

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 558 229**

51 Int. Cl.:

**G10L 19/20** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **08.07.2009 E 09777044 (0)**

97 Fecha y número de publicación de la concesión europea: **04.11.2015 EP 2311034**

54 Título: **Codificador y decodificador de audio para codificar tramas de señales de audio muestreadas**

30 Prioridad:

**11.07.2008 US 79851**  
**08.10.2008 US 103825**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**02.02.2016**

73 Titular/es:

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

72 Inventor/es:

**LECOMTE, JÉRÉMIE;**  
**GOURNAY, PHILIPPE;**  
**BAYER, STEFAN;**  
**MULTRUS, MARKUS y**  
**RETTELBACH, NIKOLAUS**

74 Agente/Representante:

**ARIZTI ACHA, Monica**

**ES 2 558 229 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

Codificador y decodificador de audio para codificar tramas de señales de audio muestreadas

- 5 La presente invención pertenece al campo de la codificación/decodificación de audio, en especial en los conceptos de codificación de audio en los que se utilizan múltiples dominios de codificación.

10 En la técnica, los esquemas de codificación de dominio de la frecuencia tales como MP3 o AAC son conocidos. Estos codificadores del dominio de la frecuencia se basan en una conversión del dominio de la frecuencia/dominio del tiempo, una etapa de cuantificación posterior, en la que el error de cuantificación es controlado por el uso de información de un módulo psicoacústico, y una etapa de codificación, en la que los coeficientes espectrales cuantificados y la información conexa correspondiente están codificados por entropía por el uso de tablas de códigos.

15 Por el contrario, existen codificadores que son muy adecuados para procesamiento del habla tales como AMR-WB+ como se describe en 3GPP TS 26.290. Tales esquemas de codificación del habla realizan un filtrado LP (LP = Predictivo Lineal) de una señal del dominio del tiempo. Un filtrado LP de este tipo, se deduce de un análisis de predicción lineal de la señal de entrada del dominio del tiempo. Los coeficientes resultantes del filtro LP se cuantifican/codifican a continuación y se transmiten como información conexa. El proceso es conocido como LPC (LPC = Codificación de Predicción Lineal). En la salida del filtro, la señal residual de predicción o la señal de error de predicción que se conoce también como la señal de excitación, es codificada por el uso de las etapas de análisis por síntesis del codificador ACELP o, como alternativa, es codificada por el uso de un codificador de transformación, que utiliza una transformada de Fourier con una superposición. La decisión entre la codificación por ACELP y la codificación por Excitación Codificada por Transformada, también denominada TCX, se hace por el uso de un algoritmo de bucle cerrado o un algoritmo de bucle abierto.

20 También pueden combinarse esquemas de codificación de audio por dominio de la frecuencia, tales como el esquema de codificación ACC de alta eficacia, que combina un esquema de codificación AAC y una técnica de replicación de banda espectral, con un estéreo conjunto o una herramienta de codificación de canales múltiples que se conoce bajo el término "MPEG envolvente".

Por el contrario, los codificadores del habla tales como el AMR-WB+ también tienen una etapa de potenciación de alta frecuencia y una funcionalidad estéreo.

35 Los esquemas de codificación de dominio de la frecuencia son ventajosos dado que muestran una alta calidad en velocidades de bits bajas para señales de música. Sin embargo, es problemática la calidad de las señales del habla en velocidades de bits bajas. Los esquemas de codificación del habla muestran una alta calidad para señales del habla, incluso en velocidades de bits bajas, pero muestran una calidad pobre para señales de música en velocidades de bits bajas.

40 A menudo, los esquemas de codificación de dominio de la frecuencia hacen uso de la denominada MDCT (MDCT= Transformada de Coseno Discreta Modificada). La MDCT se ha descrito inicialmente en 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 banco de filtros MDCT es ampliamente utilizada en los codificadores de audio modernos y eficaces. Esta clase de procesamiento de señales proporciona las siguientes ventajas:

50 Fundido cruzado suave entre bloques de procesamiento: incluso si la señal de cada bloque de procesamiento se altera en forma diferente (por ejemplo, debido a la cuantificación de coeficientes espectrales), no tienen lugar artefactos de bloqueo causados por transiciones abruptas de bloque a bloque debido a la operación de superposición/adición en ventanas.

Muestreo crítico: el número de valores espectrales en la salida del banco de filtros es igual al número de valores de entrada del dominio del tiempo en su entrada y tienen que transmitirse valores de tara adicionales.

55 El banco de filtros MDCT proporciona una selectividad de alta frecuencia y una ganancia de codificación.

60 Estas propiedades excelentes se logran por el uso de la técnica de cancelación de solapamiento del dominio del tiempo. La cancelación de solapamiento del dominio del tiempo se lleva a cabo en la síntesis por la superposición-adición de dos señales en ventana adyacentes. Si no se aplica cuantificación entre las etapas de análisis y de síntesis de la MDCT, se obtiene una reconstrucción perfecta de la señal original. Sin embargo, la MDCT se utiliza para esquemas de codificación, que se adaptan específicamente para señales de música. Tales esquemas de codificación del dominio de frecuencia, como se ha indicado anteriormente, reducen la calidad en velocidades de bits bajas para señales del habla, mientras que los codificadores del habla específicamente adaptados tienen una calidad superior a velocidades de bits comparables o incluso tienen velocidades de bits significativamente inferiores

para la misma calidad, en comparación con esquemas de codificación del dominio de la frecuencia.

Las técnicas de codificación del habla tales como el códec AMR-WB+ (AMR-WB+ = Banda Ancha de Múltiple Velocidad Adaptativa extendida), según lo definido en "Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec", 3GPP TS 26.290 V6.3.0, 2005-06, especificación técnica, no aplican la MDCT y, por lo tanto, no pueden aprovechar ninguna ventaja de las excelentes propiedades de la MDCT que, específicamente, se basan, por un lado, en un procesamiento críticamente muestreado, y, por el otro, de un cruce de un bloque al otro. Por lo tanto, el cruce de un bloque al otro obtenido por la MDCT sin ninguna penalización con respecto a la velocidad de bits y, por lo tanto, la propiedad de muestreo crítico de MDCT aún no ha sido obtenida en los codificadores del habla.

Cuando pretenden combinarse codificadores del habla y codificadores de audio dentro de un esquema de codificación híbrida único, aún existe el problema de cómo obtener una conmutación de un modo de codificación al otro modo de codificación a una velocidad de bits baja y a una alta calidad.

Se conoce un esquema de codificación híbrido de este tipo por ejemplo a partir del documento WO 2008/071353 A2.

Los conceptos convencionales de codificación de audio están normalmente diseñados para ser iniciados al comienzo de un archivo de audio o de una comunicación. Por el uso de estos conceptos convencionales, las estructuras de los filtros, como por ejemplo filtros de predicción, alcanzan un estado fijo en un cierto tiempo al comienzo del procedimiento de codificación o decodificación. Sin embargo, para un sistema conmutado de codificación de audio que utiliza por un lado, por ejemplo, codificación basada en transformadas y, por el otro, codificación del habla de acuerdo con un análisis previo de la entrada, las estructuras de filtros respectivas no se actualizan en forma activa y continua. Por ejemplo, puede solicitarse que los codificadores del habla sean reiniciados con frecuencia en un breve período de tiempo. Una vez reiniciados, un período de arranque es nuevamente iniciado, y los estados internos se resetean a cero. La duración necesaria para que un codificador del habla alcance, por ejemplo, un estado fijo puede ser crítico, en especial, para la calidad de las transiciones.

Los conceptos convencionales como por ejemplo el AMR-WB+, consúltese "Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec", 3GPP TS 26.290 V6.3.0, 2005-06, especificación técnica, utilizan un reseteo total del codificador del habla cuando pasa o conmuta entre el codificador basado en la transformada y el codificador del habla.

El AMR-WB+ se optimiza bajo la condición de que comience únicamente cuando la señal esté fundida en entrada, suponiendo que no haya interrupciones o reseteos intermedios. Por lo tanto, todas las memorias del codificador pueden actualizarse en una base trama a trama. En el caso de que el AMR-WB+ se utilice en el medio de una señal, debe realizarse un reseteo, y todas las memorias utilizadas en el lado de codificación o decodificación se establecen a cero. Por lo tanto, los conceptos convencionales tienen el problema de que se aplican duraciones demasiado extensas antes de alcanzar un estado fijo del codificador del habla, junto con la introducción de distorsiones intensas en las fases no fijas.

Otra desventaja de los conceptos convencionales es que utilizan largos segmentos en superposición cuando conmutan dominios de codificación que introducen taras, que afectan perjudicialmente la eficiencia de la codificación.

Es el objetivo de la presente invención proporcionar un concepto mejorado para la codificación de audio por el uso de la conmutación del dominio de la codificación.

El objetivo se logra con un codificador de audio de acuerdo con la reivindicación 1, un método para la codificación de audio de acuerdo con la reivindicación 3, un decodificador de audio de acuerdo con la reivindicación 4, un método para la decodificación de audio de acuerdo con la reivindicación 8 y un programa informático de acuerdo con la reivindicación 9.

La presente invención se basa en el hallazgo de que los problemas mencionados con anterioridad pueden resolverse en un decodificador, considerando la información de estado de un filtro acorde tras el reseteo. Por ejemplo, tras el reseteo, cuando los estados de un cierto filtro han sido establecidos a cero, puede acortarse el procedimiento de arranque o preparación del filtro, si el filtro no es iniciado desde cero, es decir, con todos los estados o memorias establecidas a cero, sino alimentado con una información de un cierto estado, a partir de lo que puede lograrse un período de arranque o preparación más corto.

Es otro hallazgo de la de la presente invención el hecho de que dicha información referente a un estado de conmutación puede generarse en el lado del codificador o del decodificador. Por ejemplo, cuando se conmuta entre un concepto de codificación basado en una predicción y un concepto de codificación basado en una transformada, puede proporcionarse información adicional antes de la conmutación, con el fin de permitir que el decodificador lleve los filtros de síntesis de predicción a un estado fijo antes de tener que utilizar realmente sus salidas.

En otras palabras, el hallazgo de la presente invención de que, en especial, cuando se conmuta, entre el dominio de la transformada al dominio de predicción en un codificador de audio conmutado, información adicional de los estados de los filtros justo antes de una conmutación real al dominio de predicción, puede resolver el problema de la generación de los artefactos de conmutación.

Es otro hallazgo de la presente invención el hecho de que tal información de la conmutación puede generarse únicamente en el decodificador, considerando sus salidas justo antes de que tenga lugar la conmutación real, y que básicamente el procesamiento del codificador se desarrolla en dicha salida, para determinar una información del filtro de los estados de memoria justo antes de la conmutación. Con ellos, algunas realizaciones pueden utilizar codificadores convencionales y reducir el problema de los artefactos de conmutación únicamente por el procesamiento del decodificador. Considerando dicha información, por ejemplo, los filtros de predicción ya pueden ir preparándose antes de la conmutación real, por ejemplo, analizando la salida de un decodificador del dominio de la transformada correspondiente.

Las realizaciones de la presente invención se detallarán usando las figuras adjuntas, en las que:

La Figura 1 muestra una realización de un codificador de audio;

La Figura 2 muestra una realización de un decodificador de audio;

La Figura 3 muestra una forma de ventana utilizada por una realización;

Las Figuras 4a y 4b ilustran la MDCT y el solapamiento del dominio del tiempo;

La Figura 5 ilustra un diagrama de bloques de una realización para la cancelación del solapamiento del dominio del tiempo;

Las Figuras 6a-6g ilustran señales que están procesadas para la cancelación del solapamiento del dominio del tiempo en una realización;

Las Figuras 7a-7g ilustran una cadena de procesamiento de señales para una cancelación del solapamiento del dominio del tiempo en una realización en que se utiliza un decodificador de la predicción lineal;

Las Figuras 8a-8g ilustran una cadena de procesamiento de señales en una realización con cancelación del solapamiento del dominio del tiempo; y

Las Figuras 9a y 9b ilustran un procesamiento de señales en los lados del codificador y decodificador en las realizaciones.

La Figura 1 muestra una realización de un codificador de audio 100. El codificador de audio 100 se adapta para codificar tramas de una señal de audio muestreada para obtener tramas codificadas, donde una trama comprende un número de muestras de audio del dominio del tiempo. La realización del codificador de audio comprende un estado de análisis de codificación predictiva 110 para determinar una información de los coeficientes de un filtro de síntesis y una información sobre una trama de dominio de predicción basándose en una trama de muestras de audio. En las realizaciones, la trama de dominio de predicción puede corresponder a una trama de excitación o a una versión filtrada de una trama de excitación. A continuación, puede referirse a la codificación del dominio de predicción cuando se codifica una información sobre los coeficientes de un filtro de síntesis y una información de una trama del dominio de predicción basándose en una trama de muestras de audio.

Además, la realización del codificador de audio 100 comprende un transformador del dominio de la frecuencia 120 para transformar una trama de muestras de audio al dominio de la frecuencia para obtener un espectro de tramas. A continuación, puede referirse a la codificación del dominio de la transformada cuando se codifica un espectro de tramas. Además, la realización del codificador de audio 100 comprende un tomador de decisiones del dominio de codificación 130 para decidir, si los datos codificados para una trama están basados en la información sobre los coeficientes y en la información de la trama del dominio de predicción, o están basados en el espectro de tramas. La realización del codificador de audio 100 comprende un controlador 140 para determinar una información de un coeficiente de conmutación, cuando el tomador de decisiones del dominio de codificación decide que los datos codificados de una trama actual están basados en la información sobre los coeficientes y en la información sobre la trama de dominio de predicción cuando los datos codificados de una trama previa se codificaron basándose en un espectro de tramas previo. La realización del codificador de audio 100 además comprende un codificador reductor de la redundancia 150 para codificar la información sobre la trama de dominio de predicción, la información de los coeficientes, la información sobre el coeficiente del dominio de conmutación y/o el espectro de tramas. En otras palabras, el tomador de decisiones del dominio de codificación 130 decide el dominio de codificación, mientras que

el controlador 140 proporciona la información sobre el coeficiente de conmutación cuando se conmuta del dominio de la transformada al dominio de predicción.

5 En la Figura 1 existen algunas conexiones mostradas por líneas discontinuas. Estas indican las diferentes opciones en las realizaciones. Por ejemplo, la información sobre los coeficientes de conmutación puede obtenerse simplemente llevando a cabo en forma permanente la etapa de análisis de codificación predictiva 110 de modo que la información sobre los coeficientes y la información de las tramas del dominio de predicción estén siempre disponibles en su salida. El controlador 140 puede a continuación indicar al codificador reductor de la redundancia 150 cuándo codificar la salida a partir de la etapa de análisis de codificación predictiva 110 y cuándo codificar la salida del espectro de tramas en un transformador del dominio de la frecuencia 120 después de que la decisión de conmutación haya sido tomada por el tomador de decisiones del dominio de codificación 130. Por lo tanto, el controlador 140 puede controlar el codificador reductor de la redundancia 150 para codificar la información sobre el coeficiente de conmutación cuando se conmuta del dominio de la transformada al dominio de predicción.

15 Si tiene lugar la conmutación, el controlador 140 puede indicar al codificador reductor de la redundancia 150 que codifique una trama en superposición, durante una trama anterior, el codificador reductor de la redundancia 150 puede ser controlado por el controlador 140 en un modo tal que una corriente de bits contenga para la trama previa, información sobre los coeficientes e información sobre la trama de dominio de predicción, así como el espectro de tramas. En otras palabras, en las realizaciones, el controlador puede controlar el codificador reductor de la redundancia 150 en un modo tal que las tramas codificadas incluyen la información anteriormente descrita. En otras realizaciones, el tomador de decisiones del dominio de codificación 130 puede decidir cambiar el dominio de codificación y conmutar entre la etapa de análisis de codificación predictiva 110 y el transformador del dominio de la frecuencia 120.

25 En estas realizaciones, el controlador 140 puede llevar a cabo algunos análisis internamente para proporcionar los coeficientes de conmutación. En las realizaciones, la información sobre un coeficiente de conmutación puede corresponder a una información de los estados de los filtros, del contenido del libro de códigos adaptativos, de los estados de memoria, de la información de una señal de excitación, de los coeficientes de LPC, etc. La información sobre el coeficiente de conmutación puede comprender cualquier información que permita la preparación o iniciación de una etapa de síntesis predictiva 220.

35 El tomador de decisiones del dominio de codificación 130 puede determinar su decisión referente a cuándo conmutar el dominio de codificación basándose en las tramas o muestras de señales de audio, que también se indica por las líneas discontinuas en la Figura 1. En otras realizaciones, dicha decisión puede tomarse basándose en los coeficientes de información, en la información de la trama del dominio de predicción y/o en el espectro de tramas.

40 En general, las realizaciones no se limitan al modo en que el tomador de decisiones del dominio de codificación 130 decide cuándo cambiar el dominio de codificación, es más importante que los cambios del dominio de codificación sean decididos por el tomador de decisiones del dominio de codificación 130, durante lo que tienen lugar los problemas anteriormente descritos, y en el que en algunas realizaciones el codificador de audio 100 está coordinado en un modo tal que los efectos desventajosos anteriormente descritos están, al menos parcialmente, compensados.

45 En las realizaciones, el tomador de decisiones del dominio de codificación 130 puede adaptarse para decidir basándose en una propiedad de señales o en las propiedades de las tramas de audio. Como ya es sabido, las propiedades de audio de una señal de audio pueden determinar la eficiencia de la codificación, es decir, para ciertas características de una señal de audio, puede ser más eficiente utilizar codificación basándose en las transformadas, para otras características puede ser más beneficioso utilizar codificación del dominio de predicción. En algunas realizaciones, el tomador de decisiones del dominio de codificación 130 puede adaptarse para decidir el uso de codificación basándose en transformadas cuando la señal sea muy tonal o muda. Si la señal es una señal transitoria o una señal similar a una voz, el tomador de decisiones del dominio de codificación 130 puede adaptarse para decidir el uso de una trama del dominio de predicción como se indica para la codificación.

55 De acuerdo con las otras líneas discontinuas y las flechas de la Figura 1, el controlador 140 puede proporcionarse con la información de los coeficientes, la información de la trama del dominio de predicción y el espectro de tramas, y el controlador 140 puede adaptarse para determinar la información sobre el coeficiente de conmutación basándose en dicha información. En otras realizaciones, el controlador 140 puede proporcionar una información a la etapa de análisis de codificación predictiva 110 para determinar los coeficientes de conmutación. En las realizaciones, los coeficientes de conmutación pueden corresponder a la información de los coeficientes y en otras realizaciones, pueden estar determinados en un modo diferente.

60 La Figura 2 ilustra una realización de un decodificador de audio 200. La realización del decodificador de audio 200 se adapta para decodificar tramas codificadas para obtener tramas de una señal de audio muestreada, en el que una trama comprende un número de muestras de audio del dominio del tiempo. La realización del decodificador de audio 200 comprende un decodificador recuperador de la redundancia 210 para decodificar las tramas codificadas

para obtener una información de una trama del dominio de predicción, una información de los coeficientes para un filtro de síntesis y/o un espectro de tramas. Además, la realización del decodificador de audio 200 comprende una etapa de síntesis predictiva 220 para determinar una trama predicha de muestras de audio basándose en la información sobre los coeficientes para el filtro de síntesis y la información sobre la trama del dominio de predicción, y un transformador del dominio del tiempo 230 para transformar el espectro de tramas al dominio del tiempo para obtener una trama transformada a partir del espectro de tramas. La realización del decodificador de audio 200 además comprende un combinador 240 para combinar la trama transformada y la trama predicha para obtener las tramas de la señal de audio muestreada.

Además, la realización del decodificador de audio 200 comprende un controlador 250 para controlar un proceso de conmutación, llevándose a cabo el proceso de conmutación cuando una trama previa está basada en la trama transformada y una trama actual está basada en una trama predicha, estando configurado el controlador 250 para proporcionar coeficientes de conmutación a la etapa de síntesis predictiva 220 para entrenar, iniciar o preparar a la etapa de síntesis predictiva 220, de modo que la etapa de síntesis predictiva 220 se inicie cuando se lleve a cabo el proceso de conmutación.

De acuerdo con las flechas discontinuas mostradas en la Figura 2, el controlador 250 puede adaptarse para controlar partes o todos los componentes del decodificador de audio 200. Por ejemplo, el controlador 250 puede adaptarse por ejemplo, para coordinar el decodificador recuperador de la redundancia 210, para recuperar información extra de los coeficientes de conmutación o información de la trama del dominio de predicción previo, etc. En otras realizaciones, el controlador 250 puede adaptarse para deducir dicha información de los coeficientes de conmutación por sí mismo, por ejemplo proporcionándose con las tramas decodificadas por el combinador 240, llevando a cabo un análisis de LP basándose en la salida del combinador 240. A continuación, el controlador 250 puede adaptarse para coordinar o controlar la etapa de síntesis predictiva 220 y un transformador del dominio del tiempo 230 para establecer las tramas en superposición anteriormente descritas, la temporización, el análisis del dominio del tiempo y la cancelación del análisis del dominio del tiempo, etc.

A continuación, se considera un códec del dominio basándose en LPC, que incluye predictores y filtros internos que, durante un arranque necesitan un cierto tiempo para alcanzar un estado que asegura una síntesis de filtro precisa. En otras palabras, en las realizaciones del codificador de audio 100, la etapa de análisis de codificación predictiva 110 puede adaptarse para determinar la información de los coeficientes del filtro de síntesis y la información de la trama de dominio de predicción basándose en un análisis de LPC. En las realizaciones del decodificador de audio 200, la etapa de síntesis predictiva 220 puede adaptarse para determinar las tramas predichas basándose en un filtro de síntesis LPC.

Es evidente que el uso de una ventana rectangular al comienzo de la primera trama de LPD (LPD = Dominio de Predicción Lineal) y el reseteo del códec basado en LPC a un estado de cero, no proporciona una opción ideal para estas transiciones, dado que el códec de LPD no cuenta con tiempo suficiente para generar una buena señal, lo que introduciría artefactos de bloqueo.

En las realizaciones, para manejar la transición de un modo no LPD a un modo LPD, pueden utilizarse ventanas en superposición. En otras palabras, en las realizaciones del codificador de audio 100, el transformador del dominio de la frecuencia 120 puede adaptarse para transformar la trama de muestras de audio, basándose en una Transformada Rápida de Fourier (FFT = Transformada Rápida de Fourier), o una MDCT (MDCT = Transformada de Coseno Discreta Modificada). En las realizaciones del decodificador de audio 200, el transformador del dominio del tiempo 230 puede adaptarse para transformar los espectros de tramas al dominio de tiempo, basándose en una FTT inversa (IFFT = FTT inversa), o una MDCT inversa (IMDCT = MDCT inversa).

Con ello, las realizaciones pueden llevarse a cabo en un modo no LPD, que también puede denominarse como el modo basado en transformadas, o en un modo LPD, que también se denomina como el análisis y síntesis predictivo. En general, las realizaciones pueden utilizar ventanas en superposición, en especial cuando se utiliza MDCT e IMDCT. En otras palabras, en el modo no LPD, puede utilizarse generación de ventanas en superposición con solapamiento del dominio del tiempo (TDA = Solapamiento del Dominio del Tiempo). Con ello, cuando se conmuta del modo no LPD al modo LPD, puede compensarse el solapamiento del dominio del tiempo de la última trama no LPD. Las realizaciones pueden introducir un solapamiento del dominio del tiempo en la señal original antes de llevar a cabo la codificación LPD, sin embargo, el solapamiento del dominio del tiempo puede no ser compatible con la codificación del dominio del tiempo basada en la predicción, tal como ACELP (ACELP = Predicción Lineal con Excitación por Libro de Códigos Algebraicos). Las realizaciones pueden introducir un solapamiento artificial al comienzo del segmento LPD y aplicar cancelación del dominio del tiempo de la misma manera que para ACELP a transiciones no LPD. En otras palabras, el análisis y la síntesis predictiva pueden basarse en una ACELP en las realizaciones.

En algunas realizaciones, el solapamiento artificial se produce a partir de la señal de síntesis en lugar de la señal original. Dado que la señal de síntesis es imprecisa, en especial al arranque del LPD, estas realizaciones pueden compensar bastante los artefactos de bloqueo por la introducción de un TDA artificial, sin embargo, la introducción

del TDA artificial puede introducir un error de imprecisión junto con la reducción de los artefactos.

La Figura 3 ilustra un proceso de conmutación dentro de una realización. En la realización mostrada en la Figura 3, se supone que el proceso de conmutación conmuta del modo no LPD, por ejemplo, el modo MDCT, al modo LPD. De acuerdo con lo indicado en la Figura 3, se considera una longitud de ventana total de 2048 muestras. En el lado izquierdo de la Figura 3, se ilustra el borde creciente de la ventana MDCT que se extiende a través de 512 muestras. Durante el proceso de MDCT e IMDCT, estas 512 muestras del borde creciente de la ventana MDCT se plegarán con las próximas 512 muestras, que se asignan en la Figura 3 al núcleo MDCT, que comprende las 1024 muestras centradas dentro de la ventana completa de 2048 muestras. Como se explicará en más detalle a continuación, el solapamiento del dominio del tiempo introducido por el proceso de MDCT e IMDCT no es crítico cuando la trama precedente también haya sido codificada en el modo no LPD, dado que una de las propiedades ventajosas de la MDCT es que el solapamiento del dominio del tiempo puede compensarse intrínsecamente por las ventanas MDCT en superposición consecutivas respectivas.

Sin embargo, cuando se conmuta al modo LPD, es decir, considerando ahora la parte a la derecha de la ventana MDCT mostrada en la Figura 3, tal cancelación del solapamiento del dominio del tiempo no se lleva a cabo automáticamente, dado que la primera trama decodificada en el modo LPD no tiene que compensar automáticamente el solapamiento del dominio del tiempo con la trama MDCT precedente. Por lo tanto, en una región en superposición, las realizaciones pueden introducir un solapamiento artificial del dominio del tiempo, como se indica en la Figura 3 en el área de las 128 muestras centradas al final de la ventana del núcleo MDCT, es decir, centradas después de 1536 muestras. En otras palabras, en la Figura 3 se supone que el solapamiento artificial del dominio del tiempo se introduce al comienzo, es decir, en esta realización, las primeras 128 muestras, de la trama del modo LPD, para compensarlo con el solapamiento del dominio de tiempo introducido al final de la última trama MDCT.

En la realización preferida, la MDCT se aplica con el fin de obtener la conmutación críticamente muestreada de una operación de codificación en un dominio a una operación de codificación en otro dominio diferente, es decir, llevándose a cabo en realizaciones del transformador del dominio de la frecuencia 120 y/o del transformador del dominio del tiempo 230. Sin embargo, también pueden aplicarse todas las otras transformadas. No obstante, dado que la MDCT es la realización preferida, se analizará la MDCT en mayor detalle con respecto a la Figura 4a y Figura 4b.

La Figura 4a ilustra una ventana 470, 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 470 tiene, como puede observarse en la figura, solo porciones de solapamiento en la situación del 50 % de superposición/adición ilustrada. Específicamente, la primera porción que tiene muestras de cero a N corresponde a las segundas porciones de una ventana precedente 469, y la segunda mitad que se extiende entre la muestra N y la muestra 2N de la ventana 470 está superpuesta con la primera porción de la ventana 471, que está en la ventana de la realización ilustrada  $i+1$ , mientras que la ventana 470 es la ventana  $i$ .

La operación de MDCT puede observarse como la puesta en cascada de la generación de ventanas y la operación de plegado y una operación de transformada posterior y, específicamente, una operación de DCT (DCT = Transformada de Coseno Discreta) posterior, donde se aplica la DCT de tipo IV (DCT-IV). Específicamente, la operación de plegado se obtiene por calculando la primera porción  $N/2$  del bloque de plegado como  $-c_R-d$ , y calculando la segunda porción de  $N/2$  muestras de la salida de plegado como  $a-b_R$ , donde R es el operador inverso. Así, la operación de plegado da como resultado N valores de salida mientras que se reciben 2N valores de entrada 2N.

Se ilustra también en la Figura 4a, en forma de ecuación, una operación de desplegado correspondiente en el lado del decodificador.

En general, una operación de MDCT sobre (a, b, c, d) da como resultado exactamente los mismos valores de salida que la DCT-IV de  $(-c_R-d, a-b_R)$ , como se indica en la Figura 4a.

Correspondientemente, y utilizando la operación de desplegado, una operación de IMDCT da como resultado la salida de la operación de desplegado aplicada a la salida de una transformada inversa por DCT-IV.

Por lo tanto, el solapamiento del tiempo se introduce realizando una operación de plegado en el lado del codificador. A continuación, el resultado de la operación de generación de ventanas y plegado se transforma en el dominio de la frecuencia por el uso de una transformada de bloque DCT-IV que requiere N valores de entrada.

En el lado del decodificador, N valores de entrada se transforman de vuelta al dominio del tiempo por el uso de una operación DCT-IV, y la salida de esta operación de transformada inversa se cambia de esta manera a una operación de desplegado para obtener 2N valores de salida que, sin embargo, son valores de salida solapados.

Con el objetivo de retirar el solapamiento que se ha introducido por la operación de plegado y que aún permanece con posterioridad a la operación de desplegado, la operación de superposición/adición puede llevar a cabo la cancelación del solapamiento del dominio del tiempo.

5 Por lo tanto, cuando el resultado de la operación de desplegado se agrega con el resultado de la IMDCT previo en la mitad en superposición, los términos invertidos se cancelan en la ecuación de la parte inferior de la Figura 4a y se obtiene de manera sencilla, por ejemplo, b y d, recuperando de esta manera los datos originales.

10 Con el objetivo de obtener una TDAC para la MDCT en ventana, existe un requisito, que se conoce como condición "Princen-Bradley", que significa que los coeficientes de la ventana se elevan a 2 para las muestras correspondientes que se combinan en el cancelador del solapamiento del dominio del tiempo, para dar como resultado la unidad (1) para cada muestra.

15 Mientras que la Figura 4a ilustra la secuencia de ventana como, por ejemplo, aplicada en la AAC-MDCT (AAC = Codificación Avanzada de Audio) para ventanas extensas o ventanas breves, la Figura 4b ilustra una función de ventana diferente que también tiene, además de las porciones solapamiento, una porción no de solapamiento también.

20 La Figura 4b ilustra una función de ventana de análisis 472 que tiene una porción cero a1 y d2, que tiene una porción de solapamiento 472a, 472b, y que tiene una porción no de solapamiento 472c.

La porción de solapamiento 472b que se extiende sobre c2, d1 tiene una porción de solapamiento correspondiente de una ventana posterior 473, que se indica en 473b. Correspondientemente, la ventana 473 comprende  
 25 adicionalmente una porción no de solapamiento 473a. La Figura 4b, cuando se compara con la Figura 4a expone con claridad que, dado que existen porciones cero a1, d1, para la ventana 472 o c1 para la ventana 473, ambas ventanas reciben una porción no de solapamiento, y que la función de ventana en la porción de solapamiento está más escalonada que en la Figura 4a. En vista de eso, la porción de solapamiento 472a corresponde a  $L_k$ , la porción no de solapamiento 472c corresponde a la porción  $M_k$ , y la porción de solapamiento 472b corresponde a  $R_k$  en la  
 30 Figura 4b.

Cuando la operación de plegado se aplica a un bloque de muestras en ventana por la ventana 472, se obtiene una situación como se ilustra en la Figura 4b. La porción izquierda que se extiende sobre las primeras N/4 muestras tiene solapamiento. La segunda porción que se extiende sobre las N/2 muestras está libre de solapamiento, dado  
 35 que la operación de plegado se aplica sobre las porciones de ventana que tienen valores cero, y las últimas N/4 muestras están, nuevamente, afectadas por el solapamiento. Debido a la operación de plegado, el número de valores de salida de la operación de plegado es igual a N, mientras que la entrada era 2N, aunque, de hecho, se establecieron N/2 valores en esta realización a cero a causa de la operación de generación de ventanas que utiliza la ventana 472.

40 Ahora, la DCT-IV se aplica al resultado de la operación de plegado, pero, en gran medida, la porción de solapamiento 472, que está en la transición de un modo de codificación a otro modo de codificación se procesa de una forma diferente a la de la porción no de solapamiento, aunque ambas porciones pertenecen al mismo bloque de muestras de audio y, en gran medida, se introducen en la misma operación de transformada de bloque.

45 La Figura 4b ilustra adicionalmente una secuencia de ventana de las ventanas 472, 473, 474, donde la ventana 473 es una ventana de transición de una situación donde no existen porciones de solapamiento a una situación donde únicamente existen porciones de solapamiento. Esto se obtiene conformando asimétricamente la función de ventana. La porción derecha de la ventana 473 es similar a la porción derecha de las ventanas en la secuencia de  
 50 ventanas de la Figura 4a, mientras que la porción izquierda tiene una porción no de solapamiento y la porción cero correspondiente (en c1). Por lo tanto, la Figura 4b ilustra una transición de MDCT-TCX a AAC, cuando la AAC va a realizarse utilizando ventanas completamente en superposición o, como alternativa, se ilustra una transición de AAC a MDCT-TCX, cuando la ventana 474 genera en ventanas un bloque de datos TCX en una forma completamente en superposición, que es, por un lado, la operación normal para MDCT-TCX y, por el otro lado, MDCT-AAC cuando no  
 55 existe razón para conmutar de un modo al otro modo.

Por lo tanto, la ventana 473 puede denominarse como una "ventana de interrupción", que tiene, además, la característica preferida de que la longitud de esta ventana es idéntica a la longitud de al menos una ventana vecina, de modo que se mantiene el patrón de bloques general o la exploración de trama, cuando se establece un bloque  
 60 para que tenga el mismo número que coeficientes de ventana, es decir, 2N muestras en el ejemplo de la Figura 4a o Figura 4b.

A continuación, se describirá en detalle el método del solapamiento artificial del dominio del tiempo y la cancelación del solapamiento del dominio del tiempo. La Figura 5 muestra un diagrama de bloques, que puede utilizarse en una

realización, que muestra una cadena de procesamiento de señales. Las Figuras 6a a 6g y 7a a 7g ilustran señales de muestra, donde las Figuras 6a a 6g ilustran un proceso principal de la cancelación del solapamiento del dominio del tiempo suponiendo que se utiliza la señal original, en el que se ilustran las muestras de señales de las Figuras 7a a 7g que se determinan basándose en la suposición de que la primera trama LPD resulta después de un reseteo completo y sin ninguna adaptación.

En otras palabras, la Figura 5 ilustra una realización de un proceso para introducir un solapamiento artificial del dominio del tiempo y una cancelación del solapamiento del dominio del tiempo para la primera trama en modo LPD, en caso de transición de un modo no LPD a un modo LPD. La Figura 5 muestra que, en primer lugar, se aplica una generación de ventanas a la trama LPD actual en el bloque 510. Como ilustran las Figuras 6a, 6b, y Figuras 7a, 7b, la generación de ventanas corresponde a un fundido de entrada de las señales respectivas. Como se ilustra en el pequeño gráfico de vista por encima del bloque de generación de ventanas 510 en la Figura 5, se supone que la generación de ventanas se aplica a  $L_k$  muestras. La generación de ventanas 510 se sigue por una operación de plegado 520, que da como resultado  $L_k/2$  muestras. El resultado de la operación de plegado se ilustra en las Figuras 6c y 7c. Puede observarse que a causa del número reducido de muestras, existe un período de cero que se extiende a través de las  $L_k/2$  muestras al comienzo de las señales respectivas.

Las operaciones de generación de ventanas en el bloque 510 y de plegado en el bloque 520 pueden resumirse como el solapamiento del dominio del tiempo que se introduce a través de MDCT. Sin embargo, surgen efectos de solapamiento adicionales cuando se transforma a la inversa a través de IMDCT. Los efectos provocados por la IMDCT se resumen en la Figura 5 en los bloques 530 y 540, que pueden nuevamente resumirse como el solapamiento invertido del dominio del tiempo. Como se muestra en la Figura 5, a continuación se lleva a cabo un desplegado en el bloque 530, que da como resultado la duplicación del número de muestras, es decir, da como resultado  $L_k$  muestras. Las señales respectivas se muestran en las Figuras 6d y 7d. Puede observarse a partir de las Figuras 6d y 7d que los números de muestras han sido duplicados, y que se ha introducido un solapamiento del tiempo. La operación de desplegado 530 se sigue por otra operación de generación de ventanas 540, para fundir en entrada las señales. Los resultados de la segunda generación de ventanas 540 se muestran en las Figuras 6e y 7e. Por último, las señales solapadas artificialmente en el tiempo mostradas en las Figuras 6e y 7e están superpuestas y se agregan a la trama previa codificada en el modo no LPD, que se indica en el bloque 550 en la Figura 5, y las señales respectivas se muestran en las Figuras 6f y 7f.

En otras palabras, en las realizaciones del decodificador de audio 200, el combinador 240 puede adaptarse para llevar a cabo las funciones del bloque 550 en la Figura 5.

Las señales resultantes se muestran en las Figuras 6g y 7g. Resumiendo, en ambos casos la parte izquierda de la trama respectiva está en ventanas, que se indica mediante las Figuras 6a, 6b, 7a, y 7b. A continuación, la parte izquierda de la ventana se pliega, que se indica en las Figuras 6c y 7c. Después del desplegado, consúltese 6d y 7d, se aplica otra generación de ventanas, consúltese Figuras 6e y 7e. Las Figuras 6f y 7f muestran la trama en proceso actual con la forma de la trama no LPD previa y las Figuras 6g y 7g muestran los resultados después de una operación de superposición y adición. A partir de las Figuras 6a a 6g, puede observarse que puede lograrse una reconstrucción perfecta mediante las realizaciones después de aplicar una TDA artificial sobre la trama LPD y de aplicar la superposición y adición con la trama previa. Sin embargo, en el segundo caso, es decir, el caso ilustrado en las Figuras 7a a 7g, la reconstrucción no es perfecta. Como ya se ha mencionado anteriormente, se supone que en el segundo caso, el modo LPD se ha reseteado completamente, es decir que los estados y memorias de la síntesis de LPC se establecieron a cero. Esto da como resultado que la señal de síntesis no sea precisa durante las primeras muestras. En este caso, la TDA artificial más la adición en superposición da como resultado distorsiones y artefactos, en lugar de una reconstrucción perfecta, consúltese Figuras 6g y 7g.

Las Figuras 6a a 6g y 8a a 8g ilustran otra comparación entre el uso de la señal original para el solapamiento artificial del dominio del tiempo y la cancelación del solapamiento del dominio del tiempo, y otro caso del uso de la señal de arranque LPD, sin embargo, en las Figuras 8a a 8g, se supuso que el período de arranque LPD toma más tiempo que el que toma en las Figuras 7a a 7g. Las Figuras 6a a 6g y 8a a 8g ilustran gráficos de señales de muestra a las que se han aplicado las mismas operaciones ya explicadas con respecto a la Figura 5. Comparando las Figuras 6g y 8g, puede observarse que las distorsiones y artefactos introducidos a la señal mostrada en la Figura 8g son incluso más significativos que aquellos de la Figura 7g. La señal mostrada en la Figura 8g contiene una gran cantidad de distorsiones durante un tiempo relativamente largo. Solo para comparación, la Figura 6g muestra la reconstrucción perfecta cuando se considera la señal original para la cancelación del solapamiento del dominio del tiempo.

Las realizaciones de la presente invención pueden acelerar el período de arranque, por ejemplo, de un códec de núcleo LPD, como una realización de la etapa de análisis de codificación predictiva 110, la etapa de síntesis predictiva 220, respectivamente. Las realizaciones pueden actualizar todas las memorias y estados involucrados para permitir la reducción de una señal sintetizada lo más cercana posible a la señal original, y reducir las distorsiones como se muestra en las Figuras 7g y 8g. Además, en las realizaciones puede posibilitarse periodos de

solapamiento y adición más largos dada la introducción mejorada del solapamiento del dominio del tiempo y de la cancelación del solapamiento del dominio del tiempo.

5 Como ya se ha descrito anteriormente, el uso de una ventana rectangular al comienzo de la primera trama o la trama LPD actual y el reajuste del códec basado en LPD a un estado de cero, puede no ser la opción ideal para transiciones. Pueden tener lugar distorsiones y artefactos, dado que puede no quede tiempo suficiente para que el códec LPD genere una buena señal. Consideraciones similares se aplican al ajuste de las variables del estado interno del códec a cualquier valor inicial definido, dado que un estado fijo de un codificador de este tipo depende de propiedades de señales múltiples, y que los tiempos de arranque de cualquier estado inicial predefinido, pero fijo, pueden ser largos.

15 En las realizaciones del codificador de audio 100, el controlador 140 puede adaptarse para determinar información de los coeficientes para un filtro de síntesis y una información de una trama de dominio de predicción de conmutación basándose en un análisis de LPC. En otras palabras, las realizaciones pueden utilizar una ventana rectangular y resetear el estado interno del códec LPD. En algunas realizaciones, el codificador puede incluir información de las memorias de filtro y/o un libro de códigos adaptativos utilizado por ACELP, acerca de muestras de síntesis de la trama no LPD previa en las tramas codificadas y proporcionarlos al decodificador. En otras palabras, las realizaciones del codificador de audio 100 pueden decodificar la trama no LPD previa, realizar un análisis de LPC, y aplicar el filtro de análisis de LPC a la señal de síntesis no LPD para proporcionar información de los mismos al decodificador.

25 Como ya se ha mencionado anteriormente, el controlador 140 puede adaptarse para determinar la información sobre el coeficiente de conmutación en un modo tal que dicha información puede representar una trama de muestras de audio que se superponen a la trama previa.

30 En las realizaciones, el codificador de audio 100 puede adaptarse para codificar tal información de coeficientes de conmutación por el uso del codificador reductor de la redundancia 150. Como parte de una realización, el procedimiento de reinicio puede mejorarse transmitiendo o incluyendo información de parámetros adicionales calculada por LPC sobre la trama previa en el flujo de bits. A continuación, el conjunto adicional de coeficientes LPC puede denominarse como LPC0.

35 En una realización, el códec puede operar en su modo de codificación principal de LPD, utilizando cuatro filtros LPC, en concreto LPC1 a LPC4, que se estiman o determinan para cada trama. En una realización, en transiciones de codificación no LPD a codificación LPD, puede también determinarse o estimarse un filtro LPC adicional LPC0, que puede corresponder a un análisis de LPC centrado al final de la trama previa. En otras palabras, en una realización, la trama de muestras de audio que superponen la trama previa puede centrarse al final de la trama previa.

40 En ejemplos del decodificador de audio 200, el decodificador recuperador de la redundancia 210 puede adaptarse para decodificar una información sobre el coeficiente de conmutación de las tramas codificadas. Por consiguiente, la etapa de síntesis predictiva 220 puede adaptarse para determinar una trama predicha de conmutación que se superpone a la trama previa. En otra realización, la trama predicha de conmutación puede centrarse al final de la trama previa.

45 En las realizaciones, el filtro LPC que corresponde al final del segmento o trama no LPD, es decir, LPC0, puede utilizarse para la interpolación de los coeficientes LPC o para el cálculo de la respuesta a la entrada cero en caso de una ACELP.

50 Como se ha mencionado anteriormente, este filtro LPC puede estimarse de una manera hacia delante, es decir, estimarse basándose en la señal de entrada, cuantificarse por el codificador y transmitirse al decodificador. En otras realizaciones, el filtro LPC puede estimarse de una manera hacia atrás, es decir, por el decodificador basándose en la señal sintetizada pasada. La estimación hacia delante puede utilizar velocidades de bits adicionales pero también puede permitir un período de arranque más eficiente y fiable.

55 En otras palabras, en otras realizaciones el controlador 250 en una realización del decodificador de audio 200 puede adaptarse para analizar la trama previa para obtener información de la trama previa sobre los coeficientes para un filtro de síntesis y/o una información de la trama previa sobre una trama del dominio de la predicción. El controlador 250 puede además adaptarse para proporcionar la información de la trama previa sobre los coeficientes a la etapa de síntesis predictiva 220 como coeficientes de conmutación. El controlador 250 puede proporcionar además la información de la trama previa sobre la trama del dominio de predicción a la etapa de síntesis predictiva 220, para entrenamiento.

60 En las realizaciones en las que el codificador de audio 100 proporciona información sobre los coeficientes de conmutación, la cantidad de bits en el flujo de bits puede aumentar ligeramente. Llevar a cabo el análisis en el decodificador puede no aumentar la cantidad de bits en el flujo de bits. Sin embargo, llevar a cabo el análisis en el

5 decodificador puede introducir complejidad extra. Por lo tanto, en las realizaciones, la resolución del análisis de LPC puede mejorarse reduciendo la dinámica espectral, es decir, las tramas de la señal pueden pre-procesarse en primer lugar a través de un filtro pre-énfasis. El énfasis de baja frecuencia inversa puede aplicarse a la realización del decodificador 200, así como en el codificador de audio 100 para permitir la obtención de una señal de excitación o trama de dominio de predicción necesaria para la codificación de las próximas tramas. Todos estos filtros pueden dar una respuesta de estado de cero, es decir, la salida de un filtro causada por la entrada actual, dado que no se han aplicado entradas pasadas, es decir, dado que la información de estado en el filtro se establece a cero después de un reseteo completo. En general, cuando se ejecuta con normalidad el modo de codificación LPD, la información de estado en el filtro se actualiza por el estado final tras el filtrado de la trama previa. En las realizaciones, para establecer el primer estado de filtro interno del codificado por LPD en un modo que, ya para la primera trama LPD, todos los filtros y predictores se inicien para ejecutarse en el modo óptimo o mejorado para la primera trama, cualquier información sobre el coeficiente/coeficientes de conmutación puede proporcionarse por el codificador de audio 100, o puede llevarse a cabo procesamiento adicional en un decodificador 200.

15 En general, los filtros y predictores para el análisis, como se lleva a cabo en el codificador de audio 100 por la etapa de análisis de codificación predictiva 110, se distinguen de los filtros y predictores utilizados en el lado del decodificador de audio 200 para la síntesis.

20 Para el análisis, como por ejemplo la etapa de análisis de codificación predictiva 110, todos o al menos uno de estos filtros pueden alimentarse con las muestras originales apropiadas de la trama previa para actualizar las memorias. La Figura 9a ilustra una realización de una estructura de filtro utilizada para el análisis. El primer filtro es un filtro pre-énfasis 1002, que puede utilizarse para mejorar la resolución del filtro de análisis de LPC 1006, es decir, la etapa de análisis de codificación predictiva 110. En las realizaciones, el filtro de análisis de LPC 1006 puede calcular o evaluar los coeficientes del filtro a corto plazo por el uso, por ejemplo, de muestras del habla filtradas de paso alto en la ventana de análisis. En otras palabras, de acuerdo con la invención, el controlador 140 está adaptado para determinar la información sobre el coeficiente de conmutación basándose en una versión filtrada de paso alto de un espectro de tramas decodificadas de la trama previa. De una manera similar, suponiendo que el análisis se lleva a cabo en la realización del decodificador de audio 200, el controlador 250 está adaptado para analizar una versión filtrada de paso alto de la trama previa.

30 Como se ilustra en la Figura 9a, el filtro de análisis LP 1006 se precede por un filtro de ponderación perceptual 1004. En las realizaciones, el filtro de ponderación perceptual 1004 puede emplearse en la búsqueda de análisis por síntesis de libros de códigos. El filtro puede aprovechar las propiedades de enmascarado del ruido de los formantes, como por ejemplo las resonancias del tracto vocal, ponderando los errores menos en regiones cercanas a las frecuencias formantes y más en las regiones distantes a ellas. En las realizaciones, el codificador reductor de la redundancia 150 puede adaptarse para codificar basándose en un libro de códigos que es adaptativo a la trama/tramas de dominio de predicción respectivo. Correspondientemente, el decodificador introductor de la redundancia 210 puede adaptarse para decodificar basándose en un libro de códigos que esté adaptado a las muestras de las tramas.

40 La Figura 9b ilustra un diagrama de bloques del procesamiento de señales en el caso de síntesis. En el caso de síntesis, en las realizaciones, todos o al menos uno de los filtros pueden alimentarse con las muestras sintetizadas apropiadas de la trama previa para actualizar las memorias. En las realizaciones del decodificador de audio 200, éste puede ser fácil dado que la síntesis de la trama no LPD previa está directamente disponible. Sin embargo, en una realización del codificador de audio 100, la síntesis puede no llevarse a cabo por defecto y, correspondientemente, las muestras sintetizadas pueden no estar disponibles. Por lo tanto, en las realizaciones del codificador de audio 100, el controlador 140 puede adaptarse para decodificar la trama no LPD previa. Una vez que se ha decodificado la trama no LPD, en ambas realizaciones, es decir, el codificador de audio 100 y el codificador de audio 200, puede llevarse a cabo la síntesis de la trama previa de acuerdo con la Figura 9b en el bloque 1012. Además, la salida del filtro de síntesis LP 1012 puede introducirse en un filtro de ponderación perceptual inverso 1014, después de lo que se aplica un filtro de des-énfasis 1016. En las realizaciones, puede utilizarse un libro de códigos adaptado y completarse con las muestras sintetizadas de la trama previa. En realizaciones adicionales, el libro de códigos adaptativo puede contener vectores de excitación que se adaptan para cada sub-trama. El libro de códigos adaptativo puede deducirse del estado de filtro a largo plazo. Puede utilizarse un valor de retardo como un índice en el libro de códigos adaptativo. En las realizaciones, para completar el libro de códigos adaptativo, puede finalmente calcularse la señal de excitación o señal residual por el filtrado de la señal ponderada cuantificada al filtro de ponderación inverso con una memoria en cero. En particular, la excitación puede ser necesaria en el codificador 100 para actualizar la memoria del predictor a largo plazo.

60 Las realizaciones de la presente invención pueden proporcionar la ventaja de que un procedimiento de reinicio de filtros puede potenciarse o acerrarse proporcionando parámetros adicionales y/o alimentando las memorias internas de un codificador o decodificador con muestras de la trama previa codificada por el codificador basado en transformadas.

- Las realizaciones pueden proporcionar la ventaja de una aceleración del procedimiento de inicio de un códec de principal de LPD actualizando todas o partes de las memorias involucradas, dando como resultado una señal sintetizada, que puede estar más cerca de la señal original que cuando se utilizan conceptos convencionales, en especial cuando se utiliza un reseteo completo. Además, las realizaciones pueden permitir una ventana de superposición y adición más larga y con ello posibilitar el uso mejorado de la cancelación del solapamiento del dominio del tiempo. Las realizaciones pueden proporcionar la ventaja de que una fase no fija de un codificador del habla puede acortarse, y que los artefactos producidos durante la transición de un codificador basado en transformadas a un codificador del habla puedan reducirse.
- 5
- 10 Dependiendo de ciertos requerimientos de implementación de los métodos inventivos, los métodos inventivos pueden implementarse en hardware o en software. La implementación puede realizarse utilizando un medio de almacenamiento digital, en particular un disco, un DVD, un CD, que tiene señales de control electrónicamente legibles almacenadas en el mismo, que coopera (o pueda cooperar) con un sistema informático programable de modo que se realicen los métodos respectivos.
- 15 En general, la presente invención es, por lo tanto, un producto de un programa informático con un código de programación almacenado en un soporte legible por máquina, siendo el código de programación operativo para realizar uno de los métodos cuando el producto de un programa informático se ejecuta en un ordenador.
- 20 En otras palabras, los métodos inventivos son, por lo tanto, un programa informático que tiene un código de programación para realizar al menos uno de los métodos inventivos cuando el programa informático se ejecuta en un ordenador.
- 25 Aunque anterior se ha mostrado y descrito particularmente con referencia a realizaciones particulares de lo mismo, se ha de entender por los expertos en la materia que pueden realizarse diversos otros cambios en la forma y detalles, sin alejarse del alcance de los mismos. Se ha de entender que pueden realizarse diversos cambios al adaptarse a diferentes realizaciones sin alejarse de los conceptos más amplios desvelados en el presente documento y comprendidos mediante las reivindicaciones que siguen.

**REIVINDICACIONES**

1. Un codificador de audio (100) adaptado para codificar tramas de una señal de audio muestreada para obtener tramas codificadas, en el que una trama comprende un número de muestras de audio del dominio del tiempo, que comprende:
- 5 una etapa de análisis de codificación predictiva (110) para determinar información sobre los coeficientes de un filtro de síntesis e información sobre una trama de dominio de predicción basándose en una trama de muestras de audio;
- 10 un transformador del dominio de la frecuencia (120) para transformar una trama de muestras de audio al dominio de la frecuencia para obtener un espectro de tramas;
- un tomador de decisiones del dominio de codificación (130) para decidir si los datos codificados para una trama están basados en la información sobre los coeficientes y en la información sobre la trama de dominio de predicción, o están basados en el espectro de tramas;
- 15 un controlador (140) para determinar la información sobre un coeficiente de conmutación cuando el tomador de decisiones del dominio de codificación decide que los datos codificados de una trama actual están basados en la información sobre los coeficientes y la información sobre la trama de dominio de predicción cuando los datos codificados de una trama previa se codificaron basándose en un espectro de tramas previo obtenido mediante el transformador del dominio de la frecuencia; y
- 20 un codificador reductor de la redundancia (150) para codificar la información sobre la trama de dominio de predicción, la información sobre los coeficientes, la información sobre el coeficiente de conmutación y/o el espectro de tramas,
- en el que la información sobre los coeficientes de conmutación comprende una información que posibilita una iniciación de una etapa de síntesis predictiva, y
- 25 el controlador (140) está adaptado para determinar la información sobre el coeficiente de conmutación basándose en una versión filtrada de paso alto de un espectro de trama decodificado de la trama previa.
2. El codificador de audio (100) de la reivindicación 1, en el que la etapa de análisis de codificación predictiva (110) está adaptada para determinar la información sobre los coeficientes del filtro de síntesis y la información sobre la trama de dominio de predicción basándose en un análisis de Codificación de Predicción Lineal, LPC, y/o en el que el transformador del dominio de la frecuencia (120) está adaptado para transformar la trama de muestras de audio basándose en una Transformada Rápida de Fourier, FFT, o en una transformada de coseno discreta modificada, MDCT.
- 30 3. Un método para codificar tramas de una señal de audio muestreada para obtener tramas codificadas, en el que una trama comprende un número de muestras de audio del dominio del tiempo, que comprende las etapas de:
- determinar información sobre los coeficientes de un filtro de síntesis e información sobre una trama de dominio de predicción basándose en una trama de muestras de audio;
- 40 transformar una trama de muestras de audio al dominio de frecuencia para obtener un espectro de tramas;
- decidir si los datos codificados para una trama están basados en la información sobre los coeficientes y en la información sobre la trama de dominio de predicción, o están basados en el espectro de tramas;
- determinar la información sobre un coeficiente de conmutación cuando se decide que los datos codificados de una trama actual están basados en la información sobre los coeficientes y la información sobre la trama de dominio de predicción cuando los datos codificados de una trama previa se han codificado basándose en un espectro de tramas previo obtenido mediante la transformación del dominio de la frecuencia; y
- 45 codificar la información sobre la trama de dominio de predicción, la información sobre los coeficientes, la información sobre el coeficiente de conmutación y/o los espectros de tramas, en el que la información sobre el coeficiente de conmutación comprende una información que posibilita una iniciación de una etapa de síntesis predictiva, y
- 50 la determinación está basada en una versión filtrada de paso alto de un espectro de trama decodificado de la trama previa.
4. Un decodificador de audio (200) para decodificar tramas codificadas para obtener tramas de una señal de audio muestreada, en el que una trama comprende un número de muestras de audio del dominio del tiempo, que comprende:
- 55 un decodificador recuperador de la redundancia (210) para decodificar las tramas codificadas para obtener información sobre una trama de dominio de predicción, información sobre los coeficientes para un filtro de síntesis y/o un espectro de tramas;
- 60 una etapa de síntesis predictiva (220) para determinar una trama predicha de muestras de audio basándose en la información sobre los coeficientes para el filtro de síntesis y la información sobre la trama de dominio de predicción;
- un transformador del dominio del tiempo (230) para transformar el espectro de tramas al dominio del tiempo para

obtener una trama transformada a partir del espectro de tramas;  
 un combinador (240) para combinar la trama transformada y la trama predicha para obtener las tramas de la  
 señal de audio muestreada; y  
 un controlador (250) para controlar un proceso de conmutación, efectuándose el proceso de conmutación cuando  
 una trama previa está basada en una trama transformada y una trama actual está basada en una trama predicha,  
 estando configurado el controlador (250) para proporcionar un coeficiente de conmutación a la etapa de síntesis  
 predictiva (220) para iniciación de la etapa de síntesis predictiva (220) realizando un análisis de LPC de una  
 versión filtrada de paso alto de la trama previa de modo que la etapa de síntesis predictiva (220) se inicia cuando  
 se efectúa el proceso de conmutación.

5. El decodificador de audio (200) de la reivindicación 4, en el que la etapa de síntesis predictiva (220) está adaptada  
 para determinar la trama predictiva basándose en una síntesis LPC y/o en el que el transformador del dominio del  
 tiempo (230) está adaptado para transformar el espectro de tramas al dominio del tiempo basándose en una FFT  
 inversa o una MDCT inversa.

6. El decodificador de audio (200) de la reivindicación 4 o 5, en el que el controlador (250) está adaptado para  
 analizar la trama previa para obtener una información de trama previa sobre los coeficientes para un filtro de síntesis  
 y una información de trama previa sobre una trama de dominio de predicción y en el que el controlador (250) está  
 adaptado para proporcionar la información de trama previa basándose en los coeficientes a la etapa de síntesis  
 predictiva (220) como coeficiente de conmutación y/o en el que el controlador (250) está adaptado para proporcionar  
 además la información de trama previa sobre la trama de dominio de predicción a la etapa de síntesis predictiva  
 (220) para iniciación.

7. El decodificador de audio (200) de una de las reivindicaciones 4-6, en el que la etapa de síntesis predictiva (220)  
 está adaptada para determinar una trama de predicción de conmutación que está centrada en el final de la trama  
 previa.

8. Un método para decodificar tramas codificadas para obtener tramas de una señal de audio muestreada, en el que  
 una trama comprende un número de muestras de audio del dominio del tiempo, que comprende las etapas de:

decodificar las tramas codificadas para obtener información sobre una trama de dominio de predicción, e  
 información sobre los coeficientes para un filtro de síntesis y/o un espectro de tramas;  
 determinar una trama predicha de muestras de audio basándose en la información sobre los coeficientes para el  
 filtro de síntesis y la información sobre la trama de dominio de predicción;  
 transformar el espectro de tramas al dominio del tiempo para obtener una trama transformada a partir del  
 espectro de tramas;  
 combinar la trama transformada y la trama predicha para obtener las tramas de la señal de audio muestreada; y  
 controlar un proceso de conmutación, efectuándose el proceso de conmutación cuando una trama previa está  
 basada en la trama transformada y una trama actual está basada en la trama predicha;  
 proporcionar un coeficiente de conmutación para iniciación realizando un análisis LPC de una versión filtrada de  
 paso alto de la trama previa de modo que una etapa de síntesis predictiva se inicia cuando se efectúa el proceso  
 de conmutación.

9. Programa informático que tiene un código de programación adaptado para realizar uno de los métodos de las  
 reivindicaciones 3 u 8, cuando un programa informático se ejecuta en un ordenador o procesador.

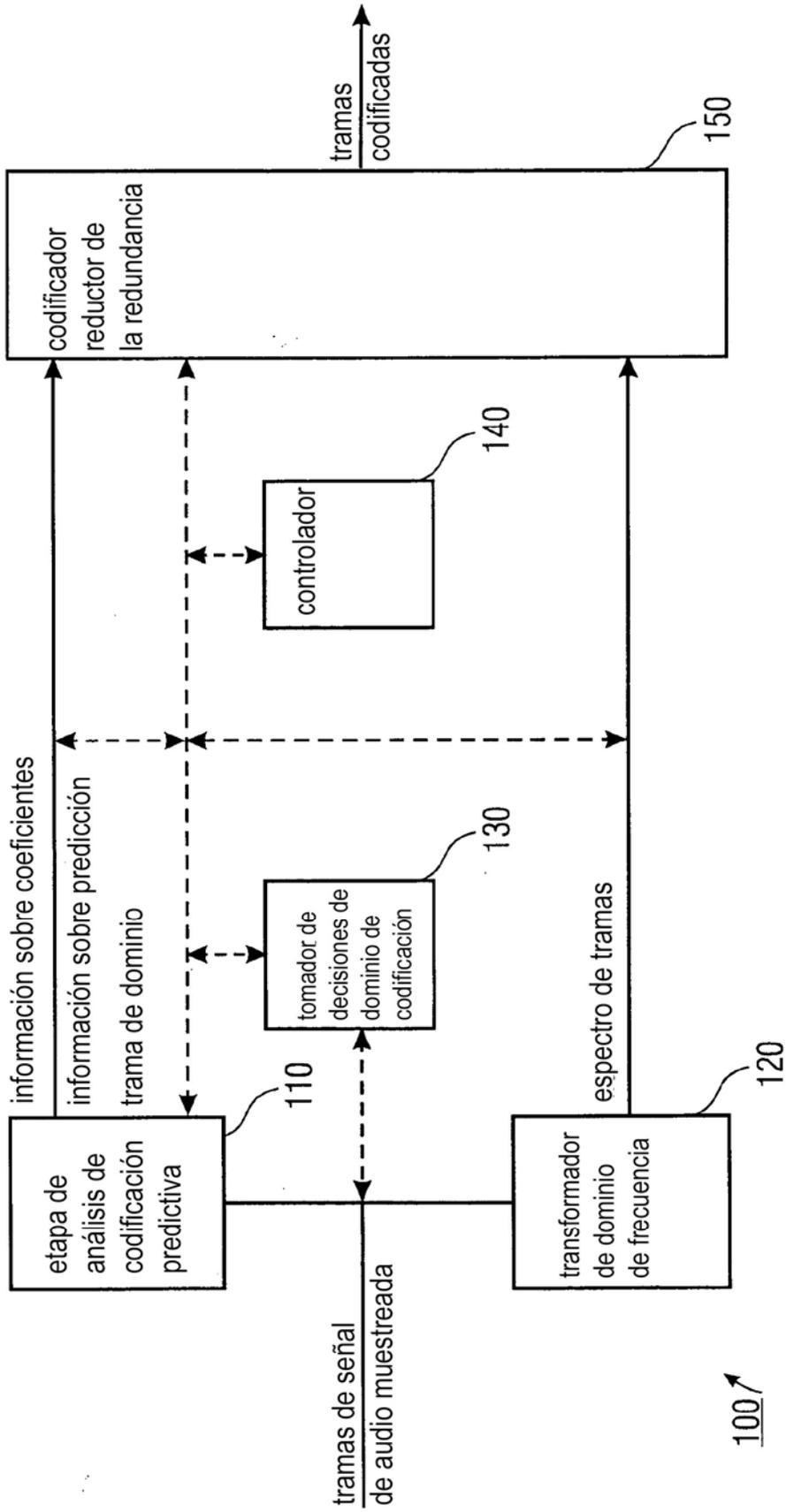


FIG 1

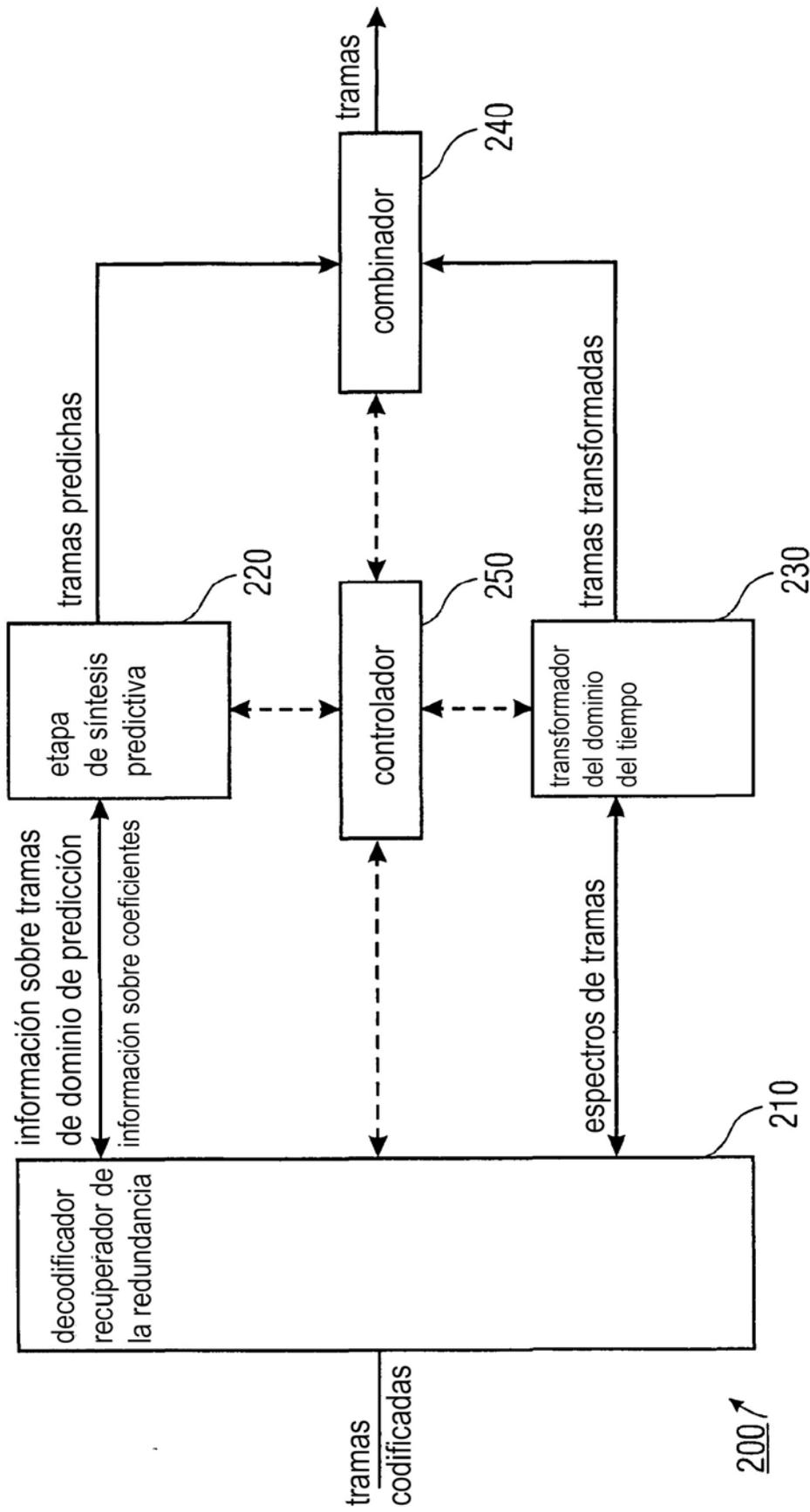


FIG 2

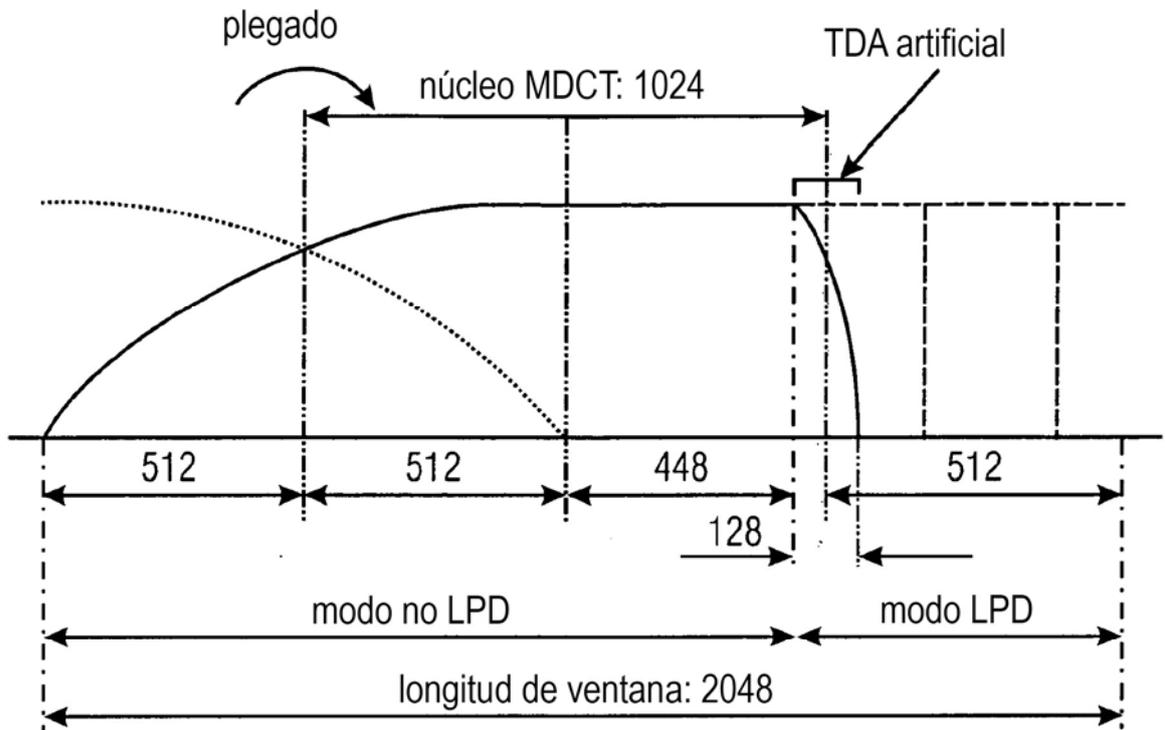
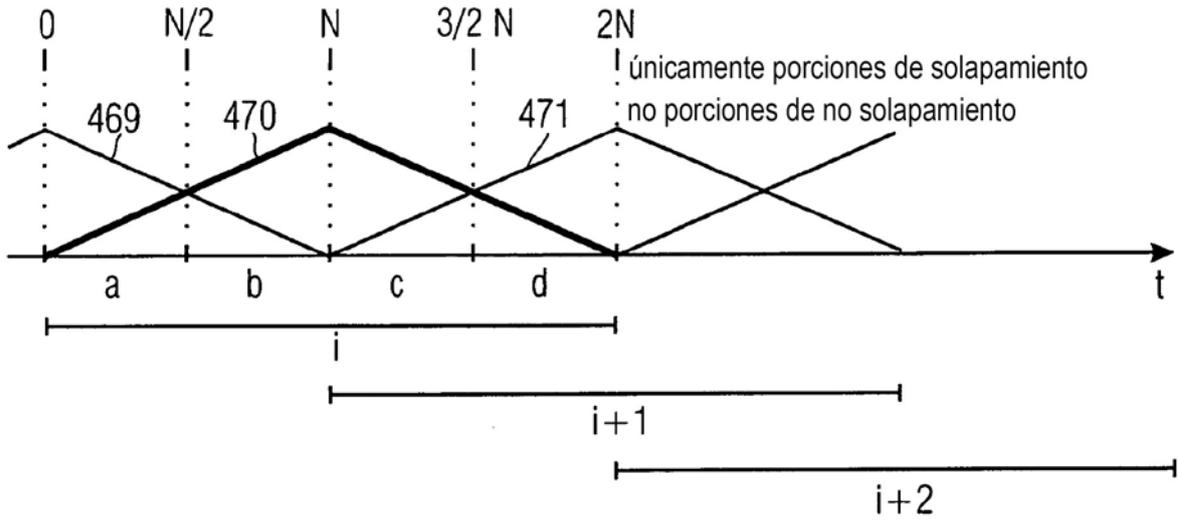


FIG 3



plegado:  $(-c_R - d, a - b_R)$  R: operador inverso

(2N val. entrada



N val. salida)

desplegado:  $(a - b_R, b - a_R, c + d_R, d + c_R) \cdot 1/2$

(N val. entrada



2N val. salida)  $\text{MDCT}(a, b, c, d) \equiv \text{DCT-IV}(-c_R - d, a - b_R)$

2N valores de salida
N valores de entrada

$$\underbrace{\text{IMDCT}(\text{MDCT}(a, b, c, d))}_{2N \text{ valores de salida}} = (a - b_R, b - a_R, c + d_R, d + c_R) 1/2$$

FIG 4A

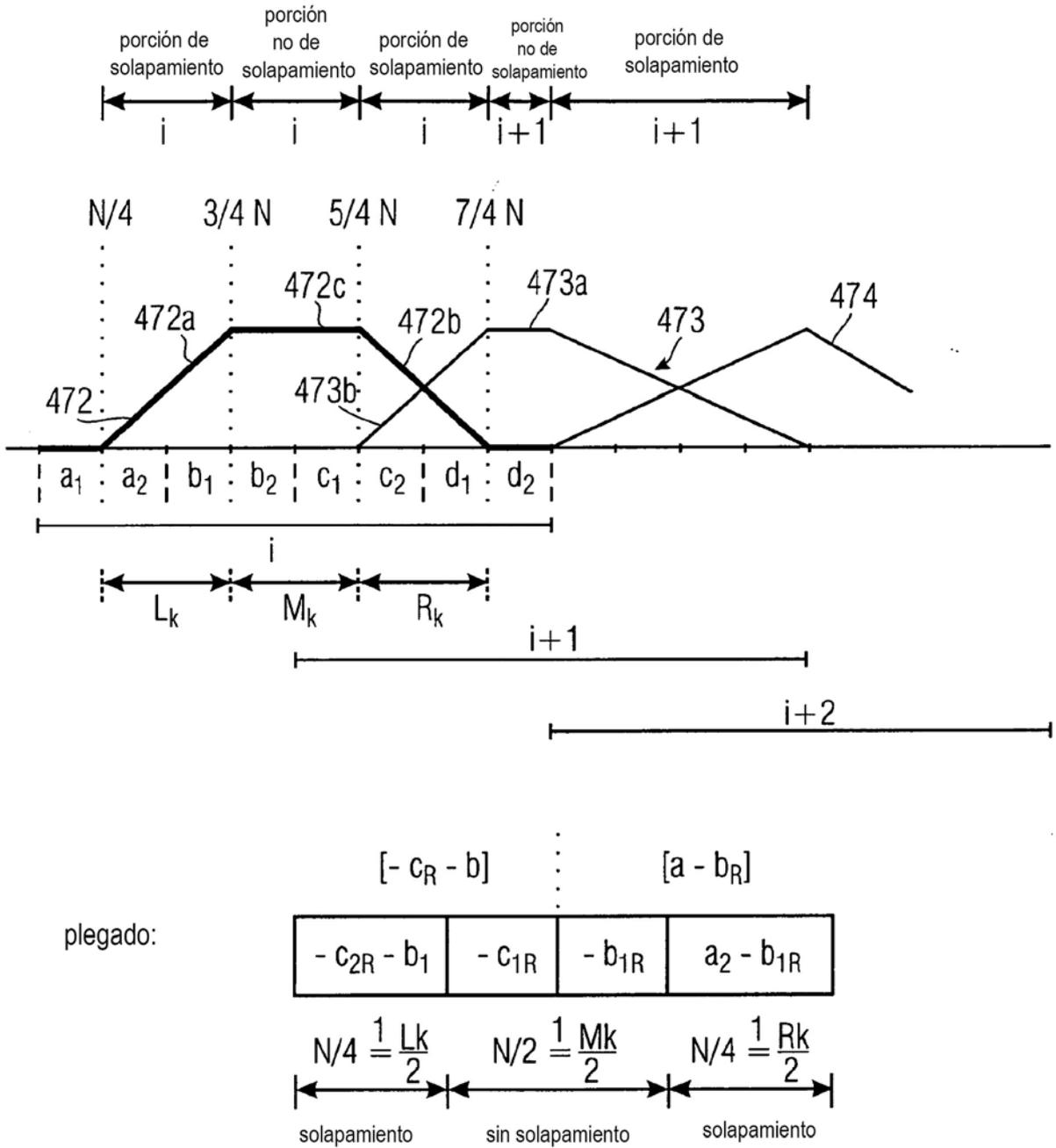


FIG 4B

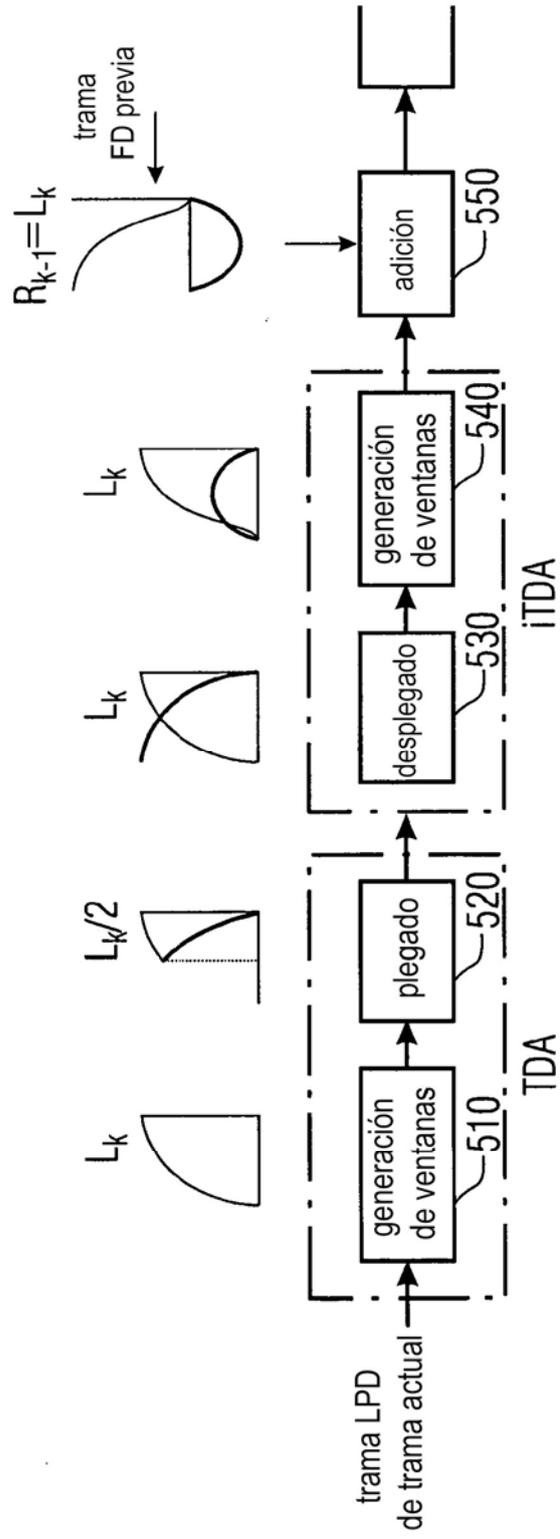


FIG 5

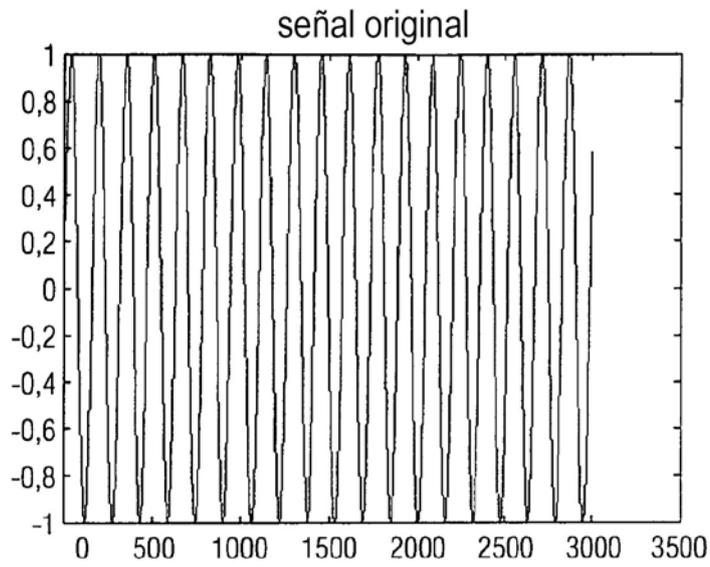


FIG 6A

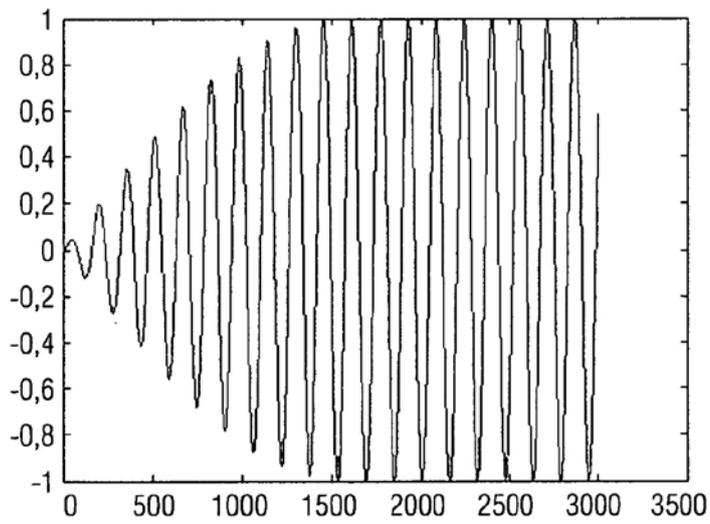


FIG 6B

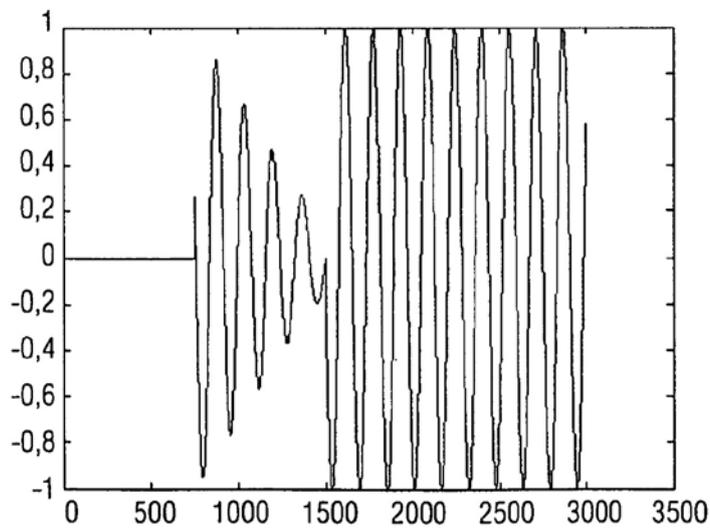


FIG 6C

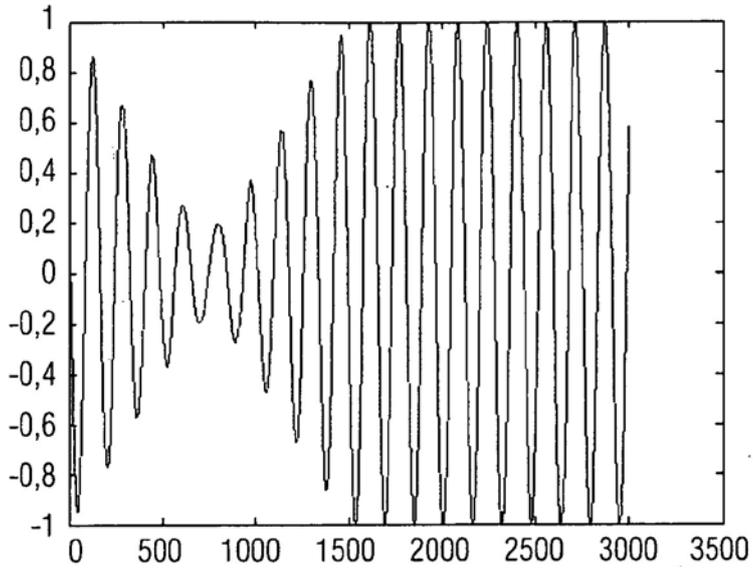


FIG 6D

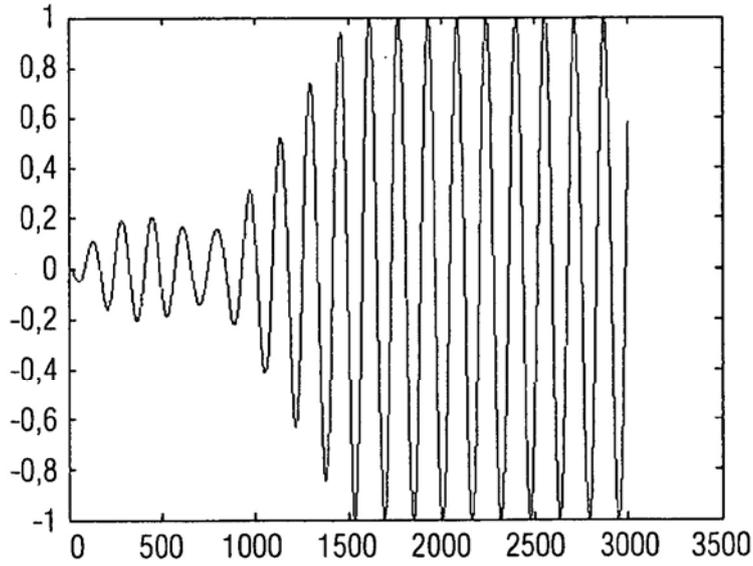


FIG 6E

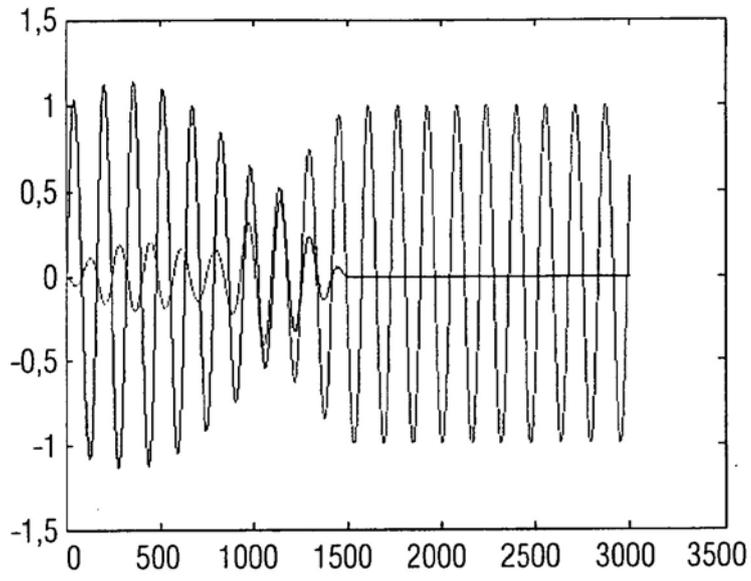


FIG 6F

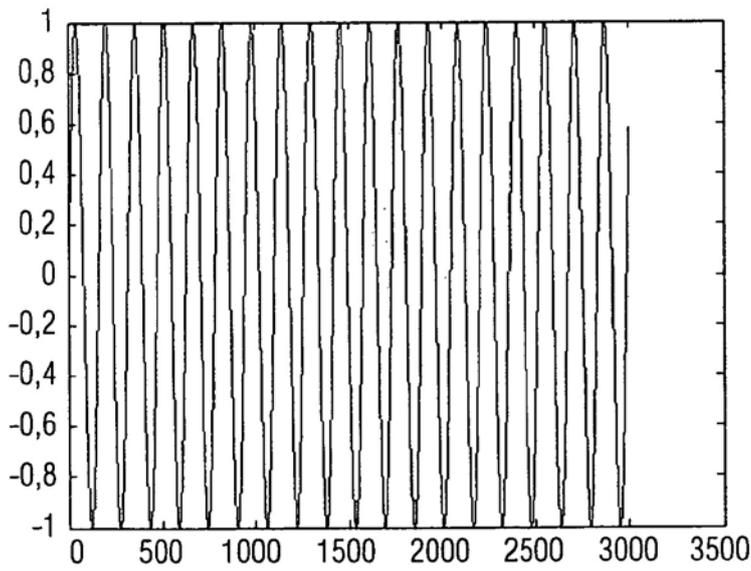


FIG 6G

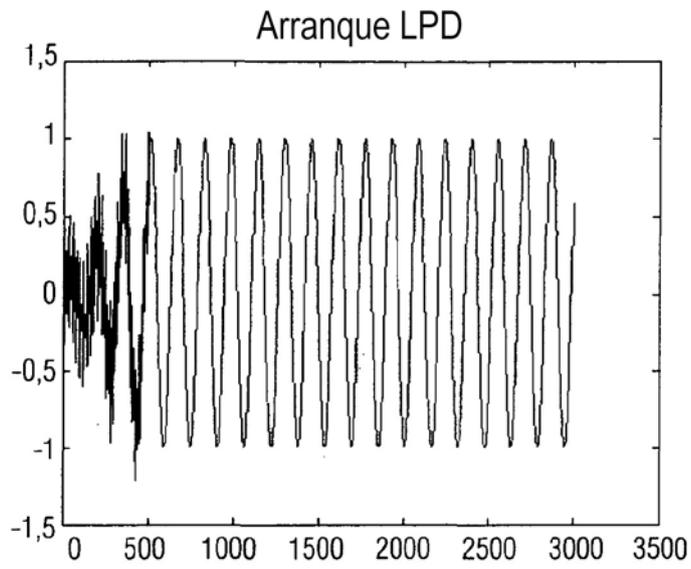


FIG 7A

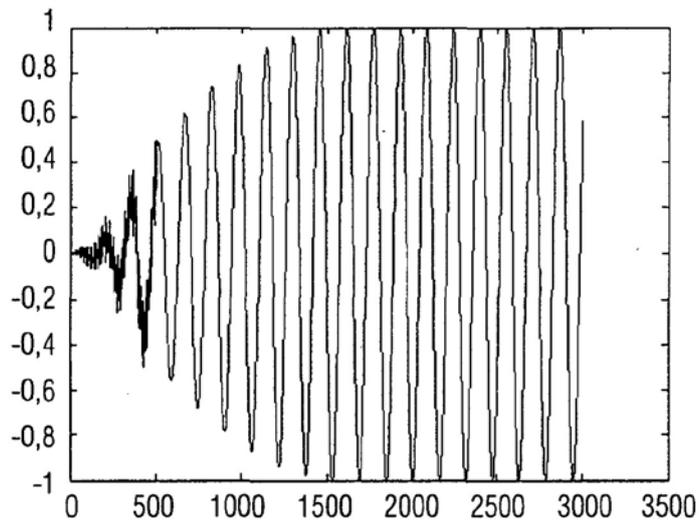


FIG 7B

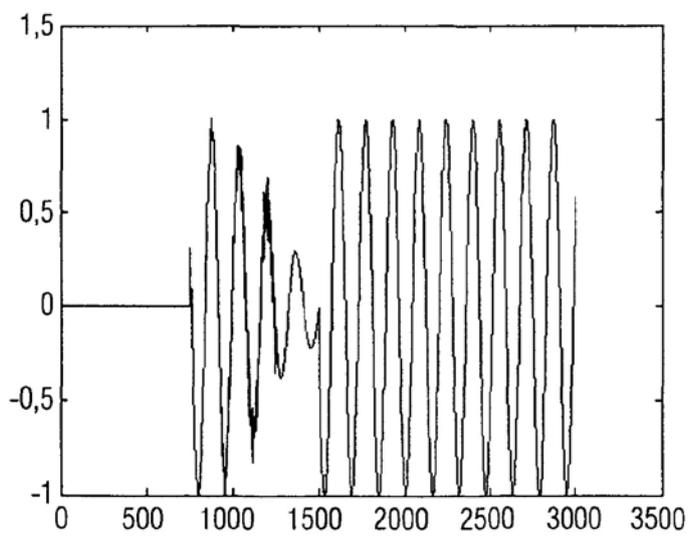


FIG 7C

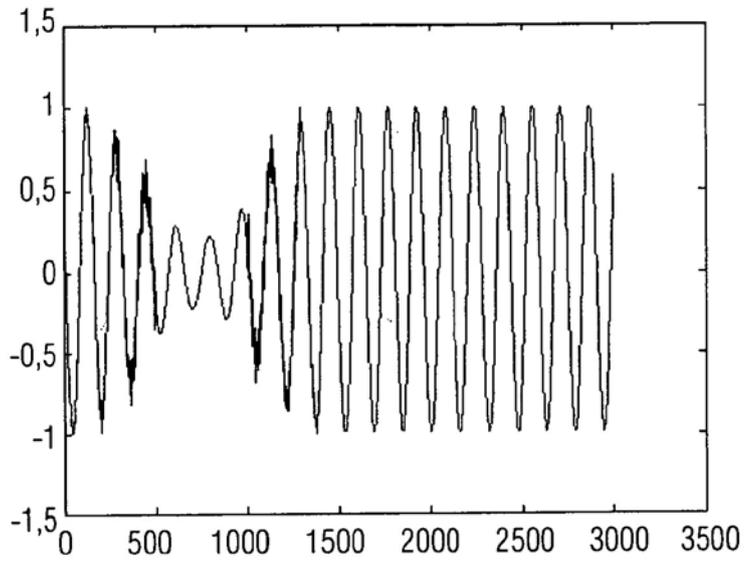


FIG 7D

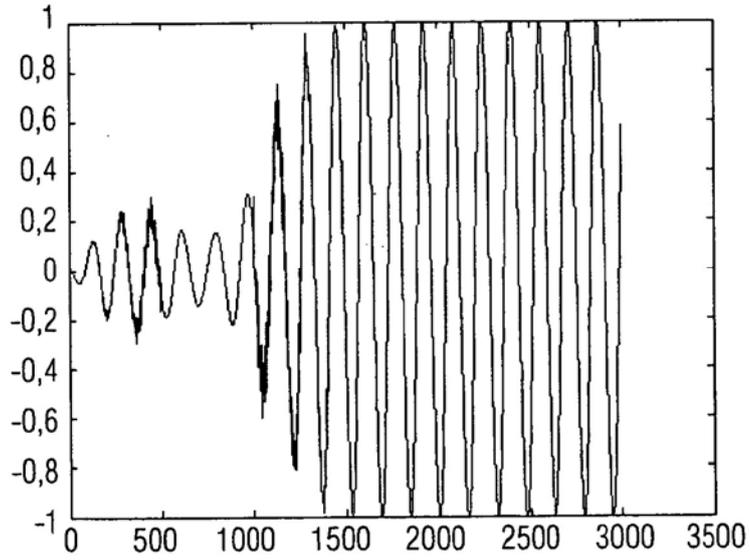


FIG 7E

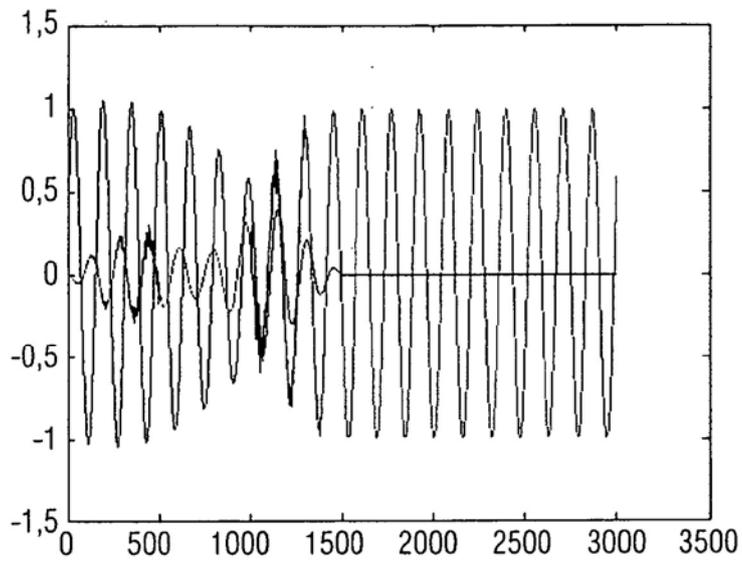


FIG 7F

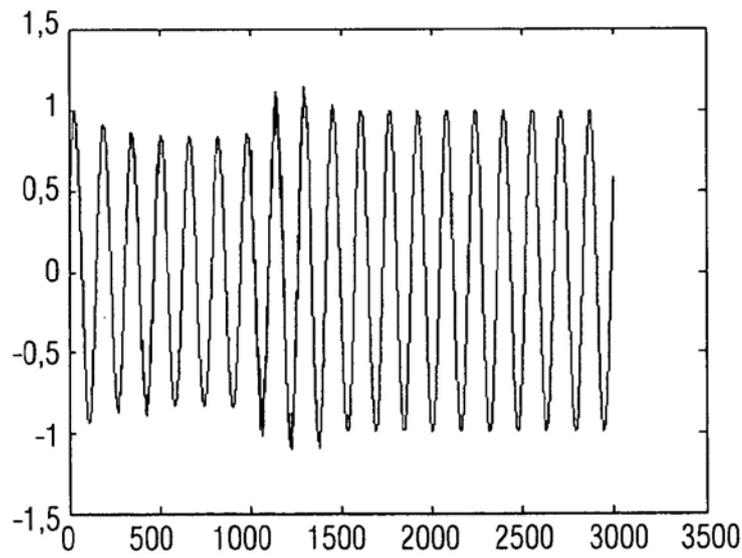


FIG 7G

Arranque LPD

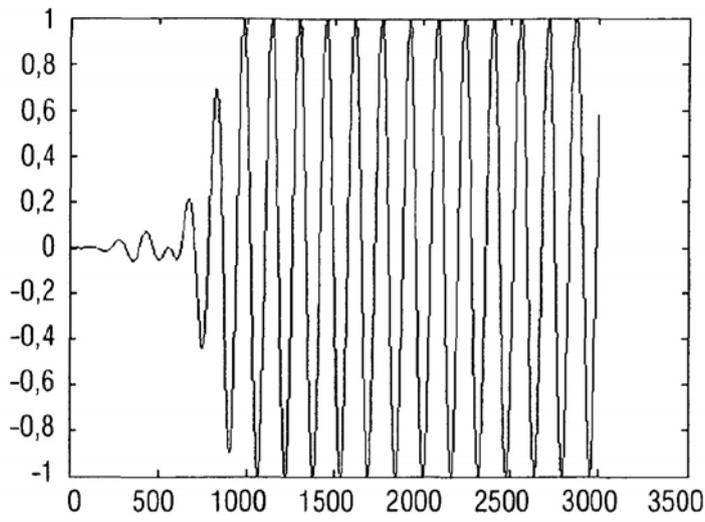


FIG 8A

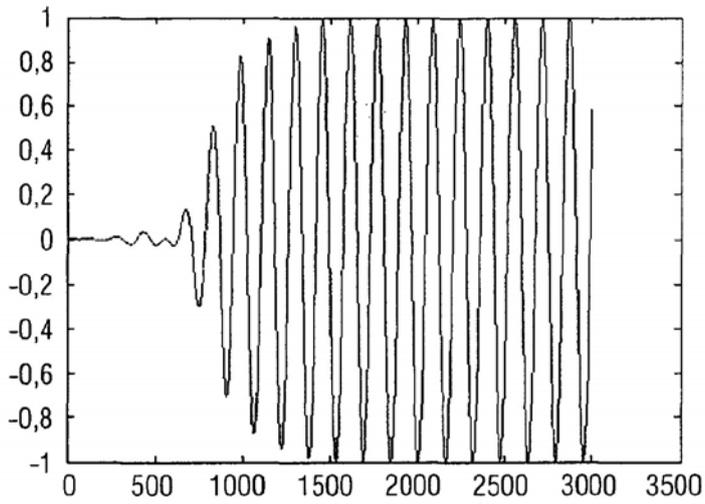


FIG 8B

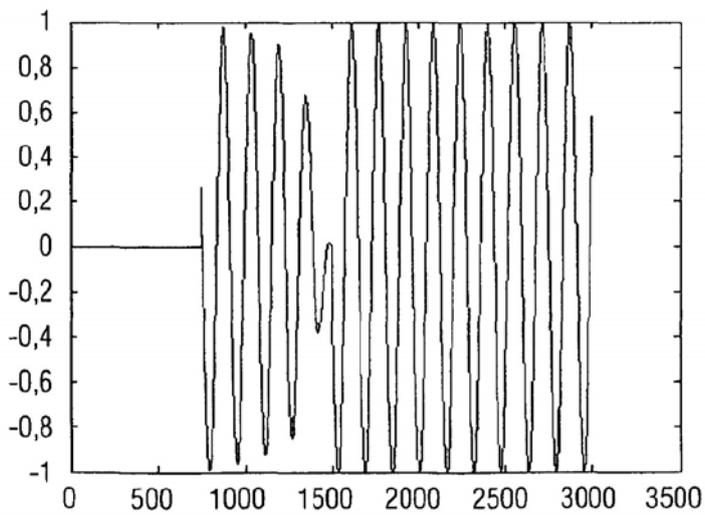


FIG 8C

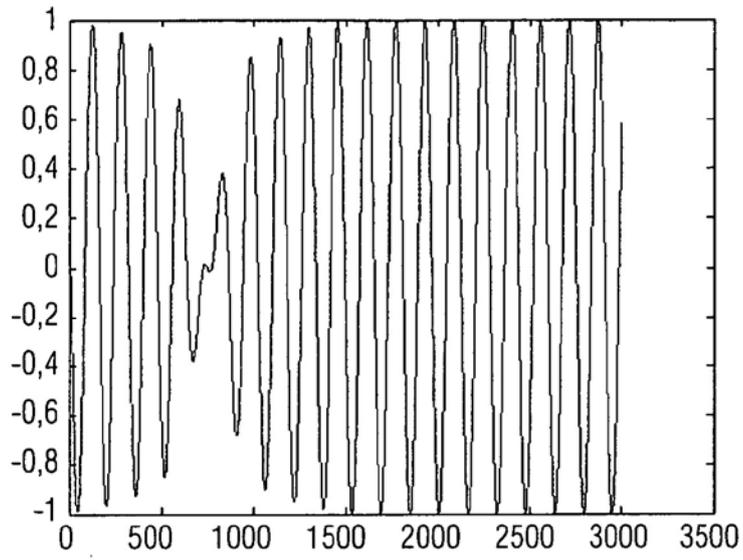


FIG 8D

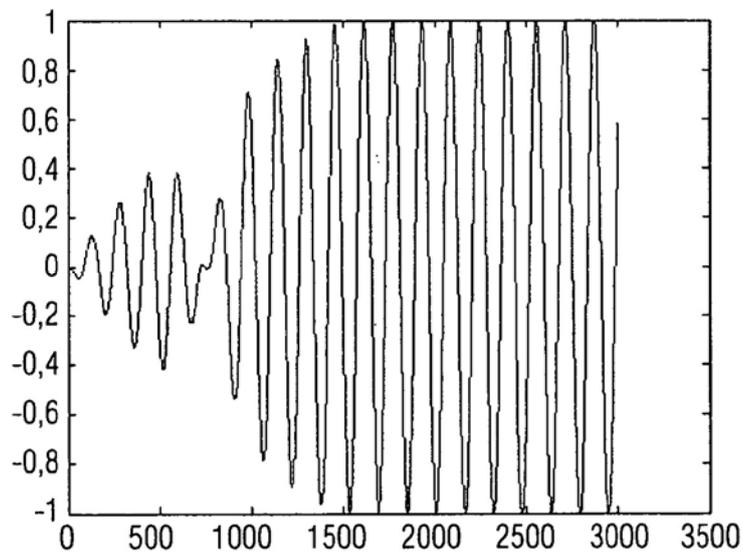


FIG 8E

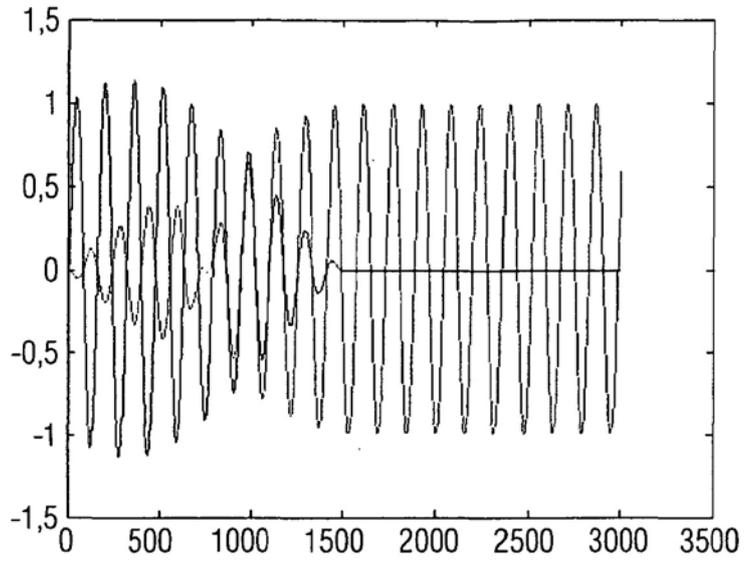


FIG 8F

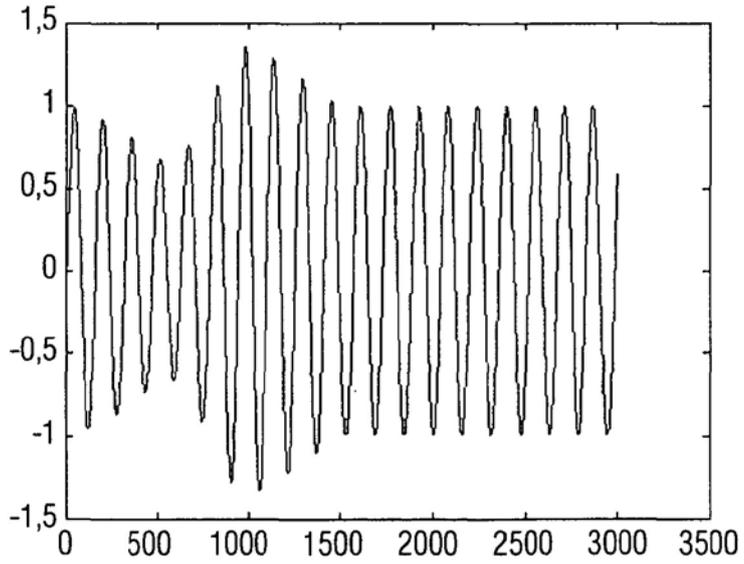


FIG 8G

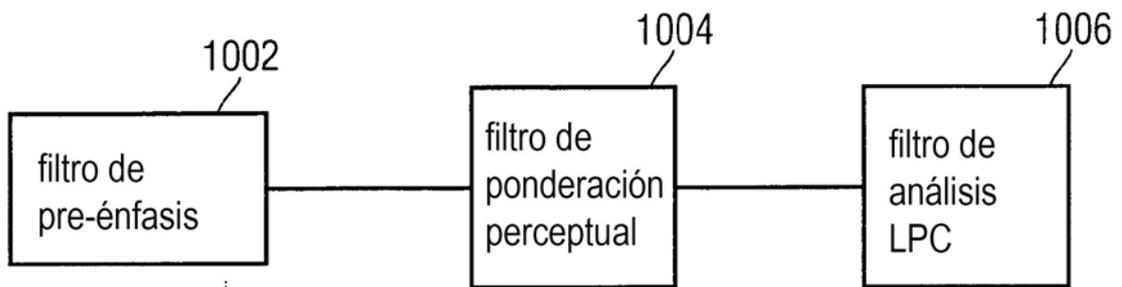


FIG 9A

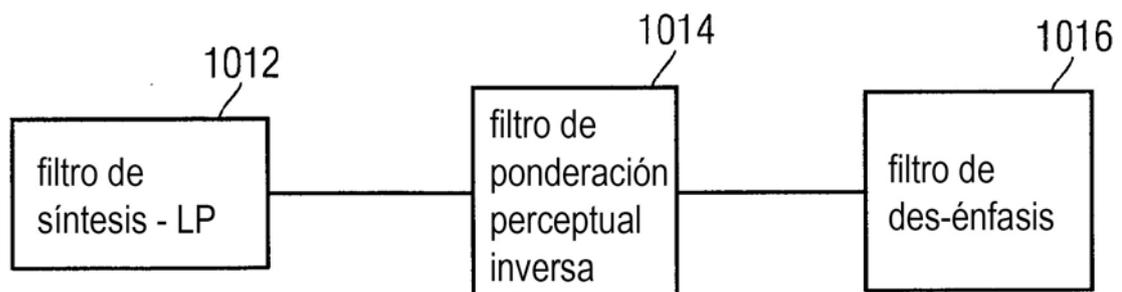


FIG 9B