

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 901 758**

51 Int. Cl.:

G10L 19/18 (2013.01)
G10L 19/02 (2013.01)
G10L 19/028 (2013.01)
G10L 19/04 (2013.01)
G10L 19/24 (2013.01)
G10L 21/038 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **24.07.2015** E 19165957 (2)

97 Fecha y número de publicación de la concesión europea: **20.10.2021** EP 3522154

54 Título: **Codificación y decodificación de audio usando un procesador de dominio de frecuencia, un procesador de dominio de tiempo y un procesador cruzado para inicialización del procesador de dominio de tiempo**

30 Prioridad:

28.07.2014 EP 14178819

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.03.2022

73 Titular/es:

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

72 Inventor/es:

**DISCH, SASCHA;
DIETZ, MARTIN;
MULTRUS, MARKUS;
FUCHS, GUILLAUME;
RAVELLI, EMMANUEL;
NEUSINGER, MATTHIAS;
SCHNELL, MARKUS;
SCHUBERT, BENJAMIN y
GRILL, BERNHARD**

74 Agente/Representante:

PONTI & PARTNERS, S.L.P.

ES 2 901 758 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 y decodificación de audio usando un procesador de dominio de frecuencia, un procesador de dominio de tiempo y un procesador cruzado para inicialización del procesador de dominio de tiempo

5

[0001] La presente invención se refiere a la codificación y decodificación de señales de audio, y en particular al procesamiento de señales de audio mediante el uso de procesadores de codificador/decodificador en el dominio de las frecuencias y en el dominio de los tiempos en paralelo.

10 **[0002]** La codificación perceptual de las señales de audio con el fin de la reducción de los datos para un almacenamiento o transmisión eficaz de estas señales es una práctica ampliamente utilizada. En particular, cuando deba lograrse una velocidad de bits lo más baja posible, la codificación utilizada conduce a una reducción de la calidad del audio que es principalmente causada por la limitación, en el lado del codificador, del ancho de banda de la señal de audio que se va a transmitir. En este caso, típicamente la señal de audio se filtra en paso bajo de tal
15 manera que no permanezca un contenido de forma de onda espectral por encima de una cierta frecuencia de corte predeterminada.

[0003] En los códecs contemporáneos existen procedimientos bien conocidos para la restauración de la señal en el lado del decodificador mediante la Extensión del Ancho de Banda de la señal de audio (BWE, Bandwidth Extension), por ejemplo, la Replicación de Banda Espectral (SBR, Spectral Band Replication) que opera en el dominio de las frecuencias o la denominada Extensión del Ancho Banda en el Dominio del Tiempo (TD-BWE, Time Domain Bandwidth Extension) que es un post-procesador en los códecs de habla que opera en el dominio de los tiempos.

25 **[0004]** Adicionalmente existen diversos conceptos de codificación en el dominio de los tiempos/dominio de las frecuencias, combinados, tales como los conceptos conocidos bajo la designación AMR-WB+ o USAC.

[0005] Todos estos conceptos de codificación en el dominio de los tiempos combinados tienen en común que el dominio de las frecuencias se basa en tecnologías de extensión de ancho de banda que incurren en una limitación
30 de banda en la señal de audio de entrada y la porción por encima de una frecuencia de cruce, o frecuencia de límite, se codifica con un concepto de codificación de baja resolución y se sintetiza en el lado del decodificador. Por ello, tales conceptos se basan principalmente en una tecnología de pre-procesador en el lado del codificador, y en una funcionalidad correspondiente del post-procesamiento en el lado del decodificador.

35 **[0006]** Típicamente, el codificador en el dominio de los tiempos se selecciona de manera que codifique señales útiles en el dominio de los tiempos tales como las señales de habla, y el codificador en el dominio de la frecuencia se selecciona para las señales que no son de habla, señales de música, etc. Sin embargo, específicamente para las señales que no son de habla que tienen armónicos prominentes en la banda de elevada frecuencia, los codificadores en el dominio de la frecuencia de la técnica anterior tienen una exactitud reducida por lo
40 que presentan una calidad de audio reducida debido al hecho de que tales armónicos prominentes pueden ser codificados solamente paramétricamente por separado o son eliminados por completo en el procedimiento de codificación/decodificación.

[0007] Además, existen conceptos donde la rama de codificación/decodificación en el dominio de los tiempos se basa adicionalmente en la extensión del ancho de banda que codifica también paramétricamente un intervalo de
45 frecuencias superior mientras que un intervalo de frecuencias inferior es codificado típicamente mediante un ACELP o con un codificador relacionado con CELP, por ejemplo un codificador del habla. Esta funcionalidad de extensión del ancho de banda aumenta la eficacia de la velocidad de bits, pero por otra parte, introduce una mayor inflexibilidad debido al hecho de que ambas ramas de codificación, es decir la rama de codificación en el dominio de las frecuencias y la rama de codificación en el dominio los tiempos, presentan una limitación de banda debido al
50 procedimiento de extensión del ancho de banda o al procedimiento de replicación por encima de una determinada frecuencia de cruce que es sustancialmente inferior a la máxima frecuencia incluida en la señal de audio introducida.

[0008] Los temas relevantes en el estado de la técnica comprenden:

55

- SBR como un post-procesador para la decodificación en forma de onda [1–3]
- Conmutación de núcleo MPEG-D USAC [4]
- MPEG-H 3D IGF [5]

60 **[0009]** En los siguientes documentos y patentes se describen procedimientos de los que se considera que constituyen la técnica anterior para la solicitud:

[1] M. Dietz, L. Liljeryd, K. Kjörling and O. Kunz, "Spectral Band Replication, a novel approach in audio coding," in 112th AES Convention, Munich, Germany, 2002.

65

[2] S. Meltzer, R. Böhm and F. Henn, "SBR enhanced audio codecs for digital broadcasting such as "Digital Radio

Mondiale" (DRM), in 112th AES Convention, Munich, Germany, 2002.

[3] T. Ziegler, A. Ehret, P. Ekstrand and M. Lutzky, "Enhancing mp3 with SBR: Features and Capabilities of the new mp3PRO Algorithm," in 112th AES Convention, Munich, Germany, 2002.

[4] MPEG-USAC Standard.

5 [5] PCT/EP2014/065109.

[0010] En el MPEG-D USAC se describe un codificador de núcleo conmutable. Sin embargo, en el USAC, el núcleo limitado en ancho de banda está restringido a transmitir siempre una señal filtrada en paso bajo. Por ello, determinadas señales de música que tienen un contenido prominente de elevada frecuencia, por ejemplo barridos de banda completa, sonido triángulo, etc. no puede reproducirse de manera fiel.

[0011] El documento WO 2011/048117 A1 describe un codificador de señal de audio, un decodificador de señal de audio y procedimientos relacionados que utilizan una cancelación de solapamiento (aliasing).

15 [0012] El documento US 2013/0030798 A1 describe un procedimiento y un aparato para codificación y decodificación de audio, donde la señal de audio incluye tramas de audio y voz genéricas. El codificador de voz utiliza dos codificadores y el decodificador de voz utiliza dos decodificadores. Durante un transitorio entre la voz y el audio genérico, los parámetros que necesita el decodificador de voz se generan mediante el procesamiento de la trama de audio genérico (sin voz) anterior para los parámetros necesarios.

20 [0013] El documento EP 2613316 A2 describe un procedimiento y un aparato para el procesamiento de tramas de audio para la transición entre diferentes códecs. El procedimiento incluye la producción, mediante el uso de un primer procedimiento de codificación, de una primera trama de muestras de audio de salida codificadas mediante la codificación de una primera trama de audio en una secuencia de tramas. Se forma una porción de solapamiento-adición mediante el uso del primer procedimiento de codificación. Además, se genera una trama de combinación de muestras de audio codificadas basándose en la combinación de la primera trama con la porción de solapamiento-adición de la primera trama, y se inicializa un estado de un segundo procedimiento de codificación basado en la combinación de la primera trama de muestras de audio codificadas.

30 [0014] La patente de EE.UU. 6.134.518 describe un procedimiento de codificación de señal de audio digital que utiliza un codificador CELP y un codificador de transformación. Se proporcionan un primer y segundo codificador para la codificación de manera digital de la señal de entrada mediante el uso del primer y segundo procedimientos de codificación respectivamente, y la disposición de conmutación dirige, en cualquier momento particular, la generación de una señal de salida mediante la codificación de la señal de entrada con el uso del primer o el segundo codificador según si la señal de entrada contiene una señal de audio de un primer tipo o un segundo tipo en ese momento.

40 [0015] El documento EP 2405426 A1 describe un procedimiento de codificación de señal de sonido, un procedimiento de decodificación de señal de sonido y los dispositivos correspondientes. Cuando una trama que precede inmediatamente a una trama de destino de codificación que va a ser codificada por una primera unidad codificada que opera bajo un esquema de codificación de predicción lineal está codificada por una segunda unidad de codificación que opera bajo un esquema de codificación diferente del esquema de codificación de predicción lineal, la trama de destino de codificación se puede codificar bajo el esquema de codificación de predicción lineal mediante la inicialización del estado de intervalo de la primera unidad de codificación.

45 [0016] Es un objeto de la presente invención proporcionar un concepto mejorado para la codificación de audio.

50 [0017] Este objeto se logra mediante un codificador de audio de la reivindicación 1, un decodificador de audio de la reivindicación 10, un procedimiento de codificación de audio de la reivindicación 16, un procedimiento de decodificación de audio de la reivindicación 17 o un programa informático de la reivindicación 18.

55 [0018] Un procesador de codificación/decodificación en el dominio de los tiempos se puede combinar con un procesador de codificación/decodificación en el dominio de las frecuencias que tenga una función de llenado de los intervalos, pero esta funcionalidad de llenado de los intervalos espectrales es operada sobre la totalidad de la banda de la señal de audio o al menos por encima de una determinada frecuencia de llenado de intervalos. Es importante observar que el procesador de codificación/decodificador en el dominio de las frecuencias es particularmente adecuado para llevar a cabo una codificación/decodificación exacta en forma de onda o de valor espectral hasta la máxima frecuencia y no solamente hasta una frecuencia de cruce. Además, la capacidad de banda completa del codificador en el dominio de las frecuencias para la codificación con la resolución elevada permite una integración de la funcionalidad del llenado de los intervalos en el codificador en el dominio de las frecuencias.

60 [0019] En un aspecto, el llenado de los intervalos en banda completa se combina con un procesador de codificación/decodificación en el dominio de los tiempos. En algunas realizaciones, las velocidades de muestreo en ambas ramas son iguales, o la velocidad del muestreo en la rama del codificador en el dominio de los tiempos es

inferior a la velocidad de muestreo en el dominio de las frecuencias.

- [0020]** En otro aspecto, un codificador/decodificador en el dominio de las frecuencias que opera sin llenado de intervalos pero que lleve a cabo una codificación/decodificación de núcleo de banda completa se combina con un procesador de codificación en el dominio de los tiempos y se proporciona un procesador de cruce para la inicialización continua del procesador de codificación/decodificación en el dominio de los tiempos. En este aspecto, las velocidades de muestreo pueden ser como en el otro aspecto, o las velocidades de muestreo en la rama en el dominio de las frecuencias son aún más bajas que en la rama en el dominio de los tiempos.
- [0021]** Mediante la utilización de un procesador codificador/decodificador espectral de banda completa, los problemas relacionados con la separación de la extensión del ancho de banda por una parte y la codificación de núcleo, por otra parte, pueden ser abordados y superados mediante la realización de la extensión del ancho de banda en el mismo dominio espectral donde opera el decodificador de núcleo. Por lo tanto, se proporciona un decodificador de núcleo de velocidad completa que codifica y decodifica el intervalo de señales de audio completo. Esto no requiere la necesidad de un muestreador descendente en el lado del codificador y de un muestreador ascendente en el lado del decodificador. En cambio, la totalidad del procesamiento es llevada a cabo con la velocidad de muestreo completa o con el dominio completo del ancho de banda. Con el fin de obtener una elevada ganancia de codificación, la señal de audio es analizada con el fin de encontrar un primer conjunto de primeras porciones espectrales que tiene que ser codificado con una elevada resolución, donde este primer conjunto de primeras porciones espectrales puede incluir, en una realización, porciones tonales de la señal de audio. Por otra parte, los componentes no tonales o ruidosos en la señal de audio que constituyen un segundo conjunto de segundas porciones espectrales son codificados paramétricamente con una baja resolución espectral. En tal caso, la señal de audio codificada solo requiere que el primer conjunto de primeras porciones espectrales sea codificado de manera que preserve la forma de onda con una elevada resolución espectral, y, adicionalmente, que el segundo conjunto de segundas porciones espectrales sea codificado paramétricamente con una baja frecuencia utilizándose "mosaicos" de frecuencia obtenidas del primer conjunto como fuente. En el lado del decodificador, el decodificador de núcleo, que es un decodificador de banda completa, reconstruye el primer conjunto de primeras porciones espectrales de manera que preserve la forma de onda, es decir, sin ningún conocimiento de que exista alguna regeneración adicional de la frecuencia. Sin embargo, el espectro generado de este modo tiene muchos intervalos espectrales. Estos intervalos son llenados posteriormente con la tecnología IGF (Intelligent Gap Filling, Llenado Inteligente de Intervalos) mediante la utilización de una regeneración de frecuencias para lo cual se utilizan datos paramétricos por una parte y se utiliza un intervalo espectral de fuente, es decir, primeras porciones espectrales reconstruidas por el decodificador de audio de velocidad completa por otra parte.
- [0022]** En otras realizaciones, las porciones espectrales, que son reconstruidas por el llenado de ruido solamente en lugar de replicación de ancho de banda o llenado de intervalos de frecuencia, constituyen un tercer conjunto de terceras porciones espectrales. Debido al hecho de que el concepto de la codificación opera en un único dominio para la codificación/decodificación por una parte y la regeneración de las frecuencias por otra parte, el IGF no está solamente restringido a rellenar un intervalo de frecuencias más elevado sino que puede rellenar intervalos de frecuencia más bajas, ya sea mediante el llenado de ruidos sin regeneración de las frecuencias como mediante la regeneración de frecuencias mediante el uso de un mosaico de frecuencias con un intervalo de frecuencias diferente.
- [0023]** Además, se hace énfasis en que una información acerca de las energías espectrales, una información sobre energías individuales o una información sobre energía individual, una información acerca de una energía de supervivencia o una información de energía de supervivencia, una información de la inercia de mosaico o una información de energía de mosaico, o una información acerca de una energía que falta o una información de energía que falta, puede comprender no solamente un valor de energía, sino también un valor de amplitud (por ejemplo, absoluto), un valor de nivel o cualquier otro valor, a partir del cual es posible derivar un valor final de la energía. Por ello, la información acerca de una energía puede comprender por ejemplo el valor de la energía como tal, y/o un valor de un nivel y/o de una amplitud y/o de amplitud absoluta.
- [0024]** Un aspecto adicional se basa en el descubrimiento de que la situación de correlación no solo es importante para el intervalo de fuente sino que también es importante para el intervalo objetivo. Además, diferentes situaciones de correlación pueden presentarse en el intervalo de fuente y en el intervalo objetivo. Si por ejemplo, se considera una señal de voz con un ruido de elevada frecuencia, puede darse la situación de que la banda de baja frecuencia que comprende la señal de voz con un bajo número de sobretonos esté estrechamente relacionada en el canal izquierdo y en el canal derecho, cuando el altavoz se coloca en el medio. Sin embargo, la porción de alta frecuencia puede estar fuertemente no correlacionada debido al hecho de que puede haber un ruido de elevada frecuencia diferente en el lado izquierdo en comparación con otro ruido de elevada frecuencia o de un ruido sin elevada frecuencia en el lado derecho. Por lo tanto, cuando deba llevarse a cabo una operación de llenado directa que ignore esta situación, entonces la porción de elevada frecuencia también estaría correlacionada, y esto podría generar serios artefactos de segregación espacial en la señal reconstruida. Con el fin de abordar este problema, se calculan los datos paramétricos para una banda de reconstrucción, o en términos generales, para el segundo conjunto de segundas porciones espectrales que tienen que ser reconstruidas, utilizando un primer conjunto de

primeras porciones espectrales con el fin de identificar ya sea una primera representación de dos canales o una segunda representación de canales, diferente para la banda de reconstrucción. En el lado del codificador, se calcula por lo tanto una identificación de dos canales para las segundas porciones espectrales, es decir para las porciones para las que adicionalmente se calcula información de energía adicional para las bandas de reconstrucción. Un regenerador de frecuencias en el lado del decodificador regenera seguidamente una segunda porción espectral en función de una primera porción del primer conjunto de porciones espectrales, es decir, el intervalo de fuente y los datos paramétricos para la segunda porción tal como información de energía de envolvente espectral o cualquier otro dato de envolvente espectral y, adicionalmente, en función de la identificación de los dos canales para la segunda porción, es decir, para esta banda de reconstrucción bajo reconsideración.

10

[0025] La identificación de dos canales se transmite preferiblemente como un indicador para cada banda de reconstrucción, y estos datos son transmitidos desde un codificador a un decodificador, y el decodificador decodifica seguidamente la señal núcleo como se indica mediante indicadores preferiblemente calculados para la banda de núcleo. A continuación, en una implementación, la señal de núcleo se almacena en ambas representaciones estéreo (por ejemplo, izquierda/derecha o centro/lado) y, para el llenado de los mosaicos de frecuencia por IGF, se elige la representación de mosaico de fuente para que concuerde con la representación de mosaico objetivo como se indica mediante los indicadores de identificación de los dos canales para el llenado inteligente de los intervalos o de las bandas de reconstrucción, es decir, para el intervalo objetivo.

15

20 **[0026]**

Se hace énfasis en que este procedimiento no solo funciona para señales estéreo, es decir, para un canal izquierdo y un canal derecho, sino que también opera para señales de múltiples canales. En el caso de las señales de múltiples canales, es posible procesar varios pares de diferentes canales de esta manera tal como un canal izquierdo y un canal derecho como primer par, un canal envolvente izquierdo y un canal envolvente derecho como el segundo par y un canal de centro y un canal LFE como un tercer par. Es posible determinar otras combinaciones para formatos de canales de salida más elevados tales como 7.1, 11.1, etc.

25

[0027]

Un aspecto adicional se basa en el descubrimiento de que la calidad de audio de la señal reconstruida se puede mejorar mediante IGF ya que la totalidad del espectro es accesible para el codificador de núcleo de tal manera que por ejemplo, las porciones tonales importantes desde el punto de vista perceptual en un intervalo espectral elevado todavía pueden ser codificadas por el codificador de núcleo en lugar de sustitución paramétrica. Adicionalmente se lleva a cabo una operación de llenado de intervalos utilizando mosaicos de frecuencia de un primer conjunto de primeras porciones espectrales que es, por ejemplo, un conjunto de porciones tonales típicamente de un intervalo de menor frecuencia, pero también de un intervalo de frecuencias más elevado, en caso de estar disponible. Sin embargo, para el ajuste de la envolvente espectral en el lado del decodificador, las porciones espectrales del primer conjunto de porciones espectrales situadas en la banda de reconstrucción no son post-procesadas posteriormente, por ejemplo, por el ajuste de la envolvente espectral. Solamente los valores espectrales remanentes en la banda de reconstrucción que no se originan en el codificador de núcleo tienen que ser ajustados por envolvente utilizando información de envolvente. Es preferible que la información de envolvente sea información de envolvente de banda completa que tenga en cuenta la energía del primer conjunto de primeras porciones espectrales en la banda de reconstrucción y el segundo conjunto de segundas porciones espectrales en la misma banda de reconstrucción, donde los valores espectrales mencionados en último término en el segundo conjunto de porciones espectrales están indicados para que sean cero, y por lo tanto, no son codificados por el codificador de núcleo, pero son codificados paraméricamente con información de energía de baja resolución.

30

35

40

45 **[0028]**

Se ha descubierto que los valores de energía absoluta, ya sean normalizados con respecto al ancho de banda de la banda correspondiente, o no normalizados, son útiles y eficaces en una aplicación en el lado del decodificador. Éste rige en especial cuándo tienen que calcularse los factores de ganancia sobre la base de una energía residual en la banda de reconstrucción, la energía que falta en la banda de reconstrucción y la información de mosaico de frecuencia en la banda de reconstrucción.

50

[0029]

Además, se prefiere que la corriente de bits codificada no solo abarque información de energía para la banda de reconstrucción, sino adicionalmente, factores de escala para bandas de factores de escala que se extienden hasta la máxima frecuencia. Esto asegura que para cada banda de reconstrucción, para la que se encuentra disponible una determinada porción tonal, es decir, una primera porción espectral, este primer conjunto de primeras porciones espectrales puede ser realmente decodificado con la amplitud correcta. Además, y en adición al factor de escala para cada banda de reconstrucción, se genera una energía para esta banda de reconstrucción en un codificador y se transmite a un decodificador. Además, se prefiere que las bandas de reconstrucción coincidan con las bandas de los factores de escala o en el caso de una agrupación de energía, que al menos los límites de una banda de reconstrucción coincidan con los límites de las bandas de factores de escala.

55

60

[0030]

Una implementación adicional aplica una operación de blanqueo de mosaicos. El blanqueo de un espectro elimina la información de envolvente espectral bruta y hace énfasis en la estructura fina espectral que es de sumo interés para evaluar la similitud de los mosaicos. Por ello, un mosaico de frecuencias por una parte y/o la señal de fuente por otra parte son blanqueadas antes de calcular una medida de correlación cruzada. Cuando se blanquee solamente el mosaico utilizando un procedimiento predefinido, se transmite un indicador de blanqueo que indica al

65

decodificador que el mismo procedimiento de blanqueo predefinido será aplicado al mosaico de frecuencia dentro del IGF. En cuanto a la selección de los mosaicos, se prefiere utilizar el retraso de la correlación para desplazar espectralmente el espectro regenerado en un número entero de tolvos de transformada. En función de la transformada subyacente, el desplazamiento espectral puede requerir correcciones de adición. En el caso de los 5 retrasos impares, el mosaico es adicionalmente modulado mediante multiplicación por una secuencia temporal alternante de $-1/1$ para compensar la representación de frecuencia inversa de cualquier otra banda dentro del MDCT. Además, se aplica el signo del resultado de correlación cuando se genera el mosaico de frecuencia.

[0031] Además, se prefiere utilizar el podado y estabilización de los mosaicos con el fin de asegurar que los 10 artefactos creados por las regiones fuente de rápido cambio para la misma región de reconstrucción o región objetivo se eviten. A tal efecto, se lleva a cabo un análisis de similitud entre las diferentes regiones fuente identificadas, y cuando un mosaico de fuente es similar a otros mosaicos de fuente con una similitud superior a un umbral, entonces este mosaico de fuente puede ser descartado del conjunto de mosaicos de fuente potenciales puesto que está estrechamente correlacionado con otros mosaicos de fuente. Además, como un tipo de 15 estabilización de la selección, se prefiere mantener el orden de los mosaicos de la trama anterior si ninguno de los mosaicos de fuente en la trama actual se correlaciona (mejor que en un umbral predeterminado) con los mosaicos objetivo en la trama actual.

[0032] Un aspecto adicional se basa en el descubrimiento de que se obtiene una calidad mejorada y una 20 velocidad de bits reducida específicamente para señales que comprenden porciones transitorias ya que se presentan muy frecuentemente en las señales de audio mediante la combinación de la tecnología de TSN (configuración temporal del ruido) o de TTS (configuración temporal de mosaicos) con una reconstrucción de elevada frecuencia. El procesamiento TNS/TTS en el lado del codificador implementado mediante una predicción sobre las frecuencias reconstruye la envolvente temporal de la señal de audio. En función de la implementación, es 25 decir, cuando se determina que el filtro de configuración de ruido temporal dentro de un intervalo de frecuencias no solamente abarca el intervalo de frecuencias fuente sino también el intervalo de frecuencia objetivo que va a ser reconstruido en un decodificador de regeneración de frecuencias, la envolvente temporal no solamente se aplica a la señal de audio núcleo hasta una frecuencia de inicio del llenado de intervalos, sino que la envolvente temporal también se aplica a los intervalos espectrales de segundas porciones espectrales reconstruidas. De esta manera se 30 reducen o eliminan los pre-ecos o post-ecos que se presentarían dentro sin la configuración temporal de los mosaicos. Esto se lleva a cabo aplicando una predicción inversa sobre las frecuencias no solamente dentro del intervalo de frecuencias del núcleo hasta una determinada frecuencia de inicio del llenado de intervalos, sino también dentro del intervalo de frecuencias superior al intervalo de frecuencias del núcleo. Con este fin, la regeneración de frecuencias o regeneración de mosaicos de frecuencia se lleva a cabo en el lado del decodificador 35 antes de aplicarse una predicción sobre las frecuencias. Sin embargo, la predicción en cuanto a las frecuencias se puede aplicar ya sea antes o posteriormente a la configuración de la envolvente espectral en función de si el cálculo de la información de la energía se ha llevado a cabo sobre los valores espectrales residuales posteriormente al filtrado o a los valores espectrales (completos) antes de la configuración de la envolvente.

40 **[0033]** El procesamiento sobre uno o más mosaicos de frecuencias (TTS) establece una continuidad de la correlación entre el intervalo de fuente y el intervalo de reconstrucción o en dos intervalos de reconstrucción adyacentes o mosaicos.

[0034] En una implementación, se prefiere utilizar el filtrado complejo TSN/TTS. De esta manera, se evitan 45 los artefactos de solapamiento (temporales) de una representación real típicamente muestreada, tal como MDCT. Es posible calcular un filtro TNS complejo en el lado del codificador mediante la aplicación no solo de una transformada de coseno discreta modificada sino también de una transformada de seno discreta modificada adicional con el fin de obtener una transformada modificada compleja. Sin embargo, solo se transmiten los valores de transformada de coseno discreta modificada, es decir, la parte real de la transformada compleja. Sin embargo, en el lado del 50 decodificador, es posible estimar la parte imaginaria de la transformada utilizando los espectros de MDCT de tramas anteriores o subsiguientes de tal manera, en el lado del decodificador, que es posible aplicar el filtro complejo de nuevo en la predicción inversa sobre las frecuencias, y específicamente, la predicción sobre el límite entre el intervalo de fuente y el intervalo de reconstrucción y también sobre el límite entre los mosaicos de frecuencia de frecuencias adyacentes dentro del intervalo de reconstrucción.

55 **[0035]** El sistema de codificación de audio ejemplar codifica de manera eficaz señales de audio arbitrarias en un amplio intervalo de velocidad de bits. En este contexto, el sistema converge hasta su transparencia, y para las velocidades de bits más bajas, se minimiza la molestia perceptiva. Por ello, la parte principal de la velocidad de bits disponible se utiliza para codificar la estructura perceptualmente más relevante de la señal en el codificador, y los 60 intervalos espectrales resultantes son llenados en el decodificador con un contenido de señales que se aproxima aproximadamente al espectro original. Se consume un presupuesto muy limitado de bits para controlar el llenado inteligente de intervalos (GFG, Gap Filling) mediante información secundaria dedicada transmitida desde el codificador al decodificador.

65 **[0036]** De conformidad con la invención, se proporciona un procesador cruzado para inicializar el

codificador/decodificador en el dominio de los tiempos con datos de inicialización derivados de la señal de codificador/decodificador en el dominio de frecuencias actualmente procesado. Esto permite que cuando se procesa la porción de señales de audio actualmente procesados por el codificador en el dominio de frecuencias, se inicializa el codificador en el dominio de los tiempos, paralelo, de tal manera que cuando tiene lugar una conmutación desde el codificador en el dominio de frecuencias a un codificador en el dominio de los tiempos, este codificador en el dominio de los tiempos puede empezar de inmediato el procesamiento ya que todos los datos de inicialización relacionados con señales anteriores ya están allí debido al procesador cruzado. Este procesador cruzado se aplica en el lado del codificador, y adicionalmente en el lado del decodificador y utiliza una transformada de frecuencia-tiempo que adicionalmente lleva a cabo un muestreo descendente muy eficaz desde la velocidad de muestreo de entrada o de salida más elevada, a la velocidad de muestreo del correspondiente núcleo en el dominio de los tiempos, más baja, mediante la mera selección de una determinada porción de baja banda de la señal del dominio junto con un determinado tamaño de transformada reducida. De esta manera se lleva a cabo de manera muy eficaz una conversión de la elevada velocidad de muestreo desde la elevada velocidad de muestreo a la baja velocidad de muestreo, y esta señal obtenida por la transformada con el tamaño de transformada reducida puede utilizarse seguidamente para inicializar el codificador/decodificador en el dominio los tiempos de tal manera que el codificador/decodificador en el dominio de los tiempos esté listo para llevar a cabo de inmediato la codificación en el dominio de los tiempos cuando esta situación es señalizada por un controlador y la porción de señal de audio inmediatamente anterior se ha codificado en el dominio de las frecuencias.

20 **[0037]** Tal como se ha destacado, la realización del procesador cruzado puede basarse en el llenado de intervalos en el dominio de frecuencias, o no. Por lo tanto, un codificador/decodificador en el dominio de los tiempos y de las frecuencias se combinan por medio del procesador cruzado, y el codificador/decodificador en el dominio de las frecuencias pueden basarse en el llenado de los intervalos, o no. Específicamente se prefieren determinadas realizaciones señaladas a grandes rasgos:

25 Estas realizaciones utilizan el llenado de intervalos en el dominio de las frecuencias y tienen los siguientes números de velocidades de muestreo y pueden basarse o no, en la tecnología del procesador cruzado:

30 Entrada SR = 8 kHz, ACELP (dominio del tiempo) SR = 12,8 kHz.
 Entrada SR = 16 kHz, ACELP SR = 12,8 kHz.
 Entrada SR = 16 kHz, ACELP SR = 16,0 kHz
 Entrada SR = 32.0 kHz, ACELP SR = 16,0 kHz
 Entrada SR = 48 kHz, ACELP SR = 16 kHz

35 **[0038]** Estas realizaciones pueden emplear o no el llenado de intervalos en el dominio de frecuencia y tienen los siguientes números de velocidad de muestreo y se basan en la tecnología de procesadores cruzados:

TCX SR es inferior al ACELP SR (8 kHz vs. 12,8 kHz), o cuando se ejecutan TCX y ACELP, ambos, a 16,0 kHz, y cuando no se utilice ningún llenado de los intervalos.

40 **[0039]** Por ello, las implementaciones preferidas permiten una conmutación continua de un codificador de audio perceptual que comprende el llenado de intervalos espectrales y un codificador en el dominio de los tiempos con o sin extensión de ancho de banda.

45 **[0040]** Por ello, se confía en procedimientos que no se restringen a eliminar el contenido de elevada frecuencia por encima de una frecuencia de corte en el codificador en el dominio de las frecuencias a partir de la señal de audio sino más bien que elimina de una manera adaptativa las señales, las regiones de pase de banda espectrales que dejan intervalos espectrales en el codificador y reconstruyen posteriormente estos intervalos espectrales en el decodificador. Es preferible utilizar una solución integrada tal como el llenado inteligente de los intervalos que de manera eficaz combina la codificación de audio de ancho de banda completo con el llenado de los intervalos espectrales particularmente en el dominio de las transformadas MDCT.

50 **[0041]** Por ello, y a diferencia de procedimientos ya existentes, el nuevo concepto utiliza la codificación de forma de onda de señales de audio de banda completa en el codificador en el dominio de las transformadas y al mismo tiempo permite una conmutación continua a un codificador de voz preferentemente seguido por una extensión del ancho de banda en el dominio de los tiempos.

60 **[0042]** Implementaciones adicionales evitan los problemas explicados que se presenten debido a una limitación de banda fija. El concepto permite la combinación conmutable de un codificador de forma de onda de banda completa en el dominio de las frecuencias equipado con un llenado de intervalos espectrales y un codificador de voz con una velocidad de muestreo más baja y una extensión del ancho de banda en el dominio de los tiempos. Un codificador de este tipo es capaz de la codificación de formas de onda de las señales problemáticas anteriormente mencionadas que proporciona un ancho de banda completo hasta la frecuencia Nyquist de la señal de entrada de audio. Sin embargo, la conmutación instantánea continua entre ambas estrategias de codificación está asegurada en particular por la realización que tiene el procesador cruzado. Para esta conmutación continua, el

procesador cruzado representa una conexión cruzada tanto en el codificador como en el decodificador entre el codificador en el dominio de las frecuencias de velocidad completa capaz de banda completa (velocidad de muestreo de entrada) y el codificador ACELP de baja velocidad que tiene una velocidad de muestreo más baja con el fin de inicializar adecuadamente los parámetros ACELP y las memorias intermedias particularmente dentro del libro de código adaptativo, el filtro de LPC o la etapa de remuestreo, cuando se conmute a partir del codificador en el dominio de las frecuencias tal como TCX al codificador en el dominio de los tiempos tal como ACELP.

[0043] A continuación se expone la presente invención con respecto a los dibujos adjuntos en los que:

- 10 Fig. 1a ilustra un aparato para la codificación de una señal de audio;
 Fig. 1b ilustra un decodificador para la decodificación de una señal de audio codificada que coincide con el decodificador de la Fig. 1a;
 Fig. 2a ilustra una implementación preferida del decodificador;
 Fig. 2b ilustra una implementación preferida del codificador
- 15 Fig. 3a ilustra una representación esquemática de un espectro generado por el decodificador en el dominio de los espectros de la Fig. 1b;
 Fig. 3b ilustra una tabla que indica la relación entre los factores de escala para las bandas de los factores de escala y la información del llenado del ruido para una banda de llenado de ruido;
 Fig. 4a ilustra la funcionalidad del codificador en el dominio de los espectros para aplicar la selección de porciones espectrales en los conjuntos primero y segundo de porciones espectrales;
 Fig. 4b ilustra una implementación de la funcionalidad de la Fig. 4a;
 Fig. 5a ilustra una funcionalidad de un codificador MDCT;
 Fig. 5b ilustra una funcionalidad del decodificador con una tecnología MDCT;
 Fig. 5c ilustra una implementación del regenerador de frecuencias;
- 25 Fig. 6 ilustra una implementación de un codificador de audio;
 Fig. 7a ilustra un procesador cruzado con el codificador de audio;
 Fig. 7b ilustra una implementación de una transformada de frecuencia-tiempo o inversa que proporciona adicionalmente una reducción de la velocidad de muestreo dentro del procesador cruzado;
 Fig. 8 ilustra una implementación preferida del controlador de la Fig. 6;
- 30 Fig. 9 ilustra otra realización del codificador en el dominio de los tiempos que tiene funcionalidades de extensión de ancho de banda;
 Fig. 10 ilustra una utilización preferida de un pre-procesador;
 Fig. 11a ilustra una implementación esquemática del decodificador de audio;
 Fig. 11b ilustra un procesador cruzado dentro del decodificador para proporcionar datos de inicialización para el
- 35 decodificador en el dominio de los tiempos;
 Fig. 12 ilustra una implementación preferida del procesador de decodificación en el dominio de los tiempos de la Fig. 11a;
 Fig. 13 ilustra una implementación adicional de la extensión del ancho de banda en el dominio de los tiempos;
 Fig. 14a ilustra una implementación preferida de un codificador de audio;
 Fig. 14b ilustra una implementación preferida de un decodificador de audio;
 Fig. 14c ilustra una implementación de un decodificador en el dominio de los tiempos con una conversión de la velocidad de muestreo y extensión de ancho de banda.

[0044] La Fig. 6 ilustra un codificador de audio para codificar una señal de audio que comprende un primer procesador de codificación 600 para codificar una primera porción de la señal de audio en un dominio de las frecuencias. El primer procesador de codificación 600 comprende un convertidor de tiempo-frecuencia 602 para convertir la primera porción de la señal de audio en una representación en el dominio de las frecuencias que tienen líneas espectrales hasta una frecuencia máxima de la señal de entrada. Por otra parte, el primer procesador de codificación 600 comprende un analizador 604 para analizar la representación en el dominio de las frecuencias hasta la frecuencia máxima con el fin de determinar primeras regiones espectrales que van a ser codificadas con una primera representación espectral y para determinar segundas regiones espectrales que van a ser codificadas con una segunda resolución espectral que es inferior a la primera resolución espectral. En particular, el analizador de banda completa 604 determina cuáles líneas de frecuencia o valores espectrales en el espectro del convertidor de tiempo-frecuencia tienen que ser codificadas a modo de espectro y cuáles otras porciones espectrales tienen que ser codificadas de una manera paramétrica, y estos últimos valores espectrales son reconstruidos seguidamente en el lado del decodificador mediante el procedimiento de llenado de intervalos. La operación de codificación real es llevada a cabo por un codificador espectral 606 para codificar las primeras regiones espectrales o porciones espectrales con la primera resolución y para codificar paraméricamente las segundas regiones espectrales o sus porciones con la segunda resolución espectral.

[0045] El codificador de audio de la Fig. 6 comprende adicionalmente un segundo procesador de codificación 610 para codificar la porción de la señal de audio en un dominio de los tiempos. Adicionalmente, el codificador de audio comprende un controlador 620 configurado para analizar la señal de audio en una señal de entrada de audio 601 y para determinar qué porción de la señal de audio es la primera porción de la señal de audio codificada en el dominio de las frecuencias y qué porción de la señal de audio es la segunda porción de la señal de audio codificada

en el dominio de los tiempos. Por otra parte, se ha proporcionado un formador de señales codificadas 630 que puede, por ejemplo, estar implementado como un multiplexador de la corriente de bits que está configurado para formar una señal de audio codificada que comprende una primera porción de la señal codificada para la primera señal de audio y una segunda porción de la señal codificada para la segunda porción de audio. Es importante
5 observar que la señal codificada solo tiene ya sea una representación en el dominio de las frecuencias o una representación en el dominio de los tiempos, de una y de la misma porción de la señal de audio.

[0046] Por lo tanto, el controlador 620 asegura que para una única porción de la señal de audio solo una representación en el dominio de los tiempos o una representación en el dominio de las frecuencias esté en la señal
10 codificada. Esto puede lograrse mediante el controlador 620 de diversas maneras. Una manera sería que, para una y la misma porción de la señal de audio, ambas representaciones lleguen al bloque 630 y el controlador 620 controle el formador de la señal codificada 630 con el fin únicamente de introducir una de ambas representaciones en la señal codificada. Sin embargo, como alternativa el controlador 620 puede controlar una entrada en el primer
15 procesador de codificación y una entrada en el segundo procesador de codificación de tal manera que, sobre la base del análisis de la correspondiente porción de la señal, solo uno de ambos bloques 600 o 610 es activado con el fin de llevar a cabo realmente la operación de codificación completa, y el otro bloque se desactiva.

[0047] Esta desactivación puede ser una desactivación, o como se ilustra, por ejemplo con respecto a la Fig. 7a, es solo un tipo de modo de "inicialización" donde el otro procesador de codificación está activo únicamente para
20 recibir y procesar datos de inicialización con el fin de inicializar las memorias internas pero ninguna operación de codificación específica en absoluto. Esta activación puede efectuarse mediante una determinada conmutación en la entrada que no se ilustra en la Fig. 6, o, lo que es preferible, mediante las líneas de control 621 y 622. Por lo tanto, en esta realización, el segundo procesador de codificación 610 no emite nada cuando el controlador 620 ha
25 determinado que la porción de la señal de audio vigente debería ser codificada por el primer procesador de codificación pero el segundo procesador de codificación se proporciona sin embargo con datos de inicialización de manera que sea activo para una conmutación instantánea en el futuro. Por otra parte, el primer procesador de codificación está configurado para no necesitar ningún dato del pasado para actualizar cualquier memoria interna, y por ello, cuando la porción de la señal de audio vigente tiene que ser codificada por el segundo procesador de
30 codificación 610, entonces el controlador 620 puede controlar el primer procesador de codificación final 600 por medio de la línea de control 620 para ser completamente inactivo. Esto significa que no es necesario que el primer procesador de codificación 600 se encuentre en un estado de inicialización o estado de espera, sino que puede estar en un estado de desactivación completa. Esto es preferible en particular para dispositivos móviles en los que el consumo de energía eléctrica, y por lo tanto, la vida de la batería es un tema que se debe tener en cuenta.

[0048] En la implementación específica adicional del segundo procesador de codificación que opera en el dominio de los tiempos, el segundo procesador de codificación comprende un muestreador descendente 900 o un
35 convertidor de velocidad de muestreo para convertir la porción de la señal de audio en una representación con una velocidad de muestreo más baja, donde la velocidad de muestreo más baja es inferior a la velocidad de muestreo en la entrada en el primer procesador de codificación. Esto se ilustra en la Fig. 9. En particular, cuando la señal de audio de entrada comprende una banda baja y una banda alta, se prefiere que la representación de la velocidad de
40 muestreo inferior en la salida del bloque 900 tenga solamente la banda baja de la porción de la señal de audio de entrada y esta banda baja es codificada seguidamente por el codificador de banda baja en el dominio de los tiempos 910 que está configurado para la codificación en el dominio de los tiempos de la representación de la velocidad de muestreo más baja proporcionada por el bloque 900. Además, se proporciona un codificador de extensión de ancho
45 de banda en el dominio de los tiempos 920 para codificar paramétricamente la banda alta. Con este fin, el codificador de la extensión de la banda baja en el dominio de los tiempos 920 recibe al menos la banda alta de la señal de audio de entrada o la banda baja y la banda alta de la señal de audio de entrada.

[0049] En otra realización de la presente invención, el codificador de audio comprende adicionalmente,
50 aunque no se ilustra en la Fig. 6 pero si en la Fig. 10, un procesador preliminar 1000 configurado para preprocesar la primera porción de la señal de audio y la segunda porción de la señal de audio. Es preferible que el procesador preliminar 100 comprenda dos ramas, donde la primera rama se ejecuta a 12,8 kHz, y lleva a cabo el análisis de la señal que más tarde se utiliza en el estimador de ruido, VAD, etc. La segunda rama se ejecuta con la velocidad de muestreo ACELP, es decir, en función de la configuración 12,8 o 16 kHz. En la práctica, cuando la velocidad de
55 muestreo ACELP es de 12,8 kHz, la mayor parte del procesamiento en esta rama se omite y, en cambio, se utiliza la primera rama.

[0050] En particular, el procesador preliminar comprende un detector transitorio 1020, y la primera rama es "abierta" por un remuestreador 1021 a por ejemplo 12,8 kHz, seguido por una etapa de énfasis preliminar 1005a, un
60 analizador de LPC 1002a, una etapa de filtrado de análisis ponderado 1022a, y un FFT/Estimador del Ruido/Detección de la Actividad de Voz (VAD, Voice Activity Detection) o etapa de Búsqueda de tono 1007.

[0051] La segunda rama es "abierta" por un remuestreador 1004 a por ejemplo 12,8 kHz o 16 kHz, es decir, a la Velocidad de Muestreo de ACELP, seguido por una etapa de énfasis preliminar 1005b, un analizador de LPC
65 1002b, una etapa de filtrado de análisis ponderado 1022b, y una etapa de extracción de parámetro TCX LTP 1024.

El bloque 1022 proporciona su salida al multiplexador de la corriente de bits. El bloque 1002 está conectado a un cuantificador LPC 1010 controlado por la decisión ACELP/TCX, y el bloque 1010 está también conectado al multiplexador de la corriente de bits.

5 **[0052]** Como alternativa, otras realizaciones pueden comprender solamente una única rama o más ramas. En una realización, este procesador preliminar comprende un analizador de predicción para determinar coeficientes de predicción. Este analizador de predicción puede ser implementado como un analizador LPC (linear prediction coding, codificación de predicción lineal) para determinar coeficientes de LPC. Sin embargo, también es posible implementar otros analizadores. Además, en la realización alternativa el procesador preliminar puede comprender un
10 cuantificador de coeficientes de predicción, donde este dispositivo recibe datos de coeficientes de predicción procedentes del analizador de predicciones.

[0053] Sin embargo, es preferible que el cuantificador de LPC no sea necesariamente parte del procesador preliminar, y que sea implementado como parte de la rutina de codificación principal, es decir, que no sea parte del
15 procesador preliminar.

[0054] Además, el procesador preliminar puede comprender adicionalmente un codificador de entropía para generar una versión codificada de los coeficientes de predicción cuantificados. Es importante observar que el formador de señales codificadas 630 o la implementación específica, es decir, el multiplexador de la corriente de bits
20 630 asegure que la versión codificada de los coeficientes de predicción cuantificados esté incluida en la señal de audio codificada 632. Es preferible que los coeficientes de LPC no sean cuantificados directamente sino que sean convertidos en una representación ISF, por ejemplo, o en cualquier otra representación mejor adecuada para la cuantificación. Es preferible que esta conversión se lleve a cabo ya sea mediante la determinación del bloque de coeficientes o dentro del bloque para cuantificar los coeficientes de LPC.

25 **[0055]** Además, el procesador preliminar puede comprender un remuestreador para remuestrear una señal de audio de entrada con una velocidad de muestreo introducida en una velocidad de muestreo más baja para el codificador en el dominio de los tiempos. Cuando el codificador en el dominio de los tiempos es un codificador ACELP que tiene una determinada velocidad de muestreo ACELP, entonces el muestreo descendente se lleva a
30 cabo a preferiblemente 12,8 kHz o 16 kHz. La velocidad de muestreo de entrada puede ser cualquiera de entre un número particular de velocidad de muestreo tales como 32 Hz o una velocidad de muestreo aún más elevada. Por otra parte, la velocidad de muestreo del codificador en el dominio de los tiempos estará predeterminada por determinadas restricciones, y el remuestreador 1004 lleva a cabo este remuestreo y emite la representación de la velocidad de muestreo más baja de la señal de entrada. Por lo tanto, el remuestreador puede llevar a cabo una
35 funcionalidad similar y puede ser incluso uno y el mismo elemento que el muestreador descendente 900 ilustrado en el contexto de la Fig. 9.

[0056] Además, se requiere aplicar un énfasis preliminar en el bloque de énfasis preliminar. El procesamiento del énfasis preliminar es bien conocido en la técnica de la codificación en el dominio de los tiempos, y se describe en
40 la bibliografía con referencia al procesamiento AMR-WEB+ y el énfasis preliminar está particularmente configurado para compensar una inclinación espectral, y por lo tanto permite un mejor cálculo de los parámetros de LPC con un orden de LPC dado.

[0057] Además, el procesador preliminar puede comprender adicionalmente una extracción de parámetros de
45 TCX-LTP para controlar un post-filtro de LTP ilustrado en 1420 en la Fig. 14b. Además, el procesador preliminar puede comprender adicionalmente otras funcionalidades ilustradas en 1007, y estas otras funcionalidades pueden comprender una funcionalidad de búsqueda de tonos, una funcionalidad de detección de la activación de voz (VAD, voice activity detection, detección de actividad de voz) o cualquier otra funcionalidad conocida en la técnica de la codificación en el dominio de los tiempos o de la voz.

50 **[0058]** Como se ilustra, el resultado del bloque 1024 se introduce en la señal codificada, es decir, en la realización de la Fig. 14a, se introduce en el multiplexador de la corriente de bits 630. Además, en caso de necesidad, los datos procedentes del bloque 1007 también pueden ser introducidos en el multiplexador de la corriente de bits o pueden, como alternativa, ser utilizados para la finalidad de la codificación en el dominio de los
55 tiempos en el codificador en el dominio de los tiempos.

[0059] Por lo tanto, para resumir, es común para ambas trayectorias un procesamiento preliminar 1000 donde se llevan a cabo las operaciones de procesamiento de las señales comúnmente utilizadas. Estas operaciones comprenden un remuestreo con una velocidad de muestreo de ACELP (de 12,8 o 16 kHz) para una trayectoria
60 paralela, y este remuestreo se lleva a cabo siempre. Además, se lleva a cabo una extracción de parámetros de TCX LTP ilustrada en el bloque 1006, y adicionalmente se lleva a cabo un énfasis preliminar y una determinación de los coeficientes de LPC. Tal como se señala a grandes rasgos, el énfasis preliminar compensa la inclinación espectral y por lo tanto hace que sea más eficaz el cálculo de los parámetros de LPC en un orden de LPC dado.

65 **[0060]** Posteriormente, se hace referencia a la Fig. 8 con el fin de ilustrar una implementación preferida del

controlador 620. El controlador recibe, en una entrada, la porción de la señal de audio en consideración. Es preferible, tal como se ilustra en la Fig. 14a, que el controlador reciba cualquier señal disponible en el procesador preliminar 1000 que puede ser la señal de entrada original con la velocidad de muestreo de entrada o una versión remuestreada con la velocidad de muestreo del codificador en el dominio en el tiempo, inferior, o una señal obtenida posteriormente al procesamiento de énfasis preliminar en el bloque 1005.

[0061] Sobre la base de esta porción de la señal de audio, el controlador 620 aborda un simulador del codificador en el dominio de las frecuencias 621 y un simulador del codificador en el dominio de los tiempos 622 con el fin de calcular para cada posibilidad de codificador una relación estimada entre señal y ruido. Posteriormente, el selector 623 selecciona el codificador que haya proporcionado la mejor relación entre señal y ruido, naturalmente bajo la consideración de una velocidad de bits predefinida. El selector identifica seguidamente el correspondiente codificador por medio de la salida de control. Cuando se determine que la porción de la señal de audio bajo consideración debe ser codificada mediante el uso del codificador en el dominio de las frecuencias, el codificador en el dominio de los tiempos se establece en una etapa de inicialización, o en otras realizaciones que no requieren una conmutación instantánea, en un estado completamente desactivado. Sin embargo, cuando se determine que la porción de la señal de audio bajo consideración debe ser codificada por el codificador en el dominio de los tiempos, en tal caso se desactiva el codificador en el dominio de las frecuencias.

[0062] A continuación se ilustra una implementación preferida del controlador ilustrado en la Fig. 8. La decisión de si debe elegirse la trayectoria ACELP o TCX se lleva a cabo en la decisión de conmutación mediante la simulación del codificador ACELP y TCX y conmutando a la rama de mejor rendimiento. Para ello, se estima el SNR de las ramas de ACELP y de TCX sobre la base de una simulación de codificador/decodificador de ACELP y TCX. La simulación del codificador/decodificador de TCX se lleva a cabo sin análisis de TNS/TTS, codificador IGF, codificador de bucle de cuantificación/aritmético, y sin ningún decodificador TCX. En cambio, se estima el SNR de TCX utilizando una estimación de la distorsión del cuantificador en el dominio de MDCT configurado. La simulación del codificador/decodificador de ACELP se lleva a cabo mediante el uso solo de una simulación del libro del código adaptativo y del libro de código informativo. El SNR de ACELP se estima simplemente computando la distorsión introducida por un filtro LTP en el dominio de las señales ponderado (libro de código adaptativo) y escalando esta distorsión mediante un factor constante (libro de código innovador). De esta manera se reduce considerablemente la complejidad en comparación con una estrategia donde la codificación de TCX y de ACELP se lleva a cabo en paralelo. La rama con el SNR más elevado se elige para la subsiguiente ejecución de codificación completa.

[0063] En el caso en que se elige la rama TCX, se ejecuta un decodificador TCX en cada trama que emite una señal con la velocidad de muestreo ACELP. Esto se utiliza para actualizar las memorias utilizadas para la trayectoria de codificación ACELP (LPC residual, Memw0, desénfasis de la memoria), de manera que se permita la conmutación instantánea de TCX a ACELP. La actualización de la memoria se lleva a cabo en cada trayectoria de TCX. Como alternativa puede llevarse a cabo un análisis completo mediante un procedimiento de síntesis, es decir, ambos simuladores de codificador 621,622 implementan las operaciones de codificación reales y los resultados son comparados por el selector 623. De nuevo, como alternativa, es posible efectuar un cálculo directo de alimentación completa mediante la realización de un análisis de las señales. Por ejemplo, cuando se determina que la señal es una señal de voz por un clasificador de señales, se selecciona el codificador en el dominio de los tiempos y cuando se determine que la señal es una señal de música, entonces se selecciona el codificador en el dominio de las frecuencias. También es posible aplicar otros procedimientos para diferenciar entre ambos codificadores sobre la base de un análisis de las señales de la porción de audio bajo consideración.

[0064] El codificador de audio comprende adicionalmente un procesador cruzado 700 ilustrado en la Fig. 7a. Cuando el codificador en el dominio de las frecuencias 600 es activo, el procesador cruzado 700 proporciona datos de inicialización al codificador en el dominio de los tiempos 610 de tal manera que el codificador en el dominio de los tiempos está listo para una conmutación continua a una porción de señal futura. En otras palabras, cuando se determine que la porción de señal actual tiene que ser codificada mediante el codificador en el dominio de las frecuencias, y cuando se determine mediante el codificador que la porción de señal de audio siguiente debe ser codificada por el codificador en el dominio de los tiempos 610, entonces, sin un procesador cruzado, una conmutación continua inmediata de ese tipo no sería posible. Sin embargo, el procesador cruzado proporciona una señal derivada del codificador en el dominio de las frecuencias 600 al codificador en el dominio de los tiempos 610 para los fines de inicializar las memorias en el codificador en el dominio de los tiempos ya que el codificador en el dominio de los tiempos 610 depende de una trama actual de la señal de entrada o codificada de una trama inmediatamente anterior en el tiempo.

[0065] Por ello, el codificador en el dominio de los tiempos 610 está configurado para ser inicializado por los datos de inicialización con el fin de codificar una porción de la señal de audio que sigue a una porción de señal de audio más temprana codificada por el codificador en el dominio de las frecuencias 600 de una manera eficaz.

[0066] En particular, el procesador cruzado comprende un convertidor de frecuencia-tiempo para convertir una representación en el dominio de las frecuencias en una representación en el dominio de los tiempos que puede ser enviada al codificador en el dominio de los tiempos directamente o después de algún otro procesamiento

adicional. Este convertidor se ilustra en la Fig. 14a como un bloque IMDCT (inverse modified discrete cosine transform, transformada de coseno discreta modificada inversa). Sin embargo, este bloque 702 tiene un tamaño de transformadas diferente en comparación con el bloque del convertidor de tiempo-frecuencias 602 indicado en el bloque de la Fig. 14a (modified discrete cosine transform block, bloque de transformadas de coseno discretas modificadas). Tal como se indica en el bloque 602, en algunas realizaciones, el convertidor de tiempo-frecuencia 602 opera con una velocidad de muestreo de entrada y la transformada de coseno discretas modificadas inversas 702 opera con la velocidad de muestreo ACELP, más baja.

[0067] En otras realizaciones, como los modos que operan con banda estrecha con una velocidad de muestreo de entrada de 8 kHz, la rama de TCX opera a 8 kHz, mientras que el ACELKP sigue ejecutándose a 12,8 kHz, es decir, el SNR del ACELP no es siempre inferior a la velocidad de muestreo de TCX. Para una velocidad de muestreo de entrada de 16 kHz (ancho de banda) también existen escenarios donde se ejecuta el ACELP con la misma velocidad de muestreo que TCX, es decir ambos se ejecutan a 16 kHz. En un modo de superancho de banda (SWB, superband mode), la velocidad de muestreo de entrada es de 32 o 48 kHz.

[0068] La relación entre la velocidad de muestreo del codificador en el dominio de los tiempos o velocidad de muestreo ACELP y la velocidad de muestreo del codificador en el dominio de las frecuencias o velocidad de muestreo de entrada puede calcularse, y es un factor de muestreo descendente DS (downsampling factor) ilustrado en la Fig. 7b. El factor de muestreo descendente es superior a 1 cuando la velocidad de muestreo de salida de la operación de muestreo descendente es inferior a la velocidad de muestreo de entrada. Sin embargo, cuando hay un muestreo ascendente real, entonces la velocidad de muestreo descendente es inferior a 1 y se lleva a cabo un muestreo ascendente real.

[0069] Para un factor de muestreo descendente superior a uno, es decir para un muestreo descendente real, el bloque 602 tiene un tamaño de transformada mayor y el bloque de IMDCT 702 tiene un tamaño de transformada pequeño. Por ello, y tal como se ilustra en la Fig. 7b, el bloque de IMDCT 702 comprende un selector 726 para seleccionar la porción espectral inferior de una entrada en el bloque de IMDCT 702. La porción del espectro de banda completa se define mediante el factor de muestreo descendente DS. Por ejemplo, cuando la velocidad de muestreo descendente es de 16 kHz y la velocidad de muestreo de entrada es de 32 kHz, entonces el factor de muestreo descendente es 2,0, y por lo tanto, el selector 726 selecciona la mitad inferior del espectro de banda completa. Cuando el espectro tiene, por ejemplo, 1024 líneas de MDCT, entonces el selector selecciona las 512 líneas inferiores.

[0070] Esta porción de baja frecuencia del espectro de banda completa se introduce en una transformada de tamaño pequeño y bloque de despliegue 720, como se ilustra en la Fig. 7b. El tamaño de la transformada se selecciona también de conformidad con el factor de muestreo descendente y representa el 50 % del tamaño de la transformada en el bloque 602. Un efecto de formación de ventana de síntesis con una ventana provista de una pequeña cantidad de coeficientes se lleva a cabo seguidamente. El número de coeficientes de la ventana de síntesis es igual a la inversa del factor de muestreo descendente multiplicado por el número de coeficientes de la ventana de análisis utilizada por el bloque 602. Finalmente, se lleva a cabo una operación de superposición-adición con el número más pequeño de operaciones por bloque, y de nuevo el número de operaciones por bloque es el número de operaciones por bloque en una implementación de velocidad completa de MDCT multiplicado por la inversa del factor de muestreo descendente.

[0071] Por lo tanto, una operación de muestreo descendente muy eficaz puede aplicarse ya que el muestreo descendente está incluido en la implementación de MDCT. En este contexto, se hace énfasis en que el bloque 602 puede ser implementado por un IMDCT pero también puede ser implementado por cualquier otra transformada o implementación de banco de filtros que puede estar adecuadamente dimensionada en un núcleo de transformadas real y en otras operaciones relacionadas con transformadas.

[0072] Para un factor de muestreo descendente inferior a uno, es decir, para un muestreo ascendente real, la notación en la Fig. 7b, bloques 720, 722, 724, 726, debe ser invertida. El bloque 722 selecciona el espectro de banda completa y adicionalmente se reduce a cero para las líneas espectrales superiores no incluidas en el espectro de banda completa. El bloque 720 tiene un tamaño de transformada superior al del bloque 710, y el bloque 722 tiene una ventana con un número de coeficientes mayor que el bloque 712, y también el bloque 724 tiene un número de operaciones mayor que en el bloque 714.

[0073] El bloque 602 tiene un tamaño de transformada pequeño y el bloque de IMDCT 702 tiene un gran tamaño de transformada. Por lo tanto y tal como se ilustra en la Fig. 7b, el bloque de IMDCT 702 comprende un selector 726 para seleccionar la porción espectral completa de una entrada en el bloque de IMDCT 702 y para la banda alta adicional requerida para la salida, se seleccionan ceros y ruido y se colocan en la banda superior requerida. La porción del espectro de banda completa está definida por el factor de muestreo descendente DS. Por ejemplo, cuando la velocidad de muestreo superior es de 16 kHz y la velocidad de muestreo de entrada es de 8 kHz, en tal caso el factor de muestreo descendente es de 0,5, y por lo tanto, el selector 726 selecciona el espectro de banda completa y selecciona adicionalmente de manera preferible ceros o ruido de baja energía para la porción

superior no incluida en el espectro del dominio de las frecuencias de banda completa. Cuando el espectro tiene por ejemplo 1024 líneas de MDCT, entonces el selector selecciona 1024 líneas de MDCT, y para las 1024 líneas de MDCT adicionales se seleccionan preferiblemente ceros.

5 **[0074]** Esta porción de frecuencia del espectro de banda completa se introduce en una transformada de gran tamaño y bloque de plegado 720, como se ilustra en la Fig. 7b. El tamaño de la transformada también se selecciona de conformidad con el factor de muestreo descendente y representa el 200 % del tamaño de la transformada en el bloque 602. A continuación se lleva a cabo un efecto de formación de ventana de síntesis con una ventana provista de un número de coeficientes más elevado. El número de coeficientes de la ventana de síntesis es igual al factor del muestreo descendente inverso dividido por el número de coeficientes de la ventana de análisis utilizada por el bloque 602. Finalmente, se lleva a cabo una operación de superposición-adición con un número más elevado de operaciones por bloque, y de nuevo el número de operaciones por bloque es el número de operaciones por bloque en un MDCT de implementación de velocidad completa multiplicado por la inversa del factor de muestreo descendente.

10
15 **[0075]** Por lo tanto, es posible aplicar una operación de muestreo ascendente muy eficaz ya que el muestreo ascendente está incluido en la implementación de IMDCT. En este contexto, se hace énfasis en que el bloque 702 puede ser implementado por un IMDCT pero también puede ser implementado por cualquier otra transformada o implementación de filtro de bancos que puede estar dimensionado de manera adecuada en el núcleo de transformada actual y en otras operaciones relacionadas con transformadas.

20 **[0076]** En términos generales, se destaca que una definición de la velocidad de muestras en el dominio de las frecuencias requiere alguna explicación. Las bandas espectrales son frecuentemente muestreadas descendentemente. Por lo tanto se utiliza la noción de una velocidad de muestreo efectiva o de una velocidad de muestra "asociada" o velocidad de muestreo. En el caso de un banco de filtros/transformada la velocidad efectiva del muestreo se definiría como $Fs_eff = \text{subbandsamplerate} * \text{num_subbands}$.

25 **[0077]** En otra realización ilustrada en la Fig. 14a, el convertidor de tiempo-frecuencia comprende funcionalidades adicionales además del analizador. El analizador 604 de la Fig. 6 puede comprender en la realización de la Fig. 14a un bloque temporal de configuración de ruido/análisis de la configuración de mosaico temporal 604a que opera como se expuso en el contexto de la Fig. 2b bloque 222 para el bloque de análisis TNS/TTS 604a y tal como se ilustra con respecto a la Fig. 2b para la máscara tonal 226 que corresponde al codificador de IGF 604b en la Fig. 14a.

30 **[0078]** Por otra parte, el codificador en el dominio de las frecuencias comprende preferiblemente un bloque de configuración de ruido 606a. El bloque de configuración de ruido 606a es controlado por coeficientes LPC cuantificados tal como son generados por el bloque 1010. Los coeficientes LPC cuantificados utilizados para la configuración de ruido 606a llevan a cabo una configuración espectral de los valores espectrales de elevada resolución o de las líneas espectrales directamente codificadas (en lugar de codificados paramétricamente), y el resultado del bloque 606a es similar al espectro de una señal subsiguiente a una etapa de filtración de LPC que opera en el dominio de los tiempos tal como un bloque de filtrado de análisis de LPC 704 que será descrito más adelante en esta invención. Por otra parte, el resultado del bloque de la configuración del ruido 606a es seguidamente cuantificado y codificado en entropía tal como se indica mediante el bloque 606b. El resultado del bloque 606b corresponde a la primera porción de la señal de audio codificada o a una porción de la señal de audio codificada en el dominio de las frecuencias (junto con otra información secundaria).

35 **[0079]** El procesador cruzado 700 comprende un decodificador espectral para calcular una versión decodificada de la primera porción de la señal codificada. En la realización de la Fig. 14a, el decodificador espectral 701 comprende un bloque de configuración de ruido inverso 703, un decodificador de llenado de intervalos opcional 704, un bloque de síntesis de TNS/TTS 705 y el bloque de IMDCT 702 anteriormente mencionados. Estos bloques deshacen las operaciones específicas llevadas a cabo por los bloques 602 a 606b. En particular, un bloque de configuración de ruidos 703 deshace la configuración de ruido llevada a cabo por el bloque 606a sobre la base de los coeficientes de LPC cuantificados 1010. El decodificador IGF 704 opera como se expuso con respecto a la Fig. 2a, los bloques 202 y 206 y el bloque de síntesis TNS/TTS 705 opera como se expuso en el contexto del bloque 210 de la Fig. 2a y el decodificador espectral comprende adicionalmente el bloque de IMDCT 702. Adicionalmente o como alternativa, el procesador cruzado 700 en la Fig. 14a comprende una etapa de retardo 707 para hacer llegar una versión retardada de la versión decodificada obtenida por el decodificador espectral 701 en una etapa de desénfasis 617 del segundo procesador de codificación con el fin de inicializar la etapa de desénfasis 617.

40 **[0080]** Además, como alternativa o adicionalmente, el procesador cruzado 700 puede comprender una etapa de filtrado de análisis de los coeficientes de predicción, ponderada, 708 para filtrar la versión decodificada y para hacer llegar una versión decodificada filtrada a un determinador de libro de código 613 indicado como "MMSE" en la Fig. 14a del segundo procesador de codificación, para inicializar este bloque. A título adicional o como alternativa, el procesador cruzado comprende la etapa de filtrado de análisis de LPC para filtrar la versión decodificada de la primera porción de la señal codificada emitida por el decodificador 700 a una etapa de libro de código adaptativo 612

para la inicialización del bloque 612. Además o como alternativa, el procesador cruzado comprende también una etapa de pre-énfasis 709 para llevar a cabo un procesamiento de pre-énfasis a la versión decodificada emitida por un decodificador espectral 701 antes del filtrado de LPC. La etapa de pre-énfasis de salida puede ser introducida también en una etapa de retardo adicional 710 con el fin de inicializar un bloque de filtrado de síntesis de LPC 616 dentro del codificador en el dominio de los tiempos 610.

[0081] El procesador codificador en el dominio de los tiempos 610 comprende, tal como se ilustra en la Fig. 14a, un pre-énfasis que opera sobre la velocidad de muestreo ACELP inferior. Tal como se ilustra, este pre-énfasis es el pre-énfasis llevado a cabo en la etapa de procesamiento preliminar 1000, y tiene el número de referencia 1005. Los datos de pre-énfasis se introducen en una etapa de filtrado de análisis de LPC 611 que opera en el dominio de los tiempos, y este filtro es controlado por los coeficientes de LPC cuantificados 1010 obtenidos por la etapa de procesamiento preliminar 1000. Como es sabido a partir de los codificadores AMR-WB+ o USAC u otros codificadores CELP, la señal residual generada por el bloque 611 se proporciona a un libro de código adaptativo 612, y además, el libro de código adaptativo 612 está conectado a una etapa de libro de código innovador 614, y los datos del libro de código 612 y del libro de código innovador se introducen en el multiplexador de la corriente de bits, tal como se ilustra.

[0082] Por otra parte, se ha proporcionado una etapa de ganancias/codificación ACELP 615 en serie a la etapa del libro de código innovador 614, el resultado de este bloque se introduce en un determinador del libro de código 613 indicado como MMSE en la Fig. 14a. Este bloque coopera con el bloque del libro de código innovador 614. Además, el codificador en el dominio de los tiempos comprende adicionalmente una porción de decodificador que tiene un bloque de filtrado de síntesis de LPC 616, un bloque de desénfasis 617 y una etapa de post-filtrado de bajos adaptativo 618 para calcular los parámetros para un post-filtro adaptativo que, sin embargo, está aplicado en el lado del decodificador. Sin ningún post-filtrado de bajos adaptativo en el lado del decodificador, los bloques 616, 617, 618 no serían necesarios para el codificador en el dominio de los tiempos 610.

[0083] Como se ilustra, varios bloques del decodificador en el dominio de los tiempos dependen de señales previas y estos bloques son el bloque del libro de código adaptativo 612, el determinador de libro del código 613, el bloque de filtración de síntesis PC L616 y el bloque del desénfasis 617. Estos bloques están provistos con datos procedentes del procesador cruzado derivado del procesador de codificación en el dominio de las frecuencias con el fin de inicializar estos bloques para que estén listos para una conmutación instantánea desde el codificador en el dominio de las frecuencias al codificador en el dominio de los tiempos. Como puede observarse en la Fig. 14a, no es necesaria ninguna dependencia con respecto a los datos anteriores para el codificador en el dominio de las frecuencias. Por ello, el procesador cruzado 700 no proporciona ningún dato de inicialización de memoria procedente del codificador en el dominio de los tiempos para el codificador en el dominio de las frecuencias. Sin embargo, para otras implementaciones del codificador en el dominio de las frecuencias, donde existen dependencias del pasado y donde se requieren datos para la inicialización de la memoria, el procesador cruzado 700 está configurado para operar en ambas direcciones.

[0084] El decodificador de audio preferido en la Fig. 14b se describe en lo sucesivo. La parte del decodificador en forma de onda consiste en una trayectoria de decodificador TCX de banda completa donde IGF opera con la velocidad de muestreo de entrada del códec. En paralelo, existe una trayectoria de decodificador ACELP alternativo con una velocidad de muestreo más baja más inferior que se refuerza más adelante a continuación de manera descendente por un TD-BWE.

[0085] Para la inicialización de ACELP cuando se conmuta de TCX a ACELP; existe una trayectoria cruzada (consistente en el extremo frontal de decodificador de TCX compartido pero que adicionalmente proporciona salida con la velocidad de muestreo más baja y algún post-procesamiento) que lleva a cabo la inicialización ACELP inventiva. Compartir la misma velocidad de muestreo y el orden del filtro entre TCX y ACELP permite una inicialización más fácil y eficaz de ACELP:

[0086] Para visualizar la conmutación, en la Fig. 14b se han esquematizado dos conmutadores. Mientras el segundo conmutador 1160 elige de manera descendente entre las salidas TCX/IGF o ACELP/TD-BWE, el primer conmutador 1480 pre-actualiza las memorias intermedias en la etapa QMF del remuestreo de manera descendente a la trayectoria de ACELP por la salida de la trayectoria cruzada o simplemente pasa a la salida de ACELP.

[0087] Posteriormente y en el contexto de las Figuras 11a-14c se exponen implementaciones del decodificador de audio de conformidad con aspectos de la presente invención.

[0088] Un decodificador de audio para la decodificación de una señal de audio codificada 1101 comprende un primer procesador de decodificación 1120 para decodificar una primera porción de la señal de audio en un dominio de las frecuencias. El primer procesador de decodificación 1120 comprende un decodificador espectral 1122 para decodificar primeras porciones espectrales con una elevada resolución espectral y para sintetizar segundas porciones espectrales utilizando una representación paramétrica de las segundas regiones espectrales y al menos una primera región espectral decodificada con el fin de obtener una representación espectral decodificada. La

representación espectral decodificada es una representación espectral decodificada de banda completa, como se expone en el contexto de la Fig. 6, y como también se expone en el contexto de la Fig. 1a. Por ello, y en términos generales, el primer procesador de decodificación comprende una implementación de banda completa con un procedimiento del llenado de intervalos en el dominio de las frecuencias. El primer procesador de decodificación
5 1120 comprende además un convertidor de frecuencia-tiempo 1124 para convertir la representación espectral decodificada en un dominio de los tiempos con el fin de obtener una primera porción decodificada de la señal de audio.

[0089] Además, el decodificador de audio comprende un segundo procesador de decodificación 1140 para la
10 decodificación de la segunda porción de la señal de audio codificada en el dominio de los tiempos con el fin de obtener una segunda porción de audio decodificada. Además, el decodificador de audio comprende un combinador 1160 para combinar la primera porción de la señal decodificada y la segunda porción de la señal decodificada con el fin de obtener una señal de audio decodificada. Las porciones de la señal decodificada se combinan en secuencia lo que también se ilustra en la Fig. 14b mediante una implementación de conmutación 1160 que representa una
15 realización del combinador 160 de la Fig. 11a.

[0090] Es preferible que el segundo procesador de decodificación 1140 contenga un procesador de extensión de ancho de banda en el dominio de los tiempos 1220 y que comprendan, como se ilustra en la Fig. 12, un decodificador de banda baja en el dominio de los tiempos 1200 para decodificar una señal en el dominio de los
20 tiempos de banda baja. Esta implementación comprende además un muestreador ascendente 1210 para muestrear ascendentemente la señal en el dominio de los tiempos de banda baja. Adicionalmente, se proporciona un decodificador de extensión de ancho de banda en el dominio de los tiempos 1220 para sintetizar una banda alta de la señal de audio de salida. Además, se proporciona un mezclador 1230 para mezclar una banda alta sintetizada de la señal de entrada en el dominio de los tiempos y una señal en el dominio de los tiempos de banda baja
25 muestreada ascendentemente con el fin de obtener la salida del decodificador en el dominio de los tiempos. Por ello, en una realización preferida el bloque 1140 en la Fig. 11a puede ser implementado por la funcionalidad de la Fig. 12.

[0091] La Fig. 13 ilustra una realización preferida del decodificador de extensión de ancho de banda en el dominio de los tiempos 1220 de la Fig. 12. Preferiblemente, se proporciona un muestreador ascendente en el
30 dominio de los tiempos 1221 que recibe, como una entrada, una señal residual procedente de un decodificador de banda baja en el dominio de los tiempos incluido dentro del bloque 1140 y como se ilustra en 1200 en la Fig. 12 y como se ilustra además en el contexto de la Fig. 14b. El muestreador ascendente en el dominio de los tiempos 1221 genera una versión muestreada ascendentemente de la señal residual de LPC. Esta versión se introduce seguidamente en un bloque de distorsión 1222 que genera, sobre la base de señal de entrada, una señal de salida
35 que tiene valores de frecuencia más elevados. Una distorsión no lineal puede ser una operación de copia ascendente, un reflejo, un desplazamiento de frecuencia o una operación de computación no lineal o dispositivo tal como un diodo o un transistor operado en la región no lineal. La señal de salida del bloque 1222 se introduce en un bloque de filtración de síntesis de LPC 1223 que es controlado por los datos de LPC utilizados para el decodificador de banda baja así como por datos de envolvente específicos generados por el bloque de extensión de ancho de
40 banda en el dominio de los tiempos 920 en el lado del decodificador de la Fig. 14a, por ejemplo. La salida del bloque de síntesis de LPC se introduce seguidamente en un paso de banda o filtro de paso alto 1224 con el fin de obtener finalmente la banda alta, la cual se introduce a continuación en el mezclador 1230, tal como se ilustra en la Fig. 12.

[0092] Posteriormente, en el contexto de la Fig. 14b se expone una implementación preferida del
45 muestreador ascendente 1210 de la Fig. 12. El muestreador ascendente comprende preferiblemente un banco de filtros de análisis que opera con una velocidad de muestreo del decodificador de banda baja en el dominio de los tiempos. Una implementación específica de un banco de filtros de análisis de este tipo es un banco de filtros de análisis de QMF 1471 ilustrado en la Fig. 14b. Además, el muestreador ascendente comprende un banco de filtros de síntesis 1473 que opera con una velocidad de muestreo de salida que es más elevada que la primera velocidad
50 de muestreo de banda baja en el dominio de los tiempos. Por ello, el banco de filtros de síntesis de QMF 1473 que es una implementación preferida del banco de filtros general opera con la velocidad de muestreo de salida. Cuando el factor de muestreo descendente DS explicado en el contexto de la Fig. 7b es de 0,5, entonces el banco de filtros de análisis de QMF 1471 tiene, por ejemplo solamente 32 canales del banco de filtros y el banco de filtros de síntesis QMF 1473 tiene por ejemplo 64 canales QMF, pero la mitad superior de los canales del banco de filtros, es
55 decir los 32 canales superiores del banco de filtros son alimentados con ceros o ruido, mientras que los 32 canales inferiores del banco de filtros son alimentados con las correspondientes señales proporcionadas por el banco de filtros de análisis de QMF 1471. Sin embargo, es preferible que se lleve a cabo un filtrado de paso de banda 1472 en el dominio del banco de filtros de QMF con el fin de asegurar que la salida de síntesis 1473 sea una versión muestreada ascendentemente de la salida del decodificador ACELP, pero sin ningún artefacto por encima de la
60 frecuencia máxima del decodificador ACELP.

[0093] Es preferible llevar a cabo operaciones de procesamiento adicionales dentro del dominio de QMF además o en lugar del filtrado de paso de banda 1472. Si no se lleva a cabo ningún procesamiento, entonces el análisis de QMF y la síntesis de QMF constituyen un muestreador ascendente eficaz 1210.

[0094] A continuación se expone con mayor detalle el diseño de los elementos individuales de la Fig. 14b.

[0095] El decodificador en el dominio de las frecuencias de banda completa 1120 comprende un primer bloque de decodificación 1222a para decodificar los coeficientes espectrales de alta resolución y para llevar a cabo adicionalmente el llenado de los ruidos en la porción de banda baja, como es conocido por ejemplo de la tecnología USAC. Además, el decodificador de banda completa comprende un procesador 1122b para llenar los huecos espectrales utilizando valores espectrales sintetizados que han sido codificados solamente paramétricamente, y por ello, codificados con una baja resolución en el lado del codificador. Seguidamente, en el bloque 1222c, se lleva a cabo una configuración inversa del ruido y el resultado se introduce en un bloque de síntesis TNS/TTS 705 que proporciona, como una salida final, una entrada al convertidor de frecuencia-tiempo 1124, que está preferiblemente implementado como una transformada de coseno discreta modificada inversa que opera con la velocidad de la salida, es decir, la velocidad de muestreo elevado.

[0096] Por otra parte, se utiliza un post-filtro armónico o de LTP que es controlado por los datos obtenidos por el bloque de extracción de parámetros TCX-LTP 1006 en la Fig. 14a. El resultado es entonces la primera porción de señal de audio decodificada con la velocidad de muestreo de salida como puede observarse en la Fig. 14b, teniendo dichos datos una elevada velocidad de muestreo, y por ello, no es necesario ningún refuerzo adicional de la frecuencia debido al hecho de que el procesador de decodificación es un decodificador de banda completa en el dominio de las frecuencias que utiliza preferiblemente la tecnología inteligente del llenado de los huecos expuesta en el contexto de las Figuras 1a-5C.

[0097] Varios elementos en la Fig. 14b son bastante similares a los correspondientes bloques en el procesador cruzado 700 de la Fig. 14a, particularmente con respecto al decodificador de IGF 704 correspondiente al procesamiento de IGF 1122 y la operación de configuración inversa del ruido controlada por los coeficientes de LPC cuantificados LPC 1145 corresponde a la configuración inversa del ruido 703 de la Fig. 14a y el bloque de síntesis de TNS/TTS 705 en la Fig. 14b corresponde al bloque de TNS/TTS 705 en la Fig. 14a. Sin embargo, lo importante es observar que el bloque de IMDCT 1124 en la Fig. 14b opera con la misma velocidad de muestreo elevada mientras que el bloque de IMDCT 702 en la Fig. 14a opera con una velocidad de muestreo baja. Por ello, el bloque 1124 en la Fig. 14b comprende la transformada de gran tamaño y el bloque de desplegado 710, la ventana de síntesis en el bloque 712 y la etapa de superposición-adición 714 con el correspondiente gran número de operaciones, gran número de coeficientes de ventana y un gran tamaño de transformada en comparación con las correspondientes características 720,72, 724 en la Fig. 7b, que son operadas en el bloque 701, como se destacará más adelante en esta invención, en el bloque 1171 del procesador cruzado 1170 en la Fig. 14b también.

[0098] Es preferible que el procesador de decodificación en el dominio de los tiempos 1140 comprenda el ACELP o que el decodificador de banda baja en el dominio de los tiempos 1200 comprenda una etapa de decodificador de ACELP 1149 para obtener ganancias decodificadas y la información innovadora del libro del código. Adicionalmente se proporciona una etapa de libro del código adaptativo de ACELP 1141 y una subsiguiente etapa de post-procesamiento de ACELP 1142 y un filtro de síntesis final tal como un filtro de síntesis LPC 1143, el que, de nuevo, es controlado por los coeficientes cuantificados 1145 obtenidos desde el demultiplexador de corriente de bits 1100 correspondiente al parser (analizador sintáctico) de señales codificadas 1100 en la Fig. 11a. La salida del filtro de síntesis de LPC 1143 se introduce en una etapa de desénfasis 1144 para cancelar o deshacer el procesamiento introducido por la etapa de pre-énfasis 1005 del pre-procesador 1000 de la Fig. 14a. El resultado es la señal de salida en el dominio de los tiempos con una baja velocidad de muestreo y una baja banda y en caso de que se requiera una salida en el dominio de frecuencias, el conmutador 1480 se encuentra en la posición indicada y la salida de la etapa de desénfasis 1144 se introduce en el muestreador ascendente 1210 y seguidamente se mezcla con las bandas elevadas procedentes del decodificador de extensión de ancho de banda en el dominio de los tiempos 1220.

[0099] De acuerdo con las realizaciones de la presente invención, el decodificador de audio comprende adicionalmente el procesador cruzado 1170 ilustrado en la Fig. 11b y en la Fig. 14b para calcular, a partir de la representación espectral decodificada de la primera porción de audio codificada, datos de inicialización del segundo procesador de decodificación de tal manera que el segundo procesador de decodificación se inicialice con el fin de decodificar la segunda porción de audio codificada que sigue en el tiempo a la primera porción de audio en la señal de audio codificada, es decir, de tal manera que el procesador de decodificación en el dominio de los tiempos 1140 esté listo para una conmutación instantánea desde una porción de señal de audio a la siguiente sin ninguna pérdida en calidad ni en eficacia.

[0100] Es preferible que el procesador cruzado 1170 comprenda un convertidor adicional de frecuencia-tiempo 1171 que opere con una velocidad de muestreo más baja que el convertidor de frecuencia del primer procesador de decodificación con el fin de obtener otra primera porción decodificada en el dominio de los tiempos que se va a utilizar como la señal de inicialización o para el que sea posible derivar cualquier dato de inicialización. Es preferible que este IMDCT o convertidor de frecuencia-tiempo de baja velocidad de muestreo sea implementado como se ilustra en la Fig. 7b, elemento 726 (selector), elemento 720 (transformada de tamaño pequeño y desplegado), formación de ventana de síntesis con un número más pequeño de coeficientes de ventana que lo

indicado en 722 y una etapa de superposición-adición con un número más pequeño de operaciones que lo indicado en 724. Por lo tanto, el bloque de IMDCT 1124 en el decodificador de banda completa en el dominio de las frecuencias se implementa como se indica mediante los bloques 710, 712, 714, y el bloque de IMDCT 1171 se implementa como se indica en la Fig. 7b mediante los bloques 726, 720, 722, 724. De nuevo, el factor de muestra descendente es la relación entre la velocidad de muestreo del codificador en el dominio de los tiempos o la baja velocidad de muestreo y la velocidad de muestreo del codificador en el dominio de frecuencias, más elevada, o la velocidad de muestreo de salida, y este factor de muestreo descendente puede ser cualquier número superior a 0 e inferior a 1. Tal como se ilustra en la Fig. 14b, el procesador cruzado 1170 comprende además, sola o en adición a otros elementos, una etapa de retardo 1172 para retardar la primera porción de señales decodificada adicional y para introducir la primera porción de señal decodificada en una etapa de desénfasis 1144 del segundo procesador de decodificación para la inicialización. Además, el procesador cruzado comprende, como adición o como alternativa, un filtro de pre-énfasis 1173 y una etapa de retardo 1175 para filtrar y retardar una primera porción de señal decodificada adicional y para proporcionar la salida retardada del bloque 4175 en una etapa de filtrado de síntesis de LPC 1143 del decodificador de ACELP con el propósito de la inicialización.

15 **[0101]** Además, el procesador cruzado puede comprender a título alternativo o como adición a los otros elementos mencionados, un filtro de análisis de LPC 1174 para generar una señal residual de predicción procedente de la primera porción de la señal adicional o de una primera porción de la señal adicional de pre-énfasis y para introducir los datos en un sintetizador de libro de códigos de un segundo procedimiento de decodificación y, lo que es preferible, en la etapa del libro del código adaptativo 1141. Además, la salida del convertidor de frecuencia-tiempo 1171 con la baja velocidad de muestreo se introduce también en la etapa de análisis de QAMF 1471 del muestreador ascendente 1210 con el fin de la inicialización, es decir, cuando la porción de la señal de audio de código actualmente decodificada se entrega al decodificador de banda completa en el dominio de las frecuencias 1120.

25 **[0102]** El decodificador de audio preferido se describe en lo sucesivo. La parte del decodificador en forma de onda consiste en una trayectoria de decodificador TCX de banda completa con un IGF y ambos operan con la velocidad de muestreo introducida del códec. En paralelo, existe una trayectoria de decodificador ACELP alternativa con una velocidad de muestreo que es baja y que es reforzada además de manera descendente por un TD-BWE.

30 **[0103]** Para la inicialización de ACELP cuando se conmute de TCX a ACELP, existe una trayectoria cruzada (consistente en un extremo frontal del decodificador TCX compartido que proporciona adicionalmente una salida con la velocidad de muestreo más baja y algún post-procesamiento) que lleva a cabo la inicialización de ACELP inventiva. Compartir la misma velocidad de muestreo y orden de filtrado entre TCX y ACELP en los LPC permite una 35 inicialización más fácil y más eficaz del ACELP.

[0104] Para visualizar la conmutación, en la Fig. 14b se han esbozado dos conmutadores. Mientras el segundo conmutador 1160 elige de manera descendente entre la salida de TCX/IGF o ACELP/TD-BWE, el primer conmutador 1480 pre-actualiza las memorias intermedias en la etapa de QMF de remuestreo de manera descendente de la trayectoria de ACELP mediante la salida de la trayectoria cruzada o simplemente pasa a la salida ACELP.

45 **[0105]** Para resumir, los aspectos preferidos se refieren a una combinación de un codificador ACELP y TD-BWE con una tecnología TCX/IGF capaz de banda completa preferiblemente asociada con la utilización de una señal cruzada.

[0106] Una característica específica adicional se refiere a una trayectoria de señal cruzada para la inicialización de ACELP con el fin de habilitar una conmutación continua.

50 **[0107]** Un aspecto adicional es que se introduce un IMDCT corto con una parte inferior de coeficientes de MDCT largos de elevada velocidad con el fin de implementar de manera eficaz una conversión de la velocidad de bits de muestras en la trayectoria cruzada.

55 **[0108]** Una característica adicional se refiere a una realización eficaz de una trayectoria cruzada parcialmente compartida con una banda completa de TCX/IGF en el decodificador.

[0109] Una característica adicional es la trayectoria de señal cruzada para la inicialización de QMF con el fin de permitir la conmutación continua de TCX a ACELP.

60 **[0110]** Una característica adicional es una trayectoria de señal cruzada para el QMF que permite compensar el intervalo de retardo entre la salida remuestreada de ACELP y una salida de banco de filtros-TCX/IGF cuando se conmute de ACELP a TCX.

[0111] Un aspecto adicional es que se proporciona un LPC tanto para el codificador de TCX como de ACELP 65 con la misma velocidad de muestreo de muestras y orden de filtro, aunque el codificador/decodificador TCX/IGF es

capaz de operar con banda completa.

[0112] Posteriormente, en la Fig. 14c se expone una implementación preferida de un decodificador en el dominio de los tiempos que opera como un decodificador autónomo o en combinación con el decodificador en el dominio de las frecuencias capaz de operar con banda completa.

[0113] En términos generales, el decodificador en el dominio de los tiempos comprende un decodificador ACELP, un remuestreador o muestreador ascendente conectados posteriormente y una funcionalidad de extensión de ancho de banda en el dominio de los tiempos. En particular, el decodificador ACELP comprende una etapa de decodificación ACELP para restaurar las ganancias y el libro del código innovador 1149, una etapa de libro de código ACELP-adaptativo 1141, un post-procesador ACELP 1142, un filtro de síntesis de LPC 1143 controlado por coeficientes LPC cuantificados procedentes de un multiplexador de corriente de bits o parser (analizador sintáctico) de señales codificadas y la etapa de desénfasis posteriormente conectada, 1124. Es preferible que la señal del dominio de los tiempos decodificada con una velocidad de muestreo ACELP sea introducida, junto con datos de control procedentes de la corriente de bits, en un decodificador de extensión de ancho de banda en el dominio de los tiempos 1220, que proporcione una banda alta en las salidas.

[0114] Con el fin de muestrear ascendentemente la salida de desénfasis 1144, se proporciona un muestreador ascendente que comprende el bloque de análisis 1471, y el bloque de síntesis 1473. Dentro del dominio del banco de filtros definido por los bloques 1471 y 1473, es preferible que se haya aplicado un filtro de paso de banda. En particular, y como ya se ha expuesto anteriormente, también es posible utilizar las mismas funcionalidades que se han expuesto con respecto a los mismos números de referencia. Además, el decodificador de extensión de ancho de banda en el dominio de los tiempos 1220 se puede implementar como se ilustra en la Fig. 13, y en términos generales, comprende un muestreo ascendente de la señal residual ACELP o de la señal residual en el dominio de los tiempos con la velocidad de muestreo de ACELP finalmente con una velocidad de muestreo de salida de la señal extendida de ancho de banda.

[0115] Posteriormente, y haciendo referencia a las Figs. 1A-5C se exponen más detalles con respecto al codificador y decodificador en el dominio de las frecuencias.

[0116] En la Fig. 1a se ilustra un aparato para codificar una señal de audio 99. La señal de audio 99 se introduce en un convertidor en el espectro de los tiempos 100 con el fin de convertir una señal de audio que tiene una velocidad de muestreo en una representación espectral 101 emitida por el convertidor en el espectro de los tiempos. El espectro 101 se introduce en un analizador espectral 102 para analizar la representación espectral 101. El analizador espectral 101 está configurado para determinar un primer conjunto de primeras porciones espectrales 103 que van a ser codificadas con una primera resolución espectral y un segundo conjunto, diferente, de porciones espectrales 105 que van a ser codificadas con una segunda resolución espectral. La segunda resolución espectral es más pequeña que la primera resolución espectral. El segundo conjunto de segundas porciones espectrales 105 se introduce en un calculador de parámetros o codificador paramétrico 104 para calcular información de envolvente espectral que tiene la segunda resolución espectral. Además, se proporciona un codificador de audio en el dominio espectral 106 para generar una primera representación codificada 107 del primer conjunto de primeras porciones espectrales que tienen la primera resolución espectral. Además, el calculador de parámetros/codificador paramétrico 104 está configurado para generar una segunda representación codificada 109 del segundo conjunto de segundas porciones espectrales. La primera representación codificada 107 y la segunda representación codificada 109 se introducen en un multiplexador de corriente de bits o formador de corriente de bits 108, y el bloque 108 emite finalmente la señal de audio codificada para su transmisión o para su almacenamiento en un dispositivo de almacenamiento.

[0117] Típicamente, una primera porción espectral tal como 306 de la Fig. 3a estará rodeada por dos porciones espectrales tales como 307a, 300b. Éste no es el caso de por ejemplo, HE-AAC, donde el intervalo de frecuencias del codificador núcleo presenta un ancho de banda limitado.

[0118] La Fig. 1b ilustra un decodificador que coincide con el codificador de la Fig. 1a. La primera representación codificada 107 se introduce en un decodificador de audio en el dominio espectral 112 para generar una primera representación decodificada de un primer conjunto de primeras porciones espectrales, teniendo la representación decodificada una primera resolución espectral. Además, la segunda representación codificada 109 se introduce en un decodificador paramétrico 114 con el fin de generar una segunda representación decodificada de un segundo conjunto de segundas porciones espectrales que tiene una segunda resolución espectral que es inferior a la primera resolución espectral.

[0119] El decodificador comprende además un regenerador de frecuencias 116 para regenerar una segunda porción espectral reconstruida que tiene la primera resolución espectral que utiliza una primera porción espectral. El regenerador de frecuencias 116 lleva a cabo una operación de llenado de mosaicos, es decir utiliza un mosaico o porción del primer conjunto de primeras porciones espectrales y copia este primer conjunto de primeras porciones espectrales en el intervalo de reconstrucción o banda de reconstrucción que tiene la segunda porción espectral y

típicamente lleva a cabo una configuración de envolvente espectral u otra operación indicada por la segunda representación decodificada emitida por el decodificador paramétrico 114, es decir, mediante la utilización de la información del segundo conjunto de segundas porciones espectrales. El primer conjunto decodificado de primeras porciones espectrales y el segundo conjunto de porciones espectrales, reconstruido, indicado como salida del regenerador de frecuencias 116 en la línea 117 se introduce en un convertidor de espectro-tiempo 118 configurado para convertir la primera configuración decodificada y la segunda porción espectral reconstruida en una representación en el tiempo 119, teniendo la representación en el tiempo una determinada velocidad de muestreo elevada.

10 **[0120]** En la Fig. 2b se ilustra una implementación del codificador de la Fig. 1a. Una señal de entrada de audio 99 se introduce en un banco de filtros de análisis 220 correspondiente al convertidor de tiempo espectro 100 de la Fig. 1a. Seguidamente, en el bloque de TNS 222 se lleva a cabo una operación de configuración temporal del ruido. Seguidamente, la entrada en el analizador espectral 102 de la Fig. 1a correspondiente a una máscara tonal de bloque 226 de la Fig. 2b puede consistir en valores espectrales completos, cuando no se aplica la configuración temporal de ruido/operación de configuración temporal de mosaicos, o puede tratarse de valores residuales espectrales, cuando se aplica la operación de TNS ilustrada en la Fig. 2b, bloque 222. Para las señales de dos canales o para las señales de múltiples canales, es posible llevar a cabo adicionalmente una codificación de canales conjuntos 228, de tal manera que el codificador en el dominio espectral 106 de la Fig. 1a pueda comprender el bloque de codificación de canales conjuntos 228. Además, se proporciona un codificador de entropía 232 para llevar a cabo una compresión sin pérdida de los datos que también es una porción del codificador en el dominio espectral 106 de la Fig. 1a.

25 **[0121]** El analizador espectral/máscara tonal 226 separa la salida del bloque TNS 122 en la banda de núcleo y en los componentes tonales correspondientes al primer conjunto de primeras porciones espectrales 103 y los componentes residuales correspondientes al segundo conjunto de segundas porciones espectrales 105 de la Fig. 1a. El bloque 224 indicado como la codificación de la extracción de los parámetros de IGF se corresponde al codificador paramétrico 104 de la Fig. 1a, y el multiplexador de la corriente de bits 230 corresponde al multiplexador de la corriente de bits 108 de la Fig. 1a.

30 **[0122]** Es preferible que el banco de filtros de análisis 222 esté implementado como un MDCT (modified discrete cosine transform filterbank, banco de filtros de transformada de coseno discreta modificada) y que se utilice el MDCT para transformar la señal 99 en el dominio de los tiempos-frecuencia, donde la transformada de coseno discreta modificada actúa como herramienta de análisis de las frecuencias.

35 **[0123]** Es preferible que el analizador espectral 226 aplique una máscara de tonalidad. La etapa de la estimación de la máscara de tonalidad se utiliza para separar los componentes tonales de los componentes similares a ruido presentes en la señal. Esto permite que el codificador de núcleo 228 codifique todos los componentes tonales con un módulo psico-acústico.

40 **[0124]** Este procedimiento presenta ciertas ventajas con respecto al SBR [1] clásico en que la red de armónicos de una señal de múltiples tonos es conservada por el codificador de núcleo, mientras que solamente los intervalos entre las sinusoides son rellenados con el "ruido configurado" de mejor concordancia procedente de la región de fuente.

45 **[0125]** En el caso de los pares de canales estéreo se aplica un procesamiento de estéreo conjunto adicional. Esto es necesario, porque para un determinado intervalo de destino la señal puede ser una fuente de sonido paneada sumamente correlacionada. En el caso en que las regiones de fuente elegidas para esta región particular no estén bien correlacionadas entre sí, aunque las energías coincidan con las regiones de destino, la imagen espacial puede sufrir debido a las regiones de fuente no correlacionadas. El codificador analiza cada banda de energía de la región de destino, llevando a cabo típicamente una correlación cruzada de los valores espectrales y si se supera un determinado umbral, establece un indicador conjunto para esta banda de energía. En el decodificador las bandas de energía de canal izquierda y derecha son tratadas individualmente si no se establece este indicador de estéreo conjunto. En el caso en que se establezca el indicador estéreo conjunto, tanto las energías como el parcheado se efectúan en el dominio estéreo conjunto. La información exterior conjunta para las regiones IGF se señala de manera similar a la información estéreo conjunta para la codificación del núcleo, lo que incluye un indicador que indica en el caso de la predicción si la dirección de la predicción es desde el mezclado descendente a residual o viceversa.

60 **[0126]** Las energías pueden calcularse a partir de las energías transmitidas en el dominio L/R.

$$\begin{aligned} Nrg_{medio}[k] &= Nrg_{izquierdo}[k] + Nrg_{derecho}[k]; \\ Nrg_{lateral}[k] &= Nrg_{izquierdo}[k] - Nrg_{derecho}[k]; \end{aligned}$$

siendo k el índice de frecuencia en el dominio de las transformadas.

[0127] Otra solución consiste en calcular y transmitir las energías directamente en el dominio estéreo conjunto para bandas donde el estéreo conjunto está activo, por lo que no se necesita información de energía adicional en el lado del decodificador.

5

[0128] Los mosaicos de fuente se crean siempre según la matriz central/lateral (Mid/Side-Matrix):

$$\begin{aligned} \text{mosaicomedio}[k] &= 0,5 \cdot (\text{mosaicoizquierdo}[k] + \text{mosaicoderecho}[k]) \\ \text{mosaicolateral}[k] &= 0,5 \cdot (\text{mosaicoizquierdo}[k] - \text{mosaicoderecho}[k]) \end{aligned}$$

10

[0129] Ajuste de la energía:

$$\begin{aligned} \text{mosaicomedio}[k] &= \text{mosaicomedio}[k] \cdot \text{Nrgmedio}[k] \\ \text{mosaicolateral}[k] &= \text{mosaicolateral}[k] \cdot \text{Nrglateral}[k] \end{aligned}$$

15

[0130] Estéreo conjunto → transformación LR:
Si no hay parámetro de predicción adicional codificado:

$$\begin{aligned} \text{mosaicoizquierdo}[k] &= \text{mosaicomedio}[k] + \text{mosaicolateral}[k] \\ \text{mosaicoderecho}[k] &= \text{mosaicomedio}[k] - \text{mosaicolateral}[k] \end{aligned}$$

20

[0131] Si se codifica un parámetro de predicción adicional y si la dirección señalada es de centro a lado:

$$\begin{aligned} \text{mosaicolateral}[k] &= \text{mosaicolateral}[k] - \text{coefpredicción} \cdot \text{mosaicomedio}[k] \\ \text{mosaicoizquierdo}[k] &= \text{mosaicomedio}[k] + \text{mosaicolateral}[k] \\ \text{mosaicoderecho}[k] &= \text{mosaicomedio}[k] - \text{mosaicolateral}[k] \end{aligned}$$

25

[0132] Si la dirección señalada es de lado a centro:

$$\begin{aligned} \text{mosaicomedio}[k] &= \text{mosaicomedio}[k] - \text{coefpredicción} \cdot \text{mosaicomedio}[k] \\ \text{mosaicoizquierdo}[k] &= \text{mosaicomedio}[k] - \text{mosaicolateral}[k] \\ \text{mosaicoderecho}[k] &= \text{mosaicomedio}[k] + \text{mosaicolateral}[k] \end{aligned}$$

30

[0133] Este procesamiento asegura que desde los mosaicos utilizados para generar regiones de destino altamente correlacionadas y regiones de destino paneadas, los canales izquierdo y derecho siguen representando una fuente de sonido correlacionada y paneada incluso si las regiones de fuente no están correlacionadas, conservándose la imagen estéreo para tales regiones.

[0134] En otras palabras, en la corriente de bits, se transmiten indicadores de estéreo conjuntos que indican si se utilizarán L/R o M/S como un ejemplo para la codificación estéreo conjunta general. En el decodificador, primero, la señal de núcleo se decodifica como se indica mediante los indicadores estéreo conjuntos para las bandas núcleo. En segundo lugar, la señal de núcleo se almacena tanto en la representación L/R como M/S. Para el llenado de mosaico IGF, se elige la presentación de mosaico fuente de manera que se ajuste a la representación de mosaico objetivo como se indica mediante la información estéreo conjunta para las bandas IGF.

[0135] La configuración temporal del ruido (TNS, Temporal Noise Shaping) es una técnica estándar y parte del AAC. El TNS puede ser considerado como una ampliación del esquema básico de un codificador perceptual, al insertar una etapa de procesamiento opcional entre el banco de filtros y la etapa de la cuantificación. La tarea principal del módulo de TNS consiste en ocultar el ruido de cuantificación producido en la región de enmascaramiento temporal de señales similares a transitorias y por lo tanto conduce a un esquema de codificación más eficaz. En primer lugar, el TNS calcula un conjunto de coeficientes de predicción mediante la utilización de la "predicción directa" en el dominio de las transformadas, por ejemplo, MDCT. Estos coeficientes se utilizan seguidamente para aplanar la envolvente temporal de la señal. Dado que la cuantificación afecta al espectro filtrado del TNS, también el ruido de cuantificación es temporalmente plano. Mediante la aplicación del filtrado de TNS inverso en el lado del decodificador, el ruido de la cuantificación se configura según la envolvente temporal del filtro de TNS y por ello el ruido de cuantificación llega a ser enmascarado por los transitorios.

[0136] El IGF se basa en una representación de MDCT. Para una codificación eficaz, deben utilizarse preferiblemente bloques largos de aproximadamente 20 ms. Si la señal con un bloque de esta longitud contiene transitorios, se presentan pre- y post-ecos audibles en las bandas espectrales de IGF debido al llenado de mosaicos.

[0137] El efecto de pre-eco se reduce utilizando TNS en el contexto IGF. En este caso, el TNS se utiliza como una herramienta de configuración temporal de mosaicos (TTS, temporal tile shaping) ya que la regeneración espectral en el decodificador se lleva a cabo en la señal residual de TNS. Los coeficientes de predicción de TTS

65

requeridos se calculan y aplican utilizando el espectro completo en el lado del codificador, como es usual. Las frecuencias de inicio y parada de TNS/TTS no se ven afectadas por la frecuencia de inicio de IGF $f_{IGFstart}$ de la herramienta de IGF. En comparación con el TNS de legado, la frecuencia de parada de TTS se incrementa a la frecuencia de parada de la herramienta IGF, que es más elevada que $f_{IGFstart}$. En el lado del decodificador los coeficientes de TNS/TTS se aplican de nuevo al espectro completo, es decir, el espectro del núcleo más el espectro regenerado más los componentes tonales de la máscara de tonalidad. La aplicación del TTS es necesaria para formar la envolvente temporal del espectro regenerada para que concuerde de nuevo con la envolvente de la señal original.

10 **[0138]** En los decodificadores de legado, el parcheado espectral en una señal de audio corrompe la correlación espectral en los bordes del parche y con ello deteriora la envolvente temporal de la señal de audio por el hecho de introducir una dispersión. Por lo tanto, otro beneficio de la realización del llenado de los mosaicos de IGF en las señales residuales es que, después de la aplicación del filtro de configuración, los bordes de mosaico son correlacionados de manera fluida, resultando una reproducción temporal más fiel de la señal.

15

[0139] En un codificador de IGF, el espectro que ha experimentado un filtrado de TNS/TFF, el procesamiento de la máscara de tonalidad y la estimación de los parámetros de IGF, carece de cualquier señal por encima de la frecuencia de inicio excepto por los componentes tonales. Este espectro escaso es ahora codificado por el codificador de núcleo utilizando principios de codificación aritmética y de codificación predictiva. Estos componentes codificados, junto con los bits de señalización, forman la corriente de bits del audio.

20

[0140] En la Fig. 2a se ilustra la correspondiente implementación del decodificador. La corriente de bits en la Fig. 2a correspondiente a la señal de audio codificada se introduce en el demultiplexador/decodificador que estaría conectado, con respecto a la Fig. 1b, a los bloques 112 y 114. El demultiplexador de la corriente de bits separa la señal de audio introducida en la primera representación codificada 107 de la Fig. 1b y la segunda representación codificada 109 de la Fig. 1b. La primera representación codificada que tiene el primer conjunto de primeras porciones espectrales se introduce en el bloque de decodificación de canales conjuntos 204 correspondiente al decodificador en el dominio espectral 112 de la Fig. 1b. La segunda representación codificada se introduce en el decodificador paramétrico 114 no representado en la Fig. 2a y seguidamente se introduce en el bloque de IGF 202 correspondiente al generador de frecuencia 116 de la Fig. 1b. El primer conjunto de porciones espectrales requerido para la regeneración de las frecuencias se introduce en el bloque 202 por medio de la línea 203. Además, posteriormente a la decodificación de los canales conjuntos 204 se aplica la decodificación de núcleo específica al bloque de máscara tonal 206 de tal manera que la salida de la máscara tonal 206 corresponda a la salida del decodificador en el dominio espectral 112. A continuación, se lleva a cabo una combinación a cargo del combinador 208, es decir, una construcción de trama donde la salida del combinador 208 tiene ahora el espectro de intervalo completo, que se encuentra todavía en el dominio filtrado de TNS/TTS. A continuación, en el bloque 210 se lleva a cabo una operación de TNS/TTS inversa utilizando la información de filtro de TNS/TTF proporcionada por medio de la línea 109, es decir, la información del lado del TTS está preferiblemente incluida en la primera representación codificada generada por el codificador en el dominio espectral 106 que puede, por ejemplo, ser un codificador de núcleo directo AAC o USAC, o también puede estar incluido en la segunda representación codificada. A la salida del bloque 210, se proporciona un espectro completo hasta la frecuencia máxima que es la frecuencia de intervalo completo definida por la velocidad de muestreo de la señal de entrada original. Seguidamente se lleva a cabo una conversión de espectro/tiempo en el banco de filtros de síntesis 212 de manera que finalmente se obtenga la señal de salida de audio.

45

[0141] En la Fig. 3a se ilustra una representación esquemática del espectro. El espectro está subdividido en bandas de factor de escala SCB donde hay siete bandas de factor de escala SCB1 a SCB7 en el ejemplo ilustrado de la Fig. 3a. Las bandas de factor de escala pueden ser bandas de factor de escala AAC que están definidas en el estándar AAC y que tienen un ancho de banda creciente hasta frecuencias superiores, tal como se ilustra esquemáticamente en la Fig. 3a. Es preferible llevar a cabo el llenado de los intervalos no desde el mismo inicio del espectro, es decir en frecuencias bajas, sino empezar la operación con una frecuencia de inicio ilustrada en 309. Por ello, la banda de frecuencia núcleo se extiende desde la frecuencia más baja a la frecuencia de inicio de IGF. Por encima de la frecuencia de inicio de IGF, se aplica el análisis de espectro para separar los componentes espectrales de elevada resolución 304, 305, 306, 307 (el primer conjunto de primeras porciones espectrales) de los componentes de baja resolución representados por el segundo conjunto de segundas porciones espectrales. La Fig. 3a ilustra un espectro que es un ejemplo de introducción en el codificador del dominio espectral 106 o del codificador de canales conjuntos 228, es decir, el codificador del núcleo opera en el intervalo completo, pero codifica una cantidad significativa de valores espectrales cero, es decir, estos valores espectrales cero se cuantifican en cero o se establecen en cero antes de la cuantificación o posteriormente a la cuantificación. Como alternativa, el codificador de núcleo opera en el intervalo completo, es decir, como si el espectro estuviese como se ilustra, es decir, el decodificador de núcleo no tiene que ser necesariamente consciente de ningún llenado de intervalo o codificación del segundo conjunto de segundas porciones espectrales con una resolución espectral más baja.

60

[0142] Es preferible que la elevada resolución esté definida por una codificación de líneas espectrales tales como las líneas de MDCT, mientras que la segunda resolución o baja resolución se decide, por ejemplo, calculando solamente un único valor espectral por banda de factor de escala, donde una banda de factor de escala abarca varias líneas de frecuencia. Por lo tanto, la segunda resolución baja es, con respecto a su resolución espectral, mucho más baja que la primera resolución, o resolución elevada, definida por la codificación de las líneas, típicamente aplicada por el codificador de núcleo tal como un codificador de núcleo AAC o USAC.

[0143] En cuanto al factor de escala o cálculo de energía, la situación se ilustra en la Fig. 3b. Debido al hecho de que el codificador es un codificador de núcleo y debido al hecho de que puede haber, pero no necesariamente, componentes del primer conjunto de porciones espectrales en cada banda, el codificador de núcleo calcula un factor de escala para cada banda no solamente en el intervalo del núcleo por debajo de la frecuencia de inicio IGF 309, sino también por encima de la frecuencia de inicio IGF hasta la máxima frecuencia $f_{IGFstop}$ que es más pequeña que, o igual a, la mitad de la frecuencia de muestreo, es decir $f_{s/2}$. Por lo tanto, las porciones tonales codificadas 302, 304, 305, 306, 307 de la Fig. 3a, y en esta realización junto con los factores de escala SCB1 a SCB7 corresponden a los datos de elevada resolución espectral. Los datos espectrales de baja resolución se calculan partiendo de la frecuencia de inicio de IGF y corresponden a los valores de información de energía E_1, E_2, E_3, E_4 , que son transmitidos conjuntamente con los factores de escala SF4 a SF7.

[0144] En particular, cuando el codificador de núcleo se encuentra en una condición de baja velocidad de bits, una operación de llenado del ruido adicional en la banda de núcleo, es decir, inferior en frecuencia a la frecuencia de inicio de IGF, es decir, en las bandas de factor de escala SCB1 a SCB3, se puede aplicar de forma adicional. En el llenado del ruido, existen varias líneas espectrales adyacentes que deben ser cuantificadas en cero. En el lado del decodificador, estos valores espectrales cuantificados en cero son resintetizados y los valores espectrales resintetizados son ajustados a sus tamaños utilizando una energía del llenado de ruido tal como NF₂ ilustrado en 308 en la Fig. 3b. La energía del llenado de ruido, que puede darse en términos absolutos o en términos nativos particularmente con respecto al factor de escala como en USAC, se corresponde a la energía del conjunto de valores espectrales cuantificados en cero. Estas líneas espectrales del llenado de ruido también pueden ser consideradas como un tercer conjunto de terceras porciones espectrales que son regeneradas mediante síntesis directa del llenado de ruido sin ninguna operación de IGF basada en la regeneración de frecuencias donde se utilizan mosaicos de frecuencia procedentes de otras frecuencias para reconstruir mosaicos de frecuencia utilizando valores espectrales tomados de un intervalo de fuente y la información de energía E_1, E_2, E_3, E_4 .

[0145] Es preferible que las bandas para las que se calcula la información de energía coincidan con las bandas de factor de escala. En otras realizaciones, se aplica un agrupamiento del valor de información de energía de tal manera que, por ejemplo, para las bandas de factor de escala 4 y 5, se transmita solamente un único valor de información de energía, pero incluso en esta realización, los límites de las bandas de reconstrucción agrupadas coinciden con los límites de las bandas de los factores de escala. Si se aplican diferentes separaciones de banda, en tal caso es posible aplicar determinados recálculos o cálculos de sincronización, y esto puede tener sentido en función de la determinada implementación.

[0146] Es preferible que el codificador en el dominio espectral 106 de la Fig. 1a sea un codificador accionado psico-acústicamente tal como se ilustra en la Fig. 4a. Típicamente, como por ejemplo en el caso ilustrado en el estándar MPEG2/4 AAC o en el MPEG1/2, estándar de Capa 3, la señal de audio que debe ser codificada después de haber sido transformada en el intervalo espectral (401 en la Fig. 4a) se envía a un calculador de factores de escala 400. El calculador de factores de escala es controlado por un modelo psico-acústico que recibe adicionalmente la señal de audio que se va a cuantificar o que recibe, como en el estándar MPEG1/2 o MPEG AAC, una representación espectral compleja de la señal de audio. El modelo psico-acústico calcula, para cada banda de factores de escala, un factor de escala que representa el umbral psico-acústico. Adicionalmente, los factores de escala son a continuación, mediante la cooperación de los bucles de iteración interiores y exteriores bien conocidos o mediante cualquier otro procedimiento de codificación adecuado, ajustados de tal manera que se satisfagan determinadas condiciones de velocidad de bits. A continuación, los valores espectrales que deben ser cuantificados por una parte y los factores de escala calculados por otra parte, son introducidos en un procesador cuantificador 404. En la operación de codificador de audio directa, los valores espectrales que deben ser cuantificados son ponderados por los factores de escala, y, los valores espectrales ponderados se introducen seguidamente en un cuantificador fijo que tiene típicamente una función de compresión, a intervalos de amplitud superiores. Seguidamente, a la salida del procesador cuantificador existen índices de cuantificación que son enviados seguidamente a un codificador de entropía que tiene típicamente una codificación específica y muy eficaz para un conjunto de índices de cuantificación en cero para valores de frecuencia adyacentes, o, como también se llama en la especialidad, una "carrera" de valores cero.

[0147] Sin embargo, en el codificador de audio de la Fig. 1a, el procesador cuantificador recibe típicamente información acerca de las segundas porciones espectrales del analizador espectral. Por lo tanto, el procesador cuantificador 404 asegura que, en la salida del procesador cuantificador 404, las segundas porciones espectrales identificadas por el analizador espectral 102 sean cero o tengan una representación reconocida por un codificador o

decodificador como una representación cero que puede ser codificada de una manera muy eficaz, específicamente cuando existan “carreras” de valores cero en el espectro.

[0148] La Fig. 4b ilustra una implementación del procesador cuantificador. Los valores espectrales de MDCT se pueden introducir en un conjunto de bloques cero 410. Seguidamente, las segundas porciones espectrales ya se han establecido en cero antes de que se lleve a cabo una ponderación por los factores de escala. En una implementación adicional, no se proporciona un bloque 410, pero la cooperación de la puesta en cero es llevada a cabo en el bloque 418 posteriormente al bloque de ponderación 412. Y en otra implementación más, la operación de la puesta en cero también se puede llevar a cabo en un conjunto de bloques cero 422 subsiguiente a una cuantificación en el bloque cuantificador 420. En esta implementación, los bloques 410 y 418 no estarían presentes. En términos generales, se proporciona al menos uno de los bloques 410, 418 y 422 en función de la implementación específica.

[0149] A continuación, a la salida del bloque 422 se obtiene un espectro cuantificado correspondiente al ilustrado en la Fig. 3a. Este espectro cuantificado se introduce seguidamente en un codificador de entropía tal como 232 en la Fig. 2b que puede ser un codificador de Huffman o un codificador aritmético tal como por ejemplo se define en el estándar USAC.

[0150] La puesta en cero de los bloques 410, 418, 422, que se proporcionan como alternativa entre sí o en paralelo son controlados por el analizador espectral 424. Es preferible que el analizador espectral comprenda cualquier implementación de un detector de tonalidad bien conocido o que comprenda cualquier tipo diferente de detector operable para separar un espectro en componentes que debe ser codificado con una elevada resolución y componentes que van a ser codificados con una baja resolución. Otros algoritmos de este tipo implementados en el analizador espectral pueden ser un detector de la actividad de voz, un detector de ruidos, o un detector de voz o cualquier otro detector que decida, en función de información espectral o de metadatos asociados, acerca de los requisitos de resolución para diferentes porciones espectrales.

[0151] La Fig. 5a ilustra una implementación preferida del convertidor del espectro de tiempo 100 de la Fig. 1a como por ejemplo se implementa en AAC o USAC. El convertidor de espectro de tiempo comprende un dispositivo de formación de ventana 502 controlado por un detector de transitorios 504. Cuando el detector de transitorios detecta un transitorio, entonces la conmutación de ventanas largas a ventanas cortas es señalada al dispositivo de formación de ventana. El dispositivo de formación de ventana 512 calcula seguidamente, para los bloques superpuestos, tramas de ventana, donde cada trama en ventana tiene típicamente N valores tal como 2048 valores. Seguidamente se lleva a cabo una transformación dentro del transformador de bloque 506, y este transformador de bloque añade típicamente de manera adicional una pérdida, de tal manera que se lleva a cabo una pérdida/transformada combinada de manera que se obtenga una trama espectral con N valores tales como valores espectrales MDCT. Por lo tanto, para una operación de ventana larga, la trama en la entrada del bloque 506 comprende dos valores N tales como 2048 valores y una trama espectral tiene entonces 1024 valores. Sin embargo, entonces se lleva a cabo una conmutación a bloques cortos, donde se llevan a cabo ocho bloques cortos donde cada bloque corto representa $1/8$ de valores de dominio de tiempos en ventana en comparación con una ventana larga y cada bloque espectral tiene un $1/8$ de valores espectrales en comparación con un bloque largo. De esta manera, cuando se combina esta pérdida con una operación de superposición del 50 % del dispositivo de ventana, el espectro es una versión críticamente muestreada de la señal de audio en el dominio del tiempo 99.

[0152] Posteriormente, se hace referencia a la Fig. 5b donde se ilustra una implementación específica del regenerador de frecuencias 116 y del convertidor de espectro/tiempo 118 de la Fig. 1b, o de la operación combinada de los bloques 208,212 de la Fig. 2a. En la Fig. 5b, se considera una banda de reconstrucción específica tal como una banda de factores de escala 6 de la Fig. 3a. La primera porción espectral en esta banda de reconstrucción, es decir, la primera porción espectral 306 de la Fig. 3a se introduce en el bloque ajustador/constructor de trama 510. Por otra parte, se introduce también una segunda porción espectral reconstruida para la banda de factor de escala 6 en el ajustador/constructor de trama 510. Además, la información de energía tal como E_3 de la Fig. 3b para una banda de factores de escala 6 se introduce en el bloque 510. La segunda porción espectral reconstruida en la banda de reconstrucción ya ha sido generada por el llenado del mosaico de frecuencia utilizando un intervalo de fuente y la banda de reconstrucción corresponde seguidamente al intervalo objetivo. Seguidamente se lleva a cabo un ajuste de energía de la trama para obtener seguidamente por último la trama reconstruida completa que tiene los N valores como por ejemplo obtenidos a la salida del combinador 208 de la Fig. 2a. A continuación, en el bloque 512, se lleva a cabo una transformada de bloque inverso/interpolación para obtener los por ejemplo 124 valores espectrales en la entrada del bloque 512. A continuación se lleva a cabo una operación de ventana de síntesis en el bloque 514 que es controlada de nuevo por una indicación de ventana larga/ventana corta transmitida como información secundaria en la señal de audio codificado. Seguidamente, en el bloque 516 se lleva a cabo una operación de superposición-adición con una trama de tiempo previo. Es preferible que el MDCT aplique un solapamiento del 50 % de tal manera que, para cada nueva trama de tiempo de $2N$ valores, se emitan finalmente N valores en el dominio del tiempo. Se prefiere ampliamente un solapamiento del 50 % debido a que proporciona un muestreo crítico y un cruce continuo de una trama a la siguiente trama debido a la operación de superposición-adición en el bloque 516.

65

[0153] Tal como se ilustra en 301 en la Fig. 3a, es posible aplicar adicionalmente una operación de llenado de ruido no solamente por debajo de la frecuencia del inicio de IGF, sino también por debajo de la frecuencia de inicio de IGF tal como para la banda de reconstrucción considerada que coincide con la banda de factores de escala 6 de la Fig. 3a. A continuación, también es posible introducir valores espectrales del llenado de ruido en el constructor/ajustador de tramas 510 y, y también es posible aplicar valores espectrales del llenado de ruido dentro de este bloque o los valores espectrales de llenado de ruido ya pueden estar ajustados utilizando la energía de llenado del ruido antes de ser introducidos en el constructor/ajustador de tramas 510.

[0154] Es preferible que se pueda aplicar una operación de IGF, es decir una operación de llenado de mosaicos de frecuencia utilizando valores espectrales procedentes de otras porciones, en el espectro completo. De esta manera, una operación de llenado del mosaico espectral no solo se puede aplicar en la banda alta por encima de una frecuencia de inicio de IGF sino que también se puede aplicar en la banda baja. Además, el llenado de ruido sin el llenado de los mosaicos de frecuencia también se puede aplicar no solo por debajo de la frecuencia de inicio de IGF sino también por encima de la frecuencia de inicio de IGF. Sin embargo, se ha descubierto que es posible obtener una codificación de audio de alta calidad y muy eficaz cuando la operación del llenado del ruido se limita a un intervalo de frecuencia inferior a la frecuencia de inicio de IGF y cuando la operación del llenado de los mosaicos de frecuencia se restringe al intervalo de frecuencia por encima de la frecuencia de inicio de IGF, tal como se ilustra en la Fig. 3a.

[0155] Es preferible que los mosaicos objetivo (TT, target tiles) (que tienen frecuencias superiores a la frecuencia de inicio de IGF) estén vinculados a límites de bandas de factores de escala del codificador de velocidad completa. Los mosaicos de fuente (ST, source tiles), de los cuales se toma la información, es decir, para frecuencias inferiores a la frecuencia de inicio de IGF no están vinculados con límites de bandas de factores de escala. El tamaño de los ST debería corresponder al tamaño del TT asociado.

[0156] Posteriormente se hace referencia a la Fig. 5c que ilustra otra realización preferida del regenerador de frecuencias 116 de la Fig. 1b o del bloque de IGF 202 de la Fig. 2a. El bloque 522 es un generador de mosaicos de frecuencia que recibe, no solamente un ID de banda objetivo, sino también un ID de banda fuente. A título de ejemplo, se ha determinado en el lado del codificador que el factor de bandas de escala 3 de la Fig. 3a es adecuado para reconstruir la banda de factores de escala 7. Por lo tanto, el ID de la banda de fuente sería 2 y el ID de la banda teórica sería 9. Sobre la base de esta información, el generador de mosaicos de frecuencia 522 aplica una copia o lleva a cabo una operación de relleno de armónicos de los mosaicos o cualquier otra operación del llenado de mosaicos con el fin de generar la segunda porción bruta con los componentes espectrales 523. La segunda porción bruta de componentes espectrales tiene una resolución de frecuencia idéntica a la resolución de frecuencias incluida en el primer conjunto de primeras porciones espectrales.

[0157] A continuación, la primera porción espectral de la banda de reconstrucción tal como 307 de la Fig. 3a se introduce en un constructor de tramas 524 y la segunda porción bruta 523 se introduce también en el constructor de tramas 524. A continuación, la trama reconstruida es ajustada por el ajustador 526 utilizando un factor de ganancias para la banda de reconstrucción calculada por el calculador de factores de ganancia 528. Sin embargo, es importante señalar que la primera porción espectral en las tramas no se ve influenciada por el ajustador 526, sino que solamente la segunda porción bruta para la trama de reconstrucción se ve influenciada por el ajustador 526. Para esta finalidad, el calculador de factores de ganancia 528 analiza la banda de fuente o la segunda porción bruta 523 y analiza adicionalmente la primera porción espectral en la banda de reconstrucción con el fin de encontrar finalmente el factor de ganancia correcto 527 de tal manera que la energía de la trama ajustada emitida por el ajustador 526 tenga la energía E_4 cuando se contempla la banda de factores de escala 7.

[0158] Por otra parte, y tal como se ilustra en la Fig. 3a, el analizador espectral está configurado para analizar la representación espectral hasta una máxima frecuencia de análisis que es solamente una parte pequeña inferior a la mitad de la frecuencia de muestreo, y preferiblemente es de al menos un cuarto de la frecuencia de muestreo o típicamente más elevada.

[0159] Como se ilustra, el codificador opera sin muestreo descendente y el decodificador opera sin muestreo ascendente. En otras palabras, el codificador de audio en el dominio espectral está configurado para generar una representación espectral que tiene una frecuencia Nyquist definida por la velocidad de muestreo de la señal de audio originalmente introducida.

[0160] Además, como se ilustra en la Fig. 3a, el analizador espectral está configurado para analizar la representación espectral empezando con una frecuencia de inicio del llenado de intervalos y terminando con una frecuencia máxima representada por una frecuencia máxima incluida en la representación espectral, donde una porción espectral que se extiende desde una frecuencia mínima hasta la frecuencia de inicio del llenado de los intervalos forma parte del primer conjunto de porciones espectrales y donde otra porción espectral tal como 304, 305, 306, 307 que tiene valores de frecuencias superiores a la frecuencia del llenado de los intervalos se incluye adicionalmente en el primer conjunto de primeras porciones espectrales.

- [0161]** Como se ha destacado a grandes rasgos, el decodificador de audio en el dominio espectral 112 está configurado de tal manera que una frecuencia máxima representada por un valor espectral en la primera representación decodificada es igual a una máxima frecuencia incluida en la representación en tiempo que tiene una velocidad de muestreo donde el valor espectral para la frecuencia máxima en el primer conjunto de primeras 5 porciones espectrales es cero o diferente de cero. De cualquier manera, para esta frecuencia máxima en el primer conjunto de componentes espectrales existe un factor de escala para la banda de factores de escala, que es generada y transmitida independientemente de si la totalidad de los valores espectrales en esta banda de factores de escala están ajustados en cero o no, como se expuso en el contexto de las Figs. 3a y 3b.
- 10 **[0162]** Por lo tanto, el IGF es ventajoso en lo que respecta a otras técnicas paramétricas para incrementar la eficacia de la compresión, es decir la sustitución del ruido y el llenado del ruido (estas técnicas sirven exclusivamente para la representación eficaz del contenido de señales locales similares de ruido), el IGF remite una reproducción exacta de las frecuencias de los componentes tonales. En la actualidad, en ninguna técnica del estado de la técnica se aborda la representación paramétrica eficaz de un contenido arbitrario de las señales mediante el 15 llenado de los intervalos espectrales sin la restricción de una división a priori fija en banda baja (LF, low band) y banda alta (HF, high band).
- [0163]** Posteriormente se exponen y definen otras características opcionales del primer procesador codificador en el dominio de las frecuencias de banda completa y del procesador decodificador en el dominio de las 20 frecuencias de banda completa que incluyen la operación del llenado de los intervalos, que se puede implementar por separado o conjuntamente.
- [0164]** En particular, el decodificador en el dominio espectral 112 correspondiente al bloque 122a está configurado para emitir una secuencia de tramas decodificadas de valores espectrales, siendo una trama 25 decodificada la primera representación decodificada, donde la trama comprende valores espectrales para el primer conjunto de porciones espectrales e indicadores cero para las segundas porciones espectrales. El aparato para decodificar comprende además un combinador 208. Los valores espectrales son generados por un regenerador de frecuencias para el segundo conjunto de segundas porciones espectrales, donde ambos, el combinador y el generador de frecuencias, están incluidos dentro del bloque 1122b. Por lo tanto, mediante la combinación de las 30 segundas porciones espectrales y de las primeras porciones espectrales se obtiene una trama espectral reconstruida que comprende valores espectrales para el primer conjunto de las primeras porciones espectrales y el segundo conjunto de porciones espectrales, y el convertidor de espectro-tiempo 118 correspondiente al bloque de IMDCT 1124 en la Fig. 14b convierte seguidamente la trama espectral reconstruida en la representación en el tiempo.
- 35 **[0165]** Como se destaca, el convertidor de espectro-tiempo 118 o 1124 está configurado para llevar a cabo una transformada de coseno discreta modificada inversa 512, 514, y además comprende una etapa de superposición-adición 516 para solapar y añadir subsiguientes tramas en el dominio del tiempo.
- 40 **[0166]** En particular, el decodificador en el dominio de audio en dominio espectral 1222 está configurado para generar la primera representación decodificada de tal manera que la primera representación decodificada tenga una frecuencia Nyquist que define una velocidad de muestreo que es igual a una velocidad de muestreo de la representación en el tiempo generado por el convertidor de espectro-tiempo 1124.
- 45 **[0167]** Por otra parte, el decodificador 1112 o 1122a está configurado para generar la primera representación decodificada de tal manera que una primera porción espectral 306 esté colocada con respecto a la frecuencia entre dos segundas porciones espectrales 307a, 307b.
- [0168]** Una frecuencia máxima representada por un valor espectral para la máxima frecuencia en la primera 50 representación decodificada es igual a la máxima frecuencia incluida en la representación en el tiempo generada por el convertidor de espectro-tiempo, donde el valor espectral para la frecuencia máxima en la primera representación es cero o diferente de cero.
- [0169]** Además, y tal como se ilustra en la Fig. 3, la primera porción de la señal de audio codificada 55 comprende además una representación codificada de un tercer conjunto de porciones espectrales que va a ser reconstruido mediante el llenado del ruido, y el primer procesador de decodificación 1120 comprende adicionalmente un llenado de ruido incluido en el bloque 1122b para extraer información del llenado de ruido 308 de una representación codificada del tercer conjunto de terceras porciones espectrales y para aplicar una operación del llenado de ruido en el conjunto de terceras porciones espectrales sin utilizar una primera porción espectral en un 60 intervalo de frecuencia diferente.
- [0170]** Además, el decodificador de audio en el dominio espectral 112 está configurado para generar la primera representación decodificada que tiene las primeras porciones espectrales, siendo los valores de frecuencia mayores que la frecuencia que es igual a una frecuencia en el medio del intervalo de frecuencias abarcado por la 65 salida de la representación en el tiempo por el convertidor de espectro-tiempo 118 o 1124.

- [0171]** Por otra parte, el analizador espectral o el analizador de banda completa 604 está configurado para analizar la representación generada por el convertidor de tiempo-frecuencia 602 con el fin de determinar un primer conjunto de primeras porciones espectrales que deben ser codificadas con la primera resolución espectral elevada y el segundo conjunto, diferente, de segundas porciones espectrales que deben ser codificadas con una segunda resolución espectral que es inferior a la primera resolución espectral, y mediante el analizador espectral, se determina una primera porción espectral 306, con respecto a la frecuencia, entre dos porciones espectrales segundas en la Fig. 3 en 307a y 307b.
- 10 **[0172]** En particular, el analizador espectral está configurado para analizar la representación espectral hasta una máxima frecuencia de análisis que es de al menos un cuarto de una frecuencia de muestreo de la señal de audio.
- [0173]** En particular, el codificador de audio en el dominio espectral está configurado para procesar una secuencia de tramas de valores espectrales para una cuantificación y codificación de la entropía, donde, en una trama, los valores espectrales del segundo conjunto de segundas porciones se establece en cero, o donde, en la trama, los valores espectrales del primer conjunto de primeras porciones espectrales y el segundo conjunto de la segunda porción de las porciones espectrales se encuentra presente, y donde, durante el procesamiento subsiguiente, los valores espectrales en el segundo conjunto de porciones espectrales se establecen en cero, como se ilustra a título de ejemplo en 410, 418, 422.
- 15 **[0174]** El codificador de audio en el dominio espectral está configurado para generar una representación espectral que tiene una frecuencia de Nyquist definida por la velocidad de muestreo de la señal de audio introducida o por la primera porción de la señal de audio procesada por el primer procesador de codificación que opera en el dominio de las frecuencias.
- 20 **[0175]** El codificador de audio en el dominio espectral 606 está configurado además para proporcionar la primera representación codificada de tal manera que, para una trama de una señal de audio muestreada, la representación codificada comprende el primer conjunto de primeras porciones espectrales y el segundo conjunto de segundas porciones espectrales, donde los valores espectrales en el segundo conjunto de porciones espectrales son codificados como valores cero o de ruido.
- [0176]** El analizador de banda completa 604 o 102 está configurado para analizar la representación espectral empezando con la frecuencia de inicio del llenado de intervalos 209 y terminando con una frecuencia máxima f_{\max} representada por una frecuencia máxima incluida en la representación espectral y una porción espectral que se extiende desde una frecuencia mínima hasta la frecuencia de inicio del llenado de intervalos 309 forma parte del primer conjunto de primeras porciones espectrales.
- 35 **[0177]** En particular, el analizador está configurado para aplicar un procesamiento de máscara tonal en al menos una porción de la representación espectral de tal manera que los componentes tonales y los componentes no tonales estén separados entre sí, donde el primer conjunto de porciones espectrales comprende los componentes tonales y donde el segundo conjunto de las segundas porciones espectrales comprende los componentes no tonales.
- 40 **[0178]** Aunque la presente invención se ha descrito en el contexto de diagramas de bloques donde los bloques representan componentes de hardware reales o lógicos, la presente invención también se puede implementar mediante un procedimiento implementado por ordenador. En este último caso, los bloques representan etapas de procedimiento correspondientes donde dichas etapas representan las funcionalidades llevadas a cabo por los correspondientes bloques de hardware lógicos o físicos.
- 45 **[0179]** 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 a una característica de una etapa de procedimiento. De manera análoga, los aspectos descritos en el contexto de una etapa de procedimiento representan también una descripción de un bloque o elemento o aspecto correspondiente de un aparato correspondiente. Algunas de las etapas de procedimiento, o todas ellas, se pueden ejecutar mediante un aparato de hardware (o mediante el uso de) un aparato de hardware tal como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, algunas de las etapas de procedimiento más importantes, o varias de ellas, pueden ser ejecutadas mediante un aparato de este tipo.
- 50 **[0180]** La señal transmitida o codificada según la invención se puede almacenar en un medio de almacenamiento digital o se puede transmitir en un medio de transmisión tal como un medio de transmisión inalámbrica o mediante un medio de transmisión por cable tal como Internet.
- 55 **[0181]** En función de determinados requisitos de implementación, las realizaciones de la invención se pueden
- 60
- 65

implementar en hardware o en software. La implementación se puede llevar a cabo utilizando un medio de almacenamiento digital, por ejemplo, un disquete, un DVD, un Blu-Ray, un CD, una ROM, una PROM, y EPROM, una EEPROM o una memoria flash, que tiene almacenadas en sí señales de control electrónicamente legibles, que cooperan (o son capaces de cooperar) con un sistema informático programable de tal manera que se lleve a cabo el respectivo procedimiento. Por ello, el medio de almacenamiento digital puede ser legible por ordenador.

[0182] 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 informático programable, de tal manera que se lleve a cabo uno de los procedimientos descritos en esta invención.

10

[0183] En términos generales, las realizaciones de la presente invención se pueden implementar como producto de programa informático con un código de programa, siendo el código de programa operable para llevar a cabo uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede, por ejemplo, ser almacenado en un soporte legible por máquina.

15

[0184] 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.

[0185] Por lo tanto, dicho en otras palabras, una realización del procedimiento inventivo consiste en un programa informático que tiene un código de programa para llevar a cabo uno de los procedimientos descritos en esta invención, cuando el programa informático se ejecuta en un ordenador.

20

[0186] Por lo tanto, otra realización del procedimiento inventivo consiste en un soporte de datos (o en un medio de almacenamiento no transitorio tal como 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 son típicamente tangibles y/o no transitorios.

25

[0187] Una realización adicional del procedimiento inventivo es por lo tanto una corriente de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención. La corriente de datos o la secuencia de señales puede por ejemplo, estar configurada para ser transferida por medio de una conexión de datos, por ejemplo, a través de Internet.

30

[0188] Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador o un dispositivo lógico programable, configurado, o adaptado para, llevar a cabo uno de los procedimientos descritos en esta invención.

35

[0189] 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.

40

[0190] Una realización adicional según la invención comprende un aparato o sistema configurados para transferir (por ejemplo, electrónicamente u ópticamente) 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.

45

[0191] En algunas realizaciones puede utilizarse un dispositivo lógico programable (por ejemplo, FPGA (field programmable gate array, matriz de puertas programable de campo) para llevar a cabo algunas de, o todas, las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, una matriz de puertas programable de campo puede cooperar con un microprocesador con el fin de llevar a cabo uno de los procedimientos descritos en esta invención. En términos generales, los procedimientos se llevan a cabo preferentemente mediante cualquier aparato de hardware.

50

[0192] Las realizaciones descritas anteriormente son meramente ilustrativas de los principios de la presente invención. Se da por entendido que para las personas expertas en la técnica serán evidentes las modificaciones y variaciones en las disposiciones y detalles descritos en esta invención. Por lo tanto, la intención es que la invención esté limitada únicamente por los alcances de las reivindicaciones de patente pendientes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones presentadas en esta invención.

55

REIVINDICACIONES

1. Codificador de audio para la codificación de una señal de audio, que comprende una banda alta y una banda baja, comprendiendo el codificador de audio:

5 un primer procesador de codificación (600) para la codificación de una primera porción de señal de audio en un dominio de frecuencia, teniendo la primera porción de la señal de audio asociada en ella una primera velocidad de muestreo, donde el primer procesador de codificación (600) comprende:

10 un convertidor de frecuencia de tiempo (602) para convertir la primera porción de señal de audio en una representación de un dominio de frecuencia que tiene líneas espectrales hasta una frecuencia máxima de la primera porción de señal de audio, donde la frecuencia máxima de la porción de señal de audio es inferior o igual a la mitad de la primera velocidad de muestreo y al menos un cuarto de la primera velocidad de muestreo o superior;

15 un codificador espectral (606) para la codificación de la representación del dominio de frecuencia para obtener una representación espectral codificada de la primera porción de la señal de audio que es una primera porción de señal codificada;

20 un segundo procesador de codificación (610) para la codificación de una segunda porción de señal de audio en un dominio de tiempo para obtener una segunda porción de señal codificada, siendo la segunda porción de la señal de audio diferente de la primera porción de la señal de audio, donde el segundo procesador de codificación (610) comprende

25 un convertidor de velocidad de muestreo (900) para convertir la segunda porción de la señal de audio a una segunda representación de velocidad de muestreo que tiene una segunda velocidad de muestreo, siendo la segunda velocidad de muestreo inferior a la primera velocidad de muestreo, donde la segunda representación de la velocidad de muestreo no incluye la banda alta de la señal de audio; codificando un codificador de banda baja de dominio de tiempo (910) para el dominio de tiempo la segunda representación de la velocidad de muestreo; y

30 un codificador de extensión de ancho de banda de dominio de tiempo (920) para codificar paramétricamente la banda alta de la señal de audio;

35 un procesador cruzado (700) para calcular, a partir de la representación espectral codificada de la primera porción de señal de audio, datos de inicialización del segundo procesador de codificación (610), de modo que el segundo procesador de codificación (610) se inicialice para codificar la segunda porción de señal de audio inmediatamente después de la primera porción de señal de audio en tiempo en la señal de audio, donde el procesador cruzado (700) está configurado para usar una transformada en tiempo de frecuencia realizando adicionalmente un muestreo descendente desde la primera velocidad de muestreo a la segunda velocidad de muestreo seleccionando una porción de banda baja de la representación de dominio de frecuencia junto con un tamaño de transformada reducido para obtener los datos de inicialización del segundo procesador de codificación (610);

40 un controlador (620) configurado para analizar la señal de audio y para determinar, la porción de la señal de audio que es la primera porción de señal de audio codificada por el primer procesador de codificación (600) y la porción de la señal de audio que es la segunda porción de la señal de audio codificada por el segundo procesador de codificación (610); y un formador de señal codificada (630) para formar una señal de audio codificada que comprende la primera porción de la señal codificada para la primera porción de la señal de audio y la segunda porción de la señal codificada para la segunda porción de la señal de audio.

2. Codificador de audio de la reivindicación 1, donde la señal de audio tiene una banda alta y una banda baja, donde el codificador espectral (606) está configurado para analizar la señal de audio para encontrar un primer conjunto de primeras porciones espectrales que se codificará con una alta resolución y un segundo conjunto de segundas porciones espectrales que se codificará paramétricamente con una resolución espectral baja y para codificar el primer conjunto de las primeras porciones espectrales de una manera de conservación de la forma de onda con la resolución espectral alta y para codificar paramétricamente el segundo conjunto de segundas porciones espectrales con la resolución espectral baja.

3. Codificador de audio de la reivindicación 1 o 2, que comprende además:

60 un pre-procesador (1000) configurado para preprocesar la primera porción de la señal de audio y la segunda porción de la señal de audio, donde el pre-procesador (1000) comprende un analizador de predicción (1002) para determinar coeficientes de predicción; donde el formador de señales codificadas (630) está configurado para introducir una versión codificada de los coeficientes de predicción en la señal de audio codificada.

65

4. Codificador de audio de las reivindicaciones 1, 2 que comprende además:

un pre-procesador (1000) configurado para preprocesar la señal de audio,
 donde el pre-procesador (1000) comprende un remuestreador (1004) para el remuestreo de la señal de audio a
 5 la segunda velocidad de muestreo del segundo procesador de codificación (610) para obtener una señal de
 audio remuestreada; y
 donde el pre-procesador (1000) comprende un analizador de predicción (1002b) configurado para determinar los
 coeficientes de predicción utilizando la señal de audio remuestreada, o
 10 donde el pre-procesador (1000) comprende además una etapa de análisis de predicción a largo plazo (1024)
 para determinar uno o más parámetros de predicción a largo plazo para la primera porción de la señal de audio.

5. Codificador de audio de una de las reivindicaciones anteriores, donde el procesador cruzado (700) comprende:

15 un decodificador espectral (701) para calcular una versión decodificada de la primera porción de la señal de
 audio;
 una etapa de retardo (707) para retrasar la versión decodificada de la primera porción de señal codificada para
 obtener una versión retrasada y para alimentar la versión retrasada en una etapa de desénfasis (617) del
 segundo procesador de codificación (610) para la inicialización;
 20 un bloque de filtrado de análisis de coeficiente de predicción ponderado (708) para filtrar la versión decodificada
 de la primera porción de señal codificada para obtener una salida de filtro y para alimentar la salida del filtro en
 un determinador de libros de códigos innovador (613) del segundo procesador de codificación (610) para la
 inicialización;
 una etapa de filtrado de análisis (706) para filtrar la versión decodificada de la primera porción de la señal
 25 codificada o una versión pre-enfatizada derivada por una etapa de pre-énfasis (709) de la versión decodificada
 de la primera porción de la señal codificada para obtener una señal residual de filtro y alimentar la señal residual
 del filtro en un determinador de libro de códigos adaptativo (612) del segundo procesador de codificación (610)
 para la inicialización; o
 un filtro de pre-énfasis (709) para filtrar la versión decodificada de la primera porción de señal codificada para
 30 obtener una versión pre-enfatizada y para alimentar la versión pre-enfatizada o una versión pre-enfatizada
 retrasada a una etapa de filtrado de síntesis (616) del segundo procesador de codificación (610) para la
 inicialización.

6. Codificador de audio de una de las reivindicaciones anteriores,

35 donde el primer procesador de codificación (600) está configurado para llevar a cabo una configuración (606a) de
 valores espectrales de la representación en el dominio de las frecuencias utilizando coeficientes de predicción (1002,
 1010) derivados de la primera porción de la señal de audio para obtener valores espectrales configurados, y donde
 el primer procesador de codificación (600) está configurado además para llevar a cabo una cuantificación y una
 operación de codificación de la entropía (606b) de valores espectrales configurados de la representación del dominio
 40 de frecuencias.

7. Codificador de audio de cualquiera de las reivindicaciones anteriores, donde el procesador cruzado (700) comprende:

45 un configurador de ruidos (703) para configurar valores espectrales cuantificados de la representación en el
 dominio de las frecuencias utilizando coeficientes LPC (1010) derivados de la primera porción de la señal de
 audio;
 un decodificador espectral (704, 705) para decodificar porciones espectrales espectralmente configuradas de la
 representación en el dominio de las frecuencias con una elevada resolución espectral de manera que se obtenga
 50 una representación espectral decodificada;
 un convertidor de frecuencia-tiempo (702) para llevar a cabo la transformada de tiempo de frecuencia de la
 representación espectral decodificada para obtener una primera porción de la señal de audio decodificada,
 donde la segunda velocidad de muestreo está asociada con la primera porción de la señal de audio decodificada.

55 8. Codificador de audio de una de las reivindicaciones anteriores, donde el segundo procesador de
 codificación (610) comprende al menos un bloque del siguiente grupo de bloques:

un filtro de análisis de predicción (611);
 una etapa del libro de código adaptativo (612);
 60 una etapa del libro de código innovadora (614);
 un estimador (613) para estimar una entrada del libro de código innovadora;
 una etapa de codificación ACELP/ganancia (615);
 una etapa de filtrado de síntesis de predicción (616);
 una etapa de desénfasis (617); y
 65 una etapa de análisis post-filtro de bajos (618).

9. Codificador de audio de una de las reivindicaciones anteriores,

5 donde el procesador cruzado (700) comprende un convertidor de tiempo de frecuencia (702) para realizar la transformada de tiempo de frecuencia en la representación espectral decodificada para generar una señal de dominio de tiempo en la segunda velocidad de muestreo, donde el convertidor de tiempo de frecuencia (702) comprende:

10 un selector (726) para seleccionar la porción de banda baja de acuerdo con una relación con la primera velocidad de muestreo y la segunda velocidad de muestreo,
un procesador de transformada (720) que tiene el tamaño de transformada reducido; y
un formador de ventana de síntesis (712) para la formación de ventana utilizando una ventana que tiene un número diferente de coeficientes de ventana en comparación con una ventana utilizada por el convertidor de frecuencia de tiempo (602).

15 10. Decodificador de audio para decodificar una señal de audio codificada, que comprende:

20 un primer procesador de decodificación (1120) para decodificar una primera porción de señal de audio codificada en un dominio de frecuencia para obtener una representación espectral decodificada, comprendiendo el primer procesador de decodificación (1120) un convertidor de tiempo de frecuencia (1124) para convertir la representación espectral decodificada en un dominio de tiempo para obtener una primera porción decodificada de la señal de audio, donde la representación espectral decodificada se extiende hasta una frecuencia máxima de una representación de tiempo de una señal de audio decodificada, siendo un valor espectral para la frecuencia máxima cero o diferente de cero, y donde la representación espectral decodificada ha asociado con el mismo una primera velocidad de muestreo;
25 un segundo procesador de decodificación (1140) para decodificar una segunda porción de señal de audio codificada en el dominio de tiempo para obtener una segunda porción decodificada de la señal de audio, donde la segunda porción decodificada de la señal de audio ha asociado con ello una segunda velocidad de muestreo, donde el segundo procesador de decodificación (1140) comprende:

30 un decodificador de banda baja de dominio de tiempo (1200) para decodificación para obtener una señal de dominio de tiempo de banda baja;
un remuestreador (1210) para remuestrear la señal de dominio de tiempo de banda baja para obtener una señal de dominio de tiempo de banda baja remuestreada;
35 un decodificador de extensión de ancho de banda de dominio de tiempo (1220) para sintetizar una banda alta de una señal de salida de dominio de tiempo para obtener una banda alta sintetizada; y
un mezclador (1230) para mezclar la banda alta sintetizada y la señal de dominio de tiempo de banda baja remuestreada;

40 un procesador cruzado (1170) para calcular, a partir de la representación espectral decodificada, datos de inicialización del segundo procesador de decodificación (1140), de modo que el segundo procesador de decodificación (1140) se inicialice para decodificar la segunda porción de señal de audio codificada siguiendo a tiempo la primera porción de señal de audio codificada en la señal de audio codificada, donde el procesador cruzado (1170) está configurado para usar una transformada de tiempo de frecuencia adicionalmente, realizando
45 un muestreo descendente a partir de la primera velocidad de muestreo a la segunda velocidad de muestreo seleccionando una porción de banda baja de la representación espectral decodificada junto con un tamaño de transformada reducido para obtener los datos de inicialización del segundo procesador de decodificación (1140);
y
50 un combinador (1160) para combinar la primera porción decodificada de la señal de audio y la segunda porción decodificada de la señal de audio para obtener una señal de audio decodificada.

11. Decodificador de audio de la reivindicación 10, donde el primer procesador de decodificación (1120) está configurado para reconstruir un primer conjunto de primeras porciones espectrales en una manera de preservación de forma de onda para generar un espectro que tiene intervalos, donde los intervalos en el espectro se
55 llenan con una tecnología de relleno inteligente de intervalos (IGF) que comprende el uso de una regeneración de frecuencia de aplicación de los datos paramétricos por una parte y el uso de las primeras porciones espectrales reconstruidas del primer conjunto de primeras porciones espectrales.

12. Decodificador de audio de una de las reivindicaciones 10 a 11,

60 donde el primer procesador de decodificación (1120) comprende un filtro posterior de predicción a largo plazo adaptativo (1420) para realizar un filtrado posterior de la primera porción de la señal de audio decodificada, donde el filtro de predicción a largo plazo adaptativo (1420) está controlado por uno o más parámetros de predicción a largo plazo incluidos en la señal de audio codificada.

65 13. Decodificador de audio de una de las reivindicaciones 10 a 12, donde el procesador cruzado (1170)

comprende además:

un convertidor de tiempo de frecuencia adicional (1171) para realizar la transformada de tiempo de frecuencia en la representación espectral decodificada, que opera a la segunda velocidad de muestreo que es diferente de la primera velocidad de muestreo asociada con el convertidor de tiempo de frecuencia (1124) del primer procesador de decodificación (1120) para obtener una primera porción de la señal de audio decodificada en el dominio de tiempo,

donde la primera porción de la señal de audio decodificada tiene la segunda velocidad de muestreo que es diferente de la primera velocidad de muestreo asociada con la primera porción de la señal de audio decodificada, donde el convertidor de tiempo de frecuencia adicional (1171) comprende:

un selector (726) para seleccionar la porción de banda baja de la representación espectral decodificada de acuerdo con una proporción de la primera velocidad de muestreo y la segunda velocidad de muestreo; un procesador de transformada (720) que tiene el tamaño de transformada reducido que es diferente de un tamaño de transformada (710) del convertidor de tiempo de frecuencia (1124); y un formador de ventana de síntesis (722) que utiliza una ventana que tiene un número diferente de coeficientes en comparación con una ventana utilizada por el convertidor de tiempo de frecuencia (1124).

14. Decodificador de audio de una de las reivindicaciones 10 a 13, donde el procesador cruzado (1170) comprende:

una etapa de retardo (1172) para retrasar la primera porción decodificada de la primera señal de audio y para alimentar una versión retrasada de la primera porción de la señal de audio decodificada adicionalmente en una etapa de desénfasis (1144) del segundo procesador de decodificación (1140) para la inicialización;

un filtro de pre-énfasis (1173) y una etapa de retardo (1175) para filtrar y retrasar la primera porción de la señal de audio decodificada y para alimentar una salida de la etapa de retardo en un filtro de síntesis de predicción (1143) del segundo procesador de decodificación (1140) para la inicialización;

un filtro de análisis de predicción (1174) para generar una señal residual de predicción de la primera porción de la señal de audio decodificada o de una primera porción de la señal de audio decodificada pre-enfatizada (1173) y para alimentar la señal residual de predicción en un sintetizador de libro de código (1141) del segundo procesador de decodificación (1140); o

un interruptor (1480) para alimentar la primera porción de la señal de audio decodificada adicionalmente en una etapa de análisis (1471) de un remuestreador (1210) del segundo procesador de decodificación (1140) para la inicialización.

15. Decodificador de audio de una de las reivindicaciones 10 a 14, donde el segundo procesador de decodificación (1140) comprende al menos un bloque del grupo de bloques que comprende:

una etapa (1149) para decodificar las ganancias de ACELP y un libro de código innovador;

una etapa de síntesis de libros de código adaptativos (1141);

un post-procesador de ACELP (1142);

un filtro de síntesis de predicción (1143); y

una etapa de desénfasis (1144).

16. Procedimiento para la codificación de una señal de audio, que comprende una banda alta y una banda baja, comprendiendo el procedimiento:

codificación (600) de una primera porción de la señal de audio en el dominio de las frecuencias, teniendo asociada la primera porción de la señal de audio una velocidad de muestreo, que comprende:

conversión (602) de la primera porción de la señal de audio en una representación en el dominio de frecuencias que tiene líneas espectrales hasta una frecuencia máxima de la primera porción de la señal de audio, donde la frecuencia máxima de la primera porción de señal de audio es inferior o igual a la mitad de la primera velocidad de muestreo y al menos un cuarto de la primera velocidad de muestreo o superior;

codificación (606) de la representación del dominio de las frecuencias para obtener una representación espectral codificada de la primera porción de señal de audio que es una primera porción de señal codificada;

codificación (610) de una segunda porción de la señal de audio en un dominio de tiempo para obtener una segunda porción de señal codificada, siendo la segunda porción de señal de audio diferente de la primera porción de señal de audio, donde la codificación (610) de la segunda porción de señal de audio comprende:

conversión de la segunda porción de la señal de audio a una segunda representación de la velocidad de muestreo que tiene una segunda velocidad de muestreo, siendo la segunda velocidad de muestreo inferior a la primera velocidad de muestreo, donde la segunda representación de la velocidad de muestreo no incluye la banda alta de la señal de audio;

dominio de tiempo que codifica la segunda representación de la velocidad de muestreo; y
 codificación de extensión de ancho de banda de dominio de tiempo que comprende la codificación
 paramétrica de la banda alta de la señal de audio;

5 cálculo (700) a partir de la representación espectral codificada de la primera porción de la señal de audio, de
 datos de inicialización para la etapa de codificación (610) de la segunda porción de señal de audio, de tal manera
 que la etapa de la codificación (610) de la segunda porción de la señal de audio se inicialice con el fin de
 10 codificar la segunda porción de la señal de audio inmediatamente después de la primera porción de la señal de
 audio en el tiempo en la señal de audio; donde el cálculo (700) comprende el uso de una transformada de tiempo
 de frecuencia realizando adicionalmente un muestreo descendente desde la primera velocidad de muestreo a
 la segunda velocidad de muestreo seleccionando una porción de banda baja de la representación del dominio de
 las frecuencias junto con un tamaño de transformada reducido para obtener los datos de inicialización de la
 etapa de codificación (610) de la segunda porción de señal de audio;
 15 análisis (620) de la señal de audio y determinación de qué porción de la señal de audio es la primera porción de
 la señal de audio codificada en el dominio de las frecuencias y qué porción de la señal de audio es la segunda
 porción de la señal de audio codificada en el dominio los tiempos; y
 configuración (630) de una señal de audio codificada que comprende la primera porción de señal codificada para
 la primera porción de la señal de audio y la segunda porción de señal codificada para la segunda porción de
 20 señal de audio.

17. Procedimiento de decodificación de una señal de audio codificada, que comprende:

25 decodificación (1120) de una primera porción de la señal de audio codificada en un dominio de las frecuencias
 para obtener una representación espectral decodificada, comprendiendo la decodificación (1120) la primera
 porción de señal de audio codificada convertir (1124) la representación espectral decodificada en un dominio de
 los tiempos para obtener una primera porción de la señal de audio decodificada, donde la representación
 espectral decodificada se extiende hasta una frecuencia máxima de una representación de tiempo de una señal
 de audio decodificada, siendo un valor espectral para la frecuencia máxima cero o diferente de cero, y donde la
 30 representación espectral decodificada ha asociado con ello una primera velocidad de muestreo;

30 decodificación (1140) de una segunda porción de la señal de audio codificada en el dominio de los tiempos para
 obtener una segunda porción de señal de audio decodificada, donde la segunda porción de señal de audio
 decodificada ha asociado con ello una segunda velocidad de muestreo, donde la decodificación (1140) de la
 segunda porción de señal de audio codificada comprende:

35 dominio de tiempo de decodificación de banda baja para obtener una señal de dominio de tiempo de banda
 baja;
 remuestreo de la señal de dominio de tiempo de banda baja para obtener una señal de dominio de tiempo de
 banda baja remuestreada;
 decodificación de extensión de ancho de banda de dominio de tiempo para sintetizar una banda alta de una
 40 señal de salida de dominio de tiempo para obtener una banda alta sintetizada; y
 mezclado de la banda alta sintetizada y la señal de dominio de tiempo de banda baja remuestreada;

45 cálculo (1170) a partir de la representación espectral decodificada de la primera porción de la señal de audio
 codificada, de datos de inicialización de la etapa de la decodificación (1140) de la segunda porción de la señal de
 audio codificada, de tal manera que la etapa de la decodificación (1140) de la segunda porción de la señal de
 audio codificada sea inicializada de manera que se decodifique la segunda porción de la señal de audio
 codificada que sigue en el tiempo a la primera porción de la señal de audio codificada en la señal de audio
 codificada, donde el cálculo (1170) comprende el uso de una transformada de tiempo de frecuencia realizando
 50 adicionalmente un muestreo descendente a partir de la primera velocidad de muestreo a la segunda velocidad de
 muestreo seleccionando una porción de banda baja de la representación espectral decodificada junto con un
 tamaño de transformada reducido para obtener los datos de inicialización de la etapa de decodificación (1140) de
 la segunda porción de señal de audio codificada; y la combinación (1160) de la primera porción de señal de
 audio decodificada y la segunda porción de señal de audio decodificada de manera que se obtenga una señal de
 55 audio decodificada.

18. Programa informático configurado para llevar a cabo, cuando se ejecuta en un ordenador o un
 procesador, el procedimiento de la reivindicación 16 o la reivindicación 17.

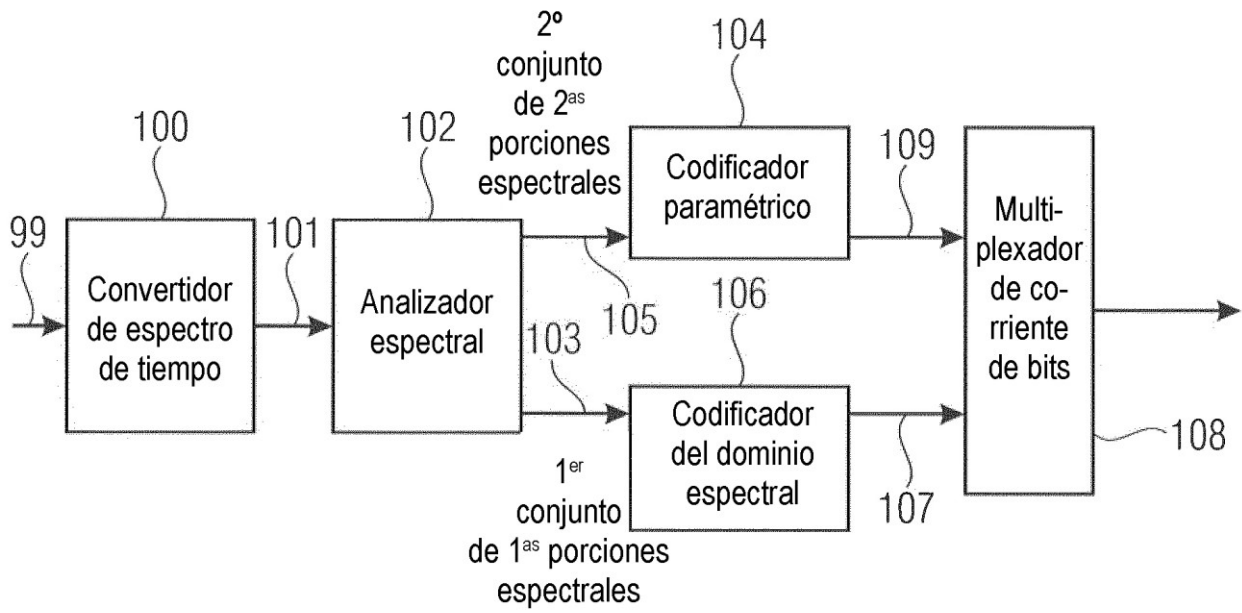


FIG 1A

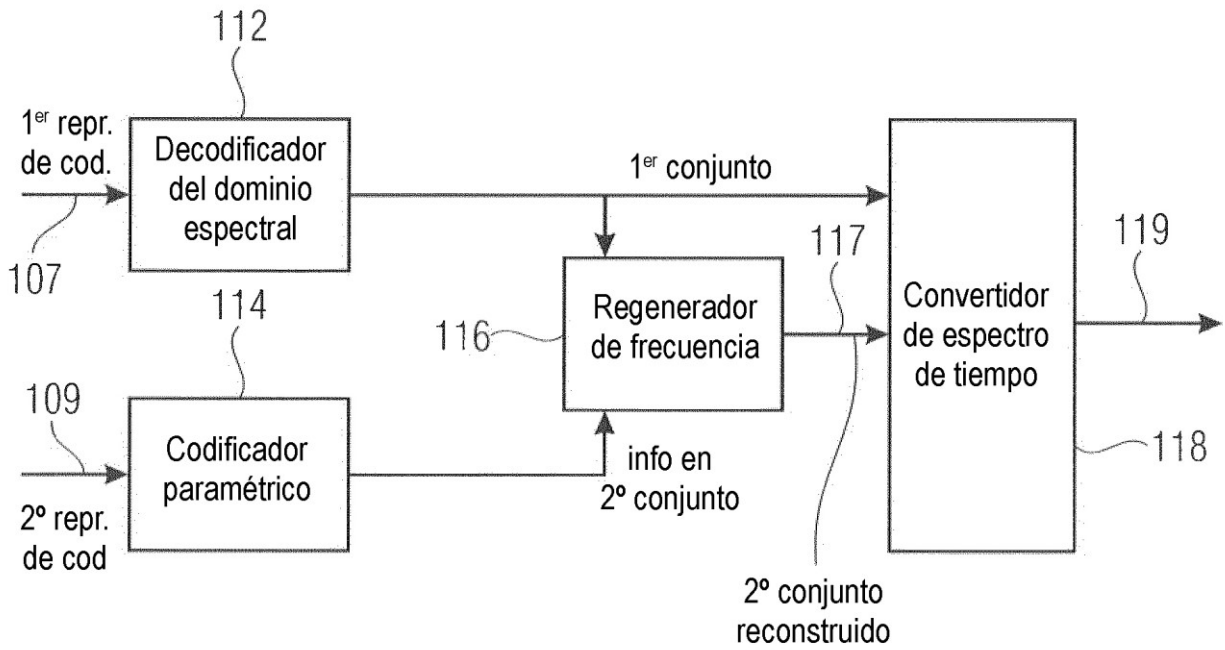


FIG 1B

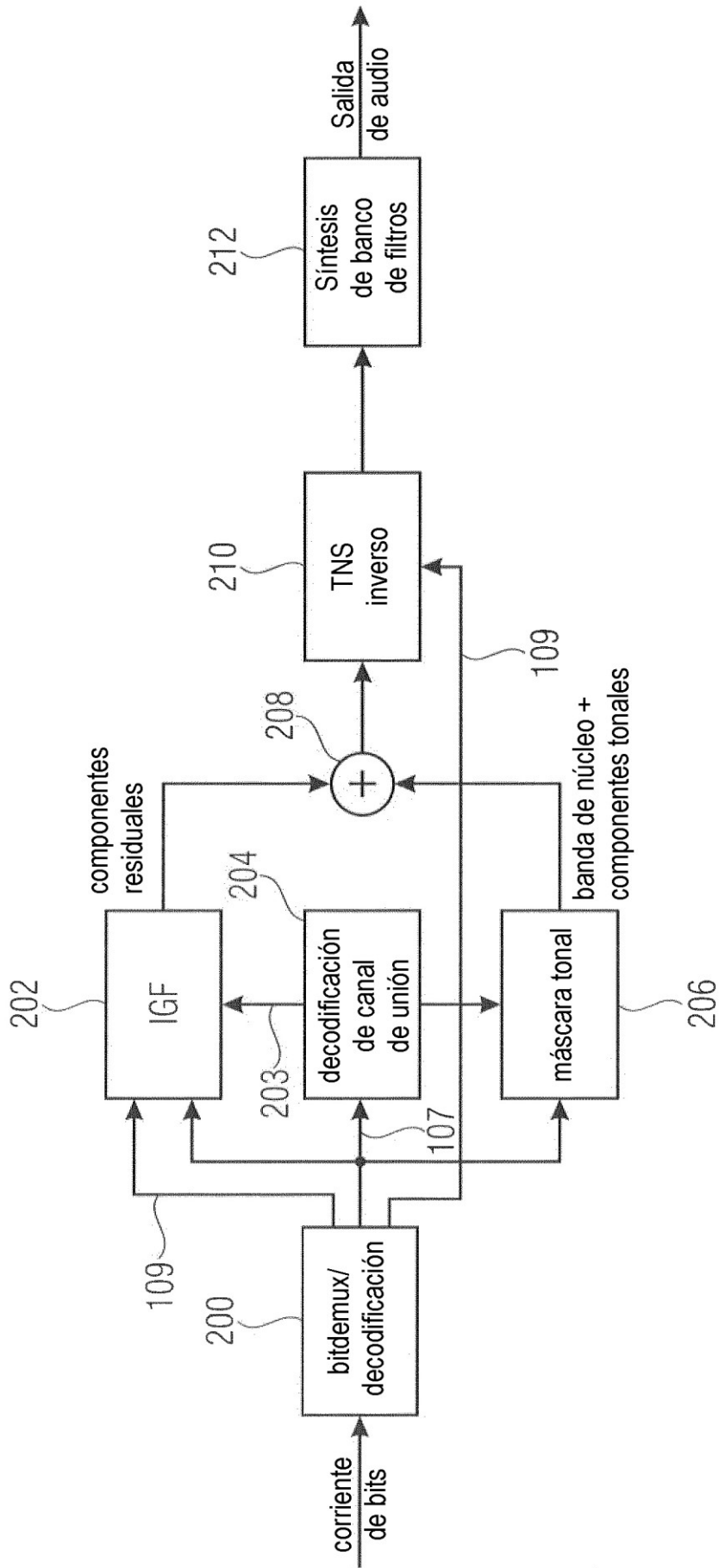


FIG 2A

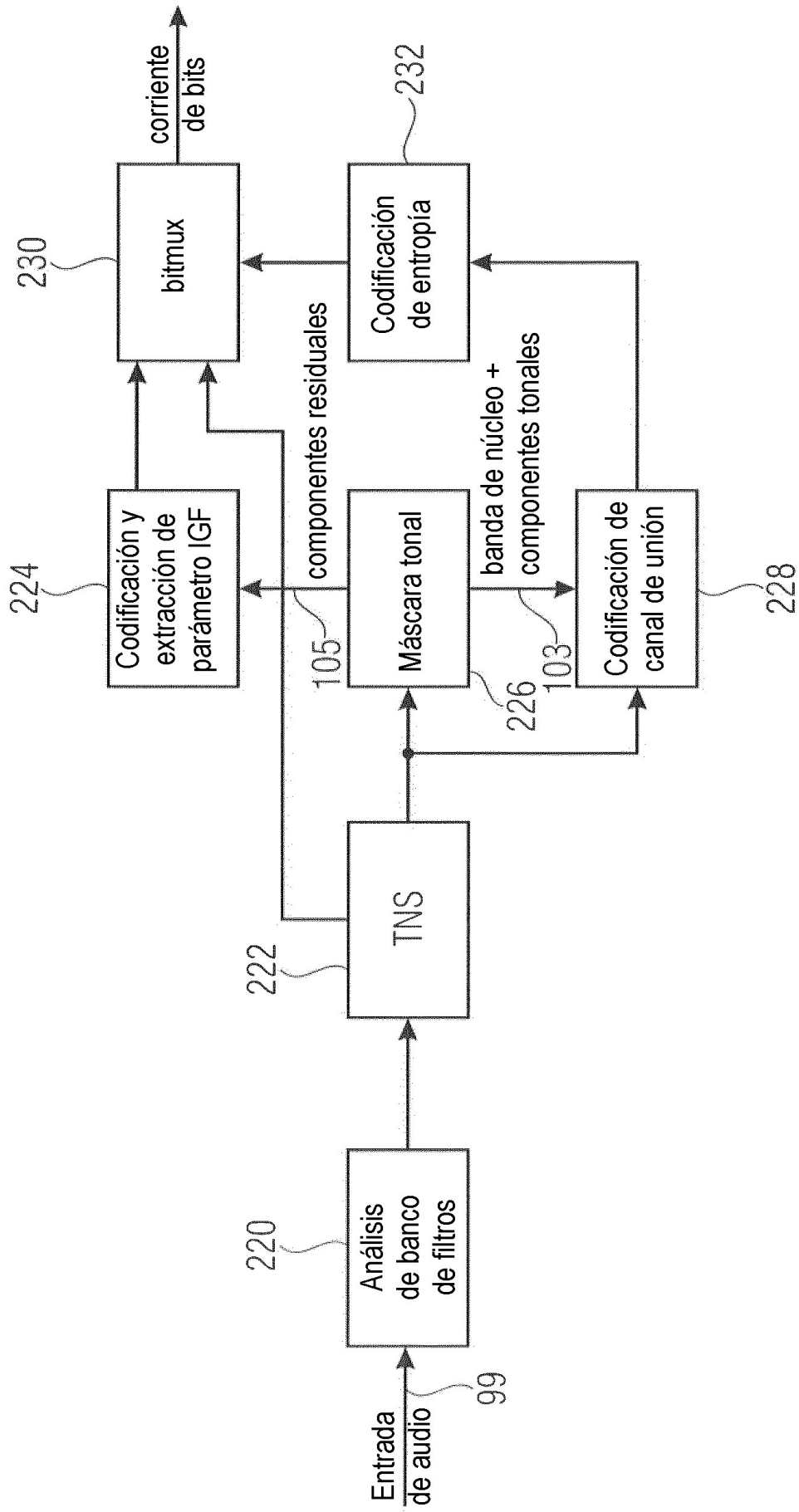


FIG 2B

- 1ª resolución (resolución alta) para “envolvente” del 1er conjunto (codificación de línea);
- 2ª resolución (resolución baja) para “envolvente” del 2º conjunto (factor de escala por SCB)

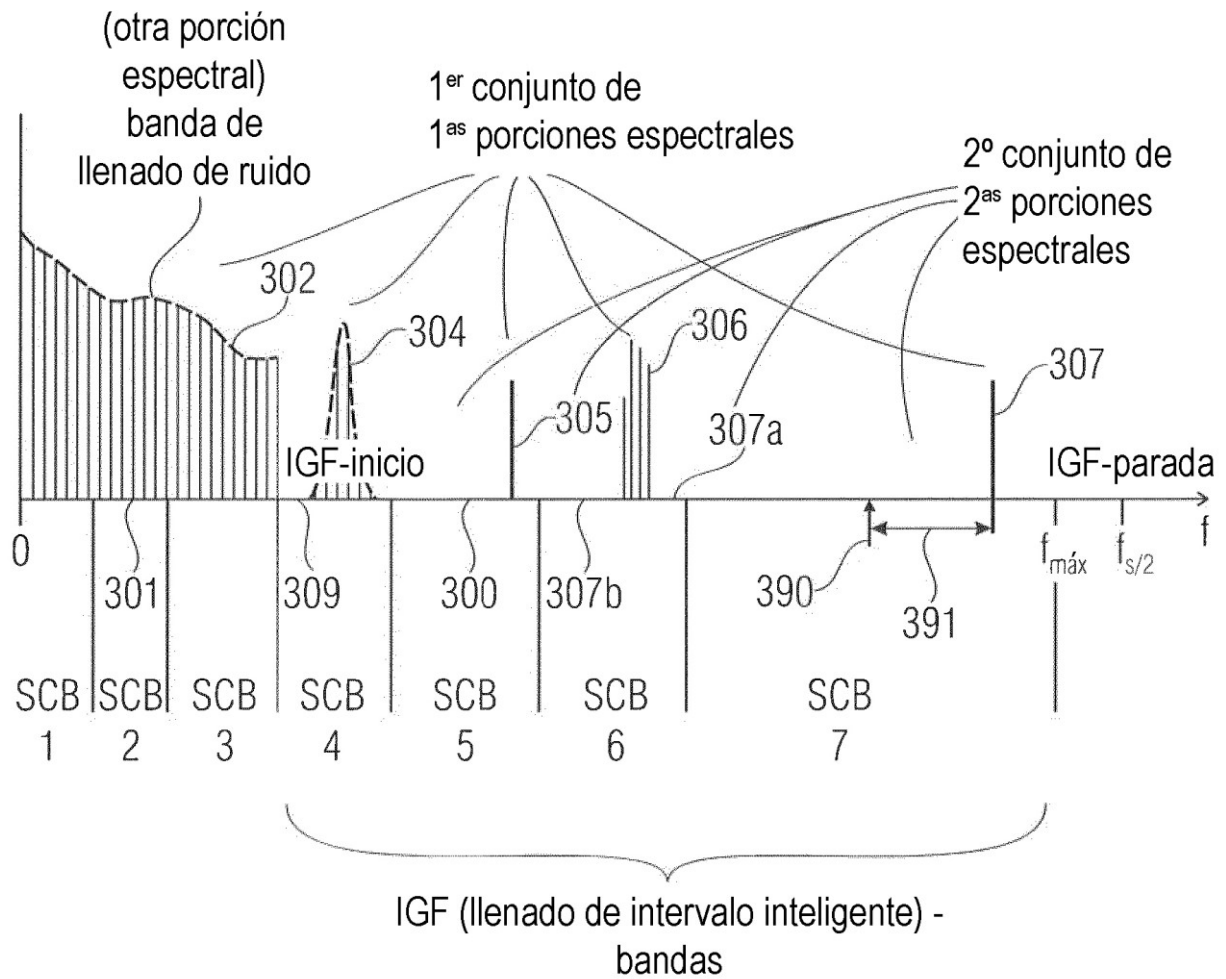


FIG 3A

SCB1	SCB2	SCB3	SCB4	SCB5	SCB6	SCB7
SF1	SF2	SF3	SF4	SF5	SF6	SF7
			E ₁	E ₂	E ₃	E ₄
	NF ₂					

308
310
312

FIG 3B

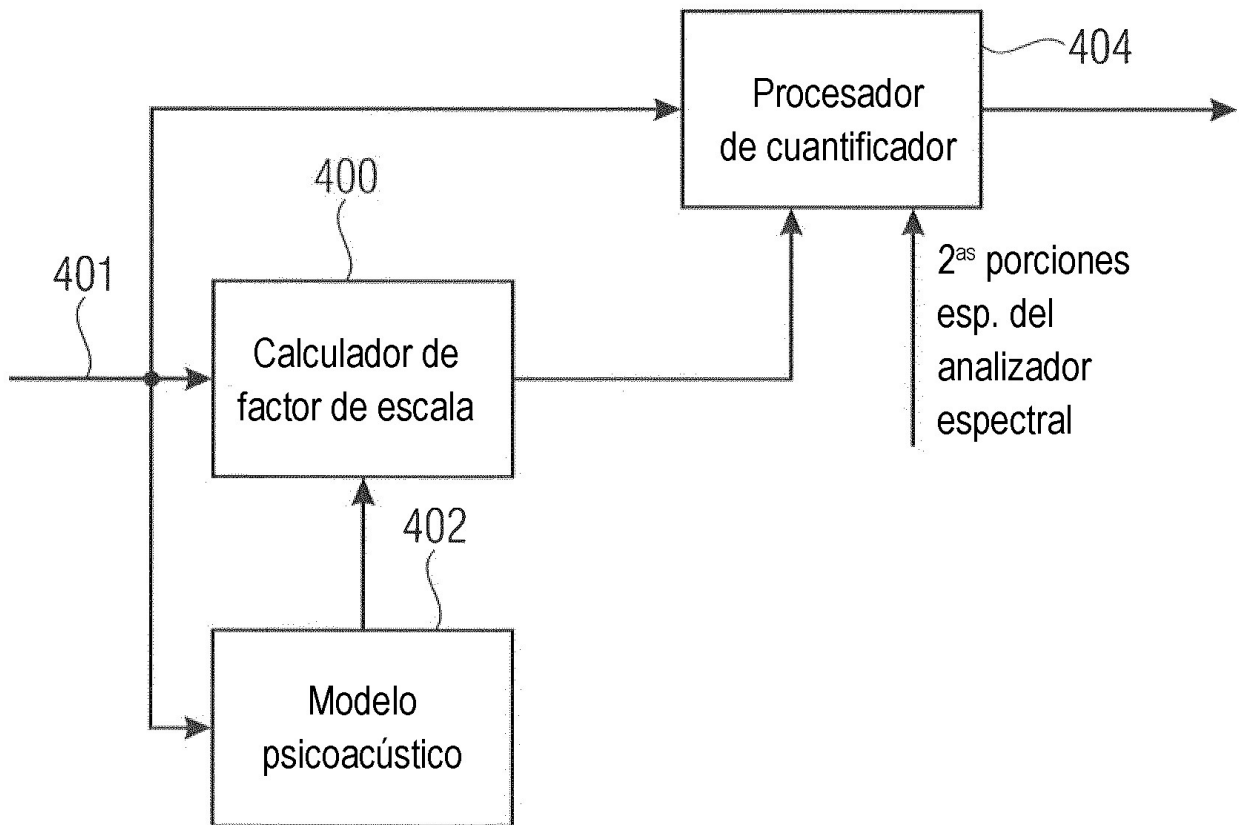


FIG 4A

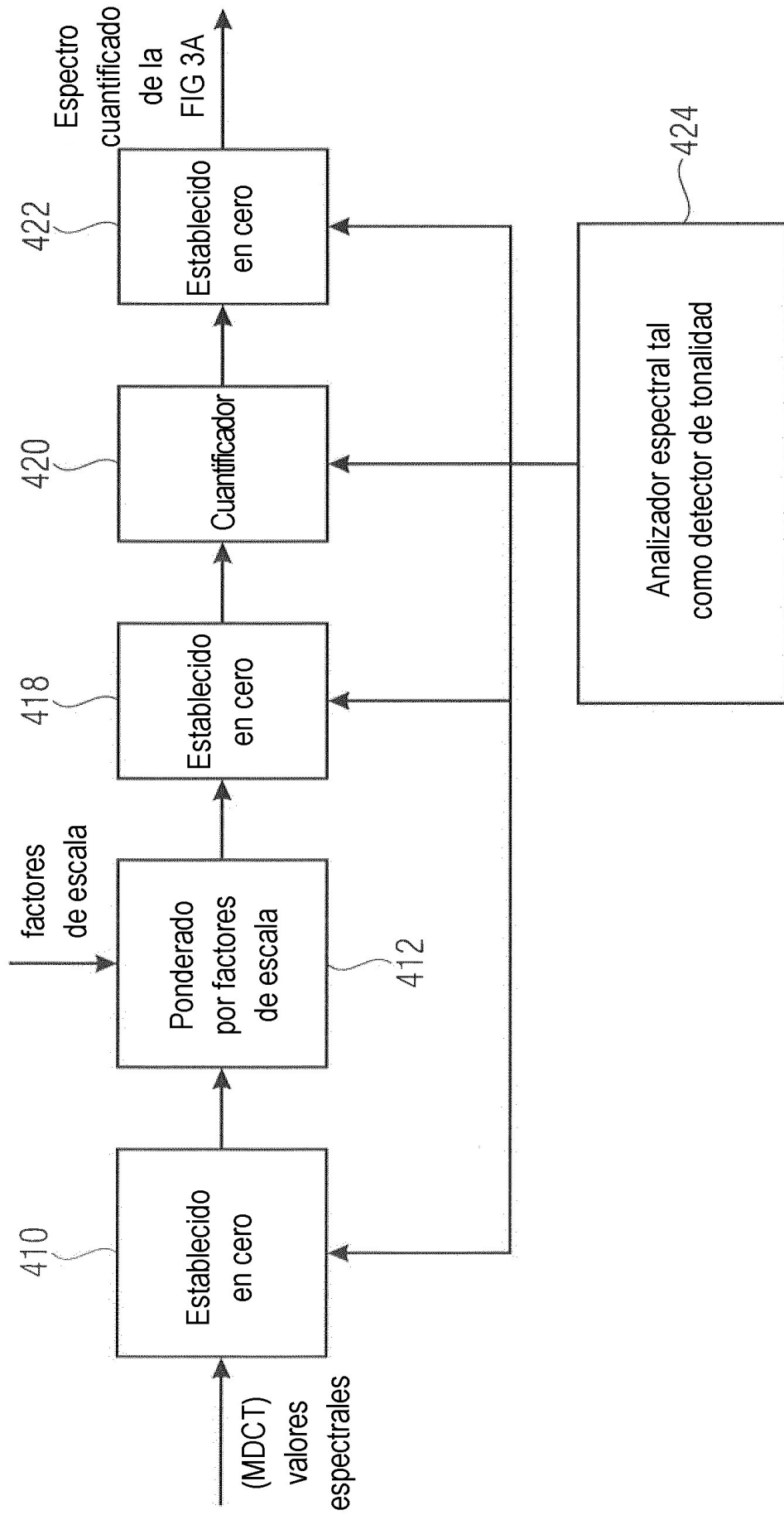


FIG 4B
(PROCESADOR DE CUANTIFICADOR)

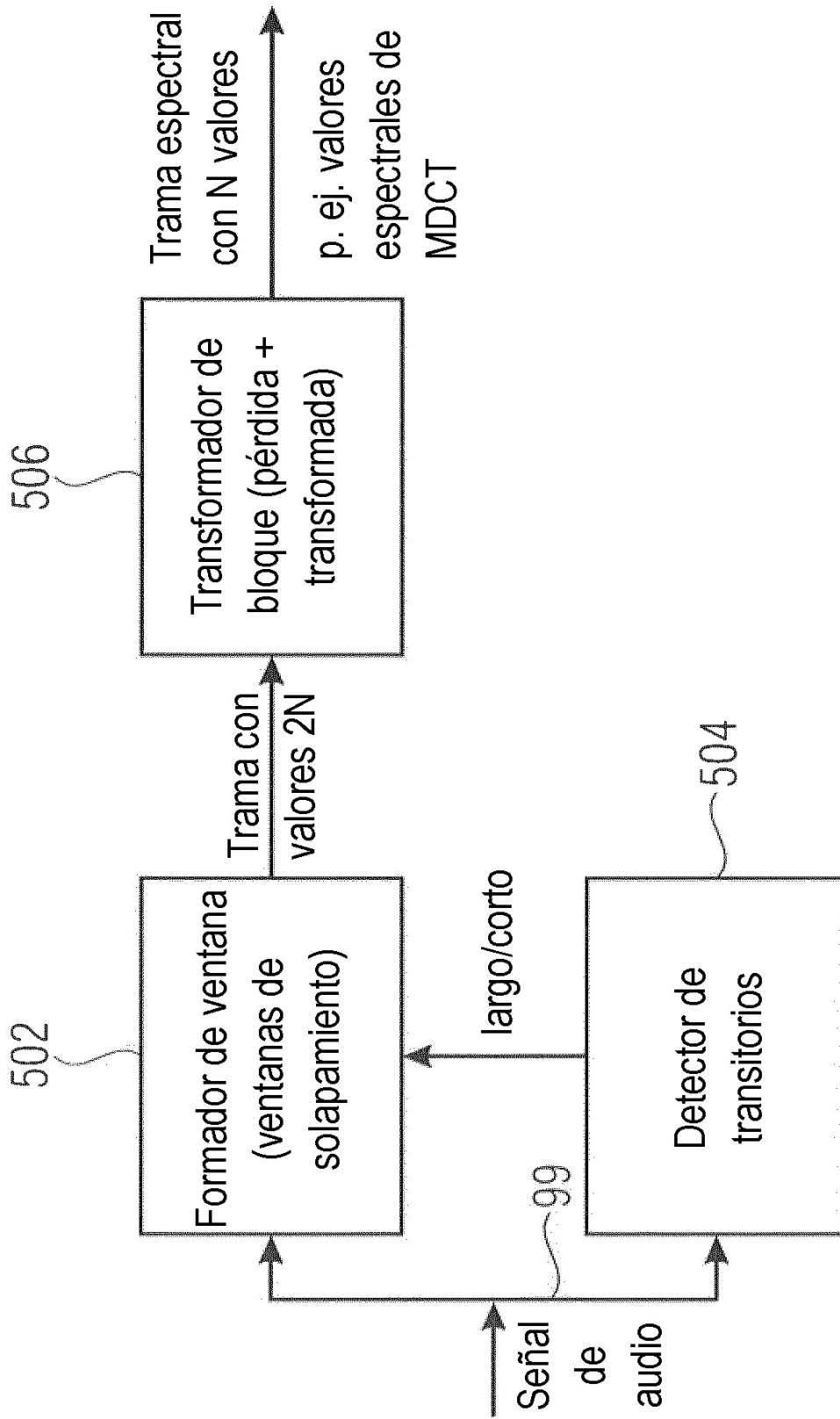


FIG 5A
(OTRAS PORCIONES ESPECTRALES)

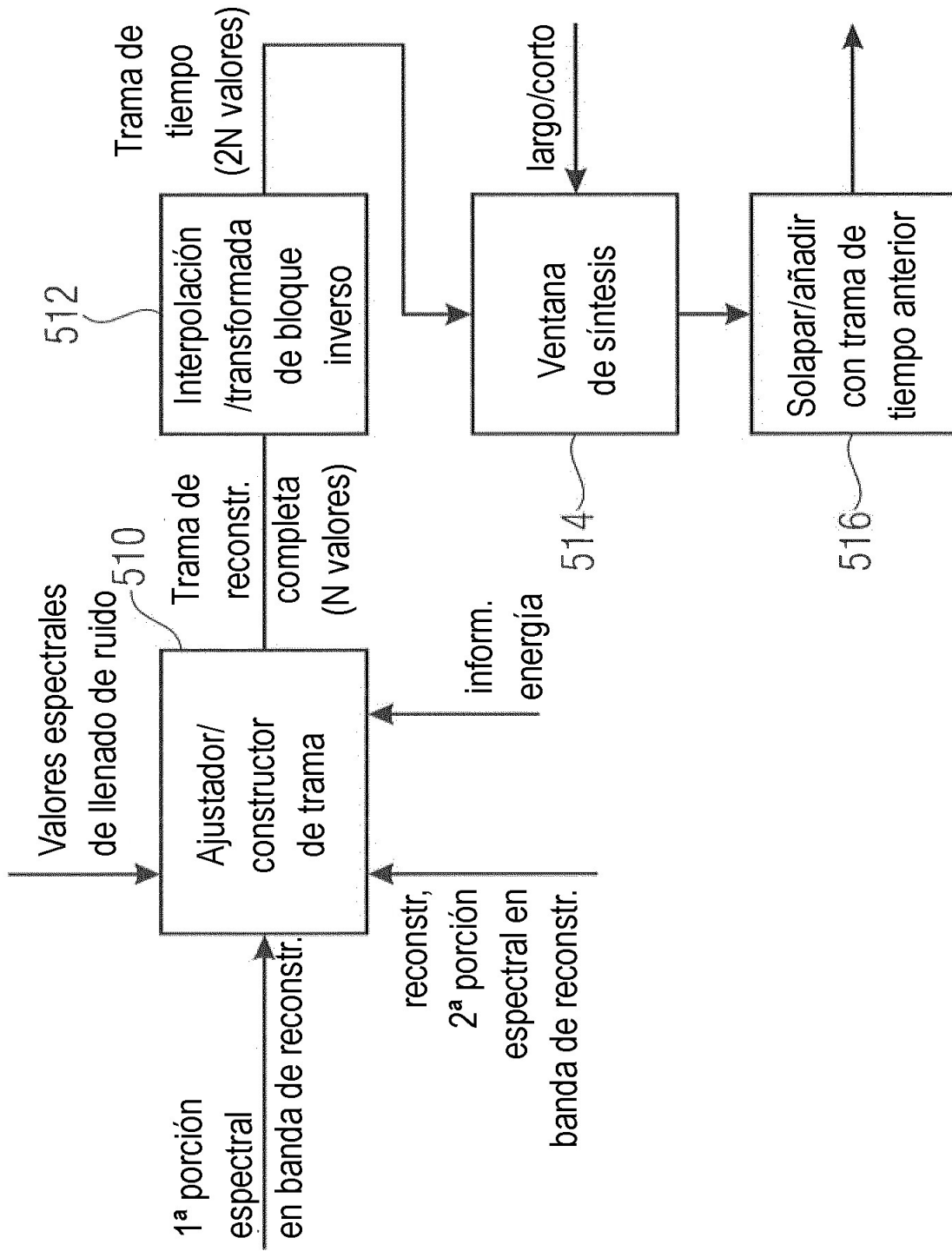


FIG 5B

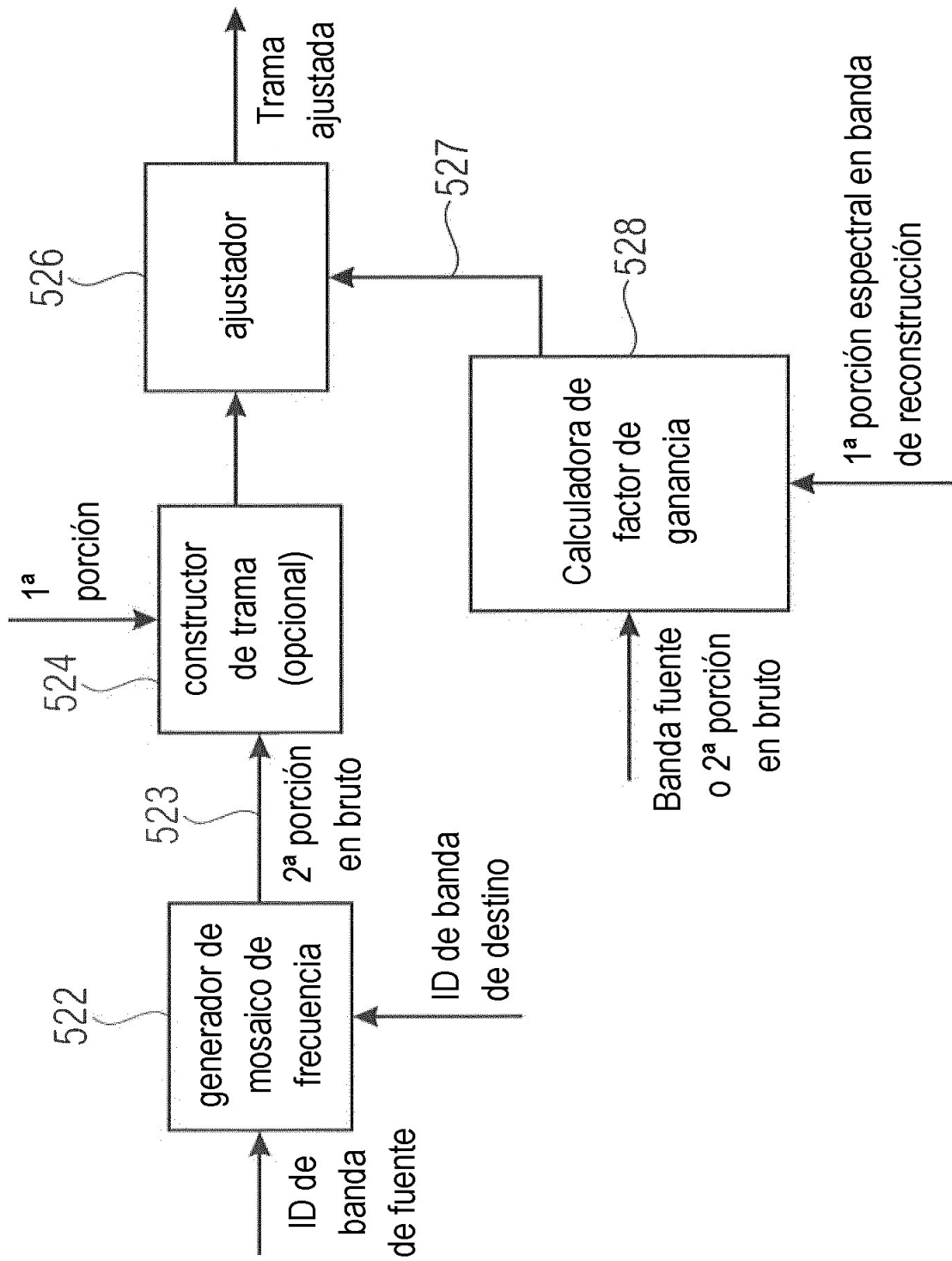


FIG 5C

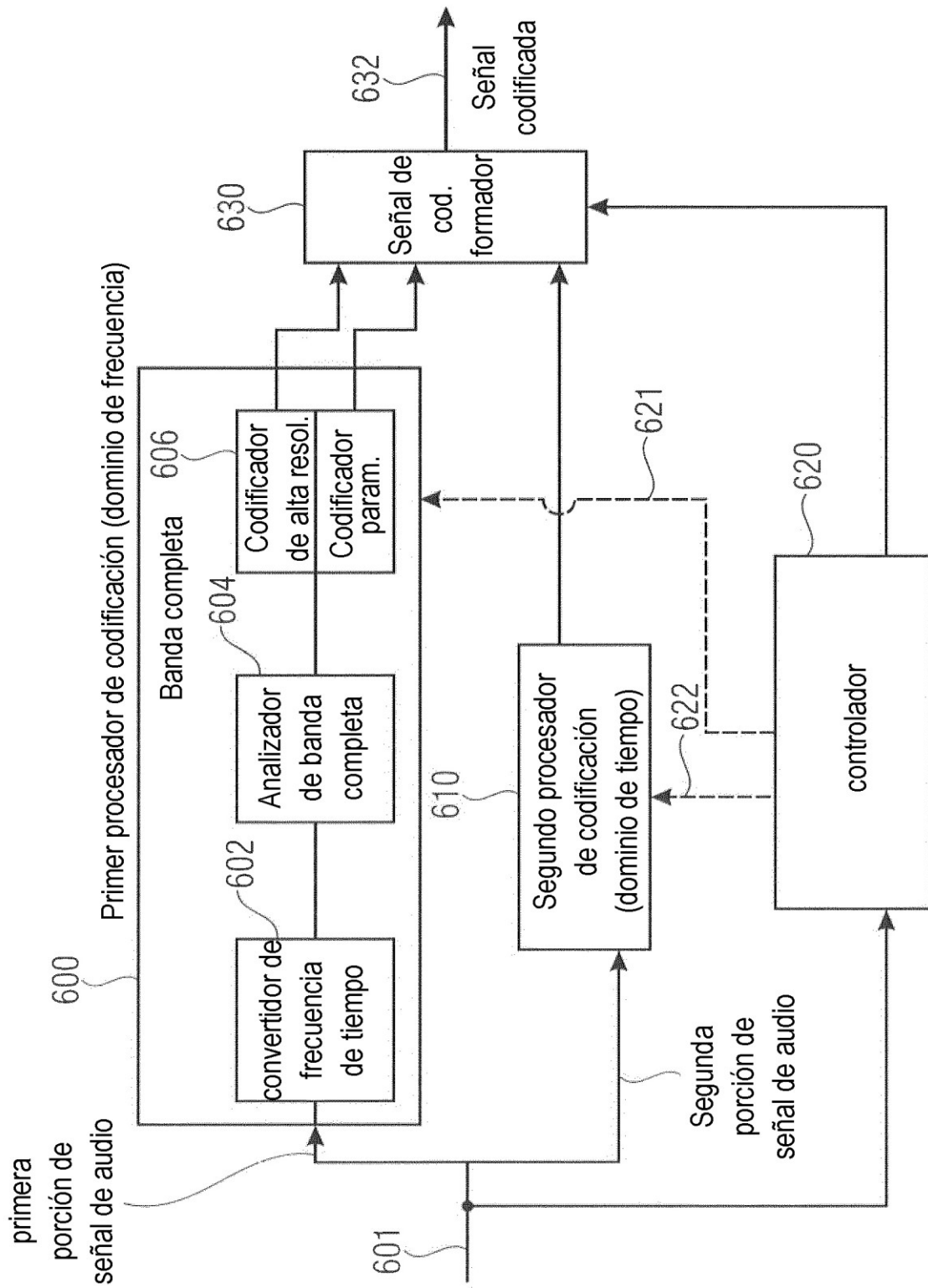


FIG 6

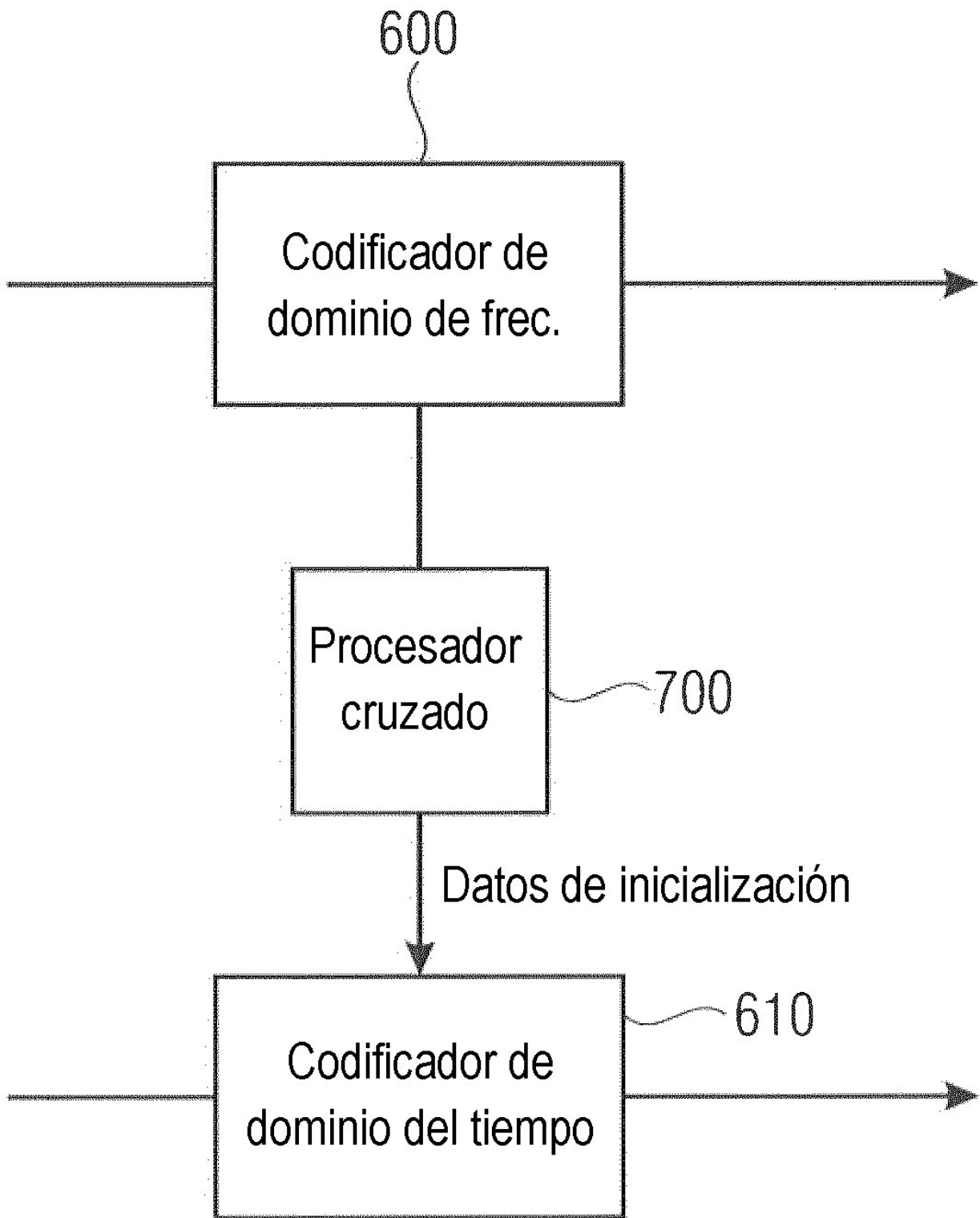
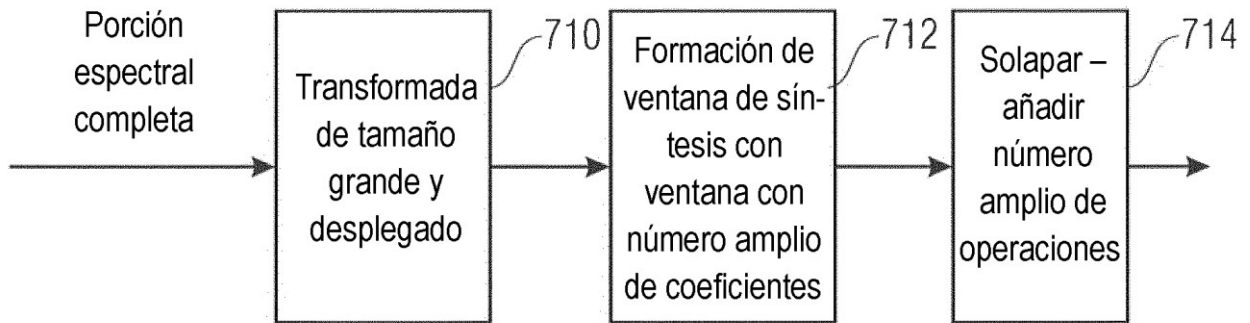
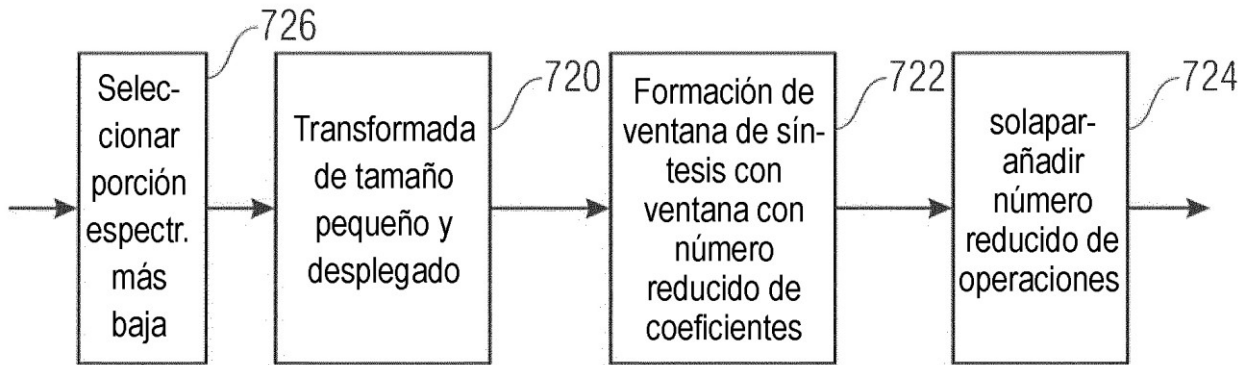


FIG 7A

VELOCIDAD DE MUESTREO DE SALIDA



VELOCIDAD DE MUESTREO DEL CODIFICADOR DE DOMINIO DEL TIEMPO



DS
Factor de muestreo descendente

$$\frac{1}{DS} = \frac{\text{Codificador T.D S.R}}{\text{Codificador F.D S.R}}$$

$$DS * \text{tamaño reducido} = \text{tamaño amplio}$$

$$DS * \text{número reducido de coef.} = \text{número amplio de coeficientes}$$

$$DS * \text{número reducido de coef.} = \text{número amplio de coeficientes}$$

FIG 7B

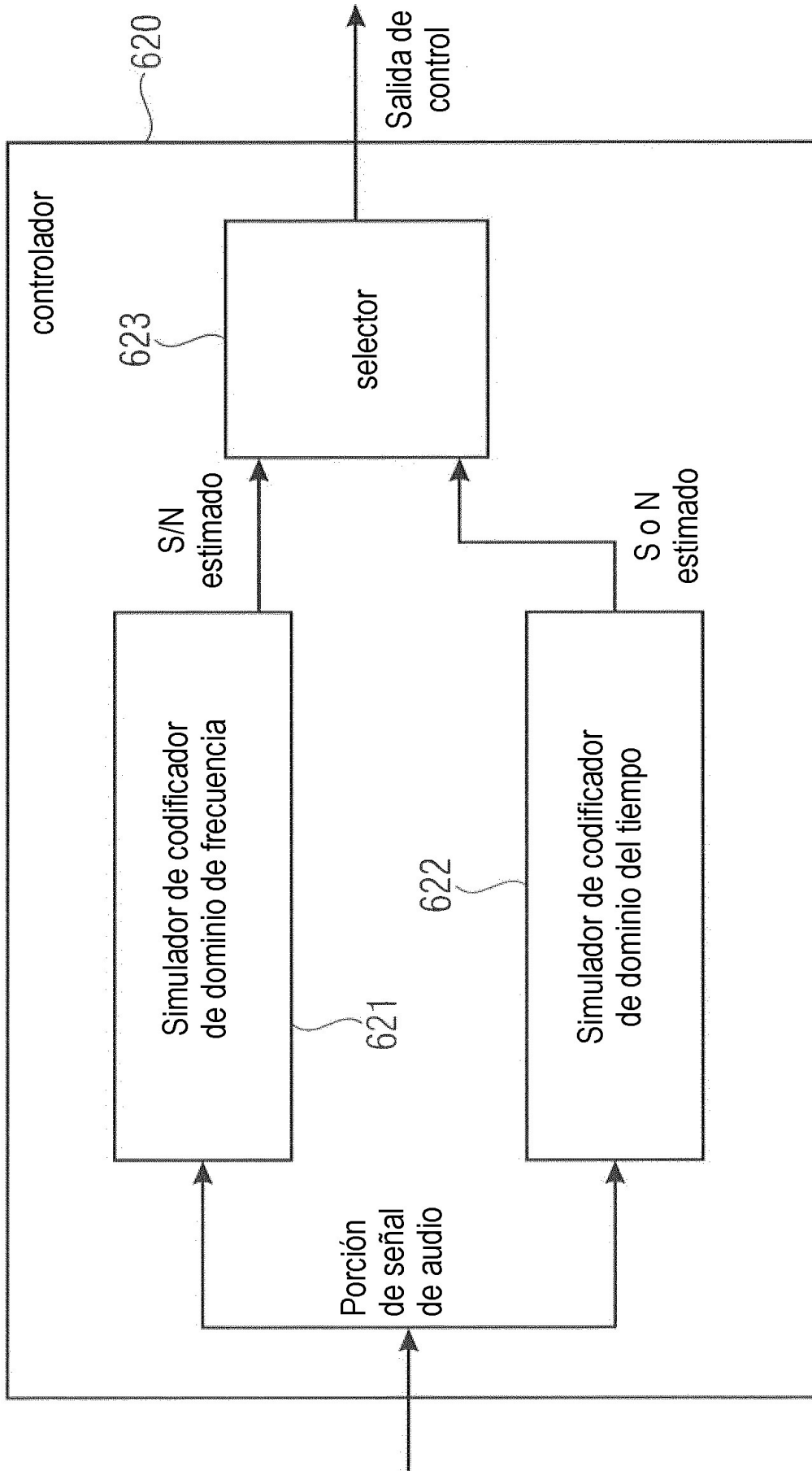


FIG 8

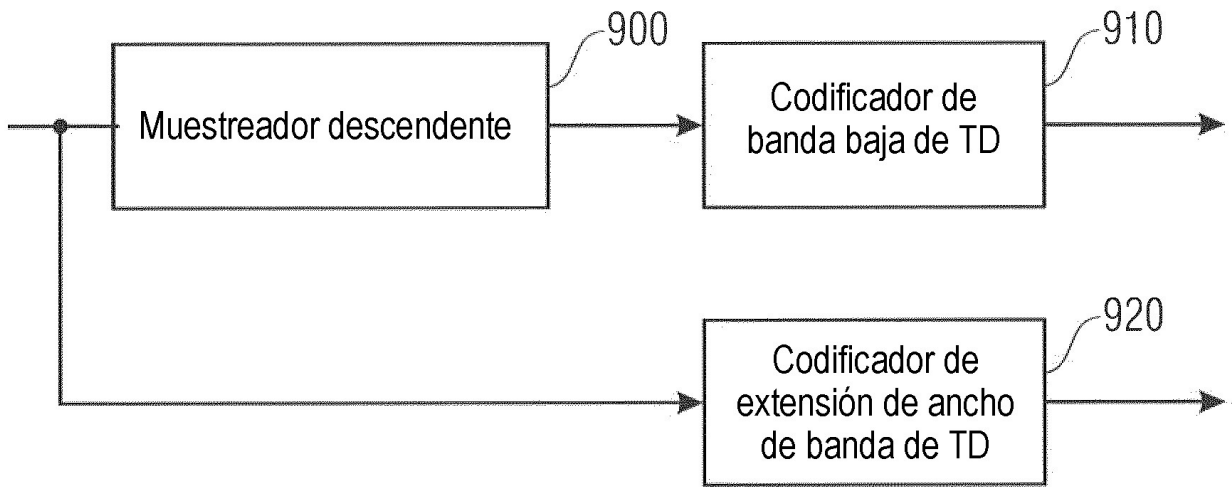


FIG 9

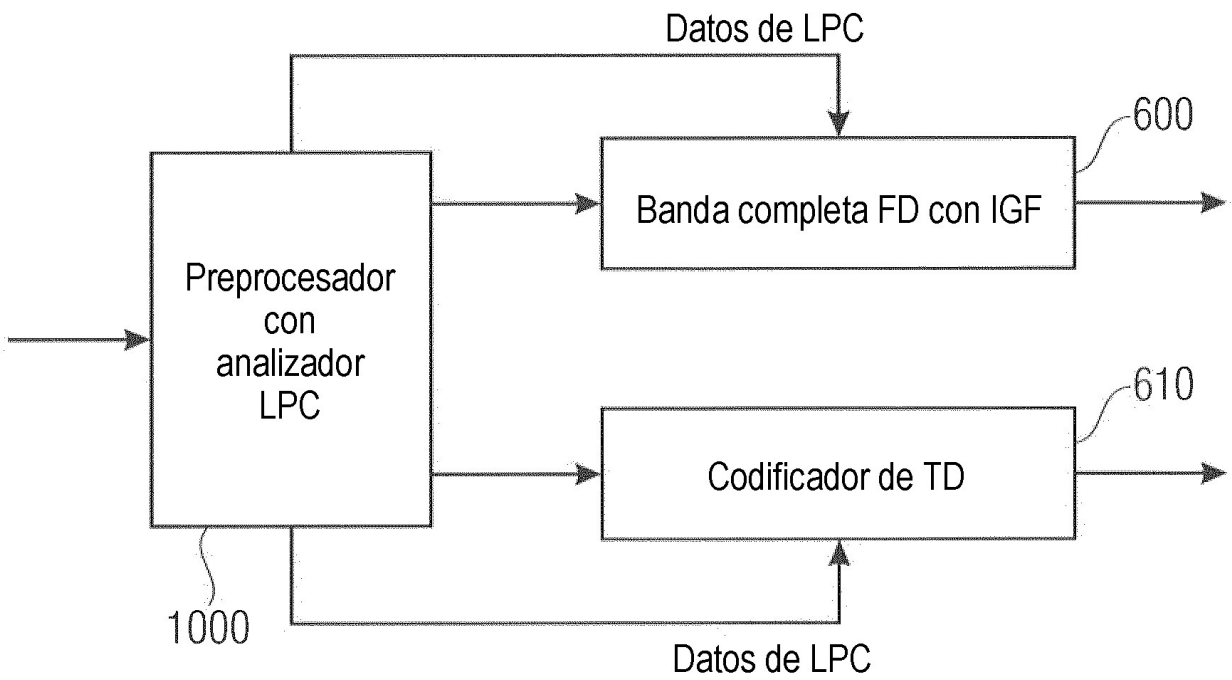


FIG 10

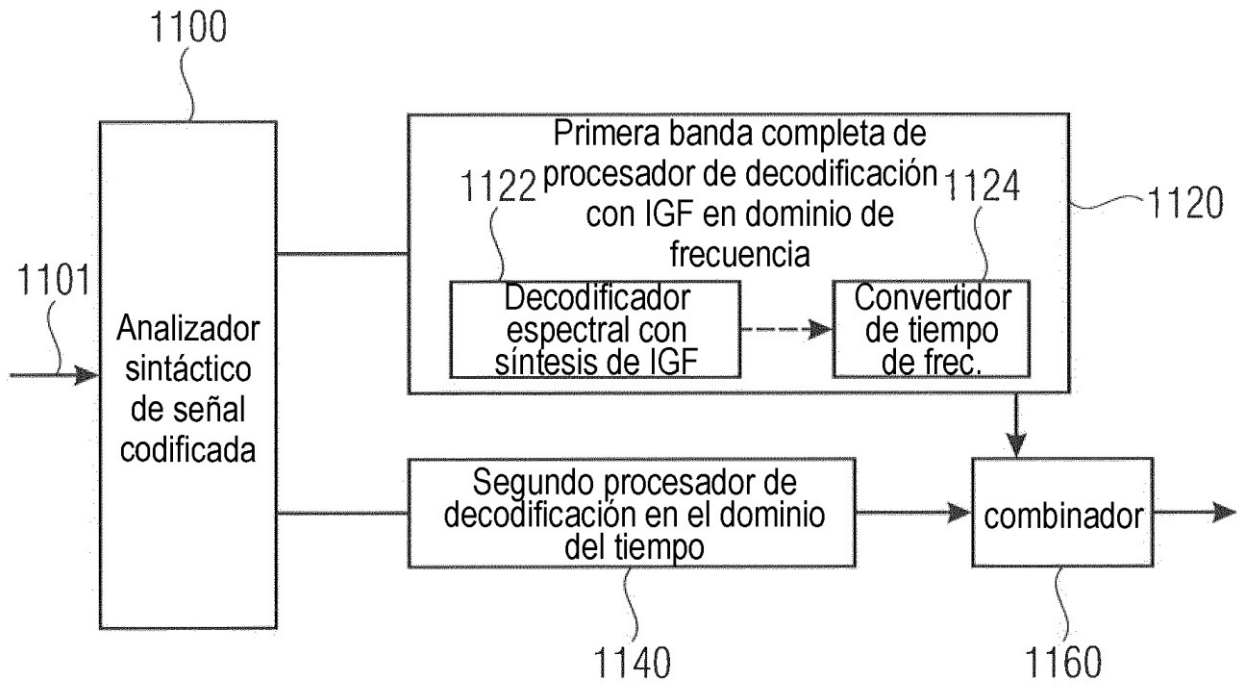


FIG 11A

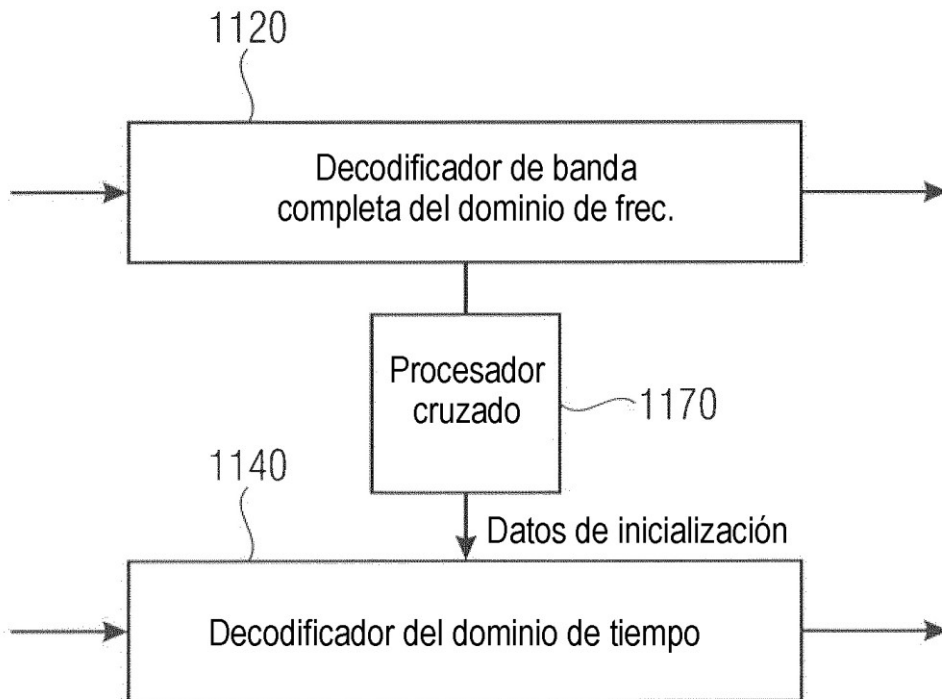


FIG 11B

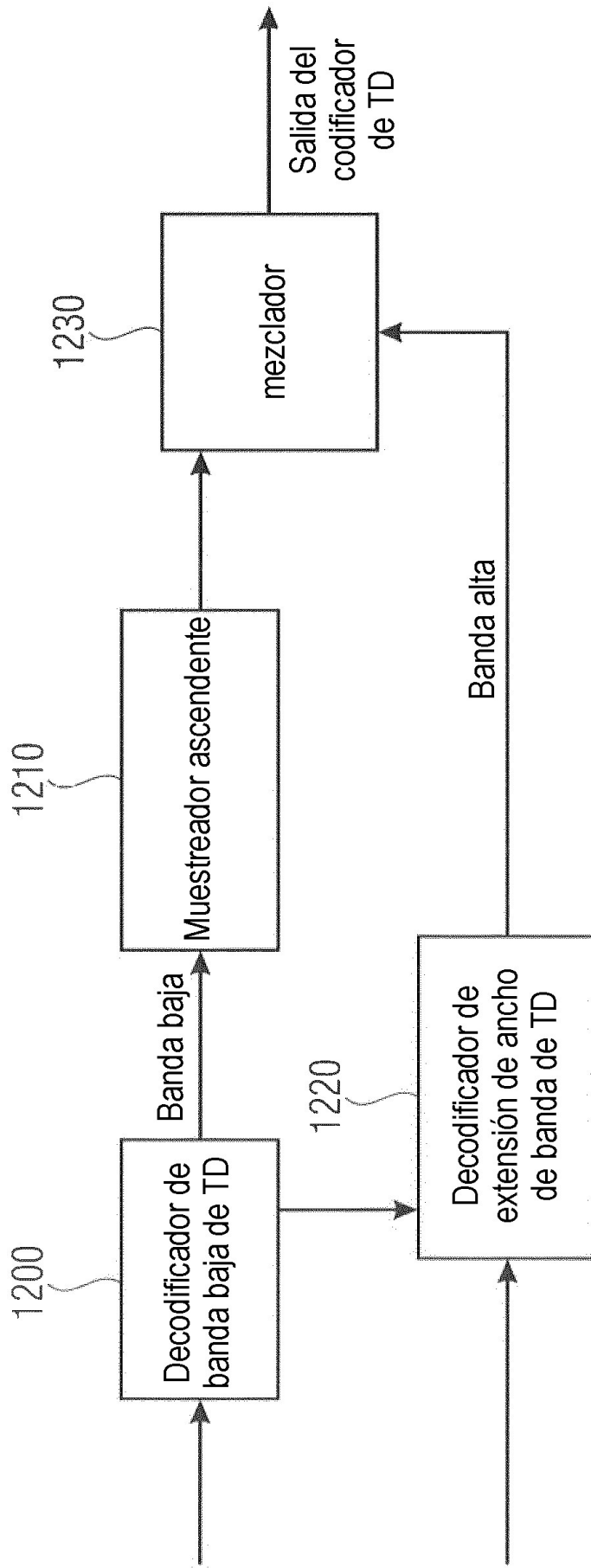


FIG 12

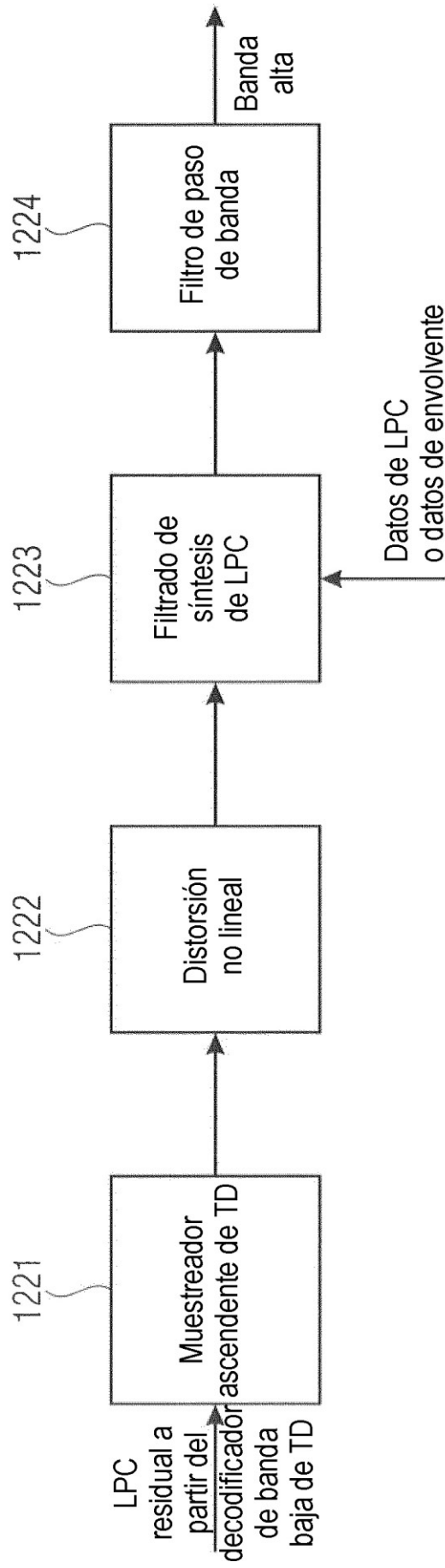


FIG 13

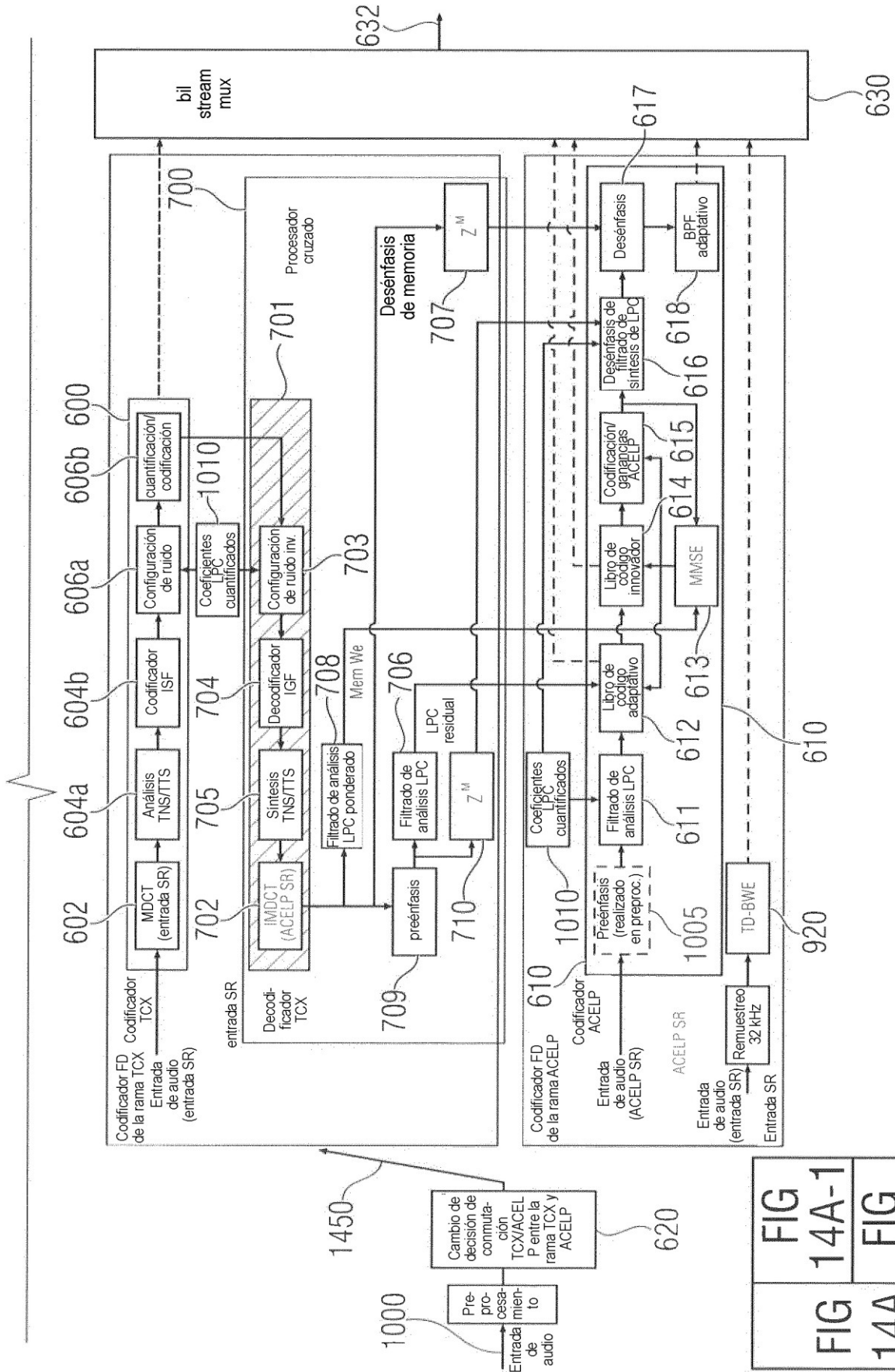
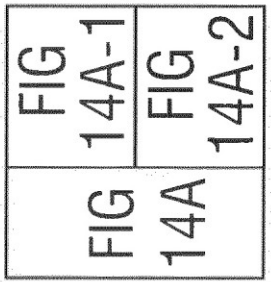


FIG 14A-2



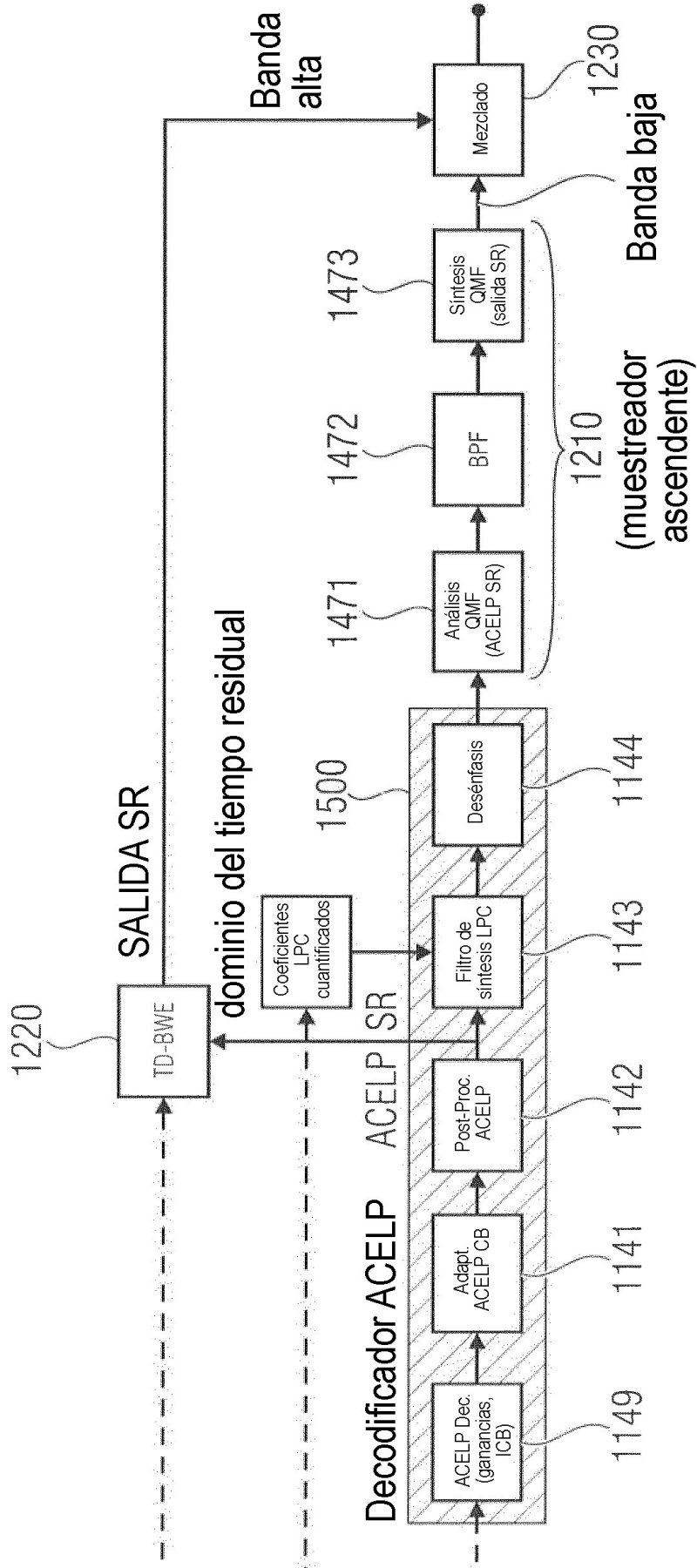


FIG 14C