

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 467 290**

51 Int. Cl.:

G10L 19/008 (2013.01)

H04S 3/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.02.2011 E 11154910 (1)**

97 Fecha y número de publicación de la concesión europea: **09.04.2014 EP 2360683**

54 Título: **Descodificación de audio usando un mezclado descendente eficaz**

30 Prioridad:

18.02.2010 US 305871 P

29.06.2010 US 359763 P

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

12.06.2014

73 Titular/es:

DOLBY LABORATORIES LICENSING CORPORATION (50.0%)
100 Potrero Avenue
San Francisco, CA 94103-4813, US y
DOLBY INTERNATIONAL AB (50.0%)

72 Inventor/es:

THESING, ROBIN;
SILVA, JAMES M. y
ANDERSEN, ROBERT L.

74 Agente/Representante:

PÉREZ BARQUÍN, Eliana

ES 2 467 290 T3

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

DESCRIPCIÓN

Descodificación de audio usando un mezclado descendente eficaz

5 **Campo de la invención**

La presente divulgación se refiere en general al procesamiento de señales de audio.

10 **Antecedentes**

15 La compresión de datos de audio digital se ha convertido en una técnica importante en la industria del audio. Se han introducido nuevos formatos que permiten una reproducción de audio de alta calidad sin necesidad de un gran ancho de banda de datos, el cual es necesario en las técnicas tradicionales. El Comité de Sistemas de Televisión Avanzada (ATSC) ha adoptado la tecnología de codificación AC-3 y, más recientemente, la AC-3 mejorada (E-AC-3) como la norma de servicio de audio para la televisión de alta definición (HDTV) en los Estados Unidos. La norma E-AC-3 también se utiliza en los sistemas multimedia de los consumidores (disco de vídeo digital) y en la radiodifusión directa por satélite. La norma E-AC-3 es un ejemplo de codificación perceptual y permite la codificación de múltiples canales de audio digital en un flujo de bits de audio codificado y metadatos.

20 Existe la necesidad de descodificar de manera eficaz un flujo de bits de audio codificado. Por ejemplo, la vida útil de las baterías de los dispositivos portátiles está limitada en gran medida por el consumo de energía de su unidad de procesamiento principal. El consumo de energía de una unidad de procesamiento está estrechamente relacionado con la complejidad computacional de sus tareas. Por tanto, reducir la complejidad computacional media de un sistema de procesamiento de audio portátil aumentará la vida útil de la batería de tal sistema.

25 El término x86 es comúnmente utilizado por los expertos en la técnica para designar una familia de arquitecturas de conjunto de instrucciones de procesador cuyos orígenes se remontan al procesador 8086 de Intel. Como resultado de la ubicuidad de la arquitectura de conjunto de instrucciones x86, existe también la necesidad de descodificar de manera eficaz un flujo de bits de audio codificado en un procesador o sistema de procesamiento que tenga una arquitectura de conjunto de instrucciones x86. Muchas implementaciones de descodificadores tienen un carácter genérico, mientras que otras están diseñadas específicamente para procesadores integrados. Nuevos procesadores tales como el procesador Geode de AMD y el nuevo Intel Atom son ejemplos de diseños de 32 bits y de 64 bits que usan el conjunto de instrucciones x86 y que están usándose en pequeños dispositivos portátiles.

35 **Sumario**

Un aspecto de la invención se refiere a un método que hace funcionar un descodificador de audio para descodificar datos de audio según la reivindicación 1. Aspectos adicionales se definen de la manera expuesta en las reivindicaciones dependientes 2 a 12.

40 Un aspecto adicional de la invención se refiere a un medio de almacenamiento legible por ordenador que almacena instrucciones de descodificación que cuando son ejecutadas por uno o más procesadores de un sistema de procesamiento hacen que el sistema de procesamiento lleve a cabo el método descrito anteriormente.

45 Un aspecto adicional de la invención se refiere a un aparato que procesa datos de audio para descodificar datos de audio que incluyen bloques codificados de N.n canales de datos de audio para formar datos de audio descodificados que incluyen M.m canales de de audio descodificado, $M \geq 1$, siendo n el número de canales de efectos de baja frecuencia en los datos de audio codificados, y siendo m el número de canales de efectos de baja frecuencia en los datos de audio descodificados, comprendiendo el aparato medios para llevar a cabo el método anterior.

50 **Breve descripción de los dibujos**

La figura 1 muestra un pseudocódigo 100 de instrucciones que, cuando se ejecutan, llevan a cabo un proceso de descodificación AC-3 típico.

55 Las figuras 2A a 2D muestran, en forma de diagrama de bloques simplificado, algunas configuraciones de descodificador diferentes que pueden usar de manera ventajosa uno o más módulos comunes.

60 La figura 3 muestra un pseudocódigo y un diagrama de bloques simplificado de una realización de un módulo de descodificación de sección de entrada.

La figura 4 muestra un diagrama de flujo de datos simplificado del funcionamiento de una realización de un módulo de descodificación de sección de entrada.

65 La figura 5A muestra un pseudocódigo y un diagrama de bloques simplificado de una realización de un módulo de descodificación de sección de procesamiento.

La figura 5B muestra un pseudocódigo y un diagrama de bloques simplificado de otra realización de un módulo de descodificación de sección de procesamiento.

5 La figura 6 muestra un diagrama de flujo de datos simplificado del funcionamiento de una realización de un módulo de descodificación de sección de procesamiento.

La figura 7 muestra un diagrama de flujo de datos simplificado del funcionamiento de otra realización de un módulo de descodificación de sección de procesamiento.

10 La figura 8 muestra un diagrama de flujo de una realización de un procesamiento de un módulo de descodificación de sección de procesamiento tal como el mostrado en la figura 7.

15 La figura 9 muestra un ejemplo de procesamiento de cinco bloques que incluye un mezclado descendente de 5.1 a 2.0 usando una realización de la presente invención para el caso de una transformada de no solapamiento que incluye un mezclado descendente de 5.1 a 2.0.

20 La figura 10 muestra otro ejemplo de procesamiento de cinco bloques que incluye un mezclado descendente de 5.1 a 2.0 usando una realización de la presente invención para el caso de una transformada de solapamiento.

La figura 11 muestra un pseudocódigo simplificado de una realización de mezclado descendente en el dominio de tiempo.

25 La figura 12 muestra un diagrama de bloques simplificado de una realización de un sistema de procesamiento que incluye al menos un procesador y que puede llevar a cabo descodificaciones, que incluye una o más características de la presente invención.

Descripción de realizaciones de ejemplo

30 Aspectos generales

Realizaciones de la presente invención incluyen un método, un aparato y lógica codificada en uno o más medios tangibles legibles por ordenador para llevar a cabo acciones.

35 Realizaciones particulares incluyen un método que hace funcionar un descodificador de audio para descodificar datos de audio que incluyen bloques codificados de N.n canales de datos de audio para formar datos de audio descodificados que incluyen M.m canales de audio descodificado, $M \geq 1$, siendo n el número de canales de efectos de baja frecuencia en los datos de audio codificados, y siendo m el número de canales de efectos de baja frecuencia en los datos de audio descodificados. El método comprende aceptar los datos de audio que incluyen bloques de N.n canales de datos de audio codificados que han sido codificados mediante un método de codificación que incluye transformar N.n canales de datos de audio digital y formar y empaquetar datos de exponente y de mantisa de dominio de frecuencia; y descodificar los datos de audio aceptados. La descodificación incluye: desempaquetar y descodificar los datos de exponente y de mantisa de dominio de frecuencia; determinar coeficientes de transformada a partir de los datos de exponente y de mantisa de dominio de frecuencia desempaquetados y descodificados; 45 someter a una transformación inversa los datos de dominio de frecuencia y aplicar un procesamiento adicional para determinar datos de audio muestreados; y mezclar de manera descendente en el dominio de tiempo al menos algunos bloques de los datos de audio muestreados determinados según datos de mezclado descendente para el caso en que $M < N$. Al menos se cumple uno de A1, B1 y C1:

50 siendo A1 que la descodificación incluye determinar bloque a bloque si hay que aplicar un mezclado descendente en el dominio de frecuencia o un mezclado descendente en el dominio de tiempo, y si para un bloque particular se determina que hay que aplicar un mezclado descendente en el dominio de frecuencia, aplicar un mezclado descendente en el dominio de frecuencia para el bloque particular,

55 siendo B1 que el mezclado descendente en el dominio de tiempo incluye comprobar si los datos de mezclado descendente han cambiado con respecto a datos de mezclado descendente usados anteriormente, y, si han cambiado, aplicar una atenuación cruzada para determinar datos de mezclado descendente atenuados de manera cruzada y un mezclado descendente en el dominio de tiempo según los datos de mezclado descendente atenuados de manera cruzada, y, si no han cambiado, realizar directamente un mezclado descendente en el dominio de tiempo 60 según los datos de mezclado descendente, y

siendo C1 que el método incluye identificar uno o más canales no contribuyentes de los N.n canales de entrada, siendo un canal no contribuyente un canal que no contribuye en los M.m canales, y que el método no lleva a cabo una transformación inversa de los datos de dominio de frecuencia ni aplica un procesamiento adicional en el uno o 65 más canales no contribuyentes identificados.

Realizaciones particulares de la invención incluyen un medio de almacenamiento legible por ordenador que almacena instrucciones de descodificación que cuando son ejecutadas por uno o más procesadores de un sistema de procesamiento hacen que el sistema de procesamiento lleve a cabo la descodificación de datos de audio que incluyen bloques codificados de N.n canales de datos de audio para formar datos de audio descodificados que incluyen M.m canales de audio descodificado, $M \geq 1$, siendo n el número de canales de efectos de baja frecuencia en los datos de audio codificados, y siendo m el número de canales de efectos de baja frecuencia en los datos de audio descodificados. Las instrucciones de descodificación incluyen: instrucciones que cuando se ejecutan hacen que se acepten los datos de audio que incluyen bloques de N.n canales de datos de audio codificados que han sido codificados mediante un método de codificación, incluyendo el método de codificación transformar N.n canales de datos de audio digital y formar y empaquetar datos de exponente y de mantisa de dominio de frecuencia; e instrucciones que cuando se ejecutan hacen que se descodifiquen los datos de audio aceptados. Las instrucciones que cuando se ejecutan provocan la descodificación, incluyen: instrucciones que cuando se ejecutan hacen que se desempaqueten y se descodifiquen los datos de exponente y de mantisa de dominio de frecuencia; instrucciones que cuando se ejecutan hacen que se determinen coeficientes de transformada a partir de los datos de exponente y de mantisa de dominio de frecuencia desempaquetados y descodificados; instrucciones que cuando se ejecutan provocan una transformación inversa de los datos de dominio de frecuencia y la aplicación de un procesamiento adicional para determinar datos de audio muestreados; e instrucciones que cuando se ejecutan hacen que se determine si $M < N$; e instrucciones que cuando se ejecutan hacen que se mezclen de manera descendente en el dominio de tiempo al menos algunos bloques de los datos de audio muestreados determinados según datos de mezclado descendente si $M < N$. Al menos se cumple uno de A2, B2 y C2:

siendo A2 que las instrucciones que cuando se ejecutan provocan la descodificación incluyen instrucciones que cuando se ejecutan hacen que se determine bloque a bloque si hay que aplicar un mezclado descendente en el dominio de frecuencia o un mezclado descendente en el dominio de tiempo, e instrucciones que cuando se ejecutan hacen que se aplique un mezclado descendente en el dominio de frecuencia si para un bloque particular se determina que hay que aplicar un mezclado descendente en el dominio de frecuencia,

siendo B2 que el mezclado descendente en el dominio de tiempo incluye comprobar si los datos de mezclado descendente han cambiado con respecto a datos de mezclado descendente usados anteriormente, y, si han cambiado, aplicar una atenuación cruzada para determinar datos de mezclado descendente atenuados de manera cruzada y un mezclado descendente en el dominio de tiempo según los datos de mezclado descendente atenuados de manera cruzada, y, si no han cambiado, realizar directamente un mezclado descendente en el dominio de tiempo según los datos de mezclado descendente, y

siendo C2 que las instrucciones que cuando se ejecutan provocan la descodificación incluyen identificar uno o más canales no contribuyentes de los N.n canales de entrada, siendo un canal no contribuyente un canal que no contribuye en los M.m canales, y que el método no lleva a cabo una transformación inversa de los datos de dominio de frecuencia ni la aplicación de un procesamiento adicional en el uno o más canales no contribuyentes identificados.

Realizaciones particulares incluyen un aparato que procesa datos de audio para descodificar datos de audio que incluyen bloques codificados de N.n canales de datos de audio para formar datos de audio descodificados que incluyen M.m canales de audio descodificado, $M \geq 1$, siendo n el número de canales de efectos de baja frecuencia en los datos de audio codificados, y siendo m el número de canales de efectos de baja frecuencia en los datos de audio descodificados. El aparato comprende: medios para aceptar los datos de audio que incluyen bloques de N.n canales de datos de audio codificados que han sido codificados mediante un método de codificación, incluyendo el método de codificación transformar N.n canales de datos de audio digital y formar y empaquetar datos de exponente y de mantisa de dominio de frecuencia; y medios para descodificar los datos de audio aceptados. Los medios de descodificación incluyen: medios para desempaquetar y descodificar los datos de exponente y de mantisa de dominio de frecuencia; medios para determinar coeficientes de transformada a partir de los datos de exponente y de mantisa de dominio de frecuencia desempaquetados y descodificados; medios para someter a una transformación inversa los datos de dominio de frecuencia y para aplicar un procesamiento adicional para determinar datos de audio muestreados; y medios para mezclar de manera descendente en el dominio de tiempo al menos algunos bloques de los datos de audio muestreados determinados según datos de mezclado descendente para el caso en que $M < N$. Al menos se cumple uno de A3, B3 y C3:

siendo A3 que los medios de descodificación incluyen medios para determinar bloque a bloque si hay que aplicar un mezclado descendente en el dominio de frecuencia o un mezclado descendente en el dominio de tiempo, y medios para aplicar un mezclado descendente en el dominio de frecuencia, donde los medios de aplicación de un mezclado descendente en el dominio de frecuencia aplican un mezclado descendente en el dominio de frecuencia para el bloque particular si para un bloque particular se determina que hay que aplicar un mezclado descendente en el dominio de frecuencia,

siendo B3 que los medios de mezclado descendente en el dominio de tiempo comprueban si los datos de mezclado descendente han cambiado con respecto a datos de mezclado descendente usados anteriormente, y, si han cambiado, aplican una atenuación cruzada para determinar datos de mezclado descendente atenuados de manera

cruzada y un mezclado descendente en el dominio de tiempo según los datos de mezclado descendente atenuados de manera cruzada, y, si no han cambiado, aplican directamente un mezclado descendente en el dominio de tiempo según los datos de mezclado descendente, y

5 siendo C3 que el aparato incluye medios para identificar uno o más canales no contribuyentes de los N.n canales de entrada, siendo un canal no contribuyente un canal que no contribuye en los M.m canales, y que el aparato no lleva a cabo una transformación inversa de los datos de dominio de frecuencia ni aplica un procesamiento adicional en el uno o más canales no contribuyentes identificados.

10 Realizaciones particulares incluyen un aparato de procesamiento de datos de audio que incluyen N.n canales de datos de audio codificados para formar datos de audio descodificados que incluyen M.m canales de audio descodificado, $M \geq 1$, siendo $n=0$ ó 1 el número de canales de efectos de baja frecuencia en los datos de audio codificados, y siendo $m=0$ ó 1 el número de canales de efectos de baja frecuencia en los datos de audio descodificados. El aparato comprende: medios para aceptar los datos de audio que incluyen N.n canales de datos de audio codificados que han sido codificados mediante un método de codificación, comprendiendo el método de codificación transformar N.n canales de datos de audio digital de manera que una transformación inversa y un procesamiento adicional puedan recuperar muestras de dominio de tiempo sin errores de solapamiento, formar y empaquetar datos de exponente y de mantisa de dominio de frecuencia y formar y empaquetar metadatos relacionados con los datos de exponente y de mantisa de dominio de frecuencia, donde los metadatos incluyen opcionalmente metadatos relacionados con un procesamiento de pre-ruido transitorio; y medios para descodificar los datos de audio aceptados. Los medios de descodificación comprenden: uno o más medios de descodificación de sección de entrada y uno o más medios de descodificación de sección de procesamiento. Los medios de descodificación de sección de entrada incluyen medios para desempaquetar los metadatos, para desempaquetar y para descodificar los datos de exponente y de mantisa de dominio de frecuencia. Los medios de descodificación de sección de procesamiento incluyen medios para determinar coeficientes de transformada a partir de los datos de exponente y de mantisa de dominio de frecuencia desempaquetados y descodificados; para someter a una transformación inversa los datos de dominio de frecuencia; para aplicar operaciones de división en ventanas y de solapamiento y suma para determinar datos de audio muestreados; para aplicar cualquier procesamiento de pre-ruido transitorio requerido de descodificación según los metadatos relacionados con el procesamiento de pre-ruido transitorio; y para realizar un mezclado descendente en el dominio de tiempo según datos de mezclado descendente, estando configurado el mezclado descendente para mezclar de manera descendente en el dominio de tiempo al menos algunos bloques de datos según datos de mezclado descendente en el caso en que $M < N$. Al menos se cumple uno de A4, B4 y C4:

35 siendo A4 que los medios de descodificación de sección de procesamiento incluyen medios para determinar bloque a bloque si hay que aplicar un mezclado descendente en el dominio de frecuencia o un mezclado descendente en el dominio de tiempo, y medios para aplicar un mezclado descendente en el dominio de frecuencia, donde los medios de aplicación de mezclado descendente en el dominio de frecuencia aplican un mezclado descendente en el dominio de frecuencia para el bloque particular si para un bloque particular se determina que hay que aplicar un mezclado descendente en el dominio de frecuencia,

45 siendo B4 que los medios de mezclado descendente en el dominio de tiempo comprueban si los datos de mezclado descendente han cambiado con respecto a datos de mezclado descendente usados anteriormente, y, si han cambiado, aplican una atenuación cruzada para determinar datos de mezclado descendente atenuados de manera cruzada y un mezclado descendente en el dominio de tiempo según los datos de mezclado descendente atenuados de manera cruzada, y, si no han cambiado, aplican directamente un mezclado descendente en el dominio de tiempo según los datos de mezclado descendente, y

50 siendo C4 que el aparato incluye medios para identificar uno o más canales no contribuyentes de los N.n canales de entrada, siendo un canal no contribuyente un canal que no contribuye en los M.m canales, y que los medios de descodificación de sección de procesamiento no llevan a cabo una transformación inversa de los datos de dominio de frecuencia ni aplican un procesamiento adicional en el uno o más canales no contribuyentes identificados.

55 Realizaciones particulares incluyen un sistema de descodificación de datos de audio que incluyen N.n canales de datos de audio codificados para formar datos de audio descodificados que incluyen M.m canales de audio descodificado, $M \geq 1$, siendo n el número de canales de efectos de baja frecuencia en los datos de audio codificados, y siendo m el número de canales de efectos de baja frecuencia en los datos de audio descodificados. El sistema comprende: uno o más procesadores; y un subsistema de almacenamiento acoplado al uno o más procesadores. El sistema acepta los datos de audio que incluyen bloques de N.n canales de datos de audio codificados que han sido codificados mediante un método de codificación, incluyendo el método de codificación transformar N.n canales de datos de audio digital, y formar y empaquetar datos de exponente y de mantisa de dominio de frecuencia; asimismo descodifica los datos de audio aceptados, lo que incluye: desempaquetar y descodificar los datos de exponente y de mantisa de dominio de frecuencia; determinar coeficientes de transformada a partir de los datos de exponente y de mantisa de dominio de frecuencia desempaquetados y descodificados; someter a una transformación inversa los datos de dominio de frecuencia y aplicar un procesamiento adicional para determinar datos de audio muestreados; y mezclar de manera descendente en el dominio de tiempo al menos algunos bloques de los datos de audio

muestreados determinados según datos de mezclado descendente para el caso en que $M < N$. Al menos se cumple uno de A5, B5 y C5:

5 siendo A5 que la descodificación incluye determinar bloque a bloque si hay que aplicar un mezclado descendente en el dominio de frecuencia o un mezclado descendente en el dominio de tiempo, y si para un bloque particular se determina que hay que aplicar un mezclado descendente en el dominio de frecuencia, aplicar un mezclado descendente en el dominio de frecuencia para el bloque particular,

10 siendo B5 que el mezclado descendente en el dominio de tiempo incluye comprobar si los datos de mezclado descendente han cambiado con respecto a datos de mezclado descendente usados anteriormente, y, si han cambiado, aplicar una atenuación cruzada para determinar datos de mezclado descendente atenuados de manera cruzada y un mezclado descendente en el dominio de tiempo según los datos de mezclado descendente atenuados de manera cruzada, y, si no han cambiado, realizar directamente un mezclado descendente en el dominio de tiempo según los datos de mezclado descendente, y

15 siendo C5 que el método incluye identificar uno o más canales no contribuyentes de los $N.n$ canales de entrada, siendo un canal no contribuyente un canal que no contribuye en los $M.m$ canales, y que el método no lleva a cabo una transformación inversa de los datos de dominio de frecuencia ni aplica un procesamiento adicional en el uno o más canales no contribuyentes identificados.

20 En algunas versiones de la realización de sistema, los datos de audio aceptados están en forma de un flujo de bits de tramas de datos codificados, y el subsistema de almacenamiento está configurado con instrucciones que cuando son ejecutadas por uno o más de los procesadores del sistema de procesamiento hacen que se descodifiquen los datos de audio aceptados.

25 Algunas versiones de la realización de sistema incluyen uno o más subsistemas que están interconectados a través de un enlace de red, incluyendo cada subsistema al menos un procesador.

30 En algunas realizaciones en las que se cumple A1, A2, A3, A4 o A5, determinar si hay que aplicar un mezclado descendente en el dominio de frecuencia o un mezclado descendente en el dominio de tiempo incluye determinar si hay un procesamiento de pre-ruido transitorio y determinar si alguno de los N canales tiene un tipo de bloque diferente, de manera que el mezclado descendente en el dominio de frecuencia solo se aplica en un bloque que tiene el mismo tipo de bloque en los N canales, si no hay ningún procesamiento de pre-ruido transitorio y si $M < N$.

35 En algunas realizaciones en las que se cumple A1, A2, A3, A4 o A5 y en las que la transformación en el método de codificación usa una transformada solapada y el procesamiento adicional incluye aplicar operaciones de división en ventanas y de solapamiento y suma para determinar datos de audio muestreados, (i) la aplicación de un mezclado descendente en el dominio de frecuencia para el bloque particular incluye determinar si el mezclado descendente para el bloque anterior se realizó mediante un mezclado descendente en el dominio de tiempo, y si el mezclado descendente del bloque anterior se realizó mediante un mezclado descendente en el dominio de tiempo, aplicar un mezclado descendente en el dominio de tiempo (o un mezclado descendente en un dominio de pseudo-tiempo) en los datos del bloque anterior que van a solaparse con los datos descodificados del bloque particular, y (ii) la aplicación de un mezclado descendente en el dominio de tiempo para un bloque particular incluye determinar si el mezclado descendente del bloque anterior se realizó mediante un mezclado descendente en el dominio de frecuencia, y si el mezclado descendente del bloque anterior se realizó mediante un mezclado descendente en el dominio de frecuencia, procesar el bloque particular de diferente manera a que si el mezclado descendente del bloque anterior no se hubiera realizado mediante un mezclado descendente en el dominio de frecuencia.

50 En algunas realizaciones en las que se cumple B1, B2, B3, B4 o B5, se usa al menos un procesador x86, cuyo conjunto de instrucciones incluye difundir en flujo continuo extensiones (SSE) de una sola instrucción y múltiples datos que comprenden instrucciones vectoriales, y el mezclado descendente en el dominio de tiempo incluye ejecutar instrucciones vectoriales en al menos un procesador del uno o más procesadores x86.

55 En algunas realizaciones en las que se cumple C1, C2, C3, C4 o C5, $n=1$ y $m=0$, de manera que la transformación inversa y la aplicación de un procesamiento adicional no se llevan a cabo en el canal de efecto de baja frecuencia. Además, en algunas realizaciones en las que se cumple C, los datos de audio que incluyen bloques codificados incluyen información que define el mezclado descendente, y donde la identificación de uno o más canales no contribuyentes usa la información que define el mezclado descendente. Además, en algunas realizaciones en las que se cumple C, la identificación de uno o más canales no contribuyentes incluye además identificar si uno o más canales tienen una cantidad insignificante de contenido con respecto a uno o más otros canales, donde un canal tiene una cantidad insignificante de contenido con respecto a otro canal si su energía o nivel absoluto es al menos 15 dB inferior a los del otro canal. En algunos casos, un canal tiene una cantidad insignificante de contenido con respecto a otro canal si su energía o nivel absoluto es al menos 18 dB inferior a los del otro canal, mientras que en otras aplicaciones un canal tiene una cantidad insignificante de contenido con respecto a otro canal si su energía o nivel absoluto es al menos 25 dB inferior a los del otro canal.

En algunas realizaciones, los datos de audio codificados se codifican según una norma del conjunto de normas que consiste en la norma AC-3, la norma E-AC-3, una norma compatible con versiones anteriores de la norma E-AC-3, la norma MPEG-2 AAC y la norma HE-AAC.

5 En algunas realizaciones de la invención, la transformación en el método de codificación usa una transformada solapada, y el procesamiento adicional incluye aplicar operaciones de división en ventanas y de solapamiento y suma para determinar datos de audio muestreados.

10 En algunas realizaciones de la invención, el método de codificación incluye formar y empaquetar metadatos relacionados con los datos de exponente y de mantisa en el dominio de frecuencia, donde los metadatos incluyen opcionalmente metadatos relacionados con un procesamiento de pre-ruido transitorio y un mezclado descendente.

15 Realizaciones particulares pueden proporcionar algunos, todos o ninguno de estos aspectos, características o ventajas. Realizaciones particulares pueden proporcionar uno o más otros aspectos, características o ventajas, donde uno o más de los cuales pueden resultar fácilmente evidentes a un experto en la técnica a partir de las figuras, descripciones y reivindicaciones de este documento.

Descodificación de un flujo codificado

20 Se describen realizaciones de la presente invención para descodificar audio que ha sido codificado según la norma AC-3 extendida (E-AC-3) en un flujo de bits codificado. La norma E-AC-3 y al versión anterior, AC-3, se describen en detalle en el documento "Digital Audio Compression Standard (AC-3, E-AC-3)", Revisión B, Documento A/52B, del Comité de Sistemas de Televisión Avanzada, Inc., (ATSC), 14 de junio de 2005, obtenido el 1 de diciembre de 2009 en el sitio de Internet www.atsc.org/standards/a_52b.pdf, (donde `^dot^` denota el periodo (".") en la dirección web real). Sin embargo, la invención no está limitada a descodificar un flujo de bits codificado en E-AC-3, y puede aplicarse a un descodificador que descodifica un flujo de bits codificado según otro método de codificación, y a métodos de tal descodificación, aparatos de descodificación, sistemas que llevan a cabo tal descodificación, a software que cuando se ejecuta hace que uno o más procesadores lleven a cabo tal descodificación y/o a medios de almacenamiento tangibles en los que está almacenado tal software. Por ejemplo, realizaciones de la presente invención también pueden aplicarse para descodificar audio que se ha codificado según las normas MPEG-2 AAC (ISO/IEC 13818-7) y MPEG-4 Audio (ISO/IEC 14496-3). La norma MPEG-4 Audio incluye la codificación AAC de alta eficacia (versión 1) (HE-AAC v1) y la codificación AAC de alta eficacia (versión 2) (HE-AAC v2), denominadas conjuntamente como HE-AAC en este documento.

35 Las normas AC-3 y E-AC-3 también se conocen como DOLBY[®] DIGITAL y DOLBY[®] DIGITAL PLUS. Una versión de HE-AAC que incorpora algunas mejoras adicionales, compatibles también se conoce como DOLBY[®] PULSE. Son marcas registradas de Dolby Laboratories Licensing Corporation, el cesionario de la presente invención, y pueden estar registradas en una o más jurisdicciones. La norma E-AC-3 es compatible con la norma AC-3 e incluye funcionalidad adicional.

40 La arquitectura x86

El término x86 es comúnmente utilizado por los expertos en la técnica para designar una familia de arquitecturas de conjunto de instrucciones de procesador cuyos orígenes se remontan al procesador 8086 de Intel. La arquitectura se ha implementado en procesadores de compañías tales como Intel, Cyrix, AMD, VIA y muchas otras. En general se entiende que el término implica una compatibilidad binaria con el conjunto de instrucciones de 32 bits del procesador 80386 de Intel. En la actualidad (principios de 2010), la arquitectura x86 se utiliza ampliamente en los ordenadores de escritorio y los ordenadores de tamaño agenda, así como en un creciente número de servidores y estaciones de trabajo. Una gran cantidad de software soporta la plataforma, incluyendo sistemas operativos tales como MS-DOS, Windows, Linux, BSD, Solaris y Mac OS X.

55 Tal y como se usa en este documento, el término x86 se refiere a una arquitectura de conjunto de instrucciones de procesador x86 que también soporta una extensión de conjunto de instrucciones (SSE) de una sola instrucción y múltiples datos (SIMD). SSE es una extensión de conjunto de instrucciones de una sola instrucción y múltiples datos (SIMD) de la arquitectura x86 original introducida en 1999 en los procesadores de la serie Pentium III de Intel y ahora común en arquitecturas x86 fabricadas por muchos vendedores.

Flujos de bits AC-3 y E-AC-3

60 Un flujo de bits AC-3 de una señal de audio multicanal está compuesto por tramas que representan un intervalo de tiempo constante de 1536 muestras moduladas por código de pulso (PCM) de la señal de audio a través de todos los canales codificados. Se proporciona un máximo de cinco canales principales y, opcionalmente, un canal de efectos de baja frecuencia (LFE) denotado como ".1", es decir, se proporciona un máximo de 5.1 canales de audio. Cada trama tiene un tamaño fijo que depende solamente de la frecuencia de muestreo y de la velocidad de los datos codificados.

- 5 En pocas palabras, la codificación AC-3 incluye usar una transformada solapada (la transformada discreta del coseno modificada (MDCT) con una ventana derivada de Kaiser Bessel (KBD) con un solapamiento del 50%) para convertir datos de tiempo en datos de frecuencia. Los datos de frecuencia se codifican perceptualmente para comprimir los datos para formar un flujo de bits comprimido de tramas que incluyen cada una datos de audio codificados y metadatos. Cada trama AC-3 es una entidad independiente que no comparte datos con tramas anteriores diferente a la transformada solapada intrínseca en la MDCT usada para convertir datos de tiempo en datos de frecuencia.
- 10 Al principio de cada trama AC-3 están los campos SI (información de sincronización) y BSI (información de flujo de bits). Los campos SI y BSI describen la configuración del flujo de bits, incluyendo la frecuencia de muestreo, la velocidad de datos, el número de canales codificados y otros diversos elementos a nivel de sistemas. También hay dos palabras de CRC (código de redundancia cíclica) por trama, una al principio y otra al final, que proporcionan un medio de detección de errores.
- 15 Dentro de cada trama hay seis bloques de audio, donde cada uno representa 256 muestras PCM por canal codificado de datos de audio. El bloque de audio contiene banderas de conmutación de bloque, coordenadas de acoplamiento, exponentes, parámetros de asignación de bits y mantisas. Los datos de una trama pueden compartirse, de manera que la información presente en el Bloque 0 puede volver a usarse en bloques siguientes.
- 20 Un campo de datos auxiliar opcional está ubicado al final de la trama. Este campo permite a los diseñadores de sistemas introducir información privada de control o de estado en el flujo de bits AC-3 para su transmisión a través del sistema.
- 25 La norma E-AC-3 conserva la estructura de trama AC-3 de seis transformadas de 256 coeficientes, permitiendo al mismo tiempo tramas más pequeñas formadas por uno, dos y tres bloques de transformadas de 256 coeficientes. Esto permite el transporte de audio a velocidades de datos superiores a 640 kbps. Cada trama E-AC-3 incluye metadatos y datos de audio.
- 30 La norma E-AC-3 permite un número mucho mayor de canales que los 5.1 de AC-3; en particular, la norma E-AC-3 puede usar los 6.1 y 7.1 canales de audio de hoy día y al menos 13.1 canales para soportar, por ejemplo, futuras pistas de sonido de audio multicanal. Los canales adicionales a los 5.1 canales se obtienen asociando el flujo de bits de programa de audio principal con un máximo de ocho subflujos dependientes adicionales, los cuales se multiplexan en un flujo de bits E-AC-3. Esto permite que el programa de audio principal pueda usar el formato de 5.1 canales de AC-3, mientras que la capacidad de canal adicional proviene de los flujos de bits dependientes. Esto significa que una versión de 5.1 canales y los diversos mezclados descendentes convencionales estén siempre disponibles y que los artefactos de codificación inducidos por sustracciones matriciales se eliminen mediante el uso de un proceso de sustitución de canal.
- 35 El poder soportar múltiples programas se debe también a la capacidad de proporcionar siete flujos de audio más independientes, cada uno con posibles subflujos dependientes asociados, para aumentar el número de canales soportado por cada programa más allá de los 5.1 canales.
- 40 La norma AC-3 usa una transformada relativamente corta y una cuantificación escalar simple para codificar perceptualmente el material de audio. Aunque la norma E-AC-3 es compatible con la AC-3, proporciona una mayor resolución espectral, una cuantificación mejorada y una codificación mejorada. Con la norma E-AC-3, la eficacia de codificación ha aumentado con respecto a la de la norma AC-3 para permitir el uso beneficioso de velocidades de datos más bajas. Esto se consigue usando un banco de filtros mejorado para convertir datos de tiempo en datos de dominio de frecuencia, una cuantificación mejorada, un acoplamiento de canal mejorado, una extensión espectral y una técnica denominada procesamiento de pre-ruido transitorio (TPNP).
- 45 Además de la transformada solapada MDCT para convertir datos de tiempo en datos de frecuencia, la norma E-AC-3 usa una transformada híbrida adaptativa (AHT) para señales de audio estacionarias. La AHT incluye la MDCT con la ventana derivada de Kaiser Bessel (KBD) seguida de, para señales estacionarias, una transformada de bloque secundaria en forma de transformada discreta del coseno (DCT) de tipo II no dividida en ventanas y no solapada. Por tanto, la AHT añade una DCT de segunda fase después del banco de filtros AC-3 MDCT/KBD existente cuando hay audio con características estacionarias para convertir los seis bloques de transformada de 256 coeficientes en un único bloque de transformada híbrida de 1536 coeficientes con una mayor resolución de frecuencia. Esta mayor resolución de frecuencia se combina con una cuantificación vectorial (VQ) de 6 dimensiones y una cuantificación adaptativa de ganancia (GAQ) para mejorar la eficacia de codificación para algunas señales, por ejemplo, señales "difíciles de codificar". La VQ se usa para codificar de manera eficaz bandas de frecuencia que necesitan una menor precisión, mientras que la GAQ proporciona una mayor eficacia cuando se necesita una cuantificación con una mayor precisión.
- 50 También se obtiene una mayor eficacia de codificación mediante el uso de un acoplamiento de canal con conservación de fase. Este método amplía el método de acoplamiento de canal de la norma AC-3, que usa un canal mono compuesto de alta frecuencia que reconstituye la parte de alta frecuencia de cada canal en la decodificación.
- 55
- 60
- 65

La adición de información de fase y el procesamiento controlado por codificador de la información de amplitud espectral enviada en el flujo de bits mejoran la fidelidad de este proceso, de manera que el canal mono compuesto puede extenderse a frecuencias más bajas de lo que era anteriormente posible. Esto reduce el ancho de banda eficaz codificado y, por tanto, aumenta la eficacia de codificación.

5 La norma E-AC-3 incluye además extensión espectral. La extensión espectral incluye sustituir coeficientes de transformada de mayor frecuencia por segmentos espectrales de menor frecuencia convertidos en frecuencia de manera ascendente. Las características espectrales de los segmentos convertidos se hacen corresponder con las originales mediante la modulación espectral de los coeficientes de transformada y también mezclando componentes de ruido conformadas con los segmentos espectrales de menor frecuencia convertidos.

10 La norma E-AC-3 incluye un canal de efectos de baja frecuencia (LFE). Es un único canal opcional con un ancho de banda limitado (<120 Hz), destinado a reproducirse en un nivel de +10 dB con respecto a los canales de ancho de banda total. El canal LFE opcional permite proporcionar altos niveles de presión de sonido para sonidos de baja frecuencia. Otras normas de codificación, por ejemplo, AC-3 y HE-AAC, también incluyen un canal LFE opcional.

15 Una técnica adicional para mejorar la calidad de audio a bajas velocidades de datos es usar el procesamiento de pre-ruido transitorio, descrito posteriormente en detalle.

20 Descodificación AC-3

En implementaciones de descodificador AC-3 típicas, para mantener los requisitos de latencia de la memoria y del descodificador tan bajos como sea posible, cada trama AC-3 se descodifica en una serie de bucles anidados.

25 Una primera etapa establece el alineamiento de trama. Esto implica hallar la palabra de sincronización AC-3 y después confirmar que las palabras de detección de errores CRC no indican ningún error. Una vez que se haya obtenido la sincronización de trama, los datos BSI se desempaquetan para determinar información de trama importante, tal como el número de canales codificados. Uno de los canales puede ser un canal LFE. El número de canales codificados se denota como N.n en este documento, donde n es el número de canales LFE y N es el número de canales principales. En las normas de codificación usadas actualmente, n=0 ó 1. En el futuro podrá haber casos en los que n>1.

30 La siguiente etapa en la descodificación es desempaquetar cada uno de los seis bloques de audio. Para minimizar los requisitos de memoria de las memorias intermedias de datos modulados por código de pulso (PCM) de salida, los bloques de audio se desempaquetan uno cada vez. Al final de cada periodo de bloque los resultados de PCM, en muchas implementaciones, se copian en memorias intermedias de salida, que para un funcionamiento en tiempo real en un descodificador hardware normalmente se duplican o se almacenan en memorias intermedias de manera circular para un acceso de interrupción directa mediante un convertidor de digital a analógico (DAC).

40 El procesamiento de bloques de audio de descodificador AC-3 puede dividirse en dos etapas diferentes, denominadas en este documento como procesamiento de entrada y procesamiento de salida. El procesamiento de entrada incluye el desempaquetado de todos los flujos de bits y la manipulación de canales codificados. El procesamiento de salida se refiere principalmente a las etapas de división en ventanas y de solapamiento y suma de la transformada MDCT inversa.

45 Esta distinción se realiza porque el número de canales de salida principales, denotado en este documento como $M \geq 1$, generado por un descodificador AC-3 no coincide necesariamente con el número de canales principales de entrada, denotado en este documento como N, $N \geq 1$, codificados en el flujo de bits, donde normalmente, aunque no necesariamente, $N \geq M$. Usando el mezclado descendente, un descodificador puede aceptar un flujo de bits con cualquier número N de canales codificados y producir un número arbitrario M, $M \geq 1$, de canales de salida. Debe observarse que, en general, el número de canales de salida se denota como M.m en este documento, donde M es el número de canales principales y m es el número de canales de salida LFE. En las aplicaciones actuales, m=0 ó 1. En el futuro es posible que m sea mayor que 1.

50 Debe observarse que en el mezclado descendente no todos los canales codificados están incluidos en los canales de salida. Por ejemplo, en un mezclado descendente de 5.1 a estéreo, la información de canal LFE se descarta normalmente. Por tanto, en algunos mezclados descendentes, n=1 y m=0, es decir, no hay ningún canal LFE de salida.

60 La figura 1 muestra un pseudocódigo 100 de instrucciones que, cuando se ejecutan, llevan a cabo un proceso de descodificación AC-3 típico.

65 El procesamiento de entrada en la descodificación AC-3 empieza normalmente cuando el descodificador desempaqueta los datos de bloque de audio fijos, que son una colección de parámetros y banderas ubicados al principio del bloque de audio. Los datos fijos incluyen elementos tales como banderas de conmutación de bloque,

información de acoplamiento, exponentes y parámetros de asignación de bits. El término "datos fijos" se refiere al hecho de que los tamaños de palabra para estos elementos de flujo de bits se conocen a priori y, por lo tanto, un proceso de decodificación de longitud variable no es necesario para recuperar tales elementos.

5 Los exponentes forman el campo más grande de la región de datos fijos, ya que incluyen todos los exponentes de cada canal codificado. Dependiendo del modo de codificación, en AC-3 solo puede haber un exponente por mantisa, hasta un máximo de 253 mantisas por canal. En lugar de desempaquetar todos estos exponentes en la memoria local, muchas implementaciones de decodificador guardan punteros que apuntan a los campos de exponente y los desempaquetan a medida que se necesiten, un canal cada vez.

10 Una vez que los datos fijos se han desempaquetado, muchos decodificadores AC-3 conocidos empiezan a procesar cada canal codificado. En primer lugar, los exponentes para el canal dado se desempaquetan de la trama de entrada. Después se lleva a cabo normalmente un cálculo de asignación de bits, que toma los exponentes y los parámetros de asignación de bits y calcula los tamaños de palabra para cada mantisa empaquetada. Después, las mantisas se desempaquetan normalmente de la trama de entrada. Las mantisas se escalan para proporcionar un control de margen dinámico apropiado y, si fuera necesario, para deshacer la operación de acoplamiento, y después se desnormalizan mediante los exponentes. Por último, se calcula una transformada inversa para determinar datos de solapamiento previo y suma, datos en lo que se denomina "dominio de ventana", y los resultados se mezclan de manera descendente en las memorias intermedias de mezclado descendente apropiadas para un procesamiento de salida subsiguiente.

20 En algunas implementaciones, los exponentes para el canal individual se desempaquetan en una memoria intermedia con una longitud de 256 muestras, denominada "memoria intermedia MDCT". Después, estos exponentes se agrupan en un máximo de 50 bandas para fines de asignación de bits. El número de exponentes en cada banda aumenta hacia frecuencias de audio más altas, siguiendo aproximadamente una división logarítmica que modela bandas críticas psicoacústicas.

30 Para cada una de estas bandas de asignación de bits, los exponentes y los parámetros de asignación de bits se combinan para generar un tamaño de palabra de mantisa para cada mantisa de esa banda. Estos tamaños de palabra se almacenan en una memoria intermedia de bandas con una longitud de 24 muestras, donde la banda de asignación de bits más amplia está formada por 24 contenedores de frecuencia. Una vez que se hayan calculado los tamaños de palabra, las mantisas correspondientes se desempaquetan de la trama de entrada y vuelven a almacenarse en su sitio en la memoria intermedia de bandas. Estas mantisas se escalan y se desnormalizan mediante el exponente correspondiente y se escriben, por ejemplo vuelven a escribirse en su sitio, en la memoria intermedia MDCT. Después de que todas las bandas se hayan procesado y todas las mantisas se hayan desempaquetado, las ubicaciones restantes de la memoria intermedia MDCT se rellenan normalmente con ceros.

35 Se lleva a cabo una transforma inversa, por ejemplo llevada a cabo en la memoria intermedia MDCT. Después, la salida de este procesamiento, los datos de dominio de ventana, puede mezclarse de manera descendente en las memorias intermedias de mezclado descendente apropiadas según parámetros de mezclado descendente determinados según los metadatos, por ejemplo extraídos de datos predefinidos según los metadatos.

40 Una vez que el procesamiento de entrada se haya completado y que las memorias intermedias de mezclado descendente se hayan generado completamente con datos mezclados de manera descendente en el dominio de ventana, el decodificador puede llevar a cabo el procesamiento de salida. Para cada canal de salida, una memoria intermedia de mezclado descendente y su memoria intermedia correspondiente de retardo de medio bloque con una longitud de 128 muestras se dividen en ventanas y se combinan para producir 256 muestras de salida PCM. En un sistema de sonido de hardware que incluye un decodificador y uno o más DAC, estas muestras se redondean al ancho de palabra DAC y se copian en la memoria intermedia de salida. Una vez que se ha realizado esto, la mitad de la memoria intermedia de mezclado descendente se copia en su memoria intermedia de retardo correspondiente, proporcionando información solapada en un 50% necesaria para una correcta reconstrucción del siguiente bloque de audio.

55 Descodificación E-AC-3

Realizaciones particulares de la presente invención incluyen un método que hace funcionar un decodificador de audio para decodificar datos de audio que incluyen un número, denotado como N.n, de canales de datos de audio codificados, por ejemplo, un decodificador de audio E-AC-3 para decodificar datos de audio codificados E-AC-3 para formar datos de audio decodificados que incluyen M.m canales de audio decodificados, $n=0$ ó 1 , $m=0$ ó 1 y $M \geq 1$. $n=1$ indica un canal LFE de entrada, $m=1$ indica un canal LFE de salida. $M < N$ indica un mezclado descendente, $M > N$ indica un mezclado ascendente.

65 El método incluye aceptar datos de audio que incluyen N.n canales de datos de audio codificados que han sido codificados mediante el método de codificación, por ejemplo mediante un método de codificación que incluye transformar N canales de datos de audio digital usando una transformada solapada, formar y empaquetar datos de exponente y de mantisa de dominio de frecuencia, y formar y empaquetar metadatos relacionados con los datos de

exponente y de mantisa de dominio de frecuencia, donde los metadatos incluyen opcionalmente metadatos relacionados con el procesamiento de pre-ruido transitorio, por ejemplo mediante un método de codificación E-AC-3.

5 Algunas realizaciones descritas en este documento están diseñadas para aceptar datos de audio codificados que han sido codificados según la norma E-AC-3 o según una norma compatible con versiones anteriores de la norma E-AC-3, y pueden incluir más de 5 canales principales codificados.

10 Tal y como se describirá posteriormente en mayor detalle, el método incluye descodificar los datos de audio aceptados, donde la descodificación incluye: desempaquetar los metadatos y desempaquetar y descodificar los datos de exponente y de mantisa de dominio de frecuencia; determinar coeficientes de transformada a partir de los datos de exponente y de mantisa de dominio de frecuencia desempaquetados y descodificados; someter a una transformación inversa los datos de dominio de frecuencia; aplicar operaciones de división en ventanas y de solapamiento y suma para determinar datos de audio muestreados; aplicar cualquier procesamiento de pre-ruido transitorio requerido de descodificación según los metadatos relacionados con el procesamiento de pre-ruido transitorio; y, en caso de que $M < N$, realizar un mezclado descendente según datos de mezclado descendente. El mezclado descendente incluye comprobar si los datos de mezclado descendente han cambiado con respecto a datos de mezclado descendente usados anteriormente, y, si han cambiado, aplicar una atenuación cruzada para determinar datos de mezclado descendente atenuados de manera cruzada y un mezclado descendente según los datos de mezclado descendente atenuados de manera cruzada, y, si no han cambiado, realizar directamente un mezclado descendente según los datos de mezclado descendente.

25 En algunas realizaciones de la presente invención, el descodificador usa al menos un procesador x86 que ejecuta un flujo continuo de instrucciones de extensiones (SSE) de una sola instrucción y múltiples datos (SIMD), que incluyen instrucciones vectoriales. En tales realizaciones, el mezclado descendente incluye ejecutar instrucciones vectoriales en al menos un procesador del uno o más procesadores x86.

30 En algunas realizaciones de la presente invención, el método de descodificación de audio E-AC-3, que puede ser audio AC-3, está dividido en módulos de operaciones que pueden aplicarse más de una vez, es decir, instanciarse más de una vez en diferentes implementaciones de descodificador. En el caso de un método que incluye descodificación, la descodificación se divide en un conjunto de operaciones de descodificación de sección de entrada (FED) y en un conjunto de operaciones de descodificación de sección de procesamiento (BED). Como se explicará posteriormente en detalle, las operaciones de descodificación de sección de entrada incluyen desempaquetar y descodificar datos de exponente y de mantisa de dominio de frecuencia de una trama de un flujo de bits AC-3 o E-AC-3 en datos de exponente y de mantisa de dominio de frecuencia desempaquetados y descodificados para la trama y los metadatos incluidos en la trama. Las operaciones de descodificación de sección de procesamiento incluyen determinar los coeficientes de transformada, someter a una transformación inversa los coeficientes de transformada determinados, aplicar operaciones de división en ventanas y de solapamiento y suma, aplicar cualquier procesamiento de pre-ruido transitorio requerido de descodificación y aplicar un mezclado descendente en caso de que haya menos canales de salida que canales codificados en el flujo de bits.

40 Algunas realizaciones de la presente invención incluyen un medio de almacenamiento legible por ordenador que almacena instrucciones que cuando son ejecutadas por uno o más procesadores de un sistema de procesamiento hacen que el sistema de procesamiento lleve a cabo la descodificación de datos de audio que incluyen $N.n$ canales de datos de audio codificados para formar datos de audio descodificados que incluyen $M.m$ canales de audio descodificado, $M \geq 1$. En las normas actuales, $n=0$ ó 1 y $m=0$ ó 1 , pero la invención no está limitada a esto. Las instrucciones incluyen instrucciones que cuando se ejecutan hacen que se acepten los datos de audio que incluyen $N.n$ canales de datos de audio codificados que han sido codificados mediante un método de codificación, por ejemplo AC-3 o E-AC-3. Las instrucciones incluyen además instrucciones que cuando se ejecutan hacen que se descodifiquen los datos de audio aceptados.

50 En algunas de tales realizaciones, los datos de audio aceptados están en forma de un flujo de bits AC-3 o E-AC-3 de tramas de datos codificados. Las instrucciones que cuando son ejecutadas hacen que se descodifiquen los datos de audio aceptados se dividen en un conjunto de módulos reutilizables de instrucciones, que incluyen un módulo de descodificación de sección de entrada (FED) y un módulo de descodificación de sección de procesamiento (BED). El módulo de descodificación de sección de entrada incluye instrucciones que cuando se ejecutan hacen que se desempaqueten y se descodifiquen los datos de exponente y de mantisa de dominio de frecuencia de una trama del flujo de bits en datos de exponente y de mantisa de dominio de frecuencia desempaquetados y descodificados para la trama y los metadatos incluidos en la trama. El módulo de descodificación de sección de procesamiento incluye instrucciones que cuando se ejecutan hacen que se determinen los coeficientes de transformada, se realice una transformación inversa, se apliquen operaciones de división en ventanas y de solapamiento y suma, se aplique cualquier procesamiento de pre-ruido transitorio requerido de descodificación y se aplique un mezclado descendente en caso de que haya menos canales de salida que canales codificados de entrada.

65 Las figuras 2A a 2D muestran en forma de diagramas de bloques simplificados algunas configuraciones de descodificador diferentes que pueden usar de manera ventajosa uno o más módulos comunes. La figura 2A muestra un diagrama de bloques simplificado de un ejemplo de descodificador E-AC-3 200 para audio 5.1 codificado AC-3 o

E-AC-3. Evidentemente, el uso del término "bloque" cuando hace referencia a los bloques de un diagrama de bloques es diferente a cuando hace referencia a un bloque de datos de audio, donde este último se refiere a una cantidad de datos de audio. El descodificador 200 incluye un módulo de descodificación de sección de entrada (FED) 201 que acepta tramas AC-3 o E-AC-3 y lleva a cabo, trama a trama, el desempaqueado de los metadatos de la trama y la descodificación de los datos de audio de la trama en datos de exponente y de mantisa de dominio de frecuencia. El descodificador 200 incluye además un módulo de descodificación de sección de procesamiento (BED) 203 que acepta los datos de exponente y de mantisa de dominio de frecuencia del módulo de descodificación de sección de entrada 201 y los descodifica para un máximo de 5.1 canales de datos de audio PCM.

- 5
- 10 La descomposición del descodificador en un módulo de descodificación de sección de entrada y un módulo de descodificación de sección de procesamiento es una elección de diseño, no una división necesaria. Tal división proporciona el beneficio de tener módulos comunes en varias configuraciones alternativas. El módulo FED puede ser común a tales configuraciones alternativas, y muchas configuraciones tienen en común el desempaqueado de los metadatos de la trama y la descodificación de los datos de audio de la trama en datos de exponente y de mantisa de dominio de frecuencia llevados a cabo por un módulo FED.
- 15

Como un ejemplo de una configuración alternativa, la figura 2B muestra un diagrama de bloques simplificado de un descodificador/convertidor E-AC-3 210 para audio 5.1 codificado E-AC-3 que descodifica audio 5.1 codificado AC-3 o E-AC-3 y que también convierte una trama codificada E-AC-3 de hasta 5.1 canales de audio en una trama codificada AC-3 de hasta 5.1 canales. El descodificador/convertidor 210 incluye un módulo de descodificación de sección de entrada (FED) 201 que acepta tramas AC-3 o E-AC-3 y que lleva a cabo, trama a trama, el desempaqueado de los metadatos de la trama y la descodificación de los datos de audio de la trama en datos de exponente y de mantisa de dominio de frecuencia. El descodificador/convertidor 210 incluye además un módulo de descodificación de sección de procesamiento (BED) 203 que es el mismo o similar al módulo BED 203 del descodificador 200 y que acepta los datos de exponente y de mantisa de dominio de frecuencia del módulo de descodificación de sección de entrada 201 y los descodifica para un máximo de 5.1 canales de datos de audio PCM. El descodificador/convertidor 210 incluye además un módulo convertidor de metadatos 205 para la conversión de metadatos y un módulo de codificación de sección de procesamiento 207 que acepta los datos de exponente y de mantisa de dominio de frecuencia del módulo de descodificación de sección de entrada 201 y codifica los datos como una trama AC-3 de hasta 5.1 canales de datos de audio sin superar la máxima velocidad de datos de 640 kbps que es posible con AC-3.

- 20
- 25
- 30

Como un ejemplo de una configuración alternativa, la figura 2C muestra un diagrama de bloques simplificado de un descodificador E-AC-3 que descodifica una trama AC-3 de hasta 5.1 canales de audio codificado y que también descodifica una trama codificada E-AC-3 de hasta 7.1 canales de audio. El descodificador 220 incluye un módulo de análisis de información de trama 221 que desempaqueta los datos BSI e identifica las tramas y los tipos de trama y proporciona las tramas a elementos de descodificador de sección de entrada apropiados. En una implementación típica que incluye uno o más procesadores y una memoria que almacena instrucciones que cuando se ejecutan hacen que se lleve a cabo la funcionalidad de los módulos, múltiples instancias de un módulo de descodificación de sección de entrada y múltiples instancias de un módulo de descodificación de sección de procesamiento pueden estar funcionando. En algunas realizaciones de un descodificador E-AC-3, la funcionalidad de desempaqueado de BSI no está presente en el módulo de descodificación de sección de entrada que consulta los datos BSI. Esto permite utilizar módulos comunes en varias implementaciones alternativas. La figura 2C muestra un diagrama de bloques simplificado de un descodificador con una arquitectura adecuada para un máximo de 7.1 canales de datos de audio. La figura 2D muestra un diagrama de bloques simplificado de un descodificador 5.1 240 con tal arquitectura. El descodificador 240 incluye un módulo de análisis de información de trama 241, un módulo de descodificación de sección de entrada 243 y un módulo de descodificación de sección de procesamiento 245. Estos módulos FED y BED pueden tener una estructura similar a la de los módulos FED y BED usados en la arquitectura de la figura 2C.

- 35
- 40
- 45
- 50

Volviendo a la figura 2C, el módulo de análisis de información de trama 221 proporciona los datos de una trama codificada AC-3/E-AC-3 independiente de hasta 5.1 canales a un módulo de descodificación de sección de entrada 223 que acepta las tramas AC-3 o E-AC-3 y que lleva a cabo, trama a trama, el desempaqueado de los metadatos de la trama y la descodificación de los datos de audio de la trama en datos de exponente y de mantisa de dominio de frecuencia. Los datos de exponente y de mantisa de dominio de frecuencia son aceptados por un módulo de descodificación de sección de procesamiento 225 que es idéntico o similar al módulo BED 203 del descodificador 200 y que acepta los datos de exponente y de mantisa de dominio de frecuencia del módulo de descodificación de sección de entrada 223 y que descodifica los datos hasta 5.1 canales de datos de audio PCM. Cualquier trama codificada AC-3/E-AC-3 dependiente de datos de canal adicionales es proporcionada a otro módulo de descodificación de sección de entrada 227 que es similar al otro módulo FED y que, por tanto, desempaqueta los metadatos de la trama y descodifica los datos de audio de la trama en datos de exponente y de mantisa de dominio de frecuencia. Un módulo de descodificación de sección de procesamiento 229 acepta los datos del módulo FED 227 y descodifica los datos en datos de audio PCM de cualquier canal adicional. Un módulo de correlación de canales PCM 231 se usa para combinar los datos descodificados de los módulos BED respectivos para proporcionar hasta 7.1 canales de datos PCM.

- 55
- 60
- 65

Si hay más de 5 canales principales codificados, es decir, el caso en que $N > 5$, por ejemplo, hay 7.1 canales codificados, el flujo de bits codificado incluye una trama independiente de hasta 5.1 canales codificados y al menos una trama dependiente de datos codificados. En realizaciones de software para este caso, por ejemplo realizaciones que comprenden un medio legible por ordenador que almacena instrucciones para su ejecución, las instrucciones están dispuestas como una pluralidad de módulos de descodificación de 5.1 canales, incluyendo cada módulo de descodificación de 5.1 canales una instancia respectiva de un módulo de descodificación de sección de entrada y una instancia respectiva de un módulo de descodificación de sección de procesamiento. La pluralidad de módulos de descodificación de 5.1 canales incluye un primer módulo de descodificación de 5.1 canales que cuando se ejecuta provoca la descodificación de la trama independiente, y uno o más otros módulos de descodificación de canal para cada trama dependiente respectiva. En algunas de estas realizaciones, las instrucciones incluyen un módulo de instrucciones de análisis de información de trama que cuando se ejecutan hacen que se desempaquete el campo de información de flujo de bits (BSI) de cada trama para identificar las tramas y los tipos de trama y se proporcionen las tramas identificadas a la instancia de módulo de descodificador de sección de entrada apropiado, y un módulo de instrucciones de correlación de canales que cuando se ejecutan, y en caso de que $N > 5$, hacen que se combinen los datos descodificados de módulos de descodificación de sección de procesamiento respectivos para formar los N canales principales de datos descodificados.

Un método para hacer funcionar un descodificador/convertidor dual AC-3/E-AC-3

Una realización de la invención adopta la forma de un descodificador/convertidor dual (DDC) que descodifica dos flujos de bits de entrada AC-3/E-AC-3, designados como "principal" y "asociado", cada uno con un máximo de 5.1 canales, en audio PCM, y en caso de conversión, convierte el flujo de bits de audio principal de E-AC-3 a AC-3, y en caso de descodificación, descodifica el flujo de bits principal y, si lo hubiera, el flujo de bits asociado. El descodificador/convertidor dual mezcla opcionalmente las dos salidas PCM usando metadatos de mezclado extraídos del flujo de bits de audio asociado.

Una realización del descodificador/convertidor dual lleva a cabo un método que hace funcionar un descodificador para llevar a cabo los procesos incluidos en la descodificación y/o conversión de los dos flujos de bits de entrada AC-3/E-AC-3. Otra realización adopta la forma de un medio de almacenamiento tangible que incluye instrucciones, por ejemplo instrucciones de software en el mismo, que cuando son ejecutadas por uno o más procesadores de un sistema de procesamiento, hacen que el sistema de procesamiento lleve a cabo los procesos incluidos en la descodificación y/o conversión de los dos flujos de bits de entrada AC-3/E-AC-3.

Una realización del descodificador/convertidor dual AC-3/E-AC-3 tiene seis subcomponentes, algunos de los cuales incluyen subcomponentes comunes. Los módulos son:

Descodificador-convertidor: El descodificador-convertidor está configurado, cuando se ejecuta, para descodificar un flujo de bits de entrada AC-3/E-AC-3 (hasta 5.1 canales) en audio PCM, y/o para convertir el flujo de bits de entrada de E-AC-3 a AC-3. El descodificador-convertidor tiene tres subcomponentes principales y puede implementar la realización 210 mostrada en la anterior figura 2B. Los subcomponentes principales son:

Descodificador de sección de entrada: El módulo FED está configurado, cuando se ejecuta, para descodificar una trama de un flujo de bits AC-3/E-AC-3 en datos de audio sin procesar de dominio de frecuencia y sus metadatos asociados.

Descodificador de sección de procesamiento: El módulo BED está configurado, cuando se ejecuta, para completar el resto del proceso de descodificación iniciado por el módulo FED. En particular, el módulo BED descodifica los datos de audio (en formato de mantisa y exponente) en datos de audio PCM.

Codificador de sección de procesamiento: El módulo de codificación de sección de procesamiento está configurado, cuando se ejecuta, para codificar una trama AC-3 usando seis bloques de datos de audio del FED. El módulo de codificación de sección de procesamiento también está configurado, cuando se ejecuta, para sincronizar, resolver y convertir los metadatos E-AC-3 en metadatos Dolby Digital usando un módulo convertidor de metadatos.

Descodificador 5.1: El módulo de descodificación 5.1 está configurado, cuando se ejecuta, para descodificar un flujo de bits de entrada AC-3/E-AC-3 (hasta 5.1 canales) en audio PCM. El descodificador 5.1 también proporciona opcionalmente metadatos de mezclado para su uso por parte de una aplicación externa para mezclar dos flujos de bits AC-3/E-AC-3. El módulo de descodificación incluye dos subcomponentes principales: un módulo FED como el descrito anteriormente en este documento y un módulo BED como el descrito anteriormente en este documento. Un diagrama de bloques de un descodificador 5.1 de ejemplo se muestra en la figura 2D.

Información de trama: El módulo de información de trama está configurado, cuando se ejecuta, para analizar una trama AC-3/E-AC-3 y desempaquetar su información de flujo de bits. Se lleva a cabo una comprobación CRC en la trama como parte del proceso de desempaquetado.

Descriptores de memoria intermedia: El módulo de descriptores de memoria intermedia contiene descripciones de

memoria intermedia AC-3, E-AC-3 y PCM, y funciones para operaciones de memoria intermedia.

Convertidor de frecuencia de muestreo: El módulo convertidor de frecuencia de muestreo es opcional y está configurado, cuando se ejecuta, para muestrear de manera ascendente audio PCM en un factor de dos.

5 Mezclador externo: El módulo mezclador externo es opcional y está configurado, cuando se ejecuta, para mezclar un programa de audio principal y un programa de audio asociado para formar un único programa de audio de salida usando metadatos de mezclado suministrados en el programa de audio asociado.

10 Diseño de un módulo de descodificación de sección de entrada

El módulo de descodificación de sección de entrada descodifica datos según los métodos de la norma AC-3 y según aspectos de descodificación E-AC-3 adicionales, incluyendo la descodificación de datos AHT para señales estacionarias, el acoplamiento de canal mejorado de la norma E-AC-3 y la extensión espectral.

15 En el caso de una realización en forma de un medio de almacenamiento tangible, el módulo de descodificación de sección de entrada comprende instrucciones de software almacenadas en un medio de almacenamiento tangible que, cuando son ejecutadas por uno o más procesadores de un sistema de procesamiento, llevan a cabo las acciones descritas en los detalles proporcionados en este documento para el funcionamiento del módulo de descodificación de sección de entrada. En una implementación en hardware, el módulo de descodificación de sección de entrada incluye elementos que están configurados en funcionamiento para llevar a cabo las acciones descritas en los detalles proporcionados en este documento para el funcionamiento del módulo de descodificación de sección de entrada.

25 En la descodificación AC-3 es posible una descodificación bloque a bloque. Con E-AC-3, el primer bloque de audio, el bloque de audio 0 de una trama, incluye las mantisas AHT de los 6 bloques. Por tanto, la descodificación bloque a bloque no se usa normalmente, sino que varios bloques se procesan a la vez. Sin embargo, el procesamiento de los datos reales se lleva a cabo, evidentemente, en cada bloque.

30 En una realización, para usar un método uniforme de descodificación/arquitectura de un descodificador independientemente de si se usa AHT, el módulo FED lleva a cabo, canal a canal, dos pasadas. Una primera pasada incluye desempaquear metadatos bloque a bloque y guardar punteros que apuntan a la ubicación en la que están almacenados los datos de exponente y de mantisa empaquetados, y una segunda pasada incluye usar los punteros guardados que apuntan a los exponentes y las mantisas empaquetados, y desempaquear y descodificar datos de exponente y de mantisa canal a canal.

40 La figura 3 muestra un diagrama de bloques simplificado de una realización de un módulo de descodificación de sección de entrada, por ejemplo implementado como un conjunto de instrucciones almacenadas en una memoria que cuando se ejecutan hacen que se lleve a cabo el procesamiento FED. La figura 3 también muestra un pseudocódigo de instrucciones para una primera pasada de un módulo de descodificación de sección de entrada de dos pasadas 300, así como un pseudocódigo de instrucciones para la segunda pasada del módulo de descodificación de sección de entrada de dos pasadas. El módulo FED incluye los siguientes módulos, cada uno incluyendo instrucciones, donde algunas de tales instrucciones son instrucciones de definición ya que definen estructuras y parámetros:

45 Canal: El módulo de canal define estructuras para representar un canal de audio en memoria y proporciona instrucciones para desempaquear y descodificar un canal de audio de un flujo de bits AC-3 o E-AC-3.

50 Asignación de bits: El módulo de asignación de bits proporciona instrucciones para calcular la curva de enmascaramiento y calcular la asignación de bits para datos codificados.

Operaciones de flujo de bits: El módulo de operaciones de flujo de bits proporciona instrucciones para desempaquear datos de un flujo de bits AC-3 o E-AC-3.

55 Exponentes: El módulo de exponentes define estructuras para representar exponentes en memoria y proporciona instrucciones configuradas, cuando se ejecutan, para desempaquear y descodificar exponentes de un flujo de bits AC-3 o E-AC-3.

60 Exponentes y mantisas: El módulo de exponentes y mantisas define estructuras para representar exponentes y mantisas en memoria y proporciona instrucciones configuradas, cuando se ejecutan, para desempaquear y descodificar exponentes y mantisas de un flujo de bits AC-3 o E-AC-3.

Matrización: El módulo de matrización proporciona instrucciones configuradas, cuando se ejecutan, para soportar la desmatrización de canales matrizados.

65 Datos auxiliares: El módulo de datos auxiliares define estructuras de datos auxiliares usadas en el módulo FED para

Llevar a cabo el procesamiento FED.

Mantisas: El módulo de mantisas define estructuras para representar mantisas en memoria y proporciona instrucciones configuradas, cuando se ejecutan, para desempaquetar y descodificar mantisas de un flujo de bits AC-3 o E-AC-3.

Transformada híbrida adaptativa: El módulo AHT proporciona instrucciones configuradas, cuando se ejecutan, para desempaquetar y descodificar datos de transformada híbrida adaptativa de un flujo de bits E-AC-3.

Trama de audio: El módulo de tramas de audio define estructuras para representar una trama de audio en memoria y proporciona instrucciones configuradas, cuando se ejecutan, para desempaquetar y descodificar una trama de audio de un flujo de bits AC-3 o E-AC-3.

Acoplamiento mejorado: El módulo de acoplamiento mejorado define estructuras para representar un canal de acoplamiento mejorado en memoria y proporciona instrucciones configuradas, cuando se ejecutan, para desempaquetar y descodificar un canal de acoplamiento mejorado de un flujo de bits AC-3 o E-AC-3. El acoplamiento mejorado amplía el acoplamiento tradicional en un flujo de bits E-AC-3 proporcionando información de fase y caos.

Bloque de audio: El módulo de bloque de audio define estructuras para representar un bloque de audio en memoria y proporciona instrucciones configuradas, cuando se ejecutan, para desempaquetar y descodificar un bloque de audio de un flujo de bits AC-3 o E-AC-3.

Extensión espectral: El módulo de extensión espectral proporciona soporte para la descodificación de extensión espectral en un flujo de bits E-AC-3.

Acoplamiento: El módulo de acoplamiento define estructuras para representar un canal de acoplamiento en memoria y proporciona instrucciones configuradas, cuando se ejecutan, para desempaquetar y descodificar un canal de acoplamiento de un flujo de bits AC-3 o E-AC-3.

La figura 4 muestra un diagrama de flujo de datos simplificado del funcionamiento de una realización del módulo de descodificación de sección de entrada 300 de la figura 3, que describe cómo el pseudocódigo y los elementos de submódulo mostrados en la figura 3 actúan conjuntamente para llevar a cabo las funciones de un módulo de descodificación de sección de entrada. Por elemento funcional se entiende un elemento que lleva a cabo una función de procesamiento. Cada elemento de este tipo puede ser un elemento de hardware, o un sistema de procesamiento y un medio de almacenamiento que incluye instrucciones que, cuando se ejecutan, llevan a cabo la función. Un elemento funcional de desempaquetado de flujos de bits 403 acepta una trama AC-3/E-AC-3 y genera parámetros de asignación de bits para un elemento funcional de asignación de bits estándar y/o AHT 405 que produce datos adicionales para el desempaquetado de flujos de bits para generar en última instancia datos de exponente y de mantisa para un elemento funcional de desacoplamiento estándar/mejorado 407. El elemento funcional 407 genera datos de exponente y de mantisa para que un elemento funcional de rematrización incluido 409 lleve a cabo cualquier rematrización necesaria. El elemento funcional 409 genera datos de exponente y de mantisa para que un elemento funcional de descodificación de extensión espectral incluido 411 lleve a cabo cualquier extensión espectral necesaria. Los elementos funcionales 407 a 411 usan datos obtenidos mediante la operación de desempaquetado del elemento funcional 403. El resultado de la descodificación de sección de entrada son datos de exponente y de mantisa, así como parámetros de trama de audio desempaquetada y parámetros de bloque de audio adicionales.

Haciendo referencia en mayor detalle al pseudocódigo de primera pasada y segunda pasada mostrado en la figura 3, las instrucciones de la primera pasada están configuradas, cuando se ejecutan, para desempaquetar metadatos de una trama AC-3/E-AC-3. En particular, la primera pasada incluye desempaquetar la información BSI y desempaquetar la información de trama de audio. Para cada bloque, empezando desde el bloque 0 hasta el bloque 5 (6 bloques por trama), se desempaquetan los datos fijos, y para cada canal se guarda un puntero que apunta a los exponentes empaquetados en el flujo de bits, se desempaquetan los exponentes y se guarda la posición del flujo de bits en la que residen las mantisas empaquetadas. Se calcula la asignación de bits y, en función de la asignación de bits, las mantisas pueden omitirse.

Las instrucciones de la segunda pasada están configuradas, cuando se ejecutan, para descodificar los datos de audio de una trama para formar datos de mantisa y de exponente. Para cada bloque, empezando por el bloque 0, el desempaquetado incluye cargar el puntero guardado que apunta a exponentes empaquetados y desempaquetar los exponentes apuntados de esta manera, calcular la asignación de bits, cargar el puntero guardado que apunta a las mantisas empaquetadas y desempaquetar las mantisas apuntadas de esta manera. La descodificación incluye llevar a cabo un desacoplamiento estándar y mejorado y generar la(s) banda(s) de extensión espectral y, para que sea independiente de otros módulos, transferir los datos resultantes a una memoria, por ejemplo una memoria externa a la memoria interna de la pasada, para que otros módulos, por ejemplo el módulo BED, puedan acceder a los datos resultantes. Por comodidad, esta memoria se denomina memoria "externa", aunque, como le resultará evidente a un

experto en la técnica, puede formar parte de una estructura de una sola memoria usada para todos los módulos.

5 En algunas realizaciones, para el desempaquetado de los exponentes, los exponentes desempaquetados durante la primera pasada no se guardan para minimizar las transferencias de memoria. Si se usa AHT para un canal, los exponentes se desempaquetan a partir del bloque 0 y se copian en los otros cinco bloques, numerados del 1 al 5. Si no se usa AHT para un canal, se guardan los punteros que apuntan a exponentes empaquetados. Si la estrategia de exponentes de canal es reutilizar los exponentes, los exponentes se desempaquetan de nuevo usando los punteros guardados.

10 En algunas realizaciones, para el desempaquetado de mantisas de acoplamiento, si se usa AHT para el canal de acoplamiento, los seis bloques de las mantisas de canal de acoplamiento AHT se desempaquetan en bloque 0 y se regeneran mediante interpolación para cada canal, es decir, un canal acoplado para producir una interpolación no correlacionada. Si no se usa AHT para el canal de acoplamiento, se guardan los punteros que apuntan a las mantisas de acoplamiento. Estos punteros guardados se usan para volver a desempaquetar las mantisas de
15 acoplamiento para cada canal que sea un canal acoplado en un bloque dado.

Diseño de un módulo de descodificación de sección de procesamiento

20 El módulo de descodificación de sección de procesamiento (BED) puede hacerse funcionar para tomar los datos de exponente y de mantisa de dominio de frecuencia y descodificarlos en datos de audio PCM. Los datos de audio PCM se representan en función de los modos seleccionados por el usuario, la compresión de margen dinámico y los modos de mezclado descendente.

25 En algunas realizaciones, en las que el módulo de descodificación de sección de entrada almacena datos de exponente y de mantisa en una memoria, la denominada memoria externa, diferente de la memoria de trabajo del módulo de sección de entrada, el módulo BED usa un procesamiento de trama bloque a bloque para minimizar los requisitos de mezclado descendente y de memoria intermedia de retardo y, para que sea compatible con la salida del módulo de sección de entrada, usa transferencias de la memoria externa para acceder a datos de exponente y de mantisa que van a procesarse.

30 En caso de una realización en forma de un medio de almacenamiento tangible, el módulo de descodificación de sección de procesamiento comprende instrucciones de software almacenadas en un medio de almacenamiento tangible que, cuando son ejecutadas por uno o más procesadores de un sistema de procesamiento, llevan a cabo las acciones descritas en los detalles proporcionados en este documento para el funcionamiento del módulo de
35 descodificación de sección de procesamiento. En una implementación en hardware, el módulo de descodificación de sección de procesamiento incluye elementos que están configurados en funcionamiento para llevar a cabo las acciones descritas en los detalles proporcionados en este documento para el funcionamiento del módulo de descodificación de sección de procesamiento.

40 La figura 5A muestra un diagrama de bloques simplificado de una realización de un módulo de descodificación de sección de procesamiento 500 implementado como un conjunto de instrucciones almacenadas en una memoria que, cuando se ejecutan, hacen que se lleve a cabo el procesamiento BED. La figura 5A muestra además un pseudocódigo de instrucciones para el módulo de descodificación de sección de procesamiento 500. El módulo BED
45 500 incluye los siguientes módulos, incluyendo cada uno instrucciones, donde algunas de tales instrucciones son instrucciones que definen:

Control de margen dinámico: El módulo de control de margen dinámico proporciona instrucciones que, cuando se ejecutan, llevan a cabo funciones para controlar el margen dinámico de la señal descodificada, incluyendo aplicar un
50 ajuste de ganancia y aplicar un control de margen dinámico.

Transformada: El módulo de transformada proporciona instrucciones que, cuando se ejecutan, llevan a cabo las transformadas inversas, incluyendo llevar a cabo una transformada discreta del coseno modificada inversa (IMDCT), que incluye llevar a cabo una rotación previa usada para calcular la transformada DCT inversa, llevar a cabo una rotación posterior usada para calcular la transformada DCT inversa y determinar la transformada rápida de Fourier
55 inversa (IFFT).

Procesamiento de pre-ruido transitorio: El módulo de procesamiento de pre-ruido transitorio proporciona instrucciones que, cuando se ejecutan, llevan a cabo el procesamiento de pre-ruido transitorio.

60 Ventana & solapamiento y suma: El módulo de ventana y de solapamiento y suma con memoria intermedia de retardo proporciona instrucciones que, cuando se ejecutan, llevan a cabo la operación de división en ventanas y de solapamiento/suma para reconstruir muestras de salida de muestras transformadas inversas.

65 Mezclado descendente en el dominio de tiempo (TD): El módulo de mezclado descendente TD proporciona instrucciones que, cuando se ejecutan, llevan a cabo un mezclado descendente en el dominio de tiempo, según sea necesario, hacia un menor número de canales.

La figura 6 muestra un diagrama de flujo de datos simplificado del funcionamiento de una realización del módulo de descodificación de sección de procesamiento 500 de la figura 5A que describe cómo el código y los elementos de submódulo mostrados en la figura 5A actúan conjuntamente para llevar a cabo las funciones de un módulo de descodificación de sección de procesamiento. Un elemento funcional de control de ganancia 603 acepta datos de exponente y de mantisa del módulo de descodificación de sección de entrada 300 y aplica cualquier control de margen dinámico requerido, una normalización de diálogo y un ajuste de ganancia según metadatos. Los datos resultantes de exponente y de mantisa son aceptados por un elemento funcional de desnormalización de mantisas mediante exponentes 605 que genera los coeficientes de transformada para la transformación inversa. Un elemento funcional de transformada inversa 607 aplica la IMDCT a los coeficientes de transformada para generar muestras de tiempo que se someten previamente a operaciones de división en ventanas y de solapamiento y suma. Tales muestras de dominio de tiempo sometidas previamente a operaciones de solapamiento y suma se denominan en este documento muestras de "dominio de pseudo-tiempo", y estas muestras están en lo que se denomina en este documento dominio de pseudo-tiempo. Éstas son aceptadas por un elemento funcional de división en ventanas y de solapamiento y suma 609 que genera muestras PCM aplicando operaciones de división en ventanas y de solapamiento y suma a las muestras de dominio de pseudo-tiempo. Un elemento funcional de procesamiento de pre-ruido transitorio 611 aplica cualquier procesamiento de pre-ruido transitorio según metadatos. Si se indica, por ejemplo, en los metadatos o de otra manera, las muestras PCM de postprocesamiento de pre-ruido transitorio resultantes se mezclan de manera descendente al número M.m de canales de salida de muestras PCM mediante un elemento funcional de mezclado descendente 613.

Haciendo referencia de nuevo a la figura 5A, el pseudocódigo para el procesamiento del módulo BED incluye, para cada bloque de datos, transferir los datos de mantisa y de exponente para bloques de un canal desde la memoria externa y, para cada canal: aplicar cualquier control de margen dinámico requerido, una normalización de diálogo y un ajuste de ganancia según los metadatos; desnormalizar mantisas mediante exponentes para generar los coeficientes de transformada para la transformación inversa; calcular una IMDCT para los coeficientes de transformada para generar muestras de dominio de pseudo-tiempo; aplicar operaciones de división en ventanas y de solapamiento y suma a las muestras de dominio de pseudo-tiempo; aplicar cualquier procesamiento de pre-ruido transitorio según los metadatos; y, si es necesario, realizar un mezclado descendente en el dominio de tiempo hacia el número M.m de canales de salida de muestras PCM.

Realizaciones de descodificación mostradas en la figura 5A incluyen llevar a cabo tales ajustes de ganancia aplicando desfases de normalización de diálogo según los metadatos y aplicando factores de ganancia de control de margen dinámico según los metadatos. Llevar a cabo tales ajustes de ganancia en la fase en la que los datos se proporcionan en forma de mantisa y exponente en el dominio de frecuencia es ventajoso. Los cambios de ganancia pueden variar en el tiempo, y tales cambios de ganancia realizados en el dominio de frecuencia dan como resultado atenuaciones cruzadas graduales una vez que se hayan llevado a cabo la transformada inversa y las operaciones de división en ventanas y de solapamiento y suma.

40 Procesamiento de pre-ruido transitorio

La codificación y la descodificación E-AC-3 se han diseñado para permitir y proporcionar una mejor calidad de audio a velocidades de datos más bajas que en AC-3. A velocidades de datos más bajas, la calidad de audio del audio codificado puede verse afectada negativamente, especialmente para un material transitorio relativamente difícil de codificar. Este impacto en la calidad de audio debe principalmente al número limitado de bits de datos disponibles para codificar de manera precisa estos tipos de señales. Los artefactos de codificación de transitorios se manifiestan como una reducción en la definición de la señal transitoria así como en el artefacto de "pre-ruido transitorio", que propaga ruido audible por toda la ventana de codificación debido a errores de cuantificación de codificación.

Tal y como se ha descrito anteriormente y en las figuras 5 y 6, la BED proporciona un procesamiento de pre-ruido transitorio. La codificación E-AC-3 incluye procesamiento de pre-ruido transitorio de codificación para reducir artefactos de pre-ruido transitorios que puedan introducirse cuando el audio que contiene transitorios se codifica sustituyendo el segmento de audio apropiado por audio que se sintetiza usando el audio situado antes del pre-ruido transitorio. El audio se procesa usando síntesis de escalado de tiempo, por lo que su duración aumenta para que tenga una longitud apropiada para sustituir el audio que contiene el pre-ruido transitorio. La memoria intermedia de síntesis de audio se analiza usando un análisis de escena de audio y un procesamiento de similitud de máximos y después se escala en el tiempo de manera que su duración aumenta lo suficiente como para sustituir el audio que contiene el pre-ruido transitorio. El audio sintetizado de longitud aumentada se usa para sustituir el pre-ruido transitorio y se atenúa de manera cruzada en el pre-ruido transitorio existente justo antes de la ubicación del transitorio para garantizar una transición gradual desde el audio sintetizado hasta datos de audio codificados originalmente. Usando el procesamiento de pre-ruido transitorio, la longitud del pre-ruido transitorio puede reducirse drásticamente o suprimirse, incluso cuando la conmutación de bloque está inhabilitada.

En una realización de codificador E-AC-3, el análisis de síntesis de escalado en el tiempo y el procesamiento para la herramienta de procesamiento de pre-ruido transitorio se llevan a cabo en datos de dominio de tiempo para determinar información de metadatos, por ejemplo parámetros de escalado en el tiempo. La información de

metadatos es aceptada por el descodificador junto con el flujo de bits codificado. Los metadatos de pre-ruido transitorios transmitidos se usan para llevar a cabo un procesamiento de dominio de tiempo en el audio descodificado para reducir o eliminar el pre-ruido transitorio introducido por la codificación de audio de baja velocidad binaria en bajas velocidades de datos.

5 El codificador E-AC-3 lleva a cabo un análisis de síntesis de escalado en el tiempo y determina parámetros de escalado en el tiempo, en función del contenido de audio, para cada transitorio detectado. Los parámetros de escalado en el tiempo se transmiten como metadatos adicionales junto con los datos de audio codificados.

10 En un descodificador E-AC-3, los parámetros óptimos de escalado en el tiempo proporcionados en los metadatos E-AC-3 se aceptan como parte de los metadatos E-AC-3 aceptados para usarse en el procesamiento de pre-ruido transitorio. El descodificador lleva a cabo una división de memoria intermedia de audio y una atenuación cruzada usando los parámetros de escalado de tiempo transmitidos obtenidos a partir de los metadatos E-AC-3.

15 Usando la información óptima de escalado en el tiempo y aplicándola con el procesamiento de atenuación cruzada apropiado, el pre-ruido transitorio introducido por la codificación de audio de baja velocidad binaria puede reducirse drásticamente o suprimirse en la descodificación.

20 Por tanto, el procesamiento de pre-ruido transitorio sobrescribe el pre-ruido con un segmento de audio que es muy parecido al contenido original. Las instrucciones de procesamiento de pre-ruido transitorio, cuando se ejecutan, mantienen una memoria intermedia de retardo de cuatro bloques para su uso en el copiado. Las instrucciones de procesamiento de pre-ruido transitorio, cuando se ejecutan, en caso de que se produzca una sobreescritura, llevan a cabo una atenuación cruzada dentro y fuera del pre-ruido sobrescrito.

25 Mezclado descendente

Denote $N.n$ el número de canales codificados en el flujo de bits E-AC-3, donde N es el número de canales principales y $n=0$ ó 1 es el número de canales LFE. Frecuentemente se desea mezclar de manera descendente los N canales principales a un número inferior, denotado como M , de canales principales de salida. Realizaciones de la presente invención soportan el mezclado descendente de N a M canales, $M < N$. El mezclado ascendente también es posible, en cuyo caso $M > N$.

30 Por tanto, en la implementación más general, las realizaciones de descodificador de audio se hacen funcionar para descodificar datos de audio que incluyen $N.n$ canales de datos de audio codificados en datos de audio descodificados que incluyen $M.m$ canales de audio descodificado, y $M \geq 1$, donde n, m indican el número de canales LFE en la entrada, salida respectivamente. El mezclado descendente es el caso en que $M < N$ y según un conjunto de coeficientes de mezclado descendente se incluye en el caso en que $M < N$.

40 Mezclado descendente de dominio de frecuencia frente a dominio de tiempo

El mezclado descendente puede realizarse completamente en el dominio de frecuencia, antes de la transformada inversa, en el dominio de tiempo después de la transformada inversa, pero en el caso de un procesamiento de bloques de solapamiento y suma antes de las operaciones de división en ventanas y de solapamiento y suma, o en el dominio de tiempo después de la operación de división en ventanas y de solapamiento y suma.

45 El mezclado descendente en el dominio de frecuencia (FD) es mucho más eficaz que el mezclado descendente en el dominio de tiempo. Su eficacia se debe, por ejemplo, a que cualquier etapa de procesamiento posterior a la etapa de mezclado descendente solo se lleva a cabo en el número de canales restante, que es generalmente inferior después del mezclado descendente. Por tanto, la complejidad computacional de todas las etapas de procesamiento posteriores a la etapa de mezclado descendente se reduce en al menos la relación de canales de entrada con respecto a los canales de salida.

50 Como un ejemplo, considérese un mezclado descendente de 5.0 canales a estéreo. En este caso, la complejidad computacional de cualquier etapa de procesamiento posterior se reducirá aproximadamente en un factor de $5/2 = 2,5$.

55 El mezclado descendente en el dominio de tiempo (TD) se usa en descodificadores E-AC-3 típicos y en las realizaciones descritas anteriormente e ilustradas en las figuras 5A y 6. Hay tres razones principales por las que los descodificadores E-AC-3 típicos usan el mezclado descendente en el dominio de tiempo:

60 Canales con diferentes tipos de bloques

Dependiendo del contenido de audio que vaya a codificarse, un codificador E-AC-3 puede elegir entre dos tipos de bloque diferentes, bloque corto y bloque largo, para segmentar los datos de audio. Datos de audio armónicos que cambian lentamente se segmentan y codifican normalmente usando bloques largos, mientras que las señales transitorias se segmentan y codifican en bloques cortos. Como resultado, la representación de bloques cortos y

bloques largos en el dominio de frecuencia es intrínsecamente diferente y no puede combinarse en una operación de mezclado descendente en el dominio de frecuencia.

5 Solamente después de que el descodificador haya deshecho las etapas de codificación específicas de tipo de bloque, los canales pueden mezclarse entre sí. Por tanto, en el caso de transformadas conmutadas por bloques se usa un proceso de transformada inversa parcial y los resultados de las dos transformadas diferentes no pueden combinarse directamente hasta justo antes de la fase de ventana.

10 Sin embargo, se conocen métodos para convertir primero datos de transformada de corta longitud en datos de dominio de frecuencia más largos, en cuyo caso el mezclado descendente puede llevarse a cabo en el dominio de frecuencia. Sin embargo, en la mayoría de implementaciones de descodificador conocidas, el mezclado descendente se lleva a cabo tras la transformación inversa según coeficientes de mezclado descendente.

15 Mezclado ascendente

Si el número de canales principales de salida es mayor que el número de canales principales de entrada, $M > N$, un enfoque de mezclado en el dominio de tiempo resulta beneficioso, ya que esto lleva la etapa de mezclado ascendente al final del procesamiento, reduciendo el número de canales en procesamiento.

20 TPNP

Los bloques que están sujetos a un procesamiento de pre-ruido transitorio (TPNP) no pueden mezclarse de manera descendente en el dominio de frecuencia, ya que el TPNP funciona en el dominio de tiempo. El TPNP requiere un histórico de hasta cuatro bloques de datos PCM (1024 muestras), que deben estar presentes para el canal en que se aplica el TPNP. Por tanto, pasar al mezclado descendente en el dominio de tiempo es necesario para rellenar el histórico de datos PCM y llevar a cabo la sustitución de pre-ruido.

30 Mezclado descendente híbrido que usa un mezclado descendente tanto en el dominio de frecuencia como en el dominio de tiempo

Los inventores reconocen que los canales en la mayoría de señales de audio codificadas usan el mismo tipo de bloque más del 90% del tiempo. Eso significa que el mezclado descendente de dominio de frecuencia más eficaz se utilizará para más del 90% de los datos en audio codificado típico, suponiendo que no hay TPNP. El 10% restante, o menos, requerirá un mezclado descendente en el dominio de tiempo, como sucede en los descodificadores E-AC-3 típicos de la técnica anterior.

40 Realizaciones de la presente invención incluyen una lógica de selección de método de mezclado descendente para determinar bloque a bloque qué método de mezclado descendente aplicar, y tanto una lógica de mezclado descendente en el dominio de tiempo como una lógica de mezclado descendente en el dominio de frecuencia para aplicar el método de mezclado descendente particular según sea apropiado. Por tanto, una realización de método incluye determinar bloque a bloque si hay que aplicar un mezclado descendente de dominio de frecuencia o un mezclado descendente de dominio de tiempo. La lógica de selección de método de mezclado descendente funciona para determinar si hay que aplicar un mezclado descendente en el dominio de frecuencia o un mezclado descendente en el dominio de tiempo, e incluye determinar si hay algún procesamiento de pre-ruido transitorio y determinar si alguno de los N canales tiene un tipo de bloque diferente. La lógica de selección determina que el mezclado descendente en el dominio de frecuencia solo se aplica para un bloque que tenga el mismo tipo de bloque en los N canales, que no hay ningún procesamiento de pre-ruido transitorio y que $M < N$.

50 La figura 5B muestra un diagrama de bloques simplificado de una realización de un módulo de descodificación de sección de procesamiento 520 implementado como un conjunto de instrucciones almacenadas en una memoria que, cuando se ejecutan, hacen que se lleve a cabo el procesamiento BED. La figura 5B también muestra un pseudocódigo de instrucciones para el módulo de descodificación de sección de procesamiento 520. El módulo BED 520 incluye los módulos mostrados en la figura 5A que solo usan mezclado descendente en el dominio de tiempo, y los módulos adicionales subsiguientes, incluyendo cada uno instrucciones, donde algunas de tales instrucciones son instrucciones que definen:

60 Módulo de selección de método de mezclado descendente que comprueba (i) cambios de tipo de bloque, (ii) si no hay un mezclado descendente verdadero ($M < N$), sino un mezclado ascendente, y (iii) si el bloque está sujeto a TPNP, y si no se cumple ninguno de estos casos, selecciona el mezclado descendente de dominio de frecuencia. Este módulo determina bloque a bloque si hay que aplicar un mezclado descendente en el dominio de frecuencia o un mezclado descendente en el dominio de tiempo.

65 Módulo de mezclado descendente en el dominio de frecuencia que lleva a cabo, después de la desnormalización de las mantisas mediante los exponentes, un mezclado descendente en el dominio de frecuencia. Debe observarse que el módulo de mezclado descendente en el dominio de frecuencia incluye además un módulo de lógica de transición de dominio de tiempo a dominio de frecuencia que comprueba si el bloque anterior usó mezclado descendente en el

domino de tiempo, en cuyo caso el bloque se trata de diferente manera a la descrita posteriormente en mayor detalle. Además, el módulo de lógica de transición también se encarga además de etapas de procesamiento asociadas a determinados eventos que no se producen regularmente, por ejemplo cambios de programa tales como canales de salida atenuados.

5 Módulo de lógica de transición de mezclado descendente de FD a TD que comprueba si el bloque anterior usó mezclado descendente en el dominio de frecuencia, en cuyo caso el bloque se trata de diferente manera a la descrita posteriormente en mayor detalle. Además, el módulo de lógica de transición también se encarga de etapas de procesamiento asociadas a determinados eventos que no se producen regularmente, por ejemplo cambios de programa tales como canales de salida atenuados.

Además, los módulos que están en la figura 5A pueden tener un comportamiento diferente en realizaciones que incluyen un mezclado descendente híbrido, es decir, un mezclado descendente tanto FD como TD, dependiendo de una o más condiciones para el bloque actual.

15 Haciendo referencia al pseudocódigo de la figura 5B, algunas realizaciones del método de descodificación de sección de procesamiento incluyen, después de transferir los datos de una trama de bloques desde la memoria externa, determinar si hay que aplicar un mezclado descendente FD o un mezclado descendente TD. Para un mezclado descendente FD, para cada canal, el método incluye (i) aplicar un control de margen dinámico y normalización de diálogo pero, como se describe posteriormente, inhabilitar el ajuste de ganancia; (ii) desnormalizar mantisas mediante exponentes; (iii) llevar a cabo un mezclado descendente FD; y (iv) determinar si hay canales de salida atenuados o si el bloque anterior se mezcló de manera descendente mediante un mezclado descendente en el dominio de tiempo, en cuyo caso el procesamiento se lleva a cabo de diferente manera a la descrita posteriormente en mayor detalle. En el caso del mezclado descendente TD, y también para datos mezclados de manera descendente en el dominio de frecuencia, el proceso incluye para cada canal: (i) procesar de diferente manera bloques que van a mezclarse de manera descendente en el dominio de tiempo en caso de que el bloque anterior se haya mezclado de manera descendente en el dominio de frecuencia, además de tratar cualquier cambio de programa; (ii) determinar la transformada inversa; (iii) llevar a cabo operaciones de visión en ventanas y de solapamiento y suma; y en caso de mezclado descendente TD, (iv) llevar a cabo cualquier TPNP y mezclado descendente en el canal de salida apropiado.

La figura 7 muestra un diagrama de flujo de datos simple. El bloque 701 corresponde a la lógica de selección de método de mezclado descendente que comprueba las tres condiciones: cambio de tipo de bloque, TPNP o mezclado ascendente, y si se cumple alguna condición, dirige el flujo de datos hacia una rama de mezclado descendente TD 721 que incluye en 723 una lógica de transición de mezclado descendente FD para procesar de diferente manera un bloque que se produce inmediatamente después de un bloque procesado mediante un mezclado descendente FFD, procesamiento de cambio de programa, y en 725 desnormaliza las mantisas mediante los exponentes. Tras el bloque 721, el flujo de datos se procesa mediante un bloque de procesamiento común 731. Si las comprobaciones del bloque de lógica de selección de método de mezclado descendente 701 determinan que el bloque es para un mezclado descendente FD, el flujo de datos se dirige hacia un procesamiento de mezclado descendente FD 711 que incluye un proceso de mezclado descendente de dominio de frecuencia 713 que inhabilita el ajuste de ganancia y, para cada canal, desnormaliza las mantisas mediante los exponentes y lleva a cabo un mezclado descendente FD, y un bloque de lógica de transición de mezclado descendente TD 715 para determinar si el bloque anterior se procesó mediante un mezclado descendente TD, para procesar tal bloque de diferente manera y también para detectar y tratar cualquier cambio en el programa, tales como canales de salida atenuados. Tras el bloque de transición de mezclado descendente TD 715, el flujo de datos se dirige al bloque de procesamiento común 731.

El bloque de procesamiento común 731 incluye una transformación inversa y cualquier procesamiento adicional en el dominio de tiempo. El procesamiento adicional en el dominio de tiempo incluye deshacer el ajuste de ganancia y un procesamiento de división en ventanas y de solapamiento y suma. Si el bloque procede del bloque de mezclado descendente TD 721, el procesamiento adicional en el dominio de tiempo incluye además cualquier procesamiento TPNP y un mezclado descendente en el dominio de tiempo.

La figura 8 muestra un diagrama de flujo de una realización de procesamiento para un módulo de descodificación de sección de procesamiento tal como el mostrado en la figura 7. El diagrama de flujo está dividido de la siguiente manera, con los mismos números de referencia usados en la figura 7 para bloques de flujo de datos funcionales respectivos similares: una sección de lógica de selección de método de mezclado descendente 701 en la que se usa una bandera lógica FD_dmx que cuando vale 1 indica que se usa un mezclado descendente de dominio de frecuencia para el bloque; una sección de lógica de mezclado descendente TD 721 que incluye una sección de lógica de cambio de programa y de lógica de transición de mezclado descendente FD 723 para procesar de diferente manera un bloque que se produce inmediatamente después de un bloque procesado mediante un mezclado descendente FD y para llevar a cabo un procesamiento de cambio de programa, y una sección para desnormalizar las mantisas mediante los exponentes para cada canal de entrada. Tras el bloque 721, el flujo de datos se procesa mediante una sección de procesamiento común 731. Si el bloque de lógica de selección de método de mezclado descendente 701 determina que el bloque es para un mezclado descendente FD, el flujo de

datos se dirige hacia una sección de procesamiento de mezclado descendente FD 711 que incluye un proceso de mezclado descendente de dominio de frecuencia que inhabilita el ajuste de ganancia, y para cada canal, desnormaliza las mantisas mediante los exponentes y lleva a cabo un mezclado descendente FD, y una sección de lógica de transición de mezclado descendente TD 715 para determinar para cada canal del bloque anterior si se produce una atenuación de canal o si el bloque anterior se procesó mediante mezclado descendente TD, y para procesar tal bloque de diferente manera. Tras la sección de transición de mezclado descendente TD 715, el flujo de datos se dirige hacia la sección de lógica de procesamiento común 731. La sección de lógica de procesamiento común 731 incluye para cada canal una transformación inversa y cualquier procesamiento adicional en el dominio de tiempo. El procesamiento adicional en el dominio de tiempo incluye deshacer el ajuste de ganancia y un procesamiento de división en ventanas y de solapamiento y suma. Si FD_dmx vale 0, lo que indica un mezclado descendente TD, el procesamiento adicional de dominio de tiempo en 731 incluye además cualquier procesamiento TPNP y un mezclado descendente en el dominio de tiempo.

Debe observarse que después del mezclado descendente FD, en la sección de lógica de transición de mezclado descendente TD 715, en 817, el número de canales de entrada N se fija que para que sea el mismo que el número de canales de salida M, de manera que el resto del procesamiento, por ejemplo el procesamiento en la sección de lógica de procesamiento común 713, se lleva a cabo solamente en los datos mezclados de manera descendente. Esto reduce la cantidad de cálculo. Evidentemente, el mezclado descendente de dominio de tiempo de los datos del bloque anterior cuando hay una transición desde un bloque que se mezcló de manera descendente en el dominio de tiempo, tal como el mezclado descendente TD mostrado como 819 en la sección 715, se lleva a cabo en todos los N canales de entrada implicados en el mezclado descendente.

Tratamiento de transiciones

En la decodificación es necesario tener transiciones graduales entre bloques de audio. La norma E-AC-3 y otros muchos métodos de codificación usan una transformada solapada, por ejemplo una MDCT con un solapamiento del 50%. Por tanto, cuando se procesa un bloque actual hay un solapamiento del 50% con el bloque anterior y, además, habrá un solapamiento del 50% con el bloque siguiente en el dominio de tiempo. Algunas realizaciones de la presente invención usan lógica de solapamiento y suma que incluye una memoria intermedia de solapamiento y suma. Cuando se procesa un bloque actual, la memoria intermedia de solapamiento y suma contiene datos del bloque de audio anterior. Puesto que es necesario tener transiciones graduales entre bloques de audio, la lógica se incluye para tratar de diferente manera transiciones desde un mezclado descendente TD hasta un mezclado descendente FD, y desde un mezclado descendente FD hasta un mezclado descendente TD.

La figura 9 muestra un ejemplo de procesamiento de cinco bloques, denotados como bloque k, k+1,..., k+4 de cinco canales de audio que incluyen, como es habitual: un canal izquierdo, un canal central, un canal derecho, un canal izquierdo envolvente y un canal derecho envolvente, denotados como L, C, R, LS y RS, respectivamente, y un mezclado descendente hacia una mezcla estéreo, usando la fórmula:

La salida izquierda se denota como $L' = aC + bL + cLS$ y la salida derecha se denota como $R' = aC + bR + cRS$.

La figura 9 supone que se usa una transformada no solapada. Cada rectángulo representa los contenidos de audio de un bloque. El eje horizontal de izquierda a derecha representa los bloques k,..., k+4 y el eje vertical de arriba abajo representa el progreso de decodificación de los datos. Supóngase que el bloque k se procesa mediante un mezclado descendente TD, que los bloques k+1 y k+2 se procesan mediante un mezclado descendente FD y que los bloques k+3 y k+4 se procesan mediante un mezclado descendente TD. Como puede observarse, para cada uno de los bloques de mezclado descendente TD, el mezclado descendente no se produce hasta después del mezclado descendente de dominio de tiempo hacia la parte inferior, tras lo cual los contenidos son los canales L' y R' mezclados de manera descendente, mientras que para el bloque mezclado de manera descendente FD, los canales izquierdo y derecho en el dominio de frecuencia ya se han mezclado de manera descendente después del mezclado descendente en el dominio de frecuencia, y los datos de los canales C, LS y RS se ignoran. Puesto que no hay ningún solapamiento entre bloques, no es necesario tratar ningún caso espacial cuando se conmuta desde un mezclado descendente TD a un mezclado descendente FD o desde un mezclado descendente FD a un mezclado descendente TD.

La figura 10 describe el caso de transformadas solapadas un 50%. Supóngase que el solapamiento y la suma se llevan a cabo mediante decodificación de solapamiento y suma usando una memoria intermedia de solapamiento y suma. En este diagrama, cuando el bloque de datos se muestra como dos triángulos, el triángulo inferior izquierdo son datos en la memoria intermedia de solapamiento y suma del bloque anterior, mientras que el triángulo superior derecho muestra los datos del bloque actual.

Tratamiento de una transición desde un mezclado descendente TD hasta un mezclado descendente FD

Considérese que el bloque k+1 es un bloque de mezclado descendente FD que sigue inmediatamente a un bloque de mezclado descendente TD. Después del mezclado descendente TD, la memoria intermedia de solapamiento y suma contiene los datos L, C, R, LS y RS del último bloque que deben incluirse en el bloque actual. También se

incluye la contribución del bloque actual $k+1$, ya mezclada de manera descendente en el dominio de frecuencia. Para determinar correctamente los datos PCM mezclados de manera descendente para la salida, es necesario incluir los datos tanto del bloque actual como del bloque anterior. Para ello, los datos del bloque anterior deben obtenerse y, puesto que aún no se han mezclado de manera descendente, mezclarse de manera descendente en el dominio de tiempo. Deben añadirse las dos contribuciones para determinar los datos PCM mezclados de manera descendente para la salida. Este procesamiento está incluido en la lógica de transición de mezclado descendente TD 715 de las figuras 7 y 8, y mediante el código en la lógica de transición de mezclado descendente TD incluida en el módulo de mezclado descendente FD mostrado en la figura 5B. El procesamiento llevado a cabo en el mismo está resumido en la sección de lógica de transición de mezclado descendente TD 715 de la figura 8. En mayor detalle, el tratamiento de una transición desde un mezclado descendente TD a un mezclado descendente FD incluye:

- Vaciar las memorias intermedias de solapamiento introduciendo ceros en la lógica de solapamiento y suma y llevando a cabo operaciones de división en ventanas y de solapamiento y suma. Copiar la salida obtenida de la lógica de solapamiento y suma. Estos son los datos PCM del bloque anterior del canal particular antes del mezclado descendente. La memoria intermedia de solapamiento incluye ahora ceros.

- Mezclar de manera descendente en el dominio de tiempo los datos PCM de las memorias intermedias de solapamiento para generar datos PCM del mezclado descendente TD del bloque anterior.

- Mezclar de manera descendente en el dominio de frecuencia los nuevos datos del bloque actual. Llevar a cabo la transformada inversa e introducir nuevos datos después del mezclado descendente FD y la transformada inversa en la lógica de solapamiento y suma. Llevar a cabo operaciones de división en ventanas y de solapamiento y suma, y así sucesivamente, con los nuevos datos para generar datos PCM del mezclado descendente FD del bloque actual.

- Añadir los datos PCM del mezclado descendente TD y del mezclado descendente FD para generar una salida PCM.

Debe observarse que en una realización alternativa, suponiendo que no se ha aplicado TPNP en el bloque anterior, los datos de las memorias intermedias de solapamiento y suma se mezclan de manera descendente y después se lleva a cabo una operación de solapamiento y suma en los canales de salida mezclados de manera descendente. Esto evita la necesidad de llevar a cabo una operación de solapamiento y suma en cada canal del bloque anterior. Además, como se ha descrito anteriormente para la decodificación AC-3, cuando una memoria intermedia de mezclado descendente y su memoria intermedia de retardo correspondiente de medio bloque con una longitud de 128 muestras se usan, se dividen en ventanas y se combinan para producir 256 muestras de salida PCM, la operación de mezclado descendente es más sencilla ya que la memoria intermedia de retardo solo tiene 128 muestras en lugar de 256. Este aspecto reduce la máxima complejidad computacional intrínseca al procesamiento de transición. Por lo tanto, en algunas realizaciones, para un bloque particular mezclado de manera descendente en el dominio de frecuencia que sigue a un bloque cuyos datos se mezclaron de manera descendente en el dominio de tiempo, el procesamiento de transición incluye aplicar un mezclado descendente en el dominio de pseudo-tiempo en los datos del bloque anterior que van a solaparse con los datos descodificados del bloque particular.

Tratamiento de una transición desde un mezclado descendente FD a un mezclado descendente TD

Considérese que el bloque $k+3$ es un bloque de mezclado descendente TD que sigue inmediatamente a un bloque de mezclado descendente FD $k+2$. Puesto que el bloque anterior era un bloque de mezclado descendente FD, la memoria intermedia de solapamiento y suma en las fases anteriores, por ejemplo antes del mezclado descendente TD, contiene los datos mezclados de manera descendente en los canales izquierdo y derecho y ningún dato en los otros canales. Las contribuciones del bloque actual no se mezclan de manera descendente hasta después del mezclado descendente TD. Para determinar correctamente los datos PCM mezclados de manera descendente para la salida, es necesario incluir los datos tanto del bloque actual como del bloque anterior. Para ello, deben obtenerse los datos del bloque anterior. Los datos del bloque actual deben mezclarse de manera descendente en el dominio de tiempo y añadirse a los datos transformados de manera inversa que se obtuvieron para determinar los datos PCM mezclados de manera descendente para la salida. Este procesamiento está incluido en la lógica de transición de mezclado descendente FD 723 de las figuras 7 y 8, y mediante el código en el módulo de lógica de transición de mezclado descendente FD mostrado en la figura 5B. El procesamiento llevado a cabo en el mismo está resumido en la sección de lógica de transición de mezclado descendente FD 723 de la figura 8. En mayor detalle, suponiendo que hay memorias intermedias PCM de salida para cada canal de salida, el tratamiento de una transición desde un mezclado descendente FD a un mezclado descendente TD incluye:

- Vaciar las memorias intermedias de solapamiento introduciendo ceros en la lógica de solapamiento y suma y llevando a cabo operaciones de división en ventanas y de solapamiento y suma. Copiar la salida en la memoria intermedia PCM de salida. Los datos obtenidos son los datos PCM del mezclado descendente FD del bloque anterior. La memoria intermedia de solapamiento incluye ahora ceros.

- Llevar a cabo una transformación inversa de los nuevos datos del bloque actual para generar datos premezclados de manera descendente del bloque actual. Introducir estos nuevos datos de dominio de tiempo (después de la transformada) en la lógica de solapamiento y suma.

5 • Llevar a cabo operaciones de división en ventanas y de solapamiento y suma, TPNP si lo hubiera, y un mezclado descendente TD con los nuevos datos del bloque actual para generar datos PCM del mezclado descendente TD del bloque actual.

10 • Añadir los datos PCM del mezclado descendente TD y del mezclado descendente FD para generar una salida PCM.

Además de las transiciones desde el mezclado descendente de dominio de tiempo al mezclado descendente de dominio de frecuencia, los cambios de programa se tratan en la lógica de transición de mezclado descendente de dominio de tiempo y en un manejador de cambios de programa. Los nuevos canales emergentes se incluyen automáticamente en el mezclado descendente y, por tanto, no necesitan ningún tratamiento especial. Los canales que ya no están presentes en el nuevo programa deben atenuarse. Esto se lleva cabo, como se muestra en la sección 715 de la figura 8 para el caso del mezclado descendente FD, vaciando las memorias intermedias de solapamiento de los canales atenuados. El vaciado se lleva a cabo introduciendo ceros en la lógica de solapamiento y suma y llevando a cabo operaciones de división en ventanas y de solapamiento y suma.

20 Debe observarse que en el diagrama de flujo mostrado y en algunas realizaciones, la sección de lógica de mezclado descendente de dominio de frecuencia 711 incluye inhabilitar la característica de ajuste de ganancia opcional para todos los canales que sean parte del mezclado descendente de dominio de frecuencia. Los canales pueden tener diferentes parámetros de ajuste de ganancia que inducirán un escalado diferente de los coeficientes espectrales de un canal, impidiendo por tanto un mezclado descendente.

En una implementación alternativa, la sección de lógica de mezclado descendente FD 711 se modifica de manera que el mínimo de todas las ganancias se use para llevar a cabo un ajuste de ganancia para un canal mezclado de manera descendente (en el dominio de frecuencia).

30 Mezclado descendente en el dominio de tiempo con coeficientes de mezclado descendente variables y necesidad de una atenuación cruzada explícita

35 El mezclado descendente puede generar varios problemas. Diferentes ecuaciones de mezclado descendente se utilizan en diferentes circunstancias, por lo que puede ser necesario modificar dinámicamente los coeficientes de mezclado descendente en función de las condiciones de señal. Hay parámetros de metadatos disponibles que permiten personalizar los coeficientes de mezclado descendente para obtener resultados óptimos.

40 Por tanto, los coeficientes de mezclado descendente pueden cambiar en el tiempo. Cuando se produce un cambio desde un primer conjunto de coeficientes de mezclado descendente a un segundo conjunto de coeficientes de mezclado descendente, los datos deben atenuarse de manera cruzada desde el primer conjunto al segundo conjunto.

45 Cuando se lleva a cabo un mezclado descendente en el dominio de frecuencia, y también en muchas implementaciones de descodificador, por ejemplo en un descodificador AC-3 de la técnica anterior, como el mostrado en la figura 1, el mezclado descendente se lleva a cabo antes de las operaciones de división en ventanas y de solapamiento y suma. La ventaja de llevar a cabo el mezclado descendente en el dominio de frecuencia, o en el dominio de tiempo antes de las operaciones de división en ventanas y de solapamiento y suma, es que hay una atenuación cruzada intrínseca como resultado de las operaciones de solapamiento y suma. Por tanto, en muchos descodificadores AC-3 y métodos de descodificación conocidos en los que el mezclado descendente se lleva a cabo en el dominio de ventana después de la transformación inversa, o en dominio de frecuencia en las implementaciones de mezclado descendente híbridas, no hay ninguna operación de atenuación cruzada explícita.

55 En el caso de un mezclado descendente en el dominio de tiempo y de un procesamiento de pre-ruido transitorio (TPNP), habrá un retardo de un bloque en procesamiento de pre-ruido transitorio de descodificación provocado por cuestiones de cambio de programa, por ejemplo en un descodificador 7.1 Por tanto, en realizaciones de la presente invención, cuando se lleva a cabo un mezclado descendente en el dominio de tiempo y se usa TPNP, el mezclado descendente en el dominio de tiempo se lleva a cabo después de las operaciones de división en ventanas y de solapamiento y suma. El orden de procesamiento en caso de que use el mezclado descendente en el dominio de tiempo es: llevar a cabo la transformada inversa, por ejemplo MDCT, llevar a cabo operaciones de división en ventanas y de solapamiento y suma, llevar a cabo cualquier procesamiento de pre-ruido transitorio de descodificación (sin retardo) y después el mezclado descendente en el dominio de tiempo.

65 En tal caso, el mezclado descendente en el dominio de tiempo requiere la atenuación cruzada de datos de mezclado descendente anteriores y actuales, por ejemplo coeficientes de mezclado descendente o tablas de mezclado descendente, para garantizar que cualquier cambio en los coeficientes de mezclado descendente sea gradual.

Una opción es llevar a cabo la operación de atenuación cruzada para calcular el coeficiente resultante. Denote $c[i]$ el coeficiente de mezclado a usar, donde i denota el índice de tiempo de 256 muestras en el dominio de tiempo, de manera que el intervalo es $i=0, \dots, 255$. Denote $w^2[i]$ una función de ventana positiva de manera que $w^2[i] + w^2[255 - i] = 1$ para $i=0, \dots, 255$. Denote c_{viejo} el coeficiente de mezclado actualizado anteriormente y c_{nuevo} el coeficiente de mezclado actualizado. La operación de atenuación cruzada a aplicar es:

$$c[i] = w^2[i] \cdot c_{viejo} + w^2[255 - i] \cdot c_{nuevo} \text{ para } i=0, \dots, 255.$$

Después de cada pasada de la operación de atenuación cruzada de coeficientes, los coeficientes viejos se actualizan con los nuevos, lo que se expresa como $c_{viejo} \leftarrow c_{nuevo}$.

En la siguiente pasada, si los coeficientes no se actualizan,

$$c[i] = w^2[i] \cdot c_{nuevo} + w^2[255 - i] \cdot c_{nuevo} = c_{nuevo}.$$

Dicho de otro modo, la influencia del conjunto de coeficientes viejos ha desaparecido completamente.

Los inventores han observado que en muchos flujos de audio y situaciones de mezclado descendente, los coeficientes de mezclado no varían frecuentemente. Para mejorar el rendimiento del proceso de mezclado descendente en el dominio de tiempo, realizaciones del módulo de mezclado descendente en el dominio de tiempo incluyen determinar si los coeficientes de mezclado descendente han cambiado con respecto a su valor anterior, y, si no han cambiado, llevan a cabo un mezclado descendente; en caso contrario, si han cambiado, llevan a cabo una atenuación cruzada de los coeficientes de mezclado descendente según una función de ventana positiva preseleccionada. En una realización, la función de ventana es la misma función de ventana que la usada en las operaciones de división en ventanas y de solapamiento y suma. En otra realización se usa una función de ventana diferente.

La figura 11 muestra un pseudocódigo simplificado para una realización de mezclado descendente. El descodificador para una realización de este tipo usa al menos un procesador x86 que ejecuta instrucciones vectoriales SSE. El mezclado descendente incluye determinar si los nuevos datos de mezclado descendente no han cambiado con respecto a los viejos datos de mezclado descendente. Si es así, el mezclado descendente incluye realizar ajustes para la ejecución de instrucciones vectoriales SSE en al menos un procesador del uno o más procesadores x86 y el mezclado descendente usando los datos de mezclado descendente no modificados, lo que incluye ejecutar al menos una instrucción vectorial SSE. En caso contrario, si los nuevos datos de mezclado descendente han cambiado con respecto a los viejos datos de mezclado descendente, el método incluye determinar datos de mezclado descendente atenuados de manera cruzada mediante una operación de atenuación cruzada.

Supresión del procesamiento de datos innecesarios

En algunas situaciones de mezclado descendente hay al menos un canal que no contribuye en la salida mezclada de manera descendente. Por ejemplo, en muchos casos de mezclado descendente de audio 5.1 a estéreo, el canal LFE no se incluye, de manera que el mezclado descendente es de 5.1 a 2.0. La exclusión del canal LFE en el mezclado descendente puede ser intrínseca al formato de codificación, como es el caso de AC-3, o puede controlarse por los metadatos, como es el caso de E-AC-3. En E-AC-3, el parámetro `lfemixlevcode` determina si el canal LFE está incluido o no en la mezcla descendente. Cuando el parámetro `lfemixlevcode` vale 0, el canal LFE no está incluido en el mezclado descendente.

Debe recordarse que el mezclado descendente puede llevarse a cabo en el dominio de frecuencia, en el dominio de pseudo-tiempo después de la transformación inversa pero antes de las operaciones de división en ventanas y de solapamiento y suma, o en el dominio de tiempo después de la transformación inversa y después de las operaciones de división en ventanas y de solapamiento y suma. Un mezclado descendente en el dominio de tiempo puro se lleva a cabo en muchos descodificadores E-AC-3 conocidos, y en algunas realizaciones de la presente invención, y es ventajoso, por ejemplo, debido a la presencia de TPNP; un mezclado descendente en el dominio de pseudo-tiempo se lleva a cabo en muchos descodificadores AC-3 y en algunas realizaciones de la presente invención, y es ventajoso porque la operación de solapamiento y suma proporciona una atenuación cruzada intrínseca que es ventajosa cuando cambian los coeficientes de mezclado descendente; y un mezclado descendente en el dominio de frecuencia se lleva a cabo en algunas realizaciones de la presente invención cuando las condiciones lo permiten.

Tal y como se describe en este documento, el mezclado descendente en el dominio de frecuencia es el método de mezclado descendente más eficaz, ya que minimiza el número de transformadas inversas y de operaciones de división en ventanas y de solapamiento y suma necesarias para producir una salida de dos canales a partir de una

entrada de 5.1 canales. En algunas realizaciones de la presente invención, cuando se lleva a cabo un mezclado descendente FD, por ejemplo, en la figura 8, en la sección de bucle de mezclado descendente FD 711 en el bucle que empieza por el elemento 813, termina en 814 y se incrementa en 815 al siguiente canal, estos canales no incluidos no incluidos en el mezclado descendente se excluyen en el procesamiento.

5 El mezclado descendente en el dominio de pseudo-tiempo después de la transformada inversa pero antes de las operaciones de división en ventanas y de solapamiento y suma, o en el dominio de tiempo después de la transformada inversa y las operaciones de división en ventanas y de solapamiento y suma es menos eficaz computacionalmente que en el dominio de frecuencia. En muchos descodificadores actuales, tales como los
10 descodificadores AC-3 actuales, el mezclado descendente se lleva a cabo en el dominio de pseudo-tiempo. La operación de transformada inversa se lleva a cabo de manera independiente de la operación de mezclado descendente, por ejemplo, en módulos distintos. La transformada inversa en tales descodificadores se lleva a cabo en todos los canales de entrada. Esto es relativamente ineficaz desde un punto de vista computacional porque, en caso de no incluirse el canal LFE, la transformada inversa sigue llevándose a cabo para este canal. Este
15 procesamiento innecesario es significativo porque, aunque el canal LFE tiene un ancho de banda limitado, aplicar la transformada inversa al canal LFE requiere un cálculo equivalente a aplicar la transformada inversa a cualquier canal de ancho de banda completo. Los inventores han reconocido esta poca eficiencia. Algunas realizaciones de la presente invención incluyen identificar uno o más canales no contribuyentes de los N.n canales de entrada, siendo un canal no contribuyente un canal que no contribuye en los M.m canales de salida de audio codificado. En algunas
20 realizaciones, la identificación usa información, por ejemplo, metadatos, que define el mezclado descendente. En el ejemplo de mezclado descendente de 5.1 a 2.0, el canal LFE es identificado como un canal no contribuyente. Algunas realizaciones de la invención incluyen llevar a cabo una transformación de frecuencia a tiempo en cada canal que contribuya en los M.m canales de salida, y no llevan a cabo ninguna transformación de frecuencia a tiempo en cada canal identificado que no contribuya en la señal de M.m canales. En el ejemplo de 5.1 a 2.0 en el
25 que el canal LFE no contribuye en el mezclado descendente, la transformada inversa, por ejemplo una IMCDT, solo se lleva a cabo en los cinco canales de ancho de banda completo, de manera que la parte de transformada inversa se lleva a cabo con una reducción aproximada del 16% de los recursos computacionales requeridos para los 5.1 canales. Puesto que la IMDCT es una fuente importante de complejidad computacional en el método de descodificación, esta reducción puede ser significativa.

30 En muchos descodificadores actuales, tales como los descodificadores E-AC-3 actuales, el mezclado descendente se lleva a cabo en el dominio de tiempo. La operación de transformada inversa y las operaciones de solapamiento y suma se llevan a cabo antes de cualquier TPNP y antes del mezclado descendente, independientemente de la operación de mezclado descendente, por ejemplo en módulos diferentes. La transformada inversa y las operaciones
35 de división en ventanas y de solapamiento y suma en tales descodificadores se llevan a cabo en todos los canales de entrada. Esto es relativamente ineficaz desde un punto de vista computacional porque, en caso de no incluirse el canal LFE, la transformada inversa y las operaciones de división en ventanas y de solapamiento y suma siguen llevándose a cabo para este canal. Este procesamiento innecesario es significativo porque, aunque el canal LFE tiene un ancho de banda limitado, aplicar la transformada inversa y las operaciones de solapamiento y suma al canal
40 LFE requiere un cálculo equivalente a aplicar la transformada inversa y las operaciones de división en ventanas y de solapamiento y suma a cualquier canal de ancho de banda completo. En algunas realizaciones de la presente invención, el mezclado descendente se lleva a cabo en el dominio de tiempo, y en otras realizaciones el mezclado descendente puede llevarse a cabo en el dominio de tiempo dependiendo del resultado de aplicar la lógica de selección de método de mezclado descendente. Algunas realizaciones de la presente invención en las que se usa el
45 mezclado descendente TD incluyen identificar uno o más canales no contribuyentes de los N.n canales de entrada. En algunas realizaciones, la identificación usa información, por ejemplo metadatos, que define el mezclado descendente. En el ejemplo de mezclado descendente de 5.1 a 2.0, el canal LFE es identificado como un canal no contribuyente. Algunas realizaciones de la invención incluyen llevar a cabo una transformada inversa, es decir, una transformación de frecuencia a tiempo en cada canal que contribuya en los M.m canales de salida, y no llevan a cabo ninguna transformación de frecuencia a tiempo ni otro procesamiento de dominio de tiempo en cada canal
50 identificado que no contribuya en la señal de M.m canales. En el ejemplo de 5.1 a 2.0 en el que el canal LFE no contribuye en el mezclado descendente, la transformada inversa, por ejemplo una IMCDT, las operaciones de solapamiento y suma y el TPNP solo se llevan a cabo en los cinco canales de ancho de banda completo, de manera que la transformada inversa y las operaciones de división en ventanas y de solapamiento y suma se llevan a cabo con una reducción aproximada del 16% de los recursos computacionales requeridos para los 5.1 canales. En el
55 diagrama de flujo de la figura 8, en la sección de lógica de procesamiento común 731, una característica de algunas realizaciones incluye que el procesamiento en el bloque que empieza por el elemento 833, sigue con el 834 y que incluye el incremento hasta el siguiente elemento de canal 835, se lleva a cabo para todos los canales excepto los canales no contribuyentes. Esto sucede intrínsecamente en un bloque que se haya mezclado de manera
60 descendente en el dominio de frecuencia.

Aunque en algunas realizaciones el canal LFE es un canal no contribuyente, es decir no se incluye en los canales de salida de mezclado descendente, como es común en AC-3 y en E-AC-3, en otras realizaciones un canal diferente al LFE es también, o en cambio, un canal no contribuyente y no se incluye en la salida mezclada de manera
65 descendente. Algunas realizaciones de la invención incluyen comprobar tales condiciones para identificar qué uno o más canales, si los hubiera, son no contribuyentes para que tales canales no se incluyan en el mezclado

descendente y, en el caso de mezclado descendente en el dominio de tiempo, no llevan a cabo el procesamiento a través de la transformada inversa ni operaciones de división en ventanas y de solapamiento y suma para cualquier canal no contribuyente identificado.

5 Por ejemplo, en AC-3 y en E-AC-3, hay determinadas condiciones en las que los canales envolventes y/o el canal central no se incluyen en los canales de salida mezclados de manera descendente. Estas condiciones se definen mediante metadatos incluidos en el flujo de bits codificado que toman valores predefinidos. Los metadatos, por ejemplo, pueden incluir información que define el mezclado descendente, incluyendo parámetros de nivel de mezclado.

10 A continuación se describen, con fines ilustrativos, algunos ejemplos de tales parámetros de nivel de mezclado para el caso de E-AC-3. En el mezclado descendente a estéreo en E-AC-3 se proporcionan dos tipos de mezclado descendente: mezclado descendente hacia un par estéreo codificado envolvente de matriz LtRt y mezclado descendente hacia una señal estéreo convencional, LoRo. La señal estéreo mezclada de manera descendente (LoRo o LtRt) puede mezclarse además para obtener una señal mono. Un código de nivel de mezclado envolvente LtRt de 3 bits denotado como ltrtsurmixlev, y un código de nivel de mezclado envolvente LoRo de 3 bits denotado como lorosurmixlev indican el nivel de mezclado descendente nominal de los canales envolventes con respecto a los canales izquierdo y derecho en un mezclado descendente LtRt o LoRo, respectivamente. El valor binario '111' indica un nivel de mezclado descendente de 0, es decir, $-\infty$ dB. Códigos de nivel de mezclado central LtRt y LoRo de 3 bits denotados como ltrcmixlev, lorocmixlev, indican el nivel de mezclado descendente nominal del canal central con respecto a los canales izquierdo y derecho en un mezclado descendente LtRt y LoRo, respectivamente. El valor binario '111' indica un nivel de mezclado descendente de 0, es decir, $-\infty$ dB.

25 Hay condiciones en las que los canales envolventes no se incluyen en los canales de salida mezclados de manera descendente. En E-AC-3, estas condiciones se identifican mediante metadatos. Estas condiciones incluyen los casos en los que surmixlev='10' (solamente en AC-3), ltrtsurmixlev='111' y lorosurmixlev='111'. Para estas condiciones, en algunas realizaciones, un decodificador incluye usar los metadatos de nivel de mezclado para identificar que tales metadatos indican que los canales envolventes no están incluidos en el mezclado descendente, y no procesa los canales envolventes a través de la fase de transformada inversa ni de la fase de división en ventanas/solapamiento y suma. Además, hay condiciones en las que el canal central no se incluye en los canales de salida mezclados de manera descendente, identificadas mediante ltrcmixlevel='111' y lorocmixlev='111'. Para estas condiciones, en algunas realizaciones, un decodificador incluye usar los metadatos de nivel de mezclado para identificar que tales metadatos indican que el canal central no está incluido en el mezclado descendente, y no procesa el canal central a través de la fase de transformada inversa ni de la fase de división en ventanas/solapamiento y suma.

40 En algunas realizaciones, la identificación de uno o más canales no contribuyentes depende del contenido. Como un ejemplo, la identificación incluye identificar si uno o más canales tienen una cantidad insignificante de contenido con respecto a uno o más otros canales. Se usa una medida de cantidad de contenido. En una realización, la medida de cantidad de contenido es energía, mientras que en otra realización la medida de cantidad de contenido es el nivel absoluto. La identificación incluye comparar la diferencia de la medida de cantidad de contenido entre pares de canales con un umbral ajustable. Como un ejemplo, en una realización, identificar uno o más canales no contribuyentes incluye determinar si la cantidad de contenido de canal envolvente de un bloque es menor que cada cantidad de contenido de canal delantero en al menos un umbral ajustable para determinar si el canal envolvente es un canal no contribuyente.

50 De manera ideal, el umbral se selecciona para que sea lo más bajo posible sin introducir artefactos notables en la versión mezclada de manera descendente de la señal para maximizar la identificación de canales no contribuyentes para reducir la cantidad de cálculo requerido, minimizando al mismo tiempo la pérdida de calidad. En algunas realizaciones se proporcionan diferentes umbrales para diferentes aplicaciones de decodificación, con la opción de un umbral para una aplicación de decodificación particular que representa un equilibrio aceptable entre la calidad de mezclado descendente (umbrales superiores) y la reducción de complejidad computacional (umbrales inferiores) para la aplicación específica.

55 En algunas realizaciones de la presente invención, un canal se considera insignificante con respecto a otro canal si su energía o nivel absoluto es al menos 15 dB inferior a los del otro canal. De manera ideal, un canal es insignificante con respecto a otro canal si su energía o nivel absoluto es al menos 25 dB inferior a los del otro canal.

60 Usar un umbral para la diferencia entre dos canales, denotados como A y B, que es equivalente a 25dB es casi lo mismo que decir que el nivel de la suma de los valores absolutos de los dos canales no supera los 0,5 dB del nivel del canal dominante. Es decir, si el canal A está a -6 dBFS (dB con respecto a la escala total) y el canal B está a -31 dBFS, la suma de los valores absolutos de los canales A y B será de -5,5 dBFS aproximadamente, o casi 0,5 dB mayor que el nivel del canal A.

65 Si el audio tiene una calidad relativamente baja, y en aplicaciones de bajo coste, puede ser aceptable sacrificar la calidad para reducir la complejidad; el umbral puede ser inferior a 25 dB. En un ejemplo se usa un umbral de 18 dB.

En tal caso, la suma de los dos canales no superará 1 dB aproximadamente del nivel del canal con el nivel superior. Esto puede ser audible en determinados casos, pero no debería ser muy excesivo. En otra realización se usa un umbral de 15 dB, en cuyo caso la suma de los dos canales no supera los 1,5 dB del nivel del canal dominante.

5 En algunas realizaciones se usan varios umbrales, por ejemplo 15 dB, 18 dB y 25 dB.

10 Debe observarse que aunque la identificación de canales no contribuyentes se ha descrito anteriormente para AC-3 y E-AC-3, la característica de identificación de canales no contribuyentes de la invención no está limitada a tales formatos. Otros formatos, por ejemplo, también proporcionan información, por ejemplo, metadatos relacionados con el mezclado descendente que pueden usarse para la identificación de uno o más canales no contribuyentes. Tanto MPEG-2 AAC (ISO/IEC 13818-7) como MPEG-4 Audio (ISO/IEC 14496-3) pueden transmitir lo que la norma denomina "coeficiente de mezclado descendente matricial". Algunas realizaciones de la invención que descodifican tales formatos usan este coeficiente para construir una señal estéreo o mono a partir de una señal 3/2, es decir, una señal izquierda, central, derecha, izquierda envolvente y derecha envolvente. El coeficiente de mezclado descendente matricial determina cómo los canales envolventes se mezclan con los canales delanteros para construir la salida estéreo o mono. Cuatro valores posibles del coeficiente de mezclado descendente matricial son posibles según cada una de estas normas, uno de las cuales es 0. Un valor de 0 da como resultado que los canales envolventes no se incluyan en el mezclado descendente. Algunas realizaciones de descodificador MPEG-2 AAC o de descodificador MPEG-4 Audio de la invención incluyen generar un mezclado descendente estéreo o mono a partir de una señal 3/2 usando los coeficientes de mezclado descendente señalizados en el flujo de bits, e incluyen además identificar un canal no contribuyente mediante un coeficiente de mezclado descendente matricial de 0, en cuyo caso la transformación inversa y el procesamiento de división en ventanas y de solapamiento y suma no se llevan a cabo.

25 La figura 12 muestra un diagrama de bloques simplificado de una realización de un sistema de procesamiento 1200 que incluye al menos un procesador 1203. En este ejemplo se muestra un procesador x86 cuyo conjunto de instrucciones incluye instrucciones vectoriales SSE. También se muestra en forma de bloque simplificado un subsistema de bus 1205 mediante el cual se acoplan los diversos componentes del sistema de procesamiento. El sistema de procesamiento incluye un subsistema de almacenamiento 1211 acoplado al (a los) procesador(es), por ejemplo a través del subsistema de bus 1205, presentando el subsistema de almacenamiento 1211 uno o más dispositivos de almacenamiento, incluyendo al menos una memoria y, en algunas realizaciones, uno o más otros dispositivos de almacenamiento, tales como componentes de almacenamiento magnéticos y/u ópticos. Algunas realizaciones incluyen además al menos una interfaz de red 1207 y un subsistema de entrada/salida de audio 1209 que puede aceptar datos PCM y que incluye uno o más DAC para convertir los datos PCM en formas de onda eléctricas para activar un conjunto de altavoces o auriculares. El sistema de procesamiento también puede incluir otros elementos, como les resultará evidente a los expertos en la técnica, no mostrados en la figura 12 por simplicidad.

40 El subsistema de almacenamiento 1211 incluye instrucciones 1213 que cuando se ejecutan en el sistema de procesamiento hacen que el sistema de procesamiento lleve a cabo la descodificación de datos de audio que incluyen N.n canales de datos de audio codificados, por ejemplo datos E-AC-3, para formar datos de audio descodificados que incluyen M.m canales de audio descodificado, $M \geq 1$, y para el caso de mezclado descendente, $M < N$. En los formatos de codificación conocidos en la actualidad, $n=0$ ó 1 y $m=0$ ó 1, pero la invención no está limitada a esto. En algunas realizaciones, las instrucciones 1211 se dividen en módulos. Otras instrucciones (otro software) 1215 también están incluidas normalmente en el subsistema de almacenamiento. La realización mostrada incluye los módulos siguientes en las instrucciones 1211: dos módulos de descodificador: un módulo de descodificador de 5.1 canales de trama independiente 1223 que incluye un módulo de descodificación de sección de entrada 1231 y un módulo de descodificación de sección de procesamiento 1233, un módulo de descodificador de trama dependiente 1225 que incluye un módulo de descodificación de sección de entrada 1235 y un módulo de descodificación de sección de procesamiento 1237, un módulo de instrucciones de análisis de información de trama 1221 que cuando se ejecutan hacen que se desempaqueten datos de campo de información de flujo de bits (BSI) de cada trama para identificar las tramas y tipos de trama y para proporcionar las tramas identificadas a instancias apropiadas de módulo de descodificación de sección de entrada 1231 ó 1235, y un módulo de instrucciones de correlación de canal 1227 que cuando se ejecutan y en caso de que $N > 5$ hacen que se combinen los datos descodificados de módulos de descodificación de sección de procesamiento respectivos para formar los N.n canales de datos descodificados.

60 Realizaciones alternativas del sistema de procesamiento pueden incluir uno o más procesadores acoplados mediante al menos un enlace de red, es decir, distribuidos. Es decir, uno o más de los módulos puede estar en otros sistemas de procesamiento acoplados a un sistema de procesamiento principal mediante un enlace de red. Tales realizaciones alternativas resultarán evidentes a un experto en la técnica. Por tanto, en algunas realizaciones, el sistema comprende uno o más subsistemas que están interconectados a través de un enlace de red, incluyendo cada subsistema al menos un procesador.

65 Por tanto, el sistema de procesamiento de la figura 12 forma una realización de un aparato de procesamiento de datos de audio que incluyen N.n canales de datos de audio codificados para formar datos de audio descodificados

que incluyen M.m canales de audio descodificado, $M \geq 1$, en el caso del mezclado descendente, $M < N$, y para el mezclado ascendente, $M > N$. Aunque en las normas actuales, $n=0$ ó 1 y $m=0$ ó 1 , otras realizaciones son posibles. El aparato incluye varios elementos funcionales expresados de manera funcional como un medio para llevar a cabo una función. Por elemento funcional se entiende un elemento que lleva a cabo una función de procesamiento. Cada elemento de este tipo puede ser un elemento de hardware, por ejemplo hardware de propósito especial, o un sistema de procesamiento que incluye un medio de almacenamiento que incluye instrucciones que cuando se ejecutan llevan a cabo la función. El aparato de la figura 12 incluye medios para aceptar los datos de audio que incluyen N canales de datos de audio codificados que han sido codificados mediante un método de codificación, por ejemplo un método de codificación E-AC-3, y en términos más generales, un método de codificación que comprende transformar, usando una transformada solapada, N canales de datos de audio digital, formar y empaquetar datos de exponente y de mantisa de dominio de frecuencia, y formar y empaquetar metadatos relacionados con los datos de exponente y de mantisa de dominio de frecuencia, donde los metadatos incluyen opcionalmente metadatos relacionados con un procesamiento de pre-ruido transitorio.

15 El aparato incluye medios para descodificar los datos de audio aceptados.

En algunas realizaciones, los medios de descodificación incluyen medios para desempaquetar los metadatos y medios para desempaquetar y descodificar los datos de exponente y de mantisa de dominio de frecuencia, medios para determinar coeficientes de transformada a partir de los datos de exponente y de mantisa de dominio de frecuencia desempaquetados y descodificados; medios para someter a una transformación inversa los datos de dominio de frecuencia; medios para aplicar operaciones de división en ventanas y de solapamiento y suma para determinar datos de audio muestreados; medios para aplicar cualquier procesamiento de pre-ruido transitorio requerido de descodificación según los metadatos relacionados con el procesamiento de pre-ruido transitorio; y medios para realizar un mezclado descendente TD según datos de mezclado descendente. Los medios de mezclado descendente TD, en el caso en que $M < N$, mezclan de manera descendente según datos de mezclado descendente, incluyendo en alguna realización comprobar si los datos de mezclado descendente han cambiado con respecto a datos de mezclado descendente usados anteriormente, y, si han cambiado, aplican una atenuación cruzada para determinar datos de mezclado descendente atenuados de manera cruzada y un mezclado descendente según los datos de mezclado descendente atenuados de manera cruzada, y, si no han cambiado, realizan directamente un mezclado descendente según los datos de mezclado descendente.

Algunas realizaciones incluyen medios para determinar si en un bloque se usa un mezclado descendente TD o un mezclado descendente FD, y medios de mezclado descendente FD que se activan si los medios que determinan si en un bloque se usa un mezclado descendente TD o un mezclado descendente FD determinan un mezclado descendente FD, incluyendo medios para un procesamiento de transición de mezclado descendente de TD a FD. Tales realizaciones incluyen además medios para un procesamiento de transición de mezclado descendente de FD a TD. El funcionamiento de estos elementos es como el descrito en este documento.

En algunas realizaciones, el aparato incluye medios para identificar uno o más canales no contribuyentes de los N.n canales de entrada, siendo un canal no contribuyente un canal que no contribuye en los M.m canales. El aparato no lleva a cabo la transformación inversa de los datos de dominio de frecuencia ni aplica el procesamiento adicional, tal como TPNP o solapamiento y suma, en el uno o más canales no contribuyentes identificados.

En algunas realizaciones, el aparato incluye al menos un procesador x86 cuyo conjunto de instrucciones incluye difundir en flujo continuo extensiones (SSE) de una sola instrucción y múltiples datos que comprenden instrucciones vectoriales. Los medios de mezclado descendente, en funcionamiento, ejecutan instrucciones vectoriales en al menos un procesador del uno o más procesadores x86.

Aparatos alternativos a los mostrados en la figura 12 también son posibles. Por ejemplo, uno o más de los elementos pueden implementarse mediante dispositivos de hardware, mientras que otros pueden implementarse haciendo funcionar un procesador x86. Tales variaciones resultarán evidentes a los expertos en la técnica.

En algunas realizaciones del aparato, los medios de descodificación incluyen uno o más medios de descodificación de sección de entrada y uno o más medios de descodificación de sección de procesamiento. Los medios de descodificación de sección de entrada incluyen los medios de desempaquetado de metadatos y los medios de desempaquetado y descodificación de los datos de exponente y de mantisa de dominio de frecuencia. Los medios de descodificación de sección de procesamiento incluyen los medios que determinan si en un bloque se usa mezclado descendente TD o mezclado descendente FD, los medios de mezclado descendente FD que incluyen los medios para el procesamiento de transición de mezclado descendente de TD a FD, los medios para el procesamiento de transición de mezclado descendente de FD a TD, los medios de determinación de coeficientes de transformada a partir de los datos de exponente y de mantisa de dominio de frecuencia desempaquetados y descodificados; los medios que someten a una transformación inversa los datos de dominio de frecuencia; los medios que aplican cualquier procesamiento de pre-ruido transitorio de descodificación según los metadatos relacionados con el procesamiento de pre-ruido transitorio; y los medios de mezclado descendente de dominio de tiempo según datos de mezclado descendente. El mezclado descendente en el dominio de tiempo, en el caso en que $M < N$, realiza un mezclado descendente según datos de mezclado descendente, incluyendo, en algunas

realizaciones, comprobar si los datos de mezclado descendente han cambiado con respecto a datos de mezclado descendente usados anteriormente, y, si han cambiado, aplicar una atenuación cruzada para determinar datos de mezclado descendente atenuados de manera cruzada y un mezclado descendente según los datos de mezclado descendente atenuados de manera cruzada, y, si no han cambiado, realizar un mezclado descendente según los datos de mezclado descendente.

Para procesar datos E-AC-3 de más de 5.1 canales de datos codificados, los medios de descodificación incluyen múltiples instancias de los medios de descodificación de sección de entrada y de los medios de descodificación de sección de procesamiento, incluyendo un primer medio de descodificación de sección de entrada y un primer medio de descodificación de sección de procesamiento para descodificar la trama independiente de hasta 5.1 canales, un segundo medio de descodificación de sección de entrada y un segundo medio de descodificación de sección de procesamiento para descodificar una o más tramas de datos dependientes. El aparato incluye además medios para desempaquetar datos de campo de información de flujo de bits para identificar las tramas y los tipos de trama y para proporcionar las tramas identificadas a medios apropiados de descodificación de sección de entrada, y medios para combinar los datos descodificados procedentes de medios respectivos de descodificación de sección de procesamiento para formar los N canales de datos descodificados.

Debe observarse que aunque la norma E-AC-3 y otros métodos de codificación usan una transformada de solapamiento y suma, y que la transformación inversa incluye operaciones de división en ventanas y de acoplamiento y suma, se sabe que otras formas de transformadas son posibles, las cuales funcionan de tal manera que la transformación inversa y el procesamiento adicional pueden recuperar muestras de dominio de tiempo sin errores de solapamiento. Por lo tanto, la invención no está limitada a las transformadas de solapamiento y suma, y cada vez que se menciona la transformación inversa de datos de dominio de frecuencia y la ejecución de operaciones de división en ventanas y de solapamiento y suma para determinar muestras de dominio de tiempo, los expertos en la técnica entenderán que, en general, estas operaciones pueden denominarse "transformación inversa de los datos de dominio de frecuencia y aplicación de procesamiento adicional para determinar datos de audio muestreados".

Aunque los términos exponente y mantisa se usan a lo largo de la descripción ya que son los términos usados en AC-3 y E-AC-3, otros formatos de codificación pueden usar otros términos, por ejemplo, factores de escala y coeficientes espectrales en el caso de HE-AAC, y el uso de los términos exponente y mantisa no limita el alcance de la invención a formatos que usan los términos exponente y mantisa.

A no ser que se indique específicamente lo contrario, como resulta evidente a partir de la siguiente descripción, debe apreciarse que a lo largo de la memoria descriptiva, los análisis que utilizan términos tales como "procesamiento", "cómputo", "cálculo", "determinación", "generación", o similares, se refieren a la acción y/o procesos de un elemento de hardware, por ejemplo un ordenador o un sistema informático, un sistema de procesamiento o dispositivo informático electrónico similar, que manipula y/o transforma datos representados de manera física, tal como datos electrónicos o cantidades en otros datos representados de manera similar como cantidades físicas.

De manera similar, el término "procesador" puede referirse a cualquier dispositivo o parte de un dispositivo que procese datos electrónicos, por ejemplo, de registros y/o de una memoria, para transformar esos datos electrónicos en otros datos electrónicos que, por ejemplo, puedan almacenarse en registros y/o memoria. Un "sistema de procesamiento" u "ordenador" o una "máquina informática" o una "plataforma informática" pueden incluir uno o más procesadores.

Debe observarse que cuando se describe un método que incluye varios elementos, por ejemplo varias etapas, no hay un orden implícito de tales elementos, por ejemplo etapas, a no ser que se indique específicamente.

En algunas realizaciones, un medio de almacenamiento legible por ordenador está configurado con, por ejemplo está codificado con, por ejemplo, almacena instrucciones que cuando son ejecutadas por uno o más procesadores de un sistema de procesamiento, tal como un dispositivo de procesamiento de señales digitales o un subsistema que incluye al menos un elemento de procesamiento y un subsistema de almacenamiento, hacen que se lleve a cabo un método como el descrito en este documento. Debe observarse que en la descripción anterior, cuando se menciona que las instrucciones están configuradas, cuando se ejecutan, para llevar a cabo un proceso, debe entenderse que esto significa que las instrucciones, cuando se ejecutan, hacen que uno o más procesadores funcionen de manera que un aparato de hardware, por ejemplo el sistema de procesamiento, lleve a cabo el proceso.

En algunas realizaciones, las metodologías descritas en este documento pueden llevarse a cabo por uno o más procesadores que aceptan lógica, instrucciones codificadas en uno o más medios legibles por ordenador. Cuando son ejecutadas por uno o más de los procesadores, las instrucciones hacen que se lleve a cabo al menos uno de los métodos descritos en este documento. Se incluye cualquier procesador que pueda ejecutar un conjunto de instrucciones (secuenciales o de otro tipo) que especifiquen las acciones que van a realizarse. Por tanto, un ejemplo es un sistema de procesamiento típico que incluye uno o más procesadores. Cada procesador puede incluir uno o más de una CPU o elemento similar, una unidad de procesamiento de gráficos (GPU) y/o una unidad DSP

programable. El sistema de procesamiento incluye además un subsistema de almacenamiento con al menos un medio de almacenamiento, que puede incluir una memoria integrada en un dispositivo semiconductor, o un subsistema de memoria aparte que incluye una RAM principal y/o una RAM estática y/o una ROM así como memoria caché. El subsistema de almacenamiento puede incluir además uno o más otros dispositivos de almacenamiento, tales como dispositivos de almacenamiento magnéticos y/u ópticos y/o de estado sólido. Un subsistema de bus puede incluirse para la comunicación entre los componentes. El sistema de procesamiento puede ser además un sistema de procesamiento distribuido con procesadores acoplados mediante una red, por ejemplo a través de dispositivos de interfaz de red o dispositivos de interfaz de red inalámbrica. Si el sistema de procesamiento requiere un dispositivo de visualización, tal dispositivo de visualización puede incluirse, por ejemplo una pantalla de cristal líquido (LCD), una pantalla de emisión de luz orgánica (OLED) o una pantalla de tubos de rayos catódicos (CRT). Si se necesita una entrada de datos manual, el sistema de procesamiento incluye además un dispositivo de entrada, tal como una o más de una unidad de entrada alfanumérica tal como un teclado, un dispositivo de control de puntero tal como un ratón, etc. Los términos dispositivo de almacenamiento, subsistema de almacenamiento o unidad de memoria usados en este documento, si resulta evidente por el contexto y a no ser que se indique explícitamente lo contrario, también abarcan un sistema de almacenamiento tal como una unidad de disco duro. En algunas configuraciones, el sistema de procesamiento puede incluir un dispositivo de salida de sonido y un dispositivo de interfaz de red.

Por tanto, el subsistema de almacenamiento incluye un medio legible por ordenador que está configurado con, por ejemplo codificado con, instrucciones, por ejemplo lógica, por ejemplo software, que cuando son ejecutadas por uno o más procesadores, hacen que se lleve a cabo una o más de las etapas de método descritas en este documento. El software puede residir en el disco duro o también puede residir, completamente o al menos parcialmente, en la memoria, tal como una RAM, y/o en la memoria interna del procesador durante la ejecución del mismo mediante el sistema informático. Por tanto, la memoria y el procesador que incluye la memoria también constituyen un medio legible por ordenador en el que hay instrucciones codificadas.

Además, un medio legible por ordenador puede formar un producto de programa informático o puede estar incluido en un producto de programa informático.

En realizaciones alternativas, el uno o más procesadores funcionan como un dispositivo autónomo o pueden estar conectados, por ejemplo conectados en red a otro(s) procesador(es), en una implantación conectada en red, el uno o más procesadores pueden funcionar en la capacidad de un servidor o una máquina cliente en un entorno de red cliente-servidor, o como una máquina homóloga en un entorno de red distribuido o de igual a igual. El término sistema de procesamiento abarca todas estas posibilidades a no ser que se excluyan explícitamente en este documento. El uno o más procesadores pueden formar un ordenador personal (PC), un dispositivo de reproducción multimedia, un PC de tableta, un descodificador (STB), un asistente digital personal (PDA), una máquina de juego, un teléfono celular, un aparato web, un encaminador de red, un conmutador o puente, o cualquier máquina capaz de ejecutar un conjunto de instrucciones (secuenciales o de otro tipo) que especifiquen acciones que deba llevar a cabo esa máquina.

Debe observarse que aunque algunos diagramas solo muestran un único procesador y un único subsistema de almacenamiento, por ejemplo una única memoria que almacena la lógica que incluye instrucciones, los expertos en la técnica entenderán que muchos de los componentes descritos anteriormente están incluidos, pero no mostrados o descritos explícitamente para no oscurecer el aspecto inventivo. Por ejemplo, aunque solo se ilustra una única máquina, también debe considerarse que el término "máquina" incluye cualquier colección de máquinas que ejecutan individual o conjuntamente un conjunto (o múltiples conjuntos) de instrucciones para llevar a cabo una cualquiera o más de las metodologías descritas en este documento.

Por tanto, una realización de cada uno de los métodos descritos en este documento está en forma de un medio legible por ordenador configurado con un conjunto de instrucciones, por ejemplo, un programa informático que cuando es ejecutado en uno o más procesadores, por ejemplo uno o más procesadores que son parte de un dispositivo multimedia, hace que se lleven a cabo las etapas de método. Algunas realizaciones adoptan la forma de la propia lógica. Por tanto, como apreciarán los expertos en la técnica, las realizaciones de la presente invención pueden realizarse como un método, un aparato tal como un aparato de propósito especial, un aparato tal como un sistema de procesamiento de datos, lógica, por ejemplo incluida en un medio de almacenamiento legible por ordenador o un medio de almacenamiento legible por ordenador codificado con instrucciones, por ejemplo un medio de almacenamiento legible por ordenador configurado como un producto de programa informático. El medio legible por ordenador está configurado con un conjunto de instrucciones que, cuando son ejecutadas por uno o más procesadores, hacen que se lleven a cabo las etapas de método. Por consiguiente, aspectos de la presente invención pueden adoptar la forma de un método, una realización completamente en hardware que incluye varios elementos funcionales, donde por elemento funcional se entiende un elemento que lleva a cabo una función de procesamiento. Cada elemento de este tipo puede ser un elemento de hardware, por ejemplo, hardware de propósito especial, o un sistema de procesamiento que incluye un medio de almacenamiento que incluye instrucciones que cuando son ejecutadas llevan a cabo la función. Aspectos de la presente invención pueden adoptar la forma de una realización completamente en software o una realización que combina aspectos de software y de hardware. Además, la presente invención puede adoptar la forma de lógica de programa, por ejemplo en un

medio legible por ordenador, por ejemplo un programa informático en un medio de almacenamiento legible por ordenador o el medio legible por ordenador configurado con código de programa legible por ordenador, por ejemplo un producto de programa informático. Debe observarse que en el caso de hardware de propósito especial, definir la función del hardware es suficiente para permitir que un experto en la técnica escriba una descripción funcional que pueda procesarse mediante programas que determinen automáticamente una descripción de hardware para generar hardware que lleve a cabo la función. Por tanto, la descripción de este documento basta para definir tal hardware de propósito especial.

Aunque el medio legible por ordenador se muestra en una realización de ejemplo como un único medio, debe tenerse en cuenta que el término "medio" incluye un único medio o múltiples medios (por ejemplo, varias memorias, una base de datos centralizada o distribuida y/o memorias caché y servidores asociados) que almacenan el uno o más conjuntos de instrucciones. Un medio legible por ordenador puede adoptar muchas formas, incluyendo, pero sin limitarse a, medios no volátiles y medios volátiles. Los medios no volátiles incluyen, por ejemplo, discos ópticos, magnéticos y discos magneto-ópticos. Los medios volátiles incluyen memorias dinámicas, tales como una memoria principal.

Debe entenderse además que las realizaciones de la presente invención no están limitadas a ninguna técnica de implementación o programación particular, y que la invención puede implementarse usando cualquier técnica apropiada para implementar la funcionalidad descrita en este documento. Además, las realizaciones no están limitadas a ningún lenguaje de programación o sistema operativo particulares.

La referencia a lo largo de esta memoria descriptiva a "una realización" significa que una propiedad, estructura o característica particular descrita en relación con la realización está incluida en al menos una realización de la presente invención. Por tanto, no todas las apariciones de la expresión "en una realización" en varias partes de esta memoria descriptiva hacen referencia necesariamente a la misma realización, aunque pueden hacerlo. Además, las propiedades, estructuras o características particulares pueden combinarse de cualquier manera adecuada, como le resultará evidente a un experto en la técnica a partir de esta memoria descriptiva, en una o más realizaciones.

Asimismo, debe apreciarse que en la descripción anterior de realizaciones de ejemplo de la invención, varias características de la invención están agrupadas algunas veces en una única realización, figura o descripción de las mismas con el fin de simplificar la divulgación y ayudar al entendimiento de uno o más de los diversos aspectos inventivos. Sin embargo, no debe interpretarse que este método de divulgación refleja la intención de que la invención reivindicada requiere más características que las mencionadas expresamente en cada reivindicación. En cambio, como las siguientes reivindicaciones reflejan, los aspectos inventivos no residen en todas las características de una sola realización descrita anteriormente. Por tanto, las reivindicaciones que siguen a la DESCRIPCIÓN DE REALIZACIONES DE EJEMPLO están incorporadas expresamente en esta DESCRIPCIÓN DE REALIZACIONES DE EJEMPLO, donde cada reivindicación representa por sí misma una realización diferente de esta invención.

Además, aunque algunas realizaciones descritas en este documento incluyen algunas pero no otras características incluidas en otras realizaciones, combinaciones de características de diferentes realizaciones están dentro del alcance de la invención y forman diferentes realizaciones, como entenderán los expertos en la técnica. Por ejemplo, en las siguientes reivindicaciones, cualquiera de las realizaciones reivindicadas puede usarse en cualquier combinación.

Además, algunas de las realizaciones están descritas en este documento como un método o combinación de elementos de un método que puede implementarse por un procesador de un sistema informático o mediante otros medios que llevan a cabo la función. Por tanto, un procesador con las instrucciones necesarias para llevar a cabo tal método o elemento de un método forma un medio para llevar a cabo el método o elemento de un método. Además, un elemento, descrito en este documento, de una realización de aparato es un ejemplo de un medio para llevar a cabo la función realizada por el elemento con el fin de llevar a cabo la invención.

En la descripción proporcionada en este documento se exponen numerosos detalles específicos. Sin embargo, debe entenderse que las realizaciones de la invención pueden llevarse a la práctica sin estos detalles específicos. En otros casos, métodos, estructuras y técnicas ampliamente conocidos no se han mostrado en detalle para no oscurecer el entendimiento de esta descripción.

Como se usa en este documento, a no ser que se especifique lo contrario, el uso de los adjetivos ordinales "primero", "segundo", "tercero", etc., para describir un objeto común, indican simplemente que se hace referencia a diferentes instancias de objetos similares y no debe considerarse que implican que los objetos descritos deben estar en una secuencia dada, ya sea temporal, espacial, jerárquica o de cualquier otra manera.

Debe apreciarse que aunque la invención se ha descrito en el contexto de la norma E-AC-3, la invención no está limitada a tales contextos y puede utilizarse para descodificar datos codificados por otros métodos que usen técnicas que sean similares a la norma E-AC-3. Por ejemplo, realizaciones de la invención también pueden aplicarse para descodificar audio codificado que sea compatible con versiones anteriores de E-AC-3. Otras realizaciones pueden aplicarse para descodificar audio codificado que haya sido codificado según la norma HE-AAC y para descodificar

audio codificado que sea compatible con versiones anteriores de HE-AAC. Otros flujos codificados también pueden descodificarse de manera ventajosa usando realizaciones de la presente invención.

5 En las siguientes reivindicaciones y en la descripción de este documento, uno cualquiera de los términos "comprender", "comprendido por" o "que comprende" es un término abierto que significa que incluye al menos los elementos/características enumerados tras el mismo, pero sin excluir otros. Por tanto, no debe interpretarse que el término "que comprende", cuando se usa en las reivindicaciones, limita los medios, elementos o etapas enumerados tras el mismo. Por ejemplo, el alcance de la expresión "un dispositivo que comprende A y B" no debe limitarse a dispositivos que consisten solamente en los elementos A y B. Uno cualquiera de los términos "incluir" o "que incluye" 10 usados en este documento es también un término abierto que también significa que incluye al menos los elementos/características que siguen al término, pero sin excluir otros. Por tanto, incluir es sinónimo de comprender.

Asimismo, debe observarse que el término "acoplado", cuando se usa en las reivindicaciones, no debe interpretarse solamente como limitado a conexiones directas. Los términos "acoplado" y "conectado", junto con sus derivados, 15 pueden usarse. Debe entenderse que esos términos no son sinónimos. Por tanto, el alcance de la expresión "un dispositivo A acopado a un dispositivo B" no debe limitarse a dispositivos o sistemas en los que una salida del dispositivo A está directamente conectada a una entrada del dispositivo B. Significa que hay una trayectoria entre una salida de A y una entrada de B, que puede ser una trayectoria que incluye otros dispositivos o medios. "Acoplado" puede significar que dos o más elementos están en contacto físico o eléctrico directo o que dos o más 20 elementos no están en contacto directo entre sí pero aún así actúan conjuntamente o interaccionan entre sí.

De este modo, aunque se ha descrito lo que se considera que son las realizaciones preferidas de la invención, los expertos en la técnica reconocerán que pueden llevarse a cabo otras modificaciones adicionales en las mismas, estando todos estos cambios y modificaciones dentro del alcance de la invención. Por ejemplo, cualesquiera fórmulas proporcionadas anteriormente son simplemente representaciones de procedimientos que pueden usarse. 25 Puede añadirse o suprimirse funcionalidad en los diagramas de bloques y pueden intercambiarse operaciones entre los elementos funcionales. Pueden añadirse o suprimirse etapas en los métodos descritos dentro del alcance de la presente invención.

REIVINDICACIONES

- 1.- Un método que hace funcionar un descodificador de audio (200) para descodificar datos de audio que incluyen bloques codificados de N.n canales de datos de audio para formar datos de audio descodificados que incluyen M.m canales de audio descodificado, $M \geq 1$, siendo n el número de canales de efectos de baja frecuencia en los datos de audio codificados, y siendo m el número de canales de efectos de baja frecuencia en los datos de audio descodificados, comprendiendo el método:
- aceptar los datos de audio que incluyen bloques de N.n canales de datos de audio codificados que han sido codificados mediante un método de codificación, incluyendo el método de codificación transformar N.n canales de datos de audio digital y formar y empaquetar datos de exponente y de mantisa de dominio de frecuencia; y
 - descodificar los datos de audio aceptados, incluyendo la descodificación:
- desempaquetar y descodificar (403) los datos de exponente y de mantisa de dominio de frecuencia,
- determinar coeficientes de transformada (605) a partir de los datos de exponente y de mantisa de dominio de frecuencia desempaquetados y descodificados,
- someter a una transformación inversa (607) los datos de dominio de frecuencia y aplicar un procesamiento adicional para determinar datos de audio muestreados, y
- mezclar de manera descendente en el dominio de tiempo (613) al menos algunos bloques de los datos de audio muestreados determinados según datos de mezclado descendente para el caso en que $M < N$;
- en el que el mezclado descendente en el dominio de tiempo incluye (1100) comprobar si los datos de mezclado descendente han cambiado en el tiempo con respecto a datos de mezclado descendente usados anteriormente, y , si han cambiado, aplicar una atenuación cruzada para determinar datos de mezclado descendente atenuados de manera cruzada y un mezclado descendente en el dominio de tiempo según los datos de mezclado descendente atenuados de manera cruzada, y, si no han cambiado, realizar directamente un mezclado descendente en el dominio de tiempo según los datos de mezclado descendente.
- 2.- El método según la reivindicación 1, donde el método incluye identificar (835) uno o más canales no contribuyentes de los N.n canales de entrada, siendo un canal no contribuyente un canal que no contribuye en los M.m canales, y donde el método no lleva a cabo una transformación inversa de los datos de dominio de frecuencia ni aplica un procesamiento adicional en el uno o más canales no contribuyentes identificados.
- 3.- El método según cualquiera de las reivindicaciones anteriores, en el que la transformación en el método de codificación usa una transformada solapada, y donde el procesamiento adicional incluye aplicar operaciones de división en ventanas y de solapamiento y suma (609) para determinar datos de audio muestreados.
- 4.- El método según cualquiera de las reivindicaciones anteriores, en el que el método de codificación incluye formar y empaquetar metadatos relacionados con los datos de exponente y de mantisa de dominio de frecuencia, donde los metadatos incluyen opcionalmente metadatos relacionados con un procesamiento de pre-ruido transitorio y el mezclado descendente.
- 5.- El método según cualquiera de las reivindicaciones anteriores, en el que el descodificador (200) usa al menos un procesador x86 cuyo conjunto de instrucciones incluye difundir en flujo continuo extensiones (SSE) de una sola instrucción y múltiples datos que comprenden instrucciones vectoriales, y en el que el mezclado descendente en el dominio de tiempo incluye ejecutar instrucciones vectoriales en al menos un procesador del uno o más procesadores x86.
- 6.- El método según la reivindicación 2, en el que $n=1$ y $m=0$, de manera que la transformación inversa y la aplicación de un procesamiento adicional no se llevan cabo en el canal de efecto de baja frecuencia.
- 7.- El método según la reivindicación 2, en el que los datos de audio que incluyen bloques codificados incluyen información que define el mezclado descendente, y en el que la identificación de uno o más canales no contribuyentes usa la información que define el mezclado descendente.
- 8.- El método según la reivindicación 7, en el que la información que define el mezclado descendente incluye parámetros de nivel de mezclado que tienen valores predefinidos que indican que uno o más canales son canales no contribuyentes.
- 9.- El método según la reivindicación 2, en el que la identificación de uno o más canales no contribuyentes incluye además identificar si uno o más canales tienen una cantidad insignificante de contenido con respecto a uno o más otros canales, en el que la identificación de si uno o más canales tienen una cantidad insignificante de contenido con

5 respecto a uno o más otros canales incluye comparar la diferencia de una medida de cantidad de contenido entre pares de canales con un umbral ajustable y/o en el que un canal tiene una cantidad insignificante de contenido con respecto a otro canal si su energía o nivel absoluto es al menos 15 dB inferior a los del otro canal o si su energía o nivel absoluto es al menos 18 dB inferior a los del otro canal o si su energía o nivel absoluto es al menos 25 dB inferior a los del otro canal.

10 10.- El método según cualquier reivindicación anterior, en el que los datos de audio aceptados están en forma de un flujo de bits de tramas de datos codificados, y en el que la descodificación se divide en un conjunto de operaciones de descodificación de sección de entrada (201) y en un conjunto de operaciones de descodificación de sección de procesamiento (203), incluyendo las operaciones de descodificación de sección de entrada el desempaquetado y la descodificación de los datos de exponente y de mantisa de dominio de frecuencia de una trama del flujo de bits en datos de exponente y de mantisa de dominio de frecuencia desempaquetados y descodificados para la trama, y los metadatos incluidos en la trama, incluyendo las operaciones de descodificación de sección de procesamiento la determinación de los coeficientes de transformada, la transformación inversa y la aplicación de un procesamiento adicional, la aplicación de cualquier procesamiento de pre-ruido transitorio requerido de descodificación y el mezclado descendente en el caso en que $M < N$.

20 11.- El método según la reivindicación 10, en el que las operaciones de descodificación de sección de entrada se llevan a cabo en una primera pasada seguida de una segunda pasada, comprendiendo la primera pasada desempaquetar metadatos bloque a bloque y guardar punteros que apuntan a la ubicación en la que están almacenados los datos de exponente y de mantisa empaquetados, y comprendiendo la segunda pasada usar los punteros guardados que apuntan a los exponentes y mantisas empaquetados, y desempaquetar y descodificar datos de exponente y de mantisa canal a canal.

25 12.- El método según cualquier reivindicación anterior, en el que los datos de audio codificados se codifican según una norma del conjunto de normas que consiste en la norma AC-3, la norma E-AC-3 y la norma HE-AAC.

30 13.- Un medio de almacenamiento legible por ordenador que almacena instrucciones de descodificación que cuando son ejecutadas por uno o más procesadores de un sistema de procesamiento hacen que el sistema de procesamiento lleve a cabo el método de cualquiera de las reivindicaciones anteriores.

35 14.- Un aparato (1200) que procesa datos de audio para descodificar los datos de audio que incluyen bloques codificados de $N.n$ canales de datos de audio para formar datos de audio descodificados que incluyen $M.m$ canales de audio descodificado, $M \geq 1$, siendo n el número de canales de efectos de baja frecuencia en los datos de audio codificados, y siendo m el número de canales de efectos de baja frecuencia en los datos de audio descodificados, comprendiendo el aparato medios para llevar a cabo el método de cualquiera de las reivindicaciones 1 a 12.

Desempaquetar datos BSI
Para bloque = 1 a B (el número de bloques)
Desempaquetar datos fijos
Guardar punteros que apuntan a exponentes empaquetados
Para chan = 1 a N (el número de canales codificados)
Desempaquetar exponentes
Para banda = 1 a L (el número de bandas)
Calcular asignación de bits
Desempaquetar mantisas
Desempaquetar canal de acoplamiento (guardar punteros)
Escalar mantisas / deshacer acoplamiento
Desnormalizar mantisas mediante exponentes
Calcular transformada inversa a dominio de ventana
Mezclar de manera descendente al número M apropiado
de canal(es) de salida
Para chan = 1 a M (el número de canales de salida)
División en ventanas & Solapamiento y suma con una memoria
intermedia de retardo
Copiar valores de memoria intermedia de mezclado descendente
en la memoria intermedia de retardo

100

FIG. 1

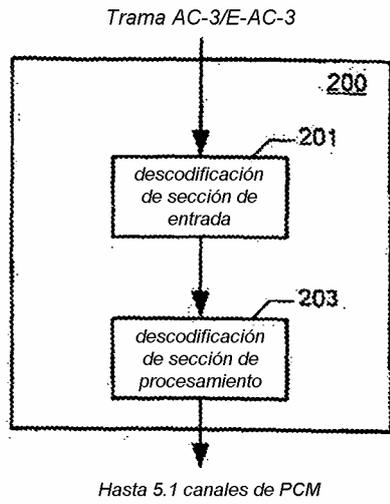


FIG. 2A

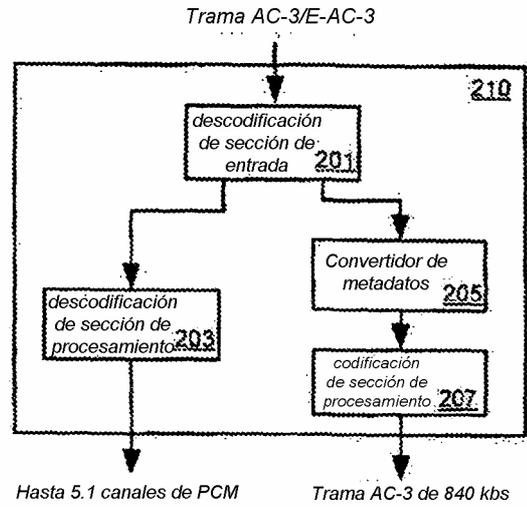


FIG. 2B

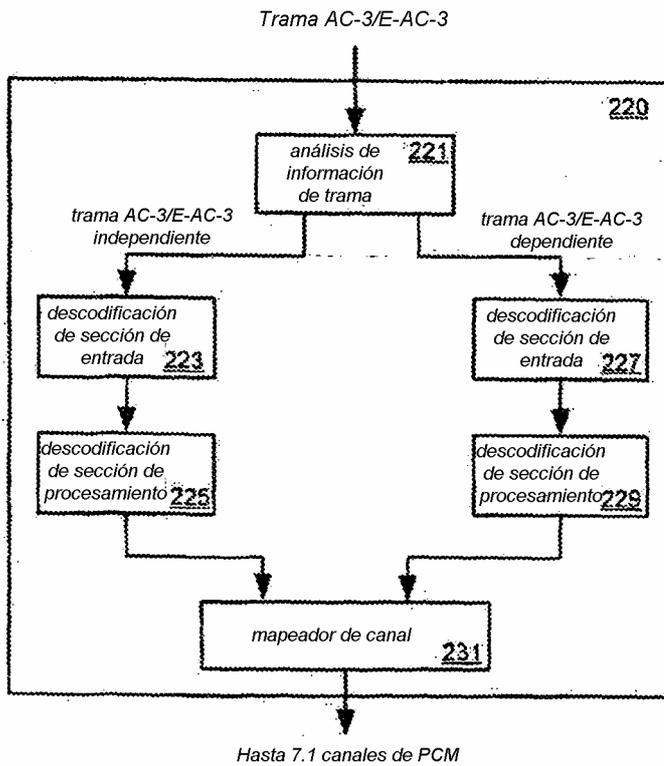


FIG. 2C

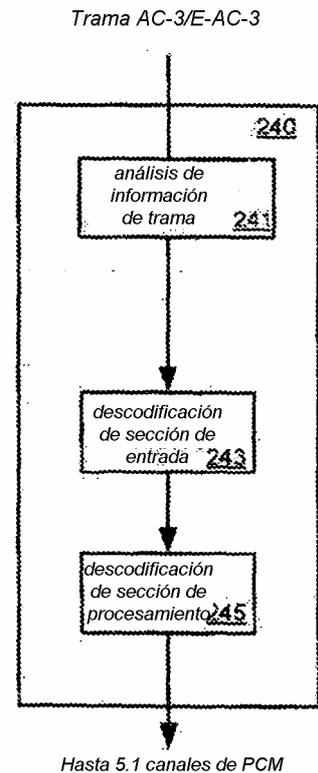


FIG. 2D

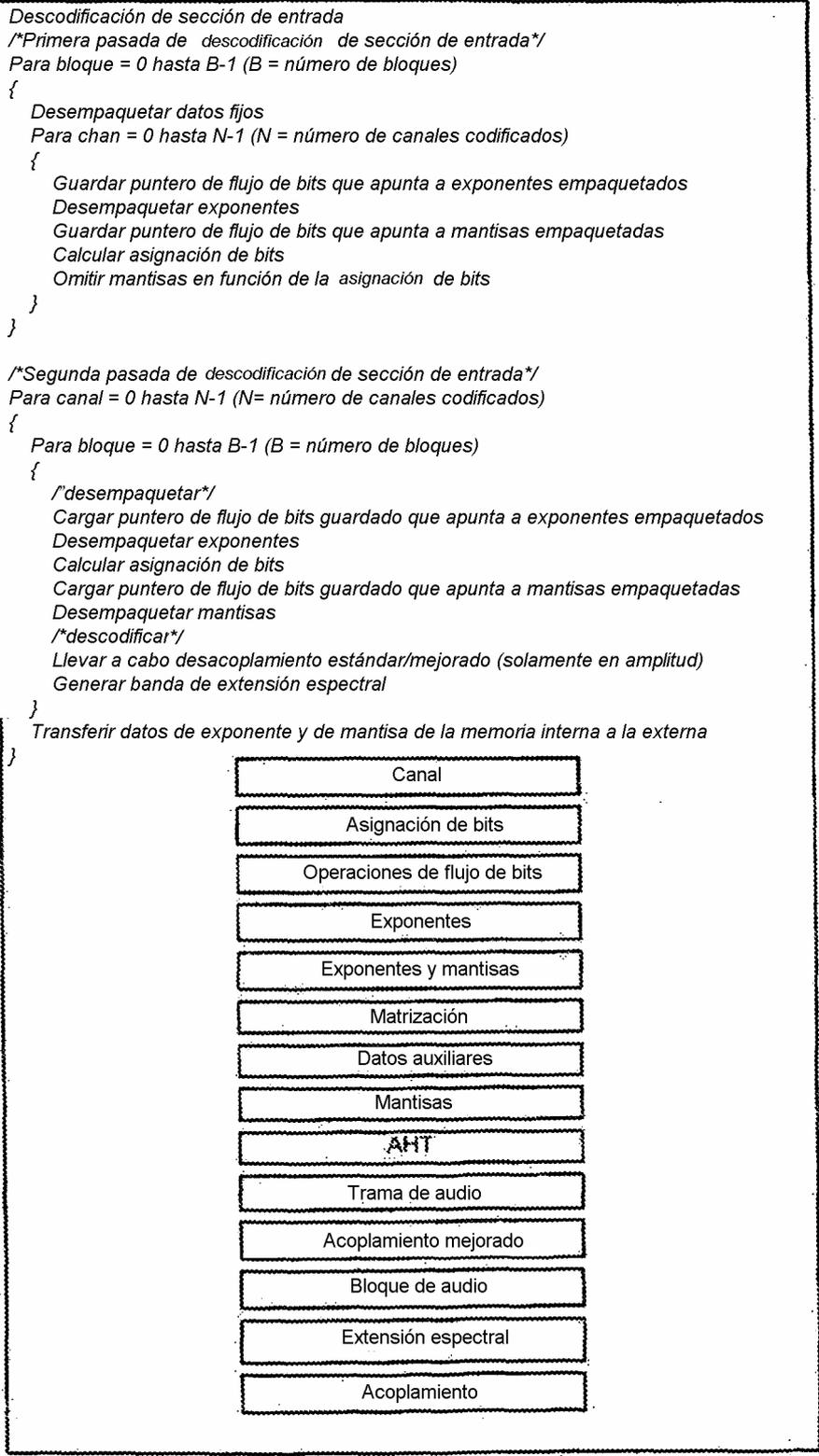


FIG. 3

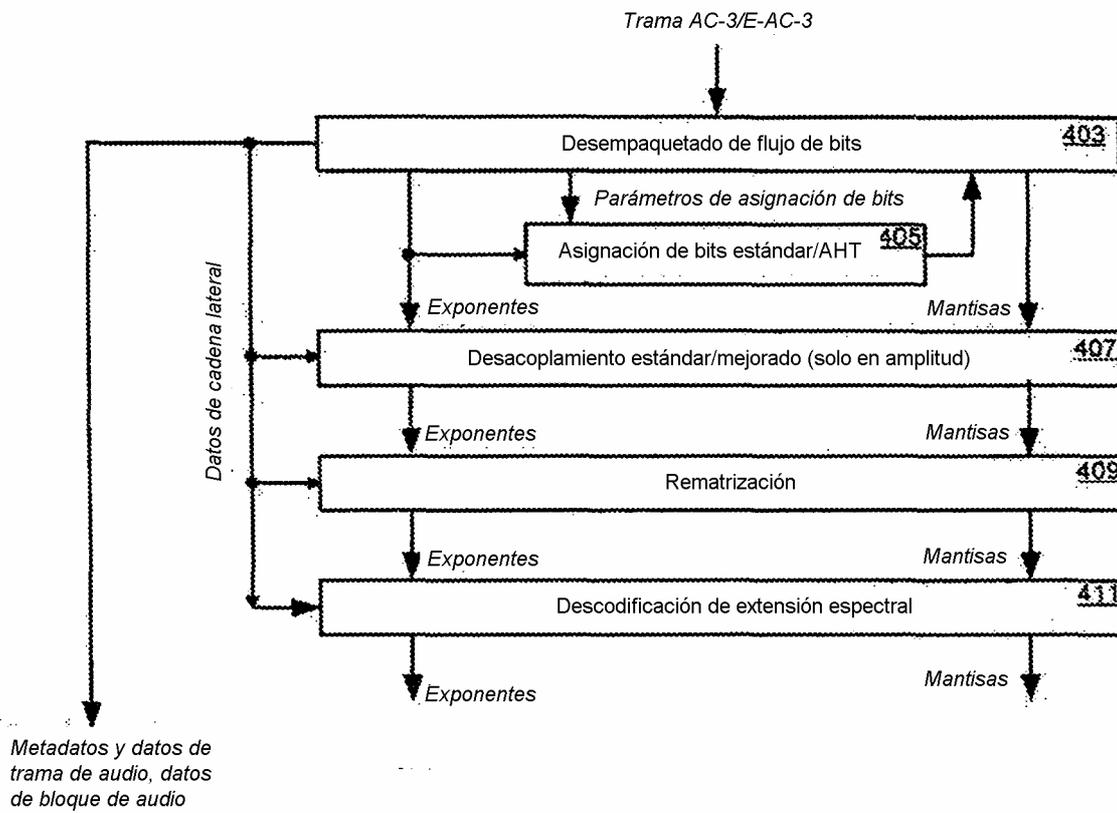


FIG. 4

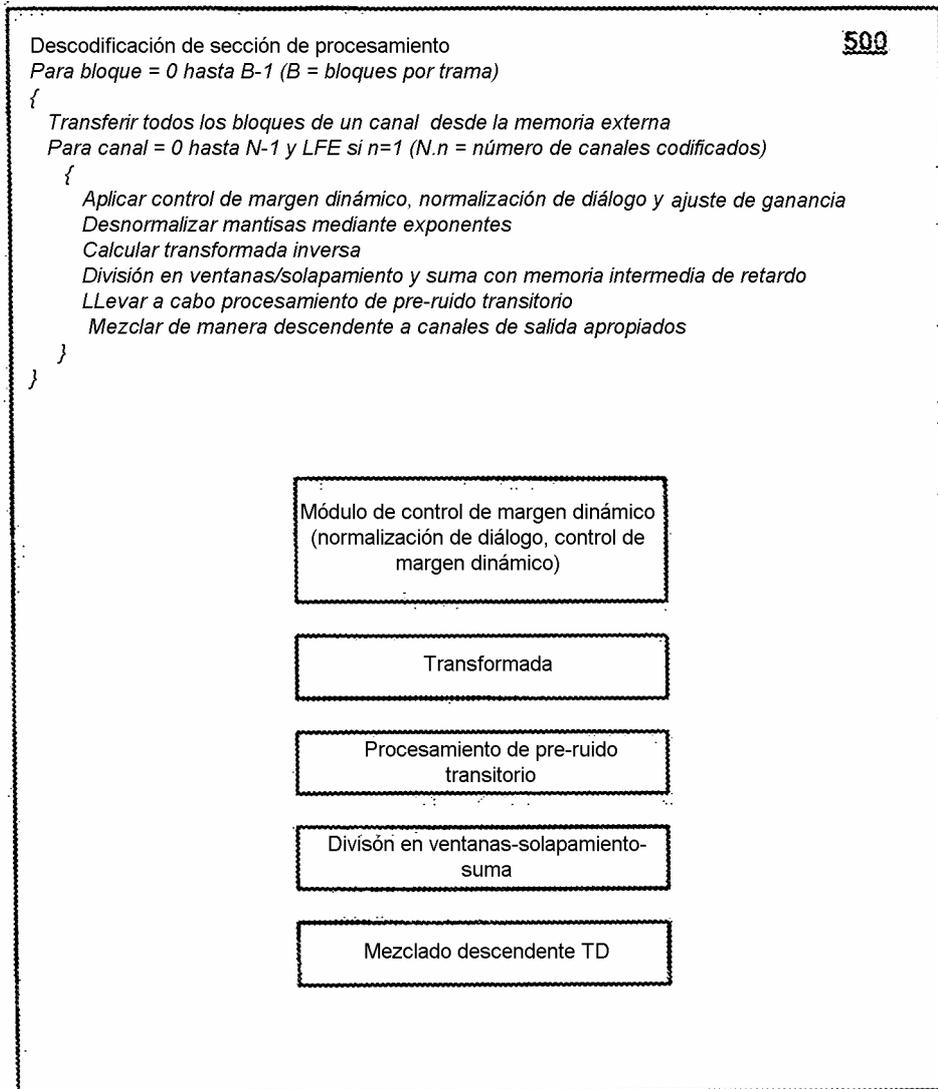


FIG. 5A

```

% Descodificación de sección de procesamiento
Para bloque = 0 hasta B-1 (B= bloques por trama)
{
  Transferir todos los bloques de un canal desde la memoria externa
  Determinar si hay que aplicar un mezclado descendente FD o TD
  Si hay que aplicar un mezclado descendente FD
  {
    Para canal = 0 hasta N-1 y LFE si n=1
    {
      Aplicar control de margen dinámico, normalización de diálogo e inhabilitar ajuste de ganancia
      Desnormalizar mantisas mediante exponentes
      Mezclar de manera descendente en el dominio de frecuencia
      Procesar de diferente manera el bloque de transición que sigue al bloque mezclado de manera
      descendente en el dominio de tiempo
      Tratar cualquier canal de salida atenuado
    }
  }
  Si no % Mezclado descendente TD
  {
    Para canal = 0 hasta N-1 y LFE si n=1 (pero solo para canales en el mezclado descendente)
    {
      Procesar de diferente manera el bloque de transición que sigue al bloque mezclado de manera
      descendente en el dominio de frecuencia
      Aplicar control de margen dinámico, normalización de diálogo e inhabilitar ajuste de ganancia
      Desnormalizar mantisas mediante exponentes
    }
  }
  Para canal = 0 hasta N-1 y LFE si n=1 (pero solo para canales en el mezclado descendente, N=M en caso de mezclado
  descendente FD)
  {
    Calcular transformada inversa
    División en ventanas / solapamiento y suma con memoria intermedia de retardo
    Si mezclado descendente TD
    {
      Llevar a cabo procesamiento de pre-ruido transitorio
      Mezclar de manera descendente a canales de salida apropiados
    }
  }
}

```

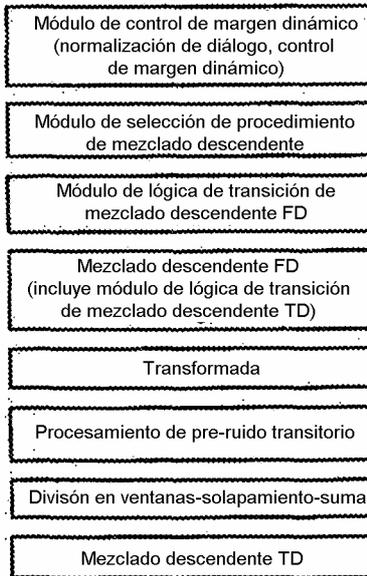


FIG. 5B

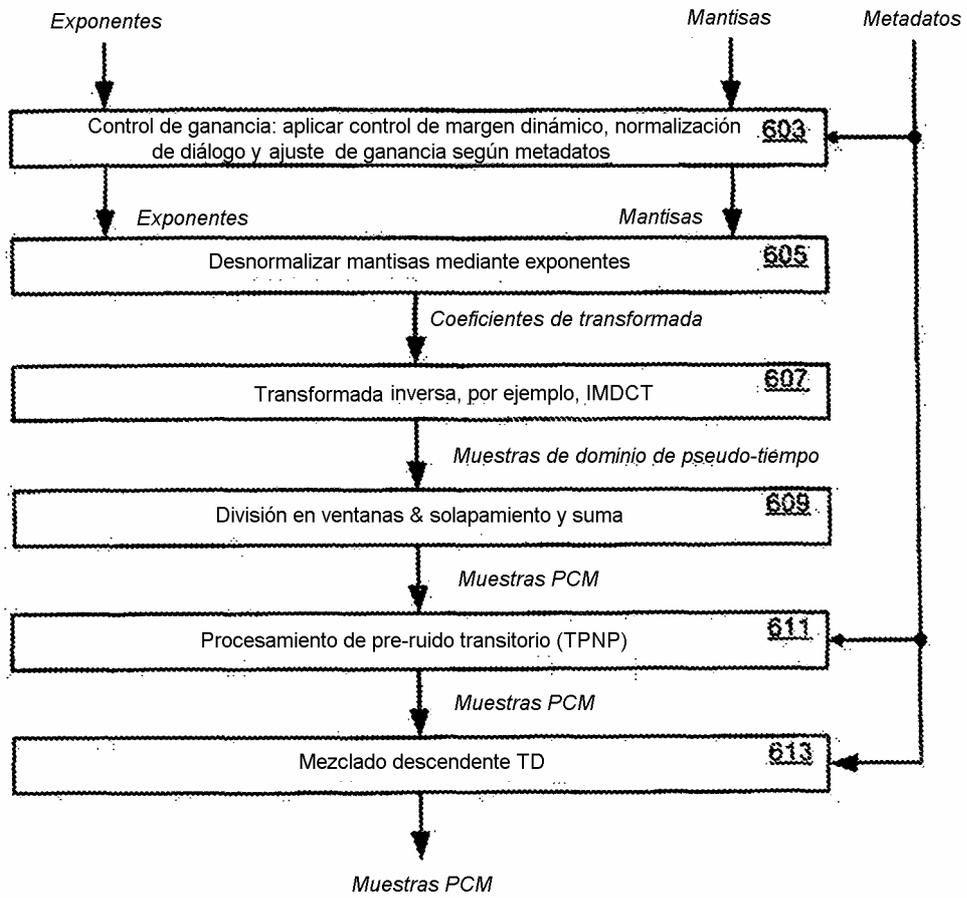


FIG. 6

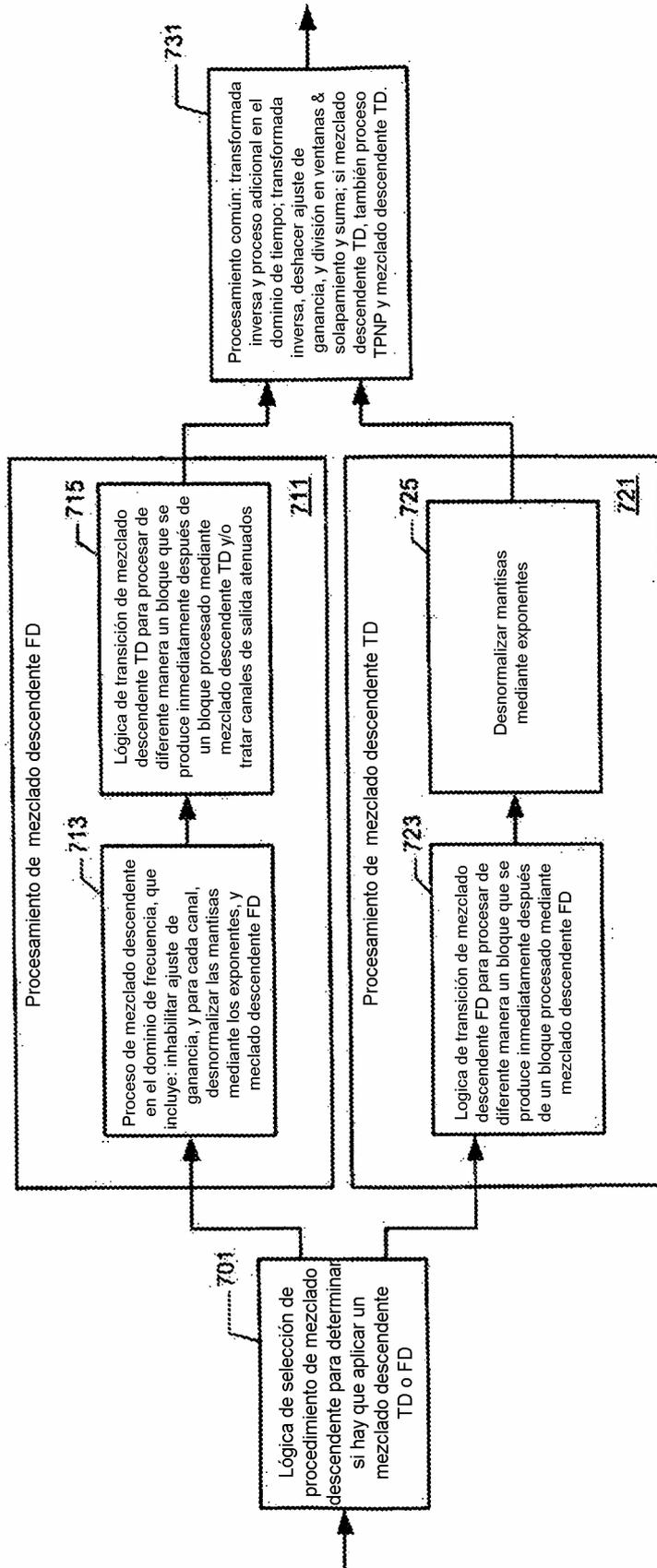


FIG. 7

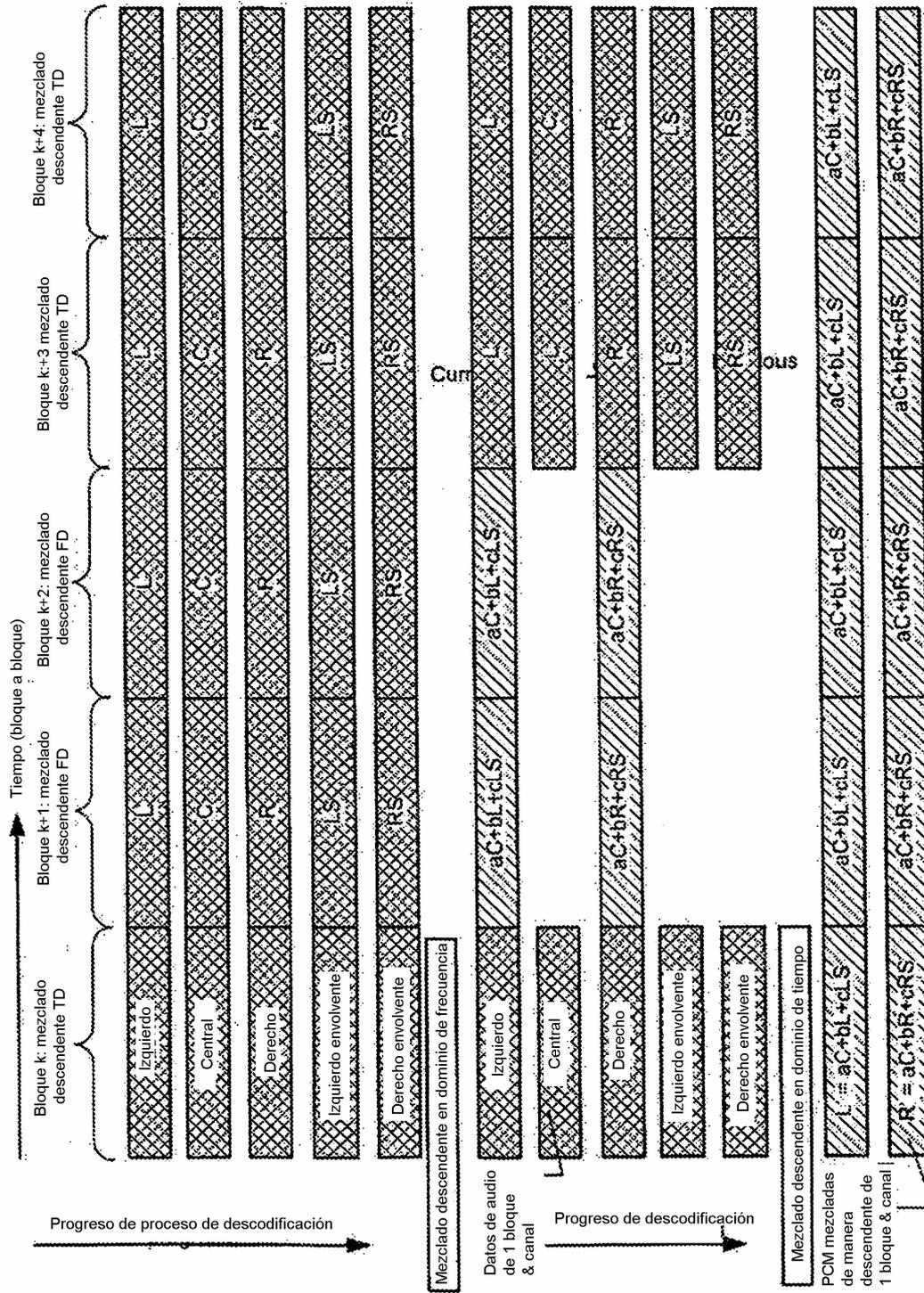


FIG. 9

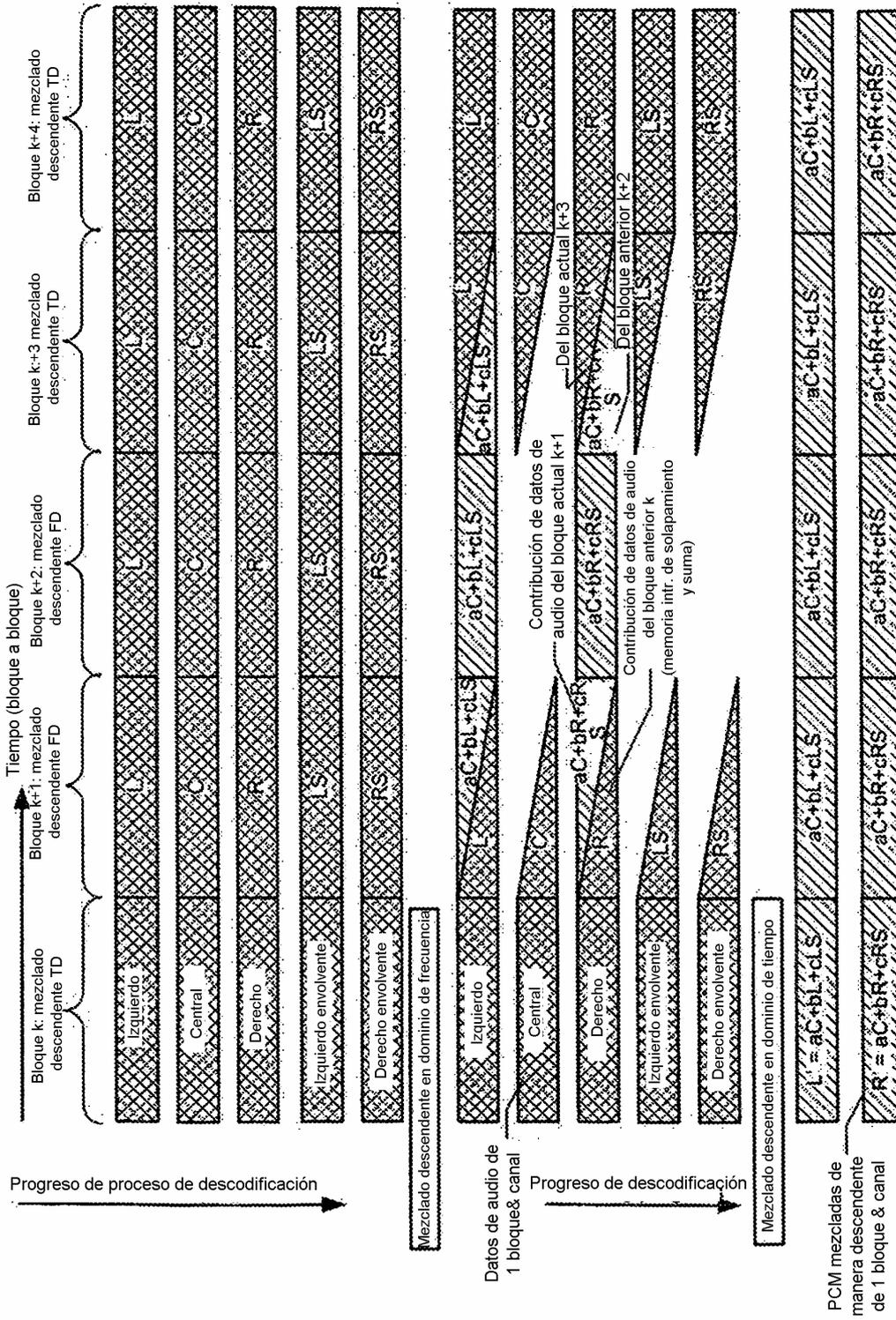


FIG. 10

```
/* Mezclado descendente TD según datos de mezclado descendente*/  
si (nuevos datos de mezclado descendente = viejos datos de mezclado  
descendente)  
{  
  configurar para instrucciones SSE;  
  mezclar de manera descendente usando datos de mezclado descendente;  
}  
si no  
{  
  atenuar de manera cruzada datos viejos con datos nuevos usando una  
  ventana  
  mezclar de manera descendente usando datos de mezclado descendente  
  atenuados de manera cruzada  
}
```

1100

FIG. 11

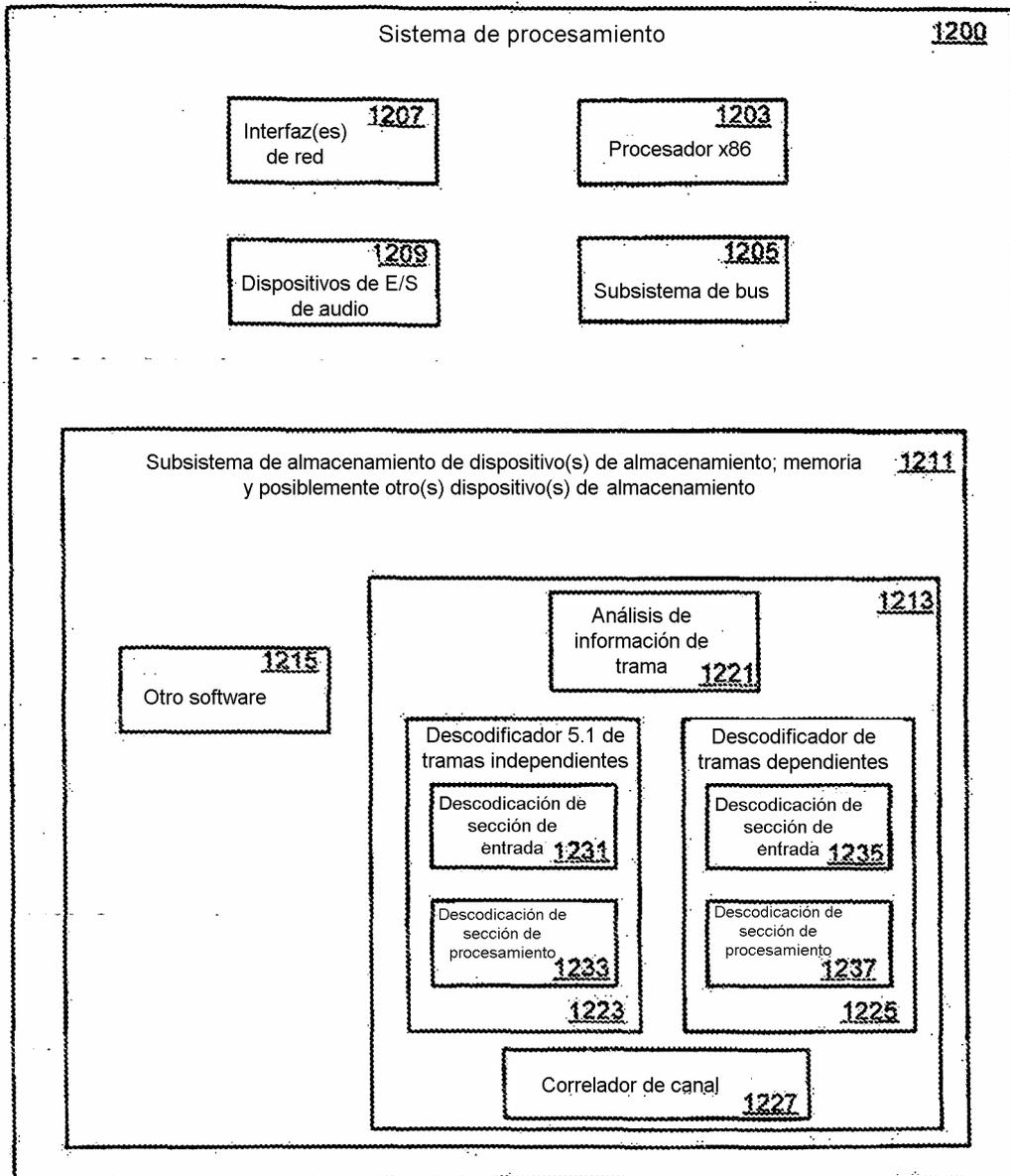


FIG. 12