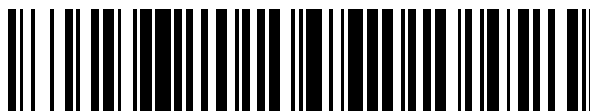


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 552 839**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/04 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.03.2011** **E 11709735 (2)**

97 Fecha y número de publicación de la concesión europea: **26.08.2015** **EP 2543038**

54 Título: **Decodificación de señales de audio multicanal usando predicción compleja**

30 Prioridad:

13.07.2010 US 363906 P

13.07.2010 EP 10169432

09.04.2010 US 322688 P

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

02.12.2015

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;

VILLEMÖES, LARS;

ROBILLARD, JULIEN;

NEUSINGER, MATTHIAS;

HELMRICH, CHRISTIAN;

HILPERT, JOHANNES;

RETTELBACH, NIKOLAUS;

DISCH, SASCHA y

EDLER, BERND

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 552 839 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).

Decodificación de señales de audio multicanal usando predicción compleja

DESCRIPCIÓN

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.

En el campo del procesamiento de estéreo o multicanal se sabe aplicar la así denominada codificación de estéreo central/lateral (*mid/side stereo coding*). De acuerdo con 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 logra una ganancia de codificación significativa, cuando la señal izquierda y la señal derecha son muy similares entre sí, debido a 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 rango de valores a cuantificar/codificar por entropía es menor. De este modo, para una PCM (modulación por codificación de pulsos, por 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 cuales la codificación central/lateral no logrará una ganancia de codificación. La situación puede ocurrir cuando las señales en ambos canales están desfasadas entre sí, por ejemplo, en 90°. Entonces, la señal central y la señal lateral pueden estar en un rango bastante similar y, por lo tanto, la codificación de la señal central y la señal lateral utilizando el codificador por entropía no logrará una ganancia de codificación e incluso puede producir un aumento de tasa de bits. Por lo tanto se puede aplicar una codificación central/lateral selectiva de frecuencia a fin de desactivar la codificación central/lateral en bandas, en donde la señal lateral no se vuelve más pequeña hasta un cierto grado con respecto a, por ejemplo, la señal izquierda original.

A pesar de que la señal lateral resultará cero, cuando las señales izquierda y derecha son idénticas, logrando 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 conformación 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 rango de valores. Cuando dicha situación ocurre 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 se puede aplicar en forma selectiva de frecuencia o se puede aplicar alternativamente en el dominio de tiempo.

Existen técnicas alternativas de codificación multicanal, las cuales no están basadas en un tipo de método de forma de onda como codificación central/lateral, pero 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, por sus siglas en inglés)". En este punto se calculan determinadas señales para una pluralidad de bandas de frecuencia. Dichas 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, pero está basada en las señales suministradas en forma selectiva de frecuencia 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.

Este método puede resultar complejo particularmente en el caso en el que el decodificador tiene que aplicar un procesamiento de descorrelación a fin de crear artificialmente señales estéreo, las cuales están descorrelacionadas entre sí, a pesar de que la totalidad de dichos canales proviene de uno y del mismo canal de mezcla descendente. Los descorreladores para este propósito 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, el cual inevitablemente produce 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 produce tasas de bits muy bajas pero puede incluir compromisos de calidad.

Existen desarrollos recientes en cuanto a la codificación de voz y audio unificada (USAC, por sus siglas en inglés) ilustrada en la Fig. 7a. Un decodificador del núcleo 700 lleva a cabo una operación de decodificación de la señal estéreo codificada en la entrada 701, la cual puede ser una señal codificada centro/lateral. El decodificador del núcleo 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, por sus siglas en inglés) por los bancos de filtros QMF

704 y 705. Luego se aplica un decodificador del Entorno de MPEG 706 para generar una señal izquierda de canal 707 y una señal derecha de canal 708. Estas señales de banda baja se introducen posteriormente en un decodificador de replicación espectral de ancho de banda (SBR, por sus siglas en inglés) 709, el cual produce señales izquierda y derecha de banda ancha en las líneas 710 y 711, las cuales luego se transforman a un dominio de tiempo por los bancos de filtros de síntesis QMF 712, 713 de modo tal que se obtienen las señales izquierda y derecha de banda ancha L, R.

La Fig. 7b ilustra una situación en la que el decodificador del Entorno de MPEG 706 llevaría a cabo una decodificación central/lateral. Alternativamente, el bloque decodificador del Entorno de MPEG 706 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 del núcleo. Naturalmente, el decodificador del Entorno de MPEG 706 también podría generar una pluralidad de señales de salida de banda baja para introducir en el bloque decodificador de SBR 709 utilizando información paramétrica tal como diferencias de nivel entre canales, medidas de coherencia entre canales u otros de dichos parámetros de información entre canales.

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

El uso de una combinación de un bloque 706 y un bloque 709 sólo genera 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, propuesta en el contexto de la USAC, produciría un aumento significativo en la complejidad computacional debido a los bancos de QMF necesarios, los cuales en este ejemplo necesitarían bancos de análisis de 64 bandas y bancos de síntesis de 64 bandas. Dichos bancos de filtros sólo deberían agregarse para los fines de la codificación estéreo.

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

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:

- 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);
- 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;
- Documento WO 2009/141775 A1 (KONINKL PHILIPS ELECTRONICS NV [NL]; SCHUIJERS ERIK G P [NL]) 26 de noviembre de 2009 (26-11-2009).

Un objetivo de la presente invención consiste en proporcionar un concepto mejorado de procesamiento de audio, el cual por una parte produce una alta ganancia de codificación y, por otra parte, produce a buena calidad de audio y/o complejidad computacional reducida.

Este objetivo se logra mediante un decodificador de audio de acuerdo con la reivindicación 1, un método de decodificación de audio de acuerdo con la reivindicación 11 o un programa informático de acuerdo con la reivindicación 12.

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 donde ambas señales de combinación provienen de las señales originales de canal utilizando una regla de combinación tal como la regla de combinación central/lateral. Se ha descubierto que esta información de predicción es calculada por un predictor en un codificador de audio de modo que se cumple un objetivo de optimización, se incurre sólo en una pequeña sobrecarga, pero logra una disminución significativa de la tasa de bits necesaria para la señal lateral sin perder ninguna calidad de audio, debido a que la predicción inventiva es sin embargo una codificación basada en forma de onda y no un método de codificación multicanal o estéreo basado en parámetros. A fin de reducir la complejidad computacional, se prefiere llevar a cabo una codificación de dominio de frecuencia, en donde 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 preferentemente 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), la cual es diferente de una transformada compleja ya que se calculan sólo valores reales o sólo valores imaginarios, en tanto que en una transformada compleja se calculan valores reales y complejos de un espectro logrando un sobremuestreo dos veces mayor.

Preferentemente se utiliza una transformada basada en introducción y cancelación de solapamiento. La MDCT, en particular, es una transformada de dichas características y permite una atenuación cruzada entre bloques subsiguientes sin ninguna sobrecarga debido a la propiedad muy conocida de cancelación de solapamiento de dominio de tiempo (TDAC, por sus siglas en inglés), la cual se obtiene mediante el procesamiento de solapar-agregar en el lado del decodificador.

Preferentemente, la información de predicción calculada en el codificador, transmitida al decodificador y utilizada en el decodificador comprende una parte imaginaria, la cual 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, ya sea un espectro real solamente o que proporciona un espectro imaginario solamente. A fin de utilizar dicha información de predicción imaginaria, la cual 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 a fin de calcular una señal residual de predicción a partir de la primera señal de combinación, la cual tiene la fase rotada con respecto a la señal de combinación original. Esta señal residual de predicción que tiene la fase rotada luego se puede combinar con la señal residual de predicción transmitida en la corriente de bits para regenerar una señal lateral, la cual finalmente se puede combinar con la señal central para obtener el canal izquierdo decodificado en una determinada banda y el canal derecho decodificado en esta banda.

A fin de aumentar la calidad de audio, el mismo convertidor real-a-imaginario o el convertidor imaginario-a-real, el cual se aplica en el lado del decodificador se implementa también en el lado del codificador, cuando la señal residual de predicción es calculada en el codificador.

La presente invención proporciona una ventaja porque mejora la calidad de audio y reduce 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 donde la SBR típicamente no se utiliza. En lugar de procesar la señal en el dominio complejo híbrido QMF, estos métodos implementan una codificación estéreo predictiva de base residual en el dominio nativo de MDCT del codificador de transformada estéreo subyacente.

De acuerdo con 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 donde la predicción compleja se lleva a cabo en el dominio de MDCT utilizando una transformada real-a-compleja, en donde dicha señal estéreo puede ser, ya sea una señal estéreo codificada en el lado del codificador o puede ser alternativamente una señal estéreo decodificada/transmitida, cuando el aparato o método para generar la señal estéreo se aplica en el lado del decodificador.

Las formas de realización preferidas de la presente invención se describen a continuación con respecto a los dibujos adjuntos, en los cuales:

La Fig. 1 es un diagrama de un ejemplo de un decodificador de audio;

La Fig. 2 es un diagrama en bloque de un ejemplo de un codificador de audio;

La Fig. 3a ilustra una implementación de la calculadora del codificador de la Fig. 2;

La Fig. 3b ilustra una implementación alternativa de la calculadora del codificador de la Fig. 2;

La Fig. 3c ilustra una regla de combinación central/lateral para aplicar en el lado del codificador;

La Fig. 4a ilustra una implementación de la calculadora del decodificador de la Fig. 1;

La Fig. 4b ilustra una implementación alternativa de la calculadora del decodificador en forma de una calculadora matriz;

- La Fig. 4c ilustra una regla de combinación inversa central/lateral correspondiente a la regla de combinación ilustrada en la Fig. 3c;
- 5 La Fig. 5a ilustra un ejemplo de un codificador de audio que opera en el dominio de frecuencia, el cual es preferentemente un dominio de frecuencia de valor real;
- La Fig. 5b ilustra una implementación de un decodificador de audio que opera en el dominio de frecuencia;
- 10 La Fig. 6a ilustra una implementación alternativa de un codificador de audio que opera en el dominio de MDCT y utiliza una transformada real-a-imaginaria;
- La Fig. 6b ilustra una realización de un decodificador de audio que opera en el dominio de MDCT y utiliza una transformada real-a-imaginaria;
- 15 La Fig. 7a ilustra un post-procesador de audio que utiliza un decodificador estéreo y un decodificador de SBR conectado posteriormente;
- La Fig. 7b ilustra una matriz de mezcla ascendente central/lateral;
- 20 La Fig. 8a ilustra una vista detallada en el bloque de MDCT en la Fig. 6a;
- La Fig. 8b ilustra una vista detallada en el bloque de MDCT⁻¹ de la Fig. 6b;
- 25 La Fig. 9a ilustra una implementación de un optimizador que opera en resolución reducida con respecto a la salida de MDCT;
- La Fig. 9b ilustra una representación de un espectro de MDCT y las bandas correspondientes de resolución inferior, en las cuales se calcula la información de predicción;
- 30 La Fig. 10a ilustra una implementación del transformador real-a-imaginario en la Fig. 6a o en la Fig. 6b; y
- La Fig. 10b ilustra una implementación posible de la calculadora de espectro imaginario de la Fig. 10a.
- La Fig. 1 ilustra un decodificador de audio para decodificar una señal de audio multicanal codificada y obtenida en una línea de entrada 100. La señal de audio multicanal codificada comprende una primera señal de combinación codificada que se genera 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, la cual tiene los 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 de entrada 102. La interfaz de entrada 102 se puede implementar como un desmultiplexor de corriente de datos, el cual 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. Preferentemente, 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 de señales 110 para decodificar la primera señal de combinación para obtener una primera señal de combinación decodificada en la línea 112. Adicionalmente, el decodificador de señales 110 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 del decodificador 115, la cual emite la primera señal de canal en la línea 117 y una segunda señal de canal en la línea 118, en donde 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. A fin de codificar totalmente dicha señal multicanal que tiene cinco canales se pueden aplicar dos decodificadores ilustrados en la Fig. 1, en donde 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 se pueden aplicar 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 una 1ª y 2ª señal utilizando dos coeficientes de predicción, en forma muy similar al módulo "dos-a-tres" en el Entorno de MPEG.

La calculadora del decodificador 116 está configurada para calcular una señal multicanal decodificada que tiene la primera señal de canal decodificada 117 y la segunda señal de canal decodificada 118 utilizando la señal residual decodificada 114, la información de predicción 108 y la primera señal de combinación decodificada 112. Particularmente, la calculadora del decodificador 116 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, las cuales están combinadas por la regla de combinación cuando se genera la primera señal de combinación y la señal residual de predicción. Específicamente, la información de predicción en la línea 108 comprende una parte de valor real distinta de cero y/o una parte imaginaria distinta de cero.

La calculadora del decodificador 116 se puede implementar de diversos modos. Una primera implementación está ilustrada en la Fig. 4a. Esta implementación comprende un predictor 1160, una calculadora de señales de combinación 1161 y un combinador 1162. El predictor recibe la primera señal de combinación decodificada 112 y la información de predicción 108 y emite una señal de predicción 1163. Específicamente, el predictor 1160 está configurado para aplicar la información de predicción 108 a la primera señal de combinación decodificada 112 o una señal proveniente de la primera señal de combinación decodificada. La regla de derivación para derivar la señal a la cual se aplica la información de predicción 108 puede ser una transformada real-a-imaginaria, o por igual, una transformada imaginaria-a-real o una operación de ponderación o, dependiendo de la implementación, una operación de desfase o una operación combinada de ponderación/desfase. La señal de predicción 1163 se introduce junto con la señal residual decodificada en la calculadora de señales de combinación 1161 a fin de calcular la segunda señal de combinación decodificada 1165. Ambas señales 112 y 1165 se introducen en el combinador 1162, el cual 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 de salida 1166 y 1167, respectivamente. Alternativamente, la calculadora del decodificador se implementa como una calculadora matriz 1168, la cual recibe, como entrada, la primera señal de combinación decodificada o señal M, la señal residual decodificada o señal D y la información de predicción a 108. La calculadora matriz 1168 aplica una matriz de transformada ilustrada como 1169 a las señales M, D para obtener las señales de salida L, R, en donde L es la primera señal de canal decodificada y R es la segunda señal de canal decodificada. La notación en la Fig. 4b se asemeja a una notación estéreo con un canal izquierdo L y un canal derecho R. Se ha aplicado esta notación a fin de proporcionar una explicación más simple 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 matriz 1169 unifica las operaciones en los bloques 1160, 1161 y 1162 de la Fig. 4a en un tipo de cálculo matriz de "disparo único", y las entradas en el circuito de la Fig. 4a y las salidas desde el circuito de la Fig. 4a son idénticas a las entradas en la calculadora matriz 1168 o a las salidas desde la calculadora matriz 1168.

La Fig. 4c ilustra un ejemplo para una regla de combinación inversa aplicada por el combinador 1162 en la Fig. 4a. Particularmente, la regla de combinación es similar a la regla de combinación del lado del decodificador en la codificación central/lateral bien conocida, en donde $L = M + S$, y $R = M - S$. Se entiende que la señal S utilizada por la regla de combinación inversa en la Fig. 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. Se entiende 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 números de referencia como tales, los cuales se han asignado a las líneas. Por lo tanto, la notación es tal que una línea que tiene una determinada señal está indicando la señal propiamente dicha. Una línea puede ser una línea física en una implementación físicamente conectada. En una implementación computarizada, sin embargo, una línea física no existe, pero la señal representada por la línea es transmitida desde un módulo de cálculo hacia el otro módulo de cálculo.

La Fig. 2 ilustra un codificador de audio para codificar una señal de audio multicanal 200 que tiene dos o más señales de canal, en donde una primera señal de canal está ilustrada en 201 y un segundo canal está ilustrado en 202. Ambas señales se introducen en una calculadora del codificador 203 para calcular una primera señal de combinación 204 y una señal residual de predicción 205 utilizando la primera señal de canal 201 y la segunda señal de canal 202 y la información de predicción 206, de modo que cuando se combina la señal residual de predicción 205 con una señal de predicción proveniente de la primera señal de combinación 204 y la información de predicción 206 se obtiene una segunda señal de combinación, en donde la primera señal de combinación y la segunda señal de combinación pueden obtenerse de la primera señal de canal 201 y la segunda señal de canal 202 utilizando una regla de combinación.

La información de predicción es generada por un optimizador 207 para calcular la información de predicción 206 de modo que la señal residual de predicción cumple un objetivo de optimización 208. La primera señal de combinación 204 y la señal residual 205 se introducen en un codificador de señales 209 para codificar la primera señal de combinación 204 para obtener una primera señal de combinación codificada 210 y para codificar la señal residual 205 para obtener una señal residual codificada 211. Ambas señales codificadas 210, 211 se introducen en una interfaz de salida 212 para combinar la primera señal de combinación codificada 210 con la señal residual de

predicción codificada 211 y la información de predicción 206 para obtener una señal multicanal codificada 213, la cual es similar a la señal multicanal codificada 100 introducida en la interfaz de entrada 102 del decodificador de audio ilustrado en la Fig. 1.

5 Dependiendo de la implementación, el optimizador 207 recibe, ya sea la primera señal de canal 201 y la segunda señal de canal 202 o, tal como se ilustra en las líneas 214 y 215, la primera señal de combinación 214 y la segunda señal de combinación 215 provenientes de un combinador 2031 de la Fig. 3a, las cuales se analizarán posteriormente.

10 Un objetivo preferido de optimización se ilustra en la Fig. 2, en la cual 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 tiene como resultado una solución para α que se ilustra la Fig. 2. Las señales S, M están proporcionadas en forma de bloque y son preferentemente señales espectrales de dominio, en donde la notación $\|\dots\|$ significa la norma 2 del argumento, y en donde $\langle \dots \rangle$ ilustra el producto de puntos en forma habitual. Cuando la primera señal de canal 201 y la segunda señal de canal 202 se introducen en el optimizador 207, el optimizador tendría entonces que aplicar la regla de combinación, en donde una regla de combinación ejemplar está ilustrada en la Fig. 3c. Sin embargo, cuando la primera señal de combinación 214 y la segunda señal de combinación 215 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 proveniente 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 a 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 dicho 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 se pueden aplicar para controlar el optimizador 207, el cual calcula la información de predicción 206.

35 La calculadora del codificador 203 en la Fig. 2 se puede implementar en distintas formas, en donde una primera implementación ejemplar está ilustrada en la Fig. 3a, en la cual una regla de combinación explícita se lleva a cabo en el combinador 2031. Una implementación alternativa ejemplar está ilustrada en la Fig. 3b, en donde se utiliza una calculadora matriz 2039. El combinador 2031 en la Fig. 3a se puede implementar para llevar a cabo la regla de combinación ilustrada en la Fig. 3c, la cual es de manera ejemplar la regla de codificación central/lateral bien conocida, en donde un factor de ponderación de 0,5 se aplica a todos los ramales. Sin embargo se pueden implementar otros factores de ponderación o ningún factor de ponderación dependiendo de la implementación. Adicionalmente, cabe destacar que se pueden aplicar 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, la cual se puede aplicar en el combinador del decodificador 1162 ilustrado en la Fig. 4a, el cual aplica una regla de combinación que es inversa a la regla de combinación aplicada por el codificador. Debido a la predicción inventiva se puede utilizar cualquier regla de predicción invertible, debido a que la influencia en la forma de onda está "equilibrada" por la predicción, es decir, cualquier error está incluido en la señal residual transmitida, debido a que la operación de predicción llevada a cabo por el optimizador 207 en combinación con la calculadora del codificador 203 es un proceso que conserva la forma de onda.

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

La calculadora residual 2034 en la Fig. 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 de señales de combinación 1161 en la Fig. 4a puede llevar a cabo una operación de adición en donde la señal residual decodificada 114 y la señal de predicción 1163 se agregan juntas para obtener la segunda señal de combinación 1165.

La Fig. 5a ilustra una implementación preferida de un codificador de audio. En comparación con el codificador de audio ilustrado en la Fig. 3a, la primera señal de canal 201 es una representación espectral de una primera señal de canal de dominio de tiempo 55a. De la misma manera, la segunda señal de canal 202 es una representación espectral de una señal de canal de dominio de tiempo 55b. La conversión desde el dominio de tiempo en la representación espectral es llevada a cabo por un convertidor de tiempo/frecuencia 50 para la primera señal de canal y un convertidor de tiempo/frecuencia 51 para la segunda señal de canal. Preferentemente, pero no necesariamente, los convertidores espectrales 50, 51 están implementados como convertidores de valor real. El algoritmo de conversión puede ser una transformada coseno discreta, una transformada FFT, en donde sólo 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 donde sólo se utiliza la parte imaginaria y se descarta la parte real. También se puede utilizar cualquier otra transformada que proporcione sólo valores imaginarios. Un motivo por el cual utilizar una transformada pura de valor real o una transformada pura imaginaria es la complejidad computacional, debido a que, para cada uno de los valores espectrales, sólo 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 una de las líneas espectrales, lo cual es un aumento de la complejidad computacional por un factor de por lo menos 2. En este punto, otro motivo para utilizar una transformada de valor real es que dicha transformada 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 estándar de "codificación de audio perceptual" implementado en "MP3", AAC, o en sistemas de codificación de audio similares).

La Fig. 5a ilustra adicionalmente la calculadora residual 2034 como un sumador, el cual recibe la señal lateral como su entrada "más" y el cual recibe la salida de la señal de predicción por el predictor 2033 como su entrada "menos". Adicionalmente, la Fig. 5a ilustra la situación en la que la información de control del predictor se transmite desde el optimizador hacia el multiplexor 212, el cual emite una corriente de datos 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 es previsible desde la señal central tal como se ilustra en las Ecuaciones a la derecha de la Fig. 5a.

Preferentemente, la información de control del predictor 206 es un factor tal como se ilustra a la derecha en la Fig. 3b. En una forma de realización en la cual la información de control de predicción sólo comprende una porción real tal como la parte real de una a de valor complejo o una magnitud de a de valor complejo, en donde esta porción corresponde a un factor distinto de cero, se puede obtener 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 sólo comprende una segunda porción, la cual puede ser la parte imaginaria de un factor de valor complejo o la información de fase del factor de valor complejo, en donde la parte imaginaria o la información de fase es distinta de cero, se logra una ganancia de codificación significativa para las señales, las cuales están desfasadas entre sí por un valor distinto de 0° o 180° , y que tienen además del desfase, características de forma de onda similares y relaciones de amplitud similares.

Preferentemente, una información de control de predicción es de valor complejo. Por lo tanto se puede obtener una ganancia de codificación significativa para las señales que tienen distinta amplitud y que están desfasadas. En una situación en la cual las transformadas de tiempo/frecuencia proporcionan espectros complejos, la operación 2034 sería una operación compleja en la cual 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 forma de 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 Fig. 5b ilustra un decodificador de audio correspondiente al codificador de audio ilustrado en la Fig. 5a. Los elementos similares con respecto al decodificador de audio de la Fig. 1 tienen números de referencia similares.

La salida de la corriente de bits por el multiplexor de la corriente de bits 212 en la Fig. 5a se introduce en un desmultiplexor de la corriente de bits 102 en la Fig. 5b. El desmultiplexor de la corriente de bits 102 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 desmultiplexor de la corriente de bits 102 desmultiplexa una información de control del predictor 108 desde la corriente de bits e introduce la misma en el predictor 1160. El predictor 1160 emite una señal lateral previsible $\alpha \cdot M$ y el combinador 1161 combina la señal residual emitida por el descuantificador 110b con la señal lateral previsible a fin de obtener finalmente la señal lateral reconstruida S. La señal luego es introducida en el combinador 1162, el cual lleva a cabo, por ejemplo, un procesamiento de suma/diferencia, tal como se ilustra en la Fig. 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 luego se convierte en una representación del dominio de tiempo por los convertidores de frecuencia/tiempo 52 y 53 correspondientes.

Dependiendo de la implementación del sistema, los convertidores de frecuencia/tiempo 52, 53 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 Fig. 6a para el codificador y en la Fig. 6b para el decodificador. Las transformadas de valor real 50 y 51 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. Debido a que ambos espectros M, S son espectros de valor real y debido a que, por lo tanto, no existe ninguna parte imaginaria del espectro, se proporciona un convertidor real-a-imaginario 2070, el cual calcula un espectro imaginario estimado 600 a partir del el espectro de valor real de la señal M. Este transformador real-a-imaginario 2070 es una parte del optimizador 207, y el espectro imaginario 600 estimado por el bloque 2070 se introduce en la etapa del optimizador de a 2071 junto con el espectro real M a fin de calcular la información de predicción 206, la cual ahora tiene un factor de valor real indicado en 2073 y un factor imaginario indicado en 2074. Aquí, de acuerdo con este ejemplo, el espectro de valor real de la primera señal de combinación M se multiplica por la parte real α_R 2073 para obtener la señal de predicción, la cual luego se sustrae del espectro lateral de valor real. Adicionalmente, el espectro imaginario 600 se multiplica por la parte imaginaria α_I ilustrada en 2074 para obtener la señal de predicción adicional, en donde esta señal de predicción luego se sustrae 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, en tanto 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 a en el cuantificador/codificador de entropía 2072 para obtener el valor de a complejo codificado, el cual se transmite al multiplexor de la corriente de bits 212 de la Fig. 5a, por ejemplo, y el cual se introduce finalmente en una corriente de bits como la información de predicción.

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

A pesar de que se podría calcular 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 Fig. 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 de ponderación 1160b y 1160c se agregan a la señal residual de predicción decodificada y descuantificada. Particularmente, los valores espectrales introducidos en el ponderador 1160c, en donde 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 real-a-imaginario 1160a, el cual se implementa preferentemente del mismo modo que el bloque 2070 de la Fig. 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 sólo los espectros de valor real codificados han sido transmitidos desde el codificador hacia el decodificador debido a las tasas de bits y por razones de complejidad.

El transformador real-a-imaginario 1160a o el bloque correspondiente 2070 de la Fig. 6a se puede implementar tal como se publica en el documento WO 2004/013839 A1 o en el documento WO 2008/014853 A1 o en la Patente de Estados Unidos N° 6.980.933. Alternativamente se puede aplicar cualquier otra implementación conocida en la técnica, y una implementación preferida se describe en el contexto de las Figs. 10a, 10b.

Específicamente, tal como se ilustra en la Fig. 10a, el convertidor real-a-imaginario 1160a comprende un selector de cuadro espectral 1000 conectado a una calculadora del espectro imaginario 1001. El selector de cuadro espectral 1000 recibe una indicación de un cuadro actual i en la entrada 1002 y, dependiendo de la implementación, información de control en una entrada de control 1003. 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 de control 1003 indica que sólo se tiene que utilizar el cuadro actual para dicho cálculo, entonces el selector de cuadro espectral 1000 sólo selecciona el cuadro actual i y transmite esta información a la calculadora de espectro imaginario. Por lo tanto, la calculadora de espectro imaginario sólo 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 cual se tiene que calcular una línea imaginaria tal como se ilustra en 1004 en la Fig. 10b. Sin embargo, cuando el selector de cuadro espectral 1000 recibe una información de control 1003 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 provenientes 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 el cual 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, la cual luego 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, si bien se realiza una sustracción en el elemento 2034b.

Cabe destacar que la información de control 1003 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" a fin de reducir el retardo sistemático.

Adicionalmente, cabe destacar que la combinación ponderada en etapas ilustrada en la Fig. 10b, en la cual, en una primera operación se combinan las líneas provenientes 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 provenientes de un número de cuadros adyacentes indicados por la información de control 103 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 a utilizar para estimar la línea imaginaria. Entonces, los resultados de dichas combinaciones "en tiempo" se someten a una combinación ponderada en la "dirección de frecuencia" para obtener finalmente la línea imaginaria k para el cuadro f_i . Los valores ponderados se fijan en un valor -1 y 1 preferentemente, y las ponderaciones se pueden implementar en una combinación directa de filtro FIR o IIR, la cual 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 Figs. 6a y 6b, el algoritmo de transformada preferido es el algoritmo de transformada MDCT, el cual se aplica en la dirección anterior en los elementos 50 y 51 en la Fig. 6a y el cual se aplica en la dirección posterior en los elementos 52, 53, a continuación de una operación de combinación en el combinador 1162 que opera en el dominio espectral.

La Fig. 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 de ventanas de análisis 500, el cual lleva a cabo una operación de ventanas que utiliza una ventana de análisis y, particularmente, lleva a cabo esta operación en un cuadro mediante un modo de cuadros 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 de transformada MDCT 501, el cual emite la secuencia de cuadros de MDCT de valor real, en donde 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 de transformada MDCT 501 emite espectros de MDCT que tienen 1024 líneas espectrales reales o valores MDCT. Preferentemente, el sistema de ventanas de análisis 500 y/o el transformador de MDCT 501 se pueden controlar mediante un control de longitud de ventana o longitud de transformada 502 de modo que, por ejemplo, para las porciones transitorias en la señal, la longitud de ventana/longitud de transformada se reduce a fin de obtener mejores resultados de codificación.

La Fig. 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. Dicho cuadro se suministra a un

sistema de ventanas de síntesis 521, el cual aplica una venta de síntesis a este cuadro de 2048 muestras. El cuadro dividido en ventanas luego se transmite a un procesador de solapar/agregar 522, el cual, a modo de ejemplo, aplica un 50 % de solapamiento entre dos cuadros subsiguientes y, entonces, lleva a cabo una adición de muestra por muestra de modo que un bloque de 2048 muestras finalmente logra 1024 muestras nuevas de la señal de salida libre de solapamiento. Nuevamente se prefiere aplicar un control de longitud de ventana/transformada utilizando información, la cual por ejemplo se transmite en la información lateral de la señal multicanal codificada según lo indicado en 523.

Los valores de predicción de a se podrían calcular para cada una de las líneas espectrales individuales 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 espectral 50 ilustrado en la Fig. 9, el cual es por ejemplo un procesador de MDCT descrito en el contexto de la Fig. 8a, proporciona un espectro de resolución de alta frecuencia que tiene determinadas líneas espectrales ilustradas en la Fig. 9b. Este espectro de resolución de alta frecuencia es utilizado por un selector de línea espectral 90 que proporciona un espectro de resolución de baja frecuencia, el cual comprende determinadas bandas B1, B2, B3, ..., BN. Este espectro de resolución de baja frecuencia es transmitido al optimizador 207 para calcular la información de predicción de modo que no se calcula una información de predicción para cada una de las líneas espectrales, sino que solamente para cada una de las bandas. 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.

Preferentemente, las bandas están configuradas en una forma psicoacústica de modo que el ancho de banda de las bandas aumenta desde frecuencias inferiores hasta frecuencias superiores tal como se ilustra en la Fig. 9b. Alternativamente, a pesar de que no se prefiere tanto como la implementación de ancho de banda en aumento también se podrían utilizar bandas de frecuencia de igual tamaño, en donde cada una de las bandas de frecuencia tiene por lo 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 a complejos, y preferentemente, más de 5 valores de a . Para los espectros con menos de 1024 líneas espectrales (por ejemplo, 128 líneas) se utilizan preferentemente menos bandas de frecuencia (por ejemplo, 6) para a .

Para calcular los valores de a no se requiere necesariamente el espectro de MDCT de alta resolución. Alternativamente también se puede utilizar un banco de filtros que tiene una resolución de frecuencia similar a la resolución necesaria para calcular los valores de a . Cuando se tienen que implementar bandas que aumentan de frecuencia, en ese caso dicho 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 se puede utilizar un banco de filtros tradicional con sub-bandas de ancho equivalente.

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

Preferentemente, los módulos 2070 en el codificador y 1160a en el decodificador deberían coincidir exactamente a fin de asegurar una correcta codificación de forma de onda. Esto se aplica preferentemente al caso en el cual dichos 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 Fig. 6a utilice el espectro de MDCT no cuantificado M como entrada, a pesar de que el módulo real-a-imaginario (R21) 1160a en el decodificador tiene sólo el espectro de MDCT cuantificado disponible como entrada. Alternativamente se puede utilizar también una implementación en la cual 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 formas de realización de la presente

invención.

La codificación de estéreo paramétrico estándar se basa en la capacidad del dominio de QMF (híbrido) complejo sobremuestreado de permitir un procesamiento de señales 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 el presente contexto), el codificador de estéreo unificado resultante actúa como un codificador de forma de onda. Esto permite la operación en un dominio críticamente muestreado, tal como el dominio de MDCT, debido a que el paradigma de codificación de forma de onda asegura que la propiedad de cancelación de solapamiento de la cadena de procesamiento de MDCT-IMDCT se conserve suficientemente bien.

Sin embargo, a 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 se puede obtener 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.

Asimismo, la parametrización de la matriz de mezcla ascendente se puede simplificar transmitiendo el coeficiente de predicción compleja α en lugar de los parámetros MPS. De este modo se transmiten sólo 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 Fig. 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 Fig. 4b.

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 debido a 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 de acuerdo con las propiedades de enmascaramiento instantáneo de la señal a expensas de una ganancia de codificación un tanto reducida.

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 como se ha analizado. Los últimos se pueden considerar como una transformada "real-a-imaginaria" (R2I). La complejidad del cómputo de frecuencia-dominio de la MDST se puede reducir de diversas formas, lo que significa que se calcula sólo una aproximación del espectro de MDST:

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

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

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 sólo del solapamiento de la mitad de ventana con el cuadro actual, es decir, para éstos es necesario un conjunto de coeficientes sólo para cada tipo de ventana (sin coeficientes adicionales para las transiciones).

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, anterior/siguiente, el filtrado bidimensional resulta más complicado. A 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.

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 donde hay menos coeficientes de MDCT circundantes para el filtrado de FIR que los necesarios. Aquí, el proceso de filtrado necesita adaptarse para computar el espectro de MDST correctamente. Esto se puede llevar a cabo, ya sea utilizando una extensión simétrica del espectro de MDCT para los coeficientes que faltan (de acuerdo con la periodicidad de los espectros de señales discretas de tiempo), o adaptando los coeficientes de filtro en forma correspondiente. Es obvio que el tratamiento de dichos casos especiales se puede simplificar a expensas de una menor exactitud en la proximidad de los límites del espectro de MDST.

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 se puede evitar utilizando una aproximación del espectro de MDST que no requiere el espectro de MDCT del cuadro siguiente como entrada.

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:

- Sólo un pequeño aumento en la complejidad computacional (cuando no se utiliza SBR).
- Amplía la capacidad hasta una reconstrucción perfecta si los espectros de MDCT no están cuantificados. Cabe destacar que esto no ocurre para la codificación estéreo unificada basada en QMF.
- Extensión natural de la codificación M/S e intensidad de la codificación estéreo.
- Arquitectura de limpieza que simplifica el ajuste del codificador, debido a que el procesamiento de la señal estéreo y la cuantificación/codificación se pueden acoplar fuertemente. Cabe destacar 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.
- Optimiza la codificación de parámetros de estéreo, debido a que sólo se tienen que transmitir dos parámetros (complejo α) en lugar de tres parámetros como en el Entorno de MPEG (ICC, CLD, IPD).
- No presenta un retardo adicional en el decodificador si el espectro de MDST se computa como una aproximación (sin utilizar el cuadro siguiente).

Las propiedades importantes de una implementación se pueden resumir 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, dicho 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.

b) Sólo 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 en forma independiente, se limitan al rango $[-3,0, 3,0]$ 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 en forma diferencial al tiempo o en 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 produce una resolución de frecuencia similar a la resolución de frecuencia del Entorno de MPEG. Este esquema de cuantificación y codificación tiene 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.

Los detalles preferidos, adicionales o alternativos de la implementación comprenden:

c) Para cada uno de los dos parámetros de α se puede elegir 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, ya sea una codificación diferencial de tiempo o frecuencia. Nuevamente, esto se puede señalar utilizando una bandera de un bit.

d) En lugar de reutilizar un código predefinido tal como el código de factor de escala AAC se puede utilizar 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.

e) El rango de los valores del parámetro α así como también el tamaño del paso de cuantificación de parámetros se pueden elegir arbitrariamente y optimizar hasta las características de señales en cuestión.

f) El número y ancho espectral y/o temporal de las bandas de parámetro activo α se pueden elegir arbitrariamente y optimizar hasta las características de señales determinadas. En particular, la configuración de banda se puede señalar por cuadro o por corriente.

g) Además de, o en lugar de los mecanismos indicados en a) anteriormente, se puede señalar explícitamente por medio de un bit por cuadro en la corriente de bits que sólo 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.

Las formas de realización 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 donde no se utiliza SBR) sin el aumento significativo en la complejidad computacional que ocurriría con un método basado en QMF.

Las dos listas siguientes resumen los aspectos preferidos de la configuración que se han descrito anteriormente, los cuales se pueden utilizar alternativamente entre sí o además de otros aspectos:

1a) concepto general: predicción compleja de MDCT lateral a partir de MDCT y MDST centrales;

1b) calcular/estimar 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;

1d) tratamiento adecuado de DC y $f_s/2$;

1e) tratamiento adecuado de conmutación de forma de ventana;

1f) no utilizar un cuadro anterior/siguiente si éste 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);

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);

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.

En una forma de realización, el codificador comprende adicionalmente: un convertidor espectral (50, 51) 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 donde el combinador (2031), el predictor (2033) y la calculadora de la señal residual (2034) están configurados para procesar cada una de las subbandas por separado de modo que la primera señal combinada y la señal residual se obtienen para una pluralidad de subbandas, en donde la interfaz de salida (212) está configurada para combinar la primera señal combinada codificada y la señal residual codificada para la pluralidad de subbandas.

A pesar de que algunos aspectos se han descrito en el contexto de un aparato, queda claro que dichos aspectos sólo representan una descripción del método correspondiente, en donde un bloque o dispositivo corresponde a un paso del método o a una característica de un paso del método. En 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.

En una forma de realización de la presente invención se aplica un tratamiento adecuado de conmutación de forma de ventana. En la Fig. 10a se ilustra que una información de forma de ventana 109 se puede introducir en la calculadora del espectro imaginario 1001. Específicamente, la calculadora del espectro imaginario, la cual 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 Fig. 6a o el elemento 1160a en la Fig. 6b) se puede implementar como un filtro FIR o IIR. Los coeficientes FIR o IIR en este módulo real-a-imaginario 1001 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 Derivada de Kaiser Bessel (KBD, por sus siglas en inglés) 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 donde una dimensión es la dimensión de tiempo en donde dos cuadros de MDCT subsiguientes se introducen en el filtro FIR y la segunda dimensión es la dimensión de frecuencia, en donde se introducen los coeficientes de frecuencia de un cuadro.

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
SÓLO_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
SÓLO_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]

Adicionalmente, la información de forma de ventana 109 proporciona información de la 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 tabla siguiente se proporcionan los coeficientes de filtro MDST correspondientes para la ventana anterior.

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
SÓLO_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]

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

La información de la forma de ventana, la cual 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 de la forma de ventana 109 se extrae de la corriente de bits mediante el desmultiplexor de la corriente de bits (por ejemplo 102 en la Fig. 5b) y se proporciona a la calculadora del espectro imaginario 1001 tal como se ilustra en la Fig. 10a.

20 Cuando la información de la forma de ventana 109 señala que el cuadro anterior tenía un distinto tamaño de transformada, entonces se prefiere que el cuadro anterior no se utilice para calcular el espectro imaginario a partir del espectro de valor real. Lo mismo ocurre cuando al interpretar la información de la forma de ventana 109 se descubre que el cuadro siguiente tiene un distinto tamaño de transformada. Entonces, el cuadro siguiente no se utiliza para calcular el espectro imaginario a partir del espectro de valor real. En dicho caso cuando, por ejemplo, el cuadro anterior tenía un distinto tamaño de transformada que el cuadro actual y cuando el cuadro siguiente nuevamente tiene un distinto tamaño de transformada en comparación con el cuadro actual, entonces sólo el cuadro actual, es decir, los valores espectrales de la ventana actual, se utilizan para estimar el espectro imaginario.

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 Fig. 3a, está basada en datos no cuantificados, entonces la calculadora residual 2034 también opera preferentemente en datos no cuantificados y la señal residual de salida de la calculadora, es decir, la señal residual 205 se cuantifica antes de ser codificada por entropía y transmitida a un decodificador. Sin embargo, en una forma de 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 Fig. 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 a continuación del combinador 2031 de tal modo que la primera señal de combinación y la segunda señal de combinación se calculen en una 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 de predicción 2035 se cuantifica antes de ser introducida en la calculadora residual. Entonces, es útil que la segunda señal de combinación 2032, la cual también se introduce en la calculadora residual 2034, también se cuantifique antes de que la calculadora residual calcule la señal residual 1070 en la Fig. 6a, la cual se puede implementar dentro del predictor 2033 en la Fig. 3a, opere en los mismos datos cuantificados que estén disponibles en el lado del decodificador. Entonces se puede asegurar que el espectro de MDST estimado en el codificador para llevar a cabo 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 Fig. 6a se cuantifica antes de ser introducida 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 Fig. 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.

En una forma de 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 se puede llevar 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 rango de valores para la parte real y la parte

imaginaria del coeficiente de predicción compleja tiene un rango desde -3 a 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.

5 Preferentemente, la parte real aplicada en el multiplicador 2073 en la Fig. 6a y la parte imaginaria 2074 aplicada en la Fig. 6a se cuantifican antes de ser aplicadas de modo que, nuevamente, se utiliza el mismo valor para la predicción en el lado del codificador así como se utiliza en el lado del decodificador. Esto garantiza que la señal residual de predicción abarque - además del error de cuantificación introducido - cualquier error que pudiera ocurrir cuando se aplica un coeficiente de predicción no cuantificado en el lado del codificador mientras se aplica un
10 coeficiente de predicción cuantificado en el lado del decodificador. Preferentemente, la cuantificación se aplica de modo que - tanto como sea posible - la misma situación y las mismas señales estén disponibles en el lado del codificador y en el lado del decodificador. De este modo se prefiere cuantificar la entrada en la calculadora real-a-imaginaria 2070 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 a 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 Fig. 6a también se puede cuantificar antes de los sumadores 2034a y 2034b. Sin embargo, no resulta problemático que el cuantificador 209b lleve a cabo la cuantificación posteriormente a la adición, en donde la adición por dichos sumadores se aplica con una señal lateral no cuantificada.

20 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. Dicha situación puede ocurrir cuando la señal central completa y la señal lateral completa no están desfasadas entre sí o están un poco desfasadas. A fin de ahorrar bits, esto es indicado 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 la corriente de bits, tal como un desmultiplexor de la corriente de bits, interpretará este indicador real y luego no buscará palabras clave para una parte imaginaria pero entenderá que todos los bits están en la sección correspondiente de la corriente de bits como bits para coeficientes de predicción de valor real. Asimismo,
25 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 Fig. 6b se desactivará y la predicción inversa sólo tendrá lugar utilizando el coeficiente de predicción de valor real aplicado en el multiplicador 1160b en la Fig. 6b. Esto mismo ocurre para el lado del codificador en donde el elemento 2070 se desactivará y la predicción sólo tendrá lugar utilizando el multiplicador 2073. Esta información lateral se utiliza preferentemente como un bit adicional bit por cuadro y el decodificador leerá este bit, cuadro por cuadro, a fin de decidir si el convertidor real-a-imaginario 1160a estará activo para un cuadro o no. De este modo, al proporcionar esta información se logra 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 cero para un cuadro y, adicionalmente, proporciona
30 menos complejidad para el decodificador para dicho cuadro logrando inmediatamente un menor consumo de batería de dicho procesador implementado, por ejemplo, en un dispositivo móvil alimentado por batería.

La predicción compleja estéreo de acuerdo con las formas de realización 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
35 utilizar un parámetro de valor complejo a , los canales izquierdo y derecho se reconstruyen por medio de la matriz siguiente. dmx_{lm} indica la MDST correspondiente a la MDCT de los canales de mezcla descendente dmx_{Re} .

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

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

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

55 **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
60 0: no se utiliza la predicción compleja, se utiliza la codificación L/R

complex_coef 1: se utiliza la predicción compleja
0: $\alpha_{lm} = 0$ para todas las bandas de predicción

use_prev_frame 1: α_{lm} se transmite para todas las bandas de predicción
0: Utilizar sólo el cuadro actual para estimar la MDST

delta_code_time 1: Utilizar el cuadro actual y anterior para estimar la MDST
0: Codificación diferencial de frecuencia de los coeficientes de predicción

hcod_alpha_q_re 1: Codificación diferencial de tiempo de los coeficientes de predicción
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 dicho cuadro.

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

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 cuales se utilizó **cplx_pred_used** = 0.

Los coeficientes de predicción inversa cuantificados **alpha_re** y **alpha_im** son proporcionados por

$$\mathbf{alpha_re} = \mathbf{alpha_q_re} * 0,1$$

$$\mathbf{alpha_im} = \mathbf{alpha_q_im} * 0,1$$

Cabe destacar que la invención no sólo se aplica a señales estéreo, es decir, señales multicanal que tienen sólo 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.

La señal de audio codificada de la invención se puede almacenar en un medio de almacenamiento digital o se puede transmitir en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable como Internet.

Dependiendo de determinados requisitos para la implementación, las formas de realización de la invención se pueden implementar en un hardware o en un software. La implementación se puede llevar a cabo utilizando un medio digital de almacenamiento, por ejemplo un disco flexible, un DVD, un CD, una ROM (memoria de sólo lectura, según sus siglas en inglés), una PROM (memoria programable de sólo lectura, según sus siglas en inglés), una EPROM (una PROM borrrable, según sus siglas en inglés), una EEPROM (una PROM borrrable eléctricamente) o una memoria FLASH, que tiene señales de control legibles electrónicamente almacenadas en el mismo, las cuales colaboran (o son capaces de colaborar) con un sistema informático programable de modo que el método respectivo se lleve a cabo.

Algunas formas de realización de acuerdo con la invención comprenden un portador de datos no temporario o tangible que tiene señales de control legibles electrónicamente, las cuales son capaces de colaborar con un sistema informático programable, de modo que uno de los métodos descritos en la presente se lleve a cabo.

En general, las formas de realización de la presente invención se pueden implementar como un producto de programa informático con un código de programa el cual es operativo para llevar a cabo uno de los métodos cuando el producto de programa informático se ejecuta en una computadora. El código de programa se puede almacenar, por ejemplo, en un portador legible por computadora.

Otras formas de realización comprenden el programa informático para llevar a cabo uno de los métodos descritos en

la presente, almacenados en un portador legible por computadora.

En otros términos, una forma de realización del método de la invención es, por lo tanto, un programa informático que tiene un código de programa para llevar a cabo uno de los métodos descritos en la presente, cuando el programa informático se ejecuta en una computadora.

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

Una forma de realización adicional del método de la invención es, por lo 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 la presente. La corriente de datos o la secuencia de señales pueden estar configuradas, por ejemplo, para ser transferidas mediante una conexión de comunicación de datos, por ejemplo, a través de Internet.

Otra forma de realización comprende un medio de procesamiento, por ejemplo, una computadora, o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los métodos descritos en la presente.

Otra forma de realización comprende una computadora que tiene instalado en la misma el programa informático para llevar a cabo uno de los métodos descritos en la presente.

En algunas formas de realización se puede utilizar un dispositivo lógico programable (por ejemplo un campo de matrices de puertas programables) para llevar a cabo algunas o todas las funcionalidades de los métodos descritos en la presente. En algunas formas de realización, un campo de matrices de puertas programables puede colaborar con un microprocesador a fin de llevar a cabo uno de los métodos descritos en la presente. En general, los métodos son llevados a cabo preferentemente por cualquier aparato de hardware.

Las formas de realización 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 la presente resultarán evidentes para los expertos en la técnica. Por consiguiente, la invención sólo 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 formas de realización incluidas en la presente.

REIVINDICACIONES

1. Un decodificador de audio para decodificar una señal de audio multicanal codificada (100), la señal de audio multicanal codificada comprende 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:
 - un decodificador de señales (110) para decodificar la primera señal de combinación codificada (104) para obtener una primera señal de combinación decodificada (112) y para decodificar la señal residual codificada (106) para obtener una señal residual decodificada (114); y
 - una calculadora del decodificador (116) para calcular una señal de audio multicanal decodificada que tiene una primera señal de audio de canal decodificada (117), y una segunda señal de audio de canal decodificada (118) utilizando la señal residual decodificada (114), la información de predicción (108) y la primera señal de combinación decodificada (112), de modo que la primera señal de audio de canal decodificada (117) y la segunda señal de audio de canal decodificada (118) son por lo menos aproximaciones de la primera señal de audio de canal y la segunda señal de audio de canal de la señal de audio multicanal, en donde la información de predicción (108) comprende un factor de valor real distinto de cero y/o un factor imaginario distinto de cero, en el cual la calculadora del decodificador (116) comprende:
 - un predictor (1160) para aplicar la información de predicción (108) a la primera señal de combinación decodificada (112) o a una señal (601) proveniente de la primera señal de combinación decodificada para obtener una señal de predicción (1163);
 - una calculadora de señales de combinación (1161) para calcular una segunda señal de combinación (1165) combinando la señal residual decodificada (114) y la señal de predicción (1163); y
 - un combinador (1162) para combinar la primera señal de combinación decodificada (112) y la segunda señal de combinación (1165) para obtener una señal de audio multicanal decodificada que tiene la primera señal de audio de canal decodificada (117) y la segunda señal de audio de canal decodificada (118),
 - en el cual el predictor (1160, 1160a) está configurado para multiplicar la primera señal de combinación decodificada por el factor real de la información de predicción (108) para obtener una primera parte de la señal de predicción, estimar una parte imaginaria de la primera señal de combinación decodificada (112) usando una parte real de la primera señal de combinación decodificada (112), comprendiendo estimar la parte imaginaria usar una pluralidad de subbandas de la primera señal de combinación decodificada adyacentes en frecuencia, en las cuales, en caso de bajas o altas frecuencias, se usa una extensión simétrica en frecuencia del cuadro actual de la primera señal de combinación para las subbandas asociadas con frecuencias más bajas o iguales a cero o más altas o iguales a una mitad de una frecuencia de muestreo en la que está basado el cuadro actual, o en el cual los coeficientes de filtro de un filtro incluido en el predictor (1160a) se establecen a diferentes valores para subbandas perdidas en comparación con subbandas no perdidas,
 - multiplicar la parte imaginaria (601) de la primera señal de combinación decodificada por el factor imaginario de la información de predicción (108) para obtener una segunda parte de la señal de predicción; y
 - en el cual la calculadora de señales de combinación (1161) está configurada para combinar en forma lineal la primera parte de la señal de predicción y la segunda parte de la señal de predicción y la señal residual decodificada para obtener la segunda señal de combinación (1165).
2. Decodificador de audio de acuerdo con la reivindicación 1, en el cual la primera señal de combinación codificada (104) y la señal residual codificada (106) se han generado utilizando un solapamiento que genera la conversión de tiempo-espectral, en donde el decodificador además comprende:
 - un convertidor espectral-tiempo (52, 53) para generar una primera señal de audio de canal de dominio de tiempo y una segunda señal de audio de canal de dominio de tiempo utilizando un algoritmo de conversión espectral-tiempo que coincide con el algoritmo de conversión de tiempo-espectral;
 - un procesador de solapar/agregar (522) para llevar a cabo un procesamiento de solapar-agregar para la primera señal de audio de canal de dominio de tiempo y para la segunda señal de audio 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.
3. Un decodificador de audio de acuerdo con una de las reivindicaciones anteriores, en el cual la primera señal de combinación codificada o decodificada (104) y la señal residual de predicción codificada o decodificada (106) comprenden, cada una, una primera pluralidad de señales de subbandas, en donde 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 donde el predictor (1160) está configurado para aplicar el mismo parámetro de predicción a por lo menos dos

- señales de subbandas diferentes de la primera señal de combinación decodificada, en donde la calculadora del decodificador (116) o la calculadora de señales de combinación (1161) o el combinador (1162) están configurados para llevar a cabo un procesamiento en subbandas; y en donde el decodificador de audio además comprende un banco de filtros de síntesis (52, 53) para combinar
- 5 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.
4. Decodificador de audio de acuerdo con la reivindicación 1,
- 10 en el cual el predictor (1160) está configurado para filtrar por lo menos dos cuadros subsiguientes de tiempo, en donde uno de los dos cuadros subsiguientes de tiempo precede o sigue un cuadro actual de la primera señal de combinación para obtener una parte imaginaria estimada de un cuadro actual de la primera señal de combinación utilizando un filtro lineal (1004, 1005, 1006, 1007).
- 15 5. Decodificador de audio de acuerdo con la reivindicación 1, en el cual la primera señal de combinación decodificada comprende una secuencia de cuadros de señales de valor real, y en el cual el predictor (1160) está configurado para estimar (1160a) una parte imaginaria del cuadro de señal actual utilizando sólo el cuadro de señal de valor real actual o utilizando el cuadro de señal de valor real actual y, ya sea
- 20 sólo uno o más cuadros de señal de valor real precedentes o sólo uno o más cuadros de valor real siguientes o utilizando el cuadro de señal de valor real actual y uno o más cuadros de señal de valor real precedentes y uno o más cuadros de señal de valor real siguientes.
6. Decodificador de audio de acuerdo con la reivindicación 1, en el cual el predictor (1160) está configurado para
- 25 recibir información de la forma de ventana (109) y para utilizar distintos coeficientes de filtro para calcular un espectro imaginario, en donde los distintos coeficientes de filtro dependen de distintas formas de ventana indicadas por la información de la forma de ventana (109).
7. Decodificador de audio de acuerdo con la reivindicación 4, 5 o 6,
- 30 en el cual la primera señal de combinación decodificada está asociada con distintas longitudes de transformadas indicadas por un indicador de longitud de transformadas incluido en la señal de audio multicanal codificada (100), y en el cual el predictor (1160) está configurado solamente para utilizar uno o más cuadros de la primera señal de combinación que tiene la misma longitud de transformada asociada para estimar la parte imaginaria para un cuadro actual para una primera señal de combinación.
- 35 8. Decodificador de audio de acuerdo con una de las reivindicaciones anteriores, en el cual la información de predicción (108) está incluida en la señal de audio multicanal codificada en una representación cuantificada y codificada por entropía, en donde el decodificador de audio además comprende un decodificador de información de predicción (65) para la
- 40 decodificación por entropía o la descuantificación para obtener una información de predicción decodificada utilizada por el predictor (1160), o en el cual la señal de audio multicanal codificada comprende una unidad de datos que indica en el primer estado que el predictor (1160) tiene que utilizar por lo menos un cuadro anterior o siguiente en el tiempo a un cuadro actual de la
- 45 primera señal de combinación decodificada, y que indica en el segundo estado que el predictor (1160) tiene que utilizar sólo un cuadro de la primera señal de combinación decodificada para estimar una parte imaginaria para el cuadro actual de la primera señal de combinación decodificada, y en el cual el predictor (1160) está configurado para detectar un estado de la unidad de datos y para operar en consecuencia.
9. Decodificador de audio de acuerdo con una de las reivindicaciones anteriores, en el cual la información de
- 50 predicción (108) comprende palabras clave de diferencias entre valores complejos secuenciales de tiempo o adyacentes de frecuencia, y en donde el decodificador de audio está configurado para llevar a cabo un paso de decodificación por entropía y un paso de decodificación de diferencia subsiguiente para obtener valores de predicción compleja cuantificada secuencial de tiempo o valores de predicción compleja para bandas de frecuencia adyacentes.
- 55 10. Decodificador de audio de acuerdo con una de las reivindicaciones 1 a 7, en el cual la señal de audio multicanal codificada comprende, como información lateral, un indicador real que indica que todos los coeficientes de predicción para un cuadro de la señal de audio multicanal codificada son de valor real, en donde el decodificador de audio está configurado para extraer el indicador real de la señal de audio multicanal
- 60 codificada (100), y en donde la calculadora del decodificador (116) está configurada para no calcular una señal imaginaria para un cuadro, para el cual el indicador real está indicando sólo coeficientes de predicción de valor real.
11. Método para decodificar una señal de audio multicanal codificada (100), 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:

- 5 decodificar (110) la primera señal de combinación codificada (104) para obtener una primera señal de combinación decodificada (112), y decodificar la señal residual codificada (106) para obtener una señal residual decodificada (114); y
- 10 calcular (116) una señal de audio multicanal decodificada que tiene una primera señal de canal decodificada (117), y una segunda señal de canal decodificada (118) utilizando la señal residual decodificada (114), la información de predicción (108) y la primera señal de combinación decodificada (112), de modo que la primera señal de audio de canal decodificada (117) y la segunda señal de audio de canal decodificada (118) son por lo menos aproximaciones de la primera señal de audio de canal y la segunda señal de audio de canal de la señal de audio multicanal, en donde la información de predicción (108) comprende un factor de valor real distinto de cero y/o un factor imaginario distinto de cero.
- 15 aplicar la información de predicción (108) a la primera señal de combinación decodificada (112) o a una señal (601) proveniente de la primera señal de combinación decodificada, por un predictor (1160), para obtener una señal de predicción (1163);
- 20 calcular una segunda señal de combinación (1165) combinando la señal residual decodificada (114) y la señal de predicción (1163), mediante un calculador de señal de combinación (1161), y
- combinar la primera señal de combinación (112) y la segunda señal de combinación (1165) decodificadas para obtener una señal de audio multicanal decodificada que tiene la primera señal de audio de canal decodificada (117) y la segunda señal de audio de canal decodificada (118), por un combinador (1162),
- 25 en el cual aplicar la información de predicción comprende:

multiplicar la primera señal de combinación decodificada por el factor real de la información de predicción (108) para obtener una primera parte de la señal de predicción,
30 estimar una parte imaginaria de la primera señal de combinación decodificada (112) usando una parte real de la primera señal de combinación decodificada (112), comprendiendo estimar la parte imaginaria usar una pluralidad de subbandas de la primera señal de combinación decodificada adyacentes en frecuencia, en las cuales, en caso de bajas o altas frecuencias, se usa una extensión simétrica en frecuencia del cuadro actual de la primera señal de combinación para las subbandas asociadas con frecuencias más bajas o iguales a cero o más altas o iguales a una mitad de una frecuencia de muestreo en la que está basado el cuadro actual, o en el cual los coeficientes de filtro de un filtro incluido en el predictor (1160a) se establecen a
35 diferentes valores para subbandas perdidas en comparación con subbandas no perdidas,
multiplicar la parte imaginaria (601) de la primera señal de combinación decodificada por el factor imaginario de la información de predicción (108) para obtener una segunda parte de la señal de predicción; y
- 40 en el cual calcular la segunda señal de combinación comprende combinar en forma lineal la primera parte de la señal de predicción, la segunda parte de la señal de predicción y la señal residual decodificadas para obtener la segunda señal de combinación (1165).
- 45 12. Un programa de computadora para llevar a cabo, cuando se ejecuta en una computadora o en un procesador, el método de la reivindicación 11.

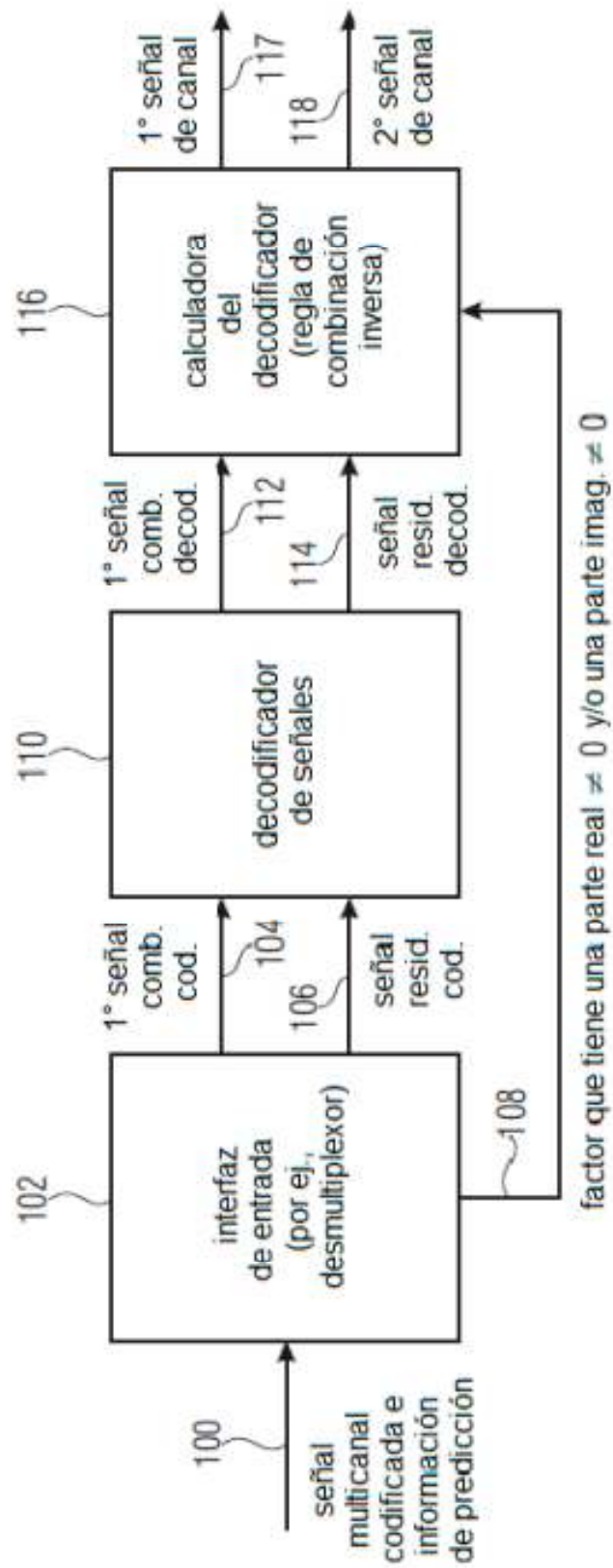


FIGURA 1
(DECODIFICADOR DE AUDIO)

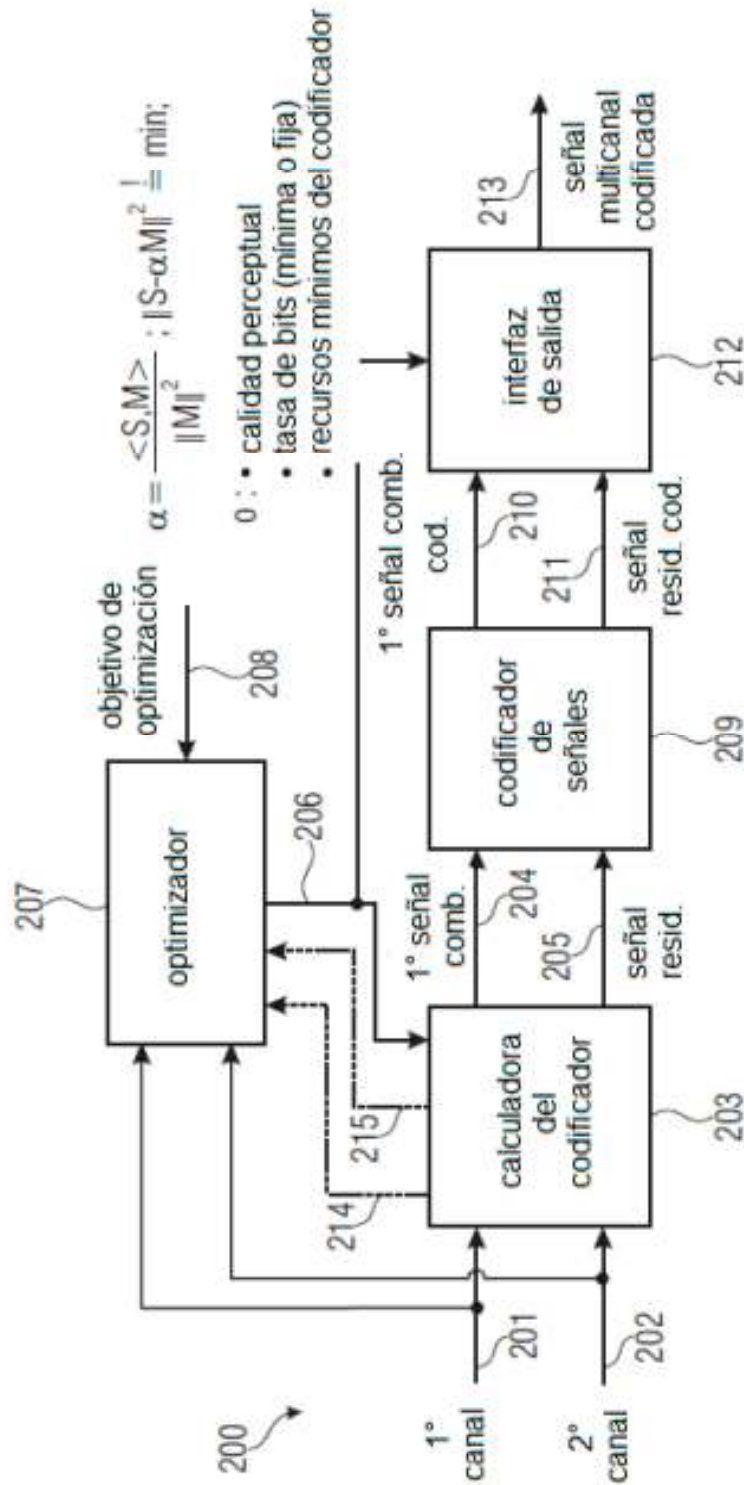


FIGURA 2
(CODIFICADOR DE AUDIO)

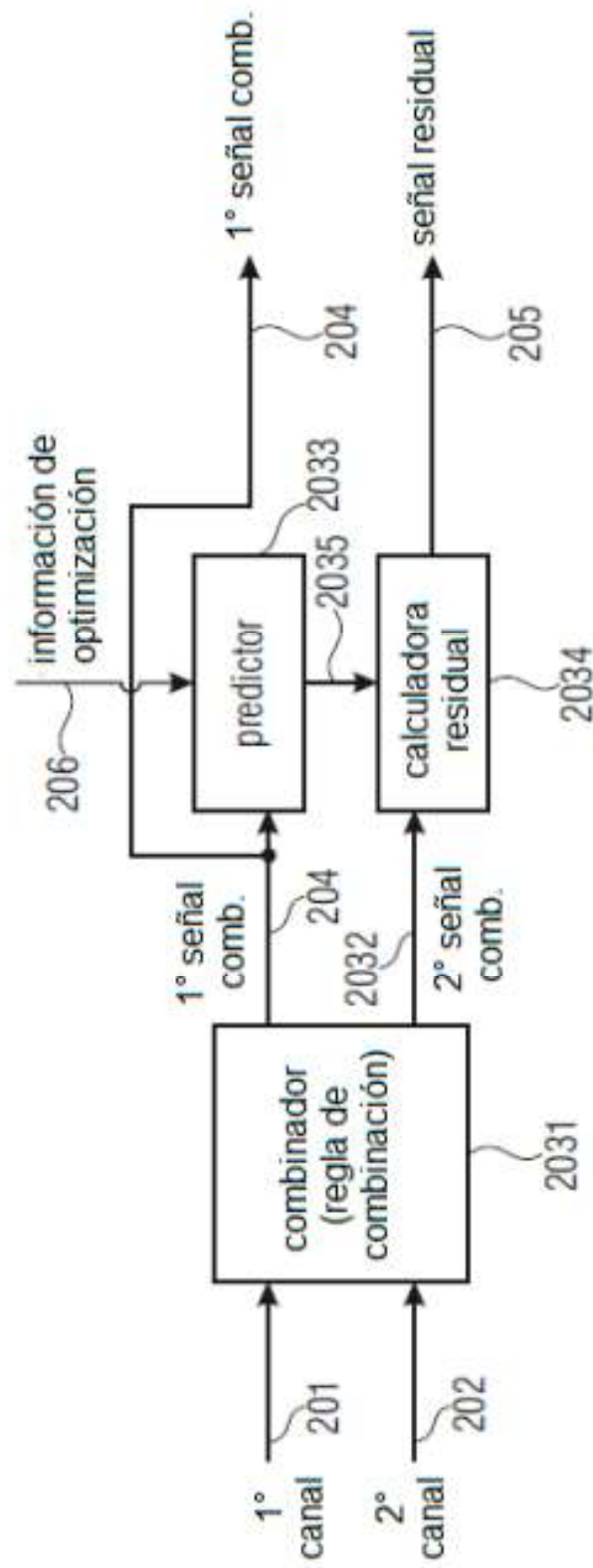


FIGURA 3A
(CODIFICADOR DE AUDIO LATERAL)

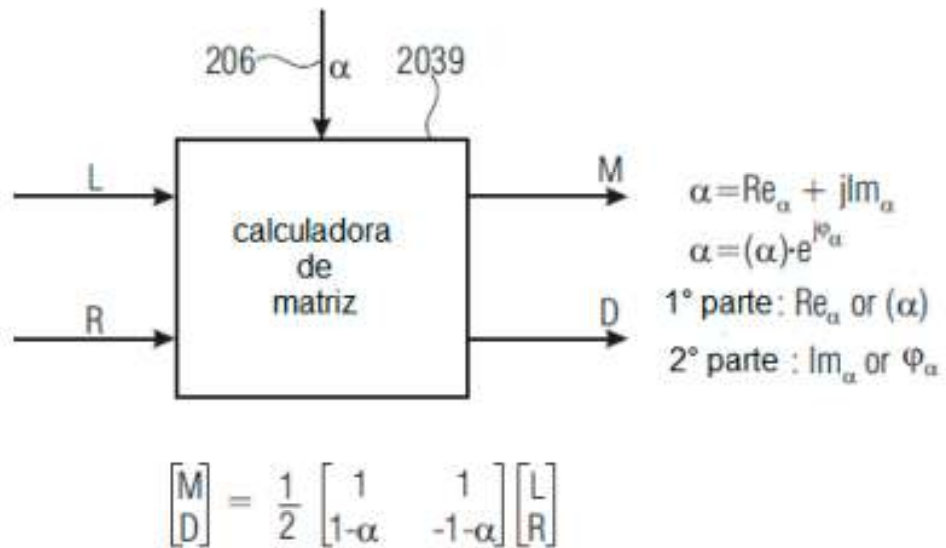


FIGURA 3B
(CODIFICADOR DE AUDIO LATERAL)

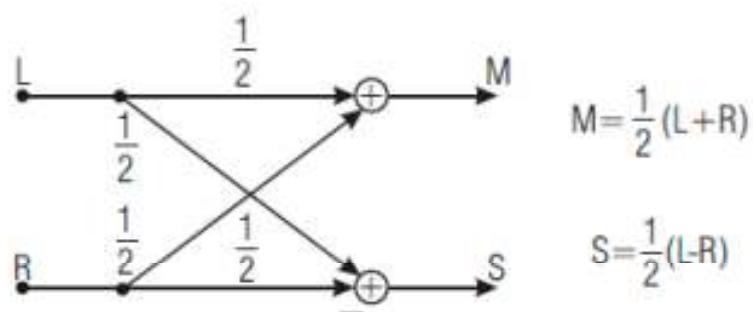


FIGURA 3C
(REGLA DE COMBINACIÓN)

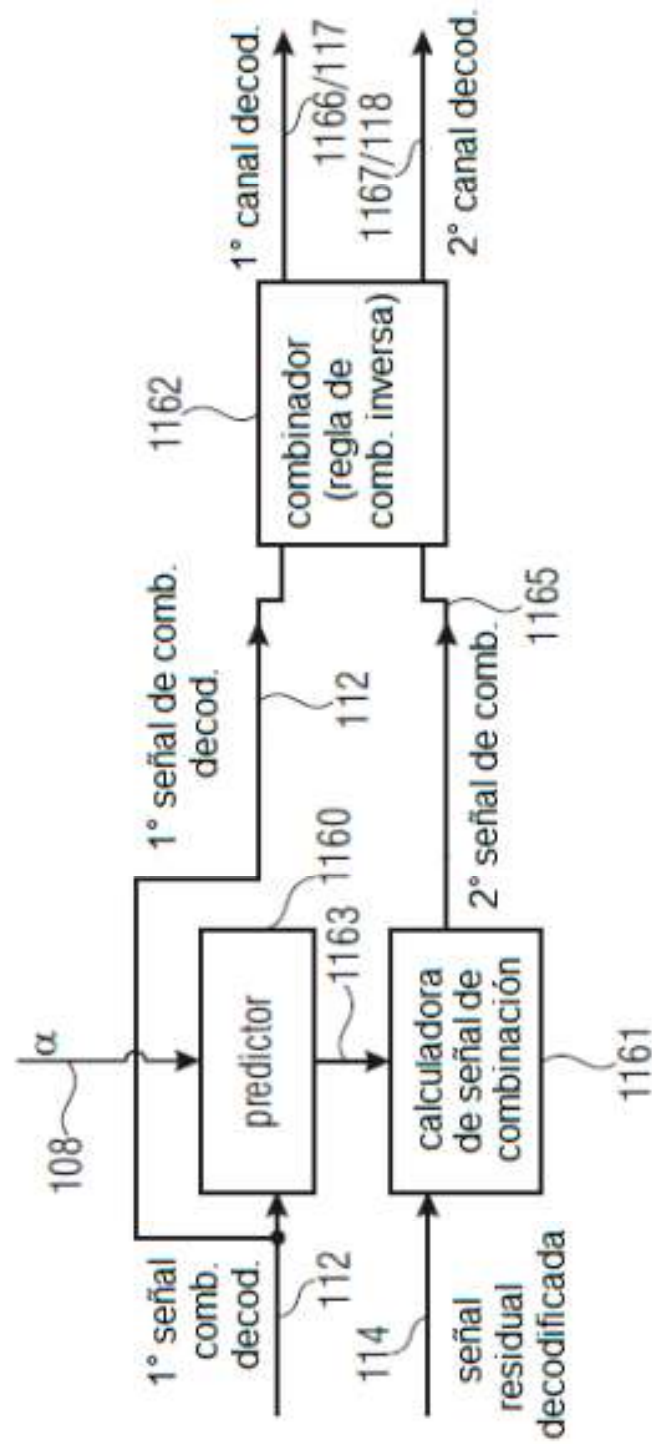


FIGURA 4A
(DECODIFICADOR DE AUDIO LATERAL)

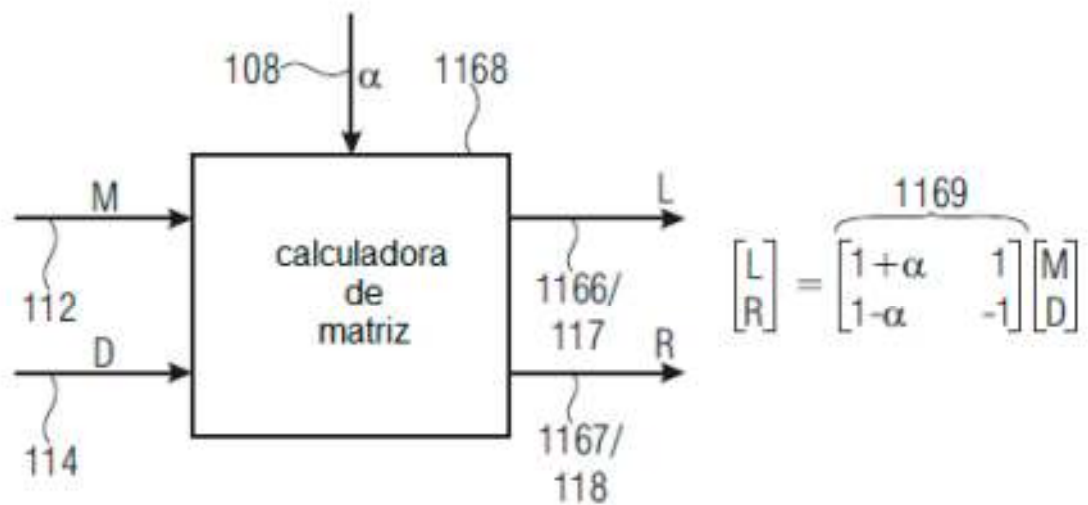


FIGURA 4B
(DECODIFICADOR DE AUDIO LATERAL)

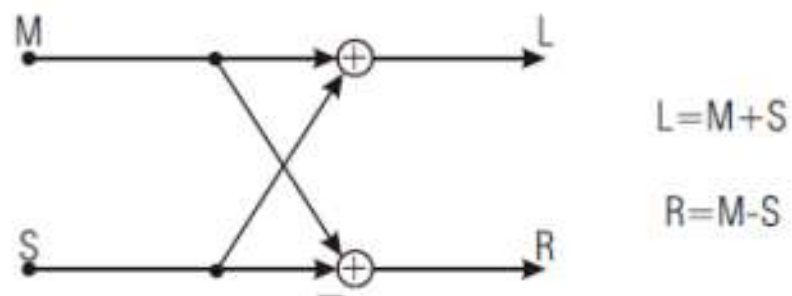


FIGURA 4C
(REGLA DE COMBINACIÓN INVERSA)

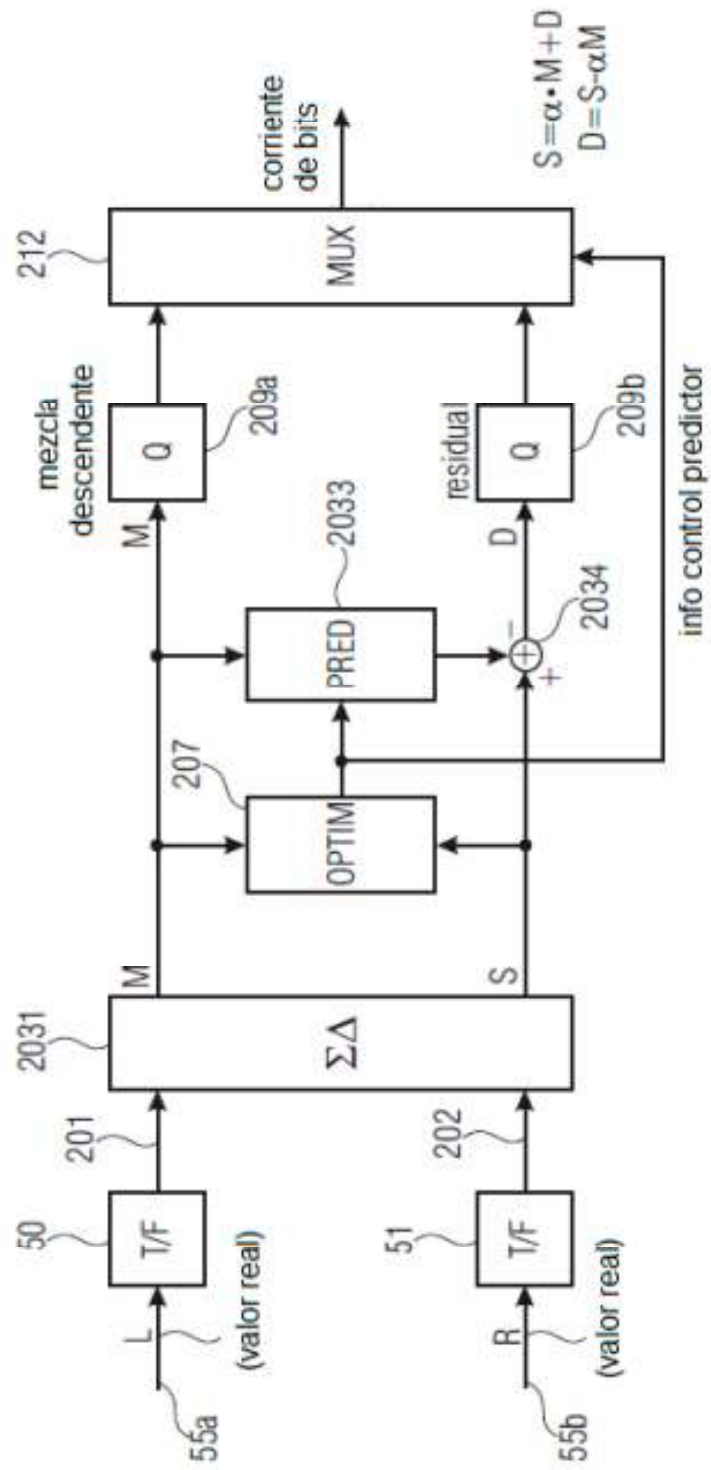


FIGURA 5A
(CODIFICADOR LATERAL)

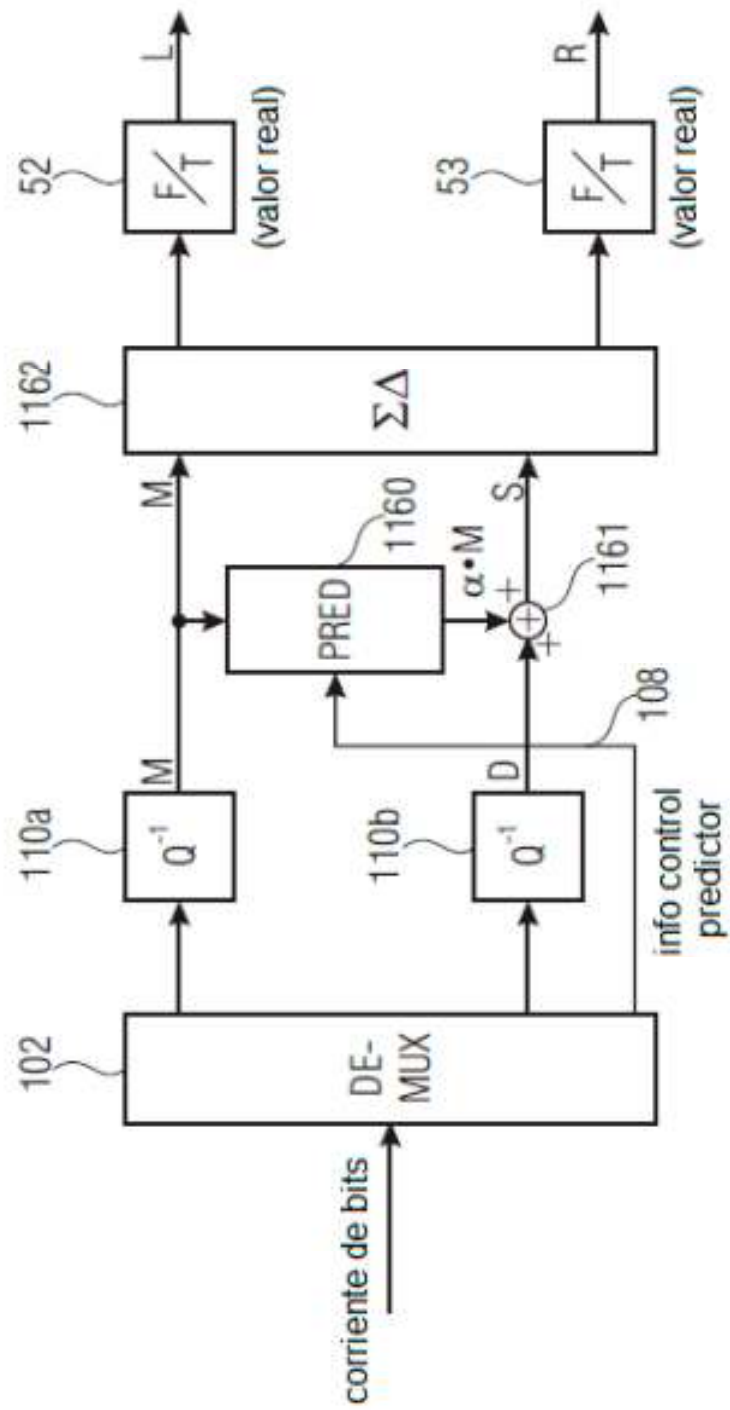


FIGURA 5B
(DECODIFICADOR LATERAL)

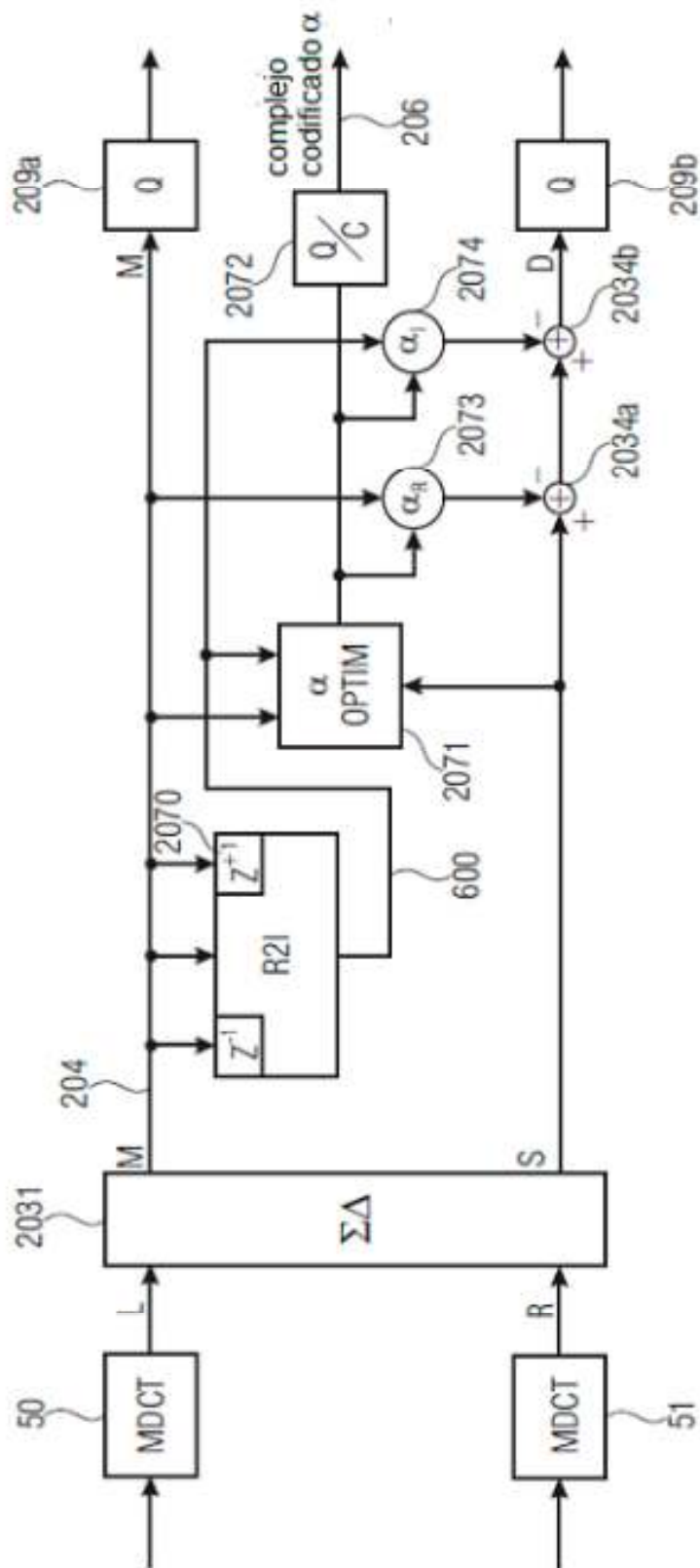


FIGURA 6A
(CODIFICADOR LATERAL)

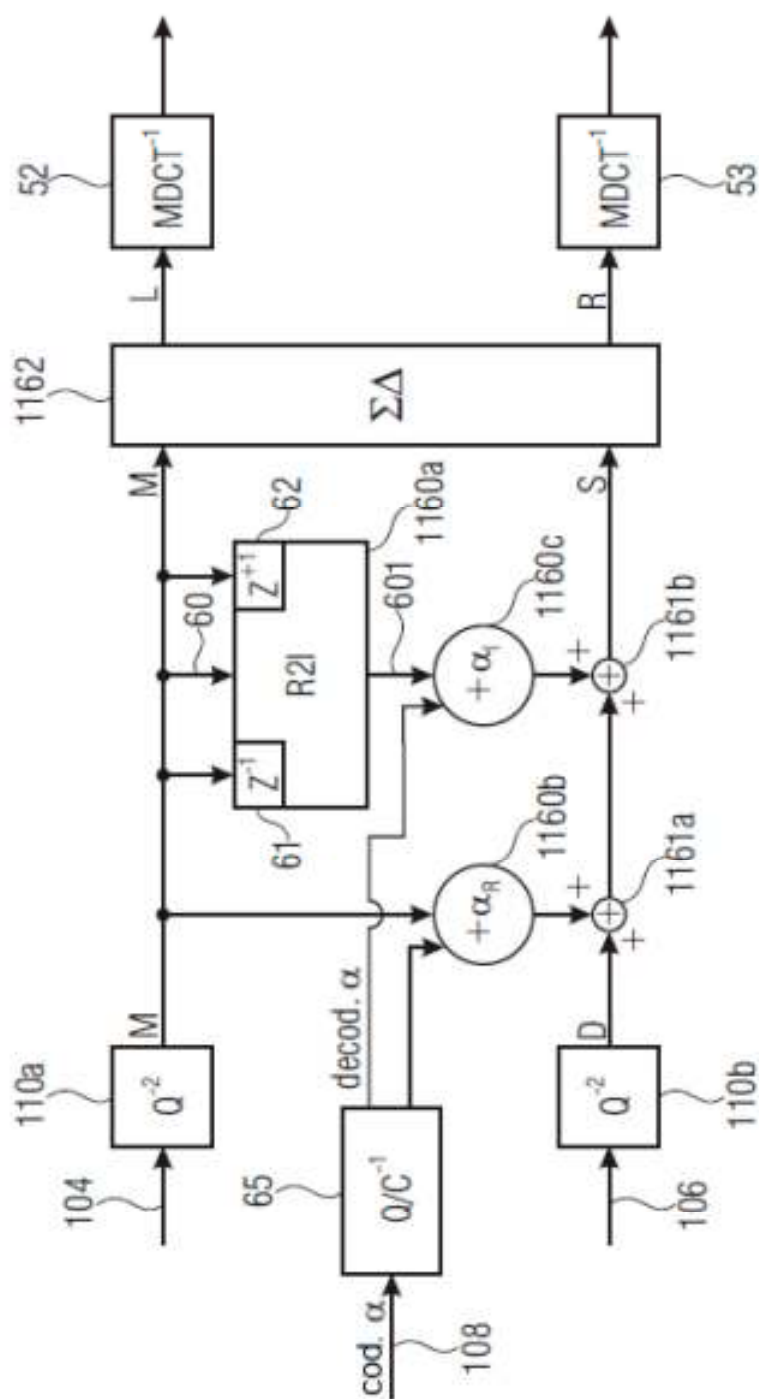


FIGURA 6B
(DECODIFICADOR LATERAL)

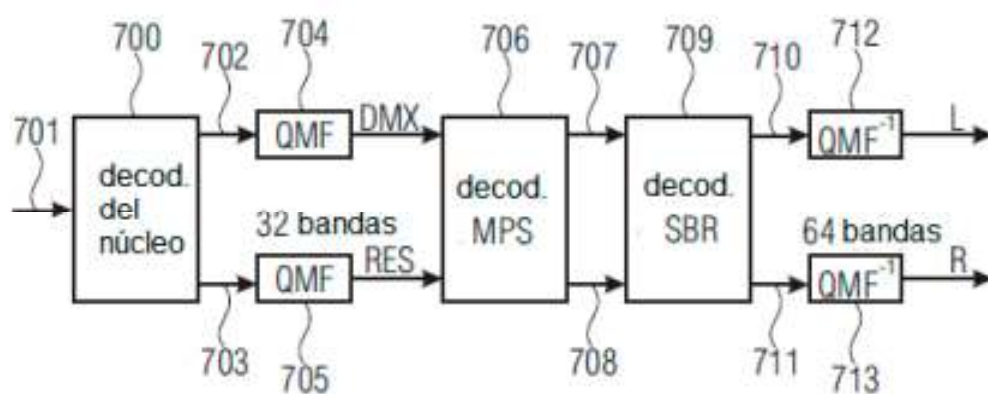


FIGURA 7A

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

FIGURA 7B

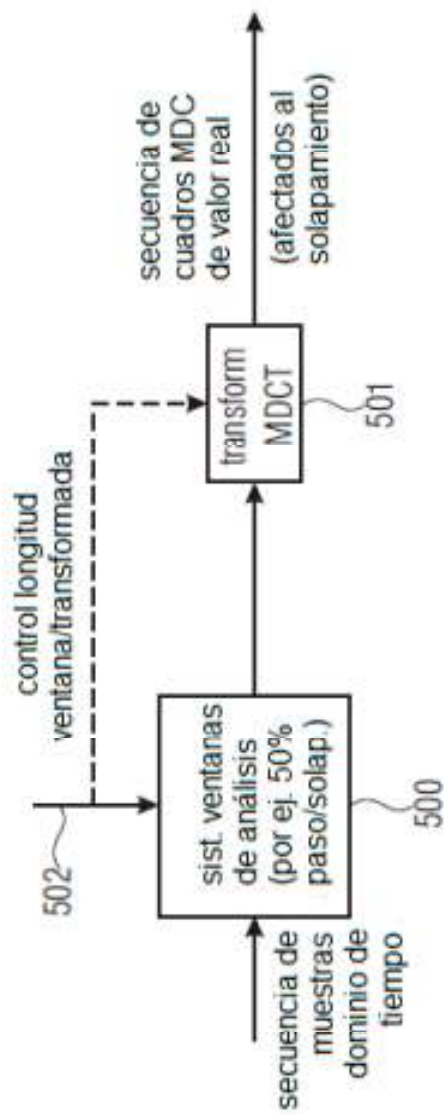


FIGURA 8A

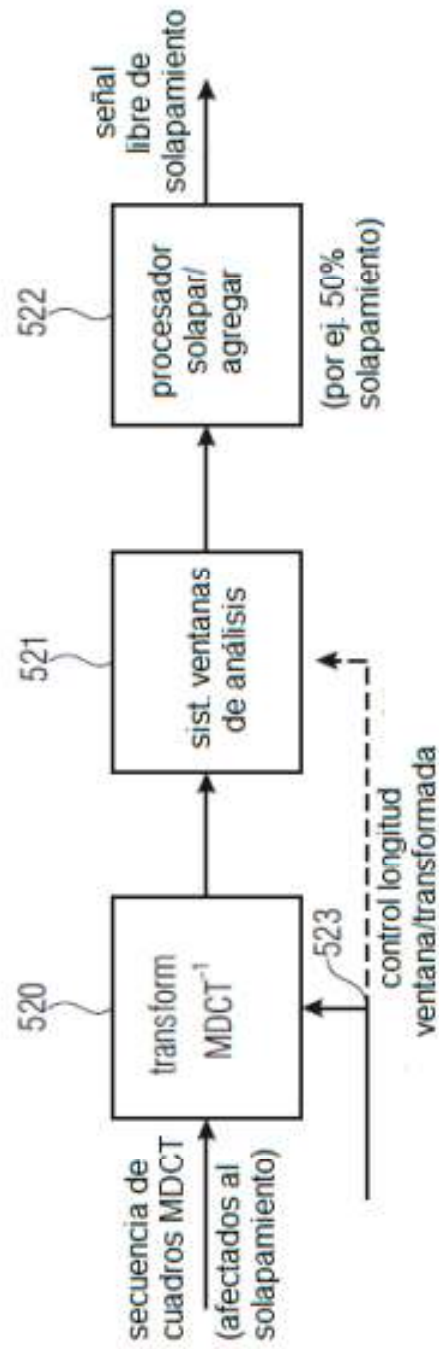


FIGURA 8B

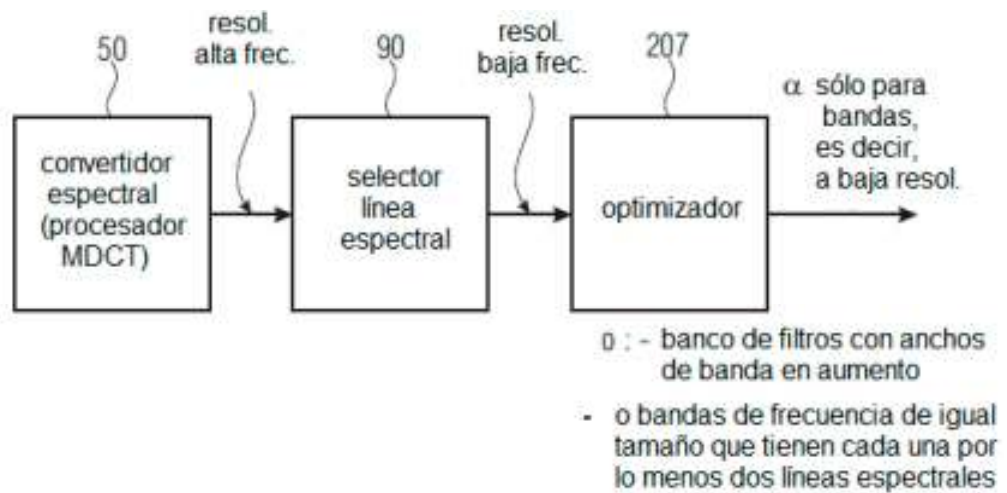


FIGURA 9A

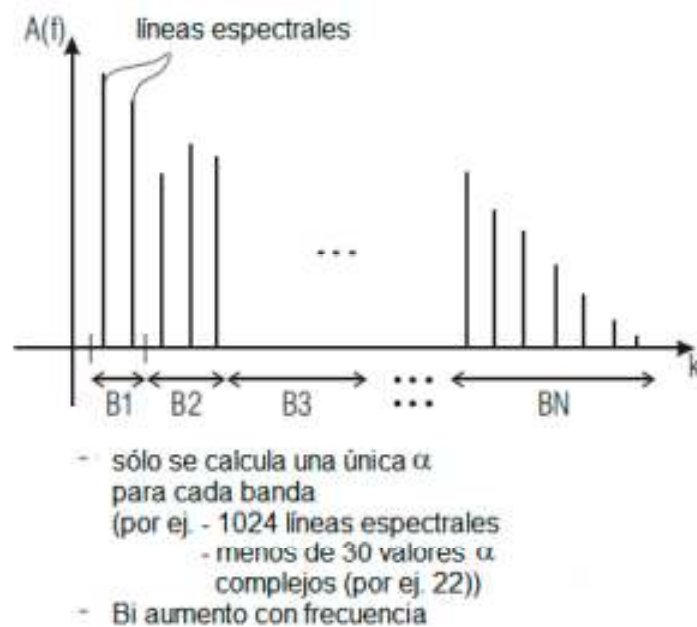


FIGURA 9B

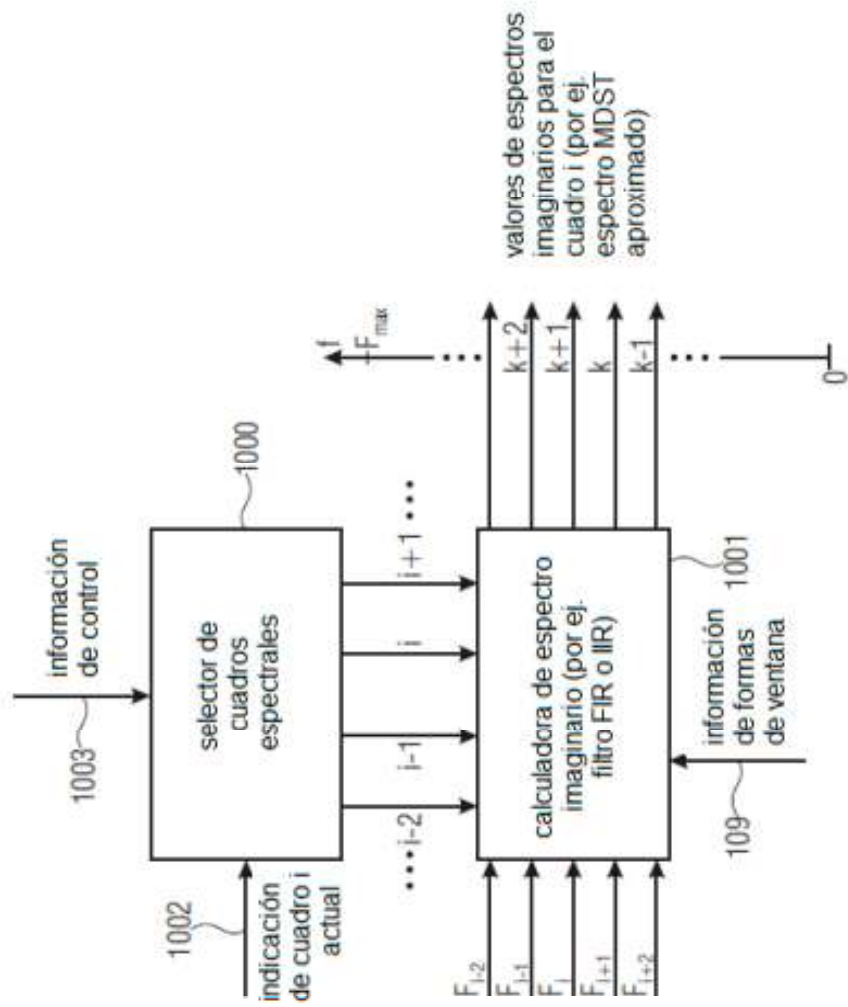


FIGURA 10A

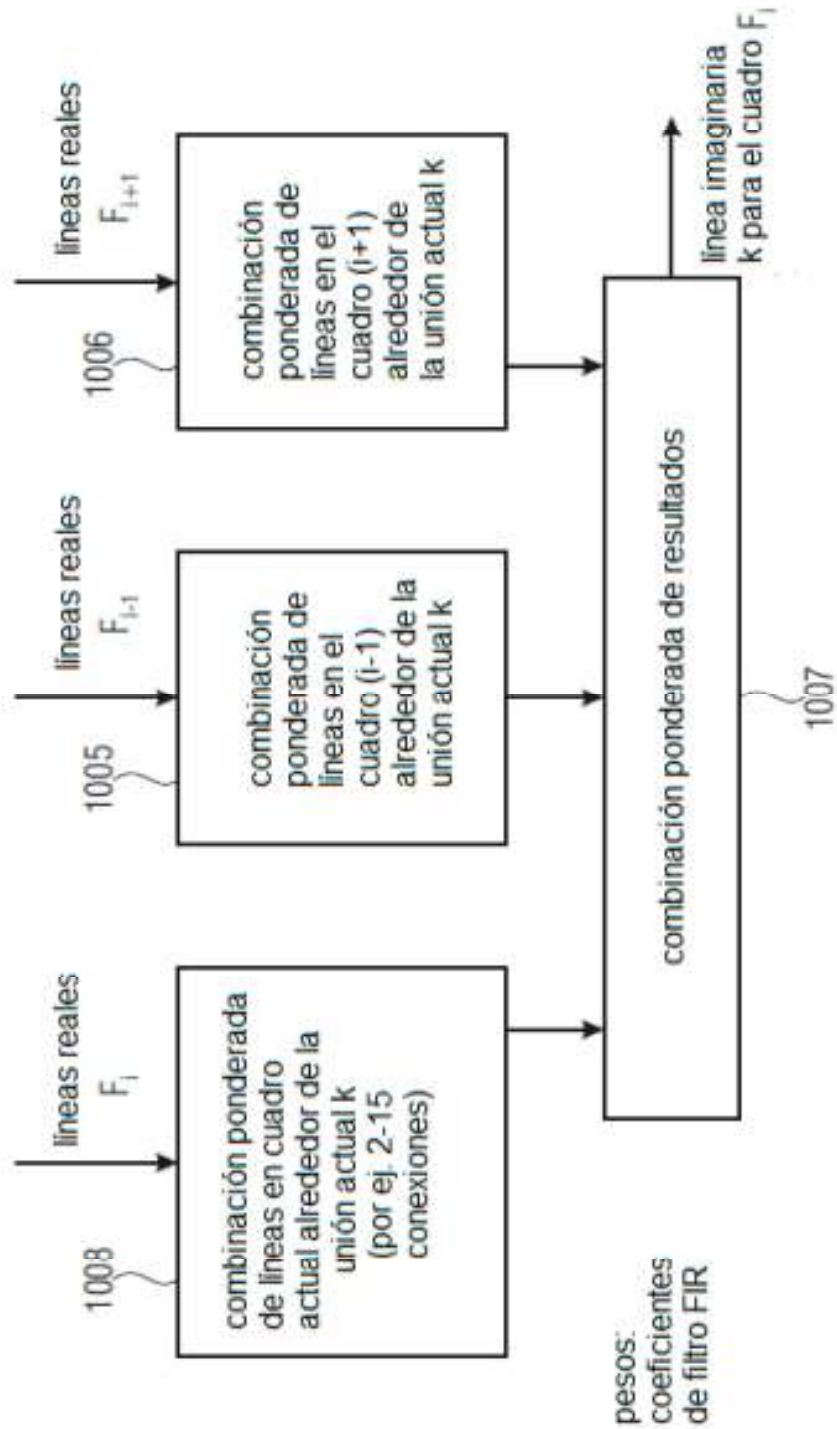


FIGURA 10B