

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 768 052**

51 Int. Cl.:

G10L 19/022 (2013.01)

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.01.2017 PCT/EP2017/051212**

87 Fecha y número de publicación internacional: **27.07.2017 WO17125562**

96 Fecha de presentación y número de la solicitud europea: **20.01.2017 E 17701669 (8)**

97 Fecha y número de publicación de la concesión europea: **13.11.2019 EP 3405951**

54 Título: **Aparatos y procedimientos para codificar o decodificar una señal de audio multicanal usando sincronización de control de trama**

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:
19.06.2020

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 2 768 052 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 codificar o decodificar una señal de audio multicanal usando sincronización de control de trama

5

[0001] La presente solicitud se refiere al tratamiento estéreo o, en términos generales al tratamiento multicanal, en el que 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 especialmente el habla estéreo de conversación, ha recibido mucho menos atención científica que el almacenamiento y radiodifusión de música estereofónica. De hecho, hoy en día en las comunicaciones la transmisión monofónica es la más usada. Sin embargo, al aumentar la anchura de banda y la capacidad de la red, se prevé que las complicaciones basadas en tecnologías estereofónicas se harán más populares y proporcionarán una mejor experiencia de escucha.

15

[0003] La codificación eficiente 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 eficientes. Con elevadas velocidades binarias, en las que la conservación de la forma de onda es esencial, el estéreo de suma-diferencia, conocido como estéreo M/S (Mid o central/Side o lateral), ha sido usado durante mucho tiempo. Para 20 bajas velocidades binarias, se ha introducido el estéreo de intensidad y más recientemente la codificación estéreo paramétrica. Esta última técnica ha sido adoptada en diferentes normas tales como HeAACv2 y Mpeg USAC. Genera una submezcla de la señal de doble canal y asocia la información compacta del lado espacial.

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

30 **[0005]** Por otra parte, el estéreo paramétrico usa un banco de filtros adicional situado en el extremo frontal del codificador como preprocesador y en el extremo posterior del decodificador como posprocesador. Por tanto, el estéreo paramétrico puede usarse 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 binarias. Sin embargo, el estéreo paramétrico, como se da por ejemplo en MPEG USAC, no está específicamente diseñado para un retardo bajo y no proporciona 35 una calidad constante para diferentes escenarios de conversaciones. En la representación paramétrica convencional de la escena espacial, la anchura de la imagen estéreo es reproducida de manera artificial por un decorrelacionador aplicado sobre los dos canales sintetizados y controlados por parámetros de coherencia entre canales (IC, Interchannel-Coherence) calculados y transmitidos por el codificador. Para la mayor parte del habla estéreo, esta manera de ensanchar la imagen estéreo no es adecuada para recrear el ambiente natural del habla que es un 40 sonido muy directo dado que es producido por una única fuente situada en una posición específica en el espacio (a veces con alguna reverberación debida al ambiente interior). En cambio, los instrumentos musicales tienen una anchura mucho más natural que el habla, que puede imitarse mejor mediante la decorrelación de los canales.

45 **[0006]** También se presentan problemas cuando se registra el habla con micrófonos no coincidentes, como en la configuración A-B cuando los micrófonos están separados entre sí o para el registro o reproducción binaurales. Pueden verse 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). En este caso el momento de la llegada de la señal es diferente de un canal a otro a diferencia de las grabaciones efectuadas en micrófonos coincidentes tales como X-Y (registro de intensidad) o M-S (registro Mid-Side, central-lateral). En 50 consecuencia, el cálculo de la coherencia de estos dos canales no alineados en el tiempo puede estimarse equivocadamente, lo que provoca un fallo de la síntesis del ambiente artificial.

[0007] Las referencias del estado anterior de la técnica relacionadas con el tratamiento estéreo son la patente de EE.UU. nº 5.434.948 o la patente de EE.UU. nº 8.811.621.

55

[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 60 reforzado genera una señal de salida multicanal provista de 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 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 subbanda. Una alineación de este tipo se lleva entonces a cabo antes de un tratamiento adicional. En el lado del decodificador, se lleva a cabo una desalineación y un tratamiento de ganancia, y las correspondientes 65 señales son sintetizadas a continuación por un banco de filtros de síntesis con el fin de generar una señal izquierda

decodificada y una señal derecha decodificada.

[0009] Por otra parte, el estéreo paramétrico usa un banco de filtros adicional situado en el extremo frontal del codificador como preprocesador y en el extremo posterior del decodificador como posprocesador. Por tanto, el estéreo paramétrico puede usarse con codificadores de estéreo convencionales tales como ACELP ya que se efectúa en MPEG USAC. Por otra parte, la parametrización de la escena de auditorio puede efectuarse con una cantidad mínima de información lateral, lo que es conveniente para bajas velocidades binarias. Sin embargo, el estéreo paramétrico como por ejemplo en MPEG USAC no está diseñado específicamente para un retardo bajo, y el sistema en su conjunto muestra un retardo algorítmico muy elevado.

[0010] A partir de la publicación de T. Jansson "Stereo coding for the ITU-T G.719 codec", UPTec F11 034, 17-05-2011, se conocen también técnicas para codificación estéreo y una exposición para reducir la complejidad y el retardo.

15 RESUMEN DE LA INVENCION

[0011] Un objetivo de la presente invención es proporcionar un concepto mejorado para la codificación/decodificación multicanal, que sea eficiente y con la capacidad de proporcionar un retardo bajo.

[0012] Este objetivo se logra mediante un aparato para codificar una señal multicanal según la reivindicación 1, un procedimiento para codificar una señal multicanal según la reivindicación 24, un aparato para decodificar una señal multicanal codificada según la reivindicación 25, un procedimiento para decodificar una señal multicanal codificada según la reivindicación 42 o un programa informático según la reivindicación 43.

[0013] La presente invención se basa en el hallazgo de que al menos una porción y preferentemente todas las partes del tratamiento multicanal, es decir, un tratamiento multicanal conjunto, se llevan a cabo en un dominio espectral. Específicamente, se prefiere llevar a cabo la operación de submezcla del tratamiento 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 tratamiento estéreo conjunto/multicanal conjunto. Además, se lleva a cabo una sincronización del control de trama para el codificador de núcleo y el tratamiento estéreo que opera en el dominio espectral.

[0014] El codificador de núcleo está configurado para operar según un primer control de trama con el fin de proporcionar una secuencia de tramas, en el que una trama está delimitada por un límite de trama inicial y un límite de trama final, y el convertidor de tiempo-espectral o el convertidor espectral-tiempo están configurados para operar según un segundo control de trama que está sincronizado con el primer control de trama, en el que el límite de trama inicial o el límite de trama final de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con un instante inicial o un instante final de una porción superpuesta de una ventana usada por el convertidor de tiempo-espectral (1000) para cada bloque de la secuencia de bloques de valores de muestreo o usada por el convertidor espectral-tiempo para cada bloque de la secuencia de bloques de salida de valores de muestreo.

[0015] En la invención, el codificador de núcleo del codificador multicanal está configurado para operar según un control de tramas, y el convertidor de tiempo-espectral y el convertidor espectral-tiempo del posprocesador estéreo y el remuestreador también están configurados para operar según un control de tramas adicional que está sincronizado con el control de tramas del codificador de núcleo. La sincronización se lleva a cabo de tal manera que un límite de trama inicial o un límite de trama final de cada trama de una secuencia de tramas del codificador de núcleo se encuentra en una relación predeterminada con un instante inicial o un instante final de una porción superpuesta de una ventana usada por el convertidor de tiempo-espectral o el convertidor espectral-tiempo para cada bloque de la secuencia de bloques de valores de muestreo o para cada bloque de la secuencia de bloques remuestreada de valores espectrales. Así, se garantiza que las operaciones de formación de tramas operan sincronizadamente entre sí.

[0016] En realizaciones adicionales, se lleva a cabo una operación de anticipación con una porción de anticipación por medio del codificador de núcleo. En esta realización, se prefiere que la porción de anticipación sea usada también por una ventana de análisis del convertidor de tiempo-espectral en la que se usa una porción superpuesta de la ventana de análisis que tiene una longitud de tiempo que es menor o igual que la longitud de tiempo de la porción de anticipación.

[0017] Así, al hacer iguales entre sí la porción de anticipación del codificador de núcleo y la porción superpuesta de la ventana de análisis o al hacer que la porción superpuesta sea menor incluso que la porción de anticipación del codificador de núcleo, el análisis tiempo-espectral del preprocesador estéreo no puede implementarse sin un retardo algorítmico adicional. Con el fin de cerciorarse de que esta porción de anticipación dividida en ventanas no influye demasiado en la funcionalidad de anticipación del codificador de núcleo, se prefiere rectificar esta porción usando una inversa de la función de ventana de análisis.

- 5 **[0018]** Con el fin de cerciorarse de que esto se hace con una buena estabilidad, se usa una raíz cuadrada de la forma de ventana sinusoidal en lugar de una forma de ventana sinusoidal como ventana de análisis y se usa el seno de la ventana de síntesis a la potencia 1,5 con el fin de realizar una división en ventanas de síntesis antes de realizar la operación de superposición en la salida del convertidor espectral-tiempo. Así, se asegura que la función de compensación asume valores que son reducidos con respecto a sus magnitudes en comparación con una función de compensación que es la inversa de una función seno.
- 10 **[0019]** Preferentemente, el remuestreo en el dominio espectral se lleva a cabo con posterioridad al tratamiento multicanal o antes incluso del tratamiento 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 requerida por un codificador de núcleo conectado posteriormente. No obstante, el procedimiento de la invención de sincronización del control de tramas del codificador de núcleo y el convertidor espectral-tiempo o tiempo-espectral también puede aplicarse en un escenario en el que no se ejecute ningún remuestreo de dominio espectral.
- 15 **[0020]** En el lado del decodificador, se prefiere de nuevo 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 submezcla en el dominio espectral y preferentemente llevar a cabo incluso el tratamiento multicanal inverso total en el dominio espectral. Además, se proporciona el convertidor de tiempo-espectral para convertir la señal decodificada de núcleo en una representación en el dominio espectral y, dentro del dominio de la frecuencia, se lleva a cabo el tratamiento multicanal inverso.
- 20 **[0021]** El decodificador de núcleo está configurado para operar según un primer control de trama con el fin de proporcionar una secuencia de tramas, en el que una trama está delimitada por un límite de trama inicial y un límite de trama final. El convertidor de tiempo-espectral o el convertidor espectral-tiempo están configurados para que operen según un segundo control de trama que está sincronizado con el primer control de trama. Específicamente, el convertidor de tiempo-espectral o el convertidor espectral-tiempo están configurados para operar según un segundo control de trama que está sincronizado con el primer control de trama, en el que el límite de trama inicial o el límite de trama final de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con un instante inicial o un instante final de una porción superpuesta de una ventana usada por el convertidor de tiempo-espectral para cada bloque de la secuencia de bloques de valores de muestreo o usada por el convertidor espectral-tiempo para cada bloque de los al menos dos secuencias de bloques de salida de valores de muestreo.
- 25 **[0022]** Se prefiere usar las mismas formas de ventanas de análisis y síntesis, ya que no se requiere, naturalmente, compensación. Por otra parte, se prefiere usar un intervalo de tiempo en el lado del decodificador, en el que el intervalo de tiempo existe entre un extremo de una porción superpuesta delantera de una ventana de análisis del convertidor de tiempo-espectral en el lado del decodificador y un instante de tiempo al final de una salida de trama por el decodificador de núcleo en el lado del decodificador multicanal. Así, las muestras de salida del decodificador de núcleo dentro de este intervalo de tiempo no son necesarias para la división en ventanas de análisis por parte del posprocesador estéreo inmediatamente, sino que se requieren solo para el tratamiento/división en ventanas de la siguiente trama. Dicho intervalo de tiempo puede ser implementado, por ejemplo, usando una porción no superpuesta normalmente en la mitad de una ventana de análisis que produce un acortamiento de la porción superpuesta. Sin embargo, pueden usarse también otras alternativas para implementar dicho intervalo de tiempo, si bien la forma preferida es la implementación del intervalo de tiempo por la porción no superpuesta en la mitad. Así, este intervalo de tiempo puede usarse para otras operaciones del decodificador de núcleo u operaciones de suavización preferentemente entre sucesos de conmutación cuando el decodificador de núcleo cambia de una trama en el dominio de la frecuencia al dominio del tiempo o para cualquier otra operación de suavización que pueda ser útil cuando se han producido cambios de los parámetros o cambios de las características de codificación.
- 30 **[0023]** En una realización se lleva a cabo un remuestreo en el dominio espectral antes del tratamiento multicanal inverso o bien después del tratamiento multicanal inverso de tal manera que, al final, un convertidor espectral-tiempo convierte una señal remuestreada espectralmente 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.
- 35 **[0024]** Por tanto, las realizaciones permiten evitar por completo cualquier operación de remuestreo en el dominio del tiempo que requiera un cálculo intensivo. Por el contrario, se combina el tratamiento multicanal con el remuestreo. En realizaciones preferidas, el remuestreo se lleva a cabo mediante truncamiento del espectro en caso de submuestreo o mediante relleno con ceros del espectro en caso de sobremuestreo. Estas operaciones fáciles, es decir, el truncamiento del espectro por una parte o el relleno con ceros del espectro por otra parte, y los cambios de escala adicionales preferentes con el fin de tener en cuenta determinadas operaciones de normalización llevadas a cabo en los algoritmos de conversión del dominio espectral/dominio del tiempo tales como el algoritmo de DFT o FFT, completan la operación de remuestreo en el dominio espectral de una manera muy eficiente y con retardo bajo.
- 40 **[0025]** Además, se ha descubierto que al menos una porción o incluso todo el tratamiento estéreo conjunto/tratamiento multicanal conjunto en el lado del codificador y el tratamiento multicanal inverso
- 45

correspondiente en el lado del decodificador son adecuados para ser ejecutados en el dominio de la frecuencia. Esto no solo es válido para la operación de submezcla como un tratamiento multicanal conjunto mínimo en el lado del codificador o para un tratamiento de sobremezcla como para un tratamiento multicanal inverso mínimo en el lado del decodificador. También puede llevarse a cabo un análisis estéreo de la 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. Rige lo mismo para la codificación de canal lateral realizada preferentemente en el lado del codificador o para síntesis de canal lateral y para el uso en la generación de los dos canales de salida decodificados en el lado del decodificador.

[0026] Por tanto, 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 realizaciones de la presente invención proporcionan un nuevo marco de trabajo para lograr un códec estéreo de retardo bajo y para integrar una herramienta estéreo común realizada en el dominio de la frecuencia tanto para un codificador de núcleo del habla como para un codificador de núcleo basado en MDCT dentro de un códec de audio conmutado.

[0027] Las realizaciones de la presente invención se refieren a un enfoque híbrido en el que se mezclan elementos de un estéreo M/S o estéreo paramétrico convencional. Las realizaciones usan algunos aspectos y herramientas de la codificación estéreo conjunta y otros de estéreo paramétrico. Más en concreto, las realizaciones adoptan el análisis y síntesis de tiempo-frecuencia 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 el uso de un banco de filtros o bien de una transformada de bloque con valores complejos. A partir de la entrada de dos canales o de múltiples canales, el tratamiento estéreo o multicanal combina y modifica los canales de entrada para los canales de salida referidos como señales central y lateral (MS, Mid and Side).

[0028] Las realizaciones de la presente invención proporcionan una solución para reducir un retardo algorítmico introducido por un módulo estéreo y especialmente a partir del entramado y división en ventanas de su banco de filtros. Proporciona una transformada inversa de velocidades múltiples 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 al producir la misma señal de tratamiento estéreo con diferentes velocidades de muestreo. Además, proporciona una división en ventanas adaptada para las diferentes restricciones del sistema de retardo bajo y baja complejidad y también para el tratamiento estéreo. Además, las realizaciones proporcionan un procedimiento para combinar y muestrear repetidamente diferentes resultados de síntesis decodificados en el dominio espectral, en el que también se aplica el tratamiento estéreo inverso.

[0029] Las realizaciones preferidas de la presente invención comprenden una multifunción en un remuestreador en el dominio espectral que no solo genera un bloque individual remuestreado en el dominio espectral de valores espectrales sino, adicionalmente, también otra secuencia remuestreada de bloques de valores espectrales correspondientes a una velocidad de muestreo diferente, más alta o más baja.

[0030] 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 primero y segundo original introducido en el convertidor de tiempo-espectral en el lado del codificador. Por tanto, en algunas realizaciones, el codificador multicanal proporciona al menos una señal de salida con la velocidad de muestreo introducida original, que se usa preferentemente para una codificación basada en MDCT. Adicionalmente, se proporciona al menos una señal de salida con una velocidad de muestreo intermedia que es útil específicamente para la codificación de ACELP y que además proporciona una señal de salida adicional a otra velocidad de muestreo de salida que también es útil para la codificación de ACELP, pero que es diferente de la velocidad de muestreo de salida adicional.

[0031] Estos procedimientos pueden llevarse a cabo sea para la señal central (Mid) o para la señal lateral (Side), o para ambas señales derivadas de las señales de los canales primera y segunda de una señal multicanal en la que 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 solo tenga dos canales (adicionalmente dos, por ejemplo, un canal de mejora de baja frecuencia).

[0032] A continuación se exponen en detalle realizaciones preferidas de la presente invención con referencia a los dibujos adjuntos, en los que:

la Figura 1 es un diagrama de bloques de una realización de codificador multicanal;

la Figura 2 ilustra realizaciones del remuestreo 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 cambios de escala correspondientes en el dominio espectral;

- la Figura 3d ilustra diferentes resoluciones de frecuencia y otros aspectos relacionados con la frecuencia, para determinadas realizaciones;
- 5 la Figura 4a ilustra un diagrama de bloques de una realización de un codificador;
- la Figura 4b ilustra un diagrama de bloques de una realización correspondiente de un decodificador;
- la Figura 5 ilustra una realización preferida de un codificador multicanal;
- 10 la Figura 6 ilustra un diagrama de bloques de una realización de un decodificador multicanal;
- la Figura 7a ilustra una realización adicional de un decodificador multicanal que comprende un combinador;
- 15 la Figura 7b ilustra una realización adicional de un decodificador multicanal que adicionalmente comprende el combinador (suma);
- la Figura 8a ilustra una ventana que ilustra diferentes características de ventana para varias velocidades de muestreo;
- 20 la Figura 8b ilustra diferentes propuestas/realizaciones para un banco de filtros DFT como una implementación del convertidor de tiempo-espectral y de convertidor espectral-tiempo;
- la Figura 8c ilustra una secuencia de dos ventanas de análisis de un DFT con una resolución de tiempo de 10 ms;
- 25 la Figura 9a ilustra una división en ventanas esquemática de codificador según una primera propuesta/realización;
- la Figura 9b ilustra una división en ventanas esquemática de decodificador según la primera propuesta/realización;
- 30 la Figura 9c ilustra las ventanas en el codificador y en el decodificador según la primera propuesta/realización;
- la Figura 9d ilustra un diagrama de flujo preferido que ilustra la realización de compensación;
- 35 la Figura 9e ilustra una realización que ilustra en mayor grado la realización de compensación;
- la Figura 9f ilustra un diagrama de flujo para explicar la realización del lado de decodificador del intervalo de tiempo;
- 40 la Figura 10a ilustra una división en ventanas esquemática de codificador según la cuarta propuesta/realización;
- la Figura 10b ilustra una división en ventanas esquemática de decodificador según la cuarta propuesta/realización;
- 45 la Figura 10c ilustra ventanas en el codificador y en el decodificador según la cuarta propuesta/realización;
- la Figura 11a ilustra una división en ventanas esquemática de un codificador según la quinta propuesta/realización;
- 50 la Figura 11b ilustra una división en ventanas esquemática de un decodificador según la quinta propuesta/realización;
- la Figura 11c ilustra ventanas en el codificador y el decodificador según la quinta propuesta/realización;
- 55 la Figura 12 es un diagrama de bloques de una implementación preferida del tratamiento multicanal que usa una submezcla en el procesador de señales;
- la Figura 13 es una realización preferida del tratamiento multicanal inverso con una operación de sobremezcla dentro del procesador de señales;
- 60 la Figura 14a ilustra un diagrama de flujo de procedimientos llevados a cabo en el aparato para codificar, con el fin de alinear los canales;
- 65 la Figura 14b ilustra una realización preferida de procedimientos llevados a cabo en el dominio de la frecuencia;

la Figura 14c ilustra una realización preferida de procedimientos llevados a cabo en el aparato para codificar usando una ventana de análisis con porciones de relleno con ceros e intervalos de superposición;

5 la Figura 14d ilustra un diagrama de flujo para otros procedimientos llevados a cabo dentro de una realización del aparato para codificar;

la Figura 15a ilustra procedimientos llevados a cabo mediante una realización del aparato para decodificar y codificar señales multicanal;

10 la Figura 15b ilustra una implementación preferida del aparato para decodificar con respecto a algunos aspectos; y

15 la Figura 15c ilustra un procedimiento llevado a cabo en el contexto de desalineación de banda ancha en el marco de la decodificación de una señal multicanal codificada.

[0033] La Figura 1 ilustra un aparato para codificar una señal multicanal que comprende al menos dos canales 1001, 1002. El primer canal 1001 es el canal izquierdo, y el segundo canal 1002 puede ser un canal 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 pueden 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 solo ejemplos, y es posible aplicar otros emparejamientos de canales según el caso.

25 **[0034]** El codificador multicanal de la Figura 1 comprende un convertidor de tiempo-espectral para convertir secuencias de bloques de valores de muestreo de los al menos dos canales en una representación en el dominio de la frecuencia a la salida del convertidor de tiempo-espectral. Cada representación en el dominio de la frecuencia tiene una secuencia de bloques de valores espectrales para uno de los al menos dos canales. En particular, un 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 de tiempo-espectral tiene valores espectrales hasta una frecuencia de entrada máxima que está relacionada con la velocidad de muestreo de entrada. En la realización ilustrada en la Figura 1, el convertidor de tiempo-espectral está conectado al procesador multicanal 1010. Este procesador multicanal está configurado para aplicar un tratamiento multicanal conjunto a las secuencias de bloques de valores espectrales con el fin de obtener al menos una secuencia de bloques de valores espectrales resultante que comprende información relacionada con los al menos dos canales. Una operación de tratamiento multicanal típica es una operación de submezcla, pero la operación multicanal preferida comprende procedimientos adicionales que serán descritos más adelante.

40 **[0035]** El codificador de núcleo 1040 está configurado para operar según un primer control de trama con el fin de proporcionar una secuencia de tramas, en el que una trama está delimitada por un límite de trama inicial 1901 y un límite de trama final 1902. El convertidor de tiempo-espectral 1000 o el convertidor espectral-tiempo 1030 están configurados para operar según un segundo control de trama que está sincronizado con el primer control de trama, en el que el límite de trama inicial 1901 o el límite de trama final 1902 de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con un instante inicial o un instante final de una porción superpuesta de una ventana usada por el convertidor de tiempo-espectral 1000 para cada bloque de la secuencia de bloques de valores de muestreo o usada por el convertidor espectral-tiempo 1030 para cada bloque de la secuencia de bloques de salida de valores de muestreo.

50 **[0036]** Tal como se ilustra en la Figura 1, el remuestreo en el dominio espectral es una característica opcional. La invención puede ejecutarse también sin ningún remuestreo, o con un remuestreo después de tratamiento multicanal o antes de tratamiento multicanal. En caso de uso, el remuestreador en el dominio espectral 1020 realiza una operación de remuestreo en el dominio de la frecuencia en los datos introducidos en el convertidor espectral-tiempo 1030 o en los datos introducidos en el procesador multicanal 1010, en el que un bloque de una secuencia de bloques remuestreada de valores espectrales tiene valores espectrales hasta una frecuencia de salida máxima 1231, 1221 que es diferente de la frecuencia de entrada máxima 1211. A continuación se describen realizaciones con remuestreo aunque debe insistirse en que el remuestreo es una característica opcional.

60 **[0037]** En una realización adicional, 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 realización alternativa, el procesador multicanal está configurado para aplicar el tratamiento multicanal conjunto no a las secuencias de bloques de valores espectrales como salida por el convertidor de tiempo-espectral, sino a secuencias remuestreadas de bloques disponibles en las líneas de conexión 1022.

65 **[0038]** El remuestreador en el dominio espectral 1020 está configurado para un remuestreo de la secuencia

resultante generada por el procesador multicanal o para remuestrear las secuencias de bloques emitidas por el convertidor de tiempo-espectral 1000 para obtener una secuencia remuestreada de bloques de valores espectrales que pueden representar una señal central tal como la ilustrada en la línea 1025. Preferentemente, el remuestreador en el dominio espectral lleva a cabo además el remuestreo hacia la señal lateral por medio del procesador multicanal y, por tanto, también emite una secuencia remuestreada correspondiente a la señal lateral como se ilustra en 1026. Sin embargo, la generación y remuestreo de la señal lateral es opcional y no se requiere para una implementación con una baja velocidad binaria. Preferentemente, el remuestreador en el dominio espectral 1020 está configurado para truncar bloques de valores espectrales para el submuestreo o para rellenar con ceros de los bloques de valores espectrales para el sobremuestreo. 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 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 realizaciones alternativas, en las que el remuestreo en el dominio espectral se lleva a cabo antes del tratamiento multicanal, el procesador multicanal proporciona la secuencia resultante por medio de la línea discontinua 1023 directamente al convertidor espectral-tiempo 1030. En esta realización alternativa, una característica opcional es que, adicionalmente, la señal lateral es generada por el procesador multicanal ya en la representación remuestreada y la señal lateral es tratada también a continuación por el convertidor espectral-tiempo.

[0039] Al final, el convertidor espectral-tiempo proporciona preferentemente una señal central en el dominio del tiempo 1031 y una señal lateral en el dominio del tiempo 1032 opcional, de manera que las dos pueden ser codificadas en núcleo por el codificador de núcleo 1040. En términos generales, el codificador de núcleo está configurado para una codificación de núcleo de la secuencia de bloques de valores de muestreo de salida para obtener la señal multicanal codificada.

[0040] La Figura 2 ilustra diagramas espectrales que son útiles para explicar el remuestreo en el dominio espectral.

[0041] El diagrama superior en la Figura 2 ilustra un espectro de canal disponible a la salida del convertidor de tiempo-espectral 1000. Este espectro 1210 tiene valores espectrales hasta la frecuencia de entrada máxima 1211. En el caso del sobremuestreo, se lleva a cabo un relleno con ceros dentro de la porción de relleno con ceros o de la región de relleno con ceros 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, ya que se pretende un sobremuestreo.

[0042] En cambio, el diagrama inferior de la Figura 2 ilustra los procedimientos en los que se incurre al submuestrear una secuencia de bloques. A tal efecto, se trunca un bloque dentro de una región truncada 1230 de manera que una frecuencia máxima de salida del espectro truncado en 1231 es inferior a la frecuencia de entrada máxima 1211.

[0043] Normalmente, la velocidad de muestreo asociada con un espectro correspondiente en la Figura 2 es de al menos 2x la frecuencia máxima del espectro. Por tanto, para el caso superior en la Figura 2, la velocidad de muestreo será al menos 2 veces la frecuencia de entrada máxima 1211.

[0044] En el segundo diagrama de la Figura 2, la velocidad de muestreo será al menos dos veces la frecuencia máxima de salida 1221, es decir, la frecuencia más elevada de la región de relleno con ceros 1220. A diferencia de lo anterior, en el diagrama inferior de la Figura 2, la velocidad de muestreo será al menos 2x la frecuencia de salida máxima 1231, es decir, el valor espectral más elevado restante posterior a un truncamiento dentro de la región truncada 1230.

[0045] Las Figuras 3a a 3c ilustran diversas alternativas que pueden usarse en el contexto de determinados 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 la que 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 un cambio de escala de N_y/N_x como se ilustra en el bloque 1321.

[0046] A diferencia de lo anterior, en la Figura 3b se ilustra una implementación, en la que la normalización se distribuye a la transformada directa 1312 y a la transformada inversa 1332. A continuación se requiere un cambio de escala como se ilustra en el bloque 1322, en el que es útil realizar la 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.

[0047] En la Figura 3c se ilustra una implementación adicional, en la que se lleva a cabo la normalización completa sobre la transformada directa en la que se realiza la transformada directa con la magnitud x . A continuación, la transformada inversa ilustrada en el bloque 1333 opera sin ninguna normalización por lo que no se requiere ningún cambio de escala como se ilustra mediante el bloque 1323 en la Figura 3c. Por tanto, en función de determinados algoritmos, se requieren determinadas operaciones de cambio de escala o incluso ninguna operación

de cambio de escala. Sin embargo, se prefiere operar según la Figura 3a.

[0048] Con el fin de mantener bajo el retardo global, la presente invención proporciona un procedimiento en el lado del codificador para evitar la necesidad de un remuestreador en el dominio del tiempo y se reemplazó por el remuestreo de las señales en el dominio de DFT. Por ejemplo, en el EVS permite ahorrar 0,9375 ms de retardo procedente del remuestreador en el dominio del tiempo. El remuestreo en el dominio de la frecuencia se logra mediante el relleno con ceros o el truncamiento del espectro y el cambio de escala correcto.

[0049] Considérese una señal dividida en ventanas 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 . En este caso el factor de muestreo es igual a:

$$f_y/f_x = N_y/N_x$$

en caso de submuestreo $N_x > N_y$. El submuestreo puede llevarse a cabo de manera sencilla en el dominio de la frecuencia mediante cambio de escala directo y truncamiento del espectro X original:

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

en caso del sobremuestreo $N_x < N_y$. El sobremuestreo puede llevarse a cabo simplemente en el dominio de la frecuencia mediante cambio de escala directo y relleno con ceros del espectro X original:

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

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

[0050] Las dos operaciones de remuestreo pueden resumirse como sigue:

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

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

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 :

$$y = iDFT(Y)$$

[0052] Para construir la señal de tiempo continua sobre diferentes tramas, se divide en ventanas la trama de salida y a continuación se superpone a la trama obtenida previamente.

[0053] La forma de ventana es igual para todas las velocidades de muestreo, pero las ventanas tienen diferentes tamaños en las muestras y se muestrean de manera diferente según las velocidades de muestreo. El número de muestras de las ventanas y sus valores pueden deducirse fácilmente dado que la forma se define de manera puramente analítica. Las diferentes partes y tamaños de la ventana pueden encontrarse en la Figura 8a como una función de la velocidad de muestreo objeto. En este caso se usa una función seno en la parte de superposición (LA) para las ventanas de análisis y síntesis. Para estas regiones, los coeficientes ascending $ovlp_size$ vienen 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 $ovlp_size$ vienen dados por:

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

en las que $ovlp_size$ es función de la velocidad de muestreo y se indica en la Figura 8a.

[0054] La nueva codificación estéreo de retardo bajo es una codificación estéreo central/lateral (M/S) conjunta que explota algunos vicios espaciales, en la que el canal central está codificado por un codificador mono núcleo primario, y el canal lateral está codificado por un codificador de núcleo secundario. Los principios de codificador y de decodificador se ilustran en las Figuras 4a y 4b.

[0055] El tratamiento estéreo se lleva a cabo principalmente en el dominio de la frecuencia (FD, frequency-domain). Opcionalmente parte del tratamiento estéreo puede llevarse a cabo en el dominio del tiempo (TD, time-domain) antes del análisis de frecuencia. Este es el caso para el cálculo de ITD, que puede calcularse y aplicarse antes del análisis de frecuencia para alinear los canales en el tiempo antes de continuar con el análisis y tratamiento

estéreo. Como alternativa, el tratamiento ITD puede llevarse a cabo directamente en el dominio de la frecuencia. Como 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 adicional 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 realización preferida, se usa un DFT sobremuestreado con una región de baja superposición. Sin embargo, en otras realizaciones, puede usarse cualquier descomposición de valor complejo de tiempo-frecuencia con una resolución temporal compleja. En lo que sigue con respecto a la banda de filtros estéreo, se hace referencia a un banco de filtros tal como QMF o a un bloque de transformada, DFT.

10 **[0056]** El tratamiento estéreo consiste en calcular los indicios espaciales y/o los parámetros estéreo tales como la ITD (diferencia de tiempo entre canales), las IPD (diferencias de fase entre canales), las ILD (diferencia de niveles entre canales) y las ganancias de predicción para predecir la señal lateral (S) con la señal central (M). Es importante observar que el banco de filtros estéreo tanto en el codificador como en el decodificador introduce un retardo adicional en el sistema de codificación.

15 **[0057]** En la Figura 4a se ilustra un aparato para codificar una señal multicanal en la que, en esta implementación, se lleva a cabo determinado tratamiento estéreo conjunto en el dominio del tiempo usando un análisis de diferencia de tiempo entre canales (ITD) y en el que el resultado de este análisis de ITD 1420 se aplica dentro del dominio del tiempo usando un bloque de desplazamiento de tiempo 1410 colocado delante de los convertidores tiempo-espectral 1000.

25 **[0058]** A continuación, dentro del dominio espectral, se lleva a cabo un tratamiento estéreo adicional 1010 que implica al menos una submezcla de izquierda y derecha de la señal central M, y, opcionalmente, el cálculo de una señal lateral S y, si bien no se ilustra explícitamente en la Figura 4a, una operación de remuestreo 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 remuestreo después del tratamiento multicanal o antes del tratamiento multicanal.

30 **[0059]** Además, en la Figura 4a se ilustran detalles adicionales de un codificador de núcleo preferido 1040. En particular, con el fin de codificar la señal central en el dominio del tiempo m a la salida del convertidor espectral-tiempo 1030, se usa un codificador EVS. Adicionalmente, se lleva a cabo una codificación MDCT 1440 y la cuantificación vectorial 1450 conectada posteriormente para la codificación de la señal lateral.

35 **[0060]** La señal central, codificada o con codificación de núcleo, y la señal lateral, con codificación de núcleo, son comunicadas a un multiplexor 1500 que multiplexa estas señales codificadas juntamente con información lateral. Un tipo de información lateral es la salida del parámetro ID en 1421 al multiplexor (y opcionalmente al elemento de tratamiento 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 relleno de estéreo como se ilustra en la línea 1422. De manera correspondiente, el aparato de la Figura 4b para decodificar una señal multicanal representada por un tren de bits 1510 comprende un demultiplexor 1520, un decodificador de núcleo que en esta realización consiste en un decodificador EVS 1602 para la señal central codificada y en un cuantificador del vector 1603 y un bloque MDCT inverso conectado posteriormente 1604. El bloque 1604 proporciona la señal lateral con decodificación de núcleo. Las señales decodificadas m, s son convertidas a continuación al dominio espectral usando los convertidores tiempo-espectral 1610, y, después, dentro del dominio espectral, se lleva a cabo el tratamiento estéreo inverso y el remuestreo. Nuevamente, en la Figura 4b se ilustra una situación en la que se lleva a cabo la sobremezcla de la señal M a la izquierda L y la derecha R, y, adicionalmente, una desalineación de banda estrecha usando parámetros IPD y, adicionalmente, otros procedimientos para calcular en la mejor medida posible los canales izquierdo y derecho usando los parámetros de diferencia de nivel entre canales ILD y los parámetros de relleno de estéreo en la línea 1605. Además, el demultiplexor 1520 no solo extrae los parámetros en la línea 1605 del tren de bits 1510, sino que también extrae la diferencia de tiempo entre canales en la línea 1606 y comunica esta información al tratamiento estéreo inverso del bloque/remuestreador y, adicionalmente, a un tratamiento del desplazamiento de tiempo inverso en el bloque 1650 que se lleva a cabo en el dominio del tiempo, es decir, después del 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 a la salida del decodificador EVS 1602 o diferente de la velocidad del bloque IMDCT 1604, por ejemplo.

60 **[0061]** A continuación, el DFT estéreo puede proporcionar diferentes versiones muestreadas de la señal que después es comunicada al codificador de núcleo conmutado. La señal para codificar puede ser el canal central, el canal lateral o los canales izquierdo y derecho, o cualquier señal resultante de una rotación o correspondencia de canales de los dos canales de entrada. Como los diferentes codificadores de núcleo del sistema conmutado aceptan diferentes velocidades de muestreo, es una característica importante que el banco de filtros para la síntesis estéreo pueda proporcionar una señal de múltiples velocidades. El principio se indica en la Figura 5.

[0062] 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 la frecuencia en las señales M y S. En el tratamiento estéreo puede establecerse

finalmente una correspondencia entre los canales de entrada o modificarse de manera que se generen dos nuevas señales M y S. M se somete a mayor codificación según la norma 3GPP EVS mono o por una versión modificada de la misma. 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 pretratamiento que se ejecuta en todo momento a 12,8 kHz y otras funciones de pretratamiento 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 a 12,8 o 16 kHz, mientras que los núcleos MDCT se ejecutan a la velocidad de muestreo de entrada. La señal S puede ser codificada por un codificador estándar EVS mono (o por una versión modificada del mismo), o por un codificador de señal lateral específico especialmente diseñado por sus características. También es posible omitir la codificación de la señal lateral S.

[0063] En la Figura 5 se ilustran detalles del codificador estéreo preferido con banco de filtros de síntesis de múltiples velocidades de las señales estéreo tratadas M y S. La Figura 5 muestra el convertidor de 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 además un bloque de análisis en el dominio del tiempo 1000a, 1000e, para cada canal. En particular, si bien la Figura 5 ilustra un bloque explícito de análisis en el dominio del tiempo, es decir, una división en ventanas para aplicar una ventana de análisis al canal correspondiente, cabe observar que en otros lugares de la presente memoria descriptiva se considera que el aspecto para la división en ventanas para aplicar el bloque de análisis en el dominio del tiempo está incluido en un bloque indicado como “convertidor de tiempo-espectral” o “DFT” con alguna velocidad de muestreo. Además, y de manera correspondiente, la mención de un convertidor espectral-tiempo normalmente incluye, en salida del algoritmo DFT real, un aspecto para la división en ventanas con el fin de aplicar una ventana de síntesis correspondiente en la que, para obtener finalmente muestras de salida, se lleva a cabo una superposición-suma de bloques de valores de muestreo divididos en ventanas con una ventana de síntesis correspondiente.

[0064] Por tanto, aun cuando el bloque 1030 solo menciona un “IDFT”, este bloque también designa una división en ventanas subsiguiente de un bloque de muestras en el dominio del tiempo con una ventana de análisis y nuevamente, una operación subsiguiente de superposición-suma para obtener finalmente la señal en el dominio del tiempo.

[0065] Además, la Figura 5 ilustra un bloque de análisis de escena estéreo específico 1011 que usa los parámetros usados en el bloque 1010 para llevar a cabo el tratamiento estéreo y de sobremezcla, y estos parámetros pueden, por ejemplo, ser los parámetros en las líneas 1422 o 1421 de la Figura 4a. Por 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, especialmente, con la secuencia de bloques de valores espectrales que no han sido remuestreadas, sino que se encuentran a la frecuencia máxima correspondiente a la velocidad de muestreo de entrada.

[0066] Además, el decodificador de núcleo 1040 comprende una rama de codificador basada en MDCT 1430a y una rama de codificación de ACELP 1430b. En particular, el codificador central para las señales centrales M y el correspondiente codificador lateral para las señales laterales lleva a cabo una codificación de conmutación entre una codificación basada en MDCT y una codificación ACELP en la que, normalmente, el codificador de núcleo tiene además un aspecto de decisión en modo de codificación que normalmente opera en una determinada porción de anticipación con el fin de determinar si un cierto bloque o trama ha de ser codificado usando procedimientos basados en MDCT o procedimientos basados en ACELP. Además, o como alternativa, el codificador de núcleo está configurado para usar la porción de anticipación con el fin de determinar otras características tales como parámetros de LPC, etc.

[0067] Además, el codificador de núcleo comprende además etapas de pretratamiento a diferentes velocidades de muestreo tales como una primera etapa de pretratamiento 1430c que opera a 12,8 kHz y una etapa de pretratamiento adicional 1430d que opera con velocidades de muestreo del grupo de las velocidades de muestreo que consiste en 16 kHz, 25,6 kHz o 32 kHz.

[0068] Por tanto, en términos generales, la realización ilustrada en la Figura 5 está configurada de manera que tiene un remuestreador en el dominio espectral para remuestreo, 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.

[0069] Además, la realización en la Figura 5 está configurada adicionalmente para tener una rama adicional que no ha sido remuestreada, es decir, la rama indicada como “IDFT con velocidad de entrada” para la señal central y, opcionalmente, para la señal lateral.

[0070] Además, el codificador en la Figura 5 comprende preferentemente un remuestreador que no solo 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 los dos preprocesadores 1430c y 1430d que, por ejemplo, pueden estar operativos para llevar a cabo algún tipo de filtrado, algún tipo de cálculo de LPC o algún tipo de otro tratamiento de

señales que se describe preferentemente en la norma 3GPP para el codificador EVS ya mencionado en el contexto de la Figura 4a.

5 **[0071]** La Figura 6 ilustra una realización para un aparato destinado a decodificar una señal multicanal codificada 1601. El aparato para decodificar comprende un decodificador de núcleo 1600, un convertidor de tiempo-espectral 1610, un remuestreador en el dominio espectral 1620, un procesador multicanal 1630 y un convertidor espectral-tiempo 1640.

10 **[0072]** 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 límite de trama inicial 1901 y un límite de trama final 1902. El convertidor de tiempo-espectral 1610 o el convertidor espectral-tiempo 1640 están configurados para operar según un segundo control de trama que está sincronizado con el primer control de trama. El convertidor de tiempo-espectral 1610 o el convertidor espectral-tiempo 1640 están configurados para operar según un segundo control de trama que está sincronizado con el primer control de trama, en el que el límite de trama
15 inicial 1901 o el límite de trama final 1902 de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con un instante inicial o un instante final de una porción superpuesta de una ventana usada por el convertidor de tiempo-espectral 1610 para cada bloque de la secuencia de bloques de valores de muestreo o usada por el convertidor espectral-tiempo 1640 para cada bloque de las al menos dos secuencias de bloques de salida de valores de muestreo.

20 **[0073]** Nuevamente, la invención con respecto al aparato para decodificar la señal multicanal codificada 1601 puede implementarse en varias alternativas. Una alternativa es no usar en absoluto el remuestreador en el dominio espectral. Otra alternativa consiste en que se usa un remuestreador y se configura para remuestrear la señal con decodificación de núcleo en el dominio espectral antes de llevar a cabo el tratamiento multicanal. Esta alternativa se
25 ilustra mediante líneas continuas en la Figura 6. Sin embargo, la alternativa adicional es que el remuestreo en el dominio espectral se lleve a cabo después del tratamiento multicanal, es decir, que el tratamiento multicanal tenga lugar con la velocidad de muestreo de entrada. Esta realización se ilustra en la Figura 6 mediante líneas discontinuas. Si se usa, el remuestreador en el dominio espectral 1620 lleva a cabo la operación de remuestreo en el dominio de la frecuencia en los datos introducidos en el convertidor espectral-tiempo 1640 o en los datos
30 introducidos en el procesador multicanal 1630, 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.

[0074] Especialmente, en la primera realización, es decir, en la que el remuestreo en el dominio espectral se lleva a cabo en el dominio espectral antes del tratamiento multicanal, la señal decodificada de núcleo que representa
35 una secuencia de bloques de valores de muestreo se convierte en una representación en el dominio de la frecuencia que tiene una secuencia de bloques de valores espectrales para la señal con decodificación de núcleo, en la línea 1611.

[0075] Adicionalmente, la señal decodificada de núcleo no solo comprende la señal M en la línea 1602, sino
40 también una señal lateral en la línea 1603, en la que se ilustra una señal lateral en 1604 en una representación con codificación de núcleo.

[0076] A continuación, el convertidor de tiempo-espectral 1610 genera adicionalmente una secuencia de bloques de valores espectrales para la señal lateral en la línea 1612.
45

[0077] A continuación, se lleva a cabo un remuestreo en el dominio espectral por el bloque 1620, y la secuencia remuestreada de bloques de valores espectrales con respecto a la señal central o canal de submezcla 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 desde el remuestreador en el
50 dominio espectral 1620 al procesador multicanal 1630 por medio de la línea 1622.

[0078] A continuación, el procesador multicanal 1630 lleva a cabo un tratamiento multicanal inverso con respecto a una secuencia que comprende una secuencia de la señal de submezcla y, opcionalmente, de la señal lateral ilustrada en las líneas 1621 y 1622 con el fin de emitir al menos dos secuencias resultantes de bloques de
55 valores espectrales ilustrados en 1631 y 1632. Estas al menos dos secuencias son convertidas a continuación en el dominio del tiempo usando el 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 de tiempo-espectral está configurado para introducir la señal con decodificación de núcleo tal como la señal central en el procesador multicanal. Adicionalmente, el convertidor de tiempo-espectral también puede conducir una señal lateral decodificada 1603 en su representación en el dominio espectral hacia el procesador multicanal 1630, si bien esta
60 opción no se ha representado en la Figura 6. A continuación, el procesador multicanal lleva a cabo el tratamiento 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 a continuación comunica lo remuestreado en estos dos canales por medio de la línea 1625 hacia el convertidor espectral-tiempo 1640.

65

[0079] Por tanto, un poco como analogía a lo que se ha expuesto en el contexto de la Figura 1, el aparato para decodificar una señal multicanal codificada también comprende dos alternativas, una en la que el remuestreo en el dominio espectral se lleva a cabo antes del tratamiento multicanal inverso o, como alternativa, una en la que remuestreo en el dominio espectral se lleva a cabo después del tratamiento multicanal a la velocidad de muestreo de entrada. Sin embargo, preferentemente se lleva 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.

[0080] Nuevamente, 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 codificado de núcleo 1601 y el decodificador de núcleo adicionalmente genera una tercera señal de salida 1603 operativa y disponible a la velocidad de muestreo de salida, es decir, la velocidad de muestreo finalmente prevista a la salida del convertidor espectral-tiempo 1640 en la Figura 7a.

[0081] La totalidad de las tres señales con decodificación de núcleo son introducidas en el convertidor de tiempo-espectral 1610 que genera tres secuencias diferentes de bloques de valores espectrales 1613, 1611 y 1612.

[0082] La secuencia de bloques de valores espectrales 1613 tiene valores de frecuencia o espectrales de hasta la máxima frecuencia de salida y, por tanto, está asociada con la velocidad de muestreo de salida.

[0083] La secuencia de bloques de valores espectrales 1611 tiene valores espectrales de hasta una máxima frecuencia diferente y, por tanto, esta señal no corresponde a la velocidad de muestreo de salida.

[0084] Además, los valores espectrales de la señal y 602 de hasta la máxima frecuencia introducida que también es diferente de la frecuencia de salida máxima.

[0085] Por tanto, se comunican las secuencias 612 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.

[0086] 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 se corresponden en situaciones superpuestas. Por tanto, normalmente 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, cuando se supera este intervalo de superposición, existe, por ejemplo, solo una señal en la señal 1603 mientras que la señal 1602, por ejemplo, no existe, y así el combinador no llevará a cabo una suma de línea espectral bloque por bloque en esta porción. Sin embargo, si posteriormente se presenta un cruce, en tal caso tendrá lugar una línea espectral bloque por bloque mediante suma de líneas espectrales durante esta región de cruce.

[0087] Además, también puede ser posible una suma continua como se ilustra en la Figura 7b, en la que se lleva a cabo una señal de salida bajo-posfiltro en el bloque 1600a, que genera una señal de error entre armónicos que podría ser por ejemplo la señal 1601 de la Figura 7a. A continuación, después de la conversión tiempo-espectral en el bloque 1610, y del remuestreo subsiguiente en el dominio espectral 1620, se lleva preferentemente a cabo una operación de filtrado adicional 1702 antes de realizar la suma en el bloque 1700 en la Figura 7b.

[0088] De manera similar, la etapa de decodificación basada en MDCT 1600d y la etapa de decodificación de extensión de la anchura 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 a continuación es convertida en la representación en el dominio espectral a la velocidad de muestreo de salida de manera que para esta señal 1613, el remuestreo en el dominio espectral no es necesario, sino que la señal puede ser comunicada directamente al combinador 1700. El tratamiento estéreo inverso o el tratamiento multicanal 1603 tienen lugar así después del combinador 1700.

[0089] Por tanto, a diferencia de la 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 en la que la secuencia, sobre la cual opera el procesador multicanal 1630, comprende además la secuencia 1613 cuyo remuestreo no fue necesario.

[0090] Como se ilustra en la Figura 7a-7b, las diferentes señales decodificadas procedentes de diferentes DFT que trabajan a 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 cambios de escala. Para armonizarlos y hacerlos compatibles, todos los espectros son

remuestreados en el dominio de la frecuencia con la velocidad de muestreo de salida deseada antes de sumarse entre sí.

[0091] Por tanto, en la Figura 7a-7b se ilustra la combinación de diferentes contribuciones de una señal sintetizada en el dominio DFT, en la que el remuestreo en el dominio espectral se lleva a cabo de una manera tal que, al final, todas las señales que deben ser sumadas 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 se obtiene así a la salida del convertidor espectral tiempo 1640.

[0092] La elección del banco de filtros estéreo es esencial para un sistema de retardo bajo, y el compromiso obtenible se ha resumido en la Figura 8b. Puede usar una DFT (transformada de bloque) o un QMF de bajo seudoretardo llamado CLDFB (banco de filtros). Cada propuesta muestra un retardo, un tiempo y unas 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 el uso de un banco de filtros de pseudo-QMF como en la propuesta 3 puede ser problemático. La resolución de frecuencia es baja. Se puede mejorar mediante enfoques híbridos como en MPS 212 de MPEG-USAC, pero tiene el inconveniente de que incrementa 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 tratamiento estéreo inverso. Cuanto mayor sea este retardo, tanto mejor. Por ejemplo, la propuesta 2 no puede proporcionar este retardo, por lo que no es una solución útil. Por las razones mencionadas anteriormente, en el resto de la descripción nos centraremos en las propuestas 1, 4 y 5.

[0093] La ventana de análisis y síntesis del banco de filtros es otro aspecto importante. En la realización preferida se usa la misma ventana para el análisis y la síntesis de la DFT. También es la misma en los lados de codificador y decodificador. Se prestó especial atención a cumplir las siguientes restricciones:

- La región de superposición debe ser igual o menor que la región de superposición del núcleo de MDCT y la anticipación de ACELP. En la realización preferida todas las magnitudes son iguales a 8,75 ms.
- El relleno con ceros 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 relleno con ceros han de 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 raíz máxima de la DFT en una implementación de FFT de tipo raíz dividida debería ser lo más baja posible.
- La resolución de tiempo se fija en 10 ms.

[0094] A partir de estas restricciones, las ventanas para las propuestas 1 y 4 se describen en la Figura 8c y en la Figura 8a.

[0095] La Figura 8c ilustra una primera ventana consistente en una porción superpuesta inicial 1801, una porción central posterior 1803 y una porción terminal superpuesta o una segunda porción superpuesta 1802. Por otra parte, la primera porción superpuesta 1801 y la segunda porción superpuesta 1802 tienen además una porción de relleno con ceros de 1804 en el inicio y 1805 en el final de la misma.

[0096] Por otra parte, en la Figura 8c se ilustra el procedimiento realizado con respecto al entramado del convertidor de 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 relleno con ceros 1814 y 1815 en el inicio y el final de la misma. Estas porciones de relleno con ceros son necesarias con el fin de estar en posición de llevar a cabo la alineación de tiempo de la anchura de banda en el dominio de la frecuencia.

[0097] 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 la 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.

[0098] Si se considera que en la Figura 8c se representa una operación de superposición-suma 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 consistente en los bloques

- 1801, 1802, 1803, 1805, 1804 corresponde a una ventana de síntesis y la segunda ventana consistente 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 realización preferida, igual a 5 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.
- 10 **[0099]** En realizaciones preferidas, se usa la misma ventana de análisis y síntesis solo para el decodificador ilustrado en la Figura 6, la Figura 7a y la Figura 7b. Por tanto, el convertidor de tiempo-espectral 1616 y el convertidor espectral-tiempo 1640 usan exactamente las mismas ventanas que las ilustradas en la Figura 8c.
- 15 **[0100]** Sin embargo, en determinadas realizaciones especialmente con respecto a la subsiguiente propuesta/realización 1, se usa una ventana de análisis que en términos generales concuerda con la Figura 8c, pero los coeficientes de ventana para las posiciones de superposición ascendente o descendente se calculan usando una raíz cuadrada de función seno, con el mismo argumento en la función seno que en la Figura 8c. De manera correspondiente, la venta de síntesis se calcula usando una función seno a la potencia 1,5, pero nuevamente con el mismo argumento que la función seno.
- 20 **[0101]** Además, cabe observar que debido a la operación de superposición-suma, la multiplicación de seno a la potencia 0,5 multiplicado por seno a la potencia 1,5 produce de nuevo un seno a la potencia 2, un resultado que es necesario para tener una situación de conservación de energía.
- 25 **[0102]** Las características principales de la propuesta 1 consisten en 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 núcleos de ACELP/MDCT y el estéreo no introduce ningún retardo adicional en el codificador. En el caso de EVS y si se usa el enfoque del banco de filtros de síntesis de múltiples velocidades descrito en la Figura 5, el retardo del codificador estéreo es bajo, de 30 aproximadamente 8,75 ms.
- [0103]** El entramado esquemático 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 en línea discontinua para el codificador y en líneas rojas continuas para el decodificador.
- 35 **[0104]** Un asunto importante en el caso de la propuesta 1 es que la anticipación en el codificador está dividida en ventanas. Se puede compensar para el tratamiento subsiguiente, o se puede dejar dividida en ventanas si el tratamiento subsiguiente está adoptado para tener en cuenta una anticipación dividida en ventanas. Podría darse el caso de que si el tratamiento estéreo llevado a cabo en la DFT modificó el canal de entrada, y 40 especialmente cuando se usan operaciones no lineales, la señal compensada o dividida en ventanas no permita lograr una reconstrucción perfecta si se omite la codificación de núcleo.
- [0105]** Cabe observar que entre la ventana de síntesis del decodificador de núcleo y la ventana de análisis del decodificador estéreo hay un intervalo de tiempo de 1,25 ms que puede ser aprovechado por el decodificador de 45 núcleo después del tratamiento, por la extensión de la anchura de banda (BWE, bandwidth extension), como BWE del dominio del tiempo usado sobre ACELP, o mediante alguna suavización en el caso de transición entre núcleos de ACELP y MDCT.
- [0106]** Dado que este intervalo de tiempo de solo 1,25 ms es inferior a los 2,3125 ms requeridos por la norma 50 EVS para dichas operaciones, la presente invención proporciona una manera de combinar, remuestrear y suavizar las diferentes partes de síntesis del descodificador conmutado dentro del dominio de DFT del módulo estéreo.
- [0107]** Como se ilustra en la Figura 9a, el codificador de núcleo 1040 está configurado para operar según un control de trama para proporcionar una secuencia de tramas, en la que una trama está delimitada por un límite inicial 55 de la trama 1901 y un límite final de trama 1902. Además, el convertidor de tiempo-espectral 1000 y/o el convertidor espectral-tiempo 1030 también están configurados para operar según un segundo control de trama que está sincronizado con el primer control de trama. El control de trama se ilustra con dos ventanas superpuestas 1903 y 1904 para el convertidor de tiempo-espectral 1000 en el codificador, y, especialmente, para el primer canal 1001 y el segundo canal 1002 que son tratados simultáneamente y completamente sincronizados. Además, el control de 60 trama también es visible en el lado del decodificador, específicamente, con dos ventanas propuestas para el convertidor de 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 submezcla 1610 de la Figura 6, por ejemplo. Además, como puede observarse en la Figura 9a, la sincronización entre el control de trama del codificador de núcleo 1040 y el convertidor de tiempo-espectral 1000 o el convertidor espectral-tiempo 65 1030 es tal que el límite inicial de la trama 1901 o el límite final de trama 1902 de cada trama de la secuencia de

tramas se encuentra en una relación predeterminada con respecto a una instancia inicial y la instancia final de una porción superpuesta de una ventana usada por el convertidor de tiempo-espectral 1000 o por el convertidor espectral-tiempo 1030 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. En la realización ilustrada en la Figura 9a, la relación

5 predeterminada es tal que el inicio de la primera porción superpuesta coincide con el límite de tiempo inicial 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 tanto, el límite de trama final 1902 coincide con el extremo de la parte central 1813 de la Figura 8c, cuando la segunda ventana en la Figura 8c corresponde a la ventana 1904 en la Figura 9a.

10

[0108] Por 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 límite de trama inicial o de detención 1902, y, por tanto, se extiende en la porción de anticipación del codificador de núcleo ilustrada en 1905.

15 **[0109]** Por tanto, el codificador de núcleo 1040 está configurado para usar una porción de anticipación tal como la porción de anticipación 1905 cuando se codifique de núcleo el bloque de salida de la secuencia de salida de los bloques de muestreo, en la que la porción de salida de anticipación está situada en el tiempo posterior al bloque de salida. El bloque de salida se corresponde con la trama delimitada por los límites de trama 1901, 1904 y la porción de salida de anticipación 1905 se presenta después de este bloque de salida para el codificador de núcleo

20 1040.

[0110] Además, como se ilustra, el convertidor de tiempo-espectral está configurado para usar una ventana de análisis, es decir, la ventana 1904 que tiene la porción de superposición con una longitud de tiempo inferior o igual a la longitud de tiempo de la porción de anticipación 1905, en la que esta porción superpuesta correspondiente

25 a la superposición 1812 de la Figura 8c que está situada en el intervalo de superposición se usa para generar la porción de anticipación dividida en ventanas.

[0111] Además, el convertidor espectral-tiempo 1030 está configurado para procesar la porción de anticipación de salida correspondiente a la porción de anticipación dividida en ventanas preferentemente usando una

30 función de compensación, en la que la función de compensación está configurada de manera que se reduzca o se elimine una influencia de la porción de superposición de la ventana de análisis.

[0112] Por tanto, el convertidor espectral-tiempo que opera entre el codificador de núcleo 1040 y el bloque de submezcla 1010/bloque de submuestreo 1020 en la Figura 9a está configurado para aplicar una función de

35 compensación con el fin de deshacer la división en ventanas aplicada por la ventana 1904 en la Figura 9a.

[0113] Por 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 cerca posible de la porción original.

40

[0114] Sin embargo, debido a las restricciones de retardo bajo, y debido a la sincronización entre el entramado del preprocesador 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 que cualquier artefacto ocasionado por este procedimiento se reduce todo lo posible.

45

[0115] En la Figura 9d y 9e se ilustra con mayor detalle una secuencia de procedimientos con respecto a esta tecnología.

[0116] En la etapa 1910, se lleva a cabo una a DFT⁻¹ de un bloque 0-ésimo^o de manera que se obtenga un

50 bloque 0-ésimo^o en el dominio del tiempo. El bloque 0-ésimo^o podría haber obtenido una ventana usada a la izquierda de la ventana 1903 en la Figura 9a. Sin embargo, este bloque 0-ésimo^o, no se ilustra explícitamente en la Figura 9a.

[0117] A continuación, en la etapa 1912, el bloque 0-ésimo^o se divide en ventanas usando una ventana de

55 síntesis, es decir, se divide en ventanas en el convertidor espectral-tiempo 1030 ilustrado en la Figura 1.

[0118] A continuación, 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 divide en ventanas nuevamente usando la ventana de síntesis en el bloque 1910.

60

[0119] A continuación, 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, a continuación la primera porción del segundo bloque es dividida en ventanas usando la ventana de síntesis como se ilustra en 1920 de la Figura 9d. Sin embargo, lo importante es que la segunda porción

65 del segundo bloque obtenido por el elemento 1918 en la Figura 9d no se divide en ventanas usando 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 usa la inversa de la función de ventana de análisis y la correspondiente porción superpuesta de función de la ventana de análisis.

5 **[0120]** Por tanto, si la ventana usada para generar el segundo bloque fue una ventana seno ilustrada en la Figura 8c, en tal caso se usa $1/\text{sen}()$ para los coeficientes de tamaño de superposición descendentes de las ecuaciones en la parte inferior de la Figura 8c, como función de compensación.

[0121] Sin embargo, se prefiere usar una raíz cuadrada de ventana seno para la ventana de análisis y, por

10 tanto, la función de compensación es una función de ventana de $\frac{1}{\sqrt{\text{sen}()}}$. Así se asegura que la porción de

anticipación compensada 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 de la suma de izquierda y derecha para obtener la señal central.

15 **[0122]** A continuación, en la etapa 1924 en la Figura 9d, se genera una trama indicada mediante los límites de trama 1901,1902 llevando a cabo una operación de superposición-suma en el bloque 1030 de manera que el codificador tenga una señal en el dominio del tiempo, y esta trama se lleva a cabo mediante una operación de superposición-suma entre el bloque correspondiente a la ventana 1903, y las muestras anteriores del bloque anterior y usando la primera porción del segundo bloque obtenido por el bloque 1920. A continuación, esta trama emitida por
20 el bloque 1924 se comunica al codificador de núcleo 1040 y, adicionalmente, el codificador de núcleo recibe la porción de anticipación compensada para la trama y, como se ilustra en la etapa 1926, el codificador de núcleo puede determinar a continuación la característica para el codificador de núcleo usando la porción de anticipación compensada obtenida por la etapa 1922. A continuación, como se ilustra en la etapa 1928, el codificador de núcleo
25 codifica de núcleo la trama usando la característica determinada en el bloque 1926 de manera que finalmente se obtenga la trama, con codificación de núcleo, correspondiente al límite de trama 1901, 1902 que tiene, en la realización preferida, una longitud de 20 ms.

[0123] Preferentemente, la porción de superposición de la ventana 1904 que se extiende en la porción de anticipación 1905 tiene la misma longitud que en la porción de anticipación, pero también puede ser más corta que
30 la porción de anticipación pero se prefiere que no sea más larga que la porción de anticipación de manera que el preprocesador estéreo no introduzca ningún retardo adicional debido a las ventanas superpuestas.

[0124] A continuación, el procedimiento continúa con la división en ventanas de la segunda porción del segundo bloque usando la ventana de síntesis ilustrada en el bloque 1930. Por tanto, la segunda porción del
35 segundo bloque es, por una parte, compensada por el bloque 1922 y por otra parte es dividida en ventanas por la 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/suma de la segunda porción dividida en ventanas del segundo bloque, un tercer bloque dividido en ventanas 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
40 más sometido a la operación de compensación como se expuso con respecto al segundo bloque en el elemento 1922 de la Figura 9d y, a continuación, se repetiría una vez más el procedimiento como se expuso en lo que precede. Además, en la etapa 1934, el codificador de núcleo determinaría las características del codificador de núcleo usado para compensar la segunda porción del cuarto bloque y, a continuación, la siguiente trama se codificaría usando las características de codificación determinadas con el fin de obtener finalmente la siguiente
45 trama codificada en el bloque 1934. Por 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 dividida en ventanas es compensada, por una parte, por 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
50 sino que aplica una influencia más pequeña, de manera que puede asegurarse que la función de compensación es más estable en comparación con el uso 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 normalmente es necesaria para determinar las características de codificación del núcleo en una porción dividida en ventanas, no es necesaria para llevar a cabo la función de compensación. Sin embargo, se ha descubierto que el uso de la función de
55 compensación es ventajoso en comparación con la modificación del codificador de núcleo.

[0125] Además, como se expuso en lo que precede, cabe observar que hay un intervalo de tiempo entre el fin de una ventana, es decir, la ventana de análisis 1914 y el límite de trama 1902 de la trama definida por el límite inicial de la trama 1901 y el límite final de la trama 1902 de la Figura 9b.

60

[0126] En particular, el intervalo de tiempo se ilustra en 1920 con respecto a las ventanas de análisis aplicadas por el convertidor de tiempo-espectral 1610 de la Figura 6, y este intervalo de tiempo también es visible

120 con respecto al primer canal de salida 1641 y el segundo canal de salida 1642.

[0127] La Figura 9f muestra un procedimiento de las etapas llevadas a cabo en el contexto del intervalo de tiempo, el decodificador de núcleo 1600 decodifica en núcleo la trama o al menos la porción inicial de la trama hasta el intervalo de tiempo 1920. A continuación, el convertidor de tiempo-espectral 1610 de la Figura 6 está configurado para aplicar una ventana de análisis a la porción inicial de la trama usando 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 solo se extiende hasta el inicio del intervalo de tiempo 1920.

[0128] Por tanto, el decodificador de núcleo tiene tiempo adicional para decodificar en núcleo las muestras en el intervalo de tiempo y/o para posprocesar las muestras en el intervalo de tiempo como se ilustra en el bloque 1940. Por tanto, el convertidor de tiempo-espectral 1610 ya emite un primer bloque como resultado de la etapa 1938 cuando el decodificador de núcleo puede proporcionar las muestras permanentes en el intervalo de tiempo y puede posprocesar las muestras en el intervalo de tiempo en la etapa 1940.

[0129] A continuación, en la etapa 1942, el convertidor de tiempo-espectral 1610 está configurado para dividir en ventanas las muestras en el intervalo de tiempo junto con muestras de la trama siguiente usando una siguiente ventana de análisis que se presentaría después de la ventana 1914 en la Figura 9b. A continuación, como se ilustra en la etapa 1944, el decodificador de núcleo 1600 está configurado para decodificar la siguiente trama o al menos la porción inicial de la trama siguiente hasta que tenga lugar el intervalo de tiempo 1920 en la siguiente trama. A continuación, en la etapa 1946, el convertidor de tiempo-espectral 1610 está configurado para dividir en ventanas las muestras en la siguiente trama hasta el intervalo 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 intervalo de tiempo de la trama siguiente y/o posprocesa estas muestras.

[0130] Por tanto, este intervalo de tiempo es de, por ejemplo, 1,25 ms cuando se considere que la realización de la Figura 9b puede ser aprovechada para el postratamiento del decodificador de núcleo, por la extensión de la anchura de banda, mediante, por ejemplo, una extensión de la anchura de banda en el dominio del tiempo usada en el contexto de ACELP, o mediante alguna suavización en el caso de una transición de la transmisión entre señales de núcleo ACELP y MDCT.

[0131] Por tanto, una vez más, el decodificador de núcleo 1600 está configurado para operar según un primer control del entramado para proporcionar una secuencia de tramas, en la que el convertidor de tiempo-espectral 1610 o el convertidor espectral-tiempo 1640 están configurados para operar según un segundo control de trama que está sincronizado con el primer control de trama, de manera que el límite inicial de trama o el límite final de trama de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con respecto a un instante inicial o con respecto a un instante final de una porción superpuesta de una ventana usada por el convertidor de tiempo-espectral o por el convertidor espectral-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.

[0132] Además, el convertidor de tiempo-espectral 1610 está configurado para usar una ventana de análisis para dividir en ventanas la trama de la secuencia de tramas que tiene un intervalo de superposición que termina antes del límite de trama final 1902 quedando un intervalo de tiempo 1920 entre el fin de la porción de superposición y el límite final de trama. El decodificador de núcleo 1600 está, por tanto, configurado para llevar a cabo el tratamiento de las muestras en el intervalo de tiempo 1920 en paralelo con la división en ventanas de la trama usando la ventana de análisis o en la que se lleva a cabo otro postratamiento en paralelo con la división en ventanas de la trama usando la ventana de análisis por medio del convertidor de tiempo-espectral.

[0133] Además, y preferentemente, la ventana de análisis para un siguiente bloque de la señal decodificada de núcleo está situada de manera que una porción central no superpuesta de la ventana se encuentra situada dentro del intervalo de tiempo como se ilustra en 1920 de la Figura 9b.

[0134] En la propuesta 4 el retardo global del sistema está incrementado con respecto al de la propuesta 1. En el codificador llega un retardo adicional debido al módulo estéreo. La cuestión de una reconstrucción perfecta ya no se aplica en la propuesta 4, a diferencia de la propuesta 1.

[0135] 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 remuestreo convencional, combinación y suavización entre las diferentes síntesis de núcleo y las señales de anchura de banda extendidas como se hace para la norma EVS.

[0136] El entramado esquemático 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.

[0137] 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 de núcleo no está dividida en ventanas, lo que supone 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 la propuesta en la Propuesta 1 (Figura 7). Las principales desventajas de esta propuesta son 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 la frecuencia.

[0138] El entramado esquemático 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.

10 **[0139]** A la vista de lo anterior, las realizaciones 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 tratada con diferentes velocidades de muestreo a los módulos de tratamiento posteriores. El módulo incluye, por ejemplo, un codificador del habla como ACELP, herramientas de pretratamiento, un codificador audio basado en MDCT tal como TCX o un codificador de extensión de anchura de banda tal como un codificador de extensión de anchura de banda
15 en el dominio del tiempo.

[0140] Con respecto al decodificador, se lleva a cabo la combinación en el remuestreo en el dominio de la frecuencia 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
20 en MDCT, un módulo de extensión de anchura de banda o una señal de error entre los armónicos procedente de un postratamiento tal como un bajo posfiltro.

[0141] Además, en lo relativo tanto al codificador como al decodificador, es útil aplicar una ventana para la DFT o un valor complejo transformado con un relleno con ceros, una región de baja superposición y un hopsize que
25 se corresponde con 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.

[0142] Algunas realizaciones son capaces de lograr una codificación con baja velocidad binaria de audio estéreo con retardo bajo. Se diseñó específicamente para combinar un esquema de codificación de audio
30 conmutado con retardo bajo, tal como EVS, con los bancos de filtros de un módulo de codificación estéreo.

[0143] Algunas realizaciones pueden encontrar un uso 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 binaria dada) tales como, por ejemplo con aplicaciones de radio digital, difusión continua en Internet y
35 comunicaciones de audio.

[0144] La Figura 12 ilustra un aparato para codificar 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
40 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 también emitidos por medio de otra línea de parámetros 14 a una 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
45 alinear los al menos dos canales de la señal multicanal 10 usando el 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 central 31 y una señal lateral 32 a partir de los canales alineados recibidos por medio de la línea 20. El aparato para codificar comprende además un
50 codificador de señales 400 para codificar la señal central de la línea 31 y la señal lateral de la línea 32 para obtener una señal central codificada en la línea 41 y una señal lateral 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 central codificada de la línea 41, la señal lateral codificada de la línea 42, los parámetros de alineación de banda ancha y los parámetros de alineación de banda ancha de la línea
55 14 y, opcionalmente, un parámetro de nivel de la línea 14 y, adicional y opcionalmente, un parámetro de relleno 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.

[0145] Preferentemente, el alineador de señales está configurado para alinear los canales de la señal multicanal usando el parámetro de alineación de banda ancha, antes de que el determinador de parámetros 100 calcule realmente los parámetros de banda estrecha. Por tanto, en esta realización, el alineador de señales 200 envía los canales alineados de banda ancha de nuevo al determinador de parámetros 100 por medio de una línea de conexión 15. A continuación, el determinador de parámetros 100 determina la pluralidad de parámetros de alineación de banda ancha de una señal multicanal ya alineada con respecto a las características de banda ancha.
60 Sin embargo, en otras realizaciones, los parámetros se determinan sin esta secuencia específica de procedimientos.
65

[0146] 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 usando 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 usando el parámetro de alineación de banda ancha. A continuación, en la etapa 17, se determinan los parámetros de banda estrecha usando 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 usando el parámetro de alineación de banda estrecha correspondiente 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 tratamiento ulterior de las señales por el procesador de señales 300 de la Figura 12.

[0147] La Figura 14b ilustra una implementación adicional del codificador multicanal de la Figura 12 en la que se llevan a cabo diversos procedimientos en el dominio de la frecuencia.

[0148] Específicamente, el codificador multicanal comprende además un convertidor de tiempo-espectral 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 la frecuencia.

[0149] 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 la frecuencia.

[0150] Además, el codificador multicanal y, específicamente, el procesador de señales comprende además un convertidor espectral-tiempo 154 para generar una representación en el dominio del tiempo de al menos la señal central.

[0151] Preferentemente, el convertidor espectral-tiempo convierte adicionalmente una representación espectral de la señal lateral 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 central y/o de la señal lateral como señales en el dominio del tiempo en función de la implementación específica del codificador de señales 400 de la Figura 12.

[0152] Preferentemente, el convertidor de tiempo-espectral 150 de la Figura 14b está configurado para implementar las etapas 155, 156 y 157 de la Figura 4c. Específicamente, la etapa 155 comprende proporcionar una ventana de análisis con al menos un paciente de relleno con ceros en uno de sus extremos, y, específicamente, una porción de relleno con ceros en la porción inicial de la ventana y una porción de relleno con ceros en la porción terminal de la ventana como se ilustra, por ejemplo, en la Figura 7 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, además, preferentemente una parte central que es un intervalo de no superposición, según sea el caso.

[0153] En la etapa 156, cada canal es dividido en ventanas usando la ventana de análisis con intervalos de superposición. Específicamente, cada canal es dividido en ventanas usando la ventana de análisis de una manera tal que se obtiene un primer bloque del canal. Posteriormente, se obtiene un segundo bloque que tiene un determinado intervalo de superposición con el primer bloque y así sucesivamente, de manera que después, por ejemplo, de cinco operaciones de división en ventanas, se dispone de cinco bloques de muestras divididas en ventanas que a continuación se transforman 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 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.

[0154] 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 usando el parámetro de alineación de banda ancha. En la etapa 160, nuevamente llevado 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 usando parámetros de alineación correspondientes de banda ancha determinados por las bandas específicas.

[0155] La Figura 14d ilustra procedimientos adicionales llevados a cabo por el procesador de señales 300. Específicamente, el procesador de señales 300 está configurado para calcular una señal central y una señal lateral

como se ilustra en la etapa 301. En la etapa 302, puede llevarse a cabo algún tipo de tratamiento adicional de la señal y a continuación, en la etapa 303, cada bloque de la señal central y la señal lateral se transforma de nuevo al 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 superposición-suma para la señal central por una parte y una
5 operación de superposición-suma para la señal lateral por otra parte de manera que finalmente se obtienen las señales central/lateral en el dominio del tiempo.

[0156] Específicamente, las operaciones de las etapas 304 y 305 producen una forma de desvanecimiento cruzado desde un bloque de la señal central o de la señal lateral en el siguiente bloque de la señal central y la señal
10 lateral se lleva a cabo de manera que incluso si se presentan algunos cambios cualesquiera de parámetros tales como el parámetro de la diferencia de tiempo entre canales o el parámetro de la diferencia de fases entre canales, sin embargo no será audible en las señales central/lateral en el dominio del tiempo obtenido por la etapa 305 en la Figura 14d.

15 **[0157]** La Figura 13 ilustra un diagrama de bloques de una realización de un aparato para decodificar una señal multicanal codificada recibida en la línea de entrada 50.

[0158] 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
20 800 conectado a un decodificador de señales 700 por una parte y conectado al desalineador de señales por otra parte.

[0159] En particular, la señal multicanal codificada comprende una señal central codificada, una señal lateral codificada, información sobre el uso del parámetro de alineación de banda ancha e información acerca de la
25 pluralidad de parámetros de banda estrecha. Por 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.

[0160] Sin embargo, es importante observar que, a diferencia de lo que se ilustra en la Figura 12, el uso del parámetro de alineación de banda ancha y de la pluralidad de parámetros de alineación de banda ancha incluidos en
30 la señal codificada en una forma de determinada pueden ser exactamente los parámetros de alineación usados 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 usados exactamente por las mismas operaciones llevadas a cabo por el alineador de señales 200 pero con valores inversos de manera que se obtiene la desalineación.

35 **[0161]** Por tanto, la información sobre los parámetros de alineación pueden ser los parámetros de alineación usados por el alineador de señales 200 en la Figura 12 o pueden ser valores inversos, es decir, "parámetros de desalineación" reales. Adicionalmente, en general estos parámetros serán cuantificados en una forma determinada como se explica más adelante con respecto a la Figura 8a-8c.

40 **[0162]** La interfaz de entrada 600 de la Figura 13 separa la información acerca del uso del parámetro de alineación de banda ancha y de la pluralidad de parámetros de alineación de banda ancha de las señales lateral/central codificadas y comunica esta información por medio de la línea de parámetros 610 al desalineador de
45 señales 900. Por otra parte, la señal central codificada se comunica al decodificador de señales 700 por medio de la línea 601 y la señal lateral codificada se comunica al decodificador de señales 700 por medio de la línea de señales 602.

[0163] El decodificador de señales está configurado para decodificar la señal central codificada y para decodificar la señal lateral codificada con el fin de obtener una señal central decodificada en la línea 701 y una señal lateral decodificada en la línea 702. Estas señales son usadas por el procesador de señales 800 para calcular una
50 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 central decodificada y la señal lateral 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 usando la información acerca del uso del parámetro de alineación de
55 banda ancha y adicionalmente usando 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.

[0164] La Figura 9a ilustra una secuencia de etapas preferida llevada a cabo por el desalineador de señales
60 900 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 desalinea subbandas individuales usando la información acerca del uso 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 usando el parámetro de alineación de banda ancha de manera que en 913a y
65 913b se obtienen canales desalineados en fase y tiempo.

- [0165]** En la etapa 914, se lleva a cabo cualquier tratamiento ulterior que comprenda usar una división en ventanas o cualquier operación de superposición-suma o, en términos generales, cualquier operación de desvanecimiento cruzado con el fin de obtener, en 915a o 915b, una señal decodificada con pocos artefactos o libre de artefactos, es decir, a señales decodificadas que no tengan ningún artefacto si bien normalmente 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.
- [0166]** La Figura 15b ilustra una implementación preferida del decodificador multicanal ilustrado en la Figura 13.
- [0167]** En particular, el procesador de señales 800 de la Figura 13 comprende un convertidor de tiempo-espectral 810.
- [0168]** El procesador de señales comprende además un convertidor central/lateral o izquierda/derecha 820 con el fin de calcular a partir de una señal central M y una señal lateral S una señal izquierda L y una señal derecha R.
- [0169]** Sin embargo, es importante observar que con el fin de calcular L y R mediante la conversión central/lateral-izquierda/derecha en el bloque 820, no es necesario usar la señal lateral S. En cambio, como se expone más adelante, las señales izquierda/derecha se calculan inicialmente usando solo un parámetro de ganancia derivado de un parámetro de diferencia de nivel entre los canales ILD. Por tanto, en esta implementación, la señal lateral S se usa solo en el actualizador de canales 830 que opera con el fin de proporcionar una mejor señal izquierda/derecha usando la señal lateral S transmitida como se ilustra mediante la línea de derivación 821.
- [0170]** Por tanto, el convertidor 820 opera usando un parámetro de nivel obtenido por medio de una entrada de parámetro de nivel 822 y sin usar realmente la señal lateral S pero el actualizador de canales 830 opera a continuación usando el lateral 821 y, en función de la implementación específica, usando un parámetro de relleno de estéreo recibido por medio de la línea 831. El alineador de señales 900 comprende así un desalineador de fase y elemento de cambio de escala de energía 910. El cambio de escala de energía lo controla un factor de cambio de escala obtenido mediante un calculador del factor de cambio de escala 940. El calculador del factor de cambio de escala 940 es alimentado por la salida del actualizador de canales 830. Basándose en 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, basándose en el uso 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 espectral-tiempo 930 con el fin de obtener finalmente la señal decodificada.
- [0171]** La Figura 15c ilustra una secuencia de etapas adicional realizadas normalmente dentro de los bloques 920 y 930 de la Figura 15b en una realización preferida.
- [0172]** 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. Después del cálculo real de las muestras en el dominio del tiempo, se lleva a cabo una división en ventanas de síntesis opcional usando una ventana de síntesis. Preferentemente, la ventana de síntesis es exactamente la misma que la ventana de análisis o se obtiene de la ventana de análisis, por ejemplo, por interpolación o reducción, pero depende en cierto modo de la ventana de análisis. Preferentemente, esta dependencia es tal que los factores de multiplicación definidos por dos ventanas de superposición se añaden a uno de cada punto en el intervalo de superposición. Por tanto, después de la ventana de síntesis en el bloque 932, se lleva a cabo una operación de superposición y una operación de suma posterior. Como alternativa, en lugar de la división en ventanas de síntesis y la operación de superposición/suma, se lleva a cabo un desvanecimiento entre bloques subsiguientes para cada canal con el fin de obtener, como ya expuso en el contexto de la Figura 15a, una señal de codificada reducida de artefacto.
- [0173]** Si se considera la Figura 6b, se pone de relieve que las operaciones de codificación reales para la señal central, es decir, el "decodificador de EVS" por una parte y, para la señal lateral, la cuantificación inversa de vector VQ^{-1} y la operación inversa de MDCT (IMDCT) corresponden al decodificador de señales 700 de la Figura 13.
- [0174]** Además, las operaciones de DFT en los bloques 810 corresponden al elemento 810 en la Figura 15b y las funcionalidades del tratamiento 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 operación correspondiente en el bloque 930 en la Figura 15b.
- [0175]** A continuación, 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. Preferentemente, el espectro de DFT o cualquier otro espectro ilustrado en la Figura 3d es un espectro complejo y cada línea es una línea de espectro

complejo que tiene magnitud y fase o que tiene una parte real y una parte imaginaria.

[0176] 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.

5 Adicionalmente, el uso de las bandas de parámetro aumenta de frecuencias más bajas a frecuencias más altas. Normalmente, el uso 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 realización en la Figura 3d.

10 **[0177]** Además, la pluralidad de parámetros de alineación de banda ancha se proporciona de manera que existe un único parámetro de alineación para cada banda de parámetro. Esto significa que el parámetro de alineación para una banda siempre se aplica a la totalidad de los valores espectrales dentro de la banda correspondiente.

15 **[0178]** Por otra parte, además del uso de los parámetros de alineación de banda ancha, también se proporcionan parámetros de nivel para cada banda de parámetros.

[0179] A diferencia de los parámetros de nivel 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
20 alineación de banda ancha solo para un número limitado de bandas inferiores tales como las bandas 1, 2, 3 y 4.

[0180] Adicionalmente, se proporcionan parámetros de relleno de estéreo para un determinado número de bandas que incluyen las bandas inferiores tales como, en el ejemplo de realización, para las bandas 4, 5 y 6, mientras que hay valores espectrales de señal lateral para las bandas de parámetro inferiores 1, 2 y 3 y, por tanto,
25 no existen parámetros de relleno de estéreo para estas bandas inferiores en los que la concordancia de forma de onda se obtiene usando la señal lateral como tal o una señal de predicción residual que representa la señal lateral.

[0181] Como ya se mencionó, existen más líneas espectrales en las bandas superiores tales como, en la realización en la Figura 3d, siete líneas espectrales en la banda de parámetros 6 frente a solo tres líneas espectrales
30 en la banda de parámetros 2. Sin embargo, por supuesto, el número de bandas de parámetros, el número de líneas espectrales y el número de líneas espectrales dentro de una banda de parámetros y también los distintos límites para determinar los parámetros serán diferentes.

[0182] Sin embargo, en la Figura 8 se ilustra una distribución del uso de los parámetros y del número de
35 bandas para los que se proporcionan parámetros en una determinada realización en la que, a diferencia de la Figura 3d, hay en realidad 12 bandas.

[0183] Como se ilustra, el ILD del nivel de los parámetros se ha proporcionado para cada una de las 12
40 bandas y se cuantifica con una exactitud de cuantificación representada por cinco bits por banda.

[0184] Además, usando los parámetros de alineación de banda ancha IPD que se proporcionan solo para las
45 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 solo 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.

[0185] Además, de una manera bastante aproximada se proporcionan parámetros de relleno de estéreo
50 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 realmente codificadas o valores residuales espectrales de señal lateral.

[0186] A continuación se resume un tratamiento 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
55 de tiempo de L y R en el dominio de la frecuencia. 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 después de la alineación usando el parámetro de alineación de banda ancha.

60 **[0187]** 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. Posteriormente, las señales central y lateral se calculan como se ilustra en la etapa 301 y, preferentemente, además,
65 con una operación de conservación de energía como se expone más adelante. Además, se lleva a cabo una

predicción de S con M como una función de ILD y opcionalmente con una señal M previa, es decir, una señal central de una trama anterior. Posteriormente, se ejecuta una DFT inversa de la señal central y de la señal lateral que corresponde a las etapas 303, 304, 305 de la Figura 14d en la realización preferida.

5 **[0188]** En la etapa final, se codifica la señal central en el dominio del tiempo y opcionalmente, la señal residual. Este procedimiento se corresponde con el que se lleva a cabo mediante el codificador de señales 400 en la Figura 12.

[0189] En el decodificador en el tratamiento estéreo inverso, se genera la señal lateral *Side* en el dominio de DFT y se predice por primera vez a partir de la señal *Mid* como:

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

en la que *g* es una ganancia calculada para cada banda de parámetros y que depende de la diferencia de niveles entre canales (ILD) transmitida.

[0190] El resto de la predicción $\widehat{Side} - g \cdot Mid$ puede afinarse seguidamente de dos maneras diferentes:

- Mediante una codificación secundaria de la señal residual:

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

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

- Mediante una predicción residual, conocida como relleno de estéreo, se predice el espectro lateral residual con el espectro de señal *Mid* decodificado previamente a partir de la trama de DFT previa:

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

en la que g_{pred} es una ganancia predictiva transmitida por banda de parámetros.

[0191] Los dos tipos de refinamiento de la codificación pueden combinarse dentro del mismo espectro de DFT. En la 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 realización preferida, la codificación residual ilustrada en la Figura 12 se lleva a cabo en el dominio de MDCT después de sintetizar la residual señal lateral en el dominio del tiempo y transformándolo mediante una MDCT. A diferencia de la DFT, la MDCT tiene un muestreo crítico y es más adecuada para la codificación 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 residual también puede ser codificada en el dominio del tiempo mediante una técnica de codificación del habla o directamente en el dominio de DFT.

[0192] A continuación se describe una realización adicional de un tratamiento de codificador estéreo/multicanal conjunto o tratamiento estéreo/multicanal inverso.

1. Análisis de tiempo-frecuencia: DFT

[0193] Es importante que la descomposición adicional de tiempo-frecuencia del tratamiento estéreo efectuada mediante DFT permita un buen análisis de la escena del auditorio sin incrementar de manera significativa el retardo global del sistema de codificación. Por omisión, se usa una resolución del tiempo de 10 ms (que es el doble del entramado 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 engendrado y que también se añade el relleno con ceros para equilibrar el desplazamiento circular cuando se aplique ITD en el dominio de la frecuencia como se explicará más adelante en la presente.

2. Parámetros estéreo

[0194] Los parámetros estéreo pueden transmitirse al máximo con la resolución temporal del DFT estéreo. En el mínimo se puede reducir a la resolución de trama del codificador de núcleo, es decir 20 ms. Por omisión, si no se detectan transitorios, los parámetros se calculan cada 20 ms en las ventanas de DFT. El uso de las bandas de parámetros constituye una descomposición no uniforme y no superpuesta del espectro después de aproximadamente 2 veces o 4 veces el ERB (Equivalent Rectangular Bandwidth, anchura de banda rectangular

equivalente). Por omisión, se usa una ERB de 4 veces para un total de 12 bandas para una anchura de banda de frecuencia de 16 kHz (32 kbps de velocidad de muestreo, estéreo de banda superancha). En la Figura 8 se resume un ejemplo de configuración, para el que la información lateral de estéreo se transmite aproximadamente a 5 kbps.

5 3. Cálculo de ITD y alineación de tiempo de canal

[0195] El ITD se calcula estimando del TDOA (Time Delay of Arrival, retardo temporal de llegada), usando la correlación cruzada generalizada (Generalized Cross Correlation) con transformada de fase (GCC-PHAT):

$$ITD = \underset{\tau}{\operatorname{argmax}} \left(\operatorname{IDFT} \left(\frac{L_l(f)R_r^*(k)}{|L_l(f)R_r^*(k)|} \right) \right)$$

en la que 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 usada para el tratamiento estéreo posterior o se puede compartir. El pseudocódigo para calcular el ITD es el siguiente:

```

15     L=fft(window(l));
        R=fft(window(r));
        tmp = L .* conj(R);
        sfm_L = prod(abs(L).^(1/length(L)))/(mean(abs(L))+eps);
20     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+eps);
        tmp = ifft(tmp);
25     tmp = tmp([length(tmp)/2+1:length(tmp) 1:length(tmp)/2+1]);
        tmp_sort = sort(abs(tmp));
        xcorr_time=abs(tmp(-(h.stereo_itd_q_max - (length(tmp)-1)/2 - 1):-
        (h.stereo_itd_q_min - (length(tmp)-1)/2 - 1)));
        %smooth output for better detection
30     xcorr_time=[xcorr_time 0];
        xcorr_time2=filter([0.25 0.5 0.25],1,xcorr_time);
        [m,i] = max(xcorr_time2(2:end));
        if m > thresh
35     itd = h.stereo_itd_q_max - i + 1;
        else
        itd = 0;
        end

```

[0196] El cálculo de ITD también puede resumirse como sigue. La correlación cruzada se calcula en el dominio de la frecuencia antes de ser suavizada en función de la medición de planeidad espectral (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 intensa. A continuación la correlación cruzada suavizada se normaliza en amplitud antes de ser transformada de nuevo al dominio del tiempo. La normalización corresponde a la transformada de fase de la correlación cruzada, y es sabido que tiene un mejor comportamiento que la 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 fija en cero.

[0197] Si se aplica la alineación del tiempo en el dominio del tiempo, se calcula el ITD en un análisis de DFT por separado. El desplazamiento se efectúa como sigue:

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

[0198] Requiere un retardo adicional en el codificador, que es igual al máximo en el ITD máximo absoluto que puede ser tratado. La variación del ITD a lo largo del tiempo se suaviza mediante la división en ventanas de análisis del DFT.

[0199] Como alternativa, la alineación en el tiempo puede llevarse a cabo en el dominio de la frecuencia. En este caso, el cálculo de ITD y el desplazamiento circular se encuentran en el mismo dominio de DFT, dominio compartido con este otro tratamiento estéreo. El desplazamiento circular viene dado por:

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

[0200] Se necesita el relleno con ceros de las ventanas de DFT para simular un desplazamiento temporal con un desplazamiento circular. La magnitud del relleno con ceros corresponde al ITD máximo absoluto que puede ser tratado. En la realización preferida, el relleno con ceros se distribuye uniformemente en los dos lados de las 10 ventanas de análisis, mediante la suma de 3,125 ms de ceros en los dos extremos. El ITD máximo absoluto posible es entonces de 6,25 ms. En la configuración de los micrófonos A-B, corresponde para el 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 división en ventanas de síntesis y superposición-suma de la DFT.

15 [0201] Es importante que al desplazamiento temporal le siga una división en ventanas de la señal desplazada. Se trata de una diferencia muy importante con respecto a la BCC (Binaural Cue Coding, codificación binaural por indicios) de la técnica anterior, en la que el desplazamiento en el tiempo se aplica a una señal dividida en ventanas pero no se divide en ventanas 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.

20 **4. Cálculo de IPD y rotación de canales**

[0202] Los IPD se calculan después de la alineación temporal de los dos canales y esto sucede para cada banda de parámetros o al menos hasta una $ipd_{i,max,band}$ dada, en función de la configuración estéreo.

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

[0203] A continuación se aplican los IPD a los dos canales para alinear sus fases:

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

[0204] En las que $\beta = atan2(\sin(IPD_i[b]), \cos(IPD_i[b]) + c)$, $c = 10^{\frac{ILD_i[b]}{20}}$ y b es el índice de la banda de parámetros al que pertenece el índice de frecuencia k . El 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 35 del nivel de amplitud relativa de los canales, ILD. Si un canal tiene una amplitud más alta, será considerado como el canal de guía y resultará menos afectado por la rotación de canal que el canal con amplitud menor.

5. Suma-diferencia y codificación de la señal lateral

40 [0205] La transformación de suma diferencia se lleva a cabo en los espectros alineados en tiempo y fase de los dos canales de una manera tal que se conserva la energía en la señal central *Mid*.

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

45 en la que $a = \sqrt{\frac{L'^2 + R'^2}{(L' + R')^2}}$ 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 fase y tiempo se han alineado de antemano. Como alternativa los límites pueden incrementarse o

disminuirse.

[0206] Además, la señal lateral S se predice con M:

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

en la que $g(ILD) = \frac{c-1}{c+1}$, siendo $c = 10^{\frac{ILD_i[b]}{20}}$. Como alternativa es posible determinar la ganancia de predicción óptima g minimizando el ECM (error cuadrático medio) del resto y los ILD deducidos mediante la ecuación anterior.

[0207] La señal residual $S'(f)$ puede modelizarse mediante dos medios: por predicción con el espectro retardado de M o mediante su codificación directa en el dominio de MDCT.

6. Decodificación estéreo

[0208] La señal central X y la señal lateral S se convierten en primer término en los canales izquierdo y derecho L y R como sigue:

$$L_i[k] = M_i[k] + gM_i[k], \text{ for } band_{limits[b]} \leq k < band_{limits[b+1]},$$

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

en la que la ganancia g por banda de parámetros se deduce del parámetro de ILD:

$$g = \frac{c-1}{c+1},$$

en la que $c = 10^{\frac{ILD_i[b]}{20}}$.

[0209] Para las bandas de parámetros inferiores a cod_max_band, los canales se actualizan con la señal lateral decodificada:

$$L_{1i}[k] = L_{1i}[k] + \mathbf{[cod_gain]}_{1i} \cdot S_{1i}[k], \text{ for } 0 \leq k < band_{limits[cod_max_band]},$$

$$R_{1i}[k] = R_{1i}[k] - \mathbf{[cod_gain]}_{1i} \cdot S_{1i}[k], \text{ for } 0 \leq k < band_{limits[cod_max_band]},$$

[0210] Para las bandas de parámetros superiores, se predice la señal lateral y los canales se actualizan como sigue:

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

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

[0211] Finalmente, se multiplican los canales por un valor complejo que tiene la finalidad de restaurar la energía original y la fase entre canales de la señal estéreo:

$$L_i[k] = a \cdot e^{j2\pi\beta} \cdot L_i[k]$$

$$R_i[k] = a \cdot e^{j2\pi\beta - IPD_i[b]} \cdot R_i[k]$$

en la que

$$2 \cdot \frac{\sum_{k=band_{limits[b]}}^{band_{limits[b+1]}} M_i^2[k]}{\sum_{k=band_{limits[b]}}^{band_{limits[b+1]}} L_i^2[k] + \sum_{k=band_{limits[b]}}^{band_{limits[b+1]}} R_i^2[k]}$$

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

de x sobre y.

[0212] Finalmente, los canales son desplazados en tiempo o en el dominio de la frecuencia en función de los ITD transmitidos. Los canales en el dominio del tiempo se sintetizan mediante DFT inversos y superposición-suma.

5

[0213] 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 por cable como Internet.

10 **[0214]** Si bien se han descrito algunos aspectos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del procedimiento correspondiente, en el que un bloque o dispositivo corresponde a una etapa de procedimiento o a una característica de una etapa de procedimiento. De manera análoga, los aspectos descritos en el contexto de una etapa de procedimiento también representan una descripción de un bloque, elemento o característica correspondientes de un aparato correspondiente.

15

[0215] En función de determinados requerimientos para la implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede llevarse a cabo usando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tienen señales de control legibles electrónicamente almacenadas en los mismos, que cooperan (o pueden cooperar) con un sistema informático programable de manera de que se lleve a cabo el procedimiento respectivo.

20

[0216] Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema informático programable, de manera que se lleva a cabo uno de los procedimientos descritos en la presente memoria.

25

[0217] En términos generales, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, estando el código de programa operativo para llevar a cabo uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede estar almacenado, por ejemplo, en un soporte legible por máquina.

30

[0218] Otras realizaciones comprenden el programa informático para llevar a cabo uno de los procedimientos descritos en la presente memoria, almacenado en un soporte legible por máquina o en un medio de almacenamiento no transitorio.

35

[0219] Por tanto, y en otras palabras, una 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 la presente memoria, cuando el programa informático se ejecuta en un ordenador.

40 **[0220]** Por tanto, una realización adicional 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 la presente memoria.

[0221] Por tanto, una realización adicional del procedimiento según la invención se refiere a un tren de datos o a una secuencia de señales que representan el programa informático para llevar a cabo uno de los procedimientos descritos en la presente memoria. El tren de datos o la secuencia de señales pueden estar configurados, por ejemplo, para su transferencia por medio de una conexión de comunicación de datos, por ejemplo, por medio de Internet.

45

50 **[0222]** Una realización adicional comprende un medio de tratamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los procedimientos descritos en la presente memoria.

[0223] Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para llevar a cabo uno de los procedimientos descritos en la presente memoria.

55

[0224] En algunas realizaciones, puede usarse un dispositivo lógico programable (por ejemplo, una matriz de puertas programable en el campo) para llevar a cabo parte o la totalidad de las funcionalidades de los procedimientos descritos en la presente memoria. En algunas realizaciones, una matriz de puertas programable en el campo puede cooperar con un microprocesador para llevar a cabo uno de los procedimientos descritos en la presente memoria. En general, los procedimientos se implementan preferentemente mediante un aparato de hardware.

60

[0225] Las realizaciones descritas anteriormente son tan solo ilustrativas con el fin de exponer los principios de la presente invención. Debe entenderse que para otros expertos en la materia serán evidentes las modificaciones

65

y variaciones de la disposición de detalles descritos en la presente memoria. Por tanto, la finalidad es que la invención esté limitada solo por el alcance de las reivindicaciones de patente en tramitación y no por los detalles específicos presentados para la descripción y explicación de las realizaciones de la presente memoria.

REIVINDICACIONES

1. Aparato para codificar una señal multicanal que comprende al menos dos canales, en el que la señal multicanal es una señal de habla o audio multicanal, que comprende:

5 un convertidor de tiempo-espectral (1000) configurado para convertir secuencias de bloques de valores de muestreo de los al menos dos canales en una representación en el dominio de la frecuencia que tiene secuencias de bloques de valores espectrales para los al menos dos canales;

10 un procesador multicanal (1010) configurado para aplicar un tratamiento multicanal conjunto a las secuencias de bloques de valores espectrales para obtener al menos una secuencia resultante de bloques de valores espectrales que comprende información relacionada con los al menos dos canales;

un convertidor espectral-tiempo (1030) configurado para convertir la secuencia resultante 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; y

15 un codificador de núcleo (1040) configurado para codificar la secuencia de bloques de salida de valores de muestreo para obtener una señal multicanal codificada (1510), en el que el codificador de núcleo (1040) está configurado para operar según un primer control de trama con el fin de proporcionar una secuencia de tramas, en el que una trama está delimitada por un límite de trama inicial (1901) y un límite de trama final (1902), y

20 **caracterizado porque** el convertidor de tiempo-espectral (1000) o el convertidor espectral-tiempo (1030) están configurados para operar según un segundo control de trama que está sincronizado con el primer control de trama, en el que el límite de trama inicial (1901) o el límite de trama final (1902) de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con un instante inicial o un instante final de una porción superpuesta de una ventana usada por el convertidor de tiempo-espectral (1000) para cada bloque de la secuencia de bloques de valores de muestreo o usada por el convertidor espectral-tiempo (1030) para cada

25 bloque de la secuencia de bloques de salida de valores de muestreo.

2. Aparato según la reivindicación 1, en el que una ventana de análisis usada por el convertidor de tiempo-espectral (1000) o una ventana de síntesis usada por el convertidor espectral-tiempo (1030) tiene, cada una,

30 una porción superpuesta creciente y una porción superpuesta decreciente, en el que el codificador de núcleo (1040) comprende un codificador en el dominio del tiempo con una porción de anticipación (1905) o un codificador en el dominio de la frecuencia con una porción superpuesta de una ventana de núcleo, y en el que la porción superpuesta de la ventana de análisis o la ventana de síntesis es menor o igual que la porción de anticipación (1905) del codificador de núcleo o la porción superpuesta de la ventana de núcleo.

3. Aparato según una de las reivindicaciones anteriores, en el que el codificador de núcleo (1040) está configurado para usar una porción de anticipación (1905) durante la codificación de núcleo de una trama derivada de la secuencia de bloques de salida de valores de muestreo que tiene asociada la velocidad de muestreo de salida, estando situada la porción de anticipación (1905) en el tiempo a

40 continuación de la trama, en el que el convertidor de tiempo-espectral (1000) está configurado para usar una ventana de análisis (1904) que tiene una porción superpuesta con una longitud de tiempo que es menor o igual que una longitud de tiempo de la porción de anticipación (1905), en el que la porción superpuesta de la ventana de análisis se usa para generar una porción de anticipación dividida en ventanas (1905).

4. Aparato según la reivindicación 3, en el que el convertidor espectral-tiempo (1030) está configurado para procesar una porción de anticipación de salida correspondiente a la porción de anticipación dividida en ventanas usando una función de compensación (1922), en el que la función de compensación está configurada de manera que la influencia de la porción

50 superpuesta de la ventana de análisis se reduce o elimina.

5. Aparato según la reivindicación 4, en el que la función de compensación es inversa a una función que define la porción superpuesta de la ventana de análisis.

6. Aparato según la reivindicación 4 o 5, en el que la porción superpuesta es proporcional a una raíz cuadrada de una función seno, en el que la función de compensación es proporcional a la inversa de la raíz cuadrada de la función seno, y en el que el convertidor espectral-tiempo (1030) está configurado para usar una porción superpuesta que es

60 proporcional a la función seno elevada a una potencia de 1,5.

7. Aparato según una de las reivindicaciones anteriores, en el que el convertidor espectral-tiempo (1030) está configurado para generar un primer bloque de salida usando una ventana de síntesis y un segundo bloque de salida usando la ventana de síntesis, en el que una segunda

65 porción del segundo bloque de salida es una porción de anticipación de salida (1905),

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

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

en el que el codificador de núcleo (1040) está configurado para codificación de núcleo de la trama usando un resultado de la operación de anticipación.

10 8. Aparato según la reivindicación 7,

en el que el convertidor espectral-tiempo (1030) está configurado para generar un tercer bloque de salida posterior al segundo bloque de salida usando la ventana de síntesis, en el que el convertidor espectral-tiempo está configurado para superponer una primera porción superpuesta del tercer bloque de salida con la segunda porción del segundo

15 bloque de salida dividido en ventanas usando la ventana de síntesis para obtener muestras de una trama adicional siguiendo la trama en el tiempo.

9. Aparato según la reivindicación 7 u 8,

en el que el convertidor espectral-tiempo (1030) se configura, cuando genera el segundo bloque de salida para la trama, para no dividir en ventanas la porción de anticipación de salida o para compensar (1922) la porción de

20 anticipación de salida para al menos deshacer en parte la influencia de una ventana de análisis usada por el convertidor de tiempo-espectral (1000), y

en el que el convertidor espectral-tiempo (1030) está configurado para realizar una operación de superposición-suma (1924) entre el segundo bloque de salida y el tercer bloque de salida para la trama adicional y para dividir en

ventanas (1920) la porción de anticipación de salida con la ventana de síntesis.

25

10. Aparato según cualquiera de las reivindicaciones anteriores,

en el que el convertidor espectral-tiempo (1030) está configurado para usar una ventana de síntesis con el fin de generar un primer bloque de muestras de salida y un segundo bloque de muestras de salida,

30 para superponer-sumar una segunda porción del primer bloque y una primera porción del segundo bloque con el fin de generar una porción de muestras de salida,

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 codificación de núcleo de 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

35 del segundo bloque.

11. Aparato según cualquiera de las reivindicaciones anteriores,

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

40 en el que el convertidor espectral-tiempo (1030) está configurado para usar la ventana de síntesis con el fin de generar bloques de muestras de salida y para llevar a cabo una operación de superposición-suma, en el que todas las muestras en una porción de anticipación del codificador de núcleo se calculan usando la operación de superposición-suma, o

en el que el convertidor espectral-tiempo (1030) está configurado para aplicar una operación de anticipación a las

45 muestras de salida para codificación de núcleo de 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.

12. Aparato según una de las reivindicaciones anteriores,

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

50 frecuencia de entrada máxima (1211) que está relacionada con la velocidad de muestreo de entrada;

en el que el aparato comprende además un remuestreador en el dominio espectral (1020) para llevar a cabo una operación de remuestreo en el dominio de la frecuencia en los datos introducidos en el convertidor espectral-tiempo (1030) o en los datos introducidos en el procesador multicanal (1010), en el que un bloque de una secuencia

55 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);

en el que la secuencia de bloques de salida de valores de muestreo tiene una velocidad de muestreo de salida asociada que es diferente de la velocidad de muestreo de entrada.

60 13. Aparato según la reivindicación 12,

en el que el remuestreador en el dominio espectral (1020) está configurado para truncar los bloques con el fin de submuestreo o para rellenar con ceros los bloques con el fin de sobremuestreo.

14. Aparato según la reivindicación 12 o 13,

65 en el que el remuestreador en el dominio espectral (1020) está configurado para cambiar de escala (1322) los

valores espectrales de los bloques de la secuencia resultante de bloques usando un factor de cambio de escala que depende de la frecuencia de entrada máxima y que depende de la frecuencia de salida máxima.

15. Aparato según la reivindicación 14,
 5 en el que el factor de cambio de escala es mayor que uno en el caso del sobremuestreo, 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 cambio de escala es inferior a uno en el caso del submuestreo, en el que la velocidad de muestreo de salida es inferior a la velocidad de muestreo de entrada, o
 en el que el convertidor de tiempo-espectral (1000) está configurado para llevar a cabo un algoritmo de transformada
 10 de tiempo-frecuencia sin usar una normalización en relación con un número total de valores espectrales de un bloque de valores espectrales (1311), y en el que el factor de cambio de 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 remuestreo, y en el que el convertidor espectral-tiempo está configurado para aplicar una normalización basada en la frecuencia de salida máxima (1331).
- 15 16. Aparato según una de las reivindicaciones anteriores,
 en el que el convertidor de 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
 20 algoritmo de transformada de Fourier inversa discreta.
17. Aparato según una de las reivindicaciones anteriores,
 en el que el procesador multicanal (1010) está configurado para obtener una secuencia resultante adicional de bloques de valores espectrales, y
 en el que el convertidor espectral-tiempo (1030) está configurado para convertir la secuencia resultante adicional de
 25 valores espectrales en una representación en el dominio del tiempo (1032) adicional que comprende una secuencia de salida adicional de bloques de valores de muestreo asociados con velocidades de muestreo de salida que son iguales a la velocidad de muestreo de entrada.
18. Aparato según una de las reivindicaciones 12 a 17,
 30 en el que el procesador multicanal (1010) está configurado para proporcionar una secuencia resultante adicional más de bloques de valores espectrales,
 en el que el remuestreador en el dominio espectral (1020) está configurado para el remuestreo de bloques de la secuencia resultante adicional en el dominio de la frecuencia con el fin de una secuencia remuestreada adicional de bloques de valores espectrales, en el que un bloque de la secuencia remuestreada adicional tiene valores
 35 espectrales hasta una frecuencia de salida máxima adicional que es diferente de la frecuencia de entrada máxima o que es diferente de la frecuencia de salida máxima,
 en el que el convertidor espectral-tiempo (1030) está configurado para convertir la secuencia remuestreada adicional de bloques de valores espectrales en una representación adicional en el dominio del tiempo que comprende una secuencia de salida adicional más de bloques de valores de muestreo que tienen asociada una velocidad de
 40 muestreo de salida adicional que es diferente de la velocidad de muestreo de salida o de la velocidad de muestreo de entrada.
19. Aparato según una de las reivindicaciones anteriores,
 en el que el procesador multicanal (1010) está configurado para generar una señal central como la al menos una
 45 secuencia resultante de bloques de valores espectrales usando solo una operación de submezcla, o una señal lateral adicional como una secuencia resultante adicional de bloques de valores espectrales.
20. Aparato según una de las reivindicaciones 12 a 19,
 en el que el procesador multicanal (1010) está configurado para generar una señal central como la al menos una
 50 secuencia resultante, en el que el remuestreador en el dominio espectral (1020) está configurado para remuestrear la señal central a dos secuencias separadas que tienen dos frecuencias de salida máxima diferentes que son 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
 55 en el que el codificador de núcleo (1030) comprende un primer preprocesador (1430c) para pretratar la primera secuencia de salida a una primera velocidad de muestreo y un segundo preprocesador (1430d) para pretratar la segunda secuencia de salida con la segunda velocidad de muestreo, y
 en el que el codificador de núcleo está configurado para codificar la primera o la segunda secuencia de salida pretratada, o
 60 en el que el procesador multicanal está configurado para generar una señal lateral como la al menos una secuencia resultante, en el que el remuestreador en el dominio espectral (1020) está configurado para remuestrear la señal lateral a dos secuencias remuestreadas que tienen dos frecuencias de salida máxima diferentes que son diferentes de la frecuencia de entrada máxima,
 en el que el convertidor espectral-tiempo (1030) está configurado para convertir las dos secuencias remuestreadas
 65 en dos secuencias de salida que tienen velocidades de muestreo diferentes, y

en el que el codificador de núcleo comprende un primer preprocesador (1430c) y un segundo preprocesador (1430d) para pretratar las secuencias de salida primera y segunda; y
 en el que el codificador de núcleo (1040) está configurado para codificar en núcleo (1430a, 1430b) la primera o segunda secuencia preprocesada.

5

21. Aparato según una de las reivindicaciones anteriores, en el que el convertidor espectral-tiempo (1030) está configurado para convertir la al menos una secuencia resultante en una representación en el dominio del tiempo sin ningún remuestreo 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 remuestreada con el fin de obtener la señal multicanal codificada, o

10

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

15

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 para obtener la señal multicanal codificada, o

20

en el que el aparato comprende además un codificador de señal lateral específico en el dominio espectral (1430e), o 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

25

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.

22. Aparato según una de las reivindicaciones anteriores,

en el que el convertidor de tiempo-espectral 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 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 relleno con ceros en una porción inicial o en una de sus porciones fin, o

30

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 relleno con ceros comprenden, cada uno de ellos, un número entero de muestras para al menos dos velocidades de muestreo que comprenden 12,8 kHz, 16 kHz, 25,6 kHz, 32 kHz, 48 kHz, o

35

en el que una raíz 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 temporal se ha fijado en un valor inferior o igual a la velocidad de trama del codificador de núcleo.

23. Aparato según una de las reivindicaciones anteriores,

en el que el procesador multicanal (1010) está configurado para procesar la secuencia de bloques con el fin de obtener una alineación en el tiempo usando un parámetro de alineación en tiempo de banda ancha (12) y de obtener una alineación de fase de banda estrecha usando una pluralidad de parámetros de alineación de fase de banda estrecha (14), y para calcular una señal central y una señal lateral como secuencias resultantes usando secuencias alineadas.

40

24. Procedimiento de codificación de una señal multicanal que comprende al menos dos canales, en el que la señal multicanal es una señal de habla o audio multicanal, que comprende:

45

conversión de tiempo-espectral (1000) de secuencias de bloques de valores de muestreo de los al menos dos canales en una representación en el dominio de la frecuencia que tiene secuencias de bloques de valores espectrales para los al menos dos canales;

50

aplicación (1010) de un tratamiento multicanal conjunto a las secuencias de bloques de valores espectrales para obtener al menos una secuencia resultante de bloques de valores espectrales que comprende información relacionada con los al menos dos canales;

conversión espectral-tiempo (1640) de la secuencia resultante 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; y

55

codificación de núcleo (1040) de la secuencia de bloques de salida de valores de muestreo para obtener una señal multicanal codificada (1510),

en el que la codificación de núcleo (1040) opera según un primer control de trama para proporcionar una secuencia de tramas, en el que una trama está delimitada por un límite de trama inicial (1901) y un límite de trama final (1902), y

60

caracterizado porque la conversión de tiempo-espectral (1000) o la conversión espectral-tiempo (1030) opera según un segundo control de trama que está sincronizado con el primer control de trama, en el que el límite de trama inicial (1901) o el límite de trama final (1902) de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con un instante inicial o un instante final de una porción superpuesta de una ventana usada por la conversión de tiempo-espectral (1000) para cada bloque de la secuencia de bloques de

65

valores de muestreo o usada por la conversión espectral-tiempo (1030) para cada bloque de la secuencia de bloques de salida de valores de muestreo.

25. Aparato para decodificar una señal multicanal codificada, en el que la señal multicanal codificada es una señal de habla o audio multicanal, que comprende:

un decodificador de núcleo (1600) configurado para generar una señal con decodificación de núcleo;
 un convertidor de tiempo-espectral (1610) configurado para convertir una secuencia de bloques de valores de muestreo de la señal con decodificación de núcleo en una representación en el dominio de la frecuencia que tiene una secuencia de bloques de valores espectrales para la señal con decodificación de núcleo;
 un procesador multicanal (1630) configurado para aplicar un tratamiento multicanal inverso a una secuencia (1615) que comprende la secuencia de bloques para obtener al menos dos secuencias resultantes (1631, 1632, 1635) de bloques de valores espectrales; y
 un convertidor espectral-tiempo (1640) configurado para convertir las al menos dos secuencias resultantes (1631, 1632) de bloques de valores espectrales en una representación en el dominio del tiempo que comprende al menos dos secuencias de bloques de salida de valores de muestreo, 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 límite de trama inicial (1901) y un límite de trama final (1902),
caracterizado porque el convertidor de tiempo-espectral (1610) o el convertidor espectral-tiempo (1640) está configurado para operar según un segundo control de trama que está sincronizado con el primer control de trama, en el que el límite de trama inicial (1901) o el límite de trama final (1902) de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con un instante inicial o un instante final de una porción superpuesta de una ventana usada por el convertidor de tiempo-espectral (1610) para cada bloque de la secuencia de bloques de valores de muestreo o usada por el convertidor espectral-tiempo (1640) para cada bloque de las al menos dos secuencias de bloques de salida de valores de muestreo.

26. Aparato según la reivindicación 25, en el que la señal con decodificación de núcleo tiene la secuencia de tramas, teniendo una trama el límite de trama inicial (1901) y el límite de trama final (1902), en el que una ventana de análisis (1914) usada por el convertidor de tiempo-espectral (1610) para división en ventanas de la trama de la secuencia de tramas tiene una porción superpuesta que termina antes del límite de trama final (1902) dejando un intervalo de tiempo (1920) entre un extremo de la porción superpuesta y el límite de trama final (1902), y en el que el decodificador de núcleo (1600) está configurado para realizar un tratamiento en las muestras en el intervalo de tiempo (1920) en paralelo con la división en ventanas de la trama usando la ventana de análisis (1914), o en el que un decodificador de núcleo postratamiento se lleva a cabo en las muestras en el intervalo de tiempo (1920) en paralelo con la división en ventanas de la trama usando la ventana de análisis.

27. Aparato según una de las reivindicaciones 25 a 26, en el que la señal con decodificación de núcleo tiene la secuencia de tramas, teniendo una trama el límite de trama inicial (1901) y el límite de trama final (1902), en el que el inicio de una primera porción superpuesta de una ventana de análisis (1914) coincide con el límite de trama inicial (1901), y en el que el fin de una segunda porción superpuesta de la ventana de análisis (1914) está situado antes del límite de trama final (1902), de manera que existe un intervalo de tiempo (1920) entre el fin de la segunda porción superpuesta y el límite de trama final, y en el que la ventana de análisis para un bloque siguiente de la señal con decodificación de núcleo está situada de manera que una porción no superpuesta central de la ventana de análisis está situada dentro del intervalo de tiempo (1920).

28. Aparato según una de las reivindicaciones 25 a 27, en el que la ventana de análisis usada por el convertidor de tiempo-espectral (1610) tiene la misma forma y longitud de tiempo que la ventana de síntesis usada por el convertidor espectral-tiempo (1640).

29. Aparato según una de las reivindicaciones 25 a 28, en el que la señal con decodificación de núcleo tiene la secuencia de tramas, en el que una trama comprende una longitud, en el que el convertidor de tiempo-espectral (1610) está configurado para usar la ventana, y en el que la longitud de la ventana excluyendo todas las porciones de relleno con ceros es menor o igual que la mitad de la longitud de la trama.

30. Aparato según una de las reivindicaciones 25 a 29, en el que el convertidor espectral-tiempo (1640) está configurado para aplicar una ventana de síntesis para obtener un primer bloque de salida de muestras divididas en ventanas para una primera secuencia de salida de las al menos dos secuencias de salida;

- para aplicar la ventana de síntesis con el fin de obtener un segundo bloque de salida de muestras divididas en ventanas para la primera secuencia de salida de las al menos dos secuencias de salida;
 para superponer-sumar el primer bloque de salida y el segundo bloque de salida con el fin de obtener un primer grupo de muestras de salida para la primera secuencia de salida;
- 5 en el que el convertidor espectral-tiempo (1640) está configurado para aplicar una ventana de síntesis con el fin de obtener un primer bloque de salida de muestras divididas en ventanas para una segunda secuencia de salida de las al menos dos secuencias de salida;
 para aplicar la ventana de síntesis con el fin de obtener un segundo bloque de salida de muestras divididas en ventanas para la segunda secuencia de salida de las al menos dos secuencias de salida;
- 10 para superponer-sumar el primer bloque de salida y el segundo bloque de salida con el fin de obtener un segundo grupo de muestras de salida para la segunda secuencia de salida;
 en el que el primer grupo de muestras de salida para la primera secuencia de salida y el segundo grupo de muestras de salida para la segunda secuencia de salida están relacionados con la misma porción temporal de la señal multicanal codificada o están relacionados con la misma trama de la señal con decodificación de núcleo.
- 15
31. Aparato según una de las reivindicaciones 25 a 30,
 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;
- 20 en el que el aparato comprende además un remuestreador en el dominio espectral (1620) para llevar a cabo una operación de remuestreo en el dominio de la frecuencia en los datos introducidos en el convertidor espectral-tiempo (1640) o en los datos introducidos en el procesador multicanal (1630), 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;
- 25 en el que las al menos dos secuencias de bloques de salida de valores de muestreo tienen asociada una velocidad de muestreo de salida que es diferente de la velocidad de muestreo de entrada.
- 32 Aparato según la reivindicación 31,
 en el que el remuestreador en el dominio espectral (1020) está configurado para truncar los bloques con el fin de
 30 submuestreo o para rellenar con ceros los bloques con el fin de sobremuestreo.
33. Aparato según la reivindicación 31 o 32,
 en el que el remuestreador en el dominio espectral (1020) está configurado para cambiar de escala (1322) los valores espectrales de los bloques de la secuencia resultante de bloques usando un factor de cambio de escala que
 35 depende de la frecuencia de entrada máxima y que depende de la frecuencia de salida máxima.
34. Aparato según una de las reivindicaciones 31 a 33,
 en el que el factor de cambio de escala es mayor que uno en el caso del sobremuestreo, 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 cambio de escala es
 40 inferior a uno en el caso del submuestreo, en el que la velocidad de muestreo de salida es inferior a la velocidad de muestreo de entrada, o
 en el que el convertidor de tiempo-espectral (1000) está configurado para llevar a cabo un algoritmo de transformada de tiempo-frecuencia sin usar una normalización en cuanto al número total de valores espectrales de un bloque de valores espectrales (1311), y en el que el factor de cambio de escala es igual a un cociente entre el número de
 45 valores espectrales de un bloque de la secuencia remuestreada y el número de valores espectrales de un bloque de valores espectrales antes del remuestreo, y en el que el convertidor espectral-tiempo está configurado para aplicar una normalización basada en la frecuencia de salida máxima (1331).
35. Aparato según una de las reivindicaciones 25 a 34,
 50 en el que el convertidor de 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 inverso de transformada de Fourier discreta.
36. Aparato según una de las reivindicaciones 25 a 35,
 55 en el que el decodificador de núcleo (1600) está configurado para generar otra señal decodificada de núcleo (1601) que tiene otra velocidad de muestreo que es diferente de la velocidad de muestreo de entrada,
 en el que el convertidor de tiempo-espectral (1610) está configurado para convertir la señal decodificada de núcleo adicional en una representación en el dominio de la frecuencia que tiene una secuencia adicional (1611) de bloques de valores para la señal decodificada de núcleo adicional, en el que un bloque de valores de muestreo de la señal
 60 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 velocidad de muestreo adicional,
 en el que el remuestreador en el dominio espectral (1620) está configurado para remuestrear la secuencia de bloques adicional para la señal decodificada de núcleo adicional en el dominio de la frecuencia con el fin de obtener una secuencia remuestreada adicional (1621) de bloques de valores espectrales, en el que un bloque de valores
 65 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
 en el que el aparato comprende además un combinador (1700) para combinar la secuencia remuestreada y la
 secuencia remuestreada adicional para obtener la secuencia (1701) que será tratada por el procesador multicanal
 (1630).

5

37. Aparato según una de las reivindicaciones 25 a 36,
 en el que el decodificador de núcleo (1600) está configurado para generar una señal decodificada de núcleo
 adicional más que tiene una velocidad de muestreo adicional más que es igual a una velocidad de muestreo de
 salida (1603),

10 en el que el convertidor de tiempo-espectral (1610) está configurado para convertir la señal decodificada de núcleo
 adicional más en una representación en el dominio de la frecuencia (1613) para obtener una secuencia adicional
 más de bloques de valores espectrales,

en el que el aparato comprende además un combinador (1700) para combinar la secuencia adicional más de
 bloques de valores espectrales y la secuencia remuestreada (1622, 1621) de bloques en un proceso de generación
 15 de la secuencia de bloques tratada por el procesador multicanal (1630).

38. Aparato según una de las reivindicaciones 25 a 37,
 en el que el decodificador de núcleo (1600) comprende al menos uno de entre una porción de decodificación basada
 en MDCT (1600d), una porción de decodificación de extensión de la anchura de banda en el dominio del tiempo
 20 (1600c), una porción de decodificación ACELP (1600b) y una porción de decodificación bajo de posfiltro (1600a).

en el que la porción de decodificación basada en MDCT (1600d) o la porción de decodificación de extensión de la
 anchura 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 bajo de posfiltro (1600a) está
 25 configurada para generar una señal decodificada de núcleo con una velocidad de muestreo que es diferente de la
 velocidad de muestreo de salida.

39. Aparato según una de las reivindicaciones 25 a 38,
 en el que el convertidor de tiempo-espectral (1610) está configurado para aplicar una ventana de análisis a al menos
 30 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,

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

35

40. Aparato según una de las reivindicaciones 25 a 39,
 en el que la secuencia tratada por el procesador multicanal (1630) corresponde a una señal central, y
 en el que el procesador multicanal (1630) está configurado para generar adicionalmente una señal lateral usando
 información en una señal lateral incluida en la señal multicanal codificada, y

40 en el que el procesador multicanal (1630) está configurado para generar las al menos dos secuencias resultantes
 usando la señal central y la señal lateral.

41. Aparato según una de las reivindicaciones 25 a 40,
 en el que el procesador multicanal (1630) está configurado para convertir (820) la secuencia en una primera
 45 secuencia para un primer canal de salida y una segunda secuencia para un segundo canal de salida usando un
 factor de ganancia por banda de parámetros;

para actualizar (830) la primera secuencia y la segunda secuencia usando una señal lateral decodificada o para
 actualizar la primera secuencia y la segunda secuencia usando una señal lateral predicha a partir de un bloque
 anterior de la secuencia de bloques para la señal central usando un parámetro de relleno de estéreo para una banda
 50 de parámetros;

para llevar a cabo (910) una desalineación de fase y un cambio de escala de energía usando información sobre la
 pluralidad de parámetros de alineación de fase de banda estrecha; y

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

55

42. Procedimiento de decodificación de una señal multicanal codificada, en el que la señal multicanal
 codificada es una señal de habla o audio multicanal, que comprende:

generación (1600) de una señal con decodificación de núcleo;

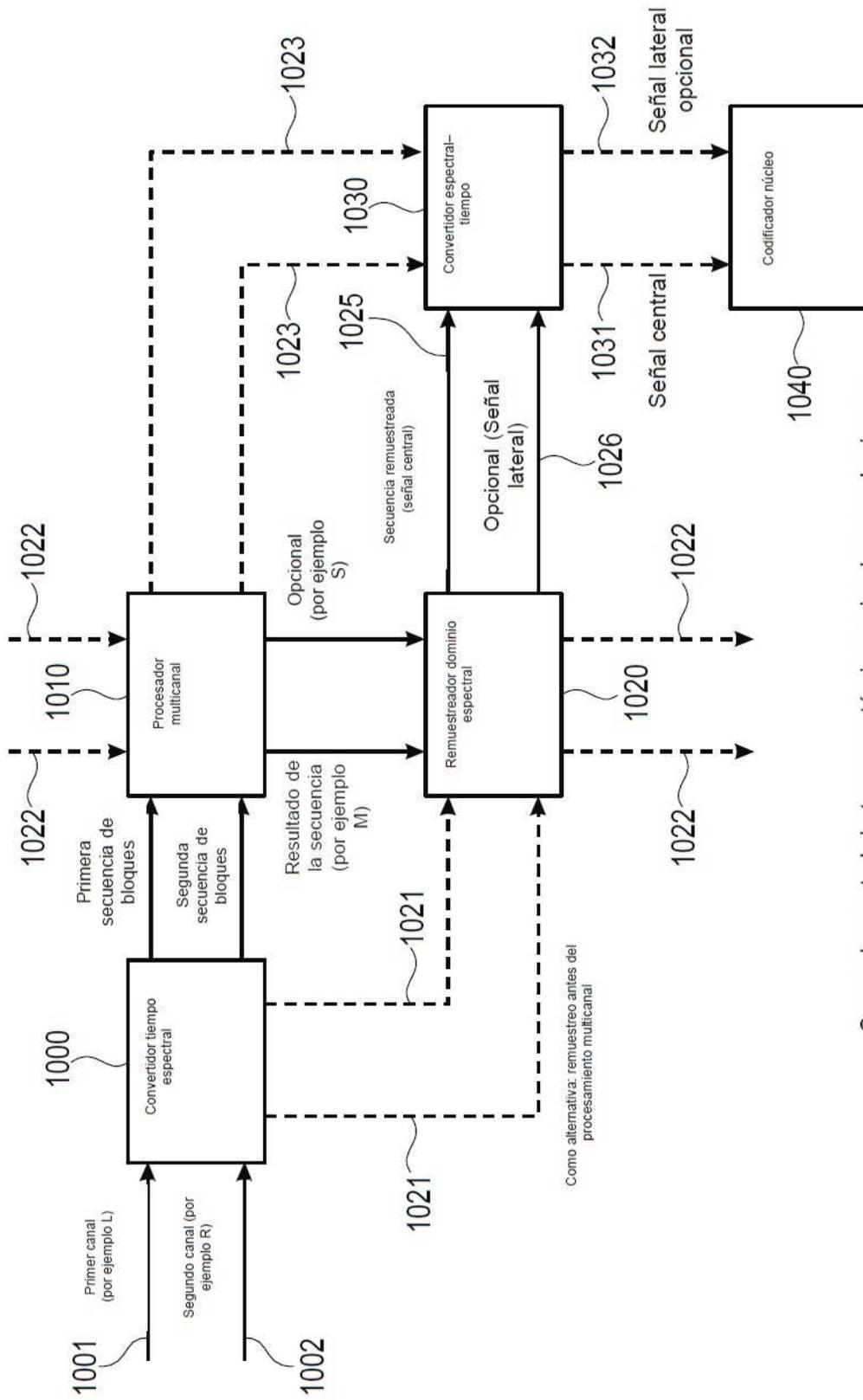
60 conversión de tiempo-espectral (1610) de una secuencia de bloques de valores de muestreo de la señal con
 decodificación de núcleo en una representación en el dominio de la frecuencia que tiene una secuencia de
 bloques de valores espectrales para la señal con decodificación de núcleo;

aplicación (1630) de un tratamiento multicanal inverso a una secuencia (1615) que comprende la secuencia de
 bloques para obtener al menos dos secuencias resultantes (1631, 1632, 1635) de bloques de valores

65

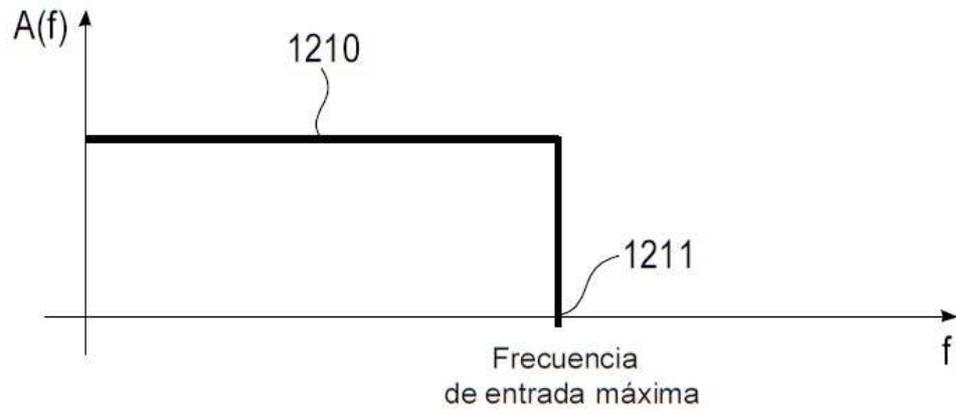
espectrales; y

- conversión espectral-tiempo (1640) de las al menos dos secuencias resultantes (1631, 1632) de bloques de valores espectrales en una representación en el dominio del tiempo que comprende al menos dos secuencias de bloques de salida de valores de muestreo,
- 5 en el que la generación de la señal con decodificación de núcleo (1600) opera según un primer control de trama para proporcionar una secuencia de tramas, en el que una trama está delimitada por un límite de trama inicial (1901) y un límite de trama final (1902),
- caracterizado porque** la conversión de tiempo-espectral (1610) o la conversión espectral-tiempo (1640) opera según un segundo control de trama que está sincronizado con el primer control de trama,
- 10 en el que el límite de trama inicial (1901) o el límite de trama final (1902) de cada trama de la secuencia de tramas se encuentra en una relación predeterminada con un instante inicial o un instante final de una porción superpuesta de una ventana usada por la conversión de tiempo-espectral (1610) para cada bloque de la secuencia de bloques de valores de muestreo o usada por la conversión espectral-tiempo (1640) para cada bloque de las al menos dos secuencias de bloques de salida de valores de muestreo.
- 15 43. Programa informático para llevar a cabo, cuando se ejecuta en un ordenador o un procesador, el procedimiento según la reivindicación 24 o el procedimiento según la reivindicación 42.

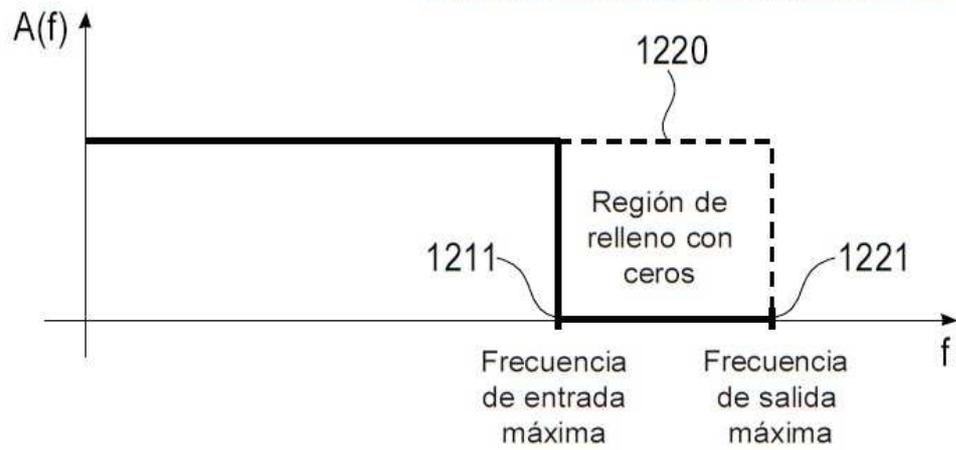


Segundo control de trama está sincronizado con el primero

Fig. 1



Sobremuestreo mediante relleno con ceros



Submuestreo por truncamiento

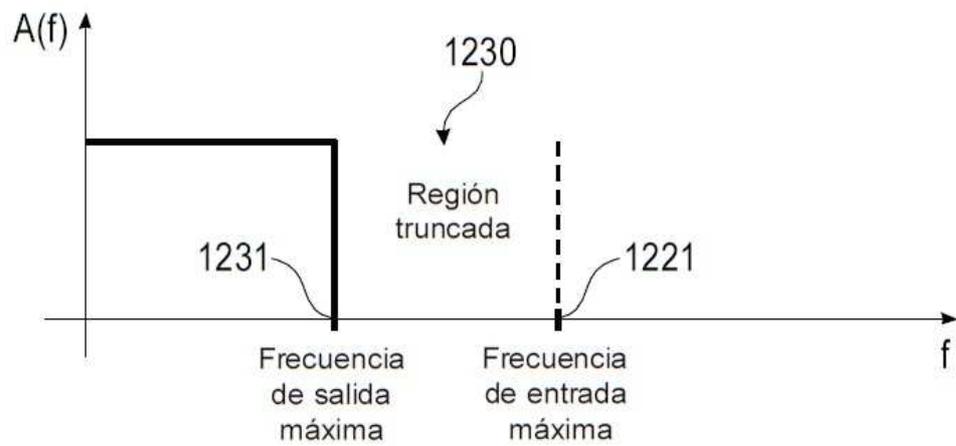


Fig. 2

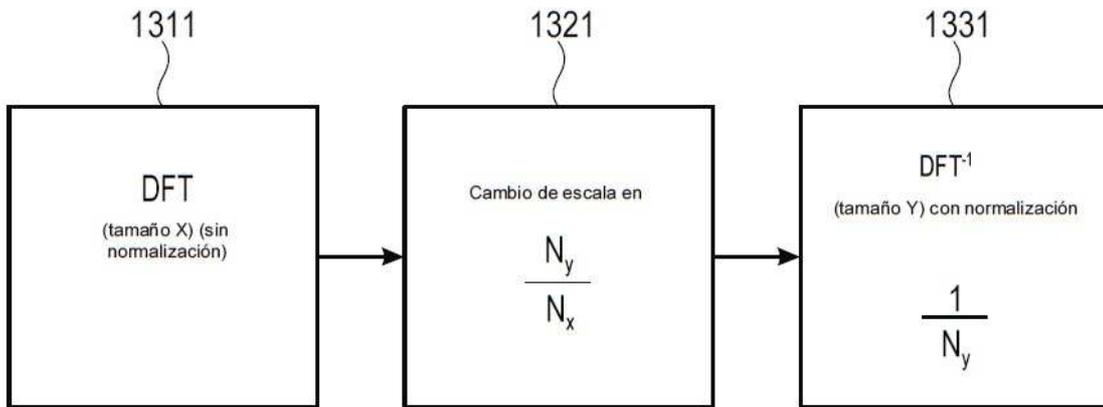


Fig. 3a

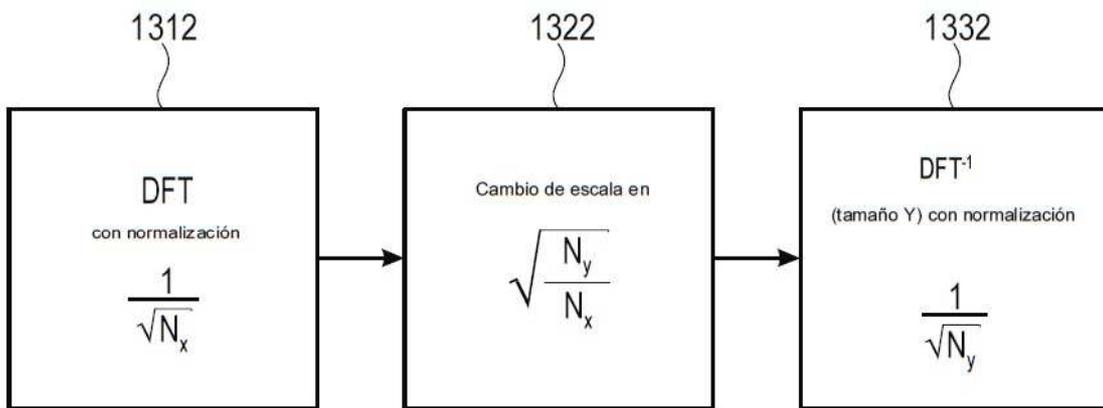


Fig. 3b

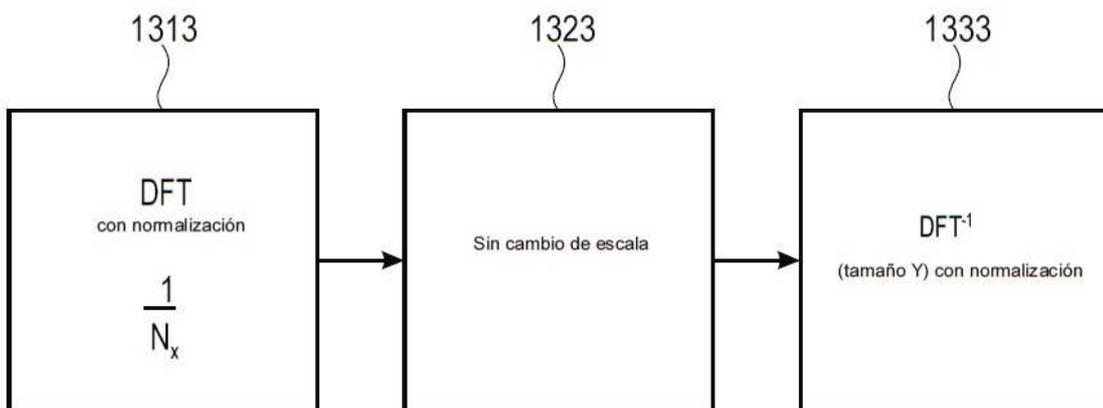
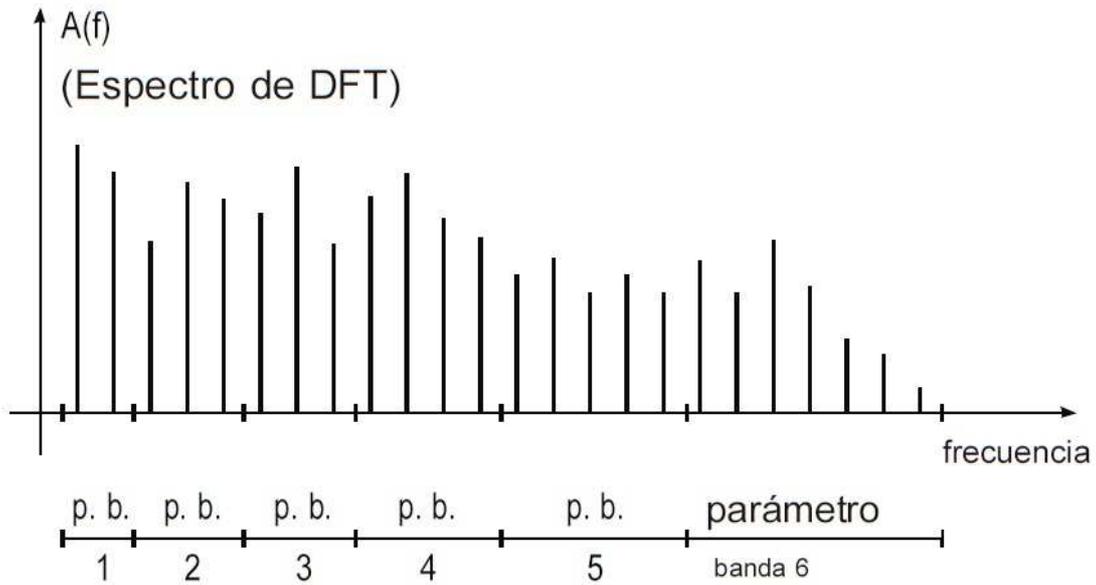


Fig. 3c



- 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 relleno estéreo para bandas de parámetros 4, 5, 6, por ejemplo parámetros de relleno estéreo;
- Señal slide (residual) para bandas de parámetro, 1, 2, 3;
- Más líneas espectrales en banda superior, por ejemplo, siete líneas espectrales en banda de parámetros 6 frente a cinco líneas espectrales en banda de parámetro 2.

Fig. 3d

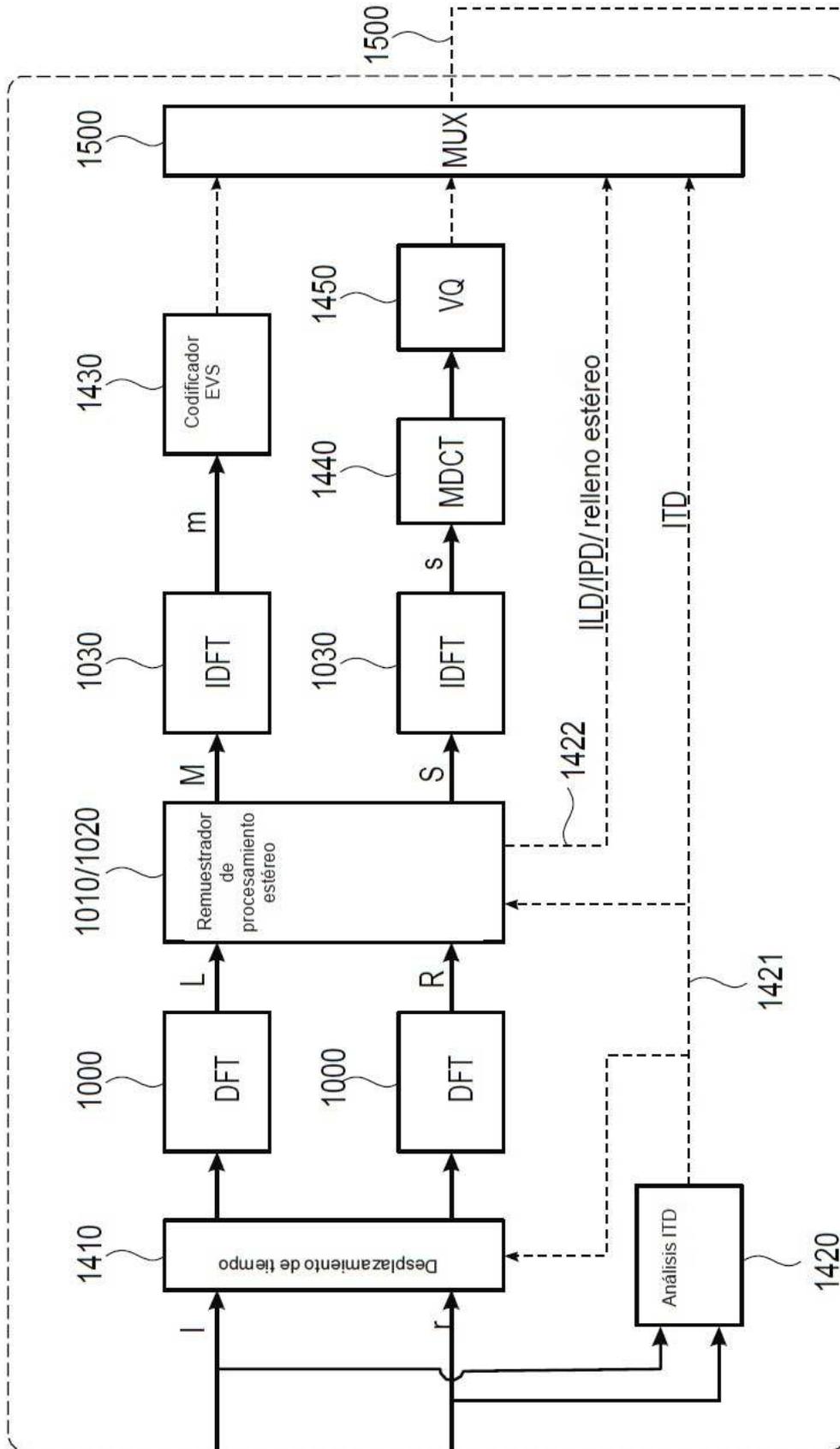


Fig. 4a

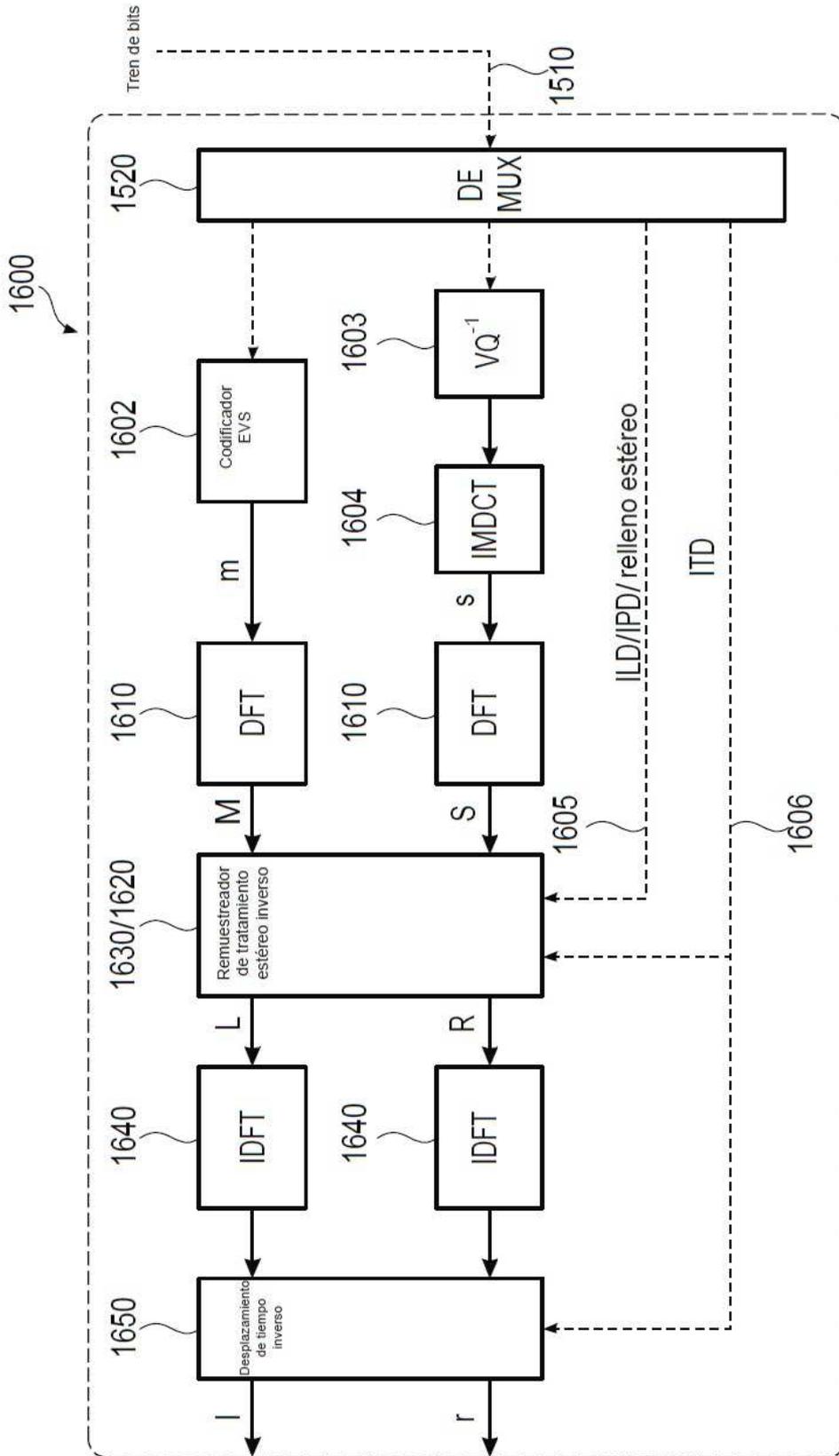


Fig. 4b

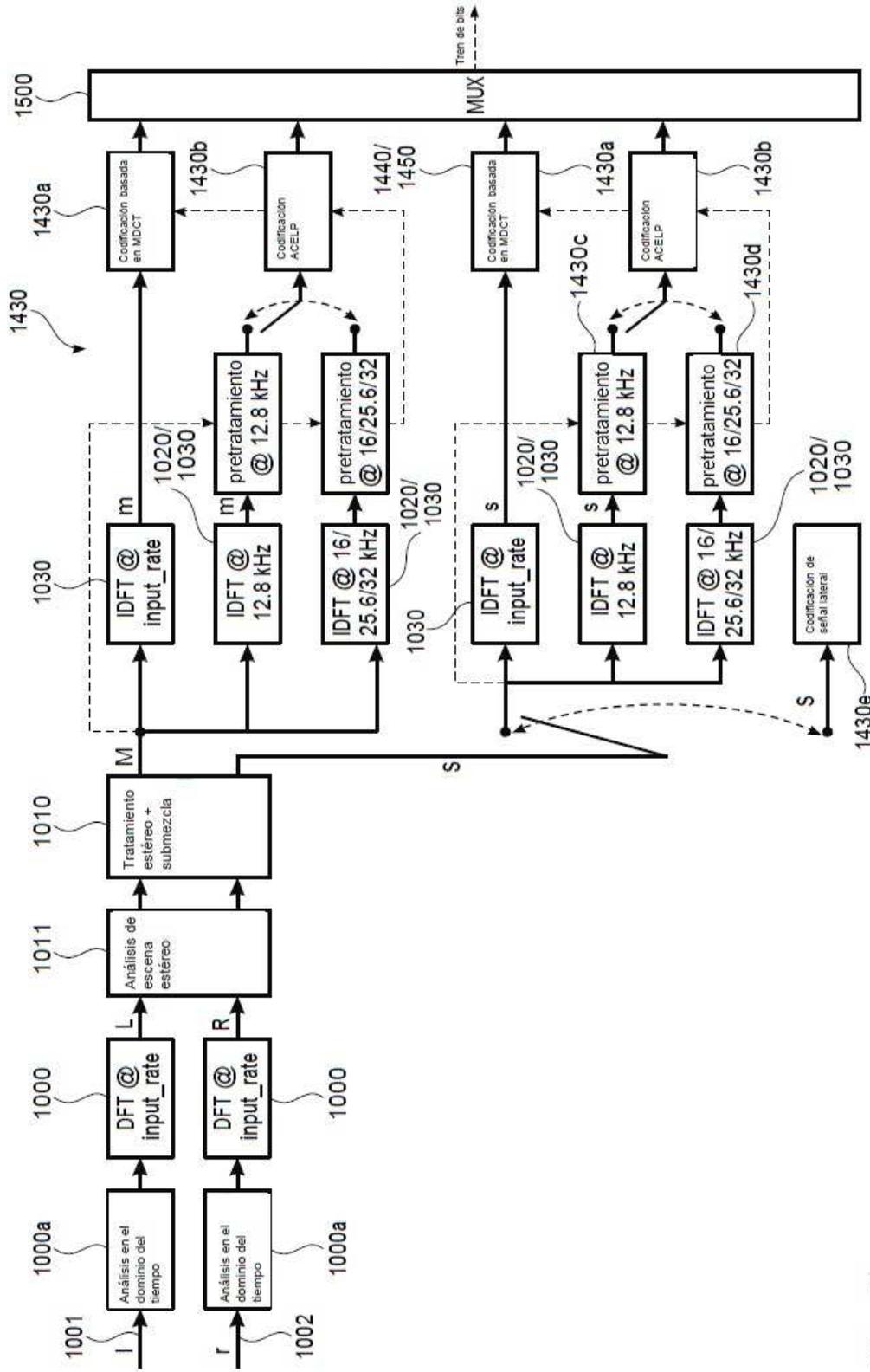


Fig. 5

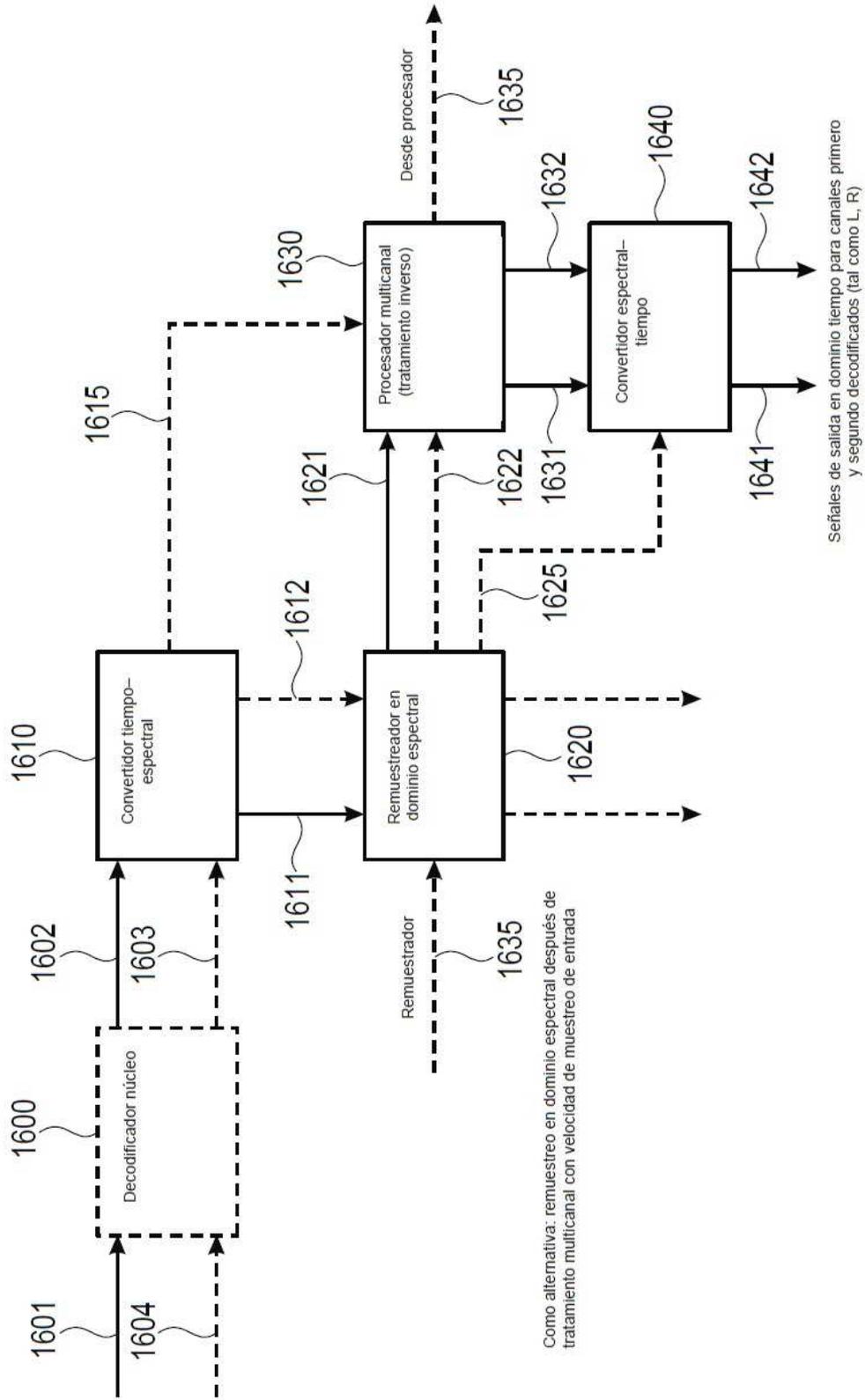


Fig. 6

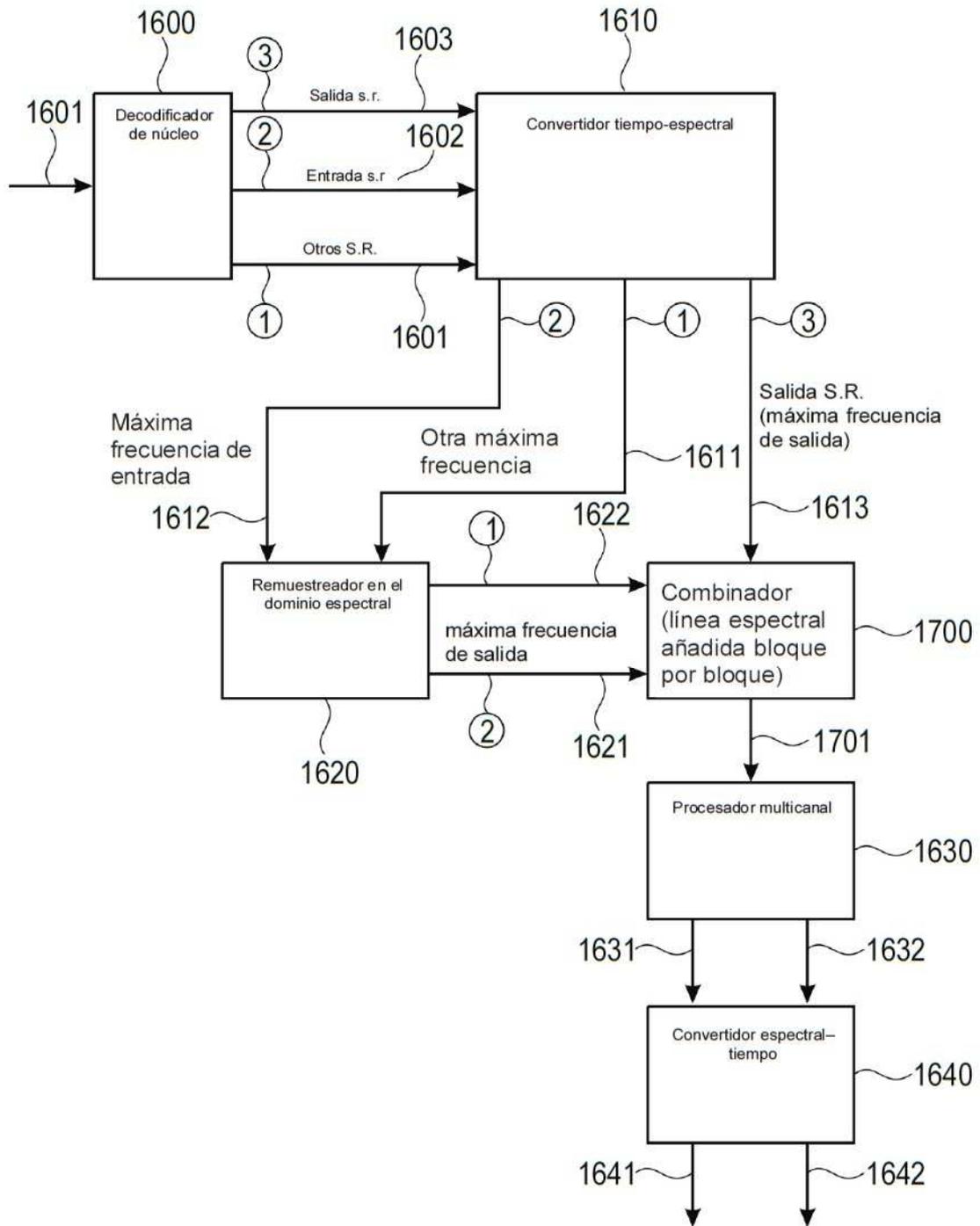


Fig. 7a

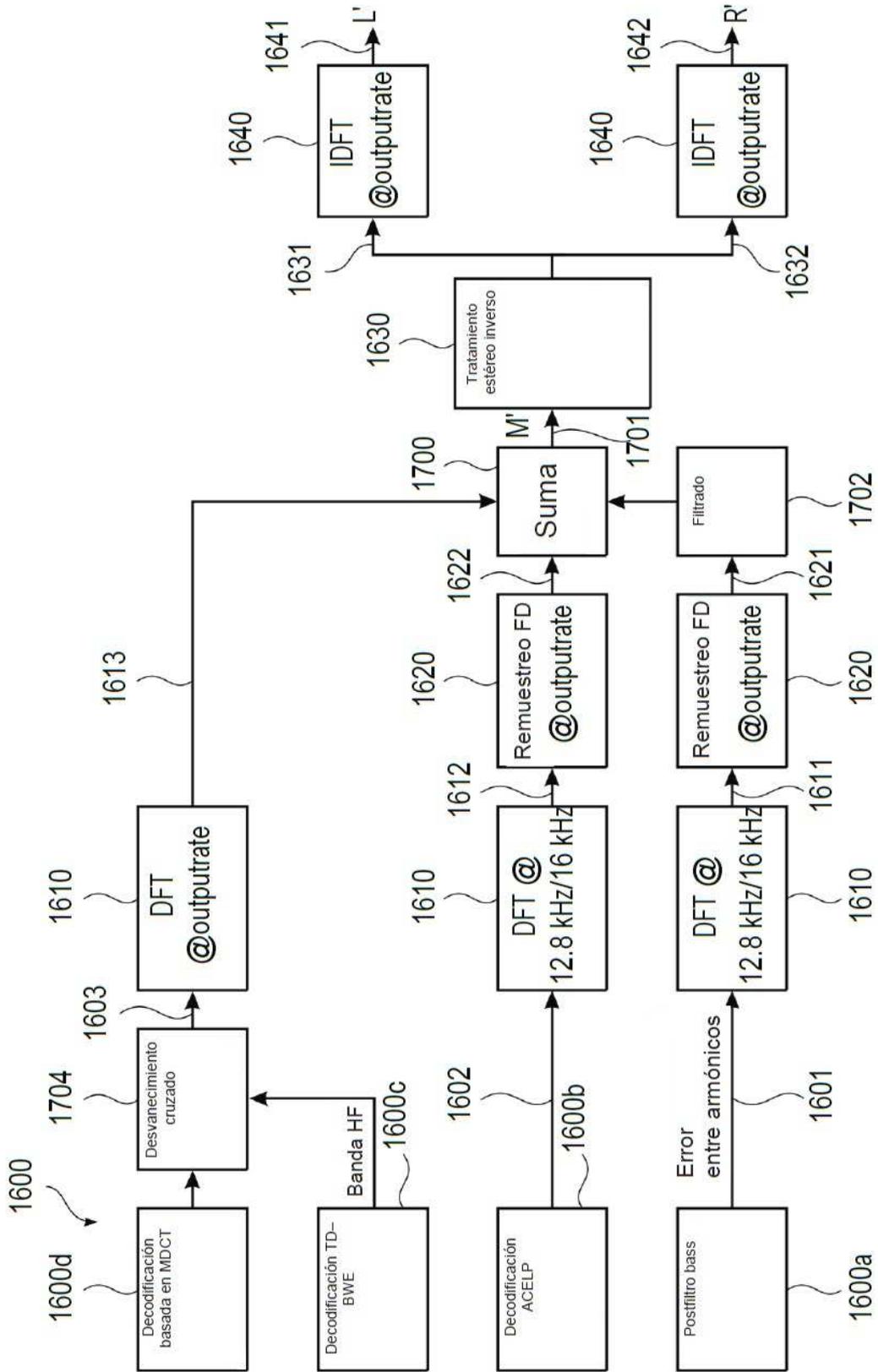


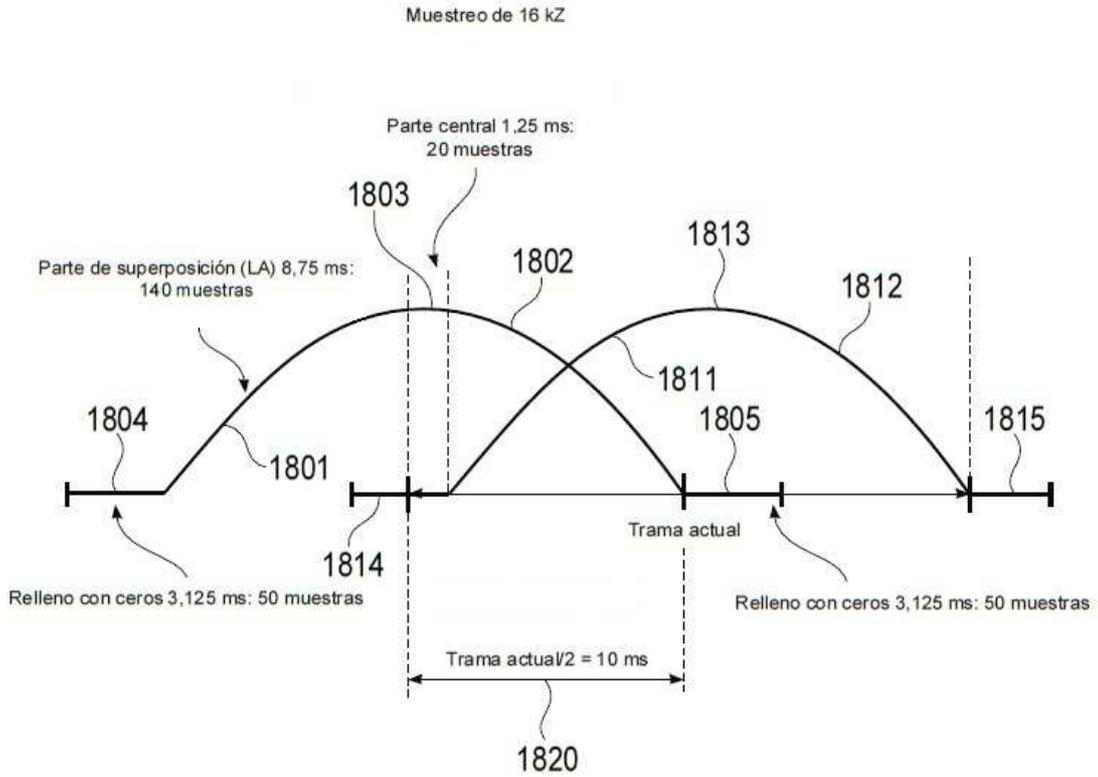
Fig. 7b

	Duración (ms)	Muestras @12.8 kHz	Muestras @16 kHz	Muestras @32 kHz	Muestras @48 kHz
Relleno con ceros	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 codificador	Retardo del decodificador	Retardo disponible en el lado del decodificador después del decodificador núcleo	Banda de filtro de resolución de frecuencia/transmada 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 ascendentes $ovlp_size$ están dados por:

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

for $k = 0..ovlp_size - 1$

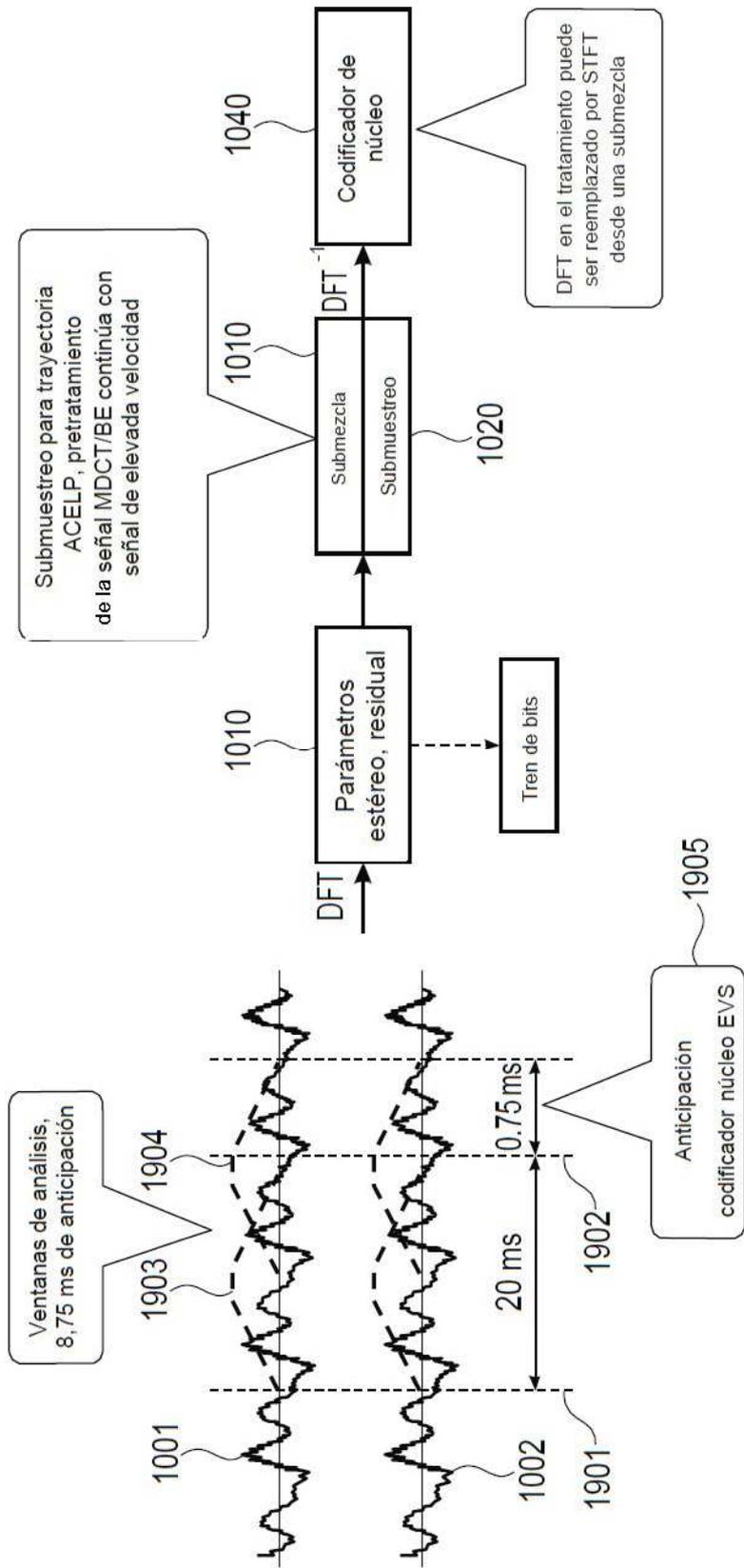
Los coeficientes descendentes $ovlp_size$ están dados por:

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

for $k = 0..ovlp_size - 1$

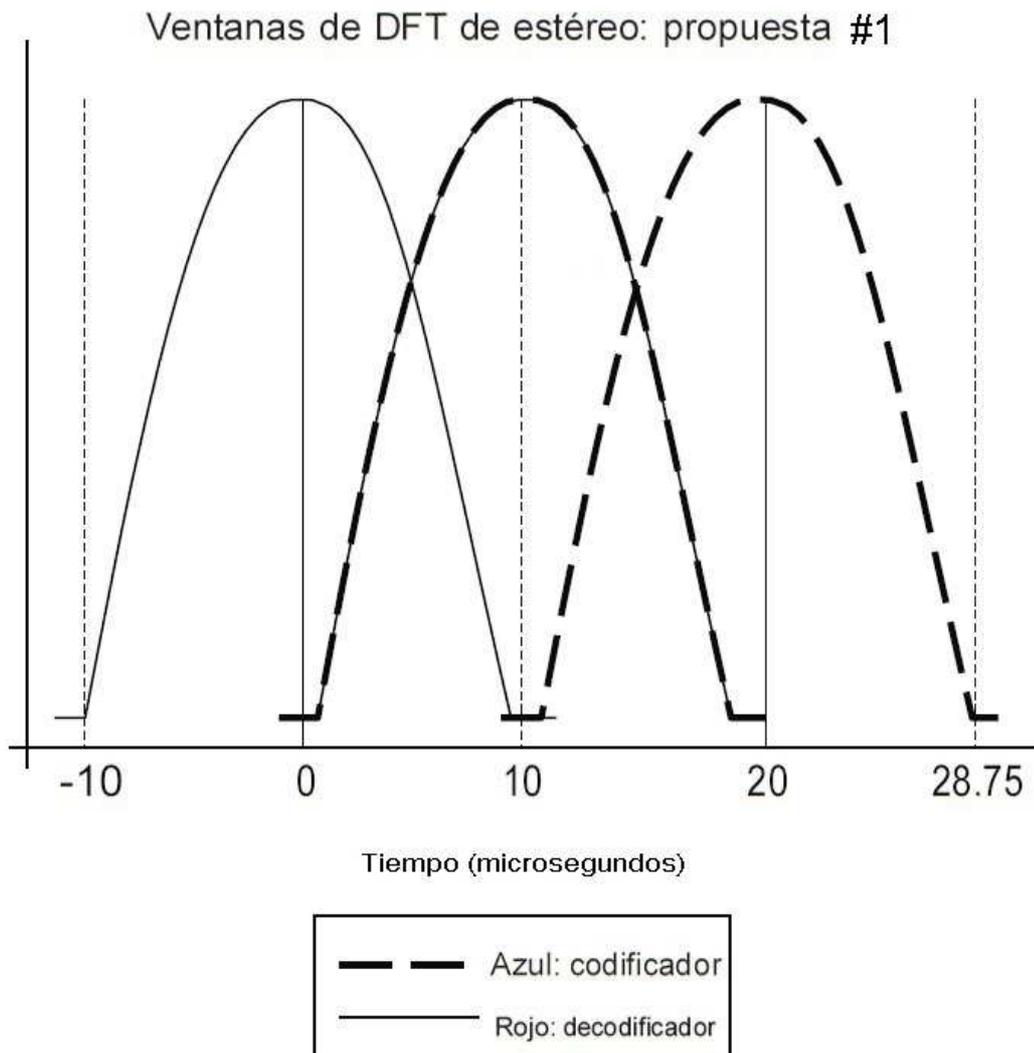
donde $ovlp_size$ es función de la velocidad de muestreo y está dada en la Figura 8A

Fig. 8c



Propuesta 1. División en ventanas esquemática del codificador

Fig. 9a



Propuesta 1. Ventanas en codificador y decodificador

Fig. 9c

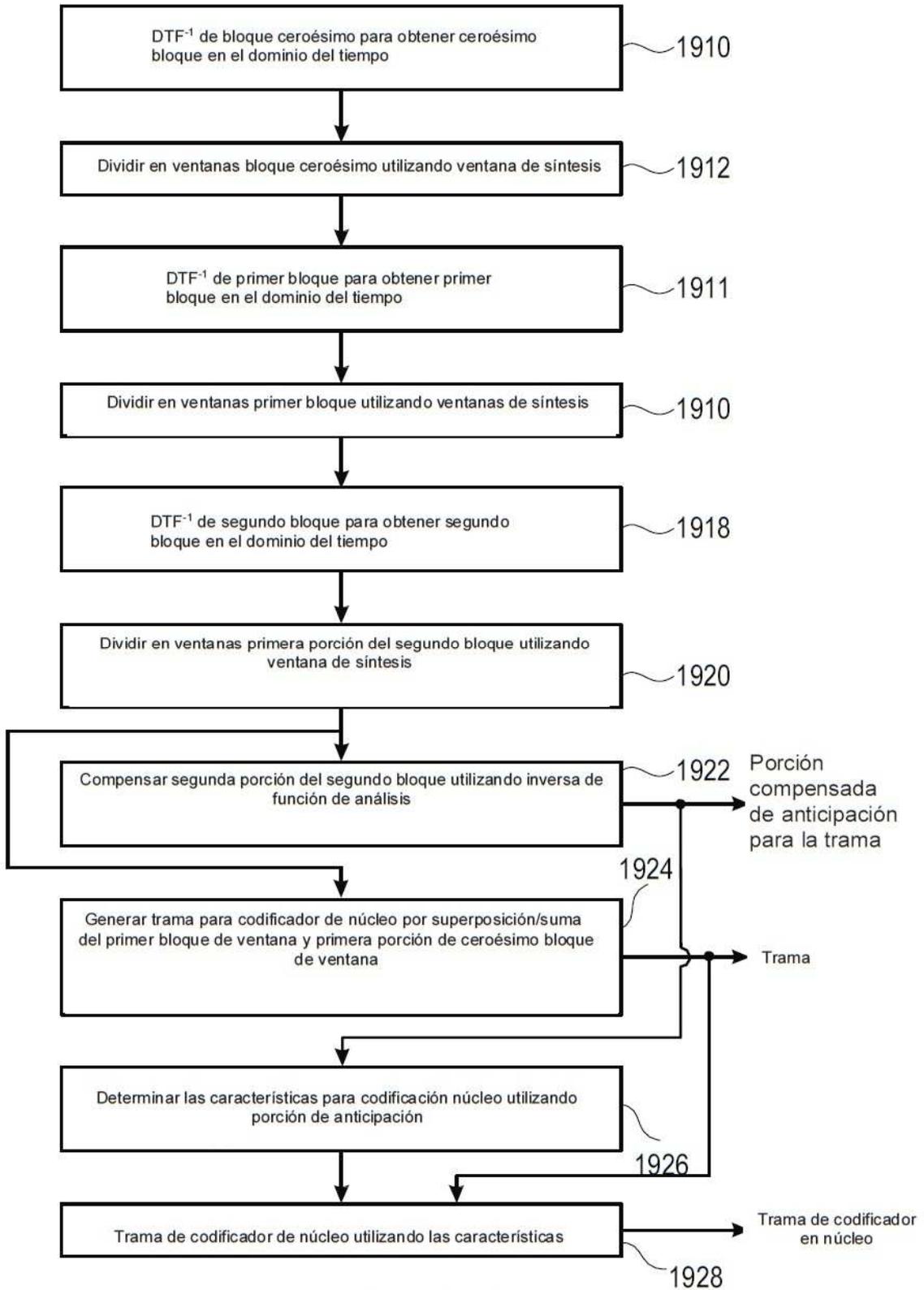


Fig. 9d

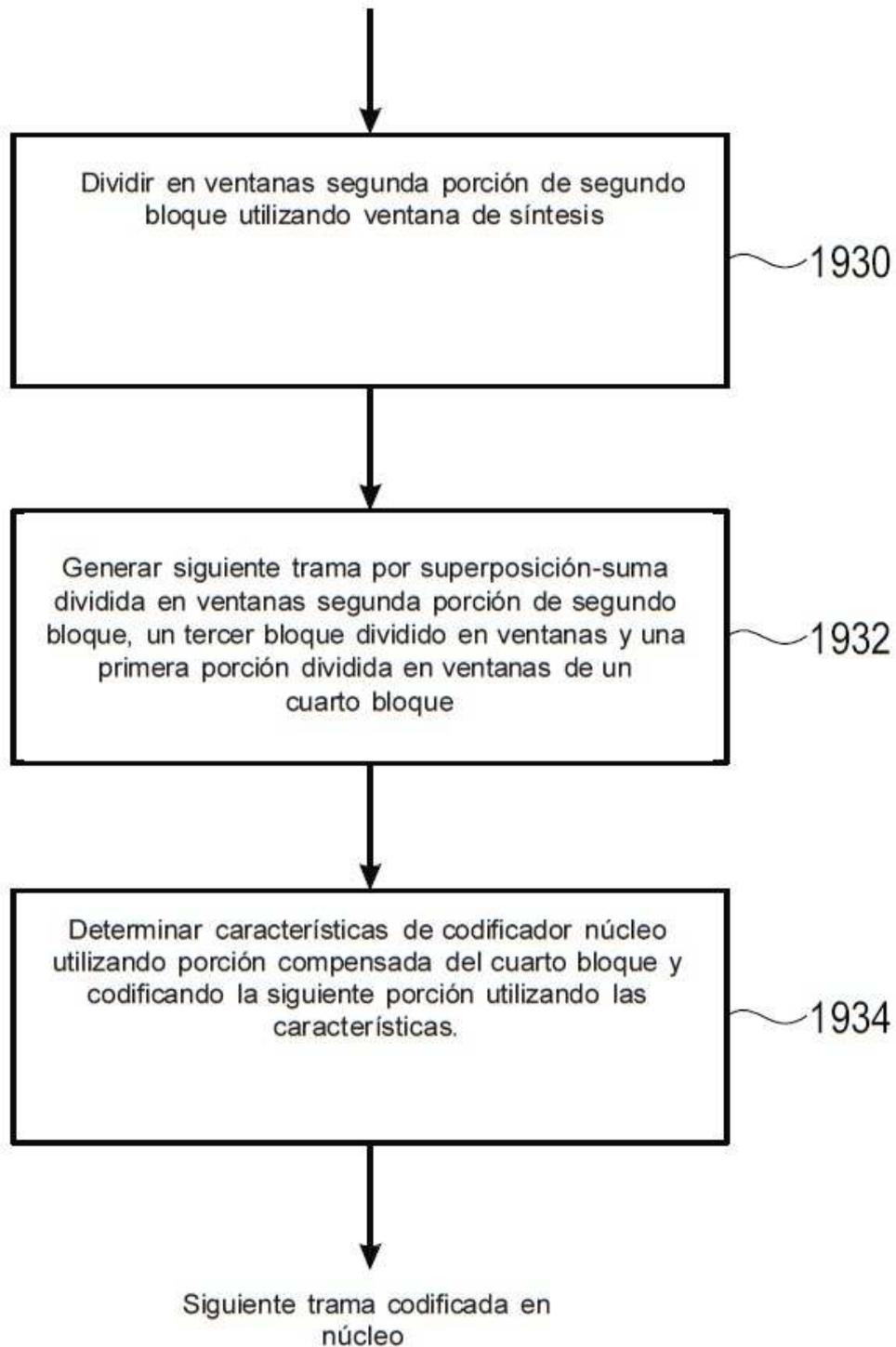


Fig. 9e

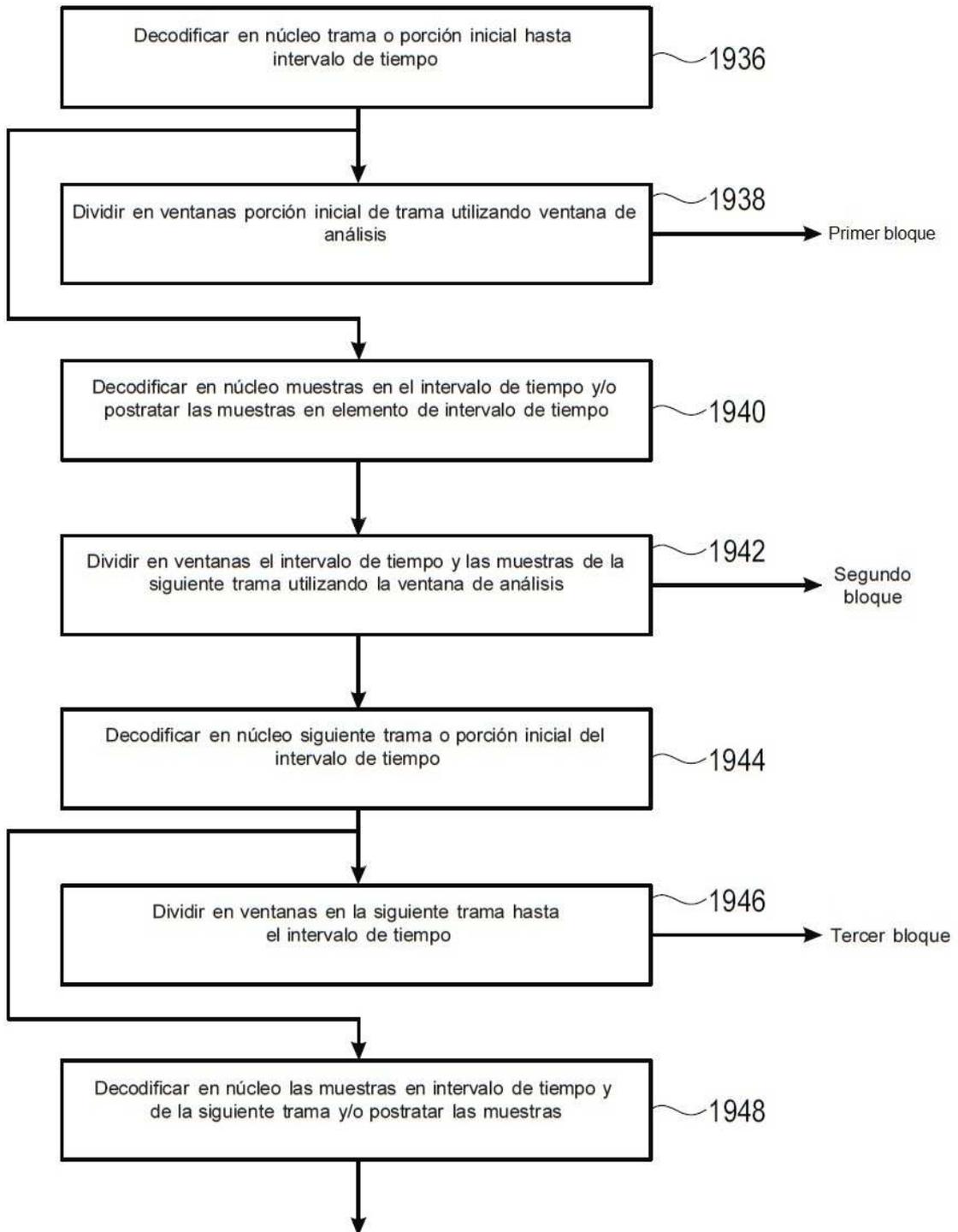
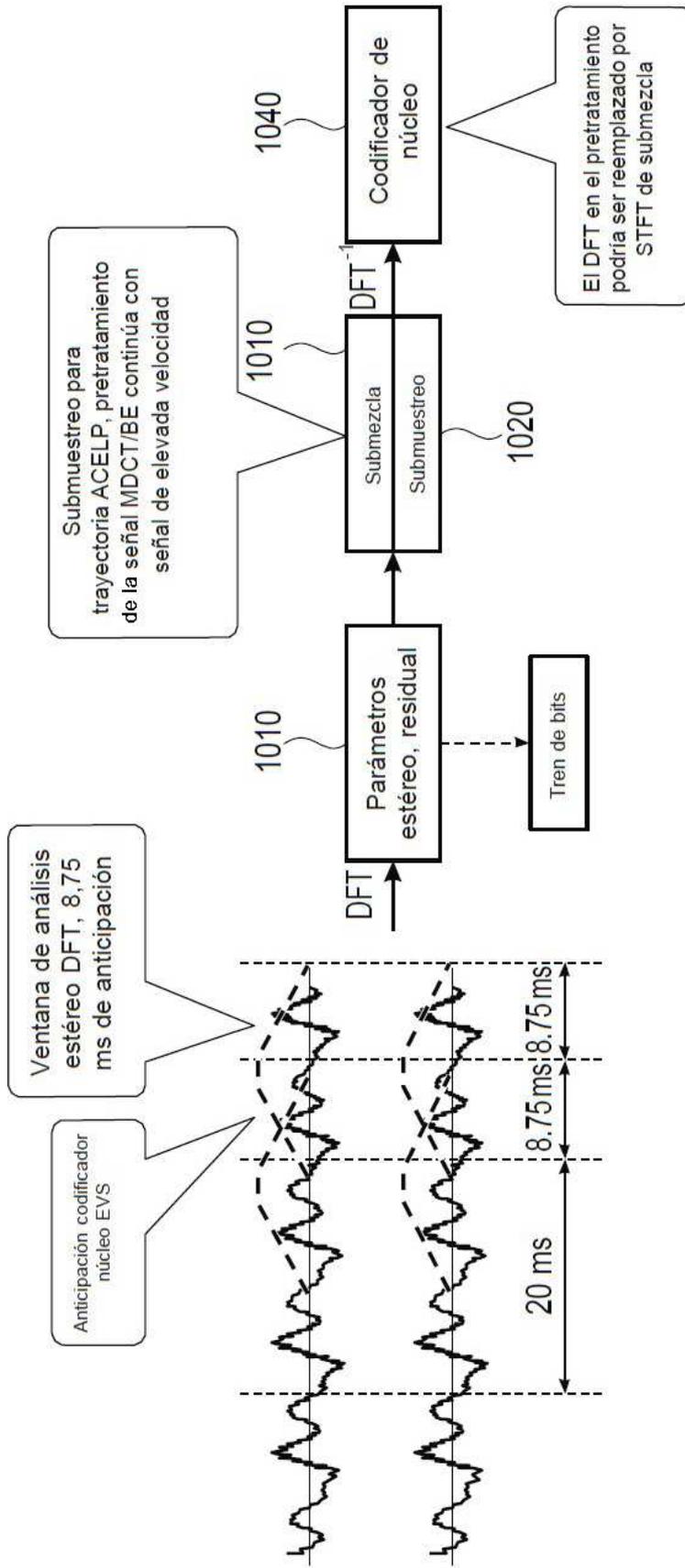
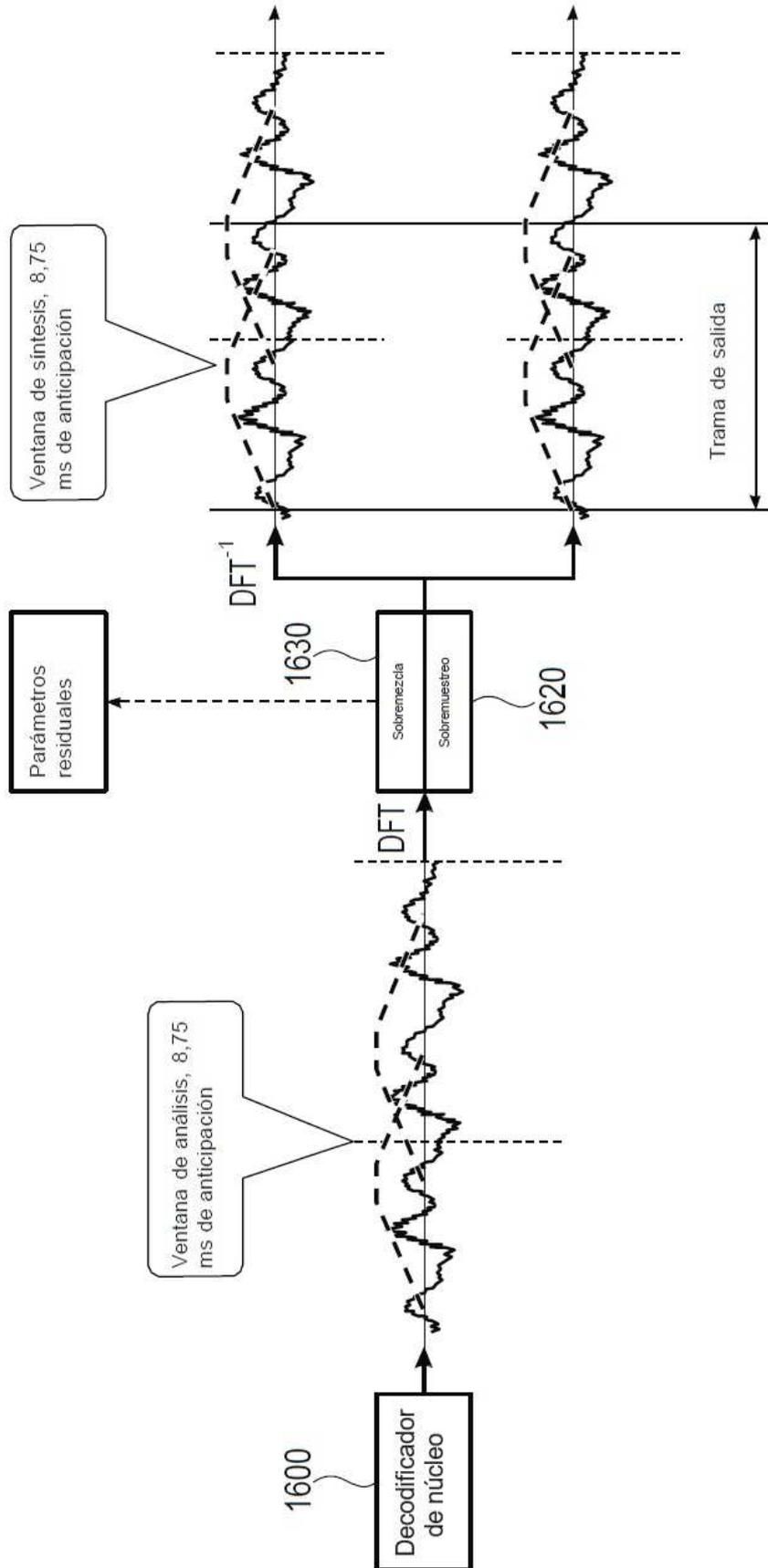


Fig. 9f



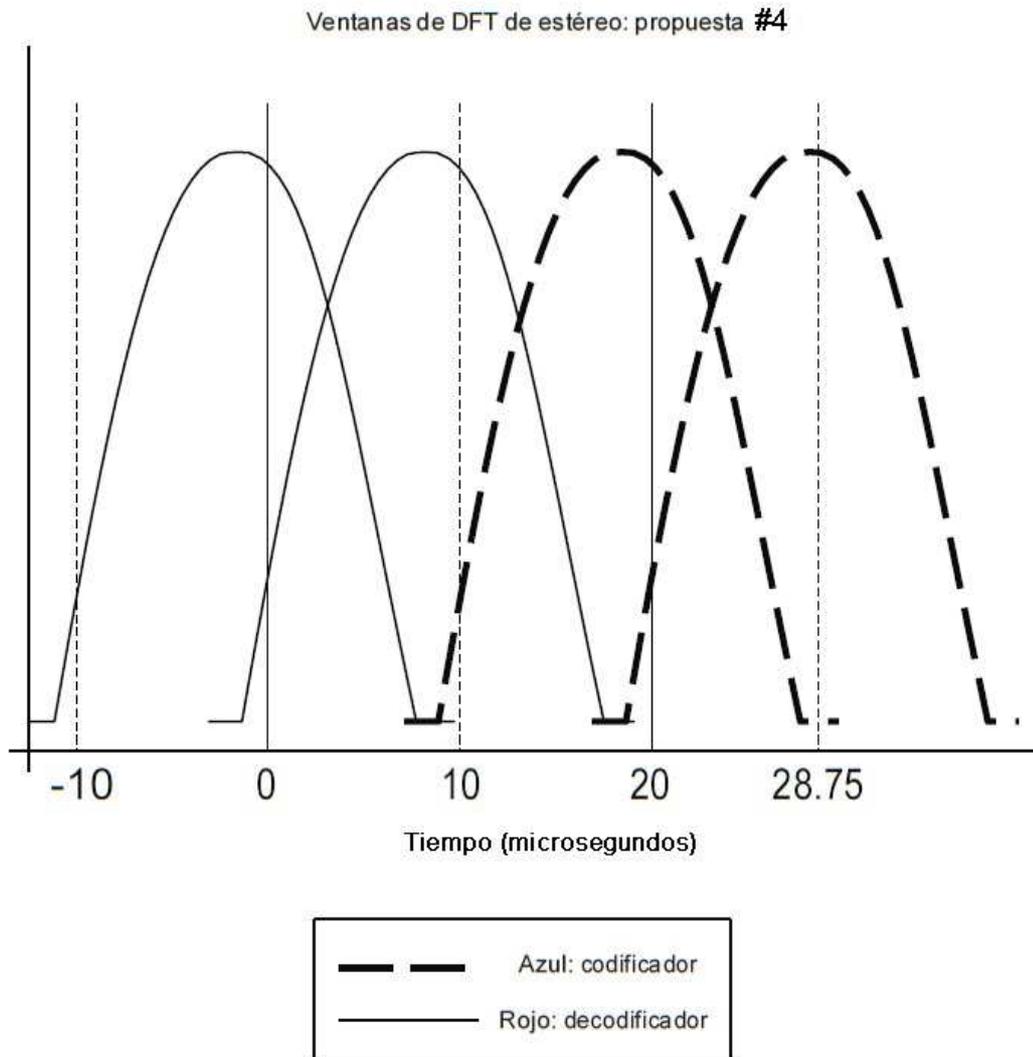
Propuesta 4. División en ventanas esquemática del codificador

Fig. 10a



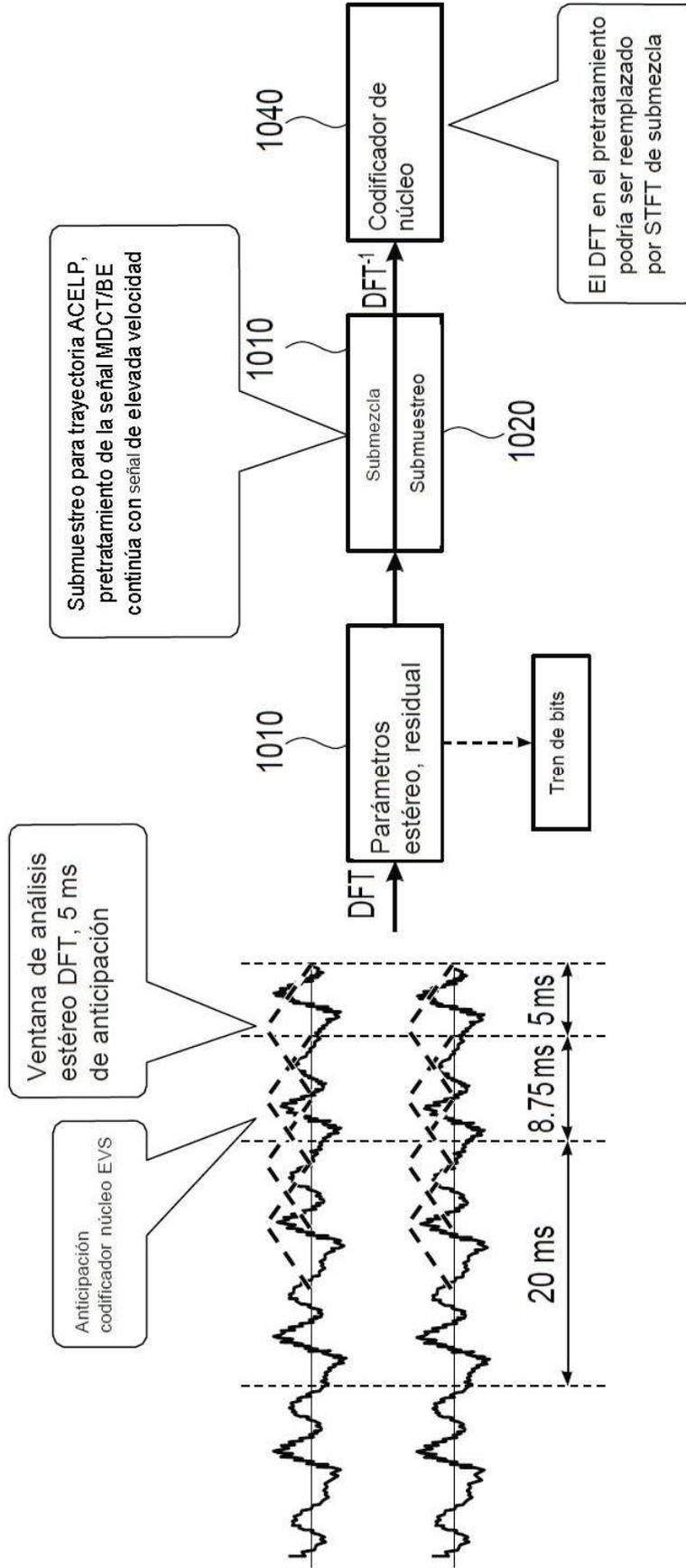
Propuesta 4. División en ventanas esquemática del decodificador

Fig. 10b



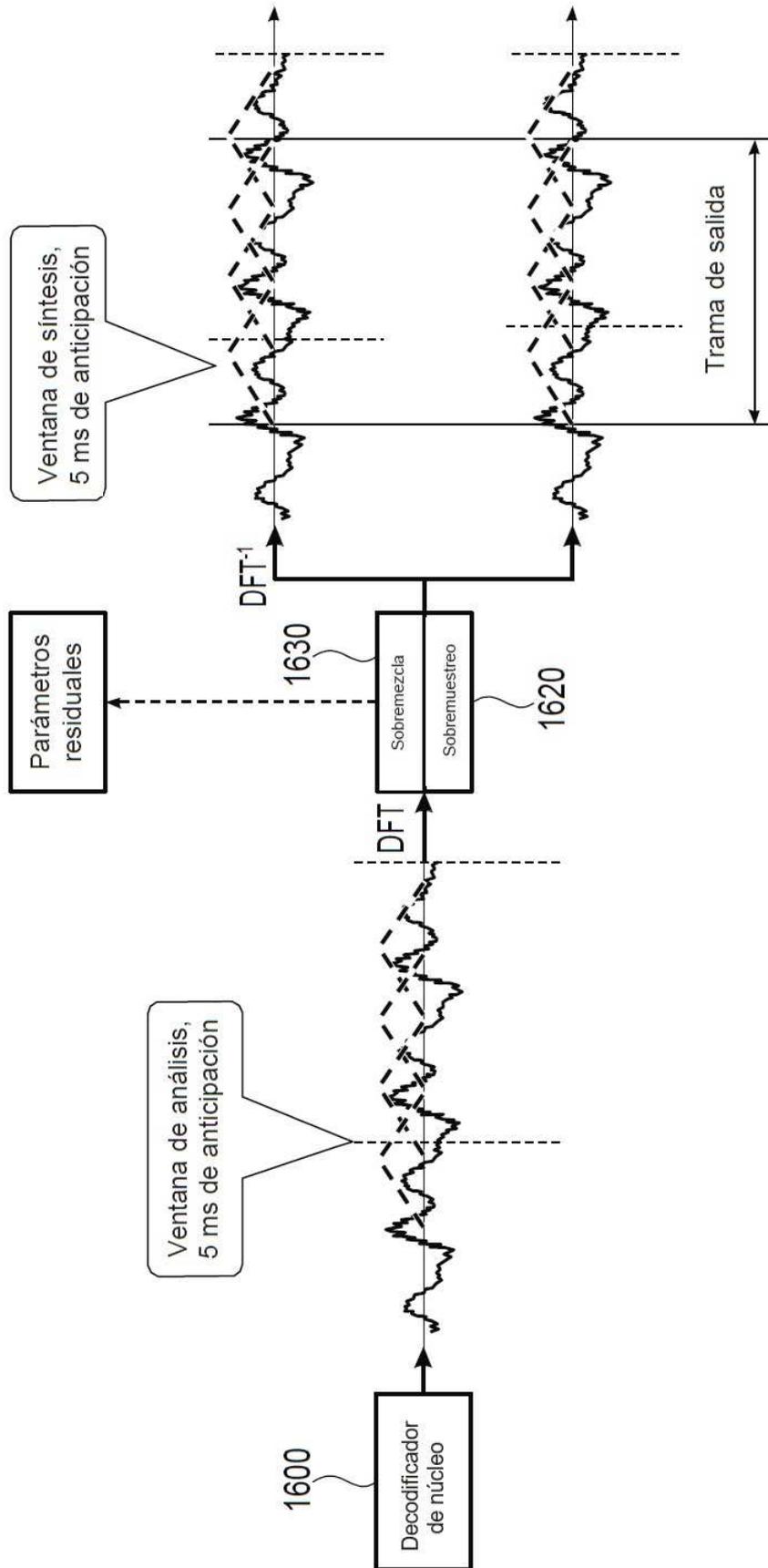
Propuesta 4. Ventanas en codificador y decodificador

Fig. 10c



Propuesta 5. División en ventanas esquemática del codificador

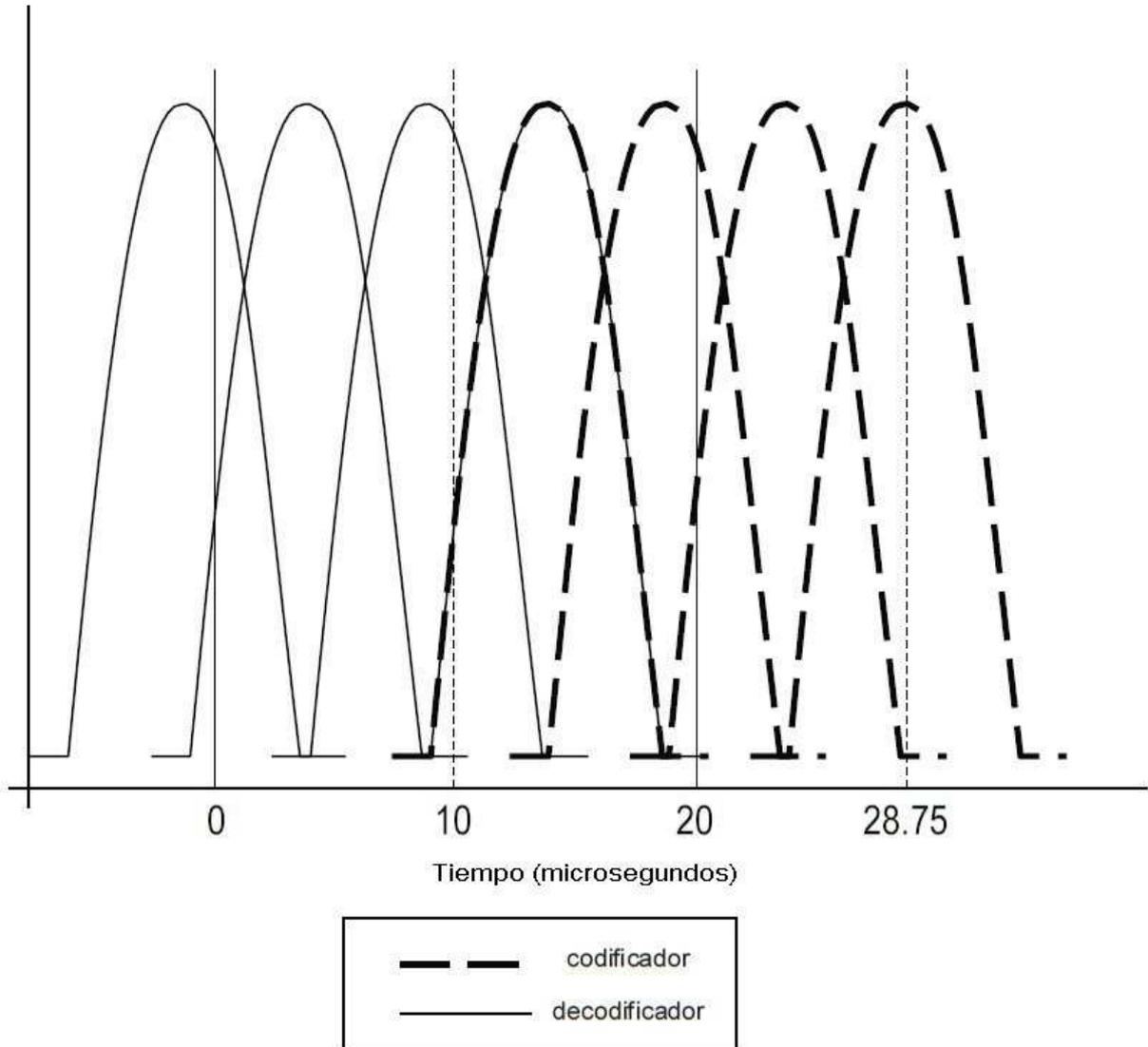
Fig. 11a



Propuesta 5. División en ventanas esquemática del decodificador

Fig. 11b

Ventana de DFT de estéreo: propuesta #5



Propuesta 5. Ventanas en codificador y decodificador

Fig. 11c

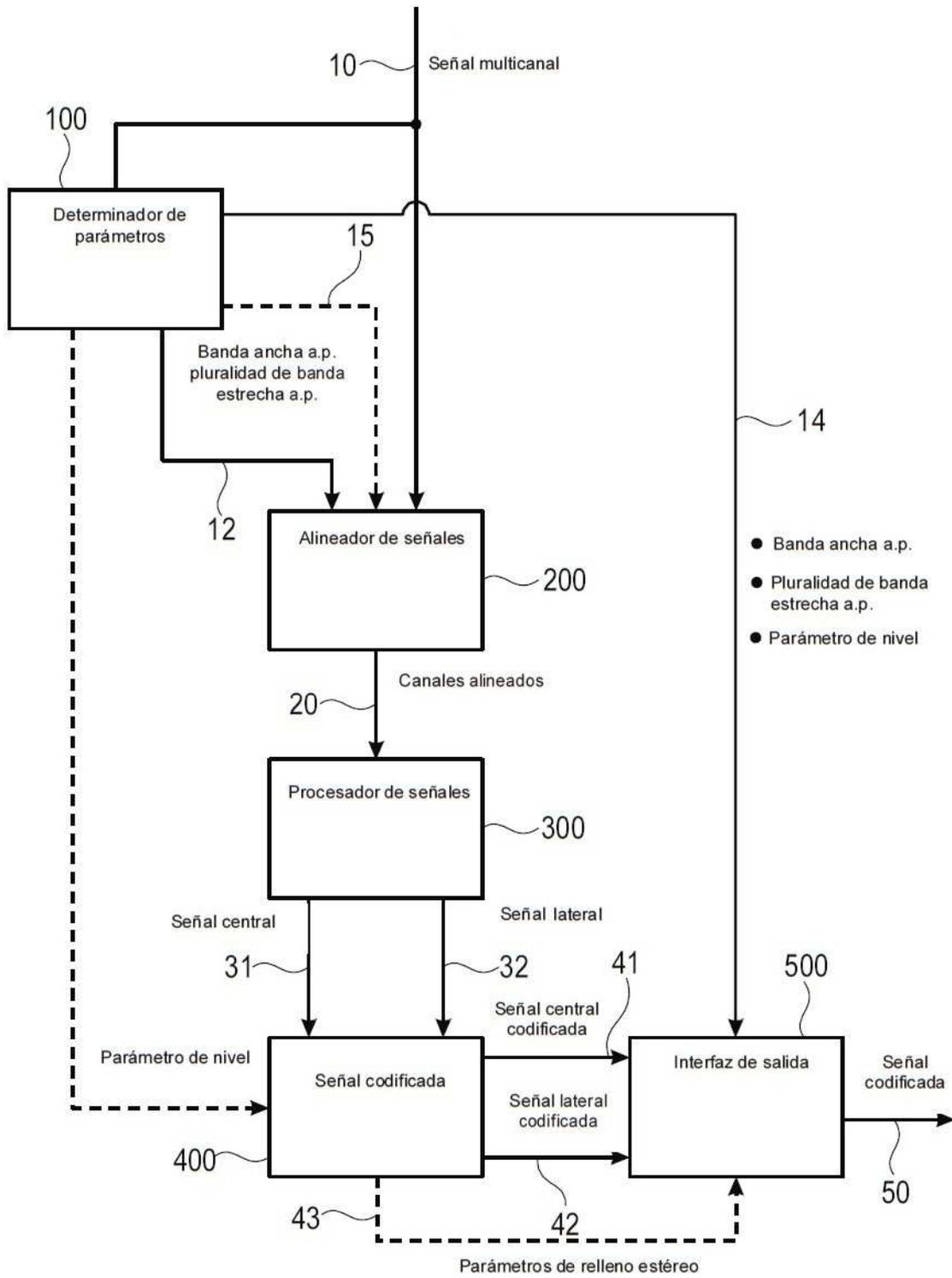


Fig. 12

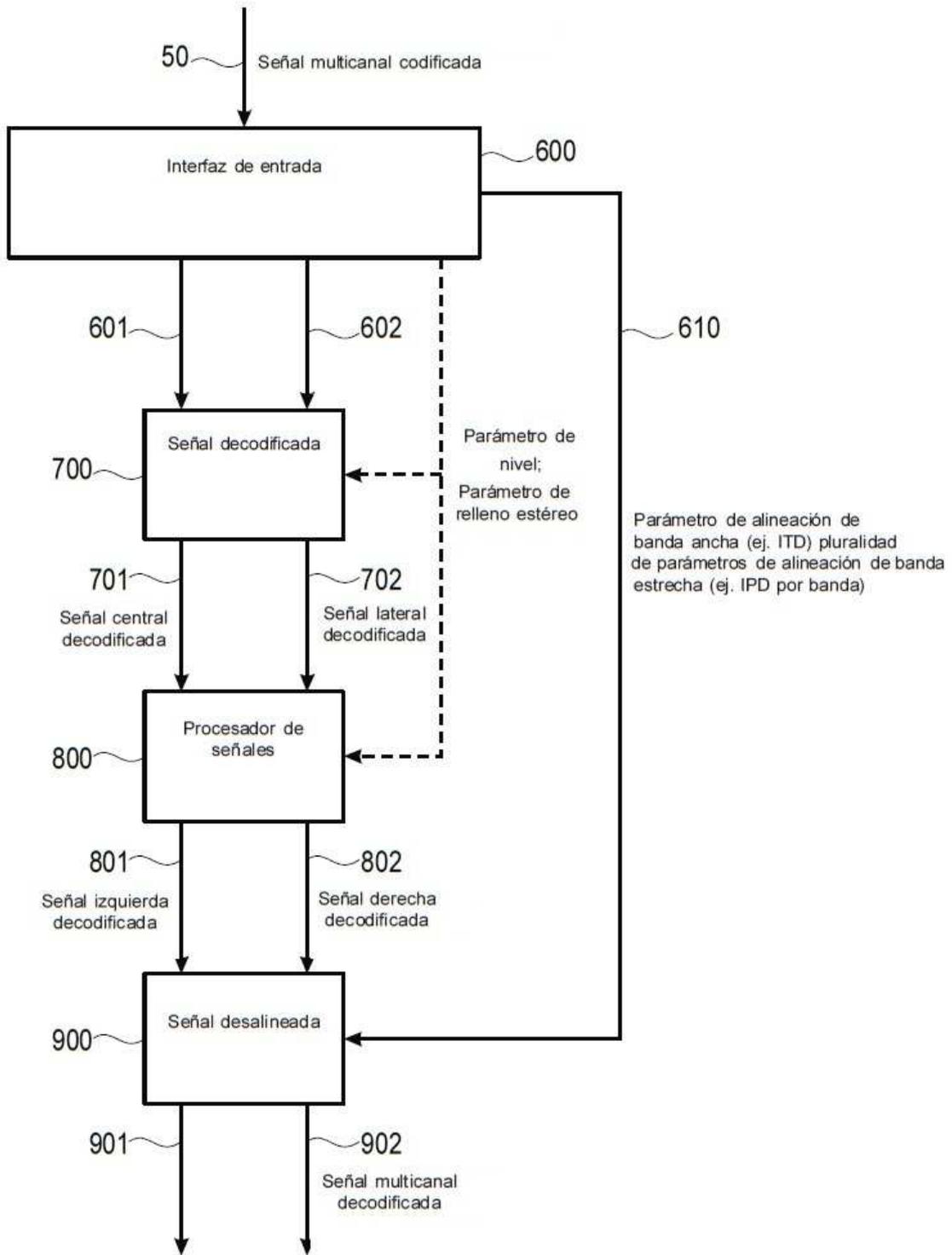


Fig. 13

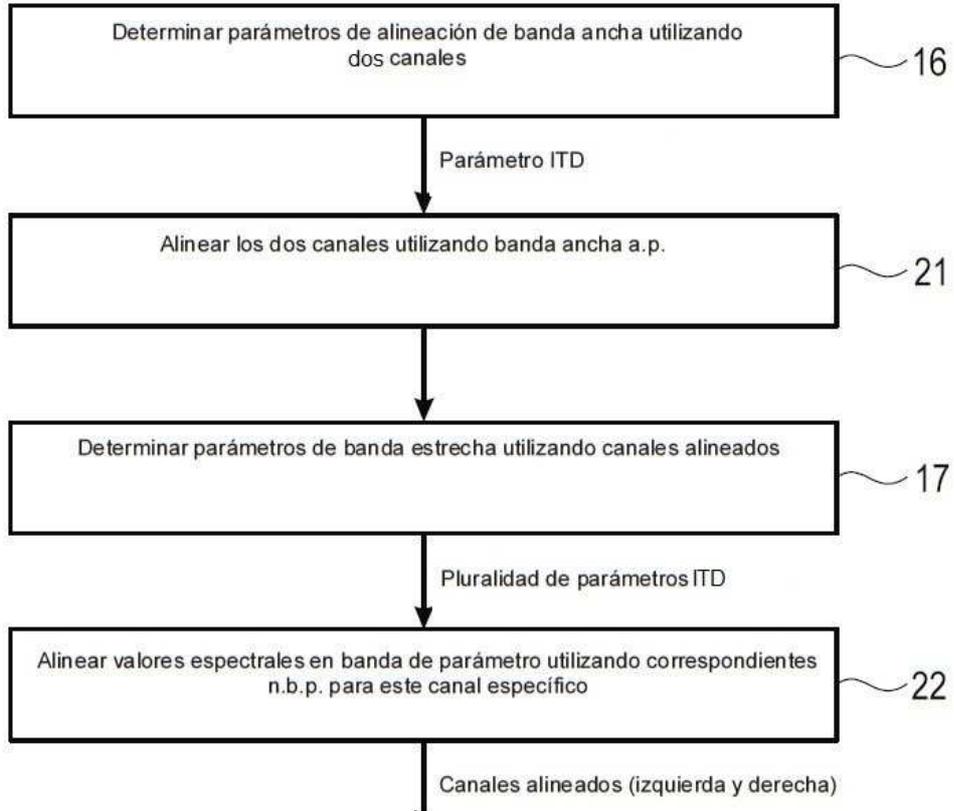


Fig. 14a

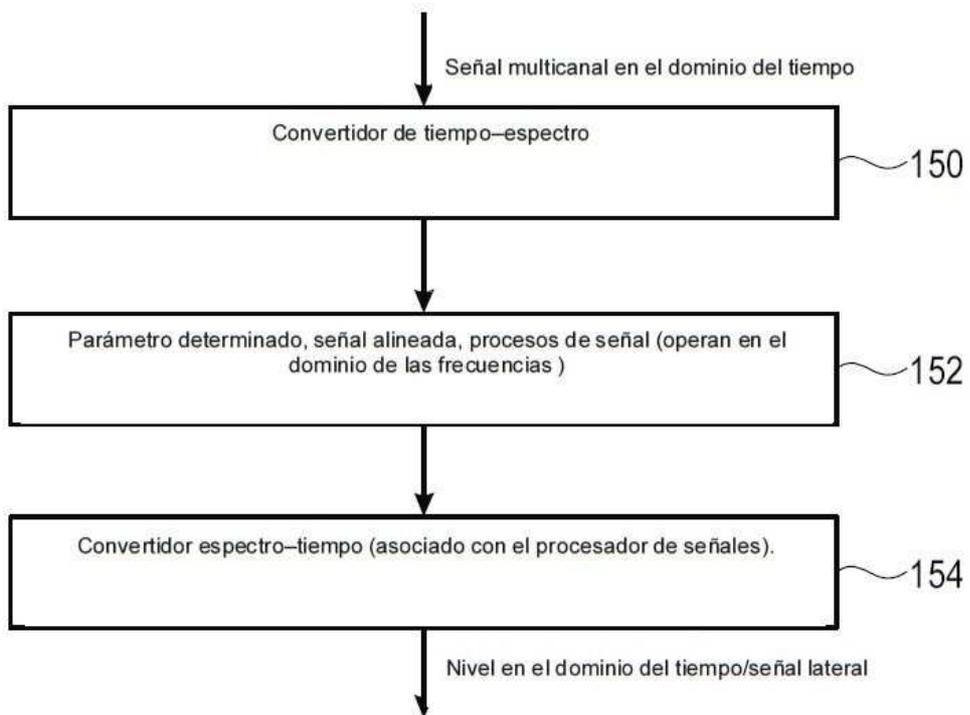


Fig. 14b

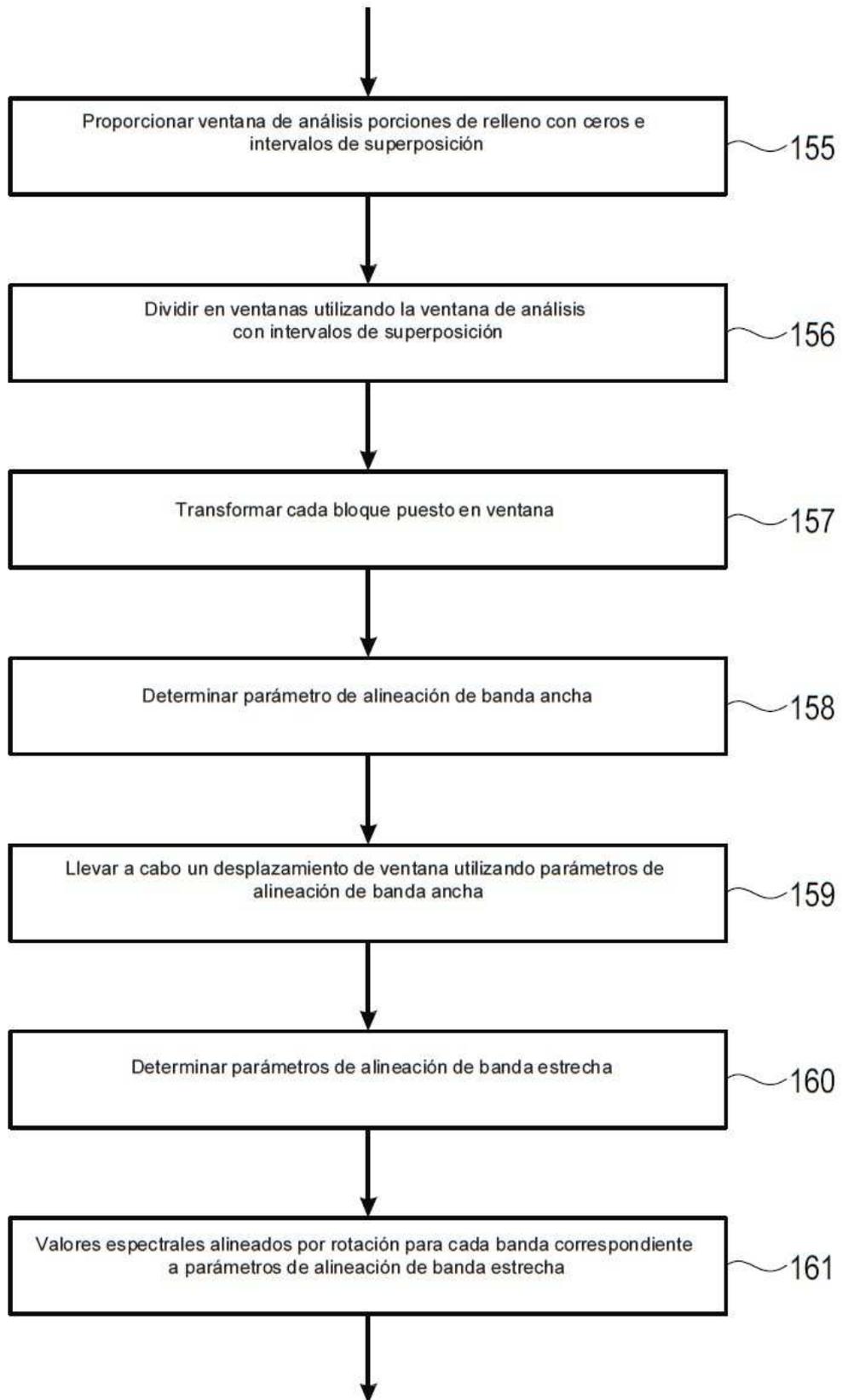


Fig. 14c

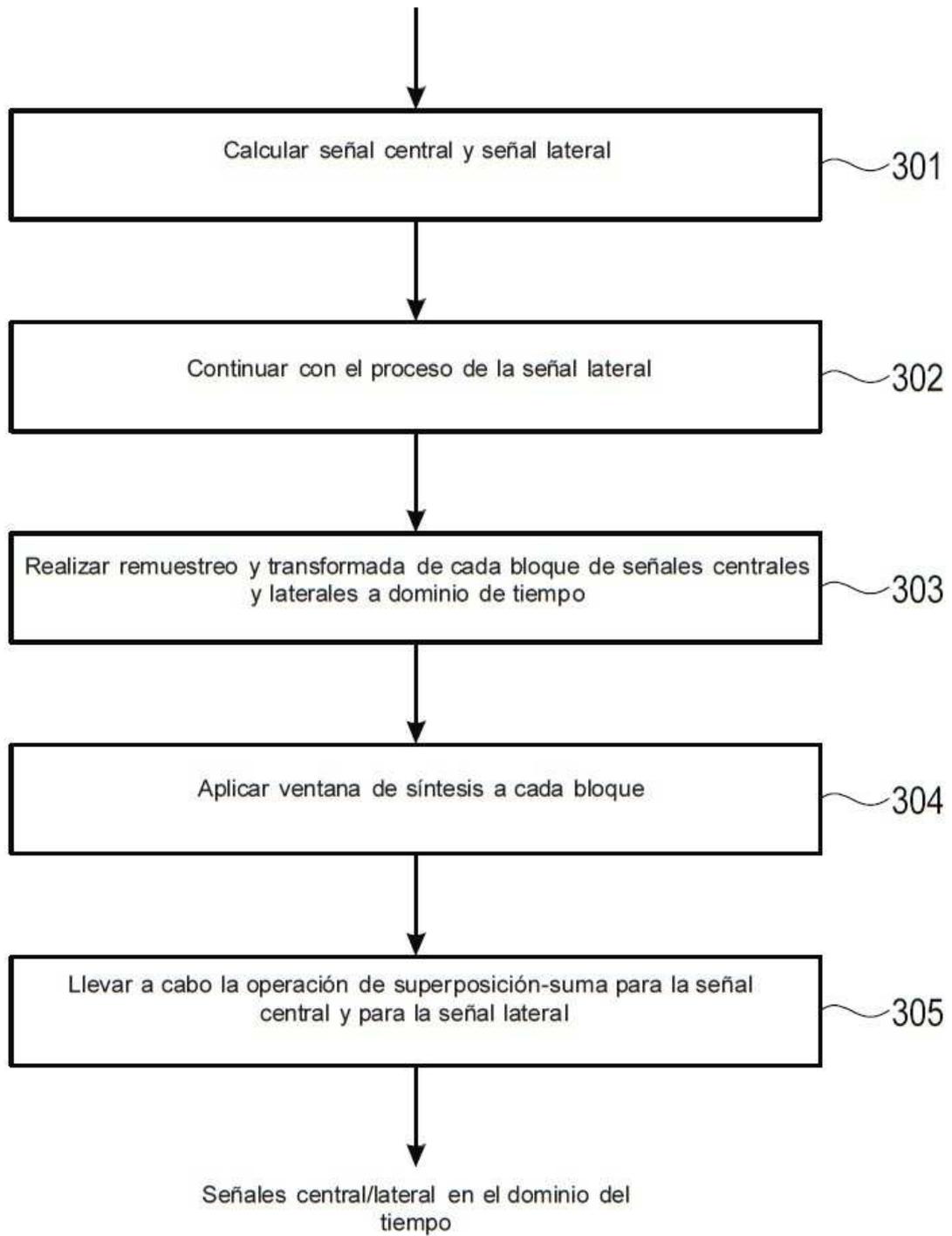


Fig. 14d

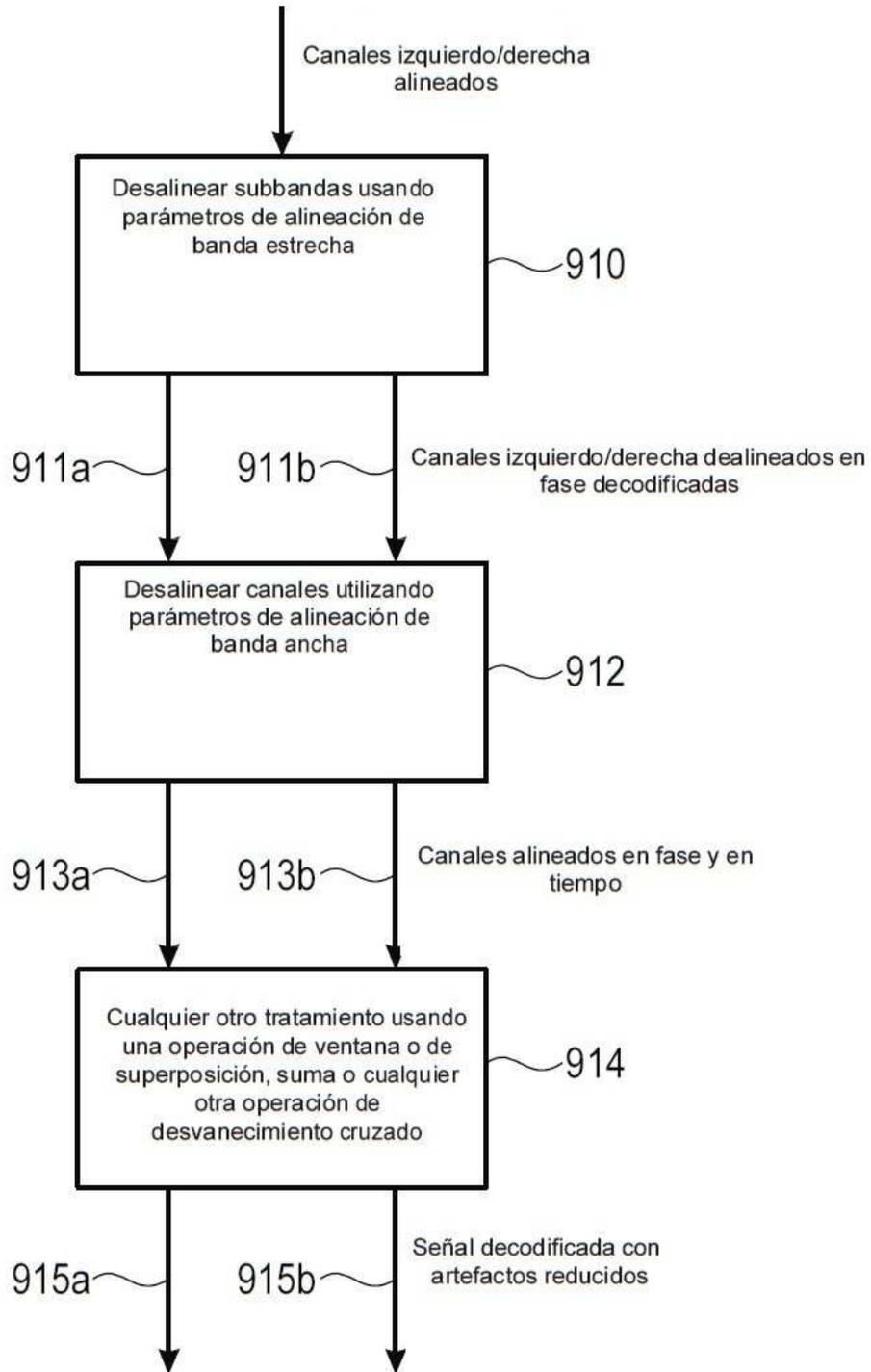


Fig. 15a

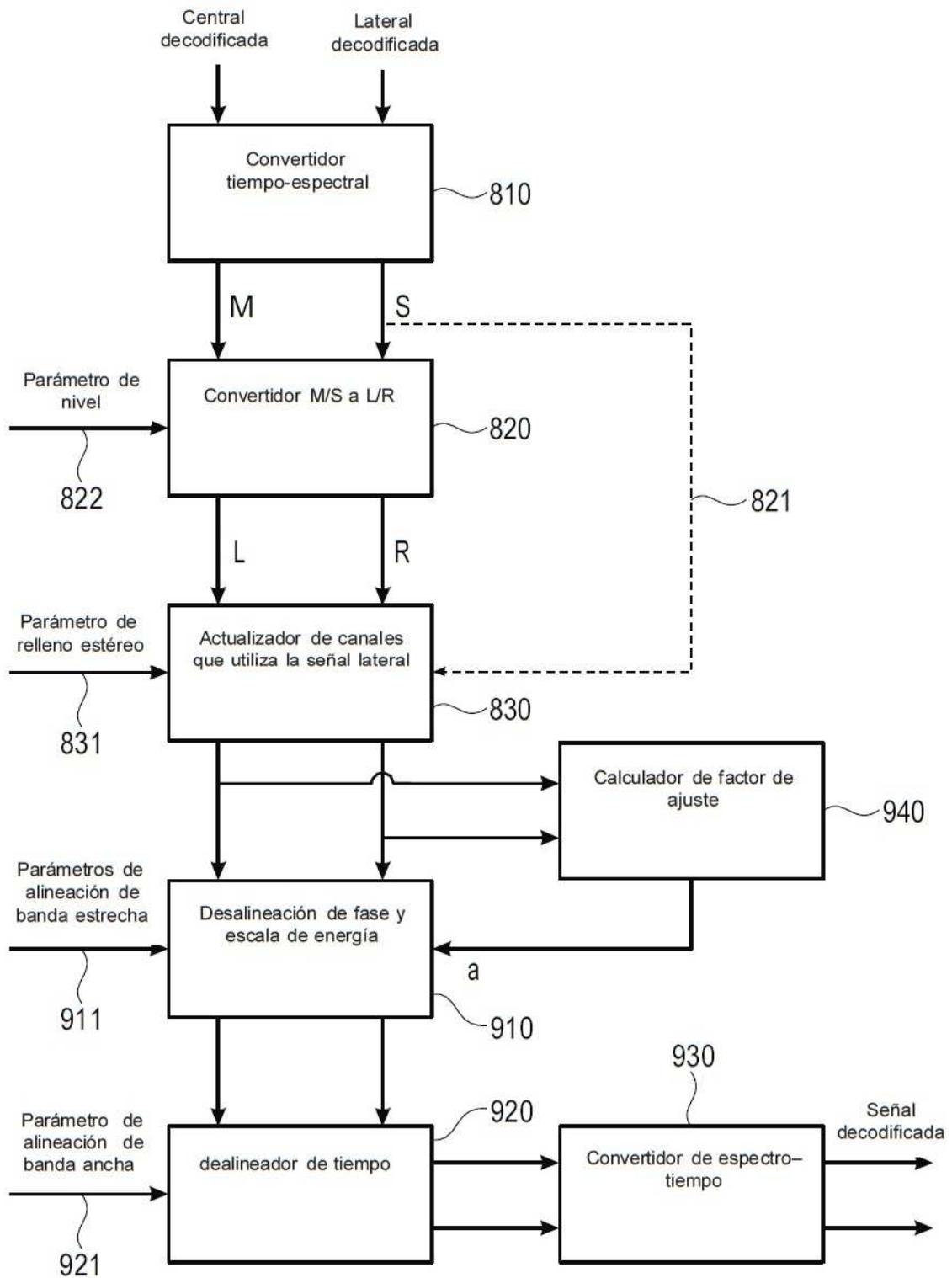


Fig. 15b

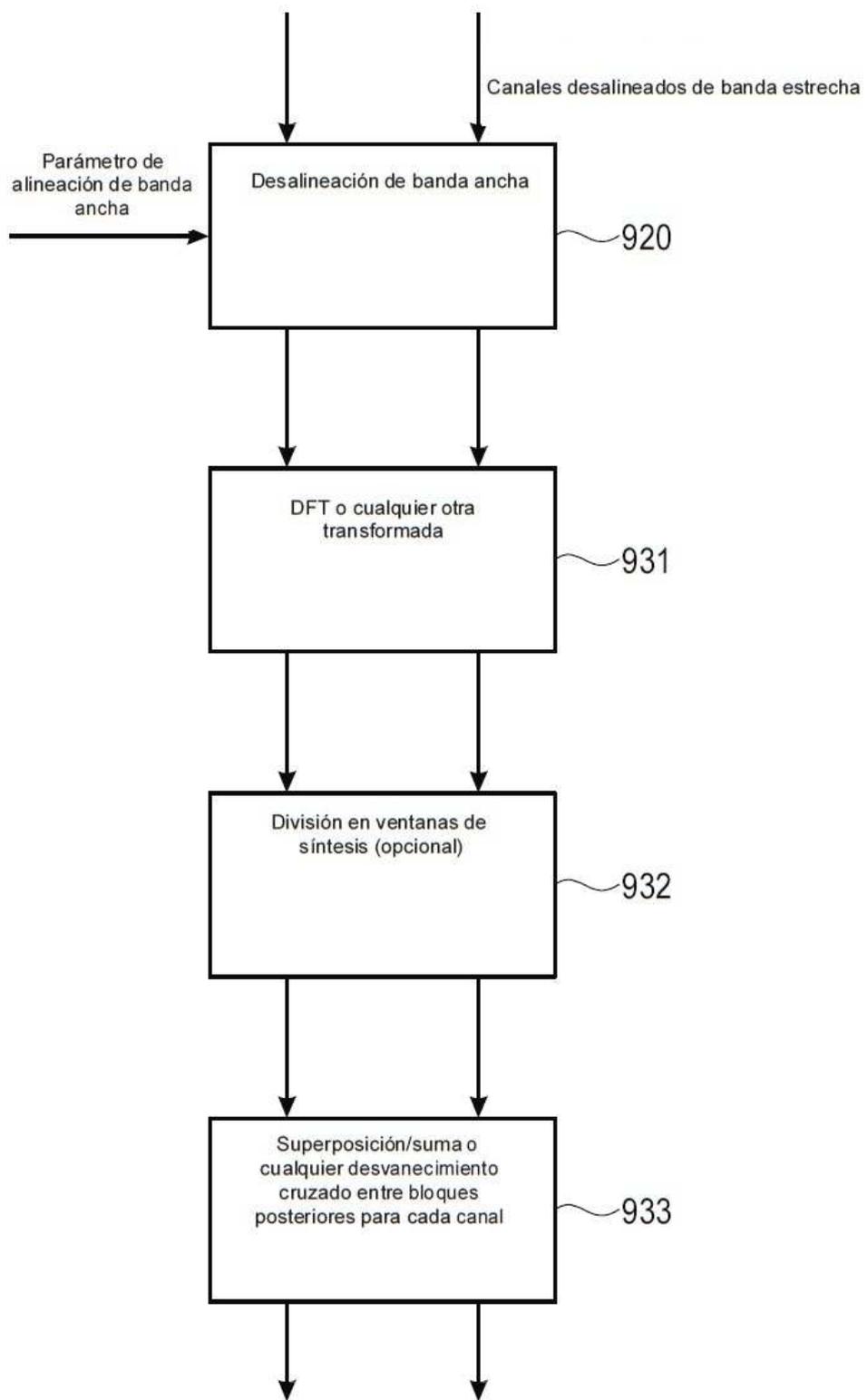


Fig. 15c