

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 645 037**

51 Int. Cl.:

**G10L 19/022** (2013.01)

G10L 19/04 (2013.01)

G10L 19/20 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.01.2010 E 10152001 (3)**

97 Fecha y número de publicación de la concesión europea: **02.08.2017 EP 2214164**

54 Título: **Decodificador de audio, procedimientos para decodificar una señal de audio y programa informático**

30 Prioridad:

**28.01.2009 US 147895 P**

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

**01.12.2017**

73 Titular/es:

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

72 Inventor/es:

**NEUENDORF, MAX;  
LECOMTE, JÉRÉMIE;  
MULTRUS, MARKUS;  
BAYER, STEFAN;  
NAGEL, FREDERIK;  
FUCHS, GUILLAUME;  
ROBILLIARD, JULIEN;  
RETTTELBACH, NIKOLAUS;  
GEIGER, RALF y  
GRILL, BERNHARD**

74 Agente/Representante:

**SALVA FERRER, Joan**

ES 2 645 037 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

Decodificador de audio, procedimientos para decodificar una señal de audio y programa informático.

## 5 ANTECEDENTES DE LA INVENCION

- [0001]** Las realizaciones de acuerdo con la presente invención se refieren a un decodificador de audio para proporcionar una información de audio decodificada a partir de una información de audio codificada. Otras realizaciones se refieren a un procedimiento para proporcionar una representación decodificada de un contenido de audio a partir de una representación codificada del contenido de audio. Otras realizaciones de acuerdo con la invención se refieren a un programa informático para llevar a cabo el procedimiento de la invención. Las realizaciones de acuerdo con la invención se refieren a mejoras de una transición desde un modo de dominio de frecuencia a un modo de dominio de predicción lineal.
- [0002]** A continuación, se ofrecerá información sobre los antecedentes de la invención con el fin de facilitar la comprensión de la invención y sus ventajas. Durante la última década, se ha invertido un gran esfuerzo para conseguir que sea posible almacenar digitalmente y distribuir contenidos de audio. En este sentido, un logro importante es el de la definición de la norma internacional ISO/IEC 14496-3. La parte 3 de esta norma se refiere a una codificación y decodificación de contenidos de audio, y el apartado 4 de la parte 3 se refiere a la codificación de audio en general. En el apartado 4 de la parte 3 de ISO/IEC 14496-3, se define un concepto para codificar y decodificar contenidos de audio en general. Además, se han propuesto otras mejoras con el fin de mejorar la calidad y/o reducir la tasa de bits requerida.
- [0003]** De acuerdo con el concepto descrito en dicha norma, una señal de dominio temporal se convierte en una representación de tiempo-frecuencia. La transformada desde el dominio temporal al dominio de tiempo-frecuencia se suele llevar a cabo mediante bloques de transformada, que también se denominan "tramas de audio" o, abreviando, "tramas".
- [0004]** Se ha descubierto que resulta ventajosa la utilización de tramas superpuestas, que están desplazadas, por ejemplo, media trama, debido a que la superposición permite eliminar artefactos eficientemente. Además, se ha descubierto que se debería llevar a cabo un enventanado con el fin de evitar los artefactos originados por el procesamiento de tramas limitadas temporalmente. Además, el enventanado permite la optimización de un proceso de superposición y adición de tramas posteriores desplazadas temporalmente, pero superpuestas.
- [0005]** Además, se han propuesto técnicas para una codificación eficiente de señales de habla. Por ejemplo, se han definido conceptos para una codificación del habla en las normas internacionales 3GPP TS 26.090, 3GPP TS 26.190 y 3GPP TS 26.290. Además, en la bibliografía se han tratado muchos más conceptos para una codificación de señales de habla.
- [0006]** No obstante, se ha descubierto que es difícil combinar los conceptos para la codificación general de audio (como se definen, por ejemplo, en la norma internacional ISO/IEC 14496-3, parte 3, apartado 4) con los conceptos para la codificación del habla (como se definen, por ejemplo, en las normas 3GPP mencionadas anteriormente).
- [0007]** En el documento WO 2010/003563 A1, se describe un codificador de audio para codificar muestras de audio que comprende un primer codificador de dominio temporal que introduce un solapamiento para codificar muestras de audio en el primer dominio de codificación. El primer codificador de dominio temporal que introduce un solapamiento posee una primera regla de generación de tramas, una ventana de inicio y una ventana de parada. El codificador de audio también cuenta con un segundo codificador para codificar muestras en un segundo dominio temporal. El segundo codificador posee una segunda regla de generación de tramas. El codificador de audio también comprende un controlador que pasa del primer codificador al segundo codificador en respuesta a una característica de las muestras de audio, y para modificar la segunda regla de generación como respuesta a la conmutación del primer codificador al segundo codificador o para modificar la ventana de inicio o la ventana de parada del primer codificador, en el que la segunda regla de generación de tramas permanece inalterada.
- [0008]** En el documento WO 2010/003532 A1, se describe un aparato y un procedimiento para codificar/decodificar una señal de audio mediante un esquema de conmutación de solapamiento. El aparato para codificar una señal de audio comprende un enventanador para enventanar un primer bloque de la señal de audio mediante una ventana de análisis que posee una porción de solapamiento y otra porción distinta. El aparato también

comprende un procesador para procesar el primer subbloque de la señal de audio relacionado con la porción de solapamiento mediante la transformación del subbloque pasando de un dominio a un dominio diferente tras enventanar el primer subbloque para obtener el primer subbloque procesado, y para procesar un segundo subbloque de la señal de audio relacionado con la otra porción mediante la transformación del segundo subbloque pasando del dominio al dominio diferente antes de enventanar el segundo subbloque para obtener un segundo subbloque procesado. El aparato también comprende un transformador para convertir dicho dominio diferente del primer subbloque procesado y el segundo subbloque procesado en otro dominio diferente mediante la misma regla de transformación de bloques para obtener un primer bloque convertido que después se puede comprimir mediante algoritmos de compresión de datos muy conocidos. De este modo, se puede obtener una conmutación muestreada críticamente entre dos modos de codificación, ya que las porciones de solapamiento pertenecientes a dos dominios diferentes se correlacionan entre sí.

**[0009]** En el artículo "*Unified Speech and Audio Coding Scheme for High Quality at Low Bitrate*" de M.s Neuendorf y col., se describe un códec de audio y habla unificado, que combina eficientemente técnicas de la codificación de habla y de la codificación de audio. El artículo ofrece una visión general de la arquitectura del códec y presenta resultados de pruebas de escucha formales.

**[0010]** En el documento WO 2008/071353 A2, se describe un codificador, un decodificador y un procedimiento para codificar y decodificar segmentos de datos que representan una corriente de datos de dominio temporal. En dicho documento se describe un aparato para decodificar segmentos de datos que representan una corriente de datos de dominio temporal, con un segmento de datos que se codifica en el dominio temporal o en el dominio de frecuencia, en el que un segmento de datos que se codifica en el dominio frecuencia posee sucesivos bloques de datos que representan bloques sucesivos y superpuestos de muestras de datos de dominio temporal. El aparato comprende un decodificador de dominio temporal para decodificar un segmento de datos que se codifica en el dominio temporal y el procesador para procesar el segmento de datos que se codifica en el dominio de frecuencia y datos de salida del decodificador de dominio temporal para obtener bloques de datos de dominio temporal superpuestos. El aparato también comprende un combinador de superposición/adición para combinar los bloques de datos de dominio temporal superpuestos con el fin de obtener un segmento de datos decodificado de la corriente de datos de dominio temporal.

**[0011]** A la vista de esta situación, se desea crear conceptos que permitan una transición suficientemente suave, pero eficiente en cuanto a la tasa de bits, entre tramas de audio codificadas en el dominio de frecuencia y tramas de audio codificadas en el dominio de predicción lineal.

## 35 RESUMEN DE LA INVENCION

**[0012]** Este problema se resuelve mediante un decodificador de audio de acuerdo con la reivindicación 1, un procedimiento para proporcionar una representación decodificada de un contenido de audio partiendo de una representación codificada del contenido de audio de acuerdo con la reivindicación 11 y mediante un programa informático de acuerdo con la reivindicación 12.

**[0013]** Una realización de acuerdo con un primer aspecto de la invención crea un decodificador de audio para proporcionar una representación decodificada de un contenido de audio a partir de una representación codificada del contenido de audio. El decodificador de audio comprende un núcleo decodificador de dominio de predicción lineal configurado para proporcionar una representación en el dominio temporal de una trama de audio a partir de un conjunto de parámetros del dominio de predicción lineal. El decodificador de audio también comprende un núcleo decodificador de dominio de frecuencia configurado para proporcionar una representación en el dominio temporal de una trama de audio a partir de un conjunto de parámetros del dominio de frecuencia, teniendo en cuenta una ventana de transformada de entre un conjunto que comprende una pluralidad de diferentes ventanas de transformada. El decodificador de audio también comprende un combinador de señales configurado para la superposición y adición de representaciones en dominio temporal de tramas de audio posteriores codificadas en diferentes dominios con el fin de suavizar una transición entre las representaciones en el dominio temporal de las tramas de audio posteriores. El conjunto de ventanas de transformada disponible para ser aplicado por el núcleo decodificador de dominio de frecuencia comprende una ventana de inserción adaptada para una generación de una representación en el dominio temporal de una única trama de audio codificada en el dominio de frecuencia e intercalada temporalmente entre una trama de audio anterior codificada en el dominio de predicción lineal y una trama de audio posterior codificada en el dominio de predicción lineal. Una pendiente de transición del lado izquierdo de la ventana de inserción está adaptada para proporcionar una transición suave entre una representación en el dominio temporal de la trama de audio anterior codificada en el dominio de predicción lineal y la representación en el

dominio temporal de la trama actual codificada en el dominio de frecuencia. Una pendiente de transición del lado derecho de la ventana de inserción está adaptada para proporcionar una transición suave entre una representación en el dominio temporal de la trama de audio actual codificada en el dominio de frecuencia y una representación en el dominio temporal de la trama posterior codificada en el dominio de predicción lineal, en la que el decodificador de audio comprende un decodificador/cuantificador inverso/reescalador configurado para recibir una información codificada de valores espectrales, para decodificar la información codificada de valores espectrales, para llevar a cabo una cuantificación inversa de la información decodificada de valores espectrales y para llevar a cabo un reescalamiento mediante la aplicación de factores de escala o factores de escala inversos a la información de valores espectrales decodificada e inversamente cuantificada, para obtener el conjunto de parámetros del dominio de frecuencia. Esta realización de la invención se basa en el hallazgo de que resulta ventajoso poder insertar una única trama de audio (denominada en la presente memoria como trama de audio "actual") entre una trama de audio anterior codificada en el dominio de predicción lineal y una trama de audio posterior también codificada en el dominio de predicción lineal, y que debería garantizarse una transición suave entre estas tres tramas de audio mediante una ventana de inserción apropiada.

15

**[0014]** Se ha descubierto que la inserción de una única trama de audio codificada en el dominio de frecuencia entre dos tramas de audio codificadas en el dominio de predicción lineal permite la posibilidad de codificar de manera realista el sonido de fondo que acompaña al habla. Aunque lo más eficiente podría ser codificar el habla propiamente dicha en el dominio de predicción lineal, una codificación en el dominio de predicción lineal suele resultar ineficiente o incluso muy ineficiente para codificar el ruido de fondo, que puede ser dominante durante las pausas comparativamente breves entre cada palabra.

**[0015]** De este modo, si no fuera posible introducir una única trama de audio codificada en el dominio de frecuencia entre dos tramas de audio codificadas en el dominio de predicción lineal, resultaría muy ineficiente en cuanto a la tasa de bits si se codificase el ruido de fondo durante las breves pausas entre dos palabras (ya que sería necesario codificar dicho ruido de fondo en el dominio de predicción lineal, que no resulta muy adecuado para codificar el ruido de fondo), o la codificación del ruido de fondo sería bastante imprecisa (si el ruido de fondo se codificase en el dominio de predicción lineal al tiempo que se limitase la tasa de bits a un valor bajo).

**[0016]** En cambio, el concepto de la invención, que proporciona una ventana de inserción apropiada, permite insertar una única trama de audio codificada en el dominio de frecuencia entre dos tramas de audio codificadas en el dominio de predicción lineal y permite, por tanto, obtener una codificación y decodificación eficiente en cuanto a los recursos, pero precisa, del ruido de fondo entre cada palabra. Aunque las partes del habla se codifican en el dominio de predicción lineal, que es el más adecuado para la codificación del habla, el ruido de fondo durante las interrupciones (es decir, las pausas) entre las palabras se codifican en el dominio de frecuencia, lo que permite una codificación eficiente en cuanto a la tasa de bits y que se adapta bien a la percepción humana de dicho ruido de fondo.

**[0017]** No obstante, se ha descubierto que una inserción suave de dicha única trama de audio codificada en el dominio de frecuencia entre múltiples tramas de audio codificadas en el dominio de predicción lineal se puede llevar a cabo con una buena eficiencia en cuanto a la tasa de bits mediante una ventana de inserción con la forma adecuada, cuyas pendientes de transición tengan la forma oportuna para suavizar las transiciones entre la trama de audio anterior codificada en el dominio de predicción lineal, la trama de audio actual codificada en el dominio de frecuencia y la trama de audio posterior codificada en el dominio de predicción lineal.

45

**[0018]** Otras realizaciones de acuerdo con la invención también crean un procedimiento para proporcionar una representación decodificada de un contenido de audio, y dicho procedimiento se basa en las ideas expuestas anteriormente.

**[0019]** Otras realizaciones de acuerdo con la invención crean un programa informático para llevar a cabo el procedimiento mencionado anteriormente.

#### BREVE DESCRIPCIÓN DE LAS FIGURAS

**[0020]** A continuación, se describirán realizaciones de la presente invención haciendo referencia a las figuras adjuntas, en las que:

la fig. 1 muestra un diagrama de bloques esquemático de un codificador de audio;

la fig. 2 muestra un diagrama de bloques esquemático de un decodificador de audio, de acuerdo con una

realización de la invención;

la fig. 3 muestra una representación gráfica de las secuencias de ventanas y ventanas de transformada para su aplicación por parte del codificador de audio de la fig. 1 y el decodificador de audio de acuerdo con la fig. 2;

la fig. 4a muestra una representación esquemática detallada de una ventana de transformada del tipo  
5 "long\_start\_window\_to\_LPD";

la fig. 4b muestra una representación esquemática detallada de una ventana de transformada del tipo  
"8\*short\_window\_to\_LPD";

la fig. 4c muestra una representación esquemática detallada de una ventana de transformada del tipo  
"long\_start\_window\_to\_LPD";

la fig. 4d muestra una representación esquemática detallada de una ventana de transformada del tipo  
10 "stop\_start\_window\_1152\_to\_LPD";

la fig. 5 muestra una representación esquemática de las secuencias de ventanas permitidas;

la fig. 6a muestra una representación esquemática de una primera secuencia de ventanas;

la fig. 6b muestra una representación esquemática de una segunda secuencia de ventanas;

la fig. 6c muestra una representación esquemática de una tercera secuencia de ventanas;

la fig. 6d muestra una representación esquemática de una cuarta secuencia de ventanas;

la fig. 7 muestra un diagrama de flujo de un procedimiento para proporcionar una representación  
15 decodificada de un contenido de audio;

la fig. 8 muestra un diagrama de flujo de un procedimiento para proporcionar una representación codificada  
20 de un contenido de audio;

la fig. 9 muestra una representación esquemática de un proceso de superposición y adición entre una  
representación en el dominio temporal de una trama de audio anterior codificada en el dominio de predicción lineal y  
una trama de audio actual codificada en el dominio de frecuencia;

la fig. 10a muestra una representación esquemática detallada de una transición entre una trama de audio  
25 codificada mediante una ventana de tipo "LPD\_start\_window" y una trama de audio codificada en el dominio de  
predicción lineal; y

la fig. 10b muestra una representación esquemática detallada de una transición entre una trama de audio  
codificada mediante una ventana de tipo "stop\_start\_window\_1125" y una trama de audio posterior codificada en el  
dominio de predicción lineal.

30

## DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES

### 1. Codificador de audio – estructura general

35 **[0021]** A continuación, se describirá un codificador de audio de acuerdo con un ejemplo, tomando como  
referencia la fig. 1, que muestra un diagrama de bloques esquemático de un codificador de audio 100. El codificador  
de audio 100 está configurado para recibir una representación de audio de entrada 110 y para proporcionar, a partir  
de la misma, una corriente de bits 112 que representa un contenido de audio que se describe mediante la  
representación de audio de entrada 110. El codificador de audio 100 comprende un núcleo codificador de dominio de  
40 frecuencia 120 y un núcleo codificador de dominio de predicción lineal 130. El núcleo decodificador de dominio de  
frecuencia 120 está configurado para recibir la representación de audio de entrada 110 u opcionalmente una versión  
preprocesada 110a de la misma. La versión preprocesada 110a se puede obtener, por ejemplo, mediante un  
preprocesador 110b opcional. El núcleo codificador de dominio de frecuencia 120 también está configurado para  
recibir una información de clasificación de señales 124, que se puede obtener a partir de la representación de audio  
45 de entrada 110 mediante un clasificador de señales 122 opcional, o de cualquier otro modo. El núcleo codificador de  
dominio de frecuencia 120 está configurado para proporcionar un conjunto de parámetros del dominio de frecuencia  
126 asociado a una trama de audio de la representación de audio de entrada 110, 110a. Por ejemplo, el núcleo  
codificador de dominio de frecuencia puede estar configurado para proporcionar un conjunto de parámetros  
espectrales (por ejemplo, valores espectrales) 126 asociado con una trama de audio de la representación de audio  
50 de entrada 110, 110a. Además, el núcleo codificador de dominio de frecuencia 120 está configurado para  
proporcionar una información de secuencia de ventanas 128 que describe qué ventana de transformada se ha  
utilizado para obtener el conjunto de parámetros del dominio de frecuencia 126. El codificador de audio 100 también  
puede comprender, opcionalmente, un procesador de modelo psicoacústico 140, que esté configurado para recibir la  
representación de audio de entrada 110, 110a y para proporcionar una información de modelo psicoacústico 142,  
55 144 a partir de la representación de audio de entrada 110, 110a.

**[0022]** El codificador de audio 100 también comprende, opcionalmente, un procesador espectral 150, que está  
configurado para recibir un conjunto de parámetros del dominio de frecuencia 126, o incluso una secuencia de  
conjuntos de parámetros del dominio de frecuencia 126, y para llevar a cabo un posprocesamiento espectral. Por

ejemplo, el posprocesador espectral 150 puede estar configurado para llevar a cabo una conformación de ruido temporal y/o una predicción a largo plazo a partir de los uno o más conjuntos de parámetros del dominio de frecuencia 126, con lo que proporciona uno o más conjuntos posprocesados de parámetros del dominio de frecuencia 152. El codificador de audio 100 también comprende un escalador/cuantificador/codificador 154 configurado para escalar, cuantificar y codificar los parámetros del dominio de frecuencia del conjunto 126 o el conjunto posprocesado 152 de parámetros del dominio de frecuencia. Por consiguiente, el escalador/cuantificador/codificador 154 proporciona un conjunto escalado, cuantificado y codificado 156 de parámetros del dominio de frecuencia.

10 **[0023]** Además, el codificador de audio 100 puede comprender un codificador opcional 160 configurado para recibir la información de la secuencia de ventanas 128, codificar la información de la secuencia de ventanas 128 y proporcionar una información de secuencia de ventanas codificada 162.

15 **[0024]** El núcleo codificador de dominio de predicción lineal 130 está configurado para recibir la representación de audio de entrada 110 (o su versión preprocesada 110a) y proporcionar un conjunto 170 de parámetros del dominio de predicción lineal a partir de la misma. El conjunto 170 de parámetros del dominio de predicción lineal puede estar asociado con una trama de audio para la que se ha determinado, por ejemplo, mediante el clasificador de señales 122, que la trama de audio es una trama de audio de habla. El conjunto 170 de parámetros del dominio de predicción lineal se puede generar mediante el núcleo codificador de dominio de predicción lineal 130, de manera que el conjunto de parámetros del dominio de predicción lineal represente coeficientes de un filtro de predicción lineal y un estímulo del filtro de predicción lineal, de manera que una salida del filtro de predicción lineal, que se puede obtener como respuesta al estímulo, se aproxima a un contenido de habla de la trama de audio introducida en el núcleo codificador de dominio de predicción lineal 130. El codificador de audio 130 también comprende un codificador opcional 180, que está configurado para codificar el conjunto 170 de parámetros de dominio de predicción lineal, con el fin de obtener un conjunto codificado 182 de parámetros de dominio de predicción lineal. El codificador de audio 100 también comprende un formateador de carga útil de corriente de bits 190 opcional, que está configurado para recibir el conjunto 126 de parámetros de dominio de frecuencia (u, opcionalmente, el conjunto 152 posprocesado espectralmente de parámetros de dominio de frecuencia, o, como otra posibilidad, el conjunto escalado, cuantificado y codificado 156 de parámetros de dominio de frecuencia), la información de la secuencia de ventanas 128 (o, como otra posibilidad, la información de la secuencia de ventanas codificada 152) y el conjunto 170 de parámetros de dominio de predicción lineal (o, como otra posibilidad, su versión codificada 182) y para proporcionar una corriente de bits 112, que representa el contenido de audio de la representación de audio de entrada 110 en una forma codificada.

35 **[0025]** En lo que respecta a la funcionalidad del codificador de audio 100, cabe señalar que el codificador de audio 100 es capaz de codificar de manera selectiva el contenido de audio de la representación de audio de entrada 110 en el dominio de frecuencia y en el dominio de predicción lineal. Por consiguiente, es posible codificar una trama de audio (por ejemplo, una trama de audio de 2048 muestras en el dominio temporal) de la representación de audio de entrada 110 en el dominio más adecuado para una codificación eficiente en cuanto a la tasa de bits. Por ejemplo, se pueden codificar contenidos de audio generales, como música instrumental y ruido ambiental, con una buena eficiencia en la codificación (es decir, con un buen equilibrio entre tasa de bits y calidad perceptual) en el dominio de frecuencia. En cambio, las tramas de audio que comprenden un contenido de audio de habla (o similar al habla) (denominadas en lo sucesivo tramas de audio de habla) se pueden codificar de manera más eficiente (es decir, con un mejor equilibrio entre la calidad del habla percibida y la tasa de bits) en el dominio de predicción lineal. Para codificar el habla en el dominio de predicción lineal, se aprovechan características del aparato fonador humano con el fin de obtener una representación eficiente en cuanto a la tasa de bits de los coeficientes del filtro de predicción lineal. Además, el núcleo codificador de dominio de predicción lineal 130 está adaptado para sacar provecho de las características de la epiglotis y las cuerdas vocales con el fin de proporcionar una representación eficiente del estímulo del filtro de predicción lineal.

50

**[0026]** El codificador de audio 100 está configurado para manejar señales en las que se incluyen tramas de audio "de no habla" (es decir, tramas en las que un contenido de habla no predomina frente a un contenido de audio general como, por ejemplo, música instrumental o ruido ambiental) y tramas de audio de habla (es decir, tramas de audio en las que un contenido de habla predomina frente a un contenido de audio de no habla). Por consiguiente, las diferentes tramas de audio, que normalmente se superponen y que pueden comprender, por ejemplo, una longitud de 2048 muestras de audio, se codifican en diferentes dominios de codificación (dominio de frecuencia o dominio de predicción lineal). Por consiguiente, se puede incluir un conjunto 126, 152, 156 de parámetros de dominio de frecuencia en la corriente de bits 112 para una trama de audio de no habla (omitiendo la inclusión de un conjunto 170, 182 de parámetros de dominio de predicción lineal para dicha trama de audio de no habla), y se puede incluir

un conjunto 170, 182 de parámetros de dominio de predicción lineal en la corriente de bits 112 para una trama de audio de habla (omitiendo la inclusión de un conjunto 126, 152, 156 de parámetros de dominio de frecuencia para dicha trama de audio de habla). Por consiguiente, cada trama de audio de la representación de audio de entrada 110 se puede codificar en el dominio que resulte más adecuado (por ejemplo, en términos de equilibrio entre calidad 5 perceptual y tasa de bits) para la codificación de la respectiva trama de audio.

**[0027]** A continuación, se expondrán algunos detalles relativos a la codificación en el dominio de frecuencia y la codificación en el dominio de predicción lineal. Cabe señalar que un aspecto importante de la codificación consiste en evitar artefactos de bloques en una transición entre tramas de audio posteriores (que suelen superponerse 10 parcialmente) codificadas en el mismo dominio o codificadas en diferentes dominios. Prestando atención a esta circunstancia en el lado del codificador, se puede facilitar una reconstrucción de toda la señal de audio evitando artefactos de bloques no deseados.

**[0028]** Para tramas de audio de no habla, la información de audio codificada, que se incluye en la corriente de 15 bits 112, se genera mediante el núcleo codificador de dominio de frecuencia 120. El núcleo codificador de dominio de frecuencia 120 comprende un enventanador/transformador 120a, que está configurado para proporcionar el conjunto 126 de parámetros de dominio de frecuencia a partir de una representación en el dominio temporal de una trama de audio de la representación de audio de entrada 110, 110a. El enventanador/transformador 120a está configurado para llevar a cabo una transformada superpuesta de una porción enventanada de la información de 20 audio de entrada, preferentemente de una porción enventanada de una trama de audio de la representación de audio de entrada 110, 110a. Por ejemplo, el enventanador/transformador 120a está configurado para llevar a cabo una transformada de coseno discreta modificada (MDCT) a partir de una representación de dominio temporal enventanada de una trama de audio dada de la representación de audio de entrada 110, 110a, con lo que se obtiene un conjunto de parámetros de transformada de coseno discreta modificada, que constituyen un conjunto 126 de 25 parámetros de dominio de frecuencia asociados con la trama de audio dada. Por ejemplo, se puede proporcionar un conjunto de 1024 coeficientes de MDCT mediante el enventanador/transformador 120a a partir de una trama de audio con 2048 muestras en el dominio temporal (aunque algunas de las 2048 muestras en el dominio temporal de la trama de audio se pueden forzar a cero mediante el enventanado). Normalmente, un cierto número de muestras de dominio temporal de la trama de audio dada que se considera para la generación del conjunto de coeficientes de 30 MDCT puede ser mayor que el número de coeficientes de MDCT, con lo que se crea un, así denominado, solapamiento de dominio temporal. No obstante, el solapamiento de dominio temporal se puede reducir, o incluso eliminar por completo, mediante una operación de superposición y adición llevada a cabo en un correspondiente decodificador de audio, superponiendo y añadiendo representaciones de dominio temporal reconstruidas de tramas de audio posteriores.

**[0029]** Es importante señalar que el enventanador/transformador puede estar configurado para aplicar 35 diferentes ventanas de transformada, de entre un conjunto que comprende una pluralidad de ventanas de transformada, antes de llevar a cabo la transformada MDCT o durante la aplicación de la transformada MDCT. La aplicación de una ventana de transformada suaviza una transición entre tramas de audio posteriores y permite lograr una superposición y adición suave de tramas de audio posteriores codificadas en el dominio temporal. Por ejemplo, 40 el enventanado también se puede llevar a cabo de manera que, en una zona de superposición, en la que hay una superposición entre tramas de audio posteriores, las muestras de audio tengan una menor ponderación al ir aumentando la distancia desde el centro de la trama de audio actual (para la que se está llevando a cabo el enventanado actualmente). De este modo, en una porción del lado izquierdo de una zona de superposición (en la que el término “del lado izquierdo” se utiliza como sinónimo de “temporalmente anterior”) entre una trama de audio 45 actual y una trama de audio posterior, las muestras de dominio temporal tienen una mayor ponderación para obtener los parámetros de MDCT de la trama de audio actual que para obtener los parámetros de MDCT de la trama de audio posterior. En cambio, en una porción del lado derecho de una zona de superposición (en la que el término “del lado derecho” se utiliza como sinónimo de “temporalmente posterior”) entre la trama de audio actual y la trama de 50 audio posterior, las muestras de dominio temporal tienen una mayor ponderación para la generación de los coeficientes de MDCT de las tramas de audio posteriores que para la generación de los coeficientes de MDCT de la trama de audio actual.

**[0030]** Normalmente, hay diferentes tipos de ventanas disponibles para codificar tramas de audio posteriores 55 que se van a codificar en el dominio de frecuencia. Por ejemplo, hay tipos de ventanas disponibles con diferentes resoluciones temporales, que comprenden unas zonas de transición (por ejemplo, en forma de pendientes de transición) de diferentes longitudes. Además, normalmente hay disponibles unos tipos de ventanas específicos para conmutar entre diferentes resoluciones temporales. Además, el codificador de audio 100 comprende una pluralidad de tipos de ventanas que se ajustan para proporcionar una transición suave entre tramas de audio codificadas en

diferentes dominios (por ejemplo, dominio de frecuencia y dominio de predicción lineal), como se expondrá detalladamente a continuación.

**[0031]** El núcleo codificador de dominio de frecuencia comprende normalmente un determinador de secuencia de ventanas/selector de ventanas 120b, que está configurado para recibir la información de clasificación de señales 124 procedente del clasificador de señales 122 y para seleccionar un tipo de ventana de entre un conjunto que comprende una pluralidad de tipos de ventanas (o una pluralidad de ventanas de transformada). Por consiguiente, el determinador de secuencia de ventanas/selector de ventanas 120b proporciona la información de secuencia de ventanas 128, que es evaluada por el enventanador/transformador 120a con el fin de llevar a cabo una información de enventanado apropiada antes de la transformada MDCT o durante la transformada MDCT.

**[0032]** El determinador de secuencia de ventanas/selector de ventanas 120b está configurado preferentemente para tener en cuenta un tipo de ventana utilizado para proporcionar los coeficientes de MDCT de una trama de audio anterior, si la trama de audio anterior se codificó en el dominio de frecuencia. Además, el determinador de secuencia de ventanas 120b tiene en cuenta una información de clasificación de señales 124 que indica si la trama de audio anterior se codificó en el dominio de frecuencia o en el dominio de predicción lineal. Además, el determinador de secuencia de ventanas 120b normalmente tiene en cuenta una información de clasificación de señales 124 que indica una resolución temporal que se debería aplicar para codificar la trama de audio actual, y además una resolución temporal que se debería utilizar para codificar una trama de audio posterior. De este modo, el determinador de secuencia de ventanas/selector de ventanas 120b está configurado preferentemente para adaptar un tipo de ventana que se va a utilizar para proporcionar los coeficientes de MDCT de la trama de audio actual al dominio en el que se ha codificado la trama de audio anterior y al dominio en el que se va a codificar la trama de audio posterior. Además, el determinador de secuencia de ventanas 120b normalmente está configurado para tener en cuenta una resolución temporal (o forma de ventana asociada) de las tramas de audio contiguas (es decir, la trama de audio anterior y la trama de audio posterior), si cualquiera de las tramas de audio contiguas se codifica en el dominio de frecuencia. Más adelante se expondrán algunos detalles específicos relativos a la selección de las ventanas de transformada tomando también como referencia las figs. 2, 3, 4a-4d, 5 y 6a-d.

**[0033]** Además, cabe señalar que la funcionalidad básica del preprocesador 110b, del núcleo codificador de dominio de frecuencia 120, del posprocesador espectral 150 y del escalador/cuantificador/codificador 154 es similar a la funcionalidad de los componentes "control de ganancia de AAC", "conmutación de bloques/banco de filtros", "TNS", "predicción a largo plazo", "intensidad", "predicción", "PNS", "M/S" y "AAC: escalamiento/cuantificación/codificación Huffman" descritos en la norma internacional ISO/IEC 14496-3, parte 3, apartado 4, y la bibliografía relacionada con la misma.

**[0034]** No obstante, el codificador de audio 100 proporciona la posibilidad añadida de codificar tramas de audio en el dominio de predicción lineal, con lo que se proporciona el conjunto 170, 182 de parámetros de dominio de predicción lineal. El conjunto 170, 182 de los parámetros de dominio de predicción lineal puede ser similar (o incluso idéntico) a un conjunto de parámetros de dominio de predicción lineal proporcionado por un codificador de audio de acuerdo con una de las normas 3GPP TS 26.090, 3GPP TS 26.190 o 3GPP TS 26.290. No obstante, el conjunto 170, 182 de parámetros de dominio de predicción lineal puede ser similar (o incluso idéntico) a cualquier otro conjunto de parámetros de dominio de predicción lineal proporcionado por un codificador de audio de habla.

## 2. Decodificador de audio – estructura general

**[0035]** A continuación, se describirá la estructura de un decodificador de audio de acuerdo con las realizaciones de la presente invención, haciendo referencia a la fig. 2, que muestra un diagrama de bloques esquemático ilustrativo de dicho decodificador de audio 200. El decodificador de audio 200 está configurado para recibir una corriente de bits 210, que puede incluir una información de audio codificada. La corriente de bits 210 puede ser idéntica a la corriente de bits 112 proporcionada por el codificador de audio 100. El decodificador de audio 200 comprende un deformateador de carga útil de corriente de bits 220 opcional, que está configurado para analizar la corriente de bits 210 y extraer diferentes elementos de información a partir de la corriente de bits 210. Por ejemplo, el deformateador de carga útil de corriente de bits 220 está configurado para extraer una información de valores espectrales codificada 222 a partir de la corriente de bits 210. Además, el deformateador de carga útil de corriente de bits 220 está configurado para extraer una información de ventana 224 a partir de la corriente de bits 210, extraer una información de estímulo codificada 226 a partir de la corriente de bits 210 y además extraer parámetros de filtro de dominio de predicción lineal codificados 228 a partir de la corriente de bits 210. No obstante, cabe señalar que no es necesario que se extraigan todos los elementos de información 222, 224, 226, 228 para cada trama de audio. En algunos casos, la información de estímulo codificada 226 también se puede omitir por



completo, si el estímulo se codifica utilizando la información de valores espectrales codificada 226 (excitación codificada por transformada).

**[0036]** El decodificador de audio 200 comprende un decodificador/cuantificador inverso/reescalador 230, que está configurado para recibir la información de valores espectrales codificada 222 y decodificar la información de valores espectrales codificada 222, por ejemplo, llevando a cabo un algoritmo de decodificación aritmética. Además, el decodificador/cuantificador inverso/reescalador 230 está configurado para llevar a cabo una cuantificación inversa de la información de valores espectrales decodificada (por ejemplo, mediante un algoritmo de cuantificación inversa) y un reescalamiento mediante la aplicación de factores de escala o factores de escala inversos a la información de valores espectrales decodificada y cuantificada inversamente. Por consiguiente, el decodificador/cuantificador inverso/reescalador 230 proporciona una información de valores espectrales decodificada, cuantificada inversamente y reescalada 230 a partir de la información de valores espectrales codificada 222.

**[0037]** El decodificador de audio 230 también comprende un preprocesador espectral 240 opcional, que está configurado para llevar a cabo un preprocesamiento espectral de la información de valores espectrales decodificada 232 con el fin de obtener un conjunto 242 preprocesado de parámetros de dominio de frecuencia asociados con una trama de audio (o una secuencia de tramas de audio). Por ejemplo, el conjunto 242 de parámetros de dominio de frecuencia puede formar parte de una representación de tiempo-frecuencia de un contenido de audio en forma de conjunto de parámetros (o valores) espectrales asociados con una trama de audio. Por ejemplo, el conjunto 242 de parámetros de dominio de frecuencia puede comprender un conjunto de coeficientes de MDCT asociados con una trama de audio y determinados, por ejemplo, por el núcleo codificador de dominio de frecuencia 120 del codificador de audio 100. El decodificador de audio 200 comprende un núcleo decodificador de dominio de frecuencia 250, que está configurado para recibir el conjunto 242 de parámetros de dominio de frecuencia, así como la información de ventana 224 (o cualquier información equivalente que describa una ventana que vaya a ser aplicada por el núcleo decodificador de dominio de frecuencia 250). El núcleo decodificador de dominio de frecuencia 250 también está configurado para proporcionar una representación en el dominio temporal 252 del contenido de audio de una trama de audio a partir del conjunto 242 de parámetros de dominio de frecuencia asociados con la trama de audio y la información de ventana 224. Para ello, el núcleo decodificador de dominio de frecuencia 250 comprende un transformador/enventanador 250a, que está configurado para llevar a cabo una transformada superpuesta del conjunto 242 de parámetros de dominio de frecuencia. El transformador/enventanador 250a puede estar configurado para llevar a cabo, además, un enventanado empleando una ventana de transformada de entre un conjunto que comprende una pluralidad de ventanas de transformada. Para seleccionar una ventana de transformada apropiada, el núcleo decodificador de dominio de frecuencia 250 comprende preferentemente un determinador de secuencia de ventanas/selector de ventana 250b, que está configurado para seleccionar una ventana de transformada apropiada a partir de la información de ventana 224 (o cualquier información equivalente). El enventanado se puede aplicar durante la transformada superpuesta (es decir, mediante una operación combinada de transformada/enventanado) o tras la transformada superpuesta (es decir, mediante otra operación de enventanado tras la operación de transformada superpuesta). A continuación, se expondrán algunos detalles relativos a la elección de la ventana de transformada apropiada, haciendo referencia a las figs. 3, 4a-4d, 5 y 6a-6d.

**[0038]** El decodificador de audio 200 también comprende un núcleo decodificador de dominio de predicción lineal 260, que está configurado para recibir los parámetros del filtro de dominio de predicción lineal codificados 228 y la información de estímulo. Por ejemplo, el núcleo decodificador de dominio de predicción lineal puede estar configurado para recibir la información de valores espectrales decodificada 232 como una información de estímulo que describe una excitación codificada por transformada de un filtro de predicción lineal. Otra posibilidad consiste en que el núcleo decodificador de dominio de predicción lineal 260 reciba la información de estímulo específico codificada 226, que puede representar un estímulo del filtro de predicción lineal empleando un concepto denominado CELP o un concepto denominado ACELP. Para obtener más detalles respecto a la codificación del estímulo, cabe hacer referencia, por ejemplo, a las normas 3GPP mencionadas anteriormente, en las que se describe una codificación de un estímulo de un filtro de predicción lineal para la codificación/decodificación de habla. El núcleo decodificador de dominio de predicción lineal 260 comprende preferentemente un procesador de estímulos 260a, que está configurado para recibir la información de estímulo 232, 226 y proporcionar una señal de estímulo de dominio temporal 260b a partir de la información de estímulo 232, 226. Por ejemplo, el procesador de estímulos 260a puede comprender un banco de filtros para proporcionar la señal de estímulo de dominio temporal 260b a partir de la información de valores espectrales decodificada 232 en el caso de una excitación codificada por transformada. Además, o como otra posibilidad, el procesador de estímulos 260a puede comprender un decodificador de dominio temporal con el fin de proporcionar la señal de estímulo de dominio temporal 260b a partir de la información de estímulo específico codificada 226, por ejemplo, en el caso de una excitación codificada por CELP o codificada por ACELP. El núcleo decodificador de dominio de predicción lineal 260 también comprende un posprocesador de

estímulos 260c opcional, que está configurado para recibir la señal de estímulo de dominio temporal 260b y proporcionar una versión posprocesada 260d de la señal de estímulo de dominio temporal. El núcleo decodificador de dominio de predicción lineal 260 también comprende un filtro de codificación por predicción lineal 260e, que está configurado para recibir los parámetros de filtro de dominio de predicción lineal codificados 228 (o parámetros de filtro de dominio de predicción lineal decodificados) y la señal de estímulo de dominio temporal 260b, 260d. El filtro de codificación por predicción lineal 260e está configurado para llevar a cabo un filtrado por predicción lineal ajustable que depende de los parámetros de filtro de dominio de predicción lineal codificados 228 (o parámetros de filtro de dominio de predicción lineal decodificados obtenidos a partir de los mismos) y proporcionar una señal de síntesis 260f filtrando la señal de excitación de dominio temporal 260b, 260d mediante una operación de filtrado por predicción lineal ajustado de acuerdo con los parámetros de filtro de dominio de predicción lineal codificados 228. Además, el núcleo decodificador de dominio de predicción lineal 260 comprende un posprocesador 260g opcional, que está configurado para recibir y posprocesar la señal de síntesis 260f y proporcionar una señal en el dominio temporal 262.

15 **[0039]** El decodificador de audio 200 también comprende un procesador de superposición y adición 270, que está configurado para recibir la representación en el dominio temporal 252 de tramas de audio codificadas en el dominio de frecuencia y la representación en el dominio temporal 262 de tramas de audio codificadas en el dominio de predicción lineal, y superponer y añadir representaciones de dominio temporal de tramas de audio posteriores, por lo que proporciona una representación continua en el dominio temporal 272 de la información de audio codificada representada por la corriente de bits 210.

**[0040]** El decodificador de audio 200 también comprende, opcionalmente, un posprocesador 280 que está configurado para recibir la representación continua en el dominio temporal 272 del contenido de audio y llevar a cabo un posprocesamiento, con lo que se obtiene una representación continua en dominio temporal posprocesada 282 del contenido de audio. El posprocesador 280 puede estar configurado, por ejemplo, para llevar a cabo un realce de tono, una ampliación de ancho de banda y/o un procesamiento de sonido envolvente.

**[0041]** El decodificador de audio 200 está configurado para proporcionar una transición suave entre representaciones de tramas de audio codificadas en diferentes dominios (dominio de frecuencia y dominio de predicción lineal) mediante una cooperación apropiada entre el núcleo decodificador de dominio de frecuencia 250, el núcleo decodificador de dominio de predicción lineal 260 y el procesador de superposición y adición 270.

**[0042]** En particular, el núcleo decodificador de dominio de frecuencia 250 está configurado para aplicar ventanas de transformada, que están adaptadas específicamente para diferentes secuencias de dominios, en las que se codifican tramas de audio posteriores. Además, las pendientes de transición de las ventanas de transformada, que están disponibles para su aplicación por parte del transformador/enventanador 250a, y que se seleccionan mediante el determinador de secuencia de ventanas/selectores de ventana 250b, están adaptadas a una secuencia específica de dominios, en los que se codifican las tramas de audio posteriores, como se expondrá detalladamente a continuación.

40  
3. Secuencias de ventanas y ventanas de transformada

**[0043]** A continuación, se describirán detalles relativos a las secuencias de ventanas y ventanas de transformada disponibles haciendo referencia a las figs. 3, 4a-4d, 5 y 6a-6d. Cabe señalar que las secuencias de ventanas y ventanas de transformada se pueden aplicar tanto en el enventanador/transformador 120a del codificador de audio 100, como en el transformador/enventanador 250a del decodificador de audio 200. No obstante, se ofrecerá una explicación más detallada respecto al caso del decodificador de audio, ya que la utilidad de las diferentes secuencias de ventanas y ventanas de transformada se puede entender con más facilidad en el contexto del decodificador de audio 200. Sin embargo, las explicaciones que se ofrecen con respecto al decodificador de audio 200 se pueden aplicar de manera análoga para el caso del codificador de audio 100.

### 3.1. Descripción general de los tipos de ventanas

**[0044]** En la fig. 3 se muestra una representación esquemática de secuencias de ventanas y ventanas de transformada, algunas de las cuales, o todas, pueden estar disponibles para ser utilizadas por el enventanador/transformador 120a en diferentes ejemplos o por el transformador/enventanador 250a en diferentes realizaciones de la invención.

**[0045]** Con respecto a la notación de la fig. 3, cabe señalar que una abscisa (por ejemplo, en una dirección

horizontal desde la izquierda del papel a la derecha del papel) describe un tiempo, y que una ordenada (por ejemplo, en una dirección vertical desde la parte inferior de la página a la parte superior de la página) describe una magnitud de un valor de una ventana.

5 **[0046]** Una porción horizontal de una forma de ventana describe una zona de una magnitud constante (al menos aproximadamente) de la forma de ventana. Una porción linealmente ascendente de la forma de ventana describe un incremento sostenido de la magnitud de la forma de ventana, en la que no es necesario que el incremento sea lineal. Más bien, un ascenso lineal de la forma de ventana describe cualquier forma de ventana con un incremento sostenido, que sea adecuada (o se adapte) para llevar a cabo una superposición y adición de  
10 reducción de solapamiento (o cancelación de solapamiento) entre representaciones en el dominio temporal de tramas de audio contiguas. Por ejemplo, una forma de ventana linealmente ascendente puede representar una forma de ventana con un incremento de tipo senoidal o de tipo cosenoidal. De forma similar, una forma de ventana linealmente descendente representa cualquier forma de ventana con una disminución sostenida, que sea adecuada (o se adapte) para llevar a cabo una superposición y adición de reducción de solapamiento (o cancelación de  
15 solapamiento) entre representaciones en el dominio temporal de tramas de audio posteriores enventanadas. Por consiguiente, una forma de ventana linealmente descendente representa, por ejemplo, una forma de ventana de tipo senoidal o de tipo cosenoidal cuya magnitud disminuye al incrementarse el tiempo. Otra posibilidad consiste en que una forma de ventana linealmente ascendente o una forma de ventana linealmente descendente represente una forma de ventana de Kaiser-Bessel creciente o decreciente. No obstante, se pueden representar otras formas de  
20 ventana con un incremento no lineal o una disminución no lineal mediante la forma de ventana linealmente ascendente y la forma de ventana linealmente descendente.

**[0047]** Además, cabe señalar que los intervalos temporales están señalados en las representaciones de las ventanas mediante líneas verticales finas continuas y líneas verticales finas discontinuas. La distancia entre dos  
25 líneas verticales finas continuas es de 1024 muestras, la distancia entre dos líneas verticales finas discontinuas es de 128 muestras y la distancia entre una línea vertical fina continua y una línea vertical fina discontinua contigua es de 128 muestras. La duración temporal de una trama de audio "normal" es de 2048 muestras. El desplazamiento temporal entre dos tramas de audio contiguas es de 1024 muestras, de tal manera que una mitad izquierda de una trama de audio actual se superpone a una mitad derecha de una trama de audio anterior, y de tal manera que una  
30 mitad derecha de la trama de audio actual se superpone a una mitad izquierda de la trama de audio posterior. En la presente memoria, los términos "izquierda" y "derecha" se utilizan para designar un elemento temporalmente anterior (izquierdo) (por ejemplo, muestra de audio o porción de una ventana de transformada) y un elemento temporalmente posterior (derecho) (por ejemplo, muestra de audio o porción de una ventana de transformada).

35 **[0048]** Tomando ahora como referencia la fig. 3, se puede observar que el enventanador/transformador 120a o el transformador/enventanador 250a pueden ser capaces de aplicar una ventana de transformada de entre, por ejemplo, hasta 11 diferentes ventanas de transformada disponibles. No obstante, hay realizaciones en las que hay más ventanas de transformada diferentes disponibles, y también hay realizaciones en las que no están disponibles todas las ventanas de transformada que se muestran en la fig. 3.

40 **[0049]** Además, en la fig. 3 se puede observar que existen 7 "tipos básicos" de ventanas de transformada, que se denominan "long\_window", "long\_start\_window", "8\*short\_window", "long\_stop\_window", "stop\_start\_window", "stop\_window\_1152" y "stop\_start\_window\_1152". No obstante, algunos de los tipos de ventanas mencionados anteriormente, en concreto los tipos de ventanas "long\_start\_window", "8\*short\_window", "stop\_start\_window" y  
45 "stop\_start\_window\_1152", se pueden aplicar antes de una trama de audio posterior codificada en el dominio de predicción lineal.

**[0050]** No obstante, se ha descubierto que se deberían utilizar diferentes formas de ventana propiamente dichas, dependiendo de si una ventana del tipo básico de ventana "long\_start\_window", "8\*short\_window",  
50 "stop\_start\_window" y "stop\_start\_window\_1152" va seguida de una trama de audio codificada en el dominio de frecuencia o una trama de audio codificada en el dominio de predicción lineal. Por consiguiente, hay dos formas de ventana diferentes disponibles para el tipo de ventana "long\_start\_window", concretamente una forma de ventana "normal", utilizada si la trama de audio posterior está codificada en el dominio de frecuencia, y una forma de ventana "modificada" (también denominada "long\_start\_window\_to\_LPD"), utilizada si la trama actual va seguida de una  
55 trama de audio codificada en el dominio de predicción lineal. De forma similar, hay dos formas de ventana propiamente dichas disponibles para el tipo de ventana "8\*short\_window", concretamente una forma de ventana "normal", utilizada si la trama de audio actual va seguida de una trama de audio codificada en el dominio de frecuencia, y una forma de ventana "modificada" (también denominada "8\*short\_window\_to\_LPD"), utilizada si la trama de audio actual va seguida de una trama de audio codificada en el dominio de predicción lineal. Además, hay

dos formas de ventana propiamente dichas disponibles para el tipo de ventana "stop\_start\_window", concretamente una forma de ventana "normal", utilizada si la trama de audio actual va seguida de una trama de audio codificada en el dominio de frecuencia, y una forma de ventana "modificada" (también denominada "stop\_start\_window\_to\_LPD"), utilizada si la trama de audio actual va seguida de una trama de audio codificada en el dominio de predicción lineal.

5 De forma similar, hay dos formas de ventana propiamente dichas disponibles para el tipo de ventana "stop\_start\_window\_1152", concretamente una forma de ventana "normal", utilizada si la trama de audio actual va seguida de una trama de audio codificada en el dominio de frecuencia, y una forma de ventana "modificada" (también denominada "stop\_start\_window\_1152\_to\_LPD"), utilizada si la trama de audio actual va seguida de una trama de audio codificada en el dominio de predicción lineal.

10

**[0051]** Para los tipos de ventana "long\_window", "long\_stop\_window" y "stop\_window\_1152", solo hay una ventana "normal" disponible en algunas realizaciones, ya que dichos tipos de ventana no resultan adecuados para tramas de audio que van seguidas de una trama de audio codificada en el dominio de predicción lineal.

15

**[0052]** A continuación, se expondrán algunos detalles relativos a los tipos de ventanas disponibles.

**[0053]** El tipo de ventana "long\_window" solo está disponible en una versión de forma de ventana "normal" 310 que comprende una pendiente de transición del lado izquierdo 310a con una longitud de 1024 muestras en el dominio temporal y una pendiente de transición del lado derecho 310b con una longitud de 1024 muestras en el dominio temporal.

**[0054]** El tipo de ventana "long\_start\_window" está disponible en una versión de forma de ventana "normal" 320 y una versión de forma de ventana "modificada" 324. La versión de forma de ventana "normal" 320 del tipo de ventana "long\_start\_window" comprende una pendiente de transición del lado izquierdo 320a con una longitud de 1024 muestras en el dominio temporal y una porción de ventana constante del lado derecho 320b con una longitud de 448 muestras en el dominio temporal, una pendiente de transición del lado derecho 320c con una longitud de transición de 128 muestras en el dominio temporal y una porción cero del lado derecho con una longitud de 448 muestras en el dominio temporal. La porción de transición del lado derecho 320c se extiende desde una mitad izquierda de la porción (o mitad) del lado derecho de la ventana de transformada hasta la mitad derecha de la porción del lado derecho de la ventana de transformada. De este modo, la pendiente de transición del lado derecho 320c está adaptada para llevar a cabo, en un decodificador de audio, una operación de superposición y adición entre dos representaciones en dominio temporal enventanadas de tramas de audio posteriores. Suponiendo además que el número de coeficientes de MDCT asociados con la ventana de transformada 320 (por ejemplo, 1024 coeficientes de MDCT) es considerablemente menor que el número de muestras del dominio temporal asociadas con la ventana de transformada 320 (por ejemplo, 2048), se puede entender que el número comparativamente menor de coeficientes de MDCT introduce un solapamiento en el dominio temporal. De este modo, una transformada MDCT de una trama de audio enventanada mediante la ventana de transformada 320 da lugar a coeficientes de MDCT que comprenden un solapamiento en el dominio temporal, ya que las muestras del dominio temporal situadas en la mitad derecha de la porción del lado derecho de la ventana de transformada 320 se pliegan sobre la mitad izquierda de la porción del lado derecho de la ventana de transformada 320. Por consiguiente, dichas muestras precedentes de la mitad derecha de la porción del lado derecho de la ventana de transformada 320 se pliegan desde una porción temporal 321a hasta una porción temporal 321b. No obstante, este solapamiento (plegamiento de muestras del dominio temporal desde una zona de la ventana de transformada a otra zona de la ventana de transformada) se puede compensar mediante una operación de superposición y adición de un decodificador de audio, como se da a conocer, por ejemplo, en la norma internacional ISO/IEC 14496-3, parte 3, apartado 4 y la bibliografía correspondiente.

**[0055]** No obstante, también existe una versión "modificada" de forma de ventana 324 del tipo de ventana "long\_start\_window", que se utiliza (por ejemplo, por parte del codificador de audio 100 o el decodificador de audio 200) para un enventanado de muestras del dominio temporal de una trama actual codificada en el dominio de frecuencia (o que se va a codificar en el dominio de frecuencia) si (y solo si) la trama de audio posterior está codificada (o se va a codificar) en el dominio de predicción lineal. La versión "modificada" de forma de ventana 324 del tipo de ventana "long\_start\_window" comprende una pendiente de transición del lado izquierdo 324a con una longitud de 1024 muestras de audio, que, por ejemplo, puede ser idéntica a la pendiente de transición del lado izquierdo 320a de la versión "normal" de forma de ventana del tipo de ventana "long\_start\_window". De forma similar, la versión "modificada" de forma de ventana 324 del tipo de ventana "long\_start\_window" comprende una porción constante (distinta de cero) 324b con una longitud de 448 muestras de dominio temporal, que puede ser idéntica a la porción constante 322b. Además, la versión "modificada" de forma de ventana del tipo de ventana

“long\_start\_window” comprende una pendiente de transición del lado derecho 324c con una longitud de transición de 64 muestras de dominio temporal (que es más corta que la longitud de la pendiente de transición del lado derecho del tipo de ventana 320). La pendiente de transición del lado derecho 324c está completamente incluida en una mitad del lado izquierdo de la porción del lado derecho de la ventana de transformada 324. Además, la ventana de transformada 324 comprende una porción cero del lado derecho 324d con una longitud de 512 muestras del dominio temporal, que ocupa por completo la mitad del lado derecho de la porción derecha de la forma de ventana 324. Por consiguiente, la forma de ventana 324 evita un solapamiento de muestras de dominio temporal desde la mitad del lado derecho de la porción derecha de la ventana de transformada 324 hacia la mitad izquierda de la porción del lado derecho de la ventana de transformada 324, si se aplica una transformada de coseno discreta modificada que proporciona, por ejemplo, 1024 coeficientes de MDCT a una trama de audio con una longitud de, por ejemplo, 2048 muestras de dominio temporal, y dicha trama de audio se enventana mediante la ventana de transformada 324, ya que las muestras del dominio temporal que quedan en la mitad derecha de la porción del lado derecho de la ventana de transformada 324, que se plegarían hacia la mitad izquierda de la porción del lado derecho de la ventana de transformada 324, se fuerzan a cero completamente mediante la porción de cero del lado derecho 324b de la ventana de transformada 324 (ya que la porción cero del lado derecho 324d se extiende a lo largo de toda la mitad derecha de la porción del lado derecho de la ventana de transformada 324). Por consiguiente, la porción temporal de la ventana de transformada 324 que comprende la pendiente de transición del lado derecho 324c no se ve afectada por un solapamiento en el dominio temporal, de manera que resulta necesario llevar a cabo, en el lado de un decodificador, una cancelación del solapamiento durante la porción temporal en la que se extiende la pendiente de transición del lado derecho 324c.

**[0056]** De este modo, se facilita una transición entre una trama de audio codificada en el dominio temporal y una trama de audio posterior codificada en el dominio de predicción lineal mediante un acortamiento de la pendiente de transición del lado derecho 324c de la versión “modificada” de forma de ventana 324 del tipo de ventana “long\_start\_window” en comparación con la pendiente de transición del lado derecho 320c de la versión “normal” de forma de ventana 320 del tipo “long\_start\_window”. La versión “modificada” de forma de ventana 324 está adaptada de manera que se evite un solapamiento del lado derecho, con lo que se elimina la necesidad de una cancelación del solapamiento en el lado del decodificador en el caso de una transición desde una trama de audio codificada en el dominio de frecuencia a una trama de audio codificada en el dominio de predicción lineal.

**[0057]** El conjunto disponible de ventanas de transformada también comprende una versión “normal” de forma de ventana 330 del tipo de ventana “8\*short\_window”. La ventana de transformada 330 comprende una porción cero del lado izquierdo 330a con una longitud de 448 muestras de dominio temporal y una pluralidad de, por ejemplo, 8 subventanas 331-338. Cada una de las subventanas 331-338 comprende una pendiente de transición del lado izquierdo (por ejemplo, indicada con el número 331a) y una pendiente de transición del lado derecho (por ejemplo, indicada con el número 331b). Como puede observarse, las subventanas 331-338 están desplazadas temporalmente unas con respecto a otras, de manera que una pendiente de transición del lado derecho (por ejemplo, indicada con el número 331b) de una subventana dada se superpone a una pendiente de transición del lado izquierdo (por ejemplo, indicada con el número 332a). Un conjunto “corto” de coeficientes de MDCT, por ejemplo, de 128 coeficientes de MDCT, se asocia con cada una de las subventanas 331-338, y cada una de las subventanas 331-338 comprende una extensión temporal de 256 muestras del dominio temporal. Por consiguiente, una representación de un contenido de audio incluido en la duración temporal de una de las subventanas 331-338, que solo comprende 128 coeficientes de MDCT, comprende un solapamiento. Sin embargo, cuando se lleva a cabo (por ejemplo, en el decodificador de audio 200) una operación de superposición y adición de representaciones en el dominio temporal de los contenidos de audio descritos por los conjuntos de coeficientes de MDCT asociados con dos subventanas (por ejemplo, las subventanas 331 y 332), se cancelan los artefactos de solapamiento dentro de las zonas de superposición de las subventanas contiguas (por ejemplo, las subventanas 332 y 332). Además, cabe señalar que una pendiente de transición del lado izquierdo 331 de una primera subventana del lado izquierdo 331 se extiende dentro de la mitad izquierda de la porción del lado izquierdo de la ventana de transformada, y que la pendiente de transición del lado derecho 338b de la última subventana del lado derecho 338 se extiende dentro de la mitad derecha de la porción del lado derecho de la ventana de transformada (de la trama de audio actual) 330. Esto se refleja en el hecho de que la pendiente de transición del lado izquierdo 331 de una primera subventana 331 comprende una duración temporal de, por ejemplo, 128 muestras del dominio temporal, que es idéntica a la duración temporal (por ejemplo, 128 muestras del dominio temporal) de la pendiente de transición del lado derecho 338b de la última subventana del lado derecho 338. En otras palabras, existe un solapamiento dentro de la zona temporal de la pendiente de transición del lado izquierdo 331a de la primera subventana 331, y también existe un solapamiento dentro de la zona temporal de la pendiente de transición del lado derecho 338b de la última subventana del lado derecho 338, de manera que se requiere una cancelación del solapamiento en el lado del decodificador, tanto en la zona temporal de la pendiente de transición del lado izquierdo 331a como en la zona de la pendiente de transición

del lado derecho 338b. Esta cancelación del solapamiento se puede obtener, por ejemplo, mediante una operación de superposición y adición.

**[0058]** A continuación, se expondrán detalles relativos a la versión “modificada” de forma de ventana 340 del tipo de ventana “8\*short\_window”. La denominada “ventana de transformada” 340 comprende una porción cero del lado izquierdo 340a que posee una duración temporal de, por ejemplo, 448 muestras de dominio temporal. La ventana de transformada 340 también comprende una pluralidad de, por ejemplo, 8 subventanas 341-348. Además, la ventana de transformada 340 comprende una porción cero del lado derecho 340b con una duración temporal de 512 muestras de dominio temporal. Cabe señalar aquí que la extensión temporal de la porción cero del lado derecho 340b es mayor (por ejemplo, 512 muestras de dominio temporal) que la extensión temporal de la porción cero del lado izquierdo 340a (por ejemplo, 448 muestras de dominio temporal). Además, cabe señalar que la porción cero del lado derecho 340b cubre por completo la mitad derecha de la porción del lado derecho de la ventana de transformada 340, mientras que la porción cero del lado izquierdo 340a es más corta temporalmente que la mitad izquierda de la porción del lado izquierdo de la ventana de transformada 340.

**[0059]** Respecto a las subventanas 341-348, cabe señalar que una primera subventana del lado izquierdo 341 comprende una extensión más larga temporalmente (por ejemplo, 256 muestras de dominio temporal) que una última subventana del lado derecho 348 (por ejemplo, 192 muestras de dominio temporal). Esto se debe al hecho de que una pendiente de transición del lado izquierdo 341 de una primera subventana del lado izquierdo 341 comprende una extensión temporal más larga (por ejemplo, 128 muestras de dominio temporal) que una pendiente de transición del lado derecho 348b (por ejemplo, 64 muestras de dominio temporal) de la última subventana del lado derecho 348. La pendiente de transición 348b de la última subventana del lado derecho 348 está completamente incluida en la mitad izquierda de la porción del lado derecho de la ventana de transformada 340 y no se extiende dentro de la mitad derecha de la porción del lado derecho de la ventana de transformada 340. Cabe señalar que la primera subventana del lado izquierdo 340 comprende preferentemente la misma forma de ventana que las subventanas centrales 342-347. Además, cabe señalar que las subventanas (preferentemente todas las subventanas 341-348) comprenden pendientes de transición que están adaptadas de tal manera que se produce una cancelación del solapamiento cuando se superponen (por ejemplo, en el decodificador de audio 200) a representaciones en el dominio temporal de contenidos de audio asociados con subventanas posteriores. Específicamente, una pendiente de transición del lado derecho 347b de la última subventana central 347 que precede a la última subventana 348 está adaptada a una pendiente de transición del lado izquierdo 348a de la última subventana 348. No obstante, mientras que las pendientes de transición del lado derecho 341b-347b de las subventanas 341-347 comprenden duraciones temporales y formas idénticas, la pendiente de transición del lado derecho 348b de la última subventana 348 comprende una duración temporal comparativamente más corta.

**[0060]** Al proporcionar una duración temporal más corta a la última subventana 348 de la forma de ventana “modificada” 340 del tipo de ventana “8\*short\_window”, se evitan artefactos de solapamiento durante la duración temporal de la pendiente de transición del lado derecho 348b de la última subventana 348. Por consiguiente, no es necesario llevar a cabo una cancelación del solapamiento durante la duración temporal de la pendiente de transición del lado derecho 348b de la última subventana 348. Por consiguiente, se puede omitir una cancelación del solapamiento cuando se superpone y añade una representación en el dominio temporal decodificada mediante la versión “modificada” de forma de ventana 340 del tipo de ventana “8\*short\_window” con una representación en el dominio temporal de una trama de audio decodificada en el dominio de predicción lineal. De hecho, se puede llevar a cabo una sencilla atenuación cruzada durante la duración temporal de la pendiente de transición del lado derecho 348b de la última subventana 348 entre una representación en el dominio temporal de una trama de audio decodificada mediante la ventana de transformada 340 y una trama de audio posterior decodificada en el dominio de predicción lineal.

**[0061]** A continuación, se describirá la versión “normal” de forma de ventana del tipo de ventana “long\_stop\_window”, que se indica con el número 350. La ventana de transformada 350 comprende una porción cero del lado izquierdo 350a, que comprende una duración temporal de 448 muestras de dominio temporal. La forma de ventana 350 también comprende una pendiente de transición del lado izquierdo 350b, que comprende una duración temporal de 128 muestras de dominio temporal, y que se extiende tanto en una mitad izquierda de la porción del lado izquierdo de la ventana de transformada 350 como en la mitad derecha de la porción del lado izquierdo de la ventana de transformada 350, de manera que normalmente se produce un solapamiento cuando se transforman 2048 muestras de dominio temporal enventanadas mediante la ventana de transformada 350 en el dominio temporal para obtener 1024 coeficientes de MDCT. Puede darse que la ventana de transformada 350 también comprenda una porción constante del lado izquierdo 350c con una duración temporal de 448 muestras de dominio temporal, que se extienda en la mitad derecha de la porción del lado izquierdo de la ventana de transformada 350. Además, la

ventana de transformada 350 comprende una pendiente de transición del lado derecho 350b con una duración temporal de 1024 muestras de dominio temporal, que se extiende en la porción del lado derecho de la ventana de transformada 350.

- 5 **[0062]** A continuación, se describirá una versión “normal” de forma de ventana del tipo de ventana “stop\_start\_window”, que se indica con el número 360. La ventana de transformada 360 comprende una porción cero del lado izquierdo 360a con una duración temporal de 448 muestras de dominio temporal y una pendiente de transición del lado izquierdo 360b con una duración temporal de 128 muestras de dominio temporal, que se extiende tanto en una mitad del lado izquierdo como en una mitad del lado derecho de la porción del lado izquierdo de la
- 10 ventana de transformada 360, de manera que se produce un solapamiento y una necesidad de una cancelación del solapamiento en el decodificador de audio 200. La ventana de transformada 360 también comprende una porción constante (distinta de cero) central 360c, que posee una duración temporal de, por ejemplo, 896 muestras de dominio temporal. La ventana de transformada 360 también comprende una pendiente de transición del lado derecho 360d con una duración temporal de 128 muestras de dominio temporal, que se extiende tanto en la mitad
- 15 del lado izquierdo como en la mitad del lado derecho de la porción del lado derecho de la ventana de transformada 360. Además, la ventana de transformada 360 comprende una porción cero del lado derecho 360e con una duración temporal de 448 muestras de dominio temporal, que se extiende en la mitad del lado derecho de la ventana de transformada 360.
- 20 **[0063]** Las pendientes de transición 360b, 360d de la ventana de transformada 360 están adaptadas de tal manera que se puede interponer una trama de audio, cuyos coeficientes de MDCT se obtienen mediante la ventana de transformada 360, entre una trama de audio anterior, cuyos coeficientes de MDCT se obtienen mediante la ventana de transformada 330, y una trama de audio posterior, cuyos coeficientes de MDCT se obtienen mediante la ventana de transformada 330, en la que se garantiza una cancelación de solapamiento apropiada mediante una
- 25 correlación de la pendiente de transición del lado derecho 338b con la pendiente de transición del lado izquierdo 360b, y además mediante una correlación de la pendiente de transición del lado derecho 360d con la pendiente de transición del lado izquierdo 331a (ya que las pendientes de transición están adaptadas para una transición suave de cancelación del solapamiento).
- 30 **[0064]** A continuación, se describirá la versión “modificada” de forma de ventana 364 del tipo de ventana “stop\_start\_window”. La ventana de transformada 364 comprende una porción cero del lado izquierdo 364a, que es idéntica a la porción cero del lado izquierdo 360a de la ventana de transformada 360, una pendiente de transición del lado izquierdo 364b, que es idéntica a la pendiente de transición del lado izquierdo 360b de la ventana de transformada 360, y una porción central 364c, que es idéntica a la porción central 360c de la ventana de transformada 360. No obstante, una pendiente de transición del lado derecho 364d de la ventana de transformada
- 35 364 está acortada hasta una duración temporal de 64 muestras de dominio temporal en comparación con la pendiente de transición del lado derecho 360d de la ventana de transformada 360, que comprende una duración temporal de 128 muestras de dominio temporal. Cabe señalar que la pendiente de transición del lado derecho 364d de la ventana de transformada 364 comprende normalmente la misma posición temporal y características que las porciones de transición del lado derecho 324c de la ventana de transformada 324 y 348b de la subventana 348. Además, la ventana de transformada 364 comprende una porción cero del lado derecho 364e, que es preferentemente idéntica en su posición temporal y características a las porciones cero del lado derecho 324d, 340b
- 40 de las ventanas de transformada 324, 340.
- 45 **[0065]** La ventana de transformada 364 está adaptada para una transformada de un contenido de audio de una trama de audio, que está insertada entre una trama de audio anterior, que está codificada en el dominio de frecuencia y cuyos coeficientes de MDCT se generan mediante la ventana de transformada 330 (es decir, una ventana de transformada de alta resolución temporal), y una trama de audio posterior codificada en el dominio de predicción lineal. Con este objeto, la pendiente de transición del lado izquierdo 364b de la ventana de transformada
- 50 364 se correlaciona con la pendiente de transición del lado derecho 338b de la subventana 338 de la ventana de transformada 330, para permitir una cancelación del solapamiento cuando se superponen y añaden representaciones en el dominio temporal de tramas de audio posteriores generadas (por ejemplo, en el decodificador de audio 200) mediante las ventanas de transformada 330, 364. Además, la pendiente de transición del lado derecho 364d de la ventana de transformada 364 está adaptada de tal manera que se puede llevar a cabo una
- 55 transición de superposición y adición sin aplicar una funcionalidad de cancelación de solapamiento explícita.
- [0066]** A continuación, se describirá la versión “normal” de forma de ventana 370 del tipo de ventana “stop\_window\_1152”. Cabe señalar aquí que la ventana de transformada 370 comprende una duración total de 2034 muestras de dominio temporal, y que hay 1152 coeficientes de MDCT asociados con una trama de audio codificada

mediante la ventana de transformada 370. La ventana de transformada 370 comprende una porción cero del lado izquierdo 370a con una duración temporal de 512 muestras de dominio temporal. Además, la ventana de transformada 370 comprende una pendiente de transición del lado izquierdo 370b con una duración temporal de 128 muestras de dominio temporal, que está dispuesta de tal manera que la pendiente de transición 370b se extiende tanto en una mitad del lado izquierdo como en una mitad del lado derecho de la porción del lado izquierdo de la ventana de transformada 370. Además, la ventana de transformada 370 comprende una porción central (distinta de cero) constante con una duración temporal de 576 muestras de dominio temporal. Además, la ventana de transformada 370 comprende una pendiente de transición del lado derecho 370d con una duración temporal de 1024 muestras de dominio temporal y una porción cero del lado derecho 370e con una duración temporal de 64 muestras de dominio temporal. La ventana de transformada 370 está adaptada para una codificación/decodificación de una trama de audio actual insertada entre una trama de audio anterior codificada en el dominio de predicción lineal y una trama de audio posterior codificada en el dominio de frecuencia con una resolución temporal baja (por ejemplo, mediante una ventana de transformada 310 o una ventana de transformada 320).

15 **[0067]** A continuación, se describirá la versión “normal” de forma de ventana 380 del tipo de ventana “stop\_start\_window\_1152”. Normalmente, hay 1152 coeficientes de MDCT asociados con una trama de audio de 2304 muestras de dominio temporal codificada mediante la forma de ventana 380. La ventana de transformada 380 comprende una porción cero del lado izquierdo 380a, que es idéntica en su posición temporal y características a la porción cero del lado izquierdo 370a de la ventana de transformada 370, y una pendiente de transición del lado izquierdo 380b, que es idéntica en su posición temporal y características a la pendiente de transición del lado izquierdo 370b de la ventana de transformada 370. La ventana de transformada 380 también comprende una porción central constante 380c con una duración temporal de 1024 muestras de dominio temporal. Además, la ventana de transformada 380 comprende una pendiente de transición del lado derecho 380d, que es similar en su posición y características a la pendiente de transición del lado derecho 360d de la ventana de transformada 360. La ventana de transformada 380 también comprende una porción cero del lado derecho 380e con una duración temporal de 512 muestras de dominio temporal. La ventana de transformada 380 está adaptada para ser utilizada para una codificación o decodificación de una trama de audio, que está intercalada entre una trama de audio anterior codificada en el dominio de predicción lineal y una trama de audio posterior codificada con una alta resolución temporal (por ejemplo, mediante la ventana de transformada 330).

30 **[0068]** A continuación, se describirá la versión “modificada” de forma de ventana 384 del tipo de ventana “stop\_start\_window\_1152”. Normalmente, hay 1152 coeficientes de MDCT asociados con una trama de audio codificada mediante la forma de ventana 384. La ventana de transformada 384 comprende una porción cero del lado izquierdo 384a, que es idéntica a la porción cero del lado izquierdo 380a de la ventana de transformada 380, así como una pendiente de transición del lado izquierdo 384b, que es idéntica a la pendiente de transición del lado izquierdo 380b de la ventana de transformada 380. La ventana de transformada 384 también comprende una porción central (distinta de cero) 384c, que es idéntica a la porción central constante 380c de la ventana de transformada 380. No obstante, una pendiente de transición del lado derecho 384d de la ventana de transformada 384 está acortada temporalmente en comparación con la pendiente de transición del lado derecho 380d de la ventana de transformada 380. Por ejemplo, la pendiente de transición del lado derecho 384d puede ser muy similar (o incluso idéntica) en su posición y características a las pendientes de transición del lado derecho 364d, 348d de la ventana de transformada 364 o la subventana 348 de la ventana de transformada 340 (teniendo en cuenta que la duración temporal de la ventana de transformada 384 es de 2304 muestras de dominio temporal, en lugar de 2048 muestras de dominio temporal). De este modo, la pendiente de transición 384d puede comprender una duración temporal de 64 muestras de dominio temporal, y puede estar incluida completamente en la mitad del lado izquierdo o la porción del lado derecho de la ventana de transformada 384, con lo que se evita el solapamiento. La ventana de transformada 384 también comprende una porción cero del lado derecho con una duración temporal de 576 muestras de dominio temporal. La ventana de transformada 384 está adaptada para ser utilizada por una codificación o decodificación de una trama de audio, que está intercalada entre una trama de audio anterior codificada en el dominio de predicción lineal y una trama de audio posterior codificada en el dominio de predicción lineal. Por consiguiente, la pendiente de transición del lado izquierdo 384b de la ventana de transformada está adaptada para una atenuación cruzada de cancelación de solapamiento entre una representación en el dominio temporal de una trama de audio anterior codificada en el dominio de predicción lineal y una representación en el dominio temporal de la trama de audio actual. En particular, una posición temporal de la pendiente de transición 384b está adaptada de tal manera que la pendiente de transición 384b se desvía hacia la derecha aproximadamente 128 muestras de dominio temporal con respecto a un centro entre los límites de dos tramas contiguas (que se muestra mediante líneas verticales finas continuas). Además, la pendiente de transición del lado derecho 384d de la ventana de transformada 384 está adaptada de tal manera que se puede llevar a cabo una transición de superposición y adición sin aplicar una funcionalidad de cancelación de solapamiento explícita.



### 3.2. Tipo de ventana "long start window to LPD"

**[0069]** A continuación, se describirán más detalles relativos a algunos tipos de ventanas de especial importancia, haciendo referencia a las figs. 4a-4d.

**[0070]** A continuación, se expondrán algunos detalles relativos a la aplicación de la versión "modificada" de forma de ventana del tipo de ventana "long\_start\_window", que también se indica mediante el número de referencia 324 y mediante la denominación de "long\_start\_window\_to\_LPD", haciendo referencia a la fig. 2a.

**[0071]** En primer lugar, cabe señalar que la ventana de transformada 324 está adaptada para su aplicación en un decodificador de audio para proporcionar una representación en el dominio temporal de una trama de audio, que está intercalada entre una trama de audio anterior codificada en el dominio de frecuencia y una trama de audio posterior codificada en el dominio de predicción lineal. Una porción de la representación en el dominio temporal de la trama de audio actual, a la que se aplica la pendiente de transición del lado izquierdo 324a, normalmente se superpone y añade a una representación en el dominio temporal de una trama de audio anterior codificada en el dominio de frecuencia, a la que se aplica una pendiente de transición del lado derecho de una ventana de transformada apropiada (por ejemplo, de la ventana de transformada 310 o de la ventana de transformada 350 o de la ventana de transformada 370). Por consiguiente, se obtiene una cancelación del solapamiento debida a la correlación de las pendientes de transición. En cambio, una porción de la representación en el dominio temporal de la trama de audio actual, a la que se aplica la pendiente de transición 324c, se superpone y añade a una versión enventanada (pero no procesada con solapamiento en el dominio temporal) de una representación en el dominio temporal de la trama de audio posterior codificada en el dominio de predicción lineal. Por consiguiente, se obtienen transiciones suavizadas entre las representaciones en el dominio temporal de la trama de audio anterior y la trama de audio actual y entre las representaciones en el dominio temporal del audio actual y la trama de audio posterior.

### 3.3 Tipo de ventana "8\*short window to LPD"

**[0072]** A continuación, se describirán algunos detalles relativos a la aplicación de la versión "modificada" de forma de ventana 340 del tipo de ventana "8\*short\_window" haciendo referencia a la fig. 4b. Como puede observarse en la fig. 4b, la ventana de transformada 340 está adaptada para su aplicación en un decodificador de audio, en el que la ventana de transformada "modificada" 340 (que en ocasiones también se denomina "8\*short\_window\_to\_LPD") está adaptada para proporcionar una representación en el dominio temporal de una trama de audio actual codificada en el dominio de predicción lineal, y dicha trama de audio actual codificada en el dominio de predicción lineal está codificada con una resolución temporal comparativamente alta, y dicha trama de audio actual codificada en el dominio de predicción lineal está adaptada para ser intercalada entre una trama de audio anterior codificada en el dominio de frecuencia y una trama de audio posterior codificada en el dominio de predicción lineal. Preferentemente, la representación en el dominio temporal de la trama de audio anterior codificada en el dominio de frecuencia se obtiene mediante una ventana de transformada 320, una ventana de transformada 360 o una ventana de transformada 380 (si la trama de audio anterior está codificada con una resolución temporal comparativamente más baja) o mediante una ventana de transformada 330 (si la trama de audio anterior está codificada con una resolución temporal comparativamente más alta). Una porción de la representación en el dominio temporal de la trama de audio actual, a la que se aplica la pendiente de transición 341a, se superpone y añade en el decodificador de audio 200 a una porción de la representación en el dominio temporal de la trama de audio anterior, a la que se aplica una de las pendientes de transición 320c, 360b, 380b, 338b. Por consiguiente, se suaviza una transición entre la representación en el dominio temporal de la trama de audio anterior y la representación en el dominio temporal de la trama de audio actual y se lleva a cabo una cancelación del solapamiento. Una porción temporal de la representación en el dominio temporal de la trama de audio actual, a la que se aplica la pendiente de transición 348b, se superpone y añade a una versión enventanada (pero no procesada con solapamiento en el dominio temporal) de una representación en el dominio temporal de la trama de audio posterior codificada en el dominio de predicción lineal. Por consiguiente, se suaviza una transición entre la trama de audio actual codificada en el dominio de predicción lineal, que comprende una resolución temporal comparativamente alta, y la trama de audio posterior codificada en el dominio de predicción lineal.

### 3.4. Tipo de ventana "stop start window to LPD"

**[0073]** A continuación, se describirá, haciendo referencia a la fig. 4c, la aplicación de la versión "modificada" de forma de ventana 364 del tipo de ventana "stop\_start\_window", que en ocasiones también se denomina "stop\_start\_window\_to\_LPD".

**[0074]** La ventana de transformada 364 es aplicada por un decodificador de audio para proporcionar una representación en el dominio temporal de una trama de audio codificada con una resolución temporal comparativamente baja, y dicha trama de audio está intercalada entre una trama de audio anterior codificada en el dominio de frecuencia, preferentemente con una resolución temporal comparativamente más alta, y una trama de audio posterior codificada en el dominio de predicción lineal. Por ejemplo, la trama de audio anterior se puede obtener mediante una versión “normal” de forma de ventana 330 del tipo de ventana “8\*short\_window”. No obstante, en algunos casos, la trama de audio anterior puede estar codificada mediante una versión “normal” de forma de ventana 320, 360 o 380 del tipo de ventana “long\_start\_window”, “stop\_start\_window” o “stop\_start\_window\_1152”.

5 En un decodificador de audio, se puede llevar a cabo una operación de superposición y adición entre muestras de dominio temporal de la trama de audio actual, a la que se aplica la pendiente de transición del lado izquierdo 364b de la ventana de transformada 364, y las muestras de dominio temporal de una representación en el dominio temporal de la trama de audio anterior codificada en el dominio de frecuencia, a la que ha aplicado una pendiente de transición 338b de la ventana de transformada 330 (o, como otra posibilidad, una pendiente de transición 320c de la ventana de transformada 320, una pendiente de transición 360b de la ventana de transformada 360 o una pendiente de transición 380b de la ventana de transformada 380). Por consiguiente, se suaviza una transición entre la representación en el dominio temporal de la trama de audio anterior y la representación en el dominio temporal de la trama de audio actual. Además, se lleva a cabo una superposición y adición en un decodificador de audio entre muestras de dominio temporal de una representación en el dominio temporal de la trama de audio actual, a la que se ha aplicado la pendiente de transición 364d de la ventana de transformada 364, y muestras de dominio temporal de una representación en el dominio temporal de la trama de audio posterior codificada en el dominio de predicción lineal (en la que se puede aplicar una ventana de transición con una correlación establecida a las muestras de dominio temporal de la trama de audio siguiente codificada en el dominio de predicción lineal antes de la operación de superposición y adición). Por consiguiente, se puede suavizar una transición entre la representación en el dominio temporal de la trama de audio actual y la representación en el dominio temporal de la trama de audio posterior sin que sea necesario aplicar un mecanismo de cancelación del solapamiento en esta transición.

### 3.5. Tipo de ventana “stop start window 1152 LPD”

**[0075]** A continuación, se describirá la aplicación de la versión “modificada” de forma de ventana 384 del tipo de ventana “stop\_start\_window\_1152” en un decodificador de audio. La ventana de transformada 384 se utiliza en un decodificador de audio para obtener una representación en el dominio temporal de una trama de audio con una resolución temporal comparativamente baja, que está intercalada entre una trama de audio anterior codificada en el dominio de predicción lineal y una trama de audio posterior también codificada en el dominio de predicción lineal. De este modo, la ventana de transformada 384 permite la inserción de una única trama de audio codificada en el dominio de frecuencia entre dos tramas de audio codificadas en el dominio de predicción lineal.

**[0076]** El decodificador de audio 200 está configurado preferentemente para llevar a cabo una operación de superposición y adición entre muestras de dominio temporal de la representación en el dominio temporal de la trama de audio actual a las cuales se les aplica la pendiente de transición 384b y muestras de dominio temporal de una representación en el dominio temporal de la trama de audio anterior codificada en el dominio de predicción lineal. Antes de llevar a cabo la operación de superposición y adición, se puede aplicar un procesamiento de solapamiento en dominio temporal a la representación en el dominio temporal de la trama de audio anterior codificada en el dominio de predicción lineal, y dicho procesamiento de solapamiento en el dominio temporal puede incluir la inserción de componentes de solapamiento en el dominio temporal y la aplicación de una pendiente de ventana a las muestras de dominio temporal en la zona de superposición temporal. Por consiguiente, se obtiene una transición suavizada entre la trama de audio anterior codificada en el dominio de predicción lineal y la trama de audio actual codificada en el dominio de frecuencia. Además, se puede llevar a cabo una operación de superposición y adición mediante el decodificador de audio 200 entre una porción temporal de la representación en el dominio temporal de la trama de audio actual, a la que se aplica la pendiente de transición 384d, y muestras de dominio temporal de una representación en el dominio temporal de la trama de audio posterior codificada en el dominio de predicción lineal (en la que se puede aplicar un enventanado a la representación en el dominio temporal de la trama de audio codificada en el dominio de predicción lineal antes de la ejecución de la operación de superposición y adición). Por consiguiente, se puede obtener una transición suavizada entre la representación en el dominio temporal de la trama de audio actual codificada en el dominio de frecuencia y la representación en el dominio temporal de la trama de audio posterior codificada en el dominio de predicción lineal.

### 3.6. Secuencias de ventanas permitidas – Visión general

**[0077]** A continuación, se ofrecerá una visión general de secuencias de ventanas (en el sentido de secuencias de tipos de ventanas 310, 320, 324, 330, 340, 350, 360, 364, 370, 380, 384) que son admitidas en un codificador de audio 100 o un decodificador de audio 200 de acuerdo con la presente invención. Sin embargo, cabe señalar que no es necesario aplicar todas las ventanas de transformada descritas con respecto a las figs. 3 y 4a-4d y todas las secuencias de ventanas descritas en referencia a la fig. 5 en un codificador de audio de acuerdo con un ejemplo o un decodificador de audio de acuerdo con la presente invención.

**[0078]** La fig. 5 muestra una representación esquemática de transiciones permitidas entre tramas de audio codificadas mediante diferentes tipos de ventanas, e incluso entre tramas de audio codificadas en diferentes dominios. Cabe señalar que las expresiones "only\_long\_sequence", "long\_start\_sequence", "eight\_short\_sequence", "long\_stop\_sequence", "stop\_start\_sequence", "stop\_1152\_sequence" y "stop\_start\_1152\_sequence" se utilizan en la fig. 5, y son equivalentes a los tipos de ventana "long\_window", "long\_start\_window", "8\*short\_window", "long\_stop\_window", "stop\_start\_window", "stop\_window\_1152" y "stop\_start\_window\_1152", como puede observarse en la fig. 3. Además, cabe señalar que la expresión "LPD\_sequence" indica una trama de audio codificada en el dominio de predicción lineal. Además, cabe señalar que se utiliza una forma de ventana "normal" 310, 320, 330, 350, 360, 370, 380 para la codificación o decodificación de una trama de audio actual, si la trama de audio actual va seguida de una trama de audio posterior codificada en el dominio de frecuencia, y que se utiliza una forma de ventana "modificada" 324, 348, 364, 384 para la codificación o decodificación de la trama de audio actual, si la trama de audio actual va seguida de una trama de audio posterior codificada en el dominio de predicción lineal.

**[0079]** Como puede observarse en la fig. 5, una trama de audio codificada mediante el tipo de ventana "long\_window" puede ir seguida de una trama de audio codificada mediante el tipo de ventana "long\_window" o "long\_start\_window". Una trama de audio codificada mediante el tipo de ventana "long\_start\_window" puede ir seguida de una trama de audio codificada mediante el tipo de ventana "8\*short\_window", "long\_stop\_window" o "stop\_start\_window". No obstante, la trama de audio codificada mediante el tipo de ventana "long\_start\_window" también puede ir seguida de una trama de audio codificada en el dominio de predicción lineal, mediante la forma de ventana "modificada" 324.

**[0080]** Una trama de audio codificada mediante el tipo de ventana "8\*short\_window" puede ir seguida de una trama de audio codificada mediante el tipo de ventana "8\*short\_window", "long\_stop\_window" o "stop\_start\_window". No obstante, al utilizar la ventana de transformada 340, una trama de audio codificada mediante el tipo de ventana "8\*short\_window" también puede ir seguida de una trama de audio codificada en el dominio de predicción lineal.

**[0081]** Una trama de audio codificada mediante el tipo de ventana "long\_stop\_window" puede ir seguida de una trama de audio codificada mediante el tipo de ventana "long\_window" o "long\_start\_window".

**[0082]** Una trama de audio codificada mediante el tipo de ventana "stop\_start\_window" puede ir seguida de una trama de audio codificada mediante el tipo de ventana "8\*short\_window", "long\_stop\_window" o "stop\_start\_window". No obstante, la trama de audio codificada mediante el tipo de ventana "stop\_start\_window" también puede ir seguida de una trama de audio codificada en el dominio de predicción lineal mediante la ventana de transformada "modificada" 364.

**[0083]** Una trama de audio codificada mediante el tipo de ventana "stop\_window\_1152" puede ir seguida de una trama de audio posterior codificada mediante el tipo de ventana "long\_window" o "long\_start\_window".

**[0084]** Una trama de audio codificada mediante el tipo de ventana "stop\_start\_window\_1152" puede ir seguida de una trama de audio posterior codificada mediante el tipo de ventana "8\*short\_window", "long\_stop\_window" o "stop\_start\_window". No obstante, una trama de audio codificada mediante el tipo de ventana "stop\_start\_window\_1152" también puede ir seguida de una trama de audio codificada en el dominio de predicción lineal mediante la ventana de transformada "modificada" 384.

**[0085]** Una trama de audio actual codificada en el dominio de predicción lineal también puede ir seguida por una trama de audio posterior codificada en el dominio de predicción lineal o por una trama de audio codificada en el dominio de frecuencia mediante el tipo de ventana "stop\_window\_1152" o "stop\_start\_window\_1152".

**[0086]** A continuación, se describirán más detalladamente algunas secuencias de tramas de audio posibles.

3.7. Transición desde una trama de audio codificada en el dominio de frecuencia con baja resolución temporal a una trama de audio codificada en el dominio de predicción lineal

**[0087]** A continuación, se describirá una secuencia de ventanas de transformada que conlleva la realización de una transición indirecta (con una trama intermedia interpuesta) desde una trama de audio codificada en el dominio de frecuencia hasta una trama de audio codificada en el dominio de predicción lineal. Cabe señalar que, en las siguientes explicaciones, las tramas se indican mediante números de trama posterior con el fin de poder identificar las tramas.

**[0088]** Tomando ahora como referencia la fig. 6a, que muestra una representación esquemática de una primera secuencia de ventanas de transformada, se describirá un caso en el que una trama de audio codificada en el dominio de predicción lineal viene precedida de una pluralidad de tramas de audio codificadas en el dominio de frecuencia con una resolución temporal comparativamente baja. Como puede observarse, una primera trama de audio 610 está codificada en el dominio de frecuencia con una resolución temporal comparativamente baja y mediante la ventana de transformada 310. Una segunda trama de audio posterior 620, que se superpone temporalmente (por ejemplo, un 50%) a la primera trama de audio 610, está codificada en el dominio de frecuencia mediante la ventana de transformada 324. Se lleva a cabo una superposición y adición (en un decodificador de audio 200) entre representaciones en el dominio temporal de los contenidos de audio de la primera y la segunda trama de audio 610, 620 en la zona de superposición temporal. Una tercera trama de audio 630, que se superpone temporalmente a la segunda trama de audio 620 (por ejemplo, un 50%) está codificada en el dominio de predicción lineal. Se lleva a cabo una operación de superposición y adición entre la representación en el dominio temporal de la segunda trama de audio 620 y la representación en el dominio temporal del contenido de audio de la tercera trama de audio 630 (representada por parámetros de dominio de predicción lineal). Con este objeto, se aplica un enventanado de pendiente de transición (representada con el número de referencia 630a) a una representación en el dominio temporal del contenido de audio de la tercera trama de audio 630. La tercera trama de audio 630 va seguida de una cuarta trama de audio 640, que puede estar codificada en el dominio de frecuencia (tal como se muestra en la fig. 6a) o en el dominio de predicción lineal.

**[0089]** La secuencia de tramas de audio codificadas que se muestra en la fig. 6a resulta útil en una situación en la que no hay transiciones escalonadas que precedan de forma cercana (en las una o dos tramas anteriores) a una trama de audio de tipo habla codificada en el dominio de predicción lineal.

**[0090]** No obstante, se describirán, haciendo referencia a las figs. 6b y 6d, secuencias de ventanas de transformada que resultan más adecuadas si una trama de audio de tipo habla viene precedida de un transitorio significativo (por ejemplo, transición escalonada) del contenido de audio.

3.8. Transición directa desde trama de audio codificada en el dominio de frecuencia con alta resolución temporal hasta trama de audio codificada en el dominio de predicción lineal

**[0091]** La fig. 6b muestra una representación esquemática de una secuencia de ventanas de transformada que aporta una eficiencia de codificación y una calidad de audio mejoradas si hay un transitorio significativo en el contenido de audio de una trama de audio que precede (directamente) a una trama de audio de tipo habla. Se ha descubierto que esta situación es relativamente frecuente, debido a que la aparición de una porción de audio de tipo habla a menudo viene precedida de una interrupción abrupta de sonidos de fondo, tales como ruido de fondo o música instrumental. Como puede observarse en la fig. 6b, una primera trama de audio 650 puede estar, por ejemplo, codificada en el dominio de frecuencia con una resolución temporal baja (tal como se muestra en la fig. 6b) o una resolución temporal alta (que no se muestra). Una segunda trama de audio posterior 652 está codificada en el dominio de frecuencia con una resolución temporal comparativamente alta. La segunda trama de audio 652 está codificada mediante la ventana de transformada 340, que se ha descrito anteriormente. La alta resolución temporal de la segunda trama de audio 652 se obtiene mediante una pluralidad de subventanas 341-348, a las que se asocian unos conjuntos diferenciados (cortos) de coeficientes de MDCT (por ejemplo, 128 coeficientes de MDCT por subventana). Es importante el hecho de que una pendiente de transición de la ventana de transformada esté adaptada para proporcionar una transición suavizada a una tercera trama de audio 654, que está codificada en el dominio de predicción lineal. Como puede observarse, se lleva a cabo una operación de superposición y adición, en un decodificador de audio, entre representaciones en el dominio temporal de un contenido de audio de la segunda trama de audio 652 (que se decodifica mediante la ventana de transformada 340) y la representación en el dominio temporal de un contenido de audio de la tercera trama de audio 654. Se aplica un enventanado a la representación en el dominio temporal del contenido de audio de la tercera trama de audio 654 que se indica con el número de referencia 654a. La tercera trama de audio 654 viene seguida de una cuarta trama de audio 656, que puede estar codificada en el dominio de predicción lineal, o que puede estar codificada en el dominio de frecuencia (por ejemplo, mediante la ventana de transformada 370, la ventana de transformada 380 o la ventana de transformada 384).

**[0092]** En resumen, la secuencia de ventanas de transformada de la fig. 6b, que comprende la ventana de transformada 340 con forma de ventana "modificada" del tipo "8\*short\_window" permite una transición directa entre la segunda trama de audio 652, que está codificada en el dominio de frecuencia con una resolución temporal comparativamente alta, y la tercera trama de audio 654 codificada en el dominio de predicción lineal.

3.9. Única trama de audio codificada en el dominio de frecuencia entre tramas de audio codificadas en el dominio de predicción lineal

10 **[0093]** A continuación, se describirá otra secuencia importante de ventanas de transformada haciendo referencia a la fig. 6c, que muestra una representación gráfica de dicha secuencia de ventanas de transformada. Como puede observarse en la fig. 6c, una primera trama de audio 660 está codificada en el dominio de predicción lineal. Una segunda trama de audio 662 está codificada en el dominio de frecuencia, en la que la ventana de transformada 384 se utiliza para codificar y decodificar la segunda trama de audio 662. La segunda trama de audio  
 15 662 va seguida de una tercera trama de audio 664, que está codificada en el dominio de predicción lineal. Como puede observarse, la pendiente de transición del lado izquierdo 384b de la ventana de transformada 384 (que se utiliza para codificar y decodificar la segunda trama de audio 662) está adaptada para llevar a cabo una operación de superposición y adición de cancelación de solapamiento entre la representación en el dominio temporal del contenido de audio de la primera trama de audio 660 y la representación en el dominio temporal del contenido de  
 20 audio de la segunda trama de audio 662. Para permitir dicha operación de superposición y adición de cancelación de solapamiento, se asocia un mayor número de coeficientes de MDCT (por ejemplo, 1152 coeficientes de MDCT) con la segunda trama de audio 662 (en comparación, por ejemplo, con los 1024 coeficientes de MDCT asociados con las tramas de audio codificadas en el dominio de frecuencia intercaladas entre dos tramas de audio contiguas codificadas en el dominio de frecuencia). Por consiguiente, se lleva a cabo una superposición y adición de  
 25 cancelación de solapamiento entre las representaciones en el dominio temporal del contenido de audio de la primera y la segunda trama de audio 660, 662, en la que se aplica un procesamiento de solapamiento en el dominio temporal y un enventanado al contenido de audio de la primera trama de audio 660. Una tercera trama de audio 664 está codificada en el dominio de predicción lineal, y se lleva a cabo una operación de superposición y adición en un decodificador de audio 200 entre las representaciones del dominio temporal de la segunda trama de audio 662 y de  
 30 la tercera trama de audio 664. Con este objeto, se aprovecha la pendiente de transición 384d de la ventana de transformada 380. Además, se aplica un enventanado a la representación en el dominio temporal de la tercera trama de audio 664.

**[0094]** La secuencia de ventanas de la fig. 6c permite la inserción de una única trama de audio codificada en el dominio de frecuencia entre tramas de audio adyacentes codificadas en el dominio de predicción lineal, en la que  
 35 se puede llevar a cabo una superposición y adición apropiada tanto en la transición desde la primera trama de audio 660 a la segunda trama de audio 662 como en la transición desde la segunda trama de audio 662 a la tercera trama de audio 664. Concretamente, las pendientes de transición de la ventana de transformada 380 están adaptadas de tal manera que se lleva a cabo una superposición y adición eficiente en cuanto a la tasa de bits con cancelación de  
 40 solapamiento entre la primera trama de audio 660 y la segunda trama de audio 662, y de tal manera que se puede llevar a cabo en la transición desde la segunda trama de audio 662 a la tercera trama de audio 664 una superposición y adición computacionalmente eficiente y con baja distorsión sin necesidad de una cancelación de solapamiento. Esto se logra mediante pendientes de transición 384b, 384d con una duración temporal diferente y asociando un mayor número de coeficientes de MDCT a la segunda trama de audio 662 (1152 coeficientes de  
 45 MDCT en lugar de 1024 coeficientes de MDCT). Por consiguiente, es posible codificar ruido de fondo estacionario en las interrupciones (o pausas) entre dos tramas de audio de tipo habla con una buena eficiencia en cuanto a la tasa de bits, al tiempo que se mantiene la posibilidad de obtener transiciones suaves en un decodificador de audio.

3.10. Transición desde trama de audio codificada en el dominio de frecuencia con alta resolución temporal a una  
 50 trama de audio codificada en el dominio de predicción lineal a través de una trama de audio interpuesta codificada en el dominio de frecuencia

**[0095]** A continuación, se describirá otra secuencia ventajosa de ventanas de transformada haciendo referencia a la fig. 6d, que muestra una representación esquemática de dicha secuencia de ventanas de transformada. Una primera trama de audio 670 está codificada en el dominio de frecuencia, por ejemplo, con una  
 55 resolución temporal comparativamente baja. Por ejemplo, la ventana de transformada 320 se puede aplicar para la codificación y la decodificación de la primera trama de audio 670. Una segunda trama de audio 672 está codificada en el dominio de frecuencia con una resolución temporal comparativamente alta. Por ejemplo, se utiliza una ventana de transformada 330 para la codificación y la decodificación de la segunda trama de audio 672. Una tercera trama

de audio 674 está codificada en el dominio de frecuencia con una resolución temporal comparativamente más baja. No obstante, en lugar de utilizar la ventana de transformada 360, se utiliza la ventana de transformada “modificada” 364 para codificar y decodificar la tercera trama de audio 674. Por consiguiente, se proporcionan unas pendientes de transición 338b, 364b correlacionadas, en la transición desde la segunda trama de audio 672 a la tercera trama de audio 674, de manera que se pueda llevar a cabo allí una superposición y adición de cancelación de solapamiento. Una cuarta trama de audio 676 está codificada en el dominio de predicción lineal. No obstante, una pendiente de transición 364d de la ventana de transformada 364 está adaptada para llevar a cabo una operación de superposición y adición con la cuarta trama de audio 676 sin que sea necesaria una cancelación de solapamiento.

10 **[0096]** La secuencia de ventanas de la fig. 6d permite una transición indirecta (con la tercera trama de audio 674 interpuesta) entre una trama de audio codificada en el dominio de frecuencia con una alta resolución temporal 672 y una trama de audio codificada en el dominio de predicción lineal 676, en la que una trama de audio codificada en el dominio de frecuencia con una baja resolución temporal 674 se encuentra entre las tramas de audio 672, 676. Dicha secuencia de ventanas resulta ventajosa, por ejemplo, si un suceso transitorio en una señal de audio, por ejemplo, una reducción rápida del volumen de un ruido de fondo, está separado de una trama de audio de tipo habla 676 por una trama de audio intermedio de no habla 674, en la que el ruido de fondo es aproximadamente estacionario. Aunque la utilización de la ventana de transformada 324 no permitiría una codificación de audio eficiente en cuanto a la tasa de bits, que represente el suceso transitorio con una buena calidad de audio, la utilización de la ventana de transformada 364 en el codificador y el decodificador permite obtener un muy buen equilibrio entre tasa de bits y calidad de audio en dichos casos.

#### 4. Detalles de la aplicación del decodificador

25 **[0097]** A continuación, se describirán algunos detalles relativos a la funcionalidad del núcleo decodificador de dominio de frecuencia 250 del decodificador de audio 200. Además, se describirán algunos detalles del procesador de superposición y adición 270. Estas funcionalidades a menudo también se denominan “banco de filtros y conmutación de bloques”.

##### 4.1. Descripción de la herramienta

30 **[0098]** Se establece una correspondencia de la representación en tiempo/frecuencia 242 de la señal al dominio temporal introduciéndola en el módulo de banco de filtros 250a. Este módulo consiste en una transformada discreta de coseno modificada (MDCT), y una ventana y una función de superposición y adición. Para adaptar la resolución en tiempo/frecuencia del banco de filtros a las características de la señal de entrada, también se adopta una herramienta de conmutación de bloques. N representa la longitud de la ventana, donde N es una función de la **window\_sequence**.

40 **[0099]** Dependiendo de la señal, el codificador puede cambiar la resolución en tiempo/frecuencia mediante tres la utilización de diferentes tamaños de ventana: 2304, 2048 y 256. Para conmutar entre ventanas, se utilizan las ventanas de transición LONG\_START\_WINDOW, LONG\_STOP\_WINDOW, STOP\_WINDOW\_1152, STOP\_START\_WINDOW y STOP\_START\_WINDOW\_1152. En la fig. 3 se ofrece una relación de las ventanas, se especifica la correspondiente longitud de la transformada y se muestra la forma de las ventanas de manera esquemática. Se utilizan tres longitudes de transformada: 1152, 1024 (o 960) (a las que se hace referencia como transformada larga) y 128 (o 120) coeficientes (a la que se hace referencia como transformada corta).

45 **[0100]** Las secuencias de ventanas están compuestas por ventanas de manera que un bloque de datos en bruto (raw\_data\_block) siempre contiene datos que representan 1024 (o 960) muestras de salida. El elemento de datos **window\_sequence** indica la secuencia de ventanas concreta que se utiliza. En la fig. 3 se ofrece una relación del modo en que las secuencias de ventanas (también denominadas “ventanas de transformada”) están compuestas por 50 ventanas individuales (también denominadas “subventanas”).

**[0101]** Para cada canal, los N/2 valores de tiempo/frecuencia  $X_{i,k}$  se transforman en los N valores en el dominio temporal  $x_{i,n}$  a través de la IMDCT. Tras aplicar la función de ventana, para cada canal, la primera mitad de la secuencia  $z_{i,n}$  se añade a la segunda mitad de la secuencia anterior enventanada en bloque  $z_{(i-1),n}$  para reconstruir 55 las muestras de salida para cada canal de salida  $out_{i,n}$ .

##### 4.2. Definiciones

**[0102]**

**window\_sequence** 2 bits que indican qué secuencia de ventanas (es decir, tamaño de bloque) se utiliza.  
**window\_shape** 1 bit que indica qué función de ventana se selecciona.

5 **[0103]** La fig. 3 muestra las once secuencias de ventanas *window\_sequences* basadas en las siete ventanas de transformada.

10 (ONLY\_LONG\_SEQUENCE, LONG\_START\_SEQUENCE,  
 EIGHT\_SHORT\_SEQUENCE, LONG\_STOP\_SEQUENCE,  
 STOP\_START\_SEQUENCE, STOP\_1152\_SEQUENCE,  
 STOP\_START\_1152\_SEQUENCE).

**[0104]** En lo sucesivo, *LPD\_SEQUENCE* hace referencia a todas las combinaciones permitidas de ventanas/modos de codificación dentro del denominado códec de dominio de predicción lineal. En el contexto de la  
 15 decodificación de una trama codificada en el dominio de frecuencia, es importante conocer únicamente si una siguiente trama está codificada con los modos de codificación del dominio LP (de predicción lineal), que está representada por una *LPD\_SEQUENCE*. No obstante, se atiende a la estructura exacta del interior de la *LPD\_SEQUENCE* cuando se decodifica la trama codificada en el dominio LP.

20 4.3. Proceso de decodificación

4.3.1. IMDCT (transformada de coseno discreta modificada inversa)

**[0105]** La expresión analítica de la IMDCT es:  
 25

$$x_{i,n} = \frac{2}{N} \sum_{k=0}^{\frac{N}{2}-1} spec[1][k] \cos\left(\frac{2\pi}{N} (n + n_0) \left(k + \frac{1}{2}\right)\right) \quad para \quad 0 \leq n < N$$

en la que:

30 n = índice de muestras  
 i = índice de ventanas  
 k = índice de coeficientes espectrales  
 N = longitud de ventana basada en el valor de *window\_sequence*

35 
$$n_0 = (N/2 + 1)/2$$

**[0106]** La longitud de la ventana de síntesis N para la transformada inversa es una función del elemento sintáctico **window\_sequence** y el contexto algorítmico. Se define del siguiente modo:

40 Longitud de ventana 2304:

**[0107]**

45 
$$N = \begin{cases} 2304, & \text{si STOP 1152 SEQUENCE} \\ 2304, & \text{si STOP\_START\_1152\_SEQUENCE} \end{cases}$$

**[0108]** Longitud de ventana 2048:

$$N = \begin{cases} 2048, & \text{si ONLY\_LONG\_SEQUENCE} \\ 2048, & \text{si LONG\_START\_SEQUENCE} \\ 256, & \text{si EIGHT\_SHORT\_SEQUENCE} \\ 2048, & \text{si LONG\_STOP\_SEQUENCE} \\ 2048, & \text{si STOP\_START\_SEQUENCE} \end{cases}$$

[0109] La relación de las transiciones de bloques significativas aparece en la tabla de la fig. 5. Una marca ( ) en una determinada celda de la tabla indica que una secuencia de ventanas que aparece en esa fila concreta puede ir seguida de una secuencia de ventanas que aparece en esa columna concreta.

#### 4.3.2. Enventanado y conmutación de bloques

[0110] Dependiendo del elemento **window\_sequence** y **window\_shape**, se utilizan diferentes ventanas de transformada. Una combinación de las mitades de ventana que se describe a continuación ofrece todas las posibles window\_sequences. La forma de ventana (*window shape*) describe la forma de las denominadas pendientes de transición.

[0111] Para **window\_shape** = 1, los coeficientes de la ventana vienen dados por la ventana derivada de Kaiser-Bessel (KBD) del siguiente modo:

$$W_{KBD\_LEFT,N}(n) = \sqrt{\frac{\sum_{p=0}^n [W'(p,\alpha)]}{\sum_{p=0}^{N/2} [W'(p,\alpha)]}} \quad \text{para } 0 \leq n < \frac{N}{2}$$

$$W_{KBD\_RIGHT,N}(n) = \sqrt{\frac{\sum_{p=0}^{N-n-1} [W'(p,\alpha)]}{\sum_{p=0}^{N/2} [W'(p,\alpha)]}} \quad \text{para } \frac{N}{2} \leq n < N$$

20 en las que:

$W'$ , función de ventana de núcleo de Kaiser-Bessel, se define del siguiente modo:

$$25 \quad W'(n,\alpha) = \frac{I_0 \left[ \pi \alpha \sqrt{1,0 - \left( \frac{n - N/4}{N/4} \right)^2} \right]}{I_0[\pi \alpha]} \quad \text{para } 0 \leq n < \frac{N}{2}$$



$$I_0[x] = \sum_{k=0}^{\infty} \left[ \frac{\left(\frac{x}{2}\right)^k}{k!} \right]^2$$

5  $\alpha =$  factor alfa de ventana de kernel,  $\alpha = \begin{cases} 4 & \text{para } N = 2048 \text{ (1920)} \\ 6 & \text{para } N = 256 \text{ (240)} \end{cases}$

[0112] En caso contrario, para **window\_shape** = 0, se emplea una ventana senoidal del siguiente modo:

10 
$$W_{SIN\_LEFT,N}(n) = \sin\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)\right) \quad \text{para } 0 \leq n < \frac{N}{2}$$

$$W_{SIN\_RIGHT,N}(n) = \sin\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)\right) \quad \text{para } \frac{N}{2} \leq n < N$$

15 [0113] La longitud de ventana N puede ser 2048 (1920) o 256 (240) para la KBD y la ventana senoidal. En el caso de STOP\_1152\_SEQUENCE y STOP\_START\_1152\_SEQUENCE, N aún puede ser 2048 o 256, las pendientes de ventana (o pendientes de transición) son similares, pero las zonas con la parte superior plana son más largas.

[0114] En las partes a)-g) de este apartado, se explica cómo obtener las posibles secuencias de ventanas.

20 [0115] Para todos los tipos de secuencias de ventanas, se determina la window\_shape de la mitad izquierda de la primera ventana de transformada mediante la forma de ventana del bloque (o trama de audio) anterior. Este hecho viene expresado por la siguiente fórmula:

25 
$$W_{LEFT,N}(n) = \begin{cases} W_{KBD\_LEFT,N}(n), & \text{si } window\_shape\_previous\_block = 1 \\ W_{SIN\_LEFT,N}(n), & \text{si } window\_shape\_previous\_block = 0 \end{cases}$$

en la que:

30 *window\_shape\_previous\_block*: **window\_shape** del bloque o trama de audio anterior (i-1).

[0116] Para el primer raw\_data\_block (bloque de datos en bruto) (o trama de audio) que se va a decodificar, la **window\_shape** de las mitades izquierda y derecha de la ventana son idénticas.

35 [0117] En el caso de que el bloque o trama de audio anterior estuviera codificado mediante el modo LPD (dominio de predicción lineal), *window\_shape\_previous\_block* se fija en 0.  
a) ONLY\_LONG\_SEQUENCE:

40 [0118] La **window\_sequence** == ONLY\_LONG\_SEQUENCE es igual a una LONG\_WINDOW con una longitud total de ventana *N\_I* de 2048 (1920).

[0119] Para **window\_shape** = 1, la ventana para ONLY\_LONG\_SEQUENCE viene dada de la siguiente manera:

45 
$$W(n) = \begin{cases} W_{LEFT,N_I}(n), & \text{para } 0 \leq n < N_I/2 \\ W_{KBD\_RIGHT,N_I}(n), & \text{para } N_I/2 \leq n < N_I \end{cases}$$

**[0120]** Si **window\_shape** = 0, la ventana para ONLY\_LONG\_SEQUENCE se puede describir de la siguiente manera:

$$W(n) = \begin{cases} W_{LEFT, N_l}(n), & \text{para } 0 \leq n < N_l/2 \\ W_{SIN\_RIGHT, N_l}(n), & \text{para } N_l/2 \leq n < N_l \end{cases}$$

**[0121]** Tras el enventanado, los valores del dominio temporal ( $z_{i,n}$ ) se pueden expresar como:

$$z_{i,n} = w(n) \cdot x_{i,n}$$

**b) LONG\_START\_SEQUENCE:**

**[0122]** La LONG\_START\_SEQUENCE se puede utilizar para obtener una correcta superposición y adición para una transición de bloques desde una ONLY\_LONG\_SEQUENCE a cualquier bloque con una mitad de ventana con una baja superposición (pendiente de ventana corta) en la izquierda (EIGHT\_SHORT\_SEQUENCE, LONG\_STOP\_SEQUENCE, STOP\_START\_SEQUENCE o LPD\_SEQUENCE).

**[0123]** Las longitudes de ventana  $N_l$  y  $N_s$  están fijadas en 2048 (1920) y 256 (240) respectivamente.

**[0124]** En caso de que la siguiente secuencia de ventanas no sea una LPD\_SEQUENCE:

si **window\_shape** = 1, la ventana para LONG\_START\_SEQUENCE viene dada de la siguiente manera:

$$W(n) = \begin{cases} W_{LEFT, N_l}(n), & \text{para } 0 \leq n < N_l/2 \\ 1,0, & \text{para } N_l/2 \leq n < \frac{3N_l - N_s}{4} \\ W_{KBD\_RIGHT, N_s}(n + \frac{N_s}{2} - \frac{3N_l - N_s}{4}), & \text{para } \frac{3N_l - N_s}{4} \leq n < \frac{3N_l + N_s}{4} \\ 0,0, & \text{para } \frac{3N_l + N_s}{4} \leq n < N_l \end{cases}$$

si **window\_shape** = 0, la ventana para LONG\_START\_SEQUENCE tiene este aspecto:

$$W(n) = \begin{cases} W_{LEFT, N_l}(n), & \text{para } 0 \leq n < N_l/2 \\ 1,0, & \text{para } N_l/2 \leq n < \frac{3N_l - N_s}{4} \\ W_{SIN\_RIGHT, N_s}(n + \frac{N_s}{2} - \frac{3N_l - N_s}{4}), & \text{para } \frac{3N_l - N_s}{4} \leq n < \frac{3N_l + N_s}{4} \\ 0,0, & \text{para } \frac{3N_l + N_s}{4} \leq n < N_l \end{cases}$$

**[0125]** Los valores del dominio temporal enventanados se pueden calcular con la fórmula explicada en a).

**[0126]** En caso de que la siguiente secuencia de ventanas sea una LPD\_SEQUENCE:

si **window\_shape** = 1, la ventana para LONG\_START\_SEQUENCE viene dada de la siguiente manera:

$$W(n) = \begin{cases} W_{LEFT,N_l}(n), & \text{para } 0 \leq n < \frac{N_l}{2} \\ 1,0, & \text{para } \frac{N_l}{2} \leq n < \frac{3N_l - N_s}{4} \\ W_{KBD\_RIGHT, N_s/2} \left( n + \frac{N_s}{4} - \frac{3N_l - N_s}{4} \right), & \text{para } \frac{3N_l - N_s}{4} \leq n < \frac{3N_l}{4} \\ 0,0, & \text{para } \frac{3N_l}{4} \leq n < N_l \end{cases}$$

si **window\_shape** = 0, la ventana para LONG\_START\_SEQUENCE tiene este aspecto:

$$W(n) = \begin{cases} W_{LEFT,N_l}(n), & \text{para } 0 \leq n < \frac{N_l}{2} \\ 1,0, & \text{para } \frac{N_l}{2} \leq n < \frac{3N_l - N_s}{4} \\ W_{SIN\_RIGHT, N_s/2} \left( n + \frac{N_s}{4} - \frac{3N_l - N_s}{4} \right), & \text{para } \frac{3N_l - N_s}{4} \leq n < \frac{3N_l}{4} \\ 0,0, & \text{para } \frac{3N_l}{4} \leq n < N_l \end{cases}$$

5

c) EIGHT\_SHORT

**[0127]** La **window\_sequence** = EIGHT\_SHORT comprende ocho SHORT\_WINDOWS (ventanas cortas) superpuestas y añadidas con una longitud  $N_s$  de 256 (240) cada una. La longitud total de la **window\_sequence** junto con los ceros a la izquierda y a la derecha es de 2048 (1920). Cada uno de los ocho bloques cortos se enventana por separado en primer lugar. El número de bloques cortos se indexa con la variable  $j = 0, \dots, M-1$  ( $M=N_l/N_s$ ).

15 **[0128]** En el caso de que la siguiente secuencia de ventanas sea una LPD\_SEQUENCE:

la **window\_shape** del bloque anterior influye solamente en el primero de los ocho bloques cortos ( $W_0(n)$ ). Si **window\_shape** = 1, las funciones de ventana pueden venir dadas de la siguiente manera:

$$W_0(n) = \begin{cases} W_{LEFT, N_s}(n), & \text{para } 0 \leq n < N_s/2 \\ W_{KBD\_RIGHT, N_s}(n), & \text{para } N_s/2 \leq n < N_s \end{cases}$$

20

$$W_j(n) = \begin{cases} W_{LEFT, N_s}(n), & \text{para } 0 \leq n < N_s/2 \\ W_{KBD\_RIGHT, N_s}(n), & \text{para } N_s/2 \leq n < N_s, \quad 0 < j < (M-1) \end{cases}$$

25

$$W_{M-1}(n) = \begin{cases} W_{KBD\_LEFT, N_s}(n), & \text{para } 0 \leq n < N_s/2 \\ W_{KBD\_RIGHT, N_s/2}(n), & \text{para } N_s/2 \leq n < 3N_s/4 \end{cases}$$

**[0129]** En caso contrario, si **window\_shape** = 0, las funciones de ventana se pueden describir como:

30

$$W_0(n) = \begin{cases} W_{LEFT, N_s}(n), & \text{para } 0 \leq n < N_s/2 \\ W_{SIN\_RIGHT, N_s}(n), & \text{para } N_s/2 \leq n < N_s \end{cases}$$

$$W_j(n) = \begin{cases} W_{SIN\_LEFT, N_s}(n), & \text{para } 0 \leq n < N_s/2 \\ W_{SIN\_RIGHT, N_s}(n), & \text{para } N_s/2 \leq n < N_s, 0 < j < (M-1) \end{cases}$$

$$W_{M-1}(n) = \begin{cases} W_{SIN\_LEFT, N_s}(n), & \text{para } 0 \leq n < N_s/2 \\ W_{SIN\_RIGHT, N_s/2}(n), & \text{para } N_s/2 \leq n < 3N_s/4 \end{cases}$$

5

**[0130]** La superposición y adición entre la **window sequence** EIGHT\_SHORT que da lugar a los valores de dominio temporal enventanados  $Z_{i,n}$  se describe de la siguiente manera:

$$Z_{i,n} = \begin{cases} 0, & \text{para } 0 \leq n < \frac{N_l - N_s}{4} \\ x_{0,n-\frac{N_l-N_s}{4}} \cdot W_0\left(n - \frac{N_l - N_s}{4}\right), & \text{para } \frac{N_l - N_s}{4} \leq n < \frac{N_l + N_s}{4} \\ x_{j-1,n-\frac{N_l+(2j-3)N_s}{4}} \cdot W_{j-1}\left(n - \frac{N_l+(2j-3)N_s}{4}\right) + x_{j,n-\frac{N_l+(2j-1)N_s}{4}} \cdot W_j\left(n - \frac{N_l+(2j-1)N_s}{4}\right), & \text{para } 1 \leq j < M, \frac{N_l+(2j-1)N_s}{4} \leq n < \frac{N_l+(2j+1)N_s}{4} \\ x_{M-1,n-\frac{N_l+(2M-3)N_s}{4}} \cdot W_{M-1}\left(n - \frac{N_l+(2M-3)N_s}{4}\right), & \text{para } \frac{N_l+(2M-1)N_s}{4} \leq n < \frac{N_l+(2M)N_s}{4} \\ 0, & \text{para } \frac{N_l+(2M)N_s}{4} \leq n < N_l \end{cases}$$

10

**[0131]** En todos los demás casos:

La **window\_shape** del bloque anterior influye solamente en el primero de los ocho bloques cortos ( $W_0(n)$ ). Si **window\_shape** = 1, las funciones de ventana pueden venir dadas de la siguiente manera:

15

$$W_0(n) = \begin{cases} W_{LEFT, N_s}(n), & \text{para } 0 \leq n < N_s/2 \\ W_{KBD\_RIGHT, N_s}(n), & \text{para } N_s/2 \leq n < N_s \end{cases}$$

$$W_j(n) = \begin{cases} W_{KDB\_LEFT, N_s}(n), & \text{para } 0 \leq n < N_s/2 \\ W_{KBD\_RIGHT, N_s}(n), & \text{para } N_s/2 \leq n < N_s, 0 < j < M-1 \end{cases}$$

20

**[0132]** En caso contrario, si **window\_shape** = 0, las funciones de ventana se pueden describir de la siguiente manera:

25

$$W_0(n) = \begin{cases} W_{LEFT, N_s}(n), & \text{para } 0 \leq n < N_s/2 \\ W_{SIN\_RIGHT, N_s}(n), & \text{para } N_s/2 \leq n < N_s \end{cases}$$

$$W_j(n) = \begin{cases} W_{SIN\_LEFT, N_s}(n), & \text{para } 0 \leq n < N_s/2 \\ W_{SIN\_RIGHT, N_s}(n), & \text{para } N_s/2 \leq n < N_s, 0 < j \leq M-1 \end{cases}$$

**[0133]** La superposición y adición entre la **window\_sequence** EIGHT\_SHORT que da lugar a los valores de dominio temporal eventanados  $Z_{i,n}$  se describe de la siguiente manera:

5

$$Z_{i,n} = \begin{cases} 0, & \text{para } 0 \leq n < \frac{N_l - N_s}{4} \\ x_{0,n-\frac{N_l-N_s}{4}} \cdot W_0\left(n - \frac{N_l - N_s}{4}\right), & \text{para } \frac{N_l - N_s}{4} \leq n < \frac{N_l + N_s}{4} \\ x_{j-1,n-\frac{N_l+(2j-3)N_s}{4}} \cdot W_{j-1}\left(n - \frac{N_l+(2j-3)N_s}{4}\right) + x_{j,n-\frac{N_l+(2j-1)N_s}{4}} \cdot W_j\left(n - \frac{N_l+(2j-1)N_s}{4}\right), & \text{para } 1 \leq j < M, \frac{N_l+(2j-1)N_s}{4} \leq n < \frac{N_l+(2j+1)N_s}{4} \\ x_{M-1,n-\frac{N_l+(2M-3)N_s}{4}} \cdot W_{M-1}\left(n - \frac{N_l+(2M-3)N_s}{4}\right), & \text{para } \frac{N_l+(2M-1)N_s}{4} \leq n < \frac{N_l+(2M+1)N_s}{4} \\ 0, & \text{para } \frac{N_l+(2M+1)N_s}{4} \leq n < N_l \end{cases}$$

d) LONG\_STOP\_SEQUENCE

10 **[0134]** Esta secuencia de ventanas es necesaria para volver a conmutar de una EIGHT\_SHORT\_SEQUENCE a una ONLY\_LONG\_SEQUENCE.

**[0135]** Si **window\_shape** = 1, la ventana para LONG\_STOP\_SEQUENCE viene dada de la siguiente manera:

$$W(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l - N_s}{4} \\ W_{LEFT, N_s}\left(n - \frac{N_l - N_s}{4}\right), & \text{para } \frac{N_l - N_s}{4} \leq n < \frac{N_l + N_s}{4} \\ 1,0, & \text{para } \frac{N_l + N_s}{4} \leq n < N_l/2 \\ W_{KBD\_RIGHT, N_l}(n), & \text{para } N_l/2 \leq n < N_l \end{cases}$$

15

**[0136]** Si **window\_shape** = 0, la ventana para LONG\_START\_SEQUENCE se determina mediante:

$$W(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l - N_s}{4} \\ W_{LEFT, N_s}\left(n - \frac{N_l - N_s}{4}\right), & \text{para } \frac{N_l - N_s}{4} \leq n < \frac{N_l + N_s}{4} \\ 1,0, & \text{para } \frac{N_l + N_s}{4} \leq n < N_l/2 \\ W_{SIN\_RIGHT, N_l}(n), & \text{para } N_l/2 \leq n < N_l \end{cases}$$

20

**[0137]** Los valores de dominio temporal eventanados se pueden calcular con la fórmula explicada en a).

e) STOP\_START\_SEQUENCE

**[0138]** La STOP\_START\_SEQUENCE se puede utilizar para obtener una superposición y adición correcta para una transición de bloques desde cualquier bloque con una mitad de ventana con baja superposición (pendiente de 5 ventana corta) en la derecha hasta cualquier bloque con una mitad de ventana con una baja superposición (pendiente de ventana corta) en la izquierda y si se desea una única transformada larga para la trama actual.

**[0139]** Las longitudes de ventana  $N_l$  y  $N_s$  están fijadas en 2048 (1920) y 256 (240) respectivamente.

10 **[0140]** En caso de que la siguiente secuencia de ventanas no sea una LPD\_SEQUENCE:

Si **window\_shape** = 1, la ventana para STOP\_START\_SEQUENCE viene dada de la siguiente manera:

$$W(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l - N_s}{4} \\ W_{LEFT, N_s} \left( n - \frac{N_l - N_s}{4} \right), & \text{para } \frac{N_l - N_s}{4} \leq n < \frac{N_l + N_s}{4} \\ 1,0, & \text{para } \frac{N_l + N_s}{4} \leq n < \frac{3N_l - N_s}{4} \\ W_{KBD\_RIGHT, N_s} \left( n + \frac{N_s}{2} - \frac{3N_l - N_s}{4} \right), & \text{para } \frac{3N_l - N_s}{4} \leq n < \frac{3N_l + N_s}{4} \\ 0,0, & \text{para } \frac{3N_l + N_s}{4} \leq n < N_l \end{cases}$$

15

si **window\_shape** = 0, la ventana para STOP\_START\_SEQUENCE tiene este aspecto:

$$W(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l - N_s}{4} \\ W_{LEFT, N_s} \left( n - \frac{N_l - N_s}{4} \right), & \text{para } \frac{N_l - N_s}{4} \leq n < \frac{N_l + N_s}{4} \\ 1,0, & \text{para } \frac{N_l + N_s}{4} \leq n < \frac{3N_l - N_s}{4} \\ W_{SIN\_RIGHT, N_s} \left( n + \frac{N_s}{2} - \frac{3N_l - N_s}{4} \right), & \text{para } \frac{3N_l - N_s}{4} \leq n < \frac{3N_l + N_s}{4} \\ 0,0, & \text{para } \frac{3N_l + N_s}{4} \leq n < N_l \end{cases}$$

20 **[0141]** En caso de que la siguiente secuencia de ventanas sea una LPD\_SEQUENCE:

si **window\_shape** = 1, la ventana para STOP\_START\_SEQUENCE viene dada de la siguiente manera:

$$W(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l - N_s}{4} \\ W_{LEFT, N_s} \left( n - \frac{N_l - N_s}{4} \right), & \text{para } \frac{N_l - N_s}{4} \leq n < \frac{N_l + N_s}{4} \\ 1,0, & \text{para } \frac{N_l + N_s}{4} \leq n < \frac{3N_l - N_s}{4} \\ W_{KBD\_RIGHT, \frac{N_s}{2}} \left( n + \frac{N_s}{2} - \frac{3N_l - N_s}{4} \right), & \text{para } \frac{3N_l - N_s}{4} \leq n < \frac{3N_l}{4} \\ 0,0, & \text{para } \frac{3N_l}{4} \leq n < N_l \end{cases}$$

si **window\_shape** = 0, la ventana para STOP\_START\_SEQUENCE tiene este aspecto:

$$5 \quad W(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l - N_s}{4} \\ W_{LEFT, N_s} \left( n - \frac{N_l - N_s}{4} \right), & \text{para } \frac{N_l - N_s}{4} \leq n < \frac{N_l + N_s}{4} \\ 1,0, & \text{para } \frac{N_l + N_s}{4} \leq n < \frac{3N_l - N_s}{4} \\ W_{SIN\_RIGHT, \frac{N_s}{2}} \left( n + \frac{N_s}{2} - \frac{3N_l - N_s}{4} \right), & \text{para } \frac{3N_l - N_s}{4} \leq n < \frac{3N_l}{4} \\ 0,0, & \text{para } \frac{3N_l}{4} \leq n < N_l \end{cases}$$

**[0142]** Los valores del dominio temporal enventanados se pueden calcular con la fórmula explicada en a).

f) STOP\_1152\_SEQUENCE

10

**[0143]** La STOP\_1152\_SEQUENCE es necesaria para obtener una superposición y adición correcta para una transición de bloques desde una LPD\_SEQUENCE a una ONLY\_LONG\_SEQUENCE.

**[0144]** Las longitudes de ventana  $N_l$  y  $N_s$  están fijadas en 2048 (1920) y 256 (240) respectivamente.

15

Si **window\_shape** = 1, la ventana para STOP\_1152\_SEQUENCE viene dada de la siguiente manera:

$$W(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l}{4} \\ W_{LEFT, N_s} \left( n - \frac{N_l}{4} \right), & \text{para } \frac{N_l}{4} \leq n < \frac{N_l + 2N_s}{4} \\ 1,0, & \text{para } \frac{N_l + 2N_s}{4} \leq n < \frac{2N_l + 3N_s}{4} \\ W_{KBD\_RIGHT, N_l} \left( n + \frac{N_l}{2} - \frac{2N_l + 3N_s}{4} \right), & \text{para } \frac{2N_l + 3N_s}{4} \leq n < N_l + \frac{3N_s}{4} \\ 0,0, & \text{para } N_l + \frac{3N_s}{4} \leq n < N_l + N_s \end{cases}$$

20 **[0145]** Si **window\_shape** = 0, la ventana para STOP\_1152\_SEQUENCE tiene este aspecto:

$$W(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l}{4} \\ W_{LEFT, N_s}(n - \frac{N_l}{4}), & \text{para } \frac{N_l}{4} \leq n < \frac{N_l + 2N_s}{4} \\ 1,0, & \text{para } \frac{N_l + 2N_s}{4} \leq n < \frac{2N_l + 3N_s}{4} \\ W_{SIN_RIGHT, N_l}(n + \frac{N_l}{2} - \frac{2N_l + 3N_s}{4}), & \text{para } \frac{2N_l + 3N_s}{4} \leq n < N_l + \frac{3N_s}{4} \\ 0,0, & \text{para } N_l + \frac{3N_s}{4} \leq n < N_l + N_s \end{cases}$$

5 [0146] Los valores del dominio temporal inventanados se pueden calcular con la fórmula explicada en a).

g) STOP\_START\_1152\_SEQUENCE

10 [0147] La STOP\_START\_1152\_SEQUENCE se puede utilizar para obtener una superposición y adición correcta para una transición de bloques desde una LPD\_SEQUENCE a cualquier bloque con una mitad de ventana con una baja superposición (pendiente de ventana corta) en la izquierda.

[0148] Las longitudes de ventana  $N_l$  y  $N_s$  están fijadas en 2048 (1920) y 256 (240) respectivamente.

15 [0149] En caso de que la siguiente secuencia de ventanas no sea una LPD\_SEQUENCE:

Si **window\_shape** = 1, la ventana para STOP\_START\_1152\_SEQUENCE viene dada de la siguiente manera:

$$W(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l}{4} \\ W_{LEFT, N_s}(n - \frac{N_l}{4}), & \text{para } \frac{N_l}{4} \leq n < \frac{N_l + 2N_s}{4} \\ 1,0, & \text{para } \frac{N_l + 2N_s}{4} \leq n < \frac{3N_l}{4} + \frac{N_s}{2} \\ W_{KBD_RIGHT, N_s}(n + \frac{N_s}{2} - \frac{3N_l}{4} + \frac{N_s}{2}), & \text{para } \frac{3N_l}{4} + \frac{N_s}{2} \leq n < \frac{3N_l}{4} + N_s \\ 0,0, & \text{para } \frac{3N_l}{4} + N_s \leq n < N_l + N_s \end{cases}$$

20

Si **window\_shape** = 0, la ventana para STOP\_START\_1152\_SEQUENCE tiene este aspecto:



$$W(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l}{4} \\ W_{LEFT, N_s}\left(n - \frac{N_l}{4}\right), & \text{para } \frac{N_l}{4} \leq n < \frac{N_l + 2N_s}{4} \\ 1,0, & \text{para } \frac{N_l + 2N_s}{4} \leq n < \frac{3N_l}{4} + \frac{N_s}{2} \\ W_{SIN\_RIGHT, N_s}\left(n + \frac{N_s}{2} - \frac{3N_l}{4} + \frac{N_s}{2}\right), & \text{para } \frac{3N_l}{4} + \frac{N_s}{2} \leq n < \frac{3N_l}{4} + N_s \\ 0,0, & \text{para } \frac{3N_l}{4} + N_s \leq n < N_l + N_s \end{cases}$$

**[0150]** En caso de que la siguiente secuencia de ventanas sea una LPD\_SEQUENCE:

5 Si **window\_shape = 1**, la ventana para STOP\_START\_1152\_SEQUENCE viene dada de la siguiente manera:

$$W(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l}{4} \\ W_{LEFT, N_s}\left(n - \frac{N_l}{4}\right), & \text{para } \frac{N_l}{4} \leq n < \frac{N_l + 2N_s}{4} \\ 1,0, & \text{para } \frac{N_l + 2N_s}{4} \leq n < \frac{3N_l}{4} + \frac{N_s}{2} \\ W_{KBD\_RIGHT, N_s/2}\left(n + \frac{N_s}{2} - \frac{3N_l}{4} + \frac{N_s}{2}\right), & \text{para } \frac{3N_l}{4} + \frac{N_s}{2} \leq n < \frac{3N_l}{4} + \frac{3N_s}{4} \\ 0,0, & \text{para } \frac{3N_l}{4} + \frac{3N_s}{4} \leq n < N_l + N_s \end{cases}$$

Si **window\_shape = 0**, la ventana para STOP\_START\_1152\_SEQUENCE tiene este aspecto:

10

$$W(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l}{4} \\ W_{LEFT, N_s}\left(n - \frac{N_l}{4}\right), & \text{para } \frac{N_l}{4} \leq n < \frac{N_l + 2N_s}{4} \\ 1,0, & \text{para } \frac{N_l + 2N_s}{4} \leq n < \frac{3N_l}{4} + \frac{N_s}{2} \\ W_{SIN\_RIGHT, N_s/2}\left(n + \frac{N_s}{2} - \frac{3N_l}{4} + \frac{N_s}{2}\right), & \text{para } \frac{3N_l}{4} + \frac{N_s}{2} \leq n < \frac{3N_l}{4} + \frac{3N_s}{4} \\ 0,0, & \text{para } \frac{3N_l}{4} + \frac{3N_s}{4} \leq n < N_l + N_s \end{cases}$$

**[0151]** Los valores del dominio temporal enventanados se pueden calcular con la fórmula explicada en a).

15 4.3.3. Superposición y adición con secuencia de ventana anterior

**[0152]** Aparte de la superposición y adición en el interior de la **window\_sequence** EIGHT\_SHORT, la primera

(izquierda) parte (o "porción") de cada **window\_sequence** anterior se superpone y añade a la segunda (derecha) parte (o "porción") de la **window\_sequence** anterior que da lugar a los valores finales de dominio temporal  $out_{i,n}$ . La expresión matemática para esta operación se puede describir de la siguiente manera.

- 5 **[0153]** En el caso de que ONLY\_LONG\_SEQUENCE, LONG\_START\_SEQUENCE, EIGHT\_SHORT\_SEQUENCE, LONG\_STOP\_SEQUENCE y STOP\_START\_SEQUENCE:

$$out_{i,n} = z_{i,n} + z_{i-1, n + \frac{N}{2}}; \quad para 0 \leq n < \frac{N}{2}, \quad N = 2048 \quad (1920)$$

- 10 **[0154]** En el caso de STOP\_1152\_SEQUENCE, STOP\_START\_1152\_SEQUENCE:

$$out_{i,n} = z_{i,n} + z_{i-1, n + \frac{N-l}{2} + \frac{3N-s}{4}}; \quad para 0 \leq n < \frac{N-l}{2}, \quad N-l = 2048, N-s = 256$$

- 15 **[0155]** En caso de ir desde el modo FD (dominio de frecuencia) al modo LPD (dominio de predicción lineal), dependiendo de la **window\_sequence** del último bloque de modo FD, se aplica una ventana SIN (seno) (si **window\_sequence** es 0) o KBD (si **window\_sequence** es 1) en la parte izquierda de la primera LPD\_SEQUENCE para tener una superposición y adición correcta con la trama anterior.

20

$$W_{SIN\_LEFT, N}(n) = \sin\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)\right) \quad para 0 \leq n < \frac{N}{2} \quad con N = 128$$

$$W_{KBD\_LEFT, N}(n) = \sqrt{\frac{\sum_{p=0}^n [W'(p, \alpha)]}{\sum_{p=0}^{N/2} [W'(p, \alpha)]}} \quad para 0 \leq n < \frac{N}{2} \quad con N = 128$$

- [0156]** En caso de STOP\_1152\_SEQUENCE, STOP\_START\_1152\_SEQUENCE, la secuencia anterior es una LPD\_SEQUENCE. En este caso, es necesario añadir artificialmente componentes de solapamiento de dominio temporal a la señal de dominio temporal decodificada con el fin de cancelar los correspondientes componentes de TDA de la siguiente trama de modo FD. Para facilitararlo, es necesario enventanar el extremo derecho de la LPD\_SEQUENCE anterior con una ventana SIN (indicada por *window\_shape\_previous\_block*), plegarlo, desplegarlo y volverlo a enventanar por medio de MDCT/IMDCT antes de la operación de superposición y adición con la siguiente trama de acuerdo con la fig. 9.

#### 5. Procedimiento para proporcionar una representación de audio codificada

- [0157]** A continuación, se describirá un procedimiento para proporcionar una representación de audio codificada haciendo referencia a la fig. 7, que muestra un diagrama de flujo de dicho procedimiento.

- [0158]** El procedimiento 700 de la fig. 7 para proporcionar una representación codificada de un contenido de audio a partir de una representación de audio de entrada del contenido de audio comprende la provisión selectiva 710 de un conjunto de parámetros de dominio de predicción lineal o un conjunto de parámetros de dominio de frecuencia a partir de una representación en el dominio temporal de una trama de audio que se va a codificar y dependiendo de una información que indique si una trama de audio actual se va a codificar en el dominio de predicción lineal o en el dominio de frecuencia. El procedimiento 700 comprende la codificación 720 de tramas de audio posteriores en diferentes dominios y la consideración 730 de una ventana de transformada de entre un conjunto que comprende una pluralidad de diferentes ventanas de transformada para proporcionar un conjunto de parámetros de dominio de frecuencia asociado con una trama de audio que se va a codificar en el dominio de frecuencia.

[0159] En un primer ejemplo preferido del procedimiento 700, se utiliza una **ventana de inserción** 384 para una generación de un conjunto de parámetros de dominio de frecuencia de una trama de audio actual que se va a codificar en el dominio de frecuencia, si la trama de audio actual está intercalada entre una trama de audio anterior que se va a codificar en el dominio de predicción lineal y una trama de audio posterior que se va a codificar en el dominio de predicción lineal. Una pendiente de transición del lado izquierdo de la ventana de inserción está específicamente adaptada para proporcionar una transición suave entre una representación en el dominio temporal de la trama de audio anterior codificada en el dominio de predicción lineal y una representación en el dominio temporal de la trama de audio actual codificada en el dominio de frecuencia. Una pendiente de transición del lado derecho de la ventana de inserción está adaptada para proporcionar una transición suave entre la representación en el dominio de frecuencia de la trama de audio actual codificada en el dominio de frecuencia y la representación en el dominio temporal de la trama de audio posterior codificada en el dominio de predicción lineal. En otras palabras, la ventana de transformada 384 se utiliza en un primer ejemplo del procedimiento 700. Por consiguiente, es posible obtener la secuencia de tramas de audio y ventanas de transformada, lo cual se ha explicado en referencia a la fig. 6c. Por consiguiente, se pueden obtener las ventajas expuestas anteriormente.

[0160] En un segundo ejemplo del procedimiento 700, el conjunto de ventanas de transformada comprende tipos de ventanas con diferentes resoluciones temporales adaptadas para una generación de un conjunto de parámetros de dominio de frecuencia de una trama de audio que se va a codificar en el dominio de frecuencia y que comprende una transición hacia una trama de audio posterior que va a codificar en el dominio de predicción lineal. Por ejemplo, ambas ventanas de transformada 324 y 340 pueden estar disponibles. Por consiguiente, las secuencias de tramas de audio y ventanas de transformada que se muestran en las figs. 6a y 6b pueden ser ambas obtenibles, de manera que se puede obtener una codificación eficiente en cuanto a la tasa de bits con una buena calidad de audio en diferentes situaciones, independientemente de si se produce o no un suceso transitorio en una trama de audio codificada en el dominio de frecuencia que preceda a una trama de audio codificada en el dominio de predicción lineal.

[0161] En un tercer ejemplo del procedimiento 700, el conjunto de ventanas de transformada comprende una **ventana de transición** 364 adaptada para una generación de un conjunto de parámetros de dominio de frecuencia a partir de una representación en el dominio temporal de una trama de audio actual, si la trama de audio actual viene precedida de una trama de audio anterior que se va a codificar en el dominio de frecuencia mediante un conjunto de alta resolución temporal de parámetros de dominio de frecuencia y si la trama de audio actual comprende una transición hacia una representación en el dominio temporal de una trama de audio posterior que se va a codificar en el dominio de predicción lineal. En otras palabras, el tercer ejemplo del procedimiento 700 utiliza la ventana de transformada 364 con el fin de obtener la secuencia de tramas de audio y ventanas de transformada que se muestra en la fig. 6d. De este modo, se puede obtener una codificación eficiente, aunque se produzca un suceso transitorio durante la penúltima trama de audio antes de una primera trama de audio codificada en el dominio de predicción lineal (trama de audio de tipo habla).

6. Procedimiento para proporcionar una representación de audio decodificada

[0162] A continuación, se describirá un procedimiento para proporcionar una representación de audio decodificada, haciendo referencia a la fig. 8, que muestra un diagrama de flujo de dicho procedimiento 800. El procedimiento 800 comprende la provisión selectiva 810 de representaciones en el dominio temporal de tramas de audio a partir de un conjunto de parámetros de dominio de predicción lineal o a partir de un conjunto de parámetros de dominio de frecuencia, en el que se tiene en cuenta una ventana de transformada de entre un conjunto que comprende una pluralidad de diferentes ventanas de transformada cuando se proporciona la representación en el dominio temporal de una trama de audio. El procedimiento 800 también comprende la realización 820 de una operación de superposición y adición de las representaciones en el dominio temporal de tramas de audio posteriores codificadas en diferentes dominios para suavizar una transición entre las representaciones en el dominio temporal de las tramas de audio posteriores codificadas en diferentes dominios.

[0163] En una primera realización del procedimiento 800, se selecciona una **ventana de inserción** 384 como ventana de transformada para la generación de una representación en el dominio temporal de una trama de audio codificada en el dominio de frecuencia e intercalada temporalmente entre una trama anterior codificada en el dominio de predicción lineal y una trama de audio posterior codificada en el dominio de predicción lineal. Una pendiente de transición del lado izquierdo de la ventana de inserción está adaptada para proporcionar una transición suave entre una representación en el dominio temporal de la trama anterior codificada en el dominio de predicción lineal y la representación en el dominio temporal de la trama de audio actual codificada en el dominio de frecuencia.

Una pendiente de transición del lado derecho de la ventana de inserción está adaptada para proporcionar una transición suave entre la representación en el dominio temporal de la trama actual codificada en el dominio de frecuencia y una representación en el dominio temporal de la trama de audio posterior codificada en el dominio de predicción lineal. Por consiguiente, se puede decodificar la secuencia de tramas de audio y ventanas de transformada que se muestra en la fig. 6c.

**[0164]** En una segunda realización del procedimiento 800, se utilizan tipos de ventanas con diferentes resoluciones temporales para la generación de las representaciones en el dominio temporal de tramas de audio codificadas en el dominio de frecuencia que comprenden una transición hacia una representación en el dominio temporal de una trama de audio codificada en el dominio de predicción lineal. Por consiguiente, de acuerdo con la segunda realización del procedimiento 800, se decodifican secuencias de tramas de audio como las que se muestran en las figs. 6a y 6b.

**[0165]** En una tercera realización del procedimiento 800, se utiliza una ventana de transición adaptada para la generación de una representación en el dominio temporal de una trama de audio actual codificada en el dominio de frecuencia, con el fin de proporcionar una representación en el dominio temporal de una trama de audio actual codificada en el dominio de frecuencia que viene precedida de una trama de audio anterior codificada en el dominio de frecuencia mediante un conjunto de alta resolución temporal de parámetros de dominio de frecuencia y que comprende una transición hacia una representación en el dominio temporal de una trama de audio posterior codificada en el dominio de predicción lineal. Por consiguiente, se decodifica una secuencia de tramas de audio como la que se muestra en la fig. 6d.

**[0166]** Cabe señalar aquí que los procedimientos 700, 800 se pueden complementar mediante cualquiera de las características y funcionalidades expuestas en la presente memoria con respecto a los aparatos de la invención y las ventanas de transformada de la invención.

## 7. Conclusión

**[0167]** Las realizaciones de acuerdo con la presente invención crean una mejora de la transición desde un modo de codificación en el dominio de frecuencia a un modo de codificación en el dominio de predicción lineal. En algunas realizaciones sencillas, la transición desde la codificación en el dominio de frecuencia al modo de codificación predictiva lineal se lleva a cabo introduciendo la denominada "LPD\_start\_sequence", que actúa como ventana de transición para la trama inmediatamente anterior a la primera trama de una "LPD\_sequence". La "LPD\_start\_sequence" es, en la práctica, una "long\_start\_sequence" con una mitad de ventana derecha modificada.

**[0168]** No obstante, técnicamente, una "LPD\_sequence" también puede ir precedida de una "stop\_start\_sequence", una "eight\_short\_sequence" o una "stop\_start\_1152\_sequence". En estos casos, las ventanas de transformada se ajustan, de acuerdo con la presente invención, de manera similar al caso de la "LPD\_start\_sequence".

**[0169]** Por ejemplo, una transición en la que se utiliza la "LPD\_start\_sequence" se maneja correctamente, tal como se muestra en la fig. 10a.

**[0170]** No obstante, es importante señalar que la superposición y adición se debería aplicar correctamente en el caso de una transición desde una "stop\_start\_1152\_sequence" a una "LPD\_sequence". La pendiente derecha de ventana de la "stop\_start\_1152\_sequence" no debería ser demasiado larga con el fin de evitar componentes de solapamiento en el dominio temporal que no se puedan cancelar mediante la contribución de la "LPD\_sequence".

**[0171]** En la fig. 10b se muestra una representación gráfica de una transición correcta entre una trama de audio codificada mediante el tipo de ventana "stop\_start\_window\_1152" y una trama de audio posterior codificada en el dominio de predicción lineal.

**[0172]** En el caso de una pendiente derecha corta de ventana, el enventanado tras la transformada discreta de coseno modificada inversa (IMDCT) y el plegado no se aplican de manera simultánea para la mitad de ventana derecha e izquierda. En lugar de ello, el enventanado de la derecha se aplica justo antes de realizar la superposición-adición con la siguiente trama.

**[0173]** Cuando se utiliza un banco de filtros con alineamiento temporal, la aplicación de la pendiente de transición del dominio de frecuencia (FD) al dominio de predicción lineal (LPD) se debe realizar de manera ligeramente distinta,

ya que el enventanado de la parte derecha se debe aplicar antes de remuestrear y, por lo tanto, no se puede posponer. En este caso, cuando una trama de dominio de predicción lineal viene precedida de una trama de dominio de predicción no lineal, se aplica una proporción entre una pendiente corta normal y la pendiente de transición de dominio de frecuencia a dominio de predicción lineal antes de superponer-añadir sobre los datos de la trama previa para lograr los mismos resultados.

**[0174]** Cabe señalar aquí que, de acuerdo con la presente invención, la transición desde el modo de dominio de frecuencia al modo de codificación de dominio de predicción lineal se describe de forma más general que en otras estrategias, lo cual hace que la transición sea más constata, más flexible y más fácil de comprender.

10

#### 8. Alternativas de aplicación

**[0175]** Aunque se han descrito algunos aspectos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del procedimiento correspondiente, en el que un bloque o dispositivo corresponde a una etapa del procedimiento o una característica de una etapa del procedimiento. De manera análoga, los aspectos descritos en el contexto de una etapa de un procedimiento también representan una descripción de un bloque correspondiente o elemento o característica de un aparato correspondiente. Algunas o todas las etapas del procedimiento se pueden ejecutar mediante (o utilizando) un aparato físico, como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, una o más de las más importantes etapas del procedimiento se pueden ejecutar mediante este tipo de aparato.

15

20

**[0176]** La señal de audio codificada ejemplar se puede almacenar en un medio de almacenamiento digital o se puede transmitir por un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable, como, por ejemplo, Internet.

25

**[0177]** Dependiendo de ciertos requisitos de aplicación, las realizaciones de la invención se pueden aplicar en forma de equipo físico o de programa informático. La aplicación se puede llevar a cabo mediante un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, en el que se almacenen señales de control legibles por ordenador, que cooperen (o sean capaces de cooperar) con un sistema informático programable, de tal manera que se lleve a cabo el procedimiento respectivo. Por lo tanto, el medio de almacenamiento digital puede ser leído por ordenador.

30

**[0178]** Algunas realizaciones de acuerdo con la invención comprenden un soporte de datos que cuenta con unas señales de control legibles por ordenador, que son capaces de cooperar con un sistema informático programable, de manera que se lleve a cabo uno de los procedimientos descritos en la presente memoria.

35

**[0179]** En general, las realizaciones de la presente invención se pueden aplicar en forma de producto informático con un código de programa, siendo el código de programa capaz de llevar a cabo uno de los procedimientos cuando el producto informático se ejecuta en un ordenador. El código de programa puede estar almacenado, por ejemplo, en un soporte legible por máquina.

40

**[0180]** Otras realizaciones comprenden el producto informático para llevar a cabo uno de los procedimientos descritos en la presente memoria, almacenado en un soporte legible por máquina.

**[0181]** En otras palabras, una realización del procedimiento de la invención consiste, por tanto, en un programa informático que posee un código de programa para llevar a cabo uno de los procedimientos descritos en la presente memoria, cuando el programa informático se ejecuta en un ordenador.

45

**[0182]** Otra realización del procedimiento de la invención consiste, por tanto, en un soporte de datos (o un medio de almacenamiento digital o un medio legible por ordenador) que comprende, registrado en el mismo, el programa informático para llevar a cabo uno de los procedimientos descritos en la presente memoria.

50

**[0183]** Otra realización del procedimiento de la invención consiste, por tanto, en una corriente de datos o secuencia de señales que representan el programa informático para llevar a cabo uno de los procedimientos descritos en la presente memoria. La corriente de datos o secuencia de señales puede estar, por ejemplo, configurada para ser transferida a través de una conexión de comunicación de datos, por ejemplo, a través de Internet.

55

**[0184]** Otra realización comprende unos medios de procesamiento, por ejemplo, un ordenador, o un dispositivo

lógico programable, configurados o adaptados para llevar a cabo uno de los procedimientos descritos en la presente memoria.

**[0185]** Otra realización comprende un ordenador que tiene instalado en el mismo el programa informático para llevar a cabo uno de los procedimientos descritos en la presente memoria.

**[0186]** En algunas realizaciones, se puede utilizar un dispositivo lógico programable (por ejemplo, una matriz de puertas programables o FPGA) para llevar a cabo algunas o todas las funcionalidades de los procedimientos descritos en la presente memoria. En algunas realizaciones, una matriz de puertas programables puede cooperar con un microprocesador con el fin de llevar a cabo uno de los procedimientos descritos en la presente memoria. Por lo general, los procedimientos se llevan a cabo preferentemente mediante cualquier aparato físico.

**[0187]** Las realizaciones descritas anteriormente tienen un carácter meramente ilustrativo de los principios de la presente invención. Se entiende que, para otros expertos en la materia, resultarán evidentes ciertas modificaciones y variaciones de las disposiciones y los detalles descritos en la presente memoria. Por lo tanto, se pretende que la única limitación sea el alcance de las siguientes reivindicaciones de patente y no los detalles concretos presentados a modo de descripción y explicación de las realizaciones del presente documento.

## REIVINDICACIONES

1. Un decodificador de audio (200) para proporcionar una representación decodificada (282) de un contenido de audio a partir de una representación codificada (210) del contenido de audio, comprendiendo el decodificador de audio:
- un núcleo decodificador de dominio de predicción lineal (260) configurado para proporcionar una representación en el dominio temporal (262) de una trama de audio a partir de un conjunto (226, 232, 228) de parámetros de dominio de predicción lineal asociados con la trama de audio;
  - un núcleo de decodificador de dominio de frecuencia (250) configurado para proporcionar una representación en el dominio temporal (252) de una trama de audio a partir de un conjunto (242) de parámetros de dominio de frecuencia, teniendo en cuenta una ventana de transformada de entre un conjunto que comprende una pluralidad de ventanas de transformada diferentes (310, 320, 324, 330, 340, 350, 360, 364, 370, 380, 384); y
  - un combinador de señales (270) configurado para superponer y añadir representaciones en el dominio temporal (252, 262) de tramas de audio posteriores codificadas en diferentes dominios, con el fin de suavizar una transición entre las representaciones en el dominio temporal de las tramas posteriores;
  - en el que el conjunto de ventanas de transformada disponibles para ser aplicadas por el núcleo decodificador de dominio de frecuencia (250) comprende una ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD) adaptada para una generación de una representación en el dominio temporal de una única trama de audio codificada en el dominio de frecuencia (662) e intercalada temporalmente entre una trama de audio anterior (660) codificada en el dominio de predicción lineal y una trama de audio posterior (664) codificada en el dominio de predicción lineal,
  - en el que una pendiente de transición del lado izquierdo (384b) de la ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD) está adaptada para proporcionar una transición suave entre una representación en el dominio temporal de la trama anterior (660) codificada en el dominio de predicción lineal y la representación en el dominio temporal de la única trama (662) codificada en el dominio de frecuencia, y en el que una pendiente de transición del lado derecho de la ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD) está adaptada para proporcionar una transición suave entre la representación en el dominio temporal de la única trama (662) codificada en el dominio de frecuencia y una representación en el dominio temporal de la trama posterior (664) codificada en el dominio de predicción lineal;
  - en el que el decodificador de audio comprende un decodificador/cuantificador inverso/reescalador (230) configurado para recibir una información de valores espectrales codificada (222), decodificar la información de valores espectrales codificada (222), llevar a cabo una cuantificación inversa de la información de valores espectrales decodificada y llevar a cabo un reescalamiento aplicando factores de escala o factores de escala inversos a la información de valores espectrales decodificada y cuantificada inversamente, para obtener el conjunto (242) de parámetros de dominio de frecuencia.
2. El decodificador de audio (200) de acuerdo con la reivindicación 1, en el que la pendiente de transición del lado derecho (384d) de la ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD) comprende una duración temporal más corta que las pendientes de transición del lado derecho (310b, 320c, 338b, 350d, 360d, 370d, 380d) de ventanas de transformada adicionales (long\_window: 310, long\_start\_window: 320, 8\*short\_window: 330, long\_stop\_window: 350, stop\_start\_window: 360, stop\_window\_1152: 370 y stop\_start\_window\_1152: 380) adaptadas para una generación de una representación en el dominio temporal de una trama de audio codificada en el dominio de frecuencia que comprende una transición hacia una representación en el dominio temporal de una trama de audio posterior codificada en el dominio de frecuencia.
3. El decodificador de audio (200) según la reivindicación 1 o 2, en el que el núcleo decodificador de dominio de frecuencia (250) está configurado para llevar a cabo una transformada superpuesta y aplicar una ventana de transformada (310, 320, 324, 330, 340, 350, 360, 364, 370, 380, 384) con el fin de proporcionar la representación en el dominio temporal (252) de una trama de audio codificada en el dominio de frecuencia; y en el que la pendiente de transición del lado derecho (384d) de la ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD) se escoge de manera que la pendiente de transición del lado derecho queda completamente incluida en una porción sin solapamiento perteneciente a la correspondiente trama de audio codificada en el dominio de frecuencia (662).
4. El decodificador de audio (200) según una de las reivindicaciones 1 a 3, en el que la pendiente de transición del lado derecho (384d) de la ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD) se escoge de manera que la pendiente de transición del lado derecho queda restringida a una mitad interior de la extensión del lado derecho de la trama de audio (662), de manera que los artefactos de solapamiento en una mitad exterior de la extensión del lado derecho de la trama de audio (662) se suprimen mediante la ventana de transformada (384, stop\_start\_window\_1152\_to\_LPD).

5. El decodificador de audio (200) según una de las reivindicaciones 1 a 4, en el que la pendiente de transición del lado izquierdo (384b) de la ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD) se escoge de manera que la pendiente de transición del lado izquierdo de la ventana de inserción se extiende dentro de una porción con solapamiento perteneciente a la correspondiente trama de audio codificada en el dominio de frecuencia (662).
6. El decodificador de audio (200) según una de las reivindicaciones 1 a 5, en el que la pendiente de transición del lado izquierdo (384b) de la ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD) se escoge de tal manera que la pendiente de transición del lado izquierdo se extiende dentro de una mitad exterior de una extensión del lado izquierdo de la trama de audio (662).
7. El decodificador de audio (200) según una de las reivindicaciones 1 a 6, en el que la pendiente de transición del lado izquierdo (384b) de la ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD) comprende una extensión temporal más larga que la pendiente de transición del lado derecho (384d) de la ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD).
8. El decodificador de audio (200) según una de las reivindicaciones 1 a 7, en el que el combinador de señales (270) está configurado para procesar una porción final de una representación de una señal en el dominio temporal de una trama de audio anterior (660) codificada en el dominio de predicción lineal, con el fin de introducir componentes de solapamiento en el dominio temporal dentro de la porción final y de aplicar una ventana de transición a la porción final; y  
 en el que el combinador de señales (270) también está configurado para llevar a cabo una operación de superposición y adición para superponer y añadir la porción final procesada de la representación en el dominio temporal de la trama de audio anterior (660), codificada en el dominio de predicción lineal, con una porción de inicio de una representación en el dominio temporal enventanada de la única trama de audio codificada en el dominio de frecuencia (662), de manera que la porción final procesada de la representación en el dominio temporal de la trama de audio anterior (660), en la que se han introducido los componentes de solapamiento en el dominio temporal, se combina con una porción de la representación en el dominio temporal enventanada de la única trama de audio codificada en el dominio de frecuencia (662), porción a la cual se ha aplicado la pendiente de transición del lado izquierdo (384b) de la ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD);  
 en el que una longitud de la pendiente de transición del lado izquierdo de la ventana de inserción es idéntica a una longitud de un enventanado aplicado a la porción final de la representación en el dominio temporal de la trama de audio anterior.
9. El decodificador de audio (200) según una de las reivindicaciones 1 a 8, en el que el combinador de señales (270) está configurado para aplicar una ventana a una porción del lado izquierdo de la representación en el dominio temporal de la trama de audio posterior codificada en el dominio de predicción lineal (664); y  
 en el que el combinador de señales (270) está configurado para llevar a cabo una operación de superposición y adición para superponer y añadir una porción final de una representación en el dominio temporal enventanada de la única trama de audio codificada en el dominio de frecuencia (662) y una porción del lado izquierdo enventanada de la representación en el dominio temporal de la trama de audio posterior codificada en el dominio de predicción lineal (662),  
 en el que una longitud de la pendiente de transición del lado derecho de la ventana de inserción es idéntica a una longitud de un enventanado aplicado a una porción del lado izquierdo de la representación en el dominio temporal de la trama de audio posterior codificada en el dominio de predicción lineal.
10. El decodificador de audio (200) según una de las reivindicaciones 1 a 9, en el que el combinador de señales está configurado para llevar a cabo una operación de superposición y adición que comprende una cancelación del solapamiento en el dominio temporal en una transición desde la representación en el dominio temporal de la trama de audio anterior codificada en el dominio de predicción lineal (660) a la representación en el dominio temporal de la única trama de audio codificada en el dominio de frecuencia (662); y en el que el combinador de señales (270) está configurado para llevar a cabo una operación de superposición y adición sin una cancelación del solapamiento en el dominio temporal en una transición desde la representación en el dominio temporal de la única trama de audio codificada en el dominio de frecuencia (662) a la representación en el dominio temporal de la trama de audio posterior codificada en el dominio de predicción lineal (664); y  
 en el que la pendiente de transición del lado izquierdo de la ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD) está adaptada para la operación de superposición y adición con cancelación de solapamiento en el dominio temporal de las representaciones en el dominio temporal de la trama de audio anterior



codificada en el dominio de predicción lineal (660) y la única trama de audio codificada en el dominio de frecuencia (662),

5 en el que la pendiente de transición del lado derecho de la ventana de inserción (384, stop\_start\_window\_1152\_to\_LPD) está adaptada para una operación de superposición y adición sin cancelación de solapamiento en el dominio temporal de las representaciones en el dominio temporal de la única trama de audio codificada en el dominio de frecuencia (662) y la trama de audio posterior codificada en el dominio de predicción lineal (664).

11. Un procedimiento (800) para proporcionar una representación decodificada de un contenido de audio a partir de una representación codificada del contenido de audio, comprendiendo el procedimiento:

10 la provisión selectiva (810) de representaciones en el dominio temporal de tramas de audio a partir de un conjunto de parámetros de dominio de predicción lineal asociados con una trama de audio o a partir de un conjunto de parámetros de dominio de frecuencia asociado con una trama de audio,

15 se tiene en cuenta en el que una ventana de transformada de entre un conjunto que comprende una pluralidad de ventanas de transformada diferentes (310, 320, 324, 330, 340, 350, 360, 364, 370, 380, 384) cuando se proporciona la representación en el dominio temporal de una trama de audio codificada en el dominio de frecuencia; y

20 la realización (820) de una operación de superposición y adición de representaciones en el dominio temporal de tramas de audio posteriores (660, 662, 664) codificadas en diferentes dominios para suavizar una transición entre las representaciones en el dominio temporal de las tramas de audio posteriores codificadas en diferentes dominios;

25 en el que una ventana de inserción (384) se selecciona como ventana de transformada para la generación de una representación en el dominio temporal de una única trama de audio codificada en el dominio de frecuencia (662) e intercalada temporalmente entre una trama de audio anterior (660) codificada en el dominio de predicción lineal y una trama de audio posterior (664) codificada en un dominio de predicción lineal,

en el que una pendiente de transición del lado izquierdo (384b) de la ventana de inserción (384) (stop\_start\_window\_1152\_to\_LPD) está adaptada para proporcionar una transición suave entre una representación en el dominio temporal de la trama de audio anterior (660) codificada en el dominio de predicción lineal y la representación en el dominio temporal de la única trama de audio (662) codificada en el dominio de frecuencia, y

30 en el que una pendiente de transición del lado derecho (384d) de la ventana de inserción está adaptada para proporcionar una transición suave entre la representación en el dominio temporal de la única trama de audio (662) codificada en el dominio de frecuencia y una representación en el dominio temporal de la trama de audio posterior (664) codificada en el dominio de predicción lineal;

35 en el que el procedimiento comprende la recepción de una información de valores espectrales codificada (222), la decodificación de la información de valores espectrales codificada (222), la realización de una cuantificación inversa de la información de valores espectrales codificada y la realización de un reescalamiento aplicando factores de escala o factores de escala inversos a la información de valores espectrales decodificada e inversamente cuantificada, para obtener el conjunto (242) de parámetros de dominio de frecuencia.

40 12. Un programa informático adaptado para llevar a cabo el procedimiento según la reivindicación 11, cuando el programa informático se ejecuta en un ordenador.

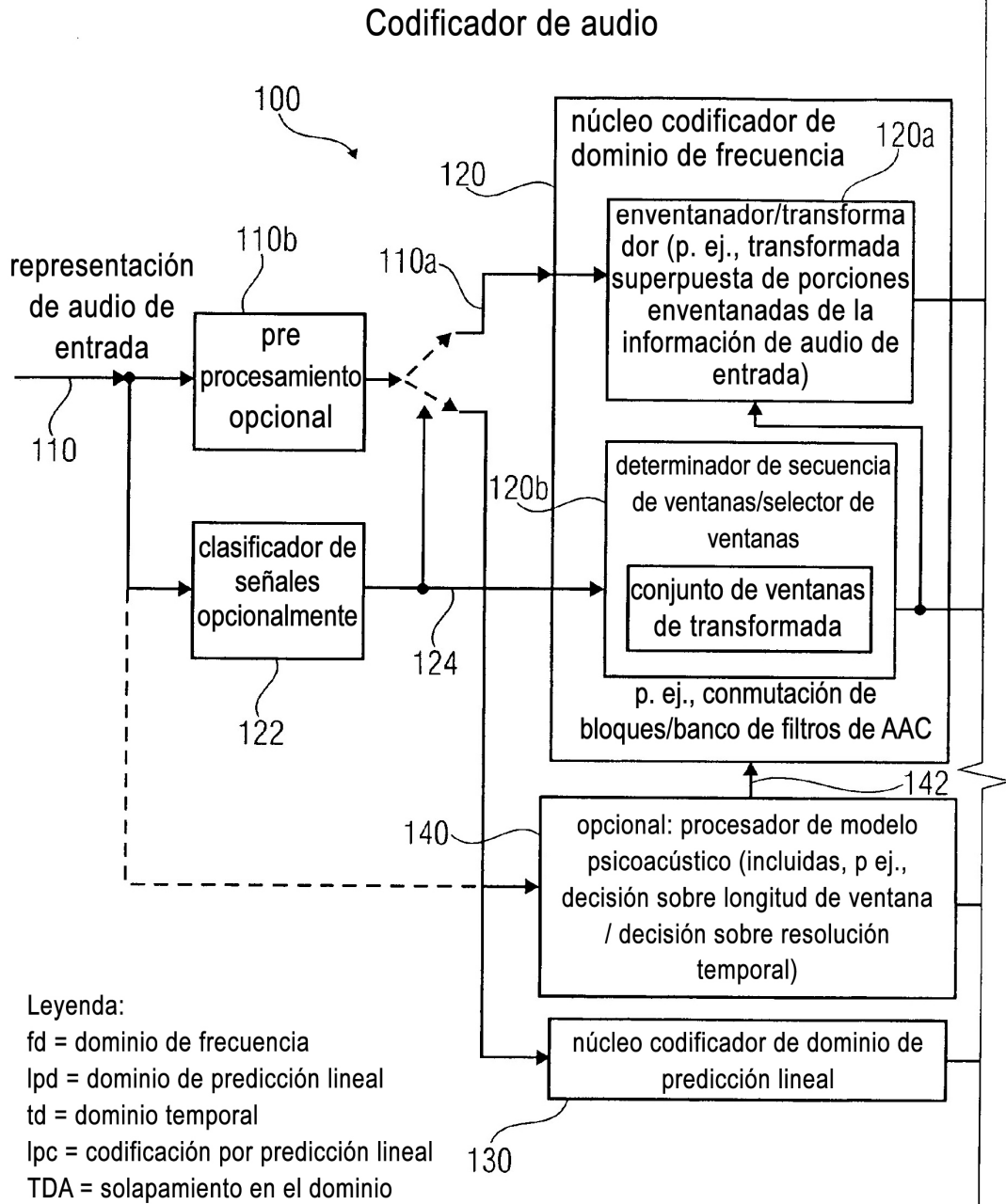
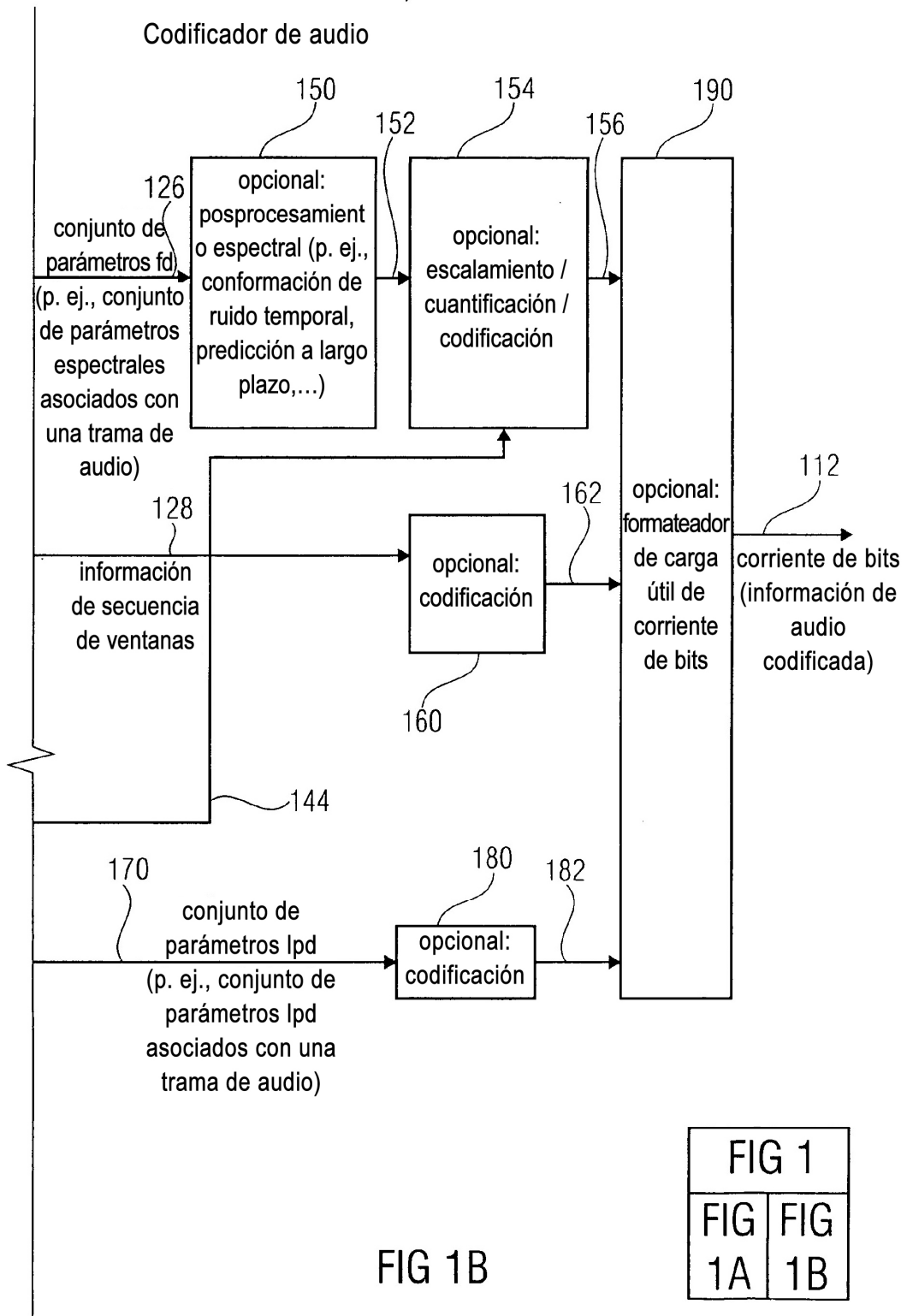


FIG 1	
FIG 1A	FIG 1B

FIG 1A



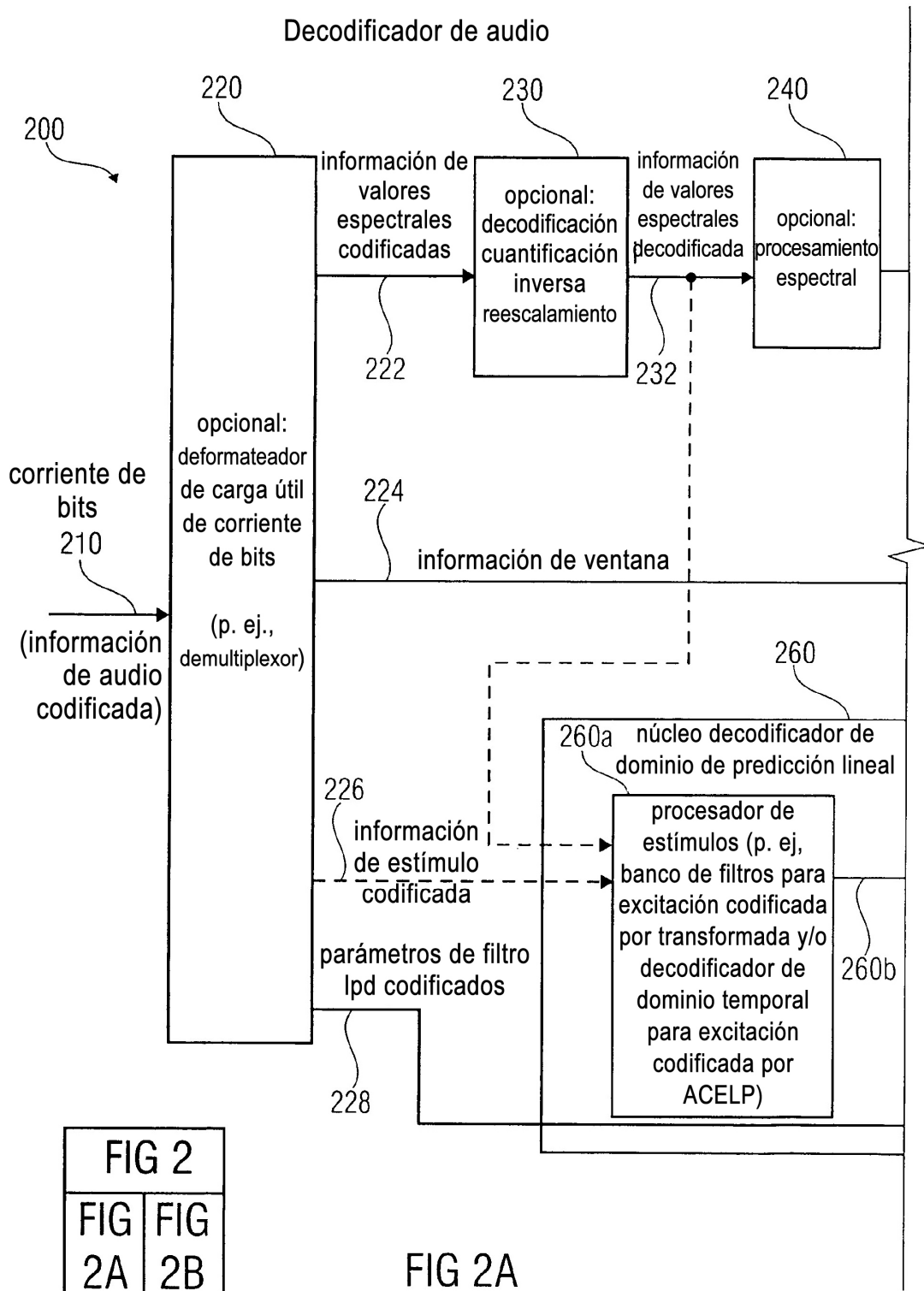
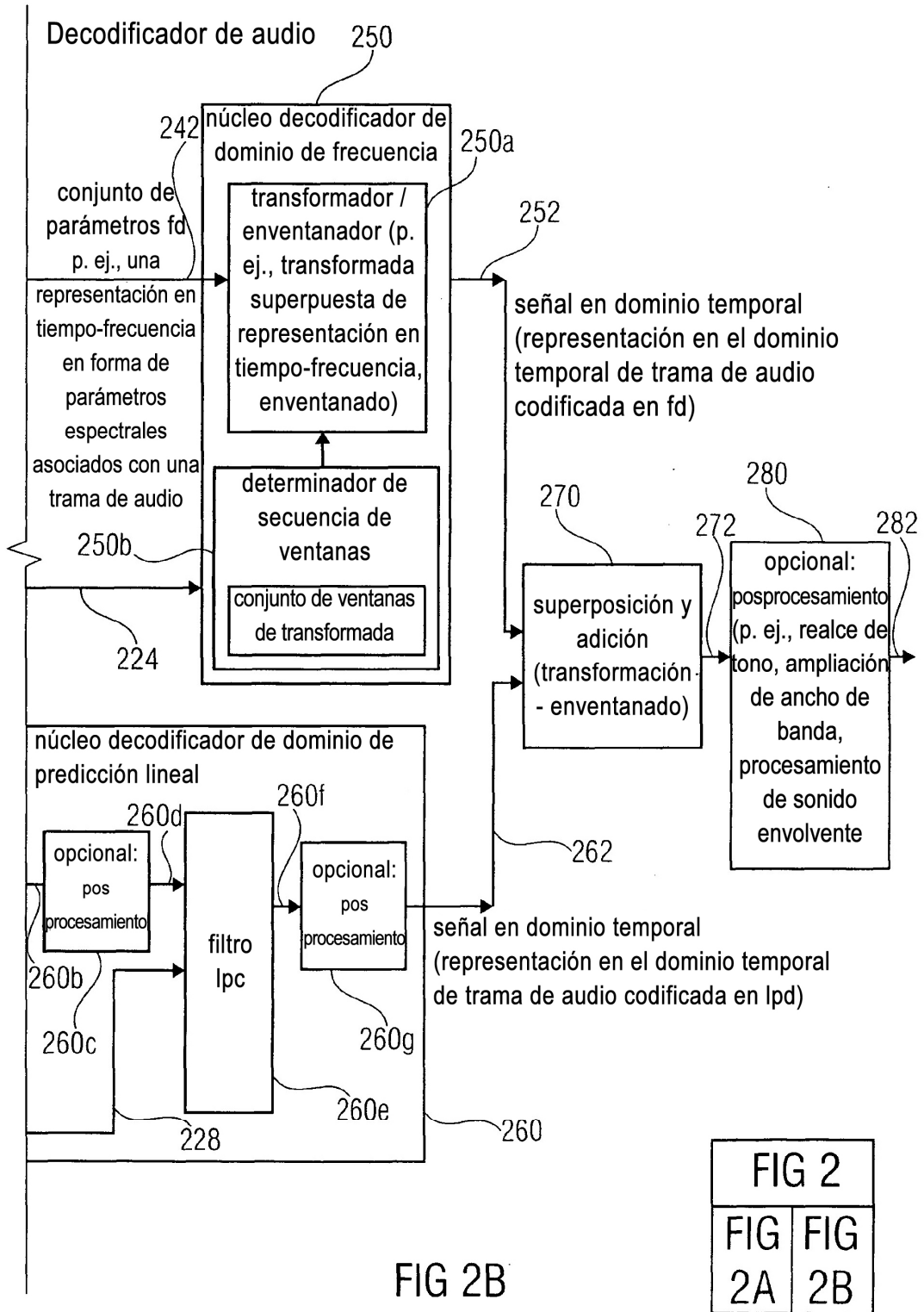


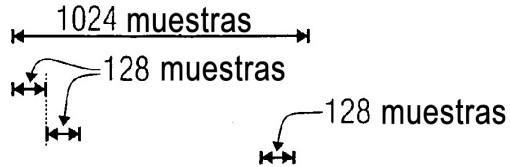
FIG 2	
FIG 2A	FIG 2B

FIG 2A



- Secuencias de ventanas y ventanas de transformada

Leyenda:



Valor	Ventana	N.º de coefs
0	ONLY_LONG_SEQUENCE =LONG_WINDOW	1024/960
1	LONG_START_SEQUENCE =LONG_START_WINDOW	1024/960
2	EIGHT_SHORT_SEQUENCE =8*_SHORT_WINDOW	8*(128/120)
3	LONG_STOP_SEQUENCE =LONG_STOP_WINDOW	1024/960
1	STOP_START_SEQUENCE =STOP_START_WINDOW	1024/960
3	STOP_1152_SEQUENCE =STOP_WINDOW_1152	1152/1080
1	STOP_START_1152_SEQUENCE =STOP_START_WINDOW_1152	1152/1080

FIG 3	
FIG 3A	FIG 3B

FIG 3A

- Secuencias de ventanas y ventanas de transformada

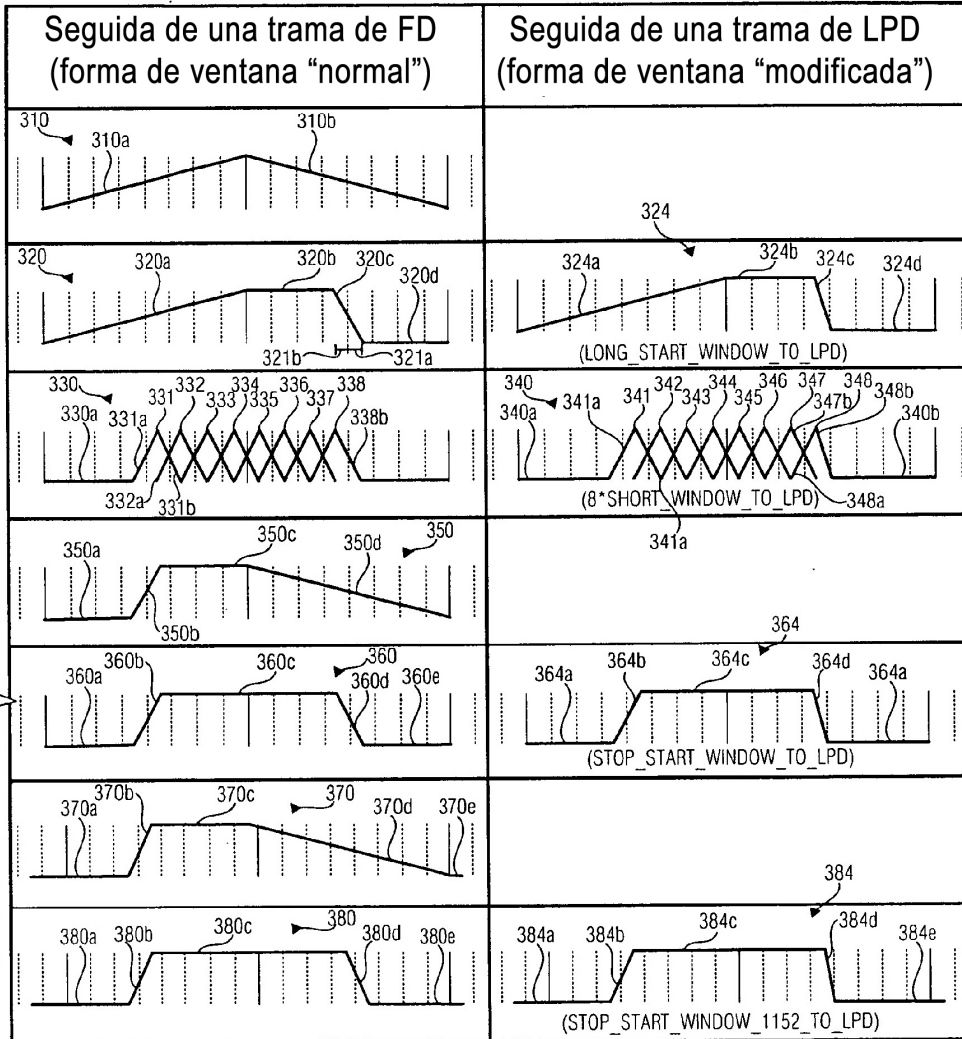


FIG 3B

FIG 3	
FIG 3A	FIG 3B

FIG 4A

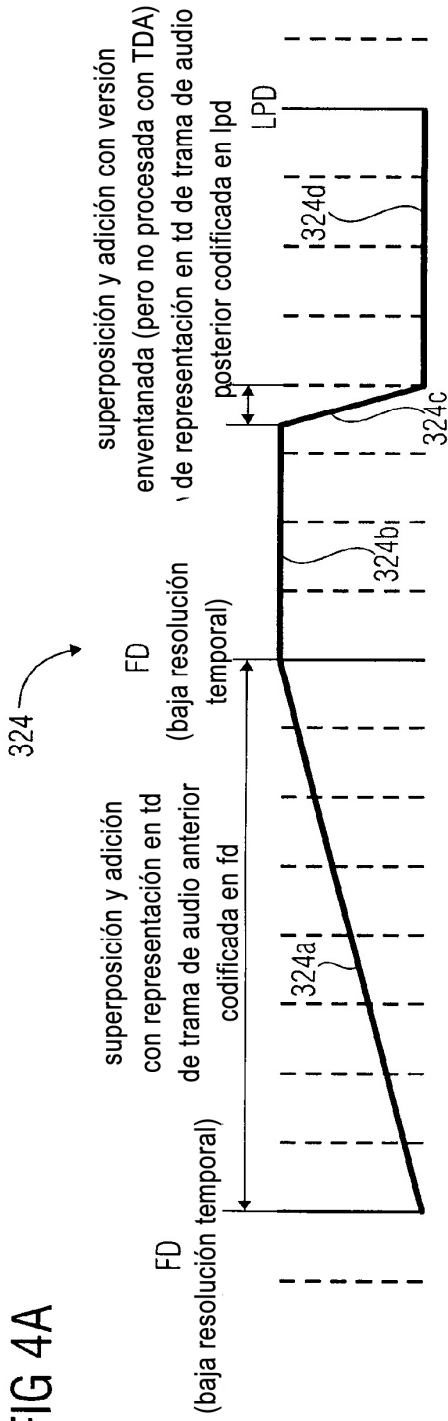


FIG 4B

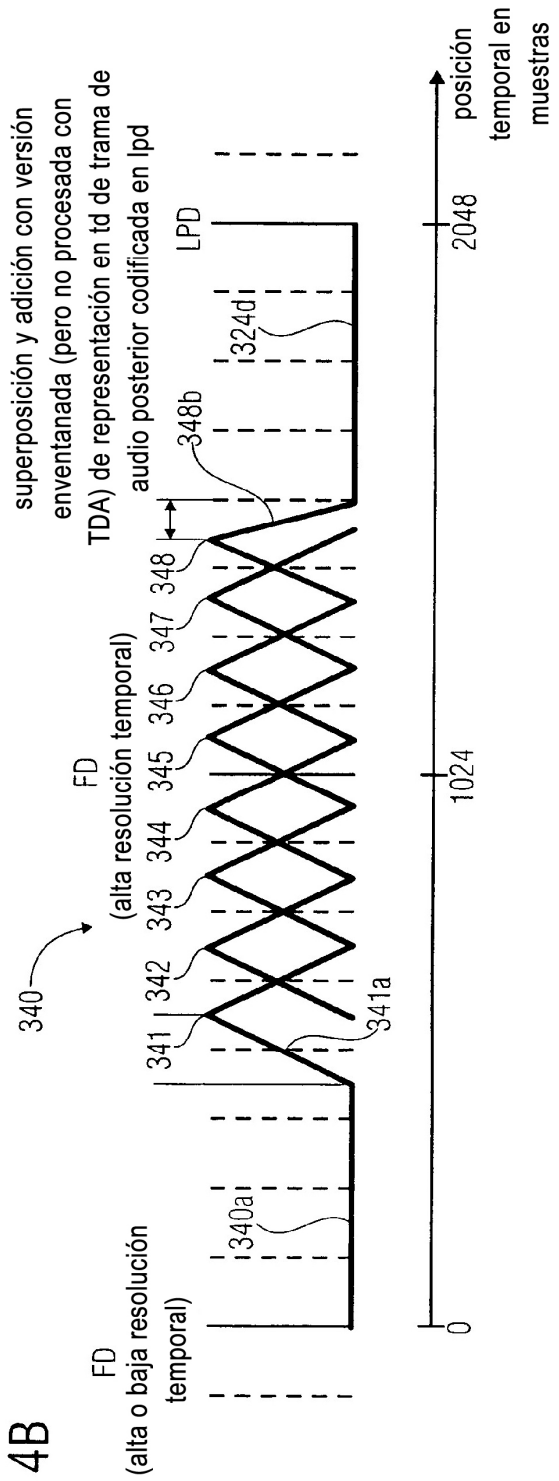




FIG 4C

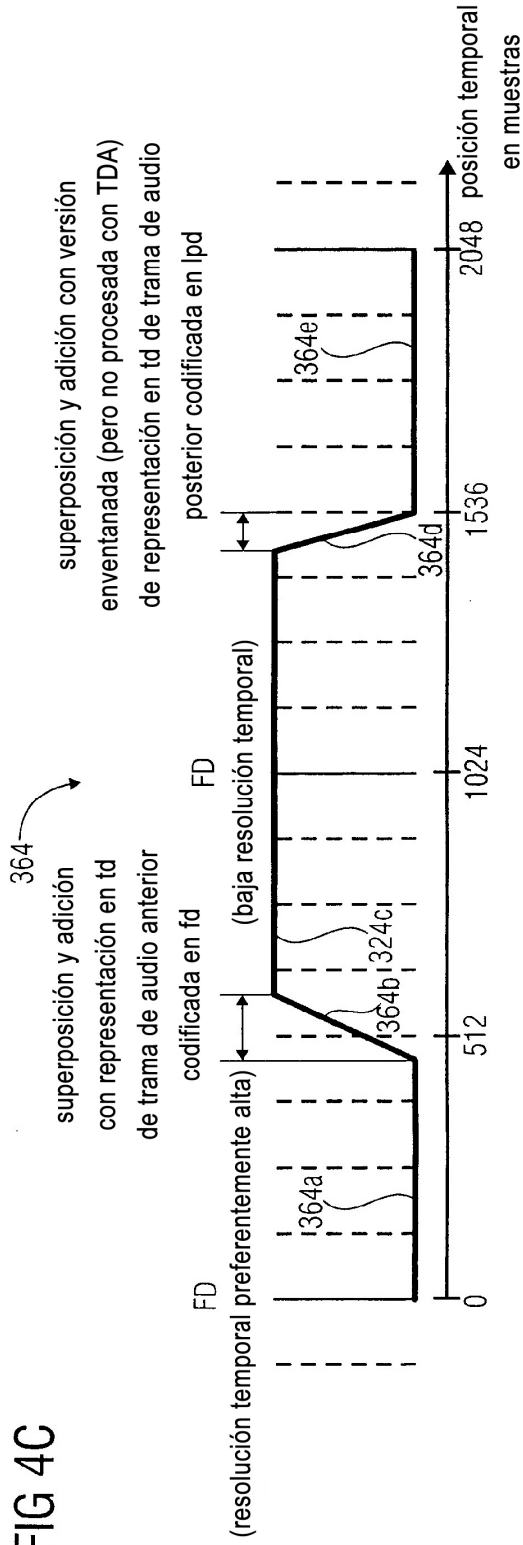
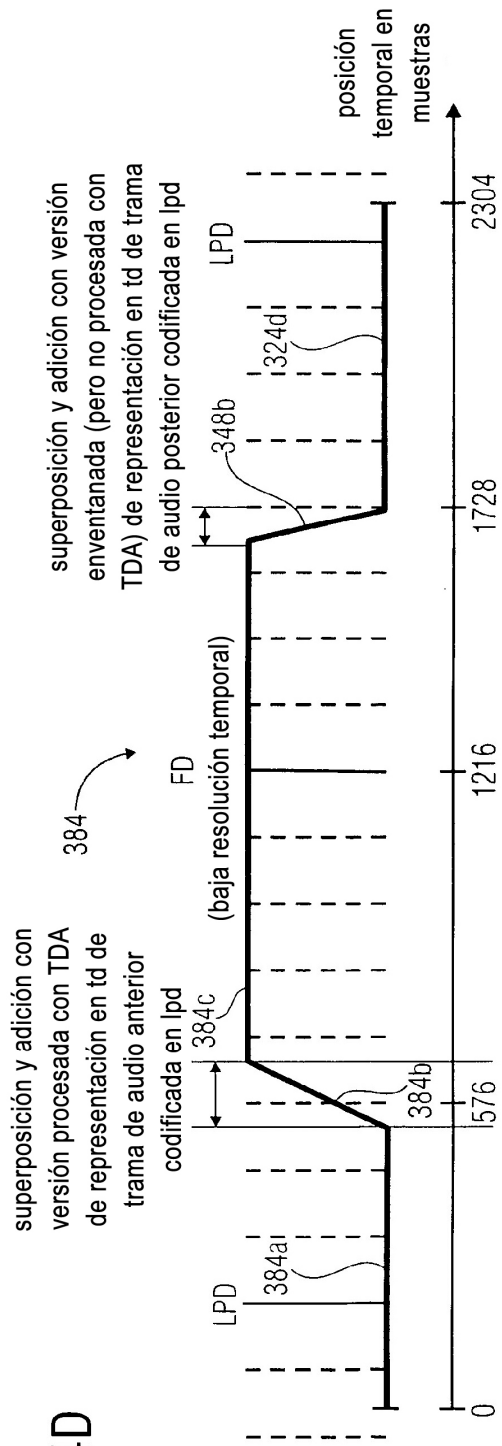


FIG 4D



secuencias de ventanas permitidas

secuencia de ventanas desde ↓ a →	ONLY_LONG_SEQUENCE	LONG_START_SEQUENCE	EIGHT_SHORT_SEQUENCE	LONG_STOP_SEQUENCE	STOP_START_SEQUENCE	LPD_SEQUENCE	STOP_1152_SEQUENCE	STOP_START_1152_SEQUENCE
ONLY_LONG_SEQUENCE	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>						
LONG_START_SEQUENCE			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
EIGHT_SHORT_SEQUENCE			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
LONG_STOP_SEQUENCE	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>						
STOP_START_SEQUENCE			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
LPD_SEQUENCE						<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
STOP_1152_SEQUENCE	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>						
STOP_START_1152_SEQUENCE			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		

FIG 5

FIG 6A

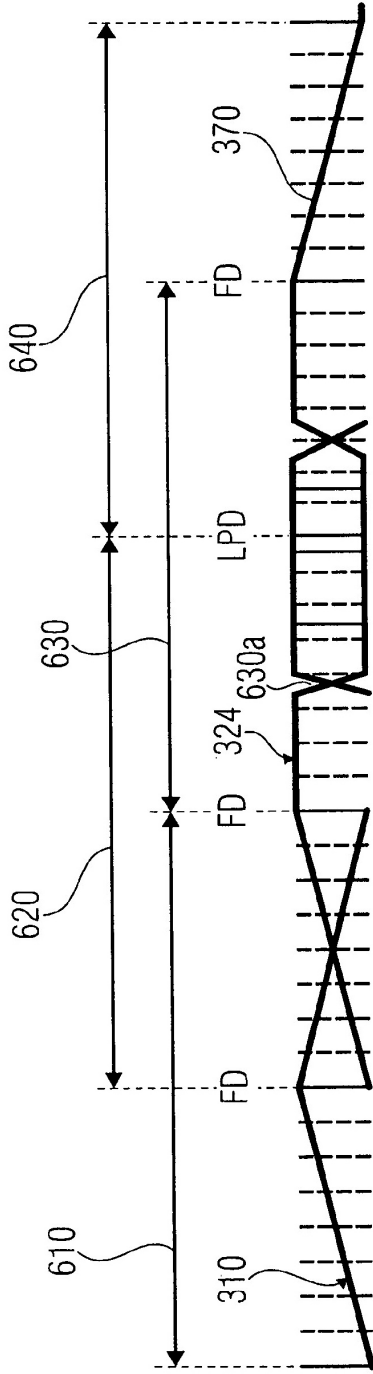


FIG 6B

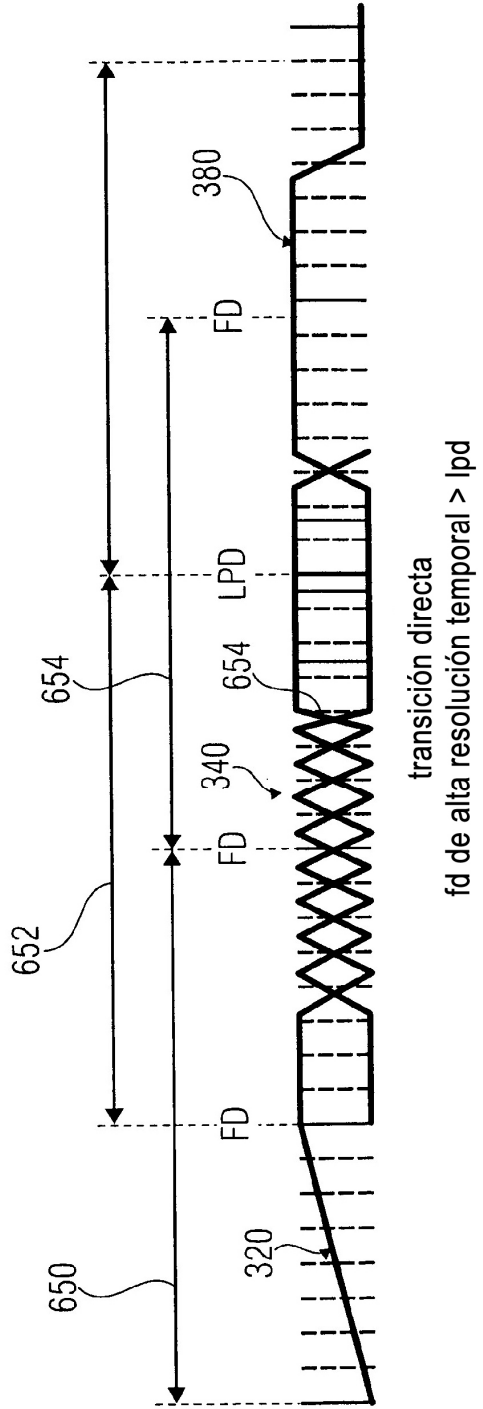


FIG 6C

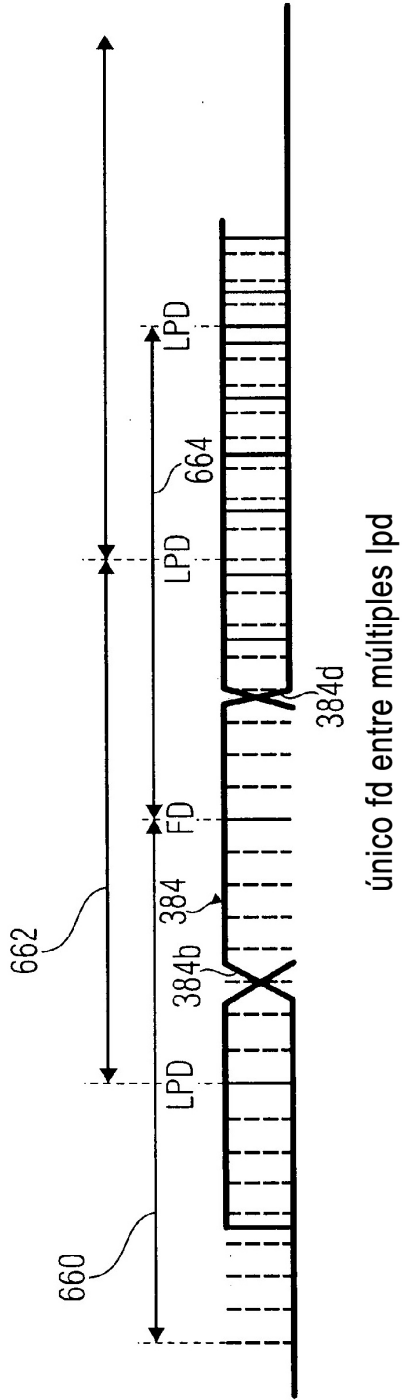
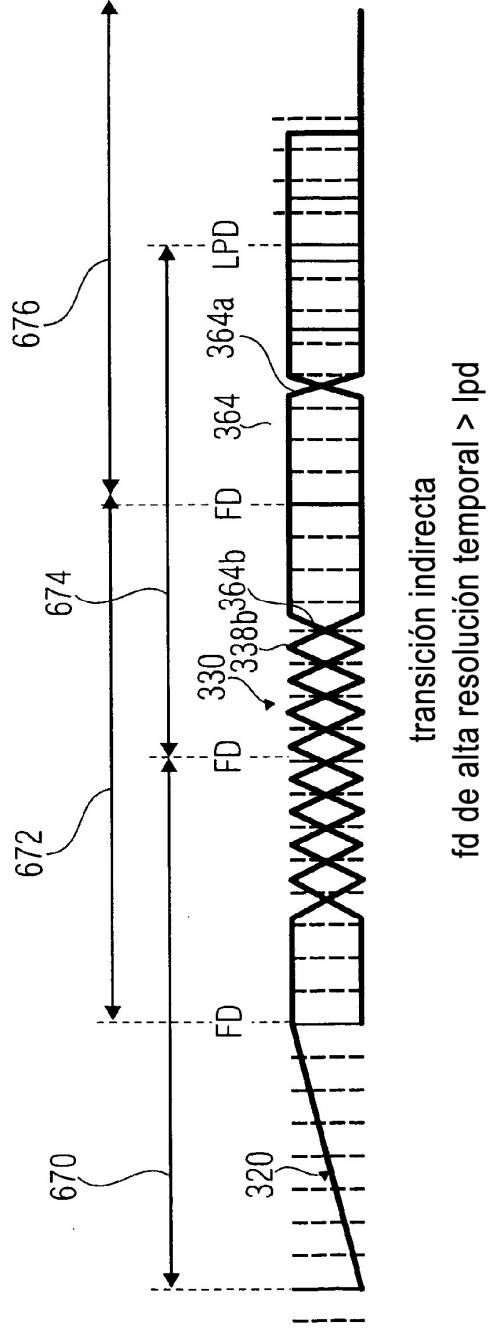


FIG 6D



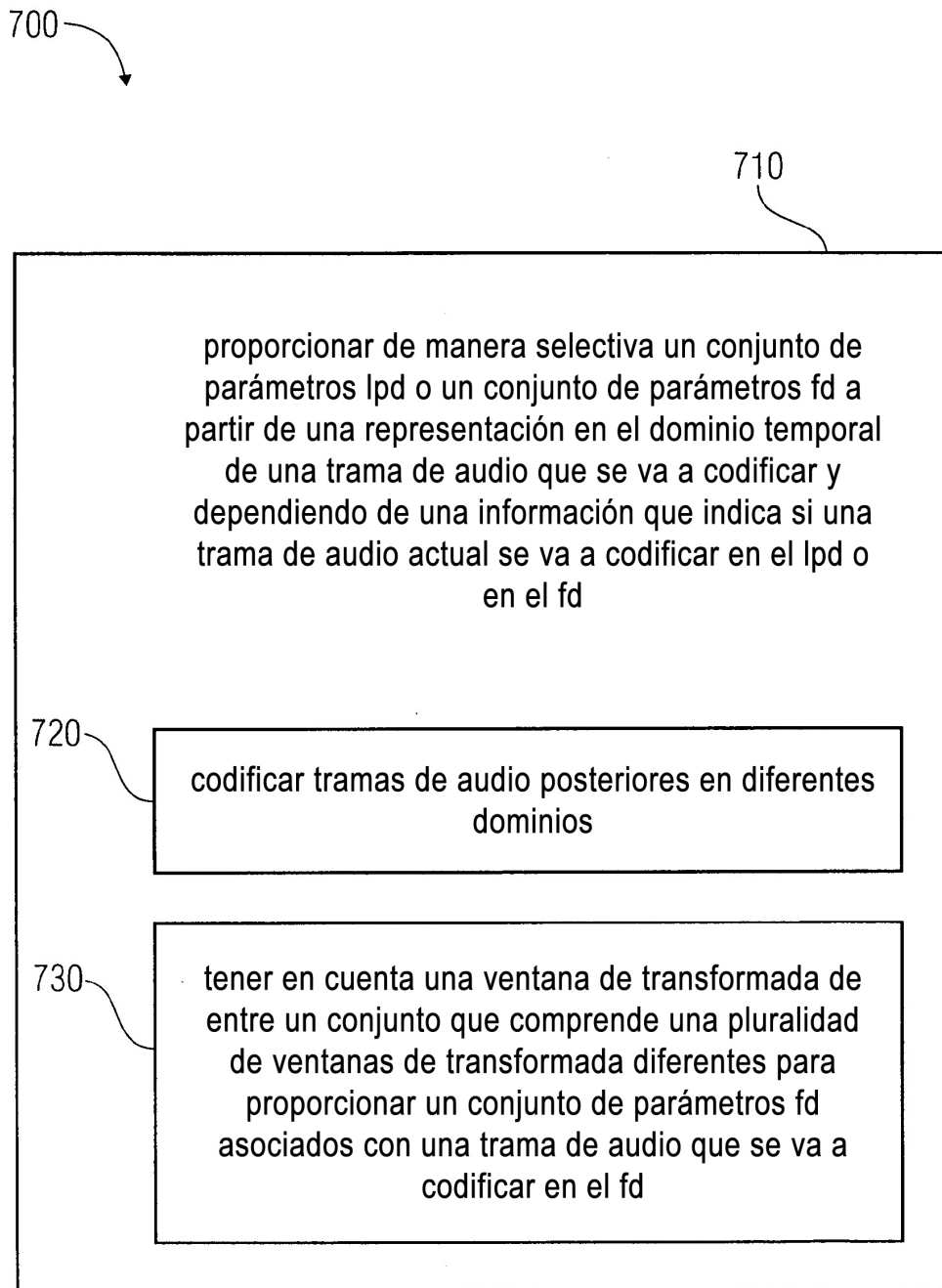


FIG 7

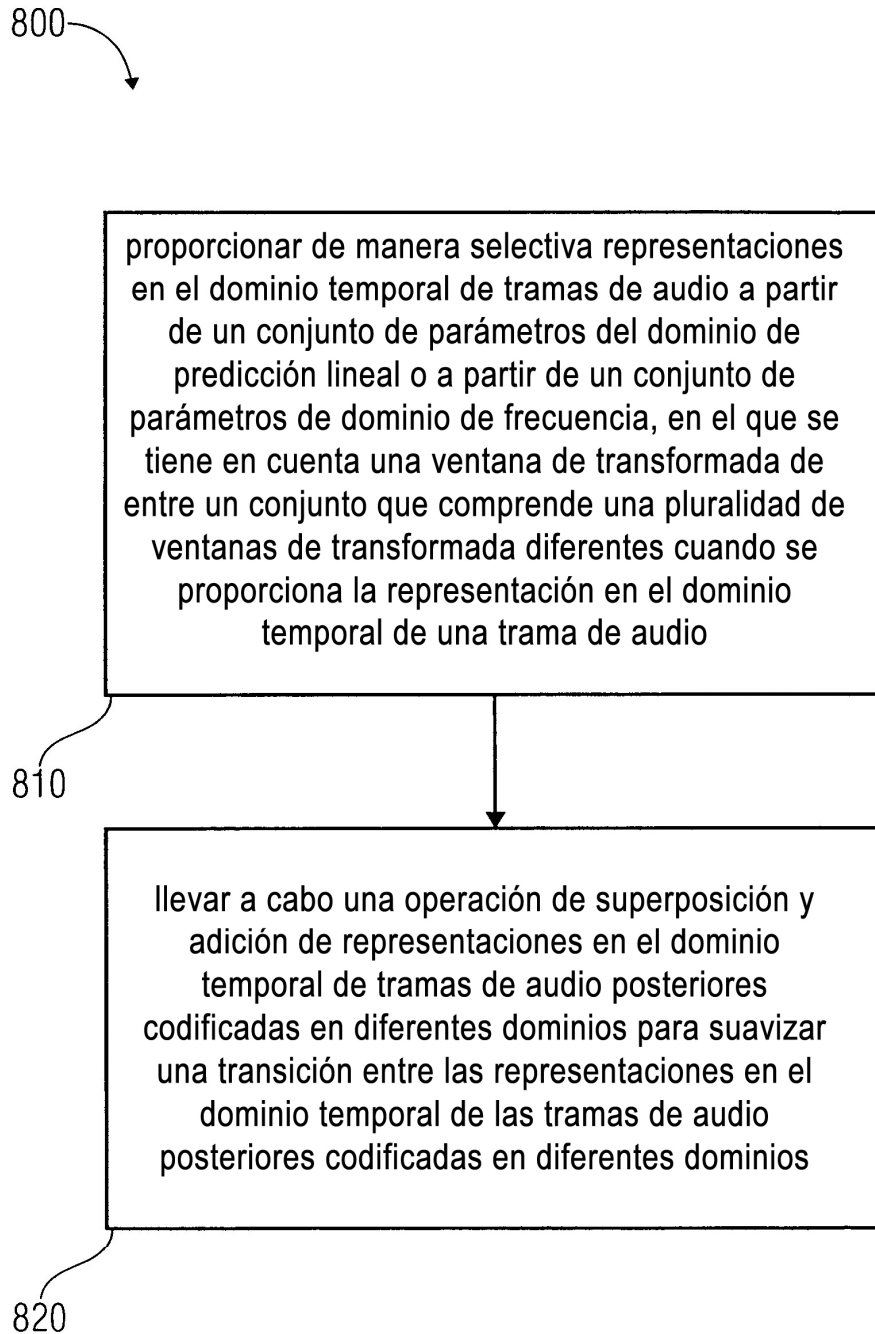
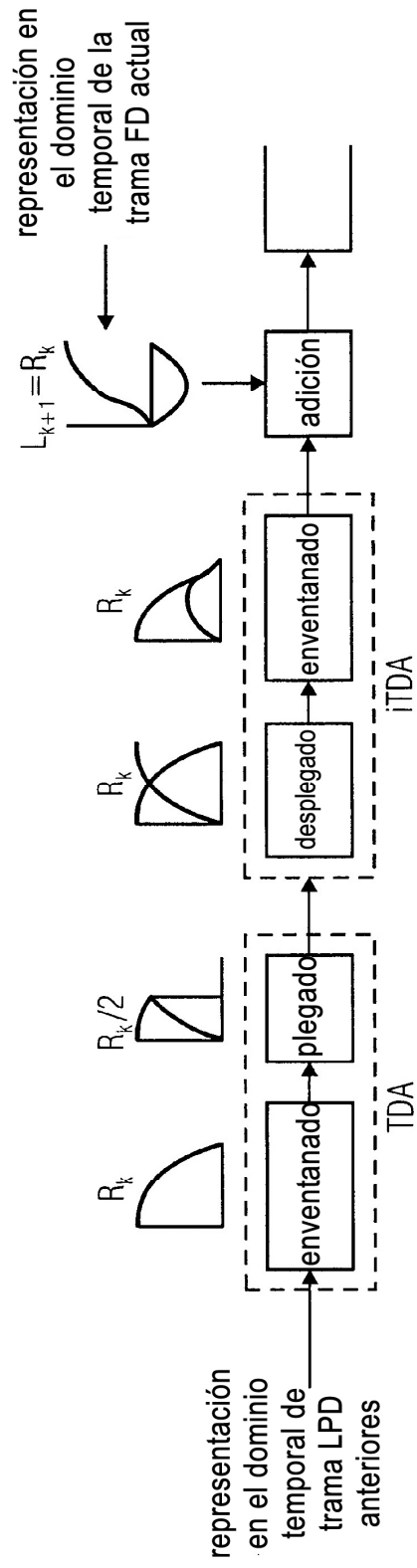


FIG 8



cancelación de solapamiento en dominio temporal

FIG 9

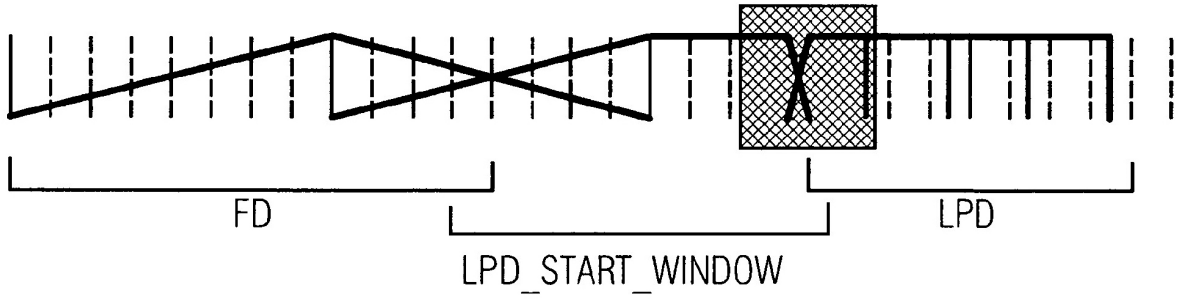


FIG 10A

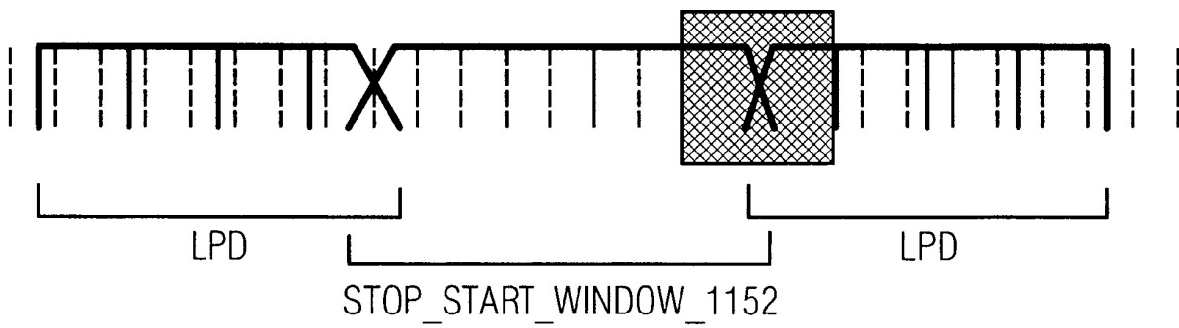


FIG 10B