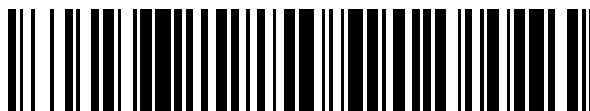


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 720 871**

51 Int. Cl.:

G10L 19/04 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.09.2007** **E 10173938 (1)**

97 Fecha y número de publicación de la concesión europea: **19.12.2018** **EP 2264699**

54 Título: **Dispositivo y método para procesamiento posterior de valores espectrales y codificador y decodificador para señales de audio**

30 Prioridad:

02.11.2006 DE 102006051673

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

25.07.2019

73 Titular/es:

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

72 Inventor/es:

**GEIGER, RALF y
EDLER, BERND**

74 Agente/Representante:

SALVÀ FERRER, Joan

ES 2 720 871 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

Dispositivo y método para procesamiento posterior de valores espectrales y codificador y decodificador para señales de audio

5

1. Campo de la invención:

[0001] La presente invención está relacionada con la codificación/decodificación de audio y, en particular, con los conceptos de codificador/decodificador escalable que tienen una capa base y una capa de extensión.

10

2. Descripción de la técnica relacionada:

[0002] Los codificadores/decodificadores de audio se conocen desde hace mucho tiempo. En particular, los codificadores/decodificadores de audio que funcionan según la norma ISO/IEC 11172-3, en los que esta norma también es conocida como el patrón de MP3, se llaman codificadores de transformación. Tal codificador de MP3 recibe una secuencia de muestras temporales como una señal de entrada que se somete a un ventanaje. El ventanaje da como resultado bloques secuenciales de muestras temporales que son convertidas a continuación en un bloque de representación espectral mediante el bloque. Según el patrón de MP3, se lleva a cabo una conversión con el llamado banco de filtros híbrido. La primera etapa del banco de filtros híbrido es un banco de filtros que tiene 32 canales con el fin de generar 32 señales de sub-banda. Los filtros de sub-banda de esta primera etapa comprenden bandas de paso solapadas, que es la razón por la cual esta filtración es propensa a la distorsión por plegado del espectro. La segunda etapa es una etapa MDCT (Transformada Discreta del Coseno Modificada, MDCT por sus siglas en inglés) para dividir las 32 señales de sub-banda en 576 valores espectrales. Los valores espectrales son cuantificados a continuación teniendo en cuenta el modelo psico-acústico y posteriormente codificados mediante Huffman con el fin de obtener finalmente una secuencia de bits e incluso un flujo de palabras de código de Huffman e información conexa necesaria para la decodificación.

[0003] En el lado del decodificador, las palabras del código de Huffman son calculadas de nuevo entonces en índices de cuantificación. Una nueva cuantificación da como resultado valores espectrales que son alimentados a continuación en un banco de filtros de síntesis híbrido que se implementa análogo al banco de filtros de análisis para obtener de nuevo bloques de muestras temporales de la señal de audio codificada y decodificada de nuevo. Todas las etapas en el lado del codificador y en el lado del decodificador se presentan en el patrón de MP3. En cuanto a la terminología, se observa esto en la referencia que sigue como "una cuantificación inversa". Aunque una cuantificación no sea reversible, cuando implica una pérdida de datos irrecuperable, a menudo se utiliza la expresión de cuantificación reversible, que debe indicar una nueva cuantificación antes presentada.

[0004] También es conocido en la técnica un algoritmo codificador/decodificador de audio llamado AAC (AAC, por sus siglas en inglés = Codificación Audio Avanzada). Tal codificador estandarizado en la norma internacional ISO/IEC 13818-7 funciona de nuevo sobre la base de muestras temporales de una señal de audio. Las muestras temporales de la señal de audio son sometidas de nuevo a un ventanaje con el fin de obtener bloques secuenciales de muestras temporales con ventanaje. A diferencia del codificador MP3 en el cual se usa un banco de filtros híbrido, en el codificador de AAC se lleva a cabo una transformación MDCT individual con el fin de obtener una secuencia de bloques de valores espectrales MDCT. Estos valores espectrales MDCT son cuantificados de nuevo entonces sobre la base de un modelo psico-acústico y los valores espectrales cuantificados son finalmente codificados mediante Huffman. En el lado del decodificador, el procesamiento es proporcional. Las palabras del código de Huffman son decodificadas y los índices de cuantificación o valores espectrales cuantificados, respectivamente, que se obtienen a partir de eso, son cuantificados entonces de nuevo para ser posteriormente o inversamente cuantificados, respectivamente, para obtener finalmente valores espectrales que puedan ser suministrados a un banco de filtros para síntesis MDCT con el fin de obtener finalmente muestras temporales codificadas/decodificadas de nuevo.

[0005] Ambos métodos funcionan con bloques solapados y funciones de ventana adaptables como se describe en la publicación de expertos "Codierung von Audiosignalen mit überlappender Transformation und adaptiven Fensterfunktionen", Bernd Edler, Frequenz, vol. 43, 1989, pp. 252-256.

55

[0006] En particular, cuando las áreas de transitorios se determinan en la señal de audio, se lleva a cabo una permutación de funciones de ventana larga a funciones de ventana corta con el fin de obtener una resolución de la frecuencia reducida a favor de una mejor resolución de tiempo. Una secuencia de ventanas cortas es introducida por una ventana de inicio y una secuencia de ventanas cortas es terminada por la ventana de finalización. Así, puede lograrse una transición sin intersticio entre funciones de ventana larga solapadas a funciones de ventana corta solapadas. Dependiendo de la implementación, el área de solapamiento con ventanas cortas es más pequeña que el área de solapamiento con ventanas largas, que es razonable en cuanto al hecho de que las porciones de señal transitorias se encuentran en la señal de audio aunque, sin embargo, esto no necesariamente tiene que ser el caso. Así, las secuencias de ventanas cortas como las secuencias de ventanas largas pueden ponerse en práctica con un solapamiento del 50 por ciento. En particular con ventanas cortas, sin embargo, para mejorar la codificación de

60

porciones de señal transitorias, puede ser seleccionada una anchura de solapamiento reducida como, por ejemplo, solo el 10 por ciento o incluso menos, en lugar del 50 por ciento.

[0007] Tanto en el patrón MP3 como en la norma AAC, existe el ventanaje con ventanas largas y cortas y las 5 ventanas de inicio o las ventanas de finalización, respectivamente, se ponen a escala de tal forma que en general siempre se puede mantener la misma trama del bloque. Para el patrón de MP3 esto significa, que para cada bloque largo se generan 576 valores espectrales y que tres bloques cortos corresponden a un bloque largo. Esto significa, que un bloque corto genera 192 valores espectrales. Con un solapamiento del 50 por ciento, se usa para el 10 ventanaje de una longitud de ventana de 1152 muestras temporales, cuando debido al principio de solapamiento y añadidura de un solapamiento del 50 por ciento, dos bloques de muestras temporales siempre dan como resultado un bloque de valores espectrales.

[0008] Tanto con codificadores MP3 como con codificadores de AAC, tiene lugar una compresión con pérdida de datos. Las pérdidas son introducidas por una cuantificación de los valores espectrales que se producen. Los 15 valores espectrales son cuantificados, en particular, de modo que las distorsiones introducidas por la cuantificación también referidas como la distorsión de cuantificación tienen una energía que está por debajo del umbral de enmascaramiento psico-acústico.

[0009] Cuanto más aproximada es la señal de audio que se cuantifica, es decir, cuanto mayor es el tamaño de la etapa de cuantificación, más elevada es la distorsión de cuantificación. Por otra parte, no obstante, para una 20 cuantificación más aproximada, un conjunto más pequeño de valores de salida de cuantificación debe considerarse, de modo que los valores cuantificados con menos aproximación pueden codificarse por entropía al usar menos bits. Esto significa que una cuantificación más aproximada da como resultado una compresión de datos más elevada, sin embargo simultáneamente da como resultado pérdidas de señal más elevadas.

25 **[0010]** Estas pérdidas de señal no son problemáticas si están por debajo del umbral de enmascaramiento. Incluso si el umbral de enmascaramiento psico-acústico solo es excedido ligeramente, no se podría obtener como resultado posiblemente a interferencias audibles para oyentes no cualificados. De todos modos, sin embargo, tiene lugar una pérdida de información que puede ser no deseada por ejemplo, debido a los artefactos que pueden ser 30 audibles en ciertas situaciones.

[0011] En particular, con conexiones de datos de banda ancha o cuando la velocidad de transferencia de datos no es el parámetro decisivo, respectivamente, o cuando tanto el ancho de banda como las redes de datos de banda estrecha están disponibles, puede ser conveniente no tener una presentación comprimida con pérdidas, sino 35 sin pérdidas o casi sin pérdidas, de una señal de audio.

[0012] Tal codificador escalable esquemáticamente ilustrado en la Figura 7 y un decodificador asociado esquemáticamente ilustrado en la Figura 8 es conocido a partir de la publicación de expertos "INTMDCT - A Link 40 Between Perceptual And Lossless Audio Coding", Ralf Geiger, Jürgen Herre, Jürgen Koller, Karlheinz Brandenburg, Int. Conference on Acoustics Speech and Signal Processing (ICASSP), 13 - 17 de mayo, 2002, Orlando, Florida. Una tecnología similar se describe en el documento de Patente Europea EP 1 495 464 B1. Los elementos 71, 72, 73, 74 ilustran un codificador de AAC con el fin de generar un flujo de bits codificado con pérdidas referido como "flujo de bits perceptivamente codificado" en la Figura 7. Este flujo de bits representa la capa base. En términos particulares, el bloque 71 en la Figura 7 designa el banco de filtros de análisis incluso el ventanaje con ventanas 45 largas y cortas según la norma de AAC. El bloque 73 representa la cuantificación/codificación según la norma de AAC y el bloque 74 representa la generación de flujo de bits de modo que el flujo de bits en el lado de salida no solo incluya palabras de código de Huffman de valores espectrales cuantificados sino también la información lateral necesaria, como por ejemplo factores de escala, etc., de modo que pueda llevarse a cabo la decodificación. La cuantificación con pérdidas en el bloque 73 es controlada aquí por el modelo psico-acústico designado como "el 50 modelo perceptual" 72 en la Figura 7.

[0013] Como ya se ha indicado, la señal de salida del bloque 74 es una capa escalable baja que requiere relativamente pocos bits y es, sin embargo, solo una representación con pérdidas de la señal de audio original y puede comprender artefactos de codificador. Los bloques 75, 76, 77, 78 representan los elementos adicionales que 55 son necesarios para generar un flujo de bits de extensión que sea sin pérdida o virtualmente sin pérdida, como se indica en la Figura 7. En particular, la señal de audio original se somete a una MDCT de enteros (IntMDCT) en la entrada 70, como se ilustra por el bloque 75. Adicionalmente, los valores espectrales cuantificados, generados por el bloque 73, en los cuales las pérdidas de codificador ya son introducidas, se someten a una cuantificación inversa y a un doblamiento subsecuente con el fin de obtener valores espectrales plegados. Aquellos son suministrados a un 60 formador de diferencias 77, formando una diferencia con función de valor espectral que es sometida a continuación a una codificación entrópica en el bloque 78 con el fin de generar un flujo de bits de potenciación sin pérdidas del esquema escalable en la Figura 7. Un espectro de valores diferenciales a la salida de bloque 77 representa así la distorsión introducida por la cuantificación psico-acústica en el bloque 73.

65 **[0014]** En el lado del decodificador el flujo de bits codificado con pérdidas o el flujo de bits perceptivamente

codificado son suministrados a un decodificador del flujo de bits 81. En el lado de salida, el bloque 81 proporciona una secuencia de bloques de valores espectrales cuantificados que son sometidos a continuación a una cuantificación inversa en un bloque 82. A la salida de bloque 82 de este modo los valores espectrales inversamente cuantificados se encuentran que ahora, en contraste con los valores a la entrada del bloque 82, ya no representan
 5 índices de cuantificación, pero que ahora son por así decir valores espectrales "correctos" que, sin embargo, son diferentes de los valores espectrales antes de la codificación en el bloque 73 de la Figura 7 debido a la cuantificación con pérdidas. Estos valores espectrales cuantificados son suministrados ahora a un banco de filtros de síntesis o una transformación de MDCT inversa (MDCT inversa), respectivamente, en el bloque 83 para obtener una señal de audio psico-acústicamente codificada y de nuevo decodificada (audio perceptual) que es diferente de la señal de
 10 audio original a la entrada 70 de la Figura 7 debido a los errores de codificación introducidos por el codificador de la Figura 7. Con el fin de obtener no solo una compresión con pérdidas sino incluso sin pérdidas, la señal de audio de bloque 82 es suministrada a un doblamiento en un bloque 84. En un circuito sumador 85, los valores espectrales inversamente cuantificados y redondeados son añadidos a los valores diferenciales que son generados por el formador de diferencias 77, donde en un bloque 86 una decodificación entrópica se lleva a cabo para decodificar las
 15 palabras de código de entropía contenidas en el flujo de bits de extensión que contiene la información sin pérdida o virtualmente sin pérdida, respectivamente.

[0015] A la salida de bloque 85, los valores espectrales de IntMDCT se encuentran así que son en el caso óptimo idéntico a los valores espectrales MDCT a la salida del bloque 75 del codificador de la Figura 7. El mismo es
 20 sometido a continuación a una MDCT de enteros inversa (IntMDCT inversa), para obtener una señal de audio sin pérdidas codificada o señal de audio prácticamente sin pérdidas (audio sin pérdidas) a la salida de bloque 87.

[0016] La MDCT de enteros (IntMDCT) es una aproximación de la MDCT, sin embargo, que genera valores de salida de número entero. Se deriva de la MDCT la utilización del esquema de elevación. Esto funciona
 25 particularmente cuando la MDCT se divide en las denominadas rotaciones Givens. A continuación, un algoritmo de dos etapas con rotaciones Givens y subsecuente resultado DCT-IV como en la MDCT de enteros en el lado de codificador y con un DCT-IV y un número descendente de rotaciones Givens en el lado decodificador. En el esquema de la Figura 7 y la Figura 8, así el espectro MDCT cuantificado generado en el codificador de AAC se utiliza para afirmar el espectro de MDCT de enteros. En términos generales, la MDCT de enteros es así un ejemplo
 30 para una transformación de número entero que genera valores espectrales de número entero y de nuevo muestras temporales de los valores espectrales de número entero, sin pérdidas introducidas por errores de redondeo. Otras transformaciones de número entero existen aparte de la MDCT de enteros.

[0017] El esquema escalable indicado en las Figuras 7 y 8 solo es lo suficientemente eficiente cuando las
 35 diferencias en la salida del formador de diferencias 77 son pequeñas. En el esquema ilustrado en la Figura 7 es el caso, cuando la MDCT y la MDCT de enteros son similares y cuando el IntMDCT en el bloque 75 es derivado de la MDCT en el bloque 71, respectivamente. Si no fuera el caso, el esquema ilustrado allí no sería adecuado cuando los valores diferenciales serían en muchos casos mayores que los valores de MDCT originales o incluso mayores que los valores de IntMDCT originales. A continuación, el esquema escalable en la Figura 7 ha perdido su valor cuando
 40 la capa escalable de extensión enviada por el bloque 78 tiene una redundancia elevada en cuanto a la capa escalable base.

[0018] Los esquemas de escalabilidad son siempre óptimos cuando la capa base comprende varios bits y cuando la capa de extensión comprende varios bits y cuando la suma de los bits en la capa base y en la capa de
 45 extensión es igual a varios bits que se obtendrían si la capa base ya fuera una codificación sin pérdidas. Este caso óptimo no se logra nunca en esquemas de escalabilidad prácticos, en cuanto a la capa de extensión, se requieren los bits de señalización adicionales. Sin embargo, se trata de lograr este grado óptimo en la medida de lo posible. Cuando las transformaciones en bloques 71 y 75 son relativamente similares en la Figura 7, el concepto ilustrado en la Figura 7 está cerca del grado óptimo.
 50

[0019] Sin embargo, este concepto de escalabilidad simple puede no obstante, no solo así ser aplicado a la
 señal de salida de un codificador MP3, cuando el codificador MP3, como se ilustra, no comprende ningún banco de filtros MDCT puro como un banco de filtros, pero el banco de filtros híbrido que tiene una primera etapa de banco de
 55 filtros para generar diferentes señales de sub-banda y MDCT descendente para la descomposición adicional de las señales de sub-banda, donde además, como se indica también en el patrón de MP3, se pone en práctica una etapa de cancelación de distorsión por plegado del espectro adicional del banco de filtros híbrido. Cuando la MDCT de enteros en el bloque 75 tiene pocas semejanzas con el banco de filtros híbrido según el patrón de MP3, una aplicación directa del concepto mostrado en la Figura 7 a una señal de salida MP3 daría como resultado valores diferenciales muy elevados en la salida del formador de diferencias 77, que da como resultado un concepto
 60 de escalabilidad muy ineficaz, cuando la capa de extensión requiere demasiados bits con el fin de codificar razonablemente los valores diferenciales a la salida del formador de diferencias 77.

[0020] Una posibilidad para generar el flujo de bits de extensión para una señal de salida MP3 se ilustra en la
 Figura 9 para el codificador y en la Figura 10 para el decodificador. Un codificador MP3 90 codifica una señal de
 65 audio y proporciona una capa base 91 en el lado de salida. La señal de audio codificada del MP3 es suministrada

entonces a un decodificador de MP3 92 que proporciona una señal de audio con pérdidas en el intervalo de tiempo. Esta señal es suministrada entonces a un bloque IntMDCT que se puede establecer en principio como el bloque 75 en la Figura 7, donde este bloque 75 proporciona a continuación valores espectrales IntMDCT en el lado de salida que son suministrados a un formador de diferencias 77 que también incluye valores espectrales de IntMDCT como valores de entrada adicionales, que no fueron, sin embargo, generados por la señal de audio decodificada del MP3, pero por la señal de audio original que fue suministrada al codificador MP3 90.

[0021] En el lado del decodificador, la capa base es suministrada de nuevo a un decodificador de MP3 92 para proporcionar una señal de audio decodificada con pérdidas a una salida 100 que correspondería a la señal a la salida del bloque 83 de la Figura 8. Esta señal tendría que someterse entonces a una MDCT de enteros 75 para ser codificada a continuación conjuntamente con la capa de extensión 93 que se generó a la salida del formador de diferencias 77. El espectro sin pérdidas se encontraría entonces en una salida 101 del circuito sumador 102 y solo tendría que ser convertido por medio de IntMDCT inverso 103 en el intervalo de tiempo con el fin de obtener una señal de audio sin pérdidas decodificada que correspondería "al audio sin pérdidas" al comienzo del bloque 87 de la Figura 8.

[0022] El concepto ilustrado en la Figura 9 y en la Figura 10, que proporciona una capa de extensión relativamente eficazmente codificada justo como los conceptos ilustrados en las Figuras 7 y 8, es costoso tanto en el lado del codificador (Figura 9) como en el lado del decodificador (Figura 10), respectivamente. En contraste con el concepto en la Figura 7, se requiere un decodificador de MP3 92 completo e IntMDCT adicional 75.

[0023] Otra desventaja en este esquema es, que un decodificador de MP3 con exactitud de bits tendría que definirse. No se pretende esto, sin embargo, cuando el patrón de MP3 no representa una especificación con exactitud de bits, pero solo tiene que ser realizado dentro del alcance "de una conformidad" por un decodificador.

[0024] En el lado del decodificador, se requiere adicionalmente una etapa 75 IntMDCT adicional completa. Ambos elementos adicionales causan sobrecarga computacional y son desventajosos particularmente para el uso en dispositivos móviles ambos con respecto a consumo de circuito integrado y también con respecto a consumo actual y también con respecto al retraso asociado.

[0025] En resumen, las ventajas del concepto ilustrado en la Figura 7 y la Figura 8 son, que en comparación con métodos de dominio de tiempo no se requiere una decodificación completa de la señal codificada adaptada para audio, y que se obtiene una codificación eficiente mediante una representación del error de cuantificación en el rango de frecuencia que se va a codificar adicionalmente. Así, el método estandarizado por la Codificación escalable sin pérdidas de ISO/IEC MPEG-4 (SLS) usa esta estrategia, como se describe en R. Geiger, R. Yu, J. Herre, S. Rahardja, S. Kim, X. Lin, M. Schmidt, "ISO/IEC MPEG-4 High-Definition Scalable Advanced Audio Coding", 120va AES meeting, 20 - 23 de mayo de 2006, París, Francia, Pre-impresión 6791. Así, una extensión compatible, sin pérdidas atrasada de métodos de codificación de audio, por ejemplo MPEG-2/4 AAC, se obtiene que utiliza la MDCT como un banco de filtros.

[0026] Esta estrategia no puede aplicarse, sin embargo, directamente al método extensamente usado para MPEG-1/2 de Capa 3 (MP3), cuando el banco de filtros híbrido usado en este método, en contraste con la MDCT, no es compatible con el IntMDCT u otra transformación de número entero. Así, una formación diferencial entre los valores espectrales decodificados y los valores de IntMDCT correspondientes en general no da como resultado pequeños valores diferenciales y así no a una codificación eficiente de los valores diferenciales. El núcleo del problema aquí es los desplazamientos de tiempo entre las funciones de modulación correspondientes del IntMDCT y el banco de filtros híbrido MP3. Éstos dan como resultado desplazamientos de fase que en casos desfavorables incluso conducen al hecho de que los valores diferenciales comprenden valores más elevados que los valores de IntMDCT. También una aplicación de los principios que son la base del IntMDCT, como por ejemplo el esquema de elevación, al banco de filtros híbrido de MP3 es problemática, como en cuanto a su estrategia básica - en contraste con MDCT - el banco de filtros híbrido es un banco de filtros que no proporciona ninguna reconstrucción perfecta.

[0027] El documento WO 2004/013839 A1 describe una disposición de banco de filtros para la generación de una representación espectral compleja de una señal discreta en el tiempo. La disposición del banco de filtros comprende un dispositivo para la generación de una serie temporal de coeficientes espectrales reales. La serie se introduce en un dispositivo para el procesamiento posterior de las series temporales de bloques de coeficientes espectrales reales con el fin de proporcionar una representación espectral compleja aproximada con bloques en serie. Un coeficiente espectral aproximado complejo puede ser representado por un primer coeficiente espectral parcial y un segundo coeficiente espectral parcial, en el que al menos el primer o el segundo coeficiente espectral parcial está determinado por una combinación de al menos dos coeficientes espectrales reales.

[0028] La publicación técnica "Codificación de audio sin pérdida mediante el uso de IntMDCT y formación de error de redondeo", Yokotani, et al, IEEE Transactions on Audio Speech, and Language Processing, (Transacciones IEEE sobre el discurso de audio y procesamiento de idioma), vol. 14, N. ° 6, noviembre de 2006, Págs. 2201-2211 describe una construcción de una MDCT entera como una aproximación entera de la MDCT mediante el uso del

esquema de elevación. Además, se deriva la forma de error de redondeo del IntMDCT. Mediante el uso de una extensión multidimensional de una factorización triangular de la matriz de rotación de Givens de un esquema de elevación, los valores del dominio de tiempo se procesan para transformar esos valores en un dominio IntMDCT.

5 RESUMEN DE LA INVENCION

[0029] Es el objetivo de la presente invención proporcionar un concepto eficiente para procesar datos de audio y particularmente para codificar o decodificar datos de audio.

10 **[0030]** Este objetivo se logra por un dispositivo para el procesamiento posterior de valores espectrales según la reivindicación 1, un método de valores espectrales de procesamiento posterior según la reivindicación 14 o un programa informático según la reivindicación 16.

[0031] La presente invención se basa en el descubrimiento, de que valores espectrales, por ejemplo los que
15 representan la capa base de un esquema escalable, es decir, por ejemplo valores espectrales de MP3, se someten al procesamiento posterior, para obtener valores a partir de ello que son compatibles con valores correspondientes obtenidos según un algoritmo de transformación alternativo. Según la invención, así tal procesamiento posterior se lleva a cabo usando adiciones ponderadas de valores espectrales de modo que el resultado del procesamiento posterior sea tan similar como sea posible a un resultado que se obtiene cuando la misma señal de audio no es
20 convertida en una representación espectral usando el primer algoritmo de transformación, pero usando el segundo algoritmo de transformación, que es, en formas de realización preferidas de la presente invención, un algoritmo de transformación de número entero.

[0032] Se ha descubierto, incluso con un primer algoritmo de transformación y un segundo algoritmo de
25 transformación muy incompatibles, que por una adición ponderada de ciertos valores espectrales del primer algoritmo de transformación, se logra una compatibilidad de los valores con procesamiento posterior con los resultados de la segunda transformación que es una capa de extensión eficiente que puede formarse con valores diferenciales, sin la costosa y así desventajosa codificación y decodificación del concepto en la Figura 9 y la Figura 10. En particular, la adición ponderada se lleva a cabo de modo que un valor espectral con procesamiento posterior
30 se genere de una adición ponderada de un valor espectral y un valor espectral adyacente a la salida del primer algoritmo de transformación, donde se usan preferentemente ambos valores espectrales de intervalos de frecuencias adyacentes y también los valores espectrales de bloques temporales adyacentes o períodos de tiempo, respectivamente. Por la adición ponderada de valores espectrales adyacentes se considera que en el primer algoritmo de transformación, los filtros adyacentes de un banco de filtros se solapan, como es el caso prácticamente
35 con todos los bancos de filtros. Por el uso de valores espectrales temporalmente adyacentes, es decir por la adición ponderada de valores espectrales (por ejemplo del mismo o solo una frecuencia ligeramente diferente) de dos bloques subsecuentes de valores espectrales de la primera transformación se considera adicionalmente que por lo común los algoritmos de transformación se usan en los que se utiliza un solapamiento de bloques.

40 **[0033]** Preferentemente, los factores de ponderación son permanentemente programados tanto en el lado del codificador como en el lado del decodificador, de modo que ningún bit adicional sea necesario para transferir factores de ponderación. En cambio, los factores de ponderación se configuran una vez y por ejemplo se almacenan como una tabla o firmemente implementados en hardware, cuando los factores de ponderación no son dependientes de la señal, sino solo dependientes del primer algoritmo de transformación y del segundo algoritmo de
45 transformación. En particular, se prefiere configurar los factores de ponderación de modo que una respuesta a impulsos de construcción de primer algoritmo de transformación y procesamiento posterior sea igual a una respuesta a impulsos del segundo algoritmo de transformación. A este respecto, una optimización de los factores de ponderación puede emplearse con métodos de optimización conocidos manuales o automatizados, por ejemplo, usando ciertas señales de prueba representativas o, como se indica, usando directamente las respuestas a impulsos
50 de los filtros resultantes.

[0034] El mismo dispositivo de procesamiento posterior puede usarse tanto en el lado del codificador como también en el lado del decodificador con el fin de adaptar valores espectrales de hecho incompatibles del primer algoritmo de transformación a valores espectrales del segundo algoritmo de transformación, de modo que ambos
55 bloques de valores espectrales puedan someterse a una formación diferencial con el fin de proporcionar finalmente una capa de extensión para una señal de audio que es por ejemplo una señal codificada de MP3 en la capa base y comprende la extensión sin pérdidas como la capa de extensión.

[0035] Cabe destacar, que la presente invención no está limitada a la combinación de MP3 y MDCT de
60 enteros, sino que la presente invención tiene uso en todas partes, cuando los valores espectrales de algoritmos de transformación de hecho incompatibles, deben ser procesados conjuntamente, por ejemplo para una formación diferencial, una adición o cualquier otro funcionamiento de combinación en un codificador de audio o un decodificador de audio. El uso preferido del dispositivo de procesamiento posterior inventivo debe proporcionar, sin embargo, una capa de extensión a una capa base en la cual una señal de audio es codificada con cierta calidad,
65 donde la capa de extensión, conjuntamente con la capa base, sirve para lograr una calidad más elevada al

decodificar, donde esta decodificación de calidad más elevada ya es preferentemente una decodificación sin pérdidas, pero también puede ser, sin embargo, una decodificación prácticamente sin pérdidas, mientras que la calidad de la señal de audio decodificada se mejora mediante el uso de la capa de extensión en comparación con la decodificación mediante el uso únicamente de la capa base.

5

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0036] A continuación, las formas de realización preferidas de la presente invención son explicadas más detalladamente en referencia a los dibujos adjuntos, en los cuales:

10

La Figura 1 muestra un dispositivo inventivo para el procesamiento posterior de valores espectrales;

La Figura 2 muestra un lado del codificador de un concepto de codificador inventivo;

La Figura 3 muestra un lado del decodificador de un concepto de decodificador inventivo;

15 La Figura 4 muestra una ilustración detallada de una forma de realización preferida del procesamiento posterior inventivo y formación diferencial para bloques largos;

La Figura 5a muestra una forma de realización preferida del dispositivo de procesamiento posterior inventivo para bloques cortos según una primera variante;

La Figura 5b muestra una ilustración esquemática de bloques de valores que pertenecen conjuntamente al concepto mostrado en la Figura 5a;

20 La Figura 5c muestra una secuencia de ventanas para la variante mostrada en la Figura 5a;

La Figura 6a muestra una forma de realización preferida del dispositivo de procesamiento posterior inventivo y formación diferencial para bloques cortos según una segunda variante de la presente invención;

La Figura 6b muestra una ilustración de valores diversos para la variante ilustrada en la Figura 6a;

La Figura 6c muestra una trama del bloque para la variante ilustrada en la Figura 6a;

25 La Figura 7 muestra una ilustración de codificador anterior para generar un flujo de datos a escala;

La Figura 8 muestra una ilustración del decodificador anterior para procesar un flujo de datos a escala;

La Figura 9 muestra una variante de codificación ineficaz; y

La Figura 10 muestra una variante del decodificador ineficaz.

30 DESCRIPCIÓN DE LAS FORMAS DE REALIZACIÓN PREFERIDAS

[0037] La Figura 1 muestra un dispositivo inventivo para el procesamiento posterior de valores espectrales que son preferentemente una representación con pérdidas de una señal de audio, en la que los valores espectrales tienen un primer algoritmo de transformación subyacente para convertir la señal de audio en una representación

35 spectral independiente del hecho de si tienen pérdidas o no tienen pérdidas. El dispositivo inventivo ilustrado en la Figura 1 o el método ilustrado también esquemáticamente en la Figura 1, respectivamente, se distingue - en cuanto al dispositivo - por un medio 12 para proporcionar una secuencia de bloques de valores espectrales que representan una secuencia de bloques de muestras de la señal de audio. En una forma de realización preferida de la presente invención que se ilustrará más adelante, la secuencia de bloques proporcionada por el medio 12 es una secuencia

40 de bloques generada por un banco de filtros MP3. La secuencia de bloques de valores espectrales es suministrada a un combinador inventivo 13, donde el combinador se pone en práctica para llevar a cabo una adición ponderada de valores espectrales de la secuencia de bloques de valores espectrales para obtener, en el lado de salida, una secuencia de bloques de valores espectrales con procesamiento posterior, como se ilustra por la salida 14. En particular, el combinador 13 se pone en práctica para usar, para el cálculo de un valor spectral con procesamiento

45 posterior para una banda de frecuencia y un período de tiempo, un valor spectral de la secuencia de bloques para la banda de frecuencia y el período de tiempo y un valor spectral para una banda de frecuencia adyacente y/o un período de tiempo adyacente. Adicionalmente, el combinador se pone en práctica para usar tales factores de ponderación para ponderar los valores espectrales usados, de modo que los valores espectrales con procesamiento posterior son una aproximación a valores espectrales obtenidos por un segundo algoritmo de transformación para

50 convertir la señal de audio en una representación spectral, donde, sin embargo, el segundo algoritmo de transformación es diferente del primer algoritmo de transformación.

[0038] Esto se ilustra por la ilustración esquemática en la Figura 1 en la parte inferior. Un primer algoritmo de transformación es representado por un número de referencia 16. El procesamiento posterior, cuando se lleva a cabo por el combinador, es representado por el número de referencia 13, y el segundo algoritmo de transformación es representado por un número de referencia 17. De los bloques 16, 13 y 17, los bloques 16 y 17 se fijan y son por lo común obligatorios debido a condiciones externas. Solo los factores de ponderación de los medios de procesamiento posterior 13 o el combinador 13, respectivamente, representado por el número de referencia 18, pueden ser configurados por el usuario. En esta conexión, esto no es dependiente de la señal, pero depende del

60 primer algoritmo de transformación y el segundo algoritmo de transformación, sin embargo. Por los factores de ponderación 18 se puede configurar adicionalmente, cuántos valores espectrales adyacentes en cuanto a frecuencia o valores espectrales adyacentes en tiempo son combinados entre sí. Si un factor de ponderación, cuando será explicado en cuanto a las Figuras 4 a 6, se configura en 0, el valor spectral asociado con este factor de ponderación no se considera en la combinación.

65

[0039] En formas de realización preferidas de la presente invención, para cada valor espectral se proporciona un conjunto de factores de ponderación. Así, resulta una cantidad considerable de factores de ponderación. Esto no es problemático, sin embargo, dado que los factores de ponderación no tienen que ser transferidos, sino solo tienen que ser permanentemente programados al lado del codificador y el lado del decodificador. Si el codificador y el decodificador acordaron así el mismo conjunto de factores de ponderación para cada valor espectral y, de ser aplicable, durante cada período de tiempo, o, como se ilustrará en lo sucesivo, para cada sub-bloque o pedido de la posición, respectivamente, no es necesario utilizar ninguna señalización para la presente invención, de modo que el concepto inventivo logre una reducción sustancial de la velocidad de transferencia de datos en la capa de extensión sin cualquier señalización de la información adicional, sin cualquier pérdida de calidad adjunta.

10

[0040] La presente invención proporciona de este modo una compensación de los desplazamientos de fase entre valores de frecuencia, a medida que se obtienen por el primer algoritmo de transformación, y valores de frecuencia, a medida que se obtienen por el segundo algoritmo de transformación, donde esta compensación de los desplazamientos de fase puede presentarse vía una representación espectral compleja. Para este fin, el concepto descrito en el documento DE 10234130 es incluido por motivos de la claridad, en el cual para calcular partes imaginarias de verdaderos valores de salida de banco de filtros, se obtienen las combinaciones lineales de valores espectrales temporalmente y espectralmente adyacentes. Si este procedimiento se utilizara para valores espectrales MP3 decodificados, se obtendría una representación espectral de valores complejos. Cada uno de los valores espectrales complejos resultantes puede ser modificado ahora en su posición de fase por una multiplicación por un factor de corrección de valores complejos de modo que, según la presente invención, esté tan cerca del segundo algoritmo de transformación como sea posible, es decir, preferentemente, el valor de IntMDCT correspondiente, y es así adecuado para una formación diferencial. Adicionalmente, según la invención, también se lleva a cabo una corrección de amplitud posiblemente requerida. Según la invención, estas etapas para la formación de la representación espectral de valores complejos y la fase o corrección de suma, respectivamente, son resumidas de tal forma que por la combinación lineal de valores espectrales sobre la base del primer algoritmo de transformación y sus vecinos temporales y espectrales se forme un nuevo valor espectral que minimiza la diferencia del valor de IntMDCT correspondiente. Según la invención, en contraste con el documento DE 10234130, un procesamiento posterior de valores de salida de banco de filtros no es llevado a cabo usando factores de ponderación con el fin de obtener partes verdaderas e imaginarias. En cambio, según la invención un procesamiento posterior se lleva a cabo usando tales factores de ponderación que, como se ilustra en la Figura 1 en la parte inferior, una combinación del primer algoritmo de transformación 16 y el procesamiento posterior 13 se configuran por los factores de ponderación de modo que el resultado corresponda a un segundo algoritmo de transformación en la medida de lo posible.

[0041] La Figura 2 y la Figura 3 muestran un campo preferido del uso del concepto inventivo ilustrado en la Figura 1 ambos en el lado del codificador (Figura 2) y también en el lado del decodificador (Figura 3) de un codificador escalable. Un flujo de bits del MP3 20 o - generalmente - un flujo de bits, respectivamente, como se puede obtener por un primer algoritmo de transformación, es alimentado a un bloque 21 con el fin de generar los valores espectrales del flujo de bits que son por ejemplo valores espectrales de MP3. La decodificación de los valores espectrales en el bloque 21 incluirá así por lo general una decodificación entrópica y una cuantificación inversa.

[0042] A continuación, en el bloque 10, se lleva a cabo un cálculo de valores de aproximación, donde el cálculo de valores de aproximación o de bloques de valores espectrales con procesamiento posterior, respectivamente, se lleva a cabo como se ilustra en la Figura 1. En ese momento, se lleva a cabo una formación diferencial en un bloque 22, mediante el uso de valores espectrales de IntMDCT, cuando se obtienen por una conversión IntMDCT en un bloque 23. El bloque 23 obtiene así una señal de audio como una señal de entrada a partir de la cual se obtiene el flujo de bits del MP3, como es alimentado en la entrada 20, mediante codificación. Preferentemente, los espectros diferenciales según son obtenidos por el bloque 22 se someten a una codificación sin pérdidas 24 que incluye por ejemplo una codificación de delta, una codificación Huffman, una codificación aritmética o cualquier otra codificación entrópica por la cual la velocidad de transferencia de datos es reducida, ninguna pérdida es introducida en una señal, no obstante.

[0043] En el lado del decodificador, el flujo de bits del MP3 20, como también se alimentó a la entrada 20 de la Figura 2, es sometido de nuevo a una decodificación de los valores espectrales por un bloque 21, que puede corresponder al bloque 21 de la Figura 2. En ese momento, los valores espectrales MP3 obtenidos en la salida del bloque 21 son procesados de nuevo según la Figura 1 o el bloque 10. En el lado del decodificador, sin embargo, los bloques de valores espectrales con procesamiento posterior, cuando son enviados por el bloque 10, son suministrados a una etapa de adición 30, que obtiene valores diferenciales IntMDCT a su otra entrada, cuando se obtienen por una decodificación sin pérdidas 31 del flujo de bits de extensión sin pérdidas que fue emitida por el bloque 24 en la Figura 2. Mediante la adición de los valores diferenciales IntMDCT enviados por el bloque 31 y los valores espectrales procesados enviados por el bloque 10, entonces, en una salida 32 de la etapa de adición 30 se obtienen bloques de valores espectrales IntMDCT que son una representación sin pérdidas de la señal de audio original, es decir, de la señal de audio que fue introducida en el bloque 23 de la Figura 2. La señal de salida de audio sin pérdidas es generada a continuación por un bloque 33 que lleva a cabo un IntMDCT inverso con el fin de obtener una señal de salida de audio sin pérdida o virtualmente sin pérdida. En términos generales, la señal de salida de

65

audio a la salida de bloque 33 tiene una mejor calidad que la señal de audio que se obtendría si la señal de salida del bloque 21 fuera procesada con un banco de filtros híbrido de síntesis MP3. Dependiendo de la implementación, la señal de salida de audio en la salida 33 puede ser así una reproducción idéntica de la señal de audio que fue introducida en el bloque 23 de la Figura 2, o una representación de esta señal de audio, que no es idéntica, es decir
 5 no completamente sin pérdidas, que tiene, sin embargo, ya una mejor calidad que una señal de audio codificada de MP3 normal.

[0044] A este punto cabe destacar, que como un primer algoritmo de transformación se prefiere el algoritmo de transformación MP3 con su banco de filtros híbrido, y que como un segundo algoritmo de transformación se
 10 prefiere el algoritmo IntMDCT como un algoritmo de transformación de número entero. La presente invención es ventajosa ya en todas partes, sin embargo, donde dos algoritmos de transformación son diferentes entre sí, donde ambos algoritmos de transformación no necesariamente tienen que ser algoritmos de transformación de número entero dentro del alcance de la transformación IntMDCT, pero también pueden ser algoritmos de transformación normales que son, dentro del alcance de una MDCT, no necesariamente una transformación de número entero
 15 reversible. Según la invención se prefiere, sin embargo, que el primer algoritmo de transformación sea un algoritmo de transformación de número no-entero y que el segundo algoritmo de transformación sea un algoritmo de transformación de número entero, donde el procesamiento posterior inventivo es particularmente ventajoso cuando el primer algoritmo de transformación proporciona espectros que, comparado con los espectros proporcionados por el segundo algoritmo de transformación, cambiados y/o desplazados en fase con respecto a sus cantidades. En
 20 particular, cuando el primer algoritmo de transformación no reconstruye incluso perfectamente, el procesamiento posterior simple inventivo por una combinación lineal es especialmente ventajoso y puede utilizarse eficazmente.

[0045] La Figura 4 muestra una forma de realización preferida del combinador 13 dentro de un codificador. La forma de realización dentro de un decodificador es idéntica, sin embargo, si el circuito sumador 22 no lleva a cabo,
 25 como en la Figura 4, una formación diferencial, como se ilustra por el signo menos por encima del circuito sumador 22, sino cuando un funcionamiento de adición se lleva a cabo, como se ilustra en el bloque 30 de la Figura 3. En cada caso los valores que son alimentados en una entrada 40 son valores a medida que se obtienen por el segundo algoritmo de transformación 23 de la Figura 2 para la forma de realización de codificador o a medida que se obtienen por el bloque 31 de la Figura 3 en la forma de realización del decodificador.

[0046] En una forma de realización preferida de la presente invención, el combinador incluye tres secciones 41, 42, 43. Cada sección incluye tres multiplicadores 42a, 42b, 42c, en los que cada multiplicador se asocia con un valor espectral con un índice de frecuencias $k-1$, k o $k+1$. Así, el multiplicador 42a se asocia con el índice de frecuencias $k-1$. El multiplicador 42b se asocia con el índice de frecuencias k y el multiplicador 42c se asocia con el
 35 índice de frecuencias $k+1$.

[0047] Cada rama sirve así para ponderar valores espectrales de un bloque actual con el índice de bloques $n+1$, n o $n-1$, respectivamente, con el fin de obtener valores espectrales ponderados para el bloque actual.

[0048] Así, la segunda sección 42 sirve para ponderar valores espectrales de un bloque temporalmente anterior o bloque temporalmente posterior. En cuanto a la sección 41, la sección 42 sirve para ponderar valores espectrales del bloque n temporalmente después del bloque $n+1$, y la sección 43 sirve para ponderar el bloque $n-1$ después del bloque n . Con el fin de indicar esto, los elementos de retraso 44 se indican en la Figura 4. Por razones de claridad, solo un elemento de retraso " z^{-1} " es designado por el número de referencia 44.
 45

[0049] En particular, cada multiplicador se proporciona con un factor de ponderación dependiente del índice espectral $c_0(k)$ a $c_8(k)$. Así, en la forma de realización preferida de la presente invención, resultan nueve valores
 50 espectrales ponderados, de los cuales un valor espectral con procesamiento posterior \hat{y} se calcula para el índice de frecuencia k y el bloque temporal n . Estos nueve valores espectrales ponderados son resumidos en un bloque 45.

[0050] El valor espectral con procesamiento posterior para el índice de frecuencia k y el índice de tiempo n se calcula así mediante la adición de valores espectrales posiblemente ponderados de forma diferente del bloque temporalmente anterior ($n-1$) y el bloque temporalmente posterior ($n+1$) y mediante el uso respectivamente de los valores espectrales adyacentes hacia arriba ($k+1$) y hacia abajo ($k-1$). Las implementaciones más simples solo
 55 pueden ser, sin embargo, que un valor espectral para el índice de frecuencias k es combinado únicamente con un valor espectral adyacente $k+1$ o $k-1$ del mismo bloque, donde este valor espectral que es combinado con el valor espectral del índice de frecuencias k , no necesariamente tiene que ser directamente adyacente, sino que también puede ser un valor espectral diferente del bloque. Debido al solapamiento común de bandas adyacentes, se prefiere, sin embargo, llevar a cabo una combinación con el valor espectral directamente adyacente a la parte superior y/o a
 60 la parte inferior.

[0051] Adicionalmente, alternativamente o además, cada valor espectral con un valor espectral para una duración de tiempo diferente, es decir un índice de bloques diferente, puede ser combinado con el valor espectral correspondiente del bloque n , donde este valor espectral de un bloque diferente no necesariamente tiene que tener

el mismo índice de frecuencias, sino que puede tener, por ejemplo, un índice de frecuencias adyacente diferente. Preferentemente, sin embargo, al menos el valor espectral con el mismo índice de frecuencias de un bloque diferente es combinado con el valor espectral del bloque actualmente considerado. Este otro bloque, de nuevo, no necesariamente tiene que ser el adyacente temporalmente directo, aunque esto sea especialmente preferible cuando el primer algoritmo de transformación y/o el segundo algoritmo de transformación tienen una característica de solapamiento de bloques, como es común para codificadores MP3 o codificadores de AAC.

[0052] Esto significa, cuando se consideran los factores de ponderación de la Figura 4, que al menos el factor de ponderación $c_4(k)$ es desigual a 0, y que al menos un segundo factor de ponderación es desigual a 0, mientras todos los demás factores de ponderación también pueden ser iguales a 0, los cuales también pueden proporcionar ya un procesamiento, que puede, sin embargo, debido al bajo número de factores de ponderación desiguales a 0 ser únicamente una aproximación relativamente aproximada del segundo algoritmo de transformación, si se considera de nuevo la mitad inferior de la Figura 1. Con el fin de tomar en cuenta más de nueve valores espectrales, las ramas adicionales para bloques adicionalmente en el futuro o adicionalmente en el pasado pueden ser añadidas. Adicionalmente, también los multiplicadores adicionales y los factores de ponderación correspondientes adicionales para valores espectrales que apoyan espectralmente pueden ser añadidos aparte, para generar un campo 3x3 en el campo de la Figura 4, que comprende más de tres líneas y/o más de tres columnas. Se ha encontrado, sin embargo, que cuando nueve factores de ponderación son admitidos para cada valor espectral, comparado con un número inferior de factores de ponderación, se logran mejoras sustanciales, mientras que cuando el número de factores de ponderación se aumenta, no se obtiene ninguna mejora adicional sustancial en cuanto a los valores diferenciales decrecientes en las salidas del bloque 22, de modo que una mayor cantidad de factores de ponderación con algoritmos de transformación comunes con un solapamiento de filtros de sub-banda adyacentes y un solapamiento temporal de bloques adyacentes no proporcione ninguna mejora sustancial.

[0053] En cuanto al solapamiento del 50 por ciento usado en la secuencia de bloques largos, se hace referencia a la ilustración esquemática de la Figura 5c a 45 a la izquierda de la figura, donde dos bloques largos posteriores se ilustran esquemáticamente. El concepto de combinador ilustrado en la Figura 4 se utiliza siempre de este modo, según la invención, cuando se utiliza una secuencia de bloques largos, donde la longitud de bloque del algoritmo IntMDCT 23 y el grado de solapamiento del algoritmo IntMDCT se configura igual al grado de solapamiento del filtro de análisis MP3 y la longitud de bloque del filtro de análisis MP3. En general, se prefiere que el solapamiento de bloques y la longitud de bloque de ambos algoritmos de transformación se configuren de igual modo, lo que no presenta ninguna limitación especial, como el segundo algoritmo de transformación, es decir por ejemplo el IntMDCT 23 de la Figura 2, puede configurarse fácilmente en cuanto a aquellos parámetros, mientras lo mismo no es fácilmente posible con el primer algoritmo de transformación, particularmente cuando el primer algoritmo de transformación es estandarizado como en cuanto al ejemplo de MP3 y se utiliza con frecuencia y no puede ser cambiado de este modo.

[0054] Cuando ya se ha ilustrado en cuanto a la Figura 2 y la Figura 3, el decodificador asociado en la Figura 3 revierte la formación diferencial de nuevo por una adición de los mismos valores de aproximación, es decir los valores diferenciales IntMDCT a la salida del bloque 22 de la Figura 2 o a la salida del bloque 31 de la Figura 3.

[0055] Según la invención, este método puede aplicarse así generalmente a la formación diferencial entre representaciones espectrales obtenidas al usar diferentes bancos de filtros, es decir cuando un banco de filtros/transformación que es la base del primer algoritmo de transformación es diferente de un banco de filtros/transformación que es la base del segundo algoritmo de transformación.

[0056] Un ejemplo para la aplicación concreta es el uso de los valores espectrales MP3 "del bloque largo" en relación con un IntMDCT, como se ha descrito en cuanto a la Figura 4. Como la resolución de la frecuencia del banco de filtros híbrido en este caso es 576, el IntMDCT también comprenderá una resolución de la frecuencia de 576, de modo que la longitud de la ventana pueda comprender un máximo de 1152 muestras temporales.

[0057] En el ejemplo descrito en lo sucesivo, solo se utilizan los vecinos temporales y espectrales directos, mientras que en el caso general también (o alternativamente) se pueden utilizar los valores que están más alejados.

[0058] Si el valor espectral de la banda k-th en el bloque MP3 de n-número es designado por $x(k, n)$ y el valor espectral correspondiente del IntMDCT es designado por $(y k, n)$, la diferencia se calcula como se ilustra en la Figura 4 para $d(k, n)$. $\hat{y}(k, n)$ es el valor de aproximación para $(y k, n)$ obtenido por la combinación lineal, y se determina cuando como se ilustra por la larga ecuación debajo de la Figura 4.

[0059] Cabe destacar aquí, que debido a la diferencia de fase diferente para cada una de las 576 sub-bandas puede requerirse un conjunto de coeficiente distinto. En la forma de realización práctica, como se ilustra en la Figura 4, para un acceso a valores espectrales temporalmente adyacentes se utilizan los retrasos 44 cuyos valores de salida corresponden respectivamente a valores de entrada en un bloque anterior correspondiente. Con el fin de permitir un acceso a valores espectrales temporalmente posteriores, así también los valores espectrales IntMDCT

como se aplican a la entrada 40 son retrasados por un retraso 46.

[0060] La Figura 5a muestra un procedimiento algo modificado cuando el banco de filtros híbrido MP3 proporciona bloques cortos en los que tres sub-bloques se generan respectivamente por 192 valores espectrales, en los que aquí aparte de la primera variante de la Figura 5a también se prefiere una segunda variante en la Figura 6a según la invención.

[0061] La primera variante se basa en una aplicación triple de un IntMDCT con una resolución de la frecuencia 192 para formar bloques correspondientes de los valores espectrales. Aquí, los valores de aproximación pueden formarse a partir de los tres valores que pertenecen a un índice de frecuencias y sus vecinos espectrales correspondientes. Para cada sub-bloque, se requiere aquí un conjunto distinto de coeficientes. Para describir el procedimiento se introduce de este modo un índice de sub-bloque u , de modo que n corresponda de nuevo al índice de un bloque completo de la longitud 576. Expresado como una ecuación, resulta de este modo el sistema de ecuaciones de la Figura 5a. Tal secuencia de bloques se ilustra en la Figura 5b en cuanto a los valores y y en la Figura 5c en cuanto a las ventanas. El codificador MP3 proporciona bloques MP3 cortos, como se ilustran en 50. La primera variante proporciona también bloques IntMDCT cortos $y(u_0)$, $y(u_1)$ y $y(u_2)$, como se ilustra en 51 en la Figura 5b. Por esto, tres bloques diferenciales cortos 52 pueden calcularse tal como un 1:1 la representación resulta entre un valor espectral correspondiente a la frecuencia k en los bloques 50, 51 y 52.

[0062] En contraste con la Figura 4 cabe destacar, que en la Figura 5a no se indican los retrasos 44. Esto resulta a partir del hecho de que el procesamiento posterior solo puede llevarse a cabo cuando se han calculado los tres sub-bloques 0, 1, 2 para un bloque n . Si el sub-bloque con el índice 0 es temporalmente el primer sub-bloque, y si el siguiente sub-bloque con el índice 1 es el bloque temporalmente posterior, y si el índice $u=2$ es de nuevo el bloque corto temporalmente posterior, entonces el bloque diferencial para el índice $u=0$ se calcula mediante el uso de los valores espectrales del sub-bloque u_0 , el sub-bloque u_1 y el sub-bloque u_2 . Esto significa, que solo con referencia al sub-bloque actualmente calculado con el índice 0 se utilizan los futuros sub-bloques 1 y 2, sin embargo ningún valor espectral del pasado. Esto es prudente, cuando se llevó a cabo una permutación a bloques cortos, ya que había un resultado transitorio en la señal de audio como es conocido y por ejemplo ilustrado en la publicación del experto anteriormente mencionado de Edler. Los valores con procesamiento posterior para el sub-bloque mediante el uso del índice 1 para ganar los valores diferenciales que tienen el índice de sub-bloque 1 son, sin embargo, calculados a partir de un sub-bloque temporalmente anterior, temporalmente actual y temporalmente posterior, mientras que los valores espectrales con procesamiento posterior para el tercer sub-bloque con el índice 2 no son calculados mediante el uso de futuros sub-bloques, sino solo mediante el uso de sub-bloques pasados que tienen el índice 1 y el índice 0, lo que también es técnicamente prudente en la medida en que de nuevo, como se indica en la Figura 5c, se puede iniciar fácilmente una permutación de ventanas a ventanas largas por una ventana de finalización, de modo que más tarde de nuevo se pueda llevar a cabo un cambio directamente en el esquema de bloque largo de la Figura 4.

[0063] La Figura 5 aclara de este modo que en particular con bloques cortos, sin embargo también generalmente, puede ser prudente mirar solo en el pasado o en el futuro y no siempre, como se indica en la Figura 4, tanto en el pasado como en el futuro, para obtener valores espectrales que proporcionan un valor espectral con procesamiento posterior después de una ponderación y una adición.

[0064] En lo sucesivo, con referencia a la Figura 6a, 6b y 6c, se ilustra la segunda variante para bloques cortos. En la segunda variante, la resolución de la frecuencia del IntMDCT todavía es 576, de modo que tres valores espectrales IntMDCT espectralmente adyacentes cada uno reside en el intervalo de frecuencias de un valor espectral MP3. Así, para cada uno de los tres valores espectrales IntMDCT, para una formación diferencial una combinación lineal distinta se forma a partir de los tres valores espectrales de sub-bloque temporalmente posteriores y sus vecinos espectrales, donde el índice s que también es referido como un índice de orden indica ahora la posición dentro de cada grupo de tres. Así, resulta la ecuación como se ilustra en la Figura 6a debajo del diagrama de bloques. Esta segunda variante es especialmente adecuada si una función de ventana con una pequeña área de solapamiento se usa en el IntMDCT, dado que entonces la sección de señal considerada corresponde bien a la de los tres sub-bloques. En este caso, como con la primera variante, se prefiere adaptar las formas de ventana del IntMDCT de los bloques largos anteriores o posteriores, respectivamente, de modo que resulta una reconstrucción perfecta. Un diagrama de bloques correspondiente para la primera variante se ilustra en la Figura 5c. Un diagrama correspondiente para la segunda variante se ilustra en la Figura 6c, en el que ahora solo un bloque IntMDCT largo individual se genera por la ventana larga 63, en el que este bloque de IntMDCT largo comprende ahora k bloques triples de valores espectrales, donde la amplitud de banda de tal bloque triple que resulta de $s=0$, $s=1$ y $s=2$ es igual a la amplitud de banda de un bloque k de los bloques MP3 cortos 60 en la Figura 6b. A partir de la Figura 6a se puede observar que para una substracción del primer valor espectral con $s=0$ para un bloque triple que tiene el índice k se utilizan de nuevo los valores del sub-bloque actual, el futuro sub-bloque y el siguiente futuro sub-bloque 0, 1, 2, sin embargo, no se utiliza ningún valor del pasado. Para calcular un valor diferencial para el segundo valor $s=1$ de un grupo triple, sin embargo, se utilizan los valores espectrales del sub-bloque anterior y el futuro sub-bloque, mientras que para calcular un valor espectral diferencial que tiene el índice de orden $s=2$, solo se utilizan sub-bloques anterior, como se ilustra por las ramas 41 y 42 que están en el pasado en cuanto a la rama 43 en la Figura

6a.

[0065] En este punto cabe destacar que, con todas las regulaciones de cálculo, los términos que exceden los límites del intervalo de frecuencias es decir, por ejemplo, el índice de frecuencia -1 ó 576 ó 192, respectivamente, se omiten cada uno. En estos casos, en el ejemplo general en las Figuras 4 a 6 la combinación lineal se reduce así a 6 términos en vez de 9 términos.

[0066] A continuación, se hace referencia detallada a las secuencias de ventanas en la Figura 5c y la Figura 6c. Las secuencias de ventanas consisten en una secuencia de bloques largos, como son procesados por el caso en la Figura 4. En ese momento, una ventana de inicio 56 que tiene una forma asimétrica, ya que es "convertida" a partir de un área de solapamiento larga a principios de la ventana de inicio a un área de solapamiento corta al final de la ventana de inicio. Análogo a esto, existe una ventana de finalización 57 que es convertida de nuevo de una secuencia de bloques cortos a una secuencia de bloques largos y comprende así un área corta de solapamiento al principio y un área larga de solapamiento al final.

[0067] Una permutación de ventanas es, como ilustra en la publicación del experto mencionado de Edler, seleccionada si un período de tiempo en la señal de audio es detectado por un codificador que comprende una señal transitoria.

[0068] Tal señalización se localiza en el flujo de bits de MP3, de modo que cuando el IntMDCT, según la Figura 2 y según la primera variante de la Figura 5c, también cambia a bloques cortos, ninguna detección transitoria distinta sea necesaria, sino que tiene lugar una detección transitoria basada solo en un aviso de ventana corta en el flujo de bits de MP3. Para el procesamiento posterior de valores en la ventana de inicio se prefiere, debido al área larga de solapamiento con la ventana anterior, el uso de bloques con el índice de bloques anterior n-1, mientras los bloques con el índice de bloques posterior solo son ligeramente ponderados o generalmente no usados debido al área corta de solapamiento. Análogo a esto, la ventana de finalización para el procesamiento posterior solo tendrá en cuenta valores con un futuro índice de bloques n+1 además de los valores para el bloque actual n, pero solo llevará a cabo una ponderación débil o una ponderación igual a 0, es decir ningún uso del pasado, es decir, por ejemplo del tercer bloque corto.

[0069] Cuando, como se muestra en la Figura 6c, la secuencia de ventanas como es implementado por el IntMDCT 23, es decir el segundo algoritmo de transformación, no lleva a cabo ninguna permutación a ventanas cortas, sin embargo pone en práctica la permutación de ventanas preferentemente usada, a continuación se prefiere iniciar o terminar, respectivamente, la ventana con el solapamiento corto, designado por 63 en la Figura 6c, también por una ventana de inicio 56 y por una ventana de finalización 57.

[0070] Aunque en la forma de realización ilustrada en la Figura 6c el IntMDCT de la Figura 2 no cambia en el modo de ventana corta, la señalización de ventanas cortas en el flujo de bits del MP3 puede ser de todos modos utilizada para activar la permutación de ventanas con una ventana de inicio, ventana con solapamiento corto, como se indica en la Figura 6c a 63, y ventana de finalización.

[0071] Adicionalmente cabe destacar, que en particular las secuencias de ventanas ilustradas en la norma de AAC, adaptada a la longitud de bloque MP3 o la alimentación de MP3, respectivamente, de 576 valores para bloques largos y 192 valores para bloques cortos, y en particular también las ventanas de inicio y las ventanas de finalización ilustradas allí, son sobre todo adecuadas para una forma de realización del IntMDCT en el bloque 23 de la presente invención.

[0072] En lo sucesivo, se hace referencia a la exactitud de la aproximación del primer algoritmo de transformación y el procesamiento posterior.

[0073] Para las 576 señales de entrada respectivamente que tienen un impulso en la posición 0 ... 575 dentro de un bloque, se llevaron a cabo las etapas siguientes:

- cálculo del banco de filtros híbrido + aproximación
- cálculo de la MDCT
- cálculo de la suma de cuadrados de los componentes espectrales de la MDCT
- cálculo de la suma de cuadrados de las desviaciones entre componentes espectrales de la MDCT y la aproximación. Aquí, se determina la desviación de cuadrados máxima a través de 576 señales.

[0074] La desviación de cuadrados relativa máxima a través de todas las posiciones era, al utilizar

- bloques largos según la Figura 4, aprox. el 3,3%
- bloques cortos (híbrido) y bloques largos (MDCT) según la Figura 6, aprox. el 20,6%.

[0075] Uno podría mencionar de este modo, que con un impulso en las entradas de las dos transformaciones,

la suma de cuadrados de las desviaciones entre la aproximación y los componentes espectrales de la segunda transformación no deberían ser más del 30% (y preferentemente no incluso más del 25% o el 10% respectivamente) de la suma de cuadrados de los componentes espectrales de la segunda transformación, independiente de la posición del impulso en el bloque de entrada. Para calcular las sumas cuadradas, todos los bloques de componentes
5 espectrales deberían considerarse que están bajo la influencia del impulso.

[0076] Cabe destacar, que en la inspección de error superior (MDCT contra FB híbrido + procesamiento posterior) siempre se consideró el error relativo que es la señal independiente.

10 **[0077]** En el IntMDCT (en comparación con la MDCT), sin embargo, el error absoluto es la señal independiente y reside en un intervalo de aproximadamente -2 a 2 de los valores de número entero redondeados. De esto resulta que el error relativo se vuelve dependiente de la señal. Con el fin de eliminar esta dependencia de señal, preferentemente se asume un impulso completamente controlado (por ejemplo, valor 32767 a 16 bits PCM).

15 **[0078]** Esto dará como resultado entonces un espectro prácticamente plano con una amplitud media de aproximadamente $32767/\sqrt{576} = 1365$ (conservación de energía). El error cuadrático medio sería de aproximadamente $2^2/1365^2=0,0002\%$, es decir, insignificante.

[0079] Con un impulso muy bajo en la entrada, el error sería drástico, sin embargo. Un impulso de la amplitud
20 1 ó 2 sería prácticamente perdido por completo en el error de aproximación IntMDCT.

[0080] El criterio de error de la exactitud de la aproximación, es decir el valor deseado para los factores de ponderación, es así el mejor, cuando esto se indica para un impulso completamente controlado.

25 **[0081]** La forma de realización de la invención comprende un codificador para la codificación de una señal de audio, que comprende: un dispositivo 10 para el procesamiento posterior de los valores espectrales basados en un primer algoritmo de transformación 16 para convertir una señal de audio en una representación espectral, que comprende: un medio 12 para proporcionar una secuencia de bloques de los valores espectrales que representan una secuencia de bloques de muestras de la señal de audio; y un combinador 13 para agregar valores espectrales
30 de forma ponderada de la secuencia de bloques de valores espectrales con el fin de obtener una secuencia de bloques de valores espectrales con procesamiento posterior, en donde el combinador 13 se implementa para utilizar, para el cálculo de un valor espectral con procesamiento posterior para una banda de frecuencia y una duración de tiempo, un valor espectral de la secuencia de bloques para la banda de frecuencia y la duración de tiempo, y un valor espectral para otra banda de frecuencia u otra duración de tiempo, y en el que el combinador 13 se implementa
35 para utilizar tales factores de ponderación cuando se añaden de forma ponderada, que los valores espectrales con procesamiento posterior son una aproximación a los valores espectrales a medida que se obtienen mediante un segundo algoritmo de transformación 17 para convertir la señal de audio en una representación espectral, en donde el segundo algoritmo de transformación 17 es diferente del primer algoritmo de transformación 16, en donde el codificador comprende además un medio 23 para calcular una secuencia de bloques de valores espectrales según
40 el segundo algoritmo de transformación 17 de la señal de audio; y un medio 22 para una formación de diferencia de valor espectral entre la secuencia de bloques debida al segundo algoritmo de transformación y la secuencia de bloques de valores espectrales con procesamiento posterior.

[0082] Una forma de realización del codificador comprende además: un medio 24 para generar un flujo de
45 bits de extensión debido a un resultado generado por los medios 22 para una formación de diferencia de valor espectral.

[0083] En una forma de realización del codificador, los medios 24 para generación comprenden un
50 codificador de entropía.

[0084] En una forma de realización del codificador, la secuencia de bloques debida al primer algoritmo de transformación 16 se basa en una compresión con pérdida, y en la que la secuencia de bloques debida al segundo algoritmo de transformación 17 se basa en una compresión sin pérdida o prácticamente sin pérdida.

55 **[0085]** Una forma de realización del codificador comprende una memoria para almacenar los factores de ponderación en los que los factores de ponderación son almacenables independientemente de una señal.

[0086] En una forma de realización del codificador, los medios 23 para generar la secuencia de bloques mediante el uso del segundo algoritmo de transformación 17 se implementan para realizar una ventana con una
60 secuencia de ventana que depende de una secuencia de ventana en la que se basa la secuencia de bloques de los valores espectrales el cual es dado debido al primer algoritmo de transformación 16.

[0087] En una forma de realización del codificador, los medios 23 para proporcionar una secuencia de bloques mediante el uso del segundo algoritmo de transformación 17 se implementan para cambiar de una ventana
65 larga con un área de solapamiento larga a una ventana larga con un área de solapamiento corta o a una pluralidad

de ventanas cortas, cuando en la secuencia de bloques de los valores espectrales debido al primer algoritmo de transformación 16 tiene lugar un cambio a ventanas cortas.

[0088] La forma de realización de la invención comprende un decodificador para la decodificación de una señal de audio codificada, que comprende: un dispositivo 10 para el procesamiento posterior de los valores espectrales basados en un primer algoritmo de transformación 16 para convertir una señal de audio en una representación espectral, que comprende: un medio 12 para proporcionar una secuencia de bloques de los valores espectrales que representan una secuencia de bloques de muestras de la señal de audio; y un combinador 13 para agregar valores espectrales de forma ponderada de la secuencia de bloques de valores espectrales con el fin de obtener una secuencia de bloques de valores espectrales con procesamiento posterior, en donde el combinador 13 se implementa para usar, para el cálculo de un valor espectral con procesamiento posterior para una banda de frecuencia y una duración de tiempo, un valor espectral de la secuencia de bloques para la banda de frecuencia y la duración de tiempo, y un valor espectral para otra banda de frecuencia u otra duración de tiempo, y en el que el combinador 13 se implementa para utilizar tales factores de ponderación cuando se añaden de forma ponderada, que los valores espectrales con procesamiento posterior son una aproximación a los valores espectrales a medida que se obtienen mediante un segundo algoritmo de transformación 17 para convertir la señal de audio en una representación espectral, en donde el segundo algoritmo de transformación 17 es diferente del primer algoritmo de transformación 16; en el que el decodificador comprende además un medio 31 para proporcionar valores diferenciales de valor espectral entre una secuencia de bloques de valores espectrales con procesamiento posterior debido al primer algoritmo de transformación 16 y una secuencia de bloques debido al segundo algoritmo de transformación 17; un medio 30 para combinar la secuencia de bloques de los valores espectrales con procesamiento posterior y los valores diferenciales con el fin de obtener una secuencia de bloques de valores espectrales de combinación; y un medio 33 para transformar de manera inversa la secuencia de bloques de valores espectrales de combinación según el segundo algoritmo de transformación 17 para obtener una señal de audio decodificada.

[0089] Las formas de realización comprenden además un método para la codificación de una señal de audio, que comprende las siguientes etapas: procesamiento posterior de valores espectrales que se basan en un primer algoritmo de transformación 16 para convertir una señal de audio en una representación espectral, que comprende las siguientes etapas: proporcionar 12 una secuencia de bloques de los valores espectrales que representan una secuencia de bloques de muestras de la señal de audio; y agregar de forma ponderada 13 los valores espectrales de la secuencia de bloques de valores espectrales para obtener una secuencia de bloques de valores espectrales con procesamiento posterior, en donde para calcular un valor espectral con procesamiento posterior para una banda de frecuencia y una duración de tiempo se utilizan un valor espectral de la secuencia de bloques para la banda de frecuencia y la duración de tiempo y un valor espectral para otra banda de frecuencia u otra duración de tiempo, y en el que tales factores de ponderación se utilizan cuando se agregan de forma ponderada de manera que los valores espectrales con procesamiento posterior sean una aproximación a los valores espectrales a medida que se obtienen por segundo algoritmo de transformación 17 para convertir la señal de audio en una representación espectral, en donde el segundo algoritmo de transformación 17 es diferente del primer algoritmo de transformación 16; en el que el método comprende además el cálculo 23 de una secuencia de bloques de valores espectrales según el segundo algoritmo de transformación 17 a partir de la señal de audio; y la formación de diferencia de valor espectral 22 entre la secuencia de bloques de valores espectrales debido al segundo algoritmo de transformación y la secuencia de bloques de valores espectrales con procesamiento posterior.

[0090] Las formas de realización comprenden además un método para la decodificación de una señal de audio codificada, que comprende las siguientes etapas: procesamiento posterior de los valores espectrales que se basan en un primer algoritmo de transformación 16 para convertir una señal de audio en una representación espectral, que comprende las siguientes etapas: proporcionar 12 una secuencia de bloques de los valores espectrales que representan una secuencia de bloques de muestras de la señal de audio; y adición de forma ponderada 13 de valores espectrales de la secuencia de bloques de valores espectrales para obtener una secuencia de bloques de valores espectrales con procesamiento posterior, en donde para calcular un valor espectral con procesamiento posterior para una banda de frecuencia y una duración de tiempo se utilizan un valor espectral de la secuencia de bloques para la banda de frecuencia y la duración de tiempo y un valor espectral para otra banda de frecuencia u otra duración de tiempo, y en el que tales factores de ponderación se utilizan cuando se agregan de forma ponderada de manera que los valores espectrales con procesamiento posterior sean una aproximación a los valores espectrales a medida que se obtienen por el segundo algoritmo de transformación 17 para convertir la señal de audio en una representación espectral, en donde el segundo algoritmo de transformación 17 es diferente del primer algoritmo de transformación; en el que el método de decodificación comprende además proporcionar 31 los valores diferenciales de valor espectral entre una secuencia de bloques de los valores espectrales con procesamiento posterior debido al primer algoritmo de transformación 16 y una secuencia de bloques de valores espectrales debido al segundo algoritmo de transformación 17; combinar 30 la secuencia de bloques de los valores espectrales con procesamiento posterior y los valores diferenciales para obtener una secuencia de bloques de valores espectrales de combinación; y transformar de manera inversa 33 la secuencia de bloques de valores espectrales de combinación según el segundo algoritmo de transformación 17 para obtener una señal de audio decodificada.

[0091] Dependiendo de las circunstancias, el método inventivo puede implementarse en el hardware o en el software. La implementación puede tener lugar en un medio de almacenamiento digital, en particular un disquete o un CD que tienen señales de control legibles por medios electrónicos, que pueden cooperar con un sistema informático programable de modo que el método se lleve a cabo. En general, la invención consiste también de este modo en un producto de programa informático que tiene un código de programa en un soporte legible por máquina para llevar a cabo el método inventivo, cuando el producto de programa informático se ejecuta en un ordenador. En otras palabras, la invención puede ser realizada de este modo como un programa informático que tiene un código de programa para llevar a cabo el método, cuando el programa informático se ejecuta en un ordenador.

10

REIVINDICACIONES

1. Un dispositivo (10) para el procesamiento posterior de valores espectrales basado en un primer algoritmo de transformación (16) para convertir una señal de audio en una representación espectral, que comprende:

5 un medio (12) para proporcionar una secuencia de bloques de los valores espectrales que representan una secuencia de bloques de muestras de la señal de audio; y

un combinador (13) para agregar de manera ponderada valores espectrales de la secuencia de bloques de valores
10 espectrales con el fin de obtener una secuencia de bloques de valores espectrales con procesamiento posterior, donde el combinador (13) se implementa para utilizar, para el cálculo de un valor espectral con procesamiento posterior para una banda de frecuencia y un período de tiempo, un valor espectral de la secuencia de bloques para la banda de frecuencia y el período de tiempo, y un valor espectral para otra banda de frecuencia u otro período de tiempo, y donde el combinador (13) se implementa para utilizar tales factores de ponderación cuando se agregan de
15 manera ponderada, de modo que los valores espectrales con procesamiento posterior sean una aproximación a valores espectrales a medida que se obtienen por un segundo algoritmo de transformación (17) para convertir la señal de audio en una representación espectral,

en el que el primer algoritmo de transformación es tal que los filtros adyacentes de un banco de filtros se solapan y
20 en el que el primer algoritmo de transformación comprende un solapamiento de bloques,

en el que el segundo algoritmo de transformación (17) es diferente del primer algoritmo de transformación (16), y

en el que el primer algoritmo de transformación (16) y el segundo algoritmo de transformación (17) son tales que
25 proporcionan señales de salida reales,

en el que el combinador (13) se implementa para implementar la siguiente ecuación:

$$\begin{aligned} \hat{Y}(k,n) = & c_0(k)x(k-1,n-1)+c_1(k)x(k-1,n)+c_2(k)x(k-1,n+1) \\ & + c_3(k)x(k,n-1)+c_4(k)x(k,n)+c_5(k)x(k,n+1) \\ & + c_6(k)x(k+1,n-1)+c_7(k)x(k+1,n)+c_8(k)x(k+1,n+1) \end{aligned}$$

30

en la que $\hat{Y}(k,n)$ es un valor espectral con procesamiento posterior para un índice de frecuencia k y un índice de tiempo n, en la que x(k,n) es un valor espectral de un bloque de valores espectrales con un índice de frecuencia k y un índice de tiempo n, donde $c_0(k), \dots, c_8(k)$ son factores de ponderación, asociados con el índice de frecuencia k, en la que k-1 es un índice de frecuencia disminuido, en la que k + 1 es un índice de frecuencia incrementado, en la
35 que n-1 es un índice de tiempo disminuido y en la que n + 1 es un índice de tiempo incrementado, o

en el que el combinador (13) se implementa para implementar la siguiente ecuación:

$$\begin{aligned} \hat{Y}(k,n,u) = & c_0(k,u)x(k-1,n,0)+c_1(k,u)x(k-1,n,1)+c_2(k,u)x(k-1,n,2) \\ & + c_3(k,u)x(k,n,0)+c_4(k,u)x(k,n,1)+c_5(k,u)x(k,n,2) \\ & + c_6(k,u)x(k+1,n,0)+c_7(k,u)x(k+1,n,1)+c_8(k,u)x(k+1,n,2) \end{aligned}$$

40 en la que $\hat{Y}(k,n,u)$ es un valor espectral con procesamiento posterior para un índice de frecuencia k y un índice de tiempo n y un índice de sub-bloque u, en la que x(k,n,u) es un valor espectral de un bloque de valores espectrales con un índice de frecuencia k y un índice de tiempo n y un índice de sub-bloque u, en donde $c_0(k), \dots, c_8(k)$ son factores de ponderación asociados con el índice de frecuencia k, en la que k-1 es un índice de frecuencia reducida, en la que k + 1 es un índice de frecuencia incrementada, en la que n-1 es un índice de tiempo disminuido y en la que
45 n + 1 es un índice de tiempo incrementado, en la que u es un índice de sub-bloque que indica una posición de un sub-bloque en una secuencia de sub-bloques, y en la que el índice de tiempo especifica un bloque largo y el índice de sub-bloques especifica un bloque comparativamente corto, o

en el que el combinador (13) se implementa para implementar la siguiente ecuación:

50

$$\hat{y}(3k+s,n) = c_0(k,s)x(k-1,n,0)+c_1(k,s)x(k-1,n,1)+c_2(k,s)x(k-1,n,2) \\ + c_3(k,s)x(k,n,0)+c_4(k,s)x(k,n,1)+c_5(k,s)x(k,n,2) \\ + c_6(k,s)x(k+1,n,0)+c_7(k,s)x(k+1,n,1)+c_8(k,s)x(k+1,n,2)$$

en la que $\hat{y}(k,n)$ es un valor espectral con procesamiento posterior para un índice de frecuencia k y un índice de tiempo n , en la que $x(k,n,u)$ es un valor espectral de un bloque de valores espectrales con un índice de frecuencia k y un tiempo índice n y un sub-bloque índice u , donde $c_0(k), \dots, c_8(k)$ son factores de ponderación asociados con el índice de frecuencia k , en la que $k-1$ es un índice de frecuencia disminuido, en la que $k+1$ es un índice de incrementado, en la que $n-1$ es un índice de tiempo disminuido y en la que $n+1$ es un índice de tiempo incrementado, en la que s es un índice de orden que indica una posición de un sub-bloque en una secuencia de sub-bloques, y en la que el índice de tiempo especifica un bloque largo y el índice de sub-bloques especifica un bloque comparativamente corto.

2. El dispositivo según la reivindicación 1, en el que el primer algoritmo de transformación (16) es un algoritmo de transformación híbrida que tiene dos etapas, y el segundo algoritmo de transformación (17) es un algoritmo de transformación en una etapa.

3. El dispositivo según la reivindicación 1, en el que el primer algoritmo de transformación (16) comprende un banco de filtros polifásicos y una transformación discreta en coseno modificada, y en el que el segundo algoritmo de transformación (17) es una MDCT de enteros.

4. El dispositivo según la reivindicación 1, en el que el combinador (13) se implementa para utilizar factores de ponderación tales que el primer algoritmo de transformación (16) y un procesamiento posterior realizado por el combinador (13) proporcionan juntos una respuesta de impulso que se aproxima a una respuesta de impulso del segundo algoritmo de transformación (17).

5. El dispositivo según la reivindicación 4, en el que en una aproximación del primer algoritmo de transformación y procesamiento posterior, los factores de ponderación son seleccionados de tal manera que con un impulso en la entrada de las dos transformaciones, la suma de cuadrados de las desviaciones entre la aproximación y los componentes espectrales de la segunda transformación no sea de más del 30% de la suma de cuadrados de los componentes espectrales de la segunda transformación.

6. El dispositivo según la reivindicación 1, en el que el medio (12) para proporcionar una secuencia de bloques se implementa para proporcionar bloques que son una representación con pérdidas de la señal de audio.

7. El dispositivo según la reivindicación 1, en el que el combinador (13) para un cálculo de un valor espectral con procesamiento posterior para una banda de frecuencia k comprende:

una primera sección (41, 42, 43) para valores espectrales ponderados de un bloque actual para la banda de frecuencia k , una banda de frecuencia $k-1$ o una banda de frecuencia $k+1$, con el fin de obtener valores espectrales ponderados para el bloque actual;

una segunda sección (41, 42, 43) para valores espectrales ponderados de un bloque temporalmente anterior $k-1$ o bloque temporalmente posterior $k+1$, con el fin de obtener valores espectrales ponderados para el bloque temporalmente anterior o el bloque temporalmente posterior; y

un medio (45) para añadir los valores espectrales ponderados para obtener un valor espectral con procesamiento posterior para la banda de frecuencia k de un bloque actual o bloque anterior o bloque posterior de valores espectrales con procesamiento posterior.

8. El dispositivo según la reivindicación 7, que comprende además:

una tercera sección (43) para ponderar valores espectrales de un bloque anterior, en el que la primera sección se implementa para ponderar valores espectrales de un bloque posterior, y el que la segunda sección (42) se implementa para ponderar valores espectrales de un bloque actual, y en el que el medio (45) se implementa para añadir valores espectrales ponderados de las tres secciones con el fin de obtener un valor espectral con procesamiento posterior para el bloque actual de valores espectrales con procesamiento posterior.

9. El dispositivo según la reivindicación 1,

en el que el primer algoritmo de transformación comprende una función de solapamiento de bloques, en el que los

bloques de las muestras de la señal de audio temporal en la que la secuencia de bloques de valores espectrales se basa en solapamiento.

10. El dispositivo según la reivindicación 1, en el que el combinador (13) se implementa para utilizar un conjunto independiente de señal de factores de ponderación para cada valor espectral.

11. El dispositivo según la reivindicación 1, en el que la secuencia de bloques de los valores espectrales comprende un conjunto de bloques de valores espectrales que son más cortos que un bloque largo de valores espectrales que sigue después del conjunto de bloques o que precede al conjunto de bloques, y en el que el combinador (13) se implementa para utilizar la misma banda de frecuencia o una banda de frecuencia adyacente de varios bloques del conjunto de bloques cortos para calcular un valor espectral con procesamiento posterior para el conjunto de bloques de valores espectrales.

12. El dispositivo según la reivindicación 11, en el que el combinador (13) se implementa solo para utilizar valores espectrales de bloques cortos y ningún valor espectral de un bloque largo anterior o un bloque largo posterior para calcular valores espectrales con procesamiento posterior debido a bloques cortos de valores espectrales.

13. Un decodificador para decodificar una señal de audio codificada o un codificador para codificar una señal de audio, que comprende:

un dispositivo para valores espectrales con procesamiento posterior basado en un primer algoritmo de transformación (16) según la reivindicación 1;

en el que el primer y el segundo algoritmos de transformación son incompatibles entre sí, en el que el dispositivo para el procesamiento posterior está configurado para generar los valores espectrales con procesamiento posterior de modo que los valores espectrales con procesamiento posterior sean compatibles con el segundo algoritmo de transformación; y

en el que los valores espectrales con procesamiento posterior y los valores espectrales obtenidos por el segundo algoritmo de transformación se procesan juntos.

14. Un método (10) para el procesamiento posterior de valores espectrales que se basan en un primer algoritmo de transformación (16) para convertir una señal de audio en una representación espectral, que comprende las etapas siguientes:

el suministro (12) de una secuencia de bloques de los valores espectrales que representan una secuencia de bloques de muestras de la señal de audio; y

adición de manera ponderada (13) de los valores espectrales de la secuencia de bloques de valores espectrales para obtener una secuencia de bloques de valores espectrales con procesamiento posterior, en el que para calcular un valor espectral con procesamiento posterior para una banda de frecuencia y un período de tiempo, un valor espectral de la secuencia de bloques para la banda de frecuencia y el período de tiempo y un valor espectral para otra banda de frecuencia u otro período de tiempo son utilizados, y en el que tales factores de ponderación se usan al agregar de manera ponderada de modo que los valores espectrales con procesamiento posterior sean una aproximación a valores espectrales cuando se obtienen por un segundo algoritmo de transformación (17) para convertir la señal de audio en una representación espectral,

en el que el primer algoritmo de transformación es tal que los filtros adyacentes de un banco de filtros se solapan y en el que el primer algoritmo de transformación comprende un solapamiento de bloques,

en el que el segundo algoritmo de transformación (17) es diferente del primer algoritmo de transformación (16), y

en el que el primer algoritmo de transformación (16) y el segundo algoritmo de transformación (17) son tales que proporcionan señales de salida reales,

en el que la suma (13) comprende la implementación de la siguiente ecuación:

$$\begin{aligned} \hat{Y}(k,n) = & c_0(k)x(k-1,n-1)+c_1(k)x(k-1,n)+c_2(k)x(k-1,n+1) \\ & + c_3(k)x(k,n-1)+c_4(k)x(k,n)+c_5(k)x(k,n+1) \\ & + c_6(k)x(k+1,n-1)+c_7(k)x(k+1,n)+c_8(k)x(k+1,n+1) \end{aligned}$$

en la que $\hat{Y}(k,n)$ es un valor espectral con procesamiento posterior para un índice de frecuencias k y un índice de tiempo n , en la que $x(k,n)$ es un valor espectral de un bloque de valores espectrales con un índice de frecuencias k y un índice de tiempo n , en la que $c_0(k), \dots, c_8(k)$ son factores de ponderación, asociados con el índice de frecuencias k , en la que $k-1$ es un índice de frecuencias disminuido, en la que $k+1$ es un índice de frecuencia incrementado, en la que $n-1$ es un índice de tiempo disminuido y en la que $n+1$ es un índice de tiempo incrementado, o

10 en el que la suma (13) comprende la implementación de la siguiente ecuación:

$$\begin{aligned} \hat{Y}(k,n,u) = & c_0(k,u)x(k-1,n,0)+c_1(k,u)x(k-1,n,1)+c_2(k,u)x(k-1,n,2) \\ & + c_3(k,u)x(k,n,0)+c_4(k,u)x(k,n,1)+c_5(k,u)x(k,n,2) \\ & + c_6(k,u)x(k+1,n,0)+c_7(k,u)x(k+1,n,1)+c_8(k,u)x(k+1,n,2) \end{aligned}$$

15 en la que $\hat{Y}(k,n,u)$ es un valor espectral con procesamiento posterior para un índice de frecuencias k y un índice de tiempo n y un índice de sub-bloque u , en la que $x(k,n,u)$ es un valor espectral de un bloque de valores espectrales con un índice de frecuencias k y un índice de tiempo n y un índice de sub-bloque u , en la que $c_0(k), \dots, c_8(k)$ son factores de ponderación asociados con el índice de frecuencias k , en la que $k-1$ es un índice de frecuencias disminuido, en la que $k+1$ es un índice de frecuencia incrementado, en la que $n-1$ es un índice de tiempo disminuido y en la que $n+1$ es un índice de tiempo incrementado, en la que u es un índice de sub-bloque que indica una posición de un sub-bloque en una secuencia de sub-bloques, y en la que el índice de tiempo especifica un bloque largo y el índice de sub-bloque especifica un bloque comparativamente corto, o

en el que la suma (13) comprende la implementación de la siguiente ecuación:

$$\begin{aligned} \hat{Y}(3k+s,n) = & c_0(k,s)x(k-1,n,0)+c_1(k,s)x(k-1,n,1)+c_2(k,s)x(k-1,n,2) \\ & + c_3(k,s)x(k,n,0)+c_4(k,s)x(k,n,1)+c_5(k,s)x(k,n,2) \\ & + c_6(k,s)x(k+1,n,0)+c_7(k,s)x(k+1,n,1)+c_8(k,s)x(k+1,n,2) \end{aligned}$$

25 en la que $\hat{Y}(k,n)$ es un valor espectral con procesamiento posterior para un índice de frecuencias k y un índice de tiempo n , en la que $x(k,n,u)$ es un valor espectral de un bloque de valores espectrales con un índice de frecuencias k y un índice de tiempo n y un índice de sub-bloque u , en la que $c_0(k), \dots, c_8(k)$ son factores de ponderación asociados con el índice de frecuencias k , en la que $k-1$ es un índice de frecuencias disminuido, en la que $k+1$ es un índice de frecuencia incrementado, en la que $n-1$ es un índice de tiempo disminuido y en la que $n+1$ es un índice de tiempo incrementado, en la que s es un índice de orden que indica una posición de un sub-bloque en una secuencia de sub-bloques, y en la que el índice de tiempo especifica un bloque largo y el índice de sub-bloque especifica un bloque comparativamente corto.

35 15. Un método para la decodificación de una señal de audio codificada o un método para la codificación de una señal de audio, que comprende las etapas siguientes:

40 procesamiento posterior de valores espectrales basados en un primer algoritmo de transformación (16) según la reivindicación 14;

en el que el primer y el segundo algoritmos de transformación son incompatibles entre sí, en el que el dispositivo para el procesamiento posterior está configurado para generar los valores espectrales con procesamiento posterior de modo que los valores espectrales con procesamiento posterior sean compatibles con el segundo algoritmo de transformación; y

en el que los valores espectrales con procesamiento posterior y los valores espectrales obtenidos por el segundo algoritmo de transformación se procesan juntos.

16. Un programa informático que tiene un código de programa para llevar a cabo el método de la
5 reivindicación 13 ó 15, cuando el programa informático se ejecuta en un ordenador.

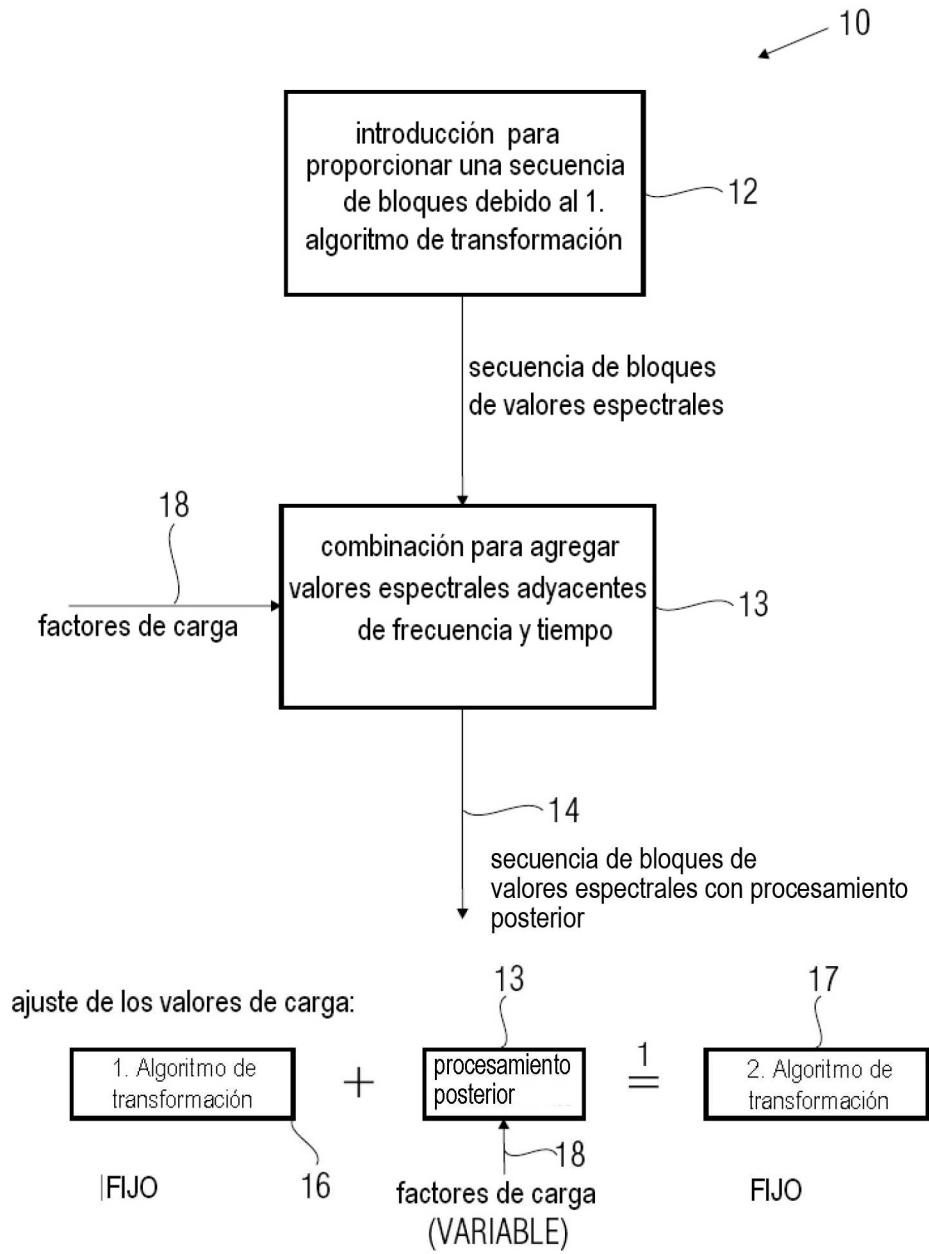


FIGURA 1

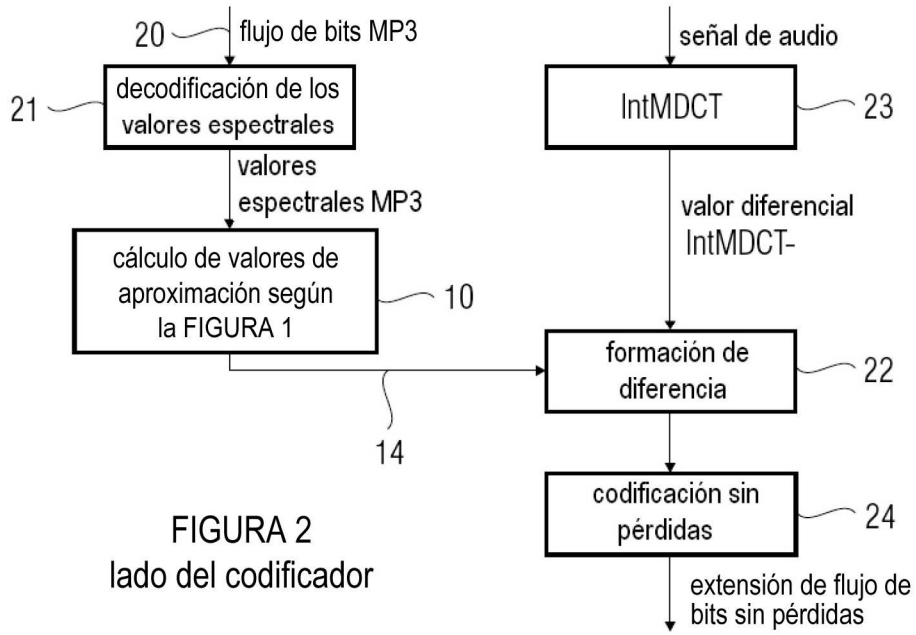


FIGURA 2
lado del codificador

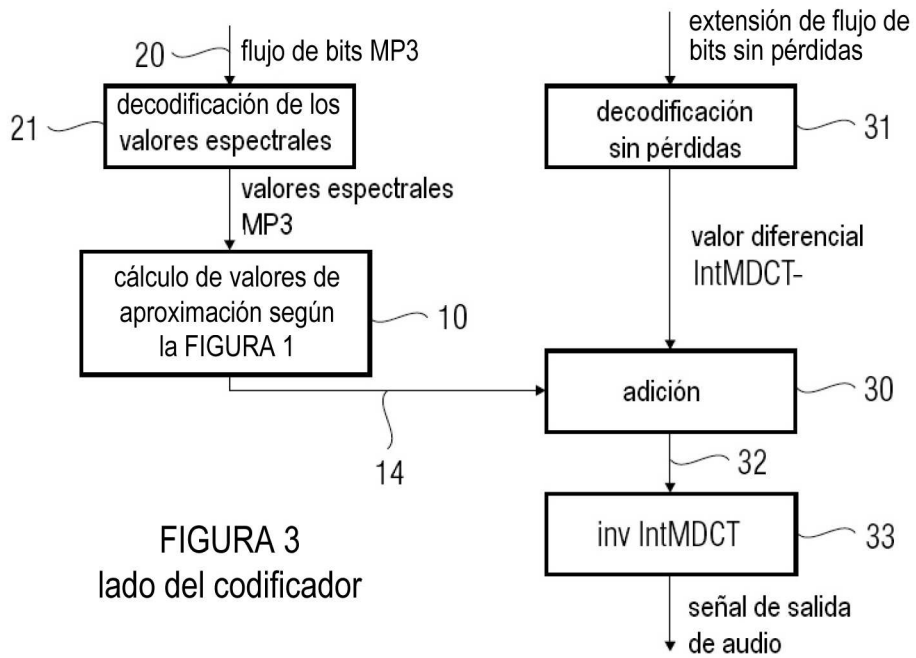
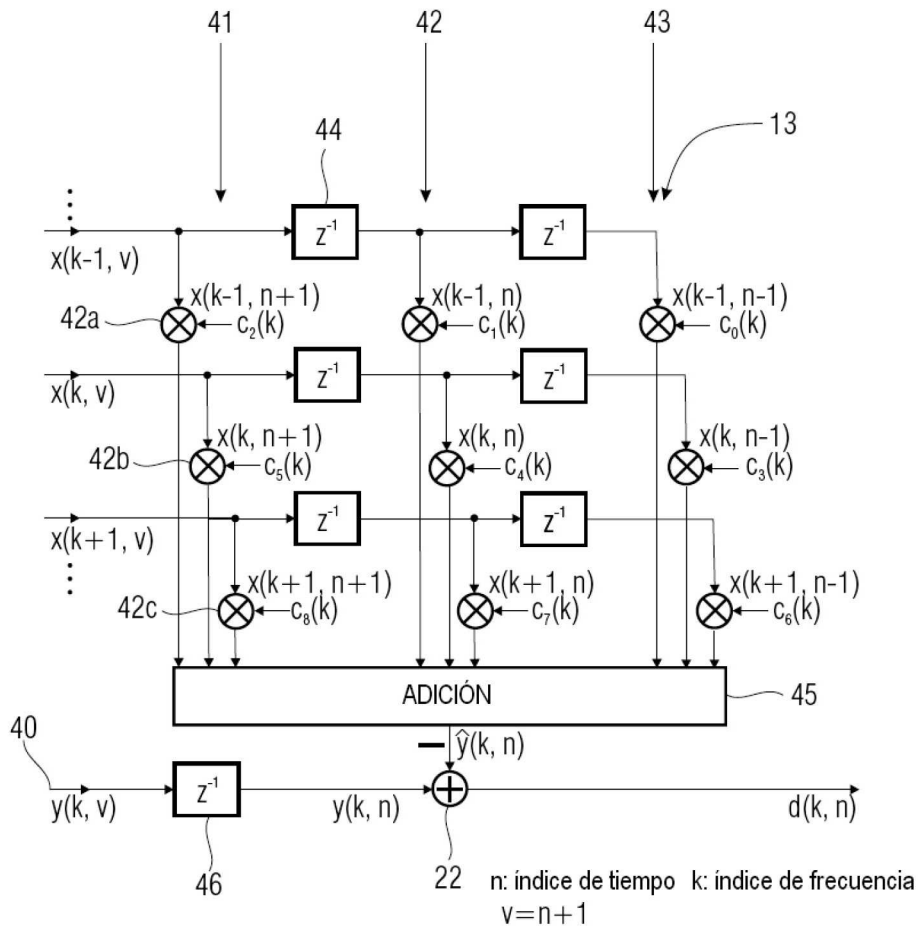


FIGURA 3
lado del codificador

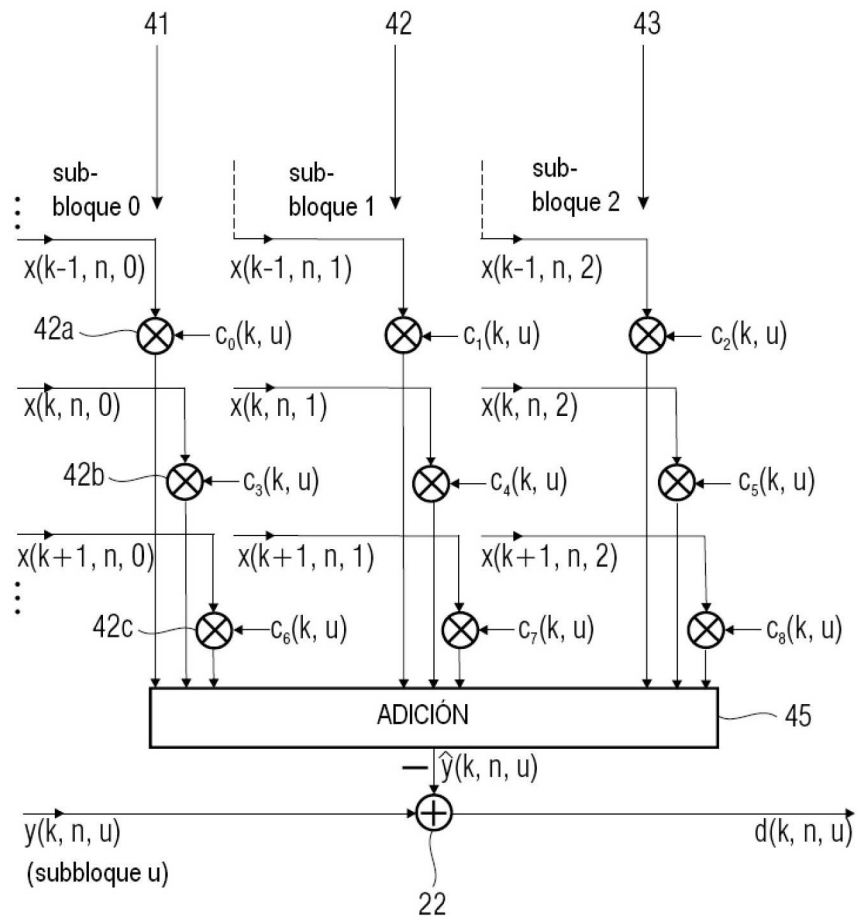


$$d(k, n) = y(k, n) - \hat{y}(k, n)$$

$$\hat{y}(k, n) = c_0(k)x(k-1, n-1) + c_1(k)x(k-1, n) + c_2(k)x(k-1, n+1) + c_3(k)x(k, n-1) + c_4(k)x(k, n) + c_5(k)x(k, n+1) + c_6(k)x(k+1, n-1) + c_7(k)x(k+1, n) + c_8(k)x(k+1, n+1)$$

en donde $k = 0 \dots 575$

FIGURA 4
BLOQUES LARGOS



n: índice de tiempo k: índice de frecuencia n: índice de subbloque

$$d(k, n, u) = y(k, n, u) - \hat{y}(k, n, u), \text{ en donde}$$

$$y(k, n, u) = c_0(k, u)x(k-1, n, 0) + c_1(k, u)x(k-1, n, 1) + c_2(k, u)x(k-1, n, 2)$$

$$+ c_3(k, u)x(k, n, 0) + c_4(k, u)x(k, n, 1) + c_5(k, u)x(k, n, 2)$$

$$+ c_6(k, u)x(k+1, n, 0) + c_7(k, u)x(k+1, n, 1) + c_8(k, u)x(k+1, n, 2)$$

en donde $k = 0 \dots 191, u = 0 \dots 2$

FIGURA 5A

BLOQUES CORTOS - 1. variante

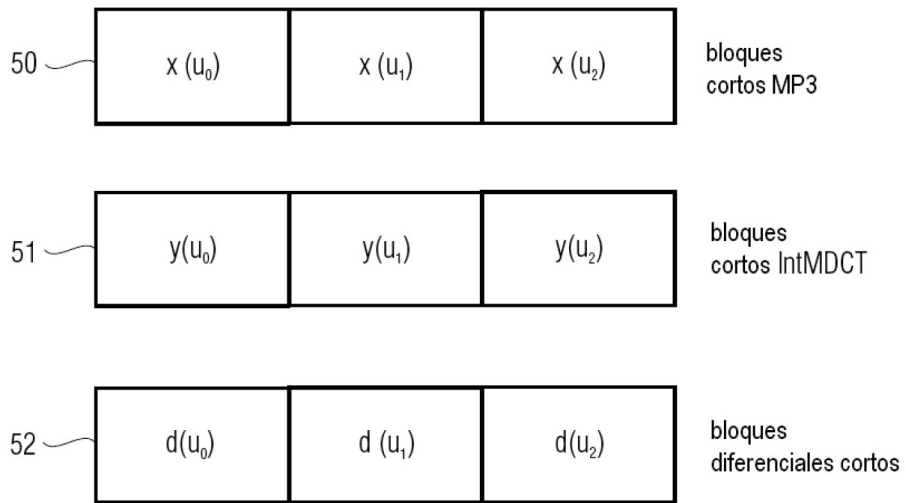


FIGURA 5B

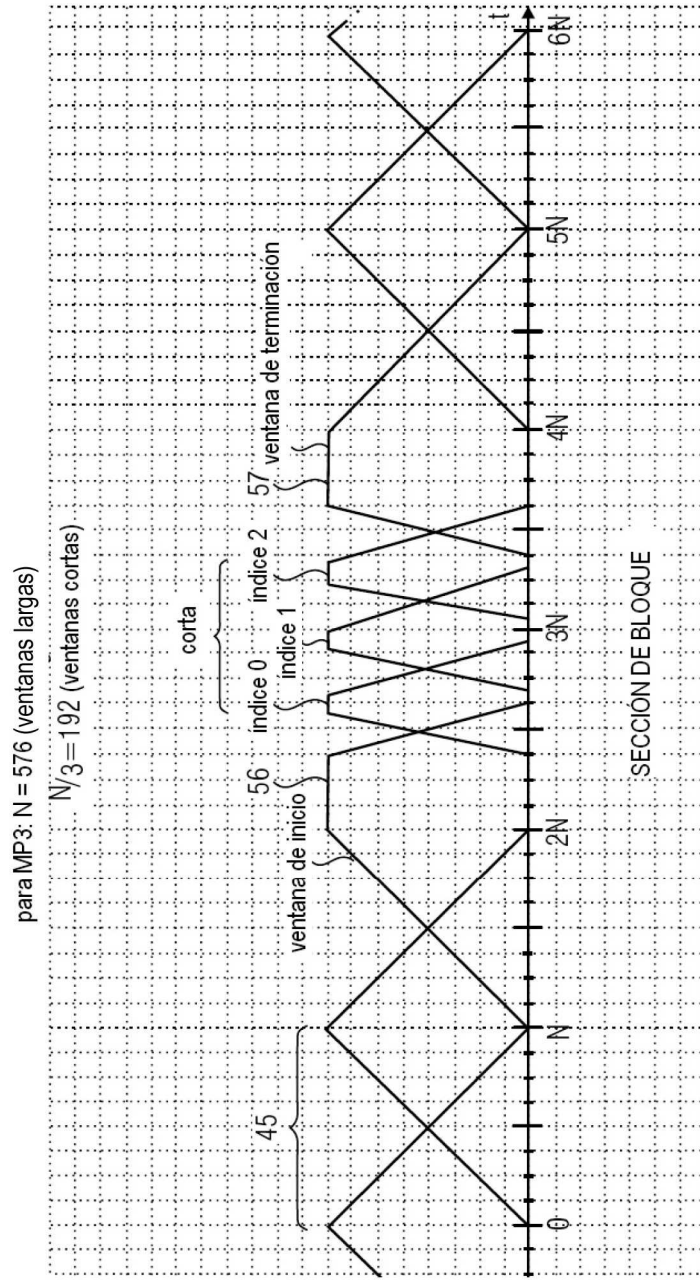
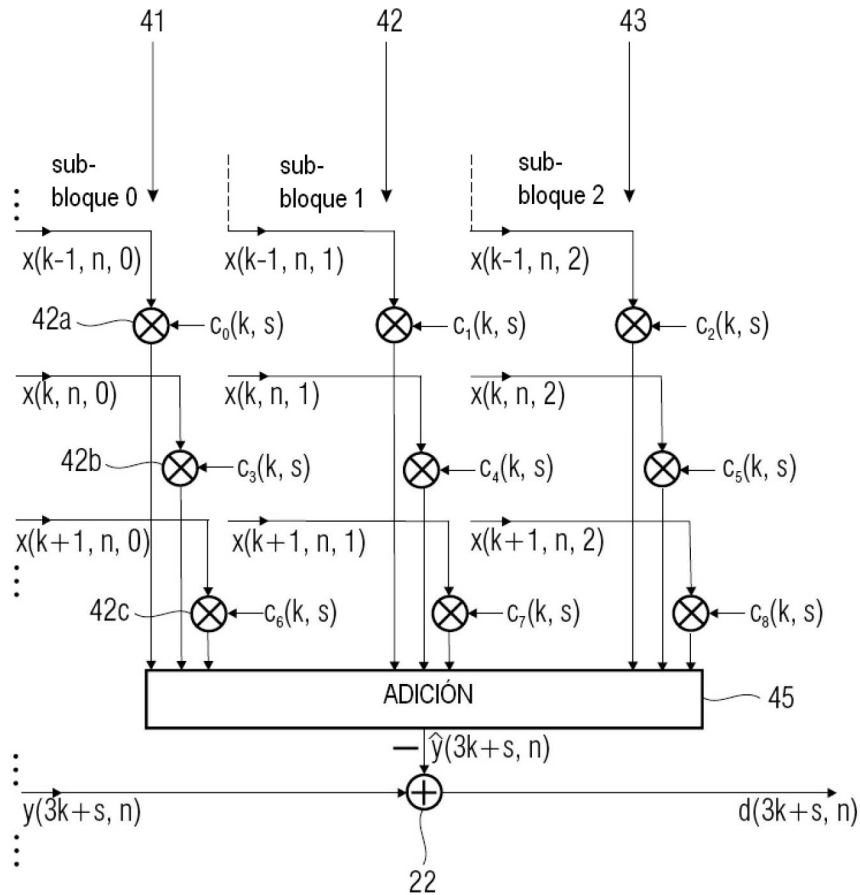


FIGURA 5C



n: índice de bloque (índice de tiempo) k: índice de frecuencia s: índice de orden

$$d(3k+s, n) = y(3k+s, n) - \hat{y}(3k+s, n)$$

$$\hat{y}(3k+s, n) = c_0(k, s)x(k-1, n, 0) + c_1(k, s)x(k-1, n, 1) + c_2(k, s)x(k-1, n, 2) + c_3(k, s)x(k, n, 0) + c_4(k, s)x(k, n, 1) + c_5(k, s)x(k, n, 2) + c_6(k, s)x(k+1, n, 0) + c_7(k, s)x(k+1, n, 1) + c_8(k, s)x(k+1, n, 2)$$

en donde $k = 0 \dots 191, s = 0 \dots 2$

FIGURA 6A

BLOQUES CORTOS - 2. variante

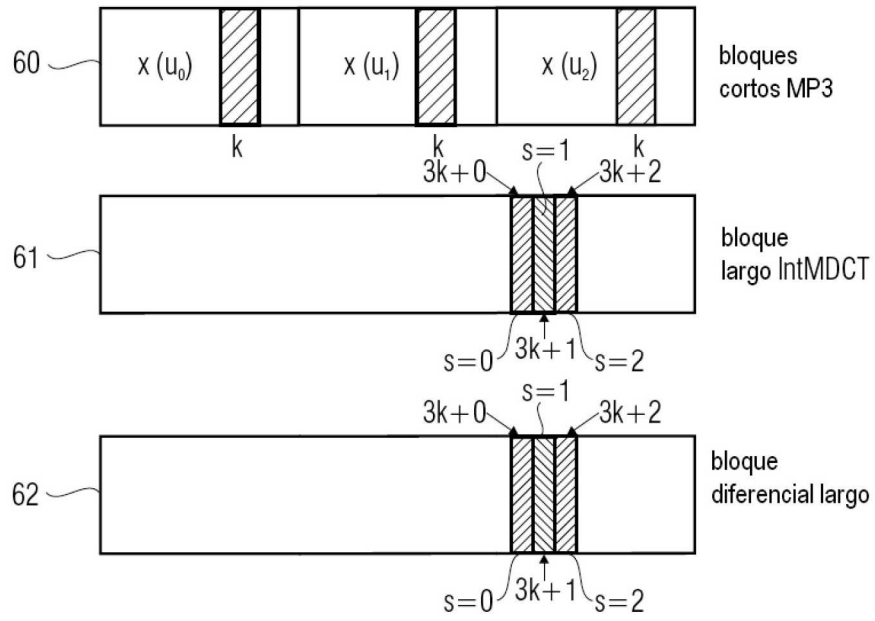


FIGURA 6B

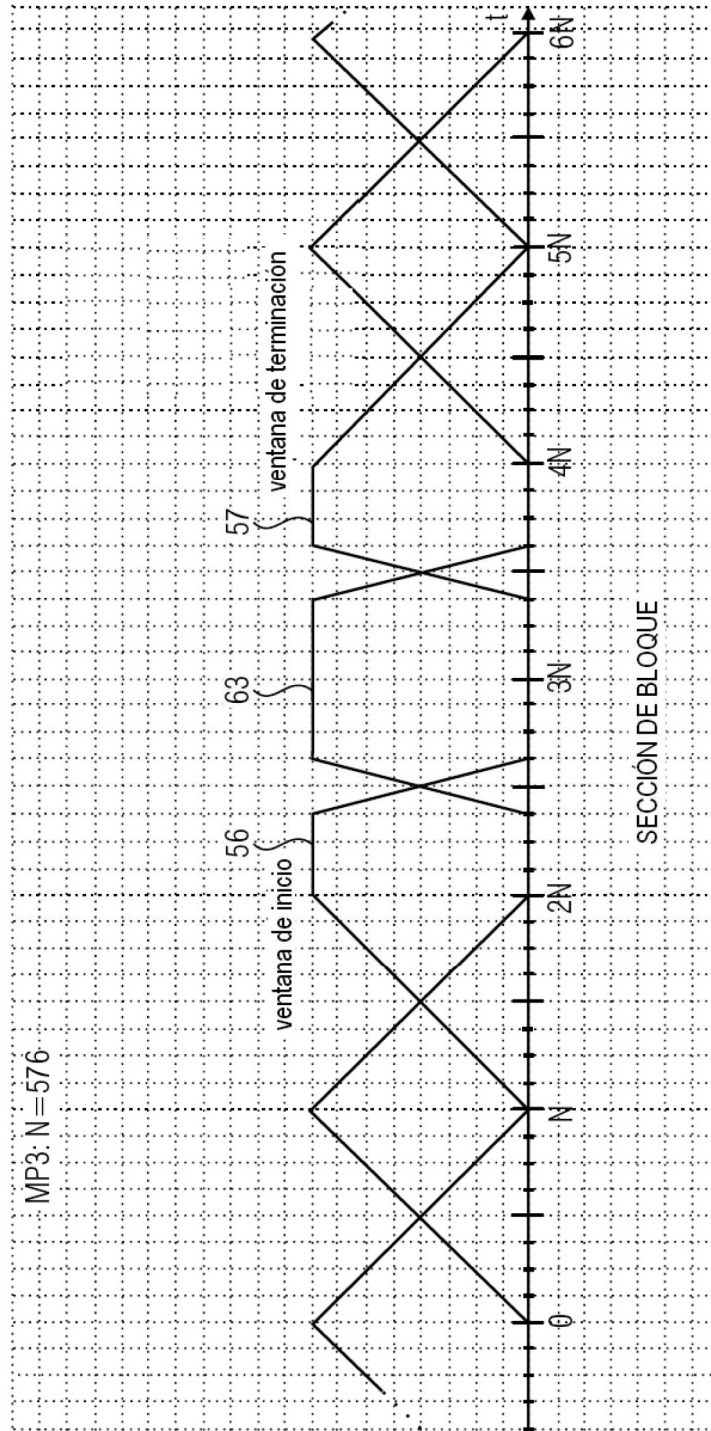


FIGURA 6C

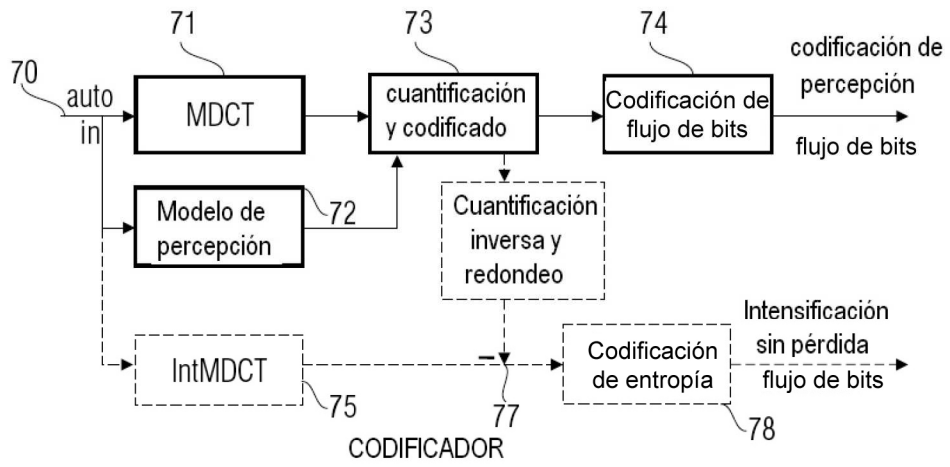


FIGURA 7
(TÉCNICA ANTERIOR)

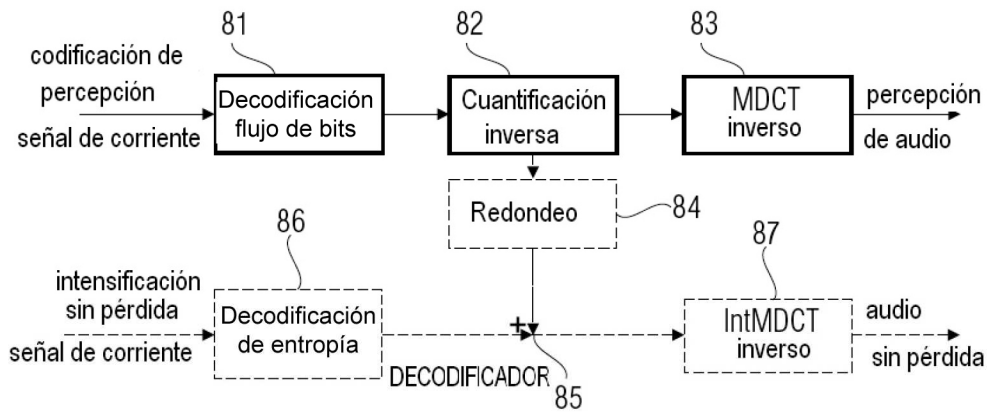


FIGURA 8
(TÉCNICA ANTERIOR)

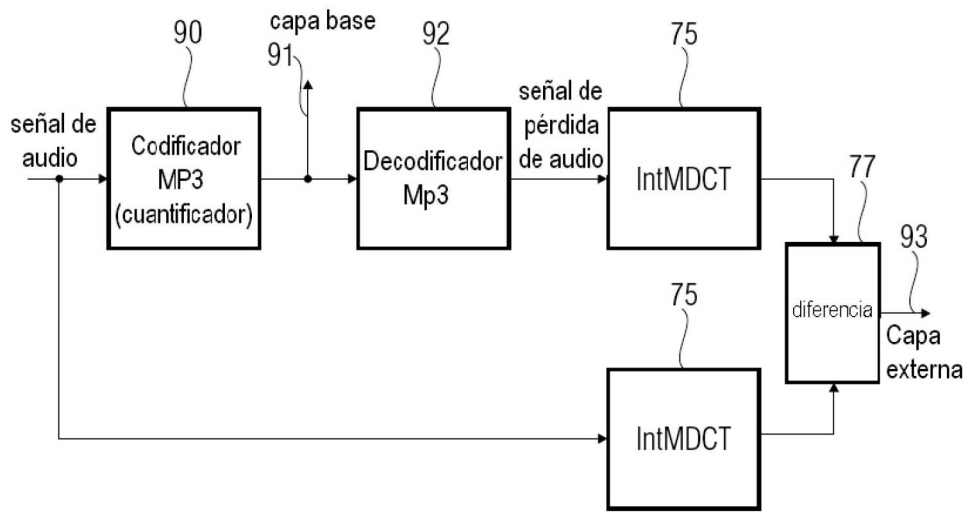


FIGURA 9
Codificador

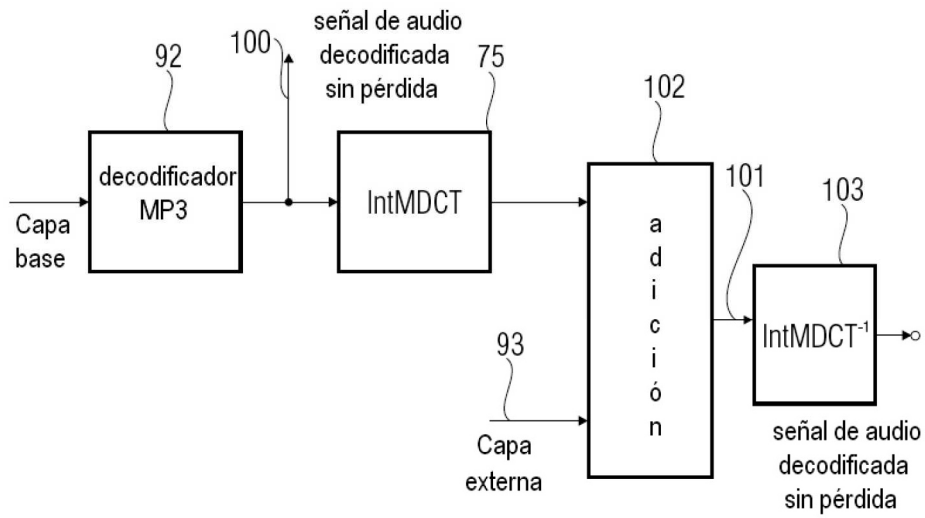


FIGURA 10
Decodificador