

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 439 549**

51 Int. Cl.:

G10L 19/20 (2013.01)

G10L 21/038 (2013.01)

G10L 21/02 (2013.01)

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.06.2009** **E 11162255 (1)**

97 Fecha y número de publicación de la concesión europea: **04.09.2013** **EP 2352147**

54 Título: **Un aparato y un método para decodificar una señal de audio codificada**

30 Prioridad:

11.07.2008 US 79841 P

08.10.2008 US 103820 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.01.2014

73 Titular/es:

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

72 Inventor/es:

**NEUENDORF, MAX;
KRÄMER, ULRICH;
GRILL, BERNHARD;
MULTRUS, MARKUS;
POPP, HARALD;
RETTTELBACH, NIKOLAUS;
NAGEL, FREDERIK;
LOHWASSER, MARKUS;
GAYER, MARC;
JANDER, MANUEL y
BACIGALUPO, VIRGILIO**

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 439 549 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Un aparato y un método para decodificar una señal de audio codificada

5 **[0001]** La presente invención se refiere a un aparato y un método para codificar.

[0002] En la técnica, se conocen esquemas de codificación de dominio de frecuencia como MP3 o AAC (codificación de audio avanzada). Estos codificadores de dominio de frecuencia se basan en una conversión de dominio de tiempo/dominio de frecuencia, una etapa de cuantización subsecuente, donde el error de cuantización es controlado utilizando información desde un módulo psicoacústico, y una etapa de codificación, donde los coeficientes espectrales cuantizados y la información lateral correspondiente son codificados por entropía utilizando tablas de códigos.

[0003] Por otro lado, existen codificadores muy apropiados para el procesamiento de voz como AMR-WB+ (velocidad múltiple adaptativa - banda ancha) como se describe en 3GPP TS 26.290. Dichos esquemas de codificación de voz realizan un filtro Predictivo Lineal de una señal de dominio del tiempo. Dicho filtro LP deriva de un análisis Predictivo Lineal de la señal de entrada de dominio del tiempo. Los coeficientes del filtro LP resultantes son luego cuantizados/codificados y transmitidos como información lateral. El proceso se conoce como Codificación de Predicción Lineal (LPC). En la salida del filtro, la señal residual de predicción o señal de error de predicción también conocida como señal de excitación es codificada utilizando las etapas de análisis-por-síntesis del codificador ACELP (predicción lineal con excitación por código algebraico) o, en forma alternativa, es codificada utilizando un codificador de transformación, que utiliza una transformada de Fourier con solapamiento. La decisión entre la codificación ACELP y la codificación de excitación Codificada de Transformación también llamada codificación TCX se realiza utilizando un algoritmo de bucle cerrado o bucle abierto. Los esquemas de codificación de audio en dominio de frecuencia como el esquema de codificación de alta eficiencia AAC, que combina un esquema de codificación AAC y una técnica de replicación de banda espectral puede combinarse también con una herramienta de codificación de estero conjunto o multi-canal conocida bajo la denominación "MPEG surround" (Grupo de Expertos en Imágenes en Movimiento). Por otra parte, codificadores de voz tales como AMR-WB+ también tienen un estado de mejora de frecuencia alta y una funcionalidad estéreo.

[0004] Dicha replicación de banda espectral (SBR) comprende una técnica que obtuvo popularidad como complemento de la codificación de audio de percepción popular como MP3 y la codificación de audio avanzada (AAC). SBR comprende un método de ampliación de ancho de banda (BWE) donde la banda baja (banda base o banda núcleo) del espectro es codificada utilizando una codificación existente, en tanto que la banda superior (o banda alta) es parametrizada en forma grosera utilizando pocos parámetros. SBR hace uso de una correlación entre la banda baja y la banda alta con el fin de predecir la señal de banda alta extrayendo rasgos de la banda inferior.

[0005] SBR es, por ejemplo, utilizada en HE-AAC o AAC+SBR. En SBR es posible cambiar en forma dinámica la frecuencia de cruce (frecuencia de inicio de BWE) como también la resolución temporal que implica la cantidad de conjuntos de parámetros (envolvente) por cuadro. AMR-WB+ implementa una ampliación de ancho de banda en dominio de tiempo combinada con un decodificador de núcleo de dominio de tiempo/frecuencia conmutada, otorgando una buena calidad de audio especialmente para señales de voz. Un factor limitante de la calidad de audio de AMR-WB+ es el ancho de banda de audio común a ambos codificadores de núcleo y frecuencia de inicio BWE que represente un cuarto de la frecuencia de muestreo interna del sistema. Mientras que el modelo de voz ACELP es capaz de modelar señales de voz lo suficientemente bien por todo el ancho de banda, el codificador de audio en dominio de frecuencia fracasa en el envío de una calidad decente para algunas señales de audio generales. En consecuencia, los esquemas de codificación de voz muestran una alta calidad para las señales de voz aún a baja tasa de bits, pero muestran una pobre calidad para las señales de música a baja tasa de bits.

[0006] Esquemas de codificación en frecuencia de dominio como HE-AAC son ventajosos porque muestran una alta calidad a baja tasa de bits para señales de música. Sin embargo, es problemático, la calidad de señales de voz a baja tasa de bits.

55 **[0007]** Por lo tanto, las diferentes clases de señal de audio demandan diferentes características de la herramienta de ampliación de ancho de banda.

[0008] Propuestas de extensión de ancho de banda son por ejemplo las descritas en las patentes WO 02/41302 A y WO 2008/031458 A1.

[0009] El objetivo de la presente invención es proveer un concepto mejorado de codificación.

[0010] El objetivo se logra por medio de un decodificador de audio de acuerdo con la reivindicación 1, un método para decodificar de acuerdo con la reivindicación 2 o un programa de computación de acuerdo con la reivindicación 3.

[0011] La presente invención se basa en descubrir que la frecuencia de cruce o la frecuencia de inicio BWE es un parámetro que ejerce influencia en la calidad de audio. Mientras que los codificadores de dominio de tiempo (voz) usualmente codifican todo el rango de frecuencia para una velocidad de muestreo dada, el ancho de banda del audio es un parámetro de sintonía (por ejemplo codificadores para música), que disminuye la cantidad total de líneas espectrales para codificar y al mismo tiempo aumentará la cantidad de bits por línea espectral disponible para codificar, significando que se realiza un intercambio de calidad versus ancho de banda de audio. Por lo tanto, en el nuevo enfoque, se combinan diferentes codificadores de núcleo con anchos de banda de audio variable con un sistema conmutado con un módulo BWE común, donde el módulo BWE debe representar los diferentes anchos de banda de audio.

[0012] Un modo directo sería hallar el menor de todos los anchos de bandas de los codificadores de núcleo y utilizarlo como frecuencia de inicio de BWE, pero esto deterioraría la calidad de audio percibida. Además, la eficiencia de codificación sería reducida, ya que en secciones de tiempo donde un codificador de núcleo es activo el cual tiene un mayor ancho de banda que la frecuencia de inicio BWE, algunas regiones de frecuencia serían representadas dos veces, por el codificador de núcleo sí como BWE que introduce redundancia. Una mejor solución consiste por lo tanto en adaptar la frecuencia de inicio BWE al ancho de banda de audio del codificador de núcleo utilizado.

[0013] Por lo tanto, de acuerdo con las formas de realización de la presente invención un sistema de codificación de audio combina una herramienta de ampliación de ancho de banda con un codificador de núcleo que depende de la señal (por ejemplo codificador de voz/audio conmutado), donde la frecuencia de cruce comprende un parámetro variable. Una salida clasificadora de señal que controla la conmutación entre diferentes modos de codificación de núcleo puede también ser utilizada para conmutar las características del sistema BWE como la resolución temporal y borrosidad, resolución espectral y la frecuencia de cruce.

[0014] En contraposición con las formas de realización, SBR en técnicas previas se aplica a un codificador de audio sin conmutador que solamente dan por resultado las siguientes desventajas. La resolución temporal así como la frecuencia de cruce podrían aplicarse en forma dinámica, pero implementaciones de vanguardia como la fuente 3GPP aplican sólo un cambio de resolución temporal para tasa de bits como, por ejemplo, castañuelas. Asimismo, una resolución temporal total más fina podría escogerse velocidades mayores como un parámetro de sintonía que depende de o un valor umbral de decisión que controla la resolución temporal, que combina de la mejor manera la tasa de bits de la señal. No se lleva a cabo una clasificación explícita que determina el tipo de resolución temporal, como por ejemplo, música tonal estacionaria versus voz. Las formas de realización de la presente invención superan estas desventajas. Las formas de realización permiten especialmente una frecuencia de cruce adaptada combinada con una opción flexible para el codificador de núcleo utilizado para que la señal codificada provea una calidad significativamente mayor en comparación con un codificador/decodificador de técnicas previas

Breve descripción de los dibujos

[0015] Las formas de realización preferidas de la presente invención se describen a continuación con respecto a los dibujos adjuntos, donde:

- Fig. 1 muestra un diagrama de bloque de un aparato para decodificar;
- Fig. 2 muestra un diagrama de bloque de un aparato para decodificar de acuerdo con el primer aspecto de la presente invención;
- Fig. 3 muestra un diagrama de bloque de un esquema de codificación con mayor detalle;
- Fig. 4 muestra un diagrama de bloque de un esquema de decodificación con mayor detalle;

- Fig. 5 muestra un diagrama de bloque de un esquema de codificación de acuerdo con un segundo aspecto;
- Fig. 6 es un diagrama esquemático de un esquema de decodificación de acuerdo con el segundo aspecto;
- 5 Fig. 7 ilustra una etapa LPC del lado del codificador que provee información de predicción a corto plazo y la señal de error de predicción;
- Fig. 8 ilustra otra forma de realización de un dispositivo LPC para generar una señal ponderada;
- 10 Figs. 9a-9b muestran un codificador que comprende un conmutador de audio/voz que da por resultado diferentes resoluciones temporales para una señal de audio;
- Fig. 10 ilustra una representación para una señal de audio codificada.

15

Descripción detallada de la invención

[0016] La Fig. 1 muestra un aparato decodificador 100 para decodificar una señal de audio codificada 102. La señal de audio codificada 102 comprende una primera porción 104a codificada de acuerdo con el primer algoritmo de codificación, una segunda porción 104b codificada de acuerdo con un segundo algoritmo de codificación, un parámetro BWE 106 para la primera porción de tiempo 104a y la segunda porción de tiempo 104b y una información de modo de codificación 108 que indican un primer algoritmo de decodificación o un segundo algoritmo de decodificación para las respectivas porciones de tiempo. El aparato para decodificar 100 comprende un primer decodificador 110a, un segundo decodificador 110b, un módulo BWE 130 y un controlador 140. El primer decodificador 110a está adaptado para decodificar la primera porción 104a de acuerdo con el primer algoritmo de decodificación para la primera porción de tiempo de la señal codificada 102 para obtener una primera señal decodificada 114a. El segundo decodificador 110b está configurado para decodificar la segunda porción 104b de acuerdo con el segundo algoritmo de decodificación para una segunda porción de tiempo de la señal codificada para obtener una segunda señal decodificada 114b. El módulo BWE 130 posee una frecuencia de cruce controlable f_x que ajusta la conducta del módulo BWE 130. El módulo BWE 130 está configurado para desarrollar un algoritmo de ampliación de ancho de banda para generar componentes de la señal de audio en la banda de frecuencia superior basados en la primera señal de decodificación 114a y los parámetros BWE 106 para la primera porción, y para generar componentes de la señal de audio en la banda de frecuencia superior basados en la segunda señal decodificada 114b y el parámetro de ampliación de ancho de banda 106 para la segunda porción. El controlador 140 está configurado para controlar la frecuencia de cruce f_x del módulo BWE 130 de acuerdo con la información del modo de codificación 108.

[0017] El módulo BWE 130 puede comprender también un combinador que combina los componentes de la señal de audio de la banda de frecuencia inferior y superior y emisiones de la señal de audio resultante 105.

[0018] La información del modo de codificación 108 indica, por ejemplo cual porción de tiempo de la señal de audio codificada 102 es codificada por cual algoritmo de codificación. Esta información puede al mismo tiempo identificar el decodificador a utilizar para las diferentes porciones de tiempo. Además, la información del modo de codificación 108 puede controlar un conmutador para conmutar entre diferentes decodificadores para diferentes porciones de tiempo.

[0019] Por ende, la frecuencia de cruce f_x es un parámetro ajustable que se ajusta de acuerdo con el decodificador utilizado el cual puede, por ejemplo, comprender un codificador de voz como primer decodificador 110a y un decodificador de audio como segundo decodificador 110b. Como se dijo anteriormente, la frecuencia de cruce f_x para un decodificador de voz (como por ejemplo basada en LPC) puede ser mayor que la frecuencia de cruce usada para un decodificador de audio (por ejemplo para música). En consecuencia, en otros ejemplos el controlador 220 está configurado para aumentar la frecuencia de cruce f_x o para disminuir la frecuencia de cruce f_x dentro de una de las porciones de tiempo (por ejemplo la segunda porción de tiempo) para que la frecuencia de cruce pueda ser cambiada sin cambiar el algoritmo de decodificación. Esto significa que un cambio en la frecuencia de cruce puede no estar relacionado con un cambio en el decodificador utilizado: la frecuencia de cruce puede ser cambiada sin cambiar el decodificador utilizado o vice versa el decodificador puede ser cambiado sin cambiar la frecuencia de cruce.

[0020] El módulo BWE 130 puede también comprender un conmutador controlado por el controlador 140 y/o por

el parámetro BWE 106 para que la primera señal decodificada 114a es procesada por el módulo BWE 130 durante la primera porción de tiempo y la segunda señal de decodificación 114b es procesada por el módulo BWE 130 durante la segunda porción de tiempo. Este conmutador puede ser activado por un cambio en la frecuencia de cruce f_x o por un bit explícito dentro de la señal de audio codificada 102 indicando el algoritmo de codificación utilizado durante la respectiva porción de tiempo.

[0021] En otros ejemplos el conmutador está configurado para conmutar entre la primera y segunda porción de tiempo desde el primer decodificador al segundo decodificador para que el algoritmo de ampliación de ancho de banda se aplique a la primera señal decodificada o a la segunda señal decodificada. De manera alternativa, el algoritmo de ampliación de ancho de banda se aplica a la primera y/o segunda señal decodificada y el conmutador de ubica después de esto para que caiga una de las señales extendidas de ancho de banda.

[0022] La Fig. 2 muestra un diagrama de bloque para un aparato 200 codificador de una señal de audio 105. El aparato codificador 200 comprende un primer codificador 210a, un segundo codificador 210b, una etapa de decisión 220 y un módulo de ampliación de ancho de banda (módulo BWE) 230. El primer codificador 210a es operativo para codificar de acuerdo con un primer algoritmo de codificación con un primer ancho de banda de frecuencia. El segundo codificador 210b es operativo para codificar de acuerdo con un segundo algoritmo de codificación con un segundo ancho de banda de frecuencia menor al primer ancho de banda de frecuencia. El primer codificador es un codificador de voz como el codificador basado en LPC, mientras que el segundo codificador 210b comprende un codificador de audio (música). La etapa de decisión 220 está configurada para indicar el primer algoritmo de codificación para la primera porción 204a de la señal de audio 105 y para indicar el segundo algoritmo de codificación para una segunda porción 204b de la señal de audio 105, donde la segunda porción de tiempo es diferente de la primera porción de tiempo. La primera porción 204a corresponde a la primera porción de tiempo y la segunda porción 204b corresponde a la segunda porción de tiempo la cual es diferente de la primera porción de tiempo.

[0023] El módulo BWE 230 está configurado para calcular los parámetros BWE 106 para la señal de audio 105 y está configurado para ser controlado por la etapa de decisión 220 para calcular el parámetro BWE 106 para una primera banda que no incluye el primer ancho de banda de frecuencia en la primera porción de tiempo 204a de la señal de audio 105. El módulo BWE 230 está además configurado para calcular el parámetro BWE 106 para una segunda banda que no incluye el segundo ancho de banda en la segunda porción de tiempo 204b de la señal de audio 105. La primera (segunda) banda comprende por lo tanto componentes de frecuencia de la señal de audio 105 que están fuera del primer (segundo) ancho de banda de frecuencia y están limitados hacia el extremo inferior del espectro por la frecuencia de cruce f_x . El primer o el segundo ancho de banda por lo tanto está definido por una frecuencia de cruce variable controlada por la etapa de decisión 220.

[0024] Además, el módulo BWE 230 puede comprender un conmutador controlado por la etapa de decisión 220. La etapa de decisión 220 puede determinar un algoritmo de codificación preferido para una porción de tiempo dada y controla el conmutador para que durante la porción de tiempo dada se utilice el codificador preferido. La información del modo de codificación modificada 108' comprende la señal de conmutación correspondiente. Además, el módulo BWE 230 puede también comprender un filtro para obtener componentes de la señal de audio 105 en la banda de frecuencia inferior/superior que están separados por la frecuencia de cruce f_x que puede comprender un valor de aproximadamente 4 kHz o 5 kHz. Finalmente el módulo BWE 130 puede también comprender una herramienta de análisis para determinar el parámetro BWE 106. La información del modo de codificación modificado 108' puede ser equivalente (o igual) a la información del modo de codificación 108. La información del modo de codificación 108 indica, por ejemplo, el algoritmo de codificación utilizado para las respectivas porciones de tiempo en la secuencia de bits de la señal de audio codificada 105.

[0025] De acuerdo con otras formas de realización, la etapa de decisión 220 comprende una herramienta de clasificación de señal que analiza la señal de entrada original 105 y genera la información de control 108 que provoca la selección de los diferentes modos de codificación. El análisis de la señal de entrada 105 depende del objetivo de elegir el óptimo modo de codificación de núcleo para un cuadro de señal de entrada dado. La salida del clasificador de señal puede (opcionalmente) ser utilizada también para influenciar en el comportamiento de otras herramientas, por ejemplo, sonido envolvente MPEG, SBR ampliado, banco de filtro con intercalación en el tiempo y otras. La entrada a la herramienta clasificadora de señal comprende, por ejemplo, la señal de entrada no modificada original 105, pero también parámetros que dependen de la implementación adicional en forma opcional. La salida de la herramienta clasificadora de la señal comprende la señal de control 108 para controlar la selección del codificador de núcleo (por ejemplo dominio de frecuencia no filtrada LP o tiempo filtrado LP o codificación de dominio de frecuencia o demás algoritmos de codificación).

[0026] De acuerdo con las formas de realización, la frecuencia de cruce f_x es una señal ajustada dependiente combinada con la decisión de conmutación para utilizar un algoritmo de codificación diferente. Por lo tanto, una señal de conmutación simple puede ser simplemente un cambio (un salto) en la frecuencia de cruce f_x . Además, la información del modo de codificación 108 puede también comprender el cambio de frecuencia de cruce f_x que indica al mismo tiempo un esquema de codificación preferido (por ejemplo voz/audio/música).

[0027] De acuerdo con otras formas de realización la etapa de decisión 220 es operativa para analizar la señal de audio 105 o una primera salida del primer codificador 210a o una segunda salida del segundo codificador 210b o una señal obtenida por la decodificación de una señal de salida del codificador 210a o el segundo codificador 210b con respecto a la función de destino. La etapa de decisión 220 puede opcionalmente ser operativa para realizar una discriminación de voz/música de modo tal que una decisión de voz es favorecida con respecto a la decisión de música para que se tome la decisión de voz, por ejemplo, aún cuando una porción menor al 50% de un cuadro para el primer conmutador es de voz y una porción mayor al 50% del cuadro para el primer conmutador es de música. Por lo tanto, la etapa de decisión 220 puede comprender una herramienta de análisis que analiza la señal de audio para decidir si la señal de audio es principalmente una señal de voz o principalmente una señal de música para que teniendo en cuenta el resultado de la etapa de decisión pueda decidir cual es el mejor codificador a utilizar para la porción de tiempo analizada de la señal de audio.

[0028] Las Figs. 1 y 2 no muestran muchos de estos detalles para el codificador/decodificador. Algunos posibles ejemplos detallados para el codificador/decodificador se muestran en las siguientes figuras. Además, con respecto al primer y segundo decodificador 110a, b de la Fig. 1 otros decodificadores pueden estar presentes los que pueden o no utilizar por ejemplo otros algoritmos de codificación. Del mismo modo, también el codificador 200 de la Fig. 2 puede comprender codificadores adicionales que pueden utilizar algoritmos de codificación adicionales. En el siguiente ejemplo con dos codificadores/decodificadores serán explicados con mayor detalle.

[0029] La Fig. 3 ilustra con mayor detalle un codificador con dos conmutadores en cascada. Una señal mono, una señal estéreo o señal multi-canal es ingresada a la etapa de decisión 220 y en un conmutador 232 que forma parte del módulo BWE 230 de la Fig. 2. El conmutador 232 es controlado por la etapa de decisión 220. De manera alternativa, la etapa de decisión 220 puede también recibir una información lateral incluida en la señal mono, la señal estéreo o la señal multi-canal o es al menos asociada con dicha señal, donde la información es existente, la cual fue, por ejemplo, generada al producir originalmente la señal mono, la señal estéreo o la señal multi-canal.

[0030] La etapa de decisión 220 activa el conmutador con el fin de suministrar una señal en una porción de codificación de frecuencia 210b ilustrada en una ramificación superior de la Fig. 3 o una porción de codificación de dominio LPC- 210a ilustrada en una ramificación inferior en la Fig. 3. Un elemento clave de la ramificación de codificación de dominio de frecuencia consiste en 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 algoritmo FFT, un análisis Wavelet o un banco de filtro como un banco de filtro tomado por muestreo en forma crítica con una cierta cantidad de canales de banco de filtro, donde las señales de sub-bandas en este banco de filtro pueden ser señales de valor real o 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 del esquema de codificación AAC.

[0031] Por lo general, el procesamiento en la ramificación 210b es un procesamiento en un modelo basado en la percepción o modelo de información de sumidero. De esa manera, esta ramificación modela el sonido de sumidero del sistema auditivo humano. En contraposición, el procesamiento en la ramificación 210a debe generar una señal en el dominio de excitación, residual o LPC. Por lo general, el procesamiento en la ramificación 210a es un procesamiento de un modelo de voz o modelo generador de información. Para señales de voz, este modelo es un modelo de voz humana/sistema generador del sonido. Sin embargo, si un sonido de una fuente diferente que requiera un modelo generador de sonido diferente debe ser codificado, entonces el procesamiento en la ramificación 210a puede ser diferente. Además de las ramificaciones de codificación mostradas, otras formas de realización comprenden ramificaciones adicionales o codificadores de núcleo. Por ejemplo, diferentes codificadores pueden estar opcionalmente presentes para diferentes fuentes, para que el sonido de cada fuente pueda estar codificado al emplear un codificador preferido.

[0032] En la ramificación de codificación inferior 210a, un elemento clave consiste en un dispositivo LPC 510, que emite una información LPC utilizada 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 señal ponderada.

5 **[0033]** El dispositivo LPC generalmente emite una señal de dominio LPC, la cual puede ser cualquier señal en el dominio LPC u otra señal, que haya sido generada aplicando coeficientes de filtro LPC a una señal de audio. Asimismo, un dispositivo LPC puede también determinar estos coeficientes y puede además cuantizar/codificar estos coeficientes.

10 **[0034]** La decisión en la etapa de decisión 220 puede ser una señal adaptable para que la etapa de decisión desarrolle una discriminación de música/ voz y controle el conmutador 232 de tal modo que las señales musicales sean ingresadas en la ramificación superior 210b, y las señales de voz sean ingresadas en la ramificación inferior 210a. En una forma de realización, la etapa de decisión 220 suministra su información de decisión en una secuencia 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. Esta información de decisión puede, por ejemplo, comprender la información del modo de codificación 108 que también puede comprender información sobre la frecuencia de cruce 15 fx o un cambio de la frecuencia de cruce fx.

20 **[0035]** Dicho decodificador se ilustra en la Fig. Fig. 4. La emisión de la señal por el codificador de audio espectral 421 es, luego de la transmisión, ingresada en un decodificador de audio espectral 431. La emisión de señal del decodificador de audio espectral 431 es ingresada en un conversor de dominio de tiempo 440. (el conversor de tiempo-dominio puede en general ser un conversor del primer al segundo dominio). De forma análoga, la emisión de la ramificación de codificación de dominio de LPC 210a de la Fig. 3 es recibida en el lado del decodificador y procesada por los elementos 531, 533, 534, y 532 para obtener una señal de excitación LPC. La excitación de la señal LPC es ingresada en una etapa de síntesis LPC 540, que recibe, como entrada adicional, la información LPC generada por la etapa de análisis LPC correspondiente 510. La salida del conversor de tiempo-dominio 440 y/o la salida de la etapa de síntesis LPC 540 son ingresadas en un conmutador 132 que puede ser parte del módulo BWE 130 in Fig. 1. El conmutador 132 es controlado a través de una señal de control del conmutador (como la información del modo de codificación 108 y/o parámetro BWE 106) que fue, por ejemplo, generada por la etapa de decisión 220, o que fue proporcionada en forma externa por medio de un creador de la señal mono, señal estéreo o señal multi-canal original. 30

35 **[0036]** En la Fig. 3, la señal de entrada en el conmutador 232 y la etapa de decisión 220 puede ser una señal mono, señal estéreo o señal multi-canal o generalmente una señal de audio. Dependiendo de la decisión que puede derivar de la señal de entrada del conmutador 232 o de cualquier fuente externa como un productor de la señal de audio original subyaciendo la emisión de señal en la etapa 232, el conmutador conmuta entre la ramificación de codificación de frecuencia 210b y la ramificación de codificación LPC 210a. La ramificación de codificación de frecuencia 210b comprende una etapa de conversión espectral 410 y una etapa de cuantización/codificación subsecuentemente conectada 421. La etapa de cuantización/codificación puede incluir cualquiera de las funcionalidades conocidas de los codificadores de dominio de frecuencia modernos como el codificador AAC. 40 Además, la operación de cuantización en la etapa de cuantización/codificación 421 puede ser controlada a través de un módulo psicoacústico que genera información psicoacústica como el valor umbral del enmascaramiento psicoacústico sobre la frecuencia, donde esta información es ingresada en la etapa 421.

45 **[0037]** En la ramificación de codificación LPC 210a, la señal de salida del conmutador es procesada a través de la etapa de análisis LPC 510 que genera información del lado LPC y una señal de dominio LPC. El codificador de excitación inventivamente comprende un conmutador adicional para conmutar el procesamiento adicional de la señal de dominio LPC entre la operación de cuantización/codificación 522 en el dominio LPC o una etapa de cuantización/codificación 524, que procesa valores en el dominio espectral LPC. Para este fin, un conversor espectral 523 se provee en la entrada de la etapa de cuantización/codificación 524. El conmutador 521 es controlado de modo de bucle abierto o modo de bucle cerrado dependiendo de las configuraciones específicas como, por ejemplo, se describen en la especificación técnica AMR-WB+. 50

55 **[0038]** Para el modo de control de bucle cerrado, el codificador además incluye un cuantizador/codificador inverso 531 para la señal de dominio LPC, un cuantizador/codificador inverso 533 para a señal de dominio espectral LPC y un conversor espectral inverso 534 para la salida el ítem 533. Ambas señales codificadas y nuevamente decodificadas en las ramificaciones de procesamiento de la segunda ramificación de codificación son ingresadas 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 de destino o se calcula una función de destino la cual puede basarse en una comparación de la distorsión en ambas señales para que la señal con menor distorsión

se use para decidir, que posición debería tomar el conmutador 521. De manera alternativa, en caso de que ambas ramificaciones provean velocidades binarias no-constantes, la ramificación que provea la tasa de bits menor podría ser seleccionada aún cuando la distorsión o la distorsión perceptual de esta ramificación es menor a la distorsión o la distorsión perceptual de la otra ramificación (un ejemplo para la distorsión puede ser una relación señal-ruido).

5 De manera alternativa, la función de destino podría utilizar, como 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 de destino se basaría en gran medida en la tasa de bits de la emisión de las dos señales por medio de 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, por ejemplo, descartar cada señal que se encuentre por sobre la tasa de bits permitida cuando ambas señales se encuentran por debajo de la tasa de bits permitida, el control del conmutador seleccionaría la señal con mejor calidad subjetiva estimada o una mejor relación señal-ruido, es decir, con la menor distorsión de cuantización/codificación.

15 **[0039]** El esquema de codificación de acuerdo con un ejemplo, como se ha establecido previamente, se ilustra en la Fig. 4. Para cada tipo de las tres señales de salida, existe una etapa específica de decodificación/re-cuantización 431, 531 o 533. Mientras que la etapa 431 emite un espectro de tiempo que se convierte en dominio de tiempo utilizando el conversor frecuencia/tiempo 440, la etapa 531 emite una señal de dominio LPC, y el ítem 533 emite un espectro LPC. Con el fin de asegurarse que las señales de entrada en el conmutador 532 se encuentran 20 ambas en el dominio LPC, se proporciona el espectro LPC/conversor LPC 534. Los datos de salida del conmutador 532 se transforman nuevamente en el dominio de tiempo utilizando una etapa de síntesis LPC 540, controlada a través de información LPC generada y transmitida del lado del codificador. Luego, subsecuente al bloque 540, ambas ramificaciones poseen información de dominio de tiempo conmutada de acuerdo con una señal de control de conmutación con el fin de finalmente obtener una señal de audio como por ejemplo una señal mono, señal estéreo, o 25 señal multi-canal, que depende de la entrada de señal en el esquema de codificación de la Fig. 3.

[0040] Las Figs. 5 y 6 muestran otra forma de realización para el codificador, y otro ejemplo para el decodificador, donde las etapas BWE como parte de los módulos BWE 130, 230 representan una unidad de procesamiento común.

30 **[0041]** La Fig. 5 ilustra un esquema de codificación, donde el esquema de preprocesamiento común conectado a la entrada del conmutador 232 puede comprender un bloque estéreo envolvente/conjunto 101 el cual genera, como salida, parámetros de estéreo conjunto y una señal de salida mono, generada al mezclar en forma descendente la señal de entrada que es una con dos o más canales. Generalmente, la señal en la salida del bloque 101 puede también ser una señal con 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 de la entrada del bloque 101.

[0042] El esquema de preprocesamiento común puede comprender además del bloque 101 una etapa de ampliación de ancho de banda 230. En la forma de realización de la Fig. 5 la salida del bloque 101 es ingresada al bloque de de ancho de banda 230 el cual emite una señal de banda limitada como la señal de banda baja o señal de paso bajo en su salida. Preferentemente, esta señal es también sometida a submuestreo (por ejemplo por un factor de dos). Los parámetros de ampliación de ancho de banda como los parámetros de envolvente espectral, parámetros de filtrado inverso, parámetros de piso de ruido etc. como se conocen del perfil HE-AAC de MPEG-4 son generados y emitidos a un multiplexor de secuencia de bits 800.

[0043] Preferentemente, la etapa de decisión 220 recibe la entrada de la señal en el bloque 101 o entrada en el bloque 230 con el fin de decidir entre, por ejemplo, un modo de música o modo de voz. En el modo de música, se selecciona la ramificación de codificación superior 210b (segundo codificador en Fig. 2) 210a mientras que en el modo de voz, se selecciona la ramificación de codificación inferior 101 y/o el bloque de ampliación de ancho de banda 230 para adaptar la funcionalidad de estos bloques a la señal específica. De esta manera, cuando la etapa de decisión 220 determina que cierta porción de tiempo de la señal de entrada corresponde al primer modo como el modo música, las características específicas el bloque 101 y/o bloque 230 pueden ser controladas por la etapa de decisión 220. Alternativamente, cuando la etapa de decisión 220 determina que la señal se encuentra en un modo de voz o, generalmente, en un modo de segundo dominio LPC, se pueden controlar las características específicas de los bloques, 101 y 230 de acuerdo con la salida de la etapa de decisión. La etapa de decisión 220 produce además la información de control 108 y/o la frecuencia de cruce f_x que también puede transmitirse al bloque BWE 230 y, además, al multiplexor de secuencia de bits 800 para que sea transmitida al lado del decodificador.

[0044] Preferentemente, la conversión espectral de la ramificación de codificación 210b se realiza utilizando una operación MDCT la cual, con mayor preferencia, consiste en la operación MDCT distorsionada en el tiempo, donde la fuerza o, generalmente, la fuerza de distorsión (warp) puede ser controlada entre cero y una alta fuerza de distorsión (warp). En una fuerza de distorsión (warp) cero, la operación MDCT en el bloque 411 es una operación MDCT sencilla conocida en la técnica. La fuerza de intercalación en el tiempo junto con la información lateral de intercalación en el tiempo puede ser transmitida/ingresada en un multiplexor de secuencia de bits 800 como información lateral.

[0045] 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 información como un índice del libro de información y ganancia del código El modo TCX como se conoce de 3GPP TS 26.290 incurre en un procesamiento de una señal perceptivamente ponderada en el dominio de transformación. Una señal ponderada transformada de Fourier es cuantizada utilizando una cuantización entramada con velocidad múltiple dividida (algebraica VQ) con cuantización de factor ruido. Una transformación se calcula en 1024, 512, o ventanas de muestra 256. La señal de excitación se recupera por filtrado inverso de la señal ponderada cuantizada a través de un filtro de ponderación. El modo TCX puede también ser utilizado en forma modificada en la cual MDCT es utilizado con solapado ampliado, cuantización escalar, y un codificador aritmético para codificar líneas espectrales.

[0046] En la ramificación de codificación de "música" 210b, un conversor espectral preferentemente comprende una operación MDCT específicamente adaptada MDCT con ciertas funciones de ventana seguidas por una etapa de codificación de cuantización/entropía la cual puede consistir en una etapa de cuantización con un solo vector, pero preferentemente es un codificador combinado escalar de cuantización/entropía similar al codificador de cuantización/entropía en la ramificación de codificación en dominio de frecuencia, es decir, en el ítem 421 de la Fig. 5.

[0047] En la ramificación de codificación de "voz" 210a, existe el bloque LPC 510 seguido de un conmutador 521, nuevamente seguida por un bloque ACELP 526 o 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 calculada por un procedimiento descrito en la Fig. 7. El bloque TCX 527 recibe una señal ponderada generada por la Fig. 8.

[0048] En el lado del decodificador ilustrado en la Fig. 6, después que la transformación espectral inversa en el bloque 537, se aplica la inversa del filtro de ponderación, es decir $(1 - \mu z^{-1}) / (1 - A(z/\gamma))$. Luego, la señal es filtrada 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 TCX-1 537 incluyen transformación inversa y luego la filtrado a través de $\frac{(1 - \mu z^{-1})}{(1 - A(z/\gamma))} (1 - A(z))$ para convertirse del dominio ponderado al dominio de excitación.

[0049] Aunque el ítem 510 en las Figs. 3, 5 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 como el modo de señal de excitación o el modo de señal ponderada puede depender del estado real del conmutador. En forma alternativa, el bloque 510 puede tener dos dispositivos de procesamiento paralelos, donde un dispositivo es implementado de manera similar a la Fig. 7 y el otro dispositivo es implementado como la Fig. 8. Por lo tanto, el dominio LPC en la salida de 510 puede representar la señal de excitación LPC o la señal ponderada LPC, o u otra señal de dominio LPC.

[0050] En la segunda ramificación de codificación (ACELP/TCX) de la Fig. 5, la señal es preferentemente pre-enfatizada a través de un filtro $1 - \mu z^{-1}$ antes de la codificación. En el decodificador ACELP/TCX en la Fig. 6 la señal sintetizada es desenfatisada con el filtro $1 / (1 - \mu z^{-1})$. En una forma de realización preferida, el parámetro μ posee el valor 0,68. El pre-énfasis puede ser parte del bloque LPC 510 donde la señal es pre-enfatizada antes del análisis y cuantización LPC. De manera similar, la desenfatisación puede ser parte del bloque de síntesis bloque LPC-1 540.

[0051] La Fig. 6 ilustra un esquema de decodificación correspondiente al esquema de codificación de la Fig. 5. La secuencia de bits generada por un multiplexor de secuencia de bits 800 (o interfaz de salida) de la Fig. 5 es ingresada en un multiplexor de secuencia de bits 900 (o interfaz de entrada). Dependiendo de la información derivada por ejemplo de una secuencia de bits a través de un bloque de detección de modo 601 (por ejemplo parte del controlador 140 en Fig. 1), es controlado para emitir señales desde la ramificación superior o señales desde la

ramificación inferior al bloque de ampliación de ancho de banda 701. El bloque de ampliación de ancho de banda 701 recibe, desde el desmultiplexor de secuencia de bits 900, información lateral y, basado en esta información lateral y la emisión de la decisión de modo 601, reconstruye la banda alta basado en la salida de banda baja por el conmutador 600. La señal de control 108 controla la frecuencia de cruce f_x utilizada.

5

[0052] La señal de banda completa generada por el bloque 701 es ingresada 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 que fueron ingresados en este bloque. Dependiendo de la aplicación, el ingreso en el bloque 702 puede aún incluir dos canales como en el modo estéreo y puede aún incluir más canales siempre que la emisión por este bloque tenga más canales que el ingreso en este bloque.

10

[0053] Se ha observado que el conmutador 232 en la Fig. 5 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 forma de realización alternativa, sin embargo, el conmutador 232 conmutador puede también estar dispuesto en forma subsecuente a por ejemplo el codificador de audio 421 y el codificador de excitación 522, 523, 524, lo que significa que ambas ramificaciones 210a, 210b procesan la misma señal en paralelo. Con el fin de no duplicar la tasa de bits, sin embargo, sólo la salida de señal por una de esas ramificaciones de codificación 210a o 210b es seleccionada 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 costo, donde la función de costo puede ser la secuencia de bits generada o la distorsión perceptiva generada una función de costo combinada de velocidad/distorsión. En consecuencia, ya sea en este modo o en el modo ilustrado en las Figuras, la etapa de decisión puede también operar en el modo de bucle cerrado con el fin de asegurarse que, finalmente, sólo la salida de ramificación de codificación es escrita en la secuencia de bits al cual tiene para una distorsión perceptivo dada la tasa de bits más baja o, para una tasa de bits dada, la distorsión perceptivo más baja. En el modo de bucle cerrado, la entrada de retroalimentación puede ser derivada de las emisiones de los tres bloques cuantizadores/escaladores 421, 522 y 424 en la Fig. 3.

15

20

25

[0054] Además en el ejemplo de la Fig. 6, el conmutador 132 puede en ejemplos alternativos estar dispuesto después del módulo BWE 701 para que la ampliación de ancho de banda se realice en paralelo para ambas ramificaciones y el conmutador selecciona una de las dos señales extendidas de ancho de banda.

30

[0055] En la implementación con dos conmutadores, es decir, el primer conmutador 232 y el segundo conmutador 521, es preferible 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 ser conmutados a través de una operación del conmutador son mayores que los bloques conmutados por el segundo conmutador 521 operando en el dominio LPC. Por ejemplo, el conmutador del dominio de frecuencia/dominio LPC 232 puede conmutar bloques de una longitud de 1024 muestras, y el segundo conmutador 521 puede conmutar bloques con 256 muestras cada uno.

35

[0056] La Fig. 7 ilustra una implementación más detallada del bloque de análisis de LPC 510. La señal de audio es ingresada en un bloque de determinación de filtro que determina la información del filtro $A(z)$. Esta información es emitida como información de predicción a corto plazo necesaria para un decodificador. La información de predicción a corto plazo es requerida por el filtro de predicción real 85. En un subtractor 86, una muestra corriente de la señal de audio es ingresada y un valor predecible para la muestra corriente es sustraído para que para esta muestra, la señal de error de predicción es generada en la línea 84.

40

45

[0057] Mientras que la Fig. 7 ilustra un modo preferido para calcular la señal de excitación, la Fig. 7f ilustra un modo preferido de calcular la señal ponderada. En contraposición a la Fig. 7e, el filtro 85 es diferente, cuando α es diferente de 1. Un valor menor a 1 se prefiere para α . Asimismo, el bloque 87 está presente, y α es preferible 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.

50

[0058] En forma subsecuente, 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 "Codificación de la voz: Una Revisión Tutorial", Andreas Spanias, Procedimientos de IEEE, Vol. 82, No. 10, Octubre 1994, páginas 1541-1582.

55

[0059] Para casos específicos, cuando un cuadro es una mezcla del discurso vocalizado y no vocalizado o cuando ocurre voz 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.

5

[0060] En la codificación tipo AMR-WB+, se lleva a cabo una selección entre los diferentes modos TCX y ACELP como se conoce de la describió AMR-WB+. Los modos TCX so diferentes ya 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.

10

[0061] Como se analizó en relación a las Fig. 5 y 6, la etapa de pre-procesamiento común 100 preferentemente incluye un dispositivo multi-canal conjunto (envolvente/estéreo conjunto) 101 y, además, una etapa de ampliación de ancho de banda 230. De manera correspondiente, el decodificador incluye una etapa de ampliación de ancho de banda 701y una etapa de multicanal conjunto subsecuentemente conectada 702. Preferentemente, la etapa de multicanal conjunto 101 está, con respecto al codificador, conectada antes de la etapa de ampliación de ancho de banda 230, y, del lado del decodificador, la etapa de ampliación de ancho de banda 701 está conectada antes de la etapa de multicanal conjunto 702 con respecto a la dirección de procesamiento de señal. De manera alternativa, sin embargo, la etapa de pre-procesamiento común puede incluir una etapa de multicanal conjunto sin la ampliación de ancho de banda subsecuentemente conectada o etapa de ampliación de ancho de banda sin una etapa de multicanal conjunto conectada.

15

20

[0062] Las Figs. 9a a 9b muestran una vista simplificada en el codificador de la Fig. 5, donde el codificador comprende la unidad de decisión de conmutación 220 y la unidad de codificación estéreo 101. Además, el codificador también comprende herramientas de ampliación de ancho de banda 230 como, por ejemplo, un calculador de datos envolventes y módulos relacionados con SBR. La unidad de decisión de conmutación 220 provee una señal de decisión de conmutación 108' que conmuta entre el codificador de audio 210b y el codificador de voz 210a. El codificador de voz 210a puede además dividirse en un codificador vocalizado y uno no vocalizado. Cada codificador puede codificar la señal de audio en la banda de frecuencia del núcleo utilizando diferentes números de valores de muestra (por ejemplo, 1024 para una resolución mayor o 256 para una resolución menor). La señal de decisión de conmutación 108' es también suministrada a la herramienta de la ampliación de ancho de banda (BWE) 230. La herramienta BWE 230 utilizará la decisión de comulación 108' con el fin de, por ejemplo, ajustar el número de envolventes espectrales 104 y para encender/apagar un detector de tasa de bits opcional y ajustar la frecuencia de cruce f_x . La señal de audio 105 es ingresada a la unidad de decisión de comulación 220 y es ingresada en la codificación estéreo 101 para que la codificación estéreo 101 pueda producir los valores de muestra ingresados en la unidad de ampliación de ancho de banda 230. Dependiendo de la decisión 108' generada por unidad de decisión de unidad de conmutación 220, la herramienta de ampliación de ancho de banda 230 generará los datos de replicación de banda espectral los cuales, a su vez, son enviados al codificador de audio 210b o al codificador de voz 210a.

25

30

35

40

[0063] La señal de decisión de conmutación 108' depende de una señal y puede obtenerse de la unidad de decisión de conmutación 220 analizando la señal de audio, por ejemplo, utilizando un detector de tasa de bits u otros detectores que pueden o no comprender un valor umbral variable. De manera alternativa, la señal de decisión de conmutación 108' puede ajustare en forma manual (por ejemplo por un usuario) u obtenerse de una secuencia de datos (incluida en la señal de audio).

45

[0064] La salida del codificador de audio 210b y el codificador de voz 210a puede nuevamente ser ingresada en el formateador se secuencia de bits 800 (ver Fig. 5).

50

[0065] La Fig. 9b muestra un ejemplo para la señal de decisión de conmutación 108' que detecta una señal de audio para un período de tiempo antes de un primer tiempo t_a después de un segundo tiempo t_b . Entre el primer tiempo t_a y el segundo tiempo t_b , la unidad de decisión de conmutación 220 detecta una señal de voz que da por resultado diferentes valores discretos para la señal de decisión de conmutación 108'.

55

[0066] La decisión que utiliza una frecuencia de cruce f_x mayor es controlada por la unidad de decisión de conmutación 220. Esto significa que el método descrito es además utilizable dentro de un sistema donde el módulo SBR se combina con un codificador con un solo núcleo y frecuencia de cruce f_x variable.

[0067] Aunque algunas de las Figs. 1 a 9 se ilustran como diagramas de bloque de un aparato, estas figuras simultáneamente son una ilustración de un método, donde las funcionalidades del bloque corresponden a los pasos

del método.

[0068] La Fig. 10 ilustra una representación para una señal de audio codificada 102 que comprende una primera porción 104a, la segunda porción 104b, una tercera porción 104c y una cuarta porción 104d. En esta representación la señal de audio codificada 102 es una secuencia de bits transmitida por un canal de transmisión que comprende además la información del modo de codificación 108. Cada porción 104 de la señal de audio codificada 102 puede representar una diferente porción de tiempo, aunque diferentes porciones 104 pueden estar en el dominio de frecuencia como en el dominio de tiempo para que la señal de audio codificada 102 no pueda representar una línea de tiempo.

[0069] En este ejemplo la señal de audio codificada 102 comprende además una primera información del modo de codificación 108a identificando el algoritmo de codificación utilizado para la primera porción 104a; una segunda información del modo de codificación 108b identificando el algoritmo de codificación utilizado para la segunda porción 104b; una tercera información del modo de codificación 108d identificando el algoritmo de codificación utilizado para la cuarta porción 104d. La primera información del modo de codificación 108a puede además identificar la primera frecuencia de cruce fx1 utilizada dentro de la primera porción 104a, y la segunda información del modo de codificación 108b puede además identificar la segunda frecuencia de cruce fx2 utilizada dentro de la segunda porción 104b. Por ejemplo, dentro de la primera porción 104a el modo de codificación de "voz" puede ser utilizado dentro de la segunda porción 104b el modo de codificación de "música" puede ser utilizado para que la primera frecuencia de cruce fx1 pueda ser mayor a la segunda frecuencia de cruce fx2.

[0070] En este ejemplo la señal de audio codificada 102 no comprende información del modo de codificación para la tercera porción 104c que indica que no hay cambio en el codificador y/o frecuencia de cruce fx utilizados entre la primera y la tercera porción 104a, c. Por lo tanto, la información del modo de codificación 108 puede aparecer como encabezado sólo para las porciones 104 que utilizan un codificador de núcleo y/o frecuencia de cruce diferentes comparados con la porción precedente. En otros ejemplos en vez de señalar los valores de las frecuencias de cruce para las diferentes porciones 104, la información del modo de codificación 108 puede comprender un sólo bit que indica el codificador de núcleo (primer o segundo codificador 210a,b) utilizado para la respectiva porción 104.

[0071] Por lo tanto, la señalización del comportamiento del conmutador entre las diferentes herramientas SBR puede realizarse presentando, por ejemplo, como bit específico dentro de la secuencia de bits, para que este bit específico pueda encender o apagar un comportamiento específico en el decodificador. En forma alternativa, en sistemas con dos codificadores de núcleo de acuerdo con ejemplos la señalización del conmutador puede también iniciarse analizando el codificador del núcleo. En este caso la presentación de la adaptación de las herramientas SBR se realiza en forma implícita, lo que significa que se determina por la actividad del codificador de núcleo correspondiente.

[0072] Más detalles de la descripción estándar de los elementos de secuencia de bits para la carga útil de SBR puede encontrarse en ISO/IEC 14496-3, sub-cláusula 4.5.2.8. Una modificación de esta secuencia de bits estándar comprende una ampliación del índice a la tabla de frecuencia maestra (para identificar la frecuencia de cruce utilizada). El índice usado es codificado, por ejemplo, con cuatro bits que permiten la banda de cruce variable sobre un rango de 0 a 15 bandas.

[0073] Las formas de realización de la presente invención pueden entonces ser sintetizadas de la siguiente manera. Las diferentes señales con diferentes características de tiempo/frecuencia poseen diferentes demandas en la característica en la ampliación de ancho de banda. Las señales de tasa de bits (por ejemplo dentro de la señal de voz) necesitan una resolución temporal fina de BWE y la frecuencia de cruce fx (el borde de frecuencia superior del codificador de núcleo) debería ser lo más alto posible (por ejemplo 4 kHz o 5 kHz o 6 kHz). Especialmente en discursos vocalizados, una estructura temporal distorsionada puede disminuir la calidad percibida. Las señales tonales necesitan una reproducción estable de los componentes espectrales y un patrón armónico combinado de las porciones de frecuencia alta reproducidas. La reproducción estable de las partes tonales limita el ancho de banda del codificador de núcleo pero no necesita un BWE con resolución temporal fina sino resolución espectral más fina. En un diseño de codificador de núcleo de voz/audio conmutado, es posible usar la decisión del codificador de núcleo también para adaptar las características temporales y espectrales de BWE así como adaptar la frecuencia de inicio BWE (frecuencia de cruce) a las características de la señal. En consecuencia, las formas de realización proveen una ampliación de ancho de banda donde la decisión del codificador de núcleo actúa como criterio de adaptación a las características de ampliación de ancho de banda.

- [0074]** La señalización de la frecuencia de inicio BWE cambiada (cruce) puede realizarse en forma explícita enviando información adicional (como, por ejemplo, la información del modo de codificación 108) en la secuencia de bits implícitamente derivando la frecuencia de cruce f_x directamente desde el codificador de núcleo usado (en caso que el codificador de núcleo esté, por ejemplo señalizado dentro de la secuencia de bits). Por ejemplo, una frecuencia BWE f_x menor para el codificador de transformación (por ejemplo codificador de audio/música) y una mayor para un codificador en domino de tiempo (voz). En este caso, la frecuencia de cruce puede estar en el rango de 0 Hz a frecuencia Nyquist.
- [0075]** Aunque se han descriptos algunos aspectos en el contexto de un aparato, es claro que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a un paso del método o rasgo de un paso del método. En forma análoga, los aspectos descriptos en el contexto de un paso del método también representan una descripción de un bloque o ítem correspondiente o rasgo de un aparato correspondiente.
- [0076]** La señal de audio codificada ejemplar puede almacenarse en un medio de almacenamiento digital o transmitida en un medio de transmisión como un medio de transmisión inalámbrico o por cable como Internet. Dependiendo de ciertos requisitos de implementación, las formas de realización de la invención pueden implementarse en hardware o software. La implementación puede realizarse utilizando un medio de almacenamiento digital por ejemplo un disquete, DVD, CD, memoria ROM, PROM, EPROM, EEPROM o FLASH, con señales de control capaces de ser leídas en forma electrónica almacenados en ellos, que cooperan (o son capaces de cooperar) con un sistema de computación programable tal que se lleve a cabo el método respectivo.
- [0077]** Algunos ejemplos comprenden un portador de datos con señales de control capaces de ser leídas en forma electrónica, capaces de cooperar con un sistema de computación programable tal que se lleve a cabo uno de los métodos descriptos.
- [0078]** Generalmente, las formas de realización de la presente invención pueden implementarse como un producto de programa de computación con un código de programa, el código de programa es operativo para desarrollar uno de los métodos cuando el producto de programa de computación es utilizado en una computadora. El código de programa puede por ejemplo ser almacenado en un portador capaz de ser leído por una máquina.
- [0079]** Otras formas de realización comprenden el programa de computación para desarrollar uno de los métodos descriptos en la presente, almacenado en un portador capaz de ser leído por una máquina.
- [0080]** En otras palabras, una forma de realización del método de invención consiste, por lo tanto, en un programa de computación con un código de programa para desarrollar uno de los métodos descriptos en la presente, cuando el programa de computación es utilizado en una computadora.
- [0081]** Otro ejemplo consiste, por lo tanto, en un portador de datos (o medio almacenador digital, o medio capaz de ser leído por una computadora) que comprende, el programa de computación grabado en los mismos, para desarrollar uno de los métodos descriptos en la presente.
- [0082]** Otro ejemplo consiste, por lo tanto, en una secuencia de datos o secuencia de señales que representan el programa de computación para desarrollar uno de los métodos descriptos en la presente. La secuencia de datos o secuencia de señales puede por ejemplo ser configurada para ser transferida a través de una conexión de comunicación de datos, por ejemplo por Internet.
- [0083]** Otra forma de realización comprende un medio de procesamiento, por ejemplo una computadora, o un dispositivo lógico programable, configurado o adaptado para desarrollar uno de los métodos descriptos en la presente.
- [0084]** Otra forma de realización comprende una computadora con programa de computación instalado en la misma, para desarrollar uno de los métodos descriptos en la presente.
- [0085]** En algunas formas de realización, un dispositivo lógico programable (por ejemplo una matriz de puertas programables por campo) puede utilizarse para desarrollar una de las funcionalidades de los métodos descriptos en la presente. En algunas formas de realización, una matriz de puertas programables por campo puede cooperar con un microprocesador con el fin de desarrollar uno de los métodos descriptos en la presente. Generalmente, los métodos son preferentemente desarrollados por cualquier aparato de hardware.

[0086] Las formas de realización descritas anteriormente son meramente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en 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 detalle específicos presentados a modo de descripción y explicación de las formas de realización de la presente.

REIVINDICACIONES

1. Un aparato para codificar (200) una señal de audio (105) que comprende:

5

un primer codificador (210a) configurado para codificar de acuerdo con un primer algoritmo de codificación, el primer algoritmo de codificación posee un primer ancho de banda de frecuencia, en el que el primer codificador (210a) es un codificador de voz;

10

un segundo codificador (210b) configurado para codificar de acuerdo con un segundo algoritmo de codificación, el segundo algoritmo de codificación posee un ancho de banda de frecuencia menor que el primer ancho de banda de frecuencia, en el que el segundo codificador (210b) comprende un codificador de música;

15

una etapa de decisión (220) para indicar el primer algoritmo de codificación para una primera porción de tiempo (204a) de la señal de audio (105) y para indicar el segundo algoritmo de codificación para una segunda porción de tiempo (204b) de la señal de audio (105), siendo la segunda porción de tiempo (204b) diferente a la primera porción de tiempo (204a); y

20

un módulo de ampliación de ancho de banda (230) para calcular parámetros de ampliación de ancho de banda (106) para la señal de audio (105), en el que el módulo de ampliación de ancho de banda (230) está configurado para ser controlado por la etapa de decisión (220) para calcular los parámetros de ampliación de ancho de banda (106) para una primera banda que no incluye el primer ancho de banda de frecuencia en la primera porción de tiempo (204a) de la señal de audio (105) y para una segunda banda que no incluye el segundo ancho de banda de frecuencia en la segunda porción (204b) de la señal de audio (105), en el que el primer ancho de banda de frecuencia o el segundo ancho de banda de frecuencia está definido por una frecuencia de cruce (fx) variable controlada por la etapa de decisión (220),

30

en el que la primera banda comprende componentes de frecuencia de la señal de audio (105) que están fuera del primer ancho de banda de frecuencia y están limitados hacia el extremo inferior del espectro por la frecuencia de cruce,

35

en el que la segunda banda comprende componentes de frecuencia de la señal de audio (105) que están fuera del segundo ancho de banda de frecuencia y están limitados hacia el extremo inferior del espectro por la frecuencia de cruce, y en el que la frecuencia de cruce para la primera banda es mayor que la frecuencia de cruce de la segunda banda.

40

2. Un método para codificar una señal de audio (105) que comprende:

45

la codificación de acuerdo con un primer algoritmo de codificación, el primer algoritmo de codificación posee un primer ancho de banda de frecuencia, en el que la etapa de codificación de acuerdo con un primer algoritmo de codificación usa un codificador de voz;

50

la codificación de acuerdo con un segundo algoritmo de codificación, el segundo algoritmo de codificación posee un segundo ancho de banda de frecuencia menor al primer ancho de banda de frecuencia, en el que la etapa de codificación de acuerdo con un segundo algoritmo de codificación comprende el uso de un codificador de música;

55

la indicación, mediante una etapa de decisión (220), del primer algoritmo de codificación para una primera porción de tiempo (204a) de la señal de audio (105) y el segundo algoritmo de codificación para una segunda porción de tiempo (204b) de la señal de audio (105), siendo la segunda porción de tiempo (204b) diferente de la primera porción de tiempo (204a); y

el cálculo de parámetros de ampliación de ancho de banda (106) para la señal de audio (105) de modo tal que los parámetros de ampliación de ancho de banda (106) se calculan para una banda que no incluye el primer ancho de banda de frecuencia en la primera porción de tiempo (204a) de la señal de

audio (105) y para una banda que no incluye el segundo ancho de banda de frecuencia en la segunda porción de tiempo (204b) de la señal de audio (105),

5 en el que la primera banda comprende componentes de frecuencia de la señal de audio (105) que están fuera del primer ancho de banda de frecuencia y están limitados hacia el extremo inferior del espectro por la frecuencia de cruce,

10 en el que la segunda banda comprende componentes de frecuencia de la señal de audio (105) que están fuera del segundo ancho de banda de frecuencia y están limitados hacia el extremo inferior del espectro por la frecuencia de cruce, en el que el primer ancho de banda de frecuencia o el segundo ancho de banda de frecuencia está definido por una frecuencia de cruce variable (f_x) controlada por la etapa de decisión (220), y en el que la frecuencia de cruce para la primera banda es mayor que la frecuencia de cruce de la segunda banda.

15

3. Un programa de computación adaptado para desarrollar, al ser utilizado en una computadora, el método de la reivindicación 2.

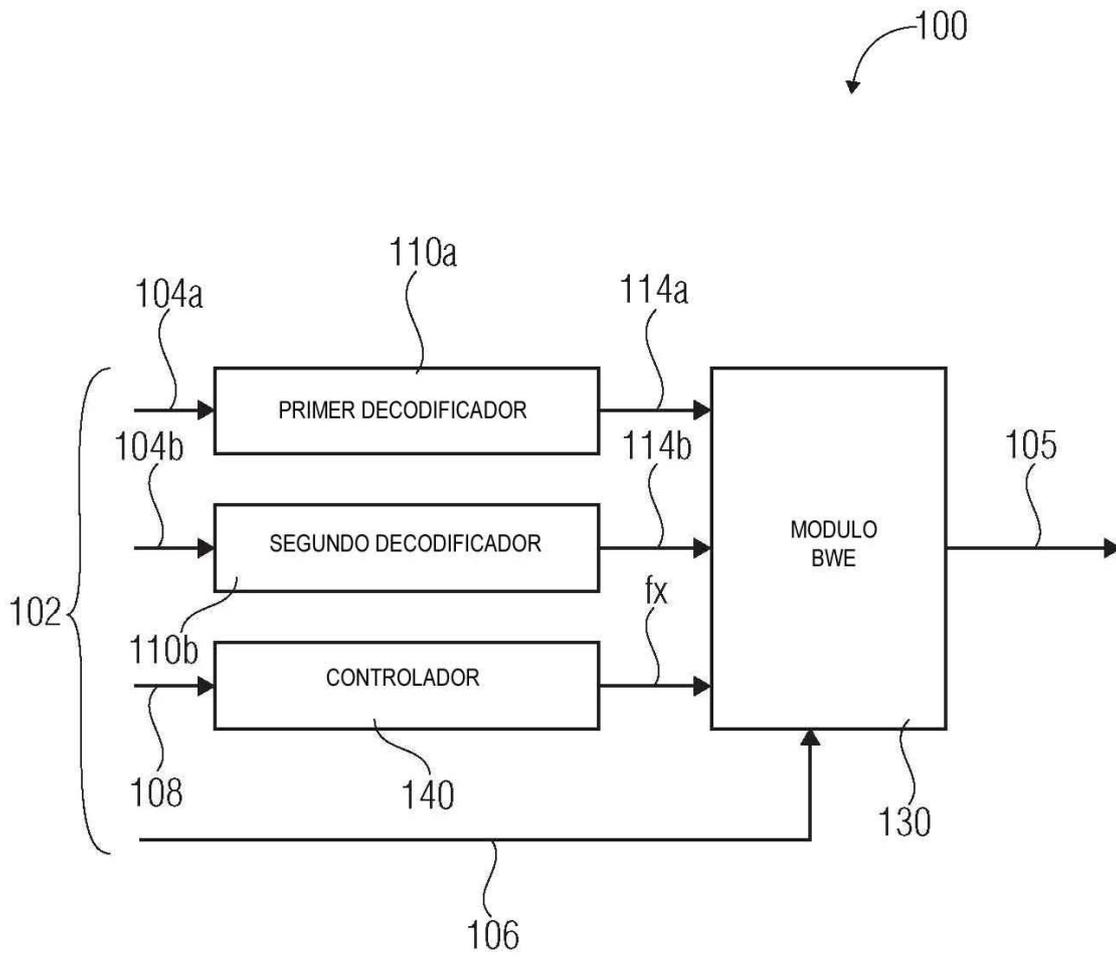


FIG 1

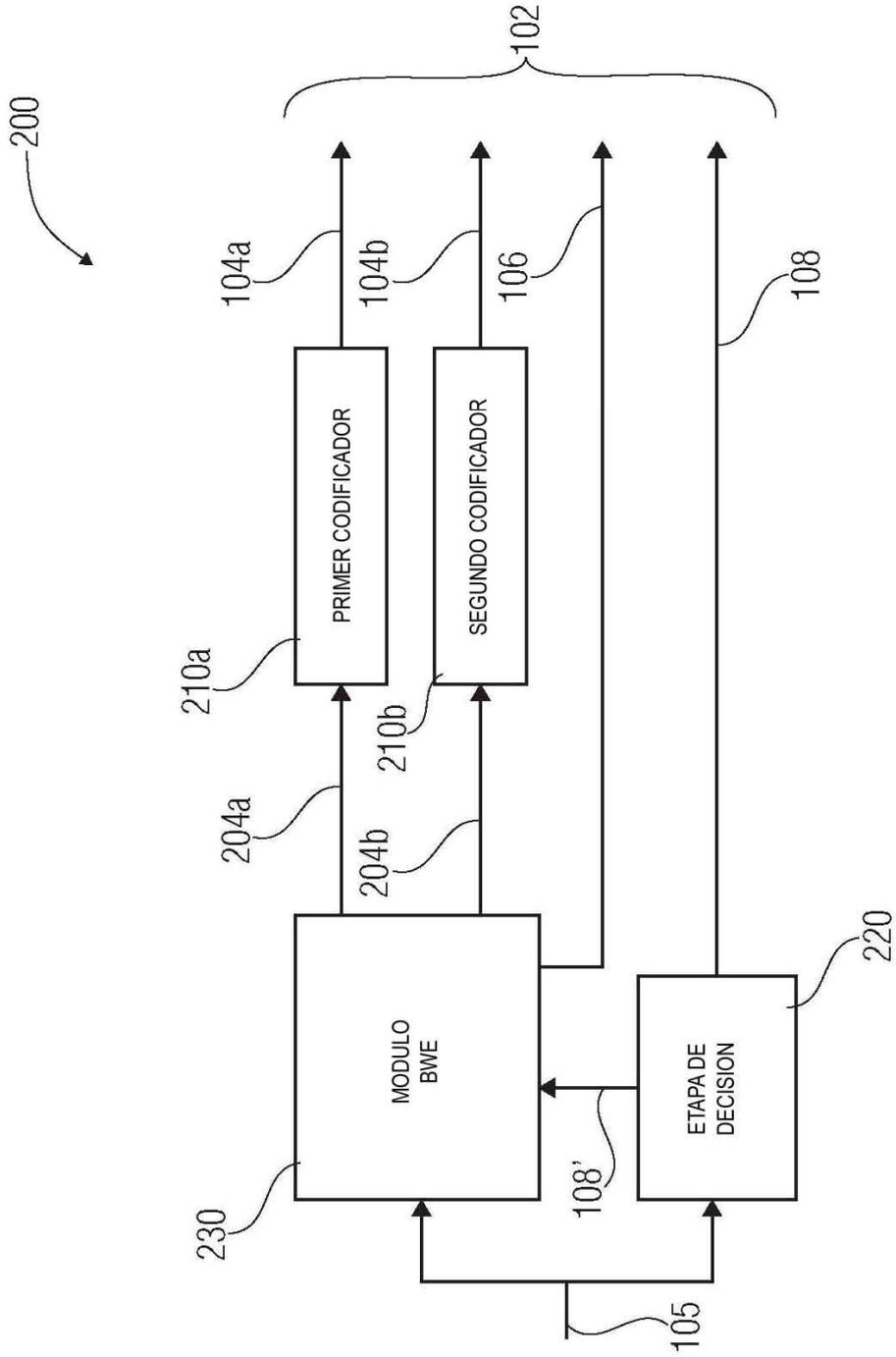


FIG 2

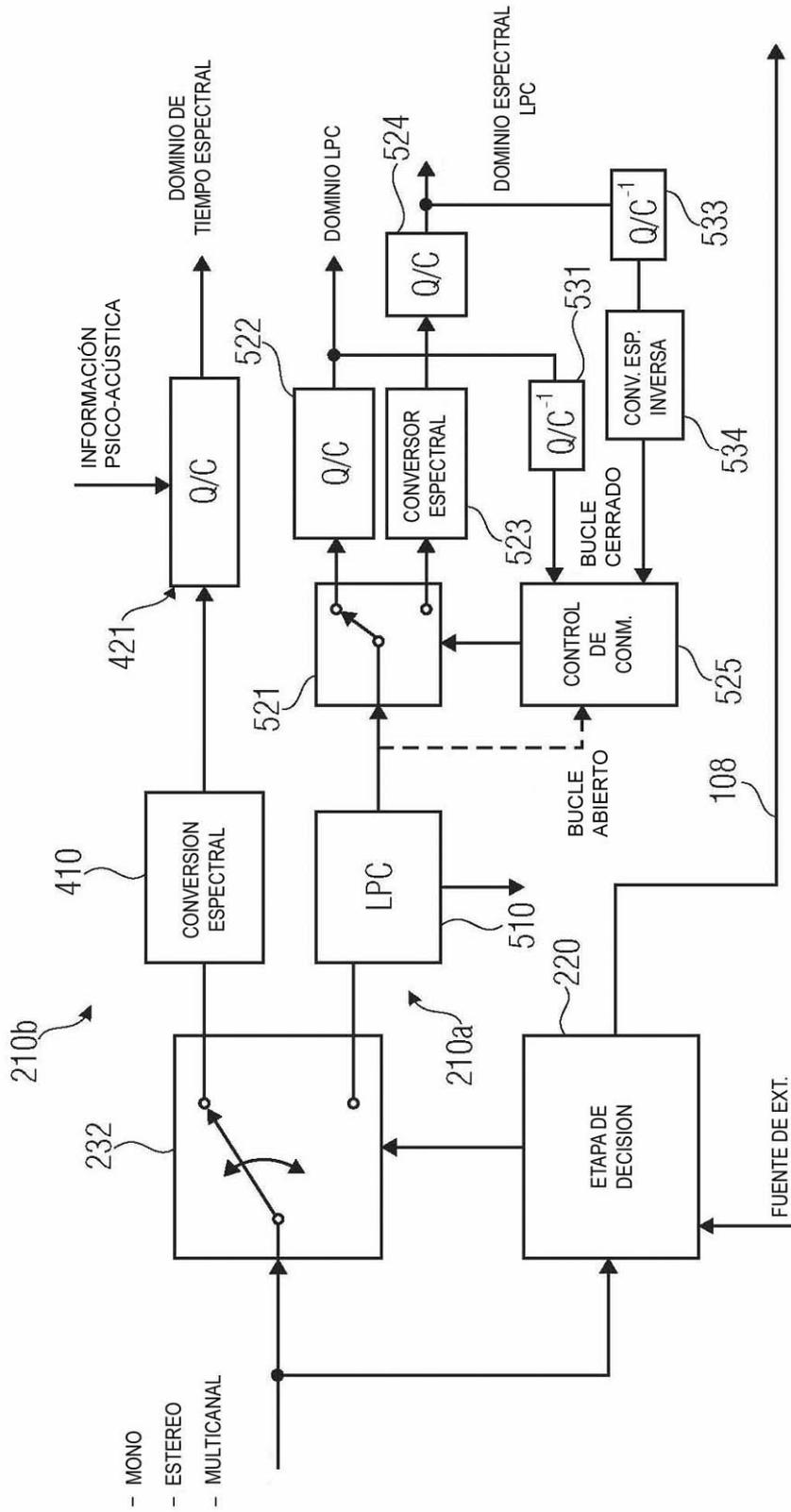


FIG 3
(CODIFICADOR)

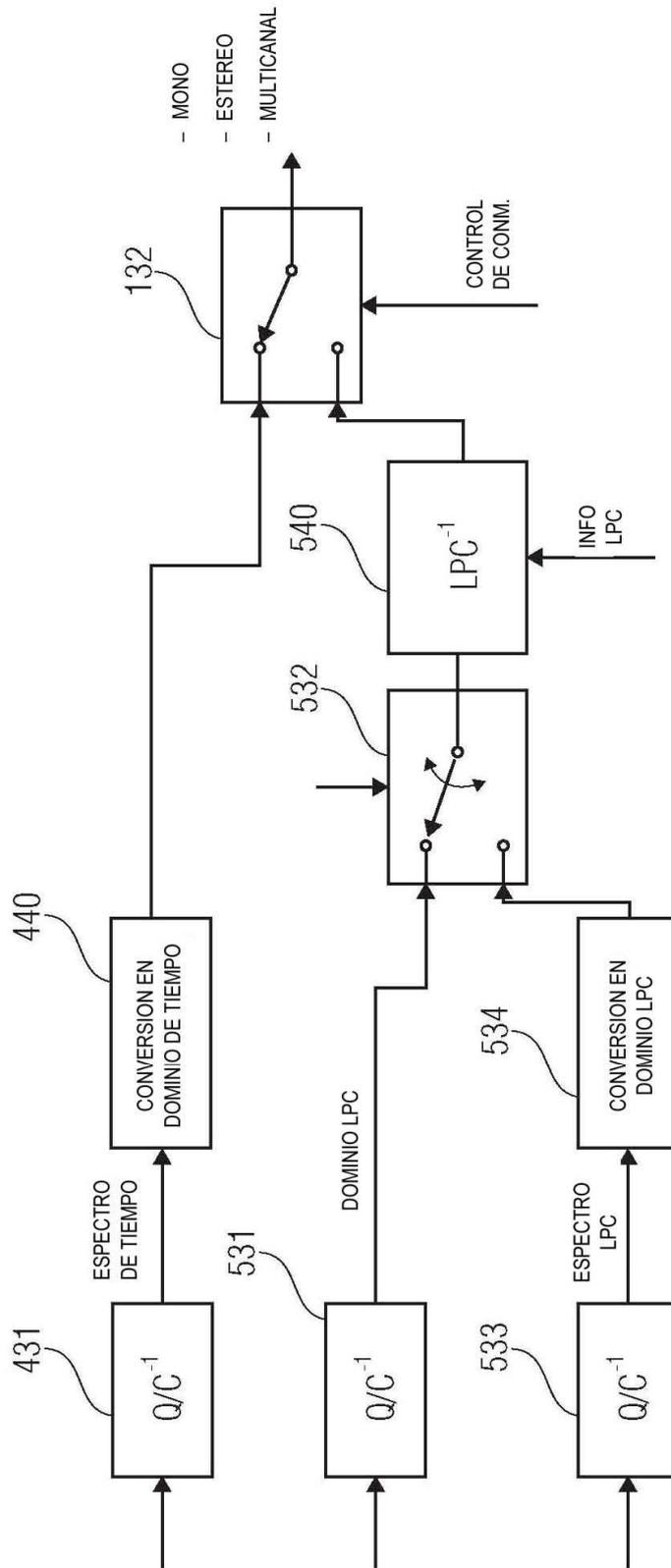


FIG 4
(DECODIFICADOR)

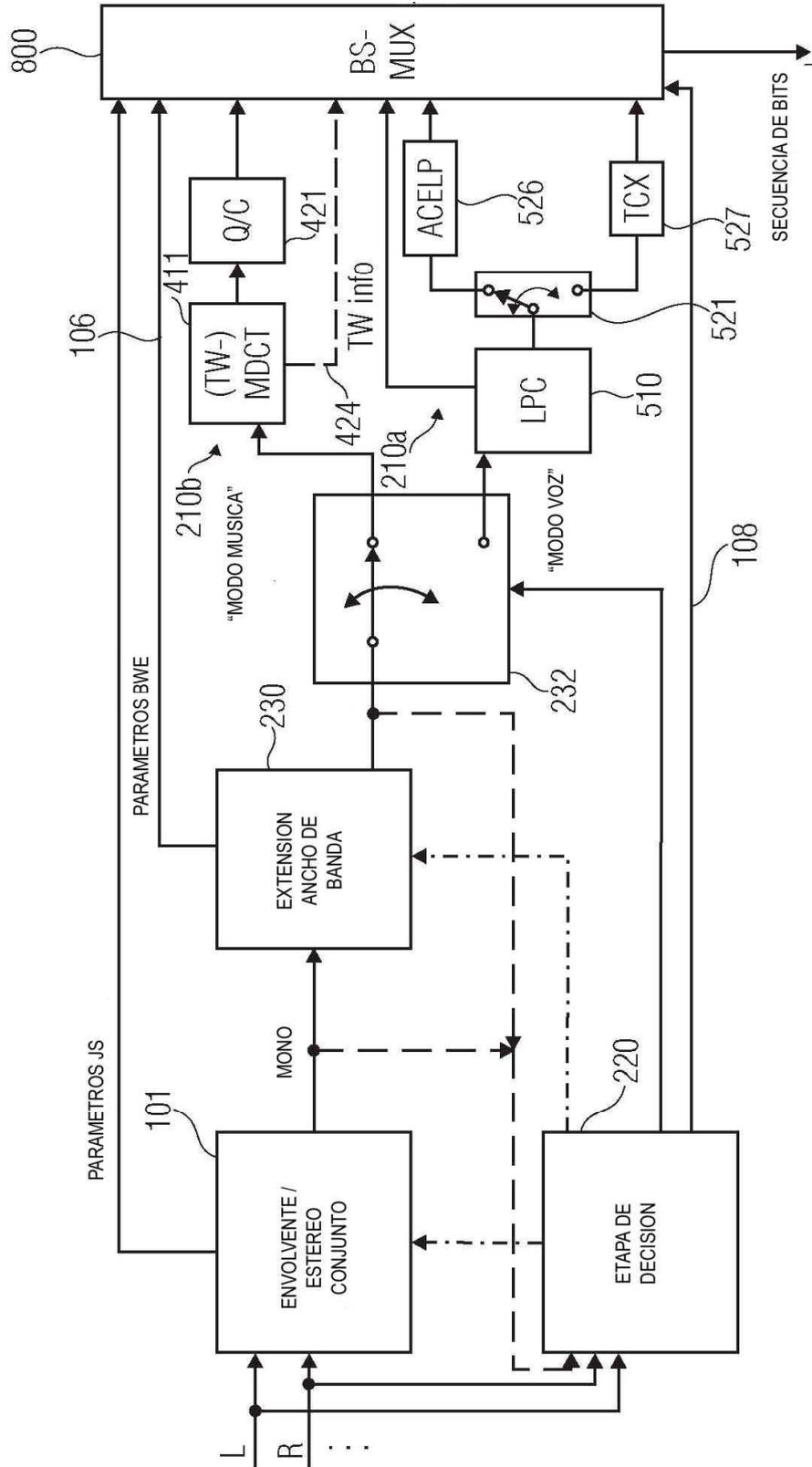


FIG 5
(CODIFICADOR)

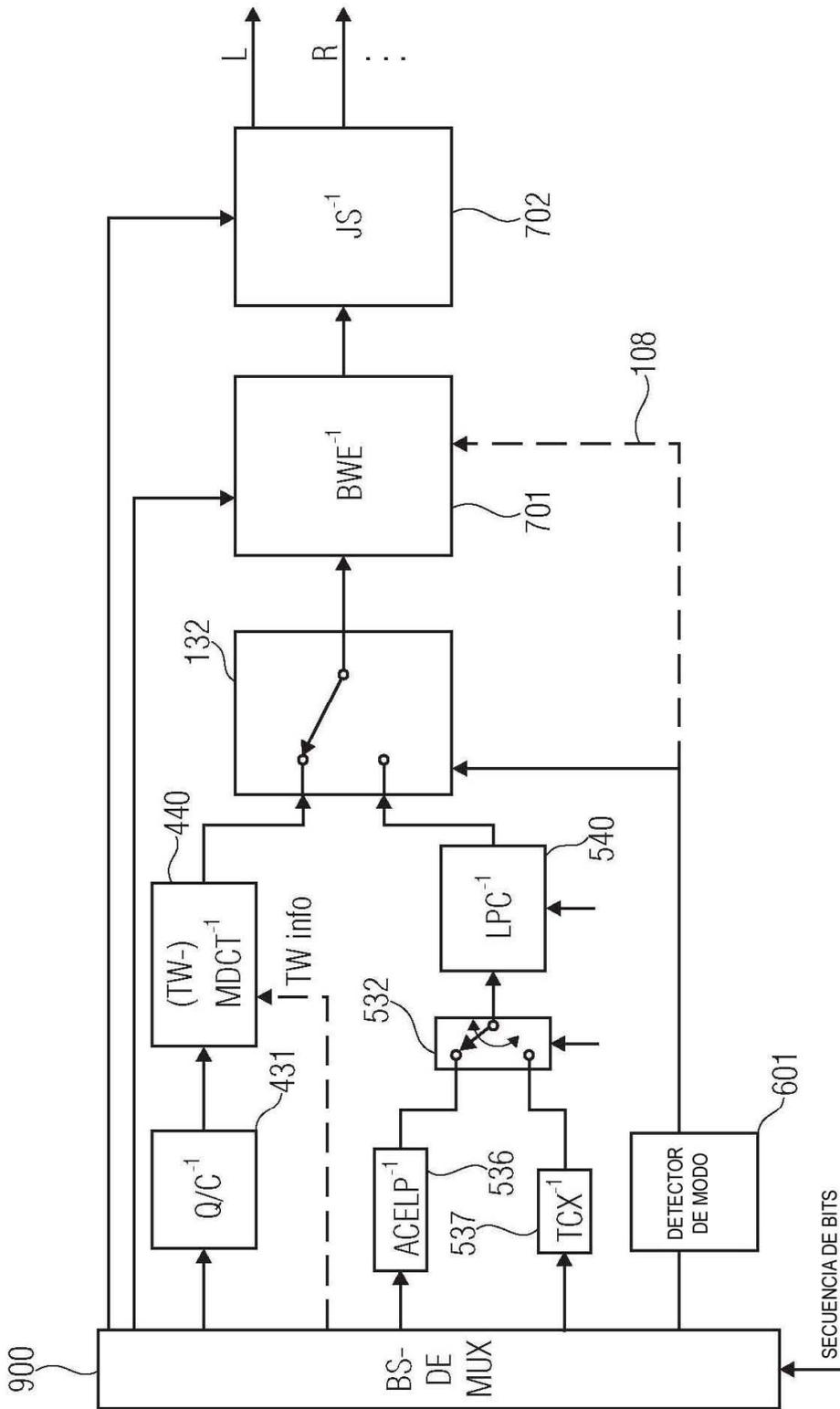


FIG 6
(DECODIFICADOR)

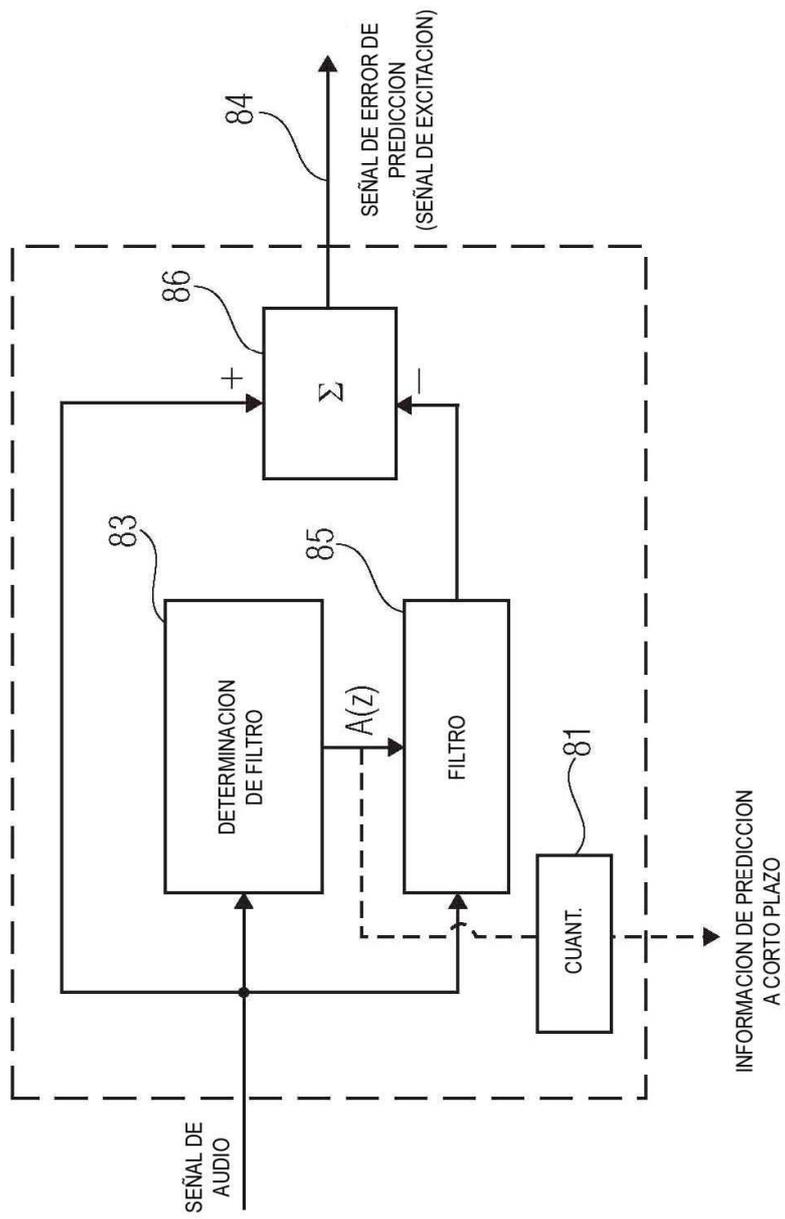


FIG 7

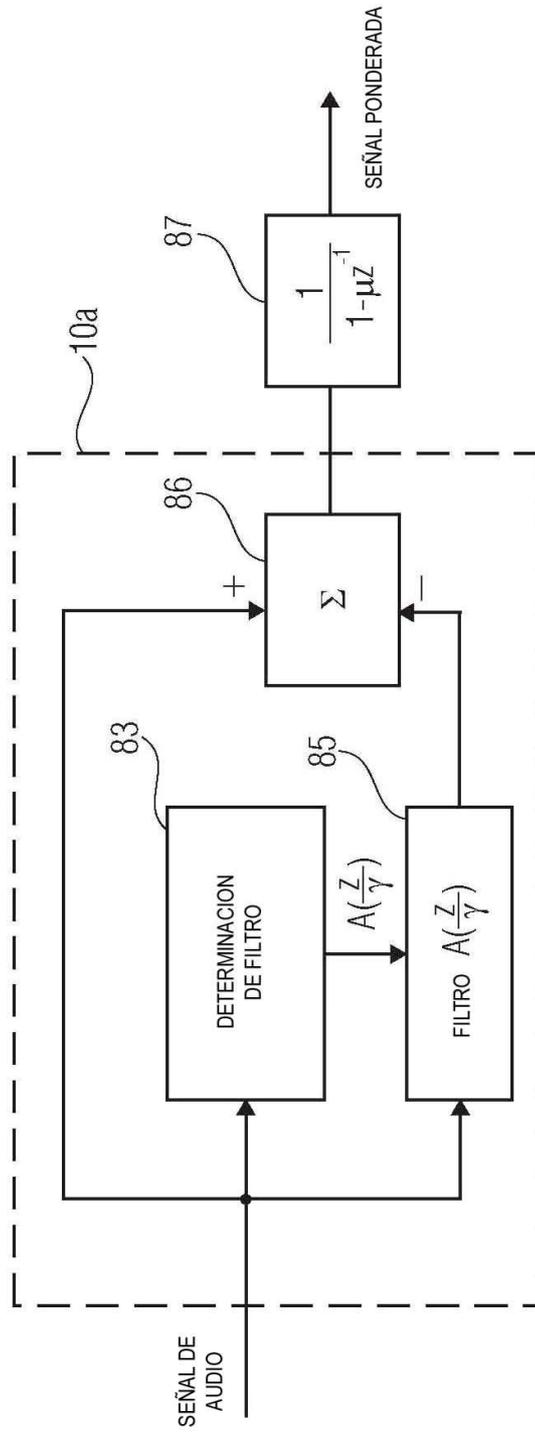


FIG 8

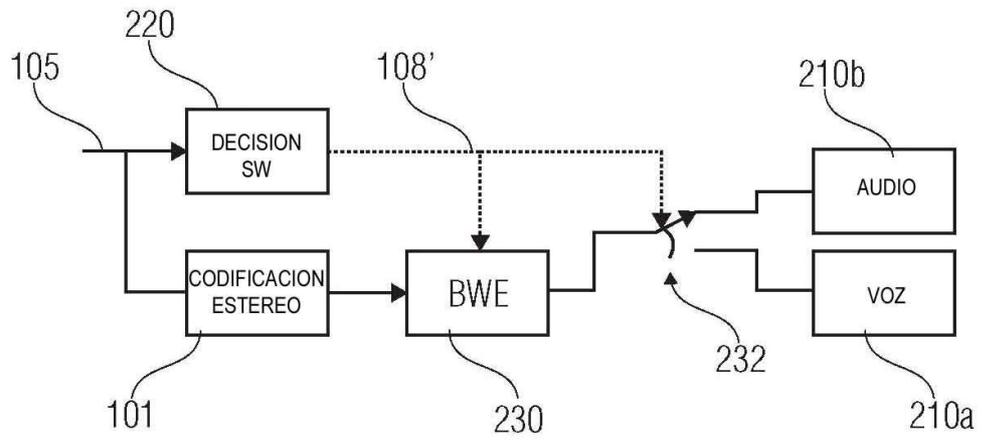


FIG 9A

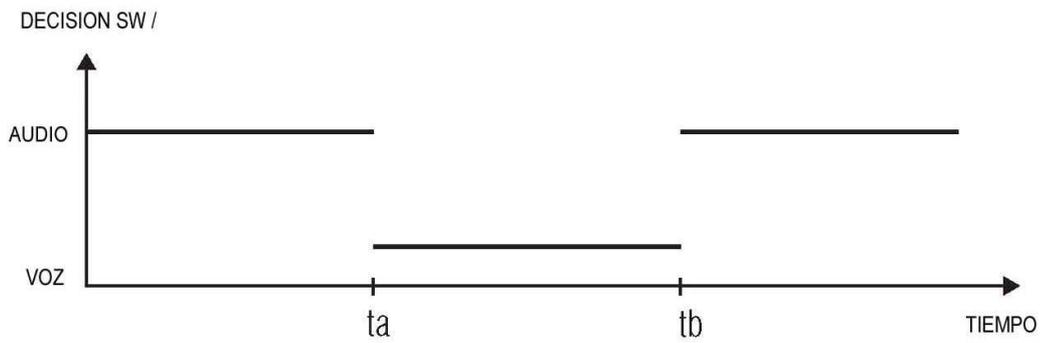


FIG 9B

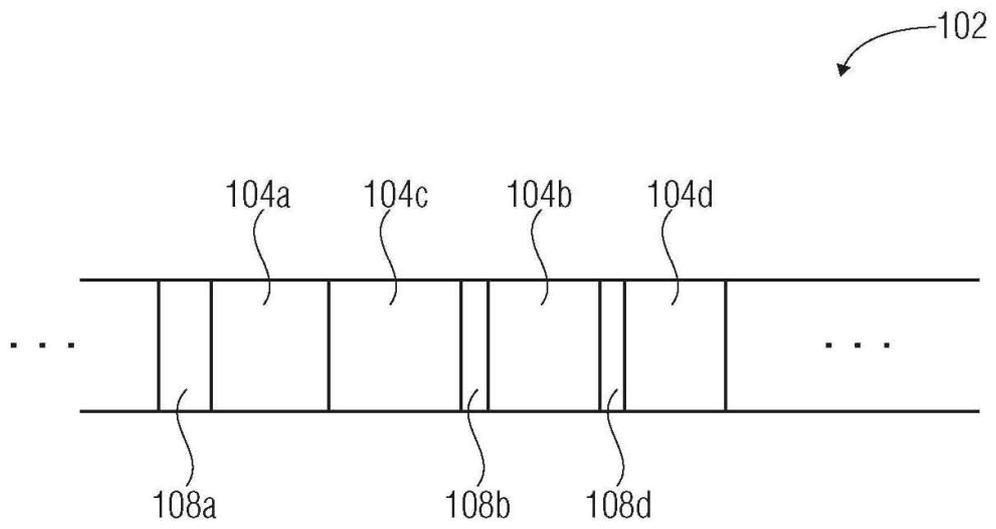


FIG 10