

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 727 462**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.01.2017 PCT/EP2017/051208**

87 Fecha y número de publicación internacional: **27.07.2017 WO17125559**

96 Fecha de presentación y número de la solicitud europea: **20.01.2017 E 17700706 (9)**

97 Fecha y número de publicación de la concesión europea: **06.03.2019 EP 3284087**

54 Título: **Aparatos y procedimientos para la codificación o decodificación de una señal multicanal de audio mediante el uso de repetición de muestreo de dominio espectral**

30 Prioridad:

22.01.2016 EP 16152450
22.01.2016 EP 16152453

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
16.10.2019

73 Titular/es:

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

72 Inventor/es:

FUCHS, GUILLAUME;
RAVELLI, EMMANUEL;
MULTRUS, MARKUS;
SCHNELL, MARKUS;
DÖHLA, STEFAN;
DIETZ, MARTIN;
MARKOVIC, GORAN;
FOTOPOULOU, ELENI;
BAYER, STEFAN y
JÄGERS, WOLFGANG

74 Agente/Representante:

SALVÀ FERRER, Joan

ES 2 727 462 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

Aparatos y procedimientos para la codificación o decodificación de una señal multicanal de audio mediante el uso de repetición de muestreo de dominio espectral

5

[0001] La presente solicitud se refiere al procesamiento estéreo o, en términos generales al procesamiento multicanal, en donde una señal multicanal tiene dos canales tales como un canal izquierdo y un canal derecho en el caso de una señal estéreo o más de dos canales, tales como tres, cuatro, cinco o cualquier otro número de canales.

10 **[0002]** El habla estéreo y, particularmente, el habla estéreo de habla de conversación, ha recibido mucha menos atención científica que el almacenamiento y radiodifusión de música estereofónica. De hecho, hoy en día en las comunicaciones de habla la transmisión monofónica es la más utilizada. Sin embargo, al aumentar el ancho de banda y la capacidad de la red, se prevé que las comunicaciones basadas en tecnologías estereofónicas se volverán más populares y ocasionarán una mejor experiencia de escucha.

15

[0003] La codificación eficaz del material de audio estereofónico ha sido objeto de estudio desde hace mucho tiempo en la codificación audio perceptual de música, con vistas a un almacenamiento o difusión eficaces. Con elevadas velocidades de bits, en las que la conservación de la forma de onda es esencial, el estéreo de suma-diferencia, conocido como estéreo M/S (lado/medio), ha sido utilizado durante mucho tiempo. Para bajas
20 velocidades de bits, se ha introducido el estéreo de intensidad y, más recientemente, la codificación estéreo paramétrica. Esta última técnica ha sido adoptada en diferentes estándares tales como HeAACv2 y Mpeg USAC. Genera un mezclado de forma descendente de la señal de doble canal y asocia la información compacta del lado espacial.

25 **[0004]** Usualmente, la codificación estéreo conjunta se efectúa con resolución de elevada frecuencia, es decir, con una baja resolución en tiempo, y una transformación de la señal en tiempo-frecuencia, por lo que no es compatible con un bajo retardo ni con el procesamiento de dominio de tiempo llevado a cabo en la mayoría de los codificadores de habla. Además, la velocidad de bits generada es usualmente elevada.

30 **[0005]** Por otra parte, el estéreo paramétrico utiliza un banco de filtros extra posicionado en el extremo frontal del codificador como pre-procesador y en el extremo posterior del decodificador como post-procesador. Por lo tanto, el estéreo paramétrico puede utilizarse con codificadores de habla convencionales como ACELP como se hace en MPEG USAC. Además, la parametrización de la escena del auditorio puede lograrse con una mínima cantidad de información lateral, lo que es conveniente para bajas velocidades de bits. Sin embargo, el estéreo paramétrico, como
35 se da por ejemplo en MPEG USAC, no está específicamente diseñado para un bajo retardo y no proporciona una calidad constante para diferentes escenarios de conversaciones. En la representación paramétrica convencional de la escena espacial, el ancho de la imagen estéreo se reproduce de manera artificial por un decorrelacionador aplicado sobre los dos canales sintetizados y controlados por parámetros de coherencia entre canales (ICs, Interchannel Coherence) computados y transmitidos por el codificador. Para la mayoría del habla estero, esta
40 manera de ensanchar la imagen estéreo no es adecuada para recrear el ambiente natural del habla que es un sonido bastante directo dado que es producido por una única fuente situada en una posición específica en el espacio (con a veces alguna reverberación debida al ambiente interior). En cambio, los instrumentos musicales tienen un ancho mucho más natural que el habla, el cual puede ser mejor imitado mediante la decorrelación de los canales.

45 **[0006]** También se presenten problemas cuando se registra el habla con micrófonos no coincidentes, tales como en la configuración A-B cuando los micrófonos están separados entre sí o para el registro o renderización binaurales. Pueden preverse estos escenarios para capturar el habla en teleconferencias o para crear una escena de auditorio virtual con locutores distantes en la unidad de control multipunto (MCU, multipoint control unit). El momento de la llegada de la señal es en tal caso diferente de un canal a otro a diferencia de las grabaciones efectuadas en
50 micrófonos coincidentes tales como X-Y (registro de intensidad) o M-S (registro lado-medio). El cálculo de la coherencia de tales dos canales no alineados en el tiempo puede en tal caso estimarse equivocadamente, lo que hace que la síntesis del ambiente artificial falle.

[0007] Las referencias del estado anterior de la técnica relacionadas con el procesamiento estéreo son la
55 Patente de EE. UU N.º 5.434.948 o la Patente de EE.UU. N.º 8.811.621.

[0008] En el documento WO 2006/089570 A1 se describe un esquema de codificador/decodificador multicanal casi transparente o transparente. Adicionalmente, un esquema de codificador/decodificador multicanal genera una señal residual de tipo forma de onda. Esta señal residual se transmite junto con uno o más parámetros multicanal a un decodificador. A diferencia de un decodificador multicanal puramente paramétrico, el decodificador reforzado genera una señal de salida multicanal que tiene una calidad de salida mejorada debido a la señal residual adicional. En el lado del codificador, un canal izquierdo y un canal derecho son filtrados, ambos, por un banco de filtros de análisis. En tal caso, para cada señal de subbanda, se calcula un valor de alineación y un valor de ganancia para una superbanda. Una alineación de este tipo se lleva entonces a cabo antes de un procesamiento
60 ulterior. En el lado del decodificador, se lleva a cabo una desalineación y un procesamiento de ganancia, y las
65

correspondientes señales son seguidamente sintetizadas por un banco de filtros de síntesis con el fin de generar una señal izquierda decodificada y una señal derecha decodificada.

Por otra parte, el estéreo paramétrico utiliza un banco de filtros extra posicionado en el extremo frontal del codificador como pre-procesador y en el extremo posterior del decodificador como post-procesador. Por ello, el estéreo paramétrico puede utilizarse con codificadores de estéreo convencionales tales como ACELP ya que se efectúa en MPEG USAC. Además, la parametrización de la escena de auditorio puede efectuarse con una cantidad mínima de información lateral, lo que es conveniente para bajas velocidades de bits. Sin embargo, el estéreo paramétrico como por ejemplo en MPEG USAC, no está diseñado para un bajo retardo, y el sistema en su conjunto muestra un retardo algorítmico muy elevado. Se conoce según la solicitud de patente internacional WO2016108655A1 un procedimiento de codificación con conversión de tasa de muestreo antes de un codificador envolvente MPEG o después del codificador envolvente MPEG.

[0009] Se conoce según la solicitud de patente de EE. UU2014/0032226A1 un procedimiento que implica el remuestreo en un dominio de frecuencia.

[0010] Es un objetivo de la presente invención proporcionar un concepto mejorado para la codificación/decodificación multicanal, que sea eficaz y con la capacidad de proporcionar un bajo retardo.

Este objetivo se logra mediante un aparato para la codificación de una señal multicanal según la reivindicación 1, un procedimiento de codificación de una señal multicanal según la reivindicación 24, un aparato para la decodificación de una señal multicanal codificada según la reivindicación 25, un procedimiento de decodificación de una señal multicanal codificada según la reivindicación 42 o un programa informático según la reivindicación 43.

La presente invención se basa en el hallazgo de que al menos una porción y preferentemente todas las partes del procesamiento multicanal, es decir, un procesamiento multicanal conjunto, se llevan a cabo en un dominio espectral. Específicamente, se prefiere llevar a cabo la operación de mezclado de forma descendente del procesamiento multicanal conjunto en el dominio espectral y, adicionalmente, las operaciones temporales y de alineación de fase o incluso los procedimientos para analizar los parámetros para el procesamiento estéreo conjunto/multicanal conjunto. Adicionalmente, el muestreo repetido en el dominio espectral se lleva a cabo ya sea subsiguientemente al procesamiento multicanal o incluso antes del procesamiento multicanal con el fin de proporcionar una señal de salida desde un convertidor espectral-tiempo adicional que ya se encuentre en una velocidad de muestreo de salida requerido por un codificador de núcleo subsiguientemente conectado.

En el lado del decodificador, se prefiere una vez más llevar a cabo al menos una operación para generar una señal del primer canal y una señal de un segundo canal a partir de una señal de mezclado de forma descendente en el dominio espectral y, es preferible para llevar a cabo incluso el procesamiento multicanal inverso total en el dominio espectral. Además, se proporciona el convertidor tiempo-espectral para convertir la señal decodificada de núcleo en una representación en el dominio espectral y, dentro del dominio de las frecuencias, se lleva a cabo el procesamiento multicanal inverso. Se lleva a cabo un muestreo repetido en el dominio espectral ya sea antes del procesamiento multicanal inverso o se lleva a cabo subsiguientemente al procesamiento multicanal inverso de tal manera que, al final, un convertidor espectral-tiempo convierte una señal espectralmente muestreada en el dominio del tiempo con una velocidad de muestreo de salida que está destinada a la señal de salida en el dominio del tiempo.

Por ello, la presente invención permite evitar por completo cualquier operación de muestreo repetido en el dominio del tiempo, computacionalmente intensivas. En cambio, se combina el procesamiento multicanal con el muestreo repetido. En formas de realización preferidas, el muestreo repetido se lleva a cabo ya sea truncando el espectro en el caso del muestreo descendente o se lleva a cabo mediante el padding cero del espectro en el caso del muestreo ascendente. Estas operaciones fáciles, es decir, la truncación del espectro por una parte o el padding cero del espectro por otra parte, y las puestas en escala adicionales preferibles con el fin de tener en cuenta determinadas operaciones de normalización llevadas a cabo en los algoritmos de conversión dominio espectral/dominio del tiempo tales como el algoritmo de DFT o FFT, completan la operación de muestreo repetido en el dominio espectral de una manera muy eficaz y con bajo retardo.

Además, se ha descubierto que al menos una porción o incluso el procesamiento estéreo conjunto/procesamiento multicanal conjunto, completo, en el lado del codificador y el correspondiente procesamiento multicanal inverso en el lado del decodificador, es adecuado para ser ejecutado en el dominio de las frecuencias. Esto es válido no solamente para la operación de mezclado de forma descendente como un procesamiento multicanal conjunto mínimo en el lado del codificador o para un procesamiento de mezclado de forma ascendente como para un procesamiento multicanal inverso mínimo en el lado del decodificador. En cambio, también pueden llevarse a cabo un análisis estéreo de escena y alineaciones del tiempo/fase en el lado del codificador o desalineaciones de fase y tiempo en el lado del decodificador, en el dominio espectral. Lo mismo se aplica para la codificación de canal lateral preferentemente llevada a cabo en el lado del codificador o para síntesis de canal lateral y para la utilización en la generación de los dos canales de salida decodificados en el lado del decodificador.

[0013] Por ello, una ventaja de la presente invención es la de proporcionar un nuevo esquema de codificación de estéreo mucho más adecuado para la conversión de un habla estéreo que los esquemas de codificación estéreo existentes. Las formas de realización de la presente invención proporcionan un nuevo armazón para lograr un códec estéreo de bajo retardo y para integrar una herramienta estéreo en común llevada a cabo en el dominio de las 5 frecuencias tanto para un codificador de núcleo del habla como para un codificador de núcleo basado en MDCT dentro de un códec de audio conmutado.

[0014] Las formas de realización de la presente invención se refieren a una estrategia híbrida en la que se mezclan elementos de un estéreo M/S o estéreo paramétrico, convencional. Las formas de realización utilizan 10 algunos aspectos y herramientas de la codificación estéreo conjunta y otros de estéreo paramétrico. Más particularmente, las formas de realización adoptan el análisis y síntesis de tiempo-frecuencia extra efectuados en el extremo frontal del codificador y en el extremo posterior del decodificador. La descomposición de tiempo-frecuencia y la transformada inversa se logran mediante la utilización ya sea de un banco de filtros o de una transformada de bloque con valores complejos. Desde la entrada de dos canales o de múltiples canales, el procesamiento estéreo o 15 multicanal combina y modifica los canales de entrada a los canales de salida que llevan la designación de señales media y lado (MS, Mid and Side).

[0015] Las formas de realización de la presente invención proporcionan una solución para reducir un retardo algorítmico introducido por un módulo estéreo y particularmente a partir de la estructura y puesta en ventana de su 20 banco de filtros. Proporciona una transformada inversa multi-coeficiente para alimentar un codificador conmutado tal como 3GPP EVS o una conmutación de codificador entre un codificador de habla tal como ACELP y un codificador de audio genérico tal como TCX por el hecho de producir la misma señal de procesamiento estéreo con diferentes velocidades de muestreo. Además, proporciona una puesta en ventana adaptada para las diferentes restricciones del sistema de bajo retardo y baja complejidad así como para el procesamiento estéreo. Además, las formas de 25 realización proporcionan un procedimiento para combinar y muestrear repetidamente diferentes resultados de síntesis decodificados en el dominio espectral, donde también se aplica el procesamiento estéreo inverso.

[0016] Las formas de realización preferidas de la presente invención comprenden una multifunción en un remuestreador en el dominio espectral que no solamente genera un bloque individual remuestreado en el dominio 30 espectral de valores espectrales sino, adicionalmente, también otra secuencia remuestreada de bloques de valores espectrales correspondientes a una velocidad de muestreo diferente, más elevada o más baja.

[0017] Además, el codificador multicanal está configurado para proporcionar adicionalmente una señal de salida a la salida del convertidor espectral-tiempo que tiene la misma velocidad de muestreo que la señal de canal 35 primero y segundo original introducida en el convertidor tiempo-espectral en el lado del codificador. Por lo tanto, en algunas formas de realización, el codificador multicanal proporciona al menos una señal de salida con la velocidad de muestreo introducida original, que se utiliza preferentemente para una codificación basada en MDCT. Adicionalmente, se proporciona al menos una señal de salida con una velocidad de muestreo intermedia que es específicamente útil para la codificación de ACELP y que proporciona adicionalmente otra señal de salida a otra 40 velocidad de muestreo de salida que también es útil para la codificación de ACELP, pero que es diferente de la otra velocidad de muestreo de salida.

[0018] Estos procedimientos pueden llevarse a cabo ya sea para la Señal media (Mid) o para la señal lado (Side), o para ambas señales derivadas de las señales de los canales primero y segundo de una señal multicanal en 45 donde la primera señal también puede ser una señal izquierda y la segunda señal puede ser una señal derecha en el caso de una señal estéreo que solamente tiene dos canales (adicionalmente dos, por ejemplo, un canal de refuerzo de baja frecuencia).

[0019] En otras formas de realización, el codificador de núcleo del codificador multicanal está configurado 50 para operar según un control de estructura, y el convertidor tiempo-espectral y el convertidor espectro-tiempo del post-procesador estéreo y remuestreador están configurados también para operar según otro control de estructura que está sincronizado con el control de control de estructura del codificador de núcleo. La sincronización se lleva a cabo de tal manera que un borde de inicio de la trama o un borde final de la trama de cada trama de una secuencia de tramas del codificador de núcleo se halla en una relación predeterminada con respecto a un instante de inicio o 55 con un instante final de una porción superpuesta de una ventana utilizada por el convertidor tiempo-espectral o por el convertidor espectral tiempo para cada bloque de la secuencia de bloques de valores de muestreo para cada bloque de la secuencia remuestreada de bloques de valores espectrales. Por lo tanto, se asegura que las operaciones de estructura subsiguientes operen de manera sincronizada entre sí.

[0020] En otras formas de realización, el codificador de núcleo lleva a cabo una operación de anticipación 60 mediante una porción de anticipación. En esta forma de realización, se prefiere que la porción de anticipación sea utilizada también por una ventana de análisis del convertidor tiempo-espectral en donde se utiliza una porción superpuesta de la ventana de análisis que tiene una longitud en tiempo que es inferior o igual a la longitud en tiempo de la porción de anticipación.

65

[0021] Por lo tanto, haciendo que la porción de anticipación del codificador de núcleo y la porción de superposición de la ventana de análisis sean iguales entre sí o haciendo que la porción de superposición sea aún más pequeña que la porción de anticipación del codificador de núcleo, no puede implementarse el análisis tiempo-espectral del pre-procesador estéreo sin algún retardo algorítmico adicional. Con el fin de asegurar que esta porción
5 puesta en ventana de anticipación no influya excesivamente sobre la funcionalidad de anticipación del codificador de núcleo, se prefiere compensar esta porción utilizando una inversa de la función de la ventana de análisis.

[0022] Con el fin de asegurar que esto se lleve a cabo con una buena estabilidad, se utiliza una raíz cuadrada de forma de ventana seno en lugar de una forma de ventana seno como una ventana de análisis y se utiliza un seno
10 a la potencia de 1,5 ventana de síntesis a los efectos de la puesta en ventana de síntesis antes de llevar a cabo la porción de superposición a la salida del convertidor espectral-tiempo. Por lo tanto, se asegura que la función de compensación supone valores que están reducidos con respecto a sus magnitudes en comparación con una función de compensación que es la inversa de una función seno.

[0023] Sin embargo, en el lado del decodificador se prefiere utilizar las mismas formas de ventana de análisis y de síntesis, ya que no se requiere ninguna compensación, por supuesto. Por otra parte, se prefiere utilizar un espacio de tiempo en el lado del decodificador, existiendo el espacio de tiempo entre un extremo de una porción superpuesta principal de una ventana de análisis del convertidor tiempo-espectral en el lado del decodificador y un instante de tiempo en el extremo de una salida de trama por el decodificador de núcleo en el lado del decodificador
20 multicanal. Por lo tanto, las muestras de salida del decodificador de núcleo dentro de este espacio de tiempo no se requieren de inmediato para los fines de la puesta en ventana de análisis por el post-procesador estéreo, pero se requieren únicamente para el procesamiento/puesta en ventana de la trama siguiente. Un espacio de tiempo de este tipo puede implementarse por ejemplo mediante el uso de una porción no superpuesta típicamente situada en el medio de una ventana de análisis, lo que resulta en un acortamiento de la porción superpuesta. Sin embargo,
25 también pueden utilizarse otras alternativas para implementar un espacio de tiempo de este tipo, pero se prefiere implementar el espacio de tiempo mediante la porción no superpuesta en el medio. Por lo tanto, este espacio de tiempo puede utilizarse para otras operaciones del decodificador de núcleo o para suavizar las operaciones entre acontecimientos que preferentemente son de conmutación cuando el decodificador de núcleo conmuta de un dominio de las frecuencias a una trama en el dominio del tiempo o para cualquier otra operación de suavización que
30 puedan ser útiles cuando han tenido lugar los cambios de parámetros o los cambios de características de la codificación.

[0024] A continuación, se exponen en detalle formas de realización preferidas de la presente invención con respecto a los dibujos adjuntos, en los que:

- 35 la figura 1 es un diagrama de bloques de una forma de realización del codificador multicanal;
- la figura 2 ilustra formas de realización del muestreo repetido en el dominio espectral;
- las figuras 3a-3c ilustran diferentes alternativas para llevar a cabo conversiones de tiempo/frecuencia o de frecuencia/tiempo con diferentes normalizaciones y correspondientes puestas en escala en el dominio espectral;
- la figura 3d ilustra diferentes resoluciones de frecuencia y otros aspectos relacionados con la frecuencia, para
40 determinadas formas de realización;
- la figura 4a ilustra un diagrama de bloques de una forma de realización de un codificador;
- la figura 4b ilustra un diagrama de bloques de una correspondiente forma de realización de un decodificador;
- la figura 5 ilustra una forma de realización preferida de un codificador multicanal;
- la figura 6 ilustra un diagrama de bloques de una forma de realización de un decodificador multicanal;
- 45 la figura 7a ilustra otra forma de realización de un decodificador multicanal que comprende un combinador;
- la figura 7b ilustra otra forma de realización de un decodificador multicanal que comprende adicionalmente el combinador (adición);
- la figura 8a ilustra una ventana que ilustra diferentes características de ventana para varias velocidades de muestreo;
- 50 la figura 8b ilustra diferentes propuestas/formas de realización para un banco de filtros DFT como una implementación del convertidor tiempo-espectral y de convertidor espectro-tiempo;
- la figura 8c ilustra una secuencia de dos ventanas de análisis de un DFT con una resolución del tiempo de 10 ms;
- la figura 9a ilustra una puesta en ventana esquemática de codificador según una primera propuesta/forma de realización;
- 55 la figura 9b ilustra una puesta en ventana esquemática de decodificador según la primera propuesta/forma de realización;
- la figura 9c ilustra las ventanas en el codificador y en el decodificador según la primera propuesta/forma de realización;
- la figura 9d ilustra un diagrama de flujo preferido que ilustra la forma de realización de compensación;
- 60 la figura 9e ilustra una forma de realización que ilustra en mayor grado la forma de realización de compensación;
- la figura 9f ilustra un diagrama de flujo para explicar la forma de realización del lado de decodificador del espacio de tiempo;
- la figura 10a ilustra una puesta en ventana esquemática de codificador según la cuarta propuesta/forma de realización;
- 65 la figura 10b ilustra una puesta en ventana esquemática de decodificadora según la cuarta propuesta/forma de

realización;

la figura 10c ilustra ventanas y en el codificador y en el decodificador según la cuarta propuesta/forma de realización;
la figura 11a ilustra una puesta en ventana esquemática de un codificador según la quinta propuesta/forma de realización;

- 5 la figura 11b ilustra una puesta en ventana esquemática de un decodificador según la quinta propuesta/forma de realización;
la figura 11c ilustra el codificador y el decodificador según la quinta propuesta/forma de realización;
la figura 12 es un diagrama de bloques de una implementación preferida del procesamiento multicanal que utiliza un mezclado de forma descendente en el procesador de señales;
- 10 la figura 13 es una forma de realización preferida del procesamiento multicanal inverso con una operación de mezclado de forma ascendente dentro del procesador de señales;
la figura 14a ilustra un diagrama de flujo de procedimientos llevados a cabo en el aparato para codificación con fines de alineación de los canales;
la figura 14b ilustra una forma de realización preferida de procedimientos llevados a cabo en el dominio de las
- 15 frecuencias;
la figura 14c ilustra una forma de realización preferida de procedimientos llevados a cabo en el aparato para la codificación mediante el uso de una ventana de análisis con porciones de padding cero e intervalos de superposición;
la figura 14d ilustra un diagrama de flujo para otros procedimientos llevados a cabo dentro de una forma de
- 20 realización del aparato para codificación;
la figura 15a ilustra procedimientos llevados a cabo mediante una forma de realización del aparato para decodificación y codificación de señales multicanal;
la figura 15b ilustra una implementación preferida del aparato para decodificación con respecto a algunos aspectos;
y
- 25 la figura 15c ilustra un procedimiento llevado a cabo en el contexto de desalineación de banda ancha en la estructura de la decodificación de una señal multicanal codificada.

[0025] La figura 1 ilustra un aparato para la codificación de una señal multicanal que comprende al menos dos canales 1001, 1002. El primer canal 1001 es el canal izquierdo, y el segundo canal 1002 pueden ser un canal

30 derecho en el caso de un escenario estéreo de dos canales. Sin embargo, en el caso de un escenario multicanal, el primer canal 1001 y el segundo canal 1002 puede ser cualquiera de los canales de la señal multicanal tales como, por ejemplo, el canal izquierdo por una parte y el canal envolvente izquierdo por otra parte o el canal derecho por una parte y el canal envolvente derecho por otra parte. Sin embargo, estos emparejamientos de canales, son solamente ejemplos, y es posible aplicar otros emparejamientos de canales en función del caso.

[0026] El codificador multicanal de la figura 1 comprende un convertidor tiempo-espectral para convertir secuencias de bloques de valores de muestreo de los al menos dos canales en una representación en el dominio de las frecuencias a la salida del convertidor tiempo-espectral. Cada representación en el dominio de las frecuencias tiene una secuencia de bloques de valores espectrales para uno de los al menos dos canales. Particularmente, un

40 bloque de valores de muestreo del primer canal 1001 o del segundo canal 1002 tiene una velocidad de muestreo de entrada asociada, y un bloque de valores espectrales de las secuencias de la salida del convertidor tiempo-espectral tiene valores espectrales hasta una frecuencia de entrada máxima que está relacionada con la velocidad de muestreo de entrada. En la forma de realización ilustrada en la figura 1, el convertidor tiempo-espectral está conectado al procesador multicanal 1010. Este procesador multicanal está configurado para aplicar un

45 procesamiento multicanal conjunto a las secuencias de bloques de valores espectrales para obtener al menos una secuencia de bloques de valores espectrales resultado que comprende información relacionada con los al menos dos canales. Una operación de procesamiento multicanal típico es una operación de mezclado de forma descendente, pero la operación multicanal preferida comprende procedimientos adicionales que serán descritos más abajo.

[0027] En una forma de realización alternativa, el procesador multicanal 1010 está conectado a un remuestreador en el dominio espectral 1020, y una salida del remuestreador en el dominio espectral 1020 se introduce en el procesador multicanal. Esto se ilustra mediante las líneas de conexión discontinuas 1021, 1022. En esta forma de realización alternativa, el procesador multicanal está configurado para aplicar el procesamiento

55 multicanal conjunto no a las secuencias de bloques de valores espectrales como salida por el convertidor tiempo-espectral, sino secuencias remuestreadas de bloques disponibles sobre las líneas de conexión 1022.

[0028] El remuestreador en el dominio espectral 1020 está configurado para un muestreo repetido de la secuencia resultado generada por el procesador multicanal o para remuestrear las secuencias de bloques emitidas

60 por el convertidor tiempo-espectral 1000 para obtener una secuencia remuestreada de bloques de valores espectrales que pueden representar una señal media (Mid) tal como la ilustrada en la línea 1025. Es preferible que el remuestreador en el dominio espectral lleve a cabo adicionalmente el muestreo repetido hacia la señal lateral (Side) por medio del procesador multicanal y, por ello, también emita una secuencia remuestreada correspondiente a la señal lateral (Side) como se ilustra en 1026. Sin embargo, la generación y muestreo repetido de la señal lateral

65 (Side) es opcional y no se requiere para una implementación con una baja velocidad de bits. Es preferible que el

remuestreador en el dominio espectral 1020 esté configurado para truncar bloques de valores espectrales a los fines del muestreo descendente o para el padding cero de los bloques de valores espectrales con fines del muestreo ascendente. El codificador multicanal comprende adicionalmente un convertidor espectral-tiempo para convertir la secuencia remuestreada de bloques de valores espectrales en una representación en el dominio del tiempo que
 5 comprende una secuencia de salida de bloques de valores de muestreo asociados con una velocidad de muestreo de salida que es diferente de la velocidad de muestreo de entrada. En formas de realización alternativas, en las que el muestreo repetido en el dominio espectral se lleva a cabo antes del procesamiento multicanal, el procesador multicanal proporciona la secuencia resultado por medio de la línea discontinua 1023 directamente al convertidor espectral-tiempo 1030. En esta forma de realización alternativa, una característica opcional es que, adicionalmente,
 10 la señal lateral (Side) es generada por el procesador multicanal ya en la representación muestreada y la señal lateral (Side) es procesada también seguidamente por el convertidor espectral-tiempo.

[0029] Al final, el convertidor espectral-tiempo proporciona preferentemente una señal media (Mid) en el dominio del tiempo 1031 y una señal lateral (Side) en el dominio del tiempo, opcional, 1032, las cuales pueden ser
 15 ambas codificadas en núcleo por el codificador de núcleo 1040. En términos generales, el codificador de núcleo está configurado para una codificación núcleo de la secuencia de bloques emitida de valores de muestreo para obtener la señal multicanal codificada.

[0030] La figura 2 ilustra diagramas espectrales que son útiles para explicar el muestreo repetido en el
 20 dominio espectral.

[0031] El diagrama superior en la figura 2 ilustra un espectro de canal disponible a la salida del convertidor tiempo-espectral 1000. Este espectro 1210 tiene valores espectrales hasta la frecuencia de entrada máxima 12w11. En el caso del muestreo ascendente, se lleva a cabo un padding cero dentro de la porción de padding cero o de la
 25 región de padding cero 1220 que se extiende hasta la frecuencia de salida máxima 1221. La frecuencia de salida máxima 1221 es mayor que la frecuencia de entrada máxima 1211, dado que se prevé un muestreo ascendente.

[0032] A diferencia de lo anterior, el diagrama más bajo en la figura 2 ilustra los procedimientos en los que se
 30 incurre por medio del muestreo descendente de una secuencia de bloques. A tal efecto, se trunca un bloque dentro de una región truncada 1230 de tal manera que una frecuencia máxima de salida del espectro truncado en 1231 es inferior a la frecuencia de entrada máxima 1211.

[0033] Típicamente, la velocidad de muestreo asociada con un espectro correspondiente en la figura 2 es de
 35 al menos 2x la frecuencia máxima del espectro. Por lo tanto, para el caso superior en la figura 2, la velocidad de muestreo será de al menos 2 veces la frecuencia de entrada máxima 1211.

[0034] En el segundo diagrama de la figura 2, la velocidad de muestreo será de al menos dos veces la
 40 frecuencia máxima de salida 1221, es decir, la frecuencia más elevada de la región de padding cero 1220. A diferencia de lo anterior, en el diagrama más bajo en la figura 2, la velocidad de muestreo será de al menos 2x la frecuencia de salida máxima 1231, es decir, el valor espectral más elevado restante subsiguiente a una truncación dentro de la región truncada 1230.

[0035] Las figuras 3a a 3c ilustran diversas alternativas que pueden utilizarse en el contexto de determinados
 45 algoritmos de transformada DFT directa o inversa. En la figura 3a, se considera una situación en la que se lleva a cabo una DFT de magnitud x , y en la que no tiene lugar ninguna normalización en el algoritmo de transformada directa 1311. En el bloque 1331, se ilustra una transformada inversa con una magnitud y diferente, en donde se lleva a cabo una normalización con $1/N_y$. N_y es el número de valores espectrales de la transformada inversa con la magnitud y . En tal caso, se prefiere llevar a cabo una puesta en escala por N_y/N_x como se ilustra en el bloque 1321.

[0036] A diferencia de lo anterior, en la figura 3b se ilustra una implementación, en la que la normalización
 50 está distribuida a la transformada directa 1312 y a la transformada inversa 1332. Seguidamente se requiere una puesta en escala como se ilustra en el bloque 1322, en la que es útil una raíz cuadrada de la relación entre el número de valores espectrales de la transformada inversa y el número de valores espectrales de la transformada directa.

[0037] En la figura 3c se ilustra otra implementación, en la que se lleva a cabo la normalización completa
 55 sobre la transformada directa en donde se lleva a cabo la transformada directa con la magnitud x . Seguidamente, la transformada inversa ilustrada en el bloque 1333 opera sin ninguna normalización por lo que no se requiere ninguna puesta en escala como se ilustra mediante el bloque 1323 en la figura 3c. Por lo tanto, en función de determinados algoritmos, se requieren determinadas operaciones de puesta en escala o incluso ninguna operación de puesta en escala. Sin embargo, se prefiere actuar según la figura 3a.

[0038] Con el fin de mantener bajo el retardo global, la presente invención proporciona un procedimiento en
 65 el lado del codificador para evitar la necesidad de un remuestreador en el dominio del tiempo y mediante su reemplazo por el muestreo repetido de las señales en el dominio de DFT. Por ejemplo, en el EVS permite ahorrar

0,9375 ms de retardo debido al remuestreador en el dominio del tiempo. El muestreo repetido en el dominio de las frecuencias se logra mediante el padding cero o truncación del espectro y su puesta en escala correcta.

[0039] Consideremos una señal puesta en ventana de entrada x muestreada con una velocidad f_x con un espectro X de magnitud N_x y una versión de la misma señal remuestreada con la velocidad f_y con un espectro de magnitud N_y . El factor de muestreo es en tal caso igual a:

$$f_y/f_x = N_y/N_x$$

10 en el caso del muestreo descendente $N_x > N_y$. El muestreo descendente puede llevarse a cabo de manera simple en el dominio de las frecuencias mediante puesta en escala directa y truncación del espectro X original:

$$Y[k] = X[k \cdot N_y/N_x] \text{ para } k=0..N_y$$

15 en el caso del muestreo ascendente $N_x < N_y$. El muestreo de forma ascendente puede llevarse a cabo simplemente en el dominio de las frecuencias mediante puesta en escala directa y padding cero del espectro X original:

$$Y[k] = X[k \cdot N_y/N_x] \text{ para } k=0..N_x$$

$$Y[k] = 0 \text{ para } k= N_x..N_y$$

20 **[0040]** Ambas operaciones de remuestreo pueden resumirse del siguiente modo:

$$Y[k] = X[k \cdot N_y/N_x] \text{ para todos los } k=0..min(N_y, N_x)$$

$$Y[k] = 0 \text{ para todos los } k= min(N_y, N_x)..N_y \text{ para si } N_y > N_x$$

25 **[0041]** Una vez que se ha obtenido el nuevo espectro Y , es posible obtener la señal en el dominio del tiempo y mediante la aplicación de la transformada inversa $iDFT$ asociada de magnitud N_y :

$$y = iDFT(Y)$$

30 **[0042]** Para construir la señal de tipo continua sobre diferentes tramas, se pone en ventana la trama de salida y y a continuación se superpone a la trama previamente obtenida.

[0043] La forma de ventana es igual para todas las velocidades de muestreo, pero las ventanas tienen diferentes tamaños en cuanto a las muestras y velocidades de muestreo. El número de las muestras de las ventanas y sus valores pueden derivarse fácilmente dado que la forma se define puramente de manera analítica. Las diferentes partes y tamaños de la ventana pueden encontrarse en la figura 8a como una función de la velocidad de muestreo específica. En este caso se utiliza una función seno en la parte de solapamiento (LA) para las ventanas de análisis y síntesis. Para estas regiones, los coeficientes ascending $ovlp_size$ están dados por:

$$win_ovlp(k) = \sin(\pi \cdot (k+0.5)/(2 \cdot ovlp_size));, \text{ para } k=0..ovlp_size-1$$

mientras que los coeficientes descending $ovlp_size$ están dados por:

$$win_ovlp(k) = \sin(\pi \cdot (ovlp_size-1-k+0.5)/(2 \cdot ovlp_size));, \text{ para } k=0..ovlp_size-1$$

45 donde $ovlp_size$ es en función de la velocidad de muestreo y se indica en la figura 8a.

[0044] La nueva codificación estéreo de bajo retardo es una codificación estéreo media/lateral (Mid/Side (M/S)) conjunta que aprovecha algunas claves espaciales, en donde el canal medio (Mid) está codificado por un codificador mono núcleo primario, y el canal lateral (Side) está codificado por un codificador de núcleo contrario. Los principios de codificador y de decodificador se ilustran en las figuras 4a y 4b.

[0045] El procesamiento estéreo se lleva a cabo principalmente en el dominio de las frecuencias (FD, frequency-domain). Opcionalmente puede llevarse a cabo parte del procesamiento estéreo en el dominio del tiempo (TD, Dominio del Tiempo (TD)) antes del análisis de frecuencia. Este es el caso para la computación de ITD, que puede computarse y aplicarse antes del análisis de frecuencia para alinear los canales en el tiempo antes de continuar con el análisis y procesamiento estéreo. Como alternativa, el procesamiento ITD puede llevarse a cabo directamente en el dominio de las frecuencias. Dado que los codificadores de habla usuales como ACELP no contienen ninguna descomposición interna de tiempo-frecuencia, la codificación estéreo añade un banco de filtros modulado complejo extra mediante un banco de filtros de análisis y síntesis antes del codificador de núcleo y de otra etapa de banco de filtros de análisis-síntesis después del codificador de núcleo. En la forma de realización preferida, se utiliza un DFT sobremuestreado con una región de baja superposición. Sin embargo, en otras formas de realización, pueden utilizarse cualquier descomposición valorada compleja de tiempo-frecuencia con una resolución temporal compleja. En lo que sigue con respecto a la banda de filtros estéreo, se hace referencia ya sea a un banco de filtros tal como QMF o a un bloque de transformada, DFT.

65

[0046] El procesamiento estéreo consiste en computar los indicios espaciales y/o los parámetros estéreo tales como la ITD (diferencia de tiempo entre canales, Diferencia de tiempo inter-canal), las IPDs (diferencias de fase inter-canal, Diferencias de tiempo entre canales), las ILDs (Diferencia de niveles entre canales, Diferencias de nivel inter-canal) y las ganancias de predicción para predecir la señal lateral (Side (S)) con la señal media (Mid(M)). Es importante observar que el banco de filtros estéreo tanto en el codificador como en el decodificador introduce un retardo extra en el sistema de codificación.

[0047] En la figura 4a se ilustra un aparato para la codificación de una señal multicanal en donde, en esta implementación, se lleva a cabo un cierto procesamiento estéreo conjunto en el dominio del tiempo mediante el uso de un análisis de diferencia de tiempo entre canales (ITD) y en donde el resultado de este análisis de ITD 1420 se aplica dentro del dominio del tiempo mediante el uso de un bloque de cambio de tiempo 1410 colocado delante de los convertidores de tiempo-espectros 1000.

[0048] A continuación, dentro del dominio espectral, se lleva a cabo otro procesamiento estéreo 1010 que implica, al menos un mezclado de forma descendente de izquierda y derecha de la señal media (Mid M), y, opcionalmente, el cálculo de una señal lateral (Side S) y, si bien no se ilustra explícitamente en la figura 4a, una operación de muestreo repetido llevada a cabo por el remuestreador en el dominio espectral 1020 ilustrado en la figura 1 que puede aplicar una de las dos alternativas diferentes, es decir, llevar a cabo el muestreo repetido subsiguiente al procesamiento multicanal o antes del procesamiento multicanal.

[0049] Además, en la figura 4a se ilustran detalles adicionales de un codificador de núcleo preferido 1040. En particular, con fines de codificación de la señal media (Mid) en el dominio del tiempo m a la salida del convertidor espectral-tiempo 1030, se utiliza un codificador EVS. Adicionalmente, se lleva a cabo una codificación MDCT 1440 y la cuantificación de vector subsiguientemente conectado 1450 con fines de la codificación de la señal lateral (Side).

[0050] La señal media (Mid), codificada o codificada por núcleo, y la señal lateral (Side), codificada por núcleo, son comunicadas a un multiplexador 1500 que multiplexa estas señales codificadas junto con información lateral. Un tipo de información lateral es la salida del parámetro ID en 1421 al multiplexador (y opcionalmente al elemento de procesamiento estéreo 1010), y otros parámetros se encuentran en los parámetros de diferencias/predicción de nivel de canal, diferencias de fase entre canales (parámetros IPD) o parámetros de llenado de estéreo como se ilustra en la línea 1422. De manera correspondiente, el aparato de la figura 4b para la decodificación de una señal multicanal representada por una corriente de bits 1510 comprende un demultiplexador 1520, un decodificador de núcleo que en esta forma de realización consiste en un decodificador EVS 1602 para la señal media (Mid) codificada y en un cuantificador del vector 1603 y un bloque MDCT inverso subsiguientemente conectado 1604. El bloque 1604 proporciona la señal lateral (Side) decodificada por núcleo. Las señales decodificadas m , s son seguidamente convertidas en el dominio espectral mediante el uso de los convertidores tiempo-espectral 1610, y, seguidamente, dentro del dominio espectral, se lleva a cabo el procesamiento estéreo inverso y muestreo repetido. De nuevo, en la figura 4b se ilustra una situación en la que se lleva a cabo el mezclado de forma ascendente de la señal M a izquierda L y derecha R y, adicionalmente, una desalineación de banda estrecha mediante el uso de parámetros IPD y, adicionalmente, otros procedimientos para calcular lo mejor posible los canales izquierdo y derecho mediante el uso de los parámetros de diferencia de nivel entre canales ILD y los parámetros de llenado de estéreo en la línea 1605. Además, el demultiplexador 1520 no solo extrae los parámetros en la línea 1605 de la corriente de bits 1510, sino que también extrae la diferencia de tiempo entre canales en la línea 1606 y comunica esta información al procesamiento estéreo inverso del bloque/remuestreador y, adicionalmente, a un procesamiento del cambio de tiempo inverso en el bloque 1650 que se lleva a cabo en el dominio del tiempo es decir, subsiguientemente al procedimiento llevado a cabo por los convertidores espectral-tiempo que proporcionan las señales izquierda y derecha decodificadas con la velocidad de salida, que es diferente de la velocidad en la salida del decodificador EVS 1602 o diferente de la velocidad del bloque MDCT 1604, por ejemplo.

[0051] El DFT estéreo puede proporcionar seguidamente diferentes versiones muestreadas de la señal que es comunicada seguidamente al codificador de núcleo conmutado. La señal para la codificación puede ser el canal medio (Mid), el canal lateral (Side), o los canales izquierdo y derecho, o cualquier señal resultante de una rotación o mapeo de canales de los dos canales de entrada. Dado que los diferentes codificadores de núcleo del sistema conmutado aceptan diferentes velocidades de muestreo, es un rasgo importante que el banco de filtros para la síntesis estéreo puede proporcionar una señal multivelocidad. El principio se indica en la figura 5.

[0052] En la figura 5, el módulo estéreo toma como entrada los dos canales de entrada, l y r , y los transforma en el dominio de las frecuencias en las señales M y S . En el procesamiento estéreo los canales de entrada pueden finalmente ser mapeados o modificados de manera que generen dos nuevas señales M y S . M es objeto de mayor codificación por el estándar 3GPP EVS mono o por una versión modificada de éste. Un codificador de este tipo es un codificador conmutado, que conmuta entre núcleos de MDCT (TCX y HQ-Core en el caso de EVS) y un codificador de habla (ACELP en EVS). También tiene una función de pre-procesamiento que se ejecuta todo el tiempo a 12,8 kHz y otras funciones de pre-procesamiento que se ejecutan con una velocidad de muestreo que varía según los modos operativos (12,8, 16, 25,6 o 32 kHz). Además, el ACELP se ejecuta ya sea a 12,8 o a 16 kHz,

mientras que los núcleos MDCT se ejecutan a la velocidad de muestreo de entrada. La señal S puede ser ya sea codificada por un codificador estándar EVS mono (o por una versión modificada del mismo), o por un codificador de señal lateral (Side) específico especialmente diseñado por sus características. También puede ser posible omitir la codificación de la señal lateral (Side S).

5

[0053] En la figura 5 se ilustran detalles del codificador estéreo preferido con banco de filtros de síntesis de múltiples velocidades de las señales estéreo procesadas M y S. La figura 5 muestra el convertidor tiempo-espectral 1000 que lleva a cabo una transformada de tiempo frecuencia con la velocidad de entrada, es decir, la velocidad que tienen las señales 1001 y 1002. Explícitamente, la figura 5 muestra adicionalmente un bloque de análisis en el dominio del tiempo 1000a, 1000e, para cada canal. En particular, aunque la figura 5 ilustra un bloque explícito de análisis en el dominio del tiempo, es decir, una puesta en ventana para aplicar una ventana de análisis al canal correspondiente, cabe observar que en otros lugares en esta memoria descriptiva, se considera que el aspecto para la puesta en ventana para aplicar el bloque de análisis en el dominio del tiempo está incluido en un bloque indicado como "convertidor tiempo-espectral" o "DFT" con alguna velocidad de muestreo. Además, y de manera correspondiente, la mención de un convertidor espectral-tiempo incluye típicamente, en la salida del algoritmo DFT real, un aspecto para la puesta en ventana para aplicar una correspondiente ventana de síntesis en donde, con el fin de obtener finalmente muestras de salida, se lleva a cabo una superposición-adición de bloques de valores de muestreo puestos en ventana con una correspondiente ventana de síntesis. Por ello, aunque el bloque 1030 solamente menciona un "IDFT" este bloque también designa una subsiguiente puesta en ventana de un bloque de muestras en el dominio del tiempo con una ventana de análisis y, de nuevo, una subsiguiente operación de superposición-adición con el fin de obtener finalmente la señal en el dominio del tiempo m.

[0054] Además, la figura 5 ilustra un bloque de análisis de escena estéreo específico 1011 que utiliza los parámetros utilizados en el bloque 1010 para llevar a cabo el procesamiento estéreo y de mezclado de forma ascendente, y estos parámetros pueden, por ejemplo, ser los parámetros en las líneas 1422 o 1421 de la figura 4a. Por lo tanto, en la implementación, el bloque 1011 puede corresponder al bloque 1420 en la figura 4a, en la que incluso el análisis de los parámetros, es decir, el análisis de la escena estéreo tiene lugar en el dominio espectral y, particularmente, con la secuencia de bloques de valores espectrales que no han sido remuestreadas, sino que se están en la frecuencia máxima correspondiente a la velocidad de muestreo de entrada.

30

[0055] Además, el decodificador de núcleo 1040 comprende una rama de codificador basada en MDCT 1430a y una rama codificadora de ACELP 1430b. En particular, el codificador medio (Mid) para las señales medias (Mid M) y, el correspondiente codificador lateral (Side) para las señales lateral (Side) lleva a cabo una codificación de conmutación entre una codificación basada en MDCT y una codificación ACELP donde, típicamente, el codificador de núcleo tiene adicionalmente un aspecto de decisión en modo de codificación que opera típicamente en una determinada porción de anticipación con el fin de determinar si un determinado bloque o trama debe ser codificado mediante el uso de procedimientos basados en MDCT o Procedimientos basados en ACELP. Además, o como alternativa, el codificador de núcleo está configurado para utilizar la porción de anticipación con el fin de determinar otras características tales como parámetros de LPC, etc.

40

[0056] Además, el codificador de núcleo comprende adicionalmente etapas del pre-procesamiento a diferentes velocidades de muestreo tales como una primera etapa de pre-procesamiento 1430c que opera a 12,8 kHz y otra etapa de pre-procesamiento 1430d que opera con velocidades de muestreo del grupo de las velocidades de muestreo consistente en 16 kHz, 25,6 kHz o 32 kHz.

45

[0057] Por ello, en términos generales, la forma de realización ilustrada en la figura 5 está configurada para tener un remuestreador en el dominio espectral para muestreo repetido, a partir de la velocidad de entrada, que puede ser 8 kHz, 16 kHz o 32 kHz en cualquiera de las velocidades de salida que sean diferentes de 8, 16 o 32.

[0058] Además, la forma de realización en la figura 5 está configurada adicionalmente para tener una rama adicional que no se ha remuestreado, es decir, la rama ilustrada como "IDFT con velocidad de entrada" para la señal media (Mid) y, opcionalmente, para la señal lateral (Side).

[0059] Además, el codificador en la figura 5 comprende preferentemente un remuestreador que no solamente remuestrea con una primera velocidad de muestreo de salida, sino también con una segunda velocidad de muestreo de salida con el fin de tener datos para ambos, los pre-procesadores 1430c y 1430d que pueden, por ejemplo, ser operativos para llevar a cabo algún tipo de filtrado, algún tipo de cálculo de LPC o algún tipo de otro procesamiento de señales que se describe preferentemente en el estándar 3GPP para el Codificador EVS ya mencionado en el contexto de la figura 4a.

60

[0060] La figura 6 ilustra una forma de realización para un aparato para decodificación de una señal multicanal codificada 1601. El aparato para decodificación comprende un decodificador de núcleo 1600, un convertidor tiempo-espectral 1610, un remuestreador en el dominio espectral 1620, un procesador multicanal 1630 y un convertidor espectral-tiempo 1640.

65

- [0061]** De nuevo, la invención con respecto al aparato para decodificación de la señal multicanal codificada 1601 puede implementarse en dos alternativas. Una alternativa es que el remuestreador en el dominio espectral está configurado para remuestrear la señal decodificada por núcleo en el dominio espectral antes de llevar a cabo el procesamiento multicanal. Esta alternativa se ilustra mediante líneas continuas en la figura 6. Sin embargo, la otra alternativa es que el muestreo repetido en el dominio espectral se lleve a cabo subsiguientemente al procesamiento multicanal, es decir, que el procesamiento multicanal tenga lugar con la velocidad de muestreo de entrada. Esta forma de realización se ilustra en la figura 6 mediante líneas discontinuas.
- [0062]** Particularmente, en la primera forma de realización, es decir, en la que el muestreo repetido en el dominio espectral se lleva a cabo en el dominio espectral antes del procesamiento multicanal, la señal decodificada de núcleo que representa una secuencia de bloques de valores de muestreo se convierte en una representación en el dominio de las frecuencias que tiene una secuencia de bloques de valores espectrales para la señal decodificada por núcleo, en la línea 1611.
- [0063]** Adicionalmente, la señal decodificada por núcleo no solo comprende la señal M en la línea 1602, sino también una señal lateral (Side) en la línea 1603, en la que se ilustra una señal lateral (Side) en 1604 en una representación codificada por núcleo.
- [0064]** Entonces, el convertidor tiempo-espectral 1610 genera adicionalmente una secuencia de bloques de valores espectrales para la señal lateral (Side) en la línea 1612.
- [0065]** Entonces, se lleva a cabo un muestreo repetido en el dominio espectral por el bloque 1620, y la secuencia remuestreada de bloques de valores espectrales con respecto a la señal media (Mid) o canal de mezclado de forma descendente o primer canal se comunica al procesador multicanal en la línea 1621 y, opcionalmente, también se comunica una secuencia remuestreada de bloques de valores espectrales para la señal lateral (Side) desde el remuestreador en el dominio espectral 1620 al procesador multicanal 1630 por medio de la línea 1622.
- [0066]** Entonces, el procesador multicanal 1630 lleva a cabo un procesamiento multicanal inverso con respecto a una secuencia que comprende una secuencia de la señal de mezclado de forma descendente y, opcionalmente, de la señal lateral (Side) ilustrada en las líneas 1621 y 1622 con el fin de emitir al menos dos secuencias resultado de bloques de valores espectrales ilustrados en 1631 y 1632. Estas al menos dos secuencias son convertidas seguidamente en el dominio del tiempo mediante el uso del convertidor espectral-tiempo con el fin de emitir las señales del canal en el dominio del tiempo 1641 y 1642. En la otra alternativa, ilustrada en la línea 1615, el convertidor tiempo-espectral está configurado para introducir la señal decodificada por núcleo tal como la señal media (Mid) en el procesador multicanal. Adicionalmente, el convertidor tiempo-espectral también puede conducir una señal lateral (Side) decodificada 1603 en su representación en el dominio espectral hacia el procesador multicanal 1630, aunque esta opción no se ha representado en la figura 6. Entonces, el procesador multicanal lleva a cabo el procesamiento inverso y la salida de al menos dos canales se comunica por medio de la línea de conexión 1635 al remuestreador en el dominio espectral que comunica seguidamente lo remuestreado en estos dos canales por medio de la línea 1625 hacia el convertidor espectral-tiempo 1640.
- [0067]** Por lo tanto, de una manera un tanto análoga a lo que se ha expuesto en el contexto de la figura 1, el aparato para decodificación de una señal multicanal codificada comprende también dos alternativas, es decir, una en la que el muestreo repetido en el dominio espectral se lleva a cabo antes del procesamiento multicanal inverso o, como alternativa, una en la que el muestreo repetido en el dominio espectral se lleva a cabo subsiguientemente al procesamiento multicanal a la velocidad de muestreo de entrada. Sin embargo, es preferible llevar a cabo la primera alternativa dado que permite una alineación ventajosa de las diferentes contribuciones de señal ilustradas en la figura 7a y en la figura 7b.
- [0068]** De nuevo, en la figura 7a se ilustra el decodificador de núcleo 1600 que, sin embargo, emite tres señales de salida diferentes, es decir, una primera señal de salida 1601 con una velocidad de muestreo diferente de la velocidad de muestreo de salida, una segunda señal decodificada de núcleo 1602 a la velocidad de muestreo de entrada, es decir, la velocidad de muestreo subyacente a la señal codificada por núcleo 1601 y el decodificador de núcleo adicionalmente genera una tercera señal de salida 1603 operable y disponible a la velocidad de muestreo de salida, es decir, la velocidad de muestreo finalmente prevista en la salida del convertidor espectral-tiempo 1640 en la figura 7a.
- [0069]** La totalidad de las tres señales decodificadas por núcleo son introducidas en el convertidor tiempo-espectral 1610 que genera tres secuencias diferentes de bloques de valores espectrales 1613, 1611 y 1612.
- [0070]** La secuencia de bloques de valores espectrales 1613 tiene valores de frecuencia o espectrales de hasta la máxima frecuencia de salida y, por ello, está asociada con la velocidad de muestreo de salida.
- [0071]** La secuencia de bloques de valores espectrales 1611 tiene valores espectrales de hasta una máxima

frecuencia diferente y, por ello, esta señal no corresponde a la velocidad de muestreo de salida.

[0072] Además, los valores espectrales de la señal 1612 de hasta la máxima frecuencia introducida que también es diferente de la frecuencia de salida máxima.

5

[0073] Por lo tanto, se comunican las secuencias 1612 y 1611 al remuestreador en el dominio espectral 1620 mientras que la señal 1613 no se comunica al remuestreador en el dominio espectral 1620, dado que esta señal ya está asociada con la velocidad de muestreo de salida correcta.

10 **[0074]** El remuestreador en el dominio espectral 1620 comunica las secuencias remuestreadas de valores espectrales a un combinador 1700 que está configurado para llevar a cabo una combinación bloque por bloque con líneas espectrales para señales que corresponden en situaciones superpuestas. Por lo tanto, típicamente habrá una región de cruce entre una conmutación desde una señal basada en MDCT a una señal ACELP, y en este intervalo de superposición, existen valores de señales que se combinan entre sí. Sin embargo, una vez superado este intervalo de superposición, y una señal existe solamente en la señal 1603 por ejemplo mientras que la señal 1602, por ejemplo, no existe, entonces el combinador no llevará a cabo una adición de línea espectral bloque por bloque en esta porción. Sin embargo, cuando se presenta un cruce posteriormente, en tal caso tendrá lugar una línea espectral bloque por bloque mediante adición de línea espectral durante esta región de cruce.

20 **[0075]** Además, también puede ser posible una adición continua como se ilustra en la figura 7b, en la que se lleva a cabo una señal de salida de post-filtro bajo en el bloque 1600a, que genera una señal de error entre armónicas que podría ser, por ejemplo, la señal 1601 de la figura 7a. Entonces, subsiguientemente a la conversión tiempo-espectral en el bloque 1610, y del muestreo repetido subsiguiente en el dominio espectral 1620, se lleva preferentemente a cabo una operación de filtrado adicional 1702 antes de llevar a cabo la adición en el bloque 1700 en la figura 7b.

[0076] De manera similar, la etapa de decodificación basada en MDCT 1600d y la etapa de decodificación de extensión del ancho de banda en el dominio del tiempo 1600c pueden acoplarse por medio de un bloque de desvanecimiento cruzado 1704 con el fin de obtener la señal decodificada de núcleo 1603 que es convertida seguidamente en la representación en el dominio espectral a la velocidad de muestreo de salida de tal manera que para esta señal 1613, el muestreo repetido en el dominio espectral no es necesario, sino que la señal puede ser comunicada directamente al combinador 1700. El procesamiento estéreo inverso o el procesamiento multicanal 1603 seguidamente tiene lugar subsiguientemente hacia el combinador 1700.

35 **[0077]** Por lo tanto, a diferencia de la forma de realización ilustrada en la figura 6, el procesador multicanal 1630 no opera sobre la secuencia remuestreada de valores espectrales, sino que opera sobre una secuencia que comprende la al menos una secuencia remuestreada de valores espectrales tales como 1622 y 1621 donde la secuencia, sobre la cual opera el procesador multicanal 1630, comprende adicionalmente la secuencia 1613 cuyo muestreo no era necesario.

40

[0078] Como se ilustra en la figura 7, las diferentes señales decodificadas procedentes de diferentes DFTs que trabajan con diferentes velocidades de muestreo ya están alineadas en el tiempo dado que las ventanas de análisis con diferentes velocidades de muestreo comparten la misma forma. Sin embargo, los espectros muestran diferentes magnitudes y puestas en escala. Para armonizarlos y hacerlos compatibles, todos los espectros son remuestreados en el dominio de las frecuencias con la deseada velocidad de muestreo de salida antes de ser adicionados entre sí.

[0079] Por lo tanto, en la figura 7 se ilustra la combinación de diferentes contribuciones de una señal sintetizada en el dominio DFT, en donde el muestreo repetido en el dominio espectral se lleva a cabo de tal manera que, al final, todas las señales que deben ser añadidas por el combinador 1700 ya están disponibles con valores espectrales que se extienden hasta la frecuencia de salida máxima que corresponde a la velocidad de muestreo de salida, es decir, es inferior o igual a la mitad de la velocidad de muestreo de salida que seguidamente se obtiene a la salida del convertidor espectral tiempo 1640.

55 **[0080]** La elección del banco de filtros estéreo es esencial para un sistema de bajo retardo, y el compromiso obtenible se ha resumido en la figura 8b. Puede utilizarse ya sea una DFT (transformada de bloque) o un QMF de bajo pseudo-retardo llamado CLDFB (banco de filtros). Cada propuesta muestra un retardo, tiempo y resoluciones de frecuencia diferentes. Para el sistema debe elegirse el mejor compromiso entre estas características. Es importante tener una buena frecuencia y buenas resoluciones de tiempo. Esta es la razón por la que mediante la utilización de un banco de filtros de pseudo QMF como en la propuesta 3 puede ser problemática. La resolución de frecuencia es baja. Se puede reforzar mediante estrategias híbridas como en MPS 212 de MPEG-USAC, pero tiene el inconveniente de incrementar de manera significativa tanto la complejidad como el retardo. Otro punto importante es el retardo disponible en el lado del decodificador entre el decodificador de núcleo y el procesamiento estéreo inverso. Cuanto mayor sea este retardo, mejor será. Por ejemplo, la propuesta 2 no puede proporcionar un retardo de éste, por lo que no es una solución útil. Por estas razones mencionadas más arriba, nos enfocaremos en el resto

65

de la descripción en las propuestas 1, 4 y 5.

[0081] La ventana de análisis y síntesis del banco de filtros es otro aspecto importante. En la forma de realización preferida se utiliza la misma ventana para el análisis y síntesis de la DFT. También es el mismo en los 5 lados de codificador y decodificador. Se prestó especial atención a satisfacer las siguientes restricciones:

- la región de superposición tiene que ser igual o más pequeña que la región de superposición del núcleo de MDCT y de la vista adelantada de ACELP. En la forma de realización preferida todas las magnitudes son iguales a 8,75 ms
- El padding cero debería ser al menos de aproximadamente 2,5 ms con el fin de permitir la aplicación de un desplazamiento lineal de los canales en el dominio de DFT.
- El tamaño de la ventana, el tamaño de la región de superposición y la magnitud del padding cero deben expresarse en números enteros de muestras para diferentes velocidades de muestreo 12,8, 16, 25,6, 32 y 48 kHz
- La complejidad de la DFT debe ser lo más baja posible, es decir la base máxima de la DFT en una implementación de FFT de tipo base dividida debería ser lo más baja posible.
- Se fija la resolución de tiempo en 10 ms.

[0082] En base a estas restricciones, las ventanas para las propuestas 1 y 4 se describen en la figura 8c y en la figura 8a.

[0083] La figura 8c ilustra una primera ventana que consiste en una porción superpuesta inicial 1801, una porción central subsiguiente 1803 y una porción superpuesta terminal o una segunda porción superpuesta 1802. Además, la primera porción superpuesta 1801 y la segunda porción superpuesta 1802 tienen adicionalmente una porción de padding cero de 1804 en el inicio y 1805 al final de ella.

[0084] Por otra parte, en la figura 8c se ilustra el procedimiento llevado a cabo con respecto a la estructura del convertidor tiempo-espectral 1000 de la figura 1 o como alternativa, 1610 de la figura 7a. La ventana de análisis adicional consistente en los elementos 1811, es decir, una primera porción superpuesta, una porción central no superpuesta 1813 y una segunda porción superpuesta 1812 se superpone a la primera ventana en un 50%. Adicionalmente, la segunda ventana tiene las porciones de padding cero 1814 y 1815 en el inicio y final de ella. Estas porciones de padding cero son necesarias con el fin de estar en la posición de llevar a cabo la alineación de tiempo del ancho de banda en el dominio de las frecuencias.

[0085] Además, la primera porción superpuesta 1811 de la segunda ventana empieza al final de la parte central 1803, es decir, la parte no superpuesta de la primera ventana, y parte superpuesta de la segunda ventana, es decir, la parte no superpuesta 1813 empieza al final de la segunda porción superpuesta 1802 de la primera ventana, como se ilustra.

[0086] Si se considera que la figura 8c representa una operación de superposición-adición en un convertidor espectral-tiempo tal como el convertidor espectral-tiempo 1030 de la figura 1 para el codificador o para el convertidor espectral-tiempo 1640 para el decodificador, entonces la primera ventana que consiste en los bloques 1801, 1802, 1803, 1805, 1804 corresponde a una ventana de síntesis y la segunda ventana que consiste en las partes 1811, 1812, 1813, 1814, 1815 corresponde a la ventana de síntesis para el bloque siguiente. En tal caso, la superposición entre la ventana ilustra la porción superpuesta, y la porción superpuesta se ilustra en 1820, y la longitud de la porción superpuesta es igual a la trama actual dividida por dos y es, en la forma de realización preferida, igual a 10 ms. Además, en la parte inferior de la figura 8c, la ecuación analítica para calcular los coeficientes de ventana ascendentes dentro del intervalo de superposición 1801 o 1811 se ilustra como una función seno, y, de manera correspondiente, los coeficientes de magnitud de superposición descendente de la porción superpuesta 1802 y 1812 también se ilustran como una función seno.

[0087] En formas de realización preferidas, se utiliza la misma ventana de análisis y síntesis solamente para el decodificador ilustrado en la figura 6, figura 7a, y figura 7b. Por lo tanto, el convertidor tiempo-espectral 1616 y el convertidor espectral-tiempo 1640 utilizan exactamente las mismas ventanas que las ilustradas en la figura 8c.

[0088] Sin embargo, en determinadas formas de realización particularmente con respecto a la subsiguiente propuesta/forma de realización 1, se utiliza una ventana de análisis que en términos generales concuerda con la figura 1c, pero los coeficientes de ventana para las posiciones de superposición ascendente o descendente se calculan utilizando una raíz cuadrada de función seno, con el mismo argumento en la función seno que en la figura 8c. De manera correspondiente, la ventana de síntesis se calcula mediante el uso de una función seno a la potencia de 1,5, pero de nuevo con el mismo argumento que la función seno.

[0089] Además, cabe observar que debido a la operación de superposición-adición, la multiplicación de seno a la potencia 0,5 multiplicado por seno a la potencia 1,5 resulta de nuevo en un seno a la potencia 2, resultado que es necesario con el fin de tener una situación de conservación de energía.

[0090] La propuesta 1 tiene como características principales que la región de superposición de la DFT tiene

el mismo tamaño y está alineada con la anticipación de ACELP y con la región de superposición del núcleo de MDCT. El retardo del codificador es en tal caso el mismo para los ACELP/Núcleos de MDCT y el estéreo no introduce ningún retardo adicional en el codificador. En el caso de EVS y en caso de utilizarse la estrategia del banco de filtros de síntesis multivelocidad descrito en la figura 5, el retardo del codificador estéreo es bajo, de 5 aproximadamente 8,75 ms.

[0091] La estructura esquemática del codificador se ilustra en la figura 9a, mientras que el decodificador se ilustra en la figura 9e. En la figura 9c, las ventanas han sido dibujadas en azul de puntos para el codificador y en rojo sólido para el decodificador.

10

[0092] Un tema importante en el caso de la propuesta 1 es que la anticipación en el codificador está puesta en ventana. Se puede compensar para el procesamiento subsiguiente, o se puede dejar puesta en ventana si el procesamiento subsiguiente está adaptado para tener en cuenta una anticipación puesta en ventana. Podría darse el caso de que si el procesamiento estéreo llevado a cabo en la DFT modificó el canal de entrada, y especialmente 15 cuando se utilizan operaciones no lineales, que la señal compensada o puesta en ventana no permita lograr una reconstrucción perfecta en el caso en que se omita la codificación de núcleo.

[0093] Cabe observar que entre la ventana de síntesis del decodificador de núcleo y la ventana de análisis del decodificador estéreo hay un espacio de tiempo de 1,25ms que puede ser aprovechado por el decodificador de 20 núcleo después del procesamiento, por la extensión del ancho de banda (BWE, bandwidth extension), como BWE del dominio del tiempo utilizado sobre ACELP, o mediante alguna suavización en el caso de transición entre ACELP y núcleos de MDCT.

[0094] Dado que este espacio de tiempo de solamente 1,25 ms es inferior a los 2,3125 ms requeridos por el 25 estándar EVS para tales operaciones, la presente invención proporciona una manera de combinar, remuestrear y suavizar las diferentes partes de síntesis del decodificador conmutado dentro del dominio de DFT del módulo estéreo.

[0095] Como se ilustra en la figura 9a, el codificador de núcleo 1040 está configurado para operar según un 30 control de estructura para proporcionar una secuencia de tramas, en donde una trama está delimitada por un borde de inicio de la trama 1901 y un borde final de la trama 1902. Además, el convertidor tiempo-espectral 1000 y/o el convertidor espectral-tiempo 1030 están configurados también para operar según un segundo control de estructura que está sincronizado con el primer control de estructura. El control de estructura se ilustra con dos ventanas superpuestas 1903 y 1904 para el convertidor tiempo-espectral 1000 en el codificador, y, particularmente, para el 35 primer canal 1001 y el segundo canal 1002 que son procesados simultáneamente y completamente sincronizados. Además, el control de estructura también es visible en el lado del decodificador, específicamente, con dos ventanas propuestas para el convertidor tiempo-espectral 1610 de la figura 6 que se ilustran en 1913 y 1914. Estas ventanas. 1913 y 1914 se aplican a la señal del decodificador de núcleo que es preferentemente una única señal mono o de 40 mezclado de forma descendente 1610 de la figura 6, por ejemplo. Además, como puede observarse en la figura 9a, la sincronización entre el control de estructura del codificador de núcleo 1040 y el convertidor tiempo-espectral 1000 o el convertidor espectral-tiempo 1030 es tal que el borde de inicio de la trama 1901 el borde final de la trama 1902 de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con respecto a una instancia de inicio y la instancia final de una porción superpuesta de una ventana utilizada por el convertidor tiempo-espectral 1000 o por el convertidor espectral-tiempo 1030 para cada bloque de la secuencia de bloques de valores de 45 muestreo o para cada bloque de la secuencia remuestreada de bloques de valores espectrales. En la forma de realización ilustrada en la figura 9a, la relación predeterminada es tal que el inicio de la primera porción superpuesta coincide con el borde de tiempo de inicio con respecto a la ventana 1903, y el inicio de la porción superpuesta de la ventana adicional 1904 coincide con el extremo de la parte central tal como la parte 1803 de la figura 8c, por ejemplo. Por lo tanto, el borde de tramo final 1902 coincide con el extremo de la parte central 1813 de la figura 8c, 50 cuando la segunda ventana en la figura 8c corresponde a la ventana 1904 en la figura 9a.

[0096] Por lo tanto, es evidente que la segunda porción superpuesta tal como 1812 de la figura 8c de la segunda ventana 1904 en la figura 9a se extiende sobre el borde de trama de inicio o de parada 1902, y, por ello, se extiende en la porción codificador-núcleo de anticipación ilustrada en 1905.

55

[0097] Por lo tanto, el codificador de núcleo 1040 está configurado para utilizar una porción de anticipación tal como la porción de anticipación 1905 cuando se codifique por núcleo el bloque de salida de la secuencia de salida de los bloques de muestreo, en donde la porción de salida de anticipación está situada en el tiempo subsiguiente al bloque de salida. El bloque de salida se corresponde con la trama delimitada por los límites de trama 1901, 1904 y la 60 porción de salida de anticipación 1905 se presenta después de este bloque de salida para el codificador de núcleo 1040.

[0098] Además, como se ilustra, el convertidor tiempo-espectral está configurado para utilizar una ventana de análisis, es decir, la ventana 1904 que tiene la porción de superposición con una longitud en tiempo inferior o igual a 65 la longitud en tiempo de la porción de anticipación 1905, en donde esta porción superpuesta correspondiente a la

superposición 1812 de la figura 8c que está situada en el intervalo de superposición, se utiliza para generar la porción puesta en ventana de anticipación.

5 **[0099]** Además, el convertidor espectral-tiempo 1030 está configurado para procesar la porción de salida de anticipación correspondiente a la porción puesta en ventana de anticipación preferentemente mediante el uso de una función de compensación, en donde la función de compensación está configurada de manera que reduzca o elimine una influencia de la superposición de la ventana de análisis.

10 **[0100]** Por lo tanto, el convertidor espectral-tiempo que opera entre el codificador de núcleo 1040 y el bloque de mezclado de forma descendente 1010/bloque de muestreo descendente 1020 en la figura 9a está configurado para aplicar una función de compensación con el fin de deshacer la puesta en ventana aplicada por la ventana 1904 en la figura 9a.

15 **[0101]** Por lo tanto, se asegura que el codificador de núcleo 1040, cuando aplica su funcionalidad de anticipación a la porción de anticipación 1095, lleve a cabo la función de anticipación no porción sino a una porción que esté lo más cercana posible a la porción original.

20 **[0102]** Sin embargo, debido a las restricciones de bajo retardo, y debido a la sincronización entre la estructura del pre-procesador estéreo y el codificador de núcleo, no existe una señal original en el dominio del tiempo para la porción de anticipación. Sin embargo, la aplicación de la función de compensación asegura que cualquier artefacto ocasionado por este procedimiento se reduzca lo máximo posible.

25 **[0103]** En la figura 9d y 9e se ilustra con mayor detalle una secuencia de procedimientos con respecto a esta tecnología.

[0104] En la etapa 1910, se lleva a cabo una a DFT⁻¹ de un bloque 0^{ésimo} de manera que se obtenga un bloque 0^{ésimo} en el dominio del tiempo. El bloque 0^{ésimo} podría haber obtenido una ventana utilizada a la izquierda de la ventana 1903 en la figura 9a. Sin embargo, este bloque 0-^{ésimo}, no se ilustra explícitamente en la figura 9a.

30 **[0105]** Seguidamente, en la etapa 1912, el bloque zero^{ésimo} se pone en ventana mediante el uso de una ventana de síntesis, es decir, se pone en ventana en el convertidor espectral-tiempo 1030 ilustrado en la figura 1.

35 **[0106]** Entonces, como se ilustra en el bloque 1911, se lleva a cabo una DFT⁻¹ del primer bloque obtenido por la ventana 1903 para obtener un primer bloque en el dominio del tiempo, y este primer bloque se pone de nuevo en ventana mediante el uso de la ventana de síntesis en el bloque 1910.

40 **[0107]** Entonces, como se indica en 1918 en la figura 9d, se lleva a cabo una DFR inversa del segundo bloque, es decir, el bloque obtenido por la ventana 1904 de la Figura 9a, para obtener un segundo bloque en el dominio del tiempo, y, seguidamente la primera porción del segundo bloque es puesta en ventana mediante el uso de la ventana de síntesis como se ilustra por 1920 de la Figura 9d. Sin embargo, lo importante es que la segunda porción del segundo bloque obtenido por el elemento 1918 en la Figura 9d no es puesta en ventana mediante el uso de la ventana de síntesis, sino que es compensada como se ilustra en el bloque 1922 de la Figura 9d, y, para la función de compensación, se utiliza la inversa de la función de ventana de análisis y la correspondiente porción superpuesta de función de la ventana de análisis.

45 **[0108]** Por lo tanto, si la ventana utilizada para generar el segundo bloque fue una ventana seno ilustrada en la Figura 8c, en tal caso se utiliza 1/sin() para los coeficientes de tamaño de superposición descendentes de las ecuaciones a la parte inferior de la Figura 8c, como función de compensación.

50 **[0109]** Sin embargo, se prefiere utilizar una raíz cuadrada de ventana seno para la ventana de análisis y,

por ello, la función de compensación es una función de ventana de $\frac{1}{\sqrt{\sin(\theta)}}$. Esto asegura que la porción compensada de anticipación obtenida por el bloque 1922 es lo más cercana posible a la señal original, pero por supuesto no es la señal izquierda original ni la señal derecha original sino la señal que se habría obtenido mediante la adición de izquierda y derecha para obtener la señal media (Mid).

55 **[0110]** Entonces, en la etapa 1924 en la figura 9d, una trama indicada mediante los límites de trama 1901,1902 se genera llevando a cabo una operación de superposición-adición en el bloque 1030 de tal manera que el codificador tenga una señal en el dominio del tiempo, y esta trama se lleva a cabo mediante una operación de superposición-adición entre el bloque correspondiente a la ventana 1903, y las muestras anteriores del bloque anterior y mediante el uso de la primera porción del segundo bloque obtenido por el bloque 1920. Entonces, Esta trama emitida por el bloque 1924 se comunica al codificador de núcleo 1040 y, adicionalmente, el codificador de núcleo recibe adicionalmente la porción compensada de anticipación para la trama y, como se ilustra en la etapa 1926, el codificador de núcleo puede determinar seguidamente la característica para el codificador de núcleo

mediante el uso de la porción compensada de anticipación obtenida por la etapa 1922. Entonces, como se ilustra en la etapa 1928, el codificador de núcleo codifica por núcleo la trama mediante el uso de la característica determinada en el bloque 1926 de manera que finalmente se obtenga la trama, codificada por núcleo, correspondiente al borde de trama 1901, 1902 que tiene, en la forma de realización preferida, una longitud de 20 ms.

5

[0111] Es preferible que la porción superpuesta de la ventana 1904 que se extiende en la porción de anticipación 1905 tenga la misma longitud que en la porción de anticipación, pero también puede ser más corta que la porción de anticipación pero se prefiere que no sea más larga que la porción de anticipación de tal manera que el pre-procesador estéreo no introduzca ningún retardo adicional debido a las ventanas superpuestas.

10

[0112] Seguidamente, el procedimiento continúa con la puesta en ventana de la segunda porción del segundo bloque mediante el uso de la ventana de síntesis ilustrada en el bloque 1930. Por lo tanto, la segunda porción del segundo bloque es, por una parte, compensada por el bloque 1922 y por otra parte es puesta en ventana por ventana de síntesis ilustrada en el bloque 1930, dado que, a continuación, se requiere esta porción para regenerar la siguiente trama para el codificador de núcleo por superposición/adición de la segunda porción puesta en ventana del segundo bloque, un tercer bloque puesto en ventana y una primera porción de ventana del cuarto bloque ilustrado en el bloque 1932. Por supuesto, el cuarto bloque y en particular la segunda porción del cuarto bloque sería una vez más sometido a la operación de compensación como se expuso con respecto al segundo bloque en el elemento 1922 de la figura 9d y, entonces, se repetiría una vez más el procedimiento como se ha expuesto más arriba. Además, en la etapa 1934, el codificador de núcleo determinaría las características del codificador de núcleo utilizado para compensar la segunda porción del cuarto bloque y, entonces, la siguiente trama se codificaría mediante el uso de las características de codificación determinadas con el fin de obtener finalmente la siguiente trama codificada en el bloque 1934. Por lo tanto, la alineación de la segunda porción superpuesta de la ventana de análisis (en la síntesis correspondiente) con la porción de anticipación del codificador de núcleo 1905 asegura que es posible obtener una implementación de retardo muy baja y que esta ventaja se debe al hecho de que la porción de anticipación puesta en ventana es compensada, por, por una parte, la realización de la función de compensación y por otra parte por la aplicación de una ventana de análisis que no es igual a la ventana de síntesis sino que aplica una influencia más pequeña, de tal manera que puede asegurarse que la función de compensación es más estable en comparación con la utilización de la misma ventana de análisis/síntesis. Sin embargo, en el caso en que el codificador del núcleo haya sido modificado para operar esta función de anticipación que típicamente es necesaria para determinar las características de codificación del núcleo en una porción puesta en ventana, no es necesaria para llevar a cabo la función de compensación. Sin embargo, se ha descubierto que la utilización de la función de compensación es ventajosa en comparación con la modificación del codificador de núcleo.

35

[0113] Además, como se ha expuesto más arriba, cabe observar que hay un espacio de tiempo entre el fin de una ventana, es decir, la ventana de análisis 1914 y el borde de trama 1902 de la trama definida por el borde de inicio de la trama 1901 y el límite final de la trama 1902 de la figura 9b.

40

[0114] En particular, el espacio de tiempo se ilustra en 1920 con respecto a las ventanas de análisis aplicadas por el convertidor de tiempo-espectro 1610 de la figura 6, y este espacio de tiempo también es visible 120 con respecto al primer canal de salida 1641 y el segundo canal de salida 1642.

45

[0115] La figura 9f muestra un procedimiento de las etapas llevadas a cabo en el contexto del espacio de tiempo, el decodificador de núcleo 1600 decodifica en núcleo la trama o al menos la porción inicial de la trama hasta el espacio de tiempo 1920. Entonces, el convertidor de tiempo-espectro 1610 de la figura 6 está configurado para aplicar una ventana de análisis a la porción inicial de la trama utilizando la ventana de análisis 1914 que no se extiende hasta el final de la trama, es decir, hasta el instante de tiempo 1902, sino que solamente se extiende hasta el inicio del espacio de tiempo 1920.

50

[0116] Por lo tanto, el decodificador de núcleo tiene tiempo adicional para decodificar en núcleo las muestras en el espacio de tiempo y/o para post-procesar las muestras en el espacio de tiempo como se ilustra en el bloque 1940. Por lo tanto, el convertidor de tiempo-espectro 1610 ya emite un primer bloque como el resultado de la etapa 1938 allí el decodificador de núcleo puede proporcionar las muestras permanentes en el espacio de tiempo y puede post-procesar las muestras en el espacio de tiempo en la etapa 1940.

55

[0117] Seguidamente, en la etapa 1942, el convertidor de tiempo-espectro 1610 está configurado para poner en ventana las muestras en el espacio de tiempo junto con muestras de la trama siguiente mediante el uso de una siguiente ventana de análisis que se presentaría subsiguientemente a la ventana 1914 en la figura 9b. A continuación, como se ilustra en la etapa 1944, el decodificador de núcleo 1600 está configurado para decodificación de la siguiente trama o al menos la porción inicial de la trama siguiente hasta que tenga lugar el espacio de tiempo 1920 en la siguiente trama. Entonces, en la etapa 1946, el convertidor de tiempo-espectro 1610 está configurado para poner en ventana las muestras en la siguiente trama hasta el espacio de tiempo 1920 de la trama siguiente y, en la etapa 1948, el decodificador de núcleo podría entonces decodificar en núcleo las muestras restantes en el espacio de tiempo de la trama siguiente y/o post-procesar estas muestras.

65

[0118] Por lo tanto, este espacio de tiempo de, por ejemplo, 1,25 ms cuando se considera que la forma de realización de la figura 9b puede ser aprovechada para el post-procesamiento del decodificador de núcleo, por la extensión del ancho de banda, mediante, por ejemplo, una extensión del ancho de banda en el dominio del tiempo utilizada en el contexto de ACELP, o mediante alguna suavización en el caso de una transición de la transmisión entre señales ACELP y MDCT de núcleo.

[0119] Por lo tanto, una vez más, el decodificador de núcleo 1600 está configurado para operar según un primer control de la estructura para proporcionar una secuencia de tramas, en la que el convertidor de tiempo-espectro 1610 o el convertidor espectro-tiempo 1640 están configurados para operar según un segundo control de estructura que está sincronizado con el primer control de estructura, de tal manera que el borde de inicio de la trama o el borde final de la trama de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con respecto a un instante de inicio o con respecto a un instante final de una porción superpuesta de una ventana utilizada por el convertidor de tiempo-espectro o por el convertidor espectro-tiempo para cada bloque de la secuencia de bloques de valores de muestreo o para cada bloque de la secuencia remuestreada de bloques de valores espectrales.

[0120] Además, el convertidor de tiempo-espectro 1610 está configurado para utilizar una ventana de análisis para poner en ventana la trama de la secuencia de tramas que tiene un intervalo de superposición que termina antes del borde de trama final 1902 quedando un espacio de tiempo 1920 entre el fin de la porción de superposición y el borde final de la trama. El decodificador de núcleo 1600 está configurado, por lo tanto, para llevar a cabo el procesamiento de las muestras en el espacio de tiempo 1920 en paralelo con la puesta en ventana de la trama mediante el uso de la ventana de análisis o en donde se lleva a cabo otro post-procesamiento en paralelo con la puesta en ventana de la trama mediante el uso de la ventana de análisis por el convertidor tiempo-espectral.

[0121] Además, es preferible que la ventana de análisis para un siguiente bloque de la señal decodificada de núcleo esté situada de tal manera que una porción central no superpuesta de la ventana se encuentre situada dentro del espacio de tiempo como se ilustra en 1920 de la figura 9b.

[0122] En la propuesta 4 el retardo global del sistema está incrementado con respecto al de la propuesta 1. En el codificador llega un retardo extra debido al módulo estéreo. El tema de una reconstrucción perfecta ya no se aplica en la propuesta 4, a diferencia de la propuesta 1.

[0123] En el decodificador, el retardo disponible entre el decodificador de núcleo y el análisis del primer DFT es de 2,5 ms, lo que permite llevar a cabo un muestreo repetido convencional, combinación y suavización entre las diferentes síntesis de núcleo y las señales de ancho de banda extendidas como se hace para el estándar EVS.

[0124] La estructura esquemática del codificador se ilustra en la figura 10a mientras que el decodificador se ilustra en la figura 10b. Las ventanas se indican en la figura 10c.

[0125] En la propuesta 5, la resolución de tiempo de la DFT está disminuida a 5 ms. La región de anticipación y de superposición del codificador núcleo no está puesta en ventana, lo que es una ventaja compartida con respecto a la propuesta 4. Por otra parte, el retardo disponible entre la decodificación de codificador y el análisis estéreo es pequeño y se necesita una solución como se propone en la Propuesta 1 (Figura 7). Las principales desventajas de esta propuesta es la baja resolución de frecuencia de la descomposición de tiempo-frecuencia y la pequeña región de superposición reducida a 5 ms, lo que impide un gran desplazamiento del tiempo en el dominio de las frecuencias.

[0126] La estructura esquemática del codificador se ilustra en la figura 11a mientras que el decodificador se ilustra en la figura 11b. Las ventanas se indican en la figura 11c.

[0127] En vista de lo anterior, las formas de realización preferidas se refieren, en cuanto al lado del codificador, a una síntesis de tiempo-frecuencia de múltiples velocidades que proporciona al menos una señal estéreo procesada con diferentes velocidades de muestreo a los módulos de procesamiento subsiguientes. El módulo incluye, por ejemplo, un codificador del habla como ACELP, herramientas de pre-procesamiento, un codificador de audio basado en MDCT tal como TCX o un codificador de extensión de ancho de banda tal como un codificador de extensión de ancho de banda en el dominio del tiempo.

[0128] Con respecto al decodificador, se lleva a cabo la combinación en el muestreo repetido en el dominio de las frecuencias de estéreo con respecto a diferentes contribuciones de la síntesis de decodificador. Estas señales de síntesis pueden proceder de un decodificador del habla tal como un Decodificador ACELP, un Decodificador basado en MDCT, un módulo de extensión de ancho de banda o una señal de error entre las armónicas procedente de un post-procesamiento tal como un post-filtro bajo.

[0129] Además, con respecto tanto al codificador como al decodificador, es útil aplicar una ventana para la DFT o un valor complejo transformado con un padding cero, una región de baja superposición y un hopsize que

corresponde a un número entero de muestras con diferentes velocidades de muestreo tales como 12,9 kHz, 16 kHz, 25,6 kHz, 32 kHz o 48 kHz.

[0130] Algunas formas de realización son capaces de lograr una codificación con baja velocidad de bits de audio estéreo con bajo retardo. Se diseñó específicamente para combinar un esquema de codificación de audio conmutado con bajo retardo, tal como EVS, con los bancos de filtros de un módulo de codificación estéreo.

[0131] Algunas formas de realización pueden encontrar una utilización en la distribución o radiodifusión de todos los tipos de contenido audio estéreo o de múltiples canales (tanto habla como música con una calidad perceptual constante con una velocidad de bits baja dada) tales como, por ejemplo con aplicaciones de radio digital, streaming en Internet y comunicaciones de audio.

[0132] La figura 12 ilustra un aparato para la codificación de una señal multicanal que tiene al menos dos canales. La señal multicanal 10 es introducida en un determinador de parámetros 100 por una parte y en un alineador de señales 200 por otra parte. El determinador de parámetros 100 determina, por una parte, un parámetro de alineación de banda ancha y, por otra parte, una pluralidad de parámetros de alineación de banda ancha de la señal multicanal. Estos parámetros son emitidos por medio de una línea de parámetros 12. Además, estos parámetros son emitidos también por medio de otra línea de parámetros 14 alguna interfaz de salida 500 como se ilustra. Sobre la línea de parámetros 14, unos parámetros adicionales tales como los parámetros de nivel son comunicados desde el determinador de parámetros 100 a la interfaz de salida 500. El alineador de señales 200 está configurado para alinear los al menos dos canales de la señal multicanal 10 mediante el uso del parámetro de alineación de banda ancha y la pluralidad de parámetros de alineación de banda ancha recibida por medio de la línea de parámetros 10 para obtener canales alineados 20 en la salida del alineador de señal 200. Estos canales alineados 20 son comunicados a un procesador de señales 300 que está configurado para calcular una señal media (Mid) 31 y una señal lateral (Side) 32 a partir de los canales alineados recibidos por medio de la línea 20. El aparato para la codificación comprende además un codificador de señales 400 para la codificación la señal media (Mid) de la línea 31 y la señal lateral (Side) de la línea 32 para obtener una señal media (Mid) codificada en la línea 41 y una señal lateral (Side) codificada en la línea 42. Estas dos señales se comunican a la interfaz de salida 500 para generar una señal multicanal codificada en la línea de salida 50. La señal codificada en la línea de salida 50 comprende la señal media (Mid) codificada de la línea 41, la señal lateral (Side) codificada de la línea 42, los parámetros de alineación de banda ancha y los parámetros de alineación de banda ancha de la línea 14 y, opcionalmente, un parámetro de nivel de la línea 14 y, adicional y opcionalmente, un parámetro de llenado de estéreo generado por el codificador de señales 400 y comunicado a la interfaz de salida 500 por medio de la línea de parámetros 43.

[0133] Preferentemente, el alineador de señales está configurado para alinear los canales de la señal multicanal mediante el uso del parámetro de alineación de banda ancha, antes de que el determinador de parámetros 100 calcule realmente los parámetros de banda estrecha. Por ello, en esta forma de realización, el alineador de señales 200 envía los canales alineados de banda ancha de vuelta al determinador de parámetros 100 por medio de una línea de conexión 15. Seguidamente, el determinador de parámetros 100 determina la pluralidad de parámetros de alineación de banda ancha una señal multicanal ya alineada con respecto a las características de banda ancha. Sin embargo, en otras formas de realización, los parámetros se determinan sin esta secuencia específica de procedimientos.

[0134] En la figura 14a se ilustra una implementación preferida, en la que se lleva a cabo la secuencia específica de etapas en la que interviene la línea de conexión 15. En la etapa 16, se determina el parámetro de alineación de banda ancha mediante el uso de los dos canales y se obtiene el parámetro de alineación de banda ancha tal como una diferencia de tiempo entre canales o parámetro de ITD. A continuación, en la etapa 21, los dos canales son alineados por el alineador de señales 200 de la figura 12 mediante el uso del parámetro de alineación de banda ancha. Seguidamente, en la etapa 17, se determinan los parámetros de banda estrecha mediante el uso de los canales alineados dentro del determinador de parámetros 100 para determinar una pluralidad de parámetros de alineación de banda ancha tales como una pluralidad de parámetros de diferencia de fase entre canales para diferentes bandas de la señal multicanal. A continuación, en la etapa 22, se alinean los valores espectrales en cada banda de parámetros mediante el uso del correspondiente parámetro de alineación de banda estrecha para esta banda específica. Si se lleva a cabo este procedimiento en la etapa 22 para cada banda, para la que está disponible un parámetro de alineación de banda estrecha, en tal caso los canales primero y segundo o izquierdo/derecho, alineados están disponibles para un procesamiento de las señales adicional por el procesador de señales 300 de la figura 12.

[0135] La figura 14b ilustra otra implementación del codificador multicanal de la figura 12 en la que se llevan a cabo diversos procedimientos en el dominio de las frecuencias.

[0136] Específicamente, el codificador multicanal comprende además un convertidor de tiempo-espectro 150 para convertir una señal multicanal en el dominio del tiempo en una representación espectral de los al menos dos canales dentro del dominio de las frecuencias.

[0137] Además, como se ilustra en 152, el determinador de parámetros, el alineador de señales y el procesador de señales ilustrados en 100, 200 y 300 en la figura 12 operan, todos ellos, en el dominio de las frecuencias.

5

[0138] Además, el codificador multicanal y, específicamente, el procesador de señales comprende además un convertidor espectro-tiempo 154 para generar una representación en el dominio del tiempo, de al menos la señal media (Mid).

10 **[0139]** Preferentemente, el convertidor de espectro tiempo convierte adicionalmente una representación espectral de la señal lateral (Side) también determinada mediante los procedimientos representados por el bloque 152 en una representación en el dominio del tiempo, y a continuación se configura el codificador de señales 400 de la figura 12 para continuar con la codificación de la señal media (Mid) y/o de la señal lateral (Side) como señales en el dominio del tiempo en función de la implementación específica del codificador de señales 400 de la figura 12.

15

[0140] Preferentemente, el convertidor de tiempo-espectro 150 de la figura 14b está configurado para implementar las etapas 155, 156 y 157 de la figura 4c. Específicamente, la etapa 155 comprende el suministro de una ventana de análisis con al menos un paciente de padding cero en uno de sus extremos, y, específicamente, una porción de padding cero en la porción inicial de la ventana y una porción de padding cero en la porción terminal de la ventana como se ilustra, por ejemplo, en la figura 7a-7b más adelante. Además, la ventana de análisis tiene adicionalmente intervalos de superposición o porciones de superposición en una primera mitad de la ventana y en una segunda mitad de la ventana y, adicionalmente, preferentemente una parte central que es un intervalo de no superposición, según sea el caso.

25 **[0141]** En la etapa 156, cada canal es puesto en ventana mediante el uso de la ventana de análisis con intervalos de superposición. Específicamente, cada canal es puesto en ventana mediante el uso de la ventana de análisis de tal manera que se obtiene un primer bloque del canal. Subsiguientemente, se obtiene un segundo bloque que tiene un determinado intervalo de superposición con el primer bloque y así sucesivamente, de tal manera que subsiguientemente a, por ejemplo, cinco operaciones de puesta en ventana, se dispone de cinco bloques de muestras puestas en ventana que son seguidamente transformados en una representación espectral como se ilustra en 157 en la figura 14c. Se lleva a cabo el mismo procedimiento también para el otro canal de tal manera que al final de la etapa 157, se encuentra disponible una secuencia de bloques de valores espectrales y, específicamente, valores espectrales complejos tales valores espectrales de DFT o muestras de subbanda complejas.

35 **[0142]** En la etapa 158, que se lleva a cabo mediante el determinador de parámetros 100 de la figura 12, se determina un parámetro de alineación de banda ancha y en la etapa 159, que se lleva a cabo mediante la alineación de señales 200 de la figura 12, se lleva a cabo un desplazamiento circular mediante el uso del parámetro de alineación de banda ancha. En la etapa 160, de nuevo llevada a cabo por el determinador de parámetros 100 de la figura 12, se determinan los parámetros de alineación de banda ancha para bandas/subbandas individuales y en la etapa 161, se hacen rotar valores espectrales alineados para cada banda mediante el uso de parámetros de alineación correspondientes de banda ancha determinados por las bandas específicas.

45 **[0143]** La figura 14d ilustra otros procedimientos llevados a cabo por el procesador de señales 300. Específicamente, el procesador de señales 300 está configurado para calcular una señal media (Mid) y una señal lateral (Side) como se ilustra en la etapa 301. En la etapa 302, puede llevarse a cabo algún tipo de procesamiento adicional de la señal y a continuación, en la etapa 303, cada bloque de la señal media (Mid) y la señal lateral (Side) se transforma de vuelta en el dominio del tiempo, y, en la etapa 304, se aplica una ventana de síntesis a cada bloque obtenido por la etapa 303 y, en la etapa 305, se lleva a cabo una operación de adición de superposición para la señal media (Mid) por una parte y una operación de adición de superposición para la señal lateral (Side) por otra parte de manera que se obtengan finalmente las señales media/lateral (Mid/Side) en el dominio del tiempo.

55 **[0144]** Específicamente, las operaciones de las etapas 304 y 305 resultan en una especie de desvanecimiento cruzado desde un bloque de la señal media (Mid) o de la señal lateral (Side) en el siguiente bloque de la señal media (Mid) y la señal lateral (Side) se lleva a cabo de tal manera que incluso si se presentan algunos cambios cualquier de parámetros tales como la diferencia de tiempo entre canales parámetro o en el parámetro de la diferencia de fases entre canales, sin embargo no será audible en las señales media/lateral (Mid/Side) en el dominio del tiempo obtenido por la etapa 305 en la figura 14d.

60 **[0145]** La figura 13 ilustra un diagrama de bloques de una forma de realización de un aparato para decodificación de una señal multicanal codificada recibida en la línea de entrada 50.

65 **[0146]** En particular, la señal es recibida por una interfaz de entrada 600. Conectados a la interfaz de entrada 600 hay un decodificador de señales 700, y un desalineador de señales 900. Además, hay un procesador de señales 800 conectado a un decodificador de señales 700 por una parte y conectado al desalineador de señales por otra parte.

- [0147]** En particular, la señal multicanal codificada comprende una señal media (Mid) codificada, una señal lateral (Side) codificada, información sobre la utilización del parámetro de alineación de banda ancha e información acerca de la pluralidad de parámetros de banda estrecha. Por lo tanto, la señal multicanal codificada en la línea 50 puede ser exactamente la misma señal que la emitida por la interfaz de salida de 500 de la figura 12.
- [0148]** Sin embargo, es importante observar que, a diferencia de lo que se ilustra en la figura 12, la utilización del parámetro de alineación de banda ancha y de la pluralidad de parámetros de alineación de banda ancha incluidos en la señal codificada en una forma determinada pueden ser exactamente los parámetros de alineación utilizados por el alineador de señales 200 en la figura 12 pero, como alternativa, también pueden ser los valores inversos de los mismos, es decir, parámetros que pueden ser utilizados exactamente por las mismas operaciones llevadas a cabo por el alineador de señales 200 pero con valores inversos de manera que se obtenga la desalineación.
- [0149]** Por lo tanto, la información sobre los parámetros de alineación pueden ser los parámetros de alineación utilizados por el alineador de señales 200 en la figura 12 o pueden ser valores inversos, es decir, "parámetros de desalineación" reales. Adicionalmente, típicamente estos parámetros serán cuantificados en una forma determinada como se explica más adelante con respecto a la figura 8.
- [0150]** La interfaz de entrada 600 de la figura 13 separa la información acerca de la utilización del parámetro de alineación de banda ancha y de la pluralidad de parámetros de alineación de banda ancha de las señales media/lateral (Mid/Side) codificadas y comunica esta información por medio de línea de parámetros 610 al desalineador de señales 900. Por otra parte, la señal media (Mid) codificada se comunica al decodificador de señales 700 por medio de la línea 601 y la señal lateral (Side) codificada se comunica al decodificador de señales 700 por medio de la línea de señales 602.
- [0151]** El decodificador de señales está configurado para decodificación de la señal media (Mid) codificada y para decodificación de la señal lateral (Side) codificada para obtener una señal media (Mid) decodificada en la línea 701 y una señal lateral (Side) decodificada en la línea 702. Estas señales las utiliza el procesador de señales 800 para calcular una señal de primer canal decodificada o una señal izquierda decodificada y para calcular un segundo canal decodificado o una señal de canal derecho decodificada de la señal media (Mid) decodificada y la señal lateral (Side) decodificada, y el primer canal decodificado y el segundo canal decodificado son emitidos en las líneas 801, 802, respectivamente. El desalineador de señales 900 está configurado para desalinear el primer canal codificado en la línea 801 y el canal derecho decodificado 802 mediante el uso de la información acerca de la utilización del parámetro de alineación de banda ancha y adicionalmente mediante el uso de la información sobre la pluralidad de parámetros de alineación de banda ancha para obtener una señal multicanal decodificada, es decir, una señal decodificada que tiene al menos dos canales decodificados y desalineados en las líneas 901 y 902.
- [0152]** La figura 9a ilustra una secuencia de etapas preferida llevada a cabo por el desalineador de señales 900 de la figura 13. Específicamente, la etapa 910 recibe los canales izquierdo y derecho alineados disponibles sobre las líneas 801, 802 de la figura 13. En la etapa 910, el desalineador de señales 900 desalineará subbandas individuales mediante el uso de la información acerca de la utilización de los parámetros de alineación de banda ancha con el fin de obtener canales primero y segundo o izquierdo y derecho decodificados, desalineados en fase en 911a y 911b. En la etapa 912, los canales se desalinean mediante el uso del parámetro de alineación de banda ancha de tal manera que en 913a y 913b, se obtienen canales desalineados en fase y tiempo.
- [0153]** En la etapa 914, se lleva a cabo cualquier procesamiento adicional que comprenda el uso de una puesta en ventana o cualquier operación de superposición-adición o, en términos generales, cualquier operación del desvanecimiento cruzado con el fin de obtener, en 915a o 915b, una señal decodificada reducida en artefactos o libre de artefactos, es decir, a señales decodificadas que no tengan ningún artefacto aunque típicamente ha habido parámetros de desalineación variables en el tiempo para la banda ancha por una parte y para la pluralidad de bandas estrechas por otra parte.
- [0154]** La figura 15b ilustra una implementación preferida del decodificador multicanal ilustrado en la figura 13.
- [0155]** En particular, el procesador de señales 800 de la figura 13 comprende un convertidor de tiempo-espectro 810.
- [0156]** El procesador de señales comprende además un convertidor medio/lateral (Mid/Side) o izquierdo/derecho 820 con el fin de calcular a partir de una señal media (Mid) M y una señal lateral (Side) S una señal izquierda L y una señal derecha R.
- [0157]** Sin embargo, es importante observar que con el fin de calcular L y R mediante la conversión media/lateral Mid/Side-izquierda/derecha en el bloque 820, no es necesario utilizar la señal lateral (Side) S. En

cambio, como se expone más adelante en esta invención, las señales izquierda/derecha se calculan inicialmente mediante el uso solamente de un parámetro de ganancia derivado de un parámetro de diferencia de nivel entre los canales ILD. Por ello, en esta implementación, la señal lateral (Side) S se utiliza solamente en el actualizador de canales 830 que opera con el fin de proporcionar una mejor señal izquierda/derecha mediante el uso de la señal lateral (Side) S transmitida como se ilustra mediante la línea de desvío 821.

[0158] Por tanto, el convertidor 820 opera utilizando un parámetro de nivel obtenido por medio de una entrada de parámetro de nivel 822 y sin utilizar realmente la señal lateral (Side) S pero el actualizador de canales 830 seguidamente opera utilizando el lado (Side) 821 y, en función de la implementación específica, mediante el uso de un parámetro de llenado de estéreo recibido por medio de la línea 831. El alineador de señales 900 comprende seguidamente un desalineador de fase y escalador de energía 910. La puesta en escala de energía la controla un factor de puesta en escala derivado mediante un calculador del factor de puesta en escala 940. El calculador del factor de puesta en escala 940 es alimentado por la salida del actualizador de canales 830. En base a los parámetros de alineación de banda ancha recibidos por medio de la entrada 911, se lleva a cabo la desalineación de fase y, en el bloque 920, en base a la utilización del parámetro de alineación de banda ancha recibido por medio de la línea 921, se lleva a cabo la desalineación en tiempo. Finalmente, se incrementa una conversión de espectro-tiempo 930 con el fin de obtener finalmente la señal decodificada.

[0159] La figura 15c ilustra otra secuencia de etapas llevadas a cabo típicamente dentro de los bloques 920 y 930 de la figura 15b en una forma de realización preferida.

[0160] Específicamente, los canales de banda estrecha desalineados se introducen en la funcionalidad de desalineación de banda ancha correspondiente al bloque 920 de la figura 15b. Se lleva a cabo una DFT o cualquier otra transformada en el bloque 931. Subsiguientemente al cálculo real de las muestras en el dominio del tiempo, se lleva a cabo una puesta en ventana de síntesis opcional mediante el uso de una ventana de síntesis. La ventana de síntesis es preferentemente exactamente la misma que la ventana de análisis o se deriva de la ventana de análisis, por ejemplo, interpolación o decimación, pero depende de cierta manera de la ventana de análisis. Es preferible que esta dependencia sea tal que los factores de multiplicación definidos por dos ventanas de superposición se añadan a uno de cada punto en el intervalo de superposición. Por lo tanto, subsiguientemente a la ventana de síntesis en el bloque 932, se lleva a cabo una operación de superposición y una operación de adición subsiguiente. Como alternativa, en lugar de la puesta en ventana de síntesis y operación de superposición/adición, se lleva a cabo cualquier desvanecimiento entre bloques subsiguientes para cada canal con el fin de obtener, como ya se ha dispuesto en el contexto de la figura 15a, una señal decodificada reducida de artefacto.

[0161] Cuando se considera la figura 6b, se pone en evidencia que las operaciones de decodificación reales para la señal media (Mid), es decir, el "Decodificador de EVS" por una parte y, para la señal lateral (Side), la cuantificación inversa de vector VQ^{-1} y la operación inversa de MDCT (IMDCT) corresponden al decodificador de señales 700 de la figura 13.

[0162] Además, las operaciones de DFT en los bloques 810 corresponden al elemento 810 en la figura 15b y las funcionalidades del procesamiento estéreo inverso y el desplazamiento inverso de tiempo corresponden a los bloques 800, 900 de la figura 13 y las operaciones inversas de DFT 930 en la figura 15b corresponden a la correspondiente operación en el bloque 930 en la figura 15b.

[0163] Subsiguientemente, se expone con mayor detenimiento la figura 3d. En particular, en la figura 3d se ilustra un espectro de DFT que tiene líneas espectrales individuales. Es preferible que el espectro de DFT o cualquier otro espectro ilustrado en la figura 3d sea un espectro complejo y que cada línea sea una línea de espectro complejo que tiene magnitud y fase o que tiene una parte real y una parte imaginaria.

[0164] Adicionalmente, el espectro también se divide en diferentes bandas de parámetros. Cada banda de parámetro tiene preferentemente al menos una línea espectral y preferentemente más de una línea espectral. Adicionalmente, la utilización de las bandas de parámetro aumenta de frecuencias más bajas a frecuencias más elevadas. Típicamente, la utilización del parámetro de alineación de banda ancha es un único parámetro de alineación de banda ancha para la totalidad del espectro, es decir, para un espectro que comprende la totalidad de las bandas 1 a 6 en el ejemplo de forma de realización en la figura 3d.

[0165] Además, la pluralidad de parámetros de alineación de banda estrecha se proporciona de tal manera que exista un único parámetro de alineación para cada banda de parámetro. Esto significa que el parámetro de alineación para una banda se aplica siempre a la totalidad de los valores espectrales dentro de la banda correspondiente.

[0166] Por otra parte, además de la utilización de los parámetros de alineación de banda estrecha, también se proporcionan parámetros de nivel para cada banda de parámetros.

[0167] A diferencia de los parámetros de nivel que se proporcionan para cada una de las bandas de

parámetros y para la totalidad de ellas, de la banda 1 a la banda 6, se prefiere proporcionar la pluralidad de parámetros de alineación de banda estrecha solamente para un número limitado a de bandas inferiores tales como las bandas 1, 2, 3 y 4.

5 **[0168]** Adicionalmente, se proporcionan parámetros de llenado de estéreo para un determinado número de bandas que incluyen las bandas inferiores tales como, en el ejemplo de forma de realización, para las bandas 4, 5 y 6, mientras que hay valores espectrales de señal lateral (Side) para las bandas de parámetro inferiores 1, 2 y 3 y, por lo tanto, no existen parámetros de llenado de estéreo para estas bandas inferiores en los que la concordancia de forma de onda se obtiene mediante el uso de la señal lateral (Side) como tal o una señal de predicción residual que
10 representa la señal lateral (Side).

[0169] Como ya se ha mencionado, existen más líneas espectrales en las bandas superiores tales como, en la forma de realización en la figura 3d, siete líneas espectrales en la banda de parámetros 6 en comparación con solamente tres líneas espectrales en la banda de parámetros 2. Sin embargo, por supuesto, el número de banda de
15 parámetros, el número de líneas espectrales y el número de líneas espectrales dentro de una banda de parámetros y también los límites diferentes para determinar los parámetros serán diferentes.

[0170] Sin embargo, en la figura 8 se ilustra una distribución de la utilización de los parámetros y del número de bandas para los que se proporcionan parámetros en una determinada forma de realización en la que, a diferencia
20 de la figura 3d, hay realmente 12 bandas.

[0171] Como se ilustra, el parámetro de nivel ILD se proporciona para cada una de las 12 bandas y se cuantifica con una exactitud de cuantificación representada por cinco bits por banda.

25 **[0172]** Además, los parámetros de alineación de banda estrecha IPD se proporcionan solamente para las bandas inferiores hasta una frecuencia límite de 2,5 kHz. Adicionalmente, la diferencia de tiempo entre canales o parámetros de alineación de banda ancha se proporciona solamente como un parámetro único para la totalidad del espectro pero con una cuantificación de exactitud de cuantificación muy elevada representada por ocho bits para la banda entera.
30

[0173] Además, de una manera más bien aproximada se proporcionan parámetros de llenado de estéreo mediante los tres bits por banda y no para las bandas inferiores por debajo de 1 kHz dado que, para las bandas inferiores, se incluyen señales laterales (Side) realmente codificadas o valores residuales espectrales de señal lateral (Side).
35

[0174] A continuación se resume un procesamiento preferido en el lado del codificador. En una primera etapa, se lleva a cabo un análisis de DFT en los canales izquierdo y derecho. Este procedimiento corresponde a las etapas 155 a 157 de la figura 14c. Se calcula el parámetro de alineación de banda ancha y, en particular, la diferencia de tiempo entre canales correspondiente al parámetro de alineación preferido (ITD). Se lleva a cabo un desplazamiento de tiempo de L y R en el dominio de las frecuencias. Como alternativa, este desplazamiento de tiempo también puede llevarse a cabo en el dominio del tiempo. A continuación se lleva a cabo un DFT inverso, se ejecuta el desplazamiento de tiempo en el dominio del tiempo y lleva a cabo una DFT directa adicional con el fin de tener una vez más representaciones espectrales posteriores a la alineación mediante el uso del parámetro de alineación de banda ancha.
40
45

[0175] Los parámetros ILD, es decir, los parámetros de nivel y los parámetros de fase (parámetros IPD), se calculan para cada banda de parámetros en las representaciones L y D desplazadas. Esta etapa corresponde a la etapa 160 de la figura 14c, por ejemplo. Las representaciones L y D desplazadas en el tiempo se hacen rotar como una función de los parámetros de diferencia de fase entre canales como se ilustra en la etapa 161 de la figura 14c.
50 Subsiguientemente, las señales media (Mid) y lateral (Side) se calculan como se ilustra en la etapa 301 y, es preferible que adicionalmente con una operación de conservación de energía como se expone más adelante en esta invención. Además, se lleva a cabo una predicción de S con M como una función de ILD y opcionalmente con una señal M pasada, es decir, una señal media (Mid) de una trama anterior. Subsiguientemente, se ejecuta una DFT inversa de la señal media (Mid) y de la señal lateral (Side) que corresponde a las etapas 303, 304, 305 de la figura
55 14d en la forma de realización preferida.

[0176] En la etapa final, se codifican la señal media (Mid) en el dominio del tiempo y opcionalmente, la señal residual. Este procedimiento se corresponde al que se lleva a cabo mediante el codificador de señales 400 en la figura 12.
60

[0177] En el decodificador en el procesamiento estéreo inverso, se genera la señal lateral (Side) en el dominio de DFT y se predice por primera vez a partir de la señal media (Mid):

$$\overline{Side} = g \cdot Mid$$

donde g es una ganancia calculada para cada banda de parámetros y va en función de la Diferencia de niveles entre canales (ILDs) transmitida.

5 **[0178]** El residual de la predicción $\widehat{Side} - g \cdot Mid$ puede refinarse seguidamente de dos maneras diferentes:

- Mediante una codificación secundaria de la señal residual:

10
$$\widehat{Side} = g \cdot Mid + g_{cod} \cdot (\widehat{Side} - g \cdot Mid)$$

donde g_{cod} es una ganancia global transmitida para la totalidad del espectro

- Mediante una predicción residual, conocida como llenado estéreo, se predice el espectro lateral (side) residual con la Señal de espectro media (Mid) previamente decodificada a partir de la trama de DFT previo:

15
$$\widehat{Side} = g \cdot Mid + g_{pred} \cdot Mid \cdot z^{-1}$$

donde g_{pred} es una ganancia predictiva transmitida por banda de parámetros.

20 **[0179]** Los dos tipos de refinamiento de la codificación pueden mezclarse dentro del mismo espectro de DFT. En la forma de realización preferida, la codificación residual se aplica sobre las bandas de parámetro inferiores, mientras que la predicción residual se aplica sobre las bandas restantes. En la forma de realización preferida, la codificación residual ilustrada en la figura 12 se lleva a cabo en el Dominio de MDCT después de sintetizar la señal lateral (Side) residual en el Dominio del Tiempo y transformándolo mediante una MDCT. A diferencia de la DFT, la MDCT tiene un muestreo crítico y es más adecuado para la codificación de audio. Los coeficientes de MDCT se cuantifican directamente por vector mediante una Cuantificación Vectorial de Retícula pero como alternativa también pueden ser codificados por un Cuantificador Escalar seguido por un codificador de entropía. Como alternativa, la señal lateral (Side) residual también puede ser codificada en el Dominio del Tiempo mediante una técnica de

25

30 codificación del habla o directamente en el dominio de DFT.

[0180] Subsiguientemente se describe otra forma de realización de un procesamiento de codificador estéreo/multicanal conjunto o procesamiento estéreo/multicanal inverso.

35 **1. Análisis de tiempo-frecuencia: DFT**

[0181] Es importante que la descomposición extra de tiempo-frecuencia del procesamiento estéreo efectuada mediante DFTs permita un buen análisis de la escena del auditorio sin incrementar de manera significativa el retardo global del sistema de codificación. Por defecto, se utiliza una resolución del tiempo de 10 ms (que es el doble de la estructura de 20 ms del codificador de núcleo). Las ventanas de análisis y síntesis son iguales y simétricas. La ventana se representa con 16 kHz de velocidad de muestreo en la figura 7a-7b. Puede observarse que la región de superposición está limitada con el fin de reducir el retardo generado y que también se añade el padding cero para equilibrar el desplazamiento circular cuando se aplique ITD en el dominio de las frecuencias como se explicará más adelante en esta invención.

45 **2. Parámetros estéreo**

[0182] Los parámetros estéreo pueden transmitirse al máximo con la resolución de tiempo del DFT estéreo. Como mínimo se puede reducir a la resolución de estructura del codificador de núcleo, es decir 20 ms. Por defecto, si no se detectan transitorios, los parámetros se calculan cada 20 ms en las ventanas de DFT. La utilización de las bandas de parámetros constituye una descomposición no uniforme y no superpuesta del espectro según aproximadamente 2 veces o 4 veces el ERB (Equivalent Rectangular Bandwidth, Ancho de Banda Rectangular Equivalente). Por defecto, se utiliza un ERB de 4 veces para un total de 12 bandas para un ancho de banda de frecuencia de 16kHz (32 kbps de velocidad de muestreo, estéreo de superancho de banda). En la figura 8 se resume un ejemplo de configuración, para el que la información lateral de estéreo se transmite aproximadamente a 5 kbps.

3. Computación de ITD y alineación de tiempo de canal

[0183] El ITD se calcula estimando del TDOA (Time Delay of Arrival, Retardo Temporal de Llegada), utilizando la Correlación Cruzada Generalizada (Generalized Cross Correlación) con transformada de fase (GCC-PHAT):

60

$$ITD = \operatorname{argmax} \left(\operatorname{IDFT} \left(\frac{L_i(f)R_i^*(k)}{|L_i(f)R_i^*(k)|} \right) \right)$$

donde L y R son los espectros de frecuencia de los canales izquierdo y derecho respectivamente. El análisis de frecuencia puede llevarse a cabo independientemente de la DFT utilizada para el subsiguiente procesamiento 5 estéreo o se puede compartir. El pseudocódigo para calcular el ITD es el siguiente:

```

L =fft(ventana(l));
R =fft(ventana(r));
tmp = L .* conj(R);
10 sfm_L = prod(abs(L).^(1/length(L)))/(mean(abs(L))+eps);
    sfm_R = prod(abs(R).^(1/length(R)))/(mean(abs(R))+eps);
    sfm = max(sfm_L,sfm_R);
    h.cross_corr_smooth = (1-sfm)*h.cross_corr_smooth+sfm*tmp;
    tmp = h.cross_corr_smooth ./ abs(h.cross_corr_smooth+ep);
15 tmp = ifft(tmp);
    tmp = tmp([length(tmp)/2+1:length(tmp) 1:length(tmp)/2+1]);
    tmp_sort = sort(abs(tmp));
    thresh = 3 * tmp_sort(round(0.95*length(tmp_sort)));
    xcorr_time=abs(tmp(-(h.stereo_itd_q_max - (length(tmp)-1)/2 - 1):-
20 (h.stereo_itd_q_min - (length(tmp)-1)/2 - 1)));
    %smooth output for better detection
    xcorr_time=[xcorr_time 0];
    xcorr_time2=filter([0.25 0.5 0.25],1,xcorr_time);
    [m,i] = max(xcorr_time2(2:end));
25 si m > thresh
    itd = h.stereo_itd_q_max - i + 1;

    también

30 itd = 0;
    fin

```

[0184] La computación de ITD también puede resumirse del siguiente modo. La correlación cruzada se computa en el dominio de las frecuencias antes de ser suavizada en función de la Medición de Llaneza Espectral 35 (SFM, Spectral Flatness Measurement). La SFM está delimitada entre 0 y 1. En el caso de señal similares a ruido, la SFM será elevada (es decir, será de aproximadamente 1) y la suavización será débil. En el caso de una señal similar a tono, la SFM será baja y la suavización se hará más fuerte. La correlación cruzada suavizada se normaliza seguidamente en amplitud antes de ser transformada de vuelta al dominio del tiempo. La normalización corresponde a la transformada de fase de la correlación cruzada, y es sabido que muestra un mejor comportamiento que la 40 correlación cruzada en entornos de bajo ruido y con una reverberación relativamente elevada. La función en el dominio del tiempo, así obtenida, se filtra en primer lugar con el fin de lograr una formación de pico más robusta. El índice correspondiente a la máxima amplitud corresponde a una estimación de la diferencia de tiempo entre canal izquierdo y canal derecho (ITD). Si la amplitud del máximo es inferior a un umbral dado, entonces se considera que el valor estimado para ITD no es fiable y se establece en cero.

45 **[0185]** Si se aplica la alineación del tiempo en el dominio del tiempo, se computa el ITD en un análisis de DFT por separado. El desplazamiento se efectúa del siguiente modo:

$$\begin{cases} r(n) = r(n + ITD) \text{ if } ITD > 0 \\ l(n) = l(n - ITD) \text{ if } ITD < 0 \end{cases}$$

50 **[0186]** Requiere un retardo extra en el codificador, que es igual como máximo al ITD máximo absoluto que puede ser procesado. La variación del ITD a lo largo del tiempo se suaviza mediante la puesta en ventana de análisis del DFT.

55 **[0187]** Como alternativa, la alineación en el tiempo puede llevarse a cabo en el dominio de las frecuencias. En este caso, el cómputo de ITD y el desplazamiento circular se encuentran en el mismo dominio de DFT, dominio compartido con este otro procesamiento estéreo. El desplazamiento circular está dado por:

$$\begin{cases} L(f) = L(f)e^{-j2\pi f \frac{ITD}{2}} \\ R(f) = R(f)e^{+j2\pi f \frac{ITD}{2}} \end{cases}$$

[0188] Se necesita el padding cero de las ventanas de DFT para simular un desplazamiento temporal con un desplazamiento circular. La magnitud del padding cero corresponde al ITD máximo absoluto que puede ser procesado. En la forma de realización preferida, el padding cero se distribuye uniformemente en ambos lados de las ventanas de análisis, mediante la adición de 3,125 ms de ceros en ambos extremos. El ITD máximo absoluto posible es entonces de 6,25 ms. En el montaje de los micrófonos A-B, corresponde al peor caso de una distancia máxima de aproximadamente 2,15 metros entre los dos micrófonos. La variación en ITD a lo largo del tiempo se suaviza mediante la puesta en ventana de síntesis y superposición-adición de la DFT.

10 [0189] Es importante que el desplazamiento temporal sea seguido por una puesta en ventana de la señal desplazada. Se trata de una distinción principal con respecto a la BCC (Binaural Cue Coding, Codificación Binaural por Indicios) de la técnica anterior, en donde el desplazamiento en tiempo se aplica sobre una señal puesta en ventana pero no es puesta en ventana adicionalmente en la etapa de síntesis. Como una consecuencia, cualquier cambio en ITD a lo largo del tiempo produce un transitorio artificial/clic en la señal decodificada.

4. Computación de IPDs y rotación de canales

[0190] Los IPDs se calculan después de la alineación temporal de los dos canales y esto para cada banda de parámetros o al menos hasta una dada $ipd_{i,max,band}$, en función de la configuración estéreo.

$$IPD[b] = angle \left(\sum_{k=band_{limits}[b]}^{band_{limits}[b+1]} L[k] R^*[k] \right)$$

25 [0191] A continuación se aplican los IPDs a los dos canales para alinear sus fases:

$$\begin{cases} L'(k) = L(k)e^{-j\beta} \\ R'(k) = R(k)e^{j(IPD[b]-\beta)} \end{cases}$$

[0192] Donde $\beta = atan2(\sin(IPD_i[b]), \cos(IPD_i[b]) + c)$, $c = 10^{\frac{ILD_i[b]}{20}}$ y b está basado en el índice de la banda de parámetros al que pertenece el índice de frecuencia k . La utilización del parámetro β es responsable de distribuir la cantidad de rotación de fase entre los dos canales mientras se alinean sus fases. β depende del IPD pero también del nivel de amplitud relativa de los canales, ILD. Si un canal tiene una amplitud más elevada, será considerado como el canal de guía y estará menos afectado por la rotación de canal que el canal con la amplitud menor.

35 5. Suma-diferencia y codificación de la señal lateral (Side)

[0193] La transformación de suma diferencia se lleva a cabo en los espectros alineados en tiempo y fase de los dos canales de tal manera que se conserve la energía en la señal media (Mid).

$$\begin{cases} M(f) = (L'(f) + R'(f)) \cdot a \cdot \sqrt{\frac{1}{2}} \\ S(f) = (L'(f) - R'(f)) \cdot a \cdot \sqrt{\frac{1}{2}} \end{cases}$$

40 donde $a = \sqrt{\frac{L'^2 + R'^2}{(L' + R')^2}}$ está delimitado entre 1/1,2 y 1,2, es decir -1,58 y +1,58 dB. La limitación evita los artefactos cuando se ajusta la energía de M y S. Cabe observar que esta conservación de la energía es menos importante cuando la fase y el tiempo han sido alineados anteriormente. Como alternativa los límites pueden ser incrementados o disminuidos.

45 [0194] Además, la señal lateral (Side) S se predice con M:

$$S'(f) - S(f) - g(ILD)M(f)$$

donde $g(ILD) = \frac{c-1}{c+1}$, donde $c = 10^{\frac{ILD_i[b]}{20}}$. Como alternativa es posible hallar la ganancia de predicción óptima g minimizando el MSE (Mean Square Error, Error Cuadrado Medio) del residual e ILDs deducidos mediante la ecuación previa.

[0195] La señal residual $S'(f)$ puede modelarse mediante dos medios: ya sea prediciéndola con el espectro retardado de M o mediante su codificación directa en el dominio de MDCT.

10 **6. Decodificación estéreo**

[0196] La señal media (Mid) X y la señal lateral (Side) S se convierten en primer término en los canales izquierdo y derecho L y R del siguiente modo:

$$\begin{aligned} L_i[k] &= M_i[k] + gM_i[k], \text{ for } band_{limits[b]} \leq k < band_{limits[b+1]}, \\ R_i[k] &= M_i[k] - gM_i[k], \text{ for } band_{limits[b]} \leq k < band_{limits[b+1]}, \end{aligned}$$

donde la ganancia g por banda de parámetros se deriva del parámetro de ILD:

$$g = \frac{c-1}{c+1}, \text{ donde } c = 10^{\frac{ILD_i[b]}{20}}.$$

[0197] Para las bandas de parámetros inferiores a cod_max_band, los dos canales se actualizan con la Señal lateral (Side) decodificada:

$$\begin{aligned} L_i[k] &= L_i[k] + \mathbb{K}cod_gain_i \cdot S_i[k], \text{ for } 0 \leq k < band_limits[cod_max_band], \\ R_i[k] &= R_i[k] - \mathbb{K}cod_gain_i \cdot S_i[k], \text{ for } 0 \leq k < band_limits[cod_max_band], \end{aligned}$$

[0198] Para las bandas de parámetros superiores, se predice la señal lateral (Side) y los canales se actualizan del siguiente modo:

$$\begin{aligned} L_i[k] &= L_i[k] + cod_{pred_i}[b] \cdot M_{i-1}[k], \text{ for } band_{limits[b]} \leq k < band_{limits[b+1]}, \\ R_i[k] &= R_i[k] - cod_{pred_i}[b] \cdot M_{i-1}[k], \text{ for } band_{limits[b]} \leq k < band_{limits[b+1]}, \end{aligned}$$

[0199] Finalmente, se multiplican los canales por un valor complejo que tiene la finalidad de restaurar la energía original y la fase intercanal de la señal estéreo:

$$\begin{aligned} L_i[k] &= a \cdot e^{j2\pi\beta} \cdot L_i[k] \\ R_i[k] &= a \cdot e^{j2\pi\beta - IPD_i[b]} \cdot R_i[k] \end{aligned}$$

40 donde

$$a = \sqrt{2 \cdot \frac{\sum_{k=band_limits[b]}^{band_limits[b+1]} M_i^2[k]}{\sum_{k=band_limits[b]}^{band_limits[b+1]-1} L_i^2[k] + \sum_{k=band_limits[b]}^{band_limits[b+1]-1} R_i^2[k]}}$$

donde a se define y delimita como se ha definido anteriormente, y donde $\beta = \text{atan2}(\sin(IPD_i[b]), \cos(IPD_i[b]) + c)$, y donde atan2(x,y) es la tangente inversa de cuatro cuadrantes de x sobre y.

[0200] Finalmente, los canales son desplazados ya sea en tiempo o en el dominio de las frecuencias en función de los ITDs transmitidos. Los canales en el dominio del tiempo se sintetizan mediante DFTs inversos y superposición-adición.

[0201] Es posible almacenar una señal audio codificada según la invención en un medio de almacenamiento digital o en un medio de almacenamiento no transitorio, o se puede transmitir en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión alámbrico tal como Internet.

- 5 **[0202]** Aunque algunos aspectos se han descrito en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del procedimiento correspondiente, en donde un bloque o dispositivo corresponde a una etapa de procedimiento o a un rasgo de una etapa de procedimiento. De manera análoga, los aspectos descritos en el contexto de una etapa de procedimiento también representan una descripción de un bloque, elemento o rasgo, correspondientes, de un aparato correspondiente.
- 10 **[0203]** En función de determinados requisitos para la implementación, las formas de realización de la invención pueden implementarse en hardware o en software. La implementación puede llevarse a cabo mediante el uso de un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un CD, un ROM, un PROM, un EPROM, un EEPROM o una memoria FLASH, que tiene señales de control electrónicamente legibles almacenadas en él, que cooperan (o que son capaces de cooperar) con un sistema informático programable de tal manera que se lleve a cabo el procedimiento respectivo.
- 15 **[0204]** Algunas formas de realización 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 lleva a cabo uno de los procedimientos descritos en esta invención.
- 20 **[0205]** En términos generales, las formas de realización de la presente invención pueden implementarse como un producto programa informático con un código de programa, siendo el código de programa operativo para llevar a cabo uno de los procedimientos cuando se ejecuta el producto de programa informático en un ordenador. El código de programa puede estar almacenado, por ejemplo, en un soporte legible por máquina.
- 25 **[0206]** Otras formas de realización comprenden el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por máquina o en un medio de almacenamiento no transitorio.
- 30 **[0207]** Por lo tanto, y en otras palabras, una forma de realización de la invención consiste en un programa informático que tiene un código de programa para llevar a cabo uno de los procedimientos descritos en esta invención, cuando se ejecuta el programa informático en un ordenador.
- 35 **[0208]** Por lo tanto, otra forma de realización de los procedimientos según la invención consiste en un soporte de datos (o 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.
- 40 **[0209]** Por lo tanto, otra forma de realización del procedimiento según la invención se refiere a un flujo de datos o a una secuencia de señales que representa el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención. El flujo de datos o la secuencia de señales pueden estar configurados, por ejemplo, para ser transferidos por medio de una conexión de comunicación de datos, por ejemplo, a través de Internet.
- 45 **[0210]** Y otra forma de realización 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.
- [0211]** Otra forma de realización comprende un ordenador que tiene instalado en él el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención.
- 50 **[0212]** En algunas formas de realización, puede utilizarse un dispositivo lógico programable (por ejemplo, una matriz de puertas programable de campo) para llevar a cabo algunas o la totalidad, de las funcionalidades de los procedimientos descritos en esta invención. En algunas formas de realización, 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 temas generales, los procedimientos se implementan preferentemente mediante un aparato de hardware.
- 55 **[0213]** Las formas de realización anteriormente descritas son meramente ilustrativas para exponer los principios de la presente invención. Se da por entendido que las modificaciones y las variaciones de las disposiciones y los detalles descritos en esta invención serán evidentes para los expertos en la técnica. Por lo tanto, la finalidad es que la invención esté limitada solamente por el ámbito de las reivindicaciones de patente pendientes y no por los detalles específicos presentados a modo de descripción y explicación de las formas de realización de esta invención.
- 60

REIVINDICACIONES

1. Aparato para la codificación una señal de audio multicanal que comprende al menos dos canales, que comprende:

5

un convertidor tiempo-espectral (1000) para convertir secuencias de bloques de valores de muestra de los al menos dos canales en una representación en el dominio de las frecuencias que tiene secuencias de bloques de valores espectrales para los al menos dos canales, en el que un bloque de valores de muestreo tiene una velocidad de muestreo de entrada asociada, y un bloque de valores espectrales de las secuencias de bloques de valores

10

espectrales tiene valores espectrales hasta una frecuencia de entrada máxima (1211) que está relacionada con la velocidad de muestreo de entrada; estando el aparato **caracterizado porque** comprende:

un procesador multicanal (1010) para aplicar un procesamiento multicanal conjunto a las secuencias de bloques de valores espectrales o a secuencias remuestreadas de bloques de valores espectrales para obtener al menos una

15

secuencia resultado de bloques de valores espectrales que comprende información relacionada con los al menos dos canales;

un remuestreador en el dominio espectral (1020) para el muestreo repetido de los bloques de las secuencias resultado en el dominio de las frecuencias o para el muestreo repetido de las secuencias de bloques de valores espectrales para los al menos dos canales en el dominio de las frecuencias para obtener una secuencia

20

remuestreada de bloques de valores espectrales, en el que un bloque de la secuencia remuestreada de bloques de valores espectrales tiene valores espectrales hasta una frecuencia de salida máxima (1231, 1221) que es diferente de la frecuencia de entrada máxima (1211);

un convertidor espectral-tiempo (1030) para convertir la secuencia remuestreada de bloques de valores espectrales en una representación en el dominio del tiempo o para convertir la secuencia resultado de bloques de valores

25

espectrales en una representación en el dominio del tiempo que comprende una secuencia de salida de bloques de valores de muestreo asociados con velocidad de muestreo de salida que es diferente de la velocidad de muestreo de entrada; y

un codificador de núcleo (1040) para la codificación de la secuencia de bloques emitida de valores de muestreo para obtener una señal multicanal codificada (1510).

30

2. Aparato de la reivindicación 1,

en el que el remuestreador en el dominio espectral (1020) está configurado para truncar los bloques con fines de muestreo descendente o para padding cero de los bloques con fines de muestreo ascendente.

35

3. Aparato de la reivindicación 1 o 2,

en el que el remuestreador en el dominio espectral (1020) está configurado para poner en escala (1322) los valores espectrales de los bloques de la secuencia resultado de bloques mediante el uso de un factor de puesta en escala que depende de la frecuencia de entrada máxima y que depende de la frecuencia de salida máxima.

40

4. Aparato de la reivindicación 3,

en el que el factor de puesta en escala es mayor que uno en el caso del muestreo ascendente, en el que la velocidad de muestreo de salida es mayor que la velocidad de muestreo de entrada, o en el que el factor de puesta en escala es inferior a uno en el caso del muestreo descendente, en el que la velocidad de muestreo de salida es inferior a la velocidad de muestreo de entrada, o

45

en el que el convertidor tiempo-espectral (1000) está configurado para llevar a cabo un algoritmo de transformada de tiempo-frecuencia sin utilizar una normalización en cuanto a un número total de valores espectrales un bloque de valores espectrales (1311), y en el que el factor de puesta en escala es igual a un cociente entre el número de valores espectrales de un bloque de la secuencia remuestreada y el número de valores espectrales de un bloque de valores espectrales antes del muestreo repetido, y en el que el convertidor espectral-tiempo está configurado para aplicar una normalización basada en la frecuencia de salida máxima (1331).

55

5. Aparato de una de las reivindicaciones anteriores,

en el que el convertidor tiempo-espectral (1000) está configurado para llevar a cabo un algoritmo de transformada de Fourier discreta, o en el que el convertidor espectral-tiempo (1030) está configurado para llevar a cabo un algoritmo de transformada de Fourier inversa discreta.

60

6. Aparato de la reivindicación 1,

en el que el procesador multicanal (1010) está configurado para obtener otra secuencia resultado de bloques de valores espectrales, y

65

en el que el convertidor espectral-tiempo (1030) está configurado para convertir la otra secuencia resultado de valores espectrales en otra representación en el dominio del tiempo (1032) que comprende otra secuencia de salida de bloques de valores de muestreo asociados con velocidades de muestreo de salida que son iguales a la velocidad de muestreo de entrada.

5
7. Aparato de una de las reivindicaciones anteriores,

en el que el procesador multicanal (1010) está configurado para proporcionar e incluso otras secuencias resultado de bloques de valores espectrales,

10
en el que el remuestreador en el dominio espectral (1020) está configurado para el muestreo repetido de bloques de incluso la secuencia resultado en el dominio de las frecuencias para obtener una secuencia remuestreada adicional de bloques de valores espectrales, en el que un bloque de la secuencia remuestreada adicional tiene valores espectrales de hasta una frecuencia de salida máxima adicional que es diferente de la frecuencia de salida máxima
15 o que es diferente de la frecuencia de entrada máxima y,

en el que el convertidor espectral-tiempo (1030) está configurado para convertir la secuencia remuestreada adicional de bloques de valores espectrales en incluso una representación en el dominio del tiempo adicional que comprende incluso una secuencia de salida adicional de bloques de valores de muestreo que tienen asociada una velocidad de
20 muestreo de salida adicional que es diferente de la velocidad de muestreo de salida o de la velocidad de muestreo de entrada.

8. Aparato de una de las reivindicaciones anteriores,

25 en el que el procesador multicanal (1010) está configurado para generar una señal media (Mid) como la al menos una secuencia resultado de bloques de valores espectrales mediante el uso solamente de una operación de mezclado de forma descendente, o una señal lateral (Side) adicional como una secuencia resultado de bloques de valores espectrales adicional.

30 9. Aparato de una de las reivindicaciones anteriores,

en el que el procesador multicanal (1010) está configurado para generar una señal media (Mid) como la al menos una secuencia resultado, en el que el remuestreador en el dominio espectral (1020) está configurado para muestrear la señal media (Mid) a dos secuencias separadas que tienen dos frecuencias de salida máxima diferentes que son
35 diferentes de la frecuencia de entrada máxima,

en el que el convertidor espectral-tiempo (1030) está configurado para convertir las dos secuencias remuestreadas en dos secuencias de salida que tienen velocidades de muestreo diferentes, y

40 en el que el codificador de núcleo (1030) comprende un primer pre-procesador (1430c) para pre-procesar la primera secuencia de salida con una primera velocidad de muestreo o un segundo pre-procesador (1430d) para pre-procesar la segunda secuencia de salida con la segunda velocidad de muestreo, y

en el que el codificador de núcleo está configurado para la codificación de la primera o la segunda señal pre-procesada, o

45 en el que el procesador multicanal está configurado para generar una señal lateral (Side) como la al menos una secuencia resultado, en el que el remuestreador en el dominio espectral (1020) está configurado para remuestrear la señal lateral (Side) a dos secuencias remuestreadas que tienen dos frecuencias de salida máxima diferentes que es diferente de la frecuencia de entrada máxima,

50 en el que el convertidor espectral-tiempo (1030) está configurado para convertir las dos secuencias remuestreadas en dos secuencias de salida que tienen velocidades de muestreo diferentes, y

en el que el codificador de núcleo comprende un primer pre-procesador (1430c) y un segundo pre-procesador
55 (1430d) para pre-procesar la primera y la segunda secuencias de salida; y

en el que el codificador de núcleo (1040) está configurado para codificar en núcleo (1430a, 1430b) la primera o segunda secuencia pre-procesada.

10. Aparato de una de las reivindicaciones anteriores,

60 en el que el convertidor espectral-tiempo (1030) está configurado para convertir la al menos una secuencia resultado en una representación en el dominio del tiempo sin ningún muestreo repetido en el dominio espectral, y

en el que el codificador de núcleo (1040) está configurado para codificar en núcleo (1430a) la secuencia de salida no
65 remuestreada para obtener la señal de audio multicanal codificada, o

en el que el convertidor espectral-tiempo (1030) está configurado para convertir la al menos una secuencia resultado en una representación en el dominio del tiempo sin ningún muestreo repetido en el dominio espectral sin la señal lateral (Side), y

5

en el que el codificador de núcleo (1040) está configurado para codificar en núcleo (1430a) la secuencia de salida no remuestreada para la señal lateral (Side) para obtener la señal de audio multicanal codificada, o en el que el aparato comprende además un codificador de señal lateral (Side) específico en el dominio espectral (1430e).

10

11. Aparato de una de las reivindicaciones anteriores,

en el que la velocidad de muestreo de entrada es al menos una velocidad de muestreo de un grupo de velocidades de muestreo que comprende 8 kHz, 16 kHz, 32 kHz, o

15

en el que la velocidad de muestreo de salida es al menos una velocidad de muestreo de un grupo de velocidades de muestreo que comprende 8 kHz, 12,8 kHz, 16 kHz, 25,6 kHz y 32 kHz.

12. Aparato de una de las reivindicaciones anteriores,

20

en el que el convertidor espectral-tiempo está configurado para aplicar una ventana de análisis,

en el que el convertidor espectral-tiempo (1030) está configurado para aplicar una ventana de síntesis,

25 en el que la longitud en tiempo de la ventana de análisis es igual a un múltiplo de número entero o fracción de número entero de la longitud en tiempo de la ventana de síntesis, o

en el que la ventana de análisis y la ventana de síntesis tienen, cada una de ellas, una porción de padding cero en una porción inicial o en una de sus porciones extremas, o

30

en el que una ventana de análisis utilizada por el convertidor tiempo-espectral (1000) o una ventana de síntesis utilizada por el convertidor espectral-tiempo (1030) tienen, cada una de ellas, una porción superpuesta creciente y una porción superpuesta decreciente, en el que el codificador de núcleo (1040) comprende un codificador en el dominio de tiempo con una anticipación (1905) o un codificador en el dominio de las frecuencias con una porción superpuesta de una ventana de núcleo, y en el que la porción superpuesta de la ventana de análisis o de la ventana de síntesis es más pequeña que o igual a la porción de anticipación (1905) del codificador de núcleo o de la porción superpuesta de la ventana de núcleo, o

35

en el que la ventana de análisis y la ventana de síntesis son tales que el tamaño de la ventana, un tamaño de la región de superposición, y un tamaño de la región de padding cero comprenden, cada uno de ellos, un número entero de muestras para al menos dos velocidades de muestreo del grupo de velocidades de muestreo que comprenden 12,8 kHz, 16 kHz, 26,6 kHz, 32 kHz, 48 kHz, o

40

en el que una base máxima de una transformada digital de Fourier en una implementación de raíz dividida es inferior o igual a 7, o en el que una resolución del tiempo ha sido fijada en un valor inferior o igual a, la velocidad de trama del codificador de núcleo.

45

13. Aparato de una de las reivindicaciones anteriores,

50 en el que el codificador de núcleo (1040) está configurado para operar según un primer control de trama para proporcionar una secuencia de tramas, en el que una trama está delimitada por un borde de inicio de la trama (1901) y un borde final de la trama (1902), y

en el que el convertidor tiempo-espectral (1000) o el convertidor espectral-tiempo (1030) están configurados para operar según un segundo control de trama que está sincronizado con el segundo control de trama, en el que el borde de inicio de la trama (1901) o el borde final de la trama (1902) de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con respecto a un instante de inicio o instante final de una porción superpuesta de una ventana utilizada por el convertidor tiempo-espectral (1000) para cada bloque de la secuencia de bloques de valores de muestreo o utilizado por el convertidor espectral-tiempo (1030) para cada bloque de la secuencia de salida de bloques de valores de muestreo.

55

60

14. Aparato de una de las reivindicaciones anteriores,

en el que el codificador de núcleo (1040) está configurado para utilizar una porción de anticipación (1905) cuando se codifica en núcleo una trama derivada de la secuencia de salida de bloques de valores de muestreo que están

65

asociados con la velocidad de muestreo de salida, estando la porción de anticipación (1905) situada en el tiempo subsiguiente a la trama,

5 en el que el convertidor tiempo-espectral (1000) está configurado para utilizar una ventana de análisis (1904) que tiene una porción superpuesta con una longitud en tiempo que es igual o inferior a la longitud en tiempo de la porción de anticipación (1905), en el que la porción superpuesta de la ventana de análisis se utiliza para generar una porción de anticipación puesta en ventana (1905).

10 15. Aparato de la reivindicación 14,

en el que el convertidor espectral-tiempo (1030) está configurado para procesar una porción de salida de anticipación correspondiente a la porción de puesta en ventana de anticipación mediante el uso de una función de compensación (1922), en el que la función de compensación está configurada de tal manera que se reduce o elimina una influencia de la porción superpuesta de la ventana de análisis.

15 16. Aparato de la reivindicación 15,

en el que la función de compensación es inversa de una función que define la porción superpuesta de la ventana de análisis.

20 17. Aparato de la reivindicación 15 o 16,

en el que la porción superpuesta es proporcional a una raíz cuadrada de función seno,

25 en el que la función de compensación es proporcional a una inversa de la raíz cuadrada de la función seno, y

en el que el convertidor espectral-tiempo (1030) está configurado para utilizar una porción superpuesta que es proporcional a la función $(\sin)^{1.5}$.

30 18. Aparato de una de las reivindicaciones anteriores,

en el que el convertidor espectral-tiempo (1030) está configurado para generar un primer bloque de salida mediante el uso de una ventana de síntesis y un segundo bloque de salida mediante el uso de la ventana de síntesis, en el que una segunda porción del segundo bloque de salida es una porción de salida de anticipación (1905),

35 en el que el convertidor espectral-tiempo (1030) está configurado para generar valores de muestreo de una trama mediante el uso de una operación de superposición-adición entre el primer bloque de salida y la porción del segundo bloque de salida que excluye la porción de salida de anticipación (1905),

40 en el que el codificador de núcleo (1040) está configurado para aplicar una operación de anticipación a la porción de salida de anticipación (1905) con el fin de determinar información de codificación para la codificación en núcleo de la trama, y

45 en el que el codificador de núcleo (1040) está configurado para codificar en núcleo la trama mediante el uso de un resultado de la operación de anticipación.

19. Aparato de la reivindicación 18,

50 en el que el convertidor espectral-tiempo (1030) está configurado para generar un tercer bloque de salida subsiguiente al segundo bloque de salida mediante el uso de la ventana de síntesis, en donde el convertidor espectral-tiempo está configurado para superponerse a una primera porción de superposición del tercer bloque de salida con la segunda porción del segundo bloque de salida puesto en ventana mediante el uso de la ventana de síntesis para obtener muestras de otra trama que sigue a la trama en tiempo.

55 20. Aparato de la reivindicación 18 y 19,

60 en el que el convertidor espectral-tiempo (1030) está configurado, cuando genera el segundo bloque de salida para la trama, para no poner en ventana la porción de salida de anticipación o para compensar (1922) la porción de salida de anticipación para deshacer, al menos parcialmente, una influencia de una ventana de análisis utilizada por el convertidor tiempo-espectral (1000), y en el que el convertidor espectral-tiempo (1030) está configurado para llevar a cabo una operación de superposición-adición (1924) entre el segundo bloque de salida y el tercer bloque de salida para la trama adicional y para poner en ventana (1920) la porción de salida de anticipación con la ventana de síntesis.

65 21. Aparato de cualquiera de las reivindicaciones 13 a 20,

en el que el convertidor espectral-tiempo (1030) está configurado,

5 para utilizar una ventana de síntesis para generar un primer bloque de muestras de salida y un segundo bloque de muestras de salida,

para superponer-añadir una segunda porción del primer bloque y una primera porción del segundo bloque para generar una porción de muestras de salida,

10 en el que el codificador de núcleo (1040) está configurado para aplicar una operación de anticipación a la porción de las muestras de salida para codificar en núcleo las muestras de salida situadas en el tiempo antes de la porción de las muestras de salida, en el que la porción de anticipación no incluye una segunda porción de muestras del segundo bloque.

15 22. Aparato de la reivindicación 13,

en el que el convertidor espectral-tiempo (1030) está configurado para generar una ventana de síntesis que proporciona una resolución en tiempo que es superior a dos veces la longitud de un codificador de núcleo de trama,

20 en el que el convertidor espectral-tiempo (1030) está configurado para utilizar la ventana de síntesis para generar bloques de muestras de salida y para llevar a cabo una operación de superposición-adición, en el que la totalidad de las muestras en una porción de anticipación del codificador de núcleo se calculan mediante el uso de la operación de superposición-adición, o

25 en el que el convertidor espectral-tiempo (1030) está configurado para aplicar una operación de anticipación a las muestras de salida para codificar en núcleo muestras de salida situadas en el tiempo antes de la porción, en el que la porción de anticipación no incluye una segunda porción de muestras del segundo bloque.

23. Aparato de una de las reivindicaciones anteriores,

30

en el que el procesador multicanal (1010) está configurado para procesar la secuencia de bloques para obtener una alineación en el tiempo mediante el uso del parámetro de alineación en tiempo de banda ancha (12) y para obtener una alineación de fase de banda estrecha mediante el uso de una pluralidad de parámetros de alineación de fase de banda estrecha (14), y para calcular una señal media (Mid) y una señal lateral (Side) como secuencias resultado

35

mediante el uso de secuencias alineadas.

24. Procedimiento para la codificación de una señal de audio multicanal que comprende al menos dos canales, que comprende:

40 convertir (1000) secuencias de bloques de valores de muestras de los al menos dos canales en una representación en el dominio de las frecuencias que tiene secuencia de bloques de valores espectrales para los al menos dos canales, en el que un bloque de valores de muestreo tiene una velocidad de muestreo de entrada asociada, y un bloque de valores espectrales de las secuencias de bloques de valores espectrales tiene valores espectrales hasta una frecuencia de entrada máxima (1211) que está relacionada con la velocidad de muestreo de entrada; estando el

45

procedimiento **caracterizado porque** comprende además:

aplicar (1010) un procesamiento multicanal conjunto a las secuencias de bloques de valores espectrales o a secuencias remuestreadas de bloques de valores espectrales para obtener al menos una secuencia resultado de bloques de valores espectrales que comprende información relacionada con los al menos dos canales;

50 un muestreo repetido en el dominio espectral (1020) de los bloques de las secuencias resultado en el dominio de las frecuencias o muestreo repetido de las secuencias de bloques de valores espectrales para los al menos dos canales en el dominio de las frecuencias para obtener una secuencia remuestreada de bloques de valores espectrales, en el que un bloque de la secuencia remuestreada de bloques de valores espectrales tiene valores espectrales hasta una frecuencia de salida máxima (1231, 1221) que es diferente de la frecuencia de entrada máxima (1211);

55 convertir (1640) la secuencia remuestreada de bloques de valores espectrales en una representación en el dominio del tiempo o para convertir la secuencia resultado de bloques de valores espectrales en una representación en el dominio del tiempo que comprende una secuencia de salida de bloques de valores de muestreo asociados con la velocidad de muestreo de salida que es diferente de la velocidad de muestreo de entrada; y

60

codificar en núcleo (1040) la secuencia de salida de bloques de valores de muestreo para obtener una señal multicanal codificada (1510).

25. Aparato para decodificación de una señal de audio multicanal codificada, que comprende:

un decodificador de núcleo (1600) para generar una señal decodificada de núcleo; estando el aparato **caracterizado**

65 **porque:**

un convertidor de tiempo-espectro (1610) para convertir una secuencia de bloques de valores de muestreo de la señal decodificada de núcleo en una representación en el dominio de las frecuencias que tiene una secuencia de bloques de valores espectrales para la señal decodificada de núcleo, en el que un bloque de valores de muestreo
 5 tiene una velocidad de muestreo de entrada asociada, y en el que un bloque de valores espectrales tiene valores espectrales hasta una frecuencia de entrada máxima que está relacionada con la velocidad de muestreo de entrada; un remuestreador en el dominio espectral (1620) para el muestreo repetido de los bloques de valores espectrales de la secuencia (1621) de bloques de valores espectrales para la señal decodificada de núcleo o al menos dos secuencias resultado (1635) obtenidas mediante el procesamiento multicanal inverso en el dominio de las
 10 frecuencias para obtener una secuencia remuestreada (1631) o al menos dos secuencias remuestreadas (1625) de bloques de valores espectrales, en el que un bloque de una secuencia remuestreada tiene valores espectrales hasta una frecuencia de salida máxima que es diferente de la frecuencia de entrada máxima; un procesador multicanal (1630) para aplicar un procesamiento multicanal inverso a una secuencia (1615) que comprende la secuencia de bloques o la secuencia remuestreada (1621) de bloques para obtener al menos dos
 15 secuencias resultado (1631, 1632, 1635) de bloques de valores espectrales; y un convertidor espectral-tiempo (1640) para convertir las al menos dos secuencias resultado (1631, 1632) de bloques de valores espectrales o las al menos dos secuencias remuestreadas (1625) de bloques de valores espectrales en una representación en el dominio del tiempo que comprende al menos dos secuencias de salida de bloques de valores de muestreo asociados con una velocidad de muestreo de salida que es diferente de la velocidad
 20 de muestreo de entrada.

26. Aparato de la reivindicación 25,

en el que el remuestreador en el dominio espectral (1020) está configurado para truncar los bloques con fines de
 25 muestreo descendente o para el padding cero de los bloques con fines de muestreo ascendente.

27. Aparato de la reivindicación 25 o 26,

en el que el remuestreador en el dominio espectral (1020) está configurado para poner en escala (1322) los valores
 30 espectrales de los bloques de la secuencia resultado de bloques mediante el uso de un factor de puesta en escala que depende de la frecuencia de entrada máxima y que depende de la frecuencia de salida máxima.

28. Aparato de una de las reivindicaciones 25 a 27,

35 en el que el factor de puesta en escala es mayor que uno en el caso del muestreo ascendente, en el que la velocidad de muestreo de salida es mayor que la velocidad de muestreo de entrada, o en el que el factor de puesta en escala es inferior a uno en el caso del muestreo descendente, en el que la velocidad de muestreo de salida es inferior a la velocidad de muestreo de entrada, o

40 en el que el convertidor tiempo-espectral (1000) está configurado para llevar a cabo un algoritmo de transformada de tiempo-frecuencia sin utilizar una normalización en cuanto al número total de valores espectrales de un bloque de valores espectrales (1311), y en el que el factor de puesta en escala es igual a un cociente entre el número de valores espectrales de un bloque de la secuencia remuestreada y el número de valores espectrales de un bloque de valores espectrales antes del muestreo repetido, y en el que el convertidor espectral-tiempo está configurado para
 45 aplicar una normalización basada en la frecuencia de salida máxima (1331).

29. Aparato de una de las reivindicaciones 25 a 28,

en el que el convertidor tiempo-espectral (1000) está configurado para llevar a cabo un algoritmo de transformada de
 50 Fourier discreta, o en el que el convertidor espectral-tiempo (1030) está configurado para llevar a cabo un algoritmo inverso de transformada de Fourier discreta.

30. Aparato de una de las reivindicaciones 25 a 29,

55 en el que el decodificador de núcleo (1600) está configurado para generar una señal decodificada de núcleo adicional (1601) que tiene una velocidad de muestreo adicional que es diferente de la velocidad de muestreo de entrada,

en el que el convertidor tiempo-espectral (1610) está configurado para convertir la otra señal decodificada de núcleo
 60 en una representación en el dominio de las frecuencias que tiene otra secuencia (1611) de bloques de valores para la otra señal decodificada de núcleo, en el que un bloque de valores de muestreo de la señal decodificada de núcleo adicional tiene valores espectrales de hasta una frecuencia de entrada máxima adicional que es diferente de la frecuencia de entrada máxima y relacionado con la otra velocidad de muestreo,

65 en el que el remuestreador en el dominio espectral (1620) está configurado para remuestrear la otra secuencia de

bloques para la señal decodificada de núcleo adicional en el dominio de las frecuencias para obtener una secuencia remuestreada adicional (1621) de bloques de valores espectrales, en el que un bloque de valores espectrales de la secuencia remuestreada adicional tiene valores espectrales hasta la frecuencia de salida máxima que es diferente de la frecuencia de entrada máxima adicional; y

5 un combinador (1700) para combinar la secuencia remuestreada y la secuencia remuestreada adicional para obtener la secuencia (1701) que va a ser procesada por el procesador multicanal (1630).

31. Aparato de una de las reivindicaciones 25 a 30,

10 en el que el decodificador de núcleo (1600) está configurado para generar una señal incluso más decodificada de núcleo que tiene una velocidad de muestreo adicional que es igual a la velocidad de muestreo de salida (1603),

15 en el que el convertidor de tiempo-espectro (1610) está configurado para convertir la secuencia incluso más en una representación en el dominio de las frecuencias (1613),

20 en el que el aparato comprende además un combinador (1700) para combinar la secuencia incluso más de bloques de valores espectrales y la secuencia remuestreada (1622, 1621) de bloques en un proceso que consiste en generar la secuencia de bloques procesada por el procesador multicanal (1630).

32. Aparato de una de las reivindicaciones 25 a 31,

25 en el que el decodificador de núcleo (1600) comprende al menos una porción de decodificación basada en MDCT (1600d), una porción de decodificación de extensión del ancho de banda en el dominio del tiempo (1600c), una porción de decodificación ACELP (1600b) y una porción de decodificación de post-filtro baja (1600a).

30 en el que la porción de decodificación basada en MDCT (1600d) o la porción de decodificación de extensión del ancho de banda en el dominio del tiempo (1600c) está configurada para generar la señal decodificada de núcleo que tiene la velocidad de muestreo de salida, o

en el que la porción de decodificación ACELP (1600b) o la porción de decodificación de post-filtro baja (1600a) está configurada para generar una señal decodificada de núcleo con una velocidad de muestreo que es diferente de la velocidad de muestreo de salida.

33. Aparato de una de las reivindicaciones 25 a 32,

35 en el que el convertidor de tiempo-espectro (1610) está configurado para aplicar una ventana de análisis al menos a dos de entre una pluralidad de diferentes señales decodificadas de núcleo, teniendo las ventanas de análisis el mismo tamaño en tiempo o teniendo la misma forma con respecto al tiempo,

40 en el que el aparato comprende además un combinador (1700) para combinar al menos una secuencia remuestreada y cualquier otra secuencia que tenga bloques con valores espectrales hasta la frecuencia de salida máxima en una base de bloque por bloque para obtener la secuencia procesada por el procesador multicanal (1630).

34. Aparato de una de las reivindicaciones 25 a 33,

45 en el que la secuencia procesada por el procesador multicanal (1630) corresponde a una señal media (Mid), y

50 en el que el procesador multicanal (1630) está configurado para generar adicionalmente una señal lateral (Side) mediante el uso de información en una señal lateral (Side) incluida en la señal de audio multicanal codificada, y

55 en el que el procesador multicanal (1630) está configurado para generar las al menos dos secuencias resultado mediante el uso de la señal media (Mid) y la señal lateral (Side).

35. Aparato de una de las reivindicaciones 25 a 34,

60 en el que el procesador multicanal (1630) está configurado para convertir (820) la secuencia en una primera secuencia para un primer canal de salida y una segunda secuencia para un segundo canal de salida mediante el uso de un factor de ganancia por banda de parámetros;

65 para actualizar (830) una primera secuencia y la segunda secuencia mediante el uso de una señal lateral (Side) decodificada o para actualizar la primera secuencia y la segunda secuencia mediante el uso de una señal lateral (Side) predicha a partir de un bloque anterior de la secuencia de bloques para la señal media (Mid) mediante el uso de un parámetro de llenado de estéreo para una banda de parámetros;

para llevar a cabo (910) una desalineación de fase y una puesta en escala de energía mediante el uso de información sobre la pluralidad de parámetros de alineación de fase de banda estrecha; y

5 para llevar a cabo (920) una desalineación en tiempo mediante el uso de información sobre un parámetro de alineación en tiempo de banda ancha para obtener las al menos dos secuencias resultado.

36. Aparato de una de las reivindicaciones 25 a 35,

10 en el que el decodificador de núcleo (1600) está configurado para operar según un primer control de trama para proporcionar una secuencia de tramas, en el que una trama está delimitada por un borde de inicio de la trama (1901) y por un borde final de la trama (1902),

15 en el que el convertidor tiempo-espectral (1610) o el convertidor espectral-tiempo (1640) está configurado para operar según un segundo control de trama que está sincronizado con el segundo control de trama,

20 en el que el convertidor tiempo-espectral (1610) o el convertidor espectral-tiempo (1640) están configurados para operar según un segundo control de trama que está sincronizado con el segundo control de trama, en el que el borde de inicio de la trama (1901) o el borde final de la trama (1902) de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con respecto a un instante de inicio o con respecto a una instante final de una porción superpuesta de una ventana utilizada por el convertidor tiempo-espectral (1610) para cada bloque de la secuencia de bloques de valores de muestreo o utilizado por el convertidor espectral-tiempo (1640) para cada bloque de las al menos dos secuencias de salida de bloques de valores de muestreo.

25 37. Aparato de una de las reivindicaciones 25 a 36,

en el que la señal decodificada de núcleo tiene la secuencia de tramas, una trama que tiene el borde de inicio de la trama (1901) y el borde final de la trama (1902),

30 en el que una ventana de análisis (1914) utilizada por el convertidor de tiempo-espectro (1610) para poner en ventana la trama de secuencia de tramas tiene una porción superpuesta que termina antes del borde final de la trama (1902) dejando un espacio de tiempo (1920) entre un extremo de la porción superpuesta y el borde final de la trama (1902), y

35 en el que el decodificador de núcleo (1600) está configurado para llevar a cabo un procesamiento de las muestras en el espacio de tiempo (1920) en paralelo con la puesta en ventana de la trama mediante el uso de la ventana de análisis (1914), o en el que un post-procesamiento con decodificador de núcleo se lleva a cabo sobre las muestras en el espacio de tiempo (1920) en paralelo con la puesta en ventana de la trama mediante el uso de la ventana de análisis.

40

38. Aparato de una de las reivindicaciones 25 a 37,

en el que la señal decodificada de núcleo tiene la secuencia de tramas, una trama que tiene el borde de inicio de la trama (1901) y el borde final de la trama (1902),

45

en el que un inicio de una primera porción superpuesta de una ventana de análisis (1914) coincide con el borde de inicio de la trama (1901), y en el que un extremo de una segunda porción superpuesta de la ventana de análisis (1914) está situada antes del borde de parada de la trama (1902), de tal manera que existe un espacio de tiempo (1920) entre el extremo de la segunda porción superpuesta y el borde de parada de la trama, y

50

en el que la ventana de análisis para un bloque siguiente de la señal decodificada de núcleo está situada de tal manera que una porción central no superpuesta de la ventana de análisis está situada dentro del espacio de tiempo (1920).

55 39. Aparato de una de las reivindicaciones 25 a 38,

en el que la ventana de análisis utilizada por el convertidor de tiempo-espectro (1610) tiene la misma forma y longitud en tiempo que la ventana de síntesis utilizada por el convertidor espectral-tiempo (1640).

60 40. Aparato de una de las reivindicaciones 25 a 39,

en el que la señal decodificada de núcleo tiene una secuencia de tramas, en el que una trama tiene una longitud, en el que la longitud de la ventana que excluye cualquier porción cero aplicada por el convertidor tiempo-espectral (1610) es más pequeña que o igual a la mitad de la longitud de la trama.

65

41. Aparato de una de las reivindicaciones 25 a 40,
en el que el convertidor espectral-tiempo (1640) está configurado
- 5 para aplicar una ventana de síntesis para obtener un primer bloque de salida de muestras puestas en ventana para una primera secuencia de salida de las al menos dos secuencias de salida;
- para aplicar la ventana de síntesis para obtener un segundo bloque de salida de muestras puestas en ventana para la primera secuencia de salida de las al menos dos secuencias de salida;
- 10 para superponer-añadir el primer bloque de salida y el segundo bloque de salida para obtener un primer grupo de muestras de salida para la primera secuencia de salida;
- en el que el convertidor espectral-tiempo (1640) está configurado
- 15 para aplicar una ventana de síntesis para obtener un primer bloque de salida de muestras puestas en ventana para una segunda secuencia de salida de las al menos dos secuencias de salida;
- para aplicar la ventana de síntesis para obtener un segundo bloque de salida de muestras puestas en ventana para
- 20 la segunda secuencia de salida de las al menos dos secuencias de salida;
- para superponer-añadir el primer bloque de salida y el segundo bloque de salida para obtener un segundo grupo de muestras de salida para la segunda secuencia de salida;
- 25 en el que el primer grupo de muestras de salida para la primera secuencia y el segundo grupo de muestras de salida para la segunda secuencia están relacionadas con la misma porción de tiempo de la señal multicanal decodificada o están relacionadas con la misma trama de la señal decodificada de núcleo.
42. Procedimiento para decodificación de una señal de audio multicanal codificada, que comprende:
- 30 generar (1600) una señal decodificada de núcleo; estando el procedimiento **caracterizado porque** comprende además:
- convertir (1610) una secuencia de bloques de valores de muestreo de la señal decodificada de núcleo en una
- 35 representación en el dominio de las frecuencias que tiene una secuencia de bloques de valores espectrales para la señal decodificada de núcleo, en el que un bloque de valores de muestreo tiene una velocidad de muestreo de entrada asociada, y en el que un bloque de valores espectrales tiene valores espectrales hasta una frecuencia de entrada máxima que está relacionada con la velocidad de muestreo de entrada;
- 40 muestrear repetidamente (1620) los bloques de valores espectrales de la secuencia (1621) de bloques de valores espectrales para la señal decodificada de núcleo o al menos dos secuencias resultado (1635) obtenidas mediante procesamiento multicanal inverso en el dominio de las frecuencias para obtener una secuencia remuestreada (1631) o al menos dos secuencias remuestreadas (1625) de bloques de valores espectrales, en el que un bloque de una secuencia remuestreada tiene valores espectrales hasta una frecuencia de salida máxima que es diferente de la
- 45 frecuencia de entrada máxima;
- aplicar (1630) un procesamiento multicanal inverso a una secuencia (1615) que comprende la secuencia de bloques o la secuencia remuestreada (1621) de bloques para obtener al menos dos secuencias resultado (1631, 1632, 1635) de bloques de valores espectrales; y
- 50 convertir (1640) las al menos dos secuencias resultado (1631, 1632) de bloques de valores espectrales o las al menos dos secuencias remuestreadas (1625) de bloques de valores espectrales en una representación en el dominio del tiempo que comprende al menos dos secuencias de salida de bloques de valores de muestreo asociadas con una velocidad de muestreo de salida que es diferente de la velocidad de muestreo de entrada.
- 55
43. Programa informático para llevar a cabo, cuando se ejecuta en un ordenador o procesador, el procedimiento de la reivindicación 24 o el procedimiento de la reivindicación 42.

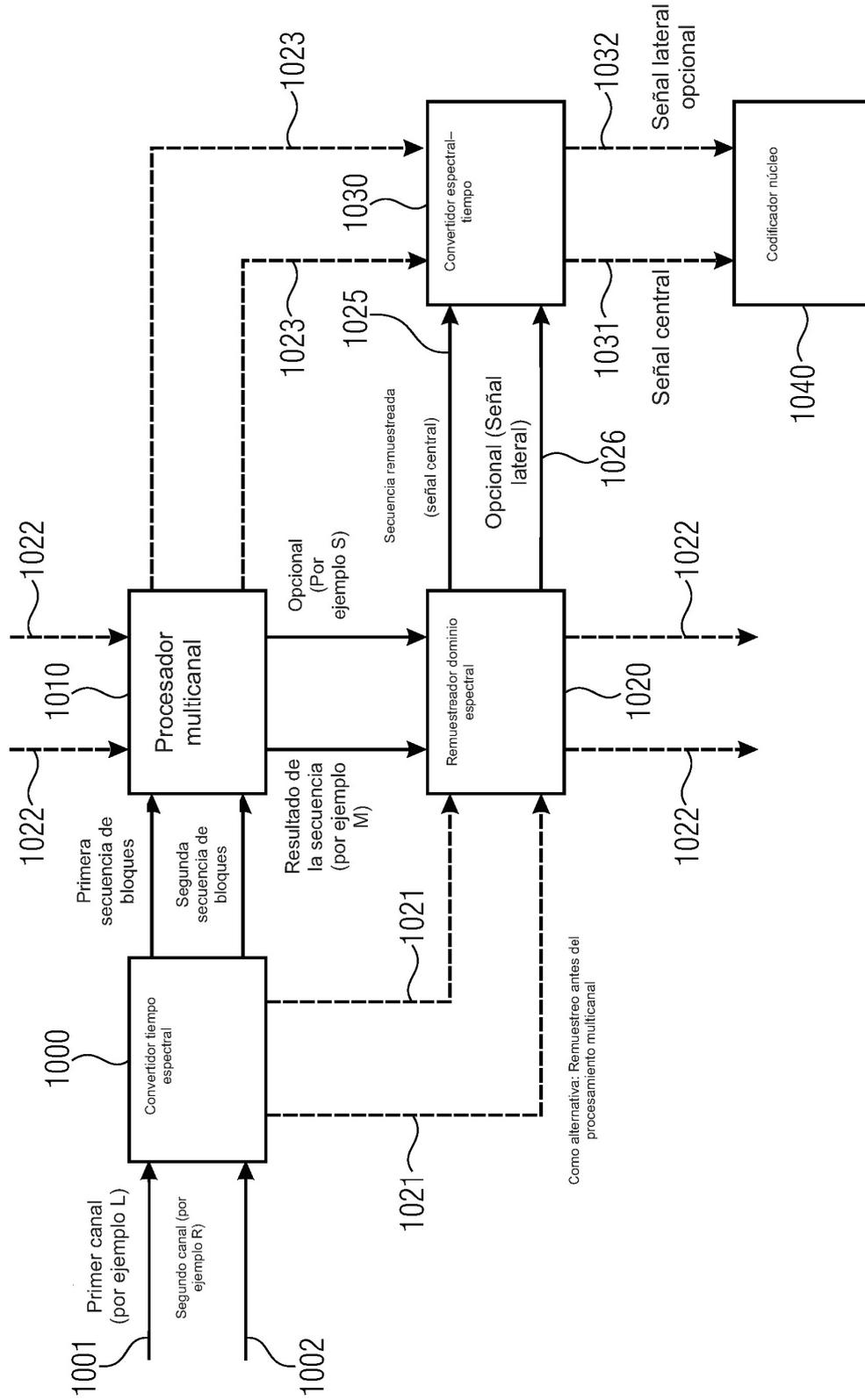
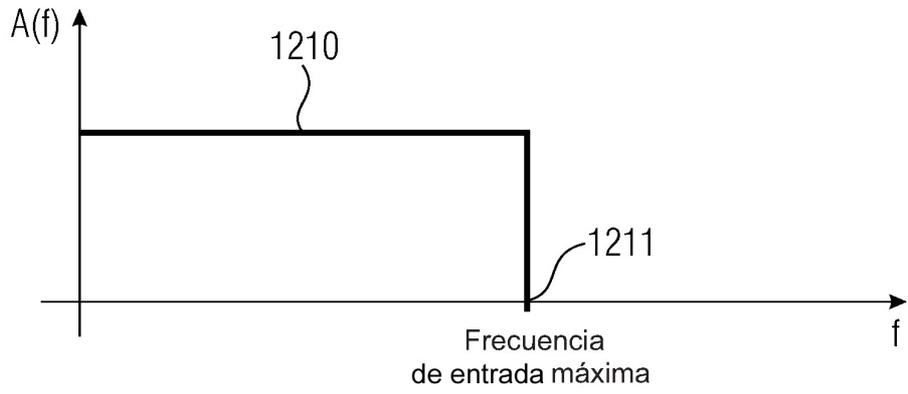
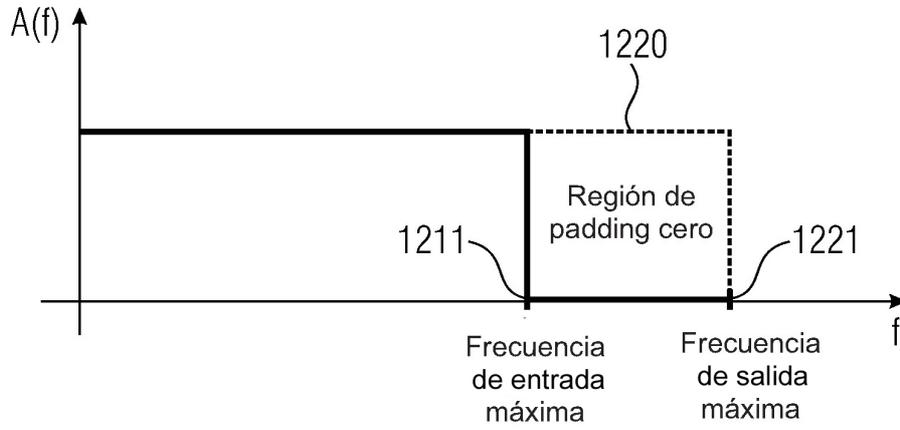


Fig. 1



Muestreo ascendente mediante padding cero



Muestreo descendente por truncación

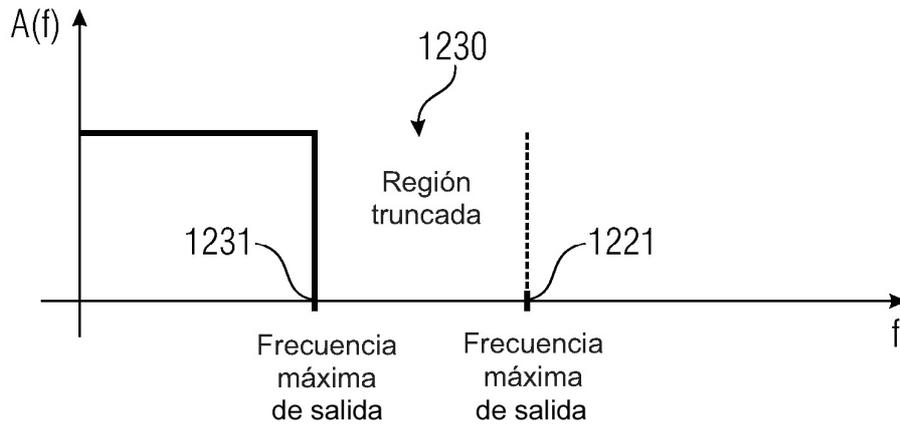


Fig. 2

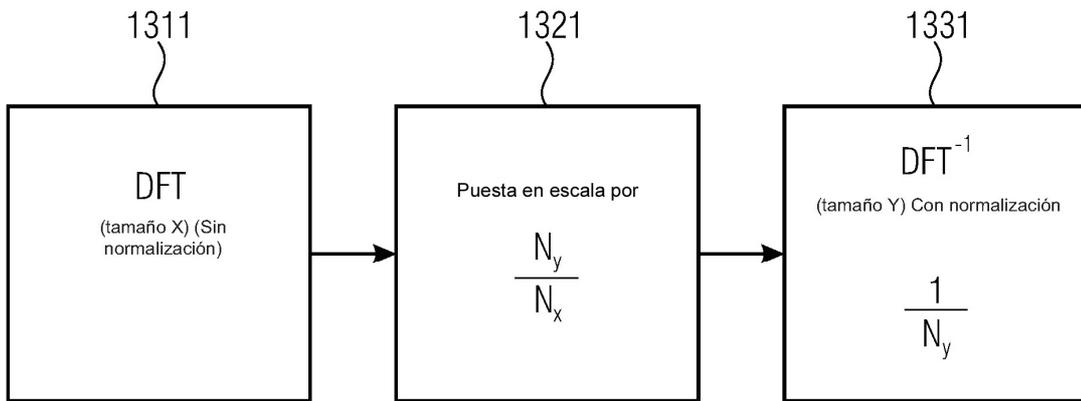


Fig. 3a

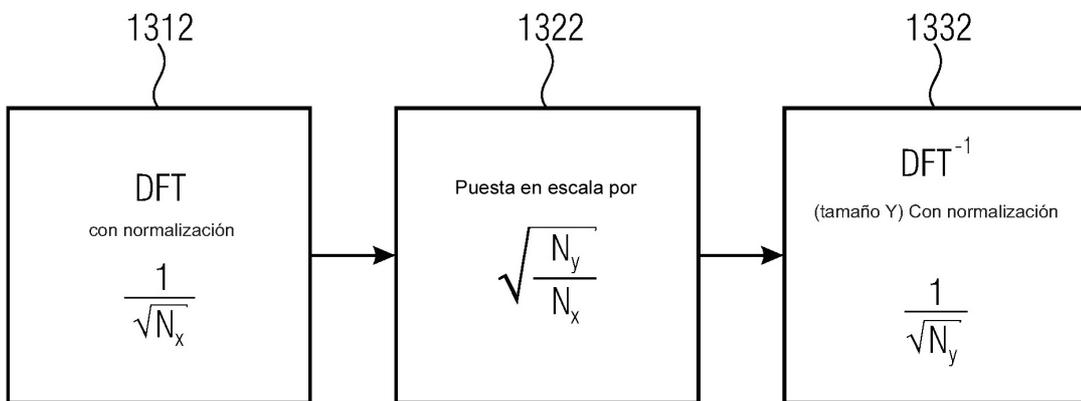


Fig. 3b

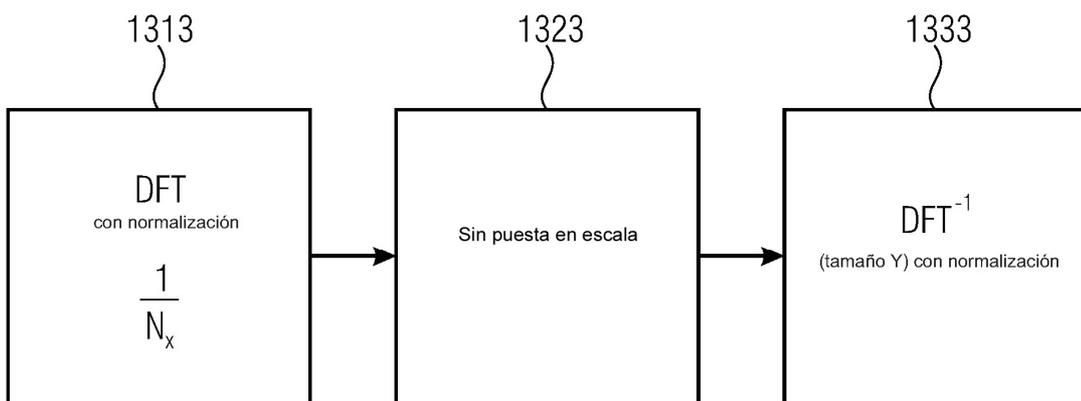
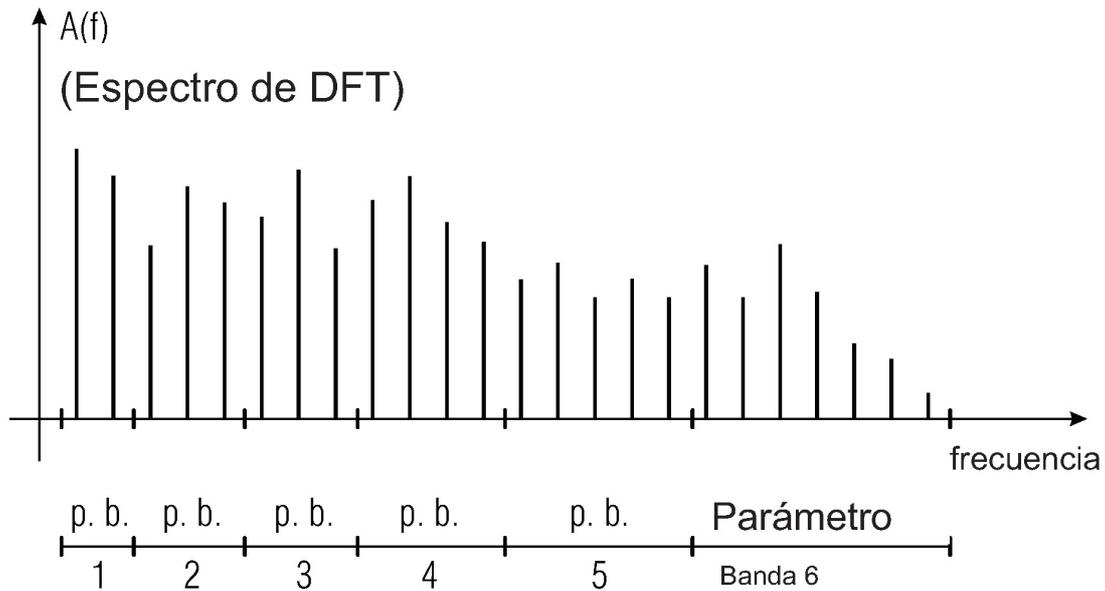


Fig. 3c



- Único parámetro de alineación de banda ancha (por ejemplo, g.p. banda 1 a banda 6);
- Pluralidad de parámetros de alineación de banda ancha para bandas de parámetro 1, 2, 3, 4, es decir, cuatro parámetros de banda estrecha);
- Parámetros de nivel para cada banda de parámetro, por ejemplo, 6 parámetros de nivel;
- Parámetros de llenado estéreo para bandas de parámetros 4, 5, 6, por ejemplo parámetros de llenado estéreo;
- Señal lateral (residual) para bandas de parámetro, 1, 2, 3;
- Más líneas espectrales en banda superior, por ejemplo, siete líneas espectrales en banda de parámetros 6 versus cinco líneas espectrales en banda de parámetro 2.

Fig. 3d

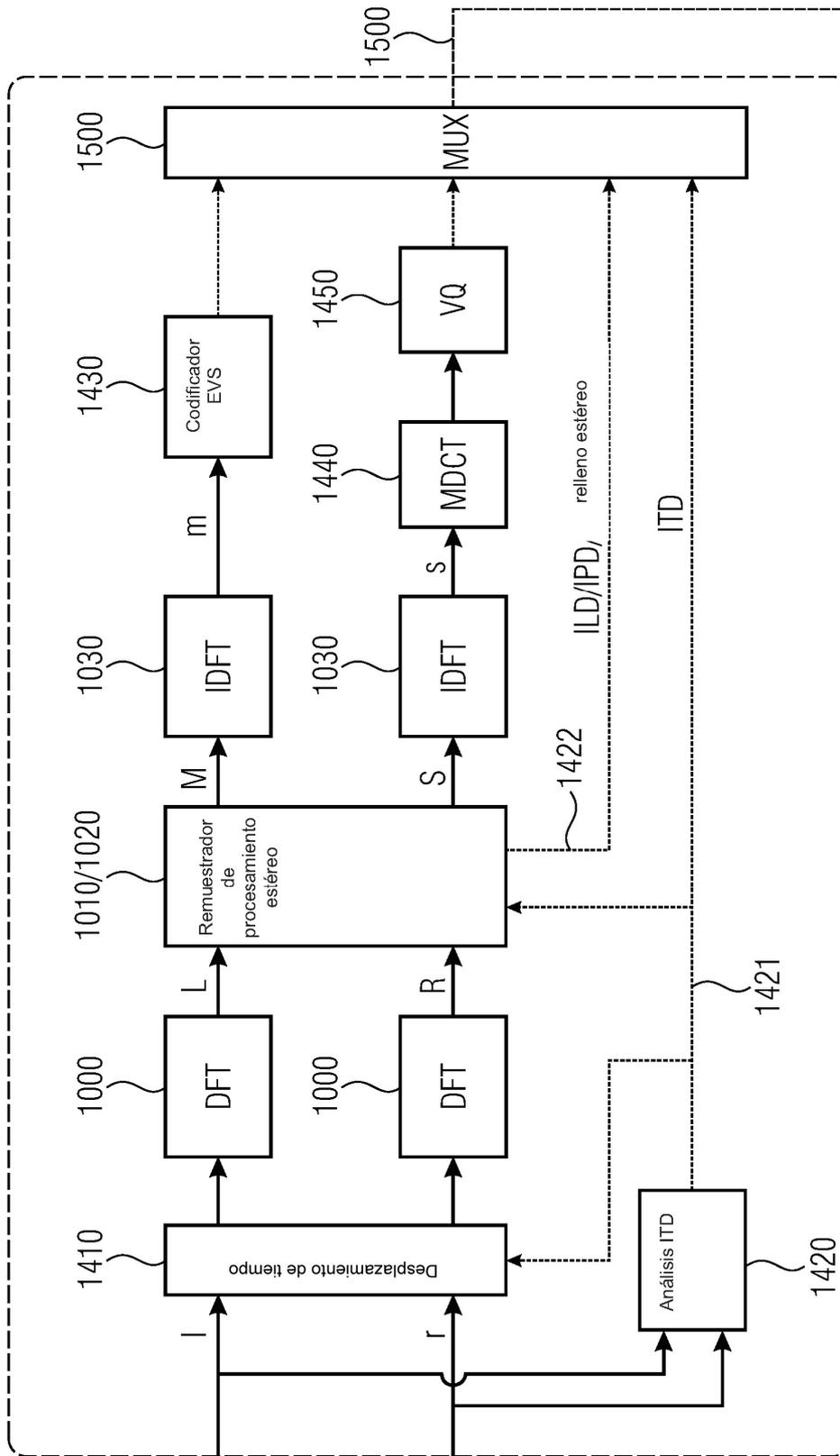


Fig. 4a

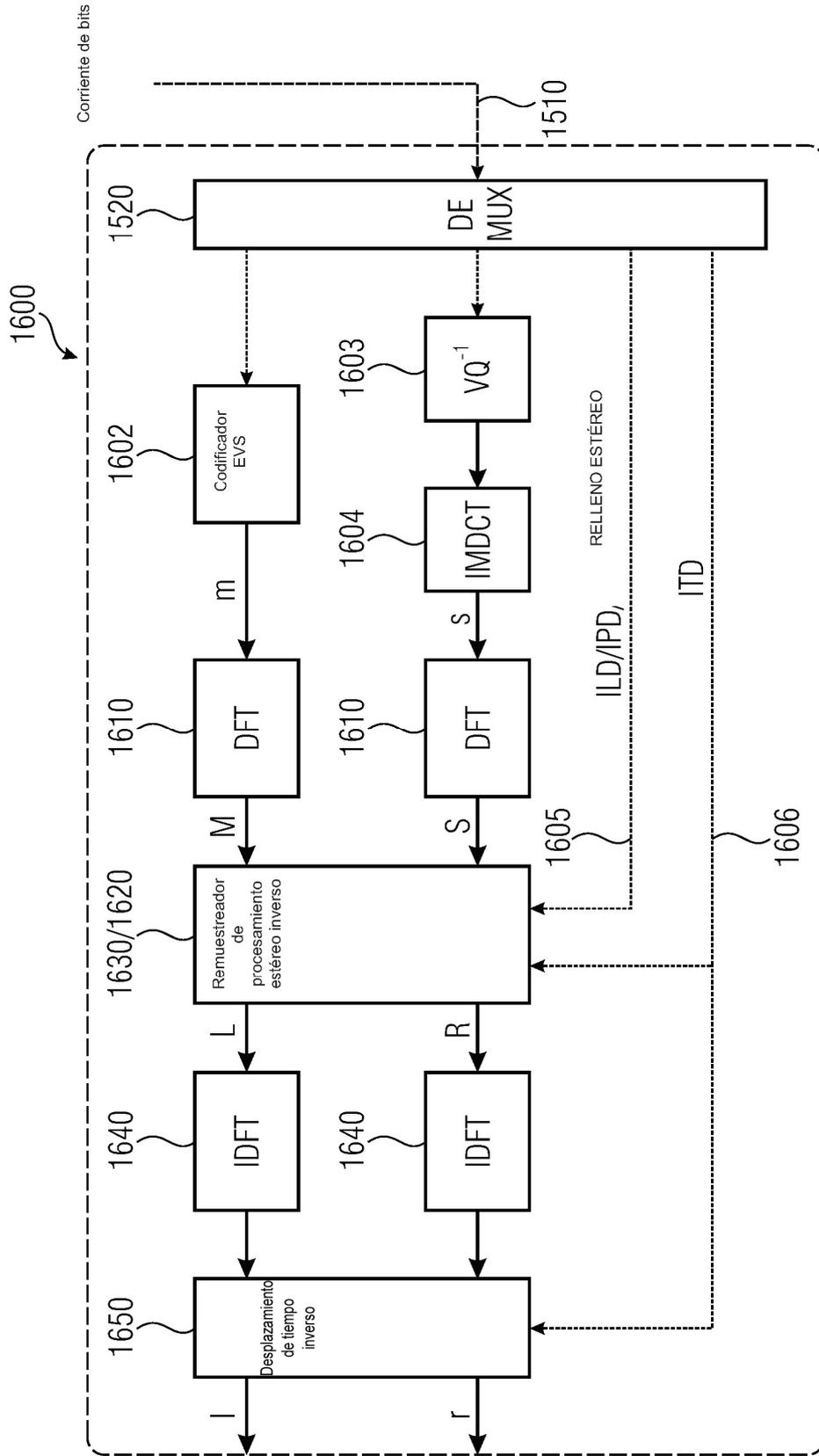


Fig. 4b

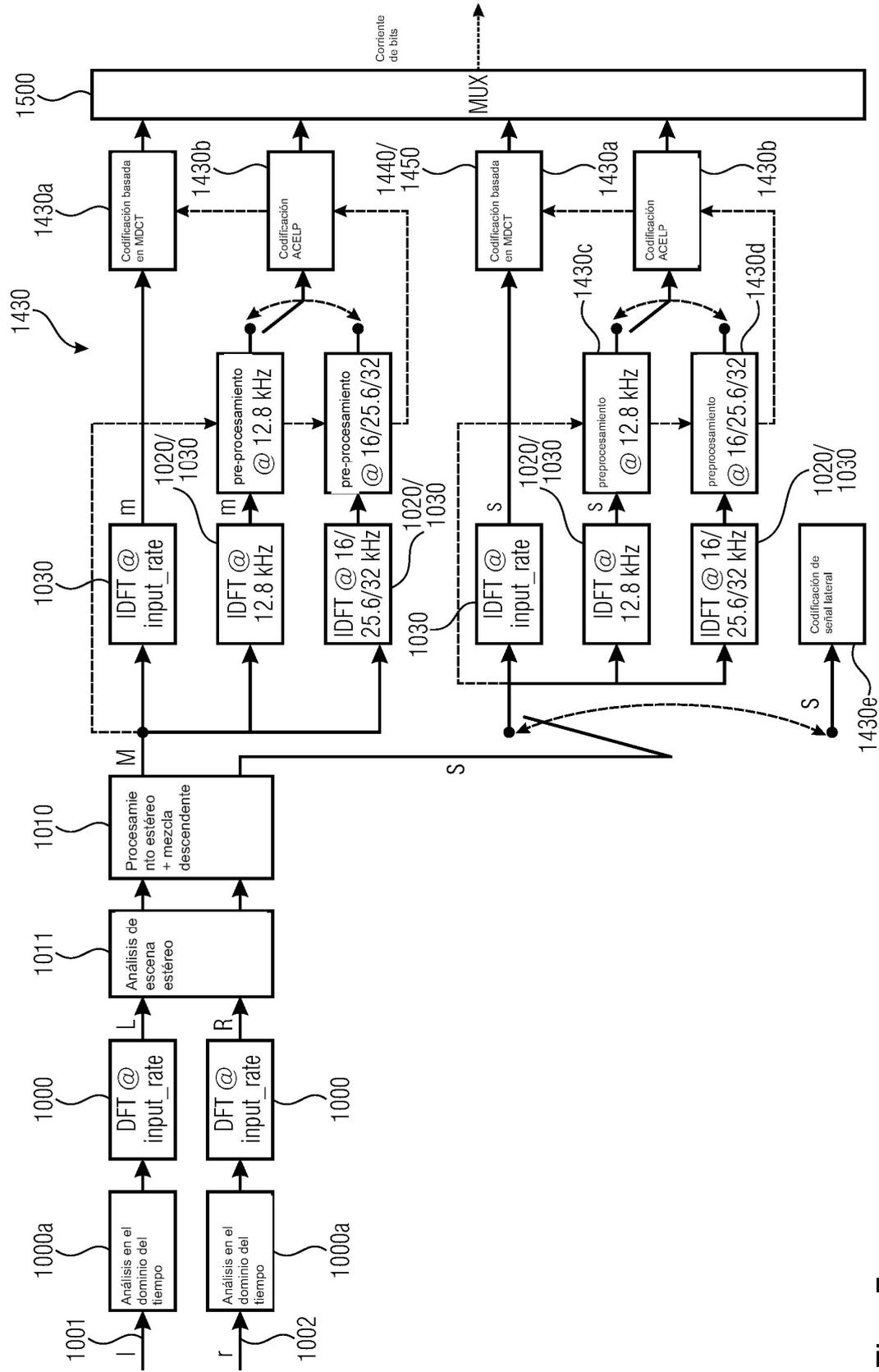


Fig. 5

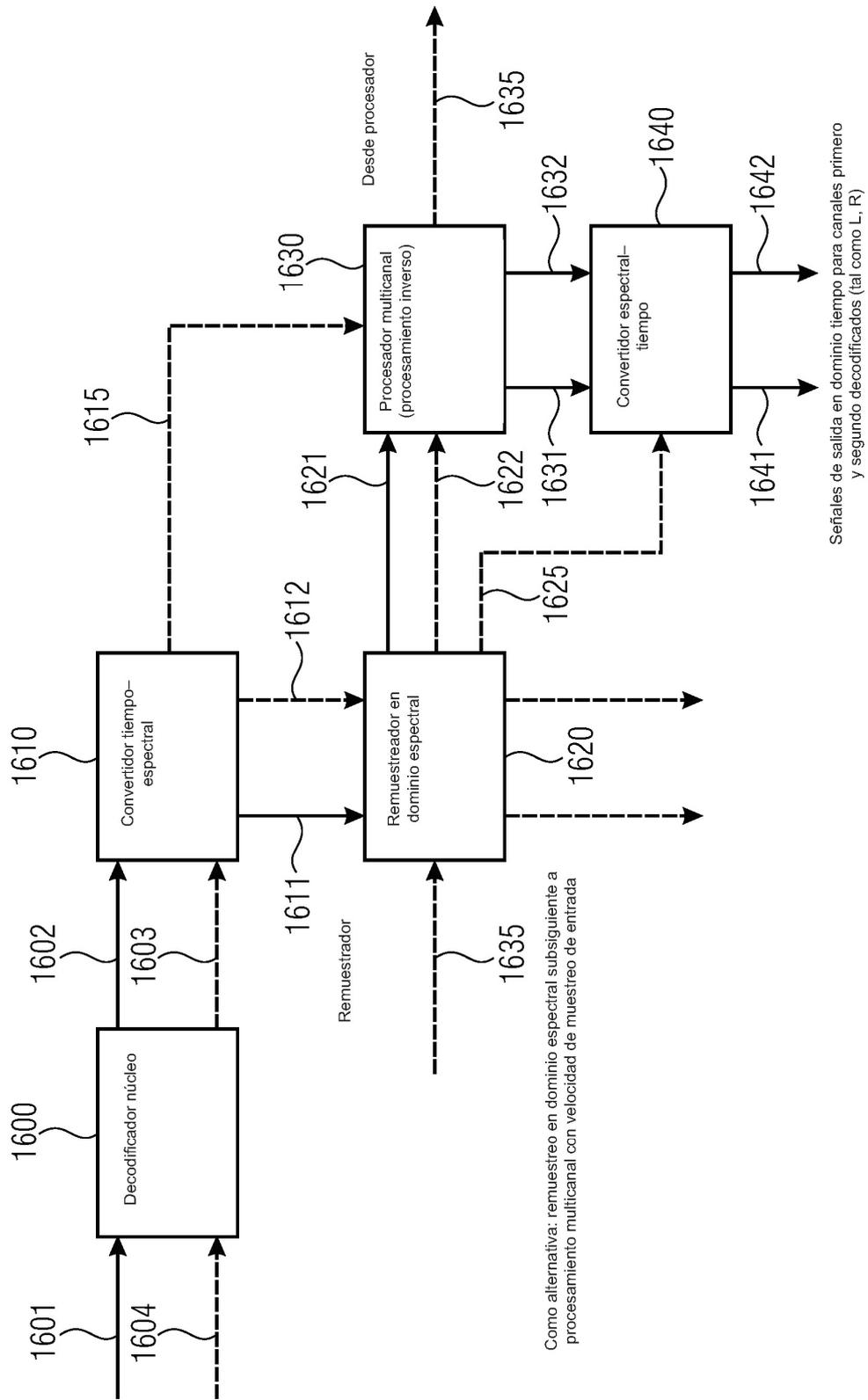


Fig. 6

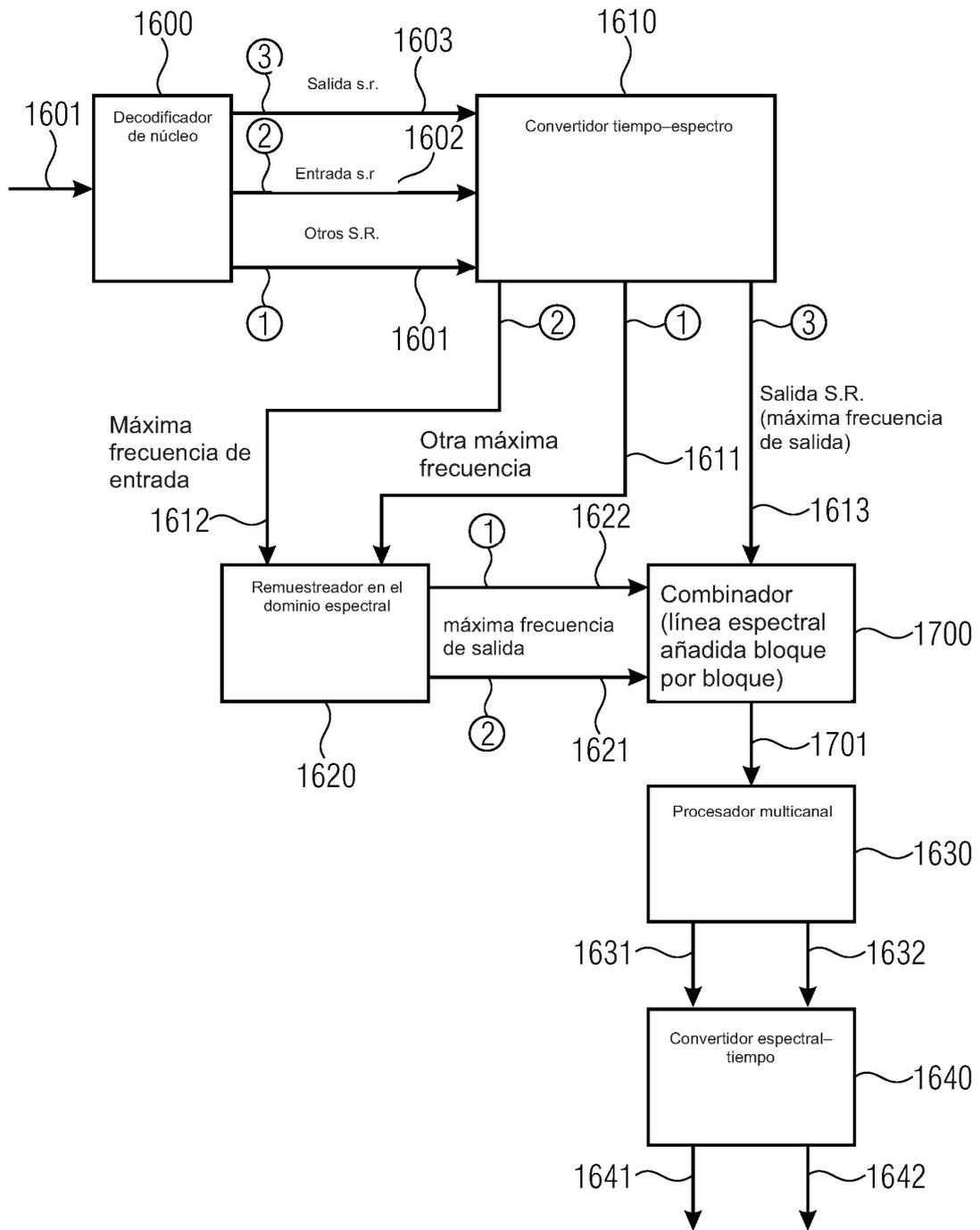


Fig. 7a

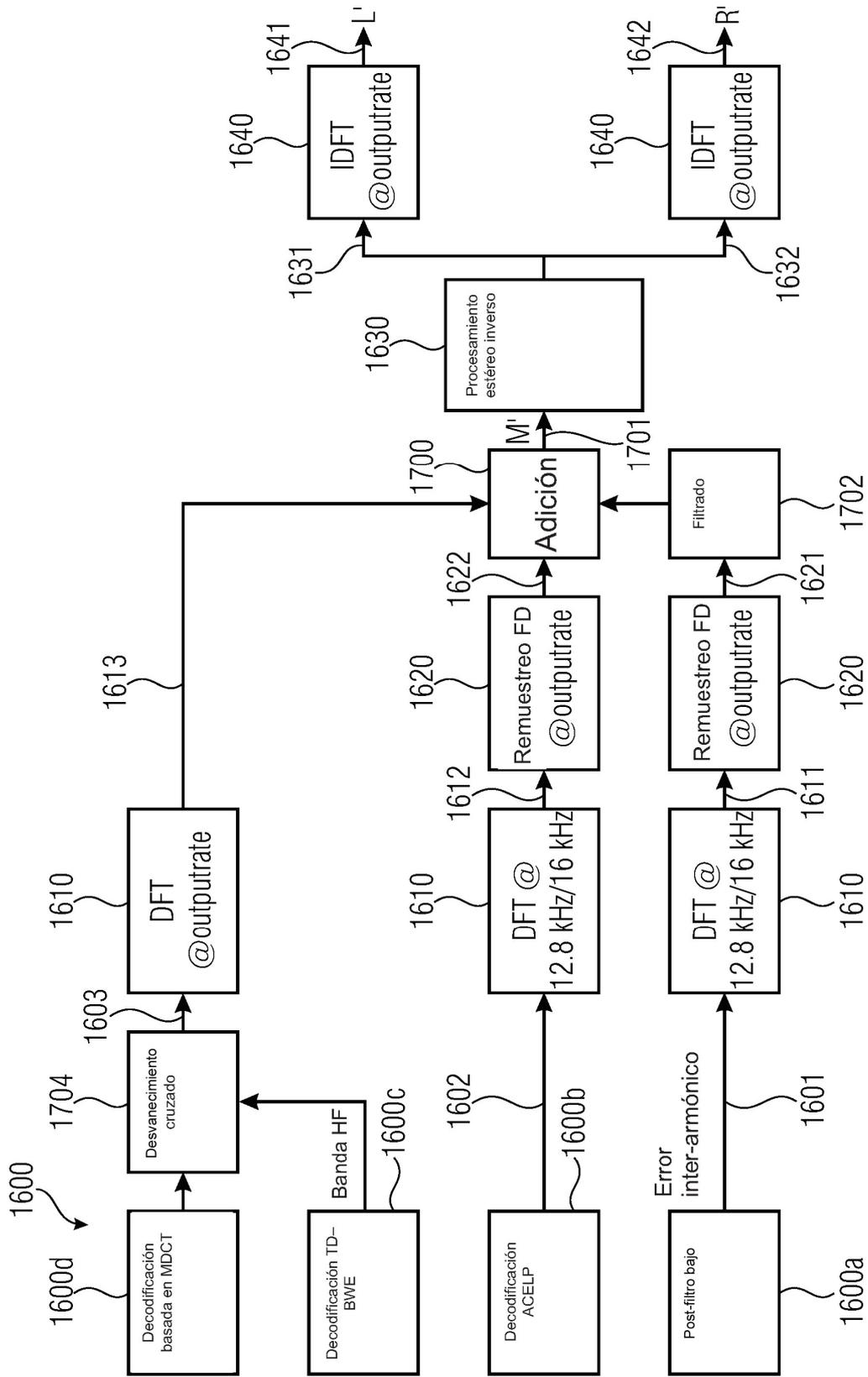


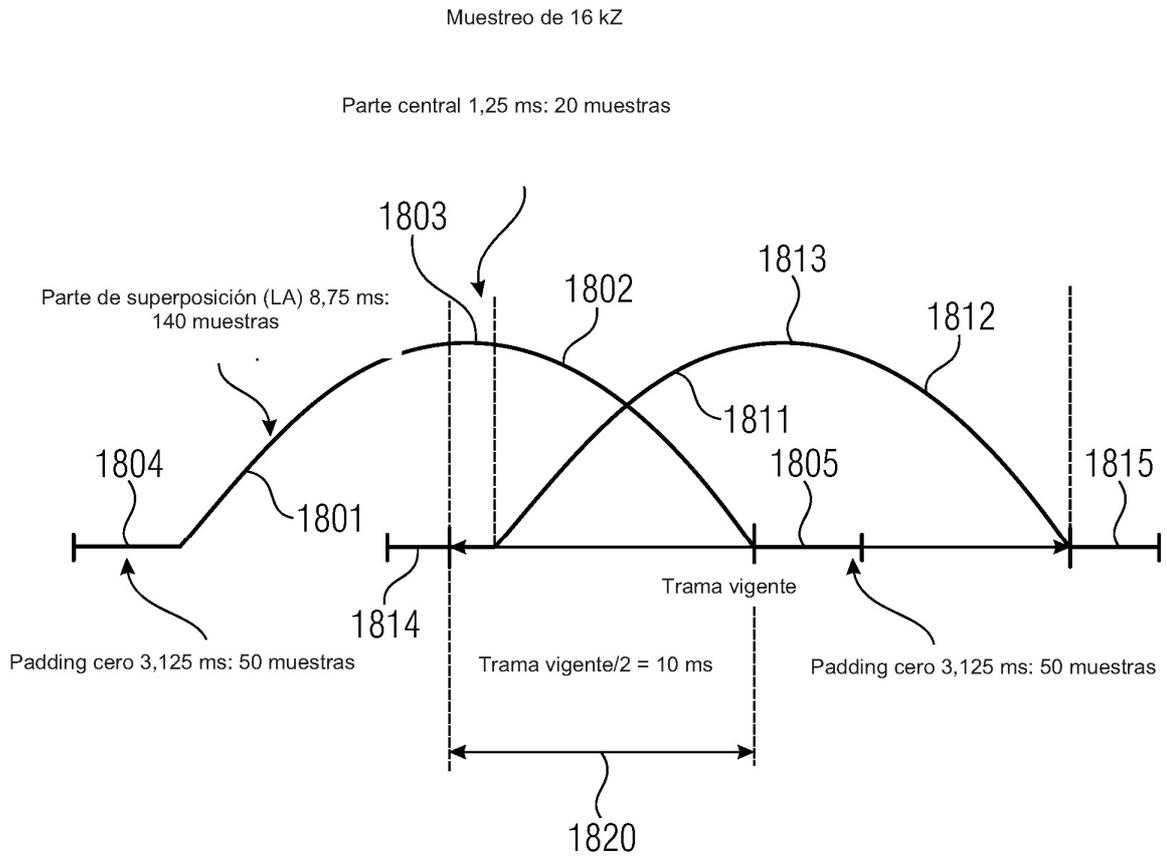
Fig. 7b

	Duración (ms)	Muestras @12.8 kHz	Muestras @16 kHz	Muestras @32 kHz	Muestras @48 kHz
Padding cero	3.125	40	50	100	150
Superposición	8.75	112	140	280	420
Parte central	1.25	16	20	40	60
Tamaño de ventana	25	320	400	800	1200
Raíz máxima de DFT		5	5	5	5

Fig. 8a

Número de propuesta	Tipo de banda de filtro/transformada de bloque	Retardo total	Retardo del codificado r	Retardo del decodificador	Retardo disponible en el lado del decodificador después del decodificador núcleo	Banda de filtro de resolución de frecuencia/transformada de bloque	Resolución de tiempo de la banda de filtro transformada de bloque
1	DFT	38.75 ms	8.75 ms	10 ms	1.25 ms	53.3 Hz	10 ms
2	DFT	38.75 ms	8.75 ms	10 ms	0 ms	53.3/34.8 Hz	10/20 ms
3	CLDFB	38.75 + x ms	13.75 ms	5 + x ms	x ms	400 Hz	1.25 ms
4	DFT	48.75 ms	17.5 ms	11.25 ms	2.5 ms	53.3 Hz	10 ms
5	DFT	40 ms	13.75 ms	6.25 ms	1.25 ms	100 Hz	5 ms

Fig. 8b



Los coeficientes ascendientes $ovlp_size$ están dados por:

$$win_ovlp(k) = \sin(\pi \cdot (k + 0.5) / (2 \cdot ovlp_size));$$

for $k = 0..ovlp_size - 1$

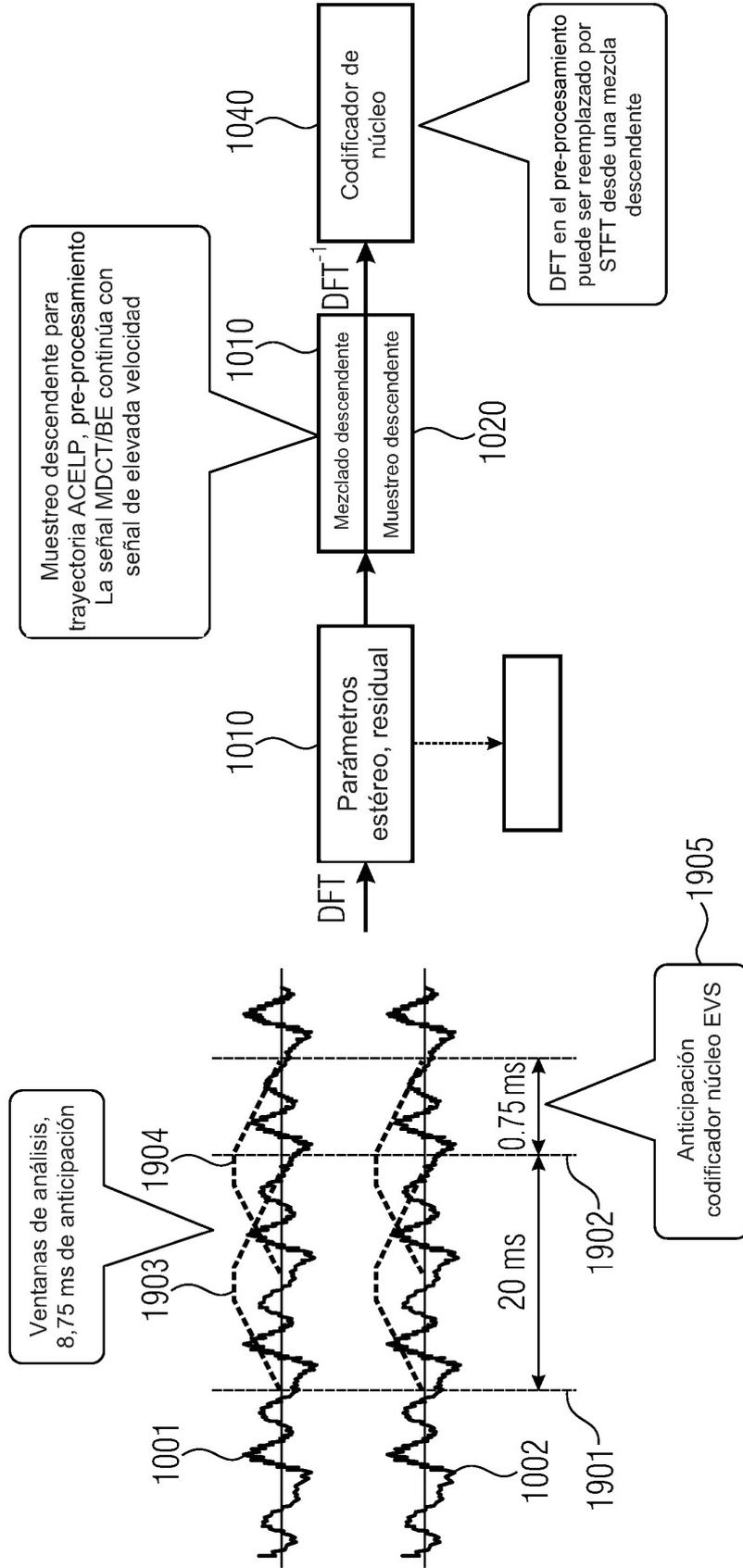
Los coeficientes descendientes $ovlp_size$ están dados por:

$$win_ovlp(k) = \sin(\pi \cdot (ovlp_size - 1 - k + 0.5) / (2 \cdot ovlp_size));$$

for $k = 0..ovlp_size - 1$

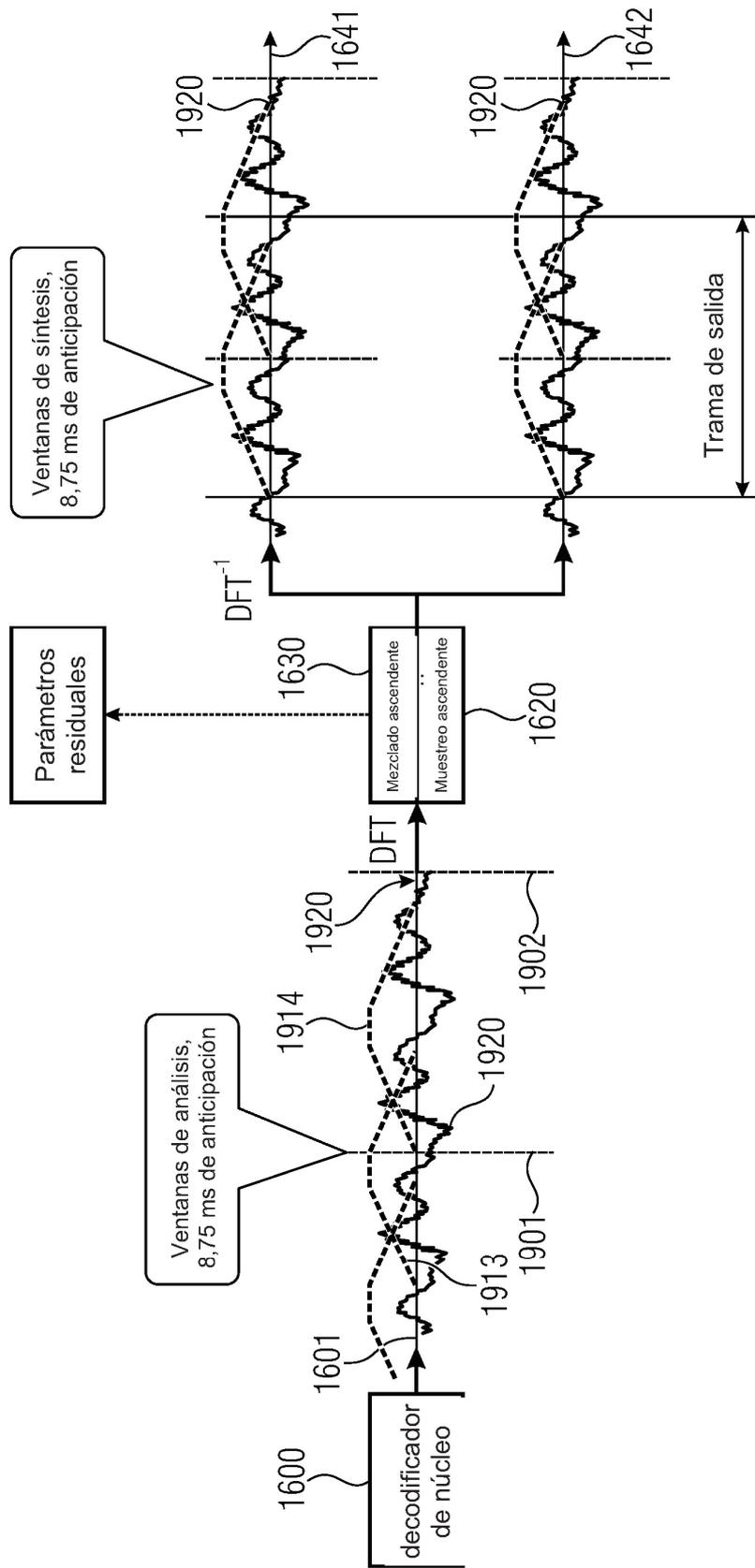
donde $ovlp_size$ es en función de la velocidad de muestreo y está dada en la Figura 8A

Fig. 8c



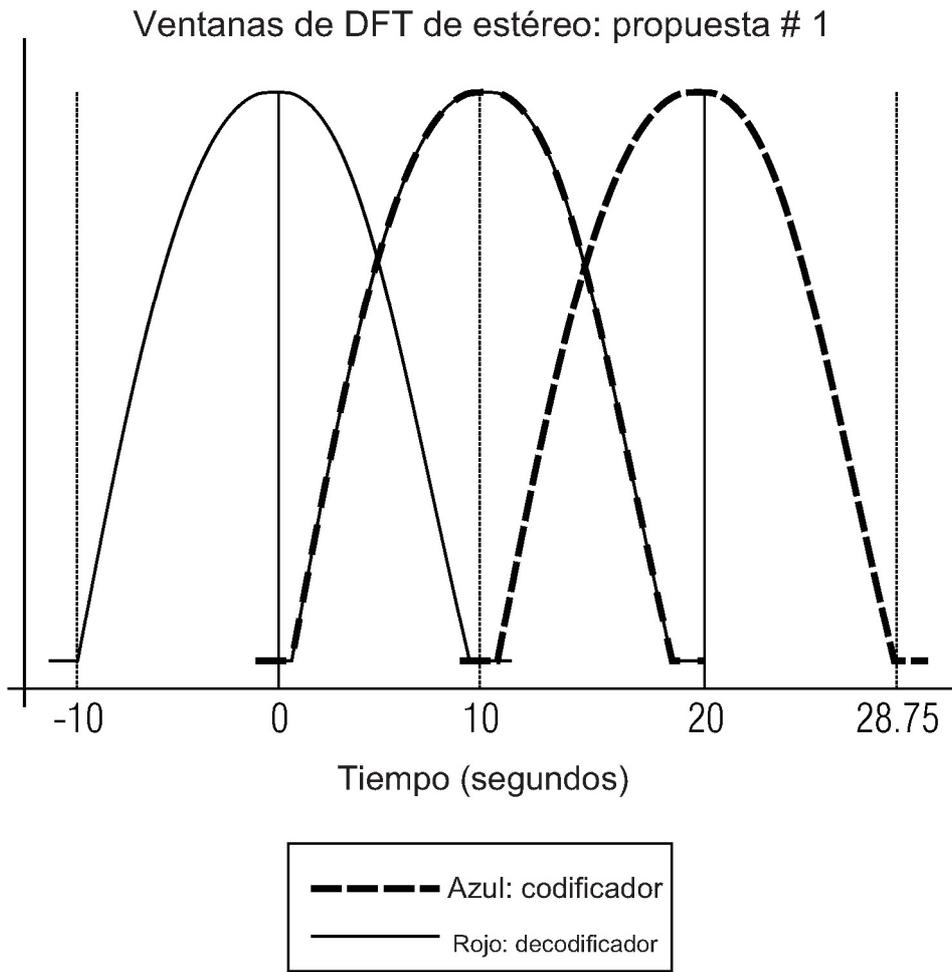
Propuesta 1. Puesta esquemática en ventana del codificador

Fig. 9a



Propuesta 1. Puesta esquemática en ventana del decodificador

Fig. 9b



Propuesta 1– Ventanas en codificador y decodificador

Fig. 9c

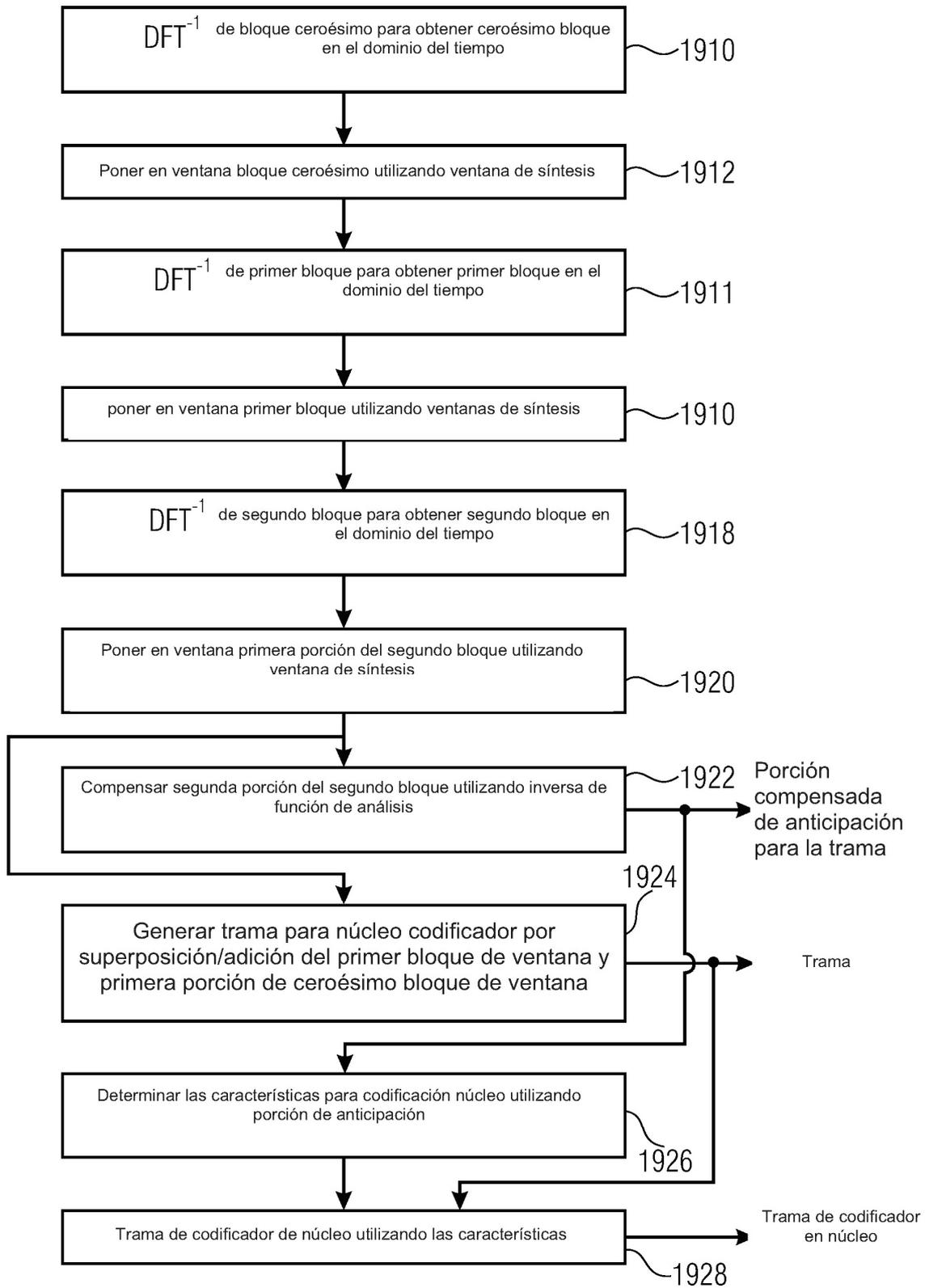


Fig. 9d

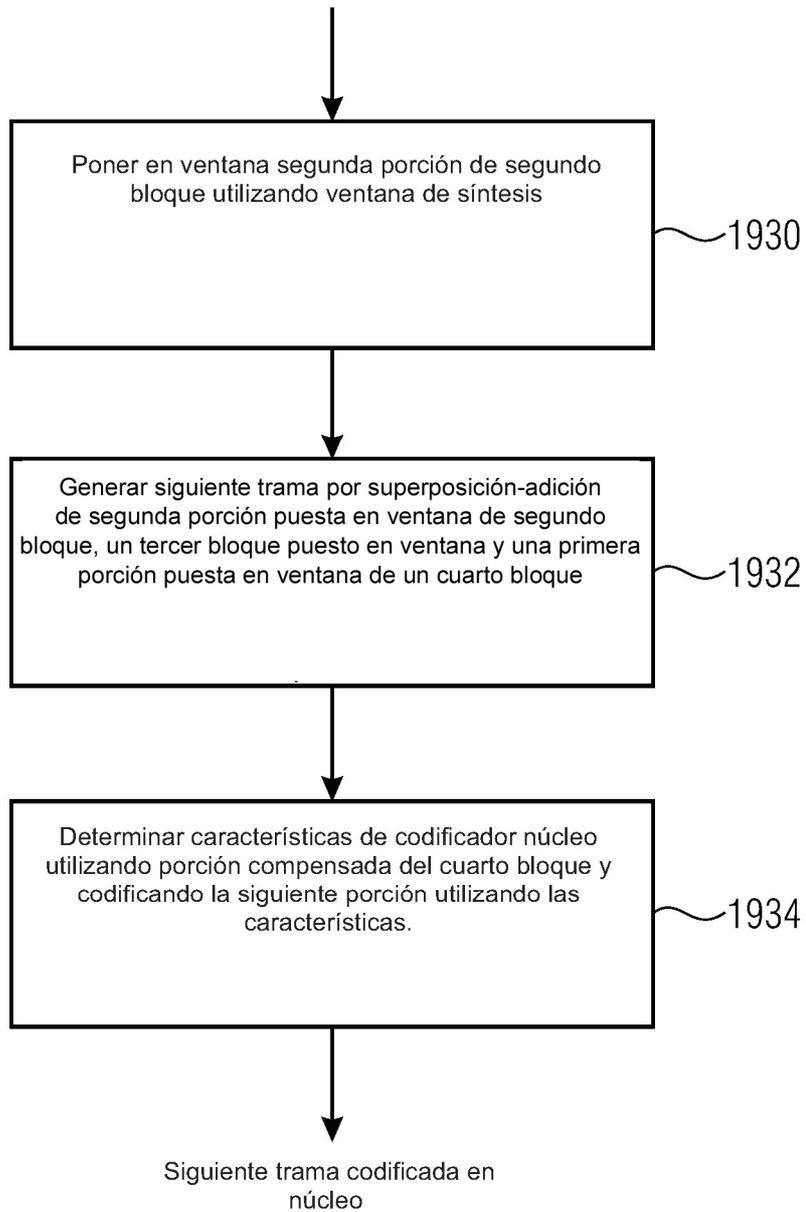


Fig. 9e

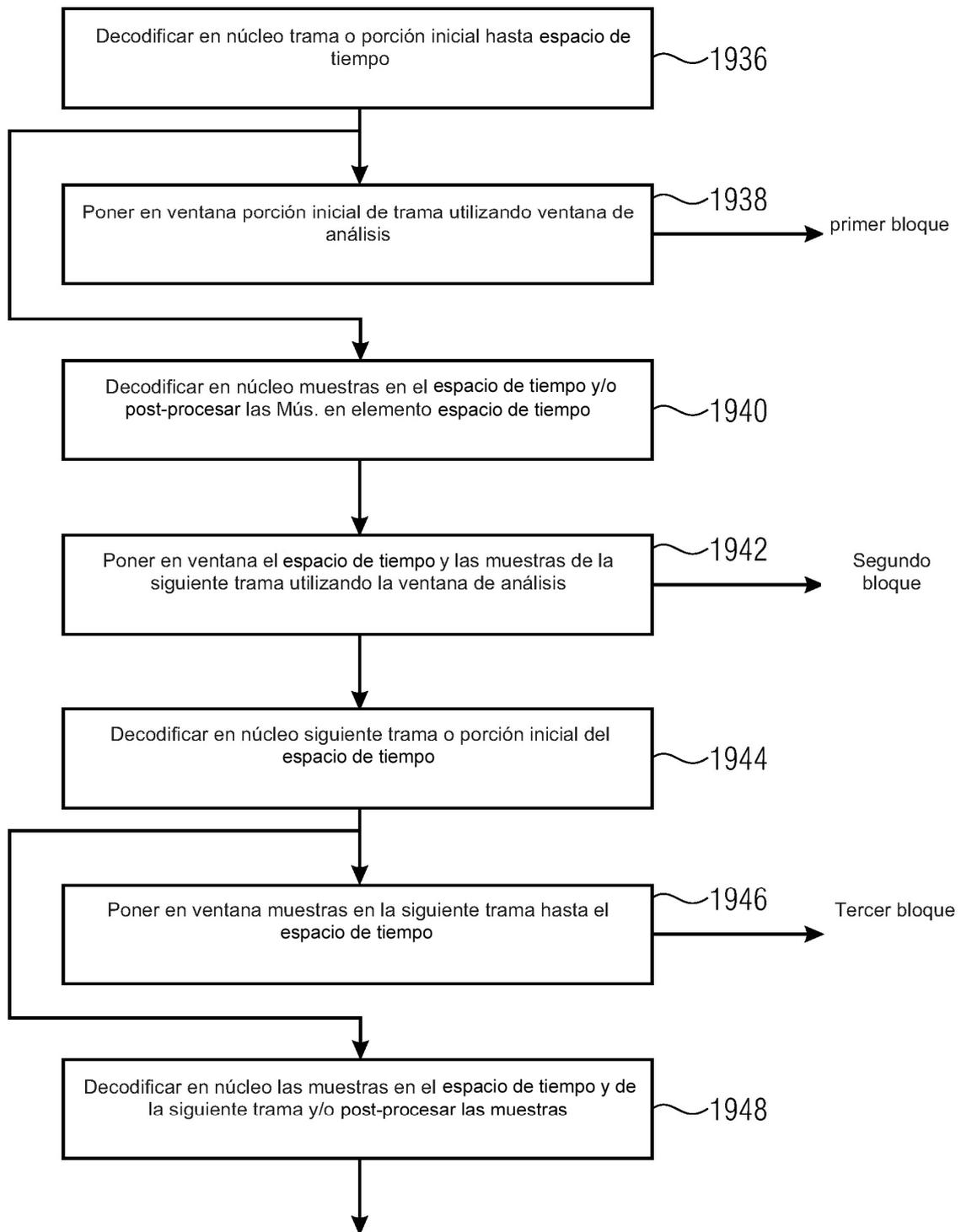
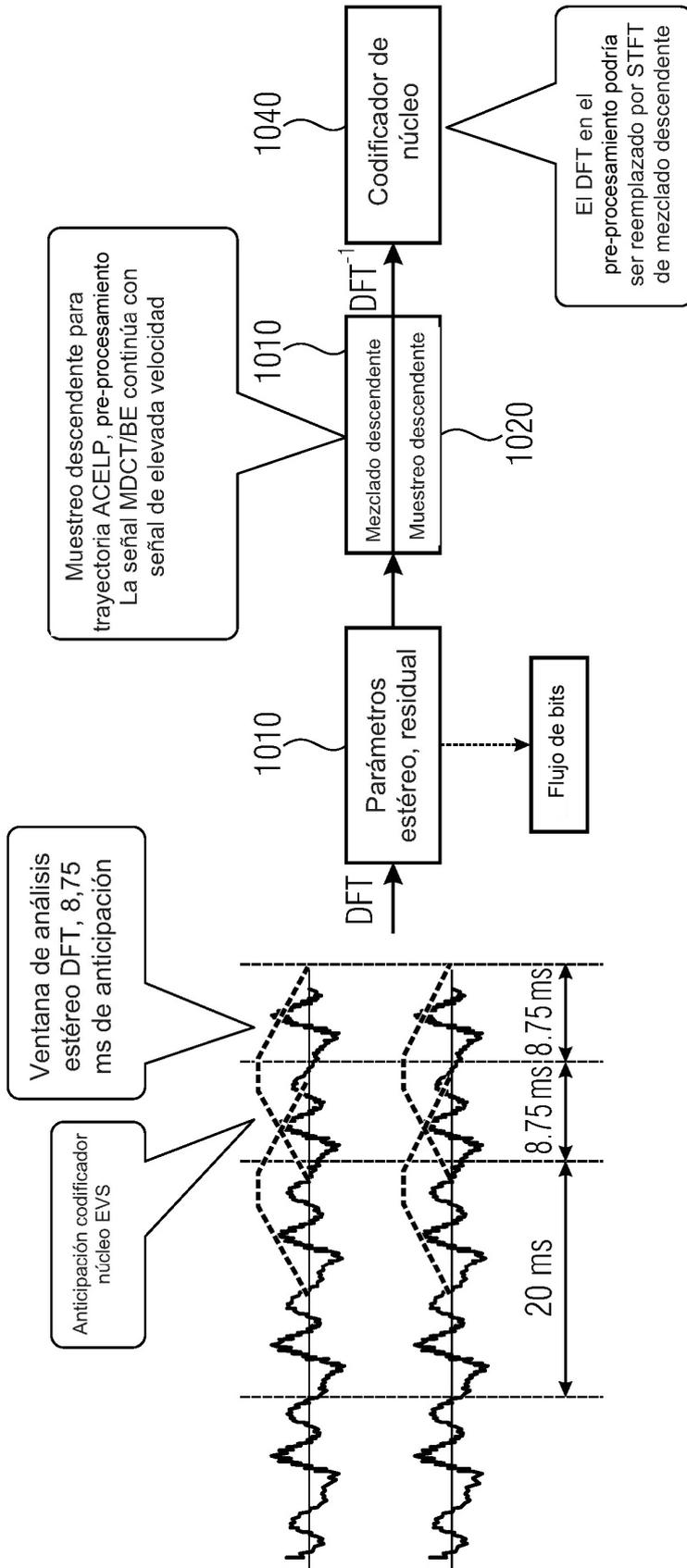
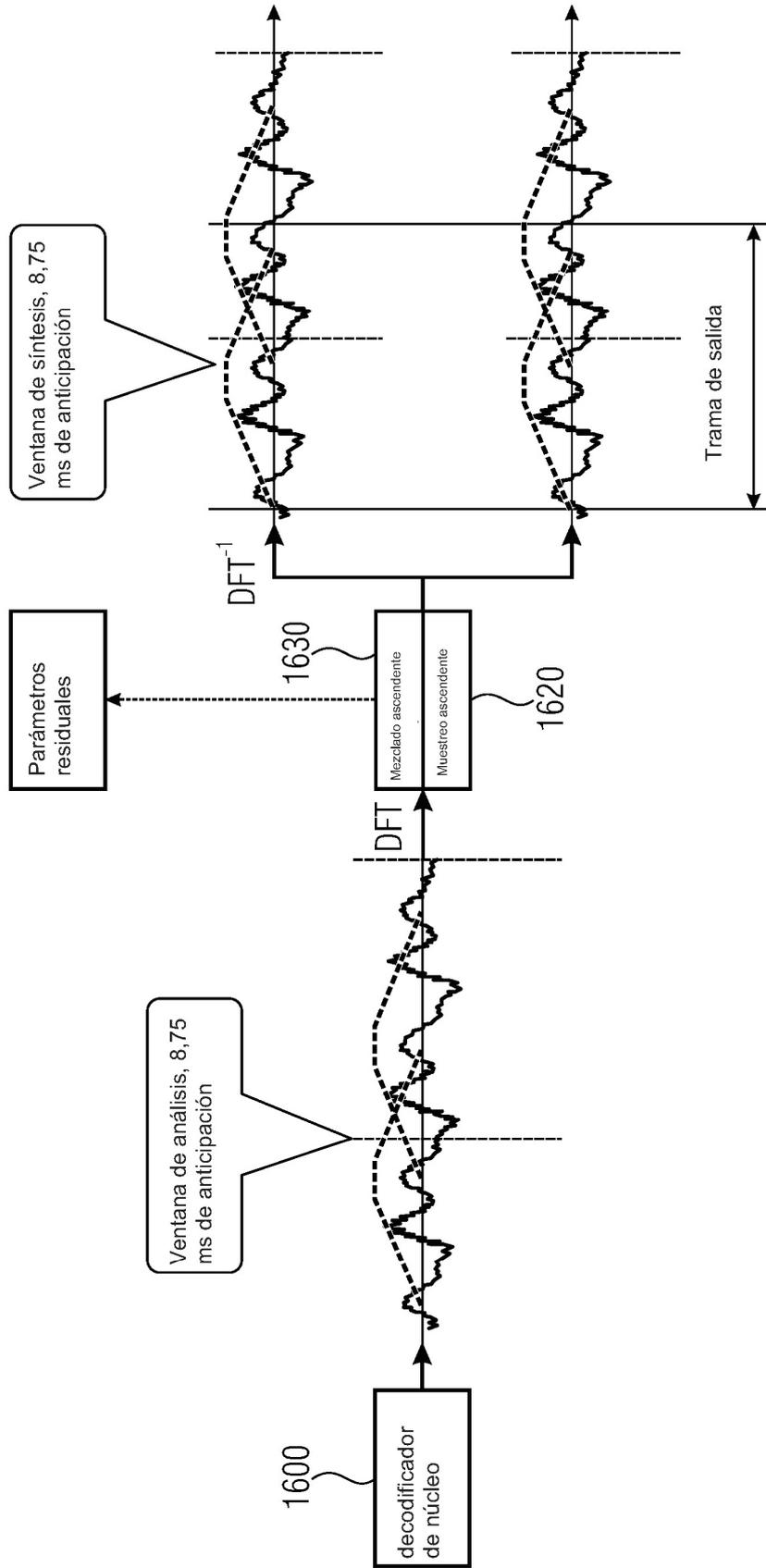


Fig. 9f



Propuesta 4. Puesta esquemática en ventana del codificador

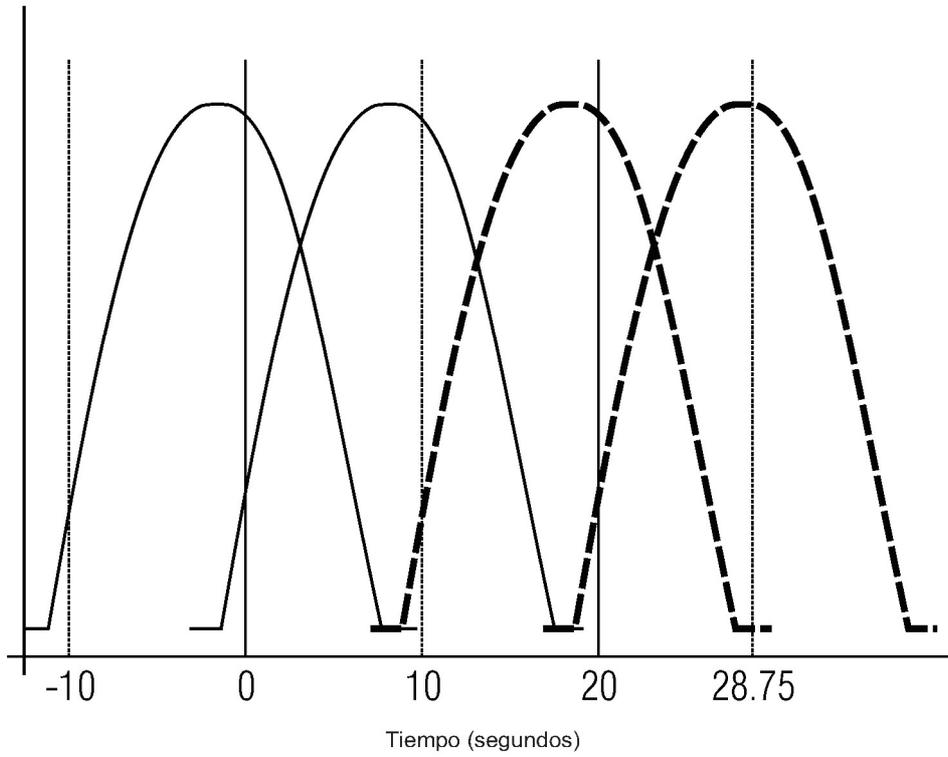
Fig. 10a



Propuesta 4. Puesta en ventana esquemática del decodificador

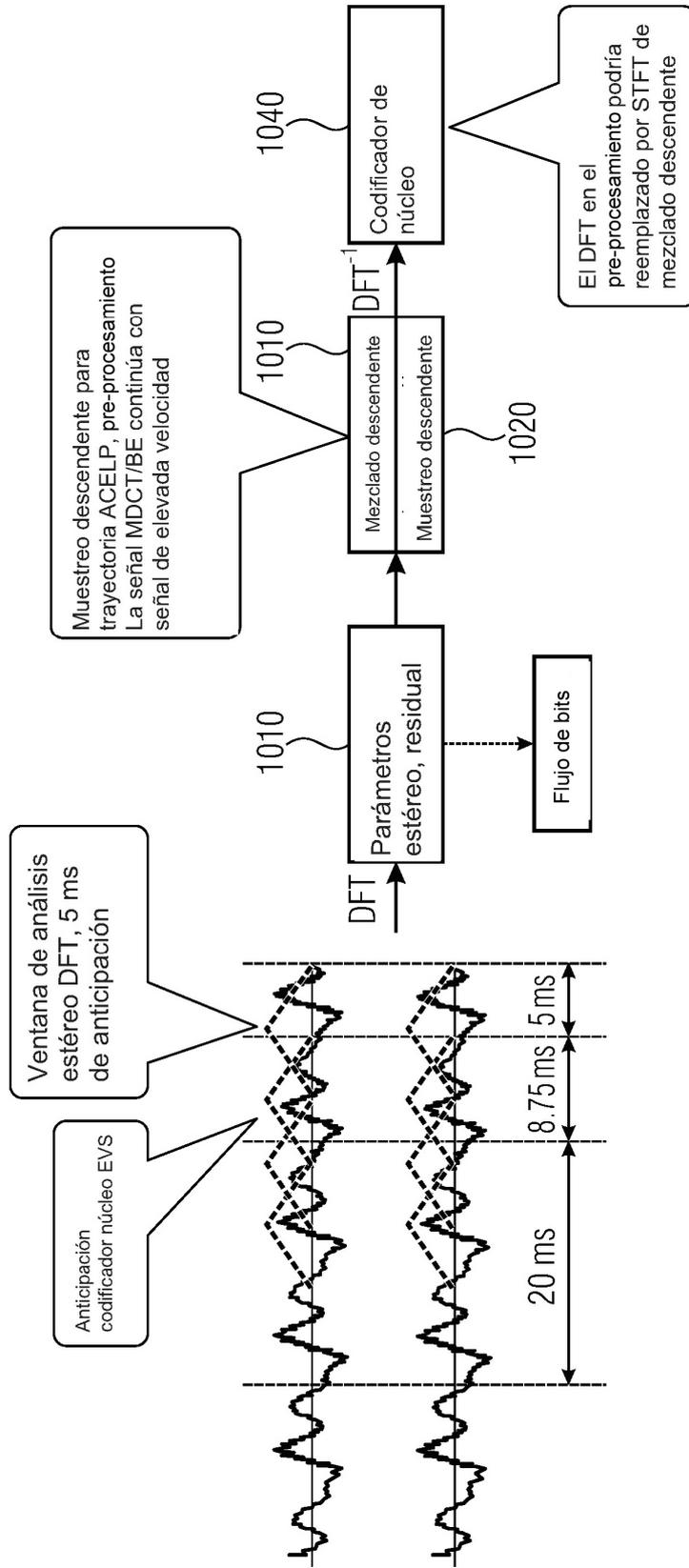
Fig. 10b

Ventanas de DFT de estéreo: propuesta # 4



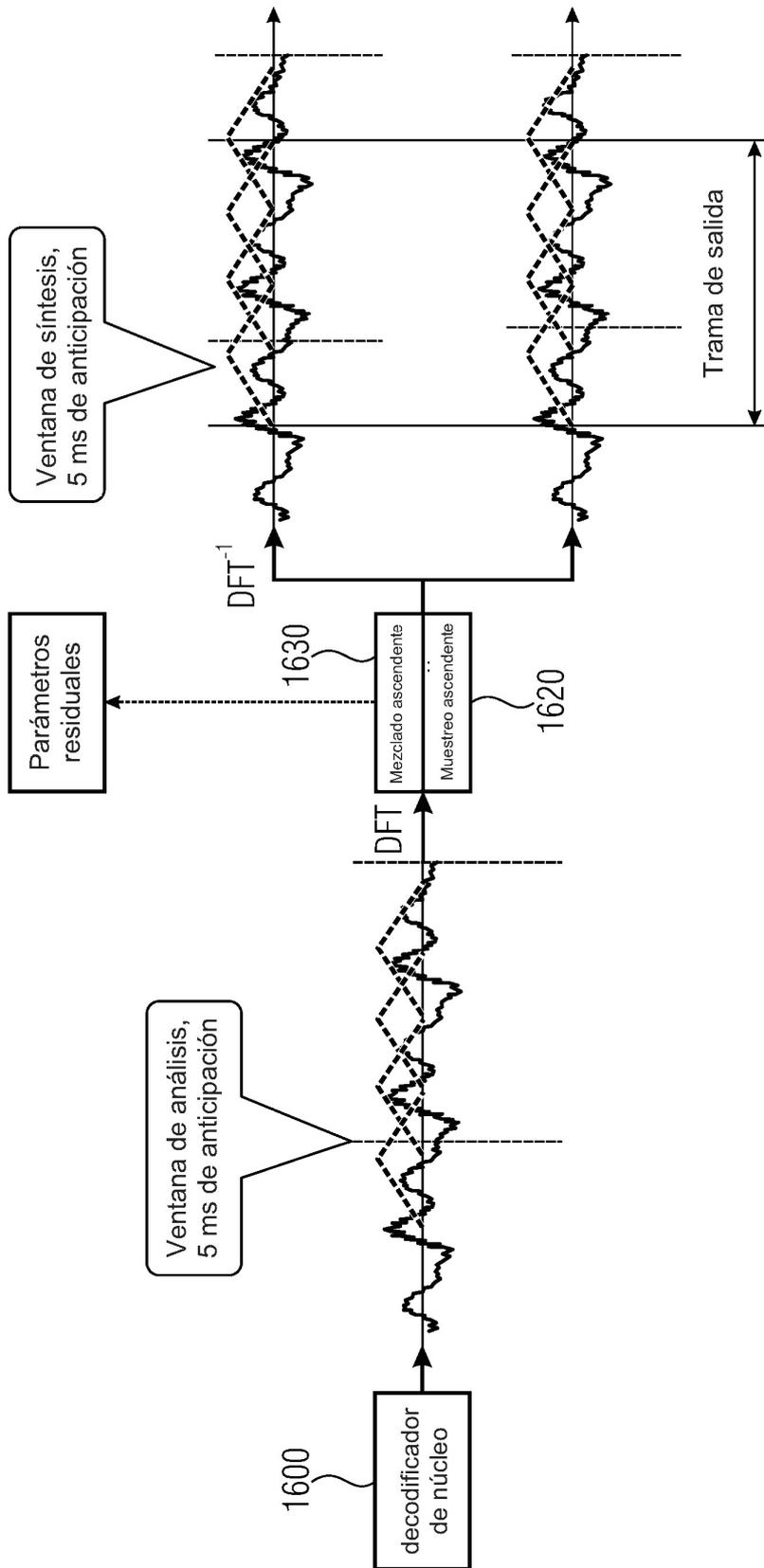
Propuesta 4- Ventanas en codificador y decodificador

Fig. 10c



Propuesta 5. Puesta esquemática en ventana del codificador

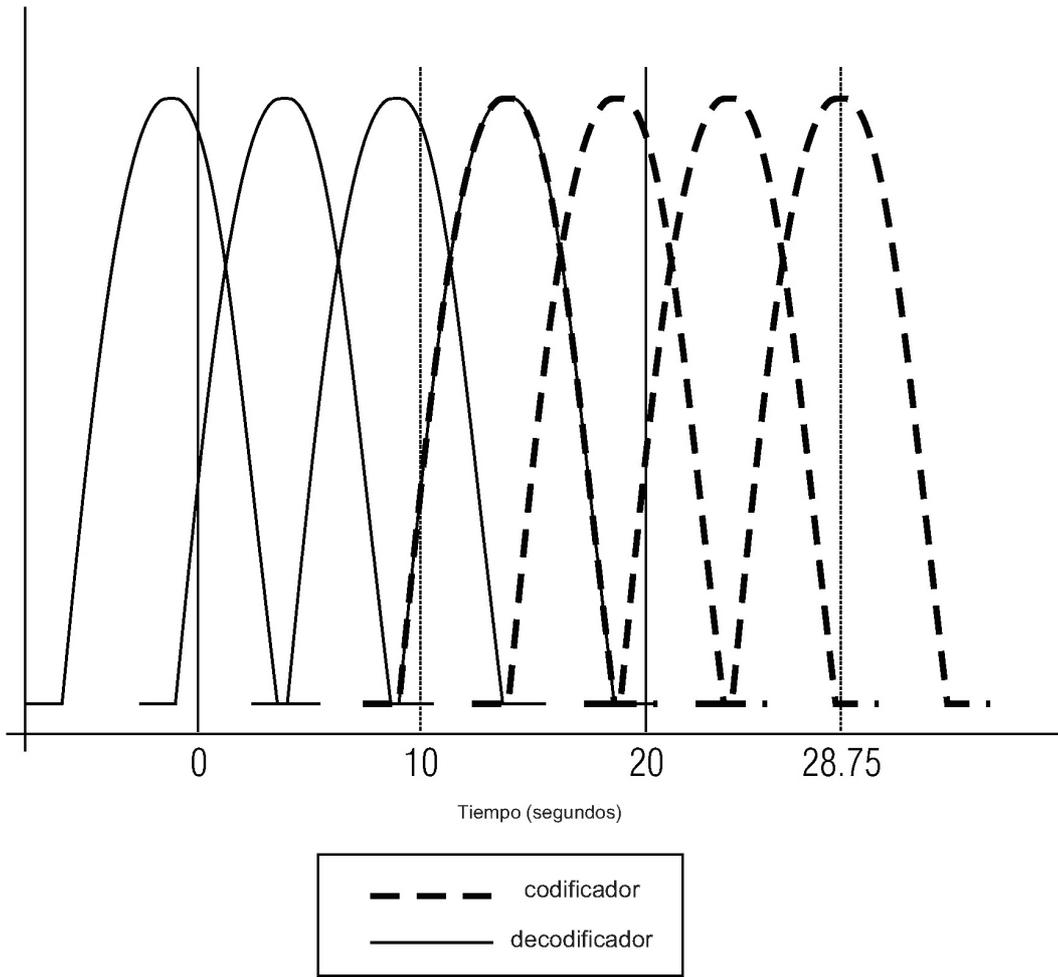
Fig. 11a



Propuesta 5. Puesta esquemática en venta del decodificador

Fig. 11b

Ventana de DFT: propuesta # 5



Propuesta 5: ventanas en codificador y decodificador

Fig. 11c

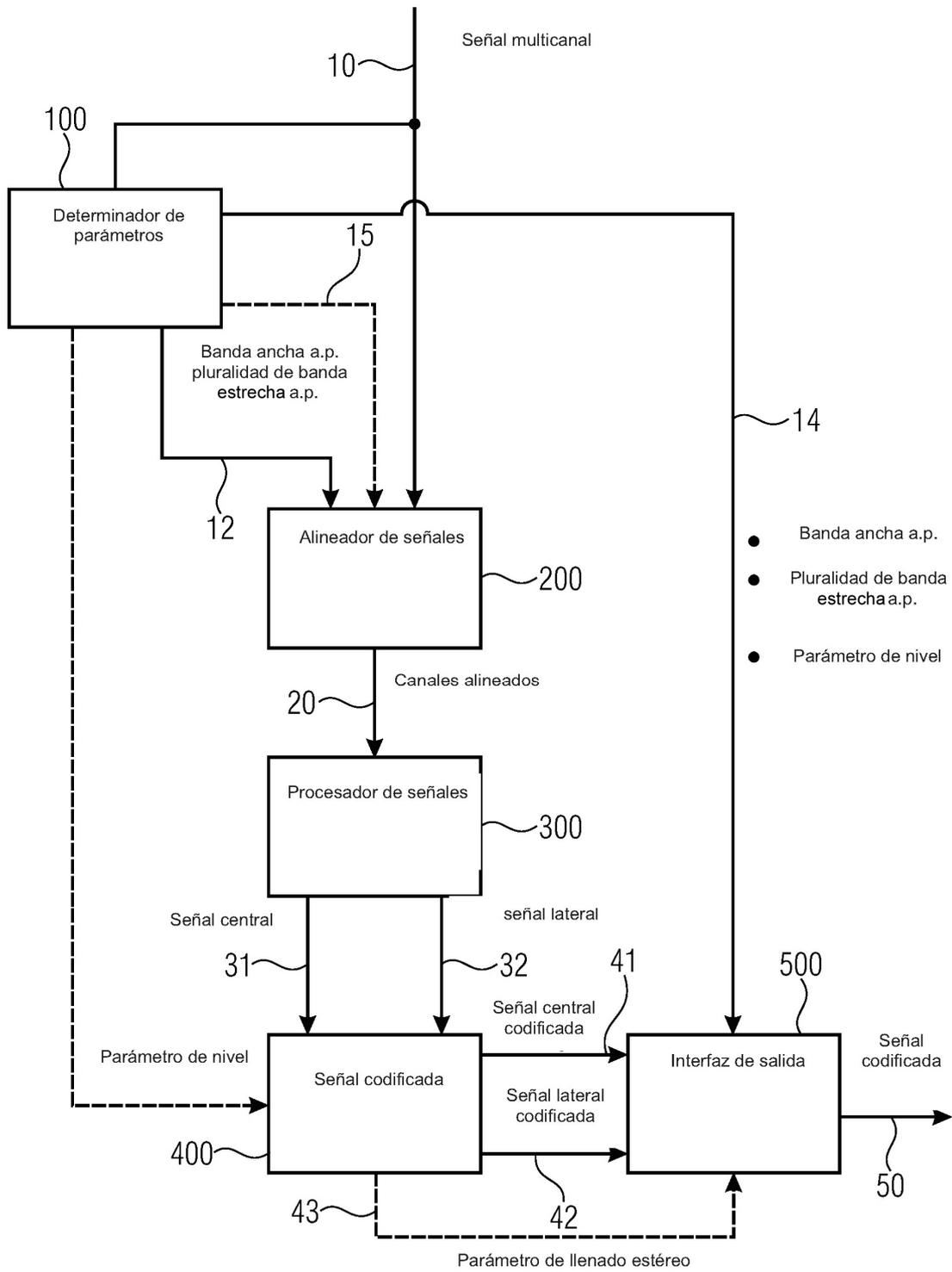


Fig. 12

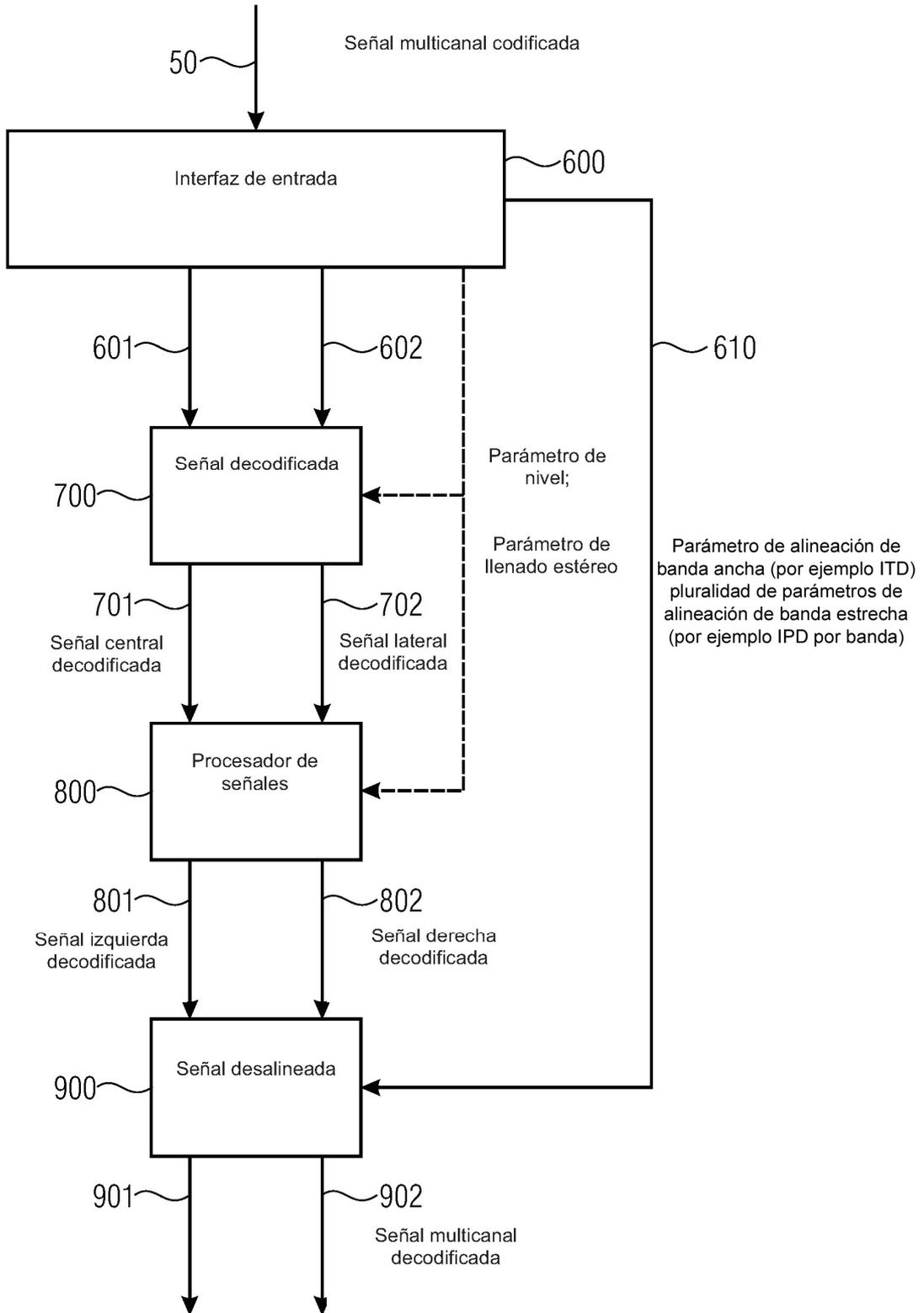


Fig. 13

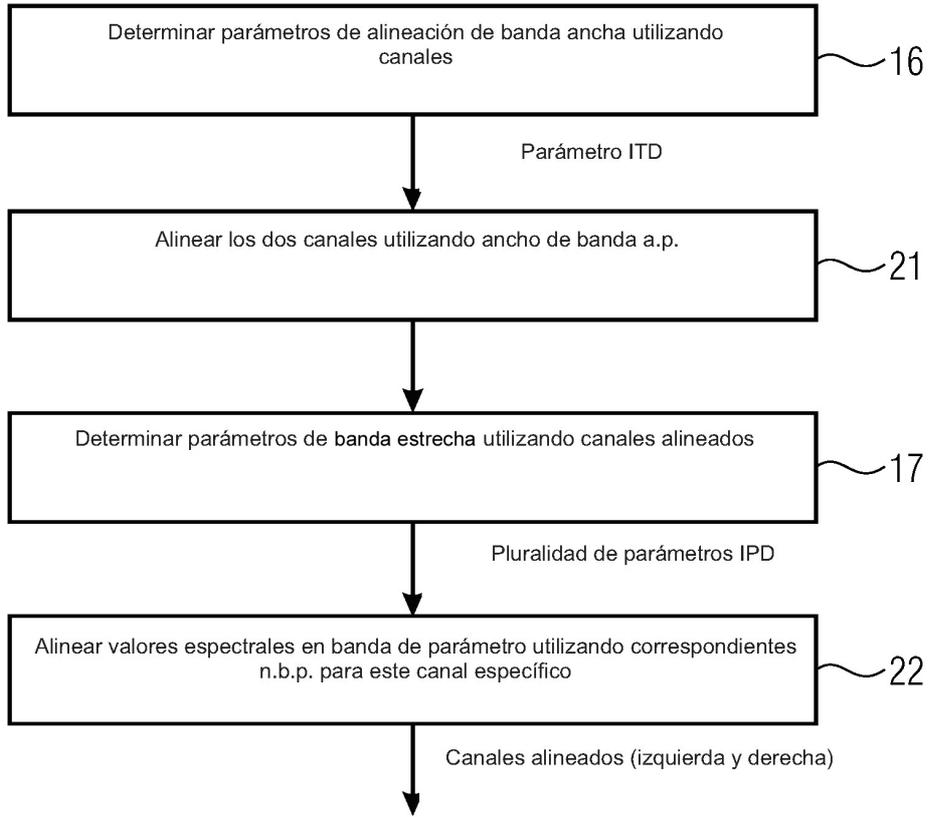


Fig. 14a

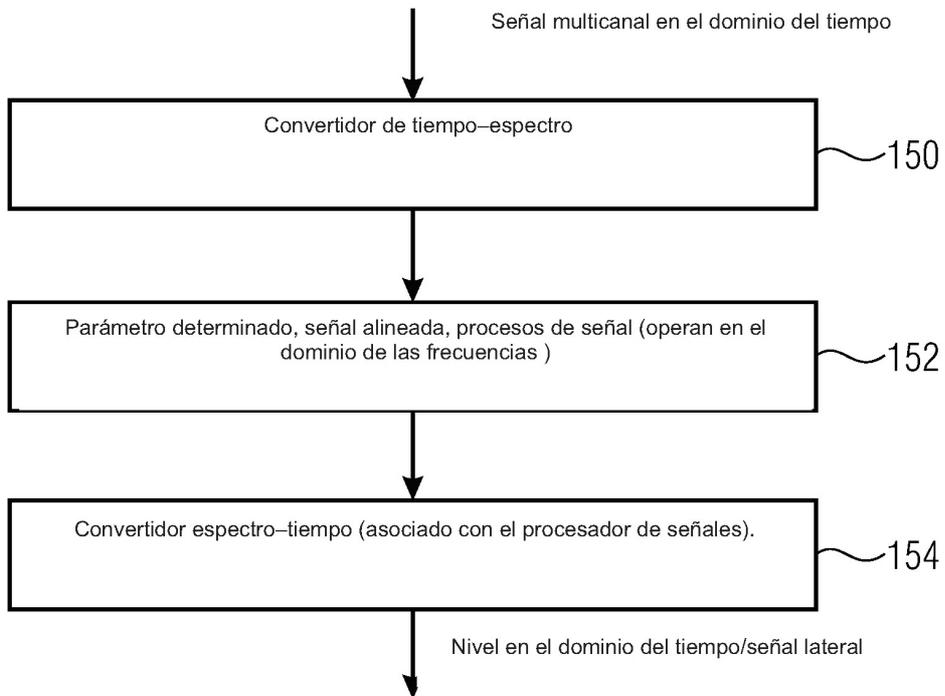


Fig. 14b

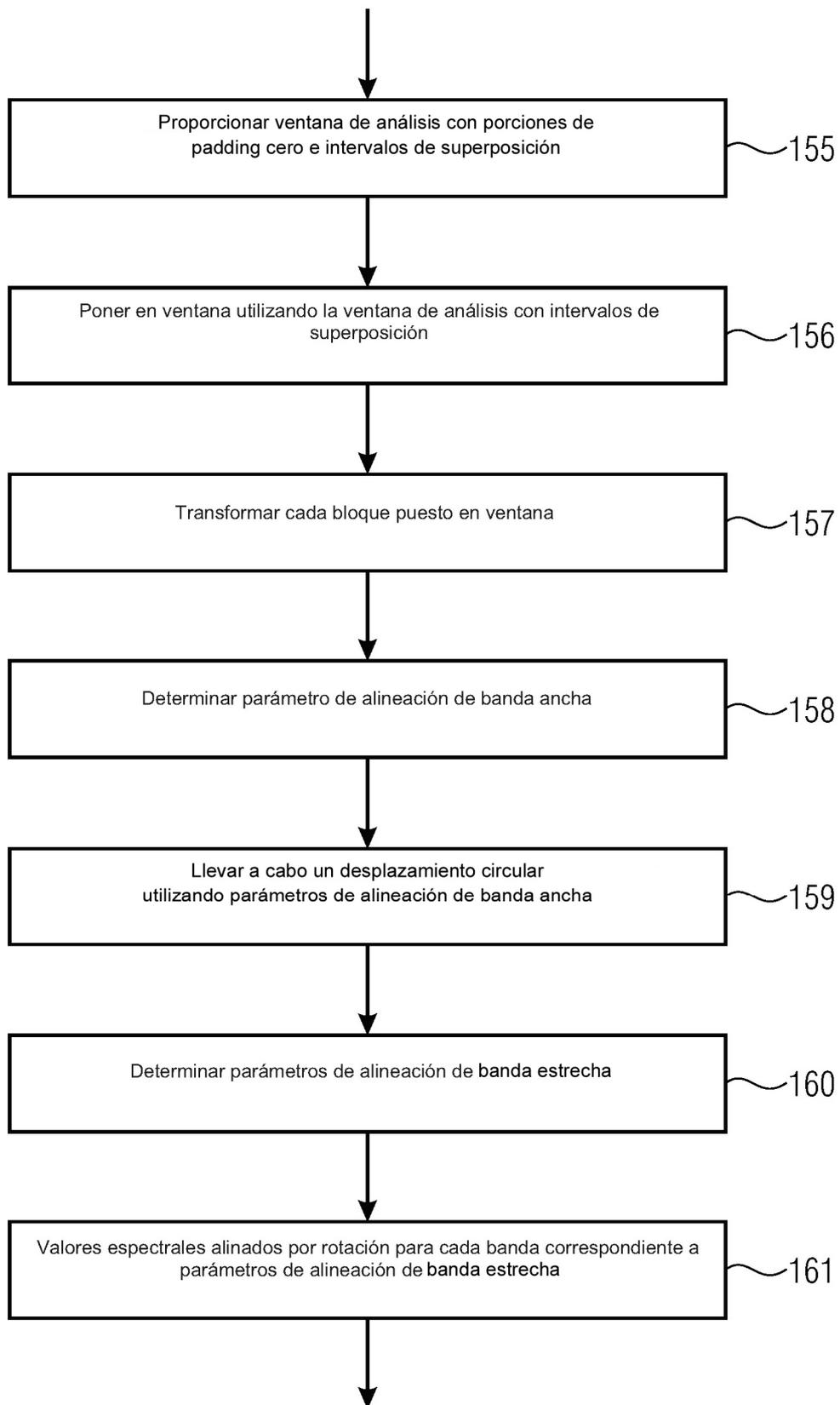


Fig. 14c

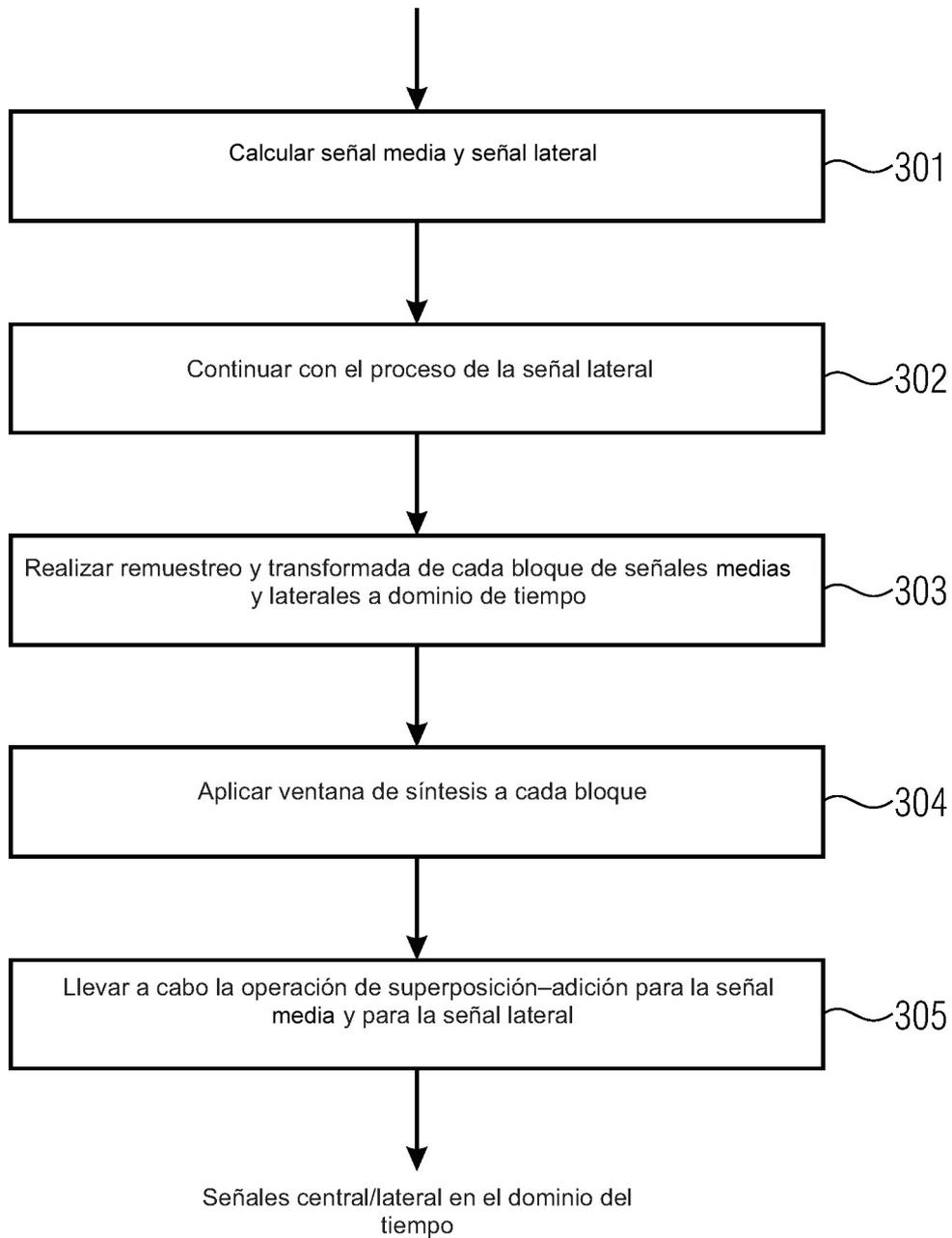


Fig. 14d

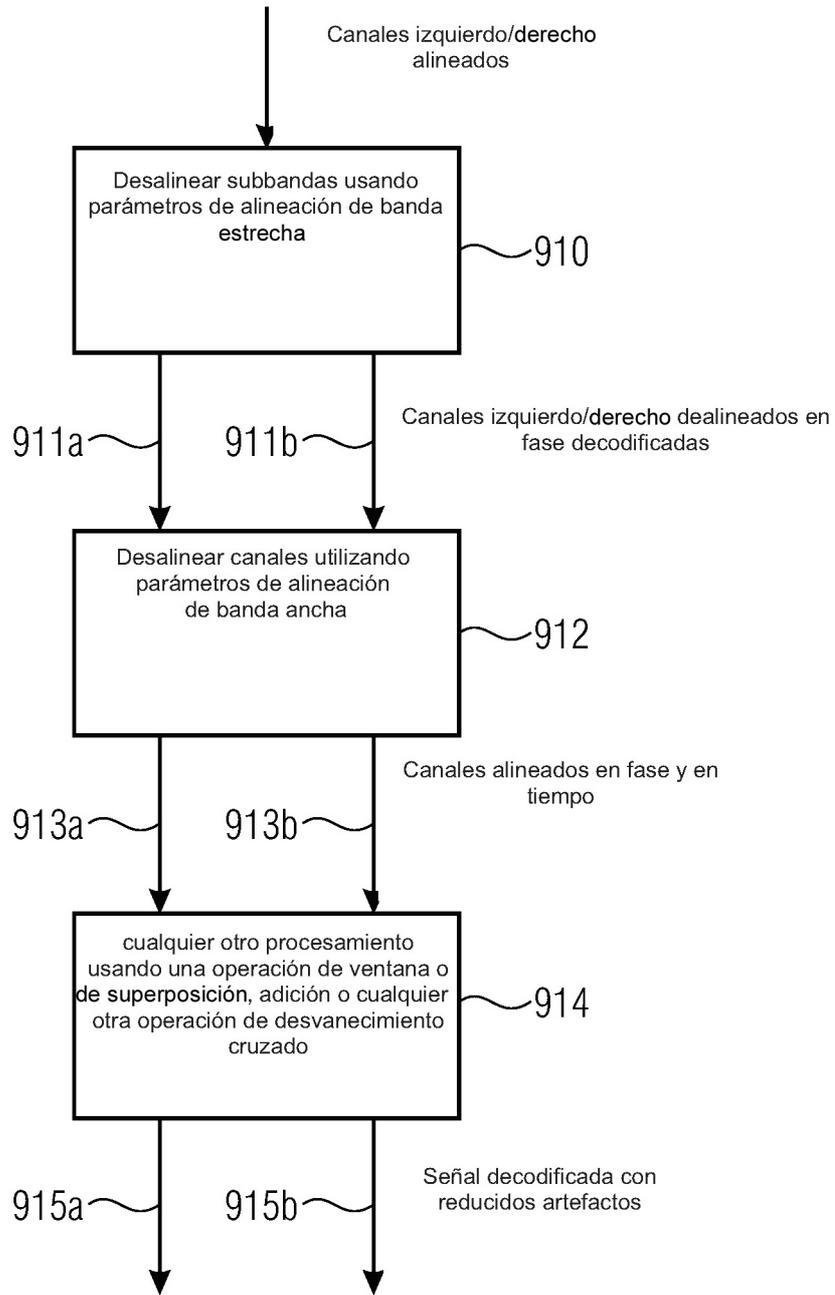


Fig. 15a

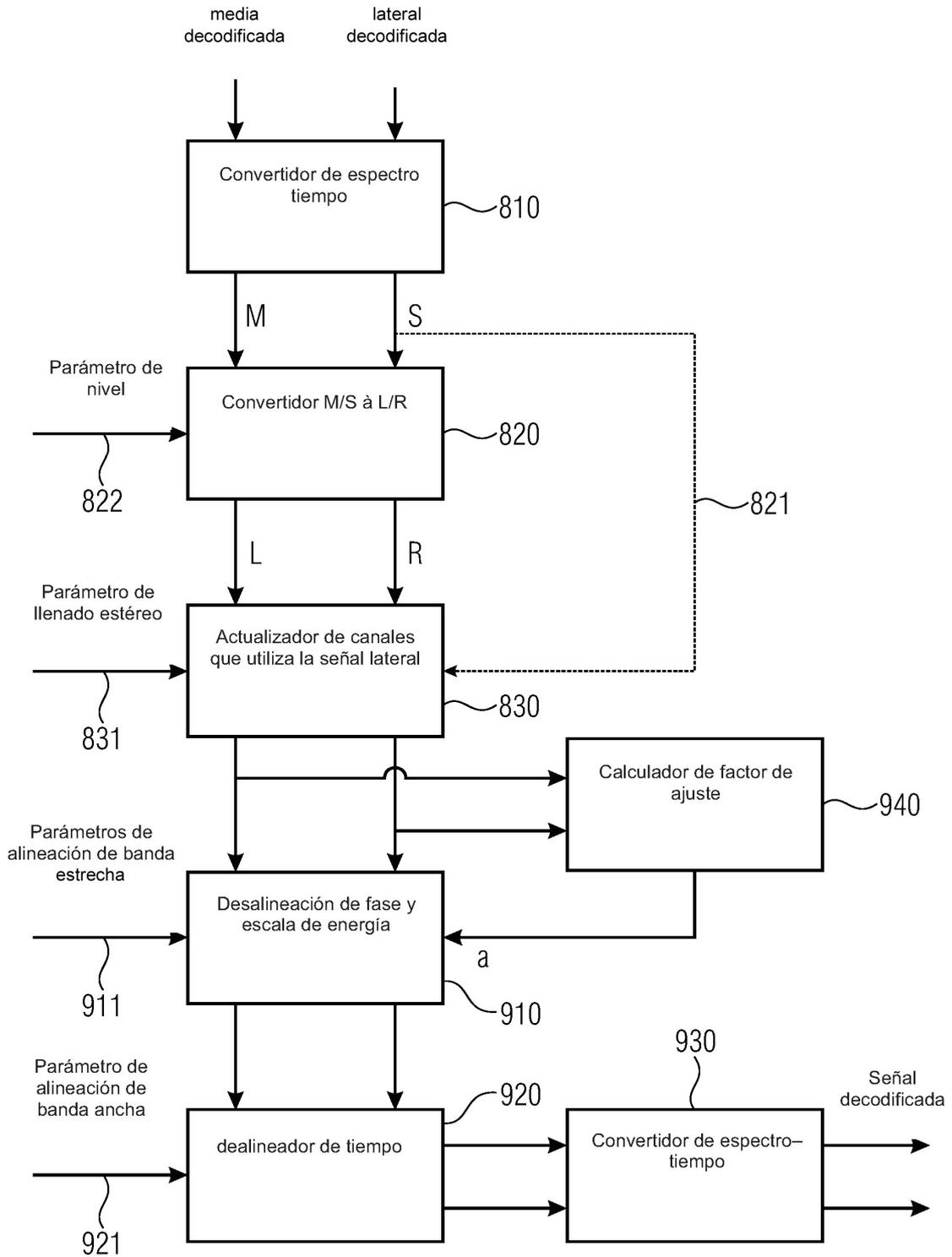


Fig. 15b

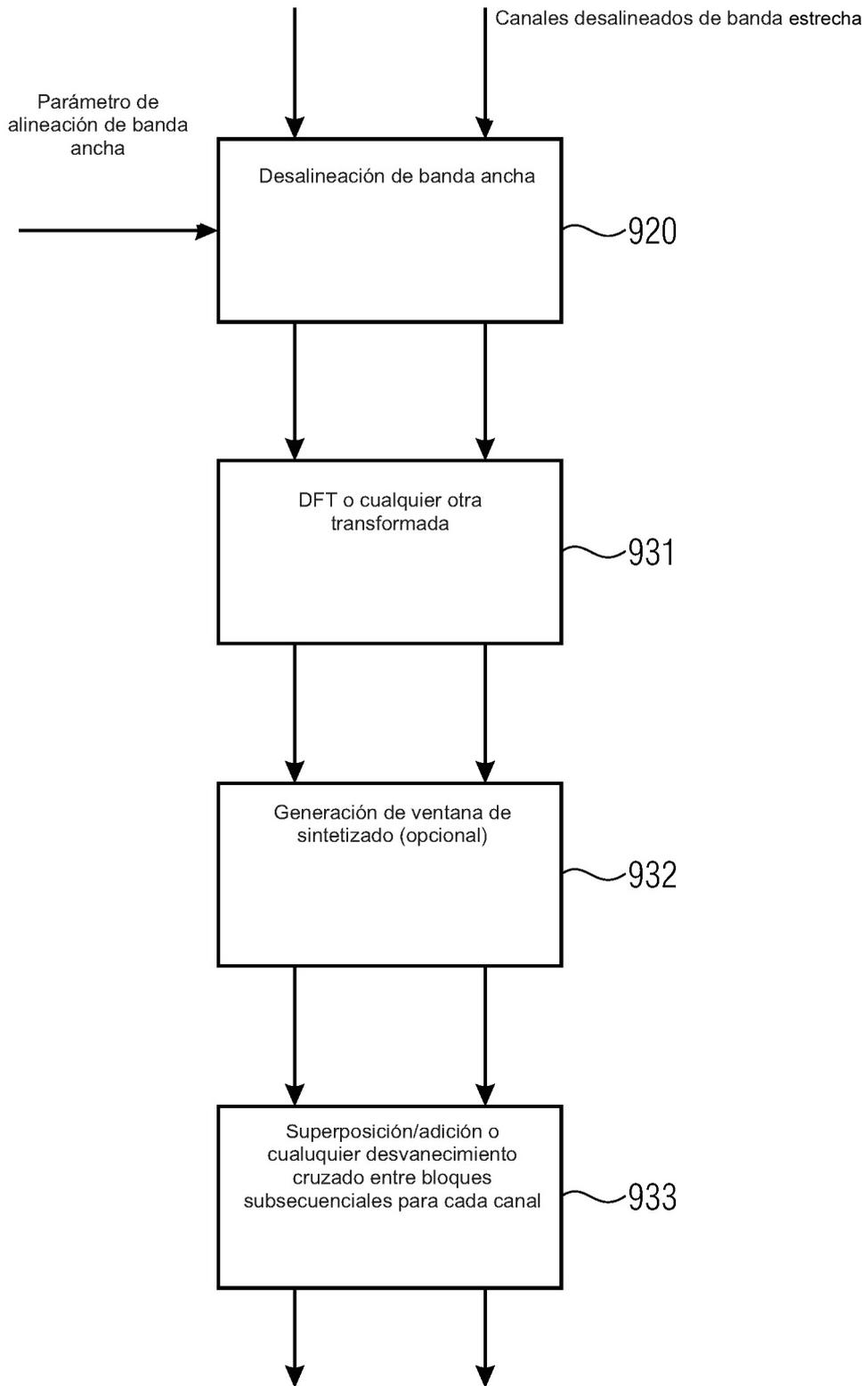


Fig. 15c