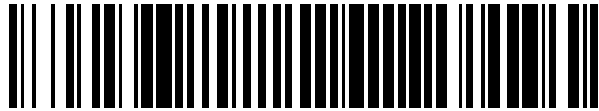


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 401 487**

51 Int. Cl.:

G10L 19/02 (2013.01)

G10L 19/04 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.06.2009 E 09776763 (6)**

97 Fecha y número de publicación de la concesión europea: **02.01.2013 EP 2301020**

54 Título: **Aparato y procedimiento para la codificación/decodificación de una señal de audio utilizando un esquema de conmutación de generación de señal ajena**

30 Prioridad:

11.07.2008 US 79852

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

22.04.2013

73 Titular/es:

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

72 Inventor/es:

**FUCHS, GUILLAUME;
LECOMTE, JÉRÉMIE;
BAYER, STEFAN;
GEIGER, RALF;
MULTRUS, MARKUS;
SCHULLER, GERALD y
HIRSCHFELD, JENS**

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 401 487 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

Aparato y procedimiento para la codificación/decodificación de una señal de audio utilizando un esquema de conmutación de generación de señal ajena

- 5 **[0001]** La presente invención se refiere a la codificación de audio, y en particular, a esquemas de codificación de audio de baja tasa de bits (en inglés, *bit rate*).
- 10 **[0002]** En el arte, se conocen esquemas de codificación de dominio de frecuencia tales como MP3 o AAC (sigla en inglés de "codificación de audio avanzada"). Estos codificadores de dominio de frecuencia se sustentan en una conversión de dominio de tiempo/dominio de frecuencia, una posterior etapa de cuantificación, en la cual el error de cuantificación es controlado usando la información de un módulo psicoacústico, y una etapa de codificación, en la cual los coeficientes espectrales cuantificados y la correspondiente información secundaria son codificados por entropía usando tablas de códigos.
- 15 **[0003]** Por otro lado, hay codificadores que son muy adecuados para el procesamiento del discurso, tales como el AMR-WB+ (sigla en inglés de "multitasa adaptativa-banda ancha"), como se describe en el documento 3GPP TS 26.290. Dichos esquemas de codificación del discurso realizan un filtrado lineal pronosticador de una señal de dominio de tiempo. Dicho filtrado LP (sigla en inglés de "predicción lineal") deriva de un análisis de predicción lineal de la señal de dominio de tiempo de entrada. Los coeficientes del filtro LP resultantes luego son cuantificados/codificados y transmitidos como información secundaria. El proceso se conoce como codificación de predicción lineal (LPC, por sus siglas en inglés). En la salida del filtro, la señal residual de predicción o la señal de error de predicción, que también se conoce como la señal de excitación, es codificada usando las etapas de análisis por síntesis del codificador ACELP (sigla en inglés de "predicción lineal excitada por código de adaptación");
- 20 **[0003]** alternativamente, es codificada usando un codificador de transformación, que utiliza una transformación de Fourier con una superposición. La decisión entre la codificación ACELP y la codificación de excitación codificada de transformación, que también se denomina codificación TCX (sigla en inglés de "excitación codificada de transformación"), se toma usando un algoritmo de bucle cerrado o de bucle abierto.
- 25 **[0004]** Los esquemas de codificación de audio de dominio de frecuencia, tales como el esquema de codificación de alta eficiencia-AAC, que combina un esquema de codificación AAC y una técnica de replicación de banda espectral, pueden también combinarse con una herramienta de codificación de multicanal o estéreo conjunto, que se conoce con el término "MPEG surround".
- 30 **[0005]** Por otro lado, los codificadores del discurso tales como el AMR-WB+ también tienen una etapa de aumento de alta frecuencia y una funcionalidad estéreo.
- [0006]** Los esquemas de codificación de dominio de frecuencia son convenientes por cuanto muestran una alta calidad a bajas tasas de bits para señales de música. Sin embargo, es problemática la calidad de las señales de discurso a bajas tasas de bits.
- 35 **[0007]** Los esquemas de codificación del discurso muestran una alta calidad para señales de discurso, aun a bajas tasas de bits, aunque exhiben una mala calidad para señales de música a bajas tasas de bits.
- [0008]** Los esquemas de codificación de dominio de frecuencia a menudo hacen uso de la así denominada MDCT (MDCT = sigla en inglés de "transformación de coseno separada modificada"). La MDCT ha sido descrita inicialmente en la referencia de J. Princen, A. Bradley: "Analysis/Synthesis Filter Bank Design Based on Time Domain Aliasing Cancellation", *IEEE Trans. ASSP*, ASSP-34 (5): 1153-1161, 1986. La MDCT o el banco de filtros de MDCT se usa ampliamente en los codificadores de audio modernos y eficientes. Esta clase de procesamiento de señales proporciona las siguientes ventajas:
- 40 **[0008]** Atenuación cruzada suave entre los bloques de procesamiento:
- aun si la señal en cada bloque de procesamiento es alterada en forma diferente (por ejemplo, debido a la cuantificación de los coeficientes espectrales), no se producen artefactos de bloqueo debido a transiciones abruptas de bloque a bloque, dada la operación de superposición/adición de ventana.
- 45 **[0009]** Muestreo crítico: la cantidad de valores espectrales en la salida del banco de filtros es igual a la cantidad de valores de entrada de dominio de tiempo en esta entrada, y los valores extra adicionales deben ser transmitidos.
- [0010]** El banco de filtros de MDCT proporciona una selectividad de frecuencia alta y una ganancia de codificación.
- 50 **[0011]** Dichas altas propiedades se logran mediante la utilización de la técnica de cancelación de aliasing (aparición de señal ajena o espúrea) de dominio de tiempo. La cancelación de aliasing de dominio de tiempo se realiza en la síntesis, por medio de la superposición-adición de dos señales de ventana adyacentes. Si no se aplica cuantificación entre las etapas de análisis y síntesis de la MDCT, se obtiene una reconstrucción perfecta de la señal original. Sin embargo, la MDCT se usa para esquemas de codificación, que se adaptan específicamente para señales de música. Dichos esquemas de codificación de dominio de frecuencia tienen, como se establece

anteriormente, menor calidad a bajas tasas de bits o señales de discurso, mientras que los codificadores del discurso específicamente adaptados tienen una mayor calidad a tasas de bits comparables, o aun, tienen tasas de bits significativamente menores para la misma calidad, en comparación con los esquemas de codificación de dominio de frecuencia.

5 **[0012]** Las técnicas de codificación del discurso, tales como la así denominada AMR–WB+ codec como se define en “Extended Adaptive Multi–Rate – Wideband (AMR–WB+) codec”, 3GPP TS 26.290 V6.3.0, 2005–06, Technical Specification, no aplican la MDCT, y por lo tanto, no pueden sacar ninguna ventaja de las excelentes propiedades de la MDCT, que, específicamente, se sustentan en un procesamiento de muestreo crítico por una parte, y un cruce de un bloque al otro, por otra parte. En consecuencia, el cruce de un bloque al otro obtenido por la MDCT sin ninguna penalidad con respecto a la tasa de bits, y por lo tanto, la propiedad de muestreo crítico de MDCT aún no se han obtenido en codificadores del discurso.

10 **[0013]** Cuando se combinan codificadores del discurso y codificadores de audio dentro de un esquema de codificación híbrido individual, se presenta todavía el problema de la manera de obtener una conmutación de un modo de codificación al otro modo de codificación, a una baja tasa de bits y una alta calidad.

15 **[0014]** El documento WO 2008/071353 A2 describe otro ejemplo de un aparato para codificar/ descodificar datos de audio.

[0015] Un objetivo de la presente invención consiste en la provisión de un concepto mejorado de codificación/decodificación.

20 **[0016]** Este objetivo se logra por medio de un aparato para la codificación de una señal de audio de acuerdo con la reivindicación 1; un aparato para la decodificación de una señal de audio codificada de acuerdo con la reivindicación 8; una señal de audio codificada de acuerdo con la reivindicación 14; un procedimiento para la codificación de una señal de audio de acuerdo con la reivindicación 15; un procedimiento para la decodificación de una señal de audio codificada de acuerdo con la reivindicación 16, o un producto de programa para computadora de acuerdo con la reivindicación 17.

25 **[0017]** Un aspecto de la presente invención consiste en la aplicación de un esquema de codificación híbrido, en el cual se aplica un primer modo de codificación adaptado específicamente para ciertas señales, y que opera en un dominio, y en el cual se usa, junto a este primer modo, otro modo de codificación adaptado específicamente para otras señales y para la operación en un dominio diferente. En este concepto de codificación/decodificación, es posible una conmutación críticamente muestreada de un modo de codificación al otro modo de codificación, por cuanto, del lado del codificador, el mismo bloque de muestras de audio que ha sido generado por una operación de ventana es procesado en forma diferente. Específicamente, se procesa una porción del aliasing del bloque de la señal de audio mediante la transformación del subbloque asociado con la porción de aliasing de la ventana de un dominio al otro dominio luego de la ventana de este subbloque, donde un subbloque diferente obtenido por la misma operación de ventana es transformado de un dominio al otro dominio antes de la ventana de este subbloque, usando una ventana de análisis.

30 **[0018]** El primer subbloque procesado y el segundo subbloque procesado, posteriormente, son transformados en un dominio adicional, usando el mismo principio de transformación de bloques, de manera de obtener un primer bloque convertido de la señal de audio, que luego puede ser adicionalmente procesado utilizando cualquiera de los bien conocidos algoritmos de compresión de información, tales como cuantificación, codificación de entropía, etc.

35 **[0019]** Del lado del decodificador, este bloque es nuevamente procesado en forma diferente, sobre la base de la elección de procesamiento de la porción de aliasing del bloque, o de la otra porción del bloque. La porción de aliasing es transformada en un dominio objetivo, antes de efectuar una ventana de síntesis, mientras que la otra porción es sometida a una ventana de síntesis, antes de efectuar la transformación en el dominio objetivo. Además, a fin de obtener la propiedad de muestreo crítico, se realiza una cancelación del aliasing de dominio de tiempo, donde la porción de aliasing de ventana y una porción de aliasing de ventana de otro bloque codificado de la información de audio se combinan luego de una transformación de la porción de aliasing del bloque de señal de audio codificada, en el dominio objetivo, de modo que se obtiene una señal de audio decodificada correspondiente a la porción de aliasing del primer bloque. En vista de lo anterior, existen, de hecho, dos subbloques/porciones en una ventana. Una porción/un subbloque (subbloque de aliasing) tiene componentes de aliasing que se superponen con un segundo bloque codificado en un dominio diferente, y un segundo subbloque/una segunda porción (subbloque adicional), que puede tener o no tener componentes de aliasing que se superponen con el segundo bloque o un bloque diferente del segundo bloque.

40 **[0020]** Preferentemente, el aliasing introducido en ciertas porciones que se corresponden entre sí, pero que son codificadas en diferentes dominios, se usa de modo conveniente a fin de obtener una conmutación muestreada críticamente de un modo de codificación al otro modo de codificación, mediante el procesamiento diferente de la porción de aliasing y la otra porción dentro del mismo bloque ventana de la muestra de audio.

- 5 **[0021]** Esta característica contrasta con el procesamiento del arte previo sobre la base de ventanas de análisis y ventanas de síntesis, ya que, hasta la fecha, un bloque de información completo obtenido mediante la aplicación de una ventana de análisis ha sido sometido al mismo procesamiento. De acuerdo con la presente invención, sin embargo, la porción de aliasing del bloque ventana es procesada en forma diferente, en comparación con la porción adicional de este bloque.
- 10 **[0022]** La otra porción, o porción adicional, puede comprender una porción sin aliasing que se produce cuando se usan ventanas de inicio/detención específicas. Alternativamente, la porción adicional puede comprender una porción de aliasing que se superpone con una porción del resultado de un proceso de ventana adyacente. Entonces, la porción adicional (aliasing) se superpone con una porción de aliasing de un marco vecino procesado en el mismo dominio, en comparación con la otra porción (aliasing) del marco actual, y la porción de aliasing se superpone con una porción de aliasing de un marco vecino procesado en un dominio diferente, en comparación con la porción de aliasing del marco actual.
- 15 **[0023]** Según la implementación, la porción adicional y la porción de aliasing juntas forman el resultado completo de una aplicación de una función de ventana para un bloque de muestras de audio. La porción adicional puede estar completamente libre de aliasing, puede ser de aliasing por completo, o puede incluir una subporción de aliasing y una subporción libre de aliasing. Además, el orden de estas subporciones y el orden de la porción de aliasing y la porción adicional pueden seleccionarse en forma arbitraria.
- 20 **[0024]** En una realización preferida del esquema de codificación de audio conmutado, los segmentos adyacentes de la señal de entrada podrían ser procesados en dos dominios diferentes. Por ejemplo, AAC computa una MDCT en el dominio de señal, y el MTPC (sigla en inglés de "codificador de predicción de transformación de multimodo") (Sean A. Ramprasad: "The Multimode Transform predictive Coding Paradigm," *IEEE Transaction on Speech and Audio Processing*, Vol. 11, Nro. 2, Marzo 2003) computa una MDCT en el dominio residual LPC. Esto podría ser problemático especialmente cuando las regiones superpuestas tienen componentes de aliasing de dominio de tiempo debido al uso de una MDCT. De hecho, el aliasing de dominio de tiempo no puede ser cancelado en las transiciones de un codificador a otro, ya que estos fueron producidos en dos dominios diferentes. Una solución consiste en la realización de las transiciones con señales de ventana de mezcla libres de aliasing. El codificador cambiado entonces no es más muestreado críticamente, y produce valores extra de información. Las realizaciones permiten mantener la ventaja de muestreo crítico, mediante la cancelación de los componentes de aliasing de dominio de tiempo computados por la operación en dos dominios diferentes.
- 25 **[0025]** En una realización preferida de la presente invención, se proveen dos conmutaciones en un orden sucesivo, donde una primer conmutación decide entre la codificación en el dominio espectral usando un codificador de dominio de frecuencia, y la codificación en el dominio LPC, es decir, el procesamiento de la señal en la salida de una etapa de análisis LPC. La segunda conmutación se provee para la conmutación en el dominio LPC, a fin de codificar la señal de dominio LPC o bien en el dominio LPC, tal como usando un codificador ACELP, o la codificación de la señal de dominio LPC en un dominio espectral de LPC, lo que requiere un conversor para la conversión de la señal de dominio LPC en un dominio espectral LPC, que es diferente de un dominio espectral, ya que el dominio espectral de LPC muestra el espectro de una señal filtrada de LPC, en lugar del espectro de la señal de dominio de tiempo.
- 30 **[0026]** La primer conmutación decide entre dos ramas de procesamiento, donde una rama es principalmente motivada por un modelo de destino y/o un modelo psicoacústico, es decir, mediante el enmascarado auditivo, y la otra es principalmente motivada por un modelo de fuente y por cálculos de SNR (sigla en inglés de "relación de señal a ruido") segmentales. A modo de ejemplo, una rama tiene un codificador de dominio de frecuencia y la otra rama tiene un codificador sobre la base de LPC, tal como un codificador del discurso. El modelo de fuente es habitualmente el procesamiento del discurso, y por lo tanto, comúnmente se usa LPC.
- 35 **[0027]** La segunda conmutación decide nuevamente entre dos ramas de procesamiento, pero en un dominio diferente del dominio de primera rama "externa". Nuevamente, una rama "interna" es principalmente motivada por un modelo de fuente o por cálculos SNR, y la otra rama "interna" puede ser motivada por un modelo de destino o un modelo psicoacústico, es decir, mediante el enmascarado, o por lo menos incluye aspectos de codificación de dominio de frecuencia/espectral. A modo de ejemplo, una rama "interna" tiene un codificador de dominio de frecuencia/ conversor espectral, y la otra rama tiene un codificador que codifica el otro dominio, tal como el dominio LPC, donde este codificador es, por ejemplo, un cuantificador/registrador de CELP (sigla en inglés de "predicción lineal excitada por código") o ACELP, que procesa una señal de entrada sin una conversión espectral.
- 40 **[0028]** Una realización preferida adicional es un codificador de audio que comprende una primera rama de codificación orientada a destino de información, tal como una rama de codificación de dominio espectral; una segunda rama de codificación orientada a SNR o fuente de información, tal como una rama de codificación de dominio LPC; y una conmutación para la conmutación entre la primera rama de codificación y la segunda rama de codificación, donde la segunda rama de codificación comprende un conversor en un dominio específico diferente del dominio de tiempo, tal como una etapa de análisis LPC que genera una señal de excitación, y donde la segunda rama de codificación además comprende un dominio específico, tal como una rama de procesamiento de dominio LPC, y un dominio espectral específico, tal como una rama de procesamiento de dominio espectral LPC; y una
- 45
- 50
- 55

conmutación adicional para la conmutación entre la rama de codificación de dominio específico y la rama de codificación de dominio espectral específico.

- 5 **[0029]** Una realización adicional de la invención es un codificador de audio que comprende un primer dominio, tal como una rama de decodificación de dominio espectral; un segundo dominio, tal como una rama de decodificación de dominio LPC, para la decodificación de una señal tal como una señal de excitación en el segundo dominio; y un tercer dominio, tal como una rama decodificadora de LPC–espectral, para la decodificación de una señal tal como una señal de excitación en un tercer dominio, tal como un dominio espectral de LPC, donde el tercer dominio se obtiene efectuando una conversión de frecuencia del segundo dominio, donde se provee una primera conmutación para la segunda señal de dominio y la tercera señal de dominio, y donde se provee una segunda conmutación para la conmutación entre el decodificador de primer dominio y el decodificador para el segundo dominio o el tercer dominio.
- 10 **[0030]** A continuación se describen realizaciones preferidas de la presente invención, con respecto a los dibujos adjuntos, en los cuales:
- 15 la Figura 1A es una representación esquemática de un aparato preferido o procedimiento para la codificación de una señal de audio;
- la Figura 1B es una representación esquemática de la transición de MDCT–TCX a AAC;
- la Figura 1C es una representación esquemática de una transición de AAC a MDCT–TCX;
- la Figura 1D es una ilustración de una realización preferida del concepto de la invención, como un diagrama de flujo;
- 20 la Figura 2 es una representación esquemática para la ilustración de cuatro dominios diferentes y sus relaciones, que se producen en realizaciones de la invención;
- la Figura 3A es un esquema que ilustra un aparato/procedimiento de la invención para la decodificación de una señal de audio;
- la Figura 3B es una ilustración adicional de esquemas de decodificación de acuerdo con realizaciones de la presente invención;
- 25 la Figura 4A ilustra detalles de transformaciones de aliasing, tales como la MDCT aplicable en ambos modos de codificación;
- la Figura 4B ilustra funciones de ventanas comparables con la función de ventana en la Figura 4A, pero con una porción de aliasing y una porción sin aliasing;
- 30 la Figura 5 es una representación esquemática de un codificador y un decodificador en un modo de codificación tal como el modo de codificación AAC–MDCT;
- la Figura 6 es una representación de un codificador y un decodificador que aplica MDCT en un dominio diferente, tal como el dominio LPC en el contexto de la codificación TCX en AMR–WB+;
- la Figura 7 es una secuencia específica de ventanas para transiciones entre AAC y AMR–WB+;
- 35 la Figura 8A es una representación de una realización preferida para un codificador y un decodificador en el contexto de la conmutación del modo TCX al modo AAC;
- la Figura 8B es una realización preferida para la ilustración de un codificador y un decodificador para una transición de AAC a TCX;
- la Figura 9A es un diagrama de bloques de un esquema de codificación cambiado híbrido preferido, en el cual se aplica la presente invención;
- 40 la Figura 9B es un diagrama de flujo que ilustra el proceso efectuado en el controlador de la Figura 9A;
- la Figura 10A es una realización preferida de un decodificador en un esquema de codificación cambiado híbrido;
- la Figura 10B es un diagrama de flujo para la ilustración del procedimiento efectuado en el controlador de transición de la Figura 10A;
- 45 la Figura 11A ilustra una realización preferida de un codificador donde se aplica, preferentemente, la presente invención; y
- la Figura 11B ilustra un decodificador preferido, donde se aplica, preferentemente, la presente invención.

- 5 [0031] La Figura 11A ilustra una realización de la invención que tiene dos conmutaciones en cascada. Una señal monoaural, una señal estéreo o una señal de multicanal se ingresa en un conmutador 200. El conmutador 200 es controlado por una etapa de decisión 300. La etapa de decisión recibe, como una entrada, una entrada de señal en el bloque 200. Alternativamente, la etapa de decisión 300 puede también recibir una información secundaria que se incluye en la señal monoaural, la señal estéreo o la señal de multicanal, o que se asocia por lo menos con dicha señal; donde existe información que, por ejemplo, fue generada cuando se produjo originalmente la señal monoaural, la señal estéreo o la señal de multicanal.
- 10 [0032] La etapa de decisión 300 acciona el conmutador 200 a fin de alimentar una señal, o bien en una porción de codificación de frecuencia 400 ilustrada en una rama superior de la Fig. 11A, o una porción de codificación de dominio LPC 500 ilustrada en una rama inferior en la Fig. 11A. Un elemento clave de la rama de codificación de dominio de frecuencia es un bloque de conversión espectral 411, que es operativo para convertir una señal de salida de etapa de preprocesamiento común (como se describe más adelante) en un dominio espectral. El bloque de conversión espectral puede incluir un algoritmo de MDCT, un QMF (sigla en inglés de "filtro espejo en cuadratura"), un algoritmo FFT (sigla en inglés de "transformación de Fourier rápida"), un análisis Wavelet o un banco de filtros, tal como un banco de filtros críticamente muestreado que tiene una cierta cantidad de canales de banco de filtros, donde las señales de subbanda en este banco de filtros pueden ser señales valoradas reales o señales valoradas complejas. La salida del bloque de conversión espectral 411 es codificada usando un codificador de audio espectral 421, que puede incluir bloques de procesamiento como se conocen del esquema de codificación AAC.
- 15 [0033] En general, el procesamiento en la rama 400 es un procesamiento en un modelo sobre la base de la percepción, o un modelo de destino de información. En consecuencia, esta rama imita el sistema auditivo humano que recibe el sonido. En contraste, el procesamiento en la rama 500 es para la generación de una señal en el dominio de excitación, residual o LPC. En general, el procesamiento en la rama 500 es un procesamiento en un modelo de discurso o un modelo de generación de información. Para señales del discurso, este modelo es un modelo del sistema de generación del discurso/sonido humano que genera sonido. Sin embargo, si debe codificarse un sonido de una fuente diferente que requiere un modelo de generación de sonido diferente, entonces el procesamiento en la rama 500 puede ser diferente.
- 20 [0034] En la rama de codificación inferior 500, un elemento clave es un dispositivo LPC 510, que emite una salida de información de LPC, que se usa para el control de las características de un filtro LPC. Esta información de LPC es transmitida a un decodificador. La señal de salida de la etapa de LPC 510 es una señal de dominio LPC que consiste en una señal de excitación y/o una señal cargada.
- 25 [0035] El dispositivo de LPC, en general, emite una salida de señal de dominio de LPC, que puede ser cualquier señal en el dominio de LPC, tal como una señal de excitación o una señal cargada (TCX), o cualquier otra señal, que ha sido generada mediante la aplicación de coeficientes de filtro de LPC a una señal de audio. Además, un dispositivo de LPC también puede determinar estos coeficientes, y puede asimismo cuantificar/codificar estos coeficientes.
- 30 [0036] La decisión en la etapa de decisión puede adaptarse a la señal, de modo que la etapa de decisión realiza una discriminación de música/discurso y controla el conmutador 200 de manera tal que las señales de música son ingresadas en la rama superior 400, y las señales de discurso son ingresadas en la rama inferior 500. En una realización, la etapa de decisión alimenta su información de decisión a una corriente de bits de salida, de modo que un decodificador puede utilizar esta información de decisión a fin de efectuar las correctas operaciones de decodificación.
- 35 [0037] Dicho decodificador se ilustra en la Figura 11B. La salida de señal del codificador de audio espectral 421, luego de la transmisión, es ingresada en un decodificador de audio espectral 431. La salida del decodificador de audio espectral 431 se ingresa en un conversor de dominio de tiempo 440. En forma análoga, la salida de la rama de codificación de dominio de LPC 500 de la Figura 11A es recibida en el lado del decodificador y procesada por los elementos 536 y 537 para obtener una señal de excitación de LPC. La señal de excitación de LPC es ingresada en una etapa de síntesis de LPC 540, que recibe, como una entrada adicional, la información de LPC generada por la correspondiente etapa de análisis de LPC 510. La salida del conversor de dominio de tiempo 440, o la salida de la etapa de síntesis de LPC 540, son ingresadas en un conmutador 600. El conmutador 600 es controlada por medio de una señal de control de conmutación que es generada, por ejemplo, por la etapa de decisión 300, o que es externamente proporcionada, tal como por un creador de la señal monoaural original, señal estéreo o señal de multicanal original. La salida del conmutador 600 es una señal monoaural completa, estéreo señal o señal de multicanal completa.
- 40 [0038] La señal de salida hacia el conmutador 200 y la etapa de decisión 300 puede ser una señal monoaural, una señal estéreo, una señal de multicanal o, generalmente, una señal de audio. De acuerdo con la decisión que puede derivar de la señal de entrada del conmutador 200, o de cualquier fuente externa, tal como un productor de la señal de audio original que subyace a la entrada de señal en etapa 200, la conmutación cambia entre la rama de codificación de frecuencia 400 y la rama de codificación de LPC 500. La rama de codificación de frecuencia 400 comprende una etapa de conversión espectral 411 y una etapa de cuantificación/codificación posteriormente conectada 421. La etapa de cuantificación/codificación puede incluir cualquiera de las funcionalidades conocidas de
- 45 [0037] Dicho decodificador se ilustra en la Figura 11B. La salida de señal del codificador de audio espectral 421, luego de la transmisión, es ingresada en un decodificador de audio espectral 431. La salida del decodificador de audio espectral 431 se ingresa en un conversor de dominio de tiempo 440. En forma análoga, la salida de la rama de codificación de dominio de LPC 500 de la Figura 11A es recibida en el lado del decodificador y procesada por los elementos 536 y 537 para obtener una señal de excitación de LPC. La señal de excitación de LPC es ingresada en una etapa de síntesis de LPC 540, que recibe, como una entrada adicional, la información de LPC generada por la correspondiente etapa de análisis de LPC 510. La salida del conversor de dominio de tiempo 440, o la salida de la etapa de síntesis de LPC 540, son ingresadas en un conmutador 600. El conmutador 600 es controlada por medio de una señal de control de conmutación que es generada, por ejemplo, por la etapa de decisión 300, o que es externamente proporcionada, tal como por un creador de la señal monoaural original, señal estéreo o señal de multicanal original. La salida del conmutador 600 es una señal monoaural completa, estéreo señal o señal de multicanal completa.
- 50 [0037] Dicho decodificador se ilustra en la Figura 11B. La salida de señal del codificador de audio espectral 421, luego de la transmisión, es ingresada en un decodificador de audio espectral 431. La salida del decodificador de audio espectral 431 se ingresa en un conversor de dominio de tiempo 440. En forma análoga, la salida de la rama de codificación de dominio de LPC 500 de la Figura 11A es recibida en el lado del decodificador y procesada por los elementos 536 y 537 para obtener una señal de excitación de LPC. La señal de excitación de LPC es ingresada en una etapa de síntesis de LPC 540, que recibe, como una entrada adicional, la información de LPC generada por la correspondiente etapa de análisis de LPC 510. La salida del conversor de dominio de tiempo 440, o la salida de la etapa de síntesis de LPC 540, son ingresadas en un conmutador 600. El conmutador 600 es controlada por medio de una señal de control de conmutación que es generada, por ejemplo, por la etapa de decisión 300, o que es externamente proporcionada, tal como por un creador de la señal monoaural original, señal estéreo o señal de multicanal original. La salida del conmutador 600 es una señal monoaural completa, estéreo señal o señal de multicanal completa.
- 55 [0038] La señal de salida hacia el conmutador 200 y la etapa de decisión 300 puede ser una señal monoaural, una señal estéreo, una señal de multicanal o, generalmente, una señal de audio. De acuerdo con la decisión que puede derivar de la señal de entrada del conmutador 200, o de cualquier fuente externa, tal como un productor de la señal de audio original que subyace a la entrada de señal en etapa 200, la conmutación cambia entre la rama de codificación de frecuencia 400 y la rama de codificación de LPC 500. La rama de codificación de frecuencia 400 comprende una etapa de conversión espectral 411 y una etapa de cuantificación/codificación posteriormente conectada 421. La etapa de cuantificación/codificación puede incluir cualquiera de las funcionalidades conocidas de
- 60 [0038] La señal de salida hacia el conmutador 200 y la etapa de decisión 300 puede ser una señal monoaural, una señal estéreo, una señal de multicanal o, generalmente, una señal de audio. De acuerdo con la decisión que puede derivar de la señal de entrada del conmutador 200, o de cualquier fuente externa, tal como un productor de la señal de audio original que subyace a la entrada de señal en etapa 200, la conmutación cambia entre la rama de codificación de frecuencia 400 y la rama de codificación de LPC 500. La rama de codificación de frecuencia 400 comprende una etapa de conversión espectral 411 y una etapa de cuantificación/codificación posteriormente conectada 421. La etapa de cuantificación/codificación puede incluir cualquiera de las funcionalidades conocidas de

los codificadores modernos de dominio de frecuencia, tales como el codificador AAC. Asimismo, la operación de cuantificación en la etapa de cuantificación/codificación 421 puede ser controlada por medio de un módulo psicoacústico, que genera información psicoacústica tal como un umbral de enmascarado psicoacústico sobre la frecuencia, donde esta información es ingresada en la etapa 421.

5 **[0039]** En la rama de codificación de LPC, la señal de salida de conmutación es procesada por medio de una etapa de análisis de LPC 510 que genera información secundaria de LPC y una señal de dominio de LPC. El codificador de excitación comprende una conmutación adicional 521 para la conmutación del procesamiento adicional de la
10 señal de dominio de LPC entre una operación de cuantificación/codificación 526 en el dominio de LPC o una etapa de cuantificación/codificación 527, que procesa valores en el dominio espectral de LPC. A estos fines, se provee un conversor espectral 527. La conmutación 521 es controlada en una forma de bucle abierto, o en una forma de bucle cerrado, de acuerdo con los parámetros específicos, como se describe en la memoria descriptiva técnica de AMR-WB+.

15 **[0040]** Para el modo de control de bucle cerrado, el codificador adicionalmente incluye un cuantificador/codificador inverso, para la señal de dominio de LPC; un cuantificador/codificador inverso, para la señal de dominio espectral de LPC; y un conversor espectral inverso, para la salida del cuantificador/codificador inverso. Ambas señales codificadas y nuevamente decodificadas en las ramas de procesamiento de la segunda rama de codificación son ingresadas en un dispositivo de control de conmutación. En el dispositivo de control de conmutación, estas dos señales de salida se comparan entre sí, o con una función objetivo, o se calcula una función objetivo, lo que puede realizarse sobre la base de una comparación de la distorsión en ambas señales, de modo que la señal que tiene la menor distorsión se usa para la decisión de la posición que debe adoptar la conmutación 521. Alternativamente, en el caso de que ambas ramas proporcionen tasas de bits no constantes, podría seleccionarse la rama que proporciona la menor tasa de bits, aun cuando la relación de señal a ruido de esta rama sea menor que la relación de señal a ruido de la otra rama. En forma alternativa, la función objetivo podría utilizar, como una entrada, la relación de señal a ruido de cada señal y una tasa de bits de cada criterio de señal o criterio adicional, a fin de hallar la mejor decisión para un objetivo específico. Por ejemplo, si el objetivo es aquel donde la tasa de bits debe ser lo más baja posible, entonces la función objetivo se sustentará fuertemente en la tasa de bits de las dos señales emitidas por el cuantificador/codificador inverso y el conversor espectral inverso. Sin embargo, cuando el objetivo principal es contar con la mejor calidad para una cierta tasa de bits, entonces el control de conmutación podría, por ejemplo, desechar cada señal superior a la tasa de bits permitida, y cuando ambas señales se encuentren debajo de la tasa de bits permitida, el control de conmutación seleccionará la señal que tiene la mejor relación de señal a ruido, es decir, que tiene las menores distorsiones de cuantificación/codificación.

20 **[0041]** Como se establece con anterioridad, el esquema de decodificación de acuerdo con la presente invención es ilustrado en la Figura 11B. Para cada una de las tres clases de señales de salida posibles, existe una etapa de decodificación/recuantificación específica 431; 536 ó 537. Mientras que la etapa 431 emite un espectro de frecuencia, que también puede denominarse "espectro de tiempo" (el espectro de frecuencia de la señal de dominio de tiempo), y que es convertido en el dominio de tiempo usando el conversor de frecuencia/tiempo 440, la etapa 536 emite una señal de dominio de LPC, y el ítem 537 recibe un espectro de frecuencia de la señal de dominio de LPC, que también puede denominarse un "espectro de LPC". A fin de garantizar que las señales de entrada en la conmutación 532 se encuentran ambas en el dominio de LPC, se provee un conversor de frecuencia/tiempo 537 en el dominio de LPC. La información de salida de la conmutación 532 es transformada nuevamente en el dominio de tiempo usando una etapa de síntesis de LPC 540, que es controlada por medio de la información secundaria de LPC generada y transmitida por el codificador. Entonces, luego del bloque 540, ambas ramas tienen información de dominio de tiempo que es cambiada de acuerdo con una señal de control de conmutación a fin de, finalmente, obtener una señal de audio, tal como una señal monoaural, una señal estéreo o una señal de multicanal, que depende de la entrada de señal en el esquema de codificación de la Figura 11A.

25 **[0042]** Por lo tanto, la Figura 11A ilustra un esquema de codificación preferido de acuerdo con la invención. Un esquema de preprocesamiento común conectado a la entrada de conmutador 200 puede comprender un bloque de estéreo conjunto/surround 101, que genera, como una salida, parámetros de estéreo conjunto y una señal monoaural de salida, generada mediante la mezcla descendente (*downmixing*) de la señal de entrada, que es una señal que tiene dos o más canales. En general, la señal en la salida del bloque 101 puede además ser una señal que tiene más canales, pero que, debido a la funcionalidad de mezcla descendente del bloque 101, la cantidad de canales en la salida del bloque 101 será menor que la cantidad de entrada de canales en el bloque 101.

30 **[0043]** El esquema de preprocesamiento común puede comprender, alternativamente al bloque 101, o además del bloque 101, una etapa de extensión de ancho de banda 102. En la realización de la Figura 11A, la salida del bloque 101 es ingresada en el bloque de extensión de ancho de banda 102, que, en el codificador de la Figura 11A, emite una señal limitada por la banda, tal como la señal de banda baja o la señal de pase baja en su salida. Preferentemente, esta señal también es muestreada descendientemente (por ejemplo, por un factor de dos). Asimismo, para la banda alta de la entrada de señal en el bloque 102, los parámetros de extensión de ancho de banda, tales como los parámetros de envolvente espectral, los parámetros de filtrado inversa, los parámetros de piso de ruido, etc., como se conocen del perfil de HE-AAC (AAC de alta eficiencia) de MPEG-4, son generados y transmitidos a un multiplexor de corriente de bits 800.

5 **[0044]** Preferentemente, la etapa de decisión 300 recibe la entrada de señal en el bloque 101, o la entrada en el bloque 102, a fin de decidir entre, por ejemplo, un modo de música o un modo de discurso. En el modo de música, se selecciona la rama de codificación superior 400, mientras que, en el modo de discurso, se selecciona la rama de codificación inferior 500. Preferentemente, la etapa de decisión además controla el bloque de estéreo conjunto 101 o el bloque de extensión de ancho de banda 102, o ambos, de manera de adaptar la funcionalidad de estos bloques a la señal específica. En consecuencia, cuando la etapa de decisión determina que una cierta porción de tiempo de la señal de entrada es del primer modo, tal como el modo de música, entonces pueden ser controlados rasgos específicos del bloque 101 y del bloque 102 por la etapa de decisión 300. Alternativamente, cuando la etapa de decisión 300 determina que la señal se encuentra en un modo de discurso, o, en general, en un segundo modo de dominio de LPC, entonces pueden ser controlados rasgos específicos de los bloques 101 y 102 de acuerdo con la salida de la etapa de decisión.

10 **[0045]** Preferentemente, la conversión espectral de la rama de codificación 400 se realiza usando una operación de MDCT que, aún más preferentemente, es la operación de MDCT deformada en el tiempo, donde la intensidad, o en general, la intensidad de deformación puede controlarse entre cero y una alta intensidad de desvío. En una intensidad de desvío de cero, la operación de MDCT en el bloque 411 es una operación de MDCT directa conocida en el arte. La intensidad de desvío de tiempo junto con la información secundaria de desvío de tiempo pueden transmitirse/ingresarse al multiplexor de corriente de bits 800, como información secundaria.

15 **[0046]** En la rama de codificación de LPC, el codificador de dominio de LPC puede incluir un núcleo ACELP 526, que calcula una ganancia de altura, un retardo de altura o una información de libro de código tal como un índice de libro de código y ganancia. El modo TCX, como se conoce de 3GPP TS 26.290, incurre en un procesamiento de una señal perceptualmente cargada en el dominio de transformación. Una señal cargada transformada de Fourier es cuantificada usando una cuantificación de reticulado de tasa múltiple dividida (VQ algebraico) con cuantificación de factor de ruido. Se calcula una transformación en 1024; 512; ó 256 ventanas de muestra. La señal de excitación es recuperada por el filtrado inversa de la señal cargada cuantificada a través de un filtro de carga inverso.

20 **[0046]** En la rama de codificación de LPC, el codificador de dominio de LPC puede incluir un núcleo ACELP 526, que calcula una ganancia de altura, un retardo de altura o una información de libro de código tal como un índice de libro de código y ganancia. El modo TCX, como se conoce de 3GPP TS 26.290, incurre en un procesamiento de una señal perceptualmente cargada en el dominio de transformación. Una señal cargada transformada de Fourier es cuantificada usando una cuantificación de reticulado de tasa múltiple dividida (VQ algebraico) con cuantificación de factor de ruido. Se calcula una transformación en 1024; 512; ó 256 ventanas de muestra. La señal de excitación es recuperada por el filtrado inversa de la señal cargada cuantificada a través de un filtro de carga inverso.

25 En la primera rama de codificación 400, un conversor espectral comprende, preferentemente, una operación de MDCT específicamente adaptada que tiene ciertas funciones de ventana, seguida de una etapa de codificación de cuantificación/entropía que puede consistir en una etapa de cuantificación de vector individual, si bien, preferentemente, es un codificador de cuantificador/entropía escalar combinado, similar al cuantificador/codificador en la rama de codificación de dominio de frecuencia, es decir, en el ítem 421 de la Figura 11A.

30 **[0047]** En la segunda rama de codificación, se presenta el bloque de LPC 510, seguido de una conmutación 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. En general, el bloque ACELP 526 recibe una señal de excitación de LPC. El bloque TCX 527 recibe una señal cargada.

35 **[0048]** En TCX, se aplica la transformación a la señal cargada computada por el filtrado de la señal de entrada a través de un filtro de carga sobre la base de LPC. El filtro de carga utilizado en realizaciones preferidas de la invención es proporcionado por $(1 - A(z/\gamma))/(1 - \mu z^{-1})$. En consecuencia, la señal cargada es una señal de dominio LPC, y su transformación es un dominio espectral de 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, si bien ambas señales se encuentran en el dominio de LPC. La señal de excitación se obtiene mediante el filtrado de la señal de entrada a través del filtro de análisis $(1 - A(z/\gamma))$.

40

[0049] En el lado del decodificador ilustrado en la Figura 11B, después de la transformación espectral inversa en el bloque 537, se aplica el inverso del filtro de carga, es decir, $(1 - \mu z^{-1})/(1 - A(z/\gamma))$. De manera opcional, la señal puede ser filtrada adicionalmente a través de $(1 - A(z))$ para ir hasta el dominio de excitación de LPC. En consecuencia, una señal del bloque TCX⁻¹ 537 puede ser convertida del dominio cargado al dominio de excitación por un filtrado a través de

45

$$\frac{(1 - \mu z^{-1})}{(1 - A(z/\gamma))} (1 - A(z)) ,$$

y luego, ser utilizada en el bloque 536. Este filtrado típica se efectúa en AMR-WB+ al final de la TCX inversa (537) para la alimentación del libro de código de adaptación de ACELP, en el caso de que esta última codificación sea seleccionada para el siguiente marco.

50 **[0050]** Si bien el ítem 510 en la Figura 11A ilustra un solo bloque, el bloque 510 puede emitir diferentes señales, siempre que estas señales se encuentren en el dominio de LPC. El modo real del bloque 510, tal como el modo de señal de excitación o el modo de señal cargada, puede depender del estado de conmutación real. En forma alternativa, el bloque 510 puede tener dos dispositivos de procesamiento paralelos. En consecuencia, el dominio

LPC en la salida de 510 puede representar o bien la señal de excitación de LPC, o la señal cargada de LPC, o cualquier otra señal de dominio de LPC.

5 **[0051]** En la segunda rama de codificación (ACELP/TCX) de la Figura 11a u 11b, la señal es preferentemente preenfatzada, a través de un filtro $1 - 0.68z^{-1}$ antes de la codificación. En el decodificador ACELP/TCX en la Figura 11B, la señal sintetizada es desenfatizada con el filtro $1/(1 - 0.68z^{-1})$. El preénfasis puede ser parte del bloque de LPC 510, donde la señal es preenfatzada antes del análisis LPC y la cuantificación. Asimismo, el desenfatizado puede ser parte del bloque de síntesis de LPC, LPC⁻¹ 540.

[0052] En una realización preferida, el primer conmutador 200 (ver Fig. 11A) es controlado a través de una decisión de bucle abierto, y la segunda conmutación es controlada a través de una decisión de bucle cerrado.

10 **[0053]** A modo de ejemplo, puede presentarse la situación donde, en la primera rama de procesamiento, el primer dominio de LPC representa la excitación de LPC, y en la segunda rama de procesamiento, el segundo dominio de LPC representa la señal cargada de LPC. Esto es, la señal de primer dominio de LPC se obtiene mediante el filtrado a través de $(1 - A(z))$ para convertirse al dominio residual de LPC, mientras que la señal de segundo dominio de LPC se obtiene mediante el filtrado a través del filtro $(1 - A(z/\gamma))/(1 - \mu z^{-1})$ para convertirse al dominio cargado de
15 LPC. En un modo preferido, μ es igual a 0,68.

[0054] La Figura 11B ilustra un esquema de decodificación que corresponde al esquema de codificación de la Figura 11A. La corriente de bits generada por el multiplexor de corriente de bits 800 de la Figura 11a es ingresada en un desmultiplexor de corriente de bits 900. De acuerdo con una información derivada, por ejemplo, de la corriente de bits por medio de un bloque de detección de modo 601, una conmutación del lado de decodificador 600 es controlado, o bien para transmitir señales de la rama superior, o señales de la rama inferior al bloque de extensión de ancho de banda 701. El bloque de extensión de ancho de banda 701 recibe, del desmultiplexor de corriente de bits 900, información secundaria, y sobre la base de esta información secundaria y la salida de la decisión de modo 601, reconstruye la banda alta, sustentada en la salida de banda baja desde el conmutador 600.

20 **[0055]** La señal de banda entera generada por el bloque 701 es ingresada en la etapa de procesamiento de estéreo conjunto/surround 702, que reconstruye dos canales estéreo o varios multicanales. En general, el bloque 702 emitirá más canales que los ingresados en este bloque. De acuerdo con la aplicación, la entrada en el bloque 702 puede aun incluir dos canales, como en un modo estéreo, y puede aun incluir más canales, siempre que la salida por este bloque tenga más canales que la entrada en este bloque.

25 **[0056]** El conmutador 200 ha demostrado cambiar entre ambas ramas, de modo que solo una rama reciba una señal para procesar, y la otra rama no reciba una señal para procesar. En una realización alternativa, sin embargo, la conmutación puede también disponerse luego de, por ejemplo, el codificador de dominio de frecuencia 421 y el codificador de dominio de LPC 510; 521; 526; 527, lo que significa que ambas ramas, 400 y 500, procesan la misma señal en paralelo. Sin embargo, a fin de no duplicar la tasa de bits, solo se selecciona la salida de señal por una de estas ramas de codificación 400 ó 500, para ser escrita en la corriente de bits de salida. La etapa de decisión entonces operará de manera que la señal escrita en la corriente de bits minimice una cierta función de costo, donde la función de costo puede ser la tasa de bits generada o la distorsión de percepción generada, o una función de costo combinada de tasa/distorsión. Por lo tanto, ya sea en este modo, ya sea en el modo ilustrado en las Figuras, la etapa de decisión puede además operar en un modo de bucle cerrado, a fin de garantizar que, finalmente, sólo se escriba una salida de rama de codificación en la corriente de bits que tiene, para una distorsión de percepción determinada, la menor tasa de bits, o para una tasa de bits determinada, tiene la menor distorsión de percepción.

30 **[0057]** En la implementación que tiene dos conmutaciones, es decir, el primer conmutador 200 y la segunda conmutación 521, se prefiere que la resolución de tiempo para la primer conmutación sea menor que la resolución de tiempo para la segunda conmutación. En otras palabras, los bloques de la señal de entrada en la primer conmutación, que pueden cambiarse por medio de una operación de conmutación, son mayores que los bloques cambiados por la segunda conmutación que opera en el dominio de LPC. A modo de ejemplo, la conmutación de dominio de frecuencia/dominio de LPC 200 puede cambiar bloques de una longitud de 1024 muestras, y la segunda conmutación 521 puede cambiar bloques que tienen 256 ó 512 muestras cada uno.

35 **[0058]** En general, el algoritmo de codificación de audio utilizado en la primera rama de codificación 400 refleja e imita la situación en un destino de audio. El destino de una información de audio normalmente es el oído humano. El oído humano puede ser imitado como un analizador de frecuencia. Por lo tanto, la primera rama de codificación emite información espectral codificada. Preferentemente, la primera rama de codificación además incluye un modelo psicoacústico para la aplicación adicional de un umbral de enmascarado psicoacústico. Este umbral de enmascarado psicoacústico se usa cuando se cuantifican valores espectrales de audio donde, preferentemente, la cuantificación se realiza de modo tal que se introduce un ruido de cuantificación mediante la cuantificación de los valores de audio espectrales, que se esconden debajo del umbral de enmascarado psicoacústico.
40
45
50
55

- 5 [0059] La segunda rama de codificación representa un modelo de fuente de información, que refleja la generación de sonido de audio. Por lo tanto, los modelos de fuente de información pueden incluir un modelo de discurso que es reflejado por una etapa de análisis de LPC, es decir, mediante la transformación de una señal de dominio de tiempo en un dominio de LPC, y el posterior procesamiento de la señal residual de LPC, es decir, la señal de excitación. Sin embargo, los modelos de fuentes de sonido alternativos son modelos de fuentes de sonido para la representación de un cierto instrumento o cualquier otro generador de sonido, tal como una fuente de sonido específica existente en el mundo real. Puede efectuarse una selección entre diferentes modelos de fuentes de sonido, cuando se dispone de varios modelos de fuentes de sonido, por ejemplo, sobre la base de un cálculo de SNR, es decir, sobre la base de un cálculo del modelo de fuente más adecuado para la codificación de una cierta porción de tiempo o una cierta porción de frecuencia de una señal de audio. Sin embargo, preferentemente, la conmutación entre las ramas de codificación se realiza en el dominio de tiempo, es decir, que una cierta porción de tiempo es codificada usando un modelo, y una cierta porción de tiempo diferente de la señal intermedia es codificada usando la otra rama de codificación.
- 10 [0060] Los modelos de fuentes de información son representados por ciertos parámetros. Con respecto al modelo de discurso, los parámetros son parámetros de LPC y parámetros de excitación codificada, cuando se considera un codificador de discurso moderno, tal como AMR-WB+. El codificador AMR-WB+ comprende un codificador ACELP y un codificador TCX. En este caso, los parámetros de excitación codificada pueden ser ganancia global, piso de ruido y códigos de longitud variable.
- 15 [0061] La señal de entrada de audio en la Figura 11A se presenta en un primer dominio que, por ejemplo, puede ser el dominio de tiempo, si bien también puede ser cualquier otro dominio tal como un dominio de frecuencia, un dominio de LPC, un dominio espectral de LPC o cualquier otro dominio. En general, la conversión de un dominio al otro dominio se realiza por medio de un algoritmo de conversión, tal como cualquiera de los bien conocidos algoritmos de conversión de tiempo/frecuencia o algoritmos de conversión de frecuencia/tiempo.
- 20 [0062] Una transformación alternativa del dominio de tiempo, por ejemplo, en el dominio de LPC, es el resultado del filtrado de LPC de una señal de dominio de tiempo, que logra una señal residual de LPC o señal de excitación. Puede usarse como un algoritmo de transformación cualquier otra operación de filtrado que produzca una señal filtrada que tenga un efecto sobre una cantidad sustancial de muestras de señal antes de la transformación, según el caso. Por lo tanto, la carga de una señal de audio usando un filtro de carga sobre la base de LPC es una transformación adicional, que genera una señal en el dominio de LPC. En una transformación de tiempo/frecuencia, la modificación de un valor espectral individual tendrá un efecto sobre todos los valores de dominio de tiempo antes de la transformación. En forma análoga, una modificación de cualquier muestra de dominio de tiempo tendrá un efecto sobre cada muestra de dominio de frecuencia. Asimismo, una modificación de una muestra de la señal de excitación en una situación de dominio de LPC tendrá, debido a la longitud del filtro de LPC, un efecto sobre una cantidad sustancial de muestras antes del filtrado de LPC. De igual modo, una modificación de una muestra antes de una transformación de LPC tendrá un efecto sobre muchas muestras obtenidas por esta transformación de LPC, debido al efecto de memoria inherente del filtro de LPC.
- 25 [0063] La Figura 1A ilustra una realización preferida de un aparato para la codificación de una señal de audio 10. La señal de audio es introducida, preferentemente, en un aparato de codificación que tiene una primera rama de codificación, tal como 400 en la Figura 11A, para la codificación de la señal de audio en un tercer dominio que, por ejemplo, puede ser el dominio de frecuencia directa. El codificador además puede comprender una segunda rama de codificación, para la codificación de una señal de audio sobre la base de un cuarto dominio que, por ejemplo, puede ser el dominio de frecuencia de LPC obtenido por el bloque TCX 527 en la Figura 11A.
- 30 [0064] Preferentemente, el aparato de la invención comprende una función de ventana 11, para la ventana del primer bloque de la señal de audio en el primer dominio, usando una primera ventana de análisis que tiene una forma de ventana de análisis, donde la ventana de análisis tiene una porción de aliasing tal como L_k o R_k , como se describe en el contexto de la Figura 8A y la Figura 8B u otras figuras, y que tiene una porción sin aliasing, tal como M_k , ilustrada en la Figura 5 u otras figuras.
- 35 [0065] El aparato además comprende un procesador 12, para el procesamiento de un primer subbloque de la señal de audio asociado con la porción de aliasing de la ventana de análisis, mediante la transformación del subbloque del primer dominio, tal como el dominio de señal o el dominio de tiempo directo, en un segundo dominio tal como el dominio de LPC luego de la ventana del primer subbloque, a fin de obtener un primer subbloque procesado, y para el procesamiento de un segundo subbloque de la señal de audio asociado con la otra porción de la ventana de análisis, mediante la transformación del segundo subbloque del primer dominio, tal como el dominio de tiempo directo, en el segundo dominio, tal como el dominio de LPC, antes de la ventana del segundo subbloque a fin de obtener un segundo subbloque procesado. El aparato de la invención además comprende un transformador 13, para la conversión del primer subbloque procesado y el segundo subbloque procesado del segundo dominio, en el cuarto dominio, tal como el dominio de frecuencia de LPC, usando el mismo principio de transformación de bloques, de manera de obtener un primer bloque convertido. Este primer bloque convertido, entonces, puede ser adicionalmente procesado en una etapa de procesamiento posterior 14, de modo de efectuar una compresión de información.
- 40 [0063] La Figura 1A ilustra una realización preferida de un aparato para la codificación de una señal de audio 10. La señal de audio es introducida, preferentemente, en un aparato de codificación que tiene una primera rama de codificación, tal como 400 en la Figura 11A, para la codificación de la señal de audio en un tercer dominio que, por ejemplo, puede ser el dominio de frecuencia directa. El codificador además puede comprender una segunda rama de codificación, para la codificación de una señal de audio sobre la base de un cuarto dominio que, por ejemplo, puede ser el dominio de frecuencia de LPC obtenido por el bloque TCX 527 en la Figura 11A.
- 45 [0064] Preferentemente, el aparato de la invención comprende una función de ventana 11, para la ventana del primer bloque de la señal de audio en el primer dominio, usando una primera ventana de análisis que tiene una forma de ventana de análisis, donde la ventana de análisis tiene una porción de aliasing tal como L_k o R_k , como se describe en el contexto de la Figura 8A y la Figura 8B u otras figuras, y que tiene una porción sin aliasing, tal como M_k , ilustrada en la Figura 5 u otras figuras.
- 50 [0065] El aparato además comprende un procesador 12, para el procesamiento de un primer subbloque de la señal de audio asociado con la porción de aliasing de la ventana de análisis, mediante la transformación del subbloque del primer dominio, tal como el dominio de señal o el dominio de tiempo directo, en un segundo dominio tal como el dominio de LPC luego de la ventana del primer subbloque, a fin de obtener un primer subbloque procesado, y para el procesamiento de un segundo subbloque de la señal de audio asociado con la otra porción de la ventana de análisis, mediante la transformación del segundo subbloque del primer dominio, tal como el dominio de tiempo directo, en el segundo dominio, tal como el dominio de LPC, antes de la ventana del segundo subbloque a fin de obtener un segundo subbloque procesado. El aparato de la invención además comprende un transformador 13, para la conversión del primer subbloque procesado y el segundo subbloque procesado del segundo dominio, en el cuarto dominio, tal como el dominio de frecuencia de LPC, usando el mismo principio de transformación de bloques, de manera de obtener un primer bloque convertido. Este primer bloque convertido, entonces, puede ser adicionalmente procesado en una etapa de procesamiento posterior 14, de modo de efectuar una compresión de información.
- 55 [0065] El aparato además comprende un procesador 12, para el procesamiento de un primer subbloque de la señal de audio asociado con la porción de aliasing de la ventana de análisis, mediante la transformación del subbloque del primer dominio, tal como el dominio de señal o el dominio de tiempo directo, en un segundo dominio tal como el dominio de LPC luego de la ventana del primer subbloque, a fin de obtener un primer subbloque procesado, y para el procesamiento de un segundo subbloque de la señal de audio asociado con la otra porción de la ventana de análisis, mediante la transformación del segundo subbloque del primer dominio, tal como el dominio de tiempo directo, en el segundo dominio, tal como el dominio de LPC, antes de la ventana del segundo subbloque a fin de obtener un segundo subbloque procesado. El aparato de la invención además comprende un transformador 13, para la conversión del primer subbloque procesado y el segundo subbloque procesado del segundo dominio, en el cuarto dominio, tal como el dominio de frecuencia de LPC, usando el mismo principio de transformación de bloques, de manera de obtener un primer bloque convertido. Este primer bloque convertido, entonces, puede ser adicionalmente procesado en una etapa de procesamiento posterior 14, de modo de efectuar una compresión de información.

- 5 **[0066]** Preferentemente, el procesamiento adicional además recibe, como una entrada, un segundo bloque de la señal de audio en el primer dominio que se superpone con el primer bloque, donde el segundo bloque de la señal de audio en el primer dominio, tal como el dominio de tiempo, es procesado en el tercer dominio, es decir, el dominio de frecuencia directo, usando una segunda ventana de análisis. Esta segunda ventana de análisis tiene una porción de aliasing que corresponde a una porción de aliasing de la primera ventana de análisis. La porción de aliasing de la primera ventana de análisis y la porción de aliasing de la segunda ventana de análisis, preferentemente, se relacionan con las mismas muestras de audio de la señal de audio original antes de la ventana, y estas porciones se someten a una cancelación de aliasing de dominio de tiempo, es decir, un procedimiento de superposición-adición del lado de decodificador.
- 10 **[0067]** La Figura 1B ilustra la situación que se produce cuando tiene lugar la transición de un bloque codificado en el cuarto dominio, por ejemplo, el dominio de frecuencia de LPC, a un tercer dominio, tal como el dominio de frecuencia. En una realización, el cuarto dominio es el dominio de MDCT-TCX, y el tercer dominio es el dominio AAC. Una ventana aplicada a la señal de audio codificada en el dominio de MDCT-TCX tiene una porción de aliasing 20 y una porción sin aliasing 21. El mismo bloque, que se denomina "primer bloque" en la Figura 1B, puede tener o no tener una porción de aliasing adicional 22. Esto también se aplica a la porción sin aliasing. Puede estar presente o ausente.
- 15 **[0068]** El segundo bloque de la señal de audio codificada en el otro dominio, tal como el dominio AAC, comprende una porción de aliasing correspondiente 23, y este segundo bloque puede incluir otras porciones, tales como una porción sin aliasing o una porción de aliasing, según el caso, que se indica en 24 en la Figura 1B. Por lo tanto, la Figura 1B ilustra un procesamiento de superposición de la señal de audio, de modo que las muestras de audio en la porción de aliasing 20 del primer bloque antes de la ventana son idénticas a las muestras de audio en la correspondiente porción de aliasing 23 del segundo bloque antes de la ventana. En consecuencia, las muestras de audio en el primer bloque se obtienen mediante la aplicación de una ventana de análisis a la señal de audio, que es una corriente de muestras de audio, y el segundo bloque se obtiene mediante la aplicación de una segunda ventana de análisis a una cantidad de muestras de audio que incluyen las muestras en la correspondiente porción de aliasing 23 y las muestras en la otra porción 24 del segundo bloque. Por lo tanto, las muestras de audio en la porción de aliasing 20 son el primer bloque de la señal de audio asociado con la porción de aliasing 20, y las muestras de audio en la otra porción 21 de la señal de audio corresponden al segundo subbloque de la señal de audio asociado con la otra porción 21.
- 20 **[0069]** La Figura 1C ilustra una situación similar a aquella de la Figura 1B, aunque como una transición de AAC, es decir, el tercer dominio en el dominio MDCT-TCX, es decir, el cuarto dominio.
- 25 **[0070]** La diferencia entre la Figura 1B y la Figura 1C es, en general, que la porción de aliasing 20 en la Figura 1B incluye muestras de audio que se producen en el tiempo posterior a las muestras de audio en la otra porción 21, mientras que, en la Figura 1C, las muestras de audio en la porción de aliasing 20 se producen, en tiempo, antes que las muestras de audio en la otra porción 21.
- 30 **[0071]** La Figura 1D ilustra una representación detallada de las etapas efectuadas con las muestras de audio en el primer subbloque y el segundo subbloque de un único bloque ventana de muestras de audio. En general, una ventana tiene una porción creciente y una porción decreciente, y de acuerdo con la forma de la ventana, puede haber o no puede haber una porción del medio relativamente constante.
- 35 **[0072]** En una primera etapa 30, se realiza una operación de formación de bloque, en la cual se toma una cierta cantidad de muestras de audio de una corriente de muestras de audio. Específicamente, el bloque que forma la operación 30 definirá las muestras de audio que pertenecen al primer bloque y las muestras de audio que pertenecen al segundo bloque de la Figura 1B y de la Figura 1C.
- 40 **[0073]** Las muestras de audio en la porción de aliasing 20 son sometidas a ventana en una etapa 31a. Sin embargo, es importante el hecho de que las muestras de audio en la porción sin aliasing, es decir, en el segundo subbloque, son transformadas en el segundo dominio, es decir, el dominio de LPC en la realización preferida en la etapa 32. Entonces, luego de la transformación de las muestras de audio en el segundo subbloque, se realiza la operación de ventana 31b. Las muestras de audio establecidas por la operación de ventana 31b forman las muestras que son ingresadas en una operación de transformación de bloque al cuarto dominio, ilustrado en la Figura 1D como el ítem 35.
- 45 **[0074]** La operación de ventana en el bloque 31a, 31b, puede incluir o no incluir una operación de plegado, como se describe en relación con las Figuras 8A, 8B, 9A, 10A. Preferentemente, la operación de ventana 31a, 31b, además comprende una operación de plegado.
- 50 **[0075]** Sin embargo, la porción de aliasing es transformada en el segundo dominio, tal como el dominio de LPC en el bloque 33. En consecuencia, el bloque de muestras por ser transformado en el cuarto dominio indicado en 34 es completado, y el bloque 34 constituye un bloque de entrada de información en una operación de transformación de bloque, tal como una operación de tiempo/frecuencia. Debido a que el segundo dominio, en la realización preferida, es el dominio de LPC, la salida de la operación de transformación de bloque como en la etapa 35 será en el cuarto
- 55

- dominio, es decir, el dominio de frecuencia de LPC. Este bloque generado por la transformación de bloque 35 será el primer bloque convertido 36, que luego es primero procesado en la etapa 37, a fin de aplicar cualquier clase de compresión de información que comprende, por ejemplo, las operaciones de compresión de información aplicadas a la información TCX en el codificador AMR-WB+. Naturalmente, todas las otras operaciones de compresión de información pueden efectuarse también en el bloque 37. Por lo tanto, el bloque 37 corresponde al ítem 14 en la Figura 1A, y el bloque 35 en la Figura 1D corresponde al ítem 13 en la Figura 1A; las operaciones de ventana corresponden a 31b y 31a en la Figura 1D, que corresponden al ítem 11 en la Figura 1A, y la programación del orden entre la transformación y la ventana que es diferente para la otra porción y la porción de aliasing es realizada por el procesador 12 en la Figura 1A.
- 5
- 10 **[0076]** La Figura 1D ilustra el caso en el cual la otra porción consiste en la subporción sin aliasing 21 y una subporción de aliasing 22 de la Figura 1B o 1C. Alternativamente, la otra porción puede incluir solo una porción de aliasing sin una porción sin aliasing. En este caso, 21 en la Figura 1B y en la Figura 1C no estaría allí, y 22 se extendería del borde del bloque al borde de la porción de aliasing 20. En cualquier caso, la otra porción/otro subbloque se procesa de la misma manera (sin consideración de estar completamente libre de aliasing o de tener completamente aliasing, o de tener una subporción de aliasing y una subporción sin aliasing), si bien en forma diferente del subbloque de aliasing.
- 15
- [0077]** La Figura 2 ilustra una reseña sobre diferentes dominios que se producen en realizaciones preferidas de la presente invención.
- 20 **[0078]** Normalmente, la señal de audio estará en el primer dominio 40 que, por ejemplo, puede ser el dominio de tiempo. Sin embargo, la invención en realidad se aplica a todas las situaciones que se producen cuando una señal de audio debe ser codificada en dos dominios diferentes, y cuando la conmutación de un dominio al otro dominio debe efectuarse en una forma óptima en relación con la tasa de bits, es decir, usando muestreado crítico.
- [0079]** En una realización preferida, el segundo dominio será un dominio de LPC 41. Una transformación del primer dominio al segundo dominio se realizará mediante un filtro de LPC/transformación, como se indica en la Figura 2.
- 25 **[0080]** En una realización preferida, el tercer dominio es el dominio de frecuencia directo 42 que se obtiene por cualquiera de las transformaciones de tiempo/frecuencia bien conocidas, tales como una DCT (transformación de coseno discreta, por sus siglas en inglés), una DST (transformación de seno discreta, por sus siglas en inglés), una transformación de Fourier o una transformación de Fourier rápida, o cualquier otra transformación de tiempo/frecuencia.
- 30 **[0081]** De manera correspondiente, una conversión del segundo dominio en un cuarto dominio 43, tal como un dominio de frecuencia de LPC o, en términos generales, el dominio de frecuencia con respecto al segundo dominio 41, también puede obtenerse por cualquiera de los algoritmos de transformación de tiempo/frecuencia bien conocidos, tales como DCT, DST, FT (sigla en inglés de "transformación de Fourier"), FFT.
- 35 **[0082]** Entonces, la Figura 2 se compara con la Figura 11A o la Figura 11B; la salida del bloque 421 tendrá una señal en el tercer dominio. Además, la salida del bloque 526 tendrá una señal en el segundo dominio, y la salida del bloque 527 comprenderá una señal en el cuarto dominio. La otra entrada de señal en el conmutador 200, o en general, la entrada en la etapa de decisión 300 o la etapa de surround/estéreo conjunto 101 estará en el primer dominio, tal como el dominio de tiempo.
- 40 **[0083]** La Figura 3A ilustra una realización preferida de un aparato de la invención para la decodificación de una señal de audio codificada que tiene un primer bloque codificado 50 de información de audio, donde el bloque codificado tiene una porción de aliasing y una porción adicional. El decodificador de la invención además comprende un procesador 51, para el procesamiento de la porción de aliasing mediante la transformación de la porción de aliasing en un dominio objetivo, para efectuar una ventana de síntesis de manera de obtener una porción de aliasing de ventana 52, y para efectuar una ventana de síntesis de la porción adicional antes de realizar una transformación de la porción adicional de ventana en el dominio objetivo.
- 45 **[0084]** Por lo tanto, del lado del decodificador, las porciones de un bloque que pertenece a la misma ventana son procesadas en forma diferente. Se ha aplicado un procesamiento similar del lado del codificador, a fin de permitir una conmutación críticamente muestreado entre diferentes dominios.
- 50 **[0085]** El decodificador de la invención además comprende un cancelador de aliasing de dominio de tiempo 53, para la combinación de la porción de aliasing de ventana del primer bloque, es decir, la entrada 52, y una porción de aliasing de ventana de un segundo bloque codificado de información de audio posterior a una transformación de la porción de aliasing del segundo bloque codificado en el dominio objetivo, a fin de obtener una señal de audio codificada 55, que corresponde a la porción de aliasing del primer bloque. La porción de aliasing de ventana del segundo bloque codificado es ingresada por medio de 54 en el cancelador de aliasing de dominio de tiempo 53.
- 55 **[0086]** Preferentemente, un cancelador de aliasing de dominio de tiempo 53 es implementado como un dispositivo de superposición/adición, que, por ejemplo, aplica una superposición del 50%. Esto significa que el resultado de una ventana de síntesis de un bloque se superpone con el resultado de un procesamiento de ventana de síntesis de un

- bloque codificado adyacente de información de audio, donde esta superposición, preferentemente, comprende 50% del bloque. Esto significa que la segunda porción de información de audio de ventana de síntesis de un bloque anterior se agrega en forma de muestra a la primera porción de un segundo bloque posterior de información de audio codificada, de modo que, al final, las muestras de audio decodificadas son la suma de correspondientes muestras de ventanas de dos bloques adyacentes. En otras realizaciones, el rango de superposición puede ser más o menos del 50%. Este rasgo de combinación del cancelador de aliasing de dominio de tiempo proporciona una mezcla continua de un bloque al siguiente, lo que elimina por completo cualquier artefacto de bloqueo que se produzca en cualquier esquema de codificación de transformación sobre la base de bloques. Debido a que las porciones de aliasing de diferentes dominios pueden ser combinadas por la presente invención, se obtiene una operación de conmutación críticamente muestreada de un bloque de un dominio a un bloque del otro dominio.
- [0087]** En comparación con un codificador de conmutación sin ninguna mezcla, donde se realiza una conmutación dura de un bloque al otro bloque, la calidad de audio mejora mediante el presente procedimiento, ya que la conmutación dura provocará, inevitablemente, artefactos de bloqueo tales como crujidos perceptibles o cualquier otro ruido indeseado en el borde del bloque.
- [0088]** En comparación con la mezcla no críticamente muestreada, que, de hecho, eliminaría dicho ruido agudo indeseado en el borde del bloque, la presente invención, sin embargo, no produce ningún incremento de tasa de información debido a la conmutación. Cuando, en el arte previo, las mismas muestras de audio son codificadas en el primer bloque por medio de la primera rama de codificación y son codificadas en el segundo bloque por medio de la segunda rama de codificación, una cantidad de muestra ha sido codificada en ambas ramas de codificación, lo que consumirá tasa de bits, cuando se procesa sin una introducción de aliasing. De acuerdo con la presente invención, sin embargo, se introduce un aliasing en los bordes de bloque. Esta introducción de aliasing que se obtiene por una reducción de muestra, no obstante, logra una posibilidad de aplicar una operación de mezcla por medio del cancelador de aliasing de dominio de tiempo 53, sin la penalidad de una mayor tasa de bits o una conmutación no críticamente muestreada.
- [0089]** En la realización de mayor preferencia, se realiza una conmutación verdaderamente críticamente muestreada. Sin embargo, en ciertas situaciones, puede haber también menos realizaciones eficientes, en las cuales se introduce solo una cierta cantidad de aliasing, y se permite una cierta cantidad de información extra de tasa de bits. Debido a que las porciones de aliasing se usan y se combinan, todas estas realizaciones menos eficientes, sin embargo, son siempre mejores que una transición libre de aliasing por completo, sin atenuación cruzada, o con respecto a la calidad, son mejores que una conmutación dura de una rama de codificación a la otra rama de codificación.
- [0090]** En este contexto, debe observarse que la porción sin aliasing en TCX aún produce muestras codificadas críticamente muestreadas. La adición de una porción sin aliasing en TCX no compromete el muestreo crítico, si bien compromete la calidad de la transición (menor transmisión) y la calidad de la representación espectral (menor compactación de energía).
- [0091]** En vista de lo anterior, es preferible tener la porción sin aliasing en TCX lo más pequeña posible, o aún cercana a cero, de modo que la otra porción sea aliasing por completo y no tenga una subporción libre de aliasing.
- [0092]** A continuación, se describirá la Figura 3B a fin de ilustrar una realización preferida del procedimiento de la Figura 3A.
- [0093]** En una etapa 56, se realiza el procesamiento del decodificador del primer bloque codificado que, por ejemplo, se encuentra en el cuarto dominio. Este procesamiento de decodificador puede ser una decodificación de entropía, tal como la decodificación de Huffman, o una decodificación aritmética que corresponde a las operaciones de procesamiento adicionales en el bloque 14 de la Figura 1A del lado del codificador. En la etapa 57, se realiza una conversión de frecuencia/tiempo del primer bloque completo, como se indica en la etapa 57. De acuerdo con la Figura 2, este procedimiento en la etapa 57 logra un primer bloque completo en el segundo dominio. Ahora, de acuerdo con la presente invención, las porciones del primer bloque son procesadas en forma diferente. Específicamente, la porción de aliasing, es decir, el primer subbloque de la salida de la etapa 57, se transformará en el dominio objetivo antes de efectuar una operación de ventana, usando una ventana de síntesis. Esto se indica mediante el orden de la etapa de transformación 58a y la etapa de ventana 59a. El segundo subbloque, es decir, el subbloque libre de aliasing, es sometido a una ventana usando una ventana de síntesis como se indica en 59b, como tal, es decir, sin la operación de transformación en el ítem 58a en la Figura 3B. La operación de ventana en el bloque 59a o 59b puede comprender o no comprender una operación de plegado (desplegado). Preferentemente, sin embargo, la operación de ventana comprende una operación de plegado (desplegado).
- [0094]** De acuerdo con el hecho de que si el segundo subbloque correspondiente a la porción adicional es, en realidad, un subbloque de aliasing o un subbloque sin aliasing, la operación de transformación en el dominio objetivo tal como se indica en 59b se realiza sin ninguna operación de TDAC (sigla en inglés de "cancelación de aliasing de dominio de tiempo")/operación de combinación en el caso de que el segundo subbloque sea un subbloque sin aliasing. Sin embargo, cuando el segundo subbloque es un subbloque de aliasing, se realiza una operación de TDAC, es decir, una operación de combinación 60b, con una porción correspondiente de otro bloque, antes de obtener la operación

de transformación en el dominio objetivo en la etapa 59b a fin de calcular la señal de audio decodificada para el segundo bloque.

- 5 [0095] En la otra rama, es decir, para la porción de aliasing correspondiente al primer subbloque, el resultado de la operación de ventana en la etapa 59a es ingresado en una etapa de combinación 60a. Esta etapa de combinación 60a además recibe, como una entrada, la porción de aliasing del segundo bloque, es decir, el bloque que ha sido codificado en el otro dominio, tal como el dominio AAC en el ejemplo de la Figura 2. Entonces, la salida del bloque 60a constituye la señal de audio decodificada para el primer subbloque.
- 10 [0096] Cuando se comparan la Figura 3A y la Figura 3B, se torna evidente que la operación de combinación 60a corresponde al procesamiento efectuado en el bloque 53 de la Figura 3A. Además, la operación de transformación y la operación de ventana efectuadas por el procesador 51 corresponden a los ítems 58a, 58b, con respecto a la operación de transformación, y 59a y 59b, con respecto a la operación de ventana, donde el procesador 51 en la Figura 3A además garantiza el mantenimiento del orden correcto para la porción de aliasing y la otra porción, es decir, el segundo subbloque.
- 15 [0097] En la realización preferida, se aplica la transformación de coseno discreta modificada (MDCT), a fin de obtener la conmutación de muestreo crítico de una operación de codificación en un dominio a una operación de codificación en un dominio diferente. Sin embargo, pueden aplicarse también todas las otras transformaciones. Sin embargo, debido a que la MDCT es la realización preferida, se describirá en más detalle la MDCT con respecto a la Figura 4A y la Figura 4B.
- 20 [0098] La Figura 4A ilustra una ventana 70, que tiene una porción creciente a la izquierda y una porción decreciente a la derecha, donde puede dividirse esta ventana en cuatro porciones: a, b, c y d. La ventana 70 tiene, como puede observarse en la figura, solo porciones de aliasing en la situación de 50% superposición/adición ilustrada. Específicamente, la primera porción que tiene muestras de cero a N corresponde a las segundas porciones de una ventana precedente 69, y la segunda mitad, que se extiende entre la muestra N y la muestra 2N de la ventana 70, se superpone con la primera porción de la ventana 71, que se encuentra en la ventana de la realización ilustrada $i + 1$, mientras que la ventana 70 es la ventana i .
- 25 [0099] La operación de MDCT puede observarse como la cascada de la operación de plegado y una operación de transformación posterior, y específicamente, una operación de DCT posterior, donde se aplica la DCT de tipo IV (DCT-IV). De manera específica, la operación de plegado se obtiene mediante el cálculo de la primera porción N/2 del bloque de plegado como $-C_R-d$, y el cálculo de la segunda porción de muestras N/2 de la salida de plegado como $a-b_R$, donde R es el operador inverso. En consecuencia, la operación de plegado logra valores de salida N, mientras que se reciben valores de entrada 2N.
- 30 [0100] En la Figura 4A se ilustra también una operación de desplegado correspondiente del lado del decodificador, en forma de ecuación.
- 35 [0101] En general, una operación de MDCT en (a, b, c, d) logra exactamente los mismos valores de salida que la DCT-IV de $(-C_R-d, a-b_R)$, como se indica en la Figura 4A.
- [0102] En forma correspondiente, y usando la operación de desplegado, una operación de IMDCT (sigla en inglés de "transformación de coseno separada modificada inversa") logra la salida de la operación de desplegado aplicada a la salida de una transformación inversa de DCT-IV.
- 40 [0103] Por lo tanto, se introduce aliasing de tiempo efectuando una operación de plegado del lado del decodificador. Entonces, el resultado de la operación de plegado se transforma en el dominio de frecuencia usando una transformación de bloque de DCT-IV que requiere valores de entrada N.
- 45 [0104] Del lado del decodificador, los valores de entrada N se transforman nuevamente en el dominio de tiempo usando una operación de DCT-IV⁻¹, y la salida de esta operación de transformación inversa es entonces cambiada en una operación de desplegado, a fin de obtener valores de salida 2N que, sin embargo, son valores de salida de aliasing.
- [0105] A fin de eliminar el aliasing que se ha introducido por la operación de plegado y que aún está allí luego de la operación de desplegado, se necesita la operación de superposición/adición por el cancelador de aliasing de dominio de tiempo 53 de la Figura 3A.
- 50 [0106] Por lo tanto, cuando el resultado de la operación de desplegado se agrega al resultado de IMDCT previo en la mitad de superposición, los términos invertidos se cancelan en la ecuación en la base de la Figura 4A, y se obtiene simplemente, por ejemplo, b y d, para recuperar de ese modo la información original.
- 55 [0107] A fin de obtener una TDAC para la MDCT de ventana, existe un requisito, que se conoce como condición de "Princen-Bradley", lo que significa que los coeficientes de ventana son elevados a ² para las muestras correspondientes que se combinan en el cancelador de aliasing de dominio de tiempo para lograr la unidad (1) para cada muestra.

- [0108]** Si bien la Figura 4A ilustra la secuencia de ventana aplicada, por ejemplo, en la AAC–MDCT para ventanas largas o ventanas cortas, la Figura 4D ilustra una función de ventana diferente que tiene, además de las porciones de aliasing, también una porción sin aliasing.
- 5 **[0109]** La Figura 4D ilustra una función de ventana de análisis 72 que tiene una porción cero a_1 y d_2 , que tiene una porción de aliasing 72a, 72b, y que tiene una porción sin aliasing 72c.
- 10 **[0110]** La porción de aliasing 72b que se extiende sobre c_2 , d_1 tiene una porción de aliasing correspondiente de una ventana siguiente 73, que es indicada en 73b. En forma correspondiente, la ventana 73 además comprende una porción sin aliasing 73a. La Figura 4B, cuando se compara con la Figura 4A, torna evidente que, debido a que hay porciones cero a_1 , d_1 , para la ventana 72, o c_1 , para la ventana 73, ambas ventanas reciben una porción sin aliasing, y la función de ventana en la porción de aliasing es más empujada que en la Figura 4A. En vista de lo anterior, la porción de aliasing 72a corresponde a K_k , la porción sin aliasing 72c corresponde a la porción M_k , y la porción de aliasing 72b corresponde a R_k en la Figura 4B.
- 15 **[0111]** Cuando se aplica la operación de plegado a un bloque de muestras de ventanas por la ventana 72, se obtiene una situación como se ilustra en la Figura 4B. La porción izquierda que se extiende sobre las primeras muestras $N/4$ tiene aliasing. La segunda porción que se extiende sobre muestras $N/2$ está libre de aliasing, ya que la operación de plegado se aplica en porciones de ventana que tienen valores cero, y las últimas muestras $N/4$ son, nuevamente, afectadas por aliasing. Debido a la operación de plegado, la cantidad de valores de salida de la operación de plegado es igual a N , mientras que la entrada fue $2N$, si bien, de hecho, los valores $N/2$ en esta realización se establecieron en cero debido a la operación de ventana usando la ventana 72.
- 20 **[0112]** Ahora, se aplica la DCT IV al resultado de la operación de plegado, si bien, es importante el hecho de que la porción de aliasing 72 que se encuentra en la transición de un modo de codificación al otro modo de codificación es procesada en forma diferente de la porción sin aliasing, si bien ambas porciones pertenecen al mismo bloque de muestras de audio, y lo que es más importante, son ingresadas en la misma operación de transformación de bloque efectuada por el transformador 30 en la Figura 1A.
- 25 **[0113]** La Figura 4B además ilustra una secuencia de ventana de las ventanas 72; 73; 74, donde la ventana 73 es una ventana de transición de una situación donde sí existen porciones sin aliasing, a una situación donde solo existen porciones de aliasing. Se obtiene por medio de la forma asimétrica la función de ventana. La porción derecha de la ventana 73 es similar a la porción derecha de las ventanas en la secuencia de ventana de la Figura 4A, si bien la porción izquierda tiene una porción sin aliasing y la correspondiente porción cero (en c_1). Por lo tanto, la Figura 4B
- 30 **[0113]** ilustra una transición de MDCT–TCX a AAC, cuando debe efectuarse AAC usando ventanas de superposición completa; o alternativamente, se ilustra una transición de AAC a MDCT–TCX, cuando la ventana 74 somete a la ventana un bloque de información TCX de una manera de completa superposición, que es la operación regular para MDCT–TCX por una parte, y MDCT–AAC por otra parte, cuando no hay razón para cambiar de un modo al otro modo.
- 35 **[0114]** Por lo tanto, la ventana 73 puede denominarse una "ventana de inicio" o una "ventana de detención", que tiene, además, la característica preferida de que la longitud de esta ventana es idéntica a la longitud de por lo menos una ventana vecina, de modo que se mantiene el cuadro de bloque general o el cuadro de marco, cuando se establece un bloque de manera de tener la misma cantidad de coeficientes de ventana, es decir, muestras $2n$ en el ejemplo de la Figura 4D o la Figura 4A.
- 40 **[0115]** A continuación, se describe el procedimiento AAC–MDCT del lado del codificador y del lado del decodificador, con respecto a la Figura 5.
- 45 **[0116]** En una operación de ventana 80, se aplica una función de ventana ilustrada en 81. La función de ventana tiene dos porciones de aliasing L_k y R_k , y una porción sin aliasing M_k . Por lo tanto, la función de ventana 81 es similar a la función de ventana 72 e la Figura 4B. La aplicación de esta función de ventana a una pluralidad correspondiente de muestras de audio logra el bloque ventana de muestras de audio que tiene un subbloque de aliasing correspondiente a R_k/L_k y un subbloque sin aliasing correspondiente a M_k .
- [0117]** La operación de plegado ilustrada por 82 se efectúa como se indica en la Figura 4B, y logra salidas N , lo que significa que las porciones L_k , R_k se reducen de manera de tener una cantidad menor de muestras.
- 50 **[0118]** Entonces, se realiza una DCT IV 83, como se describe en relación con la ecuación de MDCT en la Figura 4A. La salida de MDCT es adicionalmente procesada por cualquier compresor de información disponible, tal como un cuantificador 84 o cualquier otro dispositivo que realice cualquiera de las herramientas AAC bien conocidas.
- 55 **[0119]** Del lado del decodificador, se realiza un procesamiento inverso 85. Entonces, se realiza una transformación del tercer dominio al primer dominio por medio de la DCT^{-1} IV 86. A continuación, se realiza una operación de desplegado 87, como se describe en relación con la Figura 4A. Luego, en un bloque 88, se realiza una operación de ventana de síntesis, y los ítems 89a y 89b juntos efectúan una cancelación de aliasing de dominio de tiempo. El ítem 89b es un dispositivo de demora que aplica una demora de muestras $M_k + R_k$ a fin de obtener la superposición como se describe en relación con la Figura 4A, y el añadidor 89a realiza una combinación de la porción actual de las

muestras de audio, tal como la primera porción L_k de una salida de ventana actual, y la última porción R_{k-1} de la ventana previa. Esto logra, como se indica en 90, las porciones libres de aliasing L_k y M_k . Debe observarse que M_k estaba libre de aliasing desde el comienzo, pero el procesamiento por los dispositivos 89a, 89b, ha cancelado el aliasing en la porción de aliasing L_k .

- 5 **[0120]** En la realización preferida, la AAC–MDCT además puede aplicarse con ventanas que solo tienen porciones de aliasing como se indica en la Figura 4A, si bien, para una conmutación entre un modo de codificación y el otro modo de codificación, se prefiere la aplicación de una ventana de AAC tenga una porción de aliasing y que tenga una porción sin aliasing.
- 10 **[0121]** Se usa una realización de la presente invención en una codificación de audio cambiada que cambia entre AAC y AMR–WB+ [4].
- [0122]** AAC utiliza una MDCT como se describe en la Figura 5. AAC es muy adecuada para la señal de música. La codificación cambiada utiliza AAC cuando la señal de entrada es detectada en un procesamiento previo como música, o rotulada como música por el usuario.
- 15 **[0123]** El marco de señal de entrada k es sometido a ventana por una ventana de tres partes de tamaños L_k , M_k y R_k . La MDCT introduce componentes de aliasing de dominio de tiempo antes de la transformación de la señal en dominio de frecuencia, donde se realiza la cuantificación. Luego de añadir la señal de ventana previa superpuesta de tamaño $R_{k-1} = L_k$, las primeras muestras $L_k + M_k$ de marco de señal original podrían recuperarse si se introdujera algún error de cuantificación. El aliasing de dominio de tiempo se cancela.
- 20 **[0124]** A continuación, se describe el procedimiento de TCX–MDCT con respecto a la presente invención, en relación con la Figura 6.
- [0125]** En contraste al codificador de la Figura 5, una transformación en el segundo dominio es efectuada por el ítem 92. El ítem 92 es un transformador de LPC que o bien genera una señal residual de LPC, o una señal cargada que puede calcularse cargando una señal residual de LPC usando un filtro de carga, como se conoce del procesamiento de TCX. Naturalmente, la señal de TCX además puede calcularse con un solo filtro, mediante el filtrado de la señal de dominio de tiempo a fin de obtener la señal de TCX, que es una señal en el dominio de LPC, o en términos generales, en el segundo dominio. Por lo tanto, el conversor de primer dominio/segundo dominio 92 proporciona, en su sitio de salida, la entrada de señal en el dispositivo de ventana 80. Aparte del transformador 92, el procedimiento en el codificador de la Figura 6 es similar al procedimiento en el codificador de la Figura 5. Naturalmente, pueden aplicarse diferentes algoritmos de compresión de información en los bloques 84 de la Figura 5 y la Figura 6, que son fácilmente evidentes, cuando se comparan las herramientas de codificación de AAC con las herramientas de codificación de TCX.
- 25 **[0126]** Del lado del decodificador, se realizan las mismas etapas que se describen en relación con la Figura 5, si bien estas etapas no se efectúan en una señal codificada en el dominio de frecuencia directo (tercer dominio), sino que se realizan en una señal codificada generada en el cuarto dominio, es decir, el dominio de frecuencia de LPC.
- 30 **[0127]** Por lo tanto, un procedimiento de adición de superposición es realizado por los dispositivos 89a, 89b, en la Figura 6, en el segundo dominio en lugar del primer dominio, como se ilustra en la Figura 5.
- [0128]** AMR–WB+ se sustenta en una ACELP de codificación de discurso y una TCX de codificación sobre la base de la transformación. Para cada supermarco de 1024 muestras, AMR–WB+ selecciona, con la decisión de bucle cerrado, entre 17 combinaciones diferentes de TCX y ACELP, la mejor, de acuerdo con la decisión cerrada usando la evaluación objetiva de SegSNR. La AMR–WB+ es muy adecuada para señales de discurso y discurso sobre música. La DFT (sigla en inglés de "transformación Fourier discreta") original de la TCX fue reemplazada por una MDCT, a fin de disfrutar sus grandes propiedades. La TCX de AMR–WB+ entonces es equivalente a la codificación de MPTC, excepto la cuantificación, que se mantuvo como tal. La AMR–WB+ modificada es utilizada por el codificador de audio cambiado cuando se detecta la señal de entrada, o cuando esta es rotulada como discurso o discurso sobre música.
- 35 **[0129]** La TCX–MDCT efectúa una MDCT no directamente sobre el dominio de señal, sino luego del filtrado de la señal por un filtro de análisis $W(z)$, sobre la base de un coeficiente de LPC. El filtro se denomina filtro de análisis de carga, y permite que la TCX, al mismo tiempo, blanquee la señal y de forma al ruido de cuantificación por medio de una curva sobre la base del formante, lo que se encuentra en línea con las teorías psicoacústicas.
- 40 **[0130]** El procesamiento ilustrado en la Figura 5 se realiza para un modo de AAC–MDCT directo, sin ninguna conmutación al modo de TCX o cualquier otro modo usando las ventanas de superposición completa en la Figura 4A. Sin embargo, cuando se detecta una transición, se aplica una ventana específica, que es una ventana de inicio AAC para una transición al otro modo de codificación, o una ventana de detención de AAC para la transición desde el otro modo de codificación al modo AAC, como se ilustra en la Figura 7. Una ventana de detención AAC 93 tiene una porción de aliasing ilustrada en 93b y una porción sin aliasing ilustrada en 93a, es decir, indicada en la figura como la parte horizontal de la ventana 93. De manera correspondiente, la ventana de detención AAC 94 es ilustrada de modo de poseer una porción de aliasing 94b y una porción sin aliasing 94a. En la porción AMR–WB+, se aplica una
- 45 **[0129]** La TCX–MDCT efectúa una MDCT no directamente sobre el dominio de señal, sino luego del filtrado de la señal por un filtro de análisis $W(z)$, sobre la base de un coeficiente de LPC. El filtro se denomina filtro de análisis de carga, y permite que la TCX, al mismo tiempo, blanquee la señal y de forma al ruido de cuantificación por medio de una curva sobre la base del formante, lo que se encuentra en línea con las teorías psicoacústicas.
- 50 **[0130]** El procesamiento ilustrado en la Figura 5 se realiza para un modo de AAC–MDCT directo, sin ninguna conmutación al modo de TCX o cualquier otro modo usando las ventanas de superposición completa en la Figura 4A. Sin embargo, cuando se detecta una transición, se aplica una ventana específica, que es una ventana de inicio AAC para una transición al otro modo de codificación, o una ventana de detención de AAC para la transición desde el otro modo de codificación al modo AAC, como se ilustra en la Figura 7. Una ventana de detención AAC 93 tiene una porción de aliasing ilustrada en 93b y una porción sin aliasing ilustrada en 93a, es decir, indicada en la figura como la parte horizontal de la ventana 93. De manera correspondiente, la ventana de detención AAC 94 es ilustrada de modo de poseer una porción de aliasing 94b y una porción sin aliasing 94a. En la porción AMR–WB+, se aplica una
- 55 **[0129]** La TCX–MDCT efectúa una MDCT no directamente sobre el dominio de señal, sino luego del filtrado de la señal por un filtro de análisis $W(z)$, sobre la base de un coeficiente de LPC. El filtro se denomina filtro de análisis de carga, y permite que la TCX, al mismo tiempo, blanquee la señal y de forma al ruido de cuantificación por medio de una curva sobre la base del formante, lo que se encuentra en línea con las teorías psicoacústicas.

- 5 ventana similar a la ventana 72 de la Figura 4B, donde esta ventana tiene una porción de aliasing 72a y una porción sin aliasing 72c. Si bien solo puede observarse una ventana individual AMR-WB+ como una ventana de inicio/detención como se ilustra en la Figura 7, puede haber una pluralidad de ventanas, que, preferentemente, tienen una superposición del 50%, y por lo tanto, pueden ser similares a las ventanas de la Figura 4A. Habitualmente, TCX en AMR-WB+ no utiliza ninguna superposición del 50%. Solo se adopta una pequeña superposición para poder cambiar rápidamente hacia y desde ACELP, que emplea una ventana inherentemente rectangular, es decir, 0% de superposición.
- 10 **[0131]** Sin embargo, cuando tiene lugar la transición, se aplica una ventana de inicio AMR-WB+ ilustrada en la posición central izquierda en la Figura 7, y cuando se decide que debe efectuarse la transición desde AMR-WB+ hacia AAC, se aplica una ventana de detención AMR-WB+. La ventana de inicio tiene una porción de aliasing a la izquierda, y la ventana de detención tiene una porción de aliasing hacia la derecha, donde estas porciones de aliasing se indican como 72a, y donde estas porciones de aliasing corresponden a las porciones de aliasing de las ventanas de inicio/detención de AAC vecinas indicadas en 93b o 94b.
- 15 **[0132]** El procesamiento específico se produce en las dos regiones superpuestas de 128 muestras de la Figura 7. Para cancelar el aliasing de dominio de tiempo de AAC, el primer y el último marco del segmento AMR-WB+ son obligados a ser TCX y no ACELP. Esto se efectúa inclinando el puntaje SegSNR en la decisión de bucle cerrado. Además, las primeras 128 muestras de la TCX-MDCT se procesan en forma específica como se ilustra en la Figura 8A, donde $L_k = 128$.
- 20 **[0133]** Las últimas 128 muestras de AMR-WB+ son procesadas como se ilustra en la Figura 8B, donde $R_k = 128$.
- 25 **[0134]** La Figura 8A ilustra el procesamiento para la porción de aliasing R_k a la derecha de la porción sin aliasing para una transición de TCX a AAC, y la Figura 8B ilustra el procesamiento específico de la porción de aliasing L_k a la izquierda de una porción sin aliasing para una transición de AAC a TCX. El procesamiento es similar con respecto a la Figura 6, si bien la operación de carga, es decir, la transformación del primer dominio al segundo dominio, se posiciona en forma diferente. Específicamente, en la Figura 6, la transformación se realiza antes de la ventana, mientras que en la Figura 8B, la transformación 92 se realiza después de la ventana 80 (y el plegado 82), es decir, la operación de introducción de aliasing de dominio de tiempo indicada por "TDA" (sigla en inglés de "aliasing de dominio de tiempo").
- 30 **[0135]** Del lado del decodificador, nuevamente, se realizan etapas de procesamiento bastante similares a la Figura 6, si bien, una vez más, la posición de la carga inversa para la porción de aliasing es anterior a la ventana 88 (y anterior al desplegado 87), y posterior a la transformación del primer dominio al segundo dominio indicado por 86 en la Figura 8A.
- 35 **[0136]** Por lo tanto, de acuerdo con una realización preferida de la presente invención, la porción de aliasing de una ventana de transición para TCX es procesada como se indica en la Figura 1A o la Figura 1B, y una porción sin aliasing para la misma ventana es procesada de acuerdo con la Figura 6.
- 40 **[0137]** El procesamiento para cualquier ventana AAC-MDCT permanece igual, excepto el hecho de que se selecciona una ventana de inicio o una ventana de detención en la transición. Sin embargo, en otras realizaciones, el procesamiento de TCX puede permanecer igual, y la porción de aliasing de la ventana de AAC-MDCT es procesada en forma diferente, en comparación con la porción sin aliasing.
- 45 **[0138]** Además, ambas porciones de aliasing de ambas ventanas, es decir, una ventana AAC o una ventana TCX, pueden procesarse en forma diferente de sus porciones sin aliasing, según sea el caso. En la realización preferida, sin embargo, se prefiere que el procesamiento de AAC sea efectuado como tal, debido a que ya se encuentra en el dominio de señal posterior al procedimiento de superposición-adición, como es evidente de la Figura 5, y que la ventana de transición de TCX sea procesada como se ilustra en el contexto de la Figura 6, para una porción sin aliasing, y como se ilustra en la Figura 8A u 8B, para la porción de aliasing.
- 50 **[0139]** A continuación, se describirá la Figura 9A, en la cual el procesador 12 de la Figura 1A se ha indicado como un controlador 98.
- [0140]** Los dispositivos en la Figura 9A que tienen correspondientes números de referencia que corresponden a los ítems de la Figura 11A tienen funcionalidad similar, y no se describen nuevamente.
- 55 **[0141]** En forma específica, el controlador 98 ilustrado en la Figura 9A opera como se indica en la Figura 9B. En la etapa 98a, se detecta una transición, donde esta transición es indicada por la etapa de decisión 300. A continuación, el controlador 98 está activo para inclinar la conmutación 521, de modo que la conmutación 521 selecciona la alternativa (2b) en cualquier caso.
- [0142]** Entonces, la etapa 98b es efectuada por el controlador 98. Específicamente, el controlador opera para tomar la información en la porción de aliasing, y no alimentar la información en la LPC 510 directamente, sino alimentar la información antes del filtro de LPC 510 directamente, sin la carga de un filtro LPC, en el bloque de TDA 527a. A continuación, esta información es tomada por el controlador 98 y cargada, y luego, alimentada en el bloque DCT

- 527b, es decir, luego de haber sido cargada por el filtro de carga en la salida del controlador 98. El filtro de carga en el controlador 98 utiliza los coeficientes de LPC calculados en el bloque de LPC 510 luego de un análisis de señal. El bloque de LPC es capaz de alimentar o bien ACELP o TCX y además efectuar un análisis de LPC para obtener los coeficientes de LPC. La porción de DCT 527b del dispositivo de MDCT consiste en el dispositivo de TDA 527a y el dispositivo DCT 527b. El filtro de carga en la salida del controlador 98 tiene la misma característica que el filtro en el bloque de LPC 510 y un filtro de carga adicional potencialmente presente, tal como el filtro de percepción en el procesamiento de TCX de AMR-WB+. En consecuencia, en la etapa 98b, se realizan los procesamientos de TDA, LPC y DCT, en este orden.
- [0143]** La información en la porción adicional es alimentada al bloque de LPC 510, y a continuación, en el bloque de MDCT 527a, 527b, como es indicado por el recorrido de señal normal en la Figura 9A. En este caso, el filtro de carga de TCX no está ilustrado explícitamente en la Figura 9A, debido a que pertenece al bloque de LPC 510.
- [0144]** Como se establece con anterioridad, la información en la porción de aliasing es, como se indica en la Figura 8A, sometida a ventana en el bloque 527a, y la información de ventana generada dentro del bloque 527 es filtrada por LPC en la salida del controlador, y el resultado del filtrado de LPC es luego aplicado a la porción de transformación 527b del bloque de MDCT 527. El filtro de carga de TCX para la carga de la señal residual de LPC generada por el dispositivo de LPC 510 no se ilustra en la Figura 9A. Además, el dispositivo 527a incluye la etapa de ventana 80 y la etapa de plegado 82, y el dispositivo 527b incluye la etapa de DCT IV 83 como se describe en relación con la Figura 8A. La etapa de DCT IV 83/527b entonces recibe la porción de aliasing después del procesamiento y la porción adicional luego del procesamiento correspondiente, y realiza la operación de MDCT común, y se efectúa una compresión de información posterior en el bloque 528, como se indica en la etapa 98d de la Figura 9B. Por lo tanto, en el caso de un codificador con soporte cableado o controlado por programas informáticos, como se describe en relación con la Figura 9A, el controlador 98 realiza la programación de información como se indica en la Figura 9D, entre los diferentes bloques 510 y 527a, 527b.
- [0145]** Del lado del decodificador, se provee un controlador de transición 99, además de los bloques indicados en la Figura 11B, que ya se han descrito.
- [0146]** La funcionalidad del controlador de transición 99 se describe en relación con la Figura 10B.
- [0147]** Tan pronto como el controlador de transición 98 ha detectado una transición como se detalla en la etapa 99a de la Figura 10B, el marco entero es alimentado a la etapa de MDCT⁻¹ 537b, luego de una descompresión de información en el descompresor de información 537a. Este procedimiento se indica en la etapa 99b de la Figura 10B. A continuación, como se indica en la etapa 99c, la porción de aliasing es alimentada directamente a la etapa de LPC⁻¹ antes de efectuar un procesamiento de TDAC. Sin embargo, la porción de aliasing no se somete a un procesamiento de "MDCT" completo, sino que, como se ilustra en la Figura 8B, solo se somete a la transformación inversa del cuarto dominio al segundo dominio.
- [0148]** La alimentación de la porción de aliasing luego de la etapa 86/etapa 537b de DCT⁻¹ IV de la Figura 8B en la etapa de LPC⁻¹ adicional 537d de la Figura 10A asegura la realización de una transformación del segundo dominio al primer dominio, y luego, la operación de desplegado 87 y la operación de ventana 88 de la Figura 8B se realizan en el bloque 537c. Por lo tanto, el controlador de transición 99 recibe la información del bloque 537b luego de la operación de DCT⁻¹ de la etapa 86, y luego alimenta esta información al bloque 537d de LPC⁻¹. La salida de este procedimiento luego se alimenta al bloque 537d a fin de efectuar el desplegado 87 y la ventana 88. A continuación, el resultado de la ventana de la porción de aliasing es transmitido al bloque de TDAC 440b, a fin de efectuar una operación de superposición-adición con la correspondiente porción de aliasing de un bloque de AAC-MDCT. En vista de lo anterior, el orden del procesamiento para el bloque de aliasing es: descompresión de información en 537a; DCT⁻¹ en 537b; carga de percepción de LPC inversa y TCX inversa (lo que juntas significan la carga inversa) en 537d; procesamiento TDA⁻¹ en 537c; y luego, superposición y adición en 440b.
- [0149]** No obstante, la porción restante del marco es alimentada a la etapa de ventana antes de TDAC y el filtrado/carga inversa en 540, como se describe en relación con la Figura 6 y como se ilustra con el flujo de señal normal en la Figura 10A, cuando las flechas conectadas al bloque 99 son ignoradas.
- [0150]** En vista de lo anterior, la etapa 99c logra la señal de audio decodificada para la porción de aliasing luego de la TDAC 440b, y la etapa 99d logra la señal de audio decodificada para la porción restante/adicional luego de la TDAC 537c en el dominio de LPC y la carga inversa en el bloque 540.
- [0151]** De acuerdo con ciertos requisitos de implementación, las realizaciones de la invención pueden llevarse a cabo en soporte físico o en soporte lógico. La implementación puede efectuarse usando un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD (disco versátil digital), un CD (sigla en inglés para "disco compacto"), una memoria ROM (sigla en inglés para "memoria de solo lectura"), una memoria PROM (sigla en inglés para "memoria de solo lectura programable"), una memoria EPROM (sigla en inglés de "memoria de solo lectura programable borrrable"), una memoria EEPROM (sigla en inglés de "memoria de solo lectura programable borrrable eléctricamente"), o una memoria FLASH (EEPROM para múltiples posiciones de memoria), que tiene

señales de control para lectura electrónica allí almacenadas, que cooperan (o son capaces de cooperar) con un sistema informático programable de modo tal de efectuar el procedimiento respectivo.

5 [0152] Algunas realizaciones de acuerdo con la invención comprenden un portador de información que tiene señales de control para lectura electrónica, que son capaces de cooperar con un sistema informático programable, de modo de efectuar uno de los procedimientos descritos en esta solicitud.

10 [0153] En general, las realizaciones de la presente invención pueden implementarse como un producto de programa para computadora, o programa informático, con un código de programa, donde el código de programa es operativo para efectuar uno de los procedimientos cuando el producto de programa de computadora es ejecutado en una computadora. El código de programa, por ejemplo, puede ser almacenado en un portador para lectura en una máquina.

[0154] Otros ejemplos comprenden el programa de computadora para la realización de uno de los procedimientos descritos en esta solicitud, almacenado en un portador para lectura en una máquina.

15 [0155] En otras palabras, una realización del procedimiento de la invención, por lo tanto, es un programa para computadora que tiene un código de programa para la realización de uno de los procedimientos descritos en esta solicitud, cuando el programa para computadora se ejecuta en una computadora.

[0156] Un ejemplo adicional de los procedimientos de la invención, por lo tanto, es un portador de información (o un medio de almacenamiento digital, o un medio para lectura en computadora) que comprende, allí registrado, el programa de computadora para la realización de uno de los procedimientos descritos en esta solicitud.

20 [0157] Un ejemplo adicional del procedimiento de la invención, por lo tanto, es una corriente de información o una secuencia de señales que representan el programa de computadora para la realización de uno de los procedimientos descritos en esta solicitud. La corriente de información o la secuencia de señales, por ejemplo, puede estar configurada para ser transferida por medio de una conexión de comunicación de información, por ejemplo, por medio de Internet.

25 [0158] Un ejemplo adicional comprende un medio de procesamiento, por ejemplo, una computadora, o un dispositivo lógico programable, configurado o adaptado para efectuar uno de los procedimientos descritos en esta solicitud.

[0159] Un ejemplo adicional comprende una computadora que tiene allí instalado el programa de computadora para la realización de uno de los procedimientos descritos en esta solicitud.

30 [0160] En algunos ejemplos, puede usarse un dispositivo lógico programable (por ejemplo, una serie de puertas programables en campo) a fin de efectuar algunas o la totalidad de las funcionalidades de los procedimientos descritos en esta solicitud. En algunas realizaciones, una serie de puertas programables en campo puede cooperar con un microprocesador a fin de efectuar uno de los procedimientos que se describen en esta solicitud.

35 [0161] Las realizaciones descritas con anterioridad son sólo ilustrativas de los principios de la presente invención. Se entiende que serán evidentes para los expertos en el arte modificaciones y variaciones de las disposiciones y los detalles que se describen en esta solicitud. Por lo tanto, se tiene el propósito de limitarse solo al alcance de las reivindicaciones de patente inminentes, y no a los detalles específicos presentados a modo de descripción y explicación de las realizaciones de la presente invención.

REIVINDICACIONES

1. Un aparato para la codificación de una señal de audio, que comprende:
- una función de ventana (11), para someter a ventana un primer bloque de la señal de audio usando una ventana de análisis, donde la ventana de análisis tiene una porción de aliasing (L_k , R_k) y una porción adicional (M_k);
- 5 un procesador (12), para el procesamiento de un primer subbloque (20) de la señal de audio asociado con la porción de aliasing, mediante la transformación del primer subbloque en un dominio diferente del dominio, después de la ventana del primer subbloque para obtener un primer subbloque procesado; y para el procesamiento de un segundo subbloque (21) de la señal de audio asociado con la otra porción, mediante la transformación del segundo subbloque en el dominio diferente, antes de someter a ventana el segundo subbloque a fin de obtener un segundo subbloque procesado; y
- 10 un transformador (13), para la conversión del primer subbloque procesado y el segundo subbloque procesado del dominio diferente en un dominio adicional, usando el mismo principio de transformación de bloques, de manera de obtener un primer bloque convertido;
- 15 donde el aparato está configurado para el procesamiento adicional (14) del primer bloque convertido usando un algoritmo de compresión de información.
2. El aparato de acuerdo con la reivindicación 1, que está configurado para el procesamiento de un segundo bloque de la señal de audio que se superpone con el primer bloque, usando una segunda ventana de análisis (73) que tiene una porción de aliasing (73b) que corresponde a la porción de aliasing (72b) de la primera ventana de análisis.
- 20 3. El aparato de acuerdo con la reivindicación 1 o la reivindicación 2, donde el dominio en el cual está posicionada la señal de audio es un dominio de tiempo, donde el dominio diferente es un dominio de LPC, donde un tercer dominio, en el cual un segundo bloque de la señal de audio que se superpone con el primer bloque de la señal de audio es codificado, es un dominio de frecuencia, y donde el dominio adicional, donde el transformador (13) está configurado para la transformación, es un dominio de frecuencia de LPC; y
- 25 donde el procesador (12) comprende un filtro de LPC para la transformación del primer dominio al segundo dominio, o donde el transformador (13) comprende un algoritmo de conversión sobre la base de Fourier para la transformación de la información de entrada en un dominio de frecuencia de la información de entrada, tal como una DCT, una DST, una FFT o una DFT.
- 30 4. El aparato de acuerdo con una de las reivindicaciones precedentes, donde la función de ventana (11) comprende una función de plegado (82), para el plegado de los valores de entrada a fin de obtener valores de salida, donde la cantidad de valores de salida es menor que la cantidad de valores de entrada; donde la función de plegado es de modo tal que se introduce aliasing de tiempo en los valores de salida.
- 35 5. El aparato de acuerdo con cualquiera de las reivindicaciones precedentes, donde la función de ventana (11) es operativo de manera de efectuar la ventana, a fin de obtener los valores de entrada para una función de plegado efectuada posteriormente (82).
- 40 6. El aparato de acuerdo con una de las reivindicaciones precedentes, donde el aparato comprende una primera rama de codificación (400), para la codificación de la señal de audio en un dominio de frecuencia; y una segunda rama de codificación (500), para la codificación de la señal de audio sobre la base de un dominio de frecuencia diferente;
- 45 donde la segunda rama de codificación tiene una primera subrama (527; 528), para la codificación de la señal de audio en el otro dominio de frecuencia; y una segunda subrama (526), para la codificación de la señal de audio en el otro dominio; donde el aparato además comprende una etapa de decisión (300), para la decisión de si un bloque de información de audio es representado en una corriente de bits de salida por información generada usando la primera rama de codificación, o la primera subrama o la segunda subrama de la segunda rama de codificación; y
- 50 donde el controlador (98) está configurado para el control de la etapa de decisión (300) de manera de decidir a favor de la primera subrama, cuando debe efectuarse la transición de la primera rama de codificación a la segunda rama de codificación, o de la segunda rama de codificación a la primera rama de codificación.
7. El aparato de acuerdo con cualquiera de las reivindicaciones precedentes, donde la porción adicional comprende una porción sin aliasing (M_k) y una porción de aliasing adicional, o una porción de aliasing que se superpone con una porción de aliasing correspondiente de un bloque vecino de la señal de audio.
8. Un aparato para la decodificación de una señal de audio codificada que tiene un primer bloque codificado de información de audio, donde el bloque codificado tiene una porción de aliasing y una porción adicional, que comprende:

- un procesador (51) para el procesamiento de la porción de aliasing (L_k, R_k), mediante la transformación (86) de la porción de aliasing en un dominio objetivo, antes de efectuar una ventana de síntesis (88) a fin de obtener una porción de aliasing de ventana; y para efectuar una ventana de síntesis (88) de la porción adicional, antes de efectuar una transformación (98) en el dominio objetivo; y
- 5 un cancelador de aliasing de dominio de tiempo (53) para la combinación de la porción de aliasing de ventana y la porción de aliasing de ventana de un segundo bloque codificado de información de audio, luego de una transformación (91) de la porción de aliasing del primer bloque codificado de información de audio, en el dominio objetivo, a fin de obtener una señal de audio decodificada correspondiente a la porción de aliasing del primer bloque.
9. El aparato de acuerdo con la reivindicación 8,
- 10 donde el procesador (51) comprende un transformador (86) para la conversión de la porción de aliasing de un cuarto dominio, en un segundo dominio, y donde el procesador además comprende un transformador (91), para la conversión de la porción de aliasing representada en el segundo dominio, en el primer dominio, donde el transformador (86) es operativo de manera de efectuar un algoritmo de conversión de tiempo de frecuencia sobre la base de bloque.
- 15 10. El aparato de acuerdo con la reivindicación 8 ó 9, donde el procesador (12) es operativo de manera de efectuar una operación de desplegado (87), para la obtención de información de salida que tiene una cantidad de valores mayor que una cantidad de valores de entrada en la operación de desplegado (87).
11. El aparato de acuerdo con cualquiera de las reivindicaciones 8; 9 ó 10, donde el procesador (12) es operativo de manera de utilizar una función de ventana de síntesis (88) relacionada con una función de ventana de análisis utilizada cuando se genera la señal de audio codificada.
- 20 12. El aparato de acuerdo con cualquiera de las reivindicaciones 8–11, donde la señal de audio codificada comprende un indicador de modo de codificación que indica un modo de codificación para el primer bloque codificado y el segundo bloque codificado,
- 25 donde el aparato además comprende un controlador de transición (99), para el control del procesador (12), cuando el indicador de modo de codificación indica una conmutación de modo de codificación, de un primer modo de codificación a un segundo modo de codificación diferente, o viceversa, y para el control del procesador (12), a fin de efectuar la misma operación para un bloque de codificación completo, cuando no se señala una conmutación de modo de codificación entre dos bloques de codificación.
13. El aparato de acuerdo con cualquiera de las reivindicaciones 8–12,
- 30 donde un primer modo de codificación y un segundo modo de codificación comprenden una etapa de decodificación de entropía; una etapa de descuantificación; una etapa de conversión de tiempo de frecuencia que comprende una operación de desplegado; y una etapa de ventana de síntesis;
- 35 donde el cancelador de aliasing de dominio de tiempo (53) comprende un añadidor (89a), para la adición de correspondientes porciones de aliasing de bloques codificados obtenidos por la etapa de ventana de síntesis (88), donde las correspondientes porciones de aliasing se obtienen por medio de un procesamiento de superposición (89b) de la señal de audio; y
- 40 donde, en el primer modo de codificación, el cancelador de aliasing de dominio de tiempo (53) está configurado para la adición de porciones de bloques obtenidas por la ventana de síntesis, a fin de obtener, como una salida de la adición (89a), la señal decodificada en el dominio objetivo; y
- 45 donde, en el segundo modo de codificación, la salida de la adición (89a) es procesada por el procesador (12), a fin de efectuar una transformación (91) de la salida de la adición, al dominio objetivo.
14. Una señal de audio codificada que comprende un primer bloque codificado de una señal de audio y un segundo bloque codificado superpuesto de la señal de audio, donde el primer bloque codificado de la señal de audio comprende una porción de aliasing y una porción adicional; donde la porción de aliasing ha sido transformada de un primer dominio a un segundo dominio, luego de la ventana (80) de la porción de aliasing; y la porción adicional ha sido transformada del primer dominio en el segundo dominio, antes de la ventana (80) del segundo subbloque; donde el segundo subbloque ha sido transformado en un cuarto dominio usando el mismo principio de transformación de bloques; y
- 50 donde el segundo bloque codificado ha sido generado por medio de la ventana (80) de un bloque de superposición de muestras de audio, y mediante la transformación de un bloque ventana en un tercer dominio; donde el segundo bloque codificado tiene una porción de aliasing que corresponde a la porción de aliasing del primer bloque codificado de muestras de audio.
15. Un procedimiento para la codificación de una señal de audio, que comprende:

- la ventana (11) de un primer bloque de la señal de audio usando una ventana de análisis, donde la ventana de análisis tiene una porción de aliasing (L_k, R_k) y una porción adicional (M_k);
- 5 el procesamiento (12) de un primer subbloque (20) de la señal de audio asociado con la porción de aliasing, mediante la transformación del primer subbloque en un dominio diferente del dominio, luego de la ventana del primer subbloque, a fin de obtener un primer subbloque procesado;
- el procesamiento de un segundo subbloque (21) de la señal de audio asociado con la porción adicional, mediante la transformación del segundo subbloque, en el dominio diferente, antes de la ventana del segundo subbloque, a fin de obtener un segundo subbloque procesado;
- 10 la conversión (13) del primer subbloque procesado y el segundo subbloque procesado, del dominio diferente en un dominio adicional, usando el mismo principio de transformación de bloques, a fin de obtener un primer bloque convertido; y
- el procesamiento adicional (14) del primer bloque convertido, usando un algoritmo de compresión de información.
- 15 16. Un procedimiento de decodificación de una señal de audio codificada que tiene un primer bloque codificado de información de audio, donde el bloque codificado tiene una porción de aliasing y una porción adicional, que comprende:
- el procesamiento (51) de la porción de aliasing (L_k, R_k) mediante la transformación (86) de la porción de aliasing en un dominio objetivo, antes de efectuar una ventana de síntesis (88) a fin de obtener una porción de aliasing de ventana;
- 20 una ventana de síntesis (88) de la porción adicional, antes de efectuar una transformación (98), en el dominio objetivo; y
- 25 la combinación (53) de la porción de aliasing de ventana y la porción de aliasing de ventana de un segundo bloque codificado de información de audio, a fin de obtener una cancelación de aliasing de dominio de tiempo, luego de una transformación (91) de la porción de aliasing del primer bloque codificado de información de audio, en el dominio objetivo, a fin de obtener una señal de audio decodificada que corresponde a la porción de aliasing del primer bloque.
17. Un producto de programa para computadora que tiene un código de programa para la realización, cuando se ejecuta en una computadora, del procedimiento para la codificación de acuerdo con la reivindicación 15, o el procedimiento de decodificación de acuerdo con la reivindicación 16.

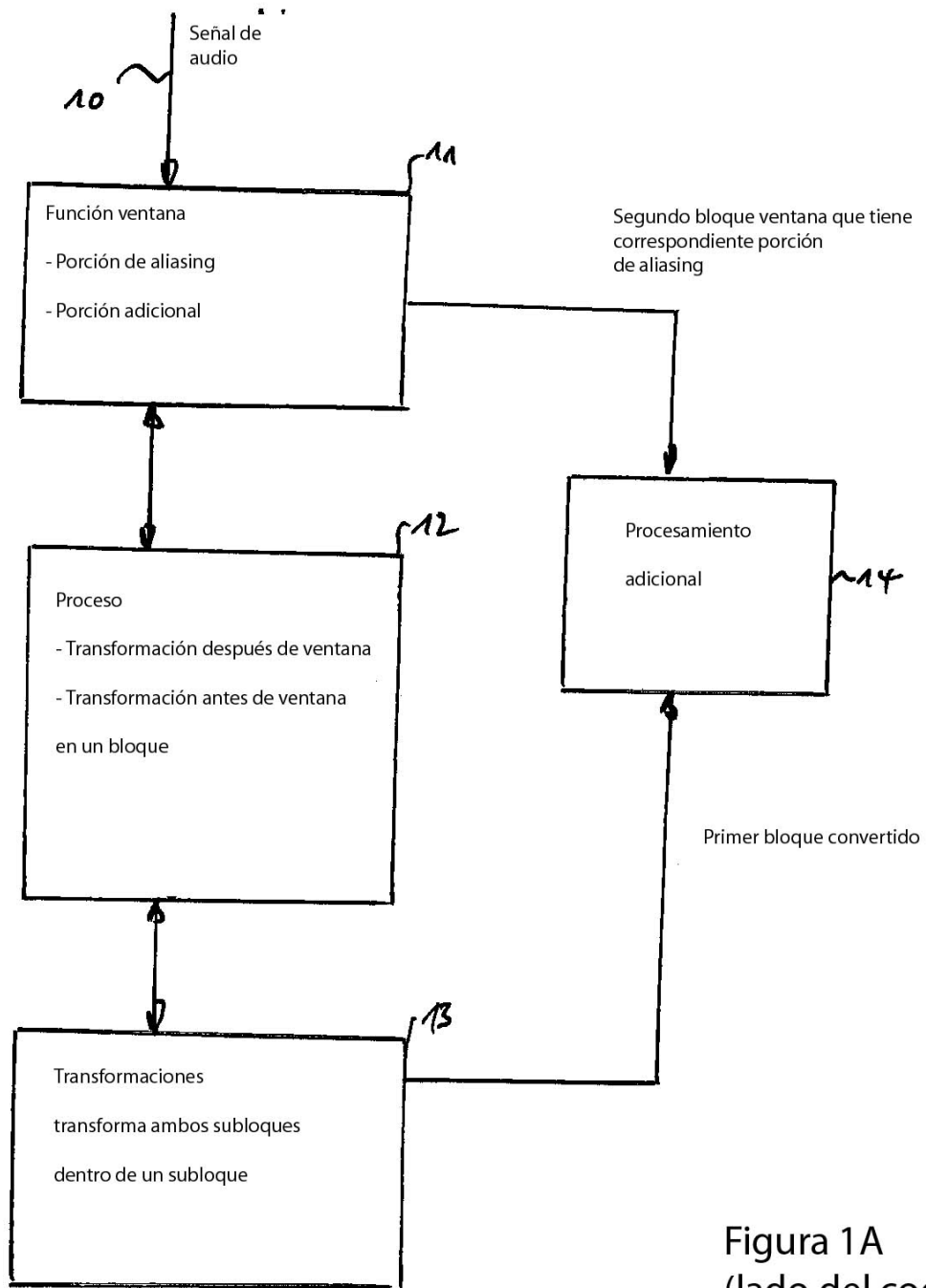


Figura 1A
(lado del codificador)

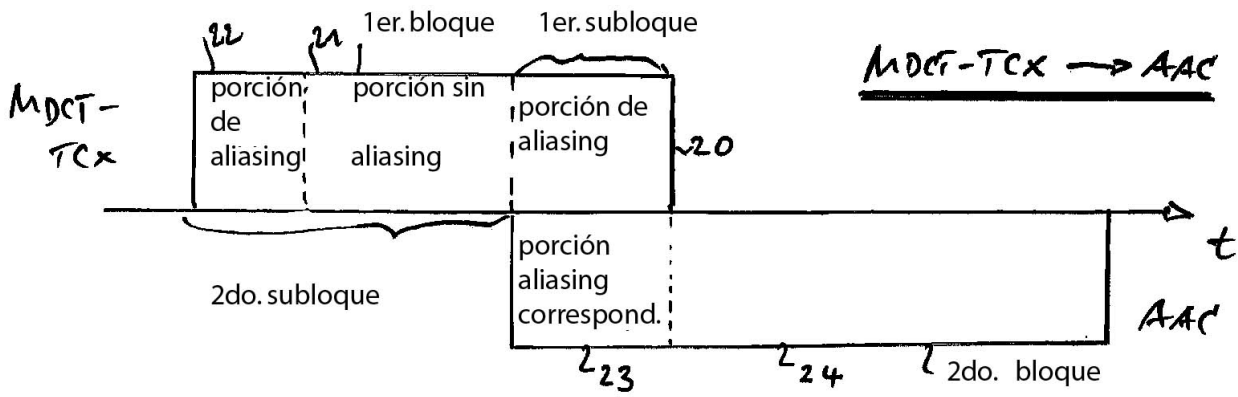


Fig. 1B

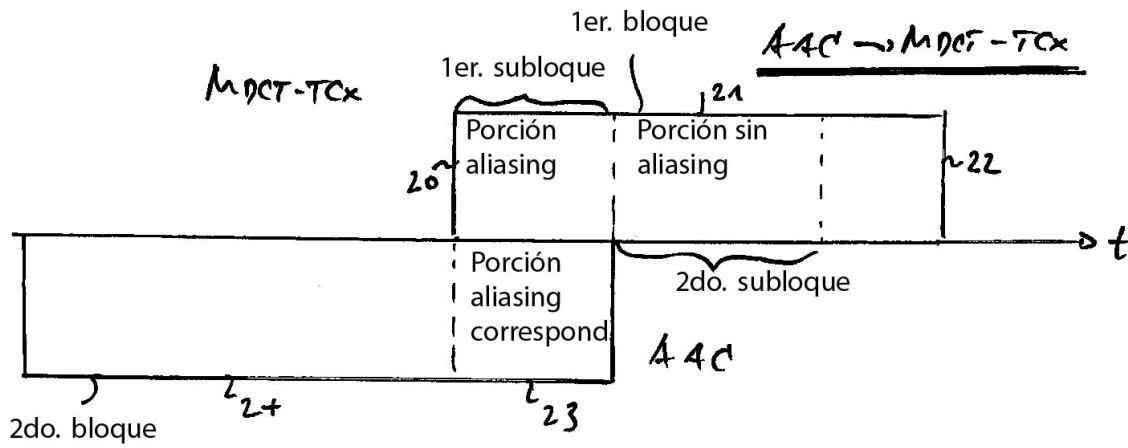


Fig. 1C

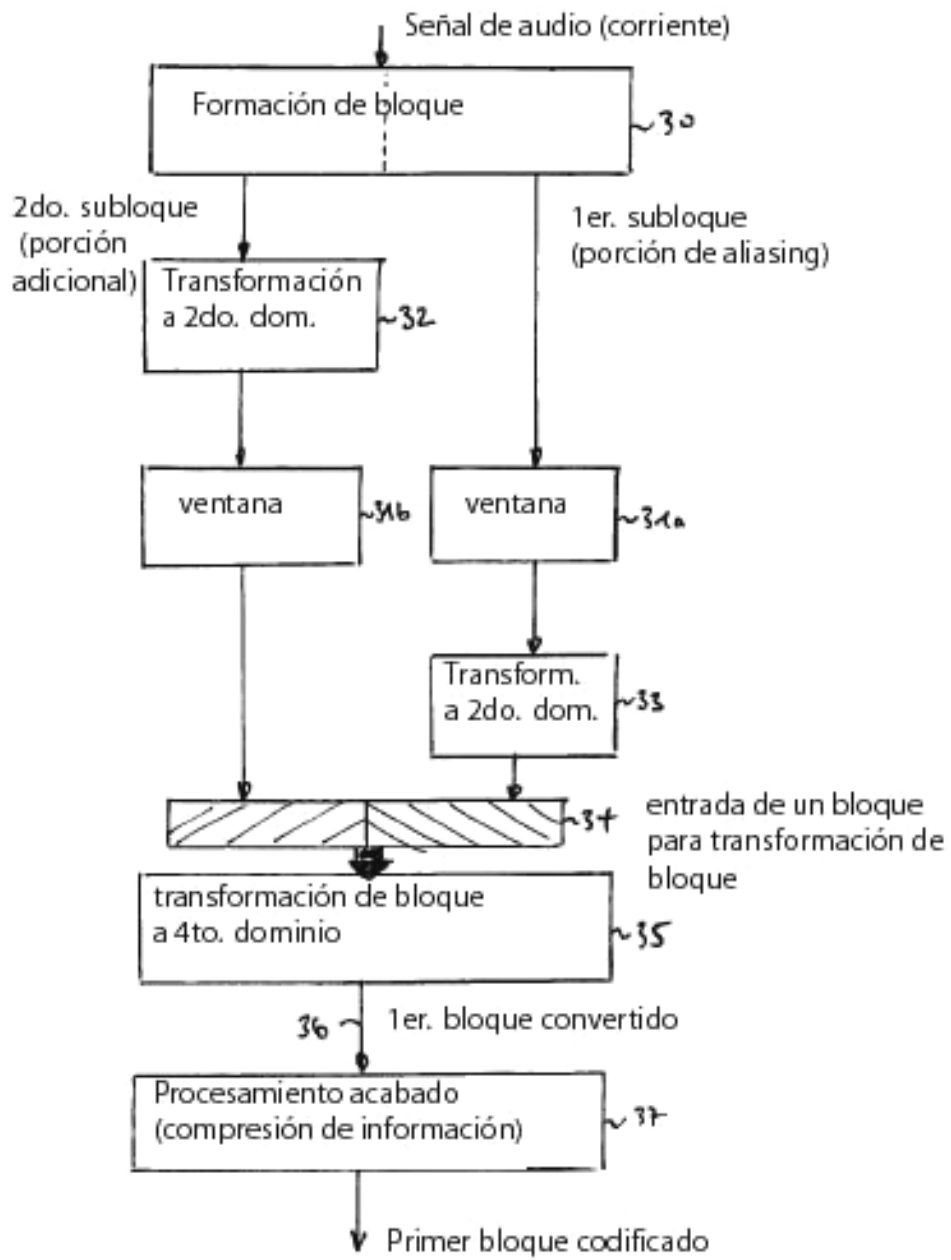


Fig. 10

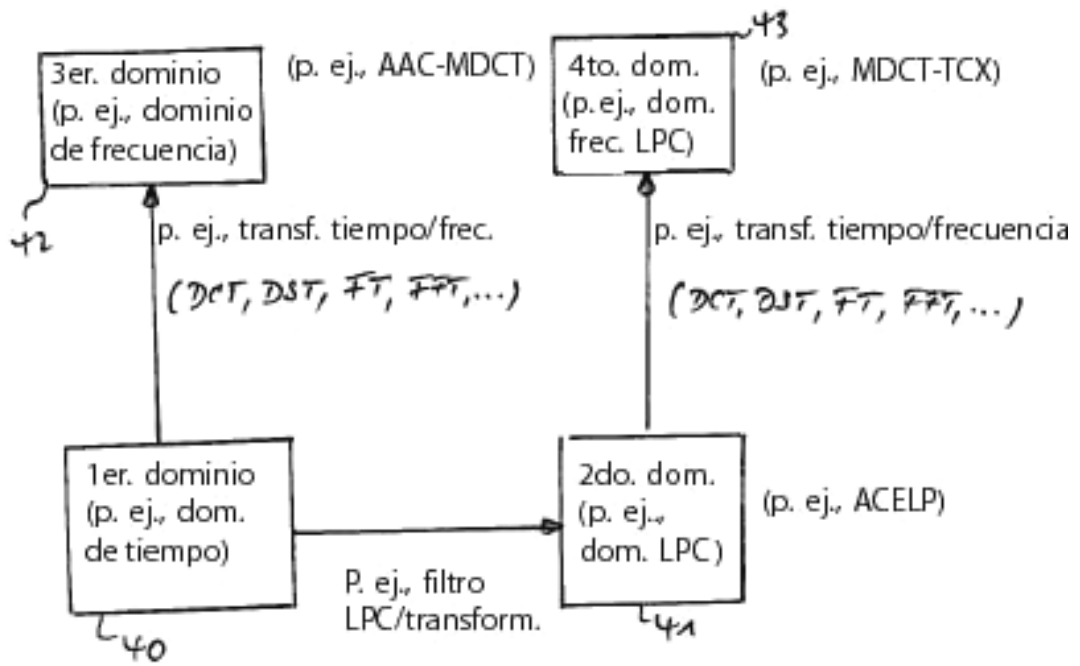


Fig. 2

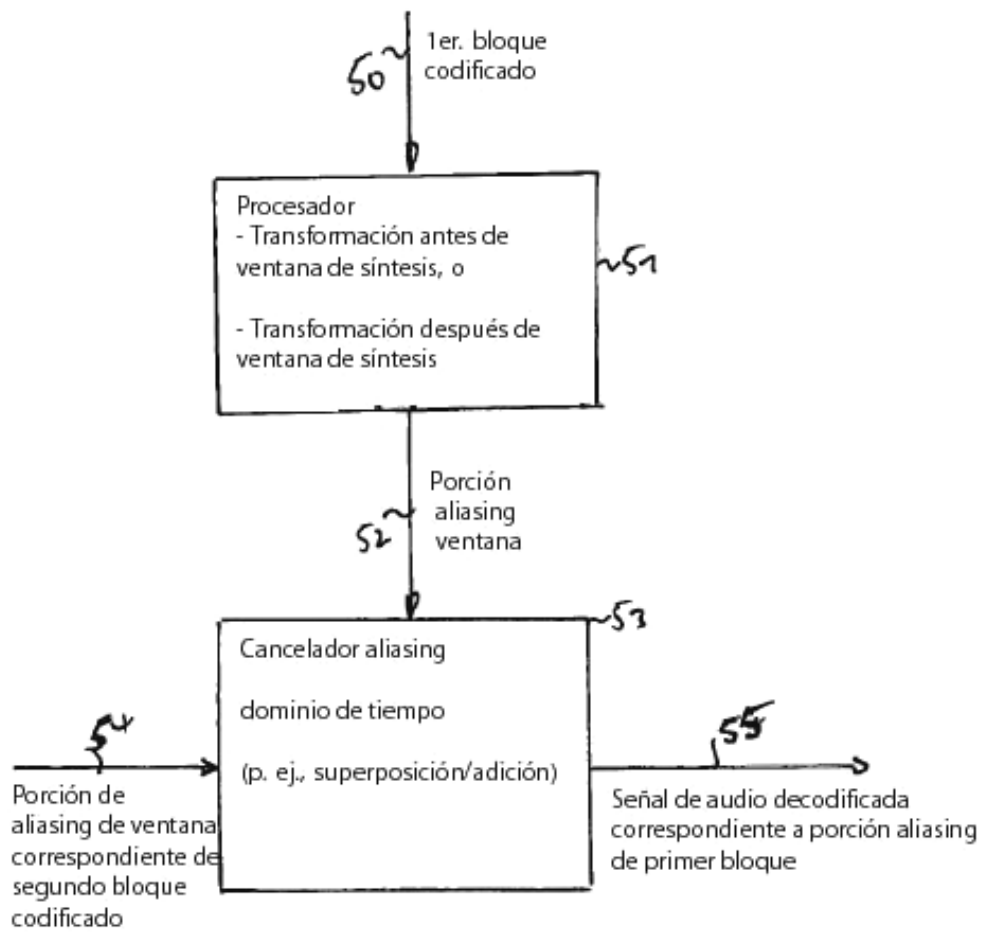


Fig. 3A (Lado del decodificador)

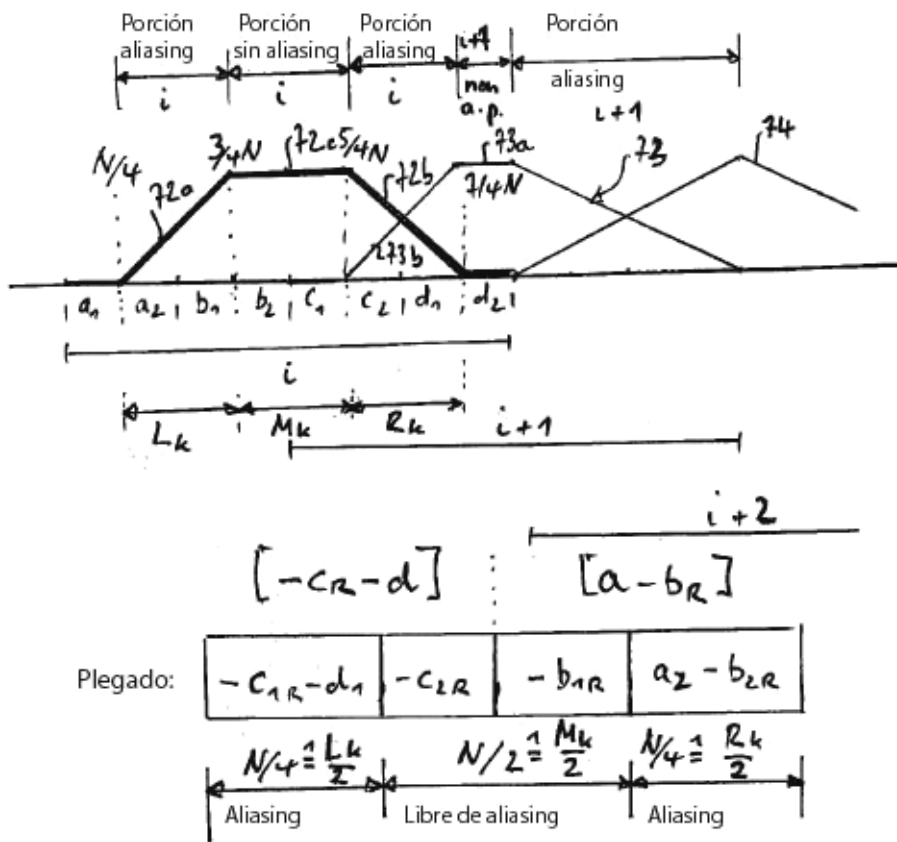


Fig. 4B

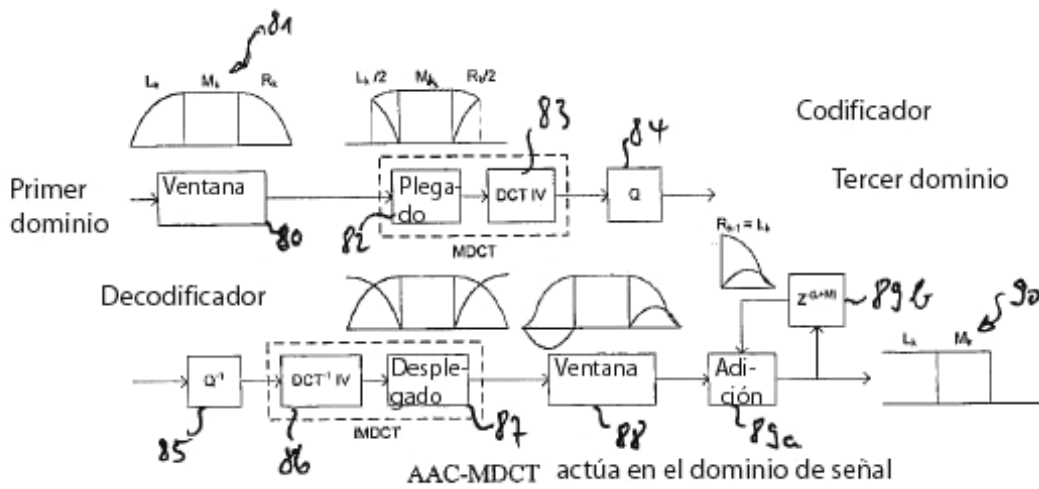


Fig.5

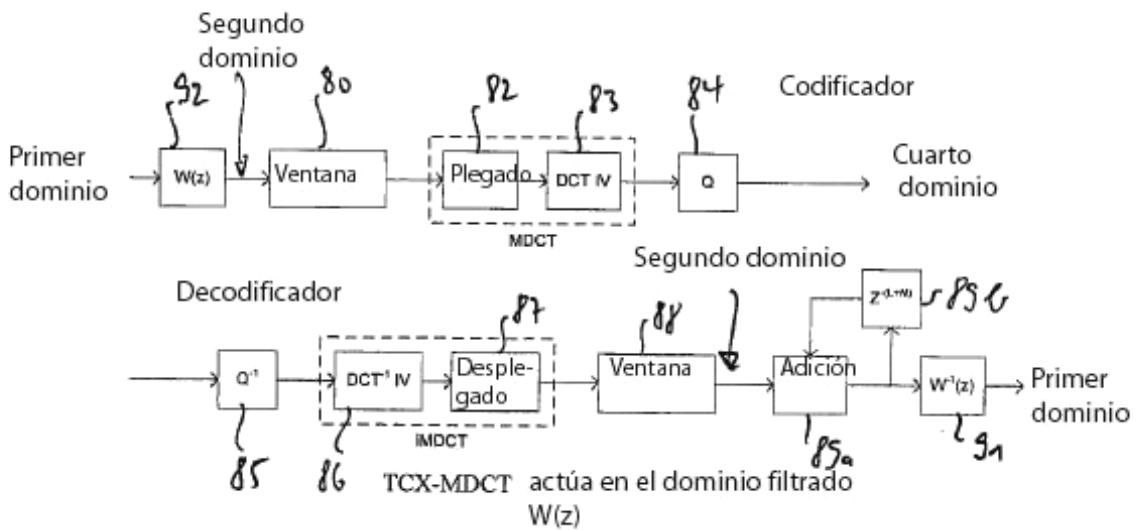
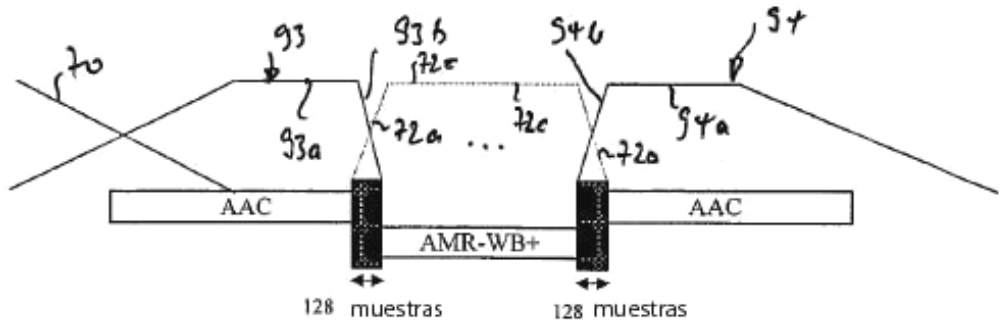


Fig.6



secuencia específica de ventanas para transiciones entre AAC y AMR-WB+

Fig. 7

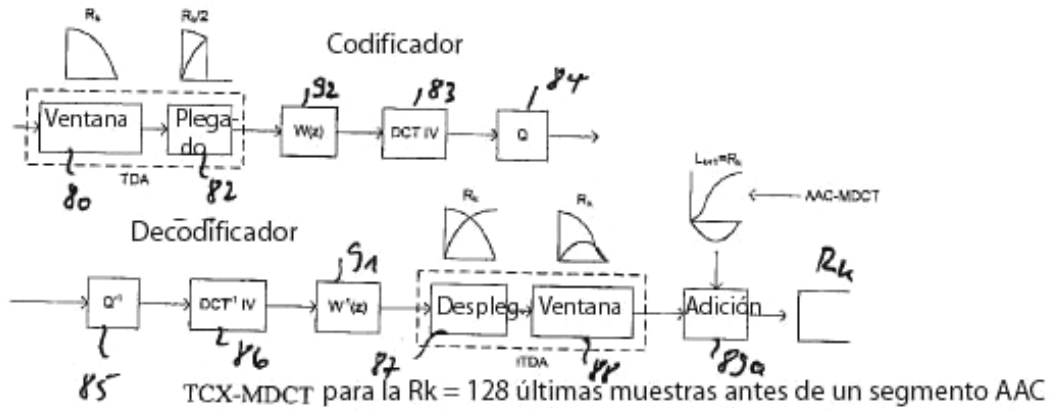


Fig. 8A

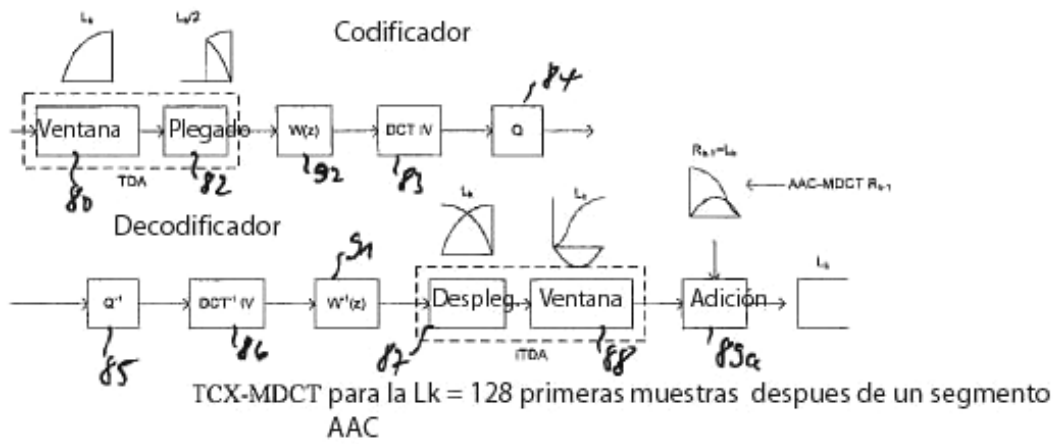


Fig. 8B

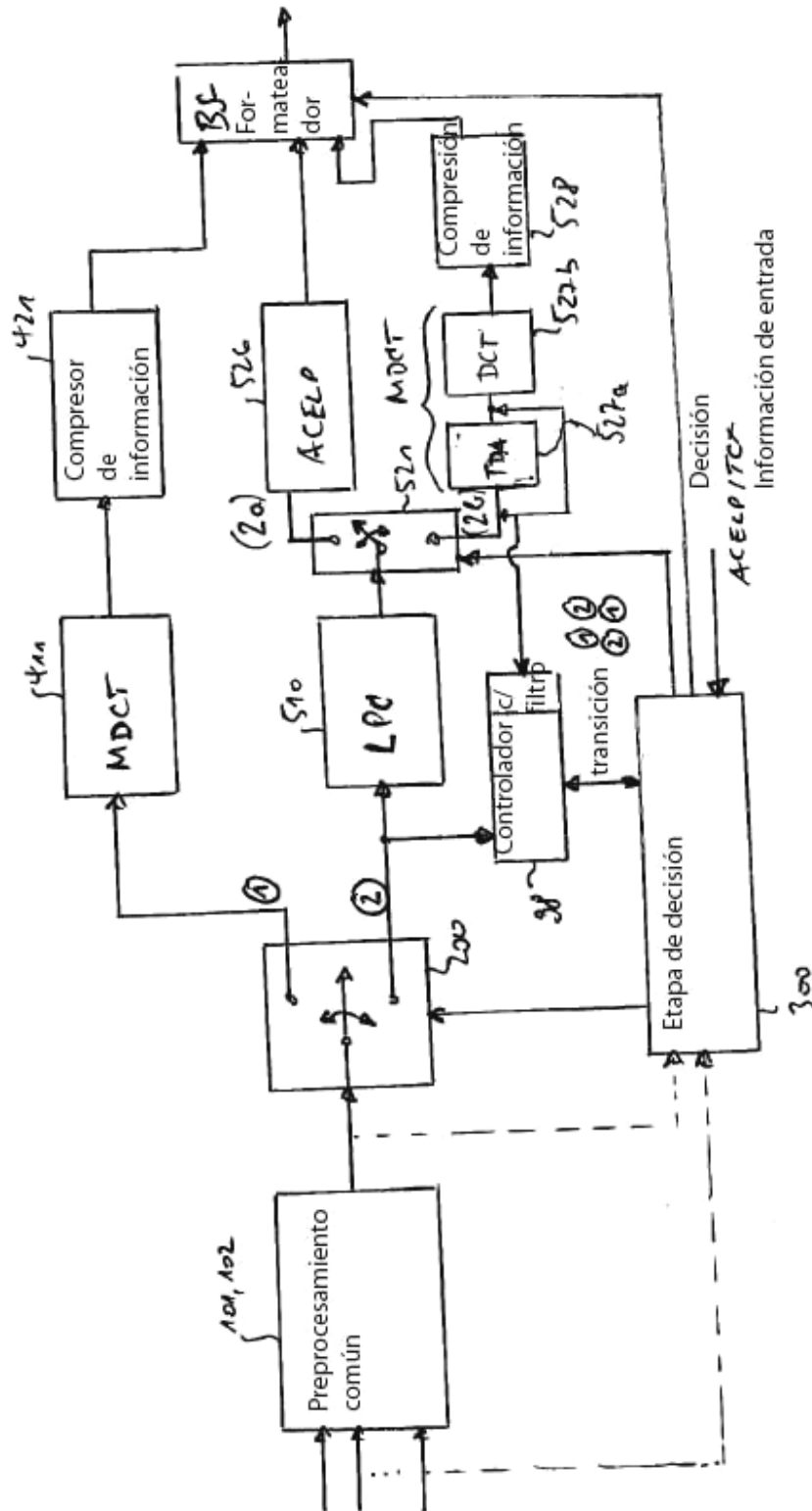


Fig. 9A

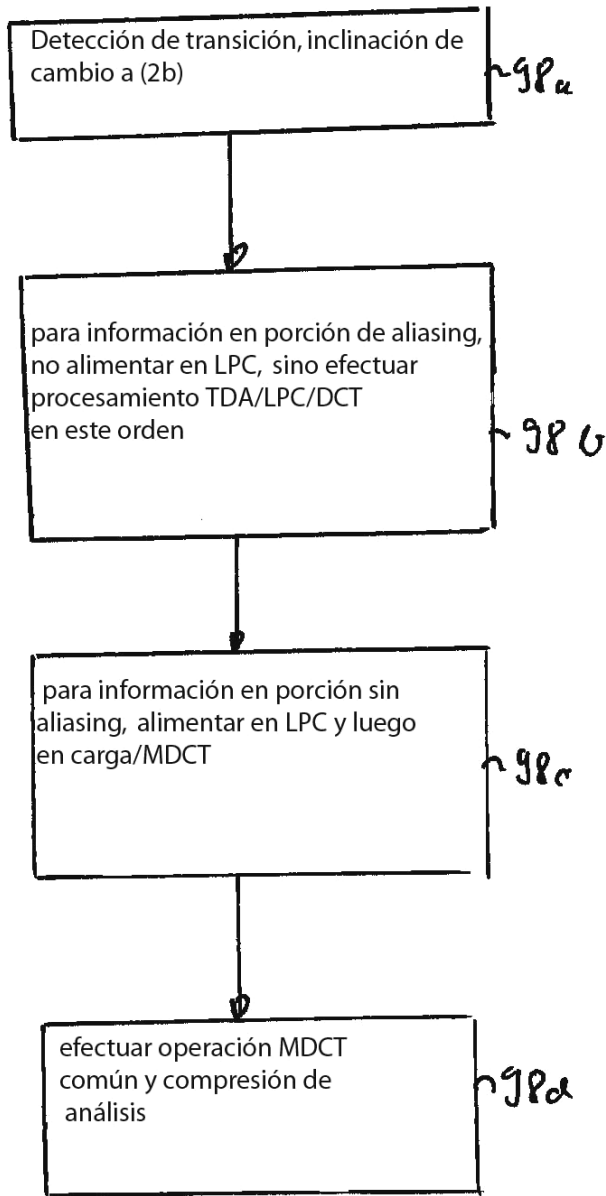


Fig. 9B

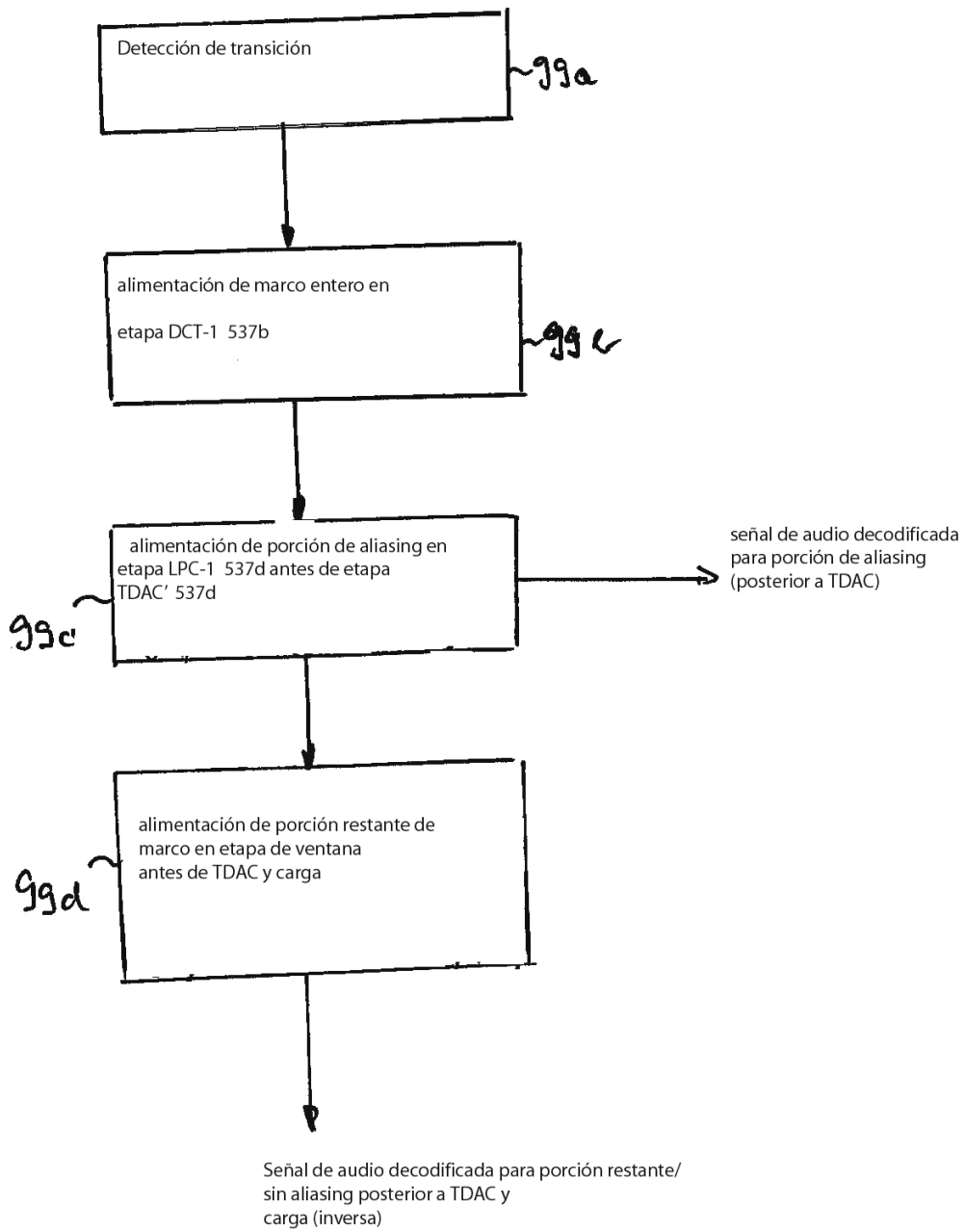


Fig. 10B

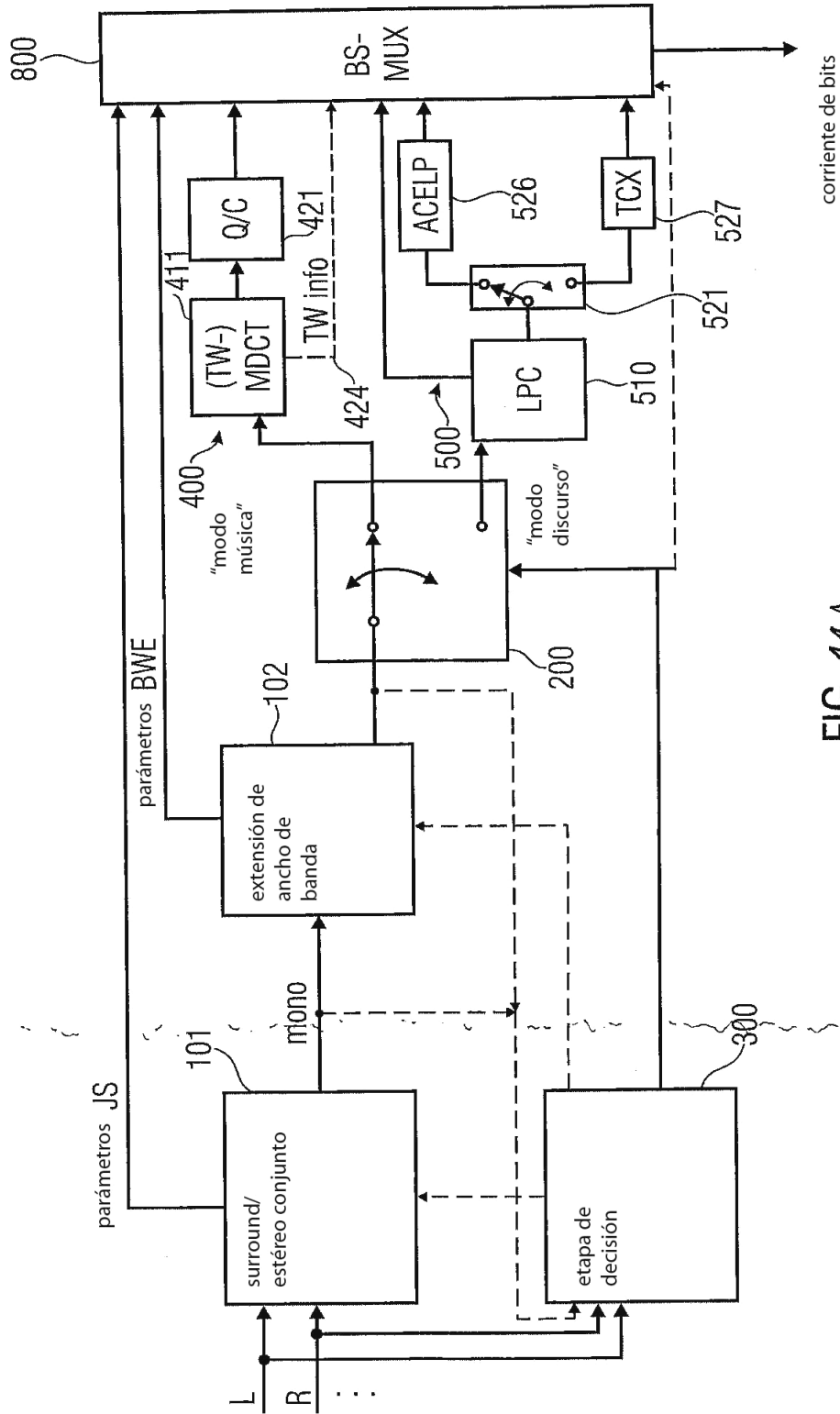


FIG 11A
(Codificador)

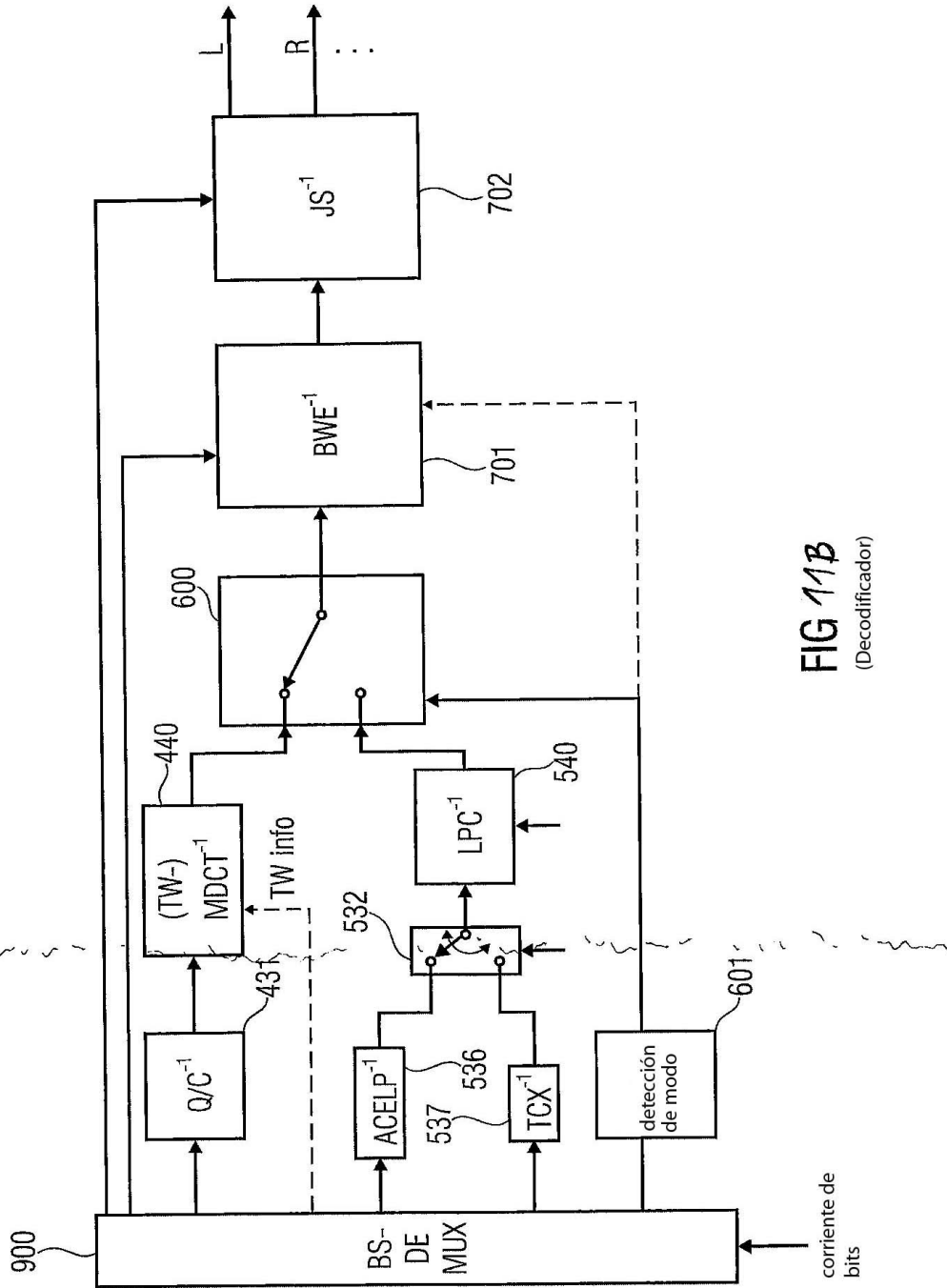


FIG 11B
(Decodificador)