

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 708 581**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/04 (2013.01)

G10L 19/022 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.03.2011 E 15176778 (7)**

97 Fecha y número de publicación de la concesión europea: **31.10.2018 EP 2947653**

54 Título: **Codificación de audio multicanal usando predicción compleja e información de forma de ventana**

30 Prioridad:

09.04.2010 US 322688 P

13.07.2010 US 363906 P

13.07.2010 EP 10169432

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

10.04.2019

73 Titular/es:

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

Hansastraße 27c

MÜNCHEN, DE y

DOLBY INTERNATIONAL AB (50.0%)

72 Inventor/es:

PURNHAGEN, HEIKO;

CARLSSON, PONTUS;

VILLEMÖES, LARS;

ROBILLIARD, JULIEN;

NEUSINGER, MATTHIAS;

HELMRICH, CHRISTIAN;

HILPERT, JOHANNES;

RETTTELACH, NIKOLAUS;

DISCH, SASCHA y

EDLER, BERND

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 708 581 T3

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

DESCRIPCIÓN

Codificación de audio multicanal usando predicción compleja e información de forma de ventana

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

10 En el campo del procesamiento de estéreo o multicanal se sabe aplicar la así denominada codificación de estéreo central / lateral. 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. Por lo general, 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 o un codificador aritmético de entropía o basado en Huffman, la ganancia de codificación aumenta, cuando la señal lateral se vuelve más pequeña. No obstante, 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 usando 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 con el 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.

25 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 resulta de nuevo 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 de forma significativa, 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 de nuevo 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, como alternativa, en el dominio de tiempo.

35 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 de acuerdo con 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". 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.

50 Este método puede resultar complejo en particular en el caso en el que el descodificador tiene que aplicar un procesamiento de descorrelación con el fin de crear, de forma artificial, 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 puesta en práctica y pueden introducir distorsiones, en particular 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 produce inevitablemente una pérdida de información no solo 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.

60 Existen desarrollos recientes en cuanto a la codificación de voz y audio unificada (*unified speech and audio coding*, USAC) que se ilustra en la figura 7a. Un descodificador de núcleo 700 lleva a cabo una operación de descodificación de la señal estéreo codificada en la entrada 701, la cual puede ser una señal codificada centro / lateral. El descodificador de 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 de QMF por los bancos de filtros QMF 704 y 705. Luego se aplica un descodificador 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 descodificador de replicación espectral

de ancho de banda (*spectral band replication*, SBR) 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.

5 La figura 7b ilustra una situación en la que el descodificador del Entorno de MPEG 706 llevaría a cabo una descodificación central / lateral. Como alternativa, el bloque descodificador del Entorno de MPEG 706 podría llevar a cabo una descodificación paramétrica basada en la señal binaural para generar señales estéreo a partir de una única señal mono del descodificador de núcleo. Naturalmente, el descodificador del Entorno de MPEG 706 también
10 podría generar una pluralidad de señales de salida de banda baja para introducir en el bloque descodificador de SBR 709 usando información paramétrica tal como diferencias de nivel entre canales, medidas de coherencia entre canales u otros de dichos parámetros de información entre canales.

15 Cuando el bloque descodificador del Entorno de MPEG 706 lleva a cabo la descodificación central / lateral que se ilustra en la figura 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, de forma respectiva, representadas en el dominio complejo híbrido QMF.

20 El uso de una combinación de un bloque 706 y un bloque 709 solo genera un pequeño aumento en la complejidad computacional en comparación con un descodificador estéreo usado como base, debido a que la representación del QMF complejo de la señal ya está disponible como parte del descodificador de SBR. No obstante, 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 solo deberían agregarse para los fines de la codificación estéreo.

25 En el sistema de USAC del MPEG en desarrollo, no obstante, también existen modos de codificación a altas tasas de bits en donde por lo general, no se usa la SBR.

Los siguientes documentos son ejemplares de esquemas de (des)codificación 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:

30 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.º M16921, 23 de octubre de 2009 (23 - 10 - 2009);

35 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.º 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).

40 Además, se hace referencia al documento WO 2011/124608, publicado posteriormente.

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.

45 Este objetivo se logra mediante un descodificador de audio de acuerdo con la reivindicación 1, un codificador de audio de acuerdo con la reivindicación 13, un método de descodificación de audio de acuerdo con la reivindicación 17, un método de codificación de audio de acuerdo con la reivindicación 18 y un programa informático de acuerdo con la reivindicación 19. 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 de forma significativa mediante una predicción de una segunda señal de combinación usando una primera señal de combinación, en donde ambas señales de combinación provienen de las señales originales de canal usando 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 tal modo que se cumple un objetivo de optimización, se incurre solo en una pequeña sobrecarga, pero logra una disminución significativa de la tasa de bits necesaria para la señal lateral sin
50 perder ninguna calidad de audio, debido a que la predicción inventiva es, no obstante, una codificación basada en forma de onda y no un método de codificación multicanal o estéreo basado en parámetros. Con el fin de reducir la complejidad computacional, se prefiere llevar a cabo una codificación en el dominio de la 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 en el dominio del tiempo en una representación espectral es preferentemente un proceso críticamente muestreado tal como una transformada de coseno discreta modificada (*modified discrete cosine transform*, MDCT) o una transformada de seno discreta modificada (*modified discrete sine transform*, MDST), la cual es diferente de una transformada compleja debido a que se calculan solo valores reales o solo valores imaginarios, en tanto que en una transformada compleja se calculan valores reales y complejos de un espectro logrando un sobremuestreo dos veces mayor.
60

Preferentemente se usa 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 en el dominio del tiempo (*time domain aliasing cancellation*, TDAC), la cual se obtiene mediante el procesamiento de solapar - agregar en el lado del descodificador.

Preferentemente, la información de predicción calculada en el codificador, transmitida al descodificador y usada en el descodificador comprende una parte imaginaria, la cual puede reflejar de forma ventajosa las diferencias de fase entre los dos canales de audio en cantidades seleccionadas de forma arbitraria entre 0° y 360°. La complejidad computacional se reduce de forma significativa cuando se aplica solo una transformada de valor real o, en general, una transformada que proporciona, ya sea solo un espectro real o que solo proporciona un espectro imaginario. Con el fin de usar 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 de real a imaginario o, dependiendo de la puesta en práctica de la transformada, un convertidor de complejo a real se proporciona en el descodificador con el 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 descodificado en una determinada banda y el canal derecho descodificado en esta banda.

Con el fin de aumentar la calidad de audio, el mismo convertidor de real a imaginario o el convertidor de complejo a real, el cual se aplica en el lado del descodificador se pone en práctica 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 por lo general, no se usa la SBR. En lugar de procesar la señal en el dominio complejo híbrido QMF, estos métodos ponen en práctica 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 usando una transformada de real a complejo, en donde dicha señal estéreo puede ser, ya sea una señal estéreo codificada en el lado del codificador o puede ser, como alternativa, una señal estéreo descodificada / transmitida, cuando el aparato o método para generar la señal estéreo se aplica en el lado del descodificador.

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

- La figura 1 es un diagrama de un descodificador de audio;
- La figura 2 es un diagrama de bloques de un codificador de audio;
- La figura 3a ilustra una puesta en práctica de la calculadora de codificador de la figura 2;
- La figura 3b ilustra una puesta en práctica alternativa de la calculadora de codificador de la figura 2;
- La figura 3c ilustra una regla de combinación central / lateral para aplicar en el lado del codificador;
- La figura 4a ilustra una puesta en práctica de la calculadora de descodificador de la figura 1;
- La figura 4b ilustra una puesta en práctica alternativa de la calculadora de descodificador en forma de una calculadora de matriz;
- La figura 4c ilustra una regla de combinación inversa central / lateral correspondiente a la regla de combinación que se ilustra en la figura 3c;
- La figura 5a ilustra una forma de realización de un codificador de audio que opera en el dominio de frecuencia, el cual es preferentemente un dominio de frecuencia de valor real;
- La figura 5b ilustra una puesta en práctica de un descodificador de audio que opera en el dominio de frecuencia;
- La figura 6a ilustra una puesta en práctica alternativa de un codificador de audio que opera en el dominio de MDCT y usa una transformada de real a imaginario de acuerdo con una forma de realización de la presente invención;
- La figura 6b ilustra un descodificador de audio que opera en el dominio de MDCT y usa una transformada de real a imaginario de acuerdo con una forma de realización de la presente invención;
- La figura 7a ilustra un post-procesador de audio que usa un descodificador estéreo y un descodificador de SBR conectado posteriormente;
- La figura 7b ilustra una matriz de mezcla ascendente central / lateral;
- La figura 8a ilustra una vista detallada en el bloque de MDCT en la figura 6a;

- La figura 8b ilustra una vista detallada en el bloque de MDCT¹ de la figura 6b;
- La figura 9a ilustra una puesta en práctica de un optimizador que opera en resolución reducida con respecto a la salida de MDCT;
- La figura 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;
- La figura 10a ilustra una puesta en práctica del transformador de real a imaginario en la figura 6a o en la figura 6b; y
- La figura 10b ilustra una puesta en práctica posible de la calculadora de espectro imaginario de la figura 10a.
- La figura 1 ilustra un descodificador de audio para descodificar 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 usando 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 poner en práctica 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 descodificador de señales 110 para descodificar la primera señal de combinación para obtener una primera señal de combinación descodificada en la línea 112. Adicionalmente, el descodificador de señales 110 está configurado para descodificar la señal residual codificada para obtener una señal residual descodificada en la línea 114. Dependiendo del procesamiento de codificación en un lado del codificador de audio, el descodificador de señales puede comprender un descodificador de entropía tal como un descodificador Huffman, un descodificador aritmético o cualquier otro descodificador 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 de descodificador 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. Con el fin de codificar totalmente dicha señal multicanal que tiene cinco canales se pueden aplicar dos descodificadores que se ilustran en la figura 1, en donde el primer descodificador procesa el canal izquierdo y el canal derecho, el segundo descodificador procesa el canal del entorno izquierdo y el canal del entorno derecho, y un tercer descodificador mono se usaría para llevar a cabo una codificación mono del canal central. No obstante, 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 3^a señal de combinación a partir de una 1^a y 2^a señal usando dos coeficientes de predicción, en forma muy similar al módulo "de dos a tres" en el Entorno de MPEG.
- La calculadora de descodificador 116 está configurada para calcular una señal multicanal descodificada que tiene la primera señal de canal descodificada 117 y la segunda señal de canal descodificada 118 usando la señal residual descodificada 114, la información de predicción 108 y la primera señal de combinación descodificada 112. En concreto, la calculadora de descodificador 116 está configurada para operar de tal modo que la primera señal de canal descodificada y la segunda señal de canal descodificada son al 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. En concreto, 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 de descodificador 116 se puede poner en práctica de diversos modos. Una primera puesta en práctica se ilustra en la figura 4a. Esta puesta en práctica comprende un predictor 1160, una calculadora de señales de combinación 1161 y un combinador 1162. El predictor recibe la primera señal de combinación descodificada 112 y la información de predicción 108 y emite una señal de predicción 1163. En concreto, el predictor 1160 está configurado para aplicar la información de predicción 108 a la primera señal de combinación descodificada 112 o una señal proveniente de la primera señal de combinación descodificada. La regla de derivación para derivar la señal a la cual se aplica la información de predicción 108 puede ser una transformada de real a imaginario, o por igual, una transformada de imaginario a real o una operación de ponderación o, dependiendo de la puesta en práctica, una operación de desfase o una operación combinada de ponderación / desfase. La señal de predicción 1163 se introduce junto con la señal residual descodificada en la calculadora de señales de combinación 1161 con el fin de calcular la segunda señal de combinación descodificada 1165. Ambas señales 112 y 1165 se introducen en el combinador 1162, el cual combina la primera señal de combinación descodificada y la segunda señal de combinación para obtener la señal de audio multicanal descodificada que tiene la primera señal de canal descodificada y la segunda señal de canal descodificada en las líneas de salida 1166 y 1167, de forma respectiva.

Como alternativa, la calculadora de descodificador se pone en práctica como una calculadora de matriz 1168, la cual recibe, como entrada, la primera señal de combinación descodificada o señal M, la señal residual descodificada o señal D y la información de predicción α 108. La calculadora de matriz 1168 aplica una matriz de transformada que se ilustra como 1169 a las señales M, D para obtener las señales de salida L, R, en donde L es la primera señal de canal descodificada y R es la segunda señal de canal descodificada. La notación en la figura 4b se asemeja a una notación estéreo con un canal izquierdo L y un canal derecho R. Se ha aplicado esta notación con el fin de proporcionar una 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 figura 4a en un tipo de cálculo matriz de “disparo único”, y las entradas en el circuito de la figura 4a y las salidas desde el circuito de la figura 4a son idénticas a las entradas en la calculadora de matriz 1168 o a las salidas desde la calculadora de matriz 1168.

La figura 4c ilustra un ejemplo para una regla de combinación inversa aplicada por el combinador 1162 en la figura 4a. En concreto, la regla de combinación es similar a la regla de combinación del lado del descodificador en la codificación central / lateral bien conocida, en donde $L = M + S$, y $R = M - S$. Se entiende que la señal S usada por la regla de combinación inversa en la figura 4c es la señal calculada por la calculadora de señales de combinación, es decir, la combinación de la señal de predicción en la línea 1163 y la señal residual descodificada 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 puesta en práctica físicamente conectada. En una puesta en práctica computarizada, no obstante, 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 figura 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 se ilustra en 201 y un segundo canal se ilustra en 202. Ambas señales se introducen en una calculadora de codificador 203 para calcular una primera señal de combinación 204 y una señal residual de predicción 205 usando la primera señal de canal 201 y la segunda señal de canal 202 y la información de predicción 206, de tal 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 usando una regla de combinación.

La información de predicción es generada por un optimizador 207 para calcular la información de predicción 206 de tal 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 descodificador de audio que se ilustra en la figura 1.

Dependiendo de la puesta en práctica, 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 figura 3a, las cuales se analizarán en lo sucesivo.

Un objetivo preferido de optimización se ilustra en la figura 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 se selecciona de tal modo que $\|S - \alpha M\|^2$ está reducida al mínimo. Esto tiene como resultado una solución para α que se ilustra la figura 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 se ilustra en la figura 3c. No obstante, 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 poner en práctica 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 puestas en práctica 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 pondría en práctica para llevar a cabo una operación de cuantificación / codificación de entropía con el fin de determinar la tasa de bits necesaria para determinados valores de α de tal modo que α se puede fijar para cumplir los requerimientos tales como una tasa mínima de bits, o como alternativa, una tasa fija de bits. Otras puestas en práctica del objetivo de optimización pueden estar relacionadas con un uso mínimo de los recursos del codificador o del descodificador. En el caso de una puesta en práctica 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.

La calculadora de codificador 203 en la figura 2 se puede poner en práctica en distintas formas, en donde una primera puesta en práctica ejemplar se ilustra en la figura 3a, en la cual una regla de combinación explícita se lleva a cabo en el combinador 2031. Una puesta en práctica alternativa ejemplar se ilustra en la figura 3b, en donde se usa una calculadora de matriz 2039. El combinador 2031 en la figura 3a se puede poner en práctica para llevar a cabo la regla de combinación que se ilustra en la figura 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. No obstante, se pueden poner en práctica otros factores de ponderación o ningún factor de ponderación dependiendo de la puesta en práctica. 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 descodificador 1162 que se ilustra en la figura 4a, el cual aplica una regla de combinación que es inversa a la regla de combinación aplicada por el codificador. Se puede usar 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 de codificador 203 es un proceso que conserva la forma de onda.

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. En concreto, 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 se ilustran en un ejemplo en la figura 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 de real a imaginario / transformada de imaginario a real, la cual se puede poner en práctica usando un filtro lineal tal como un filtro FIR que lleva a cabo adiciones ponderadas de determinados valores. La calculadora residual 2034 en la figura 3a puede llevar a cabo una operación de sustracción de tal modo que la señal de predicción se sustrae de la segunda señal de combinación. No obstante, son posibles otras operaciones en la calculadora residual. De la misma manera, la calculadora de señales de combinación 1161 en la figura 4a puede llevar a cabo una operación de adición en donde la señal residual descodificada 114 y la señal de predicción 1163 se agregan juntas para obtener la segunda señal de combinación 1165.

La figura 5a ilustra un codificador de audio. En comparación con el codificador de audio que se ilustra en la figura 3a, la primera señal de canal 201 es una representación espectral de una primera señal de canal en el dominio del tiempo 55a. De la misma manera, la segunda señal de canal 202 es una representación espectral de una señal de canal en el dominio del 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 se ponen en práctica como convertidores de valor real. El algoritmo de conversión puede ser una transformada de coseno discreta, una transformada FFT, en donde solo se usa la parte real, una MDCT o cualquier otra transformada que proporcione valores espectrales de valor real. Como alternativa, ambas transformadas se pueden poner en práctica como una transformada imaginaria, tal como una DST, una MDST o una FFT en donde solo se usa la parte imaginaria y se descarta la parte real. También se puede usar cualquier otra transformada que proporcione solo valores imaginarios. Un motivo por el cual usar 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, solo se tiene que procesar un único valor tal como magnitud o la parte real o, como alternativa, 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 al menos 2. En este punto, otro motivo para usar una transformada de valor real es que dicha transformada generalmente se muestrea críticamente, y de este modo proporciona un dominio adecuado (y comúnmente usado) para la cuantificación de señales y la codificación de entropía (el paradigma estándar de "codificación de audio perceptual" que se pone en práctica en "MP3", AAC, o en sistemas de codificación de audio similares).

La figura 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 figura 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

5 señal de audio multicanal codificada. En concreto, la operación de predicción se lleva a cabo de tal modo que la señal lateral se predice desde la señal central tal como se ilustra en las Ecuaciones a la derecha de la figura 5a. Preferentemente, la información de control del predictor 206 es un factor tal como se ilustra a la derecha en la figura 3b. En un ejemplo en el cual la información de control de predicción solo comprende una porción real tal como la parte real de un α de valor complejo o una magnitud del α de valor complejo, en donde esta porción corresponde a

10 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.

No obstante, cuando la información de control de predicción solo 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, la presente invención logra una ganancia de

15 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

20 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 predicho y un espectro imaginario predicho, y el espectro real predicho se sustraería del espectro real de la señal lateral S (en forma de banda), y el espectro imaginario predicho se sustraería de la parte imaginaria del

25 espectro de S para obtener un espectro residual complejo D.

Las señales en el dominio del tiempo L y R son señales de valor real, si bien las señales en el dominio de la frecuencia pueden ser de valor real o complejo. Cuando las señales en el dominio de la frecuencia son de valor real, entonces la transformada es una transformada de valor real. Cuando las señales en el dominio de la frecuencia son complejas, entonces la transformada es una transformada de valor complejo. Esto significa que la entrada a las

30 transformadas de tiempo a frecuencia y la salida de las transformadas de frecuencia a tiempo son de valor real, en tanto que las señales en el dominio de la frecuencia podrían ser, por ejemplo, señales en el dominio de QMF de valor complejo.

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

35 similares.

La salida de la corriente de bits por el multiplexor de la corriente de bits 212 en la figura 5a se introduce en un desmultiplexor de la corriente de bits 102 en la figura 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.

40 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 predicha $\alpha \cdot M$ y el combinador 1161 combina la señal residual emitida por el descuantificador 110b con la señal lateral predicha con el 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

45 figura 4c con respecto a la codificación central / lateral. En concreto, el bloque 1162 lleva a cabo una descodificación central / lateral (inversa) para obtener una representación en el dominio de la frecuencia del canal izquierdo y una representación en el dominio de la 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.

50 Dependiendo de la puesta en práctica 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.

No obstante, para aumentar la eficiencia se prefiere llevar a cabo una transformada de valor real tal como se ilustra en otra puesta en práctica en la figura 6a para el codificador de acuerdo con una forma de realización de la presente invención y en la figura 6b para el descodificador de acuerdo con una forma de realización de la presente invención.

55 Las transformadas de valor real 50 y 51 se ponen en práctica por medio de 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

60 imaginaria del espectro, se proporciona un convertidor de 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 de 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 α 2071 junto con el espectro real M con el 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. A continuación,

de acuerdo con la presente forma de realización, 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 que se ilustra 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 α en el cuantificador / codificador de entropía 2072 para obtener el valor de α complejo codificado, el cual se transmite al multiplexor de la corriente de bits 212 de la figura 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 α , se destaca que los multiplicadores 2073 y 2074 preferentemente usan exactamente el mismo α (cuantificado) que también se usará en el descodificador. De este modo se podría mover 2072 directamente hacia la salida de 2071, o se podría considerar que la cuantificación de α 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 de real a complejo en el bloque 2070 en el codificador de tal modo que se produzcan condiciones similares con respecto a un descodificador que se ilustra en la figura 6b. El descodificador 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 descodificación de entropía y una descuantificación en el bloque 65 para obtener la parte real α_R que se ilustra en 1160b y la parte imaginaria α_i que se ilustra en 1160c. Las señales centrales emitidas por los elementos de ponderación 1160b y 1160c se agregan a la señal residual de predicción descodificada y descuantificada. En concreto, los valores espectrales introducidos en el ponderador 1160c, en donde la parte imaginaria del factor de predicción compleja se usa como el factor de ponderación, provienen del espectro de valor real M por el convertidor de real a imaginario 1160a, el cual se pone en práctica preferentemente del mismo modo que el bloque 2070 de la figura 6a relacionada con el lado del codificador. En el lado del descodificador no hay una representación de valor complejo de la señal central o de la señal lateral, lo cual implica una diferencia con respecto al lado del codificador. El motivo es que solo los espectros de valor real codificados han sido transmitidos desde el codificador hacia el descodificador debido a las tasas de bits y por razones de complejidad.

El transformador de real a imaginario 1160a o el bloque correspondiente 2070 de la figura 6a se puede poner en práctica 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 con n.º 6.980.933. Como alternativa, se puede aplicar cualquier otra puesta en práctica conocida en la técnica, y una puesta en práctica preferida se describe en el contexto de las figuras 10a, 10b.

En concreto, tal como se ilustra en la figura 10a, el convertidor de real a imaginario 1160a comprende un selector de cuadro espectral 1000 conectado a una calculadora de 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 puesta en práctica, 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 solo se tiene que usar el cuadro actual para dicho cálculo, entonces el selector de cuadro espectral 1000 solo selecciona el cuadro actual i y transmite esta información a la calculadora de espectro imaginario. Por lo tanto, la calculadora de espectro imaginario solo usa 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 figura 10b. No obstante, 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 usar también para el cálculo del espectro imaginario, entonces la calculadora de espectro imaginario adicionalmente recibe los valores provenientes de los cuadros i - 1 e 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 descodificador. 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 usar cuadros "futuros" con el fin de reducir el retardo sistemático.

Adicionalmente, cabe destacar que la combinación ponderada en etapas que se ilustra en la figura 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 usar 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 poner en práctica 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 figuras 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 figura 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 figura 8a ilustra una puesta en práctica más detallada del bloque 50 o 51. En concreto, 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 usa una ventana de análisis y , en concreto, lleva a cabo esta operación en un cuadro mediante un modo de cuadros pero usando 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 tal modo que, por ejemplo, para las porciones transitorias en la señal, la longitud de ventana / longitud de transformada se reduce con el fin de obtener mejores resultados de codificación.

La figura 8b ilustra la operación de MDCT inversa llevada a cabo en los bloques 52 y 53. A modo de ejemplo, el bloque 52 comprende un bloque 520 para llevar a cabo una transformada de MDCT inversa de cuadro a 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 ventana 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 tal modo que un bloque de 2048 muestras finalmente logra 1024 muestras nuevas de la señal de salida libre de solapamiento. De nuevo se prefiere aplicar un control de longitud de ventana / transformada usando información, la cual por ejemplo se transmite en la información lateral de la señal multicanal codificada de acuerdo con 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. No obstante, 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 que se ilustra en la figura 9, el cual es por ejemplo un procesador de MDCT descrito en el contexto de la figura 8a, proporciona un espectro de resolución de alta frecuencia que tiene determinadas líneas espectrales que se ilustran en la figura 9b. Este espectro de resolución de alta frecuencia es usado por un selector de línea espectral 90 que proporciona un espectro de resolución de baja frecuencia, el cual comprende determinadas bandas B_1 , B_2 , B_3 , ..., B_N . Este espectro de resolución de baja frecuencia es transmitido al optimizador 207 para calcular la información de predicción de tal modo que no se calcula una información de predicción para cada una de las líneas espectrales, sino solo 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 usa el mismo valor α para todas las líneas espectrales en la banda.

Preferentemente, las bandas están configuradas en una forma psicoacústica de tal modo que el ancho de banda de las bandas aumenta desde frecuencias inferiores hasta frecuencias superiores tal como se ilustra en la figura 9b. Como alternativa, a pesar de que no se prefiere tanto como la puesta en práctica de ancho de banda en aumento también se podrían usar bandas de frecuencia de igual tamaño, en donde cada una de las bandas de frecuencia tiene al menos dos o, por lo general, muchas más, tal como al menos 30 líneas de frecuencia. Por lo general, para un espectro de 1024 líneas espectrales se calculan menos de 30 valores de α complejos, y preferentemente, más de 5 valores de α . Para los espectros con menos de 1024 líneas espectrales (por ejemplo, 128 líneas) se usan preferentemente menos bandas de frecuencia (por ejemplo, 6) para α .

Para calcular los valores de α no se requiere necesariamente el espectro de MDCT de alta resolución. Como alternativa, también se puede usar un banco de filtros que tiene una resolución de frecuencia similar a la resolución necesaria para calcular los valores de α . Cuando se tienen que poner en práctica bandas que aumentan de frecuencia, en ese caso dicho banco de filtros debería tener un ancho de banda variable. No obstante, cuando es suficiente un ancho de banda constante de frecuencias bajas a altas, entonces se puede usar un banco de filtros tradicional con sub-bandas de ancho equivalente.

Dependiendo de la puesta en práctica, el signo del valor de α indicado en la figura 3b o 4b se puede invertir. No obstante, con el fin de ser coherentes, es necesario que dicha reversión del signo se use en el lado del codificador así como también en el lado del descodificador. A diferencia de la figura 6a, la figura 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 usada en la figura 6a en los bloques 50, 51 se usa una transformada generalizada de tiempo / frecuencia en la figura 5a tal como se describe. Tal como se ha indicado en lo que antecede, la figura 6a es el proceso del codificador, el cual corresponde al proceso del descodificador en la figura 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. Como alternativa, 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 descodificador deberían coincidir exactamente con el fin de asegurar una correcta codificación de forma de onda. Esto se aplica preferentemente al caso en el cual dichos módulos usan alguna forma de aproximación tal como filtros truncados o cuando se usan solo uno o dos en lugar de los tres cuadros de MDCT, es decir, el cuadro de MDCT actual en la línea 60, el cuadro de MDCT anterior en la línea 61 y el cuadro de MDCT siguiente en la línea 62.

Adicionalmente, se prefiere que el módulo 2070 en el codificador en la figura 6a use el espectro de MDCT no cuantificado M como entrada, a pesar de que el módulo de real a imaginario (R21) 1160a en el descodificador tiene solo el espectro de MDCT cuantificado disponible como entrada. Como alternativa, se puede usar también una puesta en práctica en la cual el codificador usa los coeficientes de MDCT cuantificados como entrada en el módulo 2070. No obstante, 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, de forma más detallada, 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 motivado perceptualmente con variación de tiempo y frecuencia sin introducir distorsiones de solapamiento. No obstante, en el caso de la codificación de mezcla descendente / residual (tal como se usa 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 lo bastante bien.

No obstante, con el fin de poder aprovechar la eficiencia de codificación mejorada que se puede lograr en el caso de señales estéreo con diferencias de tiempo o fase entre canales por medio de un coeficiente de predicción de valor complejo α , se necesita una representación de valor complejo en el dominio de la frecuencia de la señal de mezcla descendente DMX como entrada en la matriz de mezcla ascendente de valor complejo. Esto se puede obtener usando 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 solo dos parámetros (parte real e imaginaria de α) en lugar de tres (ICC, CLD e 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 descodificador (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 , que se ilustra 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:

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

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

10 Existen dos opciones para calcular la señal residual de predicción en el codificador. Una opción consiste en usar 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 descodificador usan los mismos valores para generar la predicción. La otra opción consiste en usar los valores espectrales de MDCT no cuantificados. Esto implica que el codificador y el descodificador no usarán 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.

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

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

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

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

40 En caso de que el codificador de MDCT subyacente use 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. Con el fin de evitar la creciente complejidad computacional y estructural, el cuadro anterior / siguiente se puede excluir del filtrado en las transiciones de longitud de transformada, a expensas de una menor exactitud de la aproximación para los cuadros respectivos.

45 Asimismo, se debe tener especial cuidado con las partes más bajas y más altas del espectro de MDST (próximas a CC y $f_s / 2$), en donde hay menos coeficientes de MDCT circundantes para el filtrado de FIR que los necesarios. En el presente caso, el proceso de filtrado necesita adaptarse para computar el espectro de MDST correctamente. Esto se puede llevar a cabo, ya sea usando 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.

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

60 El retardo adicional se puede evitar usando 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:

- Solo un pequeño aumento en la complejidad computacional (cuando no se usa 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 solo 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 usar el cuadro siguiente).
- 20 Las propiedades importantes de una puesta en práctica 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 usados. 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) Solo dos parámetros, la parte real e imaginaria del coeficiente de predicción compleja α , se transmiten en lugar de ICC, CLD, e IPD. Las partes reales e imaginarias de α se tratan 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 usa 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 usando 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 puesta en práctica 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. De nuevo, esto se puede señalar usando una bandera de un bit.
- d) En lugar de reutilizar un código predefinido tal como el código de factor de escala AAC se puede usar 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 y optimizar de forma arbitraria 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 y optimizar de forma arbitraria 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) en lo que antecede, se puede señalar de forma explícita por medio de un bit por cuadro en la corriente de bits que solo se usa 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 usar 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 usa SBR) sin el aumento significativo en la complejidad computacional que ocurriría con un método basado en QMF.

5 Las dos listas siguientes resumen los aspectos preferidos de la configuración que se han descrito en lo que antecede, los cuales se pueden usar de forma alternativa entre sí o además de otros aspectos:

- 10 1a) concepto general: predicción compleja de MDCT lateral a partir de MDCT y MDST centrales;
- 15 1b) calcular / estimar la MDST a partir de MDCT ("R2I") en el dominio de frecuencia usando 1 o más cuadros (3 cuadros causan retardo);
- 20 1c) truncado del filtro (incluso descendiendo hasta 1 cuadro 2 conexiones, es decir, [-1 0 1]) para reducir la complejidad computacional;
- 25 1d) tratamiento adecuado de CC y $f_s / 2$;
- 30 1e) tratamiento adecuado de conmutación de forma de ventana;
- 35 1f) no usar un cuadro anterior / siguiente si este tiene un tamaño de transformada diferente;
- 40 1g) predicción basada en coeficientes de MDCT no cuantificados o cuantificados en el codificador;
- 45 2a) cuantificar y codificar la parte real e imaginaria del coeficiente de predicción compleja directamente (es decir, sin parametrización del Entorno de MPEG);
- 50 2b) usar cuantificador uniforme para esto (tamaño de paso, por ejemplo, 0,1);
- 55 2c) usar una resolución de frecuencia adecuada para los coeficientes de predicción (por ejemplo, 1 coeficiente por cada 2 Bandas de Factor de Escala);
- 60 2d) señalización económica en el caso de que todos los coeficientes de predicción sean reales;
- 65 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 sub-bandas para las dos señales de canal, en donde el combinador (2031), el predictor (2033) y la calculadora de señal residual (2034) están configurados para procesar cada una de las sub-bandas por separado de tal modo que la primera señal combinada y la señal residual se obtienen para una pluralidad de sub-bandas, 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 sub-bandas.

45 A pesar de que algunos aspectos se han descrito en el contexto de un aparato, queda claro que dichos aspectos solo 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.

50 En una forma de realización de la presente invención se aplica un tratamiento adecuado de conmutación de forma de ventana. En la figura 10a se ilustra que una información de forma de ventana 109 se puede introducir en la calculadora de espectro imaginario 1001. En concreto, la calculadora de espectro imaginario, la cual lleva a cabo la conversión de real a imaginario del espectro de valor real tal como el espectro de MDCT (como, por ejemplo, el elemento 2070 en la figura 6a o el elemento 1160a en la figura 6b) se puede poner en práctica como un filtro FIR o IIR. Los coeficientes FIR o IIR en este módulo de 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 (*Kaiser Bessel Derived*, KBD) 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 de 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 de MDST para una secuencia de ventana actual para distintas formas de ventana y distintas puestas en práctica de la mitad izquierda y la mitad derecha de la ventana.

Tabla A - Parámetros de Filtro de MDST para la Ventana Actual

5

Secuencia de Ventana Actual	Mitad Izquierda: Forma de Seno Mitad Derecha: Forma de Seno	Mitad Izquierda: Forma de KBD Mitad Derecha: Forma de KBD
SOLO_SECUENCIA_LARGA, SECUENCIA_CORTA_OCHO	[0,000000, 0,000000, 0,500000, 0,000000, -0,500000, 0,000000, 0,000000]	[0,091497, 0,581427, 0,000000, 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 de Seno Mitad Derecha: Forma de Seno	Mitad Izquierda: Forma de KBD Mitad Derecha: Forma de KBD
SOLO_SECUENCIA_LARGA, SECUENCIA_CORTA_OCHO	[0,045748, 0,057238, 0,540714, 0,000000, -0,540714, -0,057238, - 0,045748]	[0,045748, -0,057238, 0,540714, 0,000000 -0,540714, 0,057238, - 0,045748]
SECUENCIA_LARGA_DE_INICIO	[0,104763, 0,105207, 0,567861, 0,000000, -0,567861, -0,105207, - 0,104763]	[0,148406, 0,046553, 0,607863, 0,000000 -0,607863, -0,046553, - 0,148406]
SECUENCIA_LARGA_DE_PARADA	[0,148406, -0,046553, 0,607863, 0,000000, -0,607863, 0,046553, - 0,148406]	[0,104763, -0,105207, 0,567861, 0,000000, -0,567861, 0,105207, - 0,104763]
SECUENCIA_DE_PARADA_INICIO	[0,207421, 0,001416, 0,635010, 0,000000, -0,635010, -0,001416, - 0,207421]	[0,207421, -0,001416, 0,635010, 0,000000, -0,635010, 0,001416, - 0,207421]

Adicionalmente, la información de forma de ventana 109 proporciona información de la forma de ventana para la ventana anterior, cuando se usa 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 de MDST correspondientes para la ventana anterior.

10

Tabla B - Parámetros de Filtro de MDST para la Ventana Anterior

Secuencia de Ventana Actual	Mitad Izquierda de Ventana Actual: Forma de Seno	Mitad Izquierda de Ventana Actual: Forma de KBD
SOLO_SECUENCIA_LARGA, SECUENCIA_LARGA_DE_INICIO, SECUENCIA_CORTA_OCHO	[0,000000, 0,106103, 0,250000, 0,318310, 0,250000, 0,106103, 0,000000]	[0,059509, 0,123714, 0,186579, 0,213077, 0,186579, 0,123714, 0,059509]
SECUENCIA_LARGA_DE_PARADA, SECUENCIA_DE_PARADA_INICIO	[0,038498, 0,039212, 0,039645, 0,039790, 0,039645, 0,039212, 0,038498]	[0,026142, 0,026413, 0,026577, 0,026631, 0,026577, 0,026413, 0,026142]

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

15

La información de la forma de ventana, la cual se usa en el lado del descodificador 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 descodificador, la información de la forma de ventana 109 se extrae de la corriente de bits mediante el demultiplexor de la corriente de bits (por ejemplo 102 en la figura 5b) y se proporciona a la calculadora de espectro imaginario 1001 tal como se ilustra en la figura 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 use 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 usa para

5 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 de nuevo tiene un distinto tamaño de transformada en comparación con el cuadro actual, entonces solo el cuadro actual, es decir, los valores espectrales de la ventana actual, se usan para estimar el espectro imaginario.

10 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 que se ilustra por medio del elemento 2033 en la figura 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 descodificador. No obstante, en una forma

15 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 figura 3a de tal modo que un primer canal cuantificado y un segundo canal cuantificado sean la base para calcular la señal residual. Como alternativa, 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

20 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 figura 6a, la cual se puede poner en práctica dentro del predictor 2033 en la figura 3a, opere en los mismos datos cuantificados

25 que estén disponibles en el lado del descodificador. 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 descodificador usado para llevar a cabo la predicción inversa, es decir, para calcular la señal lateral desde la señal residual. Para tal fin, la primera señal de combinación tal como la señal M en la línea 204 en la figura 6a se cuantifica antes de ser introducida en el bloque 2070. Entonces, el espectro de MDST calculado usando

30 el espectro de MDCT cuantificado del cuadro actual, y dependiendo de la información de control, el espectro de MDCT cuantificado del cuadro anterior o siguiente, se introduce en el multiplicador 2074, y la salida del multiplicador 2074 de la figura 6a de nuevo 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.

35 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 usando 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 puesta en práctica, el rango de valores para la parte real y la parte

40 imaginaria del coeficiente de predicción compleja tiene un rango desde -3 a 3, lo que significa que se usan 60 o, dependiendo de los detalles de puesta en práctica, 61 pasos de cuantificación para la parte real y la parte imaginaria del coeficiente de predicción compleja.

45 Preferentemente, la parte real aplicada en el multiplicador 2073 en la figura 6a y la parte imaginaria 2074 aplicada en la figura 6a se cuantifican antes de ser aplicadas de tal modo que, de nuevo, se usa el mismo valor para la predicción en el lado del codificador así como se usa en el lado del descodificador. 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

50 coeficiente de predicción cuantificado en el lado del descodificador. Preferentemente, la cuantificación se aplica de tal 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 descodificador. De este modo se prefiere cuantificar la entrada en la calculadora de real a imaginario 2070 usando 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

55 en el elemento 2073 y en el elemento 2074. La cuantificación es la misma que se aplica en el cuantificador 2072. Adicionalmente, la señal lateral emitida por el bloque 2031 en la figura 6a también se puede cuantificar antes de los sumadores 2034a y 2034b. No obstante, 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.

60 En un ejemplo adicional se aplica 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. Con el

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 descodificador, la interfaz del descodificador 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, 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 descodificador de la figura 6b se desactivará y la predicción inversa solo tendrá lugar usando el coeficiente de predicción de valor real aplicado en el multiplicador 1160b en la figura 6b. Esto mismo ocurre para el lado del codificador en donde el elemento 2070 se desactivará y la predicción solo tendrá lugar usando el multiplicador 2073. Esta información lateral se usa preferentemente como un bit adicional por cuadro y el descodificador leerá este bit, cuadro por cuadro, con el fin de decidir si el convertidor de 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 menos complejidad para el descodificador para dicho cuadro logrando inmediatamente un menor consumo de batería de dicho procesador que se pone en práctica, 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 usar un parámetro de valor complejo α , 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} .

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

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

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

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

Estos elementos de datos se calculan en un codificador y se introducen en la información lateral de una señal de audio multicanal o estéreo. Los elementos se extraen de la información lateral en el lado del descodificador por medio de un extractor de información lateral y se usan para controlar la calculadora de descodificador 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 usa 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.

- 5 Para todos los coeficientes de predicción, la diferencia de un valor anterior (en tiempo o frecuencia) se codifica usando un código de Huffman. Los coeficientes de predicción no se transmiten para las bandas de predicción para las cuales se usó `cplx_pred_used = 0`.

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

10

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

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

- 15 Cabe destacar que la invención no solo se aplica a señales estéreo, es decir, señales multicanal que tienen solo dos canales, sino que también se aplica a dos canales de una señal multicanal que tiene tres o más canales tales como una señal 5.1 o 7.1.

20 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 puesta en práctica, las formas de realización de la invención se pueden poner en práctica en un hardware o en un software. La puesta en práctica se puede llevar a cabo usando un medio digital de almacenamiento, por ejemplo un disco flexible, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control legibles almacenadas electrónicamente en el mismo, las cuales colaboran (o son capaces de colaborar) con un sistema informático programable de tal modo que el método respectivo se lleve a cabo.

25

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 tal modo que uno de los métodos descritos en la presente se lleve a cabo.

30

En general, las formas de realización de la presente invención se pueden poner en práctica 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 un ordenador. El código de programa se puede almacenar, por ejemplo, en un portador legible por ordenador.

35

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 ordenador.

40

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 un ordenador.

- 45 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 ordenador) 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.

50

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

55

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

- 60 En algunas formas de realización se puede usar 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 con el 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 que se han descrito en lo que antecede son meramente 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 solo está limitada por el alcance de las siguientes reivindicaciones de la patente y no por los detalles específicos presentados a modo de descripción y explicación de las formas de realización incluidas en la presente.

REIVINDICACIONES

1. Descodificador de audio para descodificar una señal de audio multicanal codificada (100), comprendiendo la señal de audio multicanal codificada (100) una primera señal de combinación codificada (104) generada sobre la base de 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 (106) e información de predicción (108), que comprende:
- un descodificador de señales (110) para descodificar la primera señal de combinación codificada (104) para obtener una primera señal de combinación descodificada (112) y para descodificar la señal residual de predicción codificada (106) para obtener una señal residual descodificada (114); y una calculadora de descodificador (116) para calcular una señal de audio multicanal descodificada que tiene una primera señal de canal descodificada (117), y una segunda señal de canal descodificada (118) usando la señal residual descodificada (114), la información de predicción (108) y la primera señal de combinación descodificada (112), de tal modo que la primera señal de canal descodificada (117) y la segunda señal de canal descodificada (118) son por lo menos aproximaciones de la primera señal de canal y la segunda señal de canal de la señal de audio multicanal, en donde la información de predicción (108) comprende una porción imaginaria distinta de cero, en donde la calculadora de descodificador (116) comprende:
- un predictor (1160) para aplicar la información de predicción (108) a la primera señal de combinación descodificada (112) o a una señal (601) proveniente de la primera señal de combinación descodificada (112) 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 descodificada (114) y la señal de predicción (1163); y un combinador (1162) para combinar la primera señal de combinación descodificada (112) y la segunda señal de combinación (1165) para obtener la señal de audio multicanal descodificada que tiene la primera señal de canal descodificada (117) y la segunda señal de canal descodificada (118),
- en donde el predictor (1160) comprende un convertidor de real a imaginario (1160) para estimar (1160a) un espectro imaginario de la primera señal de combinación descodificada (112) usando una parte real de la primera señal de combinación descodificada (112) directamente en el dominio de la frecuencia usando un filtrado bidimensional, estando sometida la parte real de la primera señal de combinación descodificada (112) a una conmutación de ventana, en donde el predictor (1160) está configurado para multiplicar una parte imaginaria (601) de la primera señal de combinación descodificada (112) por la parte imaginaria de la información de predicción (108) para obtener al menos una parte de la señal de predicción (1163), y en donde el predictor (1160) está configurado para recibir una información de forma de ventana (109) y para usar diferentes coeficientes de filtro por el convertidor de real a imaginario (1160) para calcular el espectro imaginario de la primera señal de combinación descodificada (112), en donde los diferentes coeficientes de filtro dependen de diferentes formas de ventana que son indicadas por la información de forma de ventana (109), en donde los coeficientes de filtro que son usados por el predictor (1160) dependen de una ventana completa, y en donde se requiere un conjunto de coeficientes de filtro para cada tipo de ventana y para cada transición de ventana.
2. Descodificador de audio de acuerdo con la reivindicación 1, en donde la primera señal de combinación codificada (104) y la señal residual de predicción codificada (106) se han generado usando un solapamiento que genera la conversión espectral - tiempo, en donde el descodificador además comprende:
- un convertidor espectral - tiempo (52, 53) para generar una primera señal de canal en el dominio del tiempo y una segunda señal de canal en el dominio del tiempo usando un algoritmo de conversión espectral - tiempo que coincide con el algoritmo de conversión espectral - tiempo; un procesador de solapar / agregar (522) para llevar a cabo un procesamiento de solapar - agregar para la primera señal de canal en el dominio del tiempo y para la segunda señal de canal en el dominio del tiempo para obtener una primera señal en el dominio del tiempo sin solapamiento y una segunda señal en el dominio del tiempo sin solapamiento.
3. Descodificador de audio de acuerdo con una de las reivindicaciones anteriores, en donde la información de predicción (108) comprende adicionalmente un factor real distinto de cero, en donde el predictor (1160) está configurado para multiplicar la primera señal de combinación descodificada (112) por el factor real para obtener una primera parte de la señal de predicción (1163), y en donde la calculadora de señales de combinación (1161) está configurada para combinar de forma lineal la señal residual descodificada (114) y la primera parte de la señal de predicción (1163) y la al menos una parte de la señal residual de predicción.

4. Descodificador de audio de acuerdo con una de las reivindicaciones anteriores, en donde la primera señal de combinación codificada (104) o la primera señal de combinación descodificada (112) y la señal residual de predicción codificada (106) o la señal residual de predicción descodificada (114) comprenden, cada una, una primera pluralidad de señales de sub-bandas,
- 5 en donde la información de predicción (108) 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 sub-bandas diferentes de la primera señal de combinación descodificada (112),
- 10 en donde la calculadora de descodificador (116) o la calculadora de señales de combinación (1161) o el combinador (1162) están configurados para llevar a cabo un procesamiento en sub-bandas; y en donde el descodificador de audio además comprende un banco de filtros de síntesis (52, 53) para combinar señales de sub-bandas de la primera señal de combinación descodificada (112) y la segunda señal de combinación descodificada (1165) para obtener una primera señal descodificada en el dominio del tiempo y una segunda señal descodificada en el dominio del tiempo.
- 15 5. Descodificador de audio de acuerdo con la reivindicación 1, en donde la primera señal de combinación descodificada (112) comprende una secuencia de cuadros de señales de valor real, y en donde el predictor (1160) está configurado para estimar (1160a), como el espectro imaginario de la primera señal de combinación descodificada, una parte imaginaria del cuadro de señal actual usando solo el cuadro de señal de valor real actual o usando el cuadro de señal de valor real actual y, ya sea solo uno o más cuadros de señal de valor real precedentes o solo uno o más cuadros de valor real siguientes o usando 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.
- 20 6. Descodificador de audio de acuerdo con una de las reivindicaciones 1 a 5, en donde la señal de audio multicanal codificada (100) 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 (100) son de valor real, en donde el descodificador de audio está configurado para extraer el indicador real de la señal de audio multicanal codificada (100), y
- 25 30 en donde la calculadora de descodificador (116) está configurada para no calcular una señal imaginaria para un cuadro, para el cual el indicador real está indicando solo coeficientes de predicción de valor real.
7. Descodificador de audio de acuerdo con la reivindicación 1, en donde el espectro de un cuadro previo o el espectro de un cuadro siguiente es un espectro de MDCT, en donde los coeficientes de filtro que se aplican al espectro del cuadro previo o que se aplican al espectro del cuadro siguiente dependen solo de la mitad de ventana que se solapa con un cuadro actual, en donde solo se requiere un conjunto de coeficientes para cada tipo de ventana.
- 35 8. Descodificador de audio de acuerdo con la reivindicación 1, en donde un tipo de ventana es o bien una ventana de seno o bien una ventana Derivada de Kaiser Bessel y, sometida a una configuración dada de secuencia de ventana, el tipo de ventana puede ser una ventana larga, una ventana de inicio, una ventana de parada, una ventana de parada inicio o una ventana corta.
- 40 9. Descodificador de audio de acuerdo con la reivindicación 1, en donde una configuración de ventana puede ser una configuración de ventana de ventana larga, ventana corta, ventana de inicio, ventana de parada o ventana de parada inicio.
- 45 10. Descodificador de audio de acuerdo con la reivindicación 1, en donde el predictor (1160) está configurado para calcular un espectro de MDST como el espectro imaginario usando un espectro de MDCT de un cuadro actual como la parte real de la primera señal de combinación descodificada (112), en donde los coeficientes de filtro que son usados por el predictor (1160) son coeficientes de filtro de MDST y dependen de la forma de ventana de una mitad izquierda de la ventana actual y la mitad derecha de la ventana actual,
- 50 en donde: o bien la mitad izquierda es una forma de seno y la mitad derecha es una forma de seno, o bien la mitad izquierda es una forma Derivada de Kaiser Bessel y la mitad derecha es una forma Derivada de Kaiser Bessel, o bien
- 55 la mitad izquierda es una forma de seno y la mitad derecha es una forma Derivada de Kaiser Bessel, o bien la mitad izquierda es una forma Derivada de Kaiser Bessel y la mitad derecha es una forma de seno.
- 60 11. Descodificador de audio de acuerdo con la reivindicación 1, en donde el predictor (1160) está configurado para calcular un espectro de MDST como el espectro imaginario usando un espectro de MDCT de un cuadro actual usando los siguientes coeficientes de filtro de MDST que se seleccionan para una mitad izquierda y una mitad derecha correspondiente de una ventana actual y una secuencia de ventana actual correspondiente de acuerdo con la siguiente Tabla A:

Tabla A - Coeficientes de Filtro de MDST para la Ventana Actual

Secuencia de Ventana Actual	Mitad Izquierda: Forma de Seno Mitad Derecha: Forma de Seno	Mitad Izquierda: Forma de KBD Mitad Derecha: Forma de KBD
SOLO_SECUENCIA_LARGA, SECUENCIA_CORTA_OCHO	[0,000000, 0,000000, 0,500000, 0,000000, -0,500000, 0,000000, 0,000000]	[0,091497, 0,000000, 0,581427, 0,000000, -0,581427, 0,000000, - 0,091497]
SECUENCIA_LARGA_DE_INICIO	[0,102658, 0,103791, 0,567149, 0,000000, -0,567149, -0,103791, - 0,102658]	[0,150512, 0,047969, 0,608574, 0,000000, -0,608574, -0,047969, - 0,150512]
SECUENCIA_LARGA_DE_PARADA	[0,102658, -0,103791, 0,567149, 0,000000, -0,567149, 0,103791, - 0,102658]	[0,150512, -0,047969, 0,608574, 0,000000, -0,608574, 0,047969, - 0,150512]
SECUENCIA_DE_PARADA_INICIO	[0,205316, 0,000000, 0,634298, 0,000000, -0,634298, 0,000000, - 0,205316]	[0,209526, 0,000000, 0,635722, 0,000000, -0,635722, 0,000000, - 0,209526]
Secuencia de Ventana Actual	Mitad Izquierda: Forma de Seno Mitad Derecha: Forma de Seno	Mitad Izquierda: Forma de KBD Mitad Derecha: Forma de KBD
SOLO_SECUENCIA_LARGA, SECUENCIA_CORTA_OCHO	[0,045748, 0,057238, 0,540714, 0,000000, -0,540714, -0,057238, - 0,045748]	[0,045748, -0,057238, 0,540714, 0,000000, -0,540714, 0,057238, - 0,045748]
SECUENCIA_LARGA_DE_INICIO	[0,104763, 0,105207, 0,567861, 0,000000, -0,567861, -0,105207, - 0,104763]	[0,148406, 0,046553, 0,607863, 0,000000, -0,607863, -0,046553, - 0,148406]
SECUENCIA_LARGA_DE_PARADA	[0,148406, -0,046553, 0,607863, 0,000000, -0,607863, 0,046553, - 0,148406]	[0,104763, -0,105207, 0,567861, 0,000000, -0,567861, 0,105207, - 0,104763]
SECUENCIA_DE_PARADA_INICIO	[0,207421, 0,001416, 0,635010, 0,000000, -0,635010, -0,001416, - 0,207421]	[0,207421, -0,001416, 0,635010, 0,000000, -0,635010, 0,001416, - 0,207421]

- 5 12. Descodificador de audio de acuerdo con la reivindicación 11, en donde el predictor (1160) está configurado para calcular un espectro de MDST como el espectro imaginario usando, adicionalmente, un espectro de MDCT de un cuadro previo y usando los siguientes coeficientes de filtro de MDST que se seleccionan para una mitad izquierda de una ventana actual y una secuencia de ventana actual correspondiente de acuerdo con la siguiente Tabla B:

Tabla B - Coeficientes de Filtro de MDST para la Ventana Anterior

10

Secuencia de Ventana Actual	Mitad Izquierda de Ventana Actual: Forma de Seno	Mitad Izquierda de Ventana Actual: Forma de KBD
SOLO_SECUENCIA_LARGA, SECUENCIA_LARGA_DE_INICIO, SECUENCIA_CORTA_OCHO	[0,000000, 0,106103, 0,250000, 0,318310, 0,250000, 0,106103, 0,000000]	[0,059509, 0,123714, 0,186579, 0,213077, 0,186579, 0,123714, 0,059509]
SECUENCIA_LARGA_DE_PARADA, SECUENCIA_DE_PARADA_INICIO	[0,038498, 0,039212, 0,039645, 0,039790, 0,039645, 0,039212, 0,038498]	[0,026142, 0,026413, 0,026577, 0,026631, 0,026577, 0,026413, 0,026142]

13. Codificador de audio para codificar una señal de audio multicanal que tiene dos o más señales de canal, que comprende:
- 15 una calculadora de codificador (203) para calcular una primera señal de combinación (204) y una señal residual de predicción (205) usando una primera señal de canal (201) y una segunda señal de canal (202) y una información de predicción (206), de tal modo que una señal residual de predicción (205), cuando se combina con una señal de predicción (2035) proveniente de la primera señal de combinación (204) o una señal proveniente de la primera señal de combinación (204) y la información de predicción (206) da como resultado una segunda señal de combinación (2032), pudiendo derivarse la primera señal de combinación (204) y la segunda señal de combinación (2032) a partir de la primera señal de canal (201) y la segunda señal de canal (202) usando una regla de combinación;
- 20 un optimizador (207) para calcular la información de predicción (206), de tal modo que la señal residual de predicción (205) cumple un objetivo de optimización (208);

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

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

en donde la calculadora de codificador (203) comprende:

10 un combinador (2031) para combinar la primera señal de canal (201) y la segunda señal de canal (202) de dos formas diferentes para obtener la primera señal de combinación (204) y la segunda señal de combinación (2032);

un predictor (2033) para aplicar la información de predicción (206) a la primera señal de combinación (204) o una señal (600) proveniente de la primera señal de combinación (204) para obtener la señal de predicción (2035); y

15 una calculadora de señal residual (2034) para calcular la señal residual de predicción (205) por medio de la combinación de la señal de predicción (2035) y la segunda señal de combinación (2032),

en donde el predictor (2033) está configurado

20 para multiplicar la primera señal de combinación (204) por una parte real (2073) de la información de predicción (206) para obtener una primera parte de la señal de predicción (2035);

25 para estimar (2070) una parte imaginaria (600) de la primera señal de combinación usando la primera señal de combinación (204), en donde el predictor (2033) comprende un convertidor de real a imaginario (2070) para estimar, directamente en el dominio de la frecuencia, un espectro imaginario de la primera señal de combinación como la parte imaginaria (600) de la primera señal de combinación usando la primera señal de combinación (204) usando un filtrado bidimensional, estando sometida la primera señal de combinación (204) a una conmutación de ventana; y

30 para multiplicar la parte imaginaria (600) de la primera señal de combinación por una parte imaginaria (2074) de la información de predicción (206) para obtener una segunda parte de la señal de predicción (2035);

en donde la calculadora residual (2034) está configurada para combinar de forma lineal la primera parte de la señal de predicción (2035) o la segunda parte de la señal de predicción (2035) y la segunda señal de combinación (2032) para obtener la señal residual de predicción (205), y

35 en donde el predictor (2033) está configurado para recibir una información de forma de ventana (109) y para usar diferentes coeficientes de filtro para calcular, usando el convertidor de real a imaginario (2070), el espectro imaginario de la primera señal de combinación, en donde los diferentes coeficientes de filtro dependen de diferentes formas de ventana que son indicadas por la información de forma de ventana (109), en donde los coeficientes de filtro que son usados por el predictor (2033) dependen de una ventana completa, y en donde se requiere un conjunto de coeficientes de filtro para cada tipo de ventana y para cada transición de ventana.

40 14. Codificador de audio de acuerdo con la reivindicación 13, en donde el predictor (2033) comprende un cuantificador para cuantificar la primera señal de canal, la segunda señal de canal, la primera señal de combinación (204), o la segunda señal de combinación (2023) para obtener una o más señales cuantificadas, y en donde el predictor (2033) está configurado para calcular la señal residual de predicción (205) usando señales cuantificadas.

45 15. Codificador de audio de acuerdo con una de las reivindicaciones 13 a 14, en donde la primera señal de canal es una representación espectral de un bloque de muestras; en donde la segunda señal de canal es una representación espectral de un bloque de muestras, en donde las representaciones espectrales son o bien unas representaciones espectrales reales puras o bien unas representaciones espectrales imaginarias puras, en donde el optimizador (207) está configurado para calcular la información de predicción (206) como un factor de valor real distinto de cero y / o como un factor imaginario distinto de cero, y en donde la calculadora de codificador (203) está configurada para calcular la primera señal de combinación (204) y la señal residual de predicción (205), de tal modo que la señal de predicción (2035) se deriva de la representación espectral real pura o la representación espectral imaginaria pura usando el factor de valor real.

60 16. Codificador de audio de acuerdo con una de las reivindicaciones 13 a 15, en donde la primera señal de canal es una representación espectral de un bloque de muestras; en donde la segunda señal de canal es una representación espectral de un bloque de muestras, en donde las representaciones espectrales son o bien unas representaciones espectrales reales puras o bien unas representaciones espectrales imaginarias puras, en donde el optimizador (207) está configurado para calcular la información de predicción (206) como un factor de valor real distinto de cero y / o como un factor imaginario distinto de cero, y en donde el predictor (2033) de la calculadora de codificador (203) comprende el convertidor de real a imaginario

(2070) o un transformador de imaginario a real para derivar una representación espectral transformada a partir de la primera señal de combinación (204), y en donde la calculadora de codificador (203) está configurada para calcular la primera señal de combinación (204) y la señal residual de predicción (205), de tal modo que la señal residual de predicción (205) se deriva de la representación espectral transformada usando el factor imaginario.

17. Método para descodificar una señal de audio multicanal codificada (100), comprendiendo la señal de audio multicanal codificada (100) una primera señal de combinación codificada (104) generada sobre la base de 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 (106) e información de predicción (108), que comprende:

descodificar (110) la primera señal de combinación codificada (104) para obtener una primera señal de combinación descodificada (112), y descodificar la señal residual de predicción codificada (106) para obtener una señal residual descodificada (114); y

calcular (116) una señal de audio multicanal descodificada que tiene una primera señal de canal descodificada (117), y una segunda señal de canal descodificada (118) usando la señal residual descodificada (114), la información de predicción (108) y la primera señal de combinación descodificada (112), de tal modo que la primera señal de canal descodificada (117) y la segunda señal de canal descodificada (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 una porción imaginaria distinta de cero, en donde el cálculo de la señal de audio multicanal descodificada (116) comprende:

aplicar la información de predicción (108) a la primera señal de combinación descodificada (112) o a una señal (601) proveniente de la primera señal de combinación descodificada (112), para obtener una señal de predicción (1163);

calcular una segunda señal de combinación (1165) combinando la señal residual descodificada (114) y la señal de predicción (1163); y

combinar la primera señal de combinación (112) y la segunda señal de combinación (1165) descodificadas para obtener la señal de audio multicanal descodificada que tiene la primera señal de canal descodificada (117) y la segunda señal de canal descodificada (118),

en donde aplicar la información de predicción (108) comprende estimar (1160a) un espectro imaginario de la primera señal de combinación descodificada (112) usando una parte real de la primera señal de combinación descodificada (112) directamente en el dominio de la frecuencia usando un filtrado bidimensional en un convertidor de real a imaginario (1160), estando sometida la parte real de la primera señal de combinación descodificada (112) a una conmutación de ventana,

en donde la aplicación de la información de predicción (108) comprende multiplicar una parte imaginaria (601) de la primera señal de combinación descodificada (112) por la parte imaginaria de la información de predicción (108) para obtener al menos una parte de la señal de predicción (1163), y

en donde la aplicación de la información de predicción (108) comprende recibir una información de forma de ventana (109) y usar diferentes coeficientes de filtro por el convertidor de real a imaginario (1160) para calcular el espectro imaginario de la primera señal de combinación descodificada (112), en donde los diferentes coeficientes de filtro dependen de diferentes formas de ventana que son indicadas por la información de forma de ventana (109), en donde los coeficientes de filtro dependen de una ventana completa, y en donde se requiere un conjunto de coeficientes de filtro para cada tipo de ventana y para cada transición de ventana.

18. Método de codificación de una señal de audio multicanal que tiene dos o más señales de canal, que comprende:

calcular (203) una primera señal de combinación (204) y una señal residual de predicción (205) usando una primera señal de canal (201) y una segunda señal de canal (202) y una información de predicción (206), de tal modo que una señal residual de predicción, cuando se combina con una señal de predicción (2035) proveniente de la primera señal de combinación (204) o una señal proveniente de la primera señal de combinación (204) y la información de predicción (206) da como resultado una segunda señal de combinación (2032), pudiendo derivarse la primera señal de combinación (204) y la segunda señal de combinación (2032) a partir de la primera señal de canal (201) y la segunda señal de canal (202) usando una regla de combinación;

calcular (207) la información de predicción (206), de tal modo que la señal residual de predicción (205) cumple un objetivo de optimización (208);

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

combinar (212) la primera señal de combinación codificada (210), la señal residual de predicción codificada (211) y la información de predicción (206) para obtener una señal de audio multicanal codificada, en donde el cálculo (203) comprende:

- combinar la primera señal de canal (201) y la segunda señal de canal (202) de dos formas diferentes para obtener la primera señal de combinación (204) y la segunda señal de combinación (2032); aplicar la información de predicción (206) a la primera señal de combinación (204) o una señal (600) proveniente de la primera señal de combinación (204) para obtener una señal de predicción (2035); y
- 5 calcular la señal residual de predicción (205) por medio de la combinación de la señal de predicción (2035) y la segunda señal de combinación (2032),
en donde la aplicación de la información de predicción (206) comprende:
- 10 multiplicar la primera señal de combinación (204) por una parte real (2073) de la información de predicción (206) para obtener una primera parte de la señal de predicción (2035);
estimar (2070) una parte imaginaria (600) de la primera señal de combinación usando la primera señal de combinación (204) en donde la aplicación de la información de predicción (206) comprende estimar, por un convertidor de real a imaginario (2070), directamente en el dominio de la frecuencia, un espectro imaginario
- 15 de la primera señal de combinación como la parte imaginaria (600) de la primera señal de combinación usando la primera señal de combinación (204) por medio de un filtrado bidimensional, estando sometida la primera señal de combinación (204) a una conmutación de ventana; y
multiplicar la parte imaginaria (600) de la primera señal de combinación (204) por una parte imaginaria (2074) de la información de predicción (206) para obtener una segunda parte de la señal de predicción (2035);
- 20 en donde el cálculo de la señal residual (2034) comprende combinar de forma lineal la primera parte de la señal de predicción (2035) o la segunda parte de la señal de predicción (2035) y la segunda señal de combinación (2023) para obtener la señal residual de predicción (205), y
en donde la aplicación de la información de predicción (206) comprende recibir una información de forma de ventana (109) y usar, por el convertidor de real a imaginario (2070), diferentes coeficientes de filtro para calcular
- 25 el espectro imaginario de la primera señal de combinación, en donde los diferentes coeficientes de filtro dependen de diferentes formas de ventana que son indicadas por la información de forma de ventana (109), en donde los coeficientes de filtro dependen de una ventana completa, y en donde se requiere un conjunto de coeficientes de filtro para cada tipo de ventana y para cada transición de ventana.
- 30 19. Programa informático que está adaptado para llevar a cabo, cuando se ejecuta en un ordenador o en un procesador, el método de la reivindicación 17 o el método de la reivindicación 18.

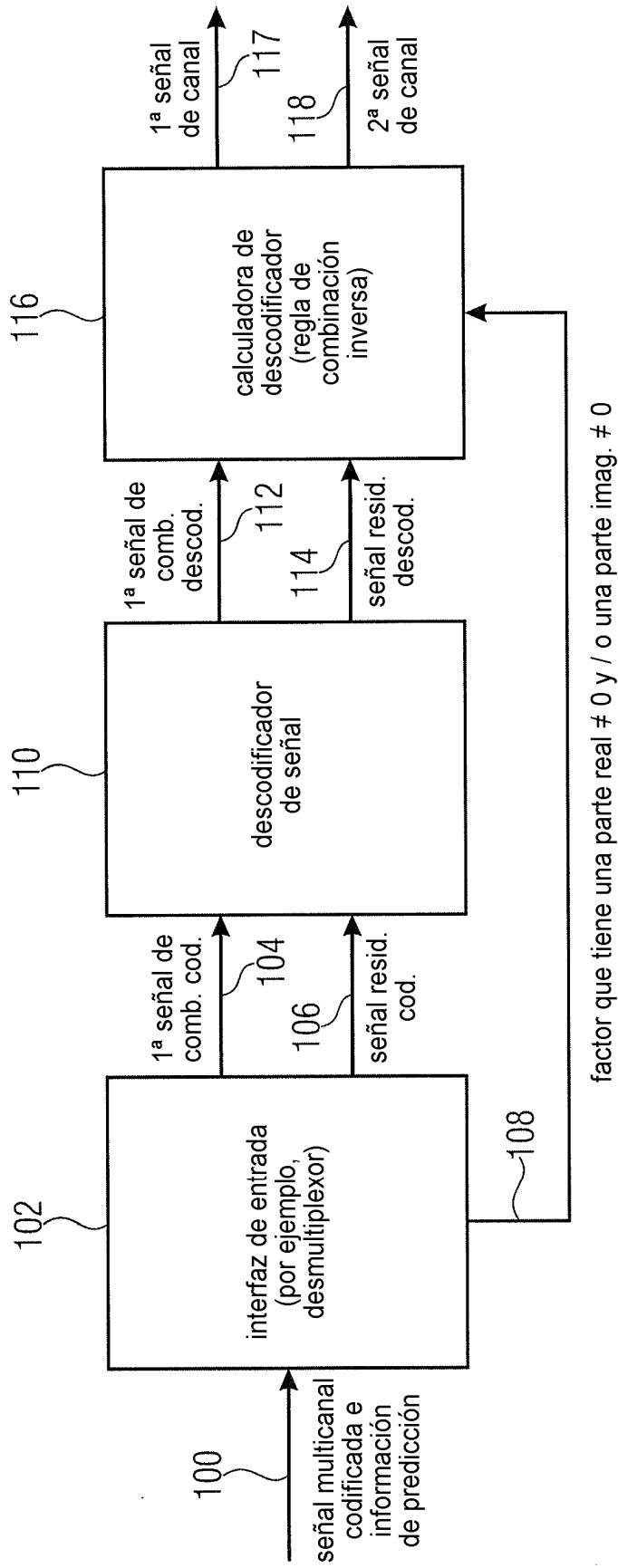


FIGURA 1
(DESCODIFICADOR DE AUDIO)

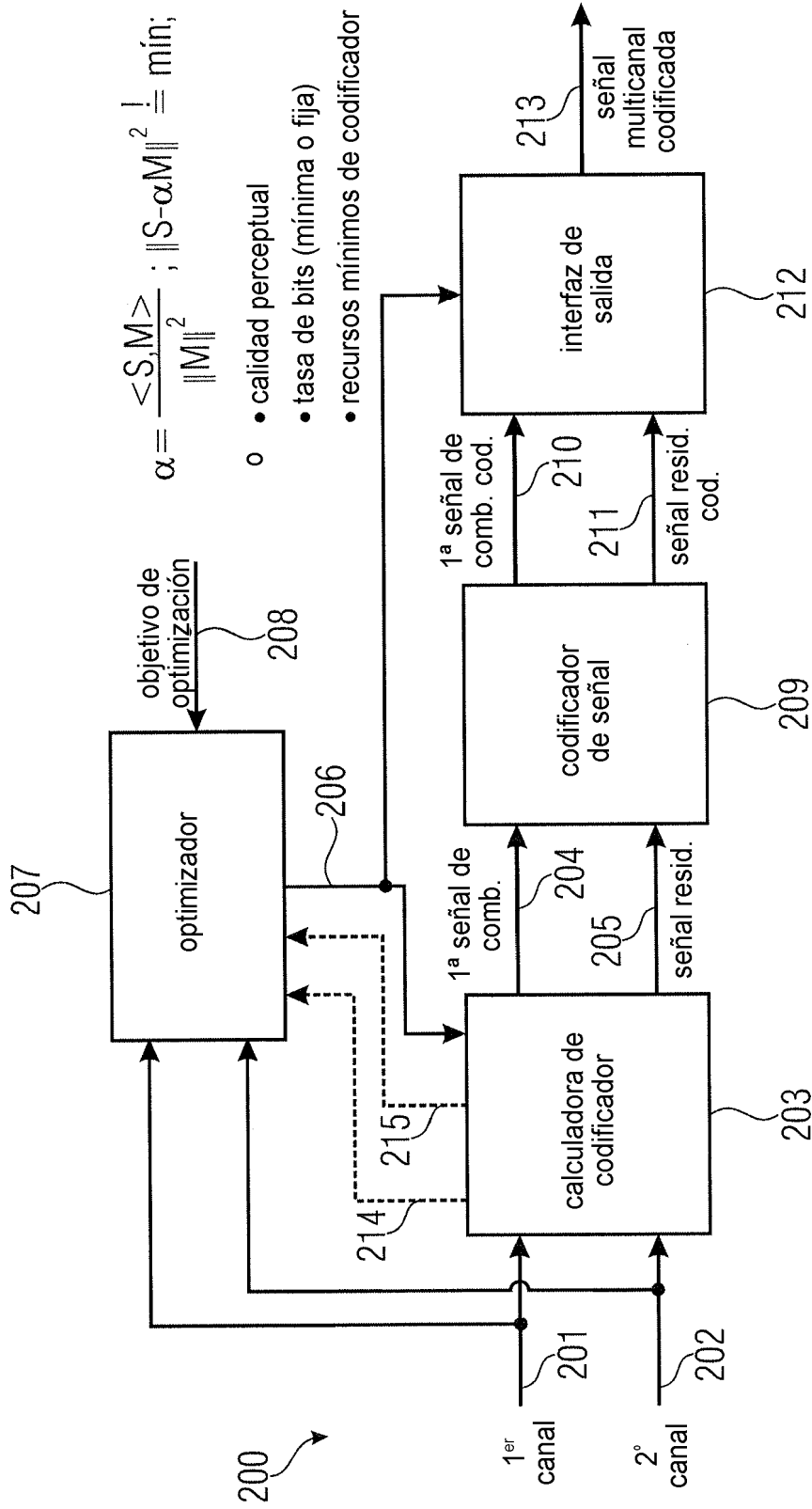


FIGURA 2
(CODIFICADOR DE AUDIO)

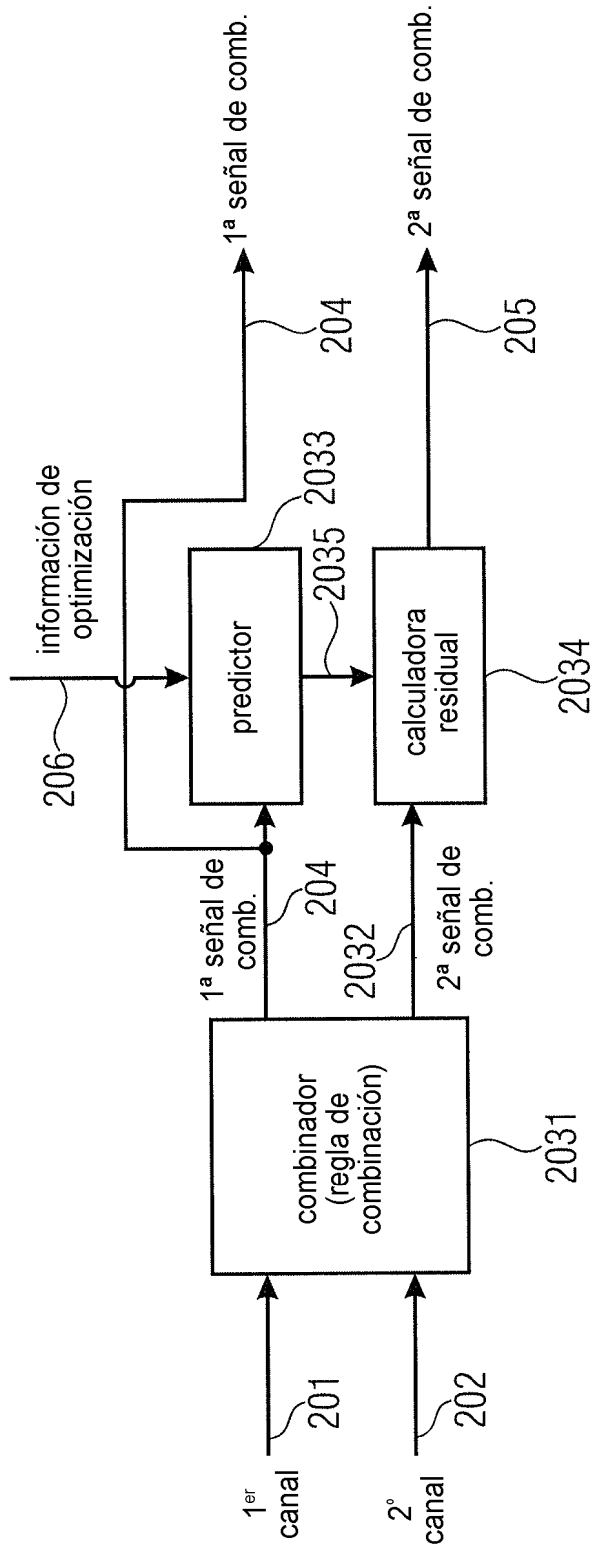
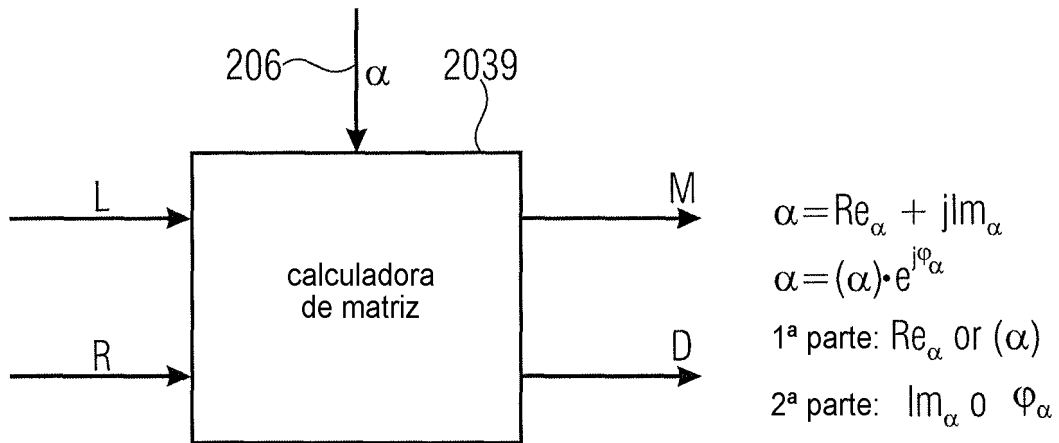


FIGURA 3A
(LADO DE CODIFICADOR DE AUDIO)



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

FIGURA 3B
(LADO DE CODIFICADOR DE AUDIO)

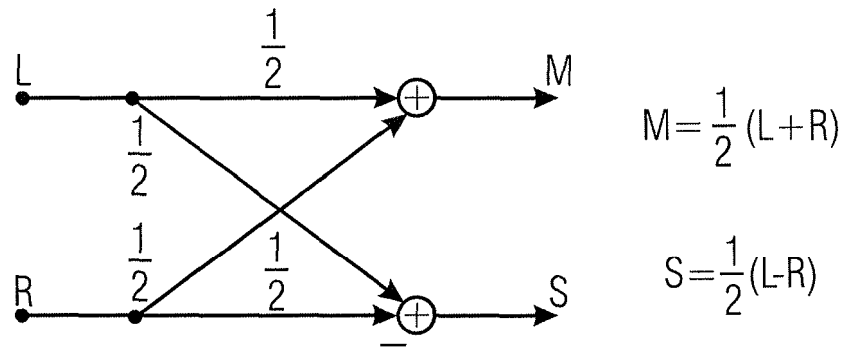


FIGURA 3C
(REGLA DE COMBINACIÓN)

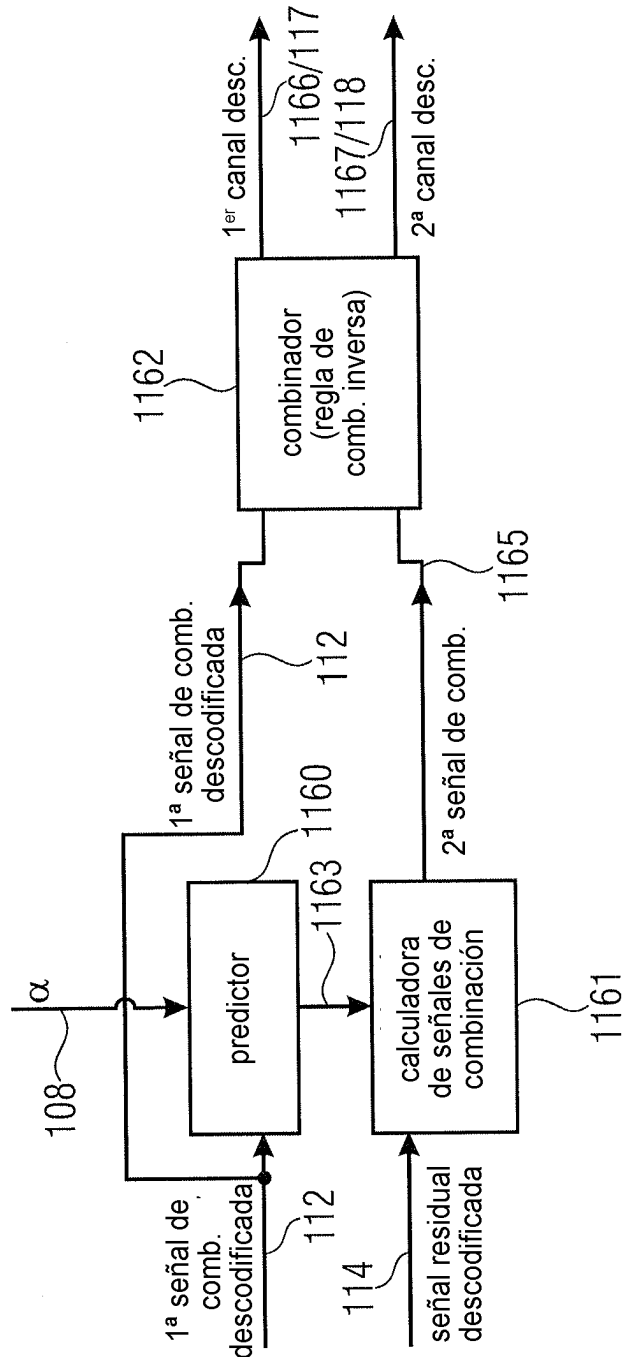


FIGURA 4A
(LADO DE DESCODIFICADOR DE AUDIO)

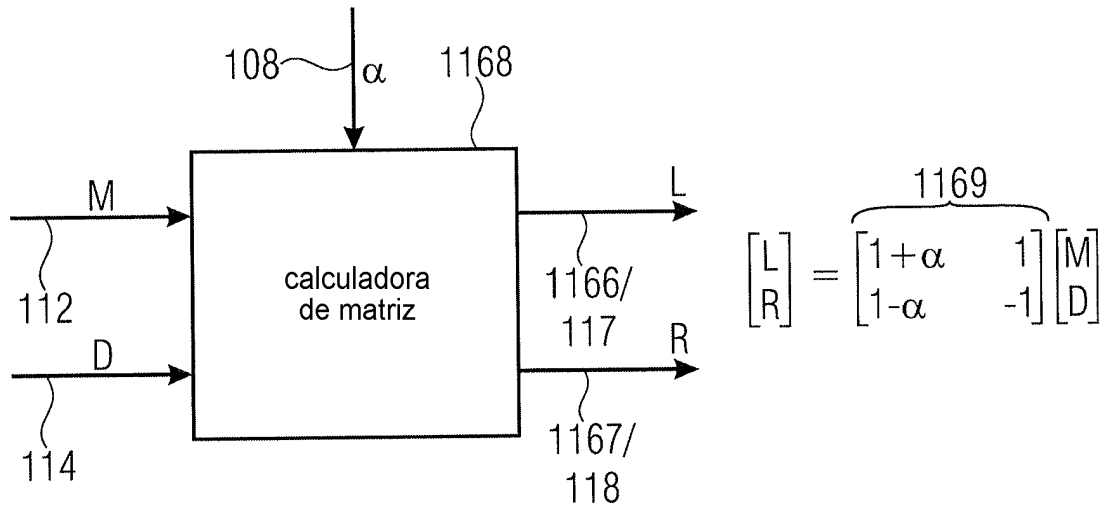


FIGURA 4B
(LADO DE DESCODIFICADOR DE AUDIO)

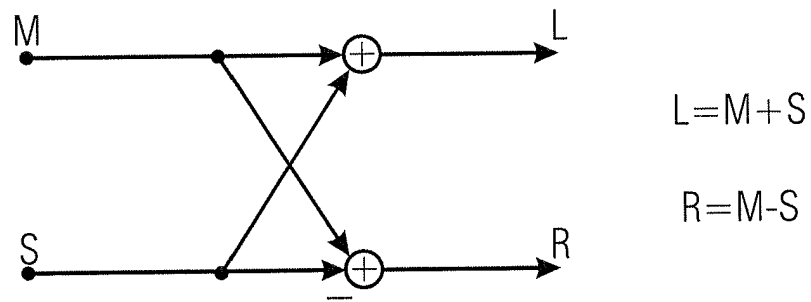


FIGURA 4C
(REGLA DE COMBINACIÓN INVERSA)

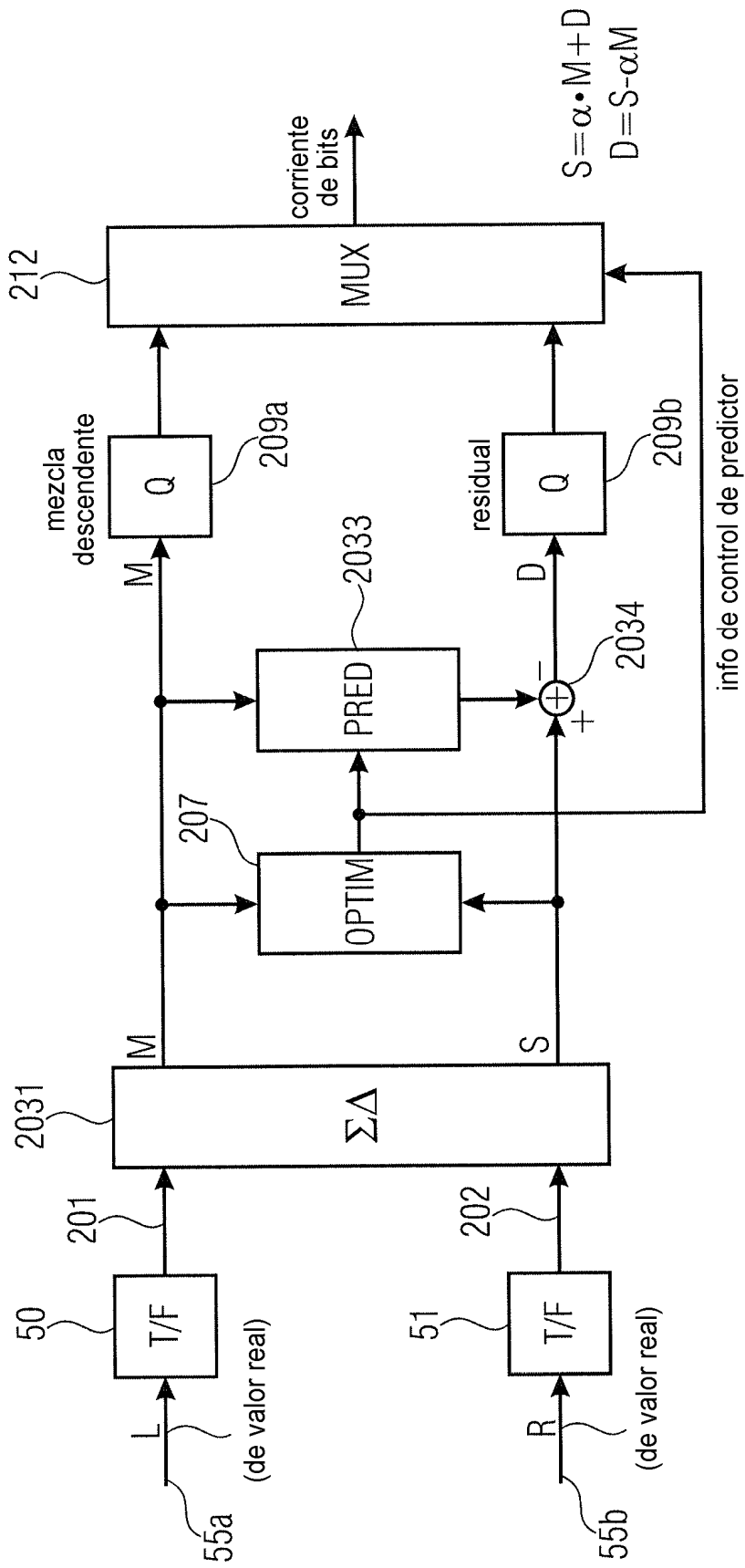


FIGURA 5A
(LADO DE CODIFICADOR)

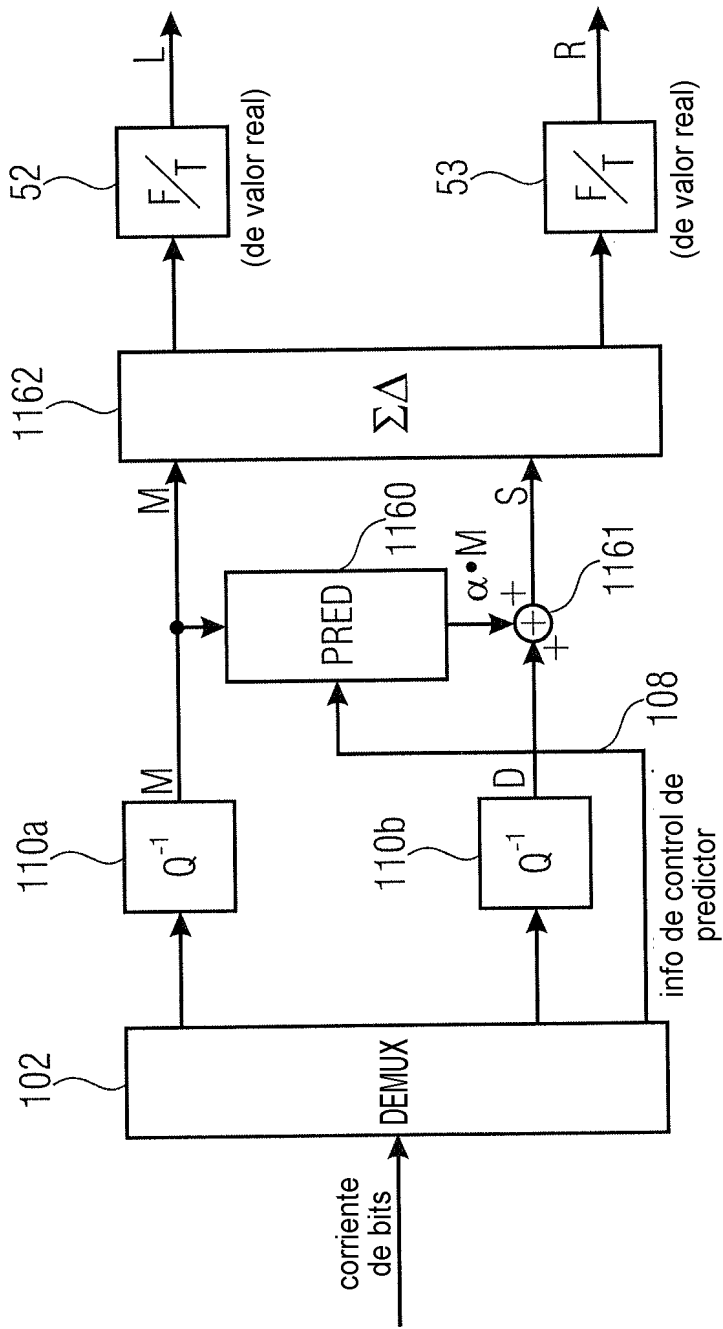


FIGURA 5B
(LADO DE DESCODIFICADOR)

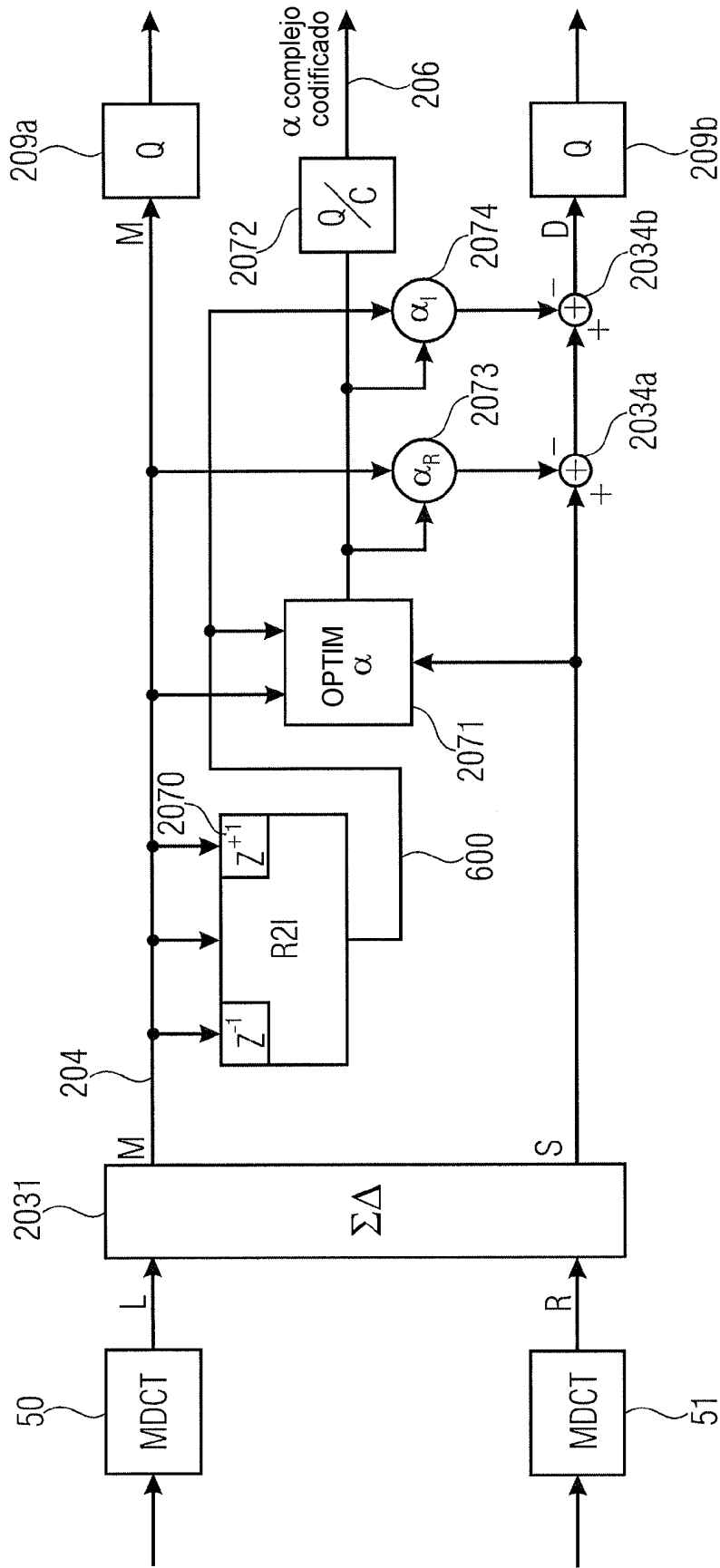


FIGURA 6A
(LADO DE CODIFICADOR)

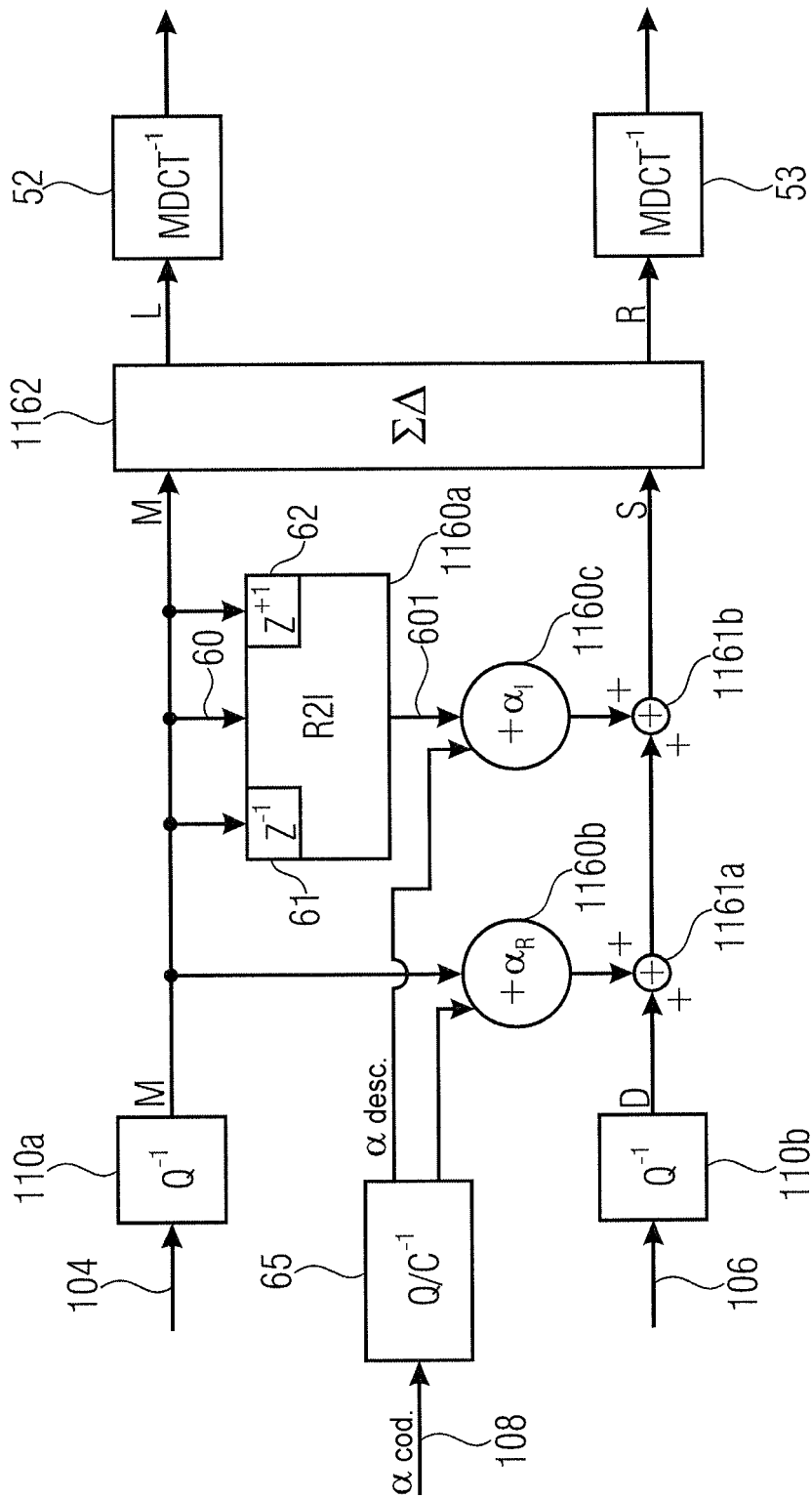


FIGURA 6B
(LADO DE DESCODIFICADOR)

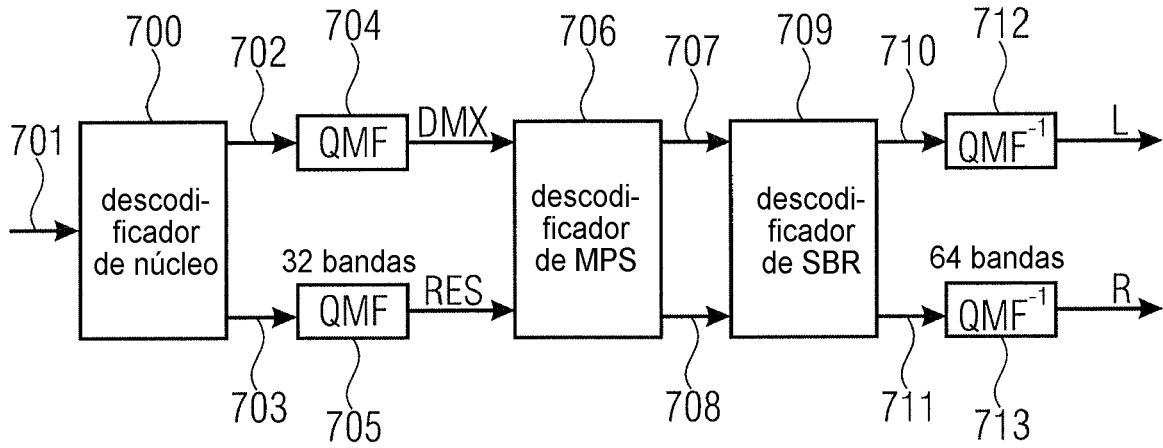


FIGURA 7A

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

FIGURA 7B

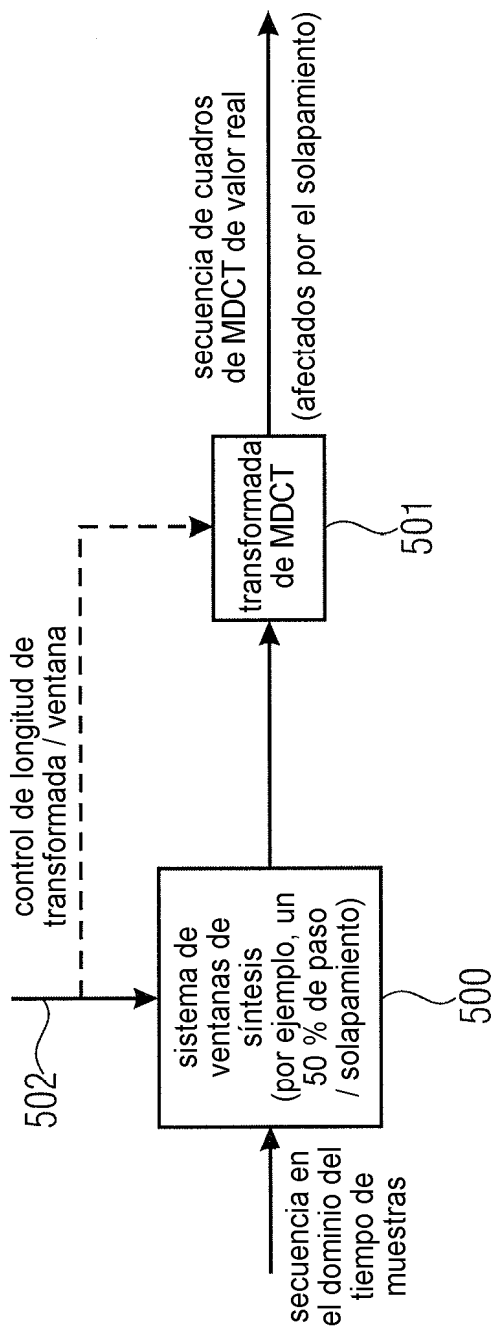


FIGURA 8A

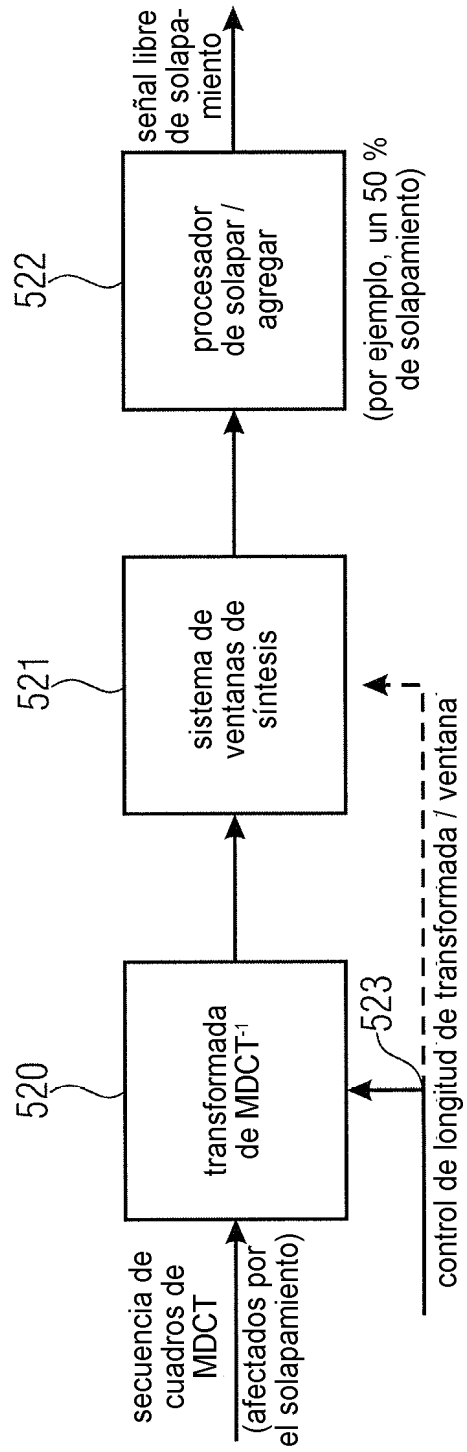
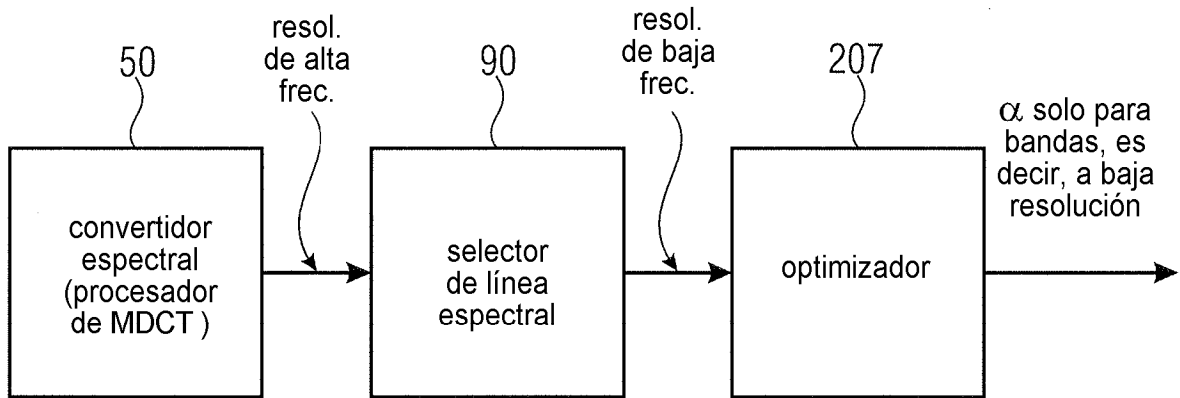
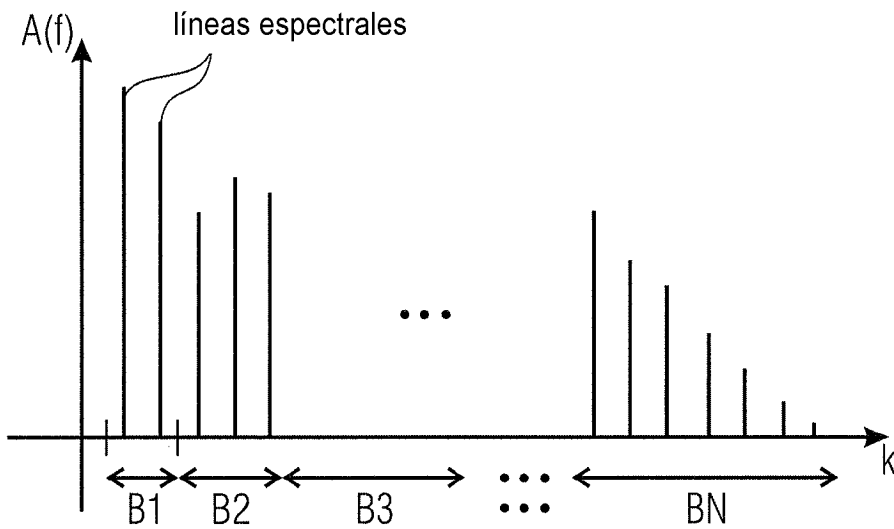


FIGURA 8B



- o: - un banco de filtros con anchos de banda en aumento
- o unas bandas de frec. de igual tamaño que tienen, cada una, al menos dos líneas espectrales

FIGURA 9A



- solo se calcula un único α para cada banda (por ejemplo, - 1024 líneas espectrales; - menos de 30 valores de α complejos (por ejemplo, 22))
- aumento de B_i con la frecuencia

FIGURA 9B

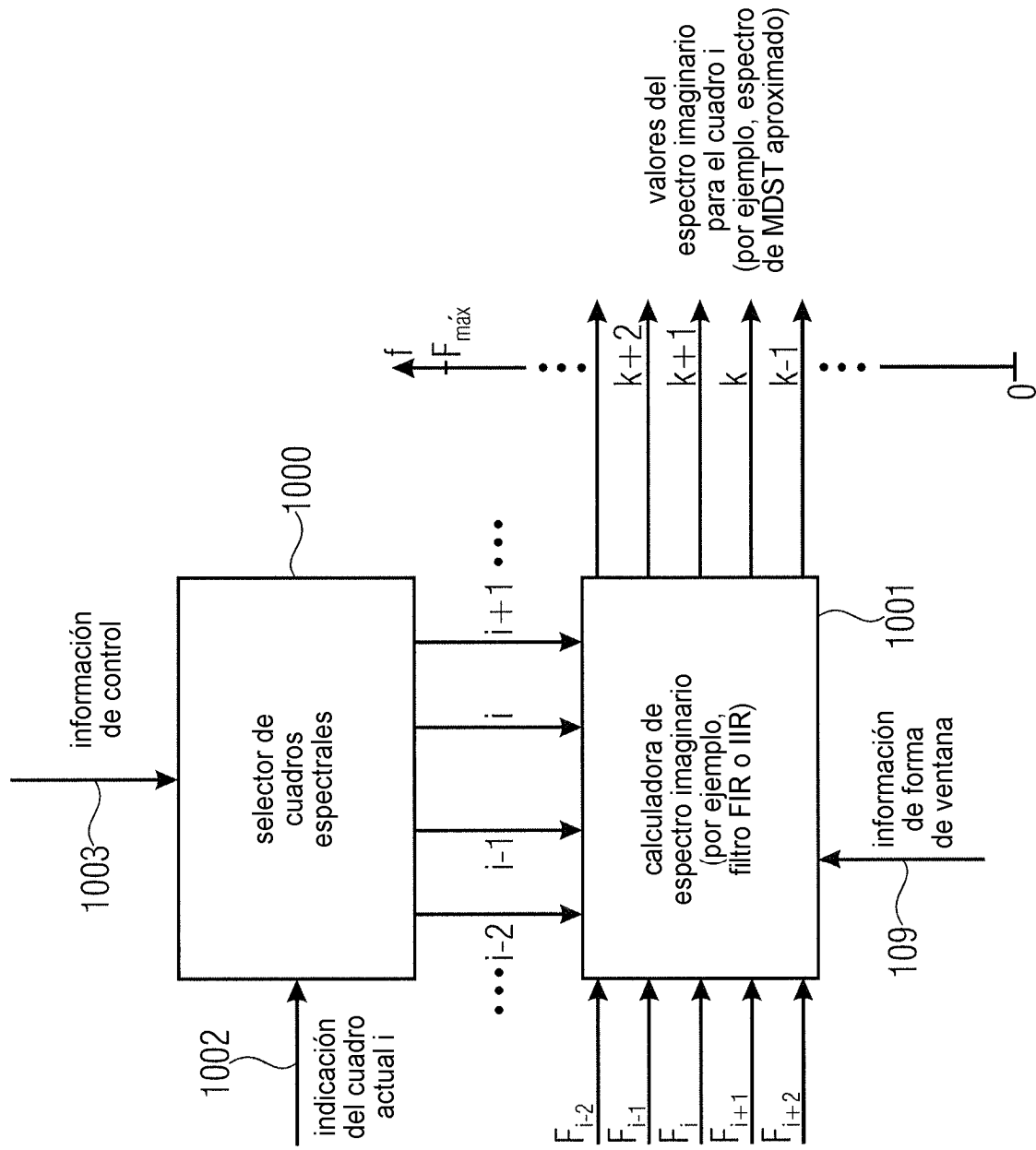


FIGURA 10A

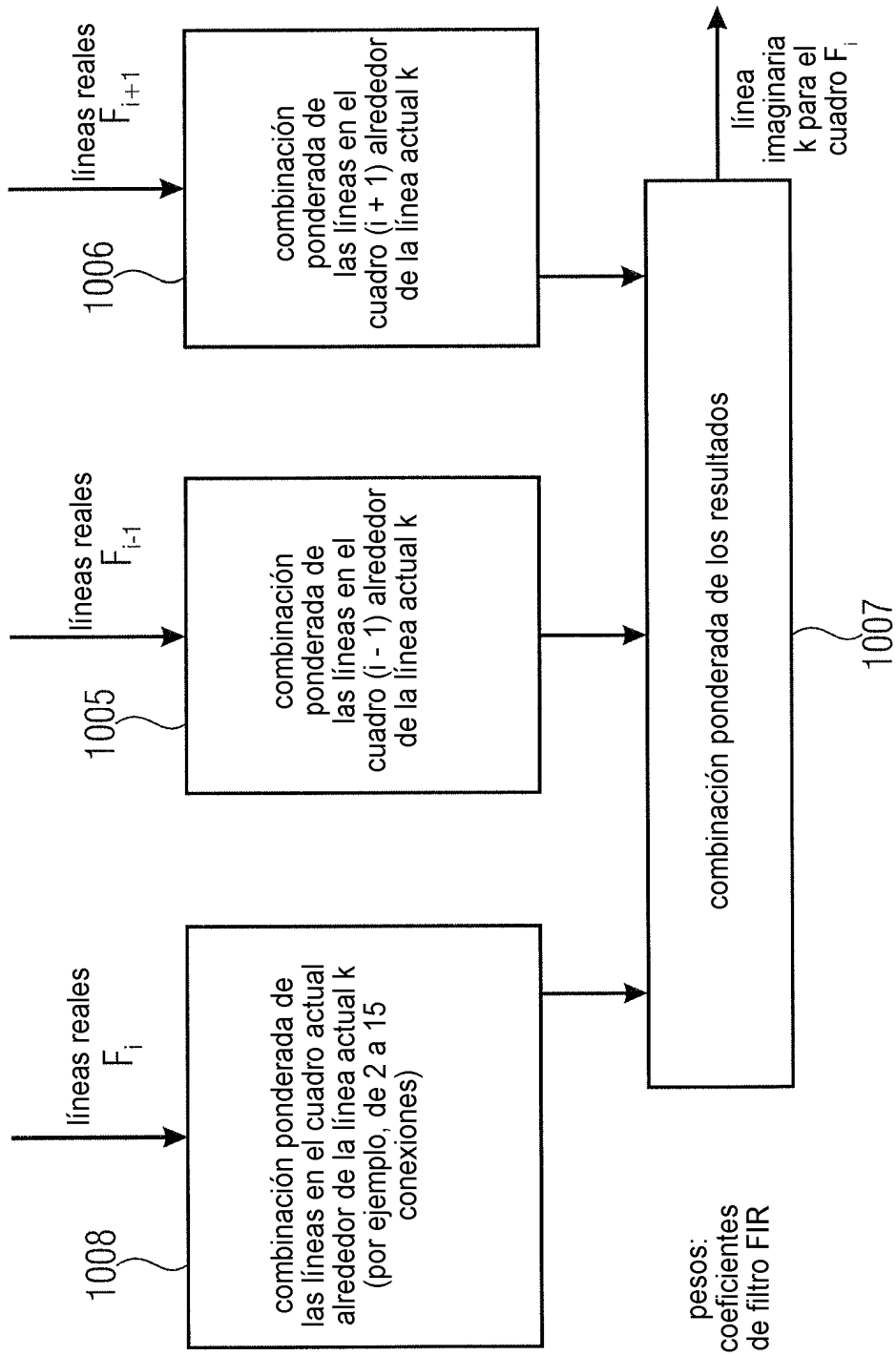


FIGURA 10B