

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 983 180**

51 Int. Cl.:

G10L 19/022 (2013.01)

G10L 19/03 (2013.01)

G10L 19/008 (2013.01)

G10L 19/028 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.07.2014** **E 23150061 (2)**

97 Fecha y número de publicación de la concesión europea: **27.03.2024** **EP 4191581**

54 Título: **Codificación de audio en el dominio de la frecuencia que soporta la conmutación de longitud de transformada**

30 Prioridad:

22.07.2013 EP 13177373

18.10.2013 EP 13189334

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

22.10.2024

73 Titular/es:

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

72 Inventor/es:

**DICK, SASCHA;
HELMRICH, CHRISTIAN y
HÖLZER, ANDREAS**

74 Agente/Representante:

PONTI & PARTNERS, S.L.P.

ES 2 983 180 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

Codificación de audio en el dominio de la frecuencia que soporta la conmutación de longitud de transformada

- 5 **[0001]** La presente solicitud se refiere a la codificación de audio en el dominio de la frecuencia que soporta la conmutación de longitud de transformada.
- [0002]** Los modernos sistemas de codificación de voz/audio en el dominio de la frecuencia tales como el códec Opus/Celt del IETF [1], MPEG-4 (HE-)AAC [2] o, en particular, MPEG-D xHE-AAC (USAC) [3], ofrecen medios para
10 codificar tramas de audio mediante el uso de una transformada larga – un bloque largo – u ocho transformadas cortas secuenciales - bloques cortos - según la estacionalidad temporal de la señal.
- [0003]** Para determinadas señales de audio tales como la lluvia o el aplauso de una gran audiencia, ni la codificación de bloques largos ni de bloques cortos produce calidad satisfactoria a bajas tasas de bits (bitrates). Esto
15 puede explicarse por la densidad de los transitorios prominentes en tales grabaciones; la codificación sólo con bloques largos puede causar deformación del tiempo, frecuente y audible del error de codificación, también conocido como preeco, mientras que la codificación solamente con bloques cortos generalmente es ineficiente debido a la mayor sobrecarga de datos, lo que conduce a agujeros espectrales.
- [0004]** Por lo tanto, sería favorable tener a mano un concepto de codificación de audio en el dominio de la frecuencia que soporte longitudes de transformada que sean también adecuadas para los tipos, que se acaban de señalar, de señales de audio. Naturalmente, sería factible construir un nuevo códec de audio en el dominio de la frecuencia que soporte la conmutación entre un conjunto de longitudes de transformada que, entre otros, abarque una determinada longitud de transformada deseada adecuada para un determinado tipo de señal de audio. Sin embargo,
20 no es una tarea fácil obtener un nuevo códec de audio en el dominio de la frecuencia adoptado en el mercado. Ya hay códecs bien conocidos disponibles y de uso frecuente. Por lo tanto, sería favorable poder tener a mano un concepto que permita que los códecs de audio en el dominio de la frecuencia existentes se amplíen de una manera que soporten adicionalmente una longitud de transformada, nueva y deseada, pero que, sin embargo, mantengan una retrocompatibilidad con los codificadores y decodificadores existentes.
25
30
- [0005]** La codificación de audio digital: Dolby AC-3 (Grant A. Davidson), enseña a exponer un bloque de audio corto a cuantificación y transmitancia de una manera idénticamente a un bloque largo único con el decodificador que procesa las secuencias entrelazadas de forma idéntica a las secuencias de bloque largas, excepto durante la transformación inversa. Es decir, un concepto de entrelazado se aplica para transmitir de forma eficiente los
35 coeficientes de transformada derivados de las transformadas de diferentes longitudes, es decir, correspondientes a diferentes resoluciones temporales espectrales dentro de la corriente de datos.
- [0006]** ITU-T G.719 sugiere un entrelazado en unidades de los denominados “subvectores”, cuya codificación de entropía se trata en el capítulo 7.4. El elemento de sintaxis es la conmutación de señales transitorias entre las transformadas cortas y largas.
40
- [0007]** La codificación de audio avanzada ISO/IEC MPEG-2 describe una técnica TNS de una manera que permite TNS en combinación con otros procedimientos que abordan el problema de TNS, como la conmutación de bloque de transformada y el control de preeco. El desentrelazado de las líneas de frecuencia en dos (o incluso más)
45 grupos de transformada es seguido de una aplicación TNS (filtrado de síntesis) independientemente para cada grupo de transformada de una manera independiente entre los grupos que utilizan filtros TNS independientes. Esto corresponde al uso o aplicación de un filtro TNS por bloque corto.
- [0008]** Por lo tanto, es un objeto de la presente invención proporcionar un concepto tal que permita ampliar los
50 códecs de audio en el dominio de la frecuencia existentes de forma retrocompatible, hacia el soporte de una longitud de transformada adicional para conmutar entre longitudes de transformada que también incluyen esta nueva longitud de transformada.
- [0009]** Este objeto se logra mediante la materia objeto según las reivindicaciones independientes adjuntas en esta invención.
55
- [0010]** La presente invención se basa en el descubrimiento de que es posible proporcionar un códec de audio en el dominio de la frecuencia con la capacidad de soportar adicionalmente una determinada longitud de transformada de una manera retrocompatible, cuando los coeficientes en el dominio de frecuencia de una respectiva trama son transmitidos de una manera entrelazada independientemente de la señalización que señala las tramas a las cuales se aplica realmente la longitud de la transformada, y cuando adicionalmente la extracción del coeficiente en el dominio de la frecuencia y la extracción del factor de escala operan independientemente de la señalización. Con esta medida, los codificadores/decodificadores de audio en el dominio de la frecuencia, anticuados, insensibles con respecto a la señalización, serían sin embargo capaces de operar sin fallos y de reproducir con una calidad razonable. Al mismo
60 tiempo, los codificadores/decodificadores de audio en el dominio de la frecuencia que responden a la conmutación
65

a/desde la longitud de transformada adicionalmente soportada lograrían una calidad aún mejor a pesar de la retrocompatibilidad. En cuanto a las penalizaciones en la eficiencia de codificación debido a la codificación de los coeficientes en el dominio de la frecuencia de una manera transparente para los decodificadores más antiguos a los que se refiere, las mismas son de una naturaleza comparativamente menor debido al entrelazado.

5

[0011] En las reivindicaciones dependientes se exponen implementaciones ventajosas de la presente solicitud. Los ejemplos y realizaciones siguientes del decodificador de audio en el dominio de la frecuencia no están según la invención y están presentes solo con fines de ilustración.

10 **[0012]** En particular, a continuación se describen realizaciones preferidas de la presente solicitud con respecto a las figuras entre las que

La Figura 1 muestra un diagrama de bloques esquemático de un decodificador de audio en el dominio de la frecuencia según una realización;

15 La Figura 2 muestra un diagrama esquemático que ilustra la funcionalidad del transformador inverso de la Figura 1; La Figura 3 muestra un diagrama esquemático que ilustra un posible desplazamiento del procedimiento de filtrado TNS inverso de la Figura 2 hacia una dirección situada corriente arriba según una realización;

La Figura 4 muestra una posibilidad para seleccionar ventanas cuando se utiliza la división de transformada para una ventana larga de parada-arranque en USAC según una realización; y

20 La Figura 5 muestra un diagrama de bloques de un codificador de audio en el dominio de la frecuencia según una realización.

[0013] La Figura 1 muestra un decodificador de audio en el dominio de la frecuencia que soporta la conmutación de longitud de transformada según una realización de la presente solicitud. El decodificador de audio en el dominio de la frecuencia de la Figura 1 se indica en términos generales mediante el uso de un número de referencia 10 y comprende un extractor de coeficiente en el dominio de la frecuencia 12, un extractor de factor de escala 14, un transformador inverso 16, y un combinador 18. En su entrada, el extractor de coeficiente en el dominio de la frecuencia y el extractor de factor de escala 12 y 14 tienen acceso a una corriente de datos de entrada 20. Las salidas del extractor de coeficiente en el dominio de la frecuencia 12 y el extractor de factor de escala 14 están conectadas a respectivas entradas del transformador inverso 16. La salida del transformador inverso 16, a su vez, está conectada a una entrada del combinador 18. Este último emite la señal de audio reconstruida hacia una salida 22 del codificador 10.

[0014] El extractor de coeficiente en el dominio de la frecuencia 12 está configurado para extraer los coeficientes en el dominio de la frecuencia 24 de las tramas 26 de la señal de audio desde la corriente de datos 20. Los coeficientes en el dominio de la frecuencia 24 pueden ser coeficientes MDCT o pueden formar parte de alguna otra transformada tal como otra transformada superpuesta. De una manera que se describe con más detenimiento en lo que sigue, los coeficientes en el dominio de la frecuencia 24 pertenecientes a una trama determinada 26 describen el espectro de la señal de audio dentro de la trama respectiva 26 en una resolución que varía según el espectro y del tiempo. Las tramas 26 representan porciones temporales donde la señal de audio se subdivide secuencialmente en el tiempo. Si se unifica la totalidad de los coeficientes en el dominio de la frecuencia 24 de todas las tramas, los mismos representan un espectrograma 28 de la señal de audio. Las tramas 26 pueden ser por ejemplo de igual longitud. Debido al tipo de contenido de audio de la señal de audio que cambia a lo largo del tiempo, puede ser desventajoso describir el espectro para cada trama 26 con una resolución espectro-temporal continua mediante el uso de por ejemplo transformadas que tienen una longitud de transformada constante que abarca por ejemplo, la longitud del tiempo de cada trama 26, es decir implica valores de muestras dentro de esta trama 26 de la señal de audio así como también muestras en el dominio del tiempo que preceden y suceden a la trama respectiva. Los artefactos de preeco pueden resultar por ejemplo de una transmisión con pérdidas del espectro de la trama respectiva en forma de los coeficientes en el dominio de la frecuencia 24. Por lo tanto, de una manera que se describe a grandes rasgos en lo que sigue, los coeficientes en el dominio de la frecuencia 24 de una trama respectiva 26 describen el espectro de la señal de audio dentro de esta trama 26 en una resolución espectro temporal conmutable mediante la conmutación entre diferentes longitudes de transformada. Sin embargo, en lo que se refiere al extractor de coeficiente en el dominio de la frecuencia 12, esta última circunstancia es transparente para el mismo. El extractor de coeficiente en el dominio de la frecuencia 12 opera independientemente de cualquier señalización que señale la conmutación que se acaba de mencionar entre diferentes resoluciones espectro temporales para las tramas 26.

55

[0015] El extractor de coeficiente en el dominio de la frecuencia 12 puede utilizar la codificación por entropía a efectos de extraer los coeficientes en el dominio de la frecuencia 24 a partir de la corriente de datos 20. Por ejemplo, el extractor de coeficiente en el dominio de la frecuencia puede utilizar la decodificación por entropía en base al contexto, tal como la decodificación aritmética variable con el contexto, a efectos de extraer los coeficientes en el dominio de la frecuencia 24 de la corriente de datos 20 con la asignación, a cada uno de los coeficientes en el dominio de la frecuencia 24, del mismo contexto independientemente de la señalización anteriormente mencionada que señala la resolución espectro temporal de la trama 26 a la que pertenece el respectivo coeficiente en el dominio de la frecuencia. Como alternativa, y como un segundo ejemplo, el extractor 12 puede utilizar la decodificación de Huffman y definir un conjunto de palabras del código de Huffman independientemente de dicha señalización que especifica la resolución de la trama 26.

65

[0016] Existen diferentes posibilidades en cuanto a la manera en que los coeficientes en el dominio de la frecuencia describen el espectrograma 28. Por ejemplo, los coeficientes en el dominio de la frecuencia 24 pueden representar meramente algún residuo de predicción. Por ejemplo, los coeficientes en el dominio de la frecuencia 5 pueden representar un residuo de una predicción que, al menos parcialmente, se ha obtenido mediante predicción estéreo desde otra señal de audio que representa un correspondiente canal de audio o mezcla descendente de una señal de audio de múltiples canales a la que pertenece el espectrograma de señales 28. Como alternativa o adicionalmente a un residuo de predicción, los coeficientes en el dominio de la frecuencia 24 pueden representar una señal de suma (media) o de diferencia (lateral) según el paradigma de estéreo M/S [5]. Además, los coeficientes en el 10 dominio de la frecuencia 24 pueden haber sido sometidos a conformación de ruido temporal.

[0017] Más allá de esto, los coeficientes en el dominio de la frecuencia 12 se cuantifican y a efectos de mantener el error de cuantificación por debajo de un umbral de detección psicoacústico (o enmascaramiento), por ejemplo, se varía espectralmente la magnitud de la etapa de cuantificación de una manera controlada por respectivos 15 factores de escala asociados con los coeficientes en el dominio de la frecuencia 24. El extractor de factor de escala 14 es responsable de la extracción de los factores de escala a partir de la corriente de datos 20.

[0018] Observando con un poco más de detalle la conmutación entre diferentes resoluciones espectro temporales de trama en trama, se observará lo siguiente. Tal como se describirá con mayor detalle en lo que sigue, la 20 conmutación entre diferentes resoluciones espectro temporales indicará que, dentro de una trama determinada 26, todos los coeficientes en el dominio de la frecuencia 24 pertenecen a una única transformada, o que los coeficientes en el dominio de la frecuencia 24 de la trama respectiva 26 pertenecen de hecho a diferentes transformadas tales como por ejemplo dos transformadas, la longitud de transformada de la que es la mitad de la longitud de transformada de la única transformada que se acaba de mencionar. La realización que se describe seguidamente con respecto a 25 las figuras supone la conmutación entre una sola transformada por una parte y dos transformadas por otra parte, pero de hecho una conmutación entre una sola transformada y más de dos transformadas sería en principio también factible así como las realizaciones indicadas en lo que sigue que son fácilmente transferibles a tales realizaciones alternativas.

[0019] La Figura 1 ilustra mediante líneas oblicuas el caso dado como ejemplo en que la trama actual es del 30 tipo representado por dos transformadas cortas, una de las cuales ha sido derivada mediante el uso de una mitad trasera de la trama actual 26, y la otra de las cuales se ha obtenido mediante la transformación de una mitad delantera de la trama actual 26 de la señal de audio. Debido a la longitud de transformada acortada, la resolución espectral a la que los coeficientes en el dominio de la frecuencia 24 describen que el espectro de la trama 26 se reduce, específicamente se reduce a la mitad en el caso de utilizarse dos transformadas cortas, mientras que se incrementa 35 la resolución temporal, específicamente se duplica en el presente caso. En la Figura 1, por ejemplo, los coeficientes en el dominio de la frecuencia 24 mostrados en líneas oblicuas formarán parte de la transformada delantera, mientras que los coeficientes no indicados con líneas oblicuas pertenecerán a la transformada trasera. Por lo tanto, los coeficientes en el dominio de la frecuencia, situados espectralmente en el mismo lugar 24, describen de este modo el mismo componente espectral de la señal de audio dentro de la trama 26, pero en momentos de tiempos ligeramente 40 diferentes, a saber, en dos ventanas de transformada consecutivas de la trama de división de transformada.

[0020] En la corriente de datos 20, los coeficientes en el dominio de la frecuencia 24 se transmiten de una manera entrelazada de tal manera que los coeficientes en el dominio de la frecuencia espectralmente correspondientes de las dos transformadas diferentes se sucedan inmediatamente. Dicho con otras palabras, los coeficientes en el 45 dominio de la frecuencia 24 de una trama de transformada dividida, es decir una trama 26 para la que se señala la división de transformada en la corriente de datos 20, se transmiten de tal manera que si los coeficientes en el dominio de la frecuencia 24 recibidos procedentes desde el extractor de coeficiente en el dominio de la frecuencia 12 estarían ordenados consecutivamente de una manera como si fuesen coeficientes en el dominio de la frecuencia de una transformada larga, entonces se disponen en la secuencia de una manera entrelazada de tal manera que los 50 coeficientes en el dominio de la frecuencia espectralmente situados en el mismo lugar 24 inmediatamente adyacentes entre sí y los pares de tales coeficientes en el dominio de la frecuencia situados respectivamente en el mismo lugar 24 están dispuestos según un orden espectral/frecuencia. Es interesante observar que, ordenada de esa manera, la secuencia de coeficientes en el dominio de la frecuencia entrelazados 24 tiene un aspecto similar a una secuencia de coeficientes en el dominio de la frecuencia 24 obtenidos mediante una única transformada larga. De nuevo, en lo que 55 se refiere al extractor de coeficiente en el dominio de la frecuencia 12, la conmutación entre diferentes longitudes de transformada o de resoluciones espectro temporales en unidades de las tramas 26 es transparente para los mismos, y, por lo tanto, la selección del contexto para la codificación por entropía de los coeficientes en el dominio de la frecuencia 24 de una manera adaptativa al contexto, tiene como resultado que se selecciona el mismo contexto - independientemente de la trama actual que en realidad es una trama de transformada larga de la trama actual que es 60 del tipo de transformada dividida sin que se conozca la localización del extractor 12. Por ejemplo, el extractor de coeficiente en el dominio de la frecuencia 12 puede seleccionar el contexto que se va a utilizar para un determinado coeficiente en el dominio de la frecuencia en base a coeficientes en el dominio de la frecuencia ya codificados/decodificados en un vecindario espectro temporal, estando dicho vecindario espectro temporal definido en el estado entrelazado ilustrado en la Figura 1. Esto tiene la siguiente consecuencia. Imagínese que un coeficiente en 65 el dominio de la frecuencia, actualmente codificado/decodificado 24 era parte de una transformada delantera indicada

mediante líneas oblicuas en la Figura 1. Un coeficiente en el dominio de la frecuencia adyacente, inmediatamente espectral, sería de hecho un coeficiente en el dominio de la frecuencia 24 de la misma transformada delantera (es decir una transformada indicada con líneas oblicuas en la Figura 1). Sin embargo, no obstante, el extractor de coeficientes en el dominio de la frecuencia 12 utiliza para la selección del contexto, un coeficiente en el dominio de la frecuencia 24 perteneciente a la transformada trasera, a saber, aquella que es inmediatamente adyacente (según una resolución espectral reducida de la transformada acortada), suponiendo que esta última sería la vecina espectral inmediata de una transformada larga del coeficiente en el dominio de la frecuencia actual 24. De manera similar, para la selección del contexto de un coeficiente en el dominio de la frecuencia 24 de una transformada trasera, el extractor de coeficiente en el dominio de la frecuencia 12 utilizaría como vecina espectralmente inmediata un coeficiente en el dominio de la frecuencia 24 perteneciente a la transformada delantera, y que de hecho está situada espectralmente en el mismo lugar que dicho coeficiente. En particular, el orden de decodificación definido entre los coeficientes 24 de la trama actual 26 conduciría, por ejemplo, de una menor frecuencia a una mayor frecuencia. Son válidas observaciones similares en el caso del extractor de coeficiente en el dominio de la frecuencia 12 que está configurado para decodificar por entropía los coeficientes en el dominio de la frecuencia 24 de una trama actual 26 en grupos/tuplas de coeficientes en el dominio de la frecuencia inmediatamente consecutivos 24 cuando están ordenados de manera no desentrelazada. En lugar de utilizar la tupla de coeficientes en el dominio de la frecuencia espectralmente adyacentes 24 solamente pertenecientes a la misma transformada corta, el extractor de coeficiente en el dominio de la frecuencia 12 seleccionaría el contexto para una determinada tupla de una mezcla de coeficientes en el dominio de la frecuencia 24 perteneciente a diferentes transformadas cortas sobre la base de una tupla espectralmente adyacente de una mezcla tal de coeficientes en el dominio de la frecuencia 24 perteneciente a las diferentes transformadas.

[0021] Debido al hecho de que, como se ha indicado más arriba, en el estado entrelazado, el espectro resultante obtenido mediante dos transformadas cortas tiene un aspecto muy similar al espectro obtenido mediante una única transformada larga, la penalización de la codificación por entropía resultante de la operación agnóstica del extractor de coeficiente en el dominio de la frecuencia 12 con respecto a la conmutación de longitud de transformada es baja.

[0022] Se reanuda la descripción del decodificador 10 con el extractor de factor de escala 14 que es, como se ha mencionado más arriba, responsable de extraer los factores de escala de los coeficientes en el dominio de la frecuencia 24 de la corriente de datos 20. La resolución espectral con la que los factores de escala son asignados a los coeficientes en el dominio de la frecuencia 24 es más ordinaria que la resolución espectral, comparativamente fina, soportada por la transformada larga. Tal como se ilustra con llaves 30, los coeficientes en el dominio de la frecuencia 24 pueden estar agrupados en múltiples bandas de factores de escala. La subdivisión en las bandas de factores de escala puede seleccionarse en base a consideraciones psicoacústicas y puede coincidir, por ejemplo, con las denominadas bandas Bark (o críticas). Dado que el extractor de factor de escala 14 es agnóstico en cuanto a la conmutación de longitud de transformada, de la misma manera que lo es el extractor de coeficiente en el dominio de la frecuencia 12, el extractor de factor de escala 14 supone que cada trama 26 está subdividida en un número de bandas de factores de escala 30 que es igual, independientemente de la señalización de conmutación de longitud de transformada, y extrae para cada banda de factor de escala 30 un factor de escala 32. En el lado del codificador, la atribución de los coeficientes en el dominio de la frecuencia 24 a estas bandas de factor de escala 30 se efectúa en un estado no desentrelazado ilustrado en la Figura 1. Como consecuencia, en lo que se refiere a las tramas 26 correspondientes a la transformada dividida, cada factor de escala 32 forma parte de un grupo poblado por coeficientes en el dominio de la frecuencia 24 de la transformada delantera, y los coeficientes en el dominio de la frecuencia 24 de la transformada trasera.

[0023] El transformador inverso 16 está configurado para recibir para cada trama 26 los correspondientes coeficientes en el dominio de la frecuencia 24 y los correspondientes factores de escala 32 y para someter los coeficientes en el dominio de la frecuencia 24 de la trama 26, escalados según los factores de escala 32, a una transformación inversa a efectos de adquirir porciones en el dominio del tiempo de la señal de audio. Un transformador inverso 16 puede utilizar una transformada solapada tal como, por ejemplo, una transformada de coseno discreta modificada (MDCT). El combinador 18 combina las porciones en el dominio del tiempo para obtener la señal de audio tal como mediante el uso de, por ejemplo, un procedimiento de superposición/adición adecuado que resulta por ejemplo en la cancelación de aliasing en el dominio del tiempo dentro de las porciones superpuestas de las porciones en el dominio del tiempo emitidas por el transformador inverso 16.

[0024] Naturalmente, el transformador inverso 16 responde a la conmutación de longitud de transformada anteriormente mencionada, señalizada dentro de la corriente de datos 20 para las tramas 26. La operación del transformador inverso 16 se describe con mayor detalle con respecto a la Figura 2.

[0025] La Figura 2 muestra una posible estructura interna del transformador inverso 16, con más detalle. Como se indica en la Figura 2, el transformador inverso 16 recibe para una trama actual los coeficientes en el dominio de la frecuencia 24 asociados con esa trama, así como los correspondientes factores de escala 32 para descuantificar los coeficientes en el dominio de la frecuencia 24. Además, el transformador inverso 16 es controlado por la señalización 34 que está presente en la corriente de datos 20 para cada trama. El transformador inverso 16 se puede controlar además mediante otros componentes de la corriente de datos 20 opcionalmente comprendidos en éste. En la siguiente

descripción, se describen detalles relacionados con estos parámetros adicionales.

5 **[0026]** Como se muestra en la Figura 2, el transformador inverso 16 de la Figura 2 comprende un descuantificador 36, un desentrelazador activable 38 y una fase de transformación inversa 40. Para mayor facilidad de entendimiento de la siguiente descripción, los coeficientes en el dominio de la frecuencia de entrada 24 derivados para la trama actual del extractor de coeficiente en el dominio de la frecuencia 12, se muestran con la numeración de 0 a N-1. De nuevo, dado que el extractor de coeficiente en el dominio de la frecuencia 12 es agnóstico a, es decir, opera independientemente de, la señalización 34, el extractor de coeficiente en el dominio de la frecuencia 12 proporciona al transformador inverso 16 los coeficientes en el dominio de la frecuencia 24 de la misma manera independiente de la trama actual que es del tipo de transformada dividida, o del tipo 1-transformada, es decir el número de coeficientes en el dominio de la frecuencia 24 es N en el presente caso ilustrado y la asociación de los índices 0 a N-1 a los N coeficientes en el dominio de la frecuencia 24 también sigue siendo la misma independientemente de la señalización 34. En el caso de que la trama actual sea del tipo de transformada larga, los índices 0 a N-1 corresponden al ordenamiento de los coeficientes en el dominio de la frecuencia 24 de la frecuencia más baja a la frecuencia más alta, y en el caso en que la trama actual sea del tipo de transformada dividida, los índices corresponden al orden de los coeficientes en el dominio de la frecuencia cuando están espectralmente dispuestos según su orden espectral, pero de una manera entrelazada de tal manera que cada segundo coeficiente en el dominio de la frecuencia 24 pertenece a la transformada trasera, mientras que los otros pertenecen a la transformada delantera.

20 **[0027]** Son ciertos hechos similares para los factores de escala 32. Dado que el extractor de factor de escala 14 opera de una manera agnóstica con respecto a la señalización 34, el número y el orden así como los valores de los factores de escala 32 procedentes del extractor de factor de escala 14 son independientes de la señalización 34, indicándose en los factores de escala 32 en la Figura 2 a título de ejemplo como S_0 a S_M , correspondiendo el índice al orden secuencial entre las bandas de factor de escala con las que se asocian estos factores de escala.

25 **[0028]** De una manera similar al extractor de coeficiente en el dominio de la frecuencia 12 y el extractor de factor de escala 14, el descuantificador 36 puede operar agnósticamente con respecto a, o independientemente de la señalización 34. El descuantificador 36 descuantifica, o escala, los coeficientes en el dominio de la frecuencia de entrada 24 utilizando el factor de escala asociado con la banda de factor de escala a la que pertenecen los respectivos coeficientes en el dominio de la frecuencia. De nuevo, la membresía de los coeficientes en el dominio de la frecuencia de entrada 24 con respecto a las bandas de factor de escala individuales y, por lo tanto, la asociación de los coeficientes en el dominio de la frecuencia de entrada 24 con respecto a los factores de escala es independiente de la señalización 34, y el transformador inverso 16 por lo tanto somete los coeficientes en el dominio de la frecuencia 24 a escalado según los factores de escala 32 a una resolución espectral que es independiente de la señalización. Por ejemplo, el descuantificador 36, independientemente de la señalización 34, asigna coeficientes en el dominio de la frecuencia con índices 0 a 3 a la primera banda de factor de escala y, por lo tanto, al primer factor de escala S_0 , los coeficientes en el dominio de la frecuencia con los índices 4 a 9 a la segunda banda de factor de escala y por lo tanto el factor de escala S_1 y así sucesivamente. Los límites de factor de escala tienen como objeto son solamente ilustrativos. El descuantificador 36 podría, por ejemplo, a efectos de descuantificar los coeficientes en el dominio de la frecuencia 24, llevar a cabo una multiplicación utilizando el factor de escala asociado, es decir, calcular el coeficiente en el dominio de la frecuencia x_0 para que sea $x_0 \cdot s_0$, x_1 para que sea $x_1 \cdot s_0$, ..., x_3 para que sea $x_3 \cdot s_0$, x_4 para que sea $x_4 \cdot s_1$, ..., x_9 para que sea $x_9 \cdot s_1$, y así sucesivamente. Como alternativa, el descuantificador 36 puede llevar a cabo una interpolación de los factores de escala realmente utilizados para la descuantificación de los coeficientes en el dominio de la frecuencia 24 procedentes de la resolución espectral más ordinaria definida por las bandas de factor de escala. La interpolación puede ser independiente de la señalización 34. Sin embargo, como alternativa, la interpolación anterior puede depender de la señalización a efectos de tener en cuenta las diferentes posiciones de muestreo espectro temporal de los coeficientes en el dominio de la frecuencia 24 dependiendo de si la trama actual es de tipo transformada dividida o de tipo transformada única/larga.

50 **[0029]** La Figura 2 ilustra que hasta el lado de entrada del desentrelazador activable 38, el orden entre los coeficientes en el dominio de la frecuencia 24 sigue siendo el mismo y se aplica el mismo, al menos sustancialmente, con respecto a la operación general hasta dicho punto. La Figura 2 muestra que corriente arriba del desentrelazador activable 38, es posible llevar a cabo otras operaciones mediante el transformador inverso 16. Por ejemplo, el transformador inverso 16 podría estar configurado para llevar a cabo el relleno de ruido sobre los coeficientes en el dominio de la frecuencia 24. Por ejemplo, en la secuencia de los coeficientes en el dominio de la frecuencia 24, las bandas de factor de escala, es decir, los grupos de coeficientes en el dominio de la frecuencia de entrada en el orden que sigue los índices 0 a N-1, se podrían identificar, donde todos los coeficientes en el dominio de la frecuencia 24 de las respectivas bandas de factor de escala están cuantificados en cero. Tales coeficientes en el dominio de la frecuencia podrían ser llenados, por ejemplo, mediante la generación de ruido artificial tal como, por ejemplo, utilizando un generador de números pseudoaleatorios. La intensidad/nivel del ruido relleno en una banda de factor de escala cuantificada en cero se podría ajustar mediante la utilización del factor de escala de la respectiva banda de factor de escala ya que la misma no es necesaria para el escalado ya que los coeficientes espectrales en la misma son todos cero. Tal relleno de ruido se muestra en la Figura 2 con el número de referencia 40 y se describe con más detalle en una realización en la patente EP2304719A1 [6].

65

[0030] La Figura 2 muestra que el transformador inverso 16 puede estar configurado para soportar la codificación conjunta-estéreo y/o la predicción estéreo entre canales. En la infraestructura de la predicción estéreo entre canales, el transformador inverso 16 podría predecir por ejemplo 42 el espectro en la disposición no desentrelazada representada por el orden de índices 0 a N-1 procedente de otro canal de la señal de audio. Es decir, 5 podría ser el caso de que los coeficientes en el dominio de la frecuencia 24 describan el espectrograma de un canal de una señal de audio estéreo, y que el transformador inverso 16 está configurado para tratar los coeficientes en el dominio de la frecuencia 24 como un residuo de predicción de una señal de predicción derivada del otro canal de esta 10 señal de audio estéreo. Esta predicción estéreo entre canales podría llevarse a cabo por ejemplo con cierta granularidad espectral independientemente de la señalización 34. Los parámetros de predicción complejos 44 que controlan la predicción estéreo compleja 42 podrían por ejemplo activar la predicción estéreo compleja 42 para determinadas bandas de entre las bandas de factor de escala anteriormente mencionadas. Para cada banda de factor de escala para la que se activa la predicción compleja mediante el parámetro de predicción complejo 44, los 15 coeficientes en el dominio de la frecuencia escalados 24, dispuestos en un orden de 0 a N-1, que residen dentro de la respectiva banda de factor de escala, serían sumados junto con la señal de predicción entre canales obtenida del otro canal de la señal de audio estéreo. Un factor complejo contenido dentro de los parámetros de predicción compleja 44 para esta respectiva banda de factor de escala podría controlar la señal de predicción.

[0031] Además, dentro de la infraestructura de codificación conjunta-estéreo, el transformador inverso 16 podría estar configurado para llevar a cabo la decodificación MS 46. Es decir, el decodificador 10 de la Figura 1 podría 20 llevar a cabo las operaciones descritas hasta aquí, dos veces, una vez para un primer canal y otra vez para un segundo canal de una señal de audio estéreo, y se podría controlar por medio de parámetros MS dentro de la corriente de datos 20, el transformador inverso 16 podría decodificar MS estos dos canales o dejarlos como están, a saber, como canales izquierdo y derecho de la señal de audio estéreo. Los parámetros MS 48 podrían conmutar entre la codificación MS en un nivel de trama o incluso a cierto nivel más fino tal como en unidades de bandas de factores de escala o grupos 25 de éstas. En el caso de una decodificación activada de MS, por ejemplo, el transformador inverso 16 podría formar una suma de los correspondientes coeficientes en el dominio de la frecuencia 24 en el orden de los coeficientes de 0 a N-1, con correspondientes coeficientes en el dominio de la frecuencia del otro canal de la señal de audio estéreo, o una diferencia de los mismos.

[0032] La Figura 2 muestra entonces que el desentrelazador activable 38 responde a la señalización 34 para la trama actual a efectos de, si la trama actual está señalizada por la señalización 34 como una trama de transformada dividida, desentrelazar los coeficientes en el dominio de la frecuencia de entrada, para obtener dos transformadas, a 30 saber una transformada delantera 50 y una transformada trasera 52, y dejar los coeficientes en el dominio de la frecuencia entrelazados para dar como resultado una única transformada 54 en caso de que la señalización 34 indique que la trama actual es una trama de transformada larga. En el caso de un desentrelazado, el desentrelazador 38 forma una transformada única a partir de 50 y 52, una primera transformada corta a partir de los coeficientes en el dominio de la frecuencia que tienen índices pares, y la otra transformada corta a partir de los coeficientes en el dominio de la frecuencia en posiciones de índices impares. Por ejemplo, los coeficientes en el dominio de la frecuencia de índice 35 par podrían formar la transformada delantera (cuando se empieza con el índice 0), mientras que los otros forman la transformada trasera. Las transformadas 50 y 52 están sometidas a transformación inversa de longitud de transformada más corta, resultando en porciones en el dominio del tiempo 56 y 58, respectivamente. El combinador 40 18 de la Figura 1 posiciona correctamente las porciones en el dominio de tiempo 56 y 58 en el tiempo, particularmente la porción en el dominio del tiempo 56 resultante de la transformada delantera 50 situada en frente de la porción en el dominio del tiempo 58 resultante de la porción trasera 52, y lleva a cabo el procedimiento de superposición y adición 45 entre las mismas y con porciones en el dominio del tiempo que resultan de tramas anteriores y sucesivas de la señal de audio. En el caso de no desentrelazado, los coeficientes en el dominio de la frecuencia que llegan al entrelazador 38 constituyen la transformada larga 54 tal como son, y la fase de transformación inversa 40 lleva a cabo una transformada inversa sobre la misma para resultar en una porción en el dominio del tiempo 60 que abarca, y se extiende más allá del intervalo de tiempo completo de la trama actual 26. El combinador 18 combina la porción en el 50 dominio del tiempo 60 con respectivas porciones en el dominio del tiempo resultantes de tramas anteriores y sucesivas de la señal de audio.

[0033] El decodificador de audio en el dominio de la frecuencia descrito hasta aquí permite la conmutación de longitud de transformada de una manera que es compatible con los decodificadores de audio en el dominio de la 55 frecuencia que no responden a la señalización 34. En particular, tales decodificadores "anticuados" supondrían erróneamente que las tramas que son actualmente señaladas por la señalización 34 como del tipo de transformada dividida, serían de un tipo de transformada larga. Es decir, erróneamente dejarían los coeficientes en el dominio de la frecuencia de tipo dividido entrelazados y llevarían a cabo una transformación inversa de la longitud de transformada larga. Sin embargo, la calidad resultante de las tramas afectadas de la señal de audio reconstruida seguiría siendo 60 bastante razonable.

[0034] A su vez, la penalización en cuanto a la eficiencia de codificación es también bastante razonable. La penalización de la eficiencia de codificación resulta de no tener en cuenta la señalización 34 como los coeficientes en el dominio de la frecuencia y los factores de escala se codifican sin tomar en cuenta el significado variante de los 65 coeficientes y explotar esta variación para incrementar la eficiencia de codificación. Sin embargo, esta última

penalización es comparativamente pequeña en comparación con la ventaja de permitir una retrocompatibilidad. Esta última afirmación también es cierta con respecto a la restricción para activar y desactivar el relleno de ruido 40, la predicción de estéreo compleja 42 y la decodificación de MS 46 meramente con porciones espectrales continuas (bandas de factor de escala) en el estado desentrelazado definido por los índices 0 a N-1 en la Figura 2. La oportunidad de renderizar el control de estas herramientas de codificación específicamente para el tipo de trama (por ejemplo, que tiene dos niveles de ruido) posiblemente podría proporcionar ventajas, pero las ventajas están sobrecompensadas por la ventaja de tener una retrocompatibilidad.

[0035] La Figura 2 muestra que el decodificador de la Figura 1 podría estar configurado incluso para soportar la codificación TNS mientras que, sin embargo, mantiene la retrocompatibilidad con decodificadores que son insensibles con respecto a la señalización 34. En particular, la Figura 2 ilustra la posibilidad de llevar a cabo el filtrado inverso de TNS después de cualquier predicción de estéreo compleja 42 y de decodificación de MS 46, en caso de haberlas. Con el fin de mantener la retrocompatibilidad, el transformador inverso 16 está configurado para llevar a cabo el filtrado TNS 62 sobre una secuencia de N coeficientes independientemente de la señalización 34 mediante la utilización de respectivos coeficientes TNS 64. Mediante esa medida, la corriente de datos 20 codifica los coeficientes de TNS 64 equitativamente, de manera independiente de la señalización 34. Es decir, el número de coeficientes de TNS y la manera de codificarlos son los mismos. Sin embargo, el transformador inverso 16 está configurado para aplicar los coeficientes TNS 64 de manera diferente. En el caso de que la trama actual sea una trama de transformada larga, el filtrado inverso de TNS se lleva a cabo sobre la transformada larga 54, es decir, estando los coeficientes en el dominio de la frecuencia en secuencia en el estado entrelazado, y en el caso de que la trama actual esté señalizada por la señalización 34 como trama de transformada dividida, el transformador inverso 16 invierte los filtros de TNS 62 una concatenación de transformada delantera 50 y de transformada trasera 52, es decir, la secuencia de coeficientes en el dominio de la frecuencia de los índices 0, 2, ..., N-2, 1, 3, 5, ..., N-1. El filtrado inverso de TNS 62 puede, por ejemplo, implicar que el transformador inverso 16 aplique un filtro, cuya función de transferencia se establece según los coeficientes de TNS 64 sobre la secuencia desentrelazada o entrelazada de coeficientes que han pasado por la secuencia de procesamiento del desentrelazador situado corriente arriba 38.

[0036] Por lo tanto, un decodificador "anticuado" que accidentalmente trate tramas de tipo transformada dividida como tramas de transformada larga, aplica coeficientes de TNS 64 que han sido generados por un codificador mediante el análisis de una concatenación de dos transformadas cortas, particularmente 50 y 52, sobre la transformada 54 y por lo tanto produce, por medio de la transformada inversa aplicada sobre la transformada 54, una porción incorrecta en el dominio del tiempo 60. Sin embargo, incluso este deterioro de calidad en tales codificadores podría ser tolerable para los oyentes en caso de restringirse el uso de tales tramas de transformada dividida a ocasiones donde la señal representa lluvia o aplausos o similares.

[0037] En aras de integridad, en la Figura 3 se muestra que el filtrado inverso de TNS 62 del transformador inverso 16 también se puede insertar en otra parte en la secuencia de procesamiento mostrada en la Figura 2. Por ejemplo, el filtrado inverso de TNS 62 podría estar posicionado corriente arriba de la predicción de estéreo compleja 42. Con el fin de mantener el dominio de desentrelazado corriente abajo y corriente arriba de filtrado inverso de TNS 62, la Figura 3 muestra que en el caso en que los coeficientes en el dominio de la frecuencia 24 son meramente desentrelazados 66 de manera preliminar, a efecto de llevar a cabo el filtrado inverso de TNS 68 dentro del estado concatenado desentrelazado donde los coeficientes en el dominio de la frecuencia 24 procesados hasta ahora tienen el orden de índices 0, 2, 4, ..., N-2, 1, 3, ..., N-3, N-1, a lo cual se invierte el desentrelazado 70 para obtener los coeficientes en el dominio de la frecuencia en la versión filtrada de TNS inversamente en su orden entrelazado 0, 1, 2, ..., N-1, de nuevo. La posición del filtrado inverso de TNS 62 dentro de la secuencia de etapas de procesamiento mostrada en la Figura 2 se podría fijar o señalar mediante la corriente de datos 20 tal como por ejemplo sobre la base de trama por trama o con alguna otra granularidad.

[0038] Cabe observar que, a efectos de facilitar la descripción, las realizaciones anteriormente descritas se centran sobre la yuxtaposición de tramas de transformada larga y tramas de transformada dividida solamente. Sin embargo, las realizaciones de la presente solicitud pueden ampliarse mucho por la introducción de tramas de otro tipo de transformada tales como tramas de ocho transformadas cortas. En este aspecto, cabe destacar que el agnosticismo anteriormente mencionado se refiere meramente a las tramas diferenciadas, por medio de una señalización adicional, con respecto a otras tramas de cualquier otro tercer tipo de transformada de tal manera que un decodificador "anticuado", al inspeccionar la señalización adicional contenida en todas las tramas, accidentalmente trate las tramas de transformada dividida como tramas de transformada larga, y meramente las tramas diferenciadas de las otras tramas (todo con excepción de tramas de transformada dividida y tramas de transformada larga) comprenderían la señalización 34. En lo que se refiere a tales otras tramas (todo con excepción de tramas de transformada dividida y tramas de transformada larga), cabe destacar que el modo operativo de los extractores 12 y 14 tales como la selección de contexto, etc., podría depender de la señalización adicional, es decir, dicho modo de operación podría ser diferente del modo de operación aplicado para las tramas de transformada dividida y tramas de transformada larga.

[0039] Antes de describir un codificador adecuado que se ajuste a las realizaciones del decodificador descritas anteriormente, se describe una implementación de las realizaciones anteriormente descritas que sería adecuada para mejorar por consiguiente los codificadores/decodificadores de audio basados en xHE-AAC a efectos de permitir el

soporte de la división de transformada de una manera retrocompatible.

[0040] Es decir, en lo sucesivo, se describe una posibilidad de cómo llevar a cabo la división de longitud de transformada en un códec de audio que está basado en MPEG-D xHE-AAC (USAC) con el objetivo de mejorar la calidad de codificación de determinadas señales de audio con bajas tasas de bits. La herramienta de división de transformada se señala con una semirretrocompatibilidad de tal manera que los decodificadores xHE-AAC heredados puedan analizar y decodificar corrientes de bits según las realizaciones anteriores sin errores de audio evidentes ni pérdidas de señal. Como se mostrará en lo sucesivo, esta señalización semirretrocompatible aprovecha posibles valores no utilizados de un elemento de sintaxis de trama que controla, de una manera condicionalmente codificada, la utilización del relleno de ruido. Mientras los decodificadores xHE-AAC heredados no son sensibles a estos posibles valores del respectivo elemento de sintaxis de relleno de ruido, si lo son los decodificadores de audio reforzados.

[0041] En particular, la implementación descrita en lo anterior permite, en sintonía con las realizaciones anteriormente descritas, ofrecer una longitud de transformada intermedia para codificar señales similares a la lluvia o aplausos, preferentemente un bloque largo dividido, es decir, dos transformadas secuenciales, cada una de ellas de la mitad o un cuarto de la longitud espectral de un bloque largo, con un máximo solapamiento de tiempo entre estas transformadas inferior a un máximo solapamiento temporal entre bloques largos consecutivos. Para permitir que las corrientes de bits codificadas con división de transformada, es decir, señalización 34, sean leídas y analizadas por decodificadores xHE-AAC heredados, debería utilizarse la división de una manera semirretrocompatible: la presencia de tal herramienta de división de transformada no debería hacer que decodificadores heredados detengan o ni siquiera empiecen la codificación. La legibilidad de tales corrientes de bits por la infraestructura de xHE-AAC también puede facilitar su adopción en el mercado. Para lograr el objetivo anteriormente mencionado de semirretrocompatibilidad para la utilización de la división de transformada en el contexto de xHE-AAC o de sus derivados potenciales, se señala una división de transformada por medio de la señalización de relleno de ruido de xHE-AAC. En cumplimiento con las realizaciones anteriormente descritas, y a efectos de incorporar la división de transformada en codificadores/decodificadores de xHE-AAC, en lugar de una secuencia de ventana de parada-arranque en el dominio de la frecuencia (FD), puede utilizarse una transformada dividida consistente en dos transformadas separadas de media longitud. Las transformadas de media longitud temporalmente secuenciales están entrelazadas en un solo bloque similar a parada-arranque en un modo de coeficiente por coeficiente para decodificadores que no soporten la división de transformada, es decir, decodificadores xHE-AAC heredados. La señalización mediante señalización de relleno de ruido se lleva a cabo como se describe en lo sucesivo. En particular, la información secundaria del relleno de ruido de 8 bits puede utilizarse para transportar la división de transformada. Esto es viable porque el estándar MPEG-D [4] establece que la totalidad de 8 bits se transmiten incluso si el nivel de ruido que se va a aplicar es cero. En esa situación, algunos de los bits de relleno de ruido se pueden reutilizar para la división de transformada, es decir, para la señalización 34.

[0042] La semirretrocompatibilidad en cuanto al análisis y reproducción de corrientes de bits por decodificadores xHE-AAC heredados puede asegurarse del siguiente modo. La división de transformada se señala mediante un nivel de ruido de cero, es decir, los primeros tres bits de relleno de ruido tienen todos un valor de cero, seguido por cinco bits no cero (que tradicionalmente representan una compensación de ruido) que contiene información secundaria con respecto a la división de transformada, así como también al nivel de ruido que falta. Ya que un decodificador xHE-AAC heredado no toma en cuenta el valor de la compensación de 5 bits si el nivel de ruido de 3 bits es cero, la presencia de señalización de división de transformada 34 solamente tiene un efecto en el relleno de ruido en el decodificador heredado: el relleno de ruido se desactiva ya que los tres primeros bits son cero, y el resto de la operación de decodificación se ejecuta como se prevé. En particular, una transformada dividida se procesa de la misma manera que un bloque de parada-arranque tradicional con una transformada inversa de longitud completa (debido al entrelazado de coeficientes anteriormente mencionado) y no se lleva a cabo ningún desentrelazado. Por ello, un decodificador heredado sigue ofreciendo una decodificación "atractiva" de la corriente de datos/corriente de bits reforzada ya que no es necesario silenciar la señal de salida o ni siquiera abortar la decodificación cuando se llegue a una trama del tipo de división de transformada. Naturalmente, tal decodificador heredado no es capaz de proporcionar una correcta reconstrucción de tramas de transformada dividida, lo que conduce a una calidad deteriorada en las tramas afectadas en comparación con la decodificación mediante un decodificador adecuado según la Figura 1, por ejemplo. No obstante, suponiendo que la división de transformada se utilice como se esperaba, es decir solamente sobre entrada transitoria o ruidosa a bajas tasas de bits, la calidad por medio de un decodificador xHE-AAC debería ser mejor que si las tramas afectadas se abandonaran debido al silenciamiento o de alguna otra manera condujera a errores de reproducción obvios.

[0043] En concreto, una extensión de un codificador/decodificador xHE-AAC hacia la división de transformada podría ser del siguiente modo.

[0044] Según la descripción anterior, la nueva herramienta que se utilizará para XHE-AAC se podría denominar división de transformada (TS). Sería una nueva herramienta en el codificador en el dominio de la frecuencia (FD) de xHE-AAC o, por ejemplo, audio 3D de MPEG-H basado en USAC [4]. La división de transformada sería entonces utilizable en determinados pasos de señal transitoria como alternativa a transformadas largas regulares (que conducen a deformación de tiempo, especialmente preeco a bajas tasas de bits) u 8 transformadas cortas (que conducen a

agujeros espectrales y a artefactos de burbujas a bajas tasas de bits). Entonces la TS se podría señalar como semirretrocompatible por entrelazado de coeficientes de FD en una transformada larga que puede ser analizada correctamente por un decodificador MPEG-D USAC heredado.

5 **[0045]** Una descripción de esta herramienta sería similar a la descripción anterior. Cuando la TS está activa en una transformada larga, se emplean dos MDCT de media longitud en lugar de una MDCT de longitud completa, y los coeficientes de las dos MDCT, es decir, 50 y 52, se transmiten de una forma entrelazada línea por línea. La transmisión entrelazada ya se utilizó, por ejemplo, en el caso de transformadas de arranque de FD (parada), con los coeficientes de la MDCT de primera vez colocados en índices pares y con los coeficientes de la MDCT de segunda vez colocados en índices impares (donde la indexación comienza en cero), pero un decodificador que no es capaz de gestionar transformadas de parada-arranque no habría sido capaz de analizar correctamente la corriente de datos. Es decir, debido a los diferentes contextos utilizados para la codificación por entropía, los coeficientes en el dominio de la frecuencia sirven para tal transformada de parada-arranque, una sintaxis variada racionalizada sobre las transformadas a la mitad, cualquier decodificador incapaz de soportar ventanas de parada-arranque habría tenido que pasar por alto las respectivas tramas de ventana de parada-arranque.

[0046] Volviendo brevemente a la realización anteriormente descrita, esto significa que el decodificador de la Figura 1 podría ser, más allá de la descripción anticipada hasta el momento, capaz de soportar alternativamente mayor longitud de transformada, es decir, una subdivisión de determinadas tramas 26 en incluso más de dos transformadas mediante una señalización que extiende la señalización 34. Sin embargo, en cuanto a la yuxtaposición de subdivisiones de transformada de tramas 26, aparte de la transformada dividida activada mediante la señalización 34, el extractor de coeficiente de FD 12 y el extractor de factor de escala 14 serían sensibles a esta señalización por el hecho de que su modo de operación cambiaría en dependencia de dicha señalización adicional, además de la señalización 34. Además, una transmisión racionalizada de los coeficientes de TNS, los parámetros de MS y los parámetros de predicción complejos, adaptados al tipo de transformada señalizada distinto del tipo de transformada dividida según 56 y 59, necesitaría que cada decodificador tuviera que ser capaz de responder, es decir, entender, la selección de señalización entre estos "tipos de transformada conocida" o tramas que incluyen el tipo de transformada larga según 60, y otros tipos de transformada tales como uno que subdivide las tramas en ocho transformadas cortas como en el caso de AAC, por ejemplo. En este caso, esta "señalización conocida" identificaría las tramas para las que la señalización 34 señala el tipo de transformada dividida, como tramas del tipo de transformada larga de tal manera que los decodificadores no capaces de entender la señalización 34 tratarían estas tramas como tramas de transformada larga en lugar de tramas de otros tipos, tales como tramas de tipo 8 transformadas cortas.

[0047] Volviendo a la descripción de una posible extensión de XHE-AAC, ciertas restricciones operativas podrían proporcionarse para incorporar una herramienta de TS en esta infraestructura de codificación. Por ejemplo, se podría permitir el uso de TS solamente en una ventana de largo-arranque o de parada-arranque de FD. Es decir, se podría requerir que la secuencia de ventana de elemento de sintaxis subyacente sea igual a 1. Además, debido a la señalización semirretrocompatible, puede ser un requisito que la TS pueda aplicarse solamente cuando el noiseFilling (relleno de ruido) del elemento de sintaxis sea uno en el contenedor de sintaxis UsacCoreConfig(). Cuando se señala que la TS está activa, todas las herramientas de FD excepto TNS y MDCT inversa operan en el conjunto entrelazado (largo) de coeficientes de TS. Esto permite la reutilización de la compensación de banda de factor de escala y tablas del codificador aritmético de transformada larga así como también de las formas de ventanas y las longitudes de solapamiento.

45 **[0048]** En lo que sigue, se presentan los términos y definiciones que se utilizan seguidamente a efectos de explicar cómo podría ampliarse el estándar USAC descrito en [4] para ofrecer la funcionalidad de TS retrocompatible, donde a veces se hace referencia a secciones dentro de ese estándar para el lector interesado.

50 **[0049]** Un nuevo elemento de datos podría ser:

split_transform bandera binaria que indica si se utiliza TS en la trama y canal actuales

[0050] Nuevos elementos de ayuda podrían ser:

window_sequence	tipo de secuencia de ventana de FD para la trama y canal actuales (sección 6.2.9)
noise_offset	compensación de relleno de ruido para modificar factores de escala de bandas cuantificadas en cero (sección 7.2)
noise_level	nivel de relleno de ruido que representa amplitud de ruido de espectro añadido (sección 7.2)
half_transform_length	una mitad de coreCoderFrameLength (ccfl, la longitud de transformada, sección 6.1.1)
half_lowpass_line	una mitad del número de líneas de MDCT transmitidas para el canal actual.

55 **[0051]** La decodificación de una transformada de arranque (parada) de FD utilizando división de transformada (TS) en la infraestructura de USAC podría llevarse a cabo en etapas puramente secuenciales, del siguiente modo:

[0052] Primero, podría llevarse a cabo una decodificación de `split_transform` y `half_lowpass_line`. `split_transform` realmente no representaría un elemento de corriente de bits independiente, sino que se deriva de elementos de relleno de ruido, `noise_offset` y `noise_level`, y en caso de un `UsacChannelPairElement()`, la bandera `common_window` en `StereoCoreToolInfo()`. Si `noiseFilling = 0`, `split_transform` es 0. En caso contrario,

```
5
si ((noiseFilling != 0) && (noise_level == 0)) {
split_transform = (noise_offset & 16) / 16;
noise_level = (noise_offset & 14) / 2;

noise_offset = (noise_offset & 1) * 16;
}
en otro caso {
split_transform = 0;
}
```

[0053] En otras palabras, si `noise_level == 0`, `noise_offset` contiene la bandera `split_transform` seguida por 4 bits de datos de relleno de ruido, los cuales se reordenan entonces. Ya que esta operación cambia los valores de `noise_level` y `noise_offset`, se debe ejecutar antes del procedimiento de relleno de ruido de la sección 7.2. Por otra parte, si `common_window == 1` en un `UsacChannelPairElement()`, `split_transform` se determina solamente en el canal izquierdo (primero), la `split_transform` del canal derecho se establece igual a (es decir, se copia de) `split_transform` del canal izquierdo, y el pseudocódigo anteriormente mencionado no se ejecuta en el canal derecho.

Se determina `half_lowpass_line` a partir de la tabla de compensación de bandas de factor de escala "larga", `swb_offset_long_window`, y `max_sfb` del canal actual, o en caso de `stereo` y `common_window == 1`, `max_sfb_ste`.

```
max_sfb_ste en elementos con StereoCoreToolInfo() y common_window == 1,
lowpass_sfb =
max_sfb en caso contrario. En base a la bandera de igFilling, se deriva half_lowpass_line:
si (igFilling != 0) {
lowpass_sfb = max(lowpass_sfb, ig_stop_sfb);
}
half_lowpass_line = swb_offset_long_window[lowpass_sfb] / 2;
```

[0054] Entonces, como segunda etapa, se llevaría a cabo el desentrelazado de espectros de media longitud para conformación de ruido temporal.

[0055] Después de la descuantificación de espectros, el relleno de ruido, y la aplicación del factor de escala y antes de la aplicación de Conformación de Ruido Temporal (TNS), se desentrelazan los coeficientes de TS en `spec[]` utilizando un búfer auxiliar (`helper buffer[]`):

```
25
para (i = 0, i2 = 0; i < half_lowpass_line; i += 1, i2 += 2) {
spec[i] = spec[i2]; /* aislar 1ª ventana */
buffer[i] = spec[i2+1]; /* aislar 2ª ventana */
}
para (j = 0; j < half_lowpass_line; j += 1) {
spec[j+half_lowpass_line] = buffer[j]; /* copiar 2ª ventana */
}
```

[0056] El desentrelazado in situ pone efectivamente los dos espectros de TS de mediana longitud uno por encima del otro, y la herramienta de TNS opera ahora como siempre sobre el pseudoespectro de longitud completa resultante.

[0057] Con referencia a lo anterior, se ha descrito un procedimiento con respecto a la Figura 3.

[0058] Entonces, como tercera etapa, se utilizaría el entrelazado temporal junto con dos MDCT inversas secuenciales.

[0059] Si `common_window == 1` en la trama actual o si se lleva a cabo la decodificación estéreo después de la decodificación de TNS (`tns_on_lr == 0` en la sección 7.8), `spec[]` debe ser reentrelazado temporalmente en un espectro de longitud completa:

40

```

para (i = 0; i < half_lowpass_line; i += 1) {
buffer[i] = spec[i];                               /* copiar 1ª ventana */
}
para (i = 0, i2 = 0; i < half_lowpass_line; i += 1, i2 += 2) {
spec[i2] = buffer[i];                             /* fusionar 1ª ventana */
spec[i2+1] = spec[i+half_lowpass_line];          /* fusionar 2ª ventana */
}

```

[0060] Se utiliza el pseudoespectro resultante para la decodificación estéreo (sección 7.7) y para actualizar `dmx_re_prev[]`

5

(secciones 7.7.2 y A.1.4). En caso de `tns_on_lr == 0`, los espectros de longitud completa estéreo-decodificados se desentrelazan de nuevo al repetir el procedimiento de la sección A.1.3.2. Finalmente se calculan las dos MDCT inversas con `ccfl` y la `window_shape` del canal de la trama actual y la última. Véase la sección 7.9 y la Figura 1.

10 **[0061]** Puede efectuarse alguna modificación a la decodificación estéreo de predicciones complejas de xHE-AAC.

[0062] Como alternativa puede utilizarse un procedimiento de señalización semirretrocompatible implícito para incorporar TS en xHE-AAC.

15

[0063] En lo anterior se describe una estrategia que utiliza un bit en una corriente de bits para señalar el uso de la división inventiva de transformada, contenida en `split_transform`, a un decodificador inventivo. En particular, tal señalización (llamémosla señalización semirretrocompatible explícita) permite que los siguientes datos de corriente de bits heredados - en este caso, la información secundaria de relleno de ruido - se utilicen independientemente de la señal inventiva: en la presente realización, los datos de relleno de ruido no dependen de los datos de división de transformada, y viceversa. Por ejemplo, los datos de relleno de ruido constituidos por todos los ceros (`noise_level = noise_offset = 0`) se pueden transmitir mientras que `split_transform` puede contener cualquier valor posible (siendo una bandera binaria, ya sea 0 o 1).

25 **[0064]** En casos donde no se requiera tal independencia estricta entre los datos de corriente de bits heredados e inventivos y la señal inventiva es una decisión binaria, la transmisión explícita de un bit de señalización se puede evitar, y dicha decisión binaria se puede señalar mediante la presencia o ausencia de lo que puede llamarse señalización semirretrocompatible implícita. Tomando de nuevo la realización anterior como un ejemplo, el uso de la división de transformada podría transmitirse simplemente mediante el uso de la señalización inventiva: si `noise_level` es cero y al mismo tiempo, `noise_offset` no es cero, entonces `split_transform` se establece igual a 1. Si tanto `noise_level` como `noise_offset` no son cero, `split_transform` se establece igual a 0. Surge una dependencia de la señal implícita inventiva con respecto a la señal de relleno de ruido heredada cuando tanto `noise_level` como `noise_offset` son cero. En este caso, no está claro si se está utilizando señalización heredada o implícita inventiva. Para evitar esta ambigüedad, el valor de `split_transform` se debe definir con anterioridad. En el presente ejemplo, es adecuado definir

30 `split_transform = 0` si los datos de relleno de ruido consisten en todos cero, ya que esto es lo que los codificadores heredados sin división de transformada señalarán cuando en una trama no debe utilizarse el relleno de ruido.

[0065] Queda por resolver el tema de cómo la señalización implícita semirretrocompatible debe señalar la señal `split_transform == 1` y no relleno de ruido al mismo tiempo. Como ya se explicó, los datos de relleno de ruido no deben ser todos cero, y si se requiere una magnitud de ruido de cero, el `noise_level` ($((\text{noise_offset} \& 14)/2)$ como en lo anterior) debe ser igual a 0. Esto deja solamente un `noise_offset` ($((\text{noise_offset} \& 1)*16)$ como anteriormente) mayor que 0 como una solución. Afortunadamente, el valor de `noise_offset` se ignora si no se lleva a cabo un relleno de ruido en un decodificador basado en USAC [4], por lo que esta estrategia resulta ser factible en la presente realización. Por ello, la señalización de `split_transform` en el pseudocódigo anteriormente descrito podría modificarse del siguiente modo, utilizándose el bit de señalización de TS guardado para transmitir 2 bits (4 valores) en lugar de 1 bit para

45 `noise_offset`:

```

si ((noiseFilling != 0) && (noise_level == 0) && (noise_offset != 0)) {
split_transform = 1;
noise_level = (noise_offset & 28) / 4;
noise_offset = (noise_offset & 3) * 8;
}
en otro caso {
split_transform = 0;
}

```

50 **[0066]** Por lo tanto, al aplicar esta alternativa, la descripción de USAC podría ampliarse mediante la siguiente descripción.

[0067] La descripción de la herramienta sería ampliamente la misma. Es decir:

[0068] Cuando la división de transformada (TS) es activa en una transformada larga, se emplean dos MDCT de media longitud en lugar de una MDCT de longitud completa. Los coeficientes de las dos MDCT se transmiten de una manera entrelazada línea por línea como transformada tradicional de dominio de frecuencia (FD), estando los coeficientes de la MDCT de primera vez colocados en índices pares y los coeficientes de la MDCT de segunda vez colocados en índices impares.

10 **[0069]** Las restricciones operativas podrían requerir que la TS se pueda utilizar solamente en una ventana de largo-arranque o parada-arranque de FD (`window_sequence == 1`) y que la TS pueda aplicarse solamente cuando `noiseFilling` es 1 en `UsacCoreConfig()`. Cuando se señala TS, todas las herramientas de FD excepto para TNS y MDCT inversa operan en el conjunto entrelazado (largo) de coeficientes de TS. Esto permite la reutilización de la compensación de banda de factor de escala y de las tablas de decodificador aritmético de transformada larga así como también de las formas de ventana y de longitudes de solapamiento.

[0070] Los términos y definiciones utilizados en lo que sigue implican los siguientes elementos de ayuda

<code>common_window</code>	indica si el canal 0 y el canal 1 de un CPE utilizan parámetros de ventana idénticos (véase ISO/IEC 23003-3:2012 sección 6.2.5.1.1).
<code>window_sequence</code>	tipo de secuencia de ventana de FD para la trama y canal actuales (véase ISO/IEC 23003-3:2012 sección 6.2.9).
<code>tns_on_lr</code>	indica el modo operativo del filtrado de TNS (véase ISO/IEC 23003-3:2012 sección 7.8.2).
<code>noiseFilling</code>	esta bandera señala la utilización de relleno de ruido de agujeros espectrales en el codificador de núcleo de FD (véase ISO/IEC 23003-3:2012 sección 6.1.1.1).
<code>noise_offset</code>	compensación de relleno de ruido para modificar factores de escala de bandas cuantificadas en cero (véase ISO/IEC 23003-3:2012 sección 7.2).
<code>noise_level</code>	nivel de relleno de ruido que representa la amplitud de ruido de espectro añadido (véase ISO/IEC 23003-3:2012 sección 7.2).
<code>split_transform</code>	bandera binaria que indica si se utiliza TS en la trama y canal actuales.
<code>half_transform_length</code>	una longitud de <code>coreCoderFrameLength</code> (<code>ccfl</code> , la longitud de transformada, véase ISO/IEC 23003-3:2012 sección 6.1.1).
<code>half_lowpass_line</code>	una mitad del número de líneas de MDCT transmitidas por el canal actual.

20 **[0071]** El procedimiento de decodificación que implica TS podría describirse del siguiente modo. En particular, la decodificación de una transformada de (parada-)arranque de FD con TS se lleva a cabo en tres etapas secuenciales del siguiente modo.

[0072] Primero, se lleva a cabo la decodificación de `split_transform` y de `half_lowpass_line`. El elemento de ayuda `split_transform` no representa un elemento de corriente de bits independiente sino que se deriva de los elementos de relleno de ruido, `noise_offset` y `noise_level`, y en el caso de un `UsacChannelPairElement()`, la bandera `common_window` en `StereoCoreToolInfo()`. Si `noiseFilling == 0`, `split_transform` es 0. En caso contrario,

```

si (noiseFilling != 0) && (noise_level == 0) && (noise_level == 0) {
split_transform = 1;
noise_level = (noise_offset & 28) / 4;
noise_offset = (noise_offset & 3) * 8;
}
en otro caso {
split_transform = 0;
}

```

30 **[0073]** En otras palabras, si `noise_level == 0`, `noise_offset` contiene la bandera `split_transform` seguido por 4 bits de datos de relleno de ruido, que son ordenados entonces. Dado que esta operación cambia los valores de `noise_level` y de `noise_offset`, se deben ejecutar antes del procedimiento de relleno de ruido de ISO/IEC 23003-3:2012 sección 7.2.

35 **[0074]** Por otra parte, si `common_window == 1` en un `UsacChannelPairElement()`, `split_transform` se determina solamente en el canal izquierdo (primero); la `split_transform` del canal derecho se establece igual a (es decir, se copia de) la `split_transform` del canal izquierdo, y el pseudocódigo anteriormente mencionado no se ejecuta en el canal derecho.

40 **[0075]** El elemento de ayuda `half_lowpass_line` se determina a partir de la tabla de compensaciones de las

bandas de factor de escala "larga", `swb_offset_long_window`, y de `max_sfb` del canal actual, o en el caso de stereo y `common_window == 1`, `max_sfb_ste`.

`lowpass_sfb = {max_sfb_ste en elementos StereoCoreToolInfo() y common_window = 1.
max_sfb de lo contrario.`

5

[0076] Según la bandera de `igFilling`, se deriva `half_lowpass_line`:

```
si (igFilling != 0) {
  lowpass_sfb = max(lowpass_sfb, ig_stop_sfb);
}
10 half_lowpass_line = swb_offset_long_window[lowpass_sfb] / 2;
```

[0077] Entonces, se lleva a cabo el desentrelazado de los espectros de media longitud para la conformación de ruido temporal.

15 **[0078]** Después de la descuantificación del espectro, el relleno de ruido, y aplicación del factor de escala y antes de la aplicación de la Conformación de Ruido Temporal (TNS), se desentrelazan los coeficientes de TN en `spec[]` mediante un búfer auxiliar (`helper buffer[]`):

```
para (i = 0, i2 = 0; i < half_lowpass_line; i += 1, i2 += 2) {
  spec[i] = spec[i2]; /* aislar 1ª ventana */
  buffer[i] = spec[i2+1]; /* aislar 2ª ventana */
}

para (i = 0; i < half_lowpass_line; i += 1) {
  spec[i+half_lowpass_line] = buffer[i]; /* copiar 2ª ventana */
}
20
```

[0079] El desentrelazado realizado in situ pone de manera efectiva los dos espectros de TS de media longitud uno por encima del otro, y la herramienta de TNS opera ahora como siempre sobre el pseudoespectro de media longitud resultante.

25 **[0080]** Finalmente, puede utilizarse el reentrelazado temporal y dos MDCT secuenciales inversas:

[0081] Si `common_window == 1` en la trama actual o la decodificación estéreo se lleva a cabo después de la decodificación de TNS (`tns_on_lr == 0` en la sección 7.8), `spec[]` debe entrelazarse temporalmente en un espectro de longitud completa:

30

```
para (i = 0; i < half_lowpass_line; i += 1) {
  buffer[i] = spec[i]; /* copiar 1ª ventana */
}
para (i = 0, i2 = 0; i < half_lowpass_line; i += 1, i2 += 2) {
  spec[i2] = buffer[i]; /* fusionar 1ª ventana */
  spec[i2+1] = spec[i+half_lowpass_line]; /* fusionar 2ª ventana */
}
35
```

[0082] El pseudoespectro resultante se utiliza para la decodificación estéreo (ISO/IEC 23003-3:2012 sección 7.7) y para actualizar `dmx_re_prev[]` (ISO/IEC 23003-3:2012 sección 7.7.2) y en el caso en que `tns_on_lr == 0`, los espectros de longitud completa decodificados en estéreo se desentrelacen de nuevo mediante la repetición del procedimiento de la sección. Finalmente, se calculan las dos MDCT inversas con `ccfl` y la `window_shape` del canal de la trama actual y la última.

40 **[0083]** El procesamiento para TS sigue la descripción dada en ISO/IEC 23003-3:2012 sección "7.9 Filterbank and block switching". Deben tomarse en cuenta las siguientes adiciones.

[0084] Los coeficientes de TS en `spec[]` son desentrelazados mediante un búfer auxiliar (`helper buffer[]`) con `N`, la longitud de ventana basada en el valor de `window_sequence`:

```

para (i = 0, i2 = 0; i < N/2; i += 1, i2 += 2) {
spec[0][i] = spec[i2]; /* aislar 1ª ventana */
buffer[i] = spec[i2+1]; /* aislar 2ª ventana */
}
para (i = 0; i < N/2; i += 1) {
spec[1][i] = buffer[i]; /* copiar 2ª ventana */
}

```

[0085] La IMDCT para el espectro de TS de media longitud se define entonces como:

$$x_{(0,1),n} = \frac{2}{N} \sum_{k=0}^{\frac{N-1}{4}} \text{spec}[(0,1)][k] \cos\left(\frac{4\pi}{N} (n+n_0) \left(k + \frac{1}{2}\right)\right) \quad \text{para } 0 \leq n < \frac{N}{2}$$

[0086] Las etapas subsiguientes de formación de ventanas y de conmutación de bloques se definen en las siguientes subsecciones.

10 [0087] La división de transformada con STOP_START_SEQUENCE sería similar a la siguiente descripción:

[0088] Una STOP_START_SEQUENCE en combinación con división de transformada se representó en la Figura 2. Comprende dos ventanas de media longitud superpuestas y añadidas 56, 58 con una longitud de N/2 que es de 1024 (960, 768). N_s se establece en 256 (240, 192) respectivamente.

15 [0089] Las ventanas (0,1) para las dos IMDCT de media longitud se dan del siguiente modo:

$$W_{(0,1)}(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_{I/2} - N_s}{4} \\ W_{(0,1),IZQUIERDA,N_s}(n - \frac{N_{I/2} - N_s}{4}), & \text{para } \frac{N_{I/2} - N_s}{4} \leq n < \frac{N_{I/2} + N_s}{4} \\ 1,0, & \text{para } \frac{N_{I/2} + N_s}{4} \leq n < \frac{3N_{I/2} - N_s}{4} \\ W_{(0,1),DERECHA,N_s}(n + \frac{N_s}{2} - \frac{3N_{I/2} - N_s}{4}), & \text{para } \frac{3N_{I/2} - N_s}{4} \leq n < \frac{3N_{I/2} + N_s}{4} \\ 0,0, & \text{para } \frac{3N_{I/2} + N_s}{4} \leq n < N_{I/2} \end{cases}$$

20 donde para la primera IMDCT se aplican las ventanas

$$W_{0,IZQUIERDA,N_s}(n) = \begin{cases} W_{KBD,IZQUIERDA,N_s}(n), & \text{si } \text{window_shape_previous_block} == 1 \\ W_{SIN,IZQUIERDA,N_s}(n), & \text{si } \text{window_shape_previous_block} == 0 \end{cases}$$

$$W_{0,DERECHA,N_s}(n) = \begin{cases} W_{KBD,DERECHA,N_s}(n), & \text{si } \text{window_shape} == 1 \\ W_{SIN,DERECHA,N_s}(n), & \text{si } \text{window_shape} == 0 \end{cases}$$

y para la segunda IMDCT se aplican las ventanas

25

$$W_{1,IZQUIERDA,N_s}(n) = \begin{cases} W_{KBD,IZQUIERDA,N_s}(n), & \text{si } \text{window_shape} == 1 \\ W_{SIN,IZQUIERDA,N_s}(n), & \text{si } \text{window_shape} == 0 \end{cases}$$

$$W_{1,DERECHA,N_s}(n) = \begin{cases} W_{KBD,DERECHA,N_s}(n), & \text{si } \text{window_shape} == 1 \\ W_{SIN,DERECHA,N_s}(n), & \text{si } \text{window_shape} == 0 \end{cases}$$

[0090] El solapamiento y la adición entre las dos ventanas de media longitud que resultan en los valores en el dominio del tiempo puestos en ventanas z_{i,n} se describe del siguiente modo. En este caso, N_I se establece en 2048

(1920, 1536), N_s a 256 (240, 192) respectivamente:

$$Z_{L(n)} = \begin{cases} 0,0, & \text{para } 0 \leq n < N_s \\ x_{2,N-N_s} \cdot W_0(n - N_s), & \text{para } N_s \leq n < \frac{2N_l - N_s}{4} \\ x_{2,N-N_s} \cdot W_0(n - N_s) + x_{1,N-(N_l/2-N_s)} \cdot W_1(n - (N_l/2 - N_s)), & \text{para } \frac{N_l + N_s}{4} \leq n < \frac{2N_l + N_s}{4} \\ x_{1,N-(N_l/2-N_s)} \cdot W_1(n - (N_l/2 - N_s)), & \text{para } \frac{2N_l + N_s}{4} \leq n < N \end{cases}$$

5 [0091] La división de transformada con LONG_START_SEQUENCE sería similar a la siguiente descripción:

[0092] La LONG_START_SEQUENCE en combinación con la división de transformada se ilustra en la Figura 4. Comprende tres ventanas que se definen del siguiente modo, donde N_l se establece en 1024 (960, 768), N_s se establece en 256 (240, 192) respectivamente.

10

$$W_0(n) = \begin{cases} 1,0, & \text{para } 0 \leq n < \frac{3N_l/2 - N_s}{4} \\ W_{0,DERECHA,N_s}(n + \frac{N_s}{2} - \frac{3N_l/2 - N_s}{4}), & \text{para } \frac{3N_l/2 - N_s}{4} \leq n < \frac{3N_l/2 + N_s}{4} \\ 0,0, & \text{para } \frac{3N_l/2 + N_s}{4} \leq n < N_l/2 \end{cases}$$

$$W_1(n) = \begin{cases} 0,0, & \text{para } 0 \leq n < \frac{N_l/2 - N_s}{4} \\ W_{1,IZQUIERDA,N_s}(n - \frac{N_l/2 - N_s}{4}), & \text{para } \frac{N_l/2 - N_s}{4} \leq n < \frac{N_l/2 + N_s}{4} \\ 1,0, & \text{para } \frac{N_l/2 + N_s}{4} \leq n < \frac{3N_l/2 - N_s}{4} \\ W_{1,DERECHA,N_s}(n + \frac{N_s}{2} - \frac{3N_l/2 - N_s}{4}), & \text{para } \frac{3N_l/2 - N_s}{4} \leq n < \frac{3N_l/2 + N_s}{4} \\ 0,0, & \text{para } \frac{3N_l/2 + N_s}{4} \leq n < N_l/2 \end{cases}$$

[0093] Las mitades de ventana izquierda/derecha vienen dadas por:

$$W_{1,IZQUIERDA,N_s}(n) = \begin{cases} W_{KBD,IZQUIERDA,N_s}(n), & \text{si } window_shape == 1 \\ W_{SIN,IZQUIERDA,N_s}(n), & \text{si } window_shape == 0 \end{cases}$$

$$W_{(0,1),DERECHA,N_s}(n) = \begin{cases} W_{KBD,DERECHA,N_s}(n), & \text{si } window_shape == 1 \\ W_{SIN,DERECHA,N_s}(n), & \text{si } window_shape == 0 \end{cases}$$

15

[0094] La tercera ventana es igual a la mitad izquierda de una LONG_START_WINDOW:

$$W_2(n) = \begin{cases} W_{IZQUIERDA,N_l}(n), & \text{para } 0 \leq n < N_l/2 \\ 1,0, & \text{para } N_l/2 \leq n < N_l \end{cases}$$

20

con

$$W_{IZQUIERDA,N_1}(n) = \begin{cases} W_{RED_IZQUIERDA,N_1}(n), & \text{si } window_shape_previous_block == 1 \\ W_{SIN_IZQUIERDA,N_1}(n), & \text{si } window_shape_previous_block == 0 \end{cases}$$

[0095] El solapamiento y adición entre las dos ventanas de media longitud que resultan en valores intermedios

5 en el dominio del tiempo puestos en ventanas $\tilde{Z}_{i,n}$ se describen del siguiente modo. En este caso, N_1 se establece en 2048 (1920, 1536), N_s se establece en 256 (240, 192) respectivamente.

$$\tilde{Z}_{i,n}(n) = \begin{cases} -i \cdot x_{0,2N_s-n-1} \cdot W_0(2N_s - n - 1), & \text{para } 0 \leq n < N_s \\ x_{0,n-N_s} \cdot W_2(n - N_s), & \text{para } N_s \leq n < \frac{2N_1-N_s}{4} \\ x_{0,n-N_s} \cdot W_0(n - N_s) + x_{1,n-\left(\frac{N_1}{2}-N_s\right)} \cdot W_1\left(n - \left(\frac{N_1}{2} - N_s\right)\right), & \text{para } \frac{N_1+N_s}{4} \leq n < \frac{2N_1+N_s}{4} \\ x_{1,n-\left(\frac{N_1}{2}-N_s\right)} \cdot W_1\left(n - \left(\frac{N_1}{2} - N_s\right)\right), & \text{para } \frac{2N_1+N_s}{4} \leq n < N \end{cases}$$

10 **[0096]** Los valores en el dominio del tiempo puestos en ventanas finales $Z_{i,n}$ se obtienen aplicando W_2 :

$$Z_{i,n}(n) = \tilde{Z}_{i,n}(n) \cdot W_2(n), \quad \text{para } 0 \leq n < N_1$$

15 **[0097]** Independientemente de si se utiliza la señalización semirretrocompatible explícita o implícita, ambas descritas anteriormente, puede ser necesaria alguna modificación en la decodificación de estéreo de predicción compleja de xHE-AAC a efectos de lograr una operación significativa de los espectros entrelazados.

[0098] La modificación efectuada en la decodificación de estéreo de predicción compleja podría implementarse del siguiente modo.

20

[0099] Dado que las herramientas de estéreo de FD operan sobre un pseudoespectro entrelazado cuando TS está activa en un par de canales, no son necesarios cambios en el procesamiento subyacente de predicción de M/S o compleja. Sin embargo, la derivación de la mezcla descendente de la trama precedente $dmx_re_prev[]$ y el cálculo de MDST de mezcla descendente $dmx_im[]$ en ISO/IEC 23003-3:2012 sección 7.7.2 se debe adoptar si se utiliza TS en
25 cualquier canal en la trama actual o última:

- use_prev_frame debe ser 0 si la actividad de TS ha cambiado en cualquiera de los canales desde la última trama hasta la actual. En otras palabras, no debe utilizarse $dmx_re_prev[]$ en este caso, debido a la conmutación de longitud de transformada.
- 30 • Si la TS estaba o está activa, $dmx_re_prev[]$ y $dmx_re[]$ especifican pseudoespectros entrelazados y deben ser desentrelazados en sus correspondientes espectros de TS para el cálculo correcto de MDST.
- Al haber actividad de TS, se calculan dos mezclas descendentes de MDST de media longitud por lo cual se utilizan coeficientes de filtro adaptados (Tablas 1 y 2) y se entrelazan en un espectro de longitud completa $dmx_im[]$ (de la misma manera que $dmx_re[]$).
- 35 • $window_sequence$: se calculan estimaciones de MDST de mezcla descendente para cada par de ventanas del grupo. use_prev_frame se evalúa solo para el primero de los dos pares de media ventana. Para el par de ventanas restante, siempre se utiliza el par de ventanas precedente en la estimación de MDST, lo que implica $use_prev_frame = 1$.
- Formas de ventanas: los parámetros de estimación de MDST para la ventana actual, que son coeficientes de
40 filtro descritos más abajo, dependen de las formas de las mitades de ventana izquierda y derecha. Para la 1ª ventana, esto significa que los parámetros de filtro son una función de las banderas de $window_shape$ actuales y anteriores. La ventana restante solo se ve afectada por la $window_shape$ actual.

45

Tabla 1 - Parámetros de filtro de MDST para Ventana Actual (filter_coefs)

Secuencia de Ventana Actual	Mitad Izquierda: Forma Sinusoidal Mitad Derecha: Forma Sinusoidal	Mitad Izquierda: Forma de KBD Mitad Derecha: Forma de KBD
LONG_START_SEQUENCE STOP_START_SEQUENCE	[0,185618f, -0,000000f, 0,627371f, 0,000000f, -0,627371f, 0,000000f, -0,185618f]	[0,203599f, -0,000000f, 0,633701f, 0,000000f, -0,633701f, 0,000000f, -0,203599f]
Secuencia de Ventana Actual	Mitad Izquierda: Forma Sinusoidal Mitad Derecha: Forma de KBD	Mitad Izquierda: Forma de KBD Mitad Derecha: Forma Sinusoidal
LONG_START_SEQUENCE STOP_START_SEQUENCE	[0,194609f, 0,006202f, 0,630536f, 0,000000f, -0,630536f, -0,006202f, -0,194609f]	[0,194609f, -0,006202f, 0,630536f, 0,000000f, -0,630536f, 0,006202f, -0,194609f]

Tabla 2 - Parámetros de filtro de MDST para Ventana Previa (filter_coefs_prev)

Secuencia de Ventana Actual	Mitad izquierda de ventana actual: Forma Sinusoidal	Mitad izquierda de ventana actual: Forma de KBD
LONG_START_SEQUENCE STOP_START_SEQUENCE	[0,038498, 0,039212, 0,039645, 0,039790, 0,039645, 0,039212, 0,038498]	[0,038498, 0,039212, 0,039645, 0,039790, 0,039645, 0,039212, 0,038498]

5

[0100] Finalmente, la Figura 5 muestra, en aras de integridad, un posible codificador de audio que opera en el dominio de la frecuencia y que soporta la conmutación de longitud de transformada que encaja en las realizaciones señaladas a grandes rasgos anteriormente. Es decir, el codificador de la Figura 5, indicado en términos generales mediante el número de referencia 100, es capaz de codificar una señal de audio 102 situada en la corriente de datos 20 de una manera tal que el decodificador de la Figura 1 y sus correspondientes variantes descritas anteriormente también sean capaces de aprovechar el modelo de la división de transformada para algunas de las tramas, mientras que los decodificadores "anticuados" siguen siendo capaces de procesar tramas de la TS sin analizar errores o similares.

15 [0101] El codificador 100 de la Figura 5 comprende un transformador 104, un escalador inverso 106, un dispositivo de inserción de coeficientes en el dominio de la frecuencia 108 y un dispositivo de inserción de factores de escala 110. El transformador 104 recibe la señal de audio 102 que se va a codificar y está configurado para someter porciones en el dominio del tiempo de la señal de audio a transformación para obtener coeficientes en el dominio de la frecuencia para tramas de señal de audio. En particular, y como se puso en claro en la exposición anterior, el transformador 104 decide sobre una base de trama por trama en cuanto a la subdivisión de estas tramas 26 en transformadas - o ventanas de transformada - que se utiliza. Tal como se ha descrito anteriormente, las tramas 26 pueden ser de igual longitud y la transformada puede ser una transformada solapada para lo que se utilizan transformadas solapadas de diferentes longitudes. La Figura 5 ilustra por ejemplo que una trama 26a está sometida a una transformada larga, una trama 26b está sometida a división de transformada, es decir a dos transformadas de media longitud, y otra trama 26c se muestra sometida a más de dos, es decir a $2^n > 2$, transformadas aún más cortas de 2^{-n} la longitud de transformada larga. Como se ha expuesto anteriormente, mediante esta disposición, el codificador 100 es capaz de adaptar la resolución de espectro temporal representada por la transformada solapada llevada a cabo por el transformador 104 al contenido de audio, o tipo de contenido de audio, de variación de tiempo, de la señal de audio 102.

30

[0102] Es decir, los coeficientes en el dominio de la frecuencia resultan en la salida del transformador 104 que representa un espectrograma de la señal de audio 102. El escalador inverso 106 está conectado a la salida del transformador 104 y está configurado para escalar de manera inversa, y al mismo tiempo cuantificar, los coeficientes en el dominio de la frecuencia según factores de escala. Es interesante observar que el escalador inverso opera sobre los coeficientes de frecuencia a medida que éstos son obtenidos por el transformador 104. Es decir, el escalador inverso 106 debe ser necesariamente consciente de la asignación de longitud de transformada o del modo de asignación de transformada a las tramas 26. Cabe observar también que el escalador inverso 106 necesita determinar los factores de escala. Para tal fin, el escalador inverso 106, por ejemplo, es parte de un bucle de retroalimentación que evalúa un umbral de enmascaramiento psicoacústico determinado por la señal de audio 102 para mantener el ruido de cuantificación introducido por la cuantificación y gradualmente establecido según los factores de escala, por debajo del umbral de detección psicoacústico en la medida de lo posible, con o sin obedecer alguna restricción en cuanto a la tasa de bits.

35

40

[0103] A la salida del escalador inverso 106, los factores de escala y los coeficientes inversamente escalados y cuantificados en el dominio de la frecuencia son emitidos y el dispositivo de inserción de factores de escala 110 está configurado para insertar los factores de escala en la corriente de datos 20, mientras que el dispositivo de inserción de coeficientes en el dominio de la frecuencia 108 está configurado para insertar los coeficientes en el dominio de la frecuencia de las tramas de la señal de audio, inversamente escalados y cuantificados según los factores de escala, en la corriente de datos 20. De una manera correspondiente al decodificador, ambos dispositivos de inserción 108 y 110 operan independientemente del modo de transformada asociado con las tramas 26 en lo que se refiere a la yuxtaposición de tramas 26a del modo de transformada larga y de las tramas 26b del modo de división de transformada.

[0104] En otras palabras, los dispositivos de inserción 110 y 108 operan independientemente de la señalización 34 anteriormente mencionada con la que el transformador 104 está configurado para señalar, o insertar en, la corriente de datos 20 para las tramas 26a y 26b, respectivamente.

[0105] En otras palabras, en la realización anteriormente descrita, es el transformador 104 que de manera adecuada dispone los coeficientes de transformada de las tramas de transformada larga y de transformada dividida, particularmente mediante disposición seriada plana o entrelazada, y el dispositivo de inserción funciona realmente de manera independiente con respecto a 109. Pero en un sentido más general basta con que la independencia del dispositivo de inserción de coeficientes en el dominio de la frecuencia tenga restringida la señalización de la inserción de una secuencia de los coeficientes en el dominio de la frecuencia de cada una de las tramas de transformada larga y de transformada dividida de la señal de audio, inversamente escalados según factores de escala, en la corriente de datos de tal manera que según la señalización, la secuencia de coeficientes en el dominio de la frecuencia esté formada mediante la disposición secuencial de los coeficientes en el dominio de la frecuencia de la transformada única de una trama respectiva de una manera no entrelazada en el caso de que la trama sea una trama de transformada larga, y mediante el entrelazado de los coeficientes en el dominio de la frecuencia de más de una transformada de la trama respectiva en caso de que la respectiva trama sea una trama de transformada dividida.

[0106] En lo que se refiere al dispositivo de inserción de coeficientes en el dominio de la frecuencia 108, el hecho de que el mismo opera independientemente de la señalización 34 que distingue entre las tramas 26a por una parte y las tramas 26b por otra parte, significa que el dispositivo de inserción 108 inserta los coeficientes en el dominio de la frecuencia de las tramas de la señal de audio, escalados inversamente según los factores de escala, en la corriente de datos 20 de una manera secuencial en el caso de una transformada llevada a cabo por la trama respectiva, de una manera no entrelazada, e inserta los coeficientes en el dominio de la frecuencia de las tramas respectivas por las cuales se utiliza el entrelazado en caso de haber más de una transformada llevada a cabo por la trama respectiva, particularmente dos en el ejemplo de la Figura 5. Sin embargo, como ya se ha expuesto anteriormente, el modo de división de transformada se puede implementar también de manera diferente tal como mediante la división de una transformada en más de dos transformadas.

[0107] Finalmente, cabe observar que el codificador de la Figura 5 también se puede adaptar para llevar a cabo la totalidad de las medidas de codificación adicionales señaladas a grandes rasgos anteriormente con respecto a la Figura 2 tales como la codificación de MS, la predicción estéreo compleja 42 y la TNS, con, para tal fin, la determinación de sus respectivos parámetros 44, 48 y 64.

[0108] Aunque algunos aspectos se han descrito en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del procedimiento correspondiente, donde un bloque o dispositivo corresponde a una etapa de procedimiento o rasgo de una etapa de procedimiento. De manera análoga, los aspectos descritos en el contexto de una etapa de procedimiento también representan una descripción de un bloque, artículo o rasgo, correspondientes de un aparato correspondiente. Algunas de las etapas del procedimiento, o todas ellas, pueden ser ejecutadas por (o utilizando) un aparato de hardware, similar a, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, una o más de las etapas del procedimiento más importantes se pueden ejecutar mediante un aparato de este tipo.

[0109] Según determinados requisitos para la implementación, las realizaciones de la invención se pueden implementar en hardware o en software. La implementación puede llevarse 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, que tienen señales de control electrónicamente legibles almacenadas en ellos, que cooperan (o que son capaces de cooperar) con un sistema de ordenador programable de tal manera que se lleve a cabo el respectivo procedimiento. Por tanto, el medio de almacenamiento digital puede ser legible por ordenador.

[0110] Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema de ordenador programable de tal manera que se lleve a cabo uno de los procedimientos descritos en esta invención.

[0111] Por lo general, las realizaciones de la presente invención pueden implementarse en forma de un producto de programa informático con un código de programación, siendo el código de programación operativo para llevar a cabo uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código de programación se puede almacenar por ejemplo en un soporte legible por máquina.

5

[0112] Otras realizaciones comprenden el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por máquina.

[0113] En otras palabras, una realización del procedimiento inventivo es, por tanto, un programa informático que tiene un código de programación para llevar a cabo uno de los procedimientos descritos en esta invención, cuando se ejecuta el programa informático en un ordenador.

[0114] Por tanto, una realización adicional de los procedimientos inventivos consiste en un soporte de datos (o en un medio de almacenamiento digital o un medio legible por ordenador) que comprende, grabado en él, el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención. El soporte de datos, el medio de almacenamiento digital o el medio grabado típicamente son tangibles y/o no transitorios.

[0115] Por tanto, una realización adicional del procedimiento inventivo es una corriente de datos o una secuencia de señales que representa el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención. La corriente de datos o la secuencia de señales por ejemplo pueden estar configuradas para ser transferidas por medio de una conexión de comunicación de datos, por ejemplo, por medio de Internet.

[0116] Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, un dispositivo lógico programable configurado o adaptado para llevar a cabo uno de los procedimientos descritos en esta invención.

[0117] Una realización adicional comprende un ordenador que tiene instalado en él el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención.

[0118] Una realización adicional según la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente o de forma óptica) un programa informático para llevar a cabo uno de los procedimientos descritos en esta invención a un receptor. El receptor puede ser, por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o sistema puede comprender, por ejemplo, un servidor de archivos para transferir el programa informático al receptor.

35

[0119] En algunas realizaciones, es posible utilizar un dispositivo lógico programable (por ejemplo, una disposición de puerta programable de campo) para llevar a cabo algunas de las funcionalidades, o todas ellas, de los procedimientos descritos en esta invención. En algunas realizaciones, la disposición de puerta programable de campo puede cooperar con un microprocesador para llevar a cabo uno de los procedimientos descritos en esta invención. En términos generales, es preferible que los procedimientos sean llevados a cabo mediante un aparato de hardware.

40

Referencias

[0120]

45

[1] Internet Engineering Task Force (IETF), RFC 6716, "Definition of the Opus Audio Codec", Proposed Standard, septiembre de 2012. Disponible online en <http://tools.ietf.org/html/rfc6716>.

[2] International Organization for Standardization, ISO/IEC 14496-3:2009, "Information Technology – Coding of audio-visual objects – Part 3: Audio", Ginebra, Suiza, agosto de 2009.

50 [3] M. Neuendorf y col., "MPEG Unified Speech and Audio Coding – The ISO/MPEG Standard for High-Efficiency Audio Coding of All Content Types", in Proc. 132nd Convention of the AES, Budapest, Hungría, abril de 2012. También aparecerá en el Journal of the AES, 2013.

[4] International Organization for Standardization, ISO/IEC 23003-3:2012, "Information Technology – MPEG audio – Part 3: Unified speech and audio coding", Ginebra, enero de 2012.

55 [5] J.D.Johnston and A.J.Ferreira, "Sum-Difference Stereo Transform Coding", in Proc. IEEE ICASSP-92, Vol. 2, marzo de 1992.

[6] N.Rettelbach, y col., European Patent EP2304719A1, "Audio Encoder, Audio Decoder, Methods for Encoding and Decoding an Audio Signal, Audio Stream and Computer Program", abril de 2011.

REIVINDICACIONES

1. Decodificador de audio en el dominio de la frecuencia que soporta la conmutación de longitud de transformada, que comprende
- 5 un extractor de coeficiente en el dominio de la frecuencia (12) configurado para extraer los coeficientes en el dominio de la frecuencia (24) de tramas de una señal de audio de una corriente de datos;
- un extractor de factor de escala (14) configurado para extraer los factores de escala de la corriente de datos;
- 10 un transformador inverso (16) configurado para someter los coeficientes en el dominio de la frecuencia de las tramas, escalados según los factores de escala, a una transformación inversa para obtener las porciones en el dominio del tiempo de la señal de audio;
- un combinador (18) configurado para combinar las porciones en el dominio del tiempo para obtener la señal de audio,
- 15 donde el transformador inverso responde a una señalización dentro de las tramas de la señal de audio para, dependiendo de la señalización,
- formar una única transformada disponiendo de manera secuencial los coeficientes en el dominio de la frecuencia de una trama respectiva, escalados según los factores de escala, de una manera no desentrelazada y someter la una transformada a una transformación inversa de una primera longitud de transformada, o
- 20 formar más de una transformada mediante el desentrelazado de los coeficientes en el dominio de la frecuencia de la trama respectiva, escalados según los factores de escala, y someter cada una de más de una transformada a una transformación inversa de una segunda longitud de transformada, más corta que la primera longitud de transformada,
- 25 caracterizado porque el extractor de coeficiente en el dominio de la frecuencia y el extractor de factor de escala operan independiente de la señalización,
- y el transformador inverso está configurado para
- 30 realizar un filtrado inverso de conformación de ruido temporal (62) en una secuencia de N coeficientes, independientemente de la señalización, mediante la aplicación de un filtro de una función de transferencia de la que se establece según los coeficientes de TNS (64) sobre la secuencia de N coeficientes, con
- en la formación de una transformada, aplicar el filtrado inverso de conformación de ruido temporal utilizando los coeficientes en el dominio de la frecuencia secuencialmente dispuestos de una manera no desentrelazada como la secuencia de N coeficientes, y
- 35 en la formación de más de una transformada, aplicar el filtrado inverso de conformación de ruido temporal en los coeficientes en el dominio de la frecuencia utilizando los coeficientes en el dominio de la frecuencia secuencialmente dispuestos de una manera según la cual más de una transformada se concatenan espectralmente como la secuencia de N coeficientes,
- 40 donde los coeficientes del dominio de la frecuencia (24) se agrupan en un número de bandas de factor de escala que es independiente de la señalización, y el extractor de factor de escala (14) está configurado para extraer para cada banda de factor de escala de ese tipo (30) un factor de escala (32).
- 45 2. Decodificador de audio en el dominio de la frecuencia según la reivindicación 1, donde el transformador inverso está configurado para someter los coeficientes en el dominio de la frecuencia a relleno de ruido, con los coeficientes en el dominio de la frecuencia secuencialmente dispuestos de una manera no desentrelazada, y en una resolución espectral independiente de la señalización.
- 50 3. Decodificador de audio en el dominio de la frecuencia según la reivindicación 1, donde el transformador inverso está configurado para soportar la codificación conjunta-estéreo con o sin predicción estéreo entre canales y para utilizar los coeficientes en el dominio de la frecuencia como un espectro de suma (medio) o de diferencia (lateral) o residual de predicción de la predicción estéreo entre canales, con los coeficientes en el dominio de la frecuencia dispuestos de una manera no desentrelazada, independientemente de la señalización.
- 55 4. Decodificador de audio en el dominio de la frecuencia según la reivindicación 1, donde e número de las más de una transformadas es igual a 2, y la primera longitud de transformada es dos veces la segunda longitud de transformada.
- 60 5. Procedimiento para la decodificación de audio en el dominio de la frecuencia que soporta la conmutación de longitud de transformada, que comprende
- extraer los coeficientes en el dominio de la frecuencia de tramas de una señal de audio de una corriente de datos;
- 65 extraer los factores de escala de la corriente de datos;
- someter los coeficientes en el dominio de la frecuencia de las tramas, escalados según los factores de escala, a

una transformación inversa para obtener las porciones en el dominio del tiempo de la señal de audio;
combinar las porciones en el dominio del tiempo para obtener la señal de audio,
donde el sometimiento a la transformación inversa responde a una señalización dentro de las tramas de la señal
de audio de modo que, dependiendo de la señalización, comprende

5

formar una única transformada al disponer secuencialmente los coeficientes en el dominio de la frecuencia de una trama respectiva de una manera no desentrelazada y someter la una transformada a una transformación inversa de una primera longitud de transformada, o

10

formar más de una transformada al desentrelazar los coeficientes en el dominio de la frecuencia de la trama respectiva y someter cada una de las más de una transformada a una transformación inversa de una segunda longitud de transformada, más corta que la primera longitud de transformada,

15

caracterizado porque la extracción de los coeficientes en el dominio de la frecuencia y la extracción de los factores de escala son independientes de la señalización,

15

y el sometimiento a la transformación inversa comprende

realizar un filtrado inverso de conformación de ruido temporal (62) en una secuencia de N coeficientes, independientemente de la señalización, mediante la aplicación de un filtro de una función de transferencia de la cual se establece según los coeficientes de TNS (64) sobre la secuencia de N coeficientes, con

20

en la formación de la una transformada, aplicar el filtrado inverso de conformación de ruido temporal utilizando los coeficientes en el dominio de la frecuencia secuencialmente dispuestos de una manera no desentrelazada como la secuencia de N coeficientes, y

25

en la formación de más de una transformada, aplicar el filtrado inverso de conformación de ruido temporal a los coeficientes en el dominio de la frecuencia utilizando los coeficientes en el dominio de la frecuencia secuencialmente dispuestos de una manera según la cual las más de una transformada se concatenan espectralmente como la secuencia de N coeficientes,

30

donde los coeficientes en el dominio de la frecuencia (24) están agrupados en un número de bandas de factor de escala que es independiente de la señalización, y un factor de escala (32) se extrae para cada banda de factor de escala de ese tipo (30).

6. Programa informático que comprende instrucciones que, cuando el programa es ejecutado por un ordenador, dan lugar a que el ordenador lleve a cabo el procedimiento según la reivindicación 5.

35

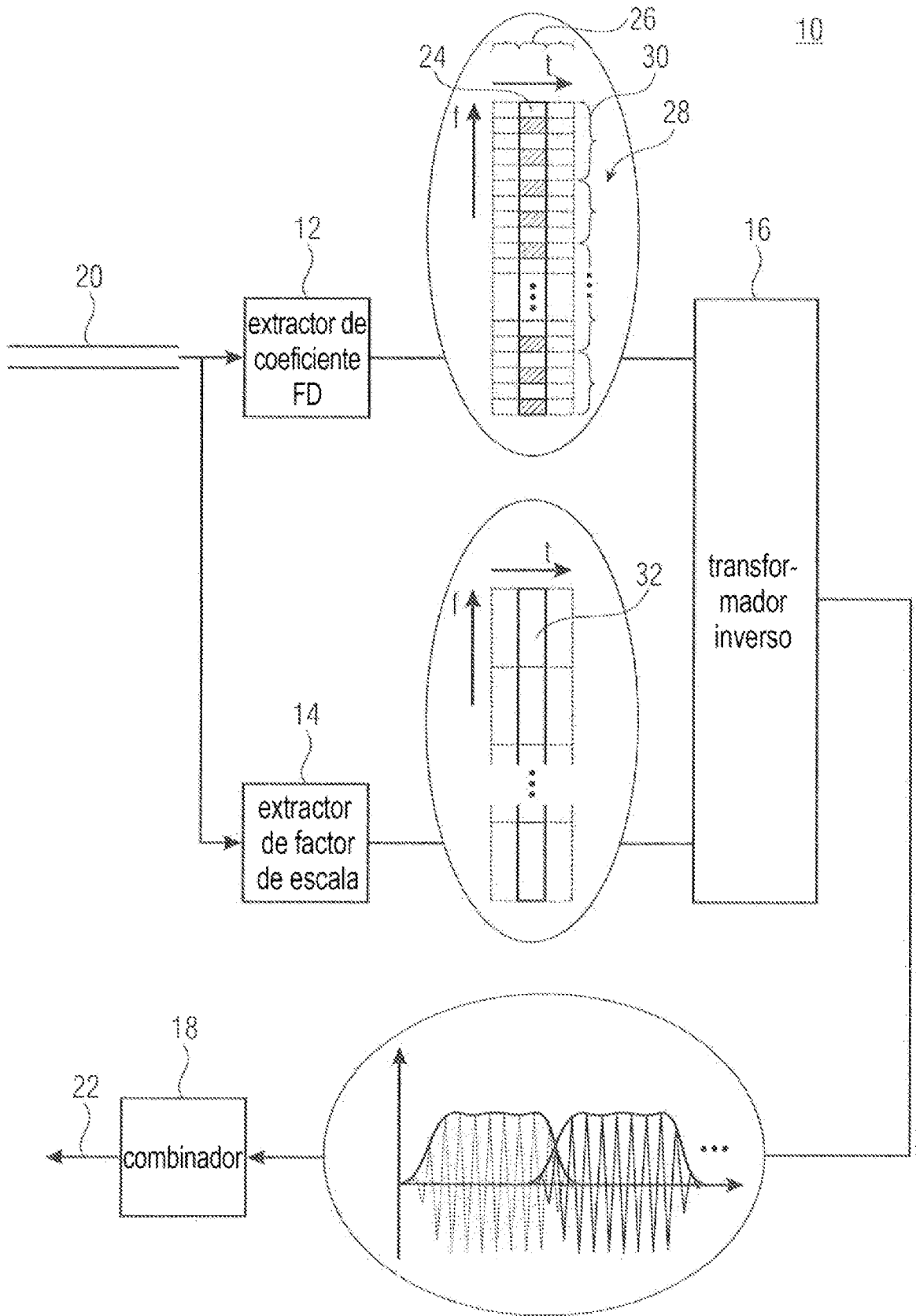


FIG 1

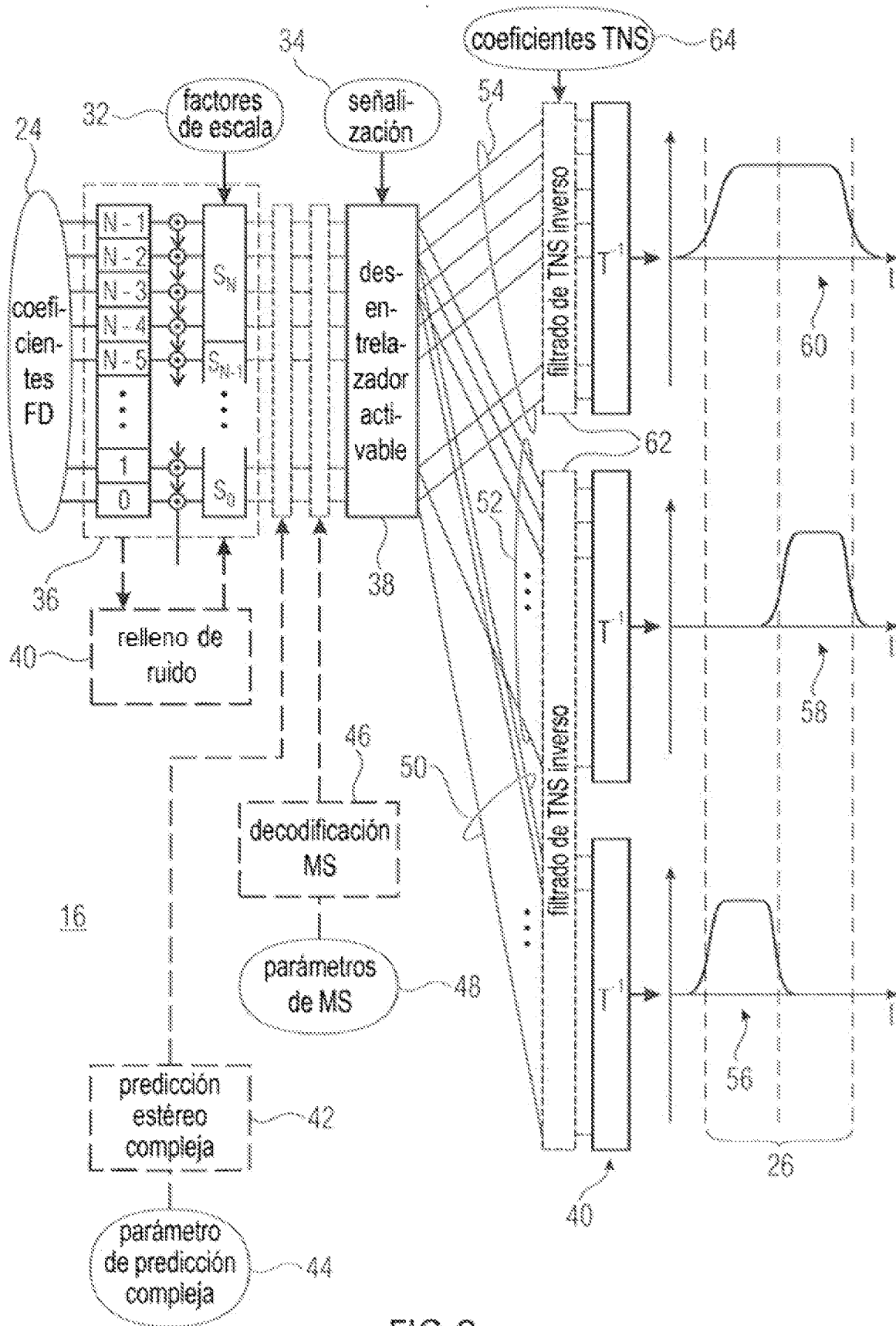


FIG 2

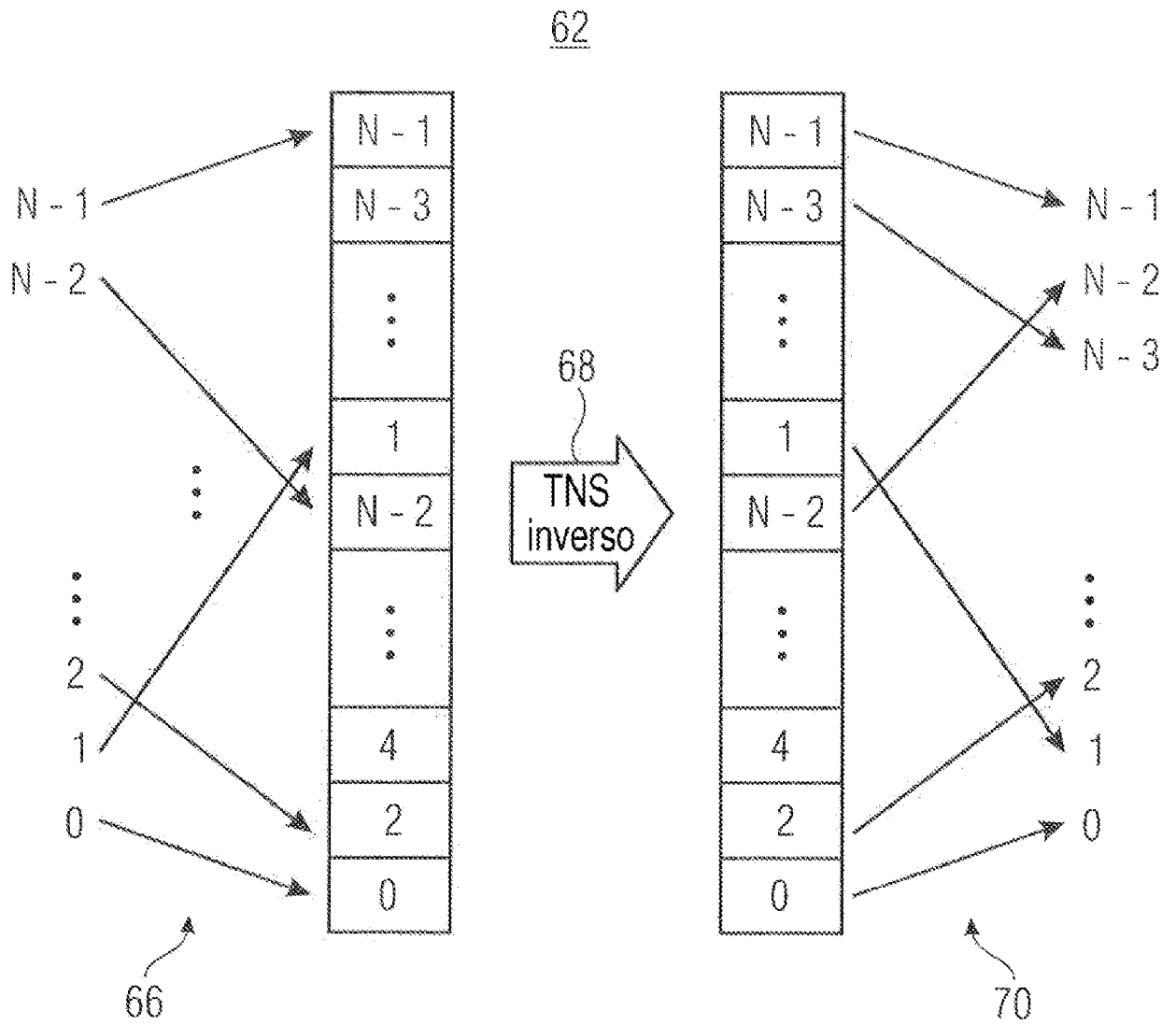


FIG 3

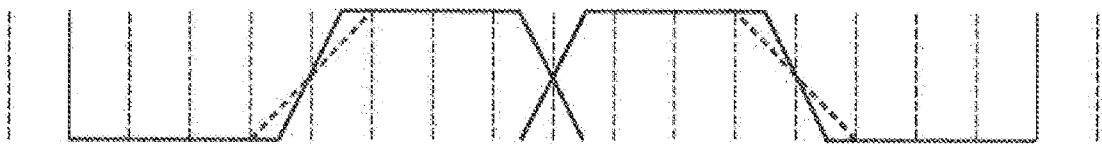


FIG 4

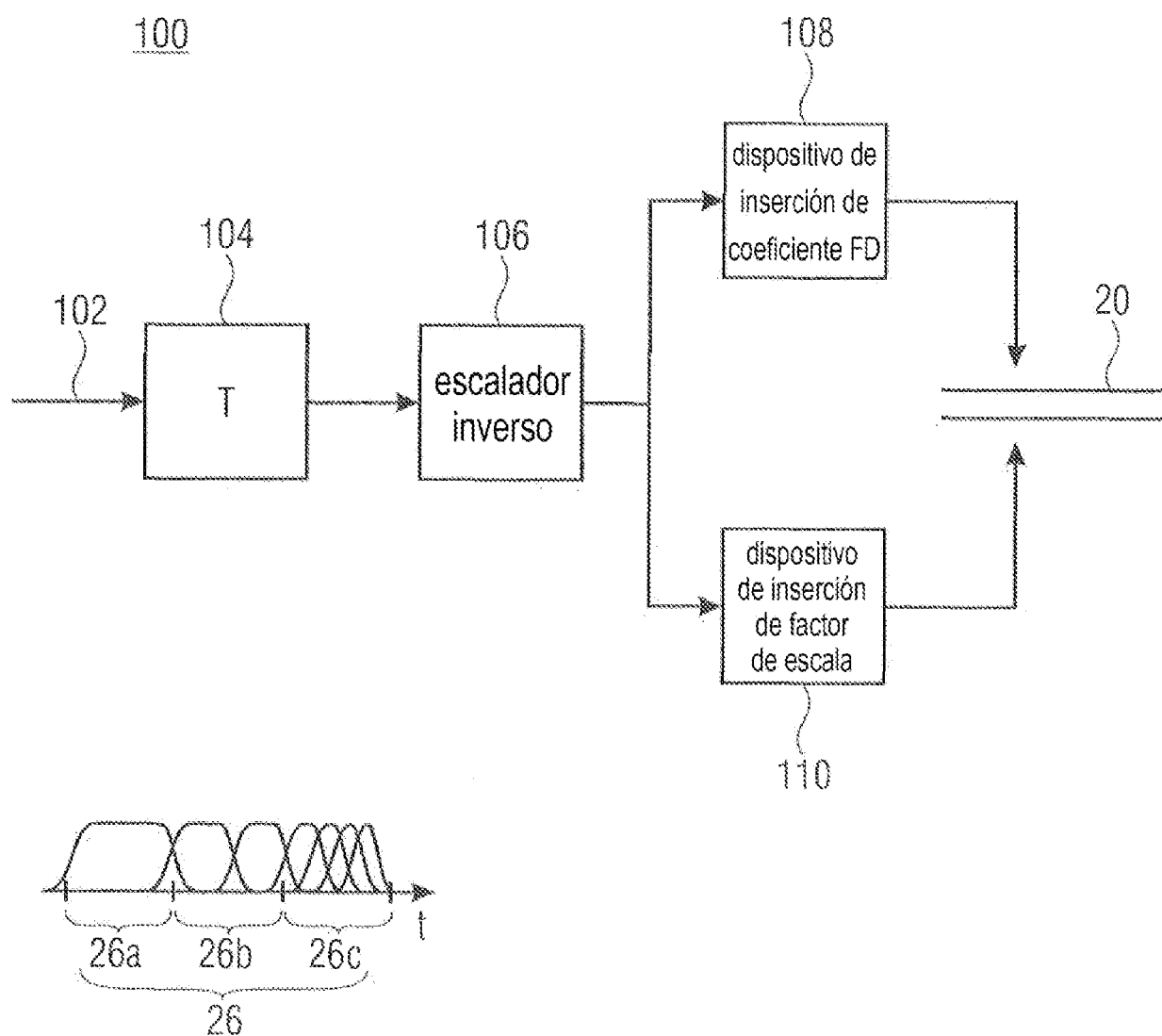


FIG 5