

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 701 456**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/04 (2013.01)

G10L 19/00 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.03.2011 E 15176784 (5)**

97 Fecha y número de publicación de la concesión europea: **19.09.2018 EP 2947656**

54 Título: **Codificación de señales de audio multicanal usando predicción compleja y codificación diferencial**

30 Prioridad:

09.04.2010 US 322688 P

13.07.2010 US 363906 P

13.07.2010 EP 10169432

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

22.02.2019

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (50.0%)**

Hansastraße 27c

80686 München, DE y

DOLBY INTERNATIONAL AB (50.0%)

72 Inventor/es:

PURNHAGEN, HEIKO;

CARLSSON, PONTUS;

VILLEMOES, LARS;

ROBILLIARD, JULIEN;

NEUSINGER, MATTHIAS;

HELMRICH, CHRISTIAN;

HILPERT, JOHANNES;

RETTELBACH, NIKOLAUS;

DISCH, SASCHA y

EDLER, BERND

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 701 456 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Codificación de señales de audio multicanal usando predicción compleja y codificación diferencial

5 La presente invención se refiere a procesamiento de audio y, particularmente, al procesamiento de audio multicanal de una señal multicanal que tiene dos o más señales de canal.

10 En el campo del procesamiento de estéreo o multicanal se conoce aplicar la denominada codificación de estéreo central/lateral (*mid/side stereo coding*). Según este concepto, se forma una combinación de la señal de canal de audio izquierda o primera y la señal de canal de audio derecha o segunda para obtener una señal central o mono M. Adicionalmente se forma una diferencia entre la señal de canal izquierda o primera y la señal de canal derecha o segunda para obtener la señal lateral S. Este método de codificación central/lateral da como resultado una ganancia de codificación significativa, cuando la señal izquierda y la señal derecha son muy similares entre sí, ya que la señal lateral se volverá bastante pequeña. Típicamente, una ganancia de codificación de una etapa de codificador de entropía/cuantificador resultará superior, cuando el intervalo de valores que van a cuantificarse/codificarse por entropía es menor. De este modo, para una PCM (modulación por codificación de pulsos, según sus siglas en inglés) o un codificador aritmético de entropía o basado en Huffman, la ganancia de codificación aumenta, cuando la señal lateral se vuelve más pequeña. Sin embargo, existen determinadas situaciones en las que la codificación central/lateral no dará como resultado una ganancia de codificación. La situación puede producirse cuando las señales en ambos canales están desfasadas una con respecto a otra, por ejemplo, en 90°. Entonces, la señal central y la señal lateral pueden estar en un intervalo bastante similar y, por lo tanto, la codificación de la señal central y la señal lateral utilizando el codificador por entropía no dará como resultado una ganancia de codificación y puede, incluso, dar como resultado un aumento de tasa de bits. Por lo tanto, puede aplicarse una codificación central/lateral de frecuencia selectiva con el fin de desactivar la codificación central/lateral en bandas, en la que la señal lateral no se vuelve más pequeña hasta un cierto grado con respecto a, por ejemplo, la señal izquierda original.

30 A pesar de que la señal lateral resultará cero, cuando las señales izquierda y derecha son idénticas, lo que dará como resultado una máxima ganancia de codificación debido a la eliminación de la señal lateral, la situación nuevamente resulta diferente cuando la señal central y la señal lateral son idénticas con respecto a la forma de la forma de onda, pero la única diferencia entre ambas señales radica en sus amplitudes generales. En este caso, cuando se presume adicionalmente que la señal lateral no tiene un desfase con respecto a la señal central, la señal lateral aumenta significativamente, a pesar de que, por otra parte, la señal central no disminuye demasiado con respecto a su intervalo de valores. Cuando dicha situación se produce en una determinada banda de frecuencia, entonces se desactivaría nuevamente la codificación central/lateral debido a la falta de ganancia de codificación. La codificación central/lateral puede aplicarse en forma de frecuencia selectiva o, alternativamente, puede aplicarse en el dominio de tiempo.

40 Existen técnicas de codificación multicanal alternativas que no están basadas en un tipo de método de forma de onda como codificación central/lateral, sino que están basadas en el procesamiento paramétrico según determinadas señales binaurales. Dichas técnicas se conocen según los términos "codificación binaural de la señal", "codificación de estéreo paramétrico" o "codificación del entorno de MPEG (Grupo de expertos en imágenes en movimiento, según sus siglas en inglés)". En este caso, se calculan determinadas señales para una pluralidad de bandas de frecuencia. Estas señales incluyen diferencias de nivel entre canales, medidas de coherencia entre canales, diferencias de tiempo entre canales y/o diferencias de fase entre canales. Estos métodos parten de la premisa de que una impresión multicanal sentida por el oyente no necesariamente está basada en las formas de onda detalladas de los dos canales, sino que está basada en las señales suministradas en forma de frecuencia selectiva o en la información exacta entre canales. Esto significa que, en una máquina de presentación se debe tener cuidado de presentar señales multicanal que reflejen las señales con exactitud, si bien las formas de onda no tienen una importancia significativa.

55 Este método puede resultar complejo particularmente en el caso en el que el decodificador tenga que aplicar un procesamiento de descorrelación con el fin de crear artificialmente señales estéreo que están descorrelacionadas una con respecto a otra, a pesar de que la totalidad de estos canales proviene de uno y del mismo canal de mezcla descendente. Los descorreladores para este fin son complejos, dependiendo de su implementación, y pueden introducir distorsiones, particularmente en el caso de porciones de señales transitorias. Adicionalmente, a diferencia de la codificación de forma de onda, el método de codificación paramétrica es un método de codificación con pérdidas, que, inevitablemente, da como resultado una pérdida de información no solamente introducida por la cuantificación típica sino también introducida al considerar las señales binaurales en lugar de las formas de onda en particular. Este método da como resultado tasas de bits muy bajas, pero puede incluir compromisos de calidad.

60 Existen desarrollos recientes en cuanto a la codificación de voz y audio unificada (USAC, según sus siglas en inglés) ilustrada en la figura 7a. Un decodificador 700 principal lleva a cabo una operación de decodificación de la señal

estéreo codificada en la entrada 701, que puede ser una señal codificada central/lateral. El decodificador principal emite una señal central en la línea 702 y una señal lateral o residual en la línea 703. Ambas señales se transforman a un dominio QMF (filtro espejo en cuadratura, según sus siglas en inglés) por los bancos 704 y 705 de filtros QMF. Luego se aplica un decodificador 706 del entorno de MPEG para generar una señal 707 izquierda de canal y una
 5 señal 708 derecha de canal. Estas señales de banda baja se introducen posteriormente en un decodificador 709 de replicación espectral de ancho de banda (SBR, según sus siglas en inglés), que produce señales izquierda y derecha de banda ancha en las líneas 710 y 711, que luego se transforman a un dominio de tiempo por los bancos de filtros 712, 713 de síntesis QMF de modo que se obtienen las señales izquierda y derecha de banda ancha L, R.

10 La figura 7b ilustra una situación en la que el decodificador 706 del entorno de MPEG llevaría a cabo una decodificación central/lateral. Alternativamente, el bloque decodificador 706 del entorno de MPEG podría llevar a cabo una decodificación paramétrica basada en la señal binaural para generar señales estéreo a partir de una única señal mono del decodificador principal. Naturalmente, el decodificador 706 del entorno de MPEG también podría
 15 generar una pluralidad de señales de salida de banda baja para introducirse en el bloque 709 decodificador de SBR que usa información paramétrica tal como diferencias de nivel entre canales, medidas de coherencia entre canales u otros parámetros de información entre canales de este tipo.

20 Cuando el bloque decodificador 706 del entorno de MPEG lleva a cabo la decodificación central/lateral ilustrada en la figura 7b puede aplicarse un factor de ganancia real g y DMX/RES y L/R son señales izquierda/derecha y de mezcla descendente/residual representadas, respectivamente, en el dominio QMF complejo híbrido.

25 El uso de una combinación de un bloque 706 y un bloque 709 solo provoca un pequeño aumento en la complejidad computacional en comparación con un decodificador estéreo utilizado como base, debido a que la representación del QMF complejo de la señal ya está disponible como parte del decodificador de SBR. Sin embargo, en una configuración que no es de SBR, la codificación estéreo basada en QMF, tal como se propone en el contexto de la USAC, daría como resultado un aumento significativo en la complejidad computacional debido a los bancos de QMF necesarios, que en este ejemplo necesitarían bancos de análisis de 64 bandas y bancos de síntesis de 64 bandas. Estos bancos de filtros solo tendrían que agregarse para el fin de la codificación estéreo.

30 En el sistema de USAC del MPEG en desarrollo, sin embargo, también existen modos de codificación a altas tasas de bits en las que la SBR típicamente no se utiliza.

35 El documento WO 9016136 da a conocer un sistema de audioconferencia (por ejemplo, estéreo) polifónico, en el que canales derecho e izquierdo de entrada están alineados en el tiempo mediante etapas de retardo variable, controladas mediante un calculador de retardo (por ejemplo, derivando el valor de correlación cruzada máximo), y entonces sumadas en un elemento de suma y restadas en un elemento de resta para formar señales de diferencia y suma. La señal de suma se transmite a una calidad relativamente elevada; la señal de diferencia se reconstruye en el decodificador mediante predicción a partir de la señal de suma usando un filtro adaptativo. El filtro adaptativo de
 40 decodificador está configurado o bien mediante coeficientes de filtro recibidos o bien usando adaptación hacia atrás, a partir de una señal residual recibida producida por un filtro adaptativo correspondiente en el codificador, o ambos. Preferiblemente, el filtro adaptativo es un filtro enrejado, que emplea un algoritmo de gradiente para actualización de coeficiente. La complejidad del filtro adaptativo se reduce mediante blanqueado previo, en el codificador, usando tanto la señal de diferencia como la señal de suma filtros de blanqueado correspondientes derivados del canal de suma.

45 Los siguientes documentos son ejemplares de esquemas de (de)decodificación de audio multicanal con lo que se estima una señal de diferencia a partir de una señal de mezcla descendente/sumador/mono mediante un coeficiente de predicción de valor complejo:

50 HEIKO PURNHAGEN *ET AL*: "Technical description of proposed Unified Stereo Coding in USAC", 90. MPEG MEETING; 26-10-2009 - 30-10-2009; XIAN; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11), N° M16921, 23 de octubre de 2009 (23-10-2009);

55 MAX NEUENDORF (EDITOR): "WD5 of USAC", 90. MPEG MEETING; 26-10-2009 - 30-10-2009; XIAN; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11), N° N11040, 8 de diciembre de 2009 (08-12-2009), páginas 1-146;

60 El documento WO 2009141775 A1 da a conocer un aparato de mezcla ascendente estéreo paramétrico que genera una señal izquierda. Dicha mezcla ascendente estéreo paramétrico está caracterizado porque comprende medios para predecir una señal de diferencia que comprende una diferencia entre la señal izquierda y la señal derecha (207) basándose en la señal mezclada ascendente mono aumentada con un coeficiente de predicción. Dicho coeficiente de predicción está derivado de parámetros espaciales. Dicho aparato de mezcla ascendente estéreo paramétrico comprende además medios aritméticos para derivar la señal izquierda y la señal derecha basándose en una suma y una diferencia de la señal de mezcla ascendente mono y dicha señal de diferencia.

El documento WO 2011124608 A1 da a conocer métodos y dispositivos para decodificar y codificar estéreo usando predicción compleja en el campo de frecuencia. En una realización, un método de decodificación, para obtener una señal estéreo de salida de una señal estéreo de entrada codificada mediante codificación de predicción compleja y que comprende primeras representaciones de campo de frecuencia de dos canales de entrada, comprende las etapas de mezcla ascendente de: (i) calcular una segunda representación de campo de frecuencia de un primer canal de entrada; y (ii) calcular un canal de salida basándose en las representaciones de campo de frecuencia primera y segunda del primer canal de entrada, la primera representación de campo de frecuencia del segundo canal de entrada y un coeficiente de predicción compleja. La mezcla ascendente puede suspenderse en respuesta a datos de control.

Un objetivo de la presente invención consiste en proporcionar un concepto de procesamiento de audio mejorado, que, por una parte, produzca una alta ganancia de codificación y, por otra parte, dé como resultado una buena calidad de audio y/o una complejidad computacional reducida.

Este objetivo se logra mediante un decodificador de audio según la reivindicación 1, un codificador de audio según la reivindicación 7, un método de decodificación de audio según la reivindicación 15, un método de codificación de audio según la reivindicación 16, o un programa informático según la reivindicación 17.

La presente invención se basa en la conclusión de que una ganancia de codificación del método de codificación de forma de onda de alta calidad se puede mejorar significativamente mediante una predicción de una segunda señal de combinación utilizando una primera señal de combinación, en la que ambas señales de combinación provienen de las señales de canal originales que usan una regla de combinación tal como la regla de combinación central/lateral. Se ha descubierto que esta información de predicción se calcula por un predictor en un codificador de audio de modo que se cumple un objetivo de optimización, se incurre solo en una pequeña sobrecarga, pero da como resultado una disminución significativa de la tasa de bits necesaria para la señal lateral sin perder ninguna calidad de audio, ya que la predicción de la invención es, sin embargo, una codificación basada en forma de onda y no un método de codificación multicanal o estéreo basado en parámetros. Con el fin de reducir la complejidad computacional, se prefiere realizar una codificación de dominio de frecuencia, en el que la información de predicción proviene de datos de entrada del dominio de frecuencia en una forma selectiva de bandas. El algoritmo de conversión para convertir la representación de dominio de tiempo en una representación espectral es preferiblemente un proceso críticamente muestreado tal como una transformada coseno discreta modificada (MDCT, según sus siglas en inglés) o una transformada seno discreta modificada (MDST, según sus siglas en inglés), que es diferente de una transformada compleja porque se calculan solo valores reales o solo valores imaginarios, mientras que, en una transformada compleja se calculan valores reales y complejos de un espectro dando como resultado un sobremuestreo dos veces mayor. Preferiblemente se utiliza una transformada basada en introducción y cancelación de solapamiento. La MDCT, en particular, es una transformada de este tipo y permite una atenuación cruzada entre bloques posteriores sin ninguna sobrecarga debido a la propiedad de cancelación de solapamiento de dominio de tiempo (TDAC, según sus siglas en inglés) que se conoce bien, que se obtiene mediante el procesamiento de solapar-agregar en el lado del decodificador.

La información de predicción calculada en el codificador, transmitida al decodificador y utilizada en el decodificador comprende una parte imaginaria que puede reflejar ventajosamente las diferencias de fase entre los dos canales de audio en cantidades seleccionadas arbitrariamente entre 0° y 360°. La complejidad computacional se reduce significativamente cuando se aplica solamente una transformada de valor real o, en general, una transformada que proporciona, o bien un espectro real solamente o bien proporciona un espectro imaginario solamente. Con el fin de utilizar esta información de predicción imaginaria que indica un desfase entre una determinada banda de la señal izquierda y una banda correspondiente de la señal derecha, un convertidor real-a-imaginario o, dependiendo de la implementación de la transformada, un convertidor imaginario-a-real se proporciona en el decodificador con el fin de calcular una señal residual de predicción a partir de la primera señal de combinación, que se hace rotar de fase con respecto a la señal de combinación original. Esta señal residual de predicción con fase rotada puede entonces combinarse con la señal residual de predicción transmitida en la corriente de bits para regenerar una señal lateral, que, finalmente, puede combinarse con la señal central para obtener el canal izquierdo decodificado en una determinada banda y el canal derecho decodificado en esta banda.

Para aumentar la calidad de audio, el mismo convertidor real-a-imaginario o el convertidor imaginario-a-real, que se aplica en el lado del decodificador también se implementa en el lado del codificador, cuando la señal residual de predicción se calcula en el codificador.

La presente invención es ventajosa porque proporciona una mejora de la calidad de audio y una reducción de la tasa de bits en comparación con los sistemas que tienen la misma tasa de bits o que tienen la misma calidad de audio.

Adicionalmente, se obtienen ventajas con respecto a una eficiencia computacional de codificación estéreo unificada que resulta útil en el sistema de USAC del MPEG a altas tasas de bits, en el que la SBR típicamente no se utiliza. En

lugar de procesar la señal en el dominio QMF híbrido complejo, estos métodos implementan una codificación estéreo predictiva de base residual en el dominio nativo de MDCT del codificador de transformada estéreo subyacente.

5 Según un aspecto de la presente invención, la presente invención comprende un aparato o método para generar una señal estéreo por predicción compleja en el dominio de MDCT, en la que la predicción compleja se lleva a cabo en el dominio de MDCT utilizando una transformada real-a-compleja, en la que esta señal estéreo puede ser, o bien una señal estéreo codificada en el lado del codificador o bien, alternativamente, puede ser una señal estéreo decodificada/transmitida, cuando el aparato o método para generar la señal estéreo se aplica en el lado del decodificador.
10

Realizaciones preferidas de la presente invención se describen a continuación con respecto a los dibujos adjuntos, en los que:

- 15 la figura 1 es un diagrama de una realización preferida de un decodificador de audio;
la figura 2 es un diagrama de bloques de una realización preferida de un codificador de audio;
la figura 3a ilustra una implementación de la calculadora del codificador de la figura 2;
20 la figura 3b ilustra una implementación alternativa de la calculadora del codificador de la figura 2;
la figura 3c ilustra una regla de combinación central/lateral que va a aplicarse en el lado del codificador;
25 la figura 4a ilustra una implementación de la calculadora del decodificador de la figura 1;
la figura 4b ilustra una implementación alternativa de la calculadora del decodificador en forma de una calculadora matriz;
30 la figura 4c ilustra una regla de combinación inversa central/lateral correspondiente a la regla de combinación ilustrada en la figura 3c;
la figura 5a ilustra una realización de un codificador de audio que opera en el dominio de frecuencia, que es preferiblemente un dominio de frecuencia de valor real;
35 la figura 5b ilustra una implementación de un decodificador de audio que opera en el dominio de frecuencia;
la figura 6a ilustra una implementación alternativa de un codificador de audio que opera en el dominio de MDCT y utiliza una transformada real-a-imaginaria;
40 la figura 6b ilustra una realización de un decodificador de audio que opera en el dominio de MDCT y utiliza una transformada real-a-imaginaria;
la figura 7a ilustra un post-procesador de audio que utiliza un decodificador estéreo y un decodificador de SBR conectado posteriormente;
45 la figura 7b ilustra una matriz de mezcla ascendente central/lateral;
la figura 8a ilustra una vista detallada en el bloque de MDCT en la figura 6a;
50 la figura 8b ilustra una vista detallada en el bloque de MDCT¹ de la figura 6b;
la figura 9a ilustra una implementación de un optimizador que opera en resolución reducida con respecto a la salida de MDCT;
55 la figura 9b ilustra una representación de un espectro de MDCT y las bandas correspondientes de resolución inferior, en que se calcula la información de predicción;
la figura 10a ilustra una implementación del transformador real-a-imaginario en la figura 6a o en la figura 6b; y
60 la figura 10b ilustra una implementación posible de la calculadora de espectro imaginario de la figura 10a.

La figura 1 ilustra un decodificador de audio para decodificar una señal de audio multicanal codificada y obtenida en una línea 100 de entrada. La señal de audio multicanal codificada comprende una primera señal de combinación

codificada generada utilizando una regla de combinación para combinar una primera señal de canal y una segunda señal de canal que representan la señal de audio multicanal, una señal residual de predicción codificada e información de predicción. La señal multicanal codificada puede ser una corriente de datos tal como una corriente de bits, que tiene las tres componentes en una forma multiplexada. Se puede incluir información lateral adicional en la señal multicanal codificada en la línea 100. La señal se introduce en una interfaz 102 de entrada. La interfaz 102 de entrada se puede implementar como un desmultiplexador de corriente de datos, que emite la primera señal de combinación codificada en la línea 104, la señal residual codificada en la línea 106 y la información de predicción en la línea 108. Preferiblemente, la información de predicción es un factor que tiene una parte real que no es igual a cero y/o una parte imaginaria distinta de cero. La señal de combinación codificada y la señal residual codificada se introducen en un decodificador 110 de señales para decodificar la primera señal de combinación para obtener una primera señal de combinación decodificada en la línea 112. Adicionalmente, el decodificador 110 de señales está configurado para decodificar la señal residual codificada para obtener una señal residual decodificada en la línea 114. Dependiendo del procesamiento de codificación en un lado del codificador de audio, el decodificador de señales puede comprender un decodificador de entropía tal como un decodificador Huffman, un decodificador aritmético o cualquier otro decodificador de entropía y una etapa de descuantificación conectada posteriormente para llevar a cabo una operación de descuantificación que coincida con una operación del cuantificador en un codificador de audio asociado. Las señales en la línea 112 y 114 se introducen en una calculadora 115 del decodificador, que emite la primera señal de canal en la línea 117 y una segunda señal de canal en la línea 118, en la que estas dos señales son señales estéreo o dos canales de una señal de audio multicanal. Cuando, por ejemplo, la señal de audio multicanal comprende cinco canales, entonces las dos señales son dos canales de la señal multicanal. Con el fin de codificar totalmente dicha señal multicanal que tiene cinco canales, pueden aplicarse dos decodificadores ilustrados en la figura 1, en los que el primer decodificador procesa el canal izquierdo y el canal derecho, el segundo decodificador procesa el canal del entorno izquierdo y el canal del entorno derecho, y un tercer decodificador mono se utilizaría para llevar a cabo una codificación mono del canal central. Sin embargo, también pueden aplicarse otros agrupamientos o combinaciones de codificadores de forma de onda y codificadores paramétricos. Una forma alternativa de generalizar el esquema de predicción a más de dos canales sería tratar tres señales (o más) al mismo tiempo, es decir, predecir una 3ª señal de combinación a partir de señales 1ª y 2ª utilizando dos coeficientes de predicción, en forma muy similar al módulo "dos-a-tres" en el entorno de MPEG.

La calculadora 116 del decodificador está configurada para calcular una señal multicanal decodificada que tiene la primera señal 117 de canal decodificada y la segunda señal 118 de canal decodificada utilizando la señal 114 residual decodificada, la información 108 de predicción y la primera señal 112 de combinación decodificada. Particularmente, la calculadora 116 del decodificador está configurada para operar de tal modo que la primera señal de canal decodificada y la segunda señal de canal decodificada son por lo menos una aproximación de una primera señal de canal y una segunda señal de canal de la señal multicanal introducidas en un codificador correspondiente, que se combinan mediante la regla de combinación cuando se genera la primera señal de combinación y la señal residual de predicción. Específicamente, la información 108 de predicción en la línea comprende una parte de valor real distinta de cero y/o una parte imaginaria distinta de cero.

La calculadora 116 del decodificador se puede implementar de diversos modos. Una primera implementación está ilustrada en la figura 4a. Esta implementación comprende un predictor 1160, una calculadora 1161 de señales de combinación y un combinador 1162. El predictor recibe la primera señal 112 de combinación decodificada y la información 108 de predicción y emite una señal 1163 de predicción. Específicamente, el predictor 1160 está configurado para aplicar la información 108 de predicción a la primera señal 112 de combinación decodificada o una señal procedente de la primera señal de combinación decodificada. La regla de derivación para derivar la señal a la que se aplica la información 108 de predicción puede ser una transformada real-a-imaginaria, o del mismo modo, una transformada imaginaria-a-real o una operación de ponderación o, dependiendo de la implementación, una operación de desfasaje o una operación combinada de ponderación/desfasaje. La señal 1163 de predicción se introduce junto con la señal residual decodificada en la calculadora 1161 de señales de combinación con el fin de calcular la segunda señal 1165 de combinación decodificada. Ambas señales 112 y 1165 se introducen en el combinador 1162, que combina la primera señal de combinación decodificada y la segunda señal de combinación para obtener la señal de audio multicanal decodificada que tiene la primera señal de canal decodificada y la segunda señal de canal decodificada en las líneas 1166 y 1167 de salida, respectivamente. Alternativamente, la calculadora del decodificador se implementa como una calculadora 1168 matriz, que recibe, como entrada, la primera señal de combinación decodificada o señal M, la señal residual decodificada o señal D y la información 108 de predicción α . La calculadora 1168 matriz aplica una matriz de transformada ilustrada como 1169 a las señales M, D para obtener las señales de salida L, R, en la que L es la primera señal de canal decodificada y R es la segunda señal de canal decodificada. La notación en la figura 4b se asemeja a una notación estéreo con un canal izquierdo L y un canal derecho R. Se ha aplicado esta notación con el fin de proporcionar una comprensión más sencilla si bien queda claro para los expertos en la técnica que las señales L, R pueden ser cualquier combinación de dos señales de canal en una señal multicanal que tiene más de dos señales de canal. La operación 1169 matriz unifica las operaciones en los bloques 1160, 1161 y 1162 de la figura 4a en un tipo de cálculo matriz de "disparo único", y las entradas en el circuito de la figura 4a y las salidas desde el circuito de la figura 4a son idénticas a las entradas en la calculadora 1168 matriz o a las salidas desde la calculadora 1168 matriz.

- La figura 4c ilustra un ejemplo para una regla de combinación inversa aplicada por el combinador 1162 en la figura 4a. Particularmente, la regla de combinación es similar a la regla de combinación del lado del decodificador en la codificación central/lateral que se conoce bien, siendo $L = M + S$, y $R = M - S$. Ha de entenderse que la señal S utilizada por la regla de combinación inversa en la figura 4c es la señal calculada por la calculadora de señales de combinación, es decir, la combinación de la señal de predicción en la línea 1163 y la señal residual decodificada en la línea 114. Ha de entenderse que, en la presente memoria descriptiva, las señales en las líneas a veces están designadas por los números de referencia para las líneas o a veces están indicadas por los propios números de referencia, que se han asignado a las líneas. Por lo tanto, la notación es de manera que una línea que tiene una determinada señal está indicando la señal propiamente dicha. Una línea puede ser una línea física en una implementación conectada mediante cables. En una implementación computarizada, sin embargo, una línea física no existe, sino que la señal representada por la línea es transmitida desde un módulo de cálculo hacia el otro módulo de cálculo.
- La figura 2 ilustra un codificador de audio para codificar una señal 200 de audio multicanal que tiene dos o más señales de canal, en la que una primera señal de canal está ilustrada en 201 y un segundo canal está ilustrado en 202. Ambas señales se introducen en una calculadora 203 del codificador para calcular una primera señal 204 de combinación y una señal 205 residual de predicción utilizando la primera señal 201 de canal y la segunda señal 202 de canal y la información 206 de predicción, de modo que cuando se combina la señal 205 residual de predicción con una señal de predicción procedente de la primera señal 204 de combinación y la información 206 de predicción se obtiene una segunda señal de combinación, en la que la primera señal de combinación y la segunda señal de combinación pueden obtenerse de la primera señal 201 de canal y la segunda señal 202 de canal utilizando una regla de combinación.
- La información de predicción es generada por un optimizador 207 para calcular la información 206 de predicción de modo que la señal residual de predicción cumple un objetivo 208 de optimización. La primera señal 204 de combinación y la señal 205 residual se introducen en un codificador 209 de señales para codificar la primera señal 204 de combinación para obtener una primera señal 210 de combinación codificada y para codificar la señal 205 residual para obtener una señal 211 residual codificada. Ambas señales 210, 211 codificadas se introducen en una interfaz 212 de salida para combinar la primera señal 210 de combinación codificada con la señal 211 residual de predicción codificada y la información 206 de predicción para obtener una señal 213 multicanal codificada, que es similar a la señal 100 multicanal codificada introducida en la interfaz 102 de entrada del decodificador de audio ilustrado en la figura 1.
- Dependiendo de la implementación, el optimizador 207 recibe, o bien la primera señal 201 de canal y la segunda señal 202 de canal o bien, tal como se ilustra en las líneas 214 y 215, la primera señal 214 de combinación y la segunda señal 215 de combinación procedentes de un combinador 2031 de la figura 3a, que se analizarán posteriormente.
- Un objetivo preferido de optimización se ilustra en la figura 2, en la que la ganancia de codificación está potenciada al máximo, es decir, la tasa de bits se reduce tanto como sea posible. En este objetivo de optimización, la señal residual D está reducida al mínimo con respecto a α . Esto significa, en otras palabras, que la información de predicción α está seleccionado de tal modo que $\|S - \alpha M\|^2$ está reducida al mínimo. Esto da como resultado una solución para α que se ilustra la figura 2. Las señales S, M están proporcionadas en forma de bloque y son preferiblemente señales espectrales de dominio, en las que la notación $\|\dots\|$ significa la norma 2 del argumento, y en las que $\langle \dots \rangle$ ilustra el producto de puntos en forma habitual. Cuando la primera señal 201 de canal y la segunda señal 202 de canal se introducen en el optimizador 207, el optimizador tendría entonces que aplicar la regla de combinación, en la que una regla de combinación ejemplar está ilustrada en la figura 3c. Sin embargo, cuando la primera señal 214 de combinación y la segunda señal 215 de combinación se introducen en el optimizador 207, entonces el optimizador 207 no necesita implementar la regla de combinación por sí mismo.
- Otros objetivos de optimización pueden estar relacionados con la calidad perceptual. Un objetivo de optimización puede contemplar que se obtenga una calidad perceptual máxima. Entonces, el optimizador necesitaría información adicional procedente de un modelo perceptual. Otras implementaciones del objetivo de optimización pueden referirse a la obtención de una mínima tasa de bits, o una tasa fija de bits. Entonces, el optimizador 207 se implementaría para llevar a cabo una operación de cuantificación/codificación de entropía con el fin de determinar la tasa de bits necesaria para determinados valores de α de modo que α se puede fijar para cumplir los requerimientos tales como una tasa mínima de bits, o alternativamente, una tasa fija de bits. Otras implementaciones del objetivo de optimización pueden estar relacionadas con un uso mínimo de los recursos del codificador o del decodificador. En el caso de una implementación de tal objetivo de optimización, la información en los recursos necesarios estaría disponible para una determinada optimización en el optimizador 207. Adicionalmente, una combinación de estos objetivos de optimización u otros objetivos de optimización pueden aplicarse para controlar el optimizador 207, que calcula la información 206 de predicción.

La calculadora 203 del codificador en la figura 2 se puede implementar en distintas formas, en las que una primera implementación a modo de ejemplo está ilustrada en la figura 3a, en la que una regla de combinación explícita se lleva a cabo en el combinador 2031. Una implementación alternativa a modo de ejemplo está ilustrada en la figura 3b, en la que se utiliza una calculadora 2039 matriz. El combinador 2031 en la figura 3a se puede implementar para llevar a cabo la regla de combinación ilustrada en la figura 3c, que es de manera ejemplar la regla de codificación central/lateral que se conoce bien, en la que un factor de ponderación de 0,5 se aplica a todos los ramales. Sin embargo, pueden implementarse otros factores de ponderación o ningún factor de ponderación dependiendo de la implementación. Adicionalmente, ha de observarse que pueden aplicarse otras reglas de combinación tales como otras reglas de combinación lineales o reglas de combinación que no son lineales, siempre que haya una regla de combinación inversa correspondiente, que puede aplicarse en el combinador 1162 del decodificador ilustrado en la figura 4a, que aplica una regla de combinación que es inversa a la regla de combinación aplicada por el codificador. Debido a la predicción inventiva puede utilizarse cualquier regla de predicción invertible, dado que la influencia en la forma de onda está “equilibrada” por la predicción, es decir, cualquier error está incluido en la señal residual transmitida, dado que la operación de predicción llevada a cabo por el optimizador 207 en combinación con la calculadora 203 del codificador es un proceso que conserva la forma de onda.

El combinador 2031 emite la primera señal 204 de combinación y una segunda señal 2032 de combinación. La primera señal de combinación se introduce en un predictor 2033, y la segunda señal 2032 de combinación se introduce en la calculadora 2034 residual. El predictor 2033 calcula una señal 2035 de predicción, que se combina con la segunda señal 2032 de combinación para obtener finalmente la señal 205 residual. Particularmente, el combinador 2031 está configurado para combinar las dos señales 201 y 202 de canal de la señal de audio multicanal en dos formas distintas para obtener la primera señal 204 de combinación y la segunda señal 2032 de combinación, en la que las dos formas distintas están ilustradas en la figura 3c. El predictor 2033 está configurado para aplicar la información de predicción a la primera señal 204 de combinación o una señal procedente de la primera señal de combinación para obtener la señal 2035 de predicción. La señal procedente de la señal de combinación puede provenir de cualquier operación no lineal o lineal, en la que se prefiere una transformada real-a-imaginaria/transformada imaginaria-a-real, que puede implementarse utilizando un filtro lineal tal como un filtro FIR que lleva a cabo adiciones ponderadas de determinados valores.

La calculadora 2034 residual en la figura 3a puede llevar a cabo una operación de sustracción de modo que la señal de predicción se sustrae de la segunda señal de combinación. Sin embargo, son posibles otras operaciones en la calculadora residual. De la misma manera, la calculadora 1161 de señales de combinación en la figura 4a puede llevar a cabo una operación de adición en la que la señal 114 residual decodificada y la señal 1163 de predicción se agregan juntas para obtener la segunda señal 1165 de combinación.

La figura 5a ilustra una implementación preferida de un codificador de audio. En comparación con el codificador de audio ilustrado en la figura 3a, la primera señal 201 de canal es una representación espectral de una primera señal 55a de canal de dominio de tiempo. De la misma manera, la segunda señal 202 de canal es una representación espectral de una señal 55b de canal de dominio de tiempo. La conversión del dominio de tiempo en la representación espectral es llevada a cabo por un convertidor 50 de tiempo/frecuencia para la primera señal de canal y un convertidor 51 de tiempo/frecuencia para la segunda señal de canal. Preferiblemente, pero no necesariamente, los convertidores 50, 51 espectrales están implementados como convertidores de valor real. El algoritmo de conversión puede ser una transformada coseno discreta, una transformada FFT, en la que solo se utiliza la parte real, una MDCT o cualquier otra transformada que proporcione valores espectrales de valor real. Alternativamente, ambas transformadas se pueden implementar como una transformada imaginaria, tal como una DST, una MDST o una FFT en las que solo se utiliza la parte imaginaria y se descarta la parte real. También puede utilizarse cualquier otra transformada que proporcione solo valores imaginarios. Un fin por el que utilizar una transformada pura de valor real o una transformada pura imaginaria es la complejidad computacional, dado que, para cada valor espectral, solo se tiene que procesar un único valor tal como magnitud o la parte real o, alternativamente, la fase o la parte imaginaria. A diferencia de una transformada totalmente compleja tal como una FFT, tendrían que procesarse dos valores, es decir, la parte real y la parte imaginaria para cada línea espectral, lo que supone un aumento de la complejidad computacional por un factor de al menos 2. En este caso, otro motivo para utilizar una transformada de valor real es que una transformada de este tipo generalmente se muestrea críticamente, y de este modo proporciona un dominio adecuado (y comúnmente utilizado) para la cuantificación de señales y la codificación de entropía (el paradigma habitual de “codificación de audio perceptual” implementado en “MP3”, AAC, o en sistemas de codificación de audio similares).

La figura 5a ilustra adicionalmente la calculadora 2034 residual como un sumador, que recibe la señal lateral como su entrada “más” y que recibe la salida de la señal de predicción por el predictor 2033 como su entrada “menos”. Adicionalmente, la figura 5a ilustra la situación en la que la información de control del predictor se transmite desde el optimizador hacia el multiplexador 212, que emite una corriente de bits multiplexada que representa la señal de audio multicanal codificada. Particularmente, la operación de predicción se lleva a cabo de tal modo que la señal lateral puede predecirse desde la señal central tal como se ilustra en las ecuaciones a la derecha de la figura 5a.

Preferiblemente, la información 206 de control del predictor es un factor tal como se ilustra a la derecha en la figura 3b. En un ejemplo en el que la información de control de predicción solo comprende una porción real tal como la parte real de una α de valor complejo o una magnitud de α de valor complejo, en la que esta porción corresponde a un factor distinto de cero, puede obtenerse una ganancia de codificación significativa cuando la señal central y la señal lateral son similares entre sí debido a su estructura de forma de onda, si bien tienen distintas amplitudes.

Sin embargo, cuando la información de control de predicción solo comprende una segunda porción, que puede ser la parte imaginaria de un factor de valor complejo o la información de fase del factor de valor complejo, en la que la parte imaginaria o la información de fase es distinta de cero, la presente invención logra una ganancia de codificación significativa para las señales, que están desfasadas una con respecto a otra por un valor distinto de 0° o 180° , y que tienen además del desfasaje, características de forma de onda similares y relaciones de amplitud similares.

Preferiblemente, una información de control de predicción es de valor complejo y, según la invención, comprende una porción imaginaria distinta de cero. Entonces, puede obtenerse una ganancia de codificación significativa para las señales que tienen distinta amplitud y que están desfasadas. En una situación en la que las transformadas de tiempo/frecuencia proporcionan espectros complejos, la operación 2034 sería una operación compleja en la que la parte real de la información de control del predictor se aplica a la parte real del espectro complejo M y la parte imaginaria de la información de predicción compleja se aplica a la parte imaginaria del espectro complejo. Entonces, en el sumador 2034, el resultado de esta operación de predicción es un espectro real previsible y un espectro imaginario previsible, y el espectro real previsible se sustraería del espectro real de la señal lateral S (en cuanto a la banda), y el espectro imaginario previsible se sustraería de la parte imaginaria del espectro de S para obtener un espectro residual complejo D.

Las señales de dominio de tiempo L y R son señales de valor real, si bien las señales de dominio de frecuencia pueden ser de valor real o complejo. Cuando las señales de dominio de frecuencia son de valor real, entonces la transformada es una transformada de valor real. Cuando las señales de dominio de frecuencia son complejas, entonces la transformada es una transformada de valor complejo. Esto significa que la entrada a las transformadas de tiempo-a-frecuencia y la salida de las transformadas de frecuencia-a-tiempo son de valor real, en tanto que las señales de dominio de frecuencia podrían ser, por ejemplo, señales de dominio QMF de valor complejo.

La figura 5b ilustra un decodificador de audio correspondiente al codificador de audio ilustrado en la figura 5a. Los elementos similares con respecto al decodificador de audio de la figura 1 tienen números de referencia similares.

La salida de la corriente de bits por el multiplexador 212 de la corriente de bits en la figura 5a se introduce en un desmultiplexador 102 de la corriente de bits en la figura 5b. El desmultiplexador 102 de la corriente de bits desmultiplexa la corriente de bits en la señal de mezcla descendente M y la señal residual D. La señal de mezcla descendente M se introduce en un descuantificador 110a. La señal residual D se introduce en un descuantificador 110b. Adicionalmente, el desmultiplexador 102 de la corriente de bits desmultiplexa una información 108 de control del predictor desde la corriente de bits e introduce la misma en el predictor 1160. El predictor 1160 emite una señal lateral previsible α M y el combinador 1161 combina la señal residual emitida por el descuantificador 110b con la señal lateral previsible con el fin de obtener finalmente la señal lateral reconstruida S. La señal se introduce entonces en el combinador 1162, que lleva a cabo, por ejemplo, un procesamiento de suma/diferencia, tal como se ilustra en la figura 4c con respecto a la codificación central/lateral. Particularmente, el bloque 1162 lleva a cabo una decodificación central/lateral (inversa) para obtener una representación de dominio de frecuencia del canal izquierdo y una representación de dominio de frecuencia del canal derecho. La representación del dominio de frecuencia se convierte entonces en una representación del dominio de tiempo por los convertidores 52 y 53 de frecuencia/tiempo correspondientes.

Dependiendo de la implementación del sistema, los convertidores 52, 53 de frecuencia/tiempo son convertidores de valor real de frecuencia/tiempo cuando la representación del dominio de frecuencia es una representación de valor real, o convertidores de frecuencia/tiempo de valor complejo cuando la representación del dominio de frecuencia es una representación de valor complejo.

Sin embargo, para aumentar la eficiencia se prefiere llevar a cabo una transformada de valor real tal como se ilustra en otra implementación en la figura 6a para el codificador y en la figura 6b para el decodificador. Las transformadas 50 y 51 de valor real están implementadas por una MDCT. Adicionalmente, la información de predicción se calcula como un valor complejo que tiene una parte real y una parte imaginaria. Dado que ambos espectros M, S son espectros de valor real y dado que, por lo tanto, no existe ninguna parte imaginaria del espectro, se proporciona un convertidor 2070 real-a-imaginario, que calcula un espectro 600 imaginario estimado a partir del espectro de valor real de la señal M. Este transformador 2070 real-a-imaginario es una parte del optimizador 207, y el espectro 600 imaginario estimado por el bloque 2070 se introduce en la etapa 2071 del optimizador de α junto con el espectro real M con el fin de calcular la información 206 de predicción, que ahora tiene un factor de valor real indicado en 2073 y un factor imaginario indicado en 2074. Ahora, según esta realización, el espectro de valor real de la primera señal de

combinación M se multiplica por la parte 2073 real α_R para obtener la señal de predicción, que entonces se sustrae del espectro lateral de valor real. Adicionalmente, el espectro 600 imaginario se multiplica por la parte imaginaria α_I ilustrada en 2074 para obtener la señal de predicción adicional, en la que esta señal de predicción se sustrae entonces del espectro lateral de valor real tal como se indica en 2034b. Entonces, la señal residual de predicción D se cuantifica en el cuantificador 209b, mientras que el espectro de valor real de M se cuantifica/codifica en el bloque 209a. Adicionalmente, se prefiere cuantificar y codificar la información de predicción α en el cuantificador/codificador 2072 de entropía para obtener el valor de α complejo codificado, que se transmite al multiplexador 212 de la corriente de bits de la figura 5a, por ejemplo, y que se introduce finalmente en una corriente de bits tal como la información de predicción.

Con respecto a la posición del módulo 2072 de cuantificación/codificación (Q/C) para α , se observa que los multiplicadores 2073 y 2074 preferiblemente utilizan exactamente la misma α (cuantificada) que también se utilizará en el decodificador. De este modo se podría mover 2072 directamente hacia la salida de 2071, o podría considerarse que la cuantificación de α ya se ha tenido en cuenta en el proceso de optimización en 2071.

A pesar de que podría calcularse un espectro complejo en el lado del codificador debido a que toda la información está disponible, se prefiere realizar la transformada real-a-compleja en el bloque 2070 en el codificador de modo que se produzcan condiciones similares con respecto a un decodificador ilustrado en la figura 6b. El decodificador recibe un espectro codificado de valor real de la primera señal de combinación y una representación espectral de valor real de la señal residual codificada. Adicionalmente se obtiene una información de predicción compleja codificada en 108 y se lleva a cabo una decodificación de entropía y una descuantificación en el bloque 65 para obtener la parte real α_R ilustrada en 1160b y la parte imaginaria α_I ilustrada en 1160c. Las señales centrales emitidas por los elementos 1160b y 1160c de ponderación se agregan a la señal residual de predicción decodificada y descuantificada. Particularmente, los valores espectrales introducidos en el ponderador 1160c, en el que la parte imaginaria del factor de predicción compleja se utiliza como el factor de ponderación, provienen del espectro de valor real M por el convertidor 1160a real-a-imaginario, que se implementa preferiblemente del mismo modo que el bloque 2070 de la figura 6a relacionada con el lado del codificador. En el lado del decodificador no hay una representación de valor complejo de la señal central o de la señal lateral, lo cual implica una diferencia con respecto al lado del codificador. El motivo es que solo los espectros de valor real codificados han sido transmitidos desde el codificador hacia el decodificador debido a las tasas de bits y por motivos de complejidad.

El transformador 1160a real-a-imaginario o el bloque 2070 correspondiente de la figura 6a puede implementarse tal como se publica en el documento WO 2004/013839 A1 o en el documento WO 2008/014853 A1 o en la patente estadounidense n.º 6.980.933. Alternativamente puede aplicarse cualquier otra implementación conocida en la técnica, y una implementación preferida se describe en el contexto de las figuras 10a, 10b.

Específicamente, tal como se ilustra en la figura 10a, el convertidor 1160a real-a-imaginario comprende un selector 1000 de cuadro espectral conectado a una calculadora 1001 del espectro imaginario. El selector 1000 de cuadro espectral recibe una indicación de un cuadro actual i en la entrada 1002 y, dependiendo de la implementación, información de control en una entrada 1003 de control. Cuando, por ejemplo, la indicación en la línea 1002 indica que se tiene que calcular un espectro imaginario para un cuadro actual i, y cuando la información 1003 de control indica que solo se tiene que utilizar el cuadro actual para dicho cálculo, entonces el selector 1000 de cuadro espectral solo selecciona el cuadro actual i y transmite esta información a la calculadora de espectro imaginario. Entonces, la calculadora de espectro imaginario solo utiliza las líneas espectrales del cuadro actual i para llevar a cabo una combinación ponderada de las líneas ubicadas en el cuadro actual (bloque 1008), con respecto a la frecuencia, próxima a, o alrededor de la línea espectral actual k, para la que se tiene que calcular una línea imaginaria tal como se ilustra en 1004 en la figura 10b. Sin embargo, cuando el selector 1000 de cuadro espectral recibe una información 1003 de control que indica que el cuadro anterior i-1 y el cuadro siguiente i+1 se tienen que utilizar también para el cálculo del espectro imaginario, entonces la calculadora de espectro imaginario adicionalmente recibe los valores procedentes de los cuadros i-1 y i+1 y lleva a cabo una combinación ponderada de las líneas en los cuadros correspondientes tal como se ilustra en 1005 para el cuadro i-1 y en 1006 para el cuadro i+1. Los resultados de las operaciones de ponderación se combinan mediante una combinación ponderada en el bloque 1007 para obtener finalmente una línea imaginaria k para el cuadro f_i que luego se multiplica por la parte imaginaria de la información de predicción en el elemento 1160c para obtener la señal de predicción para esta línea, que entonces se agrega a la línea correspondiente de la señal central en el sumador 1161b para el decodificador. En el codificador se lleva a cabo la misma operación, pero se realiza una sustracción en el elemento 2034b.

Ha de observarse que la información 1003 de control puede indicar adicionalmente el uso de más cuadros que los dos cuadros circundantes o, por ejemplo, el uso del cuadro actual y exactamente uno o más cuadros anteriores, pero sin utilizar cuadros "futuros" con el fin de reducir el retardo sistemático.

Adicionalmente, ha de observarse que la combinación ponderada en cuanto a etapas ilustrada en la figura 10b, en la que, en una primera operación se combinan las líneas procedentes de un cuadro y posteriormente se combinan por sí mismos los resultados de dichas operaciones de combinación en cuadros, también se puede llevar a cabo en el

otro orden. El otro orden se refiere a que, en un primer paso, las líneas para la frecuencia actual k procedentes de varios cuadros adyacentes indicados por la información 103 de control se combinan mediante una combinación ponderada. Esta combinación ponderada se lleva a cabo para las líneas k , $k-1$, $k-2$, $k+1$, $k+2$, etc., dependiendo del número de líneas adyacentes que van a usarse para estimar la línea imaginaria. Entonces, los resultados de estas combinaciones “en cuanto a tiempo” se someten a una combinación ponderada en la “dirección de frecuencia” para obtener finalmente la línea imaginaria k para el cuadro i . Los valores ponderados se fijan en un valor entre -1 y 1 preferiblemente, y las ponderaciones pueden implementarse en una combinación de filtro FIR o IIR directa, que lleva a cabo una combinación lineal de líneas espectrales o de señales espectrales de distintas frecuencias y distintos cuadros.

Tal como se indica en las figuras 6a y 6b, el algoritmo de transformada preferido es el algoritmo de transformada MDCT, que se aplica en la dirección hacia adelante en los elementos 50 y 51 en la figura 6a y que se aplica en la dirección hacia atrás en los elementos 52, 53, después de una operación de combinación en el combinador 1162 que opera en el dominio espectral.

La figura 8a ilustra una implementación más detallada del bloque 50 o 51. Particularmente se introduce una secuencia de muestras de audio del dominio de tiempo en un sistema 500 de ventanas de análisis, que lleva a cabo una operación de ventanas que utiliza una ventana de análisis y , particularmente, lleva a cabo esta operación en un cuadro mediante de una manera de cuadro a cuadro, pero utilizando un paso o solapamiento del 50 %. El resultado del sistema de ventanas de análisis, es decir, una secuencia de cuadros de muestras divididas en ventanas se introduce en un bloque 501 de transformada MDCT, que emite la secuencia de cuadros de MDCT de valor real, en la que dichos cuadros están afectados por el solapamiento. A modo de ejemplo, el sistema de ventanas de análisis aplica ventanas de análisis que tienen una longitud de 2048 muestras. Entonces, el bloque 501 de transformada MDCT emite espectros de MDCT que tienen 1024 líneas espectrales reales o valores MDCT. Preferiblemente, el sistema 500 de ventanas de análisis y/o el transformador 501 de MDCT pueden controlarse mediante un control 502 de longitud de ventana o longitud de transformada de modo que, por ejemplo, para las porciones transitorias en la señal, la longitud de ventana/longitud de transformada se reduce con el fin de obtener mejores resultados de codificación.

La figura 8b ilustra la operación de MDCT inversa llevada a cabo en los bloques 52 y 53. A modo de ejemplo, el bloque 52 comprende un bloque 520 para llevar a cabo una transformada de MDCT inversa de cuadro-por-cuadro. Cuando, por ejemplo, un cuadro de valores MDCT tiene 1024 valores, entonces la salida de esta transformada inversa de MDCT tiene 2048 muestras de tiempo afectadas por el solapamiento. Tal cuadro se suministra a un sistema 521 de ventanas de síntesis, que aplica una venta de síntesis a este cuadro de 2048 muestras. El cuadro dividido en ventanas se transmite entonces a un procesador 522 de solapar/agregar, que, a modo de ejemplo, aplica un 50% de solapamiento entre dos cuadros posteriores y , entonces, lleva a cabo una adición de muestra por muestra de modo que un bloque de 2048 muestras finalmente da como resultado 1024 muestras nuevas de la señal de salida libre de solapamiento. Nuevamente se prefiere aplicar un control de longitud de ventana/transformada utilizando información, que por ejemplo se transmite en la información lateral de la señal multicanal codificada según lo indicado en 523.

Los valores de predicción de α se podrían calcular para cada línea espectral individual de un espectro de MDCT. Sin embargo, se ha descubierto que esto no es necesario y que se puede ahorrar una cantidad significativa de información lateral llevando a cabo un cálculo de bandas de la información de predicción. En otros términos, un convertidor 50 espectral ilustrado en la figura 9, que es por ejemplo un procesador de MDCT descrito en el contexto de la figura 8a, proporciona un espectro de resolución de alta frecuencia que tiene determinadas líneas espectrales ilustradas en la figura 9b. Este espectro de resolución de alta frecuencia se utiliza por un selector 90 de línea espectral que proporciona un espectro de resolución de baja frecuencia, que comprende determinadas bandas B1, B2, B3, ... , BN. Este espectro de resolución de baja frecuencia se transmite al optimizador 207 para calcular la información de predicción de modo que no se calcula una información de predicción para cada línea espectral, sino solamente para cada banda. Para tal fin, el optimizador 207 recibe las líneas espectrales por banda y calcula la operación de optimización basándose en la presunción de que se utiliza el mismo valor α para todas las líneas espectrales en la banda.

Preferiblemente, las bandas están conformadas de manera psicoacústica de modo que el ancho de banda de las bandas aumenta desde frecuencias inferiores hasta frecuencias superiores tal como se ilustra en la figura 9b. Alternativamente, a pesar de que no se prefiere tanto como la implementación de ancho de banda en aumento también podrían utilizarse bandas de frecuencia de igual tamaño, en las cada banda de frecuencia tiene al menos dos o típicamente muchas más, tal como por lo menos 30, líneas de frecuencia. Típicamente, para un espectro de 1024 líneas espectrales, se calculan menos de 30 valores de α complejos, y preferiblemente, más de 5 valores de α . Para los espectros con menos de 1024 líneas espectrales (por ejemplo, 128 líneas) se utilizan preferiblemente menos bandas de frecuencia (por ejemplo, 6) para α .

Para calcular los valores de α no se requiere necesariamente el espectro de MDCT de alta resolución.

Alternativamente también puede utilizarse un banco de filtros que tiene una resolución de frecuencia similar a la resolución necesaria para calcular los valores de α . Cuando van a implementarse bandas que aumentan de frecuencia, entonces este banco de filtros debería tener un ancho de banda variable. Sin embargo, cuando es suficiente un ancho de banda constante de frecuencias bajas a altas, entonces puede utilizarse un banco de filtros tradicional con sub-bandas de ancho equivalente.

Dependiendo de la implementación, el signo del valor de α indicado en la figura 3b o 4b puede invertirse. Sin embargo, con el fin de ser coherentes, es necesario que esta reversión del signo se utilice en el lado del codificador, así como en el lado del decodificador. A diferencia de la figura 6a, la figura 5a ilustra una vista general del codificador, en la que el elemento 2033 es un predictor que está controlado por la información 206 de control del predictor, que se determina en el elemento 207 y que se incorpora como información lateral en la corriente de bits. En lugar de la MDCT utilizada en la figura 6a en los bloques 50, 51 se utiliza una transformada generalizada de tiempo/frecuencia en la figura 5a tal como se describe. Según se indicó anteriormente, la figura 6a es el proceso del codificador, que corresponde al proceso del decodificador en la figura 6b, en la que L representa la señal izquierda de canal, R representa la señal derecha de canal, M representa la señal central o la señal de mezcla descendente, S representa la señal lateral y D representa la señal residual. Alternativamente, L también se conoce como la primera señal 201 de canal, R también se conoce como la segunda señal 202 de canal, M también se conoce como la primera señal 204 de combinación y S también se conoce como la segunda señal 2032 de combinación.

Preferiblemente, los módulos 2070 en el codificador y 1160a en el decodificador deberían coincidir exactamente con el fin de asegurar una correcta codificación de forma de onda. Esto se aplica preferiblemente al caso en el que estos módulos utilizan alguna forma de aproximación tal como filtros truncados o cuando se utiliza solamente uno o dos en lugar de los tres cuadros de MDCT, es decir, el cuadro de MDCT actual en la línea 60, el cuadro de MDCT anterior en la línea 61 y el cuadro de MDCT siguiente en la línea 62.

Adicionalmente, se prefiere que el módulo 2070 en el codificador en la figura 6a utilice el espectro de MDCT no cuantificado M como entrada, a pesar de que el módulo 1160a real-a-imaginario (R21) en el decodificador tiene solo el espectro de MDCT cuantificado disponible como entrada. Alternativamente puede utilizarse también una implementación en la que el codificador utiliza los coeficientes de MDCT cuantificados como entrada en el módulo 2070. Sin embargo, el uso del espectro de MDCT no cuantificado como entrada en el módulo 2070 es el método preferido desde un punto de vista perceptual.

A continuación, se describen más detalladamente diversos aspectos de la realización de la presente invención.

La codificación de estéreo paramétrico habitual se basa en la capacidad del dominio de QMF (híbrido) complejo sobremuestreado de permitir un procesamiento de señal perceptualmente motivado con variación de tiempo y frecuencia sin introducir distorsiones de solapamiento. Sin embargo, en el caso de la codificación de mezcla descendente/residual (como se utiliza para las altas tasas de bits consideradas en este caso), el codificador de estéreo unificado resultante actúa como un codificador de forma de onda. Esto permite el funcionamiento en un dominio críticamente muestreado, tal como el dominio de MDCT, dado que el paradigma de codificación de forma de onda garantiza que la propiedad de cancelación de solapamiento de la cadena de procesamiento de MDCT-IMDCT se conserve suficientemente bien.

Sin embargo, con el fin de poder aprovechar la eficiencia de codificación mejorada que se puede lograr en el caso de señales estéreo con diferencias de tiempo o fase entre canales por medio de un coeficiente de predicción de valor complejo α , se necesita una representación de valor complejo de dominio de frecuencia de la señal de mezcla descendente DMX como entrada en la matriz de mezcla ascendente de valor complejo. Esto puede obtenerse utilizando una transformada de MDST además de la transformada de MDCT para la señal DMX. El espectro de MDST se puede computar (en forma exacta o como una aproximación) a partir del espectro de MDCT.

Además, la parametrización de la matriz de mezcla ascendente puede simplificarse transmitiendo el coeficiente de predicción compleja α en lugar de los parámetros MPS. De este modo se transmiten solo dos parámetros (parte real e imaginaria de α) en lugar de tres (ICC, CLD y IPD). Esto resulta posible debido a la redundancia en la parametrización de MPS en el caso de la codificación de mezcla descendente/residual. La parametrización de MPS incluye información acerca de la cantidad relativa de descorrelación a agregar en el decodificador (es decir, la relación de energía entre las señales RES y DMX), y esta información es redundante cuando se transmiten las señales reales DMX y RES.

Debido a la misma razón, el factor de ganancia g , ilustrado en la matriz de mezcla ascendente anterior, es obsoleto en el caso de la codificación de mezcla descendente/residual. De este modo, la matriz de mezcla ascendente para la codificación de mezcla descendente/residual con predicción compleja es la siguiente:

$$\begin{bmatrix} L \\ R \end{bmatrix} = \begin{bmatrix} 1-\alpha & 1 \\ 1+\alpha & -1 \end{bmatrix} \begin{bmatrix} DMX \\ RES \end{bmatrix}$$

En comparación con la ecuación 1169 en la figura 4b, el signo de alfa está invertido en esta ecuación, y DMX=M y RES=D. Esta es, por lo tanto, una implementación/notación alternativa con respecto a la figura 4b.

5 Existen dos opciones para calcular la señal residual de predicción en el codificador. Una opción consiste en utilizar los valores espectrales de MDCT cuantificados de la mezcla descendente. Esto tendría como resultado la misma distribución de error de cuantificación que en la codificación M/S ya que el codificador y el decodificador utilizan los mismos valores para generar la predicción. La otra opción consiste en utilizar los valores espectrales de MDCT no cuantificados. Esto implica que el codificador y el decodificador no utilizarán los mismos datos para generar la predicción, lo que permite la redistribución espacial del error de codificación según las propiedades de enmascaramiento instantáneo de la señal a expensas de una ganancia de codificación un tanto reducida.

10 Es preferible computar el espectro de MDST directamente en el dominio de frecuencia por medio del filtrado bidimensional de FIR de tres cuadros de MDCT adyacentes tal como se ha analizado. Los últimos pueden considerarse una transformada "real-a-imaginaria" (R2I). La complejidad del cómputo de frecuencia-dominio de la MDST puede reducirse de diversas formas, lo que significa que se calcula solo una aproximación del espectro de MDST:

- 20 • Limitando el número de conexiones de filtro FIR.
- Estimando la MDST a partir del cuadro de MDCT real solamente.
- 25 • Estimando la MDST a partir del cuadro de MDCT previo y actual.

Siempre y cuando se utilice la misma aproximación en el codificador y en el decodificador, las propiedades de codificación de forma de onda no resultarán afectadas. Tales aproximaciones del espectro de MDST, sin embargo, pueden conllevar una reducción en la ganancia de codificación lograda por la predicción compleja.

30 En caso de que el codificador de MDCT subyacente soporte la conmutación de forma de ventana, los coeficientes del filtro bidimensional FIR utilizado para computar el espectro de MDST tienen que adaptarse a las formas de ventana reales. Los coeficientes de filtros aplicados al espectro de MDCT del cuadro actual dependen de la ventana completa, es decir, es necesario un conjunto de coeficientes para cada tipo de ventana y para cada transición de ventana. Los coeficientes de filtros aplicados al espectro de MDCT del cuadro anterior/siguiente dependen solo del solapamiento de la mitad de ventana con el cuadro actual, es decir, para estos es necesario un conjunto de coeficientes solo para cada tipo de ventana (sin coeficientes adicionales para las transiciones).

40 En caso de que el codificador de MDCT subyacente utilice una conmutación de longitud de transformada, incluyendo el cuadro de MDCT anterior y/o siguiente en la aproximación, esto resulta más complicado alrededor de las transiciones entre las distintas longitudes de transformadas. En este caso, debido al distinto número de coeficientes de MDCT en el cuadro actual y anterior/siguiente, el filtrado bidimensional resulta más complicado. Con el fin de evitar la creciente complejidad computacional y estructural, el cuadro anterior/siguiente se puede excluir del filtrado en las transiciones de longitud de transformada, a expensas de una menor exactitud de la aproximación para los cuadros respectivos.

45 Asimismo, se debe tener especial cuidado con las partes más bajas y más altas del espectro de MDST (próximas a DC y fs/2), en las hay menos coeficientes de MDCT circundantes para el filtrado de FIR que los necesarios. En este caso, el proceso de filtrado necesita adaptarse para computar el espectro de MDST correctamente. Esto puede llevarse a cabo, o bien utilizando una extensión simétrica del espectro de MDCT para los coeficientes que faltan (según la periodicidad de los espectros de señales discretas de tiempo), o bien adaptando los coeficientes de filtro de manera correspondiente. Obviamente, el tratamiento de estos casos especiales puede simplificarse a expensas de una menor exactitud en la proximidad de los límites del espectro de MDST.

50 El cómputo del espectro de MDST exacto desde los espectros de MDCT transmitidos en el decodificador aumenta el retardo del decodificador por un cuadro (aquí se presume que son 1024 muestras).

El retardo adicional puede evitarse utilizando una aproximación del espectro de MDST que no requiere el espectro de MDCT del cuadro siguiente como entrada.

60 La siguiente lista de elementos resume las ventajas de la codificación estéreo unificada basada en MDCT en comparación con la codificación estéreo unificada basada en QMF:

- Solo un pequeño aumento en la complejidad computacional (cuando no se utiliza SBR).
 - 5 • Amplía la capacidad hasta una reconstrucción perfecta si los espectros de MDCT no están cuantificados. Ha de observarse que esto no se produce para la codificación estéreo unificada basada en QMF.
 - Extensión natural de la codificación M/S e intensidad de la codificación estéreo.
 - 10 • Arquitectura más limpia que simplifica el ajuste del codificador, dado que el procesamiento de la señal estéreo y la cuantificación/codificación se pueden acoplar estrechamente. Ha de observarse que en la codificación estéreo unificada basada en QMF, los cuadros del entorno de MPEG y los cuadros de MDCT no están alineados y dichas bandas de factor de escala no coinciden con las bandas del parámetro.
 - 15 • Codificación de parámetros de estéreo eficaz, dado que solo tienen que transmitirse dos parámetros (complejo α) en lugar de tres parámetros como en el entorno de MPEG (ICC, CLD, IPD).
 - No presenta un retardo de decodificador adicional si el espectro de MDST se computa como una aproximación (sin utilizar el cuadro siguiente).
- 20 Las propiedades importantes de una implementación pueden resumirse del siguiente modo:
- a) Los espectros de MDST se computan por medio del filtrado bidimensional de FIR a partir de los espectros de MDCT actuales, anteriores y siguientes. Se posibilitan distintas compensaciones de complejidad/calidad para el cómputo de MDST (aproximación) reduciendo el número de conexiones de filtro FIR y/o el número de cuadros de MDCT utilizados. En particular, si un cuadro adyacente no está disponible debido a la pérdida de cuadro durante la transmisión o la conmutación de longitud de transformada, ese cuadro particular se excluye del valor estimado de MDST. Para el caso de la conmutación de longitud de transformada, la exclusión se señala en la corriente de bits.
- 25 b) Solo dos parámetros, la parte real e imaginaria del coeficiente de predicción compleja α , se transmiten en lugar de ICC, CLD, e IPD. Las partes reales e imaginarias de α se tratan de forma independiente, limitadas al intervalo [-3,0, 3,0] y se cuantifican con un tamaño de paso de 0,1. Si un determinado parámetro (parte real o imaginaria de α) no se utiliza en un determinado cuadro, esto se señala en la corriente de bits, y el parámetro irrelevante no se transmite. Los parámetros se codifican de forma diferencial al tiempo o de forma diferencial a la frecuencia y finalmente se aplica la codificación de Huffman utilizando el código de factor de escala. Los coeficientes de predicción se actualizan cada dos bandas de factor de escala, lo que da como resultado una resolución de frecuencia similar a la resolución de frecuencia del entorno de MPEG. Este esquema de cuantificación y codificación da como resultado una tasa promedio de bits de aproximadamente 2 kb/s para la información lateral estéreo dentro de una configuración típica que tiene un objetivo de tasa de bits de 96 kb/s.
- 30 c) Solo dos parámetros, la parte real e imaginaria del coeficiente de predicción compleja α , se transmiten en lugar de ICC, CLD, e IPD. Las partes reales e imaginarias de α se tratan de forma independiente, limitadas al intervalo [-3,0, 3,0] y se cuantifican con un tamaño de paso de 0,1. Si un determinado parámetro (parte real o imaginaria de α) no se utiliza en un determinado cuadro, esto se señala en la corriente de bits, y el parámetro irrelevante no se transmite. Los parámetros se codifican de forma diferencial al tiempo o de forma diferencial a la frecuencia y finalmente se aplica la codificación de Huffman utilizando el código de factor de escala. Los coeficientes de predicción se actualizan cada dos bandas de factor de escala, lo que da como resultado una resolución de frecuencia similar a la resolución de frecuencia del entorno de MPEG. Este esquema de cuantificación y codificación da como resultado una tasa promedio de bits de aproximadamente 2 kb/s para la información lateral estéreo dentro de una configuración típica que tiene un objetivo de tasa de bits de 96 kb/s.
- 35 d) En lugar de reutilizar un código predefinido tal como el código de factor de escala AAC también puede utilizarse un código especializado no variable o adaptativo de señales para codificar los valores del parámetro α , o se puede volver a los códigos de longitud fija (por ejemplo, de 4 bits) sin signo o de dos complementos.
- 40 Los detalles de la implementación adicionales o alternativos preferidos comprenden:
- e) El intervalo de los valores del parámetro α así como el tamaño del paso de cuantificación de parámetros se pueden elegir arbitrariamente y optimizarse a las características de señales en cuestión.
- 45 c) Para cada uno de los dos parámetros de α puede elegirse una codificación no diferencial (PCM) o diferencial (DPCM) por cuadro o por corriente, señalada por un bit correspondiente en la corriente de bits. Para una codificación DPCM es posible, o bien una codificación diferencial de tiempo o bien de frecuencia. Nuevamente, esto puede señalarse utilizando una bandera de un bit.
- d) En lugar de reutilizar un código predefinido tal como el código de factor de escala AAC también puede utilizarse un código especializado no variable o adaptativo de señales para codificar los valores del parámetro α , o se puede volver a los códigos de longitud fija (por ejemplo, de 4 bits) sin signo o de dos complementos.
- 50 e) El intervalo de los valores del parámetro α así como el tamaño del paso de cuantificación de parámetros se pueden elegir arbitrariamente y optimizarse a las características de señales en cuestión.
- f) El número y ancho espectral y/o temporal de las bandas de parámetro activo α pueden elegirse arbitrariamente y optimizarse a las características de señales determinadas. En particular, la configuración de banda puede señalarse por cuadro o por corriente.
- 55 g) Además de, o en lugar de los mecanismos indicados en a) anteriormente, puede señalarse explícitamente por medio de un bit por cuadro en la corriente de bits que solo se utiliza el espectro de MDCT del cuadro actual para computar la aproximación del espectro de MDST, es decir, que los cuadros de MDCT adyacentes no se tienen en cuenta.
- 60

Las realizaciones se refieren a un sistema de la invención para la codificación estéreo unificada en el dominio de

MDCT. Esto permite utilizar las ventajas de la codificación estéreo unificada en el sistema de USAC del MPEG incluso a tasas de bits más altas (en las que no se utiliza SBR) sin el aumento significativo en la complejidad computacional que se produciría con un método basado en QMF.

- 5 Las dos listas siguientes resumen los aspectos de la configuración preferidos que se han descrito anteriormente, que pueden utilizarse de manera alternativa uno con respecto a otro o junto con otros aspectos:
- 1a) concepto general: predicción compleja de MDCT lateral a partir de MDCT y MDST centrales;
- 10 1b) calcular/aproximar la MDST a partir de MDCT ("R2I") en el dominio de frecuencia utilizando 1 o más cuadros (3 cuadros causan retardo);
- 1c) truncado del filtro (incluso descendiendo hasta 1-cuadro 2-conexiones, es decir, [-1 0 1]) para reducir la complejidad computacional;
- 15 1d) tratamiento adecuado de DC y $f_s/2$;
- 1e) tratamiento adecuado de conmutación de forma de ventana;
- 20 1f) no utilizar un cuadro anterior/siguiente si este tiene un tamaño de transformada diferente;
- 1g) predicción basada en coeficientes de MDCT no cuantificados o cuantificados en el codificador;
- 2a) cuantificar y codificar la parte real e imaginaria del coeficiente de predicción compleja directamente (es decir, sin parametrización del entorno de MPEG);
- 25 2b) utilizar cuantificador uniforme para esto (tamaño de paso, por ejemplo, 0,1);
- 2c) utilizar una resolución de frecuencia adecuada para los coeficientes de predicción (por ejemplo, 1 coeficiente por 2 bandas de factor de escala);
- 30 2d) señalización económica en el caso de que todos los coeficientes de predicción sean reales;
- 2e) bit explícito por cuadro para forzar la operación de R2I de 1-cuadro.
- 35 En una realización, el codificador comprende adicionalmente: un convertidor (50, 51) espectral para convertir una representación del dominio de tiempo de las dos señales de canal en una representación espectral de las dos señales de canal que tienen señales de subbandas para las dos señales de canal, en la que el combinador (2031), el predictor (2033) y la calculadora (2034) de la señal residual están configurados para procesar cada señal de subbanda por separado de modo que la primera señal combinada y la señal residual se obtienen para una pluralidad de subbandas, en la que la interfaz (212) de salida está configurada para combinar la primera señal combinada codificada y la señal residual codificada para la pluralidad de subbandas.
- 40 A pesar de que algunos aspectos se han descrito en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del método correspondiente, en la que un bloque o dispositivo corresponde a un paso del método o a una característica de un paso del método. De forma análoga, los aspectos descritos en el contexto de un paso del método también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente.
- 45 En una realización de la presente invención, se aplica un tratamiento adecuado de conmutación de forma de ventana. Cuando se tiene en consideración la figura 10a, una información 109 de forma de ventana puede introducirse en la calculadora 1001 del espectro imaginario. Específicamente, la calculadora del espectro imaginario, que lleva a cabo la conversión real-a-imaginaria del espectro de valor real tal como el espectro de MDCT (como por ejemplo el elemento 2070 en la figura 6a o el elemento 1160a en la figura 6b) puede implementarse como un filtro
- 50 FIR o IIR. Los coeficientes FIR o IIR en este módulo 1001 real-a-imaginario dependen de la forma de ventana de la mitad izquierda y de la mitad derecha del cuadro actual. Esta forma de ventana puede ser distinta para una ventana seno o una ventana KBD (derivada de Kaiser Bessel) y dependiendo de la configuración de secuencia de ventana determinada, puede ser una ventana larga, una ventana de inicio, una ventana de parada y una ventana de parada-inicio o una ventana corta. El módulo real-a-imaginario puede comprender un filtro bidimensional FIR, en el que una
- 55 dimensión es la dimensión de tiempo en la que dos cuadros de MDCT posteriores se introducen en el filtro FIR y la segunda dimensión es la dimensión de frecuencia, en la que se introducen los coeficientes de frecuencia de un cuadro.
- 60

La siguiente tabla proporciona distintos coeficientes de filtro MDST para una secuencia de ventana actual para

distintas formas de ventana y distintas implementaciones de la mitad izquierda y la mitad derecha de la ventana.

Tabla A - Parámetros de Filtro MDST para la Ventana Actual

Secuencia de ventana actual	Mitad izquierda: forma seno Mitad derecha: forma seno	Mitad izquierda: forma KBD Mitad derecha: forma KBD
SOLO_SECUENCIA_LARGA, SECUENCIA_CORTA_OCHO	[0,000000, 0,000000, 0,500000, 0,000000, -0,500000, 0,000000, 0,000000]	[0,091497, 0,000000, 0,581427, 0,000000, -0,581427, 0,000000, -0,091497]
SECUENCIA_LARGA_DE_INICIO	[0,102658, 0,103791, 0,567149, 0,000000, -0,567149, -0,103791, -0,102658]	[0,150512, 0,047969, 0,608574, 0,000000, -0,608574, -0,047969, -0,150512]
SECUENCIA_LARGA_DE_PARADA	[0,102658, -0,103791, 0,567149, 0,000000, -0,567149, 0,103791, -0,102658]	[0,150512, -0,047969, 0,608574, 0,000000, -0,608574, 0,047969, -0,150512]
SECUENCIA_DE_PARADA_INICIO	[0,205316, 0,000000, 0,634298, 0,000000, -0,634298, 0,000000, -0,205316]	[0,209526, 0,000000, 0,635722, 0,000000, -0,635722, 0,000000, -0,209526]
Secuencia de ventana actual	Mitad izquierda: forma seno Mitad derecha: forma seno	Mitad izquierda: forma KBD Mitad derecha: forma KBD
SOLO_SECUENCIA_LARGA, SECUENCIA_CORTA_OCHO	[0,045748, 0,057238, 0,540714, 0,000000, -0,540714, -0,057238, -0,045748]	[0,045748, -0,057238, 0,540714, 0,000000, -0,540714, 0,057238, -0,045748]
SECUENCIA_LARGA_DE_INICIO	[0,104763, 0,105207, 0,567861, 0,000000, -0,567861, -0,105207, -0,104763]	[0,148406, 0,046553, 0,607863, 0,000000, -0,607863, -0,046553, -0,148406]
SECUENCIA_LARGA_DE_PARADA	[0,148406, -0,046553, 0,607863, 0,000000, -0,607863, 0,046553, -0,148406]	[0,104763, -0,105207, 0,567861, 0,000000, -0,567861, 0,105207, -0,104763]
SECUENCIA_DE_PARADA_INICIO	[0,207421, 0,001416, 0,635010, 0,000000, -0,635010, -0,001416, -0,207421]	[0,207421, -0,001416, 0,635010, 0,000000, -0,635010, 0,001416, -0,207421]

5 Adicionalmente, la información 109 de forma de ventana proporciona información de forma de ventana para la ventana anterior, cuando se utiliza la ventana anterior para calcular el espectro de MDST a partir del espectro de MDCT. En la siguiente tabla se proporcionan los coeficientes de filtro MDST correspondientes para la ventana anterior.

10

Tabla B - Parámetros de filtro MDST para la ventana anterior

Secuencia de ventana actual	Mitad izquierda de ventana actual: forma seno	Mitad izquierda de ventana actual: forma KBD
SOLO_SECUENCIA_LARGA, SECUENCIA_LARGA_DE_INICIO, SECUENCIA_CORTA_OCHO	[0,000000, 0,106103, 0,250000, 0,318310, 0,250000, 0,106103, 0,000000]	[0,059509, 0,123714, 0,186579, 0,213077, 0,186579, 0,123714, 0,059509]
SECUENCIA_LARGA_DE_PARADA, SECUENCIA_DE_PARADA_INICIO	[0,038498, 0,039212, 0,039645, 0,039790, 0,039645, 0,039212, 0,038498]	[0,026142, 0,026413, 0,026577, 0,026631, 0,026577, 0,026413, 0,026142]

15 De este modo, dependiendo de la información 109 de forma de ventana, la calculadora 1001 del espectro imaginario en la figura 10a se adapta aplicando distintos conjuntos de coeficientes de filtro.

20 La información de la forma de ventana, que se utiliza en el lado del decodificador se calcula en el lado del codificador y se transmite como información lateral junto con la señal de salida del codificador. En el lado del decodificador, la información 109 de forma de ventana se extrae de la corriente de bits mediante el desmultiplexador de la corriente de bits (por ejemplo 102 en la figura 5b) y se proporciona a la calculadora 1001 del espectro imaginario tal como se ilustra en la figura 10a.

25 Cuando la información 109 de forma de ventana señala que el cuadro anterior tenía un tamaño de transformada distinto, entonces se prefiere que el cuadro anterior no se utilice para calcular el espectro imaginario a partir del espectro de valor real. Lo mismo se produce cuando al interpretar la información 109 de forma de ventana se descubre que el cuadro siguiente tiene un tamaño de transformada distinto. Entonces, el cuadro siguiente no se

utiliza para calcular el espectro imaginario a partir del espectro de valor real. En un caso de este tipo en el que, por ejemplo, el cuadro anterior tenía un tamaño de transformada distinto del cuadro actual y cuando el cuadro siguiente nuevamente tiene un tamaño de transformada distinto en comparación con el cuadro actual, entonces solo el cuadro actual, es decir, los valores espectrales de la ventana actual, se utilizan para estimar el espectro imaginario.

5 La predicción en el codificador está basada en los coeficientes de frecuencia no cuantificados o cuantificados tales como los coeficientes de MDCT. Cuando, por ejemplo, la predicción ilustrada por el elemento 2033 en la figura 3a, está basada en datos no cuantificados, entonces la calculadora 2034 residual también opera preferiblemente sobre datos no cuantificados y la señal residual de salida de la calculadora, es decir, la señal 205 residual se cuantifica antes de ser codificada por entropía y transmitirse a un decodificador. Sin embargo, en una realización alternativa se prefiere que la predicción esté basada en coeficientes de MDCT cuantificados. Entonces, la cuantificación puede ocurrir antes del combinador 2031 en la figura 3a de tal modo que un primer canal cuantificado y un segundo canal cuantificado sean la base para calcular la señal residual. Alternativamente, la cuantificación también puede ocurrir después del combinador 2031 de modo que la primera señal de combinación y la segunda señal de combinación se calculen de forma no cuantificada y se cuantifiquen antes de que se calcule la señal residual. De nuevo, como alternativa, el predictor 2033 puede operar en el dominio no cuantificado y la señal 2035 de predicción se cuantifica antes de introducirse en la calculadora residual. Entonces, es útil que la segunda señal 2032 de combinación, que también se introduce en la calculadora 2034 residual, también se cuantifique antes de que la calculadora residual calcule la señal 1070 residual en la figura 6a, lo que puede implementarse dentro del predictor 2033 en la figura 3a, opere en los mismos datos cuantificados que estén disponibles en el lado del decodificador. Entonces puede garantizarse que el espectro de MDST estimado en el codificador con el fin de realizar el cálculo de la señal residual sea exactamente el mismo que el espectro de MDST en el lado del decodificador utilizado para llevar a cabo la predicción inversa, es decir, para calcular la señal lateral desde la señal residual. Para tal fin, la primera señal de combinación tal como la señal M en la línea 204 en la figura 6a se cuantifica antes de introducirse en el bloque 2070. Entonces, el espectro de MDST calculado utilizando el espectro de MDCT cuantificado del cuadro actual, y dependiendo de la información de control, el espectro de MDCT cuantificado del cuadro anterior o siguiente, se introduce en el multiplicador 2074, y la salida del multiplicador 2074 de la figura 6a nuevamente será un espectro no cuantificado. Este espectro no cuantificado se sustraerá del espectro introducido en el sumador 2034b y finalmente se cuantificará en el cuantificador 209b.

30 En una realización, la parte real y la parte imaginaria del coeficiente de predicción compleja por banda de predicción se cuantifican y codifican directamente, es decir, por ejemplo, sin la parametrización del entorno MPEG. La cuantificación puede llevarse a cabo utilizando un cuantificador uniforme con un tamaño de paso, por ejemplo, de 0,1. Esto significa que no se aplica ningún tamaño de paso de cuantificación logarítmica o similar, sino que se aplica cualquier tamaño de paso lineal. En una implementación, el intervalo de valores para la parte real y la parte imaginaria del coeficiente de predicción compleja oscila entre -3 y 3, lo que significa que se utilizan 60 o, dependiendo de los detalles de implementación, 61 pasos de cuantificación para la parte real y la parte imaginaria del coeficiente de predicción compleja.

40 Preferiblemente, la parte real aplicada en el multiplicador 2073 en la figura 6a y la parte 2074 imaginaria aplicada en la figura 6a se cuantifican antes de aplicarse de modo que, nuevamente, se utiliza el mismo valor para la predicción en el lado del codificador, así como cuando está disponible en el lado del decodificador. Esto garantiza que la señal residual de predicción abarque, además del error de cuantificación introducido, cualquier error que pudiera producirse cuando se aplica un coeficiente de predicción no cuantificado en el lado del codificador mientras se aplica un coeficiente de predicción cuantificado en el lado del decodificador. Preferiblemente, la cuantificación se aplica de modo que, en la medida de lo posible, la misma situación y las mismas señales estén disponibles en el lado del codificador y en el lado del decodificador. Por tanto, se prefiere cuantificar la entrada en la 2070 calculadora real-imaginaria utilizando la misma cuantificación aplicada en el cuantificador 209a. Adicionalmente, se prefiere cuantificar la parte real y la parte imaginaria del coeficiente de predicción de α para llevar a cabo las multiplicaciones en el elemento 2073 y en el elemento 2074. La cuantificación es la misma que se aplica en el cuantificador 2072. Adicionalmente, la señal lateral emitida por el bloque 2031 en la figura 6a también puede cuantificarse antes de los sumadores 2034a y 2034b. Sin embargo, no resulta problemático que el cuantificador 209b lleve a cabo la cuantificación después de la adición, en la que la adición por estos sumadores se aplica con una señal lateral no cuantificada.

55 En un ejemplo adicional, puede aplicarse una señalización económica en caso de que todos los coeficientes de predicción sean reales. Puede ocurrir que todos los coeficientes de predicción para un cuadro determinado, es decir, para la misma porción de tiempo de la señal de audio se calculen como reales. Una situación de este tipo puede producirse cuando la señal central completa y la señal lateral completa no están desfasadas una con respecto a otra o están un poco desfasadas. Con el fin de ahorrar bits, esto se indica por un único indicador real. Entonces, no es necesario señalar la parte imaginaria del coeficiente de predicción en la corriente de bits con un código que represente un valor cero. En el lado del decodificador, la interfaz del decodificador de corriente de bits, tal como un demultiplexador de corriente de bits, interpretará este indicador real y entonces no buscará códigos para una parte imaginaria, sino que entenderá que todos los bits están en la sección correspondiente de la corriente de bits como

bits para coeficientes de predicción de valor real. Además, cuando el predictor 2033 recibe una indicación de que todas las partes imaginarias de los coeficientes de predicción en el cuadro son cero, no necesitará calcular un espectro de MDST, o generalmente un espectro imaginario a partir del espectro de MDCT de valor real. Por lo tanto, el elemento 1160a en el decodificador de la figura 6b se desactivará y la predicción inversa solo tendrá lugar utilizando el coeficiente de predicción de valor real aplicado en el multiplicador 1160b en la figura 6b. Esto mismo se produce para el lado del codificador en el que el elemento 2070 se desactivará y la predicción solo tendrá lugar utilizando el multiplicador 2073. Esta información lateral se utiliza preferiblemente como un bit adicional por cuadro y el decodificador leerá este bit, cuadro por cuadro, con el fin de decidir si el convertidor 1160a real-a-imaginario estará activo para un cuadro o no. De este modo, proporcionar esta información da como resultado un tamaño reducido de la corriente de bits debido a la señalización más eficiente de todas las partes imaginarias del coeficiente de predicción que son es para un cuadro y, adicionalmente, proporciona menos complejidad para el decodificador para dicho cuadro, lo que inmediatamente da como resultado un menor consumo de batería de un procesador de este tipo implementado, por ejemplo, en un dispositivo móvil alimentado por batería.

La predicción compleja estéreo según las realizaciones preferidas de la presente invención es una herramienta para la codificación eficiente de pares de canales con diferencias de nivel y/o fase entre los canales. Al utilizar un parámetro de valor complejo α , los canales izquierdo y derecho se reconstruyen por medio de la matriz siguiente. dmx_{Im} indica la MDST correspondiente a la MDCT de los canales de mezcla descendente dmx_{Re} .

$$\begin{bmatrix} r \\ l \end{bmatrix} = \begin{bmatrix} 1 - \alpha_{Re} & -\alpha_{Im} & 1 \\ 1 + \alpha_{Re} & \alpha_{Im} & -1 \end{bmatrix} \begin{bmatrix} dmx_{Re} \\ dmx_{Im} \\ res \end{bmatrix}$$

La ecuación anterior es otra representación, que se divide con respecto a la parte real y a la parte imaginaria de α y representa la ecuación para una operación combinada de predicción/combinación, en la que la señal previsible S no se calcula necesariamente.

Los siguientes elementos de datos se utilizan preferiblemente para esta herramienta:

- cplx_pred all** 0: Algunas bandas utilizan la codificación L/R, señalada por `cplx_pred_used[]`
1: Todas las bandas utilizan la predicción compleja estéreo
- cplx_pred_used[g][sfb]** Una bandera de un bit por grupo de ventanas g y banda de factor de escala sfb (después del mapeo a partir de las bandas de predicción) que indica que
0: no se utiliza la predicción compleja, se utiliza la codificación L/R
1: se utiliza la predicción compleja
- complex_coef** 0: $\alpha_{Im} = 0$ para todas las bandas de predicción
1: α_{Im} se transmite para todas las bandas de predicción
- use_prev_frame** 0: Utilizar solo el cuadro actual para estimar la MDST
1: Utilizar el cuadro actual y anterior para estimar la MDST
- delta_code_time** 0: Codificación diferencial de frecuencia de los coeficientes de predicción
1: Codificación diferencial de tiempo de los coeficientes de predicción
- hcod_alpha_q_re** Código de Huffman de α_{Re}
- hcod_alpha_q_im** Código de Huffman de α_{Im}

Estos elementos de datos se calculan en un codificador y se introducen en la información lateral de una señal de audio multicanal o estéreo. Los elementos se extraen de la información lateral en el lado del decodificador por medio de un extractor de información lateral y se utilizan para controlar la calculadora del decodificador para llevar a cabo una acción correspondiente.

La predicción compleja estéreo necesita el espectro de MDCT de mezcla descendente del par de canales actuales y, en el caso de `complex_coef = 1`, un valor estimado del espectro de MDST de mezcla descendente del par de canales actuales, es decir, la contraparte imaginaria del espectro de MDCT. El valor estimado de la MDST de mezcla descendente se computa desde la MDCT de mezcla descendente del cuadro actual y, en el caso de `use_prev_frame = 1`, desde la MDCT de mezcla descendente del cuadro anterior. La MDCT de mezcla descendente del cuadro anterior del grupo de ventanas g y del grupo de ventanas b se obtiene a partir de los espectros izquierdo y derecho reconstruidos de ese cuadro.

En el cómputo del valor estimado de la MDST de mezcla descendente se utiliza la longitud de transformada MDCT de valor uniforme, que depende de `window_sequence`, así como también de `filter_coefs` y `filter_coefs_prev`, que son las matrices que contienen los núcleos de filtro y que se obtienen según las tablas anteriores.

Según realizaciones de la invención, para todos los coeficientes de predicción, la diferencia de un valor anterior (en tiempo o frecuencia) se codifica utilizando un código de Huffman. Los coeficientes de predicción no se transmiten para las bandas de predicción para las que se utilizó $cplx_pred_used = 0$.

5 Los coeficientes de predicción inversa cuantificados α_{re} y α_{im} son proporcionados por

$$\alpha_{re} = \alpha_{q_re} * 0,1$$

$$\alpha_{im} = \alpha_{q_im} * 0,1$$

10 Ha de enfatizarse que la invención no solo puede aplicarse a señales estéreo, es decir, señales multicanal que tienen solo dos canales, sino que también se aplica a dos canales de una señal multicanal que tiene tres o más canales tales como una señal 5.1 o 7.1.

15 La señal de audio codificada de la invención puede almacenarse en un medio de almacenamiento digital o puede transmitirse en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable como Internet.

20 Dependiendo de determinados requisitos para la implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede llevarse a cabo utilizando un medio digital de almacenamiento, por ejemplo un disco flexible, un DVD, un CD, una ROM (memoria de solo lectura), una PROM (memoria programable de solo lectura), una EPROM (una PROM borrable), una EEPROM (una PROM borrable eléctricamente) o una memoria FLASH, que tiene señales de control legibles electrónicamente almacenadas en la misma, que actúan conjuntamente (o que pueden actuar conjuntamente) con un sistema informático programable de modo que se lleve a cabo el método respectivo.

30 Algunas realizaciones según la invención comprenden un portador de datos no transitorio o tangible que tiene señales de control legibles electrónicamente, que pueden actuar conjuntamente con un sistema informático programable, de modo que se lleve a cabo uno de los métodos descritos en el presente documento.

35 En general, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa que puede hacerse funcionar para llevar a cabo uno de los métodos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede almacenarse, por ejemplo, en un portador legible por máquina.

Otras realizaciones comprenden el programa informático para llevar a cabo uno de los métodos descritos en el presente documento, almacenados en un portador legible por máquina.

40 Dicho de otro modo, una realización del método de la invención es, por tanto, un programa informático que tiene un código de programa para llevar a cabo uno de los métodos descritos en el presente documento, cuando el programa informático se ejecuta en un ordenador.

45 Una realización adicional de los métodos de la invención es, por tanto, un portador de datos (o un medio digital de almacenamiento, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para llevar a cabo uno de los métodos descritos en el presente documento.

50 Una realización adicional del método de la invención es, por tanto, una corriente de datos o una secuencia de señales que representa el programa informático para llevar a cabo uno de los métodos descritos en el presente documento. La corriente de datos o la secuencia de señales pueden estar configuradas, por ejemplo, para transferirse por medio de una conexión de comunicación de datos, por ejemplo, a través de Internet.

Otra realización comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los métodos descritos en el presente documento.

55 Otra realización comprende un ordenador que tiene instalado en el mismo el programa informático para llevar a cabo uno de los métodos descritos en el presente documento.

60 En algunas realizaciones, puede utilizarse un dispositivo lógico programable (por ejemplo, matrices de puertas programables de campo) para llevar a cabo algunas o todas las funcionalidades de los métodos descritos en el presente documento. En algunas realizaciones, una matriz de puerta programable de campo puede actuar conjuntamente con un microprocesador con el fin de llevar a cabo uno de los métodos descritos en el presente documento. En general, los métodos se llevan a cabo, preferiblemente, por cualquier aparato de hardware.

Las realizaciones descritas anteriormente son simplemente ilustrativas de los principios de la presente invención.

Cabe entender que las modificaciones y variantes de las disposiciones y detalles descritos en el presente documento resultarán evidentes para los expertos en la técnica. Por consiguiente, la invención solo está limitada por el alcance de las siguientes reivindicaciones de la patente y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones incluidas en el presente documento.

5

REIVINDICACIONES

1. Decodificador de audio para decodificar una señal (100) de audio multicanal codificada, comprendiendo la señal de audio multicanal codificada una primera señal de combinación codificada generada basándose en una regla de combinación para combinar una primera señal de audio de canal y una segunda señal de audio de canal de una señal de audio multicanal, comprendiendo una señal residual de predicción codificada e información de predicción:

5 un decodificador (110) de señales para decodificar la primera señal (104) de combinación codificada para obtener una primera señal (112) de combinación decodificada y para decodificar la señal (106) residual codificada para obtener una señal (114) residual decodificada; y

10 una calculadora (116) del decodificador para calcular una señal multicanal decodificada que tiene una primera señal (117) de canal decodificada, y una segunda señal (118) de canal decodificada utilizando la señal (114) residual decodificada, la información (108) de predicción y la primera señal (112) de combinación decodificada, de modo que la primera señal (117) de canal decodificada y la segunda señal (118) de canal decodificada son al menos aproximaciones de la primera señal de canal y la segunda señal de canal de la señal multicanal, en el que la información (108) de predicción comprende una porción imaginaria distinta de cero,

15 en el que la información (108) de predicción comprende códigos de diferencias entre porciones imaginarias adyacentes de frecuencia, y

20 en el que el decodificador de audio está configurado para llevar a cabo un paso de decodificación por entropía y un paso de decodificación de diferencia posterior para obtener valores de predicción complejos que tienen las porciones imaginarias para bandas de frecuencia adyacentes.
2. Decodificador de audio según la reivindicación 1, en el que la calculadora (116) del decodificador comprende:

30 un predictor (1160) para aplicar la información (108) de predicción a la primera señal (112) de combinación decodificada o a una señal (601) procedente de la primera señal de combinación decodificada para obtener una señal (1163) de predicción;

35 una calculadora (1161) de señales de combinación para calcular una segunda señal (1165) de combinación combinando la señal (114) residual decodificada y la señal (1163) de predicción; y

40 un combinador (1162) para combinar la primera señal (112) de combinación decodificada y la segunda señal (1165) de combinación para obtener una señal de audio multicanal decodificada que tiene la primera señal (117) de canal decodificada y la segunda señal (118) de canal decodificada.
3. Decodificador de audio según la reivindicación 1 o 2,

45 en el que la primera señal (104) de combinación codificada y la señal (106) residual codificada se han generado utilizando un solapamiento que genera la conversión de tiempo-espectral,

en el que el decodificador además comprende:

50 un convertidor (52, 53) espectral-tiempo para generar una primera señal de canal de dominio de tiempo y una segunda señal de canal de dominio de tiempo utilizando un algoritmo de conversión espectral-tiempo que coincide con el algoritmo de conversión de tiempo-espectral;

55 un procesador (522) de solapar/agregar para llevar a cabo un procesamiento de solapar-agregar para la primera señal de canal de dominio de tiempo y para la segunda señal de canal de dominio de tiempo para obtener una primera señal de dominio de tiempo sin solapamiento y una segunda señal de dominio de tiempo sin solapamiento.
4. Decodificador de audio según una de las reivindicaciones anteriores, en el que la información (108) de predicción comprende un factor real diferente de cero,

60 en el que el predictor (1160) está configurado para multiplicar la primera señal de combinación decodificada por el factor real para obtener una primera parte de la señal de predicción, y

en el que la calculadora de señal de combinación está configurada para combinar linealmente la señal

residual decodificada y la primera parte de la señal de predicción.

5. Decodificador de audio según una de las reivindicaciones anteriores,

5 en el que la primera señal (104) de combinación codificada o decodificada y la señal (106) residual de predicción codificada o decodificada comprenden, cada una, una primera pluralidad de señales de subbandas,

10 en el que la información de predicción comprende una segunda pluralidad de parámetros de información de predicción, siendo la segunda pluralidad más pequeña que la primera pluralidad,

en el que el predictor (1160) está configurado para aplicar el mismo parámetro de predicción a al menos dos señales de subbandas diferentes de la primera señal de combinación decodificada,

15 en el que la calculadora (116) del decodificador o la calculadora (1161) de señales de combinación o el combinador (1162) están configurados para llevar a cabo un procesamiento en subbandas; y

20 en el que el decodificador de audio comprende además un banco (52, 53) de filtros de síntesis para combinar señales de subbandas de la primera señal de combinación decodificada y la segunda señal de combinación decodificada para obtener una primera señal decodificada de dominio de tiempo y una segunda señal decodificada de dominio de tiempo.

6. Decodificador de audio según la reivindicación 1, en el que la información de predicción comprende un coeficiente de predicción que tiene una porción real y la porción imaginaria, en el que cada una de la porción real y la porción imaginaria se codifica de manera no diferencial o se codifica de manera diferencial de modo marco por marco o corriente por corriente señalado por un bit correspondiente en una corriente de bits que representa la señal (100) de audio multicanal codificada, en el que la señal (100) de audio multicanal codificada comprende además, como información adicional, una bandera de un bit que indica, en un primer estado, que, cuando la porción real o la porción imaginaria se codifica de manera diferencial, la información (108) de predicción comprende los códigos de diferencias entre porciones imaginarias o reales adyacentes de frecuencia y que indica en un segundo estado que la información (108) de predicción comprende códigos de diferencias entre porciones imaginarias o reales de secuencia de tiempo, en el que el decodificador de audio está configurado para extraer el elemento de datos de la señal (100) de audio multicanal codificada, y en el que el decodificador de audio está configurado para llevar a cabo un paso de decodificación por entropía y un paso de decodificación diferencial posterior para obtener porciones imaginarias o reales cuantificadas en secuencias de tiempo, cuando la bandera de un bit presenta el segundo estado.

7. Codificador de audio para codificar una señal de audio multicanal que tiene dos o más señales de canal, que comprende:

40 una calculadora (203) del codificador para calcular una primera señal (204) de combinación y una señal (205) de predicción residual utilizando una primera señal (201) de canal y una segunda señal (202) de canal e información (206) de predicción, de modo que una señal de predicción residual, cuando se combina con una señal de predicción procedente de la primera señal de combinación o una señal procedente de la primera señal de combinación y la información (206) de predicción da como resultado una segunda señal (2032) de combinación, pudiendo provenir la primera señal (204) de combinación y la segunda señal (2032) de combinación de la primera señal (201) de canal y la segunda señal (202) de canal utilizando un regla de combinación;

50 un optimizador (207) para calcular la información (206) de predicción de modo que la señal (205) de predicción residual cumpla un objetivo (208) de optimización, en el que la información (206) de predicción comprende una porción imaginaria diferente de cero;

55 un codificador (209) de señal para codificar la primera señal (204) de combinación y la señal (205) de predicción residual para obtener una primera señal (210) de combinación codificada y una señal (211) codificada residual;

60 medios para calcular diferencias entre porciones imaginarias adyacentes de frecuencia de la información de predicción y para codificar por entropía todas las diferencias utilizando un código Huffman para obtener códigos tal como la información de predicción; y

una interfaz (212) de salida para combinar la primera señal (210) de combinación codificada, la señal (221) residual de predicción codificada y la información (206) de predicción para obtener una señal de audio

multicanal codificada.

8. Codificador de audio según la reivindicación 7, en el que la calculadora (203) de codificador comprende:
- 5 un combinador (2031) para combinar la primera señal (201) de canal y la segunda señal (202) de canal de dos maneras diferentes para obtener la primera señal (204) de combinación y la segunda señal (2032) de combinación;
- 10 un predictor (2033) para aplicar la información (206) de predicción a la primera señal (204) de combinación o una señal (600) procedente de la primera señal (204) de combinación para obtener una señal (2035) de predicción; y
- una calculadora (2034) de señal residual para calcular la señal (205) residual de predicción combinando la señal (2035) de predicción y la segunda señal (2032) de combinación.
- 15 9. Codificador de audio según la reivindicación 8, en el que el predictor (2033) comprende un cuantificador para cuantificar la primera señal de canal, la segunda señal de canal, la primera señal de combinación o la segunda señal de combinación para obtener una o más señales cuantificadas, y en el que el predictor (2033) está configurado para calcular la señal residual utilizando señales cuantificadas.
- 20 10. Codificador de audio según una de las reivindicaciones 7 a 9,
- en el que la primera señal de canal es una representación espectral de un bloque de muestras;
- 25 en el que la segunda señal de canal es una representación espectral de un bloque de muestras,
- en el que las representaciones espectrales son o bien representaciones espectrales puramente reales o representaciones espectrales puramente imaginarias,
- 30 en el que el optimizador (207) está configurado para calcular la información (206) de predicción como un factor de valor real diferente de cero y/o como un factor imaginario diferente de cero, y
- en el que la calculadora (203) de codificador está configurada para calcular la primera señal de combinación y la señal residual de predicción de modo que la señal de predicción procede de la representación espectral puramente real o la representación espectral puramente imaginaria utilizando el factor de valor real.
- 35 11. Codificador de audio según una de las reivindicaciones 7 a 10,
- en el que la primera señal de canal es una representación espectral de un bloque de muestras;
- 40 en el que la segunda señal de canal es una representación espectral de un bloque de muestras,
- en el que las representaciones espectrales son o bien representaciones espectrales puramente reales o representaciones espectrales puramente imaginarias,
- 45 en el que el optimizador (207) está configurado para calcular la información (206) de predicción como un factor de valor real diferente de cero y/o como un factor imaginario diferente de cero, y
- 50 en el que la calculadora (203) de codificador comprende un transformador (2070) de real a imaginario o un transformador de imaginario a real para derivar una representación espectral transformada a partir de la primera señal de combinación, y
- en el que la calculadora (203) de codificador está configurada para calcular la primera señal (204) combinada y la primera señal (2032) residual de modo que la señal de predicción procede del espectro transformado utilizando el factor imaginario.
- 55 12. Codificador según una de las reivindicaciones 7 a 11,
- 60 en el que el predictor (2033) está configurado para multiplicar la primera señal (204) de combinación mediante una porción real de la información (2073) de predicción para obtener una primera parte de la señal de predicción;
- para estimar (2070) una parte (600) imaginaria de la primera señal de combinación utilizando la primera señal (204) de combinación para multiplicar la parte imaginaria de la primera señal combinada por una

- porción imaginaria de la información (2074) de predicción para obtener una segunda parte de la señal de predicción; y
- 5 en el que la calculadora (2034) residual está configurada para combinar linealmente la primera parte de la señal de predicción o la segunda parte de la señal de predicción y la segunda señal de combinación para obtener la señal (205) residual de predicción.
13. Codificador de audio según la reivindicación 7, en el que la información de predicción comprende un coeficiente de predicción que tiene una porción real y la porción imaginaria, en el que cada una de la porción real y la porción imaginaria se codifica de manera no diferencial o se codifica de manera diferencial de modo marco por marco o corriente por corriente señalado por un bit correspondiente en una corriente de bits que representa la señal (100) de audio multicanal codificada, en el que el codificador de audio está configurado para calcular una bandera de un bit que indica, en un primer estado, que, cuando la porción real o la porción imaginaria va a codificarse de manera diferencial, la información (108) de predicción comprende los códigos de diferencias entre porciones imaginarias o reales adyacentes de frecuencia y que indica en un segundo estado que la información (108) de predicción comprende códigos de diferencias entre porciones imaginarias o reales de secuencia de tiempo, y poner la bandera de un bit en una información adicional de la señal (100) de audio multicanal codificada.
- 10
- 15
- 20 14. Codificador de audio según la reivindicación 7, en el que la predicción de información comprende coeficientes de predicción en los que el código de Huffman es un código de factor de escala, y en el que los coeficientes de predicción se actualizan cada dos bandas de factor de escala.
- 25 15. Método para decodificar una señal (100) de audio multicanal codificada, comprendiendo la señal de audio multicanal codificada una primera señal de combinación codificada generada en base a una regla de combinación para combinar una primera señal de audio de canal y una segunda señal de audio de canal de una señal de audio multicanal, una señal residual de predicción codificada e información de predicción, que comprende:
- 30 decodificar (110) la primera señal (104) de combinación codificada para obtener una primera señal (112) de combinación decodificada, y decodificar la señal (106) residual codificada para obtener una señal (114) residual decodificada; y
- 35 calcular (116) una señal multicanal decodificada que tiene una primera señal (117) de canal decodificada, y una segunda señal (118) de canal decodificada utilizando la señal (114) residual decodificada, la información (108) de predicción y la primera señal (112) de combinación decodificada, de modo que la primera señal (117) de canal decodificada y la segunda señal (118) de canal decodificada son al menos aproximaciones de la primera señal de canal y la segunda señal de canal de la señal multicanal, en el que la información (108) de predicción comprende una porción imaginaria distinta de cero.
- 40 en el que la información (108) de predicción comprende códigos de diferencias entre porciones imaginarias adyacentes de frecuencia, y
- 45 en el que el método comprende un paso de decodificación por entropía y un paso de decodificación de diferencia posterior para obtener valores de predicción complejos que tienen las porciones imaginarias para bandas de frecuencia adyacentes.
- 50 16. Método de codificación de una señal de audio multicanal que tiene dos o más señales de canal, que comprende:
- 55 calcular (203) una primera señal (204) de combinación y una señal (205) residual de predicción utilizando una primera señal (201) de canal y una segunda señal (202) de canal e información (206) de predicción, de modo que una señal residual de predicción, cuando se combina con una señal de predicción procedente de la primera señal de combinación o una señal procedente de la primera señal de combinación y la información (206) de predicción da como resultado una segunda señal (2032) de combinación,
- 60 pudiendo la primera señal (204) de combinación y la segunda señal (2032) de combinación derivarse de la primera señal (201) de canal y la segunda señal (202) de canal utilizando una regla de combinación;
- calcular (207) la información (206) de predicción de modo que la señal (205) residual de predicción cumple un objetivo (208) de optimización, en el que la información (206) de predicción comprende una porción imaginaria distinta de cero;
- codificar (209) la primera señal (204) de combinación y la señal (205) residual de predicción para obtener

una primera señal (210) de combinación codificada y una señal (211) residual codificada; y

5 calcular diferencias entre porciones imaginarias adyacentes de frecuencia de la información de predicción y codificar mediante entropía todas las diferencias utilizando un código Huffman para obtener códigos tal como la información de predicción; y

combinar (212) la primera señal (210) de combinación codificada, la señal (211) residual de predicción codificada y la información (206) de predicción para obtener una señal de audio multicanal codificada.

10 17. Programa informático adaptado para llevar a cabo, cuando se ejecuta en un ordenador o en un procesador, el método según la reivindicación 15 o el método según la reivindicación 16.

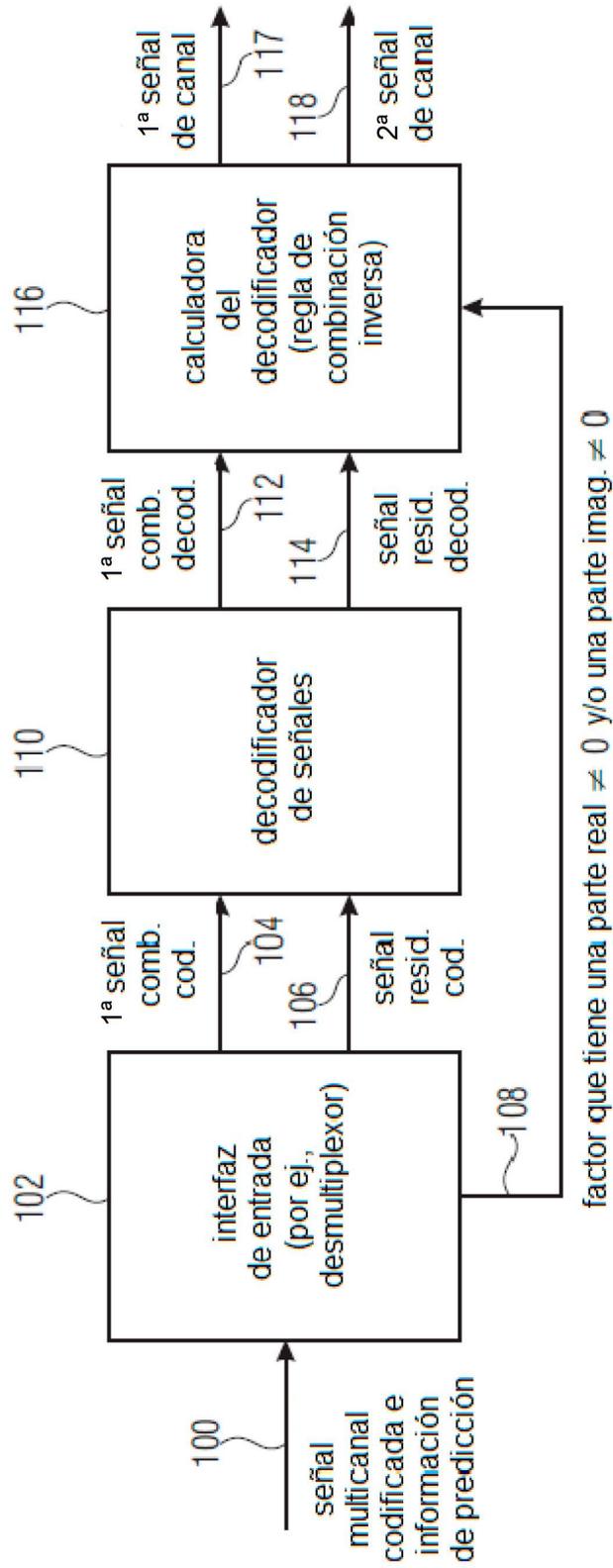


FIGURA 1
(DECODIFICADOR DE AUDIO)

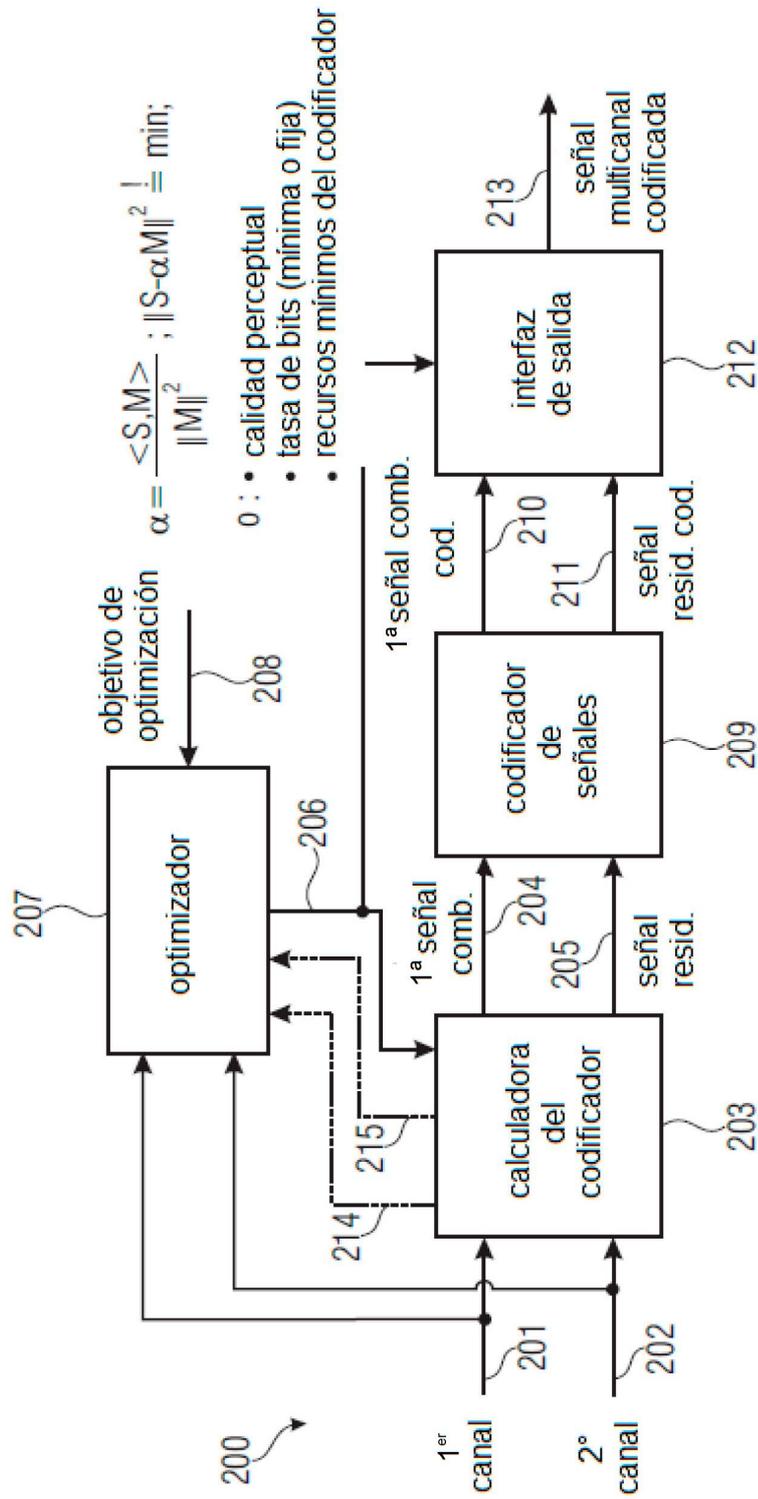


FIGURA 2
(CODIFICADOR DE AUDIO)

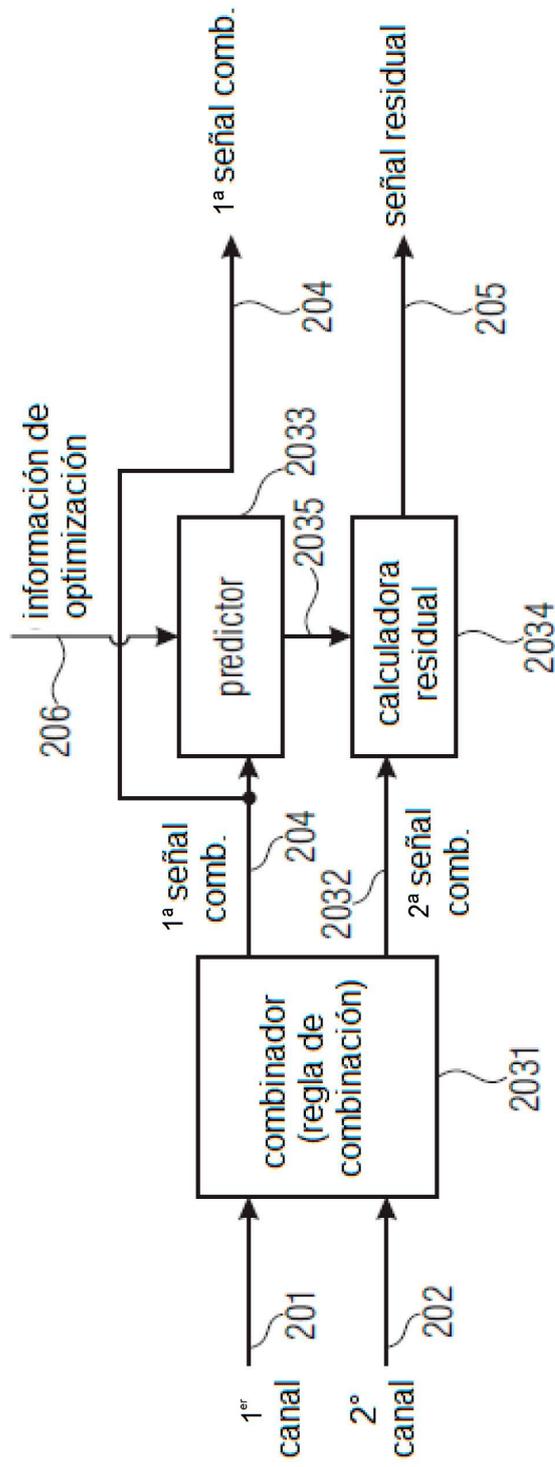
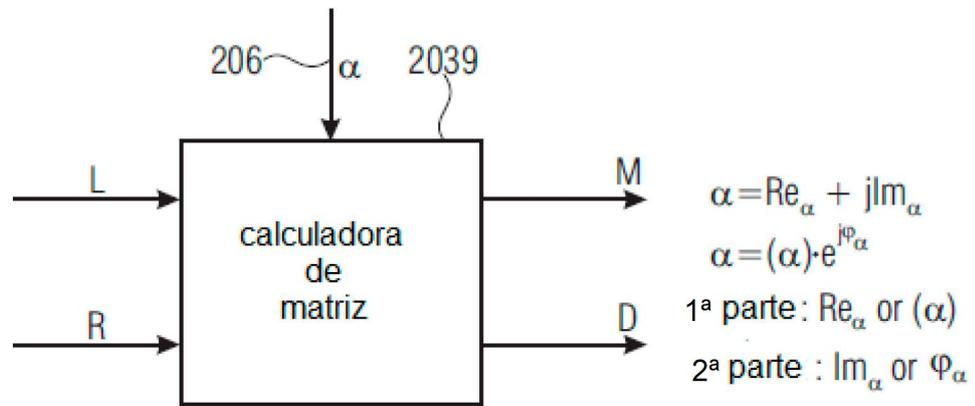


FIGURA 3A
(CODIFICADOR DE AUDIO LATERAL)



$$\begin{bmatrix} M \\ D \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1-\alpha & -1-\alpha \end{bmatrix} \begin{bmatrix} L \\ R \end{bmatrix}$$

FIGURA 3B
(CODIFICADOR DE AUDIO LATERAL)

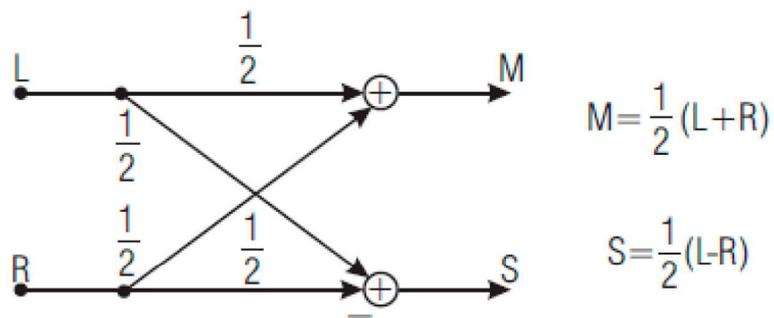


FIGURA 3C
(REGLA DE COMBINACIÓN)

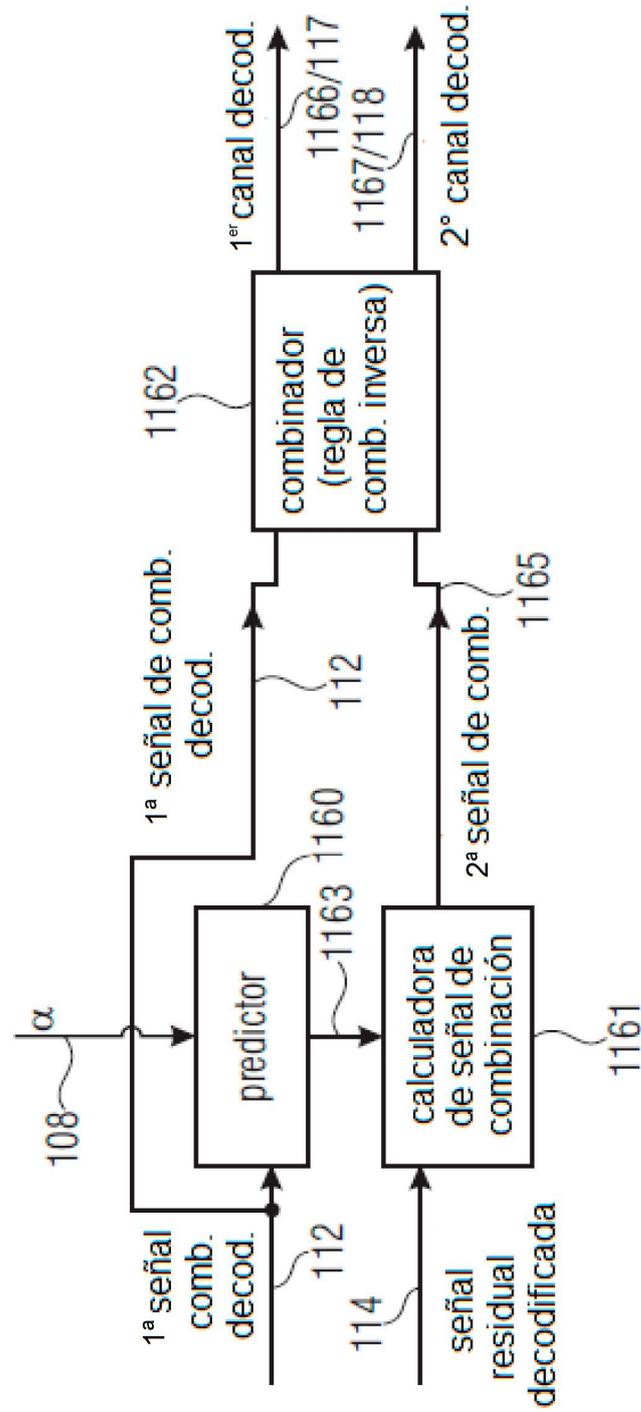


FIGURA 4A
(DECODIFICADOR DE AUDIO LATERAL)

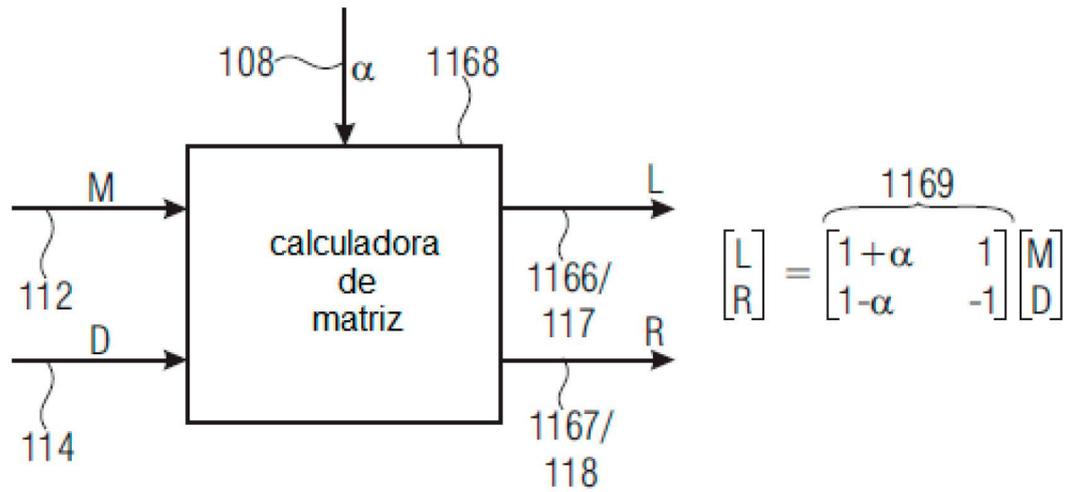


FIGURA 4B
(DECODIFICADOR DE AUDIO LATERAL)

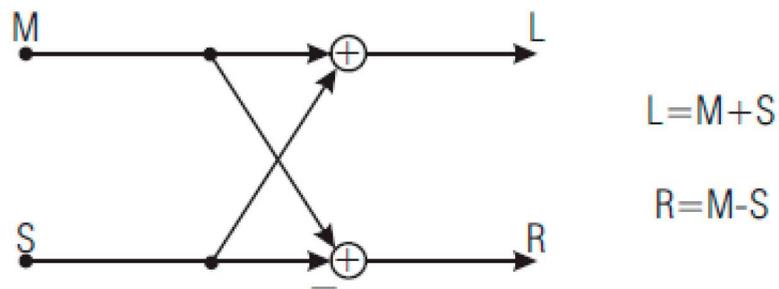


FIGURA 4C
(REGLA DE COMBINACIÓN INVERSA)

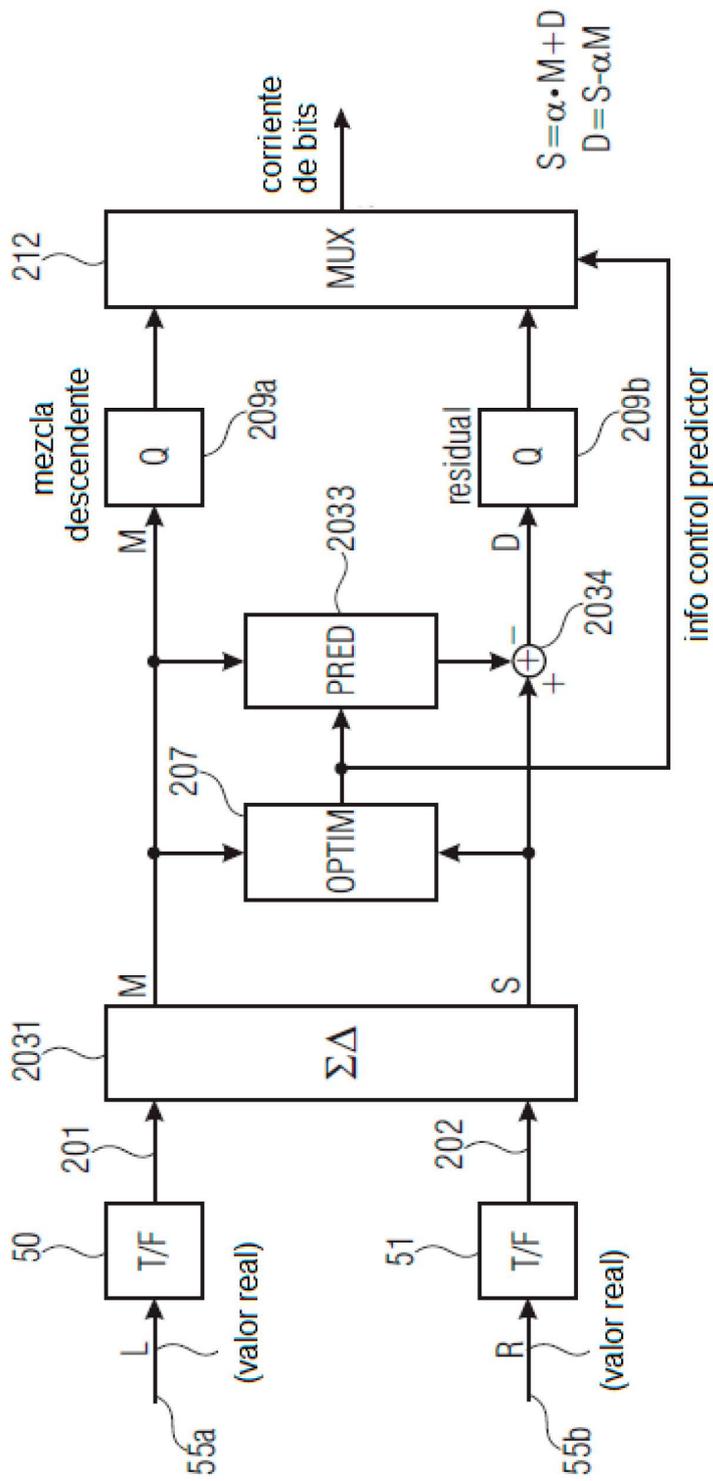


FIGURA 5A
(CODIFICADOR LATERAL)

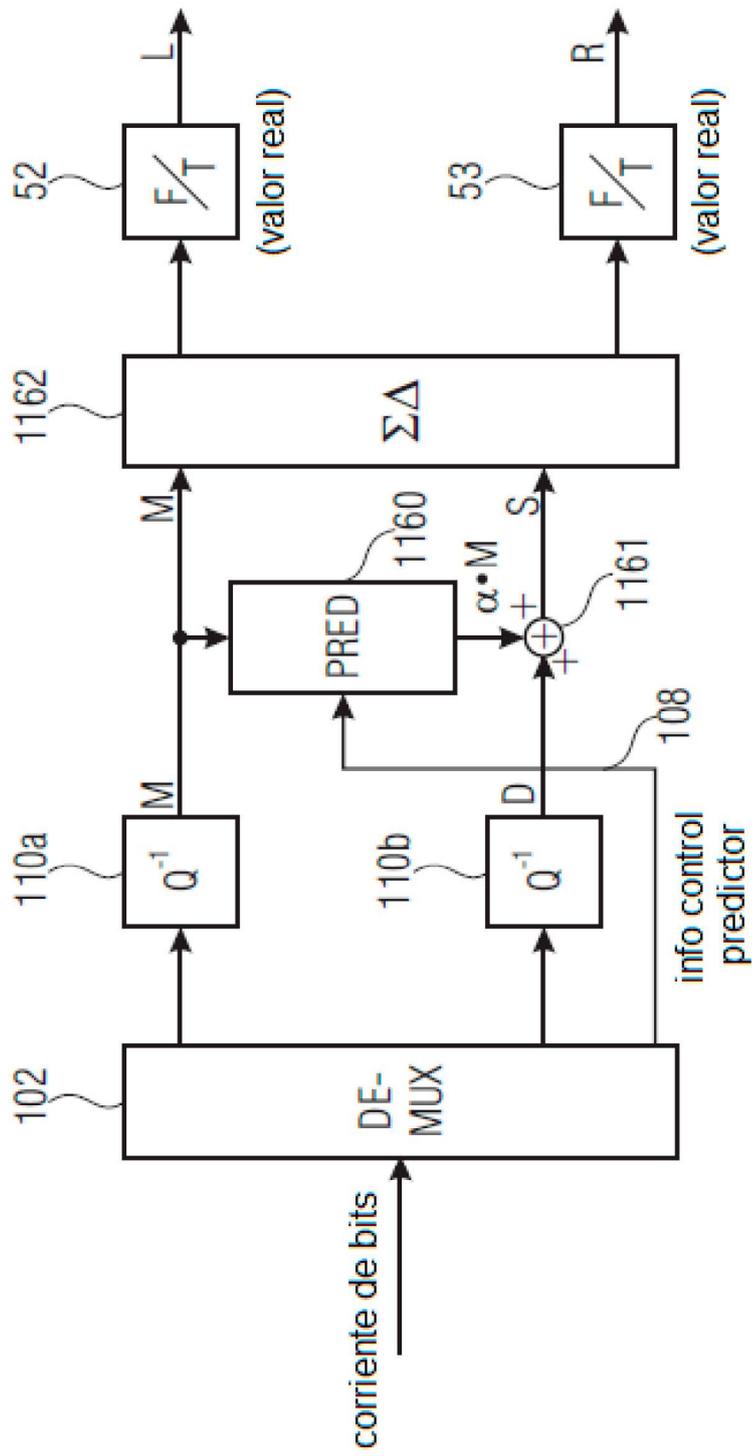


FIGURA 5B
(DECODIFICADOR LATERAL)

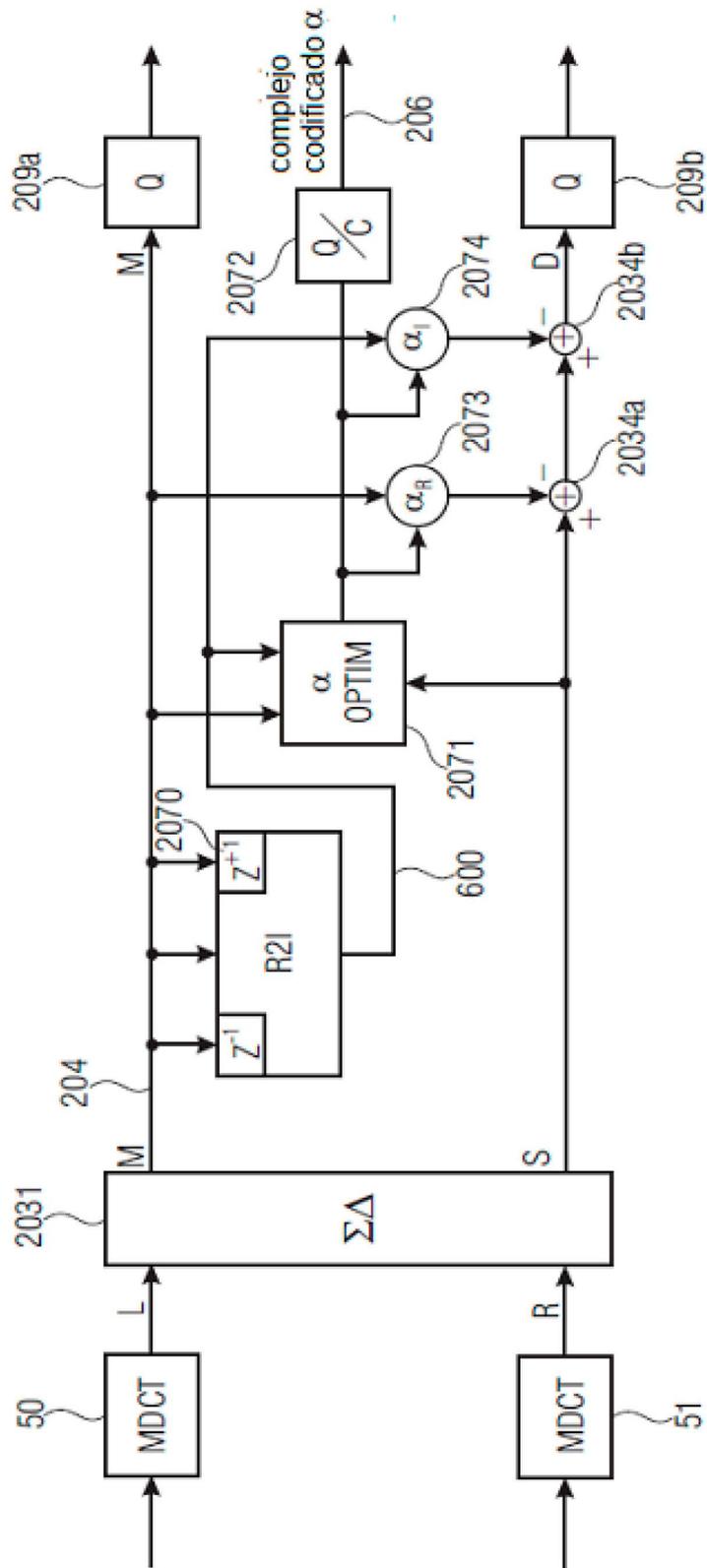


FIGURA 6A
(CODIFICADOR LATERAL)

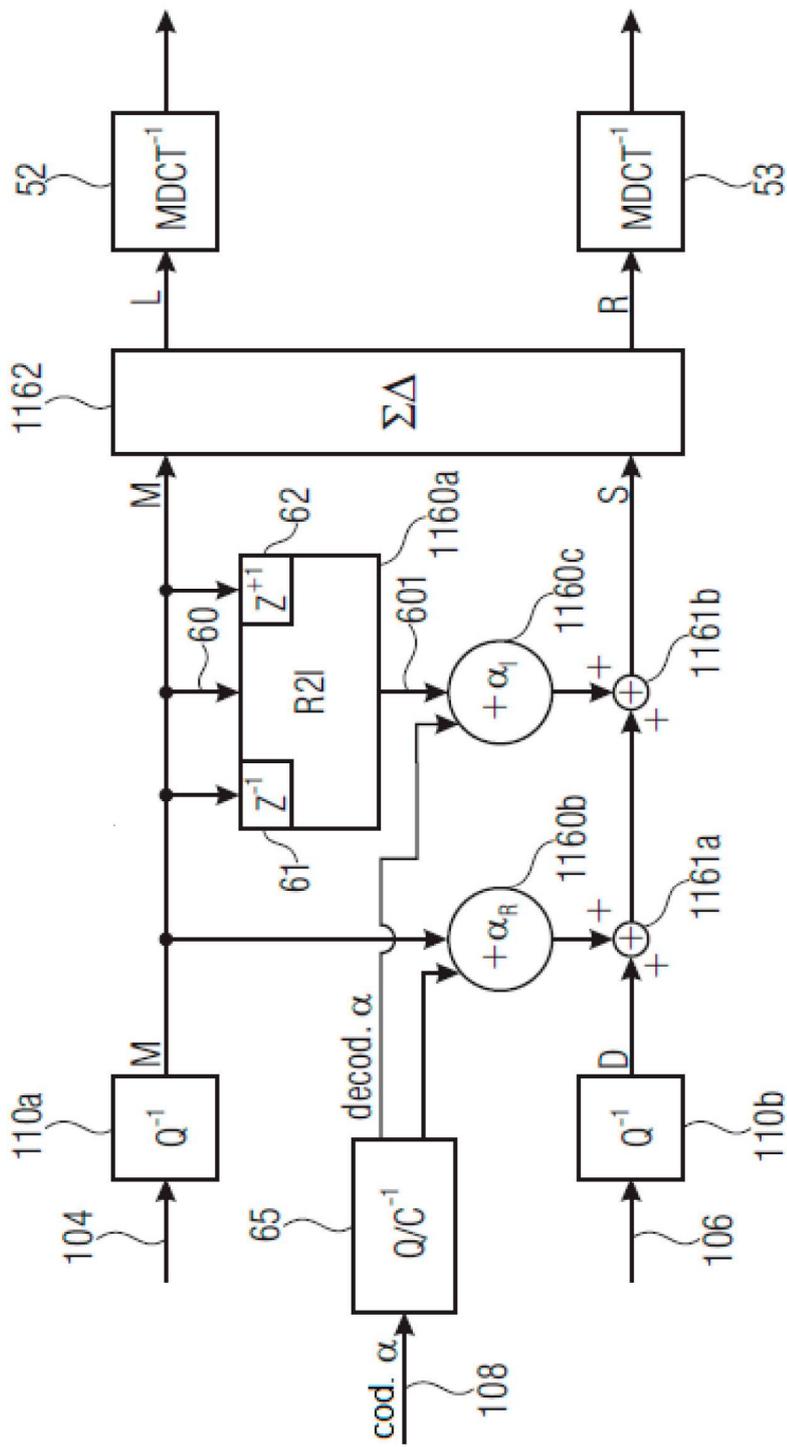


FIGURA 6B
(DECODIFICADOR LATERAL)

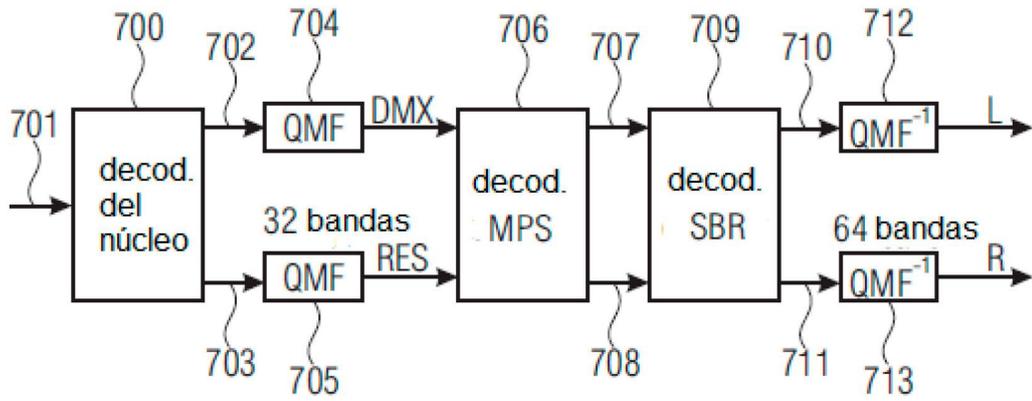


FIGURA 7A

$$\begin{bmatrix} L \\ R \end{bmatrix} = g \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} DMX \\ RES \end{bmatrix}$$

FIGURA 7B

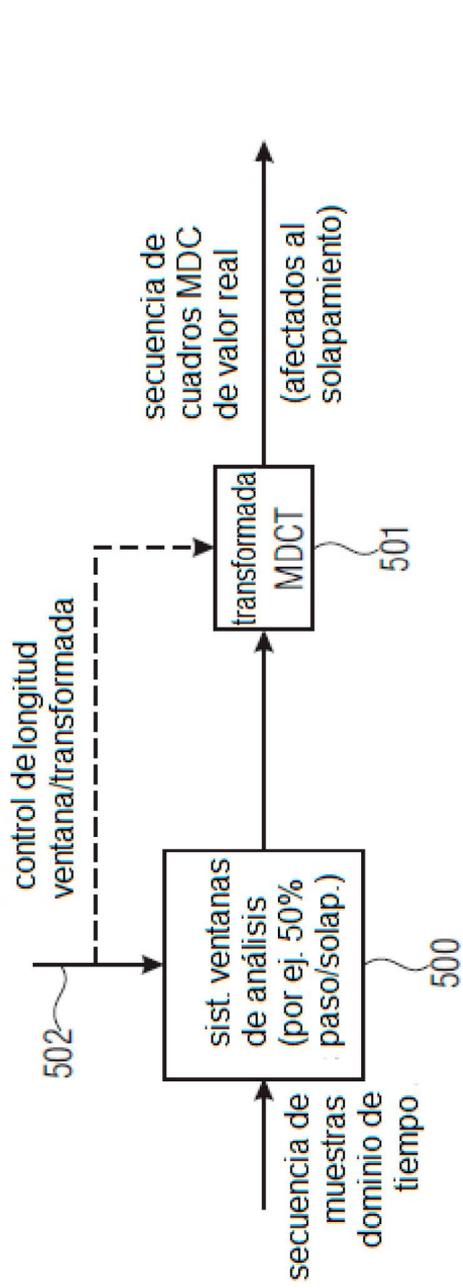


FIGURA 8A

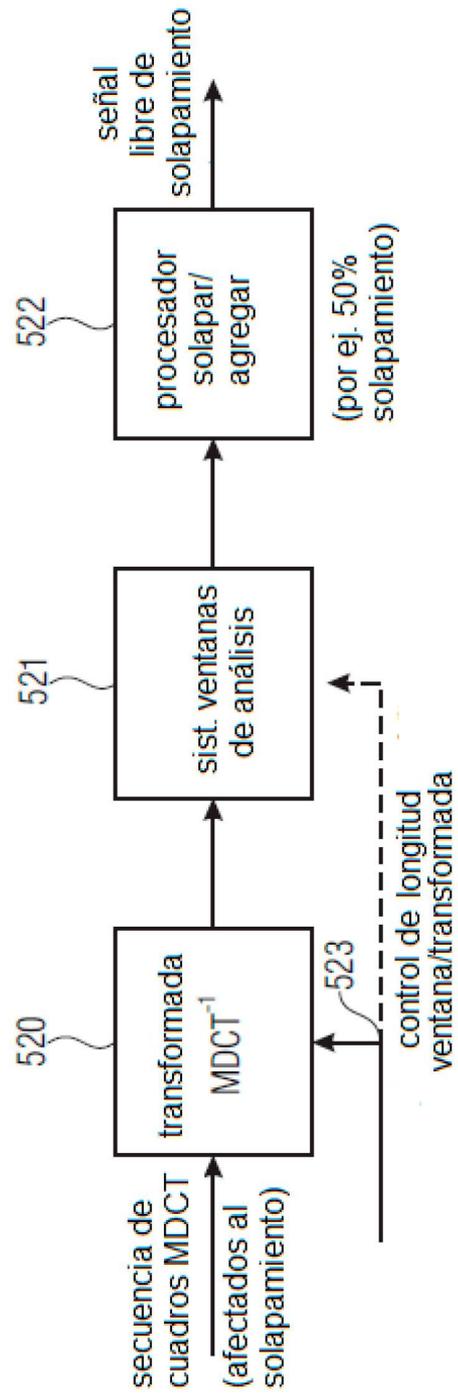
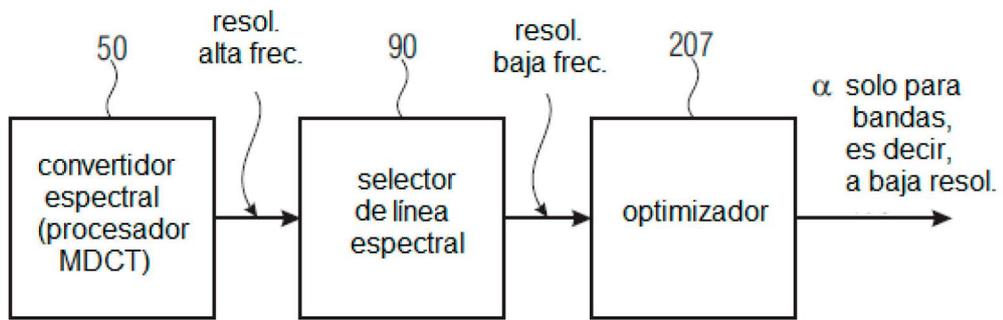


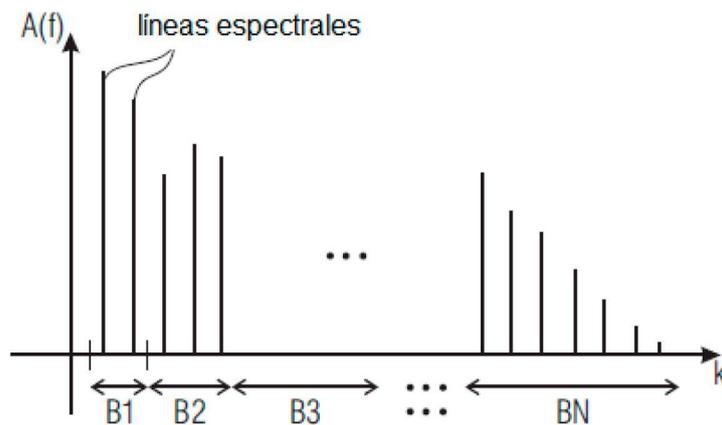
FIGURA 8B



0 : - banco de filtros con anchos de banda en aumento

- o bandas de frecuencia de igual tamaño que tienen cada una por lo menos dos líneas espectrales

FIGURA 9A



- solo se calcula una única α para cada banda
(por ej. - 1024 líneas espectrales
- menos de 30 valores α complejos (por ej. 22))
- B_i aumento con frecuencia

FIGURA 9B

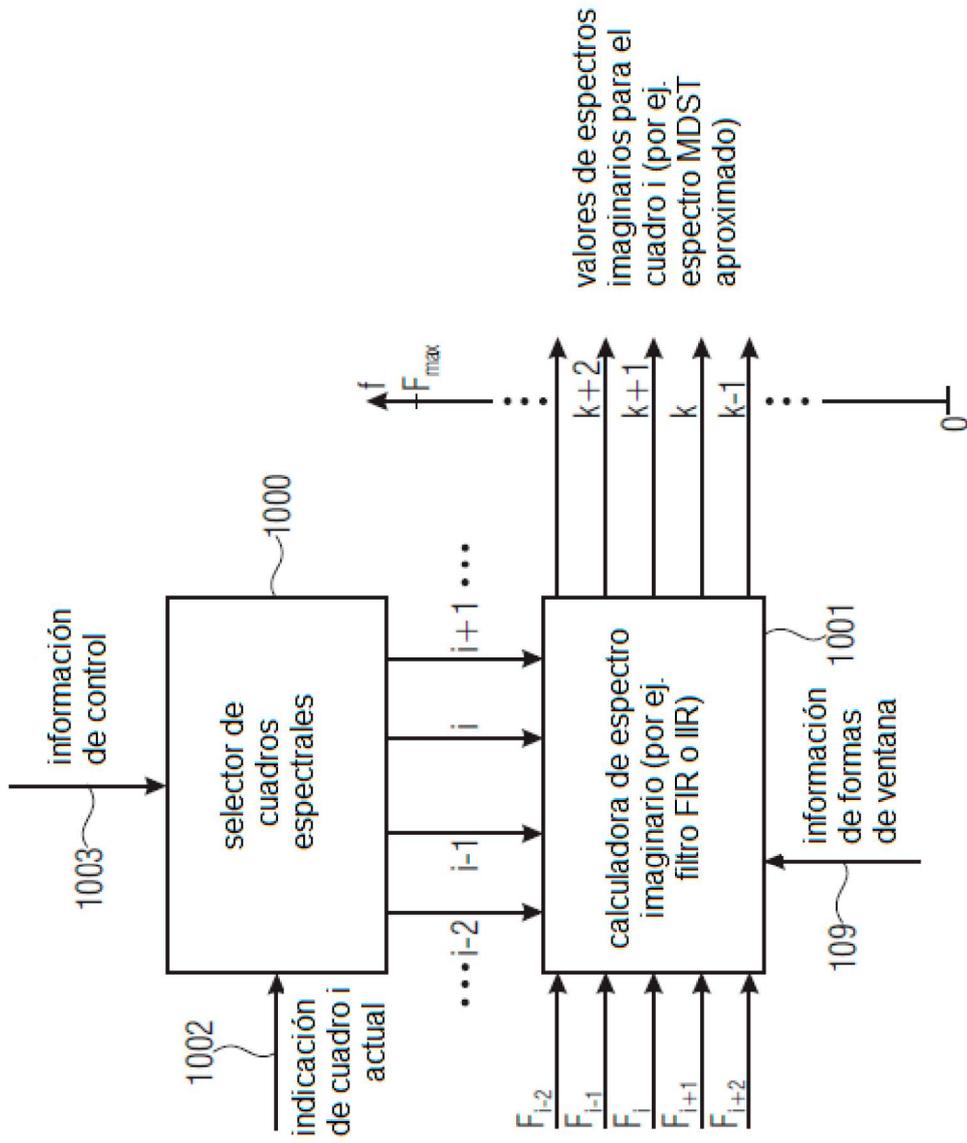


FIGURA 10A

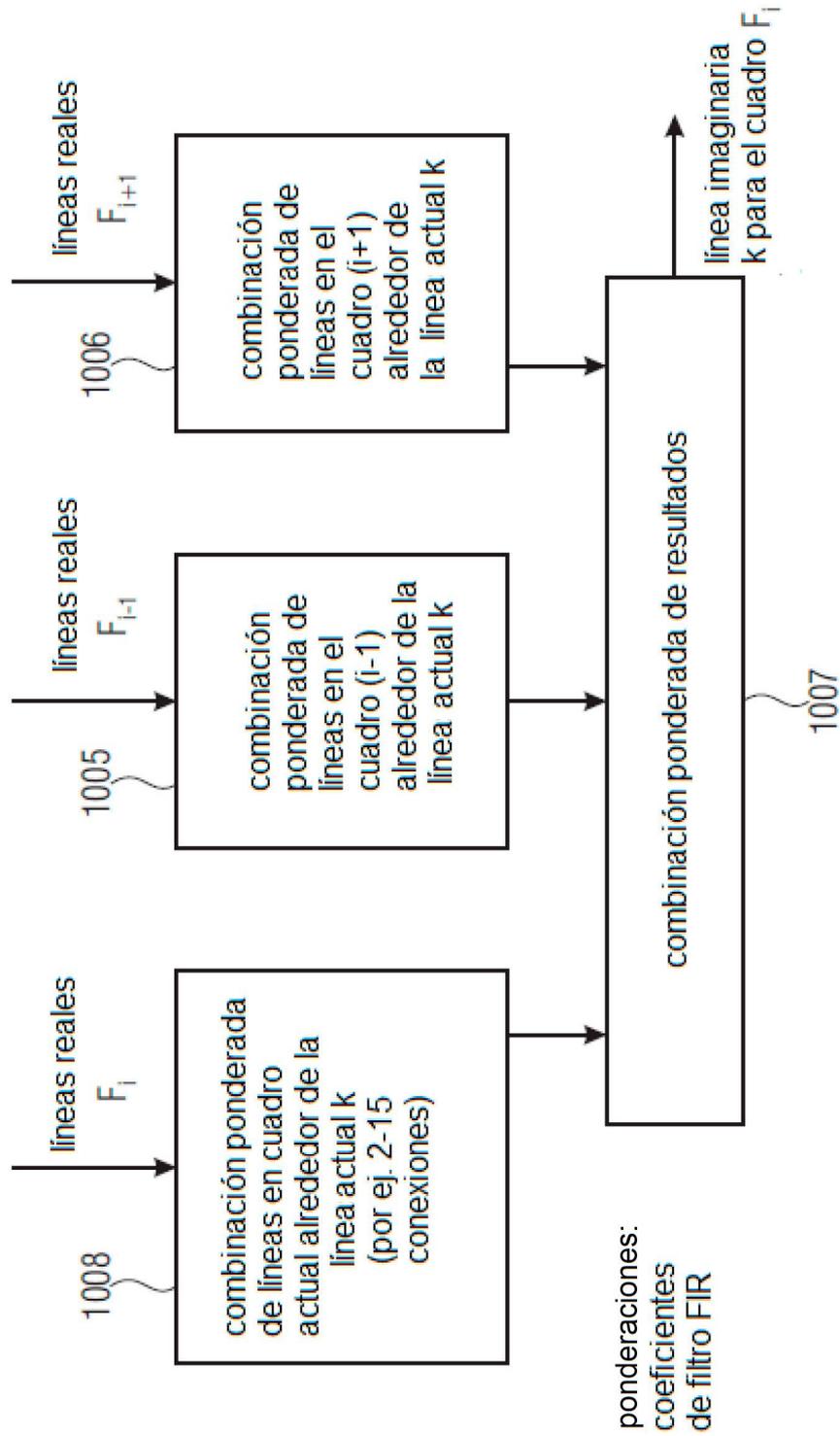


FIGURA 10B