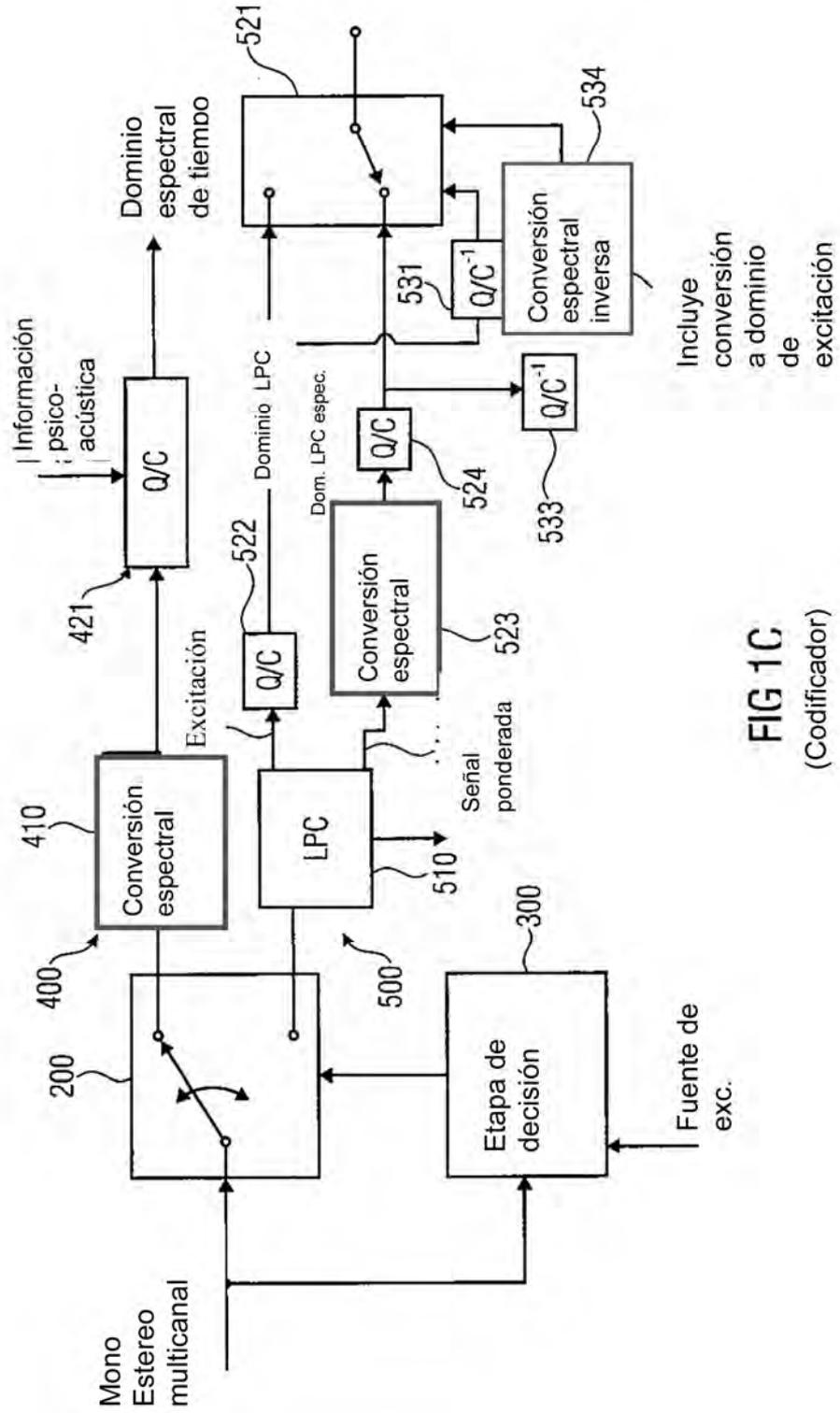


FIG 1C
(Codificador)

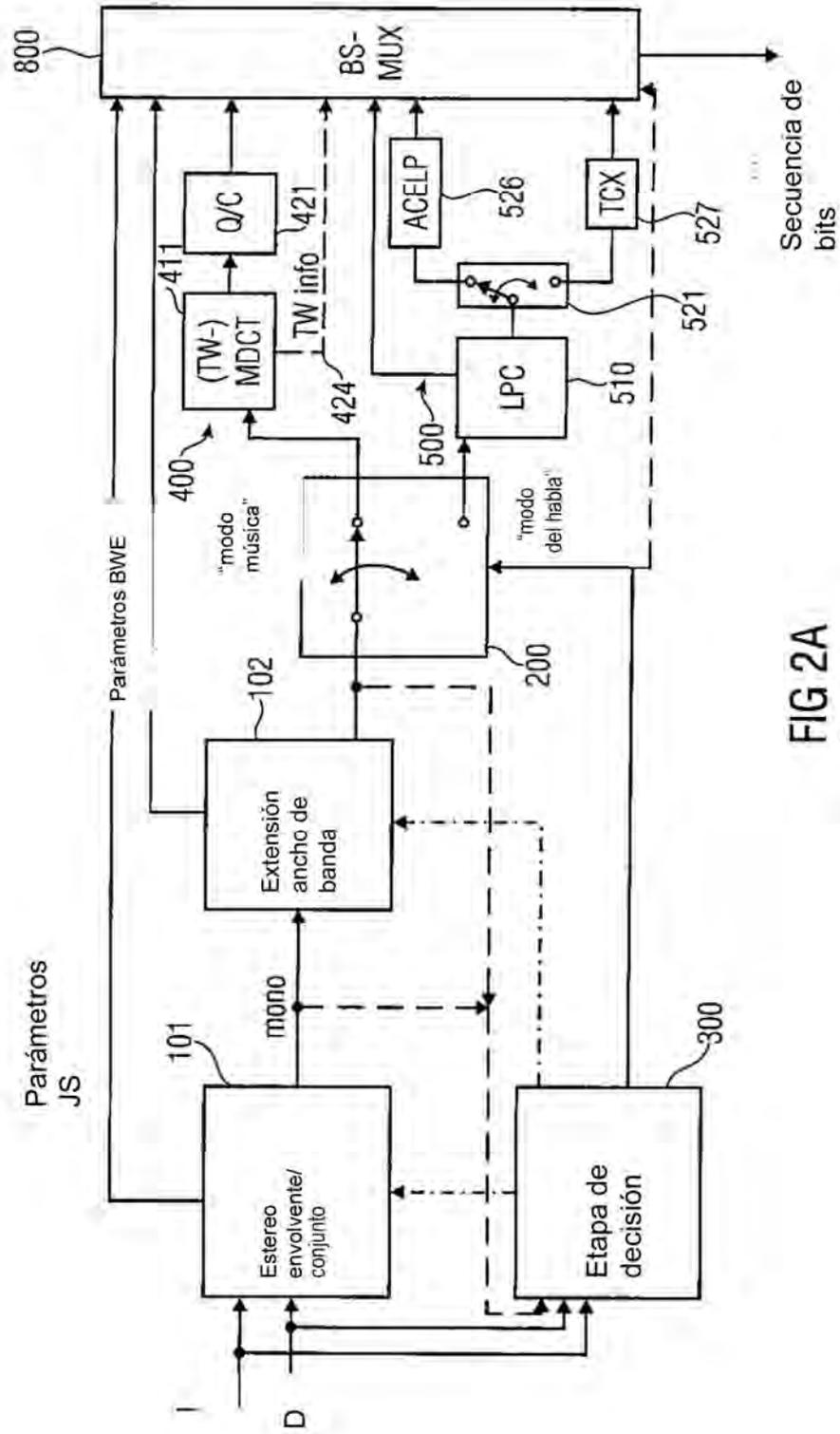


FIG 2A
(Codificador)

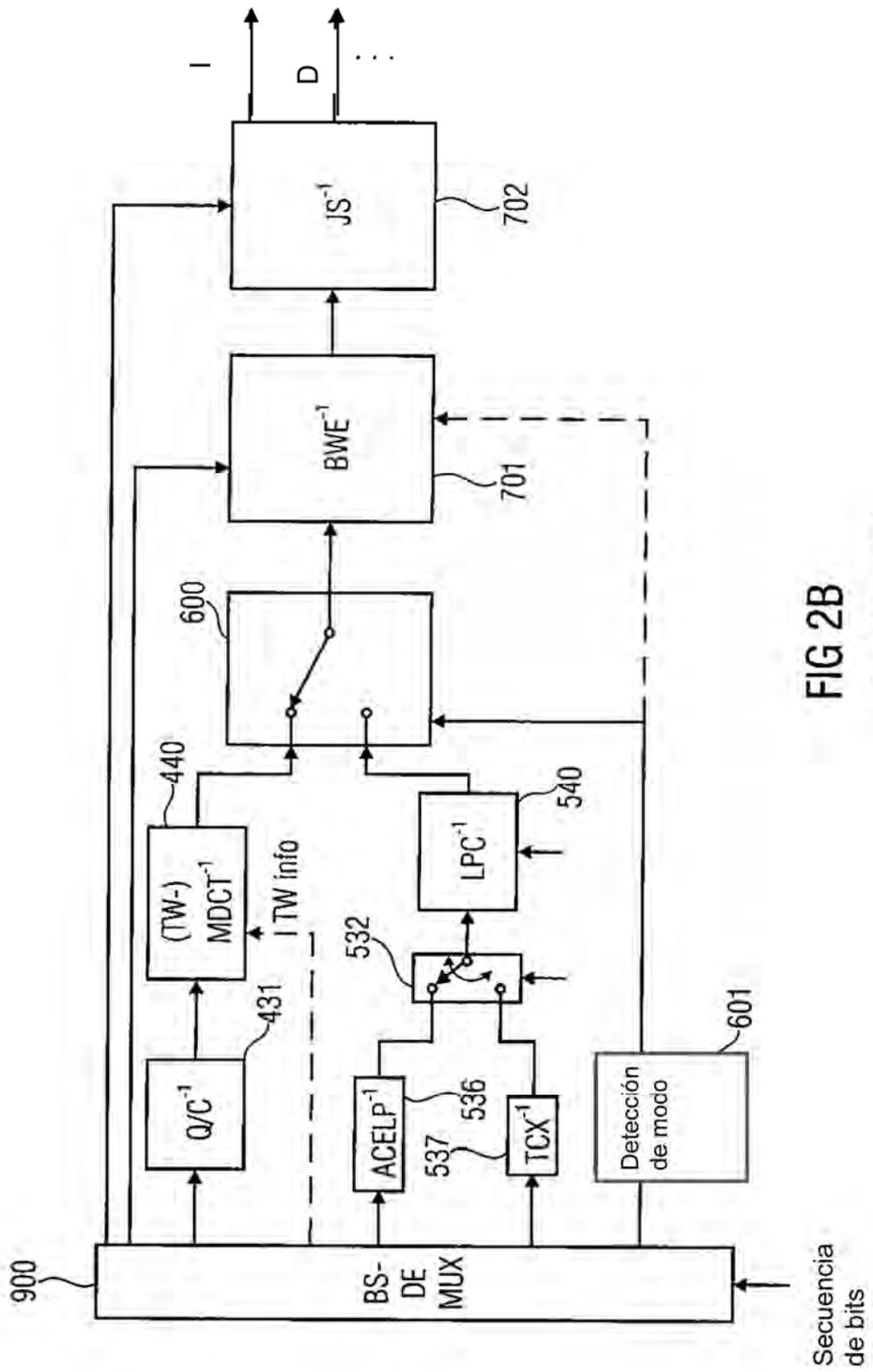


FIG 2B
(Decodificador)

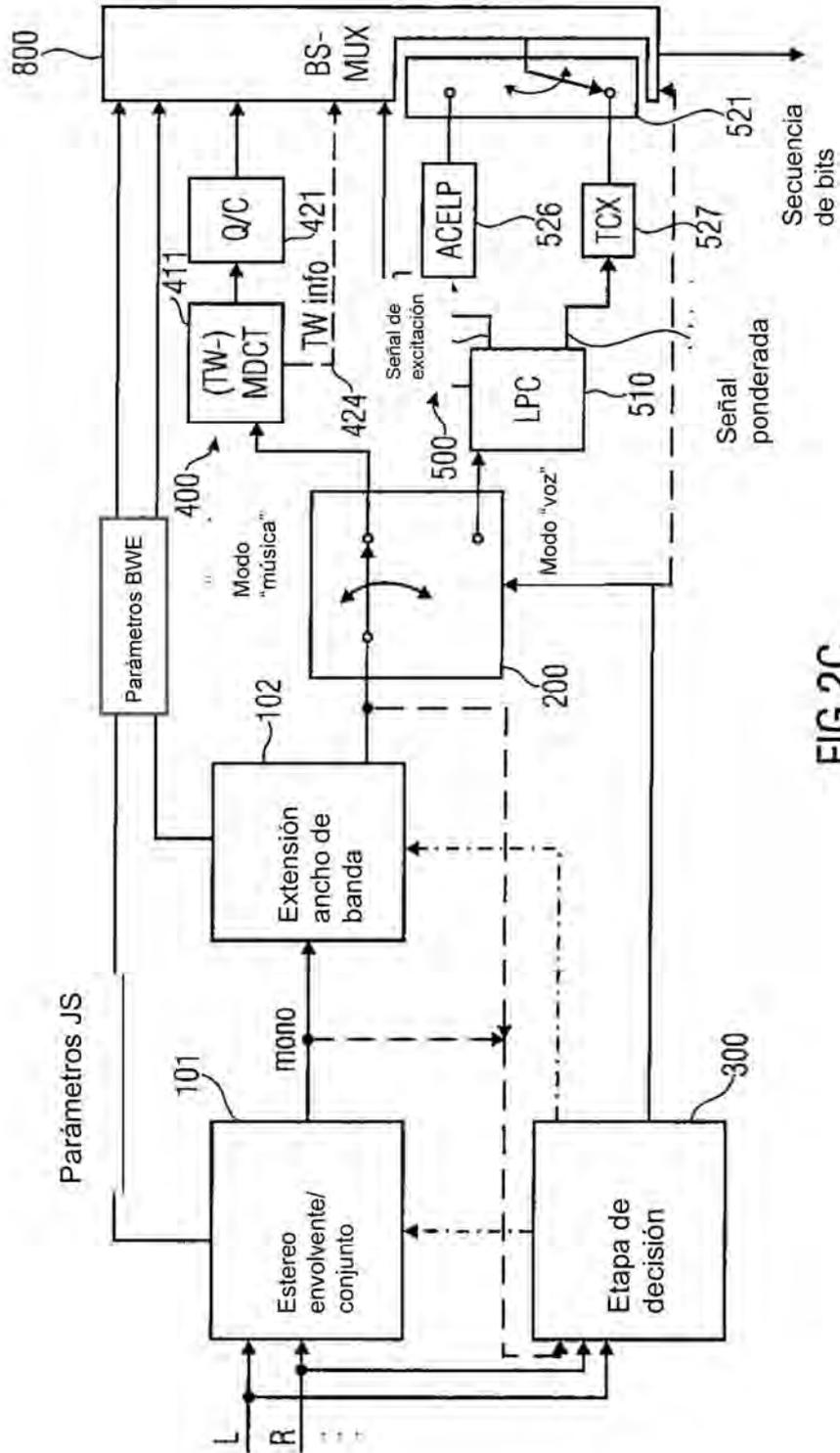


FIG 2C
(Codificador)

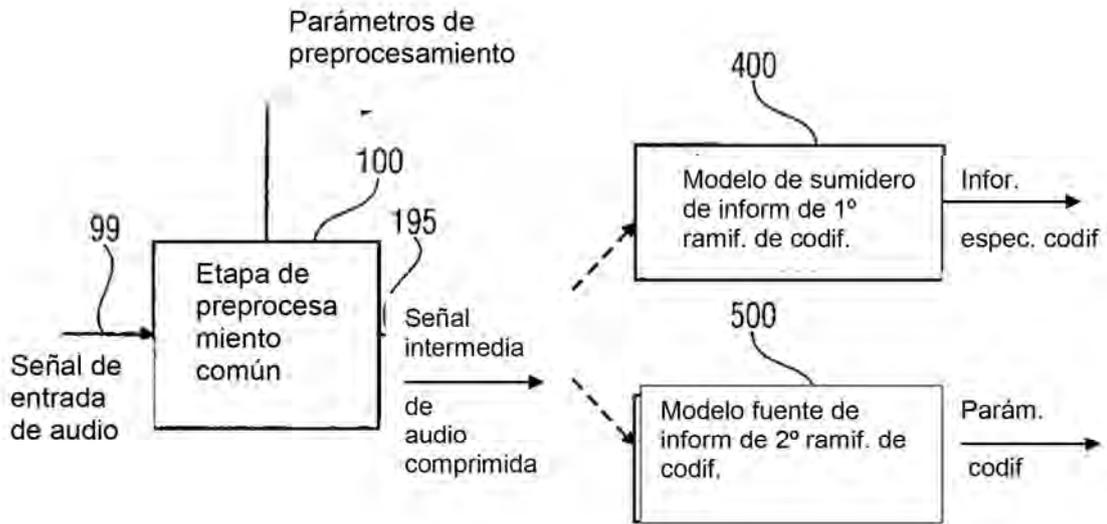


FIG 3A

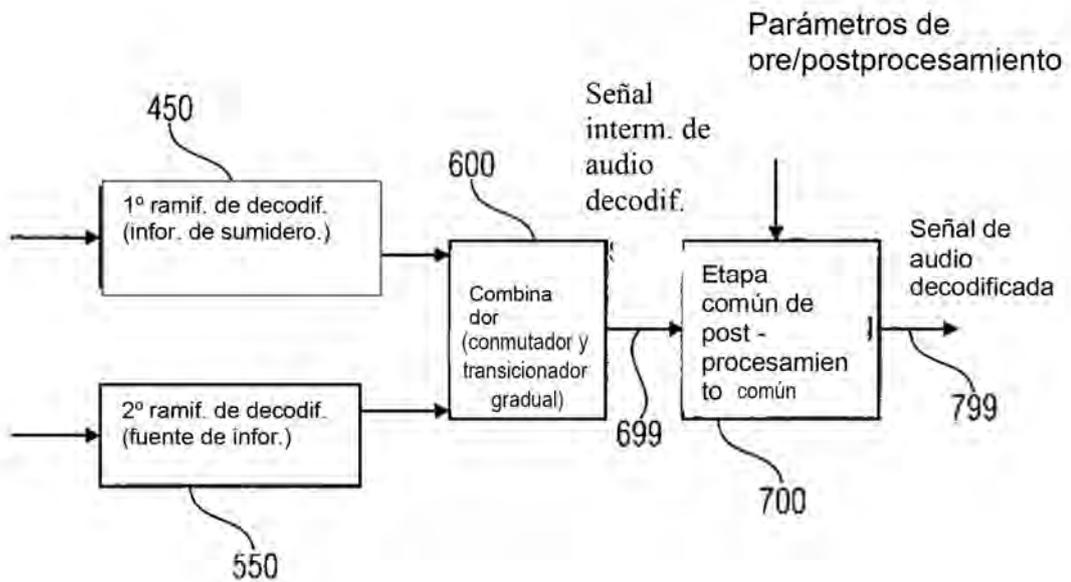
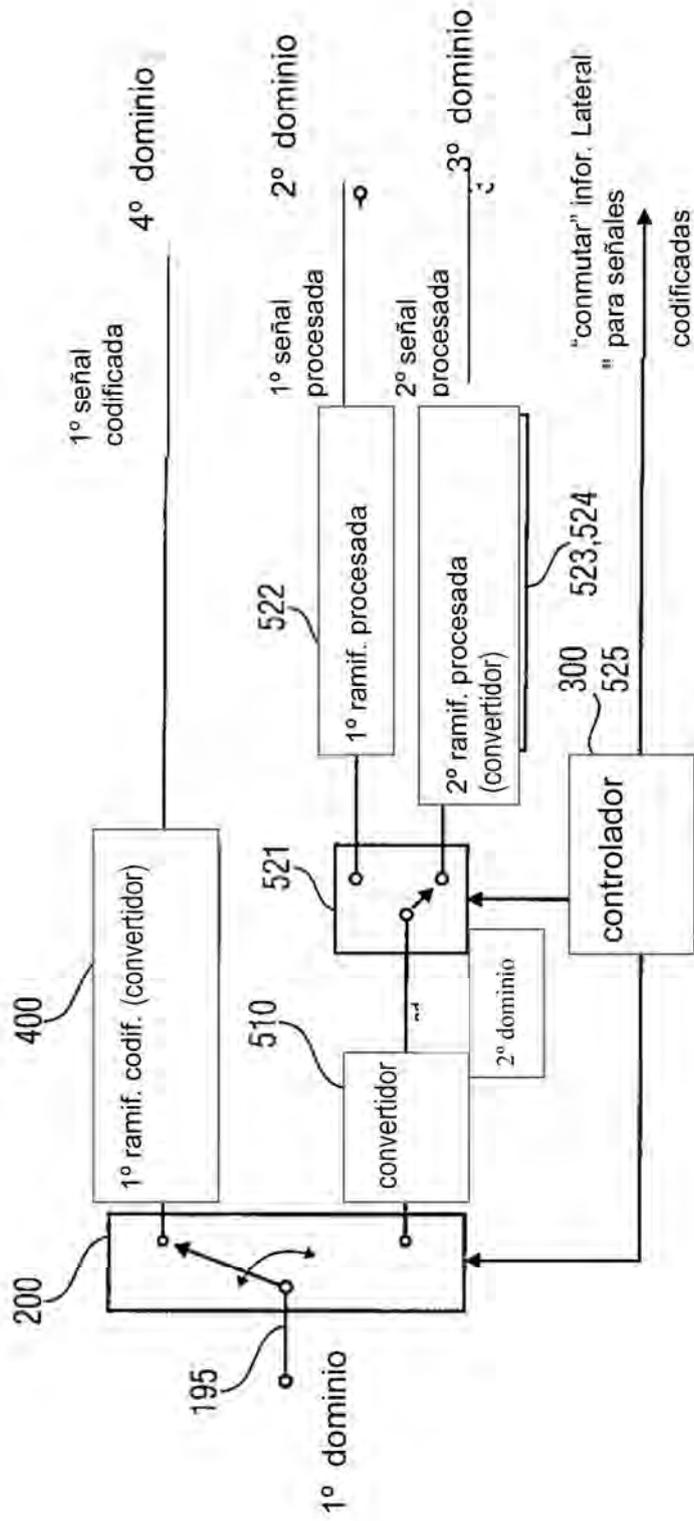


FIG 3B



Cada bloque de la señal de audio del 1º dominio se representa por un 2º dominio, 3º dominio, 4º dominio, aparte de una región de cruce opcional

FIG 3C

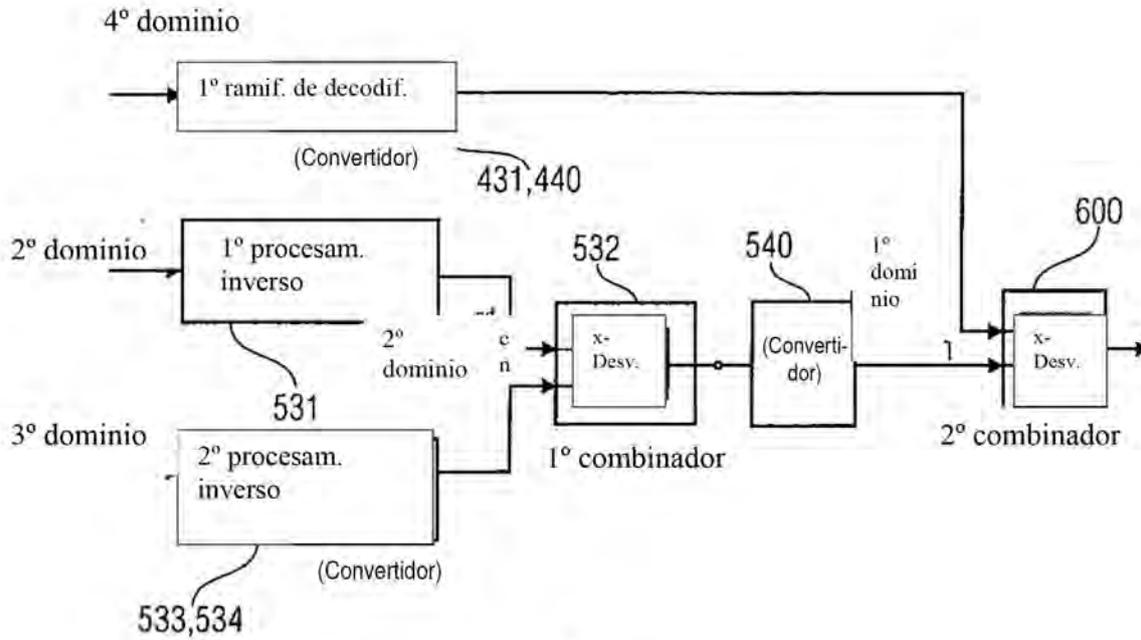


FIG 3D

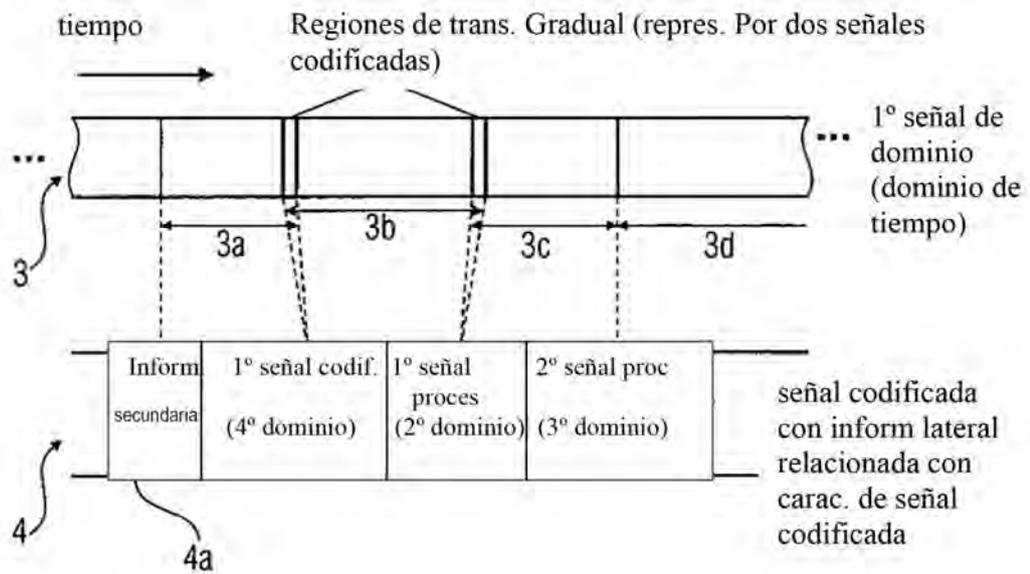


FIG 3E

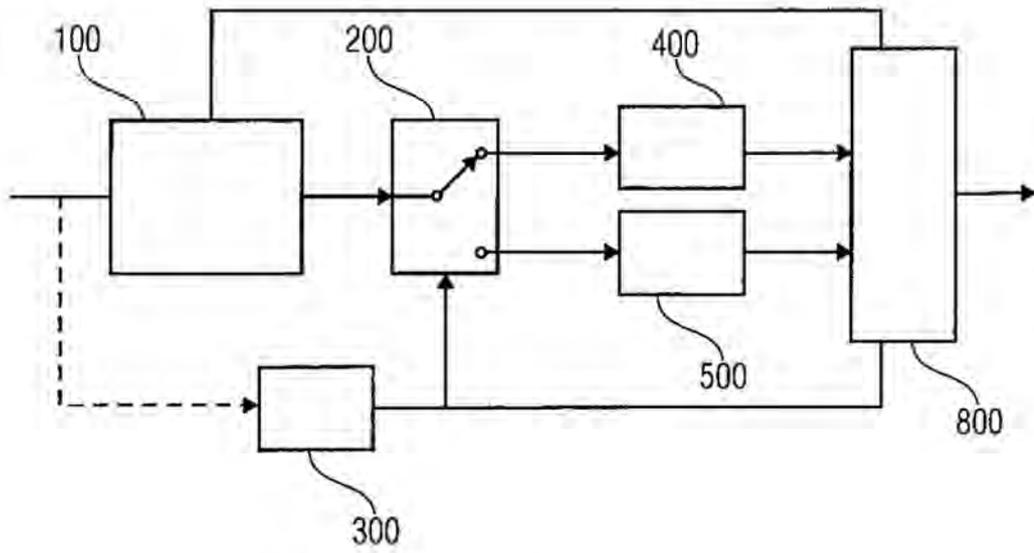


FIG 4A

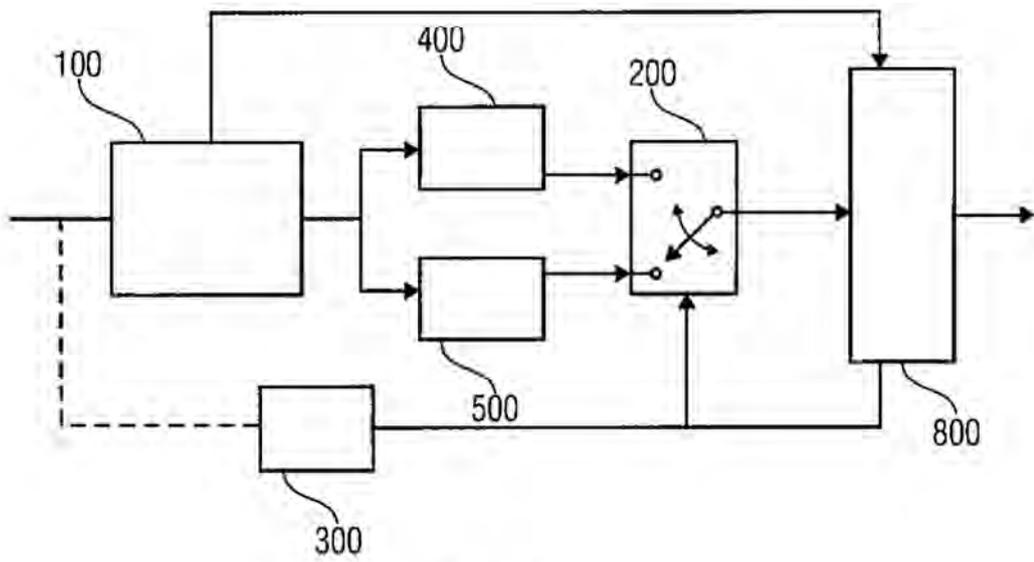


FIG 4B

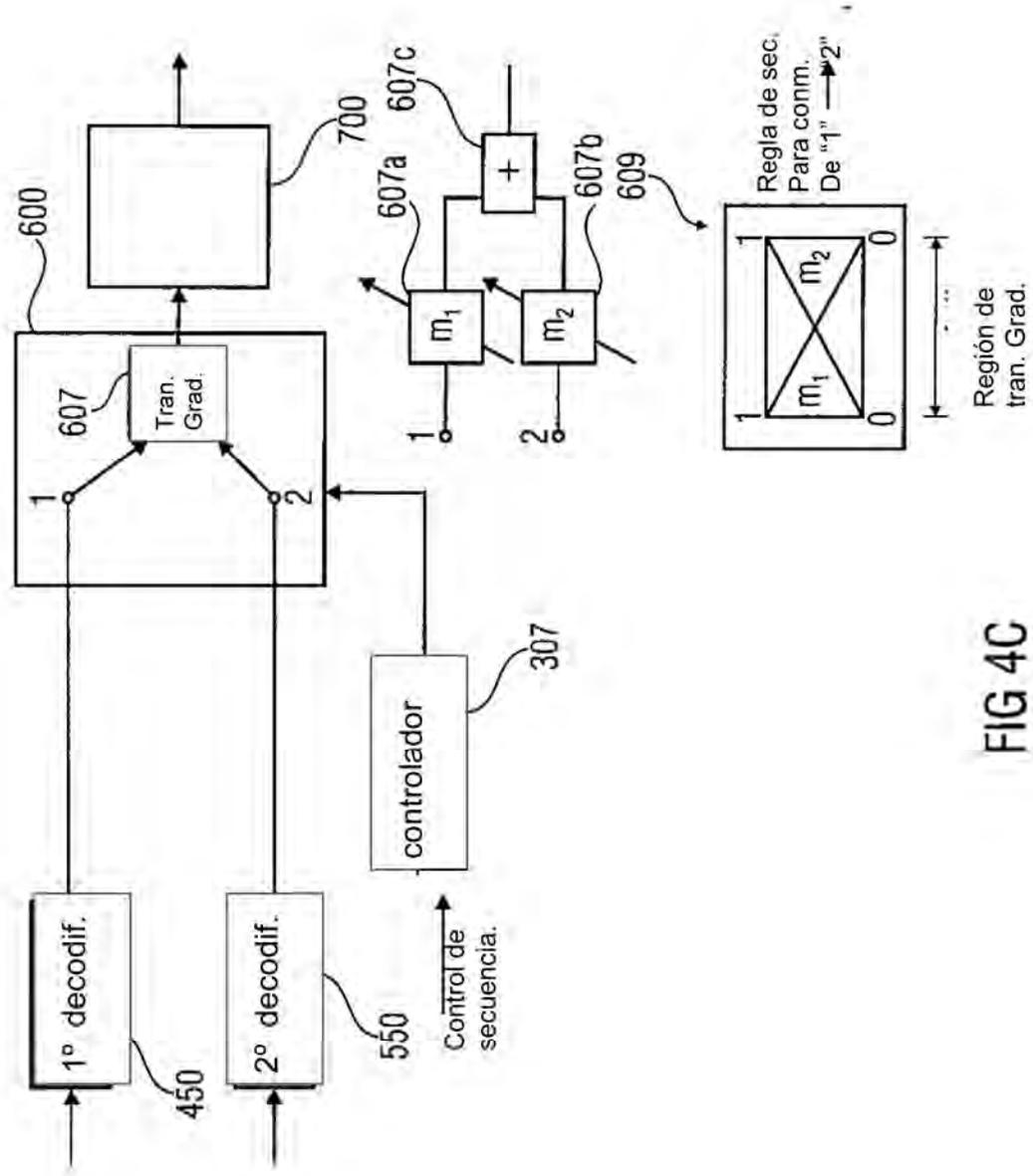


FIG 4C

Segmento de señal tipo impulso (ej. habla vocaliz.)

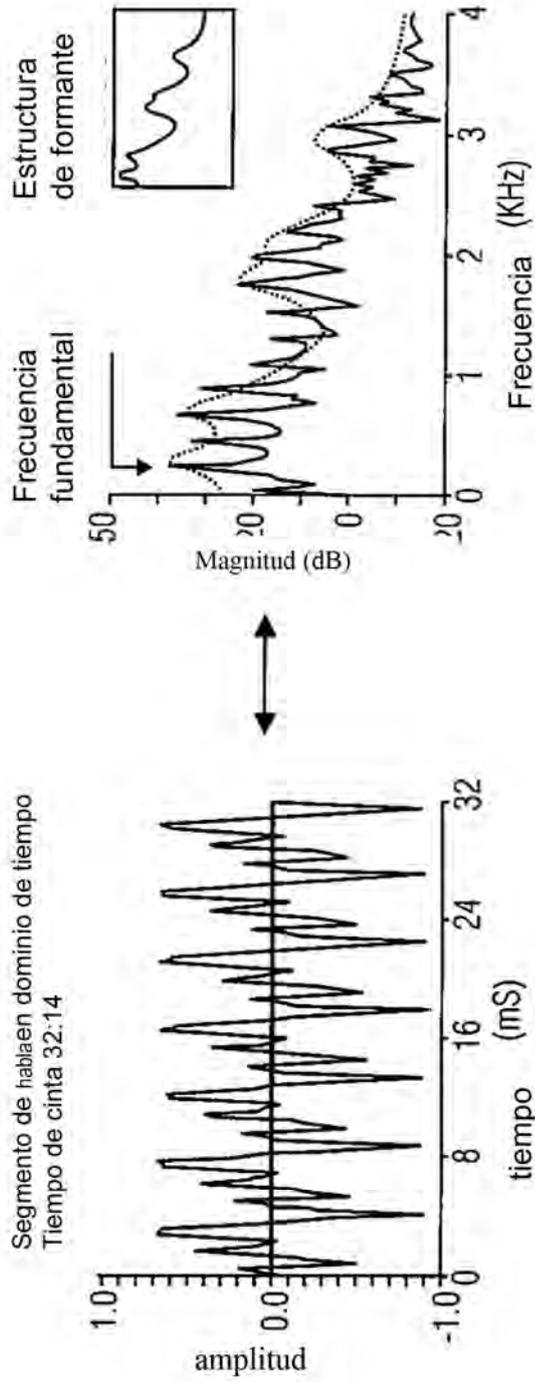


FIG 5A

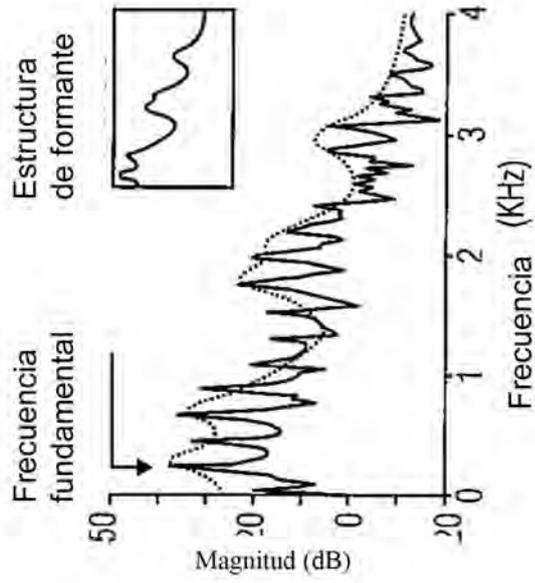


FIG 5B

Segmento estacionario (ej. habla no vocaliz.)

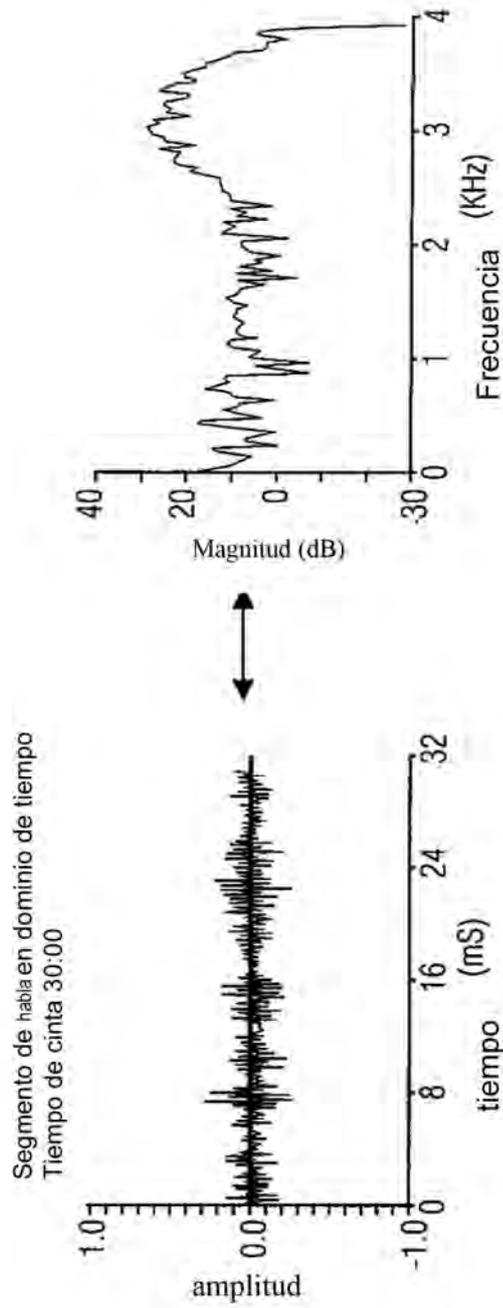
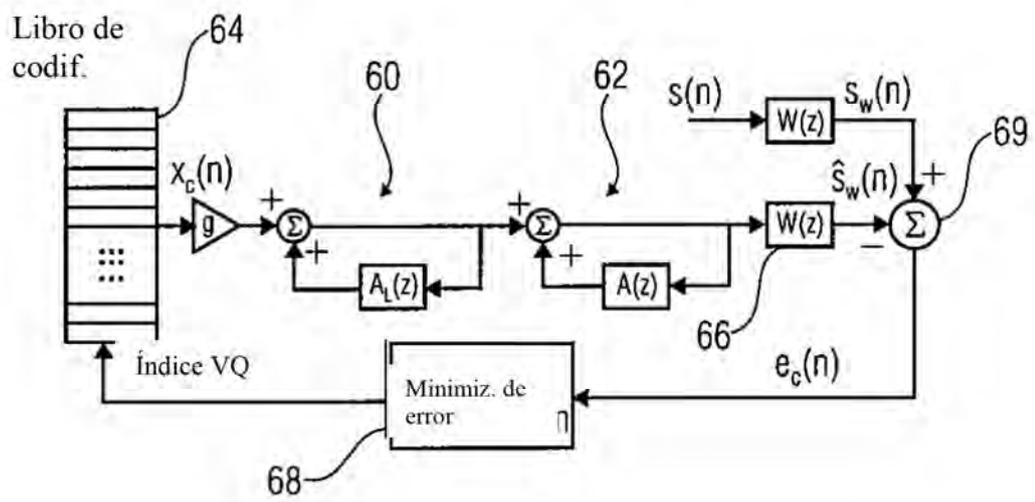


FIG 5C

FIG 5D

Análisis por síntesis CELP



$A_L(z)$ Predic. a largo plazo
 \cong Estruct. (fina) de tono

$A(z)$: Predic. a corto plazo
 \cong Estruct. Formante/envoltura spectral

FIG 6

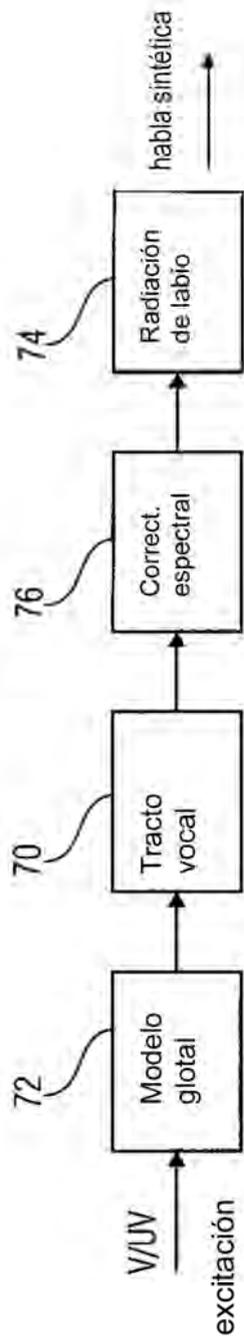


FIG 7A

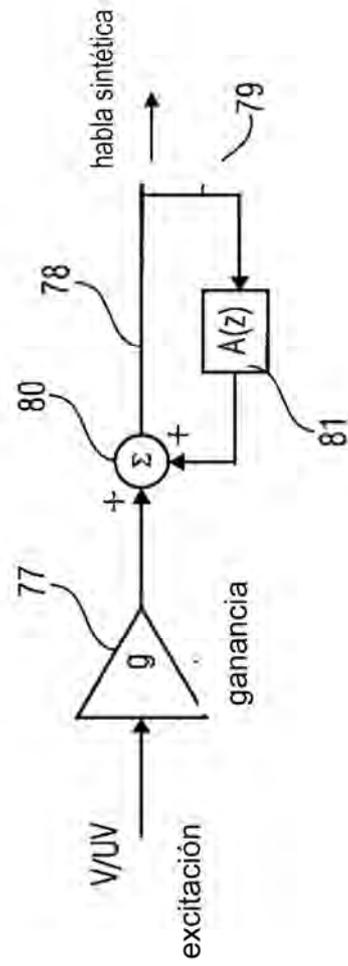


FIG 7B

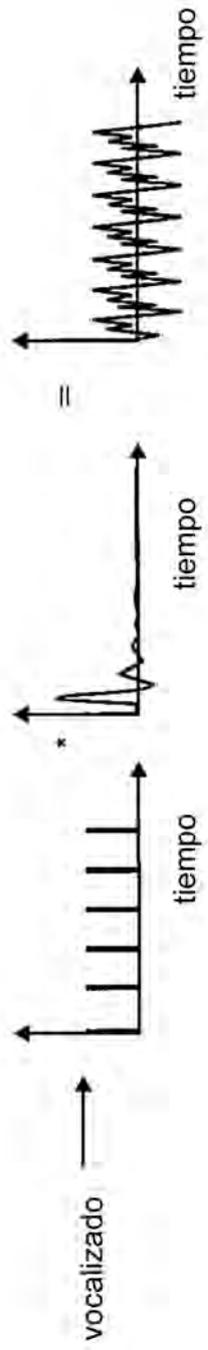


FIG 7C

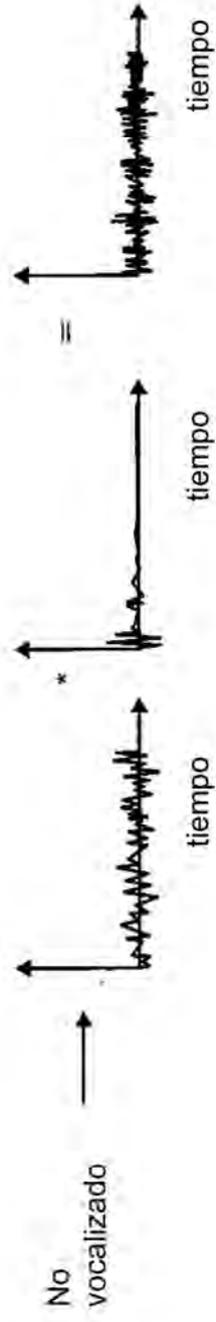


FIG 7D

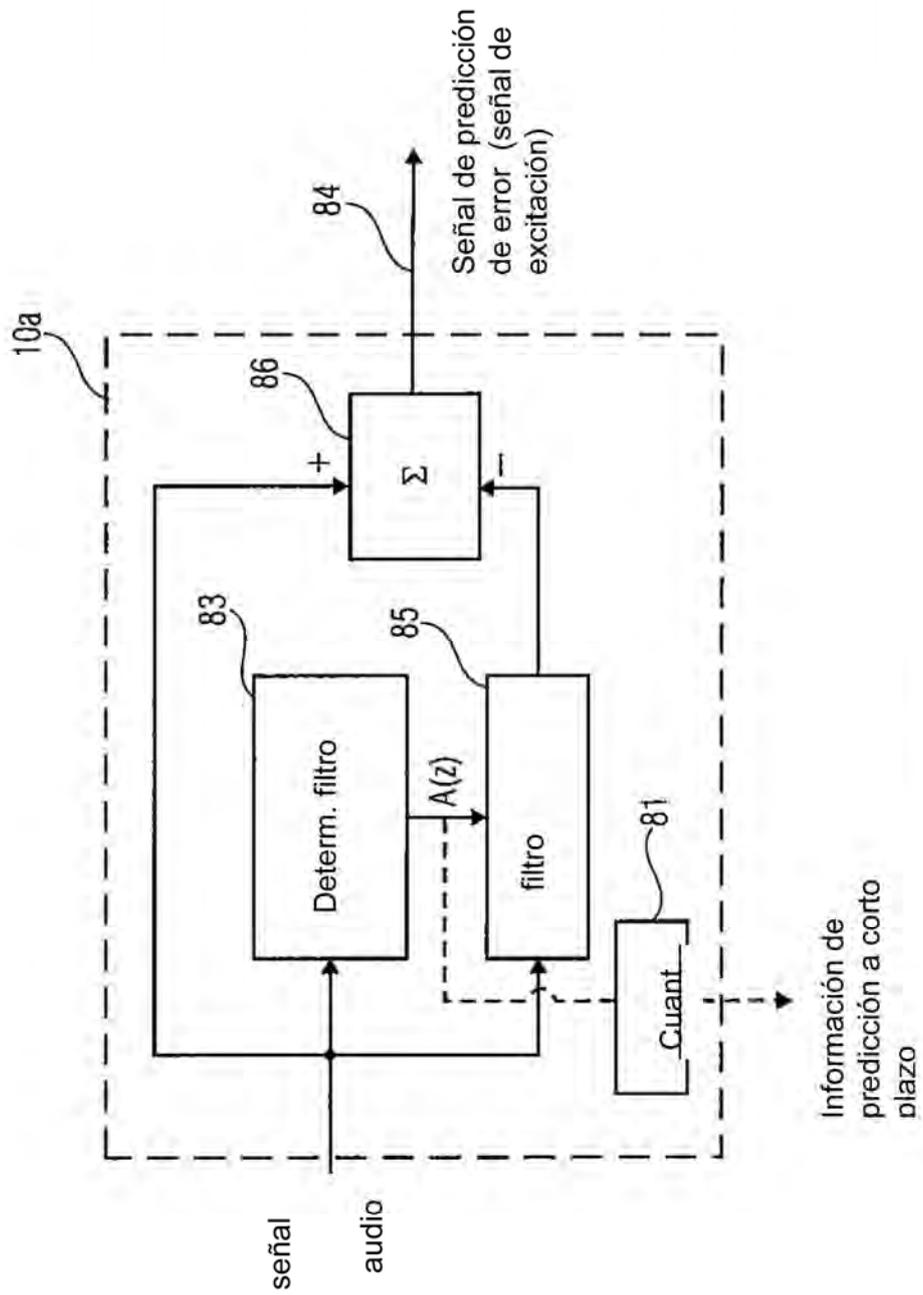


FIG 7E

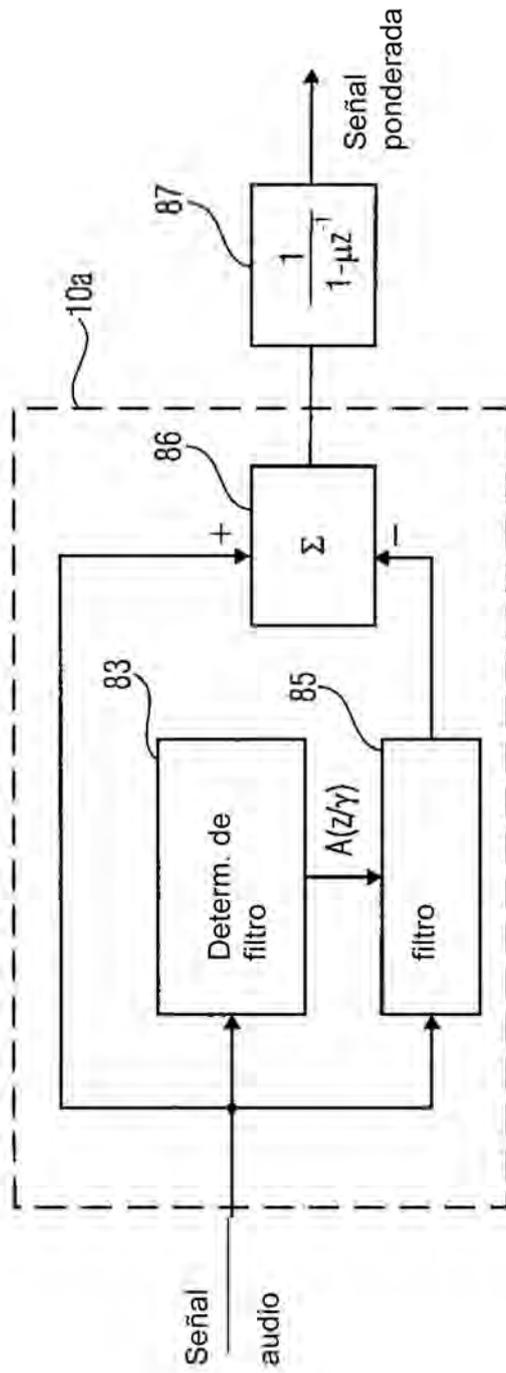


FIG 7F

(lado del codificador)

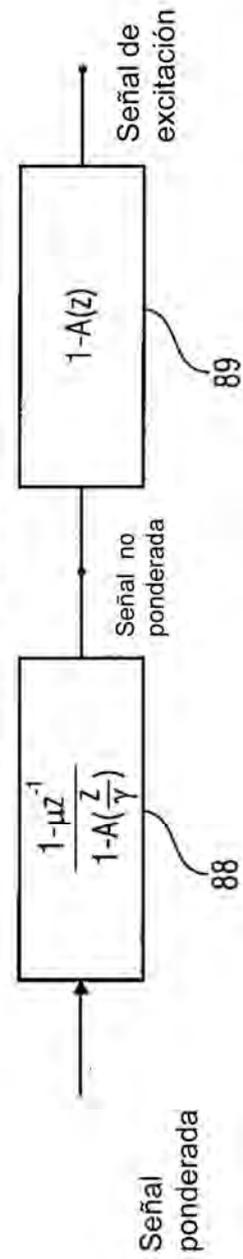


FIG 7G

(lado del codificador)

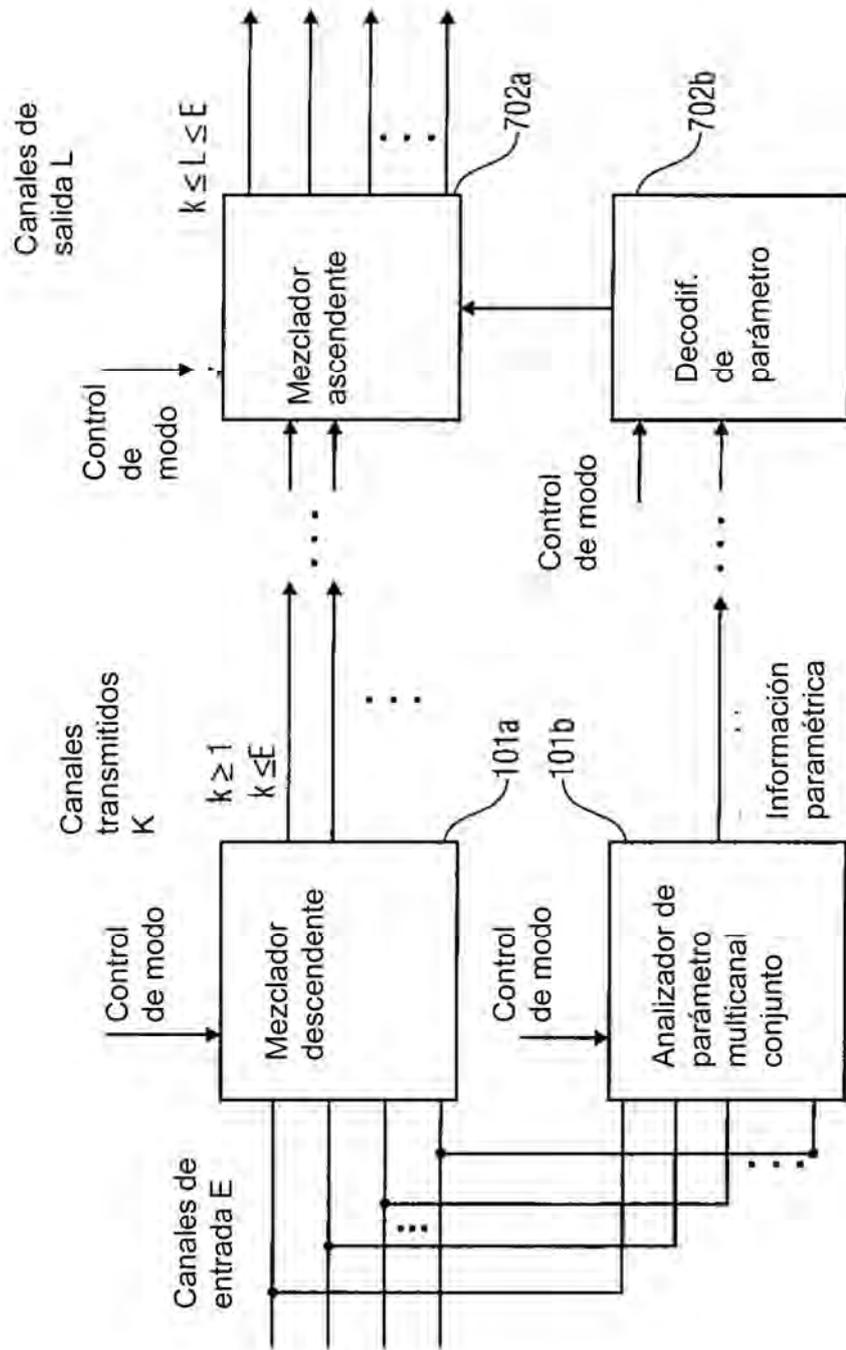


FIG 8

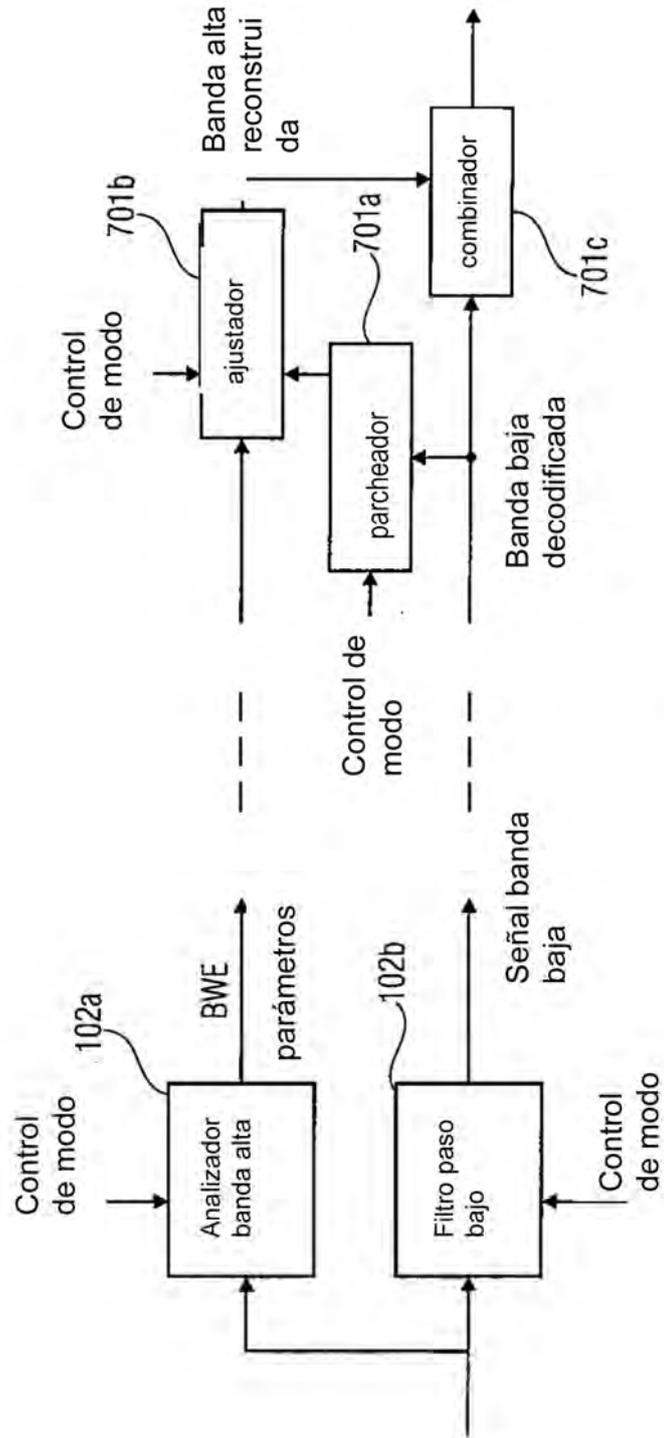
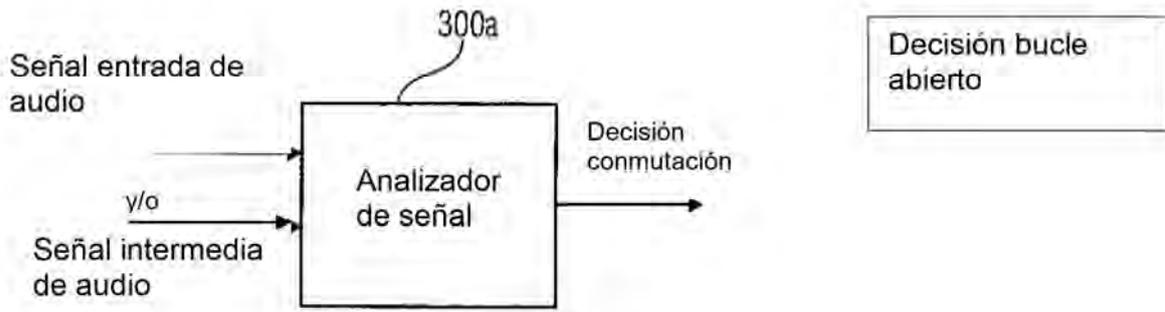


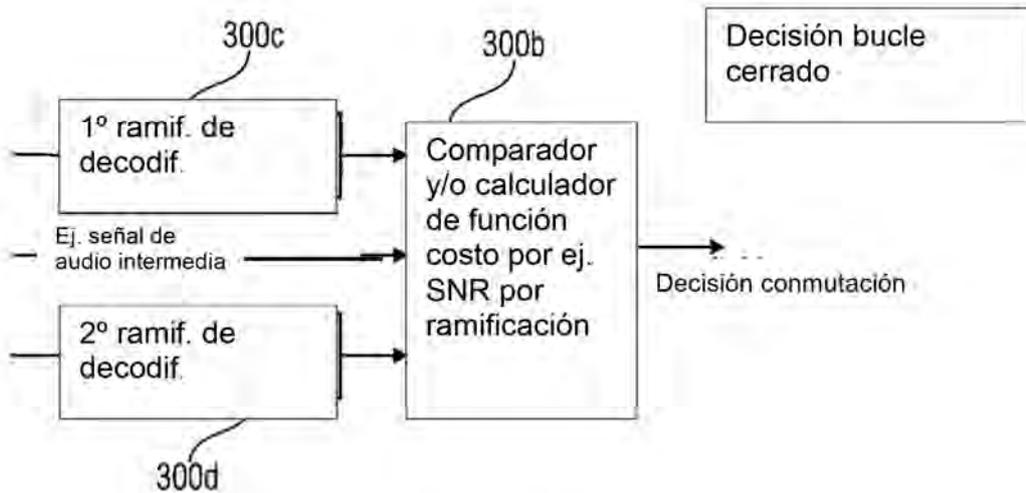
FIG 9



Decisión bucle abierto

Señal intermedia de audio
 -señal banda baja
 -señal de mezcla descendente
 o
 - porción banda baja de señal de mezcla descendente

FIG 10A



Decisión bucle cerrado

FIG 10B