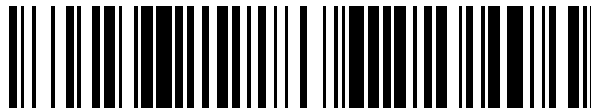


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 709 755**

51 Int. Cl.:

G10L 19/00 (2013.01)

G10L 19/008 (2013.01)

G10L 25/12 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **06.04.2011 PCT/EP2011/055369**

87 Fecha y número de publicación internacional: **13.10.2011 WO11124616**

96 Fecha de presentación y número de la solicitud europea: **06.04.2011 E 11713259 (7)**

97 Fecha y número de publicación de la concesión europea: **26.12.2018 EP 2556502**

54 Título: **Descodificación estéreo de predicción compleja basada en TCMD**

30 Prioridad:

09.04.2010 US 322458 P

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

17.04.2019

73 Titular/es:

**DOLBY INTERNATIONAL AB (100.0%)
Apollo Building, 3E, Herikerbergweg 1-35
1101 CN Amsterdam Zuidoost, NL**

72 Inventor/es:

**PURNHAGEN, HEIKO;
CARLSSON, PONTUS y
VILLEMOES, LARS**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 709 755 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

Descodificación estéreo de predicción compleja basada en TCMD

Campo técnico

5 La invención aquí descrita se refiere en general a la codificación de audio estéreo y, más precisamente, a técnicas para la codificación estéreo utilizando predicción compleja en el dominio en frecuencia.

Antecedentes de la invención

10 La codificación conjunta de los canales izquierdo (L) y derecho (R) de una señal estéreo permite una codificación más eficiente en comparación con la codificación independiente de L y R. Un enfoque común para la codificación estéreo conjunta es la codificación medio/lateral (M/S). Aquí una señal del canal medio (M) se forma agregando las señales L y R, por ejemplo, la señal M puede tener la forma

$$M = (L + R)/2$$

Además, se forma una señal lateral (S) al restar los dos canales L y R, por ejemplo, la señal S puede tener la forma

$$S = (L - R)/2$$

En el caso de la codificación M/S, las señales M y S están codificadas en lugar de las señales L y R.

15 En la norma MPEG (Grupo de Expertos en Imágenes en Movimiento) AAC (Codificación de Audio Avanzada) (véase el documento normalizado ISO/IEC 13818-7), la codificación estéreo L/R y la codificación estéreo M/S se pueden elegir en forma de variable en el tiempo y variable en frecuencia. Por lo tanto, el codificador estéreo puede aplicar la codificación L/R para algunas bandas de frecuencia de la señal estéreo, mientras que la codificación M/S se utiliza para codificar otras bandas de frecuencia de la señal estéreo (variable en frecuencia). Además, el codificador puede alternar en el tiempo entre la codificación L/R y M/S (variable en el tiempo). En MPEG AAC, la codificación estéreo se lleva a cabo en el dominio en frecuencia, más particularmente en el dominio MDCT (transformada discreta del coseno modificada). Esto permite elegir adaptativamente la codificación L/R o M/S en frecuencia y también en forma variable en el tiempo.

25 La codificación estéreo paramétrica es una técnica para codificar eficientemente una señal de audio estéreo como señal monoaural más una pequeña cantidad de información lateral para los parámetros estéreo. Forma parte de la norma de Audio MPEG-4 (consúltese el documento de norma ISO/IEC 14496-3). La señal monoaural se puede codificar utilizando cualquier codificador de audio. Los parámetros estéreo se pueden alojar en la parte auxiliar del flujo monoaural de bits, logrando así una compatibilidad total hacia adelante y hacia atrás. En el descodificador, es la señal monoaural la que se descodifica primero, después de lo cual la señal estéreo se reconstruye con la ayuda de los parámetros estéreo. Una versión no correlacionada de la señal monoaural descodificada, que tiene una correlación cruzada cero con la señal monoaural, se genera por medio de un des correlador, por ejemplo, un filtro de paso apropiado que puede incluir una o más líneas de retardo. Esencialmente, la señal no correlacionada tiene la misma distribución de energía espectral y temporal que la señal monoaural. La señal monoaural junto con la señal no correlacionada se introduce en el proceso de mezcla ascendente que se controla mediante los parámetros estéreo y que reconstruye la señal estéreo. Para obtener más información, consúltese el documento "Low Complexity Parametric Stereo Coding in MPEG-4" H. Purnhagen, Proc. de la 7ª Conferencia Internacional sobre Efectos Digitales de Audio (DAFx'04), Nápoles, Italia, 5-8, 2004 de octubre, páginas 163-168.

40 MPEG Surround (MPS; véase ISO/IEC 23003-1 y el documento "MPEG Surround - The ISO/MPEG Standard for Efficient and Compatible Multi-Channel Audio Coding", J. Herre et al., Audio Engineering Convention Paper 7084, Convención 122ª, 5-8 de mayo de 2007) permiten combinar los principios de la codificación estéreo paramétrica con la codificación residual, sustituyendo la señal no correlacionada con una residual transmitida y, por lo tanto, mejorar la calidad perceptiva. La codificación residual se puede lograr mezclando descendentemente una señal multicanal y, opcionalmente, extrayendo las señales espaciales. Durante el proceso de mezclado descendente las señales residuales que representan la señal de error se computan y posteriormente se codifican y transmiten. Pueden tomar el lugar de las señales no relacionadas en el descodificador. En un enfoque híbrido, pueden reemplazar las señales no correlacionadas en ciertas bandas de frecuencia, preferiblemente en bandas relativamente bajas.

50 De acuerdo con el actual sistema de Lenguaje Unificado y Codificación de Audio MPEG (USAC), de los cuales dos ejemplos se muestran en la figura 1, el descodificador comprende un banco de filtros espejo en cuadratura de valor complejo (QMF) situado aguas abajo del descodificador principal. La representación de QMF obtenida como salida del banco de filtros es compleja, - por lo que se muestrea en exceso por un factor dos - y puede disponerse como una señal de mezcla descendente (o, de manera equivalente, una señal del canal medio) M y una señal residual D,

a las que se aplica una matriz de mezcla ascendente con entradas complejas. Las señales L y R (en el dominio QMF) se obtienen como:

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

5 donde g es un factor de ganancia de valor real y α es un coeficiente de predicción de valor complejo. Preferiblemente, α se elige de tal manera que la energía de la señal residual D se minimice. El factor de ganancia puede determinarse mediante normalización, es decir, para garantizar que la potencia de la señal suma sea igual a la suma de las potencias de las señales izquierdo y derecho. Las partes real e imaginaria de cada una de las señales L y R son mutuamente redundantes - en principio, cada una de ellas puede ser calculada basándose en la otra - pero es beneficioso para permitir la aplicación posterior de un descodificador de replicación de banda espectral (SBR) sin que se produzcan errores de alias audibles. El uso de una representación de señal sobre muestreada también puede, por razones similares, ser elegida con el objetivo de prevenir errores conectados con otro proceso de señal adaptativa en tiempo o frecuencia (no se muestra), tal como la mezcla ascendente de monoaural a estéreo. El filtrado QMF inverso es el último paso de procesamiento en el descodificador. Se observa que la representación QMF limitada en banda de la señal permite técnicas residuales limitadas en banda y técnicas de "relleno residual", que pueden integrarse en descodificadores de este tipo.

La estructura de codificación anterior es muy adecuada para velocidades de bits bajas, generalmente por debajo de 80 kb/s, pero no es óptima para velocidades de bits más altas con respecto a la complejidad de cálculo. Más precisamente, a velocidades de bits más altas, la herramienta SBR generalmente no se utiliza (ya que no mejoraría la eficiencia de la codificación). Luego, en un descodificador sin una etapa SBR, solo la matriz de mezcla ascendente de valor complejo motiva la presencia del banco de filtros QMF, que es computacionalmente exigente e introduce un retardo (a una longitud de trama de 1024 muestras, el banco de filtros de análisis/síntesis QMF introduce un retardo de 961 muestras). Esto indica claramente la necesidad de una estructura de codificación más eficiente.

Heiko Purnhagen et al. "Technical description of proposed Unified Stereo Coding in USAC", 90. Reunión MPEG, 30-10-2009, Xian, China, describe un esquema de codificación estéreo unificado, en el que la señal de mezcla descendente y la señal residual se codifican conjuntamente por medio de un codificador principal USAC.

Resumen de la invención

Un objeto de la presente invención es proporcionar un sistema de descodificación de acuerdo con la reivindicación 1, un método de descodificación de acuerdo con la reivindicación 14 y un producto de programa de ordenador de acuerdo con la reivindicación 15 para la codificación estéreo que son computacionalmente eficientes también en el rango alto de velocidad de bits.

Las reivindicaciones dependientes definen realizaciones de la invención.

En un primer aspecto se describe un sistema descodificador, no contemplado por la invención reivindicada, para proporcionar una señal estéreo por codificación estéreo de predicción compleja, comprendiendo el sistema descodificador:

una mezcla ascendente adaptada para generar la señal estéreo basándose en las primeras representaciones en el dominio en frecuencia de una señal de mezcla descendente (M) y una señal residual (D), cada una de cuyas primeras representaciones en el dominio en frecuencia que comprenden los primeros componentes que representan el contenido espectral de la señal correspondiente expresada en un primer subespacio de un espacio multidimensional, comprendiendo la etapa de mezcla ascendente:

un módulo para calcular una segunda representación en el dominio en frecuencia de la señal de mezcla descendente basándose en la primera representación en el dominio en frecuencia de la misma, comprendiendo la segunda representación en el dominio en frecuencia los segundos componentes espectrales que representan el contenido espectral de la señal expresada en un segundo subespacio del espacio multidimensional que incluye una parte del espacio multidimensional que no está incluido en el primer subespacio;

un sumador ponderado para calcular una señal lateral (S) basándose en las primera y segunda representaciones en el dominio en frecuencia de la señal de mezcla descendente, en la primera representación en el dominio en frecuencia de la señal residual y un coeficiente de predicción complejo (α) codificado en la señal de flujo de bits; y

una etapa de suma y diferencia para calcular la señal estéreo basándose en la primera representación en el dominio en frecuencia de la señal de mezcla descendente y de la señal lateral,

en el que la etapa de mezcla ascendente funciona además en modo de paso, en el cual dichas señales de mezcla descendente y residuales se suministran directamente a la suma y diferencia.

En un segundo aspecto un ejemplo proporciona un sistema codificador, no contemplado por la invención reivindicada, para codificar una señal estéreo por medio de una señal de flujo de bits mediante codificación estéreo de predicción compleja, que incluye:

un estimador para estimar un coeficiente de predicción complejo;

una etapa de codificación que sirve para:

(a) transformar la señal estéreo en una representación en el dominio en frecuencia de una mezcla descendente y de una señal residual, en una relación determinada por el valor del coeficiente de predicción complejo;

y

un multiplexor para recibir la salida de la etapa de codificación y del estimador y codificar esto por medio de dicha señal de flujo de bits.

En un tercer y cuarto aspectos se proporcionan métodos, no contemplados por la invención reivindicada, para codificar una señal estéreo en un flujo de bits y para descodificar un flujo de bits en al menos una señal estéreo. Las características técnicas de cada método son análogas a los del sistema codificador y al sistema descodificador, respectivamente.

La invención se beneficia de las ventajas de la codificación estéreo unificada en el sistema MPEG USAC. Estas ventajas se conservan también a velocidades de bits más altas, a las que normalmente no se utiliza SBR, sin el aumento significativo de complejidad computacional que acompañaría un enfoque basado en QMF. Esto es posible porque la transformada MDCT muestreada críticamente, que es la base del sistema de codificación de la transformada MPEG USAC, puede usarse para la codificación estéreo de predicción compleja según lo que proporciona la invención, al menos en los casos en que los anchos de banda de audio del código de los canales de mezcla descendente y residuales son los mismos y el proceso de mezcla ascendente no incluye la no correlación. Esto significa que ya no se requiere una transformada QMF adicional. Una realización representativa de la codificación estéreo de predicción compleja en el dominio QMF haría en realidad aumentar el número de operaciones por unidad de tiempo significativamente en comparación con el estéreo tradicional L/R o M/S. Por lo tanto, el aparato de codificación de acuerdo con la invención aparenta ser competitivo a tales velocidades de bits, proporcionando una alta calidad de audio a un costo computacional moderado.

Como el experto en la materia puede apreciar, el hecho de que la etapa de mezcla ascendente funcione además en modo de paso permite que el descodificador descodifique de forma adaptativa de acuerdo con la codificación directa o conjunta convencional y la codificación de predicción compleja, según se determine en el lado del codificador. Por lo tanto, en aquellos casos en que el codificador no pueda aumentar positivamente el nivel de calidad más allá de la codificación estéreo directa L/R o la codificación estéreo M/S conjunta, puede al menos garantizar que se mantiene el mismo nivel. Por lo tanto, un descodificador de acuerdo con este aspecto de la invención puede, desde un punto de vista funcional, ser considerado como un superconjunto en relación con la técnica anterior.

Como ventaja sobre el estéreo con codificación de predicción basada en QMF, es posible la perfecta reconstrucción de la señal (aparte de los errores de cuantificación, que pueden hacerse arbitrariamente pequeños).

Por lo tanto, un ejemplo, no contemplado por la invención reivindicada, proporciona un aparato de codificación para la codificación estéreo basada en transformadas por predicción compleja. Preferiblemente, un aparato, no contemplado por la invención reivindicada, no se limita a la codificación estéreo de predicción compleja, sino que sirve también en una codificación estéreo L/R directa o en un régimen de codificación estéreo M/S conjunta de acuerdo con la técnica anterior, de modo que es posible seleccionar el método de codificación más adecuado para una aplicación en particular o durante un intervalo de tiempo en particular.

Una representación sobre muestreada (es decir, compleja) de la señal, que incluye tanto dichos primer y segundo componentes espectrales, se utiliza como base para la predicción compleja de acuerdo con la invención, y, por lo tanto, los módulos para calcular dicha representación de sobre muestreo se encuentran en los ejemplos de sistema codificador y de sistema descodificador no contemplados por la invención reivindicada. La composición espectral se refiere a los primero y segundo subespacios de un espacio multidimensional, que puede ser el conjunto de funciones dependientes del tiempo en un intervalo de longitud dada (por ejemplo, una longitud de trama de tiempo predefinida) muestreada a una frecuencia de muestreo finita. Es bien sabido que las funciones en este espacio multidimensional particular pueden ser aproximadas por medio de una suma ponderada finita de funciones básicas.

Como apreciarán los expertos, un codificador adaptado para cooperar con un descodificador está equipado con módulos equivalentes para proporcionar la representación sobre muestreada en la que se basa la codificación de predicción, así como para habilitar la fiel reproducción de la señal codificada. Dichos módulos equivalentes pueden ser módulos idénticos o similares o módulos con idénticas o similares características de transferencia. En particular, los módulos en el codificador y descodificador, respectivamente, pueden ser unidades de procesamiento similares o diferentes que ejecutan programas de ordenador respectivos que realizan conjuntos equivalentes de operaciones matemáticas.

En algunas realizaciones del sistema descodificador los primeros componentes espectrales tienen valores reales expresados en el primer subespacio y los segundos componentes espectrales tienen valores imaginarios expresados en el segundo subespacio. El primero y el segundo componentes espectrales juntos forman una representación espectral compleja de la señal. El primer subespacio puede ser la amplitud lineal de un primer conjunto de funciones base, mientras que el segundo subespacio puede ser la amplitud lineal de un segundo conjunto de funciones base, algunas de las cuales son linealmente independientes del primer conjunto de funciones base.

En una realización, el módulo para calcular la representación compleja es una transformada real a imaginaria, es decir, un módulo para calcular las partes imaginarias del espectro de una señal discreta en el tiempo sobre la base de una representación real de la señal. La transformada puede basarse en relaciones matemáticas exactas o aproximadas, tales como fórmulas de análisis armónico o relaciones heurísticas.

En algunas realizaciones del sistema descodificador, los primeros componentes espectrales son obtenibles por una transformada en el dominio tiempo a frecuencia, preferiblemente una transformada de Fourier, de una señal discreta en el dominio en el tiempo, tal como por una transformada discreta del coseno (DCT), una transformada discreta del coseno modificada (MDCT), una transformada discreta del seno (DST), una transformada discreta del seno modificada (MDST), una transformada rápida de Fourier (FFT), un algoritmo de Fourier basado en un factor primo o similar. En los primeros cuatro casos, los segundos componentes espectrales se pueden obtener mediante DST, MDST, DCT y MDCT, respectivamente. Como es bien sabido, el intervalo lineal de los cosenos que son periódicos en el intervalo unitario forma un subespacio que no está completamente contenido en el intervalo lineal de los senos periódicos en el mismo intervalo. Preferiblemente, los primeros componentes espectrales pueden obtenerse mediante MDCT y los segundos componentes espectrales pueden obtenerse mediante MDST.

En una realización, el sistema descodificador incluye al menos un módulo temporal de formación de ruido (módulo TNS o filtro TNS), que se dispone aguas arriba de la etapa de mezcla ascendente. En términos generales, el uso del TNS aumenta la calidad de audio percibido para señales con componentes de tipo transitorio, y esto también se aplica a las realizaciones del sistema de descodificación de la invención que incluye TNS. En la codificación estéreo L/R y M/S convencional, el filtro TNS se puede aplicar como una última etapa de procesamiento en el dominio en frecuencia, directamente antes de la transformada inversa. Sin embargo, en el caso de la codificación estéreo de predicción compleja, a menudo es ventajoso aplicar el filtro TNS en las señales de mezcla descendente y residual, es decir, antes de la matriz de mezcla ascendente. Dicho de otra manera, el TNS se aplica a combinaciones lineales de los canales izquierdo y derecho, lo que tiene diversas ventajas. En primer lugar, puede dar lugar a una situación dada que el TNS solo sea beneficioso para, por ejemplo, la señal de mezcla descendente. Entonces, el filtrado TNS se puede suprimir u omitir para la señal residual y, lo que puede significar un uso más económico del ancho de banda disponible, los coeficientes del filtro TNS solo necesitan transmitirse para la señal de mezcla descendente. En segundo lugar, el cálculo de la representación sobre muestreada de la señal de mezcla descendente (por ejemplo, los datos MDST que se derivan de los datos MDCT para formar una representación compleja en el dominio en frecuencia), que es necesaria en la codificación de predicción compleja, puede requerir que en el dominio en el tiempo la representación de la señal de mezcla descendente sea computable. Esto, a su vez, significa que la señal de mezcla descendente está disponible preferiblemente como una secuencia de tiempo de los espectros MDCT obtenidos de una manera uniforme. Si el filtro TNS se aplicara en el descodificador después de la matriz de mezcla ascendente, que convierte una representación de mezcla descendente/residual en una representación izquierdo/derecho, solo se dispondrá de una secuencia de espectros MDCT residuales del TNS de la señal de mezcla descendente. Esto haría que el cálculo eficiente de los espectros MDST correspondientes sea un gran desafío, especialmente si los canales izquierdo y derecho utilizaran filtros TNS con diferentes características.

Se enfatiza que la disponibilidad de una secuencia de tiempo del espectro MDCT no es un criterio absoluto para obtener una representación de MDST ajustada para servir de base para la codificación de predicción compleja. Además de la evidencia experimental, este hecho puede explicarse porque el TNS se aplica generalmente solo a las frecuencias más altas, como por encima de unos pocos kilohercios, de modo que la señal residual filtrada por el TNS corresponde aproximadamente a la señal residual no filtrada para frecuencias más bajas. Así, la invención puede realizarse como un descodificador para la codificación estéreo de predicción compleja, en la que los filtros TNS tienen una posición diferente que aguas arriba de la etapa de mezcla ascendente, como se indica a continuación.

5 En una realización, el sistema descodificador incluye al menos un módulo TNS adicional situado aguas abajo de la etapa de mezcla ascendente. Por medio de una disposición de selector, ya sea el(los) módulo(s) TNS aguas arriba de la etapa de mezcla ascendente o el(los) módulo(s) TNS aguas abajo de la etapa de mezcla ascendente. Bajo ciertas circunstancias, el cálculo de la representación compleja en el dominio en frecuencia no requiere que una representación en el dominio en el tiempo de la señal de mezcla descendente sea computable. Además, como se expuso anteriormente, el descodificador puede funcionar selectivamente en un modo de codificación directa o conjunta, no aplicando codificación de predicción compleja, y entonces puede ser más adecuado aplicar la localización convencional de los módulos TNS, es decir, como uno de los últimos pasos de procesamiento en el dominio en frecuencia.

10 En una realización, el sistema descodificador está adaptado para economizar recursos de procesamiento, y posiblemente energía, desactivando el módulo para calcular una segunda representación en el dominio en frecuencia de la señal de mezcla descendente cuando esta última no sea necesaria. Se supone que la señal de mezcla descendente se divide en bloques de tiempo sucesivos, cada uno de los cuales está asociado con un valor del coeficiente de predicción complejo. Este valor puede ser determinado por una decisión tomada para cada bloque de tiempo por un codificador que coopera con el descodificador. Además, en esta realización, el módulo para calcular una segunda representación en el dominio en frecuencia de la señal de mezcla descendente se adapta para desactivarse si, para un bloque de tiempo dado, el valor absoluto de la parte imaginaria del coeficiente de predicción complejo es cero o menor que una tolerancia predeterminada. La desactivación del módulo puede implicar que no se calcule una segunda representación en el dominio en frecuencia de la señal de mezcla descendente para este bloque de tiempo. Si la desactivación no tuvo lugar, la segunda representación en el dominio en frecuencia (por ejemplo, un conjunto de coeficientes MDST) se multiplicaría por cero o por una cantidad de sustancialmente el mismo orden de magnitud que la máquina epsilon (unidad de redondeo) del descodificador o algún otro valor de umbral adecuado.

25 En un desarrollo adicional de la realización anterior, la economización de los recursos de procesamiento se logra en un subnivel del bloque de tiempo en el que se divide la señal de mezcla descendente. Por ejemplo, tal subnivel dentro de un bloque de tiempo puede ser una banda de frecuencia, en la que el codificador determina un valor del coeficiente de predicción complejo para cada banda de frecuencia dentro de un bloque de tiempo. De manera similar, el módulo para producir una segunda representación en el dominio en frecuencia está adaptado para suprimir su funcionamiento para una banda de frecuencia en un bloque de tiempo para el cual el coeficiente de predicción complejo es cero o tiene magnitud menor que una tolerancia.

30 En una realización, los primeros componentes espectrales son componentes de la transformada organizados en uno o más bloques de tiempo de coeficientes de la transformada, cada bloque generado por la aplicación de una transformada a un segmento de tiempo de una señal en el dominio en el tiempo. Además, el módulo para calcular una representación en el segundo dominio en frecuencia de la señal de mezcla descendente se adapta para:

- 35 • derivar uno o más primeros componentes intermedios de al menos algunos de los primeros componentes espectrales;
- formar una combinación de dichos uno o más primeros componentes espectrales de acuerdo con al menos una parte de una o más respuestas de impulso para obtener uno o más segundos componentes intermedios; y
- 40 • derivar dicho uno o más segundos componentes espectrales de dicho uno o más segundos componentes intermedios.

45 Este procedimiento logra un cálculo de la segunda representación en el dominio en frecuencia directamente desde la primera representación en el dominio en frecuencia, tal como se describe con mayor detalle en el documento US 6.980.933 B2, en particular las columnas 8-28 y en particular su ecuación 41. Como el experto en la materia advierte, el cálculo no se realiza a través del dominio en el tiempo, a diferencia de, por ejemplo, la transformada inversa seguida de una transformada diferente.

50 Para una realización de un ejemplo de codificación estéreo de predicción compleja no contemplada por la invención reivindicada, se ha estimado que la complejidad computacional aumenta solo ligeramente (significativamente menos que el aumento causado por la codificación estéreo de predicción compleja en el dominio QMF) en comparación con el estéreo tradicional L/R o M/S. Una ilustración de este tipo incluyendo el cálculo exacto de los segundos componentes espectrales introduce un retardo que normalmente es solo de un pequeño porcentaje mayor que el introducido por una realización basada en QMF (suponiendo que la longitud del bloque de tiempo sea de 1024 muestras y comparándolo con el retardo del banco híbrido de filtros de análisis/síntesis de QMF, que es de 961 muestras).

Adecuadamente, en al menos algunas de las ilustraciones anteriores, las respuestas del impulso se adaptan a la transformada mediante la cual se puede obtener la primera representación en el dominio en frecuencia, y más precisamente de acuerdo con las características de respuesta de frecuencia de las mismas.

5 En algunos ejemplos, no contemplados por la invención reivindicada, la primera representación en el dominio en frecuencia de la señal de mezcla descendente se obtiene mediante una transformada que se aplica en conexión con una o más funciones de la ventana de análisis (o funciones de corte, por ejemplo, ventana rectangular, ventana sinusoidal, ventana derivada de Kaiser-Bessel, etc.), cuyo objetivo es lograr una segmentación temporal sin introducir una cantidad dañina de ruido o cambiar el espectro de una manera indeseada. Posiblemente, tales funciones de ventana se solapan parcialmente. Entonces, preferiblemente, las características de respuesta en frecuencia de la transformada dependen de las características de dicha una o más funciones de la ventana de análisis.

10 Todavía con referencia a las realizaciones que caracterizan el cálculo de la segunda representación en el dominio en frecuencia dentro del dominio en frecuencia, es posible disminuir la carga computacional involucrada utilizando una segunda representación aproximada en el dominio en frecuencia. Dicha aproximación se puede lograr al no requerir información completa en la que basar el cálculo. Por las enseñanzas del documento US 6.980.933 B2, por ejemplo, se requieren los primeros datos de dominio en frecuencia de tres bloques de tiempo para el cálculo exacto de la segunda representación en el dominio en frecuencia de la señal de mezcla descendente en un bloque, concretamente un bloque contemporáneo del bloque de salida, un bloque precedente y un bloque subsiguiente. Para el propósito de la codificación de predicción compleja de acuerdo con la presente invención, se pueden obtener aproximaciones adecuadas omitiendo - o reemplazando por cero - datos que emanen del bloque subsiguiente (por lo que el funcionamiento del módulo puede llegar a ser causal, es decir, no contribuye con un retardo) y/o del bloque anterior, de modo que el cálculo de la segunda representación en el dominio en frecuencia se base en datos solo de uno o dos bloques de tiempo. Obsérvese, que, aunque la omisión de los datos de entrada puede implicar un reescalado de la segunda representación en el dominio en frecuencia - en el sentido de que, por ejemplo, ya no representa la misma potencia - todavía se puede utilizar como base para la codificación de predicción compleja siempre que se calcule de manera equivalente en ambos extremos del codificador y del decodificador, como se señaló anteriormente. De hecho, un posible reescalado de este tipo será compensado por un cambio correspondiente del valor del coeficiente de predicción.

20 Otro método adicional aproximado para calcular un componente espectral que forme parte de la segunda representación en el dominio en frecuencia de la señal de mezcla descendente puede incluir una combinación de al menos dos componentes de la primera representación en el dominio en frecuencia. Los últimos componentes pueden ser contiguos con respecto al tiempo y/o a la frecuencia. Como alternativa, pueden combinarse mediante el filtrado en respuesta de impulsos finitos (FIR), con relativamente pocas tomas. Por ejemplo, en un sistema que aplique un tamaño de bloque de tiempo de 1024, dichos filtros FIR pueden incluir 2, 3, 4 etc. tomas. Descripciones de métodos aproximados de cálculo de esta naturaleza se pueden encontrar, por ejemplo, en el documento US 2005/0197831 A1. Si una función de ventana que dé ponderaciones considerablemente menores para la vecindad de cada límite de bloque de tiempo, por ejemplo, una función no rectangular, puede ser conveniente basar los segundos componentes espectrales en un bloque de tiempo solo en combinaciones de los primeros componentes espectrales en el mismo bloque de tiempo, lo que implica que no es lo mismo la cantidad de información disponible para los componentes más externos. El error de aproximación posiblemente introducido por tal práctica es, hasta cierto punto, suprimido u oculto por la forma de la función de ventana.

En una realización de un decodificador, que está diseñado para emitir una señal estéreo en el dominio en el tiempo, se incluye la posibilidad de cambiar entre codificación estéreo directa o conjunta y codificación de predicción compleja. Esto se logra mediante la provisión de:

- 45
- un conmutador que funcione de forma selectiva, ya sea como una etapa de paso (que no modifica las señales) o como una transformada de suma y diferencia;
 - una etapa de transformada inversa para realizar una transformada de frecuencia a tiempo; y
 - una disposición de selector para alimentar la etapa de transformada inversa con una señal codificada directa (o conjuntamente) o con una señal codificada por predicción compleja.

50 Como pueden advertir los expertos, tal flexibilidad sobre la parte del decodificador da la latitud del codificador para elegir entre la codificación directa o conjunta convencional y la codificación de predicción compleja. Por lo tanto, en los casos en que el nivel de calidad de la codificación estéreo L/R directa convencional o la codificación estéreo M/S conjunta no pueda ser superada, esta realización puede garantizar al menos que se mantenga el mismo nivel. Por lo tanto, el decodificador de acuerdo con esta realización se puede considerar como un superconjunto con respecto a la técnica relacionada.

55

Otro grupo de realizaciones del sistema de descodificación realiza el cálculo de los segundos componentes espectrales en la segunda representación en el dominio en frecuencia a través del dominio en el tiempo. Más precisamente, se aplica una inversa de la transformada mediante la cual se obtuvieron (o se pueden obtener) los primeros componentes espectrales y seguido por una transformada diferente que tiene como salida los segundos componentes espectrales. En particular, una TCMD inversa puede ir seguida de una MDST. Con el fin de reducir el número de transformadas y transformadas inversas, la salida de la MDCT inversa puede, en tal realización, alimentar tanto a la MDST como a los terminales de salida (posiblemente precedidos por etapas adicionales de procesamiento) del sistema de descodificación.

Para una realización de ejemplo de codificación estéreo de predicción compleja de acuerdo con la invención, se ha estimado que la complejidad computacional aumenta solo ligeramente (aun significativamente menor que el aumento causado por la codificación estéreo de predicción compleja en el dominio QMF) en comparación con el estéreo L/R o M/S tradicional.

Como un desarrollo adicional de la realización mencionada en el párrafo precedente, la etapa de mezcla ascendente puede comprender una etapa de transformada inversa adicional para procesar la señal lateral. Entonces, la etapa de suma y diferencia se suministra con una representación en el dominio en el tiempo de la señal lateral, generada por dicha etapa de transformada inversa adicional, y una representación en el dominio en el tiempo de la señal de mezcla descendente, generada por la etapa de transformada inversa ya referida. Se recuerda que, ventajosamente desde el punto de vista de la complejidad de cálculo, la última señal se suministra tanto a la etapa de suma y diferencia como a dicha etapa de transformada diferente mencionada anteriormente.

En una realización, un descodificador diseñado para enviar una señal estéreo en el dominio en el tiempo incluye la posibilidad de conmutar entre la codificación estéreo L/R directa o la codificación estéreo conjunta M/S y la codificación estéreo de predicción compleja. Esto se consigue mediante:

- un conmutador que funciona como etapa de paso o como etapa de suma y diferencia;
- otra etapa de transformada inversa para calcular una representación en el dominio en el tiempo de la señal lateral;
- una disposición de selector para conectar las etapas de transformadas inversas a una etapa adicional de suma y diferencia conectada a un punto aguas arriba de la etapa de mezcla ascendente y aguas abajo del conmutador (preferiblemente cuando el conmutador se ha activado para funcionar como un filtro de paso, como puede ser el caso en la descodificación de una señal estéreo generada por la codificación de predicción compleja) o una combinación de una señal de mezcla descendente procedente del conmutador y una señal lateral procedente del sumador ponderado (preferiblemente cuando el conmutador ha sido activado para funcionar como una etapa de suma y diferencia, como puede ser el caso en la descodificación de una señal estéreo codificada directamente).

Como pueden advertir los expertos, esto le da al codificador la libertad de elegir entre la codificación directa o conjunta convencional y la codificación de predicción compleja, lo que significa que se puede garantizar un nivel de calidad al menos equivalente al de la codificación estéreo directa o conjunta.

En un ejemplo, el sistema codificador, no contemplado por la invención reivindicada, puede comprender un estimador para estimar el coeficiente de predicción complejo con el objetivo de reducir o minimizar la potencia de la señal o la potencia media de la señal residual. La minimización puede tener lugar en un intervalo de tiempo, preferiblemente un segmento de tiempo o bloque de tiempo o trama de tiempos de la señal a codificar. El cuadrado de la amplitud puede tomarse como una medida de la potencia de señal momentánea, y se puede tomar una integral en un intervalo de tiempo de la amplitud al cuadrado (forma de onda) como una medida de la potencia media de la señal en ese intervalo. Adecuadamente, el coeficiente de predicción complejo se determina sobre una base de tiempo y banda de frecuencia, es decir, su valor se establece de tal manera que reduce la potencia media (es decir, la energía total) de la señal residual en ese bloque de tiempo y banda de frecuencia. En particular, los módulos para estimar los parámetros de codificación estéreo paramétricos, como IID, ICC e IPD o similares, pueden proporcionar resultados en los que el coeficiente de predicción complejo se puede calcular de acuerdo con las relaciones matemáticas conocidas por los expertos.

En un ejemplo, no contemplado por la invención reivindicada, la etapa de codificación del sistema codificador funciona, además, para servir como etapa de paso, así como para habilitar la codificación estéreo directa. Al seleccionar la codificación estéreo directa en situaciones donde se espera que esto proporcione una mayor calidad, el sistema codificador puede garantizar que la señal estéreo codificada

tenga al menos la misma calidad que en la codificación directa. De manera similar, en situaciones donde el mayor esfuerzo computacional que tiene lugar por la codificación de predicción compleja no está motivado por un aumento significativo de la calidad, la opción de economizar los recursos computacionales está por tanto disponible para el sistema codificador. La decisión entre la codificación conjunta, directa, predicción real y predicción compleja en el codificador se basa generalmente en una optimización racional velocidad/distorsión.

En un ejemplo, no contemplado por la invención reivindicada, el sistema codificador puede comprender un módulo para calcular una segunda representación en el dominio en frecuencia directamente (es decir, sin aplicar una transformada inversa en el dominio en el tiempo y sin usar los datos en el dominio en el tiempo de la señal) basándose en los primeros componentes espectrales. En relación con las realizaciones correspondientes del sistema descodificador descrito anteriormente, este módulo puede tener una estructura análoga, es decir, comprender concretamente las operaciones de procesamiento análogas, pero en un orden diferente, de modo que el codificador se adapte a los datos de salida adecuados como entrada en el lado del descodificador. Con el propósito de ilustrar este ejemplo, se supone que la señal estéreo a codificar comprende canales medios y laterales, o se ha transformado en esta estructura, y la etapa de codificación está adaptada para recibir una primera representación en el dominio en frecuencia. La etapa de codificación comprende un módulo para calcular una segunda representación en el dominio en frecuencia del canal medio. (La primera y segunda representaciones en el dominio en frecuencia a las que se hace referencia aquí son las definidas anteriormente; en particular, las primeras representaciones en el dominio en frecuencia pueden ser una representación MDCT y la segunda representación en el dominio en frecuencia puede ser una representación MDST.) La etapa de codificación comprende además un sumador ponderado para calcular una señal residual como una combinación lineal formada a partir de la señal lateral y las dos representaciones en el dominio en frecuencia de la señal del canal medio, ponderada por las partes real e imaginaria, respectivamente, del coeficiente de predicción complejo. La señal del canal medio, o adecuadamente la primera representación en el dominio en frecuencia de la misma, puede ser usada directamente como una señal de mezcla descendente. En este ejemplo, además, el estimador determina el valor del coeficiente de predicción complejo con el objetivo de minimizar la potencia o potencia media de la señal residual. La operación final (optimización) puede efectuarse mediante el control de realimentación, en donde el estimador puede recibir la señal residual obtenida por los valores actuales del coeficiente de predicción para ajustarse aún más si es necesario, o, de una manera de alimentación directa, por cálculos efectuados directamente sobre los canales izquierdo/derecho de una señal estéreo original o los canales medios/laterales. El método de alimentación directa es el preferido, por el cual el coeficiente de predicción complejo se determina directamente (en particular, de una manera no iterativa o sin realimentación) basándose en la primera y la segunda representaciones en el dominio en frecuencia de la señal del canal medio y en la primera representación en el dominio en frecuencia de la señal lateral. Obsérvese que la determinación del coeficiente de predicción complejo puede ser seguido por una decisión de si aplicar la codificación de predicción directa, conjunta, real o de predicción compleja, en donde se considera la calidad resultante (preferiblemente la calidad de percepción, teniendo en cuenta, por ejemplo, los efectos de señal a enmascaramiento) de cada opción disponible; por lo tanto, las declaraciones anteriores no se deben interpretar al efecto de que no exista mecanismo de realimentación en el codificador.

En un ejemplo, no contemplado por la invención reivindicada, el sistema codificador comprende módulos para calcular una segunda representación en el dominio en frecuencia de la señal del canal medio (o de mezcla descendente) a través del dominio en el tiempo. Se entiende que los detalles de ejecución relacionados con este ejemplo, al menos en lo que respecta al cálculo de la segunda representación en el dominio en frecuencia, es similar o se puede resolver de manera análoga a las correspondientes realizaciones del descodificador. En este ejemplo, la etapa de codificación comprende:

- una etapa de suma y diferencia para convertir la señal estéreo en una forma que comprende canales medios y laterales;
- una etapa de transformación para proporcionar una representación en el dominio en frecuencia del canal lateral y una representación en el dominio en frecuencia de valor complejo (y, por lo tanto, sobre muestreada) del canal medio; y
- una suma ponderada para calcular una señal residual, en la que el coeficiente de predicción complejo se utiliza como ponderación.

Aquí, el estimador puede recibir la señal residual y determinar, posiblemente en modo de control de realimentación, el coeficiente de predicción complejo para reducir o minimizar la potencia o la media de la señal residual. Preferiblemente, sin embargo, el estimador recibe la señal estéreo que se va a codificar y determina el coeficiente de predicción sobre esa base. Es ventajoso desde el punto de vista de la economía computacional usar una representación en el dominio en frecuencia muestreada críticamente en el canal lateral, ya que este último no se someterá a la multiplicación por un número complejo en esta realización. Adecuadamente, la etapa de

transformación puede comprender una etapa MDCT y una etapa MDST dispuestas en paralelo, ambas con representación en el dominio en el tiempo del canal medio como entrada. Por lo tanto, se produce una representación en el dominio en frecuencia sobre muestreada del canal medio y una representación en el dominio en frecuencia muestreada críticamente del canal lateral.

- 5 Obsérvese que los métodos y aparatos descritos en esta sección pueden aplicarse, después de las modificaciones apropiadas dentro de las capacidades del experto en la materia, incluida la experimentación rutinaria, a la codificación de señales que tienen más de dos canales. Las modificaciones en dicha funcionalidad multicanal pueden tener lugar, por ejemplo, a lo largo de las líneas de las secciones 4 y 5 en el documento de J. Herre et al. citado anteriormente.
- 10 Las características de dos o más realizaciones descritas anteriormente pueden combinarse, a menos que sean claramente complementarias, en otras realizaciones. El hecho de que dos características se invoquen en diferentes reivindicaciones no excluye que se puede combinar como una ventaja. Del mismo modo, en otras realizaciones se pueden omitir ciertas características que no sean necesarias o esenciales para el propósito deseado. Como ejemplo, el sistema de descodificación de acuerdo con la invención puede realizarse sin una etapa de des cuantificación en
- 15 los casos en que la señal codificada a procesar no esté cuantificada o esté ya disponible de una forma adecuada para ser procesada por la etapa de mezcla ascendente.

Breve descripción de los dibujos

La invención se ilustrará ahora más detalladamente mediante las realizaciones descritas en la siguiente sección, haciendo referencia a los dibujos adjuntos, en los que:

- 20 la figura 1 consiste en dos diagramas de bloques generalizados que muestran descodificadores basados en QMF de acuerdo con la técnica anterior;
- la figura 2 es un diagrama de bloques generalizado de un sistema de descodificación estéreo basado en MDCT con predicción compleja, de acuerdo con un ejemplo, no contemplado por la invención reivindicada, en el que la representación compleja de un canal de la señal a descodificar se computa en el dominio en
- 25 frecuencia;
- la figura 3 es un diagrama de bloques generalizado de un sistema descodificador estéreo basado en MDCT con predicción compleja, de acuerdo con un ejemplo no contemplado por la invención reivindicada, en el que la representación compleja de un canal de la señal a descodificar se computa en el dominio en el
- 30 tiempo;
- la figura 4 muestra una realización alternativa del sistema descodificador de la figura 2, en la que la posición de la etapa TNS activa se puede seleccionar de acuerdo con la invención reivindicada;
- la figura 5 comprende diagramas de bloques generalizados que muestran sistemas de codificador estéreo basados en MDCT con predicción compleja, de acuerdo con ejemplos no contemplados por la invención reivindicada;
- 35 la figura 6 es un diagrama de bloques generalizado de un sistema codificador estéreo basado en MDCT con predicción compleja, de acuerdo con un ejemplo no contemplado por la invención reivindicada, en el que se computa una representación compleja de un canal de la señal a codificar basándose en la representación en el dominio en el tiempo de la misma;
- 40 la figura 7 muestra un ejemplo alternativo no contemplado por la invención reivindicada del sistema codificador de la figura 6, que funciona también en un modo de codificación directa L/R;
- la figura 8 es un diagrama de bloques generalizado de un sistema codificador estéreo basado en MDCT con predicción compleja, de acuerdo con un ejemplo no contemplado por la invención reivindicada, en el que una representación compleja de un canal de la señal a ser codificada se calcula basándose en una primera
- 45 representación en el dominio en frecuencia, cuyo sistema descodificador funciona también en un modo de codificación directa L/R;
- la figura 9 muestra un ejemplo alternativo no contemplado por la invención reivindicada del sistema codificador de la figura 7, que incluye además una etapa TNS dispuesta aguas abajo de la etapa de codificación;
- 50 la figura 10 muestra ejemplos alternativos no contemplados por la invención reivindicada de la parte etiquetada A en las figuras 2 y 8;

la figura 11 muestra un ejemplo alternativo no contemplado por la invención reivindicada del sistema codificador de la figura 8, que incluye además dos dispositivos de modificación en el dominio en frecuencia dispuestos respectivamente aguas abajo y aguas arriba de la etapa de codificación;

la figura 12 es una presentación gráfica de los resultados de las pruebas de audición a 96 kb/s de seis casos que muestran opciones de diferente complejidad – equilibrio costo/calidad - para el cálculo o la aproximación del espectro MDST, en donde los puntos de datos etiquetados "+" se refieren a referencia oculta, "X" se refieren a anclaje limitado a la banda de 3.5 kHz, "*" se refiere al estéreo tradicional USAC (M/S o L/R), "□" se refiere a la codificación estéreo unificada en el dominio MDCT por predicción compleja con parte imaginaria del coeficiente de predicción desactivado (es decir, predicción de valor real, que no requiere MDST), "■" se refiere a la codificación estéreo unificada en el dominio MDCT por predicción compleja utilizando una trama MDCT actual para calcular una aproximación de la MDST, "○" se refiere a la codificación estéreo unificada en el dominio MDCT por predicción compleja usando las tramas MDCT actuales y anteriores para calcular una aproximación de la MDST y "●" se refieren a la codificación estéreo unificada en el dominio MDCT por predicción compleja utilizando las tramas MDCT actuales, anteriores y siguientes para calcular la MDST;

la figura 13 presenta los datos de la figura 12, sin embargo, como puntuaciones diferenciales relativas a la codificación estéreo unificada en el dominio MDCT por predicción compleja utilizando una trama MDCT actual para calcular una aproximación de la MDST;

la figura 14 comprende diagramas de bloques generalizados que muestran tres ejemplos de un sistema descodificador;

la figura 14C muestra una realización de la invención;

la figura 15 es un diagrama de flujo que muestra un método de descodificación de acuerdo con un ejemplo no contemplado por la invención reivindicada; y

la figura 16 es un diagrama de flujo que muestra un método de codificación de acuerdo con un ejemplo no contemplado por la invención reivindicada.

Descripción detallada de las realizaciones

I. Sistemas descodificadores

La figura 2 muestra, en forma de un diagrama de bloques generalizado, un sistema de descodificación para descodificar un flujo de bits que comprende al menos un valor de un coeficiente de predicción complejo $\alpha = \alpha_R + i\alpha_I$ y una representación MDCT de una señal estéreo que tiene canales M de mezcla descendente y D residuales. Las partes real e imaginaria α_R , $i\alpha_I$ del coeficiente de predicción pueden haber sido cuantificadas y/o codificadas conjuntamente. Sin embargo, preferiblemente, las partes real e imaginaria se cuantifican independiente y uniformemente, normalmente con un tamaño de escalón 0,1 (sin dimensiones). La resolución de banda de frecuencia utilizada para el coeficiente de predicción complejo no es necesariamente la misma que la resolución para las bandas de factores de escala (*sfb*; por ejemplo, un grupo de líneas MDCT que utilizan el mismo tamaño de escalón de cuantificación y rango de cuantificación de MDCT) de acuerdo con la norma MPEG. En particular, la resolución de la banda de frecuencia para el coeficiente de predicción puede ser una que esté psico acústicamente justificada, como la escala de Bark. Un desmultiplexor 201 está adaptado para extraer estas representaciones de la TCMD y el coeficiente de predicción (parte de la *Información de Control* como se indica en la figura) del flujo de bits que se le suministra. De hecho, se puede codificar en el flujo de bits más información de control que simplemente el coeficiente de predicción complejo, por ejemplo, instrucciones si el flujo de bits se va a descodificar en modo de predicción o no, la información del TNS, etc. La información del TNS puede incluir los valores de los parámetros TNS que se aplicarán por los filtros TNS (síntesis) del sistema descodificador. Si se van a usar conjuntos idénticos de parámetros TNS para varios filtros TNS, como para ambos canales, es económico recibir esta información en forma de un bit que indique dicha identidad de los conjuntos de parámetros en lugar de recibir los dos conjuntos de parámetros de forma independiente. También se puede incluir información sobre si aplicar el TNS antes o después de la etapa de mezcla ascendente, según corresponda basándose en, por ejemplo, una evaluación psico acústica de las dos opciones disponibles. Además, la información de control puede indicar anchos de banda individualmente limitados para las señales de mezcla descendente y residual. Para cada canal, las bandas de frecuencia por encima de un límite de ancho de banda no se descodificarán, sino que se pondrán a cero. En ciertos casos, las bandas de frecuencia más altas tienen un contenido de energía tan pequeño que ya están cuantificadas a cero. La práctica normal (véase el parámetro *max_sfb* en la norma MPEG) ha sido utilizar la misma limitación de ancho de banda para las señales de mezcla descendente y residual. Sin embargo, la señal residual, en mayor medida que la señal de mezcla descendente, tiene su contenido de energía localizado en bandas de frecuencia más bajas. Por lo tanto,

colocando un límite de ancho de banda superior especial en la señal residual, es posible una reducción de la velocidad de bits sin pérdida significativa de la calidad. Por ejemplo, esto puede ser manejado por dos parámetros max_sfb independientes codificados en el flujo de bits, uno para la señal de mezcla descendente y otro para la señal residual.

5 En este ejemplo no contemplado por la invención reivindicada, la representación MDCT de la señal estéreo está segmentada en tramas de tiempo sucesivas (o bloques de tiempo) que comprenden un número fijo de puntos de datos (por ejemplo, 1024 puntos), uno de varios números fijos de puntos de datos (por ejemplo, 128 o 1024 puntos) o un número variable de puntos. Como conocen los expertos en la técnica, la MDCT se muestrea críticamente. La salida del sistema de descodificación, indicada en la parte derecha del dibujo, es una señal estéreo en el dominio en el tiempo que tiene los canales izquierdo L y derecho R. Los módulos 202 de des cuantificación están adaptados para manejar la entrada del flujo de bits al sistema de descodificación o, cuando corresponda, dos flujos de bits obtenidos después de la desmultiplexión de un flujo de bits originales y correspondientes a cada uno de los canales de mezcla descendente y residual. Las señales del canal des cuantificado se proporcionan a un conjunto de conmutación 203 que funciona en modo de paso o en modo suma y diferencia correspondiente a las respectivas matrices de la transformada

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

y

$$\frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

20 Como se describirá más adelante en el siguiente párrafo, el sistema descodificador incluye un segundo conjunto de conmutación 205. Ambos conjuntos de conmutación 203, 205, como la mayoría de los otros conmutadores y conjuntos de conmutación en esta realización y en las realizaciones a describir, funcionan en un modo selectivo en frecuencia. Esto permite la descodificación de una gran variedad de modos de descodificación, por ejemplo, descodificación L/R o M/S dependiente de la frecuencia, como se conoce en la técnica relacionada. Por lo tanto, el descodificador de acuerdo con la invención puede considerarse como un super conjunto en relación con la técnica relacionada.

25 Suponiendo por ahora que el conjunto de conmutación 203 está en el modo de paso, las señales del canal des cuantificadas pasan, en este ejemplo, a través de los respectivos filtros TNS 204. Los filtros TNS 204 no son esenciales para el funcionamiento del sistema de descodificación y pueden ser reemplazados por elementos de paso. Después de esto, la señal se suministra al segundo conjunto de conmutación 205 que tiene la misma funcionalidad que el conjunto de conmutación 203 situado aguas arriba. Con las señales de entrada como las descritas anteriormente y con el segundo conjunto de conmutación 205 establecido en su modo de paso es, la salida del anterior es la señal del canal de mezcla descendente y la señal del canal residual. La señal de mezcla descendente, aún representada por sus espectros MDCT sucesivos en el tiempo, se suministra a una transformada 206 real a imaginaria adaptada para calcular basándose en ello, el espectro MDST de la señal de mezcla descendente. En este ejemplo, una trama MDST se basa en tres tramas MDCT, una trama anterior, una trama actual (o contemporánea) y una trama posterior. Esto se indica simbólicamente por (Z^{-1}, Z) que el lado de entrada de la transformada real a imaginaria 206 comprende los componentes de retardo.

30 La representación MDST de la señal mezcla descendente obtenida de la transformada real a imaginaria 206 es ponderada por la parte imaginaria α_I del coeficiente de predicción y se agrega a la representación MDCT de la señal de mezcla descendente ponderada por la parte real α_R del coeficiente de predicción y la representación MDCT de la señal residual. Las dos adiciones y multiplicaciones se realizan mediante multiplicadores y sumadores 210, 211, formando juntos (funcionalmente) un sumador ponderado, que se suministran con el valor del coeficiente de predicción complejo α codificado en el flujo de bits recibido inicialmente por el sistema descodificador. El coeficiente de predicción complejo se puede determinar una vez por cada trama de tiempo. También puede determinarse más a menudo, tal como una vez para cada banda de frecuencia dentro de una trama, siendo las bandas de frecuencia una partición motivada psico acústicamente. También se puede determinar con menos frecuencia, como se describirá a continuación en relación con los sistemas de codificación de acuerdo con la invención. La transformada real a imaginaria 206 se sincroniza con el sumador ponderado de tal manera que una trama MDST actual de la señal del canal de mezcla descendente se combina con unas tramas MDCT contemporáneas de cada una de las señales del canal de mezcla descendente y de la señal del canal residual. La suma de estas tres señales son una señal lateral $S = Re \{ \alpha M \} + D$. En esta expresión, M incluye las representaciones MDCT y MDST de la señal de mezcla descendente, es decir, $M = M_{MDCT} - iM_{MDST}$, mientras que $D = D_{MDCT}$ tiene un valor real. Así, se obtiene una señal estéreo que tiene un canal de mezcla descendente y un canal lateral, a partir del cual una transformada de suma y diferencia 207 restaura los canales izquierdo y derecho de la siguiente manera:

$$\begin{cases} L = M + S \\ R = M - S \end{cases}$$

Estas señales están representadas en el dominio MDCT. La última etapa del sistema de codificación consiste en aplicar una MDCT 209 inversa a cada uno de los canales, por lo que se obtiene una representación en el dominio en el tiempo de la señal estéreo izquierdo/derecho.

- 5 Una posible ejecución de la transformada real a imaginaria 206 se describe con más detalle en la patente del solicitante US 6.980.933 B2, como se indicó anteriormente. Por la fórmula 41 del documento, la transformada se puede expresar como un filtro finito de respuesta al impulso, por ejemplo, para puntos pares,

$$S(2v) = \frac{2}{N} \sum_{p=0}^{N-1} [(-1)^{p+1} X_I(p) + X_{III}(p)] h_{I,III}(2v-p) + \frac{4}{N} \sum_{l=0}^{N-1} X_{II}(2l+1) h_{II}(2v-2l-1),$$

- 10 donde $S(2v)$ es el $2v^o$ punto de datos MDST, X_I , X_{II} , X_{III} , son los datos MDCT de cada una de las tramas y N es la longitud de la trama. Además, $h_{I,III}$, h_{II} son respuestas del impulso según la función de ventana aplicada y, por lo tanto, se determinan para cada opción de función de ventana, tal como rectangular, sinusoidal y derivada de Kaiser-Bessel, y para cada longitud de trama. La complejidad de este cálculo puede reducirse omitiendo aquellas respuestas del impulso que tengan un contenido de energía relativamente menor y contribuyan relativamente menos a los datos MDST. Como alternativa o extensión a esta simplificación, las propias respuestas de impulso pueden acortarse, por ejemplo, desde la longitud completa de la trama N a un número menor de puntos. Como ejemplo, la longitud de la respuesta del impulso puede disminuirse de 1024 puntos (tomas) a 10 puntos. El truncado más extremo que aún puede considerarse significativo es

$$S(v) = X_{II}(v+1) - X_{II}(v-1).$$

Otros enfoques directos se pueden encontrar en el documento US 2005/0197831 A1.

- 20 También es posible reducir la cantidad de datos de entrada en los que se basa el cálculo. Para ilustrarlo, la transformada real a imaginaria 206 y sus conexiones aguas arriba, que se indican como una parte indicada por "A" en el dibujo, deben reemplazarse por una variante simplificada, dos de las cuales A' y A" se muestran en la figura 10. La variante A' proporciona una representación imaginaria aproximada de la señal. Aquí, el cálculo de MDST solo toma en cuenta la trama actual y la anterior. Con referencia a la fórmula anterior en este párrafo, esto puede efectuarse estableciendo $X_{III}(p) = 0$ para $p = 0, \dots, N-1$ (el índice III indica el período de tiempo subsiguiente). Dado que la variante A' no requiere el espectro MDCT de la trama subsiguiente como entrada, el cálculo de MDST no incurre en ningún retardo de tiempo. Claramente, esta aproximación reduce algo la precisión de la señal de MDST obtenida, pero también puede implicar que la energía de esta señal se reduzca; este último hecho puede ser compensado completamente por un aumento en α_I como resultado de la naturaleza de la codificación de predicción.
- 25
- 30 También se muestra en la figura 10 la variante A", que utiliza como entrada solo los datos de MDCT para la trama de tiempo actual. La variante A" podría decirse que produce una representación de MDST menos precisa que la variante A'. Por otro lado, funciona con retardo cero, al igual que la variante A', y tiene menor complejidad computacional. Como ya se mencionó, las propiedades de codificación de la forma de onda no se ven afectadas siempre que se utilice la misma aproximación en el sistema codificador y en el sistema descodificador.
- 35 Debe observarse, independientemente de qué variante A, A' o A" o cualquier otro desarrollo de lo mismo se utilice, que solo es necesario calcular aquellas partes del espectro de la MDST para las cuales la parte imaginaria del coeficiente de predicción complejo no sea cero, $\alpha_I \neq 0$. En circunstancias prácticas, esto se tomará para significar que el valor absoluto $|\alpha_I|$ de toda la parte imaginaria del coeficiente es mayor que un valor de umbral predeterminado, que puede estar relacionado con el redondeo unitario del hardware utilizado. En caso de que la parte imaginaria del coeficiente sea cero para todas las bandas de frecuencia dentro de una trama de tiempo, no hay necesidad de calcular ningún dato de la MDST para esa trama. Por lo tanto, adecuadamente, la transformada real a imaginaria 206 se adapta para responder a que existan valores $|\alpha_I|$ muy pequeños al no generar una salida MDST, por lo que los recursos de computación se pueden economizar. En ejemplos donde se utilicen más tramas que la actual para producir una trama de datos MDST, sin embargo, cualesquiera unidades aguas arriba de la transformada 206 deben continuar funcionando adecuadamente aunque no se necesite el espectro MDST, en particular, el segundo conjunto de conmutación 205 debe seguir enviando espectros MDCT, de modo que ya haya suficientes datos de entrada disponibles para la transformada real a imaginaria 206 cuando se produzca la siguiente trama de tiempo asociada con un coeficiente de predicción distinto de cero; este puede ser, por supuesto, el siguiente bloque de tiempo.
- 40
- 45

Volviendo a la figura 2, la función del sistema de descodificación ha sido descrita bajo el supuesto de que ambos conjuntos de conmutación 203, 205 se ajusten en sus respectivos modos de paso. Como se explicará ahora, el sistema de codificación también puede descodificar señales que no estén codificadas por predicción. Para este uso, el segundo conjunto de conmutación 205 se establecerá en su modo de suma y diferencia y adecuadamente, como se indica en el dibujo, se establecerá un conjunto de selector 208 en su posición inferior, asegurando así que las señales alimenten directamente a la transformada inversa 209 desde un punto de origen entre los filtros TNS 204 y el segundo conjunto de conmutación 205. Para garantizar la correcta descodificación, la señal tiene adecuadamente forma L/R en el punto de origen. Por lo tanto, para asegurar que la transformada real a imaginaria se suministre con la señal del canal medio, correcta (es decir, mezcla descendente) en todo momento (en lugar de, digamos, intermitentemente por medio de una señal izquierdo), el segundo conjunto de conmutación 205 se establece preferiblemente en su modo de suma y diferencia durante la descodificación de una señal estéreo no codificada por predicción. Como se señaló anteriormente, la codificación de predicción puede reemplazarse por codificación directa o conjunta para ciertas tramas basadas en, por ejemplo, una decisión de calidad de datos de velocidad a audio. El resultado de tal decisión puede comunicarse desde el codificador al descodificador de varias maneras, por ejemplo, mediante el valor de un bit indicador especial en cada trama, o mediante la ausencia o presencia de un valor del coeficiente de predicción. Habiendo establecido estos hechos, el rol del primer conjunto de conmutación 203 se puede realizar fácilmente. De hecho, en el modo de codificación sin predicción, el sistema descodificador puede procesar ambas señales de acuerdo con la codificación estéreo directa (L/R) o codificación conjunta (M/S), y al funcionar el primer conjunto de conmutación 203 ya sea en modo de paso o de suma y diferencia, es posible asegurar de que el punto de origen siempre esté provisto de una señal codificada directamente. Claramente, el conjunto de conmutación 203 cuando funciona como etapa de suma y diferencia convertirá una señal de entrada de forma M/S en una señal de salida (suministrada a los filtros 204 TNS opcionales) en forma L/R.

El sistema descodificador recibe una señal si una trama de tiempo en particular tiene que ser descodificada por el sistema descodificador en modo de codificación de predicción o no predicción. El modo de no predicción puede ser señalizado por el valor de un bit indicador especial en cada trama o por la ausencia (o el valor cero) del coeficiente de predicción. El modo de predicción se puede comunicar de forma análoga. Una ejecución particularmente ventajosa, que permite el repliegue sin ningún tipo de sobrecarga, utiliza un cuarto valor reservado del campo de dos bits *ms_mask_present* (véase MPEG-2 AAC, documento ISO/IEC 13818-7), que se transmite por trama de tiempo y se define de la siguiente manera:

Tabla 1: Definición de <i>ms_mask_present</i> en USAC	
Valor	Significado
00	Codificación L/R para todas las bandas de frecuencia
01	Se usa un bit de señalización por banda para indicar L/R o M/S
10	Codificación M/S para todas las bandas de frecuencia
11	Reservado

Al redefinir el valor 11 para que signifique "codificación de predicción compleja", el descodificador puede funcionar en todos los modos antiguos, particularmente en la codificación M/S y L/R, sin ninguna penalización por velocidad de bits y aún es capaz de recibir una señal que indique el modo codificación de predicción compleja para las tramas relevantes.

La figura 4 muestra un sistema descodificador de acuerdo con la invención reivindicada de la misma estructura general que la que se muestra en la figura 2 pero que incluye, sin embargo, al menos dos estructuras diferentes. En primer lugar, el sistema de la figura 4 incluye los conmutadores 404, 411 que permiten la aplicación de algunas etapas de procesamiento que involucran la modificación en el dominio de la frecuencia aguas arriba y/o aguas abajo de la etapa de mezcla ascendente. Esto se logra, por un lado, mediante un primer conjunto de modificadores en el dominio en frecuencia 403 (dibujados como filtros de síntesis TNS en esta figura) provistos junto con el primer conmutador 404 aguas abajo de los módulos de des cuantificación 401 y de un primer conjunto de conmutación 402 pero aguas arriba de un segundo conjunto de conmutación 405 dispuesto inmediatamente aguas arriba de la etapa de mezcla ascendente 406, 407, 408, 409. Por otra parte, el sistema de descodificación incluye un segundo conjunto de modificadores en el dominio en frecuencia 410 provistos junto con un segundo conmutador 411 aguas abajo de la etapa de mezcla ascendente 406, 407, 408, 409 pero aguas arriba de una etapa de transformada inversa 412. Ventajosamente, como se indica en el dibujo, cada modificador en el dominio en frecuencia se dispone en paralelo con una línea de paso que está conectada aguas arriba al lado de entrada del modificador en el dominio en frecuencia y está conectado aguas abajo al conmutador asociado. En virtud de esta estructura, el modificador en el dominio en frecuencia se suministra con los datos de la señal en todo momento, lo que permite el procesamiento en el dominio en frecuencia basándose en más tramas de tiempo que el actual. La decisión de aplicar el primer 403 o el

segundo de los conjuntos de modificadores en el dominio en frecuencia 410, puede ser tomada por el codificador (y transmitido en el flujo de bits), o se puede basar en si se aplica la codificación de predicción, o puede basarse en algún otro criterio encontrado adecuado en circunstancias prácticas. Como ejemplo, si el modificador en el dominio en frecuencia son filtros TNS, entonces el primer conjunto 403 puede ser ventajoso para usar para algunos tipos de señales, mientras que el segundo conjunto 410 puede ser ventajoso para otros tipos de señales. Si el resultado de esta selección se codifica en el flujo de bits, entonces el sistema descodificador activará el respectivo conjunto de filtros TNS en consecuencia.

Para facilitar la comprensión del sistema descodificador que se muestra en la figura 4, se señala explícitamente que la descodificación de una señal codificada directamente (L/R) tiene lugar cuando $\alpha = 0$ (lo que implica que pseudo-L/R y L/R son idénticos y que los canales lateral y residual no difieren), el primer conjunto de conmutación 402 está en modo de paso, el segundo conjunto de conmutación está en el modo de suma y diferencia, lo que hace que la señal tenga una forma M/S entre el segundo conjunto de conmutación 405 y una etapa de suma y diferencia 409 de la etapa de mezcla ascendente. Debido a que la etapa de mezcla ascendente será efectivamente una etapa de paso, es irrelevante si (usando los respectivos conmutadores 404, 411) el primero o el segundo conjunto de modificadores en el dominio en frecuencia está activado.

La figura 3 ilustra un sistema de descodificación de acuerdo con un ejemplo no contemplado por la invención reivindicada, el cual, en relación con los de las figuras 2 y 4, representa un enfoque diferente a la provisión de datos MDST requeridos para la mezcla ascendente. Semejante a los sistemas de descodificación ya descritos, el sistema de la figura 3 comprende módulos de descuantificación 301, un primer conjunto de conmutación 302 que puede funcionar en modo de paso o suma y diferencia y filtros TNS (síntesis) 303, que están todos dispuestos en serie desde el extremo de entrada del sistema descodificador. Los módulos aguas abajo de este punto se utilizan selectivamente por medio de dos segundos conmutadores 305, 310, que funcionan preferiblemente de manera conjunta de forma que ambos se encuentren en sus posiciones superior o inferior, como se indica en la figura. En el extremo de salida del sistema descodificador, hay una etapa de suma y diferencia 312 y, inmediatamente aguas arriba de ella, dos módulos MDCT inversos 306, 311 para transformar una representación en el dominio MDCT de cada canal en una representación en el dominio en el tiempo.

En la descodificación de predicción compleja, donde el sistema descodificador se suministra con un flujo de bits que codifica una señal estéreo de mezcla descendente/residual y los valores de un coeficiente de predicción complejo, el primer conjunto de conmutación 302 se establece en su modo de paso y los segundos conmutadores 305, 310 se establecen en la posición superior. Aguas abajo de los filtros TNS, los dos canales de la señal estéreo (des cuantificada, filtrada por TNS, MDCT) se procesan de diferentes maneras. El canal de mezcla descendente se proporciona, por un lado, a un multiplicador y a un sumador 308, que agrega la representación MDCT del canal de mezcla descendente ponderada por la parte real α_R del coeficiente de predicción a la representación MDCT del canal residual, y, por otro lado, a uno 306 de los módulos de transformada inversa MDCT. Se proporciona la representación en el dominio en el tiempo del canal M de mezcla descendente, que es la salida del módulo de la transformada inversa MDCT 306 tanto para la etapa final de suma y diferencia 312 como para un módulo de la transformada MDST 307. Este doble uso de la representación en el dominio en el tiempo del canal de mezcla descendente es ventajoso desde el punto de vista de la complejidad computacional. La representación MDST del canal de mezcla descendente así obtenida se suministra a un multiplicador y sumador adicional 309, que después de ponderar por medio de la parte imaginaria α_I del coeficiente de predicción agrega esta señal a la salida de combinación lineal del sumador 308; por lo tanto, la salida del sumador 309 es una señal del canal lateral, $S = \text{Re}\{\alpha M\} + D$. De manera similar al sistema de descodificación que se muestra en la figura 2, los multiplicadores y sumadores 308, 309 se pueden combinar fácilmente para formar un sumador ponderado multiseñal con entradas de las representaciones MDCT y MDST de la señal de mezcla descendente, la representación MDCT de la señal residual y el valor del coeficiente de predicción complejo. Aguas abajo de este punto en el presente ejemplo no contemplado por la invención reivindicada, solo un paso a través del módulo de la transformada inversa MDCT 311 permanece antes de que la señal del canal lateral se suministre a la etapa 312 de suma y diferencia final.

La necesaria sincronización en el sistema descodificador se puede lograr aplicando las mismas longitudes de transformada y formas de ventana en ambos módulos de la transformada inversa MDCT 306, 311, como ya es una práctica en la codificación selectiva en frecuencia M/S y L/R. Se introduce un retardo de una trama mediante la combinación de ciertas realizaciones del módulo inverso MDCT 306 y las realizaciones del módulo MDST 307. Por lo tanto, se proporcionan cinco bloques de retardo 313 opcionales (o instrucciones de software a este efecto en una ejecución de ordenador), de modo que la parte del sistema situada a la derecha de la línea discontinua se pueda retardar una trama en relación con la parte izquierda cuando sea necesario. Aparentemente, todas las intersecciones entre la línea discontinua y las líneas de conexión están provistas de bloques de retardo, con la excepción de la línea de conexión entre el módulo inverso MDCT 306 y el módulo de la transformada MDST 307, que es donde surge el retardo que requiere compensación.

El cálculo de datos de la MDST para una trama de tiempo requiere datos de una trama de la representación en el dominio en el tiempo. Sin embargo, la transformada inversa MDCT se basa en una (actual), dos (preferiblemente: anterior y actual) o tres (preferiblemente: anterior, actual y siguiente) tramas consecutivas. En virtud de la conocida cancelación de alias en el dominio en el tiempo (TDAC) asociada con la MDCT, la opción de tres tramas logra la superposición completa de las tramas de entrada y, por lo tanto, proporciona la mejor precisión (y posiblemente perfecta), al menos en tramas que contienen un alias en el dominio en el tiempo. Claramente, las tres tramas inversas MDCT funcionan con un retardo de una trama. Al aceptar usar una representación en el dominio en el tiempo aproximada como entrada a la transformada MDST, se puede evitar este retardo y por lo tanto la necesidad de compensar los retardos entre diferentes partes del sistema descodificador. En la opción de dos tramas, la función de superposición/adición de TDAC se produce en la mitad anterior de la trama, y el alias puede estar presente solo en la mitad posterior. En la opción de una trama, la ausencia de TDAC implica que puede haber alias en toda la trama; sin embargo, una representación MDST lograda de esta manera, y utilizada como una señal intermedia en la codificación de predicción compleja, todavía puede proporcionar una calidad satisfactoria.

El sistema de descodificación ilustrado en la figura 3 también puede funcionar en dos modos de descodificación de no predicción. Para descodificar una señal estéreo directamente codificada L/R, los segundos conmutadores 305, 310 se colocan en la posición inferior y el primer conjunto de conmutación 302 se establece en el modo de paso. Así, la señal tiene forma L/R aguas arriba de la etapa de suma y diferencia 304, que la convierte en forma M/S, en la que tiene lugar la transformada inversa MDCT y una operación final de suma y diferencia. Para descodificar una señal estéreo provista en forma codificada conjuntamente M/S, el primer conjunto de conmutación 302 se establece en su lugar en su modo de suma y diferencia, de modo que la señal tenga una forma L/R entre el primer conjunto de conmutación 302 y la etapa de suma y diferencia 304, que a menudo es más adecuada desde el punto de vista del filtrado TNS que lo sería una forma M/S. El procesamiento aguas arriba la etapa de suma y diferencia 304 es idéntico al del caso de la descodificación directa L/R.

La figura 14 consiste en tres diagramas de bloques generalizados de descodificadores.

A diferencia de otros diagramas de bloques que acompañan a esta aplicación, una línea de conexión en la figura 14 puede simbolizar una señal multicanal. En particular, dicha línea de conexión puede estar dispuesta para transmitir una señal estéreo que comprende canales izquierdo/derecho, medio/lateral, de mezcla descendente/residual, pseudo-izquierdo/pseudo-derecho y otras combinaciones.

La figura 14A muestra un sistema descodificador para descodificar una representación en el dominio en frecuencia (indicado, para los fines de esta figura, como una representación MDCT) de una señal de entrada. El sistema descodificador está adaptado para suministrar como salida una representación en el dominio en el tiempo de una señal estéreo, que se genera sobre la base de la señal de entrada. Para poder descodificar una señal de entrada codificada por codificación estéreo de predicción compleja, el sistema descodificador está provisto de una etapa de mezcla ascendente 1410. Sin embargo, también es capaz de manejar una señal de entrada codificada en otros formatos y posiblemente, que alterne entre varios formatos de codificación a lo largo del tiempo, por ejemplo, una secuencia de tramas de tiempo codificadas por codificación de predicción compleja puede ir seguida de una parte de tiempo codificada por codificación directa izquierdo/derecho. La capacidad del sistema descodificador para manejar diferentes formatos de codificación se logra mediante la provisión de una línea de conexión (paso) dispuesta en paralelo con dicha etapa de mezcla ascendente 1410. Por medio de un conmutador 1411 es posible seleccionar si la salida de la etapa de mezcla ascendente 1410 (posición inferior del conmutador en la figura) o la señal no procesada disponible a través de la línea de conexión (posición superior del conmutador en la figura) se suministrará a los módulos descodificadores dispuestos más adelante aguas abajo. En esta realización, un módulo inverso MDCT 1412 está dispuesto aguas abajo del conmutador, que transforma una representación MDCT de una señal en una representación en el dominio en el tiempo. Como ejemplo, la señal suministrada a la etapa de mezcla ascendente 1410 puede ser una señal estéreo en forma de mezcla descendente/residual. La etapa de mezcla ascendente 1410 está adaptada entonces para derivar una señal lateral y realizar una operación de suma y diferencia para que una señal estéreo izquierdo/derecho (en el dominio MDCT) sea la salida.

La figura 14B muestra un sistema descodificador similar al de la figura 14A. El sistema actual está adaptado para recibir un flujo de bits en su señal de entrada. El flujo de bits se procesa inicialmente mediante un módulo desmultiplexor y de des cuantificación combinado 1420, que proporciona, como primera señal de salida, una representación MDCT de una señal estéreo multicanal para un tratamiento adicional, según lo determina la posición de un conmutador 1422 con funcionalidad análoga a la del conmutador 1411 de la figura 14A. Más precisamente, el conmutador 1422 determina si la primera salida del demultiplexor y de la des cuantificación deben procesarse mediante una etapa de mezcla ascendente 1421 y un módulo inverso MDCT 1423 (posición inferior) o solo mediante el módulo inverso MDCT 1423 (posición superior). Asimismo, el módulo combinado des multiplexor y de des cuantificación 1420 generan información de control. En el presente caso, la información de control asociada a la señal estéreo puede incluir datos que indiquen si la posición superior o inferior del conmutador 1422 es adecuada

para descodificar la señal o, de manera más abstracta, de acuerdo con qué formato de codificación se tiene que descodificar la señal estéreo. La información de control también puede incluir parámetros para ajustar las propiedades de la etapa de mezcla ascendente 1421, por ejemplo, un valor del coeficiente de predicción complejo utilizado en la codificación de predicción compleja como ya se describió anteriormente.

5 La figura 14C muestra un sistema descodificador de acuerdo con la invención reivindicada que, además de las entidades análogas a las de la figura 14B, comprenden primero y segundo dispositivos modificadores en el dominio en frecuencia 1431, 1435, respectivamente, dispuestos aguas arriba y aguas abajo de una etapa de mezcla ascendente 1433. Para los fines de esta figura, cada dispositivo de modificación en el dominio en frecuencia se ilustra mediante un filtro TNS. Sin embargo, con el término dispositivo de modificación en el dominio en frecuencia también se podrían entender otros procesos distintos del filtrado TNS susceptibles de ser aplicados antes o después de la etapa de mezcla ascendente. Ejemplos de modificaciones en el dominio en frecuencia incluyen predicción, adición de ruido, ampliación del ancho de banda y procesamiento no lineal. Consideraciones psico acústicas y razones similares, que posiblemente incluyan las propiedades de la señal a procesar y/o la configuración o ajustes de dicho dispositivo de modificación en el dominio en frecuencia, a veces indican que es ventajoso aplicar dicha modificación en el dominio en frecuencia aguas arriba de la etapa de mezcla ascendente 1433 en lugar de aguas abajo. En otros casos, se puede establecer por consideraciones similares que la posición aguas debajo de la modificación en el dominio en frecuencia es preferible a la de aguas arriba. Por medio de los conmutadores 1432, 1436, los dispositivos modificadores en el dominio en frecuencia 1431, 1435 pueden activarse selectivamente de modo que, en respuesta a la información de control, el sistema descodificador pueda seleccionar la configuración deseada. Como ejemplo, la figura 14C muestra una configuración en la que la señal estéreo del módulo combinado del desmultiplexor y del módulo de des cuantificación 1430 se procese inicialmente por el primer dispositivo de modificación en el dominio en frecuencia 1431, se suministre entonces a la etapa de mezcla ascendente 1433 y finalmente se envíe directamente un módulo inverso MDCT 1437, sin pasar por el segundo dispositivo de modificación en el dominio en frecuencia 1435. Como se explica en la sección *Resumen*, esta configuración se prefiere sobre la opción de realizar el TNS después de la mezcla ascendente en la codificación de predicción compleja.

II. Sistemas codificadores

Ahora se describirá un ejemplo de sistema codificador no contemplado por la invención reivindicada

30 con referencia a la figura 5, que es un diagrama de bloques generalizado de un sistema codificador para codificar una señal estéreo izquierdo/derecho (L/R) como un flujo de bits de salida por codificación de predicción compleja. El sistema codificador recibe una representación en el dominio en el tiempo o en el dominio en frecuencia de la señal y la suministra a una etapa de mezcla descendente y a un estimador del coeficiente de predicción. Las partes real e imaginaria de los coeficientes de predicción se proporcionan a la etapa de mezcla descendente con el fin de controlar la conversión de los canales izquierdo y derecho en canales de mezcla descendente y residual, que luego se suministran a un multiplexor final MUX. Si la señal no se ha suministrado como una representación en el dominio en frecuencia al codificador, se transforma en dicha representación en la etapa de mezcla descendente o multiplexor.

Uno de los principios en la codificación de predicción es convertir la señal izquierdo/derecho

a la forma medio/lateral, es decir,

$$40 \quad \begin{cases} M = (L + R)/2 \\ S = (L - R)/2 \end{cases}$$

y luego hacer uso de la correlación restante entre estos canales, a saber, estableciendo

$$S = \text{Re}\{\alpha M\} + D,$$

45 donde α es el coeficiente de predicción complejo a determinar y D es la señal residual. Es posible elegir α para que la energía de la señal residual $D = S - \text{Re}\{\alpha M\}$ se minimice. La minimización de energía se puede realizar con respecto a la potencia instantánea, una energía de más corto o más largo periodo (potencia media), que en el caso de una señal discreta equivale a optimizar en el sentido del valor cuadrático medio.

50 Las partes real e imaginaria α_R , α_I del coeficiente de predicción pueden ser cuantificadas y/o codificadas conjuntamente. Preferiblemente, sin embargo, las partes real e imaginaria se cuantifican de forma independiente y

uniformemente, por lo general con un tamaño de paso de 0,1 (número adimensional). La resolución de la banda de frecuencia utilizada para el coeficiente de predicción complejo no es necesariamente la misma que la resolución para las bandas de factores de escala (*sfb*; es decir, un grupo de líneas MDCT que utilizan el mismo tamaño de etapa y rango de cuantificación MDCT) de acuerdo con la norma MPEG. En particular, la resolución de la banda de frecuencia para el coeficiente de predicción puede ser una que esté psico acústicamente justificada, tal como la escala Bark. Obsérvese que la resolución de la banda de frecuencia puede variar en los casos en que la longitud de la transformada varíe.

Como ya se señaló, el sistema codificador de acuerdo con el ejemplo puede tener una latitud para aplicar la predicción de codificación estéreo o no, implicando este último caso una regresión a la codificación L/R o M/S. Dicha decisión se puede tomar basándose en la trama de tiempo o mejor, en una banda de frecuencia dentro de una trama de tiempo. Como se señaló anteriormente, un resultado negativo de la decisión puede comunicarse a la entidad de descodificación de varias maneras, por ejemplo, mediante el valor de un bit indicador especial en cada trama, o por la ausencia (o valor cero) de un valor de coeficiente de predicción. Una decisión positiva puede ser comunicada de manera análoga. Una realización especialmente ventajosa, que permite la regresión sin ninguna sobrecarga, hace uso de un cuarto valor reservado del campo de dos bits *ms_mask_present* (véase MPEG-2 AAC, documento ISO/IEC 131818-7), que se transmite por trama de tiempo y se define de la siguiente manera:

Tabla 1: Definición de <i>ms_mask_present</i> en USAC	
Valor	Significado
00	Codificación L/R para todas las bandas de frecuencia
01	Se usa un bit de señalización por banda para indicar L/R o M/S
10	Codificación M/S para todas las bandas de frecuencia
11	Reservado

Al redefinir el valor 11 para que signifique "codificación de predicción compleja", el codificador puede funcionar en todos los modos antiguos, particularmente en la codificación M/S y L/R, sin cualquier penalización de velocidad de bits y aún es capaz de señalar una codificación de predicción compleja para aquellas tramas en las que sea ventajoso.

Lo sustantivo de la decisión puede basarse en una calidad de velocidad audio a datos racional. Como medida de calidad, se pueden usar los datos obtenidos utilizando un modelo psico acústico incluido en el codificador (como suele ser el caso de los codificadores de audio disponibles basados en MDCT). En particular, algunas realizaciones del codificador proporcionan una selección optimizada de velocidad-distorsión del coeficiente de predicción. De acuerdo con esto, en tales realizaciones, la parte imaginaria - y posiblemente la parte real también - del coeficiente de predicción se fija a cero si el aumento en la ganancia de predicción no economiza suficientes bits para que la codificación de la señal residual justifique el gasto de los bits necesarios para codificar el coeficiente de predicción.

Ejemplos, no contemplados por la invención reivindicada, del codificador pueden codificar la información relativa al TNS en el flujo de bits. Dicha información puede incluir valores de los parámetros del TNS para ser aplicados por los filtros TNS (síntesis) en el lado del descodificador. Si se van a utilizar conjuntos idénticos de parámetros del TNS para ambos canales, es económico incluir un bit de señalización que indique esta identidad de los conjuntos de parámetros en lugar de transmitir los dos conjuntos de parámetros de forma independiente. La información también puede ser incluida también de si aplicar TNS antes o después de la etapa mezcla ascendente, según corresponda basándose en, por ejemplo, una evaluación psico acústica de las dos opciones disponibles.

Como otra característica opcional, que es potencialmente beneficiosa desde un punto de vista de complejidad y velocidad de bits, el codificador puede adaptarse para usar un ancho de banda limitado individualmente para la codificación de la señal residual. Las bandas de frecuencia por encima de este límite no se transmitirán al descodificador, sino que se pondrán a cero. En ciertos casos, las bandas de frecuencia más altas tienen un contenido de energía tan pequeño que ya están cuantificadas a cero. La práctica normal (configurar el parámetro *max_sfb* en la norma MPEG) ha implicado el uso de la misma limitación de ancho de banda para las señales de mezcla descendente y residual. Ahora, los inventores han encontrado empíricamente que la señal residual, a un nivel mayor que la señal de mezcla descendente, tiene su contenido de energía localizado en las bandas de frecuencias más bajas. Por lo tanto, al colocar una banda superior especial con límite en la señal residual, es posible una reducción de la velocidad de bits sin una pérdida significativa de la calidad. Por ejemplo, esto se puede lograr mediante la transmisión de dos parámetros *max_sfb* independientes, uno para la señal de mezcla descendente y otro para la señal residual.

Se señala que si bien las cuestiones de determinación óptima del coeficiente de predicción, su cuantificación y codificación, la regresión al modo M/S o L/R, el filtrado TNS y la limitación del ancho de banda superior, etc., se describieron con referencia al sistema descodificador que se muestra en la figura 5, los mismos hechos son igualmente aplicables a las realizaciones que se describirán a continuación con referencia a las figuras siguientes.

5 La figura 6 muestra otro ejemplo de sistema codificador no contemplado por la invención reivindicada adaptado para realizar codificación estéreo de predicción compleja. El sistema recibe como entrada una representación en el dominio en el tiempo de una señal estéreo segmentada en sucesivas, posiblemente solapadas, tramas de tiempo y que comprenden los canales izquierdo y derecho. Una etapa de suma y diferencia 601 convierte la señal en canales medios y laterales. El canal intermedio se suministra tanto a un módulo MDCT 602 como a un módulo MDST 603, mientras que el canal lateral se suministra solo a un módulo MDCT 604. Un estimador del coeficiente de predicción 605 estima para cada trama de tiempo - y posiblemente para bandas de frecuencia individuales dentro de una trama - un valor del coeficiente de predicción complejo α como se explicó anteriormente. El valor del coeficiente α se suministra ponderado a los sumadores ponderados 606, 607, que forman una señal residual D como una combinación lineal de las representaciones MDCT y MDST de la señal del canal medio y de la representación MDCT de la señal lateral. Preferiblemente, el coeficiente de predicción complejo se suministra a los sumadores ponderados 606, 607 representado por el mismo esquema de cuantificación que se utilizará cuando se codifique en el flujo de bits; esto obviamente proporciona una reconstrucción más fiel, ya que tanto el codificador como el descodificador aplican el mismo valor del coeficiente de predicción. La señal residual, la señal del canal medio (que puede llamarse más apropiadamente señal de mezcla descendente cuando aparece en combinación con una señal residual) y el coeficiente de predicción se suministran a una etapa combinada de cuantificación y multiplexor 608, que codifica estos y la posible información adicional como un flujo de bits de salida.

La figura 7 muestra una variación del sistema codificador de la figura 6. Como se desprende de la similitud de los símbolos en la figura, tiene una estructura similar pero también la funcionalidad adicional de funcionar en un modo de regresión directa de codificación L/R. El sistema codificador actúa entre el modo de predicción compleja y el modo de regresión mediante un conmutador 710 provisto inmediatamente aguas arriba de la etapa combinada de cuantificación y multiplexor 709. En su posición superior, como se muestra en la figura, el conmutador 710 hará que el codificador opere en el modo de regresión. Desde puntos inmediatamente aguas abajo de los módulos MDCT 702, 704, la señal medio/lateral se suministra a una etapa de suma y diferencia 705, que después de convertirla en la forma izquierdo/derecho lo pasa al conmutador 710, que lo conecta a la etapa de cuantificación combinada y multiplexión 709.

La figura 8 muestra un ejemplo de codificador no contemplado por la invención reivindicada.

En contraste con los sistemas codificadores de las figuras 6 y 7, este ejemplo deriva los datos MDST necesarios para la codificación de predicción compleja directamente a partir de los datos MDCT, es decir, mediante una transformada real a imaginaria en el dominio en frecuencia. La transformada real a imaginaria aplica cualquiera de los enfoques descritos en relación con los sistemas de descodificación de las figuras 2 y 4. Es importante que coincida el método de cálculo del descodificador con el del codificador, de modo que se pueda lograr una descodificación fiel; preferiblemente, utilizándose métodos idénticos de transformada real a imaginaria en el lado del codificador y en el lado del descodificador. En cuanto a las realizaciones del descodificador, la parte A rodeada por una línea discontinua y que comprende la transformada real a imaginaria 804 se puede reemplazar por variantes aproximadas o utilizando menos tramas de tiempo de entrada como entrada. Del mismo modo, la codificación se puede simplificar utilizando cualquiera de los otros enfoques de aproximación descritos anteriormente.

En un nivel superior, el sistema codificador de la figura 8 tiene una estructura diferente de la que probablemente seguiría por una acción directa de reemplazo del módulo MDST en la figura 7 por un módulo real - imaginario (adecuadamente conectado). La arquitectura actual es limpia y logra la funcionalidad de conmutación entre la codificación de predicción y la codificación directa de L/R de una manera firme y de manera computacionalmente económica. La señal de entrada estéreo se lleva a los módulos de la transformada MDCT 801, que generan una representación en el dominio en frecuencia de cada canal. Esto se lleva tanto a un conmutador final 808 para accionar el sistema codificador entre sus modos de predicción y codificación directa y hasta una etapa de suma y diferencia 802. En la codificación directa L/R o codificación conjunta M/S - que se lleva a cabo en la trama de tiempo para la cual el coeficiente de predicción α se fija a cero - este ejemplo somete la señal de entrada solo a la transformada MDCT, cuantificación y multiplexión, efectuándose las dos últimas etapas mediante una etapa combinada de multiplexión y cuantificación 807 dispuesta en el extremo de salida del sistema, donde se suministra un flujo de bits. En la codificación de predicción, cada uno de los canales se somete a un procesamiento adicional entre la etapa de suma y diferencia 802 y el conmutador 808. De la representación MDCT de la señal del canal medio, la transformada real a imaginaria 804 deriva los datos MDST y los envía tanto a un estimador del coeficiente de predicción 803 como a un sumador ponderado 806. Al igual que en los sistemas codificadores que se muestran en las figuras 6 y 7, un sumador ponderado adicional 805 se usa para combinar la señal lateral con las

representaciones ponderadas MDCT y MDST de la señal del canal medio para formar una señal de canal residual, que se codifica junto con la señal del canal medio (es decir, mezcla descendente) y el coeficiente de predicción mediante la cuantificación combinada y el módulo multiplexor 807.

5 Volviendo ahora a la figura 9, se ilustrará que cada uno de los ejemplos del sistema codificador se puede combinar con uno o más filtros TNS (análisis). De acuerdo con las descripciones anteriores, a menudo es ventajoso aplicar el filtro TNS a la señal en su forma de mezcla descendente. Por lo tanto, como se muestra en la figura 9, la adaptación del sistema codificador de la figura 7 para incluir el TNS se efectúa agregando filtros TNS 911 inmediatamente aguas arriba del módulo de multiplexión y cuantificación combinada 909.

10 En lugar del filtro TNS derecho/residual 911 b, dos filtros TNS separados (no mostrados) pueden proporcionarse inmediatamente aguas arriba de la parte del conmutador 910 adaptado para manejar el canal derecho o residual. Por lo tanto, cada uno de los dos filtros TNS se suministrará con los datos de señal del canal respectivo en todo momento, permitiendo el filtrado TNS basado en más tramas de tiempo que solo la actual. Como ya se ha señalado, los filtros TNS son solo un ejemplo de los dispositivos modificadores en el dominio en frecuencia, en particular los dispositivos que basan su procesamiento en más tramas que la actual, lo que puede beneficiarse de tal posición tanto o más que en el filtro TNS.

15 Como otra alternativa posible al ejemplo mostrado en la figura 9, los filtros TNS para la activación selectiva se pueden organizar en más de un punto para cada canal. Esto es similar a la estructura del sistema descodificador que se muestra en la figura 4, donde se pueden conectar diferentes conjuntos de filtros TNS por medio de conmutadores. Esto permite la selección de la etapa disponible más adecuada para el filtrado TNS para cada trama de tiempo. En particular, puede ser ventajoso cambiar entre diferentes posiciones el TNS en relación con la conmutación entre la codificación estéreo de predicción compleja y otros modos de codificación.

20 La figura 11 muestra una variación basada en el sistema codificador de la figura 8, en el que una segunda representación en el dominio en frecuencia de la señal de mezcla descendente se deriva por medio de una transformada real a imaginaria 1105. De manera similar al sistema descodificador mostrado en la figura 4, este sistema codificador también incluye módulos modificadores activables selectivamente en el dominio en frecuencia, uno 1102 provisto aguas arriba de la etapa de mezcla descendente y otro 1109 provisto aguas debajo de la misma. Los módulos en el dominio en frecuencia 1102, 1109, que en esta figura han sido ejemplificados por filtros TNS, se pueden conectar a cada una de las rutas de señal utilizando los cuatro conmutadores 1103a, 1103b, 1109a y 1109b.

III. Realizaciones sin aparato

30 Las realizaciones de los aspectos tercero y cuarto mencionados anteriormente se muestran en las figuras 15 y 16. La figura 15 muestra un método para descodificar un flujo de bits en una señal estéreo, que comprende las siguientes etapas:

1. Se introduce un flujo de bits.
- 35 2. El flujo de bits se des cuantifica, por lo que se obtiene una primera representación en el dominio en frecuencia de los canales de mezcla descendente y residual de una señal estéreo.
3. Se calcula una segunda representación en el dominio en frecuencia del canal de la mezcla descendente.
4. Se calcula una señal del canal lateral basándose en las tres representaciones en el dominio en frecuencia de los canales.
- 40 5. Una señal estéreo, preferiblemente en forma izquierdo/ derecho, se calcula basándose en los canales lateral y de mezcla descendente
6. Se envía la señal estéreo así obtenida.

Las etapas 3 a 5 se pueden considerar como un proceso de mezcla ascendente. Cada una de las etapas 1 a 6 es análogo a la funcionalidad correspondiente en cualquiera de los sistemas codificadores descritos en las partes anteriores de este texto, y más adelante los detalles relacionados con su realización se pueden recuperar en las mismas partes.

La figura 16 muestra un método para codificar una señal estéreo como una señal de flujo de bits, que comprende las siguientes etapas:

1. Se recibe una señal estéreo.
2. La señal estéreo se transforma en una primera representación en el dominio en frecuencia.

3. Se determina un coeficiente de predicción complejo.

4. La representación en el dominio en frecuencia se mezcla descendentemente.

5. Los canales de mezcla descendente y residual se codifican como un flujo de bits junto con el coeficiente de predicción complejo.

5 6. Se envía el flujo de bits.

Cada una de las etapas 1 a 5 es análoga a la funcionalidad correspondiente en cualquiera de los sistemas codificadores descritos en las partes anteriores de este texto y más detalles relacionados con su realización se pueden recuperar en las mismas partes.

10 Ambos métodos pueden expresarse como instrucciones interpretables por ordenador en la forma de programas de software y ejecutados por un ordenador. El alcance de la protección de esta invención se extiende a dicho software y a los productos de programa para la distribución de dicho software.

IV. Evaluación empírica

Varias de las realizaciones aquí descritas han sido evaluadas empíricamente. Las partes más importantes del material experimental obtenido en este proceso se resumirán en esta subsección.

15 La realización utilizada para los experimentos tuvo las siguientes características:

(i) Cada espectro MDST (para una trama de tiempo) se calculó mediante un filtrado de respuesta de impulso finito bidimensional a partir del espectro TCMD actual, anterior y siguiente.

(ii) Se utilizó un modelo psico acústico del codificador estéreo USAC.

20 (iii) Las partes real e imaginaria del coeficiente de predicción complejo α se transmitieron en lugar de los parámetros PS ICC, CLD e IPD. Las partes real e imaginaria se manejaron de forma independiente, se limitaron al rango [-3,0, 3,0] y se cuantificaron utilizando un tamaño de escalón de 0,1. Luego se codificaron diferencialmente en el tiempo y finalmente se codificaron en Huffman utilizando el libro de códigos de factor de escala USAC. Los coeficientes de predicción fueron actualizados cada segunda banda de factor de escala, lo que dio lugar a una resolución en frecuencia similar a la de MPEG Surround (véase, por ejemplo, ISO/IEC 23003-1). Este esquema de cuantificación y
25 codificación dio como resultado una velocidad de bits media de aproximadamente 2 kb/s para esta información lateral estéreo en una configuración típica con una velocidad de bits objetivo de 96 kb/s.

(iv) El formato del flujo de bits se modificó sin romper el flujo de bits USAC actual, ya que el elemento de flujo de bits presente de 2 bits *ms_mask_present* tiene actualmente solo tres valores posibles. Al utilizar el cuarto valor para
30 indicar una predicción compleja, se permite un modo de regresión de la codificación básica medio/lateral sin malgastar ningún bit (para más detalles sobre este tema, consúltese la subsección previa de esta descripción).

Las pruebas de audición se realizaron de acuerdo con la metodología MUSHRA, lo que implica en particular la reproducción a través de auriculares y el uso de 8 elementos de prueba con una velocidad de muestreo de 48 kHz. Tres, cinco o seis individuos de prueba participaron en cada prueba.

35 El impacto de las diferentes aproximaciones MDST fue evaluado para ilustrar la relación práctica de complejidad a calidad que existe entre estas opciones. Los resultados se encuentran en las figuras 12 y 13, mostrando las primeras puntuaciones absolutas obtenidas y las últimas puntuaciones diferenciales en relación con *96s USAC cp1f*, es decir, codificación estéreo unificada en el dominio MDCT por predicción compleja usando una trama MDCT actual para calcular una aproximación de la MDST. Se puede ver que la ganancia de calidad lograda por la codificación estéreo unificada basada en MDCT aumenta cuando se aplican más enfoques computacionalmente complejos para
40 calcular el espectro MDST. Considerando el promedio de todas las pruebas, el sistema *96s USAC cp1f* basado en una sola trama proporciona un incremento significativo en la eficiencia de codificación sobre la codificación estéreo convencional. A su vez, incluso se obtienen resultados significativamente mejores para *96s USAC cp3f*, a saber, codificación estéreo unificada en el dominio-MDCT mediante codificación compleja usando las tramas MDCT actuales, anteriores y siguientes para calcular la MDST.

45 V. Observaciones finales

Otras realizaciones de la presente invención serán evidentes para un experto en la materia después de leer la descripción anterior. Aunque la presente descripción y los dibujos describen realizaciones y ejemplos, la invención no está restringida a estos ejemplos específicos. Pueden realizarse numerosas modificaciones y variaciones sin apartarse del alcance de la presente invención, que se define en las reivindicaciones adjuntas.

5 Se observa que los métodos y aparatos descritos en esta solicitud pueden aplicarse, después de las modificaciones apropiadas dentro de las capacidades del experto en la materia, incluida la experimentación rutinaria, a la codificación de señales que tengan más de dos canales. Se enfatiza particularmente que cualesquiera señales, parámetros y matrices mencionadas en relación con las realizaciones descritas pueden ser variantes en frecuencia o invariantes en frecuencia y/o variantes en el tiempo o invariantes en el tiempo.

10 Las etapas de computación descritas pueden llevarse a cabo en modo frecuencia o para todas las bandas de frecuencia a la vez, y todas las entidades pueden realizarse para tener una acción selectiva en frecuencia. Para los fines de la aplicación, cualquier esquema de cuantificación puede adaptarse según los modelos psico acústicos. Además, se observa que las diversas conversiones de suma y diferencia, es decir, la conversión de la forma de mezcla descendente/residual a la forma pseudo-L/R, así como la conversión L/R a M/S y la conversión M/S a L/R son todas de la forma

$$g \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

15 donde, simplemente, el factor de ganancia g puede variar. Por lo tanto, al ajustar los factores de ganancia individualmente, es posible compensar una cierta ganancia de codificación mediante una elección apropiada de la ganancia de descodificación. Además, como puede observar el experto, un número par de transformadas de suma y diferencia dispuestas en serie tienen el efecto de una etapa de paso, posiblemente con ganancia de no unidad.

20 Los sistemas y métodos descritos en este documento anteriormente pueden realizarse en software, firmware, hardware o una combinación de los mismos. Ciertos componentes o todos los componentes pueden realizarse como software ejecutado por un procesador digital de señal o microprocesador, o realizarse como hardware o como un circuito integrado de aplicación específica. Tal software puede ser distribuido en medios interpretables por ordenador, que pueden comprender medios de almacenamiento en ordenador y medios de comunicación. Como es bien conocido por los expertos en la técnica, los medios de almacenamiento en ordenador incluyen medios volátiles y no volátiles, extraíbles y no extraíbles realizados en cualquier método o tecnología para el almacenamiento de la información, como instrucciones interpretables por ordenador, estructuras de datos, módulos de programa u otros datos. Los medios de almacenamiento en ordenador incluyen, pero no se limitan a, RAM, ROM, EEPROM, memoria flash u otra tecnología de memoria, CD-ROM, discos versátiles digitales (DVD) u otro almacenamiento en disco óptico, casetes magnéticos, cinta magnética, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio que se pueda utilizar para almacenar la información deseada y que se puede acceder por un ordenador. Además, el experto en la materia sabe que los medios de comunicación normalmente incorporan instrucciones interpretables por ordenador, estructuras de datos, módulos de programa u otros datos en una señal de datos modulada, como una onda portadora u otro mecanismo de transporte, e incluye cualquier medio de entrega de información.

30

REIVINDICACIONES

1. Un sistema descodificador para proporcionar una señal estéreo por codificación estéreo de predicción compleja, que comprende:

5 una etapa de mezcla ascendente (406, 407, 408, 409; 1433) adaptada para generar la señal estéreo basada en las primeras representaciones en el dominio en frecuencia de una señal de mezcla descendente (M) y una señal residual (D), comprendiendo cada una de las primeras representaciones en el dominio en frecuencia primeros componentes espectrales que representan el contenido espectral de la señal correspondiente expresada en un primer subespacio de un espacio multidimensional, comprendiendo la etapa de mezcla ascendente:

10 un módulo (408) para calcular una segunda representación en el dominio en frecuencia de la señal de mezcla descendente basada en su primera representación en el dominio de la frecuencia, comprendiendo la segunda representación en el dominio en frecuencia los segundos componentes espectrales que representan el contenido espectral de la señal expresada en un segundo subespacio del espacio multidimensional que incluye una parte del espacio multidimensional no incluido en el primer subespacio;

15 un sumador ponderado (406, 407) para calcular una señal lateral (S) basándose en la primera y segunda representaciones en el dominio en frecuencia de la señal de mezcla descendente, en la primera representación en el dominio en frecuencia de la señal residual y en un coeficiente de predicción complejo (α) codificado en la señal del flujo de bits; y

20 una etapa de suma y diferencia (409) para calcular la señal estéreo basándose en la primera representación en el dominio en frecuencia de la señal de mezcla descendente y en la señal lateral caracterizado por:

25 una primera etapa modificadora en el dominio en frecuencia (403; 1431), dispuesta aguas arriba de la etapa de mezcla ascendente y que funciona en modo activo, en la que procesa una representación en el dominio en frecuencia de al menos una de las señales de mezcla descendente y/o la señal residual, y un modo pasivo, en la que actúa como paso; y

30 una segunda etapa modificadora en el dominio en frecuencia (410; 1435), dispuesta aguas abajo de la etapa de mezcla ascendente y que funciona en un modo activo, en la que procesa una representación en el dominio en frecuencia de al menos un canal de la señal estéreo, y un modo pasivo, en el que actúa como paso, mientras que la primera y la segunda etapas modificadoras en el dominio en frecuencia están configuradas para ser activadas selectivamente en respuesta a una información de control.

2. El sistema descodificador de la reivindicación 1, en el que al menos una de dichas etapas modificadoras en el dominio en frecuencia es una etapa, TNS, con forma de ruido temporal.

35 3. El sistema descodificador de la reivindicación 2, adicionalmente adaptado para recibir, para cada trama de tiempo, un campo de datos asociado con esa trama y para hacer funcionar, en respuesta al valor del campo de datos, la primera etapa modificadora en el dominio en frecuencia en su modo activo o en su modo de paso y la segunda etapa modificadora en el dominio en frecuencia en su modo activo o en su modo de paso.

4. El sistema descodificador de cualquiera de las reivindicaciones precedentes, que comprende, además:

40 una etapa de des cuantificación (401), dispuesta aguas arriba de la etapa de mezcla ascendente, para proporcionar dichas primeras representaciones en el dominio en frecuencia de la señal de mezcla descendente (M) y de la señal residual (D) basadas en una señal de flujo de bits.

5. El sistema descodificador de cualquiera de las reivindicaciones precedentes, en el que:

los primeros componentes espectrales tienen valores reales expresados en el primer subespacio;

los segundos componentes espectrales tienen valores imaginarios expresados en el segundo subespacio;

opcionalmente, los primeros componentes espectrales se pueden obtener mediante uno de las siguientes:

45 una transformada discreta del coseno, DCT, o

una transformada discreta del coseno modificada, MDCT, y

opcionalmente, los segundos componentes espectrales se pueden obtener mediante uno de las siguientes:

una transformada discreta del seno, DST, o

una transformada discreta del seno modificada, MDST.

6. El sistema descodificador de la reivindicación 5, en el que:

5 la señal de mezcla descendente se divide en tramas de tiempo sucesivas, cada una asociada con un valor del coeficiente de predicción complejo; y el módulo para calcular una segunda representación en el dominio en frecuencia de la señal de mezcla descendente está adaptado para desactivarse a sí mismo, en respuesta al valor absoluto de la parte imaginaria del coeficiente de predicción complejo, que sea menor que una tolerancia predeterminada para una trama de tiempo, de modo que no se genere salida para esa trama de tiempo.

10 7. El sistema descodificador de cualquiera de las reivindicaciones precedentes, en el que:

los primeros componentes espectrales son coeficientes de la transformada dispuestos en una o más tramas de tiempo de coeficientes de la transformada, cada bloque generado por la aplicación de una transformada a un segmento de tiempo de una señal en el dominio en el tiempo; y

15 el módulo para calcular una segunda representación en el dominio en frecuencia de la señal de mezcla descendente está adaptado para:

derivar uno o más primeros componentes intermedios de al menos algunos de los primeros componentes espectrales;

20 formar una combinación de dicho uno o más primeros componentes espectrales de acuerdo con al menos una parte de una o más respuestas de impulso para obtener uno o más segundos componentes intermedios; y

derivar dicho uno o más segundos componentes espectrales de dicho uno o más segundos componentes intermedios.

25 8. El sistema descodificador de la reivindicación 7, en el que las partes de la una o más respuestas de impulso se basan en las características de respuesta en frecuencia de la una o más transformadas, en donde opcionalmente las características de respuesta en frecuencia de la transformada dependen de las características de una o más más funciones de la ventana de análisis que se aplicaron con la transformada al segmento de tiempo de la señal.

9. El sistema descodificador de la reivindicación 7 u 8, en el que el módulo para calcular una segunda representación en el dominio en frecuencia de la señal de mezcla descendente está adaptado para obtener cada trama de tiempo de los segundos componentes espectrales en función de una de las siguientes:

30 (a) una trama de tiempo contemporáneo de los primeros componentes espectrales; y

(b) una trama de tiempo contemporáneo y anterior de los primeros componentes espectrales; y

(c) una trama de tiempo contemporáneo, anterior y posterior de los primeros componentes espectrales.

35 10. El sistema descodificador de cualquiera de las reivindicaciones precedentes, en el que el módulo para calcular una segunda representación en el dominio en frecuencia de la señal de mezcla descendente está adaptado para calcular una segunda representación espectral aproximada que comprende una segunda representación espectral aproximada determinada por la combinación de al menos dos primeros componentes espectrales contiguos en el tiempo y/o contiguos en frecuencia.

11. El sistema descodificador de cualquiera de las reivindicaciones precedentes, estando representada dicha señal estéreo en el dominio en el tiempo y comprendiendo adicionalmente el sistema descodificador:

40 un conjunto de conmutación (203) dispuesto entre dicha etapa de des cuantificación y dicha etapa de mezcla ascendente, que funciona como:

(a) una etapa de paso, o

(b) una etapa de suma y diferencia,

permitiendo así la conmutación entre señales de entrada estéreo codificadas directa y conjuntamente;

una etapa de transformada inversa (209) adaptada para calcular una representación en el dominio en el tiempo de la señal estéreo; y

una disposición de selector (208), situado aguas arriba de la etapa de la transformada inversa, adaptada para conectar esto de forma selectiva a:

- 5 (a) un punto aguas abajo de la etapa de mezcla ascendente, por lo que la señal estéreo obtenida por predicción compleja se suministra a la etapa de la transformada inversa; o
- (b) un punto aguas abajo del conjunto de conmutación (203) y aguas arriba de la etapa mezcla ascendente, por lo que se obtiene una señal estéreo obtenida por codificación estéreo directa que se suministra a la etapa de la transformada inversa.

- 10 12. El sistema descodificador de cualquiera de las reivindicaciones precedentes, en el que el módulo para calcular una segunda representación en el dominio en frecuencia de la señal de mezcla descendente comprende:

una etapa de transformada inversa (306) para calcular una representación en el dominio en el tiempo de la señal de mezcla descendente y/o de la señal lateral basándose en la primera representación en el dominio en frecuencia de la señal respectiva en el primer subespacio del espacio multidimensional; y

- 15 una etapa de transformada (307) para calcular la segunda representación en el dominio en frecuencia de la señal respectiva basándose en la representación de la señal en el dominio en el tiempo,
- en donde, preferiblemente, la etapa de la transformada inversa (306) realiza una transformada inversa discreta del coseno modificada, MDCT, y la etapa de la transformada realiza una transformada discreta del seno modificada, MDST.

- 20 13. El sistema descodificador de la reivindicación 12, estando representada dicha señal estéreo en el dominio en el tiempo y comprendiendo el sistema descodificador, además:

un conjunto de conmutación (302) dispuesto entre dicha etapa de des cuantificación y dicha etapa de mezcla ascendente, que funciona como:

- (a) una etapa de paso, para uso en la codificación estéreo conjunta; o

- 25 (b) una etapa de suma y diferencia, para uso en la codificación estéreo directa;

una etapa de transformada inversa adicional (311) dispuesta en la etapa de mezcla ascendente, para calcular una representación en el dominio en el tiempo de la señal lateral;

una disposición de selector (305, 310), dispuesto aguas arriba de las etapas de la transformada inversa (306, 311), adaptado para conectarlas de forma selectiva a:

- 30 (a) una etapa adicional de suma y diferencia (304) que a su vez está conectada a un punto aguas abajo del conjunto de conmutación (302) y aguas arriba de la etapa de mezcla ascendente; o

- (b) una señal de mezcla descendente obtenida del conjunto de conmutación (302) y de una señal lateral obtenida del sumador ponderado (308, 309).

- 35 14. Un método de descodificación para la mezcla ascendente de una señal estéreo de entrada por codificación estéreo de predicción compleja en una señal estéreo de salida, en la que:

dicha señal estéreo de entrada comprende una primera representación en el dominio en frecuencia

de una señal de mezcla descendente (M) y una señal residual (D) y un coeficiente de predicción complejo (α); y

- 40 cada una de dichas primeras representaciones en el dominio en frecuencia comprende primeros componentes espectrales que representan el contenido espectral de las señales correspondientes expresadas en un primer subespacio de un espacio multidimensional,

estando realizado el método en una etapa de mezcla ascendente e incluye las etapas de:

- 45 calcular una segunda representación en el dominio en frecuencia de la señal de mezcla descendente basándose en su primera representación en el dominio en frecuencia, comprendiendo la segunda representación en el dominio en frecuencia los segundos componentes espectrales que representan el

contenido espectral de la señal expresada en un segundo subespacio del espacio multidimensional que incluye una parte del espacio multidimensional no incluido en el primer subespacio;

5 calcular una señal lateral basándose en la primera y la segunda representaciones en el dominio en frecuencia de la señal de mezcla descendente, en la primera representación en el dominio en frecuencia de la señal residual y en el coeficiente de predicción complejo;

y caracterizado por

comprender además la etapa, a realizar antes de la etapa de mezcla ascendente, de aplicar la formación de ruido temporal, TNS, a dicha primera representación en el dominio en frecuencia de la señal de mezcla descendente y/o a dicha primera representación en el dominio en frecuencia de la señal residual;

10 o la etapa, a realizar después de la etapa de mezcla ascendente, de aplicar TNS a al menos una señal de dicha señal estéreo.

15. Un producto de programa de ordenador que comprende un medio interpretable por ordenador que almacena instrucciones que, cuando se ejecutan por un ordenador de propósito general, realizan el método expuesto en la reivindicación 14.

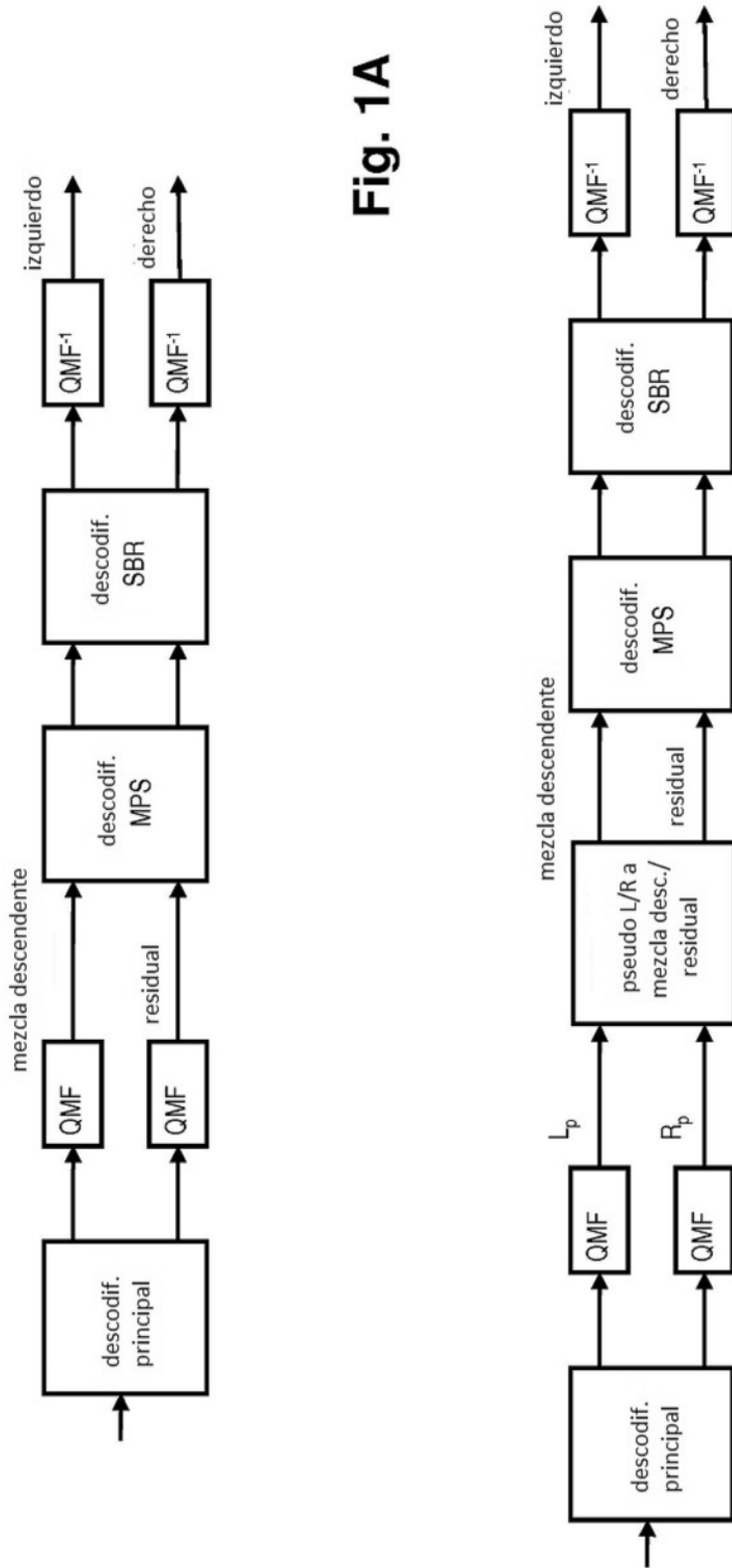


Fig. 1A

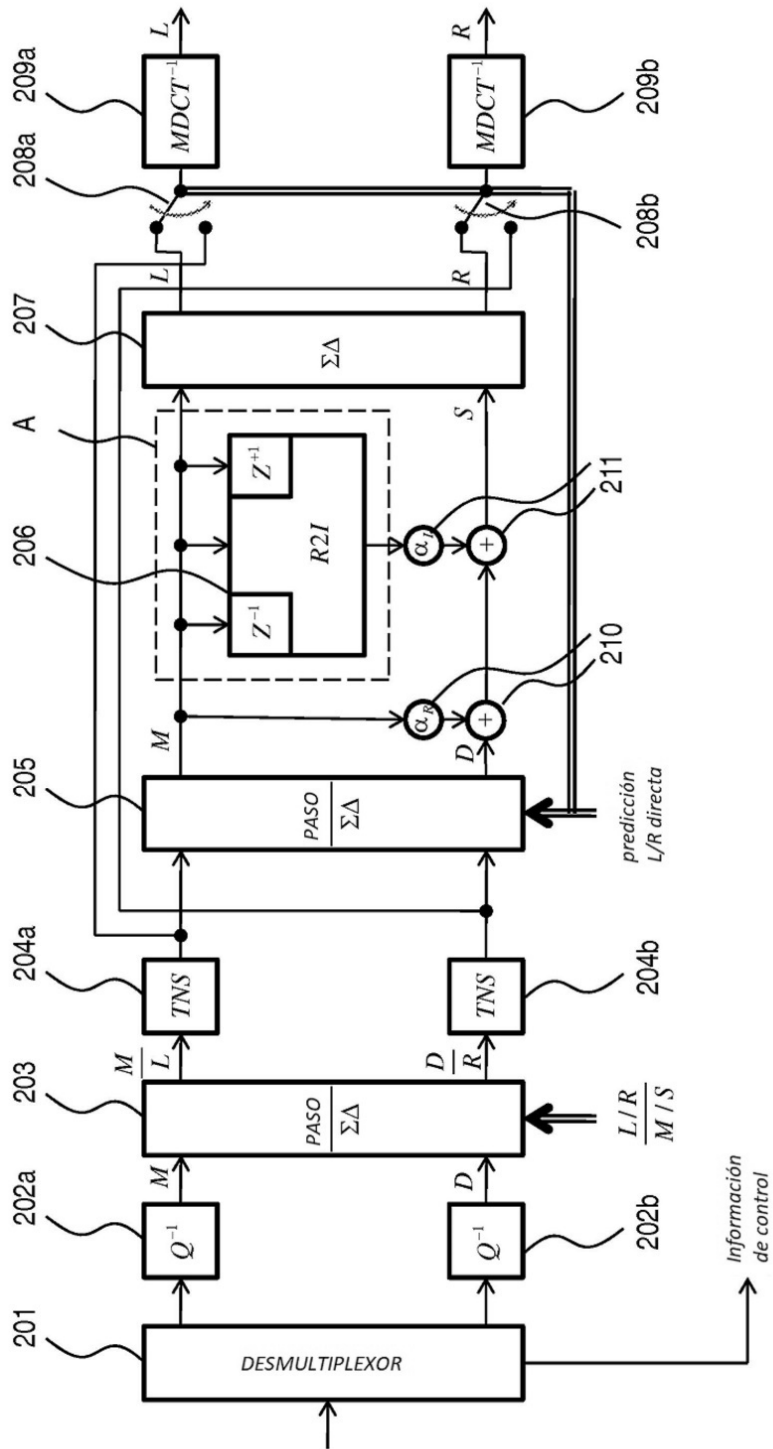


Fig. 2

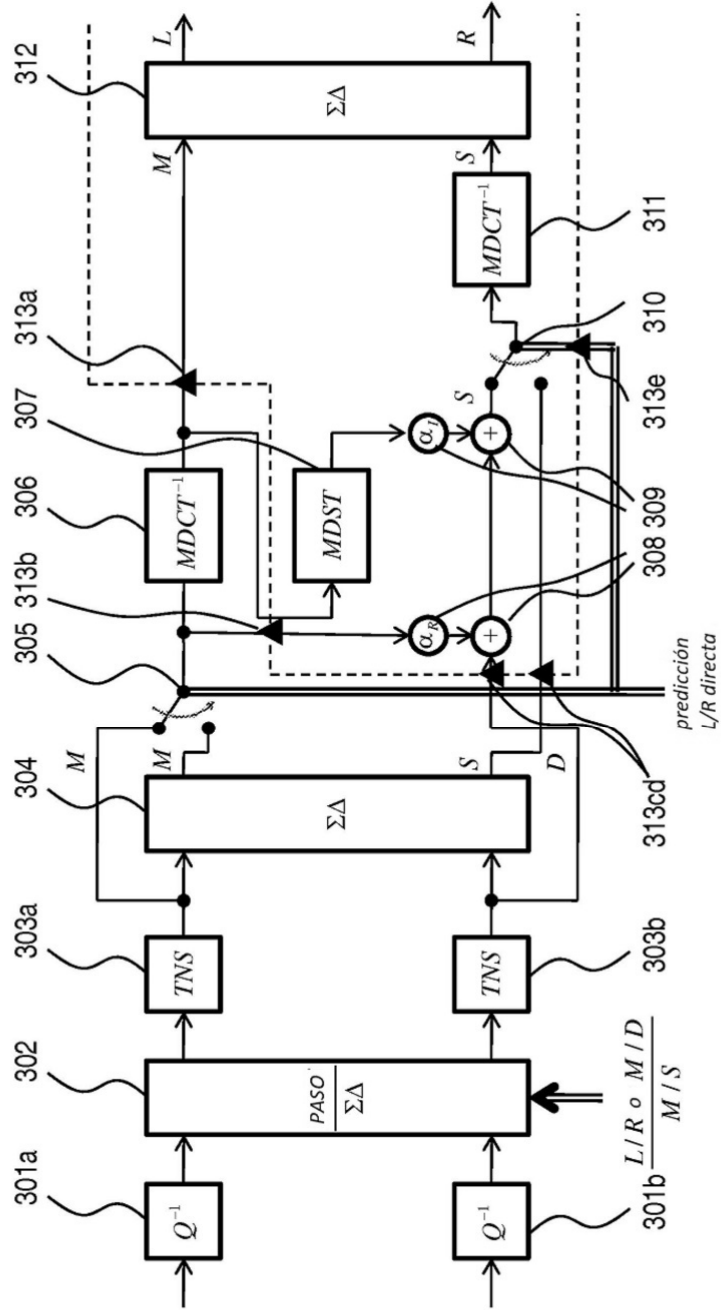


Fig. 3

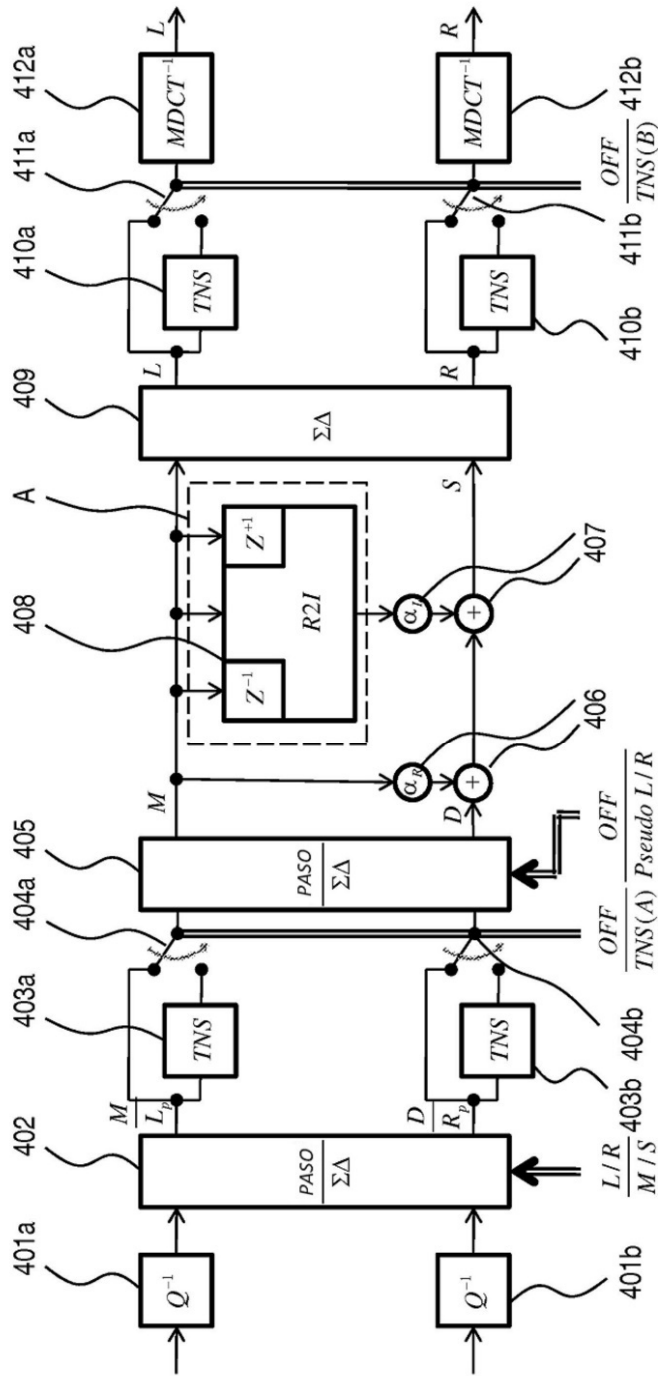


Fig. 4

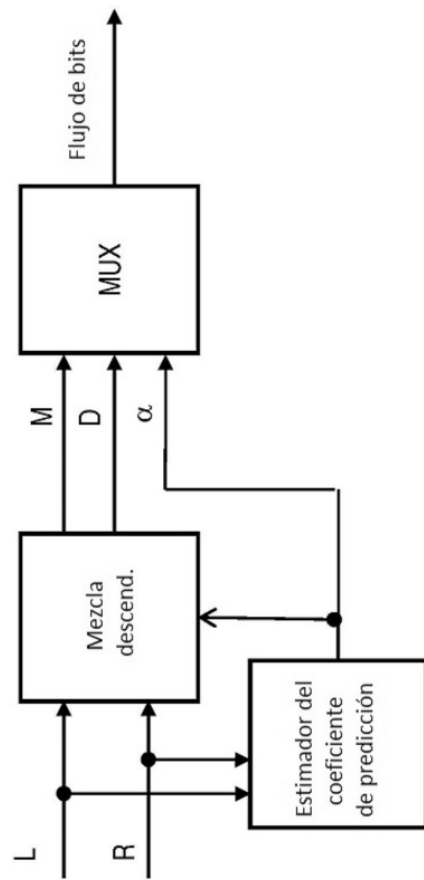


Fig. 5

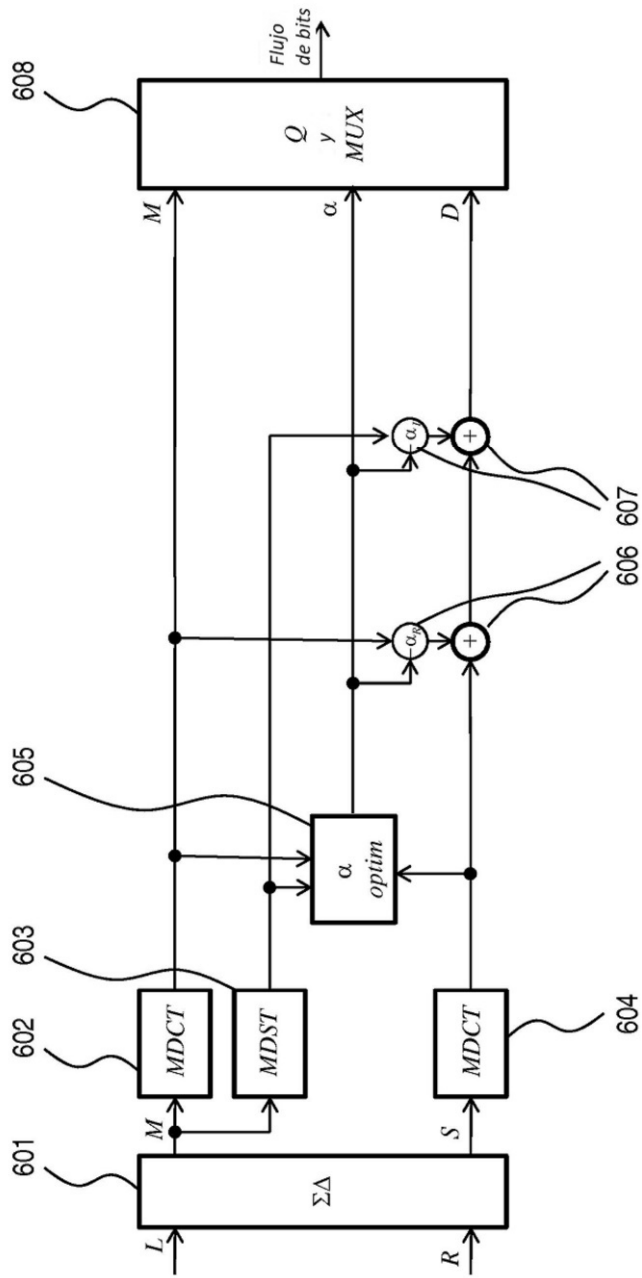


Fig. 6

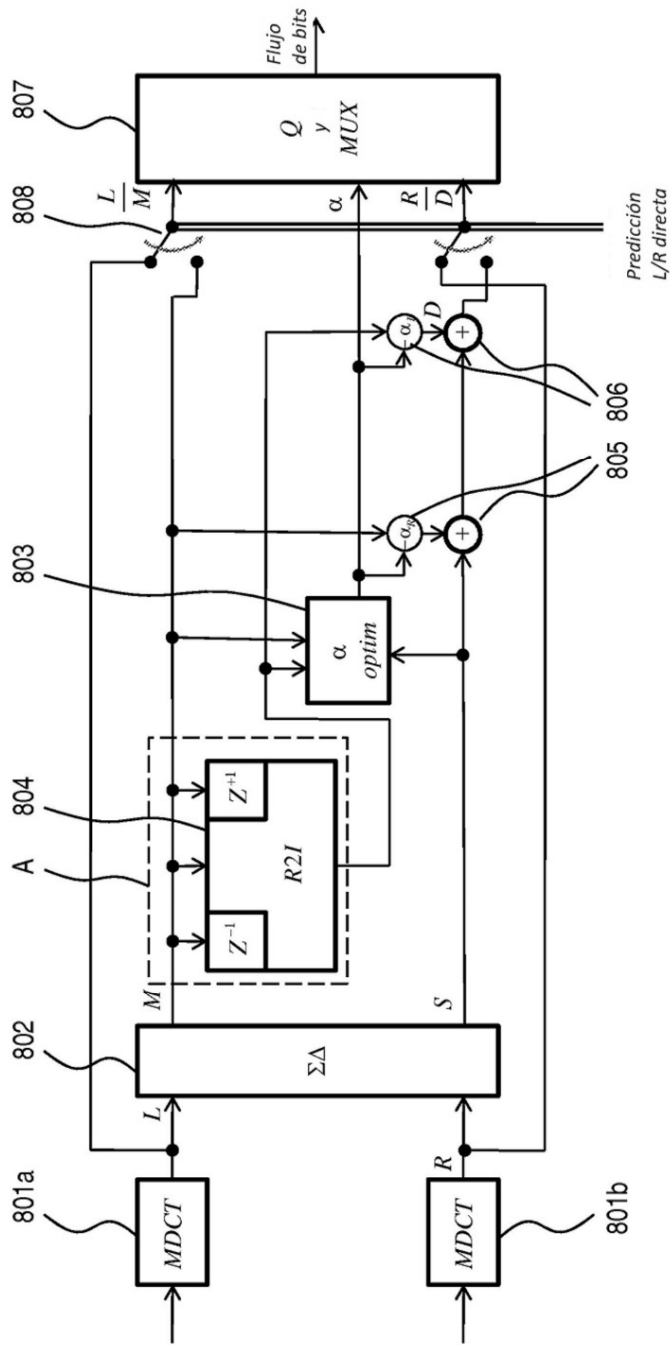


Fig. 8

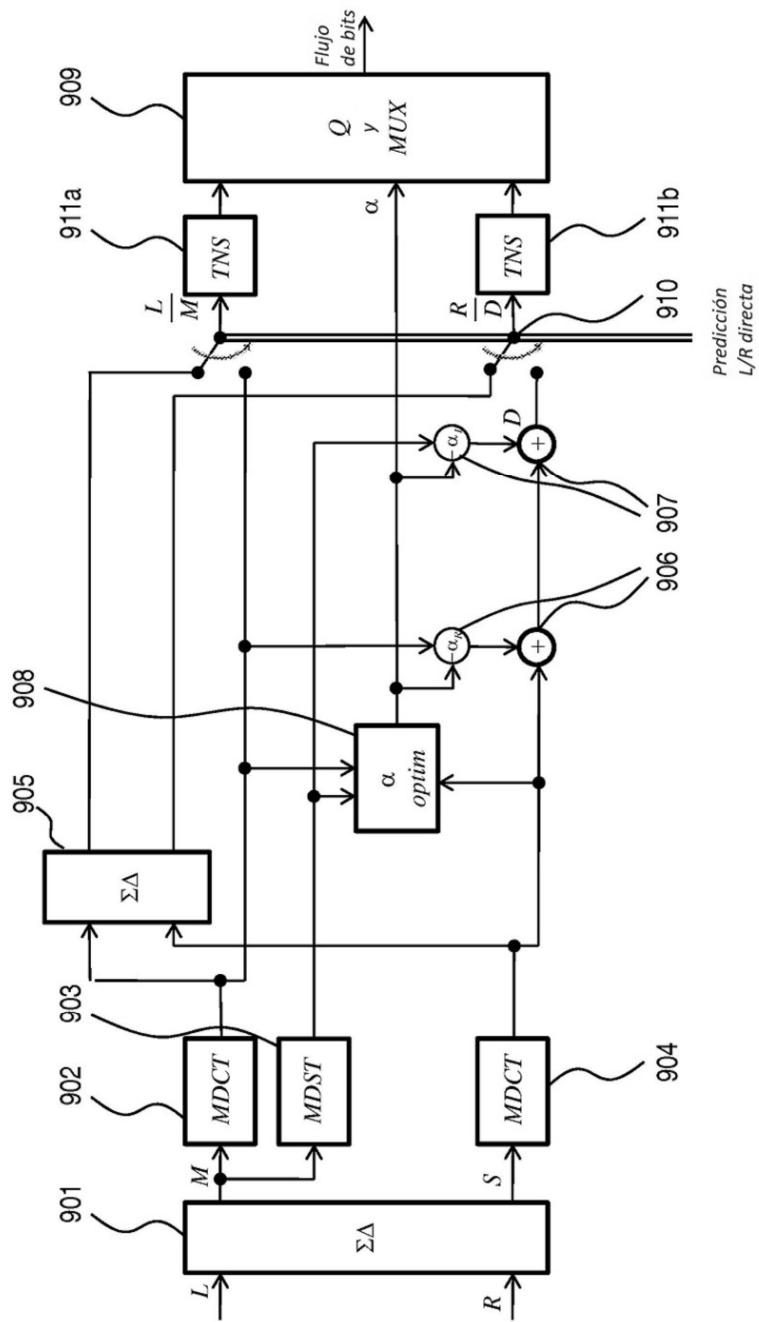


Fig. 9

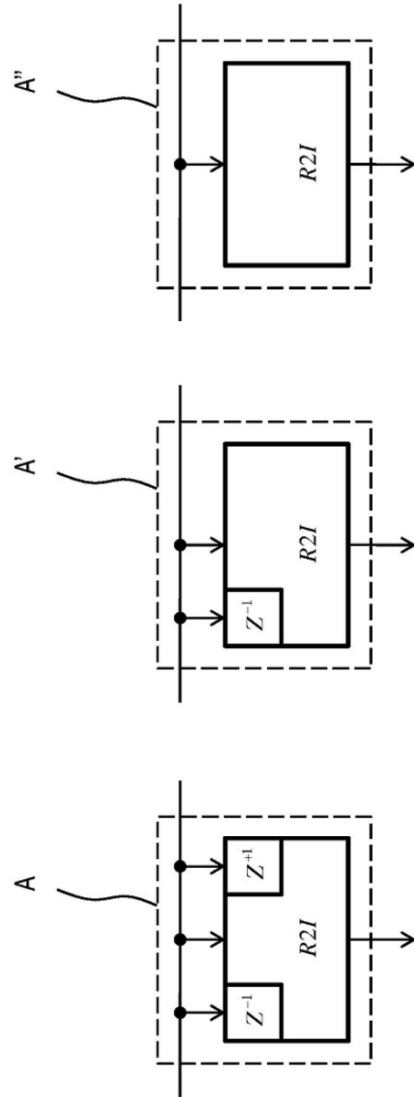


Fig. 10

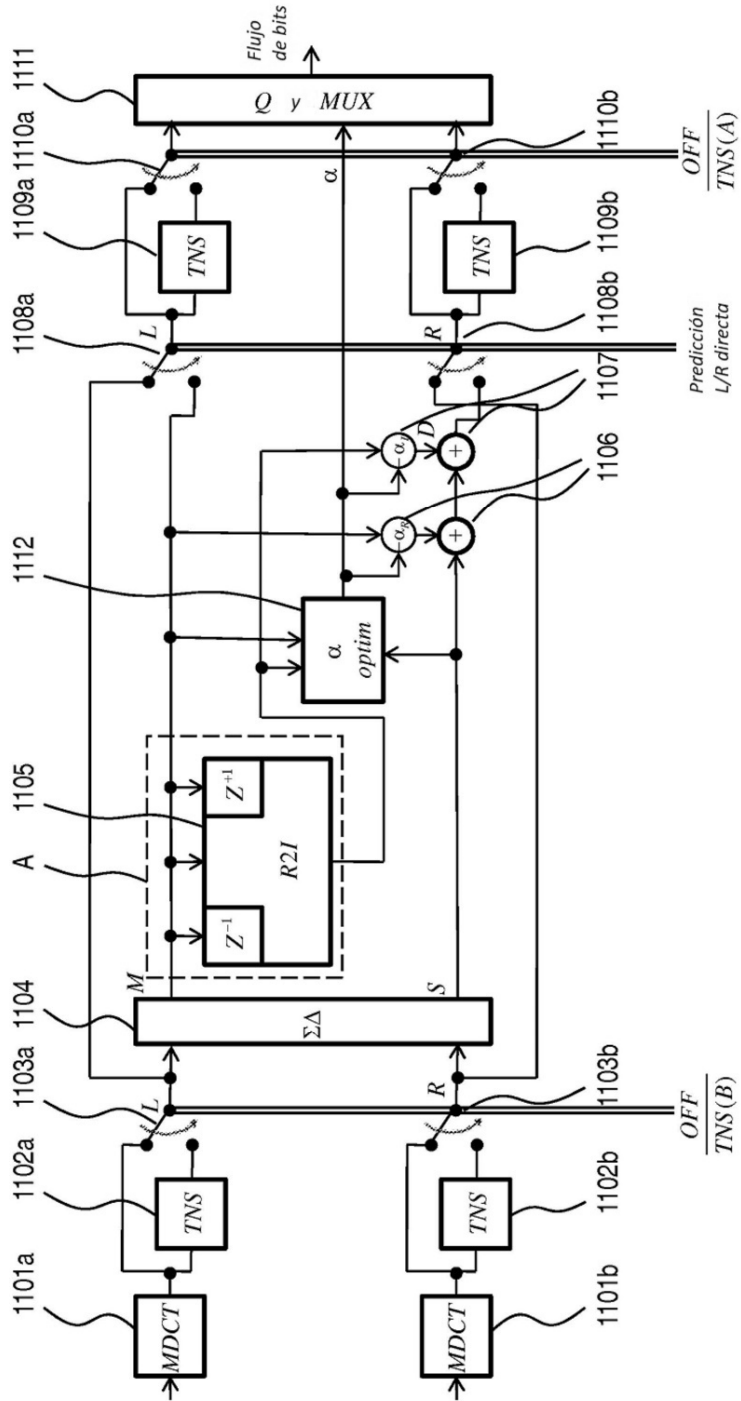
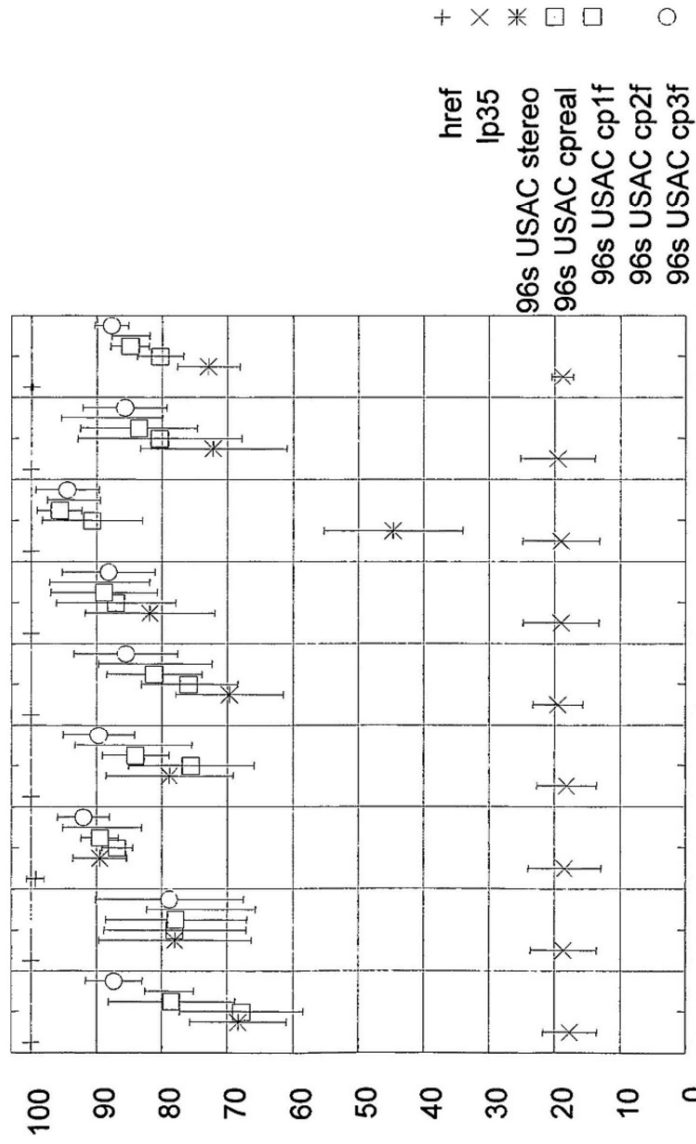


Fig. 11

MUSHRA, resultados absolutos, 6 casos



elem. 1 elem. 2 elem. 3 elem. 4 elem. 5 elem. 6 elem. 7 elem. 8 sign.

Fig. 12

HOJA SUSTITUTIVA (REGLA 6)

MUSHRA, resultados absolutos, 6 casos

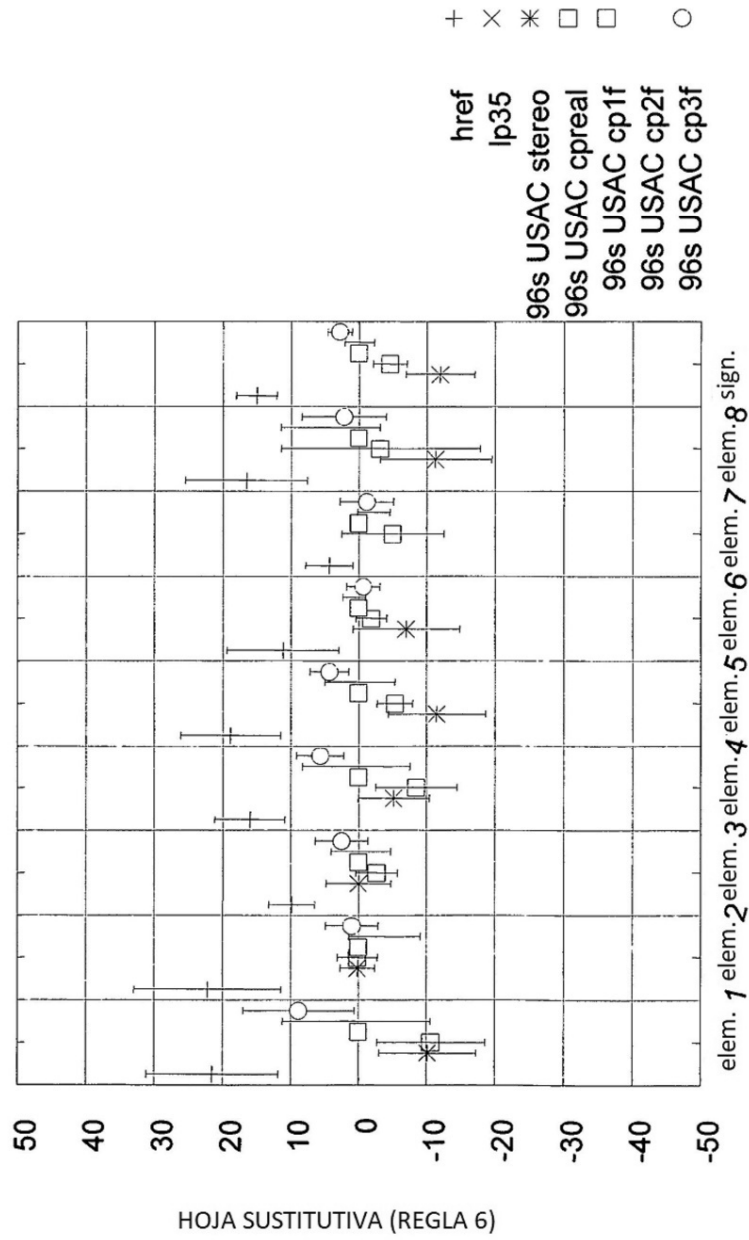


Fig. 13

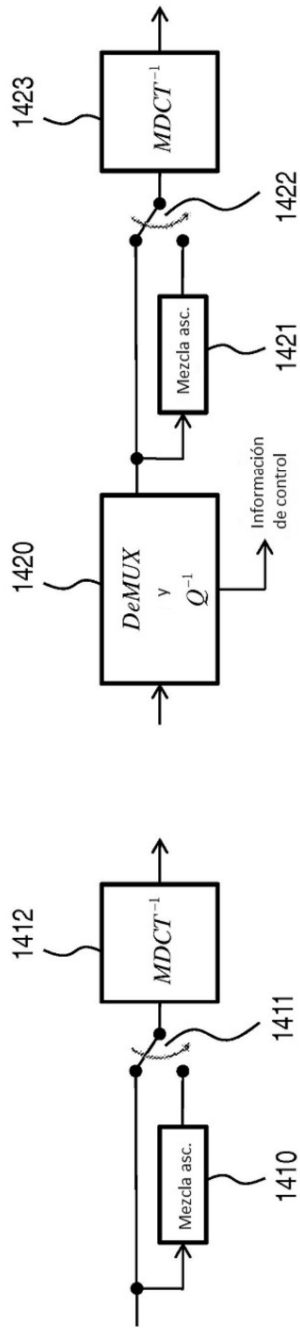


Fig. 14A

Fig. 14B

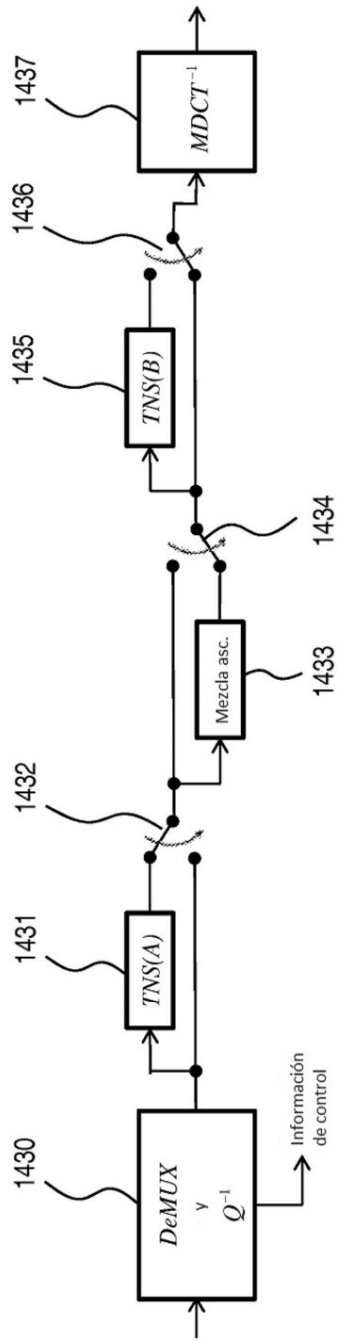


Fig. 14C

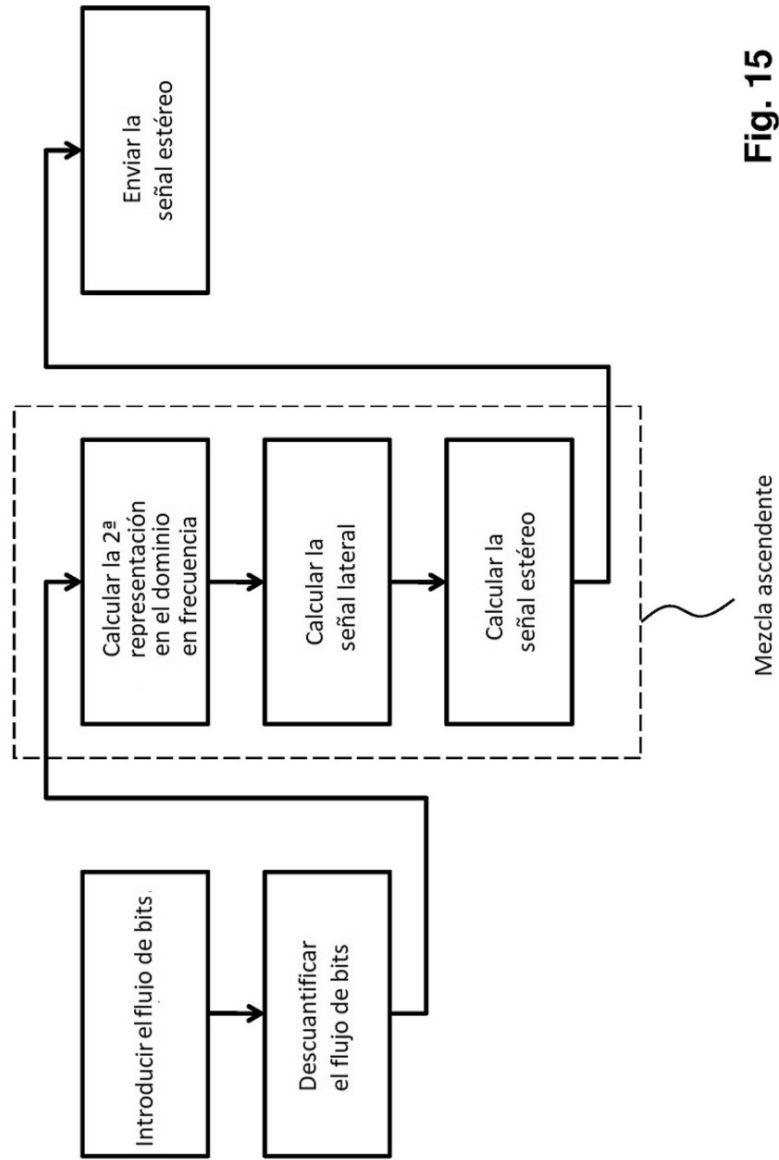


Fig. 15

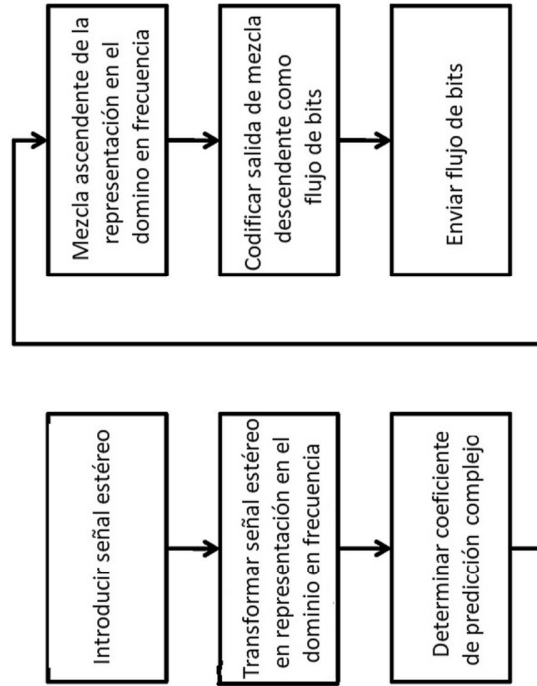


Fig. 16