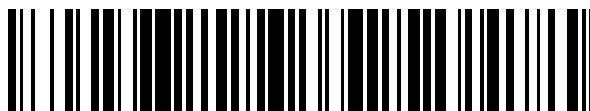


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 790 404**

51 Int. Cl.:

**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/EP2017/051205**

87 Fecha y número de publicación internacional: **27.07.2017 WO17125558**

96 Fecha de presentación y número de la solicitud europea: **20.01.2017 E 17700705 (1)**

97 Fecha y número de publicación de la concesión europea: **26.02.2020 EP 3405948**

54 Título: **Aparato y procedimiento para la codificación o decodificación de una señal de audio multi-canal mediante el uso de un parámetro de alineación de banda ancha y una pluralidad de parámetros de alineación de banda estrecha**

30 Prioridad:

**22.01.2016 EP 16152453  
22.01.2016 EP 16152450**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**27.10.2020**

73 Titular/es:

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

72 Inventor/es:

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

74 Agente/Representante:

**SALVÀ FERRER, Joan**

ES 2 790 404 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

Aparato y procedimiento para la codificación o decodificación de una señal de audio multi-canal mediante el uso de un parámetro de alineación de banda ancha y una pluralidad de parámetros de alineación de banda estrecha

5

**[0001]** La presente solicitud se refiere al procesamiento estéreo o, en general, al procesamiento multi-canal, donde una señal multi-canal 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]** La voz estéreo y, particularmente, la voz estéreo conversacional ha recibido mucha menos atención científica que el almacenamiento y la difusión de la música estereofónica. En efecto, la transmisión monofónica todavía se utiliza principalmente hoy en día en las comunicaciones de voz. No obstante, con el aumento del ancho de banda y la capacidad de las redes, se prevé que las comunicaciones basadas en tecnologías estereofónicas tendrán mayor difusión y proporcionarán una mejor experiencia auditiva.

15

**[0003]** La codificación eficaz de material de audio estereofónico ha sido estudiada durante mucho tiempo en la codificación de audio perceptual de la música para el almacenamiento y la difusión eficientes. Con altas tasas de bits, cuando es fundamental conservar la forma de onda, se ha empleado durante mucho tiempo el estéreo suma-resta, conocido como estéreo M/S (por sus siglas en inglés: *mid/side* / medio/lateral). Con bajas tasas de bits, se introdujo el estéreo en intensidad y, más recientemente, la codificación con estéreo paramétrico. Esta última técnica fue adoptada en diferentes estándares tales como HeAAC (por sus siglas en inglés: *High-Efficiency Advanced Audio Coding* / Codificador de Audio Avanzado de Alta Eficiencia)v2 y Mpeg (por sus siglas en inglés: *Moving Picture Experts Group* / Grupo de Expertos en Imágenes Móviles) USAC (por sus siglas en inglés: *Unified Speech and Audio Coding* / Codificación Unificada de Voz y Audio). Esta genera una mezcla de forma descendente de la señal de dos canales y asocia la información lateral espacial compacta.

20

**[0004]** La codificación de estéreo conjunto se construye usualmente sobre una resolución de alta frecuencia, es decir, baja resolución de tiempo, transformación tiempo-frecuencia de la señal y entonces no es compatible con el procesamiento con bajo retardo y en el dominio del tiempo que realiza la mayoría de los codificadores de voz. Además, la tasa de bits generada es usualmente alta.

30

**[0005]** Por otra parte, el estéreo paramétrico emplea un banco de filtros extra posicionado en el extremo delantero del codificador como pre-procesador y en el extremo trasero del decodificador como post-procesador. Por lo tanto, el estéreo paramétrico se puede usar con codificadores de voz convencionales como ACELP (por sus siglas en inglés: *Algebraic Code Excited Linear Prediction* / Predicción Lineal Excitada por Código Algebraico) ya que está realizado en MPEG USAC. Además, la parametrización del escenario auditivo se puede lograr con una cantidad mínima de información lateral, que es adecuada para bajas tasas de bits. No obstante, el estéreo paramétrico no está específicamente diseñado, por ejemplo, en MPEG USAC, para bajo retardo y no entrega una calidad homogénea en diferentes escenarios conversacionales. En la representación paramétrica convencional de la escena espacial, el ancho de la imagen estéreo se reproduce artificialmente mediante un decorrelacionador aplicado en los dos canales sintetizados y controlado por parámetros ICs (por sus siglas en inglés: *Inter-channel Coherence* / Coherencia entre Canales) computados y transmitidos por el codificador. Para la mayoría de la voz estéreo, esta forma de ensanchar la imagen estéreo no es adecuada para recrear el efecto ambiental natural de la voz que es un sonido bastante directo dado que está producido por una única fuente ubicada en una posición específica en el espacio (a veces, con alguna reverberación procedente de la habitación). Por el contrario, los instrumentos musicales tienen mucha más anchura natural que la voz, y se pueden imitar mejor mediante la decorrelación de los canales.

35

40

45

**[0006]** También se presentan inconvenientes cuando la voz se registra con micrófonos no coincidentes, como en la configuración A-B cuando los micrófonos están distantes o en el registro o la representación binaural. Tales escenarios se pueden prever en la captura de voz en teleconferencias o en la creación de un escenario auditivo virtual con oradores distantes en la MCU (por sus siglas en inglés: *Multipoint Control Unit* / Unidad de Control Multipunto). El momento de llegada de la señal es entonces diferente entre un canal y el otro a diferencia de los registros realizados en micrófonos coincidentes como X-Y (registro de intensidad) o M-S (registro medio-lateral). A continuación, la computación de la coherencia de esos dos canales no alineados en el tiempo se puede estimar de forma incorrecta lo cual hace fracasar la síntesis artificial del efecto ambiental.

50

55

**[0007]** Las referencias de la técnica anterior relacionadas con el procesamiento estéreo son la Patente de EE.UU. 5.434.948 o la Patente de EE.UU. 8.811.621. El documento US 2009/0313028 A1 describe un ejemplo adicional de un aparato para realizar el procesamiento de audio de una señal de audio de entrada multi-canal.

60

**[0008]** El documento WO 2006/089570 A1 describe un esquema de codificador/decodificador multi-canal transparente o casi transparente. Un esquema de codificador/decodificador multi-canal genera además una señal residual del tipo forma de onda. Esta señal residual se transmite junto con uno o más parámetros multi-canal a un decodificador. Contrariamente a un decodificador multi-canal puramente paramétrico, el decodificador mejorado genera una señal de salida multi-canal que tiene una calidad de salida mejorada debido a la señal residual adicional.

65

Del lado del codificador, tanto un canal izquierdo como un canal derecho son filtrados por un banco de filtros de análisis. A continuación, para cada señal sub-banda, se calculan un valor de alineación y un valor de ganancia para una sub-banda. Tal alineación se realiza entonces antes del procesamiento adicional. Del lado del decodificador, se realizan una desalineación y un procesamiento de ganancia y las señales correspondientes 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]** Se ha encontrado que tales procedimientos de la técnica anterior no proporcionan un óptimo para las señales de audio y, específicamente, para las señales de voz donde hay más de un orador, es decir, en un escenario de conferencia o una escena de conversación.

**[0010]** Es un objeto de la presente invención proporcionar un concepto mejorado para codificar o decodificar una señal multi-canal.

**[0011]** Este objeto se consigue mediante un aparato para codificar una señal de audio multi-canal de la reivindicación 1, un procedimiento para codificar una señal de audio multi-canal de la reivindicación 19, una señal de audio multi-canal codificada de la reivindicación 20, un aparato para decodificar una señal de audio multi-canal codificada de la reivindicación 21 o un procedimiento de decodificación de una señal de audio multi-canal codificada de la reivindicación 33 o un programa informático de la reivindicación 34. Realizaciones específicas se definen en las reivindicaciones dependientes.

**[0012]** Un aparato para codificar una señal multi-canal que tiene al menos dos canales comprende un determinador de parámetros para determinar un parámetro de alineación de banda ancha por una parte y una pluralidad de parámetros de alineación de banda estrecha por la otra. Estos parámetros son utilizados por un alineador de señales para alinear los al menos dos canales usando estos parámetros para obtener canales alineados. A continuación, un procesador de señal calcula una señal central y una señal lateral utilizando los canales alineados y la señal central y la señal lateral a continuación son codificadas y enviadas en una señal de salida codificada que además tiene, como la información lateral paramétrica, el parámetro de alineación de banda ancha y la pluralidad de parámetros de alineación de banda estrecha.

**[0013]** Del lado del decodificador, un decodificador de señal decodifica la señal central codificada y la señal lateral codificada para obtener una señal central y una señal lateral decodificada. Estas señales se procesan a continuación mediante un procesador de señales para calcular un primer canal decodificado y un segundo canal decodificado. Estos canales decodificados se desalinean a continuación usando la información sobre el parámetro de alineación de banda ancha y la información sobre la pluralidad de parámetros de banda estrecha incluidos en una señal multi-canal codificada para obtener la señal multi-canal decodificada.

**[0014]** En una implementación específica, el parámetro de alineación de banda ancha es un parámetro de diferencia de tiempos entre canales y la pluralidad de parámetros de alineación de banda estrecha son diferencias de fases entre canales.

**[0015]** La presente invención se basa en el hallazgo de que específicamente para señales de voz donde hay más de un orador, pero también para otras señales de audio donde hay varias fuentes de audio, los diferentes lugares de las fuentes de audio que se mapean en dos canales de la señal multi-canal se pueden tener en cuenta para usar un parámetro de alineación de banda ancha tal como un parámetro de diferencia de tiempos entre canales que se aplica a la totalidad del espectro tanto de uno como de ambos canales. Además de este parámetro de alineación de banda ancha, se ha descubierto que varios parámetros de alineación de banda estrecha que difieren de sub-banda a sub-banda además resultan en una mejor alineación de la señal en ambos canales.

**[0016]** Por lo tanto, una alineación de banda ancha correspondiente al mismo retardo en cada sub-banda junto con una alineación de fase correspondiente a diferentes rotaciones de fase para diferentes sub-bandas resulta en una alineación óptima de ambos canales antes de que estos dos canales se conviertan a continuación en una representación central/lateral que se codifica además a continuación. Debido al hecho de que se ha obtenido una alineación óptima, la energía en la señal central es tan alta como es posible, por una parte, y, por otra parte, la energía en la señal lateral es tan pequeña como es posible, de modo que se puede obtener un resultado óptimo de codificación con una tasa de bits tan baja como es posible o una calidad de audio tan alta como es posible para una determinada tasa de bits.

**[0017]** Específicamente para un material de voz conversacional, parece que existen generalmente oradores activos en dos lugares diferentes. Además, la situación es tal que, normalmente, solo un orador está hablando desde el primer lugar y, a continuación, el segundo orador está hablando desde el segundo lugar o ubicación. La influencia de las diferentes ubicaciones en los dos canales tales como un primer canal o canal izquierdo y un segundo canal o canal derecho se refleja mediante diferentes tiempos de llegada y, por lo tanto, un cierto retardo entre ambos canales debido a las diferentes ubicaciones, y este retardo cambia periódicamente. Generalmente, esta influencia se refleja en las señales de los dos canales como una desalineación de banda ancha que se puede corregir mediante el

parámetro de alineación de banda ancha.

**[0018]** Por otra parte, se pueden considerar otros efectos, particularmente procedentes de la reverberación o fuentes adicionales de ruido, mediante parámetros individuales de alineación de fase para bandas individuales que están superpuestas en los diferentes tiempos de llegada de banda ancha o la desalineación de banda ancha de ambos canales.

**[0019]** En vista de esto, el uso de ambos, un parámetro de alineación de banda ancha y una pluralidad de parámetros de alineación de banda estrecha además del parámetro de alineación de banda ancha resulta en una alineación de canales óptima en el lado del codificador para obtener una buena y muy compacta representación central/lateral mientras que, por otra parte, la desalineación correspondiente subsiguiente a una decodificación en el lado del decodificador resulta en una buena calidad de audio para una determinada tasa de bits o en una pequeña tasa de bits para una determinada calidad de audio requerida.

**[0020]** Una ventaja de la presente invención es que esta proporciona un nuevo esquema de codificación estéreo mucho más adecuado para una conversión de la voz estéreo que los esquemas de codificación estéreo existentes. De conformidad con la invención, las tecnologías de estéreo paramétrico y las tecnologías de codificación de estéreo conjunto se combinan de forma particular al explotar la diferencia de tiempos entre canales que se produce en los canales de una señal multi-canal específicamente en el caso de fuentes de voz, pero también en el caso de otras fuentes de audio.

**[0021]** Diversas realizaciones proporcionan ventajas útiles, tal como se expondrá más adelante.

**[0022]** El nuevo procedimiento es una estrategia híbrida que mezcla elementos de un estéreo M/S convencional y un estéreo paramétrico. En un M/S convencional, los canales se mezclan pasivamente de forma descendente para generar una señal central y una lateral. El proceso se puede extender adicionalmente al rotar el canal usando una transformada de Karhunen-Loeve (KLT), también conocida como PCA (por sus siglas en inglés: Principal Component Analysis / Análisis de Componentes Principales) antes de sumar y diferenciar los canales. La señal central está codificada en una codificación de código primario mientras que la lateral se transmite a un codificador secundario. El estéreo M/S desarrollado puede además usar la predicción de la señal lateral por el canal central codificado en la trama presente o en la anterior. El objetivo principal de la rotación y la predicción es maximizar la energía de la señal central al mismo tiempo que se minimiza la energía de la lateral. El estéreo M/S preserva la forma de onda y en este aspecto es muy sólido ante cualquier escenario estéreo, pero puede ser muy costoso en términos de consumo de bits.

**[0023]** Para una mayor eficacia con baja tasa de bits, el estéreo paramétrico computa y codifica parámetros, como ILDs (por sus siglas en inglés: Inter-channel Level Differences / Diferencias de Niveles entre Canales), IPDs (por sus siglas en inglés: Inter-channel Phase Differences / Diferencias de Fases entre Canales), ITDs (por sus siglas en inglés: Inter-channel Time Differences / Diferencias de Tiempos entre Canales) e IC. Estos representan de manera compacta la imagen estéreo y son pistas del escenario auditivo (localización de la fuente, paneo, ancho del estéreo, etc.). El objetivo es, a continuación, parametrizar el escenario estéreo y codificar solo una señal de mezcla descendente que puede estar en el decodificador y, con la ayuda de las pistas estéreo transmitidas, espacializarla de nuevo.

**[0024]** Nuestra estrategia mezcló los dos conceptos. En primer lugar, las pistas ITD e IPD se computan y aplican a los dos canales. El objetivo es representar la diferencia de tiempo en banda ancha y la diferencia de fases en bandas con diferente frecuencia. Los dos canales se alinean a continuación en tiempo y fase y, a continuación, se realiza la codificación M/S. Se ha encontrado que las ITD e IPD fueron útiles para modelar la voz estéreo y son un buen reemplazo de la rotación basada en KLT en M/S. A diferencia de una codificación puramente paramétrica, el efecto ambiental no es más modelado por las IC sino directamente por la señal lateral que se codifica y/o se predice. Se ha encontrado que esta estrategia es más sólida especialmente cuando se gestionan señales de voz.

**[0025]** La computación y el procesamiento de las ITD es una parte fundamental de la invención. Las ITD ya fueron explotadas en la técnica anterior de BCC (por sus siglas en inglés: Binaural Cue Coding / Codificación de Referencia Binaural), pero de una forma que resultaba ineficaz ya que las ITD cambian a lo largo del tiempo. Para evitar este inconveniente, se diseñó una formación de ventana específica para suavizar las transiciones entre dos ITD diferentes y poder cambiar fluidamente entre un orador y otro posicionados en lugares diferentes.

**[0026]** Realizaciones adicionales están relacionadas con el procedimiento en las que, del lado del codificador, la determinación de parámetros para determinar la pluralidad de parámetros de alineación de banda estrecha se realiza mediante el uso de canales que ya han sido alineados con el parámetro de alineación de banda ancha determinado anteriormente.

**[0027]** De manera correspondiente, la desalineación de banda estrecha en el lado del decodificador se realiza antes de realizar la desalineación de banda ancha mediante el uso del generalmente único parámetro de alineación de banda ancha.

**[0028]** En realizaciones adicionales, se prefiere que, ya sea del lado del codificador, pero incluso más importante, del lado del decodificador, se realice alguna clase de operación de formación de ventana y adición de superposición o alguna clase de desvanecimiento cruzado desde un bloque al siguiente a continuación de todas las alineaciones y, específicamente, a continuación de una alineación en el tiempo mediante el uso del parámetro de alineación de banda ancha. Esto evita cualquier defecto audible tal como clics cuando el parámetro de alineación en el tiempo o de banda ancha cambia de bloque en bloque.

**[0029]** En otras realizaciones se aplican diferentes resoluciones espectrales. En particular, las señales de los canales se someten a una conversión espectral en el tiempo con una resolución de alta frecuencia tal como un espectro DFT (por sus siglas en inglés: Discrete Fourier Transform / Transformada Discreta de Fourier) mientras que los parámetros tales como los parámetros de alineación de banda estrecha se determinan para bandas de parámetros que tienen una resolución espectral más baja. Generalmente, una banda de parámetros tiene más de una línea espectral que el espectro de la señal y generalmente tiene un conjunto de líneas espectrales del espectro DFT. Asimismo, las bandas de parámetros aumentan desde bajas frecuencias hasta altas frecuencias con el fin de considerar problemas psicoacústicos.

**[0030]** Realizaciones adicionales se refieren al uso adicional de un parámetro de nivel tal como una diferencia entre niveles u otros procedimientos para procesar la señal lateral tales como parámetros de relleno estéreo, etc. La señal codificada se puede representar mediante la señal lateral real misma, o mediante una señal residual de predicción que se realiza usando la señal central de la trama actual o cualquier otra trama, o mediante una señal lateral o una señal residual de predicción lateral solo en un subconjunto de bandas y parámetros de predicción solo para las bandas restantes, o inclusive mediante parámetros de predicción para todas las bandas sin ninguna información de señal lateral de resolución de alta frecuencia. Por lo tanto, en la última alternativa mencionada, la señal lateral codificada solo es representada por un parámetro de predicción para cada banda de parámetros o solo un subconjunto de bandas de parámetro de modo que para las bandas de parámetros restantes no existe ninguna información sobre la señal lateral original.

**[0031]** Asimismo, es preferible tener la pluralidad de parámetros de alineación de banda estrecha de modo que no todas las bandas de parámetros reflejen el ancho de banda completo de la señal de banda ancha sino solo un conjunto de bandas bajas tal como el 50 por ciento más bajo de las bandas de parámetros. Por otra parte, los parámetros de relleno estéreo no se usan para el par de bandas bajas, dado que, para estas bandas, se transmite la propia señal lateral o una señal residual de predicción a fin de asegurar que, al menos para las bandas bajas, esté disponible una representación correcta de forma de onda. Por otra parte, la señal lateral no se transmite en una representación exacta de forma de onda para las bandas altas a fin de disminuir todavía más la tasa de bits, pero la señal lateral se representa generalmente mediante parámetros de relleno estéreo.

**[0032]** Asimismo, es preferible realizar el análisis de parámetros y la alineación completos dentro de uno y el mismo dominio de la frecuencia sobre la base del mismo espectro DFT. Con este fin, es preferible usar además la tecnología GCC-PHAT con el fin de determinar la diferencia de tiempos entre canales. En una realización preferida de este procedimiento, el suavizado de un espectro de correlación sobre la base de una información sobre una forma espectral, siendo la información preferentemente una medición de aplanamiento espectral, se realiza de tal forma que un suavizado será débil en el caso de señales del tipo ruido y un suavizado será más marcado en el caso de señales del tipo tono.

**[0033]** Asimismo, es preferible realizar una rotación de fase especial, donde se tengan en cuenta las amplitudes de los canales. Particularmente, la rotación de fase se distribuye entre los dos canales con fines de alineación en el lado del codificador, y naturalmente, con fines de desalineación en el lado del decodificador, donde un canal que tiene una amplitud más alta se considera como un canal principal y se verá menos afectado por la rotación de fase, es decir, será menos rotado que un canal con una amplitud más baja.

**[0034]** Asimismo, el cálculo de suma-resta se realiza usando un escalamiento de la energía con un factor de escala que se deriva de las energías de ambos canales y además está limitado a un cierto intervalo a fin de asegurar que el cálculo centro/lateral no esté afectando demasiado a la energía. Por otra parte, no obstante, cabe destacar que, para el propósito de la presente invención, esta clase de conservación de la energía no es tan crucial como en los procedimientos de la técnica anterior, dado que el tiempo y la fase fueron alineados de antemano. Por lo tanto, las fluctuaciones de energía debido al cálculo de una señal central y una señal lateral de izquierda y derecha (del lado del codificador) o debido al cálculo de una señal izquierda y derecha desde la señal central y lateral (del lado del decodificador) no son tan importantes como en la técnica anterior.

**[0035]** A continuación, se exponen realizaciones preferidas de la presente invención con respecto a los dibujos adjuntos, en los cuales:

La Fig.1 es un diagrama de bloques de una implementación preferida de un aparato para codificar una señal multi-canal;

La Fig. 2 es una realización preferida de un aparato para decodificar una señal multi-canal codificada;

La Fig. 3 es una ilustración de diferentes resoluciones de frecuencia y otros aspectos relacionados con la frecuencia  
5 para ciertas realizaciones;

La Fig. 4a ilustra un diagrama de flujo de los procedimientos realizados en el aparato para codificar con el fin de alinear los canales;

10 La Fig. 4b ilustra una realización preferida de los procedimientos realizados en el dominio de la frecuencia;

La Fig. 4c ilustra una realización preferida de los procedimientos realizados en el aparato para codificar usando una ventana de análisis con porciones de atenuación cero e intervalos de superposición;

15 La Fig. 4d ilustra un diagrama de flujo para procedimientos adicionales realizados dentro del aparato para codificar;

La Fig. 4e ilustra un diagrama de flujo para mostrar una implementación preferida de una estimación de diferencia de tiempos entre canales;

20 La Fig. 5 ilustra un diagrama de flujo que ilustra una realización adicional de procedimientos realizados en el aparato para codificación;

La Fig. 6a ilustra un cuadro de bloques de una realización de un codificador;

25 La Fig. 6b ilustra un diagrama de flujo de una realización correspondiente de un decodificador;

La Fig. 7 ilustra un escenario de ventana preferido con ventanas seno con baja superposición con atenuación cero para un análisis y síntesis tiempo-frecuencia en estéreo;

30 La Fig. 8 ilustra una tabla que muestra el consumo de bits de diferentes valores de parámetros;

La Fig. 9a ilustra los procedimientos realizados por un aparato para decodificar una señal multi-canal codificada en una realización preferida;

35 La Fig. 9b ilustra una implementación preferida del aparato para decodificar una señal multi-canal codificada; y

La Fig. 9c ilustra un procedimiento realizado en el contexto de una desalineación de banda ancha en el contexto de la decodificación de una señal multi-canal codificada.

40 **[0036]** La Fig. 1 ilustra un aparato para la codificación de una señal multi-canal que tiene al menos dos canales. La señal multi-canal 10 es introducida en un determinador de parámetro 100 por una parte y un alineador de señales 200 por otra parte. El determinador de parámetro 100 determina, por una parte, un parámetro de alineación de banda ancha y, por otra parte, una pluralidad de parámetros de alineación de banda estrecha de la señal multi-canal. Estos parámetros se emiten a través de una línea de parámetros 12. Asimismo, estos parámetros también se emiten a través  
45 de una línea de parámetros adicional 14 a una interfaz de salida 500 tal como se ilustra. En la línea de parámetros 14, se envían parámetros adicionales tales como los parámetros de nivel desde el determinador de parámetros 100 hasta la interfaz de salida 500. El alineador de señales 200 está configurado para alinear los al menos dos canales de la señal multi-canal 10 usando el parámetro de alineación de banda ancha y la pluralidad de parámetros de alineación de banda estrecha recibidos a través de la línea de parámetros 10 para obtener canales alineados 20 en la salida del  
50 alineador de señales 200. Estos canales alineados 20 se envían a un procesador de señales 300 que está configurado para calcular una señal central 31 y una señal lateral 32 de los canales alineados recibidos a través de la línea 20. El aparato para codificación comprende además un codificador de señales 400 para codificar la señal central desde la línea 31 y la señal lateral desde la línea 32 para obtener una señal central codificada en la línea 41 y una señal lateral codificada en la línea 42. Ambas señales son enviadas a la interfaz de salida 500 para generar una señal multi-canal  
55 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 estrecha y los parámetros de alineación de banda ancha de la línea 14 y, opcionalmente, un parámetro de nivel de la línea 14 y, además opcionalmente, un parámetro de relleno estéreo generado por el codificador de señales 400 y enviado a la interfaz de salida 500 a través de la línea de parámetros 43.

60

**[0037]** Preferentemente, el alineador de señales está configurado para alinear los canales de la señal multi-canal usando el parámetro de alineación de banda ancha, antes de que el determinador de parámetros 100 realmente calcule los parámetros de banda estrecha. Por lo tanto, en esta realización, el alineador de señales 200 envía los canales alineados de banda ancha de vuelta al determinador de parámetros 100 a través de la línea de conexión 15.

65 A continuación, el determinador de parámetros 100 determina la pluralidad de parámetros de alineación de banda

estrecha con respecto a la señal multi-canal alineada con características de banda ancha. En otras realizaciones, no obstante, los parámetros se determinan sin esta secuencia específica de procedimientos.

**[0038]** La Fig. 4a ilustra una implementación preferida, donde se realiza la secuencia específica de etapas que  
5 incurrir en la línea de conexión 15. En la etapa 16, el parámetro de alineación de banda ancha se determina mediante el uso de los dos canales y se obtiene el parámetro de alineación de banda ancha tal como un parámetro de diferencia de tiempos entre canales o ITD. A continuación, en la etapa 21, los dos canales se alinean mediante el alineador de señales 200 de la Fig. 1 mediante el uso del parámetro de alineación de banda ancha. A continuación, en la etapa 17, se determinan los parámetros de banda estrecha mediante el uso de los canales alineados dentro del determinador  
10 de parámetros 100 para determinar una pluralidad de parámetros de alineación de banda estrecha tales como una pluralidad de parámetros de diferencia de fases entre canales para bandas diferentes de la señal multi-canal. A continuación, en la etapa 22, se alinean los valores espectrales en cada banda de parámetros mediante el uso del correspondiente parámetro de alineación de banda estrecha para esta banda específica. Cuando este procedimiento en la etapa 22 se realiza para cada banda, para la cual está disponible un parámetro de alineación de banda estrecha,  
15 entonces los canales alineados primeros y segundos o izquierdos/derechos están disponibles para el procesamiento adicional de señales por el procesador de señales 300 de la Fig. 1.

**[0039]** La Fig. 4b ilustra una implementación adicional del codificador multi-canal de la Fig. 1 donde se realizan diversos procedimientos en el dominio de la frecuencia.  
20

**[0040]** Específicamente, el codificador multi-canal comprende además un convertidor tiempo-espectro 150 para convertir una señal multi-canal en el dominio del tiempo en una representación espectral de los al menos dos canales dentro del dominio de la frecuencia.

**[0041]** Asimismo, 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 Fig. 1, operan todos en el dominio de la frecuencia.

**[0042]** Asimismo, el codificador multi-canal y, específicamente, el procesador de señales comprende además un convertidor espectro-tiempo 154 para generar una representación en el dominio del tiempo de la señal central al  
30 menos.

**[0043]** Preferentemente, el convertidor espectro-tiempo convierte además una representación espectral de la señal lateral también determinada por los procedimientos representados en el bloque 152 en una representación en el dominio del tiempo, y el codificador de señales 400 de la Fig. 1 está configurado a continuación para codificar  
35 adicionalmente la señal central y/o la señal lateral como señales en el dominio del tiempo según la implementación específica del codificador de señales 400 de la Fig. 1.

**[0044]** Preferentemente, el convertidor tiempo-espectro 150 de la Fig. 4b está configurado para implementar las etapas 155, 156 y 157 de la Fig. 4c. Específicamente, la etapa 155 comprende proporcionar una ventana de análisis  
40 con al menos una porción con atenuación cero en uno de sus extremos y, específicamente, una porción con atenuación cero en la porción de ventana inicial y una porción con atenuación cero en la porción de ventana terminal como se ilustra, por ejemplo, en la Fig. 7 más adelante. Además, la ventana de análisis adicionalmente tiene intervalos de superposición o porciones superpuestas 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 sin superposición, según sea el caso.  
45

**[0045]** En la etapa 156, cada canal se forma en ventana mediante el uso de la ventana de análisis con intervalos de superposición. Específicamente, cada canal se forma en ventana mediante el uso de la ventana de análisis de tal forma que se obtiene un primer bloque del canal. A continuación, se obtiene un segundo bloque del mismo canal que tiene un cierto intervalo de superposición con el primer bloque y así sucesivamente, de tal forma que después de, por  
50 ejemplo, cinco operaciones de formación de ventana, cinco bloques de muestras formadas en ventana de cada canal están disponibles que son transformados a continuación individualmente en una representación espectral tal como se ilustra en 157 en la Fig. 4c. El mismo procedimiento se realiza para los otros canales también de modo que, al final de la etapa 157, está disponible una secuencia de bloques de valores espectrales y, específicamente, valores espectrales complejos tales como valores espectrales DFT o muestras de sub-bandas complejas.  
55

**[0046]** En la etapa 158, que es realizada por el determinador de parámetros 100 de la Fig. 1, se determina un parámetro de alineación de banda ancha y en la etapa 159, que se realiza mediante el alineador de señales 200 de la Fig. 1, se realiza un desplazamiento circular usando el parámetro de alineación de banda ancha. En la etapa 160, de nuevo realizada por el determinador de parámetros 100 de la Fig. 1, se determinan los parámetros de alineación de  
60 banda estrecha para bandas individuales/sub-bandas y en la etapa 161, los valores espectrales alineados son rotados para cada banda usando los correspondientes parámetros de alineación de banda estrecha para las bandas específicas.

**[0047]** La Fig. 4d ilustra adicionalmente los procedimientos realizados mediante el procesador de señales 300.  
65 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, se puede realizar algún procesamiento adicional de la señal lateral 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 en la etapa 303 y, en la etapa 305, se realizan una operación de agregado de superposición para la señal central por un lado y una operación de agregado de superposición para la señal lateral por otro lado para obtener finalmente las señales centrales/laterales en el dominio del tiempo.

**[0048]** Específicamente, las operaciones de las etapas 304 y 305 resultan en una especie de desvanecimiento cruzado desde un bloque de la señal central o la señal lateral en el bloque siguiente de la señal central y la señal lateral se realiza de modo que, incluso cuando ocurre cualquier cambio de parámetros tal como del parámetro de diferencia de tiempos entre canales o del parámetro de diferencia de fases entre canales, esto sin embargo no será audible en las señales central/lateral en el dominio del tiempo obtenidas mediante la etapa 305 en la Fig. 4d.

**[0049]** La nueva codificación estéreo con bajo retardo es una codificación conjunta estéreo central/lateral (M/S) que explota algunas pistas espaciales, donde el canal central está codificado por un codificador principal mononúcleo, y el canal lateral está codificado en un codificador de núcleo secundario. Los principios del codificador y del decodificador se ilustran en las Figs. 6a, 6b.

**[0050]** El procesamiento estéreo se realiza principalmente en el FD (por sus siglas en inglés: *Frequency Domain* / Dominio de Frecuencia). Opcionalmente, algo del procesamiento estéreo se puede realizar en el TD (por sus siglas en inglés: *Time Domain* / Dominio del Tiempo) antes del análisis de frecuencia. Este es el caso para el cómputo ITD, que se puede computar y aplicar antes del análisis de frecuencia para alinear los canales en el tiempo antes de proseguir con el análisis y el procesamiento estéreo. Por otra parte, el procesamiento ITD se puede realizar directamente en el dominio de la frecuencia. Dado que los codificadores de la voz usuales como ACELP no contienen ninguna descomposición interna tiempo-frecuencia, la codificación estéreo agrega un banco de filtros complejo modulado extra por medio de un banco de filtros de síntesis y análisis antes del codificador núcleo y otra etapa de banco de filtros de síntesis y análisis después del decodificador núcleo. En la realización preferida, se emplea una DFT sobremuestreada con una región con baja superposición. No obstante, en otras realizaciones, se puede usar cualquier descomposición valiosa compleja de tiempo-frecuencia con resolución temporal similar.

**[0051]** El procesamiento estéreo consiste en computar las pistas espaciales: diferencia de tiempos entre canales (ITD), las diferencias de fases entre canales (IPD) y las diferencias de niveles entre canales (ILD). La ITD y las IPD se usan en la señal de entrada estéreo para alinear los dos canales L (izquierdo) y R (derecho) en tiempo y en fase. La ITD se computa en banda ancha o en el dominio del tiempo mientras que las IPD y las ILD se computan para cada una o parte de las bandas de parámetros, correspondientes a una descomposición no uniforme del espacio de la frecuencia. Una vez que los dos canales están alineados se aplica un estéreo conjunto M/S, donde la señal lateral se predice a continuación además desde la señal central. La ganancia de predicción se deriva de las ILD.

**[0052]** La señal central se codifica además mediante un codificador de núcleo primario. En la realización preferida, el codificador de núcleo primario es el 3GPP EVS estándar, o una codificación derivada de este que puede conmutar entre un modo de codificación de voz, ACELP, y un modo de música sobre la base de una transformación MDCT (por sus siglas en inglés: *Modified Discrete Cosine Transform* – Transformada de Coseno Discreta Modificada). Preferentemente, ACELP y el codificador basado en MDCT están sustentados respectivamente por módulos TD-BWE (por sus siglas en inglés: *Time Domain BandWidth Extension* - Extensión de Ancho de Banda en el Dominio del Tiempo) y/o IGF (por sus siglas en inglés: *Intelligent Gap Filling* / Cobertura Inteligente de Intervalos).

**[0053]** La predicción de la señal lateral es realizada primero por el canal central mediante el uso de ganancias de predicción derivadas de las ILD. El residuo se puede predecir además mediante una versión con retardo de la señal central o ser directamente codificado por un codificador de núcleo secundario, realizado en la realización preferida en el dominio MDCT. El procesamiento estéreo en el codificador se puede sintetizar mediante la Fig. 5 tal como se explicará más adelante.

**[0054]** La Fig. 2 ilustra un diagrama de bloques de la realización de un aparato para decodificar una señal multi-canal codificada recibida en la línea de entrada 50.

**[0055]** En particular, la señal es recibida por una interfaz de entrada 600. Están conectados con la interfaz de entrada 600 un decodificador de señales 700 y un desalineador de señales 900. Asimismo, un procesador de señales 800 está conectado con un decodificador de señales 700 por una parte y está conectado con un desalineador de señales por otra parte.

**[0056]** En particular, la señal multi-canal codificada comprende una señal central codificada, una señal lateral codificada, información sobre el parámetro de alineación de banda ancha e información sobre la pluralidad de parámetros de banda estrecha. Por lo tanto, la señal multi-canal codificada en la línea 50 puede ser exactamente la misma señal que la salida de la interfaz de salida de 500 en la Fig. 1.



**[0057]** No obstante, es importante destacar aquí que, contrariamente a lo que se ilustra en la Fig. 1, el parámetro de alineación de banda ancha y la pluralidad de parámetros de alineación de banda estrecha incluidos en la señal codificada en cierta forma pueden ser exactamente los parámetros de alineación usados por el alineador de señales 200 en la Fig. 1 pero, por otra parte, también pueden ser sus valores inversos, es decir, parámetros que pueden ser usados por exactamente las mismas operaciones realizadas por el alineador de señales 200 pero con valores inversos de modo que se obtiene la desalineación.

**[0058]** De este modo, la información sobre los parámetros de alineación puede ser los parámetros de alineación usados por el alineador de señales 200 en la Fig. 1 o puede ser valores inversos, es decir, verdaderos "parámetros de desalineación". Adicionalmente, estos parámetros generalmente serán cuantificados de una forma determinada tal como se expondrá más adelante con respecto a la Fig. 8.

**[0059]** La interfaz de entrada 600 de la Fig. 2 separa la información sobre el parámetro de alineación de banda ancha y la pluralidad de parámetros de alineación de banda estrecha de las señales centrales/laterales codificadas y envía esta información a través de la línea de parámetros 610 al desalineador de señales 900. Por otra parte, la señal central codificada se envía al decodificador de señales 700 mediante la línea 601 y la señal lateral codificada se envía al decodificador de señales 700 mediante la línea de señal 602.

**[0060]** El decodificador de señales está configurado para decodificar la señal central codificada y para decodificar la señal lateral codificada para obtener una señal central decodificada en la línea 701 y una señal lateral decodificada en la línea 702. Estas señales son utilizadas por el procesador de señales 800 para calcular una señal decodificada del primer canal o una señal izquierda decodificada y para calcular una señal decodificada del segundo canal o una señal decodificada del canal derecho a partir de la señal central decodificada y la señal lateral decodificada, y el primer canal decodificado y el segundo canal decodificado son salidas en las líneas 801, 802, respectivamente. El desalineador de señales 900 está configurado para desalinearse el primer canal decodificado en la línea 801 y el canal derecho decodificado 802 usando la información sobre el parámetro de alineación de banda ancha y adicionalmente usando la información sobre la pluralidad de parámetros de alineación de banda estrecha para obtener una señal multi-canal decodificada, es decir, una señal decodificada que tiene al menos dos canales decodificados y desalineados en las líneas 901 y 902.

**[0061]** La Fig. 9a ilustra una secuencia preferida de etapas realizadas por el desalineador de señales 900 de la Fig. 2. Específicamente, la etapa 910 recibe los canales izquierdo y derecho alineados tal como están disponibles en las líneas 801, 802 de la Fig. 2. En la etapa 910, el desalineador de señales 900 desalinea sub-bandas individuales usando la información sobre los parámetros de alineación de banda estrecha a fin de obtener canales primero y segundo o izquierdo y derecho con desalineación de fase decodificados en 911a y 911b. En la etapa 912, los canales son desalineados mediante el uso del parámetro de alineación de banda ancha de modo que, en 913a y 913b, se obtienen canales desalineados en fase y tiempo.

**[0062]** En la etapa 914, se realiza cualquier procesamiento adicional que comprende el uso de una operación de formación de ventana o cualquier agregado de superposición o, generalmente, cualquier operación de desvanecimiento cruzado a fin de obtener, en 915a o 915b, una señal decodificada con reducción de fallos o libre de fallos, es decir, obtener canales decodificados que no tienen ningún fallo aunque haya habido, generalmente, parámetros de desalineación variables en el tiempo para la banda ancha, por una parte, y para la pluralidad de bandas estrechas, por otra parte.

**[0063]** La Fig. 9b ilustra una implementación preferida del decodificador multi-canal ilustrado en la Fig. 2.

**[0064]** En particular, el procesador de señales 800 de la Fig. 2 comprende un convertidor tiempo-espectro 810.

**[0065]** El procesador de señales comprende además un convertidor central/lateral a izquierda/derecha 820 a fin de calcular a partir de una señal central M y una señal lateral S una señal izquierda L y una señal derecha R.

**[0066]** No obstante, es importante señalar que, a fin de calcular L y R mediante la conversión central/lateral-izquierda/derecha en el bloque 820, no hay que utilizar necesariamente la señal lateral S. En cambio, como se expondrá más adelante, las señales izquierda/derecha se calculan inicialmente usando solo un parámetro de ganancia derivado de un parámetro de diferencia de niveles entre canales ILD. Generalmente, también se puede considerar que la ganancia de predicción es una forma de ILD. La ganancia se puede derivar de ILD pero también se puede computar directamente. Es preferible no computar más ILD, sino computar la ganancia de predicción directamente y transmitir y usar la ganancia de predicción en el decodificador en vez del parámetro ILD.

**[0067]** Por lo tanto, en esta implementación, la señal lateral S solo se usa en el actualizador de canales 830 que opera a fin de proporcionar una mejor señal izquierda/derecha usando la señal lateral transmitida S como se ilustra mediante la línea de desvío 821.

**[0068]** Por lo tanto, el convertidor 820 opera usando un parámetro de nivel obtenido mediante una entrada de

parámetro de nivel 822 y sin realmente usar la señal lateral S, pero el actualizador de canales 830 opera a continuación mediante el uso de la lateral 821 y, según la implementación específica, mediante el uso de un parámetro de relleno estéreo recibido por la línea 831. El alineador de señales 900 comprende a continuación un desalineador de fase y escalador de energía 910. El escalador de energía se controla mediante un factor de escalamiento derivado mediante un calculador de factor de escalamiento 940. El calculador de factor de escalamiento 940 es alimentado por la salida del actualizador de canales 830. Sobre la base de los parámetros de alineación de banda estrecha recibidos mediante la entrada 911, se realiza la desalineación de fase y, en el bloque 920, sobre la base del parámetro de alineación de banda ancha recibido mediante la línea 921, se realiza la desalineación en el tiempo. Finalmente, se realiza la conversión espectro-tiempo 930 a fin de obtener finalmente la señal decodificada.

10

**[0069]** La Fig. 9c ilustra una secuencia adicional de etapas generalmente realizadas dentro de los bloques 920 y 930 de la Fig. 9b en una realización preferida.

**[0070]** Específicamente, los canales desalineados de banda estrecha son introducidos en la funcionalidad de desalineación de banda ancha correspondiente al bloque 920 de la Fig. 9b. En el bloque 931 se realiza una DFT o cualquier otra transformada. A continuación del cálculo real de las muestras en el dominio del tiempo, se realiza una formación de ventana de síntesis opcional usando una ventana de síntesis. La ventana de síntesis preferentemente es exactamente igual a la ventana de análisis, o deriva de la ventana de análisis, por ejemplo, mediante interpolación o diezmado, 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 agregan hasta uno por cada punto en el intervalo de superposición. Por lo tanto, después de la ventana de síntesis en el bloque 932, se realiza una operación de superposición y una subsiguiente operación de adición. Alternativamente, en lugar de la formación de ventana de síntesis y la operación de superposición/adición, se realiza cualquier desvanecimiento cruzado entre los bloques subsiguientes para cada canal, a fin de obtener, como ya se expuso en el contexto de la Fig. 9a, una señal decodificada con fallos reducidos.

**[0071]** Cuando se considera la Fig. 6b, queda claro que las operaciones reales de decodificación para la señal central, es decir, el "decodificador EVS" por una parte, y para la señal lateral, la cuantificación vectorial inversa VQ-1 y la operación MDCT inversa (IMDCT) corresponden al decodificador de señales 700 de la Fig. 2.

30

**[0072]** Asimismo, las operaciones DFT en el bloque 810 corresponden al elemento 810 en la Fig. 9b y las funcionalidades del procesamiento estéreo inverso y el desplazamiento inverso en el tiempo corresponden a los bloques 800, 900 de la Fig. 2 y las operaciones de la DFT inversa 930 en la Fig. 6b corresponden a la operación correspondiente en el bloque 930 de la Fig. 9b.

35

**[0073]** A continuación, se expone con mayor detalle la Fig. 3. En particular, la Fig. 3 ilustra un espectro DFT que tiene líneas espectrales individuales. Preferentemente, el espectro DFT o cualquier otro espectro ilustrado en la Fig. 3 es un espectro complejo y cada línea es una línea espectral compleja que tiene magnitud y fase o que tiene una parte real y una parte imaginaria.

40

**[0074]** Además, el espectro también está dividido en bandas de parámetros diferentes. Cada banda de parámetros tiene al menos una y preferentemente más de una línea espectral. Además, las bandas de parámetros aumentan desde frecuencias bajas a altas. Generalmente, el parámetro de alineación de banda ancha es un solo parámetro de alineación de banda ancha para todo el espectro, es decir, para un espectro que comprende todas las bandas 1 a 6 en la realización ejemplar en la Fig. 3.

45

**[0075]** Asimismo, la pluralidad de parámetros de alineación de banda estrecha se proporciona de tal forma que existe un solo parámetro de alineación para cada banda de parámetros. Esto significa que el parámetro de alineación para una banda siempre se aplica a todos los valores espectrales dentro de la banda correspondiente.

50

**[0076]** Asimismo, además de los parámetros de alineación de banda estrecha, también se proporcionan parámetros de nivel para cada banda de parámetros.

**[0077]** Contrariamente a los parámetros de nivel que se proporcionan para cada una de las bandas de parámetros desde la banda 1 hasta la banda 6, es preferible proporcionar la pluralidad de parámetros de alineación de banda estrecha solo para un número limitado de bandas bajas tales como las bandas 1, 2, 3 y 4.

55

**[0078]** Además, se proporcionan parámetros de relleno estéreo para un cierto número de bandas excluyendo las bandas bajas tales como, en la realización ejemplar, las bandas 4, 5 y 6, mientras que existen valores espectrales de señal lateral para las bandas bajas de parámetros 1, 2 y 3, y en consecuencia, no existen parámetros de relleno estéreo para estas bandas bajas donde la adaptación de forma de onda se obtiene usando tanto la señal lateral en sí misma como una señal residual de predicción que representa la señal lateral.

60

**[0079]** Como ya se señaló, existen más líneas espectrales en las bandas altas, tales como, en la realización de la Fig. 3, siete líneas espectrales en la banda de parámetros 6 frente a solo tres líneas espectrales en la banda de

65

parámetros 2. Naturalmente, sin embargo, 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 diferentes límites para ciertos parámetros serán diferentes.

5 **[0080]** No obstante, la Fig. 8 ilustra una distribución de los parámetros y el número de bandas para las cuales se proporcionan parámetros en una determinada realización donde existen, en contraste con la Fig. 3, realmente 12 bandas.

**[0081]** Como se ilustra, el parámetro de nivel ILD se proporciona para cada una de las 12 bandas y es  
10 cuantificado hasta una exactitud de cuantificación representada por cinco bits por banda.

**[0082]** Asimismo, los parámetros de alineación de banda estrecha IPD solo se proporcionan para las bandas bajas hasta una frecuencia de borde de 2,5 kHz. Además, la diferencia de tiempos entre canales o el parámetro de alineación de banda ancha solo se proporciona como un único parámetro para todo el espectro, pero con una exactitud  
15 de cuantificación muy elevada representada por ocho bits para toda la banda.

**[0083]** Asimismo, se proporcionan parámetros de relleno estéreo cuantificados de manera bastante aproximada representados por tres bits por banda y no para las bandas bajas debajo de 1 kHz dado que, para las bandas bajas, se incluyen valores realmente codificados de señal lateral o valores espectrales residuales de señal  
20 lateral.

**[0084]** A continuación, se sintetiza con respecto a la Fig. 5 un procesamiento preferido en el lado del codificador. En una primera etapa, se realiza un análisis DFT del canal izquierdo y del canal derecho. Este procedimiento corresponde a las etapas 155 a 157 de la Fig. 4c. En la etapa 158, se calcula el parámetro de alineación  
25 de banda ancha y, particularmente, la diferencia de tiempos entre canales (ITD) del parámetro de alineación de banda ancha. Como se ilustra en 170, se realiza un desplazamiento en el tiempo de L y R en el dominio de la frecuencia. Alternativamente, este desplazamiento en el tiempo se puede realizar en el dominio del tiempo. A continuación, se realiza una DFT inversa, el desplazamiento en el tiempo se realiza en el dominio del tiempo y se realiza una DFT directa adicional a fin de tener una vez más representaciones espectrales después de la alineación usando el  
30 parámetro de alineación de banda ancha.

**[0085]** Los parámetros ILD, es decir, parámetros de nivel y parámetros de fase (parámetros IPD), se calculan para cada banda de parámetros en las representaciones desplazadas L y R como se ilustra en la etapa 171. Esta etapa corresponde a la etapa 160 de la Fig. 4c, por ejemplo. Las representaciones L y R desplazadas en el tiempo se  
35 rotan como una función de los parámetros de diferencia de fases entre canales como se ilustra en la etapa 161 de la Fig. 4c o la Fig. 5. A continuación, la señal central y la señal lateral se computan como se ilustra en la etapa 301 y, preferentemente, además con una operación de conservación de energía como se expone más adelante. En una etapa siguiente 174, se realiza una predicción de S con M como una función de ILD y opcionalmente con una señal M pasada, es decir, una señal central de una trama anterior. Posteriormente, se realiza la DFT inversa de la señal central  
40 y la señal lateral que corresponde a las etapas 303, 304, 305 de la Fig. 4d en la realización preferida.

**[0086]** En la etapa final 175, la señal central en el dominio del tiempo M y, opcionalmente, la señal residual se codifica como se ilustra en la etapa 175. Este procedimiento corresponde a lo que es realizado por el codificador de  
45 señales 400 en la Fig. 1.

**[0087]** En el decodificador en el procesamiento de estéreo inverso, la señal *lateral* se genera en el dominio DFT y se predice en primer lugar a partir de la señal *Central* como:

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

50 donde  $g$  es una ganancia computada para cada banda de parámetros y va en función de la diferencia de nivel entre canales (ILD) transmitida.

**[0088]** El residual de la predicción  $Side - g \cdot Mid$  se puede refinar a continuación en dos formas diferentes:

55 - Mediante una codificación secundaria de la señal residual:

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

60 donde  $g_{cod}$  es una ganancia global transmitida para todo el espectro

- Mediante una predicción residual, conocida como relleno estéreo, que predice el espectro lateral residual con el

espectro de la señal *central* anterior decodificada de la trama DFT anterior:

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

5 donde  $g_{pred}$  es una ganancia de predicción transmitida por banda de parámetros.

**[0089]** Los dos tipos de refinamiento de codificación se pueden mezclar dentro del mismo espectro DFT. En la realización preferida, la codificación residual se aplica sobre las bandas bajas de parámetros, mientras que la predicción residual se aplica sobre las bandas restantes. La codificación residual es en la realización preferida como se ilustra en la Fig. 1 se realiza en el dominio MDCT después de sintetizar la señal lateral residual en el dominio del tiempo y transformándola mediante MDCT. Contrariamente a DFT, en MDCT se hace muestreo crítico y es más adecuada para la codificación de audio. Los coeficientes de MDCT se someten directamente a cuantificación vectorial mediante una Cuantificación Vectorial Reticular, pero, alternativamente, se pueden codificar mediante Cuantificación Escalar seguida de codificación entrópica. Por otra parte, la señal lateral residual también se puede codificar en el 15 Dominio del Tiempo mediante una técnica de codificación de la voz o directamente en el dominio DFT.

### 1. Análisis de tiempo-frecuencia: DFT

**[0090]** Es importante que la descomposición de tiempo-frecuencia extra del procesamiento estéreo realizado 20 mediante DFT permita un buen análisis del escenario auditivo mientras que no aumenta significativamente el retardo general del sistema de codificación. De manera predeterminada, se utiliza una resolución temporal de 10 ms (dos veces los 20 ms de entramado del codificador núcleo). Las ventanas de análisis y síntesis son iguales y son simétricas. En la Fig. 7 la ventana se representa con una tasa de muestreo de 16 kHz. Se puede observar que la región de superposición está limitada para reducir el retardo generado y que también se agrega atenuación cero para 25 contrarrestar el desplazamiento circular cuando se aplica ITD en el dominio de la frecuencia tal como se explicará más adelante.

### 2. Parámetros estéreo

30 **[0091]** Los parámetros estéreo se pueden transmitir como máximo a la resolución temporal de la DFT estéreo. Como mínimo estos se pueden reducir a la resolución de entramado del codificador núcleo, es decir, 20ms. De forma predeterminada, cuando no se detectan transitorios, los parámetros se computan cada 20ms sobre 2 ventanas DFT. Las bandas de parámetros constituyen una descomposición no uniforme y no superpuesta del espectro a continuación 35 aproximadamente de 2 veces o 4 veces el ERB (por sus siglas en inglés: *Equivalent Rectangular Bandwidths* / Anchos de Banda Rectangulares Equivalentes). De manera predeterminada, se usa una escala ERB de 4 veces para un total de 12 bandas para un ancho de banda de frecuencia de 16kHz (tasa de muestreo 32kbps, estéreo de súper banda ancha). La Fig. 8 sintetiza un ejemplo de configuración, en el cual la información lateral estéreo se transmite con aproximadamente 5 kbps.

### 40 3. Computación de ITD y alineación de tiempo de canales

**[0092]** Las ITD se computan por la estimación del **TDOA** (por sus siglas en inglés: *Time Delay of Arrival* / Tiempo de Retardo de Llegada) usando la Correlación Cruzada Generalizada con Transformada de Fase (**GCC-PHAT**):

45

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

donde L y R son los espectros de frecuencia de los canales izquierdo y derecho respectivamente. El análisis de frecuencia se puede realizar independientemente de la DFT usada para el procesamiento estéreo subsiguiente o se 50 puede compartir. El pseudo-código para computar la ITD es el siguiente:

```

L =fft(ventana(l));
R =fft(ventana(r));
55 tmp = L .* conj( R );
sfm_L = prod(abs(L).^(1/longitud(L)))/(media(abs(L))+eps);
60 sfm_R = prod(abs(R).^(1/longitud(R)))/(media(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 );
5 tmp = ifft( tmp );
tmp = tmp([longitud(tmp)/2+1:longitud(tmp) 1:longitud(tmp)/2+1]);
10 tmp_sort = sort( abs(tmp) );
umbral = 3 * tmp_sort( vuelta(0.95*longitud(tmp_sort)) );
xcorr_time=abs(tmp(- ( h.stereo_itd_q_max - (length(tmp)-1)/2 - 1 ):- (
15 h.stereo_itd_q_min - (length(tmp)-1)/2 - 1 )));
% suavizado de salida para mejor detección
20 xcorr_time=[xcorr_time 0];
xcorr_time2=filtro([0,25 0,5 0,25],1,xcorr_time);
[m,i] = max(xcorr_time2(2:fin));
25 si m > umbral
itd = h.stereo_itd_q_max - i + 1;
30 si no
itd = 0;
fin
35

```

**[0093]** La Fig. 4e ilustra un diagrama de flujo para implementar el pseudo-código ilustrado anteriormente a fin de obtener un cálculo robusto y eficiente de una diferencia entre canales como un ejemplo de parámetro de alineación de banda ancha.

40 **[0094]** En el bloque 451, se realiza un análisis DFT de las señales en el dominio del tiempo para un primer canal (l) y un segundo canal (r). Este análisis DFT generalmente será el mismo análisis DFT que se ha expuesto en el contexto de las etapas 155 a 157 en la Fig. 5 o la Fig. 4c, por ejemplo.

45 **[0095]** A continuación, se realiza una correlación cruzada para cada frecuencia bin como se ilustra en el bloque 452.

**[0096]** Por lo tanto, se obtiene un espectro de correlación cruzada para el intervalo espectral completo de los canales izquierdo y derecho.

50 **[0097]** En la etapa 453, se calcula a continuación una medición de aplanamiento espectral a partir de la magnitud de los espectros de L y R y, en la etapa 454, se selecciona la mayor medición de aplanamiento espectral. No obstante, la selección en la etapa 454 no necesariamente tiene que ser la selección de la mayor, sino que esta determinación de una sola SFM (por sus siglas en inglés: *Spectral Flatness Measure* / Medición de Aplanamiento Espectral) de ambos canales también puede ser la selección y el cálculo del canal izquierdo solamente o del canal 55 derecho solamente o puede ser el cálculo del promedio ponderado de ambos valores de SFM.

**[0098]** En la etapa 455, el espectro de correlación cruzada es suavizado a continuación a lo largo del tiempo según las mediciones de aplanamiento espectral.

60 **[0099]** Preferentemente, la medición de aplanamiento espectral se calcula dividiendo la media geométrica del espectro de magnitud por la media aritmética del espectro de magnitud. Por lo tanto, los valores de SFM están limitados entre cero y uno.

**[0100]** En la etapa 456, el espectro de correlación cruzada suavizado a continuación se normaliza por su 65 magnitud y en la etapa 457 se calcula una DFT inversa del espectro de correlación cruzada normalizado y suavizado.

En la etapa 458, se realiza preferentemente un determinado filtro en el dominio del tiempo, pero este filtrado en el dominio del tiempo también se puede dejar de lado según la implementación, pero se prefiere, tal como se describirá más adelante.

5 **[0101]** En la etapa 459, se realiza una estimación de la ITD mediante selección de valores máximos (*peak-picking*) de la función de filtro generalizada de correlación cruzada y mediante la realización de ciertas operaciones de umbralización.

**[0102]** Si no se obtiene un cierto umbral, entonces la ITD se establece en cero y no se realiza alineación en el tiempo para este bloque correspondiente.

**[0103]** La computación de la ITD se puede sintetizar también como sigue a continuación. La correlación cruzada se computa en el dominio de la frecuencia antes de ser suavizada según la Medición de Aplanamiento Espectral. La SFM está limitada entre 0 y 1. En el caso de señales del tipo ruido, la SFM será alta (es decir, alrededor de 1) y el suavizado será débil. En el caso de una señal del tipo tono, la SFM será baja y el suavizado se volverá más fuerte. La correlación cruzada suavizada se normaliza a continuación mediante su amplitud antes de ser transformada de nuevo al dominio del tiempo. La normalización corresponde a la transformada de fase de la correlación cruzada, y se sabe que exhibe un mejor desempeño que la correlación cruzada normal en ambientes con ruido bajo y reverberación relativamente alta. La función en el dominio de la frecuencia así obtenida, se filtra primeramente para obtener una selección más sólida de picos máximos. El índice correspondiente a la amplitud máxima corresponde a un estimado de la diferencia de tiempo entre el canal izquierdo y el derecho (ITD). Si la amplitud del máximo es más baja que un umbral dado, entonces la estimación de la ITD no se considera fiable y se establece en cero.

**[0104]** Si la alineación en el tiempo se aplica en el dominio del tiempo, la ITD se computa en un análisis DFT separado. El desplazamiento se realiza como sigue a continuación:

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

**[0105]** Este requiere un retardo extra en el codificador, que es igual en el máximo que la ITD máxima absoluta que se puede gestionar. La variación de la ITD a lo largo del tiempo se suaviza mediante la formación de ventana de análisis de DFT.

**[0106]** Alternativamente, la alineación en el tiempo se puede realizar en el dominio de la frecuencia. En este caso, la computación de ITD y el desplazamiento circular están en el mismo dominio de DFT, dominio compartido con este otro procesamiento estéreo. El desplazamiento circular está dado por:

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

**[0107]** La atenuación cero de las ventanas DFT es necesaria para simular un desplazamiento en el tiempo con un desplazamiento circular. El tamaño de la atenuación cero corresponde a la ITD máxima absoluta que se puede gestionar. En la realización preferida, la atenuación cero está dividida de manera uniforme en ambos lados de las ventanas de análisis, mediante el agregado de 3,125ms de ceros en ambos extremos. La ITD máxima absoluta posible es entonces 6,25ms. En la configuración de micrófonos A-B, esta corresponde en el peor de los casos a una distancia máxima de aproximadamente 2,15 metros entre los dos micrófonos. La variación de ITD a lo largo del tiempo se suaviza mediante la formación de ventana de síntesis y el agregado de superposición de la DFT.

**[0108]** Es importante que el desplazamiento en el tiempo esté seguido por una formación de ventana de la señal desplazada. Esta es la principal diferencia con la técnica conocida de BCC (por sus siglas en inglés: *Binaural Cue Coding* / Codificación de Referencia Binaural), donde el desplazamiento en el tiempo se aplica en una señal formada en ventana pero que no se forma en ventana adicionalmente en la etapa de síntesis. En consecuencia, cualquier cambio en ITD a lo largo del tiempo produce un transitorio/clic artificial en la señal decodificada.

#### 4. Computación de IPD y rotación de canales

55 **[0109]** Las IPD se computan después de la alineación de los dos canales y esto para cada banda de parámetros o al menos hasta una *ipd\_max\_band* dada, según la configuración estéreo.

$$IPD[b] = \text{ángulo} \left( \sum_{k=\text{band}_{\text{limits}[b]}}^{\text{band}_{\text{limits}[b+\eta]}} L[k] R^*[k] \right)$$

**[0110]** A continuación, se aplican las IPD a los dos canales para alinear sus fases:

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

5

**[0111]** Donde  $\beta = \text{atan2}(\sin(IPD_i[b]), \cos(IPD_i[b]) + c)$ ,  $c = 10^{ILD_i[b]/20}$  y  $b$  es el índice de banda de parámetros al cual pertenece el índice de frecuencia  $k$ . El parámetro  $\beta$  es responsable de distribuir la cantidad de rotaciones de fase entre los dos canales mientras hace que sus fases estén alineadas.  $\beta$  depende de IPD pero también del nivel de amplitud relativa de los canales, ILD. Si un canal tiene amplitud más elevada, será considerado como canal principal  
10 y se verá menos afectado por las rotaciones de fase que el canal con amplitud más baja.

### 5. Codificación suma-resta y señal lateral

**[0112]** La transformación suma resta se realiza en los espectros alineados de tiempo y fase de los dos canales  
15 de tal forma que la energía se conserve en la señal central.

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

donde  $a = \sqrt{\frac{L^2+R^2}{(L+R)^2}}$  está limitado entre 1/1,2 y 1,2, es decir -1,58 y +1,58 dB. Esta limitación evita fallos cuando se ajusta la energía de M y S. Cabe destacar que esta conservación de energía es menos importante cuando el tiempo y la fase  
20 se alinean de antemano. Alternativamente, los límites se pueden disminuir o aumentar.

**[0113]** Se predice adicionalmente la señal lateral con M:

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

25

donde  $g(ILD) = \frac{c-1}{c+1}$ , donde  $c = 10^{ILD_i[b]/20}$ . Alternativamente, la ganancia de predicción óptima  $g$  se puede encontrar al minimizar el MSE (por sus siglas en inglés: *Mean Square Error / Error Cuadrático Medio*) del residual y las ILD deducidos mediante la ecuación anterior.

30 **[0114]** La señal residual  $S'(f)$  se puede modelar de dos maneras: prediciéndola con el espectro de retardo de M o codificándola directamente en el dominio MDCT en el dominio MDCT.

### 6. Decodificación estéreo

35 **[0115]** La señal central X y la señal lateral S se convierten primero a los canales izquierdo y derecho L y R como sigue a continuación:

$$L_i[k] = M_i[k] + gM_i[k], \text{ para } \text{band\_limits}[b] \leq k < \text{band\_limits}[b+1],$$

$$R_i[k] = M_i[k] - gM_i[k], \text{ para } \text{band\_limits}[b] \leq k < \text{band\_limits}[b+1],$$

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

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

donde  $c = 10^{ILD_i[b]/20}$ .

45

**[0116]** Para bandas de parámetros por debajo de `cod_max_band`, los dos canales se actualizan con la señal lateral decodificada:

$$L_i[k] = L_i[k] + \text{cod\_gain}_i \cdot S_i[k], \text{ para } 0 \leq k < \text{band\_limits}[\text{cod\_max\_band}],$$

$$R_i[k] = R_i[k] - \text{cod\_gain}_i \cdot S_i[k], \text{ para } 0 \leq k < \text{band\_limits}[\text{cod\_max\_band}],$$

50

**[0117]** Para bandas de parámetros más elevadas, se predice la señal lateral y los canales se actualizan como:

$$L_i[k] = L_i[k] + \text{cod\_pred}_i[b] \cdot M_{i-1}[k], \text{ para } \text{band\_limits}[b] \leq k < \text{band\_limits}[b+1],$$

$$R_i[k] = R_i[k] - \text{cod\_pred}_i[b] \cdot M_{i-1}[k], \text{ para } \text{band\_limits}[b] \leq k < \text{band\_limits}[b+1],$$

5 **[0118]** Finalmente, los canales se multiplican por un valor complejo con el objetivo de restaurar la energía original y la fase entre canales de la señal estéreo:

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

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

10 donde

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

15 donde  $a$  está definida y limitada como se definió anteriormente, y donde  $\beta = \text{atan2}(\sin(\text{IPD}_i[b]), \cos(\text{IPD}_i[b]) + c)$ , y donde  $\text{atan2}(x,y)$  es la tangente inversa en el cuarto cuadrante de  $x$  sobre  $y$ .

**[0119]** Finalmente, los canales son desplazados en el tiempo ya sea en el dominio del tiempo o de la frecuencia, según las ITD transmitidas. Los canales en el dominio del tiempo se sintetizan mediante DFT inversas y adición de superposición.

20

**[0120]** Las características específicas de la invención se relacionan con la combinación de pistas espaciales y codificación estéreo conjunta suma-resta. Específicamente, se computan y aplican en los canales estéreo (izquierdo y derecho) las pistas espaciales IDT e IPD. Asimismo, se calcula la suma-resta (señales M/S) y, preferentemente, se aplica una predicción de S con M.

25

**[0121]** Del lado del decodificador, las pistas espaciales de banda ancha y de banda estrecha se combinan entre sí con codificación estéreo conjunta suma-resta. En particular, la señal lateral se predice con la señal central usando al menos una pista espacial tal como ILD y se calcula una suma-resta inversa para obtener los canales izquierdo y derecho, y, además, se aplican en los canales izquierdo y derecho las pistas espaciales de banda ancha y banda

30 estrecha.

**[0122]** Preferentemente, el codificador tiene una ventana y agregado de superposición con respecto a los canales alineados en el tiempo después del procesamiento con el uso de la ITD. Además, el decodificador tiene adicionalmente una operación de formación de ventana y agregado de superposición de las versiones desplazadas o

35 desalineadas de los canales después de aplicar la diferencia de tiempos entre canales.

**[0123]** La computación de la diferencia de tiempos entre canales con el procedimiento GCC-Phat es un procedimiento específicamente sólido.

40 **[0124]** El nuevo procedimiento es ventajoso con respecto a la técnica anterior ya que logra la codificación de tasa de bits de audio estéreo o audio multi-canal con bajo retardo. Está específicamente diseñado para ser sólido ante señales de entrada de distinta naturaleza y diferentes configuraciones del registro estéreo o multi-canal. En particular, la presente invención proporciona una buena calidad para codificación de tasa de bits de voz estéreo.

45 **[0125]** Los procedimientos preferidos encuentran uso en la distribución de transmisiones de todo tipo de contenido de audio estéreo o multi-canal tal como voz y música igualmente con una baja tasa de bits dada. Tales áreas de aplicación son radio digital, transmisión de Internet o aplicaciones en comunicación de audio.

**[0126]** Una señal de audio codificada de la invención se puede almacenar en un medio de almacenamiento digital o un medio de almacenamiento no transitorio o se puede transmitir en un medio de transmisión tal como un medio de transmisión inalámbrica o un medio de transmisión por cable tal como Internet.

50

**[0127]** Aunque algunos aspectos se han descrito en el contexto de un aparato, resulta claro que estos aspectos también representan una descripción del procedimiento correspondiente, donde un bloque o dispositivo corresponde a una etapa de un procedimiento o una característica de una etapa de un procedimiento. De forma análoga, los aspectos descritos en el contexto de una etapa de un procedimiento también representan una descripción de un bloque o elemento o característica de un aparato correspondiente.

55

**[0128]** Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención se pueden



implementar en un equipo físico o un soporte lógico. La implementación se puede realizar usando un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un CD, una memoria ROM, una memoria PROM, una memoria EPROM, una memoria EEPROM o una memoria FLASH, que tenga almacenadas en ella señales de control legibles electrónicamente, que cooperan (o son capaces de cooperar) con un sistema informático programable de tal forma que se realice el procedimiento respectivo.

**[0129]** Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable, de tal forma que se lleva a cabo uno de los procedimientos descritos en esta invención.

10

**[0130]** Generalmente, las realizaciones de la presente invención se pueden implementar como un producto que es un programa informático con un código de programa, siendo el código de programa operativo para realizar uno de los procedimientos cuando el programa informático se ejecuta en un ordenador. El código de programa, por ejemplo, se puede almacenar en un soporte legible por una máquina.

15

**[0131]** Otras realizaciones comprenden el programa informático para realizar uno de los procedimientos descritos en esta invención, almacenados en un soporte legible por una máquina o un medio de almacenamiento no transitorio.

20

**[0132]** En otras palabras, una realización del procedimiento de la invención es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los procedimientos descritos en esta invención, cuando el programa se ejecuta en un ordenador.

25

**[0133]** Una realización adicional de la invención es, por lo tanto, 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 realizar uno de los procedimientos descritos en esta invención.

30

**[0134]** Una realización adicional del procedimiento de la invención es, por lo tanto, una corriente de datos o una secuencia de señales que representan el programa informático para realizar uno de los procedimientos descritos en esta invención. La corriente de datos o la secuencia de señales pueden, por ejemplo, estar configuradas para ser transferidas a través de una conexión de comunicación de datos, por ejemplo, a través de Internet.

35

**[0135]** Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado o adaptado para realizar uno de los procedimientos descritos en esta invención.

**[0136]** Una realización adicional comprende un ordenador que tiene instalado el programa informático para realizar uno de los procedimientos descritos en esta invención.

40

**[0137]** En algunas realizaciones, se puede usar un dispositivo lógico programable (por ejemplo, una matriz de puertas programables por campo) para realizar algunas o todas las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, una matriz de puertas programables por campo puede cooperar con un microprocesador a fin de realizar uno de los procedimientos descritos en esta invención. Generalmente, los procedimientos se realizan preferentemente mediante cualquier aparato de hardware.

45

**[0138]** Las realizaciones anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que modificaciones y variaciones de las disposiciones y los detalles descritos en esta invención serán evidentes para otros expertos en la materia. Es la intención, por lo tanto, estar limitados únicamente por el alcance de las reivindicaciones de patente que siguen a continuación y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en esta invención.

50

**REIVINDICACIONES**

1. Aparato para la codificación de una señal de audio multi-canal que tiene al menos dos canales, que comprende:
  - 5 un determinador de parámetros (100) para determinar un parámetro de alineación de banda ancha y una pluralidad de parámetros de alineación de banda estrecha a partir de la señal de audio multi-canal;
  - un alineador de señal (200) para alinear los al menos dos canales usando el parámetro de alineación de banda ancha y la pluralidad de parámetros de alineación de banda estrecha para obtener canales alineados;
  - 10 un procesador de señal (300) para calcular una señal central y una señal lateral usando los canales alineados;
  - un codificador de señal (400) para codificar la señal central para obtener una señal central codificada y para codificar la señal lateral para obtener una señal lateral codificada; y
  - una interfaz de salida (500) para generar una señal de audio codificada multi-canal que comprende la señal central codificada, la señal lateral codificada, información sobre el parámetro de alineación de banda ancha e información
  - 15 sobre la pluralidad de parámetros de alineación de banda estrecha.
  
2. Aparato de la reivindicación 1, en el que el determinador de parámetros (100) está configurado para determinar el parámetro de alineación de banda ancha usando una representación de banda ancha de al menos dos canales, comprendiendo la representación de
  - 20 banda ancha al menos dos sub-bandas de cada uno de los al menos dos canales, y
  - en el que el alineador de señal (200) está configurado para realizar una alineación de banda ancha de la representación de banda ancha de los al menos dos canales para obtener una representación de banda ancha alineada de los al menos dos canales.
  
- 25 3. Aparato de la reivindicación 1 o la reivindicación 2, en el que el determinador de parámetros (100) está configurado para determinar un parámetro de alineación de banda estrecha separado para al menos una sub-banda de una representación de banda ancha alineada de los al menos dos canales, y
  - en el que el alineador de señal (200) está configurado para alinear individualmente cada sub-banda de la
  - 30 representación de banda ancha alineada usando el parámetro de banda estrecha para una sub-banda correspondiente para obtener una representación de banda estrecha alineada que comprende una pluralidad de sub-bandas alineadas para cada uno de los al menos dos canales.
  
4. Aparato de una de las reivindicaciones anteriores,
  - 35 en el que el procesador de señal (300) está configurado para calcular la pluralidad de sub-bandas para la señal central y una pluralidad de sub-bandas para la señal lateral usando una pluralidad de sub-bandas alineadas para cada uno de los al menos dos canales.
  
5. Aparato de una de las reivindicaciones anteriores,
  - 40 en el que el determinador de parámetros (100) está configurado para calcular, como el parámetro de alineación de banda ancha, un parámetro de diferencia de tiempo entre canales o, como la pluralidad de parámetros de alineación de banda estrecha, una diferencia de fase entre canales para cada una de una pluralidad de sub-bandas de la señal de audio multi-canal.
  
- 45 6. Aparato de una de las reivindicaciones anteriores, en el que el determinador de parámetros (100) está configurado para calcular una ganancia de predicción o una diferencia de nivel entre canales para cada una de una pluralidad de sub-bandas de la señal de audio multi-canal, y en el que el codificador de señal (400) está configurado para realizar una predicción de señal lateral en una sub-banda usando la señal central en la sub-banda y usando la diferencia de nivel entre canales o la ganancia de predicción de
  - 50 la sub-banda.
  
7. Aparato de una de las reivindicaciones anteriores, en el que el codificador de señal (400) está configurado para calcular y codificar una señal residual de predicción derivada de la señal lateral, una ganancia de predicción o una diferencia de nivel entre canales entre los al menos dos
  - 55 canales, la señal central y una señal central retardada, o en el que la ganancia de predicción en una sub-banda se calcula usando la diferencia de nivel entre canales entre los al menos dos canales en la sub-banda, o en el que el codificador de señal está configurado para codificar la señal central usando un codificador de voz o un codificador de música/voz conmutado o un codificador de extensión de ancho de banda de dominio de tiempo o un codificador de relleno de intervalo de dominio de frecuencia.
  - 60
  
8. Aparato según una de las reivindicaciones anteriores, que comprende, además:
  - un convertidor de espectro de tiempo (150) para generar una representación espectral de los al menos dos canales
  - en un dominio espectral,
  - 65 en el que el determinador de parámetros (100) y el alineador de señal (200) y el procesador de señal (300) están

- configurados para operar en el dominio espectral, y  
 en el que el procesador de señal (300) comprende además un convertidor de espectro-tiempo (154) para generar una representación en el dominio del tiempo de la señal central, y  
 en el que el codificador de señal (400) está configurado para codificar la representación en el dominio del tiempo de la señal central.
- 5
9. Aparato de una de las reivindicaciones anteriores, en el que el determinador de parámetros (100) está configurado para calcular el parámetro de alineación de banda ancha usando una representación espectral,
- 10 en el que el alineador de señal (200) está configurado para aplicar un desplazamiento circular (159) a la representación espectral de los al menos dos canales usando el parámetro de alineación de banda ancha para obtener valores espectrales alineados de banda ancha para los al menos dos canales, o en el que el determinador de parámetros (100) está configurado para calcular la pluralidad de parámetros de alineación de banda estrecha a partir de los valores espectrales alineados de banda ancha, y
- 15 en el que el alineador de señal (200) está configurado para rotar (161) los valores espectrales alineados de banda ancha usando la pluralidad de parámetros de alineación de banda estrecha.
10. Aparato de la reivindicación 8 ó 9, en el que el convertidor de espectro de tiempo (150) está configurado para aplicar una ventana de análisis a cada uno de los al menos dos canales, en el que la ventana de análisis tiene una porción con atenuación cero en un lado izquierdo o un lado derecho del mismo, en el que la porción con atenuación cero determina un valor máximo del parámetro de alineación de banda ancha o en el que la ventana de análisis tiene una región superpuesta inicial, una región central no superpuesta y una región superpuesta posterior o
- 20 en el que el convertidor de espectro de tiempo (150) está configurado para aplicar una secuencia de ventanas superpuestas, en el que una longitud de una parte superpuesta de una ventana y una longitud de una parte no superpuesta de la ventana juntas son iguales a una fracción de un marco del codificador de señal (400).
11. Aparato de una de las reivindicaciones 8 a 10, en el que el convertidor de espectro-tiempo (154) está configurado para usar una ventana de síntesis, siendo la ventana de síntesis idéntica a la ventana de análisis utilizada por el convertidor de espectro de tiempo (150) o se deriva de la ventana de análisis.
- 30
12. Aparato de una de las reivindicaciones anteriores, en el que el procesador de señal (300) está configurado para calcular una representación en el dominio del tiempo de la señal central o la señal lateral, en el que el cálculo de la representación en el dominio del tiempo comprende:
- 35 la formación de ventana (304) de un bloque actual de muestras de la señal central o la señal lateral para obtener un bloque actual formado en ventana,  
 la formación de ventana (304) de un bloque posterior de muestras de la señal central o la señal lateral para obtener un bloque posterior formado en ventana, y  
 la adición (305) de muestras del bloque actual formado en ventana y muestras del bloque posterior formado en ventana en un intervalo de superposición para obtener la representación del dominio de tiempo para el intervalo de superposición.
- 40
- 45
13. Aparato de una de las reivindicaciones anteriores, en el que el codificador de señal (400) está configurado para codificar la señal lateral o una señal residual de predicción derivada de la señal lateral y la señal central en un primer conjunto de sub-bandas, y para codificar, en un segundo conjunto de sub-bandas, diferente del primer conjunto de sub-bandas, una señal lateral derivada del parámetro de ganancia y una señal central antes, en el que la señal lateral o una señal residual de predicción no está codificada para el segundo conjunto de sub-bandas.
- 50
14. Aparato de la reivindicación 13, en el que el primer conjunto de sub-bandas tiene sub-bandas que son más bajas en frecuencia que las frecuencias en el segundo conjunto de sub-bandas.
- 55
15. Aparato de una de las reivindicaciones anteriores, en el que el codificador de señal (400) está configurado para codificar la señal lateral usando una transformada MDCT y una cuantificación tal como un vector o un escalar o cualquier otra cuantificación de coeficientes MDCT de la señal lateral.
- 60
16. Aparato de una de las reivindicaciones anteriores, en el que el determinador de parámetros (100) está configurado para determinar la pluralidad de parámetros de alineación de banda estrecha para bandas individuales que tienen ancho de banda, en el que un primer ancho de
- 65

banda de una primera banda que tiene una primera frecuencia central es menor que un segundo ancho de banda de una segunda banda que tiene una segunda frecuencia central, en el que la segunda frecuencia central es mayor que la primera frecuencia central o en el que el determinador de parámetros (100) está configurado para determinar los parámetros de alineación de banda estrecha solo para bandas hasta una frecuencia límite, siendo la frecuencia límite inferior a una frecuencia máxima de señal central o señal lateral, y en el que el alineador (200) está configurado para alinear solo los al menos dos canales en sub-bandas que tienen frecuencias por encima de la frecuencia límite usando el parámetro de alineación de banda ancha y para alinear los al menos dos canales en sub-bandas que tienen frecuencias por debajo de la frecuencia límite usando el parámetro de alineación de banda ancha y los parámetros de alineación de banda estrecha.

17. Aparato de una de las reivindicaciones anteriores, en el que el determinador de parámetros (100) está configurado para calcular el parámetro de alineación de banda ancha usando la estimación de un retraso de tiempo de llegada usando una correlación cruzada generalizada, y en el que el alineador de señal (200) está configurado para aplicar el parámetro de alineación de banda ancha en un dominio de tiempo usando un cambio de tiempo o en un dominio de frecuencia usando un cambio circular, o en el que el determinador de parámetros (100) está configurado para calcular el parámetro de banda ancha usando:

el cálculo (452) de un espectro de correlación cruzada entre el primer canal y el segundo canal;  
 el cálculo (453, 454) de una información sobre una forma espectral para el primer canal o el segundo canal o ambos canales;  
 el suavizado (455) del espectro de correlación cruzada dependiendo de la información sobre la forma espectral; opcionalmente, normalizando (456) el espectro de correlación cruzada suavizado;  
 la determinación (457, 458) de una representación en el dominio del tiempo del espectro de correlación cruzada suavizado y opcionalmente normalizado; y  
 el análisis (459) de la representación en el dominio del tiempo para obtener la diferencia de tiempo entre canales como el parámetro de alineación de banda ancha.

18. Aparato de una de las reivindicaciones anteriores, en el que el procesador de señal (300) está configurado para calcular la señal central y la señal lateral utilizando un factor de escala de energía y en el que el factor de escala de energía está limitado como máximo entre 2 y al menos 0,5, o en el que el determinador de parámetros (100) está configurado para calcular un parámetro de alineación normalizado para una banda determinando un ángulo de una suma compleja de productos de valores espectrales del primer y segundo canal dentro de la banda, o en el que el alineador de señal (200) está configurado para realizar la alineación de banda estrecha de una manera tal que tanto el primer canal como el segundo están sujetos a una rotación de canal, en el que una rotación de canal de un canal que tiene una amplitud más alta se gira en un grado menor en comparación con un canal que tiene una amplitud menor.

19. Procedimiento para la codificación de una señal de audio multi-canal que tiene al menos dos canales, que comprende:

la determinación (100) de un parámetro de alineación de banda ancha y una pluralidad de parámetros de alineación de banda estrecha a partir de la señal de audio multi-canal;  
 la alineación (200) de los al menos dos canales mediante el uso del parámetro de alineación de banda ancha y la pluralidad de parámetros de alineación de banda estrecha para obtener canales alineados;  
 el cálculo (300) de una señal central y una señal lateral mediante el uso de los canales alineados;  
 la codificación (400) de la señal central para obtener una señal central codificada y la codificación de la señal lateral para obtener una señal lateral codificada; y  
 la generación (500) de una señal de audio codificada multi-canal que comprende la señal central codificada, la señal lateral codificada, información sobre el parámetro de alineación de banda ancha e información sobre la pluralidad de parámetros de alineación de banda estrecha.

20. Señal de audio codificada multi-canal que comprende una señal central codificada, una señal lateral codificada, información sobre un parámetro de alineación de banda ancha e información sobre una pluralidad de parámetros de alineación de banda estrecha.

21. Aparato para la decodificación de una señal de audio multi-canal codificada que comprende una señal central codificada, una señal lateral codificada, información sobre un parámetro de alineación de banda ancha e información sobre una pluralidad de parámetros de alineación de banda estrecha, que comprende:

un decodificador de señal (700) para la decodificación de la señal central codificada para obtener una señal central decodificada y para la decodificación de la señal lateral codificada para obtener una señal lateral decodificada;  
 un procesador de señal (800) para el cálculo de un primer canal decodificado y un segundo canal decodificado a partir de la señal central decodificada y la señal lateral decodificada; y  
 un desalineador de señal (900) para la desalineación del primer canal decodificado y el segundo canal decodificado

mediante el uso de la información sobre el parámetro de alineación de banda ancha y la información sobre la pluralidad de parámetros de alineación de banda estrecha para obtener una señal de audio multi-canal decodificada.

- 5 22. Aparato de la reivindicación 21,  
 en el que el desalineador de señal (900) está configurado para desalinear cada una de una pluralidad de sub-bandas de los canales decodificados primero y segundo mediante el uso de un parámetro de alineación de banda estrecha asociado con la sub-banda correspondiente para obtener una sub-banda desalineada para el primero y el segundo canal, y
- 10 en el que el desalineador de señal está configurado para desalinear una representación de las sub-bandas desalineadas de los canales decodificados primero y segundo mediante el uso de la información sobre el parámetro de alineación de banda ancha.
23. Aparato de la reivindicación 21 ó 22,  
 15 en el que el desalineador de señal (900) está configurado para calcular una representación en el dominio del tiempo del primer canal decodificado o del segundo canal decodificado mediante el uso de la formación de ventana de un bloque actual de muestras del canal izquierdo o el canal derecho para obtener un bloque actual formado en ventana; la formación de ventana de un bloque posterior de muestras del primer canal y el segundo canal para obtener un bloque posterior formado en ventana; y
- 20 la adición de muestras del bloque actual formado en ventana y muestras del bloque posterior formado en ventana en un intervalo de superposición para obtener la representación del dominio de tiempo para el intervalo de superposición.
24. Aparato de una de las reivindicaciones 21 a 23,  
 en el que el desalineador de señal (900) está configurado para aplicar la información sobre la pluralidad de parámetros de alineación de banda estrecha individuales para sub-bandas individuales que tienen anchos de banda, en el que un primer ancho de banda de una primera banda que tiene una primera frecuencia central es menor que un segundo ancho de banda de una segunda banda que tiene una segunda frecuencia central,  
 25 en el que la segunda frecuencia central es mayor que la primera frecuencia central, o  
 en el que el desalineador de señal está configurado para aplicar la información sobre la pluralidad de parámetros de alineación de banda estrecha individuales para bandas individuales solo para bandas hasta una frecuencia límite, siendo la frecuencia límite inferior a una frecuencia máxima del primer canal decodificado o el segundo canal decodificado y  
 30 en el que el desalineador (900) está configurado para desalinear solo los al menos dos canales en sub-bandas que tienen frecuencias superiores a la frecuencia límite mediante el uso de la información sobre el parámetro de alineación de banda ancha y desalinear los al menos dos canales en sub-bandas que tienen frecuencias por debajo de la frecuencia límite mediante el uso de la información sobre el parámetro de alineación de banda ancha y mediante el uso de la información sobre los parámetros de alineación de banda estrecha.
25. Aparato de una de las reivindicaciones 21 a 24,  
 40 en el que el procesador de señal (800) comprende:  
 un convertidor de espectro de tiempo (810) para calcular una representación en el dominio de la frecuencia de la señal central decodificada y la señal lateral decodificada,  
 en el que el procesador de señal (800) está configurado para calcular el primer canal decodificado y el segundo canal decodificado en el dominio de frecuencia, y  
 45 en el que el desalineador de señal comprende un convertidor de espectro-tiempo (930) para convertir señales alineadas mediante el uso de la información sobre la pluralidad de parámetros de alineación de banda estrecha solamente o mediante el uso de la pluralidad de parámetros de alineación de banda estrecha y mediante el uso de la información sobre el parámetro de alineación de banda ancha en un dominio de tiempo.
- 50 26. Aparato de una de las reivindicaciones 21 a 25,  
 en el que el desalineador de señal (900) está configurado para realizar una desalineación en un dominio de tiempo mediante el uso de la información sobre el parámetro de alineación de banda ancha y para realizar una operación de formación de ventana (932) o una operación de superposición y adición (933) mediante el uso de bloques posteriores de tiempo de los canales alineados en el tiempo, o  
 55 en el que el desalineador de señal (900) está configurado para realizar una desalineación en un dominio espectral mediante el uso de la información sobre el parámetro de alineación de banda ancha y para realizar una conversión de espectro-tiempo (931) mediante el uso de los canales desalineados y realizar una formación de ventana de síntesis (932) y una operación de superposición y suma (933) mediante el uso de bloques posteriores al tiempo de canales desalineados.
- 60 27. Aparato de una de las reivindicaciones anteriores,  
 en el que el decodificador de señal está configurado para generar una señal central del dominio del tiempo y una señal lateral del dominio del tiempo,  
 65 en el que el procesador de señal (800) está configurado para realizar una formación de ventana mediante el uso de

- una ventana de análisis para generar bloques posteriores de muestras formadas en ventana para la señal central o la señal lateral,  
 en el que el procesador de señal comprende un convertidor de espectro de tiempo (810) para convertir los bloques posteriores de tiempo para obtener bloques posteriores de valores espectrales; y
- 5 en el que el desalineador de señal (900) está configurado para realizar la desalineación mediante el uso de la información sobre los parámetros de alineación de banda estrecha y la información sobre los parámetros de alineación de banda ancha en los bloques de valores espectrales.
28. Aparato de una de las reivindicaciones 21 a 27,  
 10 en el que la señal codificada comprende una pluralidad de ganancias de predicción o parámetros de nivel, en el que el procesador de señal (800) está configurado para calcular valores espectrales del canal izquierdo y el canal derecho mediante el uso de valores espectrales del canal central y un parámetro de nivel o ganancia de predicción para una banda a la que están asociados los valores espectrales (820), y mediante el uso de valores espectrales de la señal lateral decodificada (830).
- 15 29. Aparato de una de las reivindicaciones 21 a 28, en el que el procesador de señal (800) está configurado para calcular valores espectrales de los canales izquierdo y derecho mediante el uso de un parámetro de relleno estéreo para una banda a la cual están asociados los valores espectrales (830).
- 20 30. Aparato de una de las reivindicaciones 21 a 29, en el que el desalineador de señal (900) o el procesador de señal (800) está configurado para realizar un escalado de energía (910) para una banda usando un factor de escala, en el que el factor de escala depende (920) de las energías de la señal central decodificada y la señal lateral decodificada, y
- 25 en el que el factor de escala se limita entre como máximo 2,0 y como mínimo 0,5.
31. Aparato de una de las reivindicaciones 28 a 30, en el que el procesador de señal (800) está configurado para calcular los valores espectrales del canal izquierdo y el canal derecho mediante el uso de un factor de ganancia derivado del parámetro de nivel, en el que el factor de ganancia se deriva del parámetro de nivel mediante el uso de una función no lineal.
- 30 32. Aparato de una de las reivindicaciones 21 a 31, en el que el desalineador de señal (900) está configurado para desalinear una banda del primer y segundo canales decodificados mediante el uso de la información sobre el parámetro de alineación de banda estrecha para los canales mediante el uso de una rotación de los valores espectrales del primer y el segundo canal, en el que los valores espectrales de un canal que tiene una amplitud más alta se giran menos en comparación con los valores espectrales de la banda del otro canal que tiene una amplitud menor.
- 35 33. Procedimiento para la decodificación de una señal de audio multi-canal codificada que comprende una señal central codificada, una señal lateral codificada, información sobre un parámetro de alineación de banda ancha e información sobre una pluralidad de parámetros de alineación de banda estrecha, que comprende:
- 40 la decodificación (700) de la señal central codificada para obtener una señal central decodificada y la decodificación de la señal lateral codificada para obtener una señal lateral decodificada;
- 45 el cálculo (800) de un primer canal decodificado y un segundo canal decodificado a partir de la señal central decodificada y la señal lateral decodificada; y
- la desalineación (900) del primer canal decodificado y el segundo canal decodificado mediante el uso de la información sobre el parámetro de alineación de banda ancha y la información sobre la pluralidad de parámetros de alineación de banda estrecha para obtener una señal de audio multi-canal decodificada.
- 50 34. Programa informático adaptado para realizar, cuando se ejecuta en un ordenador o un procesador, el procedimiento de la reivindicación 19 o el procedimiento de la reivindicación 33.

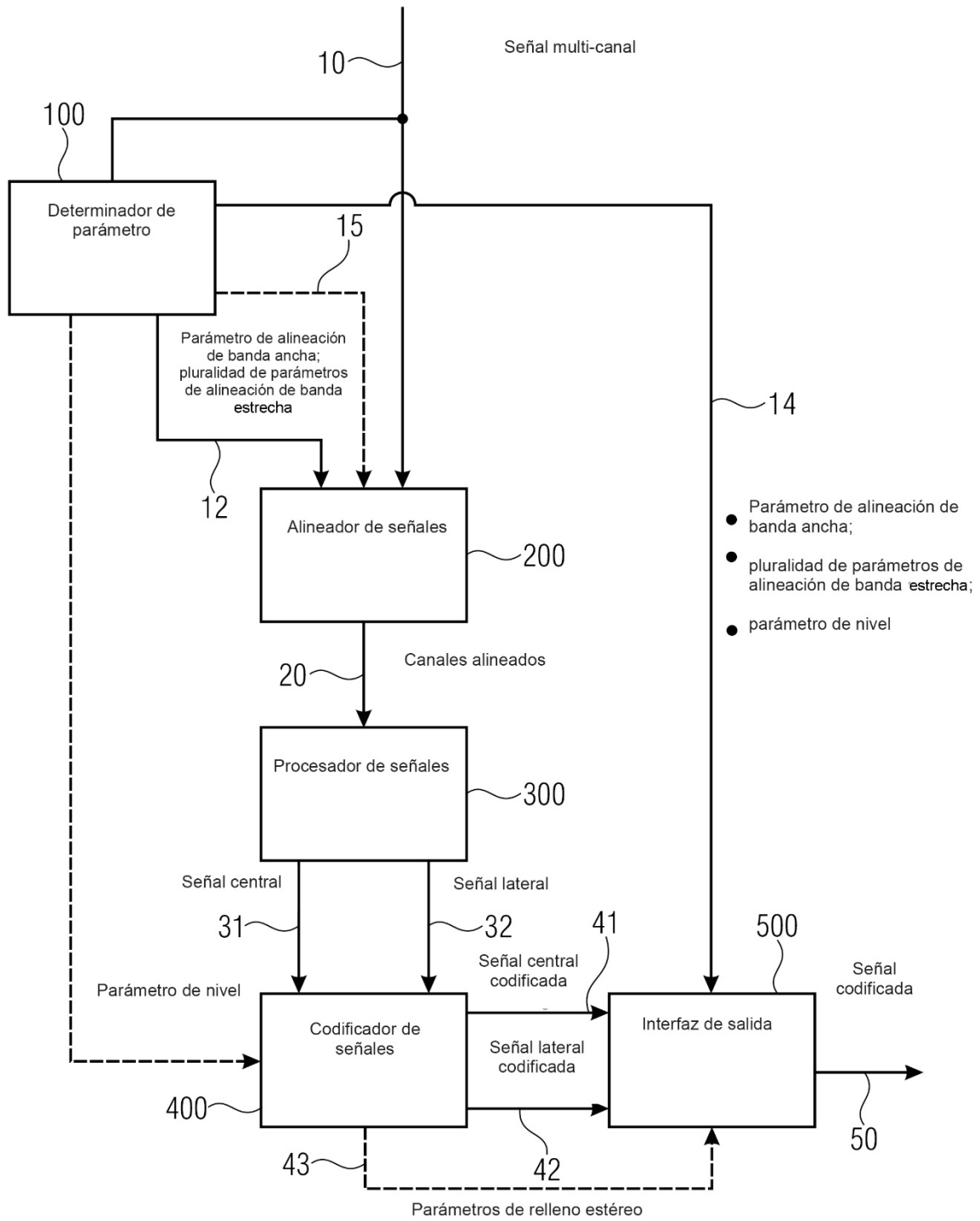


Fig. 1

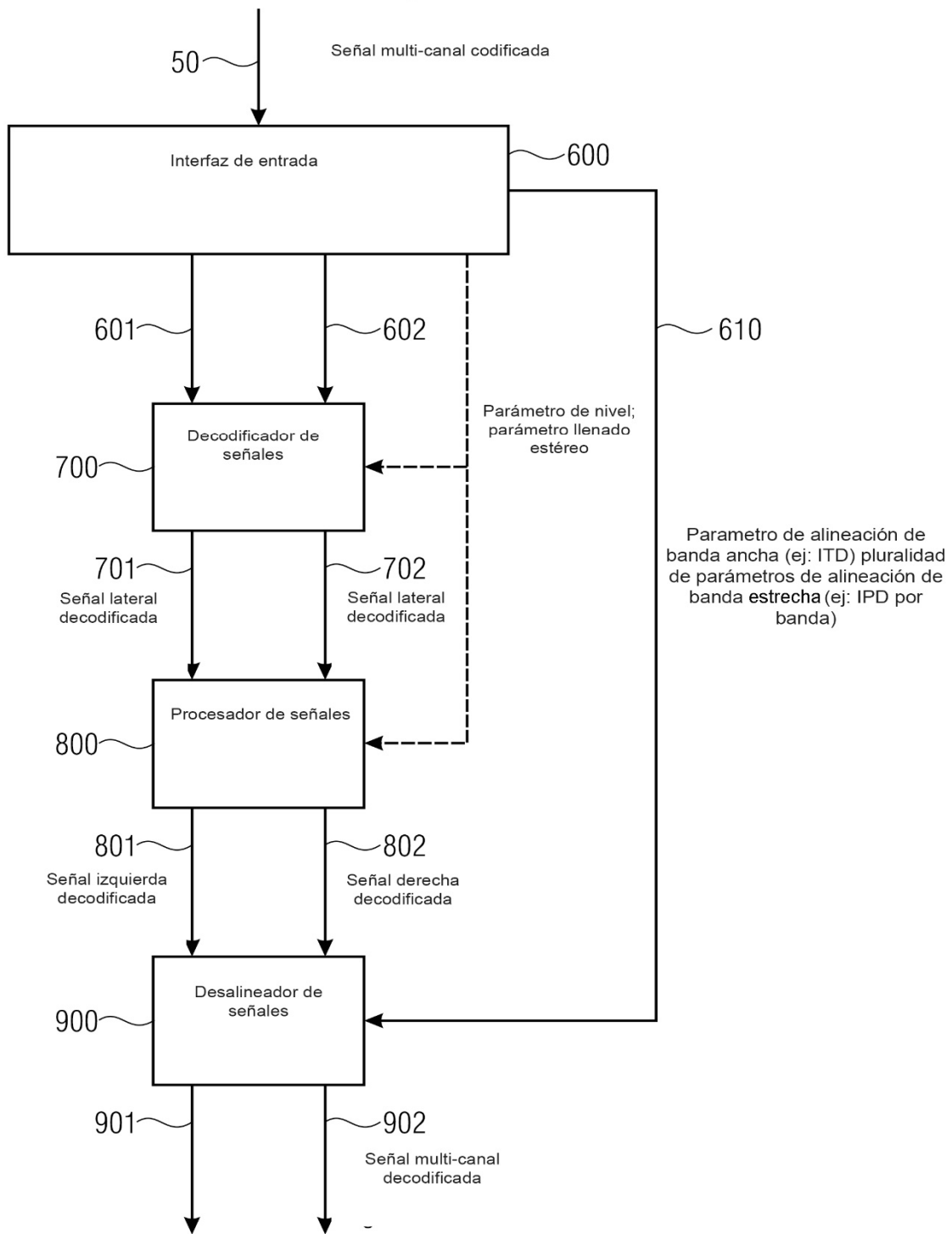
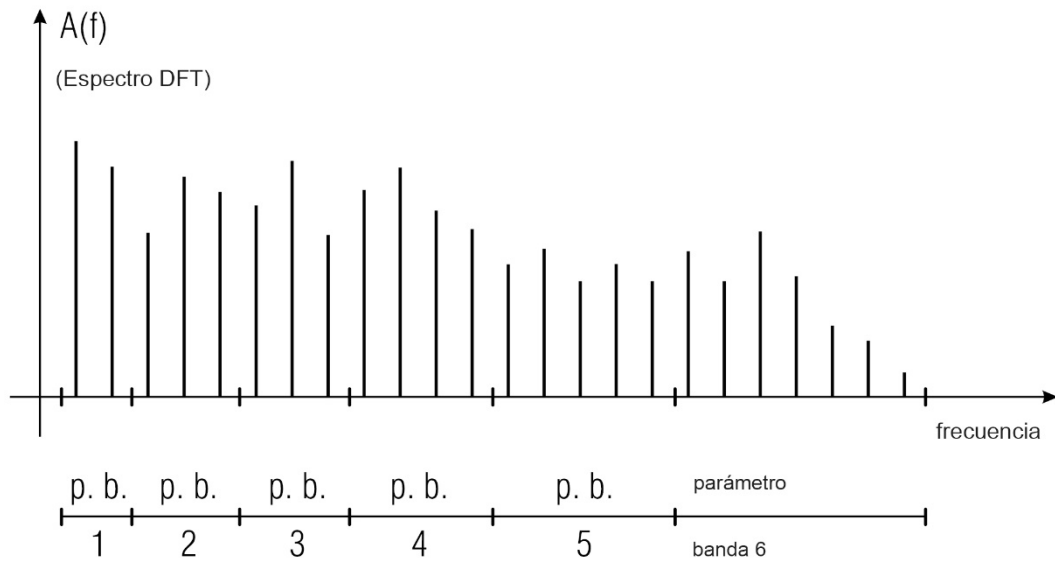


Fig. 2





- Un solo parámetro de alineación de banda ancha para todo el espectro (por ejemplo, canal 1 a banda de p. 6);
- pluralidad de parámetros de alineación de banda angosta para las bandas de parámetros 1, 2, 3, 4, es decir, cuatro parámetros de banda estrecha;
- parámetros de nivel para cada banda de parámetros, por ejemplo, 6 parámetros de nivel;
- parámetros de relleno estéreo para las bandas de parámetros 4, 5, 6, por ejemplo, tres parámetros de relleno estéreo;
- señal lateral (residual) para las bandas de parámetros 1, 2, 3;
- más líneas espectrales en las bandas altas, por ejemplo, siete líneas espectrales en la banda de parámetros 6 frente a tres líneas espectrales en la banda de parámetros 2.

Fig. 3

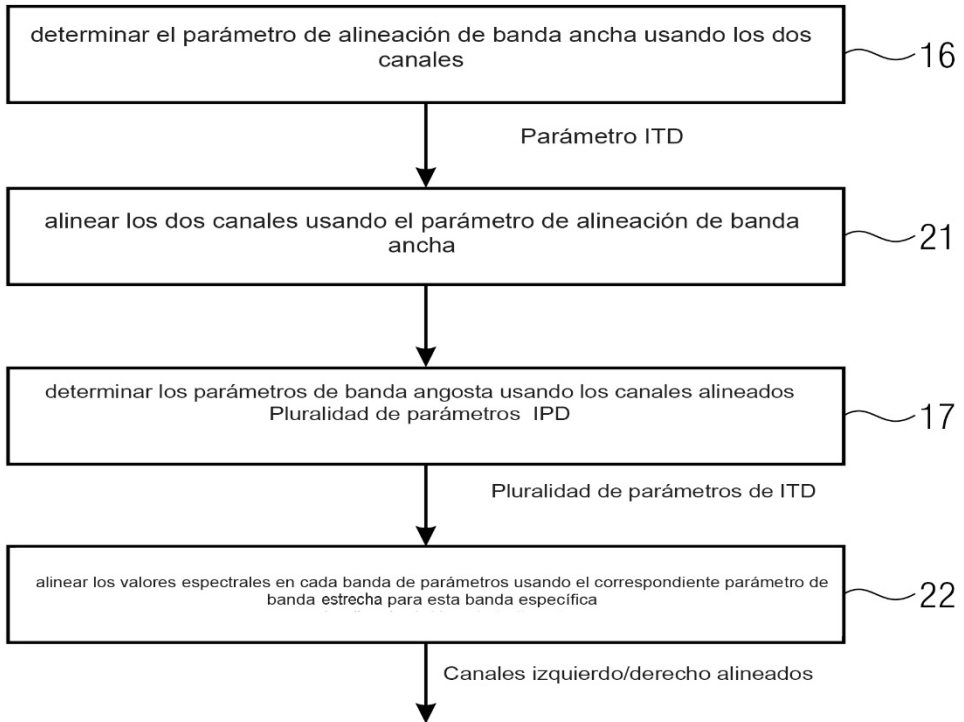


Fig. 4a

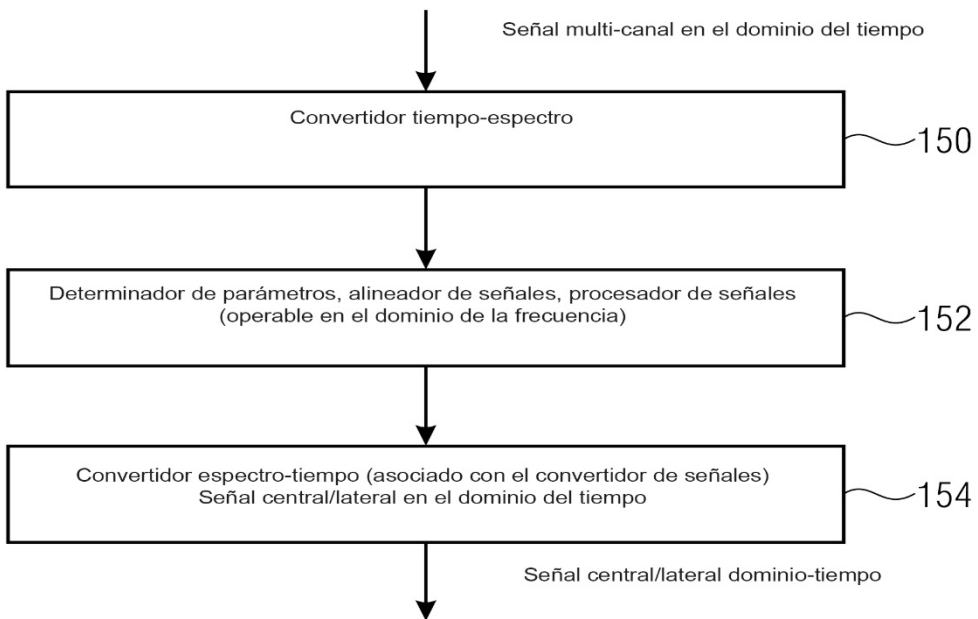


Fig. 4b

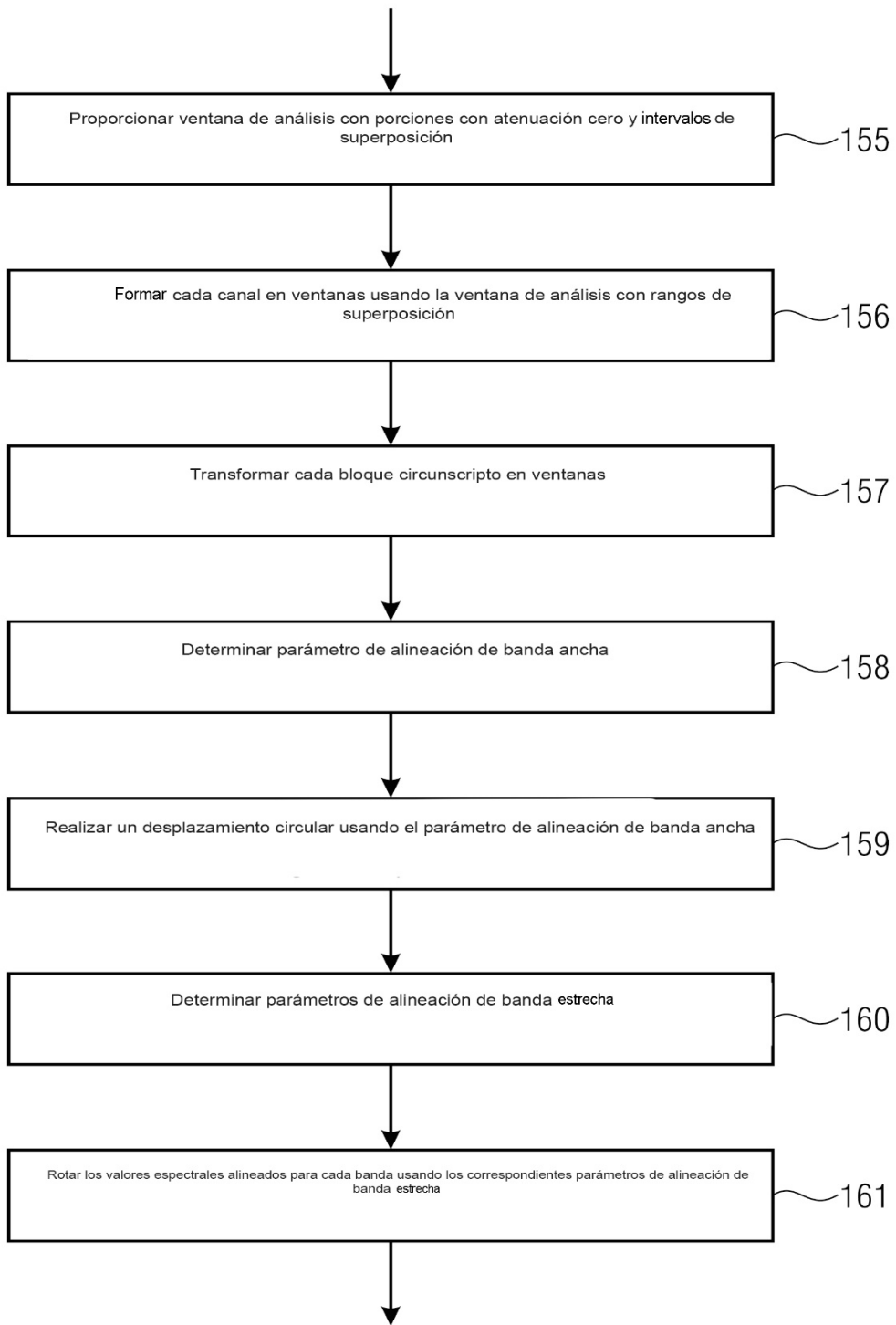


Fig. 4c

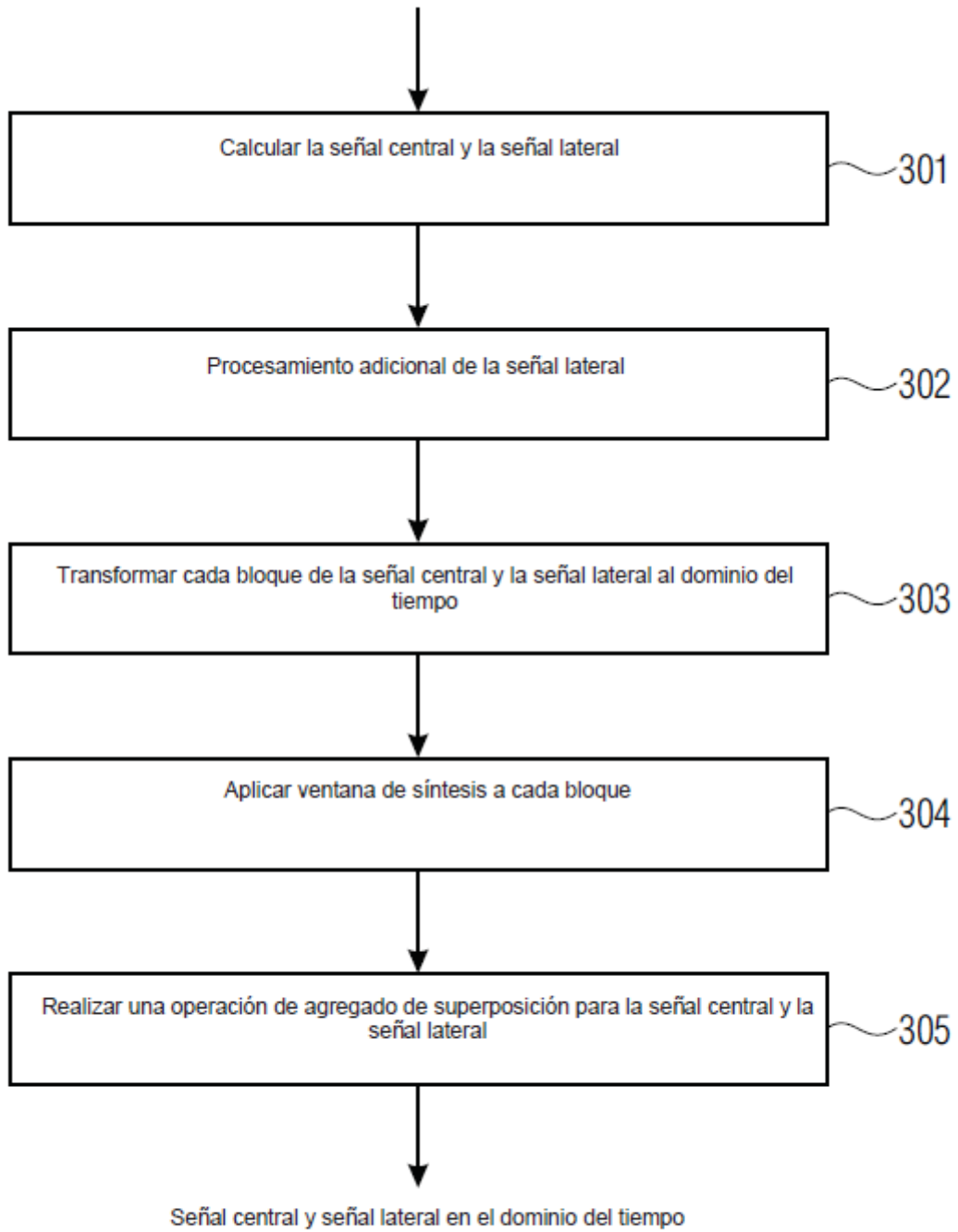
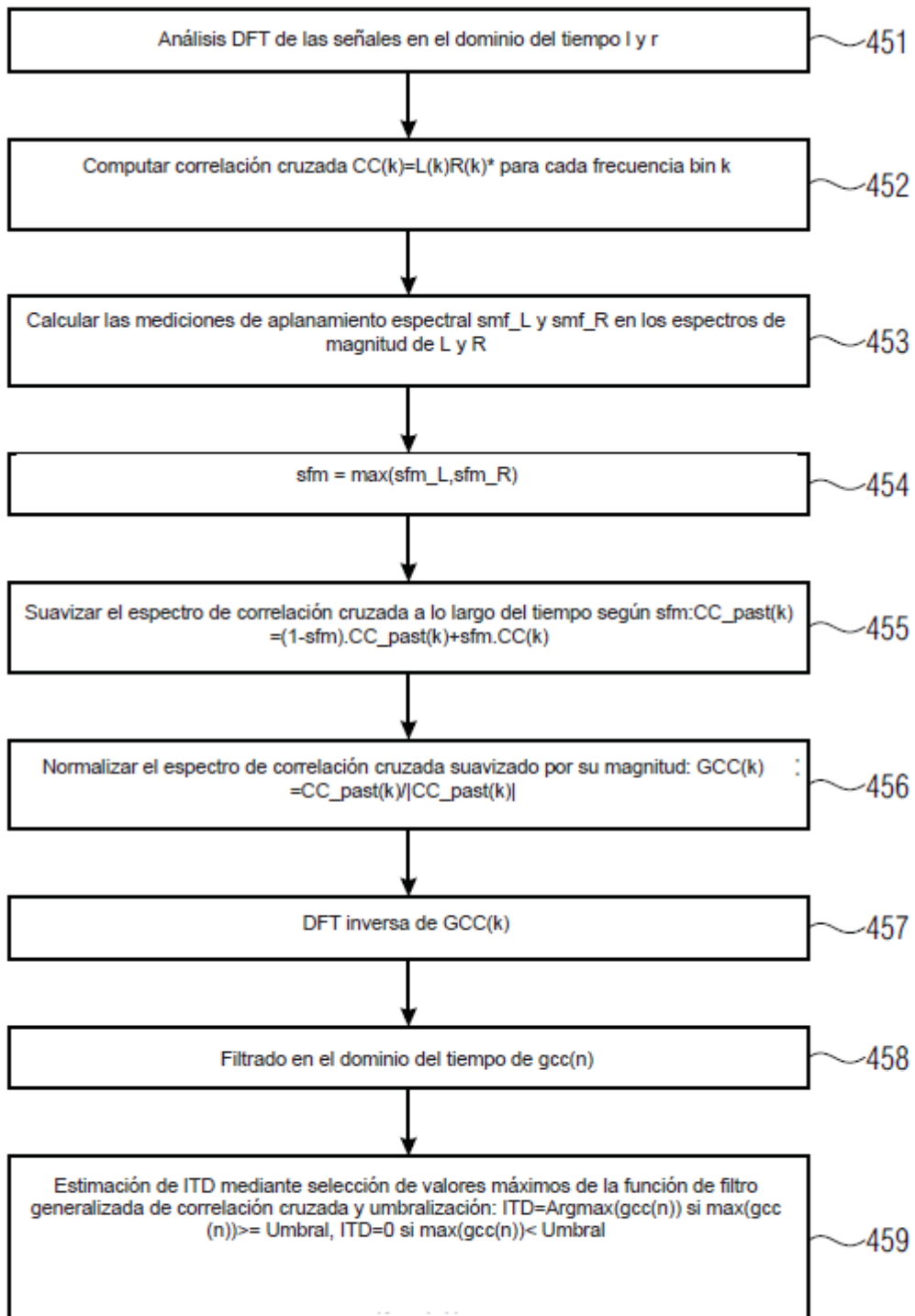


Fig. 4d



Estimación de ITD

Fig. 4e

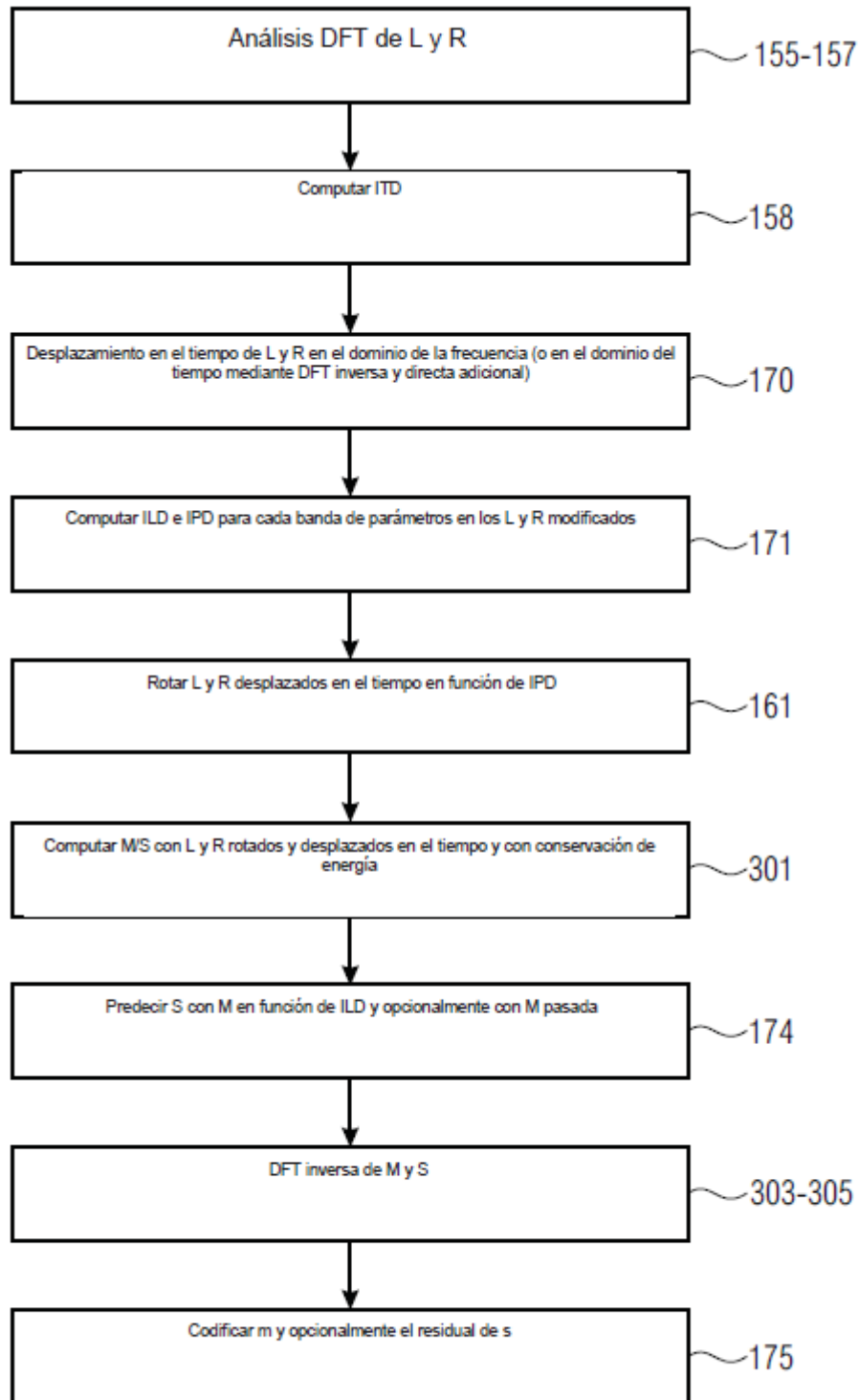


Fig. 5

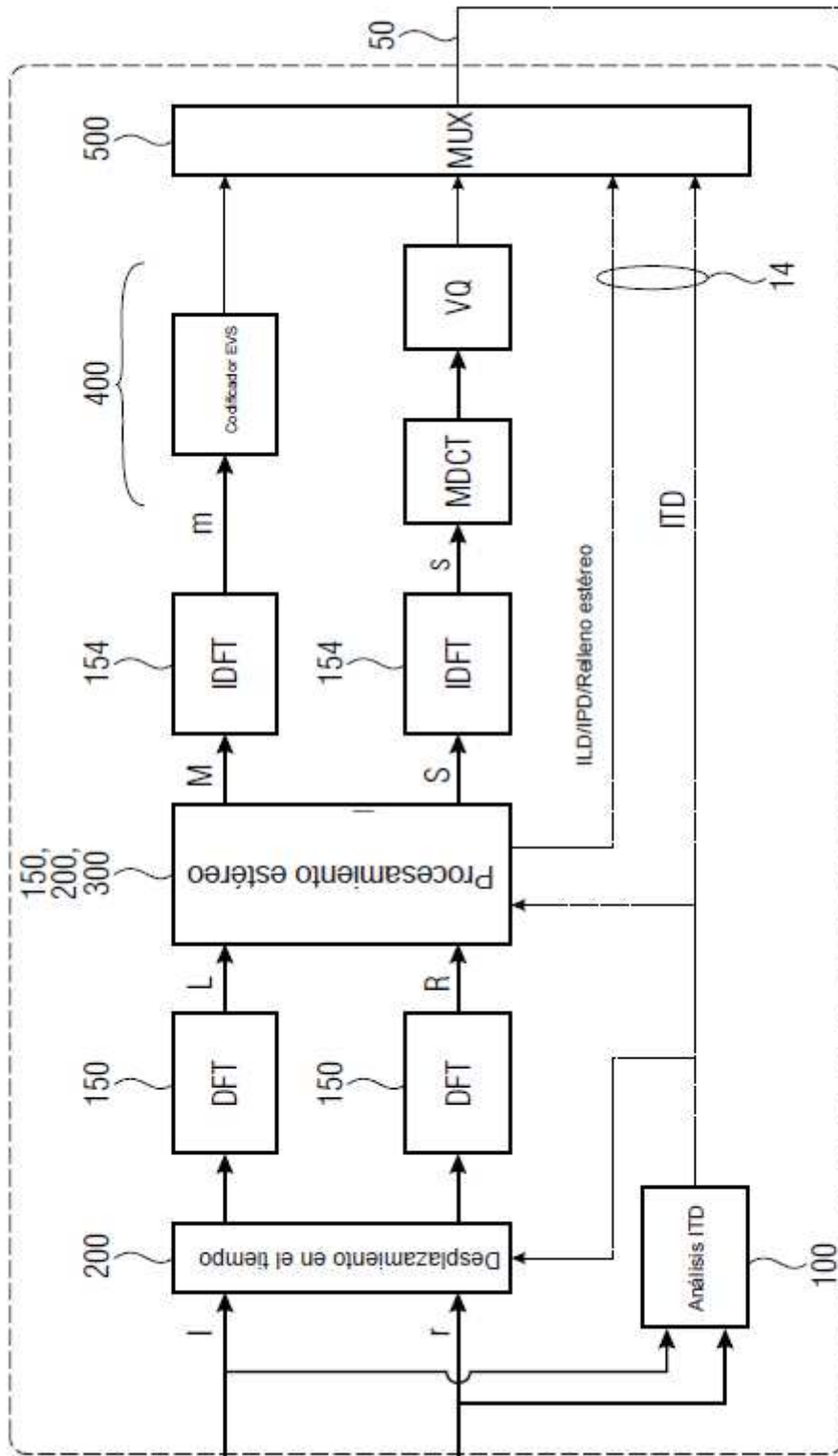


Fig. 6a

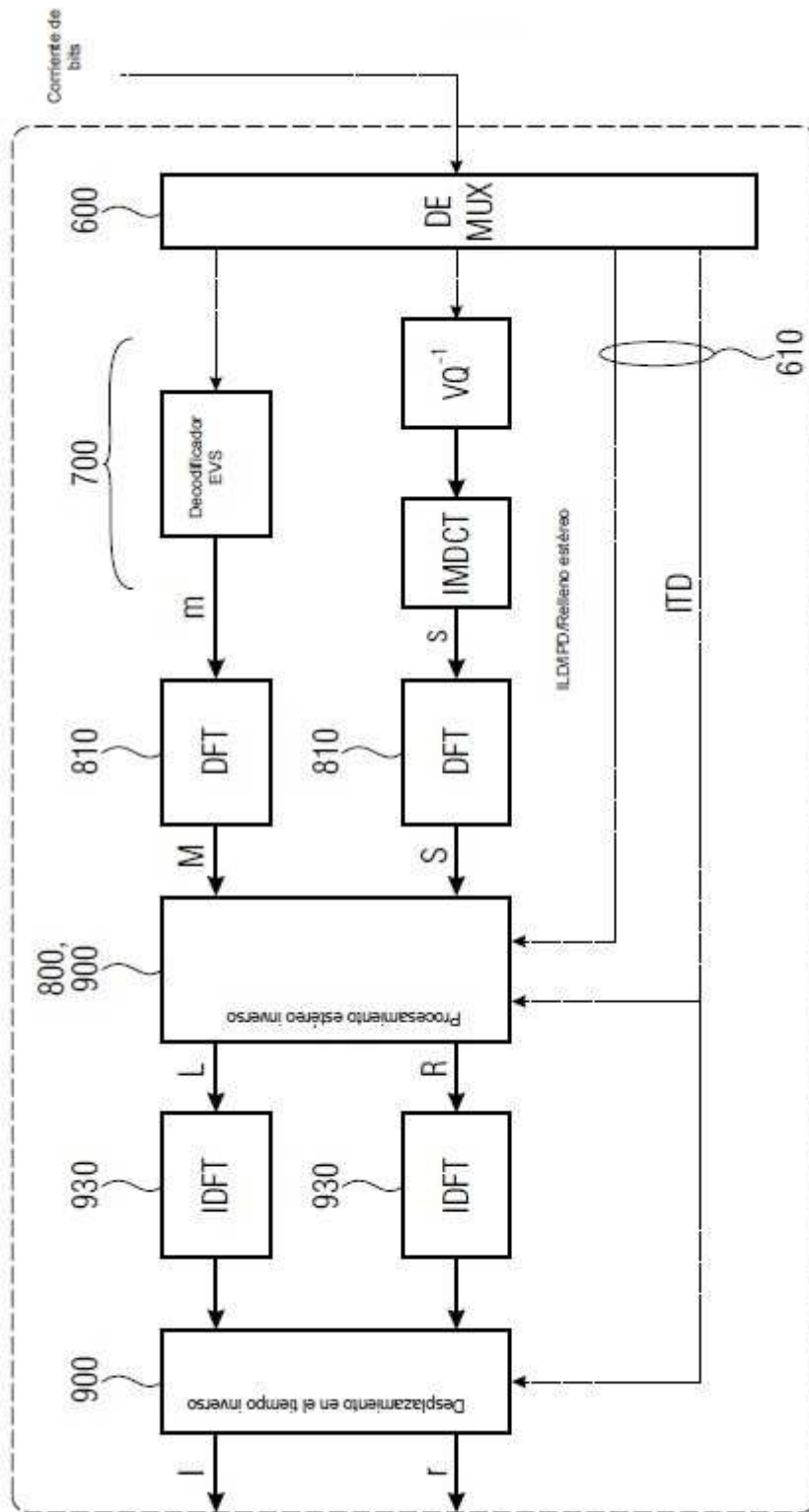


Fig. 6b



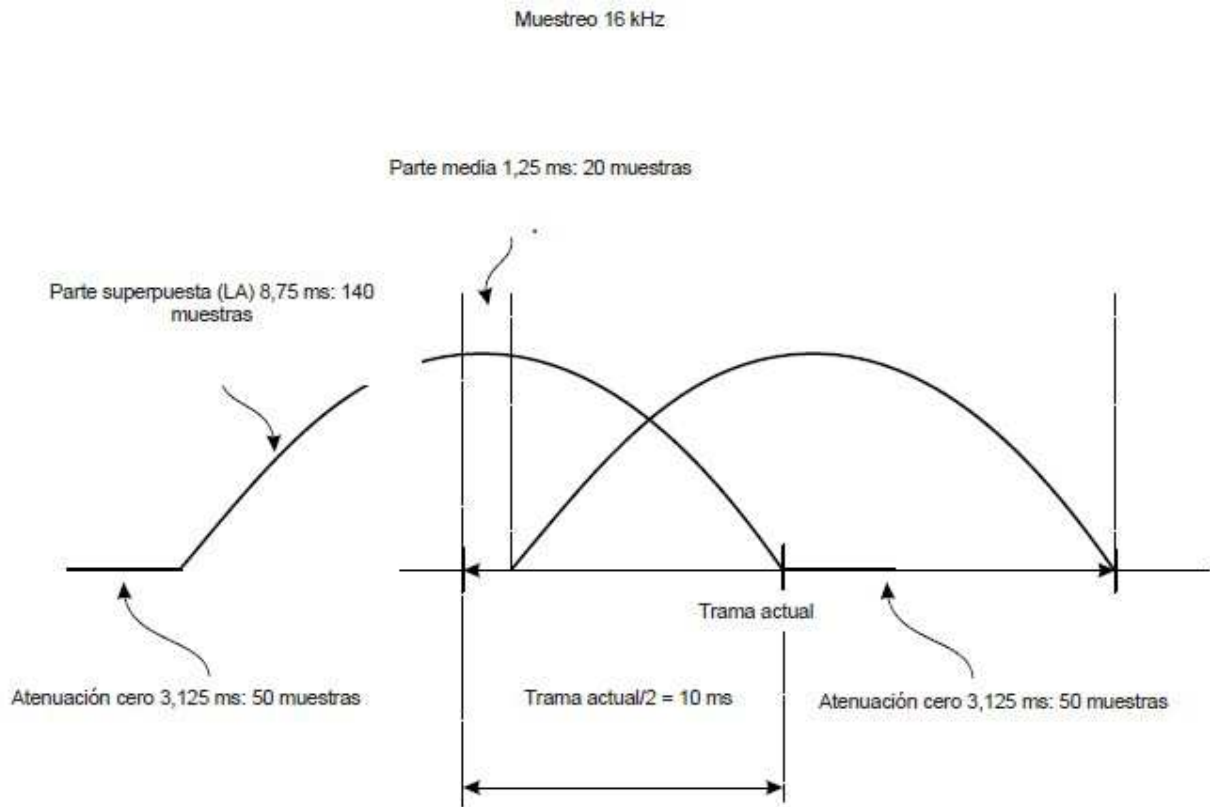


Fig. 7

Parámetros	Bits/banda	Número de bandas	Tasa de bits
ILD	5	El total de 12 bandas	3.00 kbps
IPD	3	Hasta 2,5kHz	1.05 kbps
ITD	8	Espectro completo	0.6 kbps
Relleno estéreo	3	Desde 1kHz	0.9 kbps
TOTAL			-5 kbps

Fig. 8

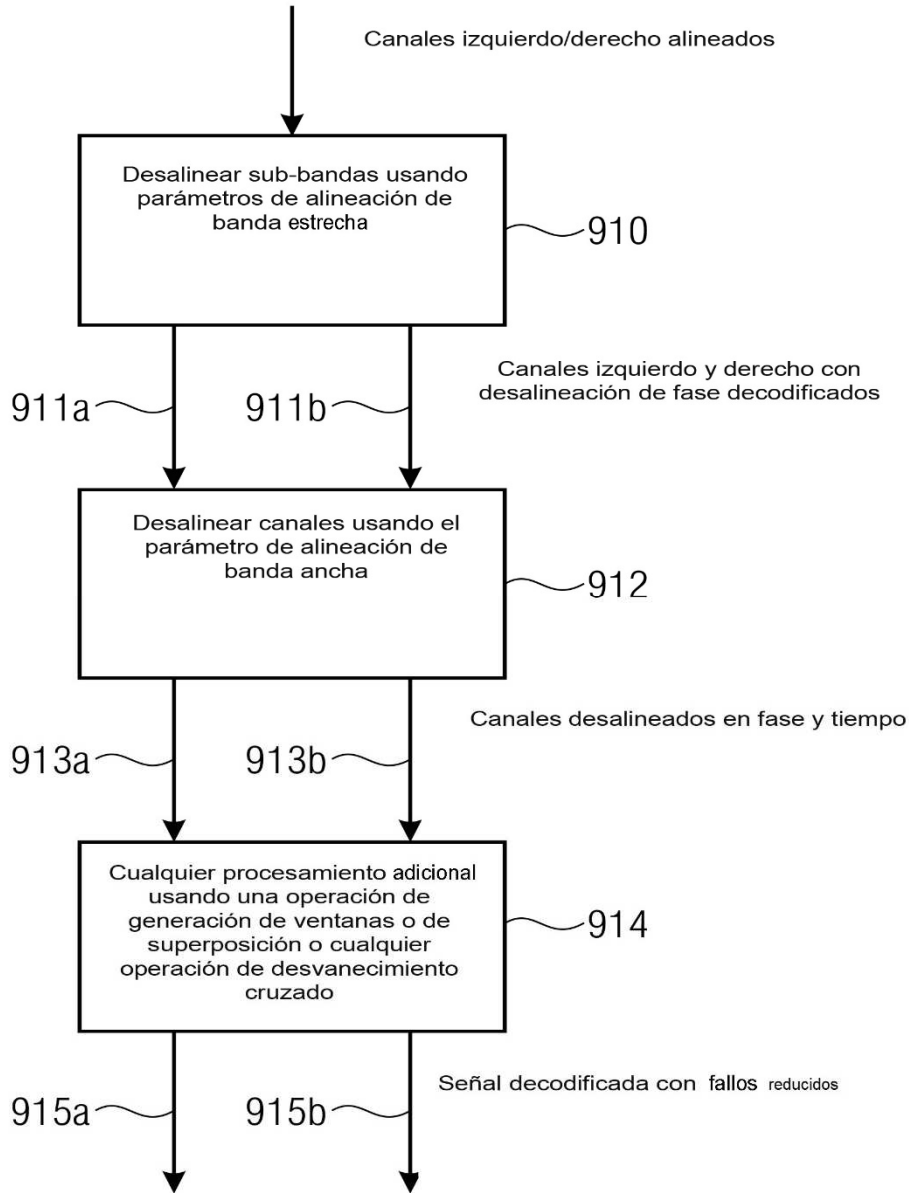


Fig. 9a

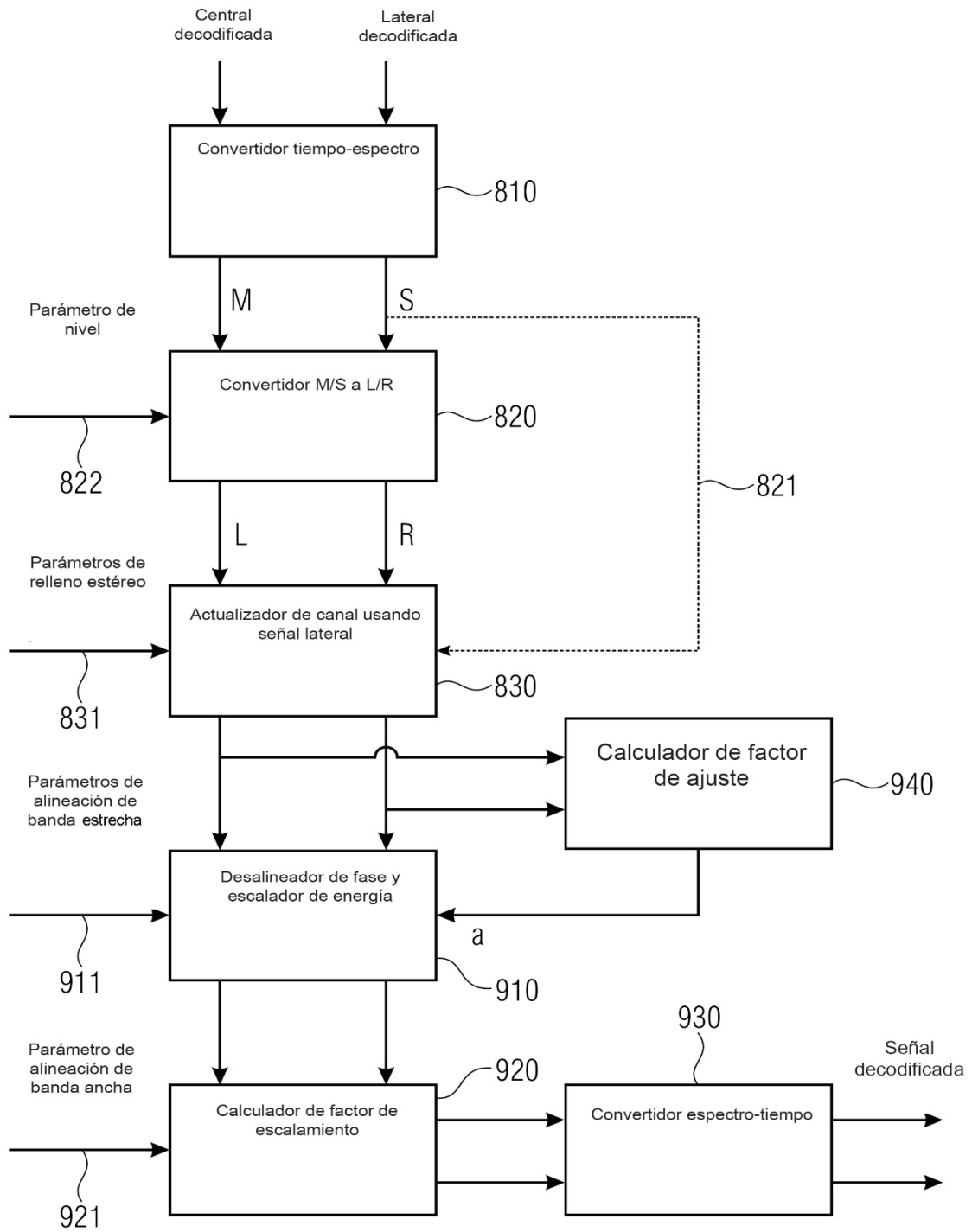


Fig. 9b

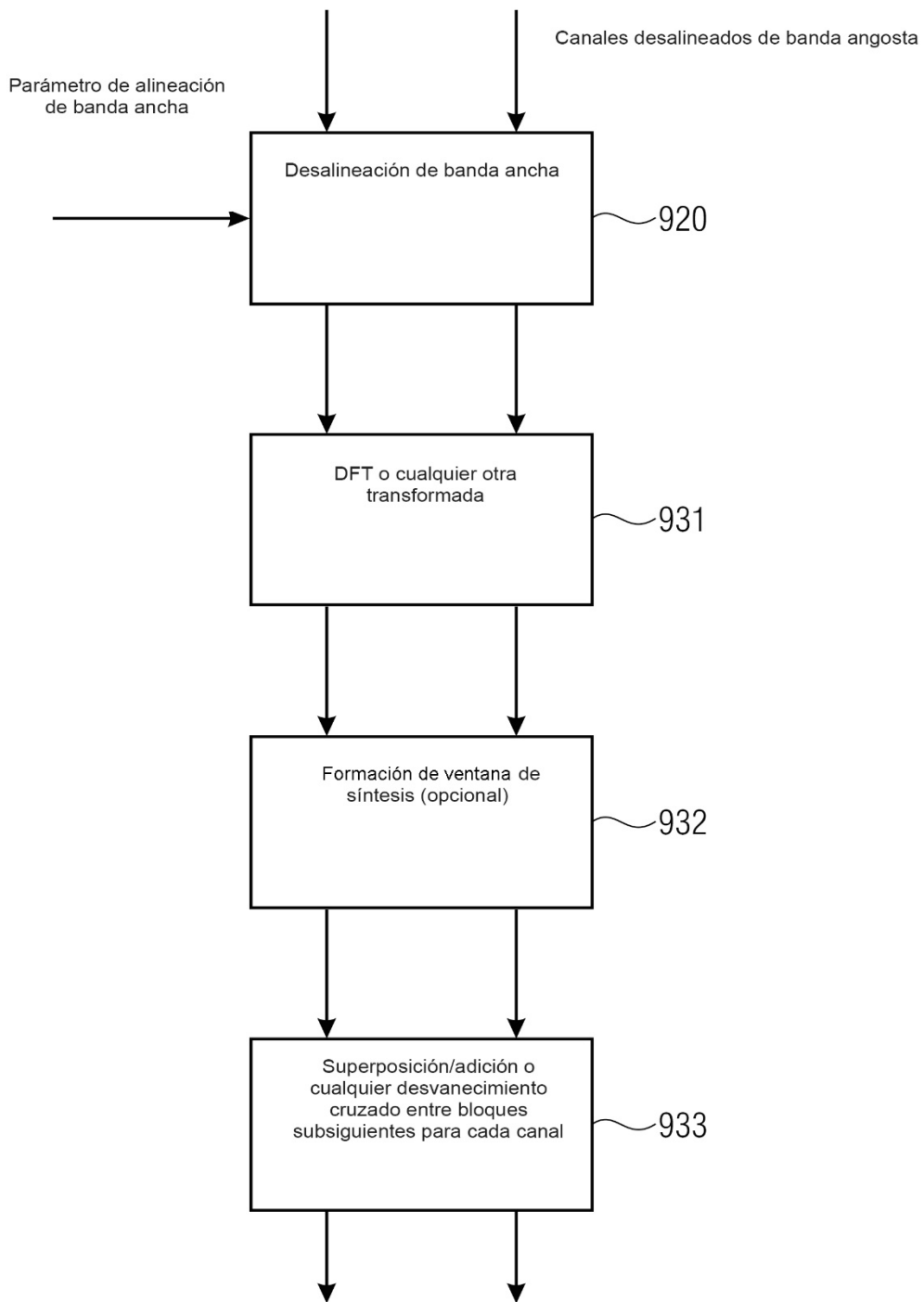


Fig. 9c