



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 313 646**

51 Int. Cl.:
G10L 19/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06727674 .1**

96 Fecha de presentación : **16.03.2006**

97 Número de publicación de la solicitud: **1866913**

97 Fecha de publicación de la solicitud: **19.12.2007**

54 Título: **Codificación y descodificación de audio.**

30 Prioridad: **30.03.2005 EP 05102515**
18.04.2005 EP 05103085
11.01.2006 EP 06100245

45 Fecha de publicación de la mención BOPI:
01.03.2009

45 Fecha de la publicación del folleto de la patente:
01.03.2009

73 Titular/es: **Koninklijke Philips Electronics N.V.**
Groenewoudseweg 1
5621 BA Eindhoven, NL

72 Inventor/es: **Hotho, Gerard, H.;**
Myburg, Francois, P. y
Oomen, Arnoldus, W., J.

74 Agente: **Zuazo Araluze, Alexander**

ES 2 313 646 T3

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

DESCRIPCIÓN

Codificación y decodificación de audio.

La invención se refiere a codificación y/o decodificación de audio para señales multicanal.

Una señal de audio multicanal es una señal de audio que tiene dos o más canales de audio. Ejemplos bien conocidos de señales de audio multicanal son las señales de audio estéreo de dos canales y las señales de audio de canal 5.1 que tienen dos canales de audio delanteros, dos canales de audio traseros, una señal de audio central y un canal de mejora de bajas frecuencias (LFE, *Low Frequency Enhancement*) adicional. Tales señales de audio de canal 5.1 se utilizan en sistemas de DVD (*Digital Versatile Disc*, disco versátil digital) y SACD (*Super Audio Compact Disc*, disco compacto de super audio). Debido a la popularidad creciente de material multicanal, está volviéndose más importante la codificación eficaz de material multicanal.

En el campo de procesamiento de audio, es bien conocido convertir un número de canales de audio en otro número de canales de audio. Tal conversión puede realizarse por diversos motivos. Por ejemplo, una señal de audio puede convertirse a otro formato para proporcionar una experiencia al usuario mejorada. Por ejemplo las grabaciones estéreo tradicionales sólo comprenden dos canales mientras que los sistemas de audio avanzados modernos normalmente utilizan cinco o seis canales, como en los populares sistemas de sonido envolvente 5.1. En consecuencia, los dos canales estéreo pueden convertirse en cinco o seis canales para aprovechar completamente el sistema de audio avanzado.

Otro motivo para una conversión de canales es la eficacia de codificación. Se ha descubierto que por ejemplo las señales de audio de sonido envolvente pueden codificarse como señales de audio de canal estéreo combinadas con un flujo de bits de parámetros que describe las propiedades espaciales multicanal de la señal de audio. El decodificador puede reproducir las señales de audio de sonido envolvente con un grado de precisión muy satisfactorio. De esta manera, pueden obtenerse ahorros sustanciales de tasa de bits.

Se conoce un sistema de codificación de audio multicanal 5.1-a-5.1. En este sistema de codificación de audio conocido se codifica una señal de audio de entrada 5.1 en y se representa por dos canales de downmix y parámetros asociados. Las señales de downmix también se denominan conjuntamente como downmix espacial. En el sistema conocido, el downmix espacial forma una señal de audio estéreo que tiene una imagen estéreo, es decir, en cuanto a calidad, comparable con un downmix ITU fijo de los canales de entrada 5.1. Los usuarios que tienen sólo equipos estéreo pueden escuchar este downmix estéreo espacial, mientras que los oyentes con equipos de canal 5.1 pueden escuchar la reproducción de canal 5.1 que se realiza utilizando este downmix estéreo espacial y los parámetros asociados. El equipo de canal 5.1 decodifica/reconstruye la señal de audio de canal 5.1 a partir del downmix estéreo espacial (es decir la señal de audio estéreo) y los parámetros asociados.

Sin embargo, un downmix estéreo espacial se considera a menudo que es de calidad reducida comparado con una señal estéreo original o una señal estéreo generada explícitamente. Por ejemplo, los ingenieros de sonido profesionales tienden a menudo a encontrar el downmix estéreo espacial algo sordo y poco interesante. Por este motivo, a menudo se genera un downmix estéreo artístico, que difiere del downmix estéreo espacial. Por ejemplo se añaden fuentes o reverberación adicionales, se ensancha la imagen estéreo, etc. Con el fin de que los usuarios puedan disfrutar del downmix estéreo artístico, este downmix artístico, en lugar del downmix espacial, puede transmitirse a través de un medio de transmisión o almacenarse en un medio de almacenamiento. Sin embargo, puesto que los datos paramétricos para generar la señal 5.1 de la señal estéreo se basan en la señal de downmix original, este enfoque afecta seriamente a la calidad de la reproducción de señal de audio de canal 5.1. Específicamente, la señal de audio de canal 5.1 de entrada se codificaba en un downmix estéreo espacial y parámetros asociados. Sustituyendo el downmix estéreo espacial por el downmix estéreo artístico, el downmix estéreo espacial ya no puede estar disponible en el extremo de decodificación del sistema y no es posible una reconstrucción de alta calidad de la señal de audio de canal 5.1.

Un posible enfoque para mejorar la calidad de la señal de audio de canal 5.1 es incluir datos adicionales de la señal de downmix estéreo espacial. Por ejemplo, además del downmix estéreo artístico, la señal de downmix estéreo espacial puede incluirse en el mismo flujo de bits o puede transmitirse en paralelo. Sin embargo, esto aumenta sustancialmente la tasa de datos y por tanto los requisitos de almacenamiento o ancho de banda de comunicación y degradará la calidad para relación de tasa de datos para una señal multicanal codificada.

Un esquema conocido para realizar downmix sobre una señal de audio multicanal se da a conocer por HERRE ET AL: "MP3 Surround: Efficient and Compatible Coding of Multi Channel Audio" AUDIO ENGINEERING SOCIETY CONVENTION PAPER, NUEVA YORK, NY, EE.UU, 8 de mayo de 2004 (08-05-2004), páginas 1-14, XP002350798.

Por tanto, sería ventajoso un sistema de codificación/decodificación mejorado para audio multicanal y en particular un sistema que permita un rendimiento, calidad y/o calidad para relación de tasa de datos mejorados.

En consecuencia, la invención según las reivindicaciones adjuntas busca preferiblemente mitigar, paliar o eliminar una a más de las desventajas mencionadas anteriormente por separado o en cualquier combinación.

Según un primer aspecto de la invención se proporciona un codificador de audio multicanal para codificar una señal de audio de N canales, comprendiendo el codificador de audio multicanal: medios para generar una primera señal de M canales para la señal de audio de N canales, siendo M menor que N; medios para generar primeros datos de mejora para la primera señal de M canales con respecto a la señal de audio de N canales; medios para generar una segunda
 5 señal de M canales para la señal de audio de N canales; medios de mejora para generar segundos datos de mejora para la segunda señal de M canales con respecto a la primera señal de M canales; medios para generar una señal de salida codificada que comprende la segunda señal de M canales, los primeros datos de mejora y los segundos datos de mejora; y en el que los medios de mejora están dispuestos para seleccionar dinámicamente entre generar los segundos
 10 datos de mejora como datos de mejora absolutos o como datos de mejora relativos con respecto a la segunda señal de M canales.

La invención puede permitir una codificación eficaz de una señal multicanal. En particular, puede conseguirse una codificación eficaz con una calidad aumentada para relación de tasa de datos. La invención puede permitir que una
 15 señal de M canales sustituya a otra señal de M canales con impacto reducido sobre la generación multicanal basándose en datos de mejora con respecto a la primera señal de M canales. Específicamente, puede transmitirse un downmix artístico en lugar de un downmix espacial mientras se permite una recreación multicanal eficaz en un descodificador basándose en datos de mejora asociados con el downmix espacial. La selección dinámica de datos de mejora permite un tamaño reducido significativamente de los datos de mejora y/o una calidad mejorada de la señal que puede generarse.

Los datos de mejora absolutos describen la primera señal de M canales sin referirse a la segunda señal de M canales mientras que los datos de mejora relativos describen la primera señal de M canales con referencia a la segunda señal de M canales.

Los medios para generar la primera y/o segunda señal de M canales pueden generar las señales procesando la señal de N canales o por ejemplo recibiendo la(s) señal(es) de M canales desde fuentes internas o externas.

Según una propiedad opcional de la invención, los medios de mejora están dispuestos para seleccionar entre los datos de mejora absolutos y los datos de mejora relativos en respuesta a una característica de la señal de N canales.

Esto puede permitir un rendimiento eficaz y en particular puede dotar a una señal codificada de calidad mejorada para relación de tasa de datos. La selección puede realizarse por ejemplo evaluando uno o más parámetros derivados a partir de una característica de un segmento de la señal de N canales y específicamente basándose en uno o más
 35 parámetros derivados a partir de la primera y/o segunda señal de M canales (que pueden derivarse a partir de la señal de N canales).

Según una propiedad opcional de la invención, los medios de mejora están dispuestos para seleccionar entre los datos de mejora absolutos y los datos de mejora relativos en respuesta a una característica relativa de los datos de mejora absolutos y los datos de mejora relativos.

Esto puede permitir un rendimiento eficaz y en particular puede dotar a una señal codificada de calidad mejorada para relación de tasa de datos. Como alternativa o adicionalmente, puede permitir una implementación de baja complejidad y/o eficaz.

Según una propiedad opcional de la invención, la característica relativa es una energía de señal de los datos de mejora absolutos con respecto a una energía de señal de los datos de mejora relativos.

Esto puede permitir un rendimiento eficaz y en particular puede dotar a una señal codificada de calidad mejorada para relación de tasa de datos. Como alternativa o adicionalmente, puede permitir una implementación de baja complejidad y/o eficaz. Específicamente, los medios de mejora pueden seleccionar el tipo de datos de mejora que tiene la
 50 energía de señal más baja.

Según una propiedad opcional de la invención, los medios de mejora están dispuestos para dividir la segunda señal de M canales en bloques de señal y para seleccionar individualmente entre los datos de mejora absolutos y los datos de mejora relativos para cada bloque de señal.

Esto puede permitir un rendimiento eficaz y en particular puede dotar a una señal codificada de calidad mejorada para relación de tasa de datos. Como alternativa o adicionalmente, puede permitir una implementación de baja complejidad y/o eficaz. Los bloques de señal pueden dividirse en el dominio del tiempo y/o la frecuencia y cada bloque de señal puede comprender específicamente un grupo de losas de tiempo/frecuencia. La división en bloques de señal puede aplicarse a la primera señal de M canales y/o la señal de N canales.

Según una propiedad opcional de la invención, los medios de mejora están dispuestos para seleccionar entre los datos de mejora absolutos y los datos de mejora relativos para un bloque de señal basándose sólo en características asociadas con el bloque de señal.

Esto puede permitir un rendimiento eficaz y en particular puede dotar a una señal codificada de calidad mejorada para relación de tasa de datos. Como alternativa o adicionalmente, puede permitir una implementación de baja com-

ES 2 313 646 T3

plejidad y/o eficaz. Específicamente, los medios de mejora pueden seleccionar el tipo de datos de mejora que tiene la energía de señal más baja.

Según una propiedad opcional de la invención, los medios de mejora están dispuestos para generar los datos de mejora como una combinación de los datos de mejora absolutos y los datos de mejora relativos durante un intervalo de tiempo de conmutación de una conmutación entre generar los datos de mejora como datos de mejora absolutos y como datos de mejora relativos.

Esto puede permitir una conmutación mejorada y puede reducir en particular los artefactos asociados con la conmutación. Puede conseguirse una calidad de sonido mejorada. La combinación durante un intervalo de tiempo de conmutación puede aplicarse cuando se conmuta desde datos de mejora absolutos a datos de mejora relativos y/o de datos de mejora relativos a absolutos. La combinación puede conseguirse utilizando una técnica de solapamiento y adición.

Según una propiedad opcional de la invención, la combinación comprende una interpolación entre los datos de mejora absolutos y los datos de mejora relativos.

Esto puede permitir una implementación práctica y eficaz con alta calidad. Puede conseguirse una calidad de sonido mejorada.

Según una propiedad opcional de la invención, los medios para generar la señal de salida codificada están dispuestos para incluir datos que indican si se utilizan datos de mejora relativos o datos de mejora absolutos.

Esto puede permitir un rendimiento eficaz y en particular puede dotar a una señal codificada de calidad mejorada para relación de tasa de datos. Como alternativa o adicionalmente, puede permitir una implementación de baja complejidad y/o eficaz. Los datos de indicación pueden incluir específicamente una indicación de selección para cada bloque de señal.

Según una propiedad opcional de la invención, los segundos datos de mejora comprenden una primera parte de datos de mejora y una segunda parte de datos de mejora, proporcionando la segunda parte una representación de calidad superior de la primera señal de M canales que la primera parte.

Esto puede permitir un rendimiento eficaz y en particular puede dotar a una señal codificada de calidad mejorada para relación de tasa de datos. La primera parte puede tener una tasa de datos inferior que la segunda parte. La segunda parte puede comprender datos que permiten de manera más precisa que un descodificador recree la primera señal de M canales.

Según una propiedad opcional de la invención, los medios de mejora están dispuestos para seleccionar dinámicamente sólo entre generar la segunda parte como datos de mejora absolutos o como datos de mejora relativos.

Esto puede permitir un rendimiento eficaz y en particular puede dotar a una señal codificada de calidad mejorada para relación de tasa de datos.

Según una propiedad opcional de la invención, los medios de mejora están dispuestos para generar datos relativos de la segunda parte con respecto a una señal de referencia generada aplicando datos de mejora de la primera parte a la primera señal de M canales.

Esto puede permitir un rendimiento eficaz y en particular puede dotar a una señal codificada de calidad mejorada para relación de tasa de datos.

Según otro aspecto de la invención, se proporciona un descodificador de audio multicanal para descodificar una señal de audio de N canales, comprendiendo el descodificador de audio multicanal: medios para recibir una señal de audio codificada que comprende una primera señal de M canales para la señal de audio de N canales, siendo M menor que N, primeros datos de mejora para la expansión multicanal, siendo los primeros datos de mejora con respecto a una segunda señal de M canales diferente de la primera señal de M canales; segundos datos de mejora para la primera señal de M canales con respecto a la segunda señal de M canales, comprendiendo los segundos datos de mejora datos de mejora absolutos y datos de mejora relativos con respecto a la primera señal de M canales, y datos de indicación indicativos de si los segundos datos de mejora para un bloque de señal son datos de mejora absolutos o datos de mejora relativos; medios de generación para generar una señal de expansión multicanal de M canales en respuesta a la primera señal de M canales y los segundos datos de mejora; y medios para generar una señal descodificada de N canales en respuesta a la señal de expansión multicanal de M canales y los primeros datos de mejora; y en el que los medios de generación están dispuestos para seleccionar entre aplicar los segundos datos de mejora como datos de mejora absolutos o datos de mejora relativos en respuesta a los datos de indicación.

La invención puede permitir una descodificación de rendimiento alto y eficaz de una señal multicanal. En particular, puede conseguirse una codificación eficaz de una señal con calidad mejorada para una tasa de datos dada. La invención también puede permitir que una señal de M canales sustituya a otra señal de M canales con impacto reducido sobre la generación multicanal basándose en datos de mejora con respecto a la primera señal de M canales.

Específicamente, puede transmitirse un downmix artístico en lugar de un downmix espacial mientras se permite una recreación multicanal eficaz en el descodificador basándose en datos de mejora asociados con el downmix espacial.

Los datos de mejora absolutos describen la segunda señal de M canales sin referirse a la primera señal de M canales mientras que los datos de mejora relativos describen la segunda señal de M canales con referencia a la primera señal de M canales.

Según una propiedad opcional de la invención, los medios de generación están dispuestos para aplicar los segundos datos de mejora a la primera señal de M canales en el dominio del tiempo.

Esto puede permitir un rendimiento eficaz y en particular puede proporcionar una señal descodificada con calidad mejorada para una tasa de datos dada. Como alternativa o adicionalmente, puede permitir una implementación de baja complejidad y/o eficaz.

Según una propiedad opcional de la invención, los medios de generación están dispuestos para aplicar los segundos datos de mejora a la primera señal de M canales en el dominio de la frecuencia.

Esto puede permitir un rendimiento eficaz y en particular puede proporcionar una señal descodificada con calidad mejorada para una tasa de datos dada. Como alternativa o adicionalmente, puede permitir una implementación de baja complejidad y/o eficaz.

En particular, en muchas realizaciones, la aplicación del dominio de la frecuencia puede reducir el número requerido de transformadas de tiempo a frecuencia. El dominio de la frecuencia puede ser por ejemplo un dominio de banco de filtros espejo en cuadratura (QMF, *Quadrature Mirror Filterbank*) o de transformada de coseno discreta modificada (MDCT, *Modified Discrete Cosine Transform*).

Según una propiedad opcional de la invención, los segundos datos de mejora comprenden una primera parte de datos de mejora y una segunda parte de datos de mejora, proporcionando la segunda parte una representación de calidad superior de la primera señal de M canales que la primera parte.

Esto puede permitir un rendimiento eficaz y en particular puede proporcionar una señal descodificada con calidad mejorada para una tasa de datos dada. Como alternativa o adicionalmente, puede permitir una implementación de baja complejidad y/o eficaz. La segunda parte puede comprender datos que permiten de manera más precisa que un descodificador recree la primera señal de M canales.

Según una propiedad opcional de la invención, los medios de generación están dispuestos para seleccionar sólo entre aplicar segundos datos de mejora de la segunda parte como datos de mejora absolutos o datos de mejora relativos.

Esto puede permitir un rendimiento eficaz y en particular puede proporcionar una señal descodificada con calidad mejorada para una tasa de datos dada. Como alternativa o adicionalmente, puede permitir una implementación de baja complejidad y/o eficaz.

Según una propiedad opcional de la invención, los medios de generación están dispuestos para generar la expansión multicanal de M canales aplicando datos de mejora relativos de la segunda parte a una señal generada aplicando datos de mejora de la primera parte a la primera señal de M canales.

Esto puede permitir un rendimiento eficaz y en particular puede proporcionar una señal descodificada con calidad mejorada para una tasa de datos dada. Como alternativa o adicionalmente, puede permitir una implementación de baja complejidad y/o eficaz.

Según otro aspecto de la invención, se proporciona un procedimiento de codificación de una señal de audio de N canales, comprendiendo el procedimiento: generar una primera señal de M canales para la señal de audio de N canales, siendo M menor que N; generar primeros datos de mejora para la primera señal de M canales con respecto a la señal de audio de N canales; generar una segunda señal de M canales para la señal de audio de N canales; generar segundos datos de mejora para la segunda señal de M canales con respecto a la primera señal de M canales; generar una señal de salida codificada que comprende la segunda señal de M canales, los primeros datos de mejora y los segundos datos de mejora; y en el que la generación de los segundos datos de mejora comprende seleccionar dinámicamente entre generar los segundos datos de mejora como datos de mejora absolutos o como datos de mejora relativos con respecto a la segunda señal de M canales.

Según otro aspecto de la invención, se proporciona un procedimiento de descodificación de una señal de audio de N canales, comprendiendo el procedimiento: recibir una señal de audio codificada que comprende una primera señal de M canales para la señal de audio de N canales, siendo M menor que N, primeros datos de mejora para la expansión multicanal, siendo los primeros datos de mejora con respecto a una segunda señal de M canales diferente de la primera señal de M canales; segundos datos de mejora para la primera señal de M canales con respecto a la segunda señal de M canales, comprendiendo los segundos datos de mejora absolutos y datos de mejora relativos con respecto a la primera señal de M canales, y datos de indicación indicativos de si los segundos datos de mejora para un bloque de señal son datos de mejora absolutos o datos de mejora relativos; generar una señal de expansión

multicanal de M canales en respuesta a la primera señal de M canales y los segundos datos de mejora; y generar una señal descodificada de N canales en respuesta a la señal de expansión multicanal de M canales y los primeros datos de mejora; y en el que la generación de la señal de expansión multicanal de M canales comprende seleccionar entre aplicar los segundos datos de mejora como datos de mejora absolutos o datos de mejora relativos en respuesta a los
5 datos de indicación.

Según otro aspecto de la invención, se proporciona una señal de audio multicanal codificada para una señal de audio de N canales que comprende: datos de señal de M canales para la señal de audio de N canales, siendo M menor que N; primeros datos de mejora para la expansión multicanal, siendo los primeros datos de mejora con respecto a
10 una segunda señal de M canales diferente de la primera señal de M canales; segundos datos de mejora para la primera señal de M canales con respecto a la segunda señal de M canales, comprendiendo los segundos datos de mejora datos de mejora absolutos y datos de mejora relativos con respecto a la primera señal de M canales; y datos de indicación indicativos de si los segundos datos de mejora para un bloque de señal son datos de mejora absolutos o datos de mejora relativos.

Según otro aspecto de la invención, se proporciona un medio de almacenamiento que tiene almacenado en el mismo una señal tal como se describió anteriormente.

Según otro aspecto de la invención, se proporciona un transmisor para transmitir una señal de audio multicanal codificada, comprendiendo el transmisor un codificador de audio multicanal tal como se describió anteriormente.

Según otro aspecto de la invención, se proporciona un receptor para recibir una señal de audio multicanal, comprendiendo el receptor un descodificador de audio multicanal tal como se describió anteriormente.

Según otro aspecto de la invención, se proporciona un sistema de transmisión que comprende un transmisor para transmitir una señal de audio multicanal codificada a través de un canal de transmisión a un receptor, comprendiendo el transmisor un codificador de audio multicanal tal como se describió anteriormente y comprendiendo el receptor un descodificador de audio multicanal tal como se describió anteriormente.

Según otro aspecto de la invención, se proporciona un procedimiento de transmisión de una señal de audio multicanal codificada, comprendiendo el procedimiento codificar una señal de audio de N canales, en el que la codificación comprende: generar una primera señal de M canales para la señal de audio de N canales, siendo M menor que N; generar primeros datos de mejora para la primera señal de M canales con respecto a la señal de audio de N canales; generar una segunda señal de M canales para la señal de audio de N canales; generar segundos datos de mejora para
35 la segunda señal de M canales con respecto a la primera señal de M canales; generar una señal de salida codificada que comprende la segunda señal de M canales, los primeros datos de mejora y los segundos datos de mejora; y en el que la generación de los segundos datos de mejora comprende seleccionar dinámicamente entre generar los segundos datos de mejora como datos de mejora absolutos o como datos de mejora relativos con respecto a la segunda señal de M canales.

Según otro aspecto de la invención, se proporciona un procedimiento de recepción de una señal de audio multicanal codificada, comprendiendo el procedimiento descodificar la señal de audio multicanal codificada, comprendiendo la descodificación recibir la señal de audio multicanal codificada que comprende una primera señal de M canales para la señal de audio de N canales, siendo M menor que N, primeros datos de mejora para la expansión multicanal, siendo los primeros datos de mejora con respecto a una segunda señal de M canales diferente de la primera señal de M canales; segundos datos de mejora para la primera señal de M canales con respecto a la segunda señal de M canales, comprendiendo los segundos datos de mejora datos de mejora absolutos y datos de mejora relativos con respecto a la primera señal de M canales, y datos de indicación indicativos de si los segundos datos de mejora para un bloque de señal son datos de mejora absolutos o datos de mejora relativos; generar una señal de expansión multicanal de M canales en respuesta a la primera señal de M canales y los segundos datos de mejora; y generar una señal descodificada de N canales en respuesta a la señal de expansión multicanal de M canales y los primeros datos de mejora; y en el que la generación de la señal de expansión multicanal de M canales comprende seleccionar entre aplicar los segundos datos de mejora como datos de mejora absolutos o datos de mejora relativos en respuesta a los datos de indicación.

Según otro aspecto de la invención, se proporciona un procedimiento de transmisión y recepción de una señal de audio, comprendiendo el procedimiento: codificar una señal de audio de N canales, en el que la codificación comprende: generar una primera señal de M canales para la señal de audio de N canales, siendo M menor que N, generar primeros datos de mejora para la primera señal de M canales con respecto a la señal de audio de N canales, generar una segunda señal de M canales para la señal de audio de N canales, generar segundos datos de mejora para la segunda señal de M canales con respecto a la primera señal de M canales, comprendiendo la generación de los segundos datos de mejora seleccionar dinámicamente entre generar los segundos datos de mejora como datos de mejora absolutos o como datos de mejora relativos con respecto a la segunda señal de M canales que genera una señal de salida codificada que comprende la segunda señal de M canales, los primeros datos de mejora y los segundos datos de mejora; transmitir la señal de salida codificada desde un transmisor hasta un receptor; recibir, en el receptor, la señal de salida codificada; descodificar la señal de salida codificada donde la descodificación comprende: generar una señal de expansión multicanal de M canales en respuesta a la segunda señal de M canales y los segundos datos de mejora, comprendiendo la generación de la señal de expansión multicanal de M canales seleccionar entre aplicar los segundos

ES 2 313 646 T3

datos de mejora como datos de mejora absolutos o datos de mejora relativos, y generar una señal descodificada de N canales en respuesta a la señal de expansión multicanal de M canales y los primeros datos de mejora.

Según otro aspecto de la invención, se proporciona un producto de programa informático operativo para hacer que un procesador realice las etapas del procedimiento descrito anteriormente.

Según otro aspecto de la invención, se proporciona a grabador de audio multicanal que comprende un codificador de audio multicanal tal como se describió anteriormente.

Según otro aspecto de la invención, se proporciona un reproductor (60) de audio multicanal que comprende un descodificador de audio multicanal tal como se describió anteriormente.

Estos y otros aspectos, propiedades, y ventajas de la invención resultarán evidentes a partir de y se aclararán con referencia a la(s) realización/realizaciones descritas a continuación en el presente documento.

Se describirán realizaciones de la invención, sólo a modo de ejemplo, con referencia a los dibujos, en los que:

La figura 1 muestra un diagrama de bloques de un codificador de audio multicanal según algunas realizaciones de la invención;

La figura 2 muestra un diagrama de bloques de un descodificador de audio multicanal según algunas realizaciones de la invención;

La figura 3 muestra un diagrama de bloques de un sistema de transmisión según algunas realizaciones de la invención;

La figura 4 muestra un diagrama de bloques de un reproductor/grabador de audio multicanal según algunas realizaciones de la invención;

La figura 5 muestra un diagrama de bloques de un codificador de audio multicanal según algunas realizaciones de la invención;

La figura 6 muestra un diagrama de bloques de un generador de datos de mejora según algunas realizaciones de la invención;

La figura 7 muestra un diagrama de bloques de un descodificador de audio multicanal según algunas realizaciones de la invención;

La figura 8 muestra un diagrama de bloques de elementos de un descodificador de audio multicanal;

La figura 9 muestra un diagrama de bloques de elementos de un descodificador de audio multicanal según algunas realizaciones de la invención;

La figura 10 muestra un diagrama de bloques de elementos de un descodificador de audio multicanal según algunas realizaciones de la invención; y

La figura 11 muestra un diagrama de bloques de elementos de un descodificador de audio multicanal según algunas realizaciones de la invención.

La siguiente descripción se centra en realizaciones de la invención aplicables a un codificador 5.1 a 2 y/o un descodificador 2 a 5.1. Sin embargo, se apreciará que la invención no está limitada a esta aplicación.

La figura 1 muestra un diagrama de bloques de una realización de un codificador 10 de audio multicanal según algunas realizaciones de la invención. Este codificador 10 de audio multicanal está dispuesto para codificar N señales 101 de audio en M señales 102 de audio y datos 104, 105 paramétricos asociados. En este caso, M y N son enteros, con $N > M$ y $M \geq 1$. Un ejemplo del codificador 10 de audio multicanal es un codificador 5.1 a 2 en el que N es igual a 6, es decir 5+1 canales, y M es igual a 2. Tal codificador de audio multicanal codifica una señal de audio de entrada de canal 5.1 en una señal de audio de salida de 2 canales, por ejemplo una señal de audio de salida estéreo, y parámetros asociados. Otros ejemplos del codificador 10 de audio multicanal son codificadores 5.1 a 1, 6.1 a 2, 6.1 a 1, 7.1 a 2 y 7.1 a 1. También son posibles codificadores que tienen otros valores para N y M siempre que N sea mayor que M y siempre que M sea mayor que o igual a 1.

El codificador 10 comprende una primera unidad 110 de codificación y acoplada a la misma una segunda unidad 120 de codificación. La primera unidad 110 de codificación recibe las N señales 101 de audio de entrada y codifica las N señales 101 de audio en las M señales 102 de audio y primeros datos 104 paramétricos asociados. Las M señales 102 de audio y los primeros datos 104 paramétricos asociados representan las N señales 101 de audio. La codificación de las N señales 101 de audio en las M señales 102 de audio según se realiza por la primera unidad 110 puede también denominarse como realización de downmix y las M señales 102 de audio también pueden

denominarse como downmix 102 espacial. La unidad 110 puede ser un codificador de audio multicanal paramétrico convencional que codifica una señal 101 de audio multicanal en una señal 102 de audio de downmix mono o estéreo y parámetros asociados 104. Los parámetros asociados 104 permiten que un descodificador reconstruya la señal 101 de audio multicanal a partir de la señal 102 de audio de downmix mono o estéreo. Se observa que el downmix 102 también puede tener más de dos canales.

La primera unidad 110 suministra el downmix 102 espacial a la segunda unidad 120. La segunda unidad 120 genera, a partir del downmix 102 espacial, segundos datos de mejora en la forma de segundos datos 105 paramétricos asociados. Los segundos datos 105 paramétricos asociados representan el downmix 102 espacial, es decir estos parámetros 105 comprenden características o propiedades del downmix 102 espacial que permiten que un descodificador reconstruya al menos parte del downmix 102 espacial, por ejemplo sintetizando una señal que se parece al downmix 102 espacial. Los datos paramétricos asociados comprenden los primeros y segundos datos 104 y 105 paramétricos asociados.

Los segundos datos 105 paramétricos asociados comprenden parámetros de modificación que permiten una reconstrucción del downmix 102 espacial a partir de $K (=M)$ señales 103 de audio adicionales. De esta manera, un descodificador puede realizar una reconstrucción incluso mejor del downmix 102 espacial. Esta reconstrucción puede realizarse basándose en un downmix 103 alternativo, es decir las K señales 103 de audio adicionales, tal como un downmix artístico. Un descodificador puede aplicar los parámetros de modificación a la señal 103 de downmix alternativo de modo que se parezca más al downmix 102 espacial.

La segunda unidad 120 puede recibir en sus entradas el downmix 103 alternativo. El downmix 103 alternativo puede recibirse desde una fuente externa al codificador 10 (tal como se muestra en la figura 1) o, como alternativa, el downmix 103 alternativo puede generarse dentro del codificador 10 (no mostrado), por ejemplo a partir de las N señales 101 de audio. La segunda unidad 120 puede comparar al menos parte del downmix 102 espacial con el downmix 103 alternativo y generar parámetros 105 de modificación que representan una diferencia entre el downmix 102 espacial y el downmix 103 alternativo, por ejemplo una diferencia entre una propiedad del downmix 102 espacial y una propiedad del downmix 103 alternativo. En el ejemplo, el downmix 103 alternativo es específicamente un downmix artístico asociado con el downmix espacial.

En el ejemplo, la segunda unidad 120 puede generar además los parámetros de modificación como valores absolutos que representan directamente el downmix 102 espacial sin ninguna referencia al downmix 103 alternativo. Además, la segunda unidad 120 comprende funcionalidad para seleccionar entre los parámetros de modificación relativos y absolutos para la señal de salida de codificador. Específicamente, esta selección se realiza dinámicamente y puede realizarse para bloques de señal individuales dependiendo de las características de los datos paramétricos y/o la señal.

Además, la segunda unidad 120 puede comprender funcionalidad para incluir una indicación de qué parámetros de modificación (absolutos o relativos) se han utilizado para diferentes secciones de la señal codificada. Por ejemplo, para cada bloque de señal, puede incluirse un bit de datos para indicar si se han incluidos datos paramétricos relativos o absolutos para ese bloque de señal.

Los parámetros 105 de modificación comprenden preferiblemente (una diferencia entre) una o más propiedades de señal estadísticas tales como varianza, covarianza y correlación, o una relación de estas propiedades, o de la(s) diferencia entre la(s) señal(es) de downmix. Se observa que la varianza de una señal es equivalente a la energía o potencia de esa señal. Estas propiedades de señal estadísticas permiten una buena reconstrucción del downmix espacial.

La figura 2 muestra un diagrama de bloques de una realización de un descodificador 20 de audio multicanal según algunas realizaciones de la invención. El descodificador 20 está dispuesto para descodificar K señales 103 de audio y datos 104, 105 paramétricos asociados en N señales de audio 203. En este caso, K y N son enteros, con $N > K$ y $K \geq 1$. Las K señales 103 de audio, es decir el downmix 103 alternativo, y los datos 104, 105 paramétricos asociados representan las N señales de audio 203, es decir la señal 203 de audio multicanal. Un ejemplo del descodificador 20 de audio multicanal es un descodificador 2 a 5.1 en el que N es igual a 6, es decir 5+1 canales, y K es igual a 2. Tal descodificador de audio multicanal descodifica una señal de audio de entrada de 2 canales, por ejemplo una señal de audio de entrada estéreo, y parámetros asociados en una señal de audio de salida de canal 5.1. Otros ejemplos del descodificador 20 de audio multicanal son descodificadores 1 a 5.1, 2 a 6.1, 1 a 6.1, 2 a 7.1 y 1 a 7.1. También son posibles descodificadores que tienen otros valores para N y K siempre que N sea mayor que K y siempre que K sea mayor que o igual a 1.

El descodificador 20 de audio multicanal comprende una primera unidad 210 y acoplada a la misma una segunda unidad 220. La primera unidad 210 recibe el downmix 103 alternativo y datos de mejora en la forma de parámetros 105 de modificación y reconstruye M señales 202 de audio adicionales, es decir el downmix 202 espacial o una aproximación del mismo, a partir del downmix 103 alternativo y los parámetros 105 de modificación. En este caso, M es un entero, con $M \geq 1$. Los parámetros 105 de modificación representan el downmix 202 espacial. La primera unidad 210 está dispuesta específicamente para determinar si los parámetros 105 de modificación son parámetros de modificación absolutos o relativos y para aplicar los parámetros en consecuencia. Específicamente, la primera unidad 210 puede determinar si los parámetros 105 de modificación para bloques de señal individuales son parámetros

relativos o absolutos basándose en datos explícitos en el flujo de bits recibido. Por ejemplo, puede incluirse un único bit de datos para cada bloque de señal que indica si los parámetros son parámetros de modificación absolutos o relativos en ese bloque de señal.

La segunda unidad 220 recibe el downmix 202 espacial desde la primera unidad 210 y parámetros de modificación 104. La segunda unidad 220 descodifica el downmix 202 espacial y parámetros de modificación 104 en la señal 203 de audio multicanal. La segunda unidad 220 puede ser un descodificador de audio multicanal paramétrico convencional que descodifica una señal 202 de audio de downmix mono o estéreo y parámetros asociados 104 en una señal 203 de audio multicanal.

La primera unidad 210 puede estar dispuesta para determinar si es necesario o deseable reconstruir la señal 202 a partir de la señal 103 de entrada. Tal reconstrucción puede no ser aplicable cuando la señal 202 de downmix espacial se suministra a la primera unidad 210 en lugar del downmix 103 alternativo. La primera unidad 210 puede determinar esto generando a partir de la señal 103 de entrada propiedades de señal similares o iguales que las comprendidas en los parámetros 105 de modificación y comparando estas propiedades de señal generadas con los parámetros 105 de modificación. Si esta comparación muestra que las propiedades de señal generadas son iguales o sustancialmente iguales a los parámetros 105 de modificación entonces la señal 103 de entrada se parece suficientemente a la señal 202 de downmix espacial y la primera unidad 210 puede reenviar la señal 103 de entrada a la segunda unidad 220. Si la comparación muestra que las propiedades de señal generadas no son iguales a o sustancialmente iguales a los parámetros 105 de modificación entonces la señal 103 de entrada no se parece suficientemente a la señal 202 de downmix espacial y la primera unidad 210 puede reconstruir/aproximar la señal 202 de downmix espacial a partir de la señal 103 de entrada y los parámetros 105 de modificación.

La primera unidad 210 puede generar, a partir del downmix alternativo, propiedades/parámetros de modificación adicionales que representan el downmix 103 alternativo. En tal caso, la primera unidad 210 puede reconstruir el downmix 202 espacial a partir del downmix 103 alternativo y (una diferencia entre) los parámetros 105 de modificación y los parámetros de modificación adicionales.

Los parámetros 105 de modificación y los parámetros de modificación adicionales, respectivamente, pueden incluir propiedades estadísticas del downmix 202 espacial y el downmix 103 alternativo, respectivamente. Estas propiedades estadísticas tales como varianza, correlación y covarianza, etc. proporcionan buenas representaciones de las señales a partir de las que se derivan. Son útiles al reconstruir el downmix 202 espacial, por ejemplo transformando el downmix alternativo de manera que sus propiedades asociadas coincidan con las propiedades comprendidas en los parámetros 105 de modificación.

La figura 3 muestra un diagrama de bloques de una realización de un sistema 70 de transmisión según algunas realizaciones de la invención. El sistema 70 de transmisión comprende un transmisor 40 para transmitir una señal de audio multicanal codificada a través de un canal 30 de transmisión, por ejemplo un enlace de comunicación cableado o inalámbrico, a un receptor 50. El transmisor 40 comprende un codificador 10 de audio multicanal tal como se describió anteriormente para codificar la señal 101 de audio multicanal en un downmix 102 espacial y parámetros 104, 105 asociados. El transmisor 40 comprende además medios 41 para transmitir una señal de audio multicanal codificada que comprende los parámetros 104, 105 y el downmix 102 espacial o el downmix 103 alternativo a través del canal 30 de transmisión al receptor 50. El receptor 50 comprende medios 51 para recibir la señal de audio multicanal codificada y un descodificador 20 de audio multicanal tal como se describió anteriormente para descodificar el downmix 103 alternativo o el downmix 102 espacial y los parámetros 104, 105 asociados en la señal 203 de audio multicanal.

La figura 4 muestra un diagrama de bloques de una realización de un reproductor/grabador 60 de audio multicanal según algunas realizaciones de la invención. El reproductor/grabador 60 de audio comprende un descodificador 20 de audio multicanal y/o un codificador 10 de audio multicanal según algunas realizaciones de la invención. El reproductor/grabador 60 de audio puede tener su propio almacenamiento por ejemplo memoria de estado sólido o disco duro. El reproductor/grabador 60 de audio también puede facilitar medios de almacenamiento separables tales como discos DVD (grabables) o discos CD (grabables). Las señales de audio multicanal codificadas almacenadas que comprenden un downmix 103 alternativo y parámetros 104, 105 pueden descodificarse mediante el descodificador 20 y pueden reproducirse mediante el reproductor/grabador 60 de audio. El codificador 10 puede codificar señales de audio multicanal para su almacenamiento en los medios de almacenamiento.

La figura 5 muestra un diagrama de bloques de un codificador 10 de audio multicanal según algunas realizaciones de la invención. El codificador de la figura 5 puede ser específicamente el codificador 10 de la figura 1. El codificador 10 comprende una primera unidad 110 y acoplada a la misma una segunda unidad 120. La primera unidad 110 recibe una señal 101 de audio multicanal 5.1 que comprende señales de audio delantera izquierda, trasera izquierda, delantera derecha, trasera derecha, central y de mejora de bajas frecuencias lf, lr, rf, rr, co y lfe, respectivamente. La segunda unidad 120 recibe un downmix 103 estéreo artístico que comprende señales de audio artística izquierda y artística derecha la y ra, respectivamente. La señal 101 de audio multicanal y el downmix 103 artístico son señales de audio en el dominio del tiempo. En la primera y segunda unidades 110 y 120 estas señales 101 y 103 se segmentan y transforman al dominio de frecuencia-tiempo.

En la primera unidad 110, se derivan datos 104 paramétricos en tres fases. En una primera fase, tres pares de señales de audio lf y rf, rf y rr, y co y lfe, respectivamente, se segmentan y las señales segmentadas se transforman al dominio

de la frecuencia en unidades 112, 113, y 114 de segmentación y transformación, respectivamente. Las representaciones resultantes en el dominio de la frecuencia de las señales segmentadas se muestran como señales en el dominio de la frecuencia L_f , L_r , R_f , R_r , Co y LFE , respectivamente. En una segunda fase, se realiza downmix sobre tres pares de estas señales en el dominio de la frecuencia L_f y L_r , R_f y R_r , y Co y LFE , respectivamente, en los elementos 115, 116, y 117 de realización de downmix, respectivamente, para generar señales de audio mono L , R , y C , respectivamente y parámetros 141, 142, y 143 asociados, respectivamente. Los elementos 115, 116, y 117 de realización de downmix pueden ser codificadores de estéreo paramétrico MPEG4 convencionales. Finalmente, en una tercera fase se realiza downmix sobre las tres señales de audio mono L , R y C en un elemento 118 de realización de downmix para obtener downmix 102 estéreo espacial y parámetros 144 asociados. El downmix 102 espacial comprende las señales Lo y Ro .

Los datos 141, 142, 143, y 144 paramétricos están comprendidos en los primeros datos de mejora en la forma de primeros datos 104 paramétricos asociados. Los datos 104 paramétricos y el downmix 102 espacial representan las señales 101 de entrada 5.1.

En la segunda unidad, la señal 103 de downmix artístico representada en el dominio del tiempo mediante señales de audio la y ra , respectivamente, se segmenta en primer lugar en una unidad 121 de segmentación. La señal 127 de audio segmentada resultante comprende las señales las y ras , respectivamente. A continuación, esta señal 127 de audio segmentada se transforma al dominio de la frecuencia mediante un transformador 122. La señal 126 resultante en el dominio de la frecuencia comprende las señales La y Ra . Finalmente, la señal 126 en el dominio de la frecuencia, que es una representación en el dominio de la frecuencia del downmix 103 artístico segmentado, y la representación en el dominio de la frecuencia del downmix 102 espacial segmentado se suministran a un generador 123 que genera además (segundos) datos de mejora en la forma de parámetros 105 de modificación que permiten que un descodificador modifique/ transforme el downmix 103 artístico de modo que se parezca más al downmix 102 espacial.

En el ejemplo específico, la señal 127 segmentada en el dominio del tiempo también se proporciona a un selector 124. Las otras dos entradas para este selector 124 son la representación en el dominio de la frecuencia del downmix 102 estéreo espacial y una señal 128 de control. La señal 128 de control determina si el selector 124 ha de emitir el downmix 103 artístico o el downmix 102 espacial como parte de la señal de audio multicanal codificada. El downmix 102 espacial puede seleccionarse cuando no está disponible el downmix artístico. La señal 128 de control puede ajustarse manualmente o puede generarse automáticamente detectando la presencia del downmix 103 artístico. La señal 128 de control puede incluirse en el flujo de bits de parámetros de modo que un descodificador 20 correspondiente puede hacer uso de esto tal como se describe posteriormente. Por tanto, el codificador a modo de ejemplo específico permite que se genere una señal que incluye el downmix 102 espacial o el downmix 103 artístico.

La señal 102, 103 de salida del selector 124 se muestra como las señales lo y ro . Si ha de emitirse el downmix 127 estéreo artístico mediante el selector 124 se combinan las señales en el dominio del tiempo segmentadas las y ras en el selector 124 mediante solapamiento-adición en las señales lo y ro . Si ha de emitirse el downmix 102 estéreo espacial tal como se indica mediante la señal 128 de control, el selector 124 transforma las señales Lo y Ro de nuevo al dominio del tiempo y las combina a través de solapamiento-adición en las señales lo y ro . Las señales en el dominio del tiempo lo y ro forman el downmix estéreo del codificador 10 5.1 a 2.

A continuación se proporciona una descripción más detallada del generador 123. La función del generador 123 es determinar segundos datos de mejora y específicamente parámetros de modificación que describen una transformación del downmix 103 artístico de modo que, en cierto sentido, se parezca el downmix 102 espacial original.

En general, esta transformación puede describirse como

$$\begin{bmatrix} \underline{L}_d & \underline{R}_d \end{bmatrix} = \begin{bmatrix} \underline{L}_a & \underline{R}_a & \underline{A}_1 & \dots & \underline{A}_N \end{bmatrix} \underline{T} \quad (1)$$

donde \underline{L}_a y \underline{R}_a son vectores que comprenden muestras de una losa de tiempo/frecuencia del canal izquierdo y derecho del downmix 103 artístico, y donde \underline{L}_d y \underline{R}_d son vectores que comprenden muestras de una losa de tiempo/frecuencia del canal izquierdo y derecho del downmix artístico modificado, donde $\underline{A}_1, \dots, \underline{A}_N$ comprenden las muestras de una losa de tiempo/frecuencia de canales auxiliares opcionales, y donde \underline{T} es una matriz de transformación. Obsérvese que cualquier vector \underline{V} se define como un vector columna. El downmix artístico modificado es el downmix 103 artístico que se transforma mediante la transformada de modo que se parece al downmix 102 espacial original. Los canales $\underline{A}_1, \dots, \underline{A}_N$ auxiliares son en el sistema descrito las señales de downmix espacial o el contenido de baja frecuencia de las mismas.

La matriz de transformación \underline{T} $(N+2) \times 2$ describe la transformación del downmix 103 artístico y los canales auxiliares al downmix artístico modificado. La matriz de transformación \underline{T} o elementos de la misma están comprendidos preferiblemente en los parámetros 105 de modificación de modo que un descodificador 20 puede reconstruir al menos parte de la matriz de transformación \underline{T} . Después, el descodificador 20 puede aplicar la matriz de transformación \underline{T} al downmix 103 artístico para reconstruir el downmix 102 espacial (como se describe posteriormente).

Como alternativa, los parámetros 105 de modificación comprenden propiedades de señal, por ejemplo valores de energía o potencia y/o valores de correlación, del downmix 102 espacial. El decodificador 20 puede generar tales propiedades de señal a partir del downmix 103 artístico. Las propiedades de señal del downmix 102 espacial y el downmix 103 artístico permiten que el decodificador 20 construya una matriz de transformación T (descrito posteriormente) y aplicarla al downmix 103 artístico para reconstruir el downmix 102 espacial (también descrito posteriormente).

Específicamente, el generador 123 está dispuesto para generar tanto datos de modificación relativos como absolutos y para seleccionar entre estos datos para bloques (o segmentos) de señal individuales. Por tanto, los parámetros 105 de modificación para la señal codificada comprenden tanto datos de modificación absolutos como datos de modificación relativos para diferentes bloques de señal. A diferencia de los datos de modificación absolutos, los datos de modificación relativos describen el downmix 102 espacial con respecto al downmix 103 artístico. Específicamente, los datos de modificación relativos pueden ser datos diferenciales que permiten modificar muestras de downmix artístico para corresponder (más) a las muestras de downmix espacial mientras que los datos de downmix absolutos pueden corresponder directamente a las muestras de downmix espacial sin ninguna referencia o dependencia de las muestras de downmix artístico.

Se apreciará que hay varias maneras de modificar el downmix 103 estéreo artístico para parecerse al downmix 102 estéreo original, incluyendo:

- I. Hacer coincidir las formas de onda.
- II. Hacer coincidir las propiedades estadísticas:
 - a. Hacer coincidir la energía o potencia del canal izquierdo y derecho.
 - b. Hacer coincidir la matriz de covarianza del canal izquierdo y derecho.
- III. Obtener la mejor coincidencia posible de la forma de onda bajo la limitación de una coincidencia de energía o potencia del canal izquierdo y derecho.
- IV. Mezclar los procedimientos I-III mencionados anteriormente.

Para mayor claridad, los canales $\underline{A}_1, \dots, \underline{A}_N$ auxiliares de (1) en primer lugar no se consideran, de modo que la matriz de transformación T puede escribirse como

$$[\underline{L}_d \ \underline{R}_d] = [\underline{L}_a \ \underline{R}_a] T \quad (2)$$

y los datos de mejora relativos pueden por ejemplo generarse según lo siguiente:

I. Coincidencia de forma de onda (procedimiento I)

Puede obtenerse una coincidencia de las formas de onda del downmix 103 artístico y el downmix 102 espacial expresando tanto la señal izquierda como derecha del downmix artístico modificado como una combinación lineal de la señal izquierda y derecha del downmix 103 estéreo artístico:

$$\underline{L}_d = \alpha_1 \underline{L}_a + \beta_1 \underline{R}_a, \quad \underline{R}_d = \alpha_2 \underline{L}_a + \beta_2 \underline{R}_a. \quad (3)$$

Entonces, la matriz T de (2) puede escribirse como:

$$T = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix}.$$

Una manera de elegir los parámetros $\alpha_1, \alpha_2, \beta_1$ y β_2 , es minimizar el cuadrado de la distancia euclídea entre las señales de downmix espacial \underline{L}_s y \underline{R}_s y sus estimaciones (es decir las señales de downmix artístico modificado \underline{L}_d y \underline{R}_d), por tanto

$$\min_{\alpha_1, \beta_1} \sum_k \|\underline{L}_s[k] - \underline{L}_d[k]\|^2 = \min_{\alpha_1, \beta_1} \sum_k \|\underline{L}_s[k] - \alpha_1 \underline{L}_a[k] - \beta_1 \underline{R}_a[k]\|^2 \quad (4)$$

y

$$\min_{\alpha_2, \beta_2} \sum_k \|R_s[k] - R_d[k]\|^2 = \min_{\alpha_2, \beta_2} \sum_k \|R_s[k] - \alpha_2 L_a[k] - \beta_2 R_a[k]\|^2. \quad (5)$$

II. Hacer coincidir las propiedades estadísticas (procedimiento II)

Procedimiento II.a: ahora se comenta la coincidencia de las energías de las señales izquierda y derecha. Las señales de downmix artístico modificado izquierda y derecha, denotadas por L_d y R_d respectivamente, se calculan ahora como

$$\underline{L}_d = \alpha \underline{L}_a, \quad \underline{R}_d = \beta \underline{R}_a, \quad (6)$$

donde, en el caso de parámetros reales, α y β vienen dados por

$$\alpha = \sqrt{\frac{\sum_k \|L_s[k]\|^2}{\sum_k \|L_a[k]\|^2}}, \quad \beta = \sqrt{\frac{\sum_k \|R_s[k]\|^2}{\sum_k \|R_a[k]\|^2}}, \quad (7)$$

de modo que la matriz de transformación T puede escribirse como

$$T = \begin{bmatrix} \sqrt{\frac{\sum_k \|L_s[k]\|^2}{\sum_k \|L_a[k]\|^2}} & 0 \\ 0 & \sqrt{\frac{\sum_k \|R_s[k]\|^2}{\sum_k \|R_a[k]\|^2}} \end{bmatrix}. \quad (8)$$

Con estas elecciones puede garantizarse que las señales L_d y R_d , respectivamente, tienen la misma energía que las señales L_s y R_s , respectivamente.

Procedimiento II.b: para hacer coincidir las matrices de covarianza del downmix 103 estéreo artístico y el downmix 102 estéreo espacial estas matrices pueden descomponerse utilizando descomposición en autovalores según lo siguiente:

$$\begin{aligned} C_a &= U_a S_a U_a^H, \\ C_0 &= U_0 S_0 U_0^H, \end{aligned} \quad (9)$$

donde la matriz de covarianza del downmix 103 estéreo artístico, C_a , viene dado por

$$C_a = [\underline{L}_a \quad \underline{R}_a]^H [\underline{L}_a \quad \underline{R}_a]. \quad (10)$$

U_a es una matriz unitaria y S_a es una matriz diagonal. C_0 es la matriz de covarianza del downmix 102 estéreo espacial, U_0 es una matriz unitaria y S_0 es una matriz diagonal. Cuando se calcula

$$X_{aw} = [\underline{L}_{aw} \quad \underline{R}_{aw}] = [\underline{L}_a \quad \underline{R}_a] U_a S_a^{-1/2}, \quad (11)$$

Se obtienen dos señales \underline{L}_{aw} y \underline{R}_{aw} descorrelacionadas entre sí (debido a la multiplicación con la matriz U_a), señales que tienen energía unidad (debido a la multiplicación con la matriz $S_a^{-1/2}$). Calculando:

$$X_d = [\underline{L}_d \ \underline{R}_d] = [\underline{L}_a \ \underline{R}_a] U_a S_a^{-1/2} U, S_0^{1/2} U_0^H, \quad (12)$$

primero la matriz de covarianza de $[\underline{L}_a \ \underline{R}_a]$ se transforma en una matriz de covarianza que es igual a la matriz identidad, es decir la matriz de covarianza de $[\underline{L}_a \ \underline{R}_a] U_a S_a^{-1/2}$. Aplicar cualquier matriz unitaria arbitraria U_r no cambiará la estructura de covarianza, y aplicar $S_0^{1/2} U_0^H$ da como resultado una estructura de covarianza igual que la del downmix 102 estéreo espacial.

Se definen la matriz S_{0w} y las señales \underline{L}_{0w} y \underline{R}_{0w} según lo siguiente:

$$S_{0w} = [\underline{L}_{0w} \ \underline{R}_{0w}] = [\underline{L}_a \ \underline{R}_a] U_0 S_0^{-1/2} \quad (13)$$

La matriz U_r puede elegirse de manera que la mejor coincidencia de forma de onda posible, en cuanto a distancia euclídea al cuadrado mínima, se obtiene entre las señales \underline{L}_{0w} y \underline{L}_{aw} y las señales \underline{R}_{0w} y \underline{R}_{aw} , donde \underline{L}_{aw} y \underline{R}_{aw} vienen dadas por (11). Con esta elección para U_r , puede utilizarse una coincidencia de forma de onda en el procedimiento estadístico.

A partir de (12) puede verse que la matriz de transformación T viene dada por

$$T = U_a S_a^{-1/2} U, S_0^{1/2} U_0^H. \quad (14)$$

III. Mejor coincidencia de forma de onda bajo una limitación de energía (procedimiento III)

Suponiendo (3) los parámetros α_1 , α_2 , β_1 y β_2 pueden obtenerse minimizando (4) y (5) bajo las limitaciones de energía

$$\sum_k \|L_s[k]\|^2 = \sum_k \|L_d[k]\|^2, \quad \sum_k \|R_s[k]\|^2 = \sum_k \|R_d[k]\|^2. \quad (15)$$

IV. Procedimiento de mezclado (procedimiento IV)

Para mezclar diferentes procedimientos, posibles combinaciones incluyen los procedimientos de mezclado II.a y II.b, o los procedimientos de mezclado II.a y III. Se puede proseguir según lo siguiente:

a) Si la coincidencia de forma de onda entre \underline{L}_s y \underline{L}_d y entre \underline{R}_s y \underline{R}_d que se obtiene cuando se utiliza el procedimiento II.b/III es buena: utilizar el procedimiento II.b/III.

b) Si la coincidencia de forma de onda es mala, utilizar el procedimiento II.a.

c) Garantizar una transición gradual entre los dos procedimientos, mezclando sus matrices de transformación, en función de la calidad de esta coincidencia de forma de onda.

Esto puede expresarse matemáticamente según lo siguiente:

utilizando (3) y (2) la matriz de transformación T puede escribirse en su forma general como

$$T = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix}. \quad (16)$$

Esta matriz se reescribe utilizando dos vectores, \underline{T}_L y \underline{T}_R , según lo siguiente

$$T = [\underline{T}_L \quad \underline{T}_R], \quad \underline{T}_L = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix}, \quad \underline{T}_R = \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix}. \quad (17)$$

La calidad de la coincidencia de forma de onda entre \underline{L}_s y \underline{L}_d obtenida mediante o bien el procedimiento II.b o bien el procedimiento III, se expresa mediante γ_L . Se define como

$$\gamma_L = \max \left(0, \frac{\sum_k L_s[k] L_d^*[k]}{\sum_k |L_s[k]| |L_d[k]|} \right). \quad (18)$$

La calidad de la coincidencia de forma de onda entre \underline{R}_s y \underline{R}_d obtenida utilizando o bien el procedimiento II.b o bien el procedimiento III, se expresa mediante γ_R . Se define como

$$\gamma_R = \max \left(0, \frac{\sum_k R_s[k] R_d^*[k]}{\sum_k |R_s[k]| |R_d[k]|} \right). \quad (19)$$

Tanto γ_L como γ_R están entre 0 y 1. El coeficiente de mezclado del canal izquierdo, δ_L , y el coeficiente de mezclado del canal derecho, δ_R , pueden definirse según lo siguiente:

$$\delta_L = \begin{cases} 1 & \gamma_L > \mu_{L,\max} \\ 0 & \gamma_L < \mu_{L,\min} \\ \frac{1}{2} - \frac{1}{2} \cos \left(\pi \frac{(\gamma_L - \mu_{L,\min})}{(\mu_{L,\max} - \mu_{L,\min})} \right) & \text{en otro caso} \end{cases}, \quad (20)$$

$$\delta_R = \begin{cases} 1 & \gamma_R > \mu_{R,\max} \\ 0 & \gamma_R < \mu_{R,\min} \\ \frac{1}{2} - \frac{1}{2} \cos \left(\pi \frac{(\gamma_R - \mu_{R,\min})}{(\mu_{R,\max} - \mu_{R,\min})} \right) & \text{en otro caso} \end{cases},$$

donde $\mu_{L,\min}$, $\mu_{L,\max}$, $\mu_{R,\min}$ y $\mu_{R,\max}$ son valores entre 0 y 1, $\mu_{L,\min} < \mu_{L,\max}$ y $\mu_{R,\min} < \mu_{R,\max}$. La ecuación (20) garantiza que los coeficientes de mezclado, δ_L y δ_R , estén entre 0 y 1.

Se define la matriz de transformación T del procedimiento II.a, II.b y III, respectivamente, como T_e , que viene dada por (8), T_a , que viene dada por (14), y T_{ce} , respectivamente. Cada matriz de transformación puede dividirse en dos vectores, de manera similar a la división de T en (17), según lo siguiente:

$$T_a = [\underline{T}_{a,L} \quad \underline{T}_{a,R}], \quad T_e = [\underline{T}_{e,L} \quad \underline{T}_{e,R}], \quad T_{ce} = [\underline{T}_{ce,L} \quad \underline{T}_{ce,R}]. \quad (21)$$

La matriz de transformación T para el procedimiento de mezclado II.a y el procedimiento II.b se obtiene como

$$T = [\underline{T}_L \quad \underline{T}_R] = [\delta_L \underline{T}_{a,L} + (1 - \delta_L) \underline{T}_{e,L} \quad \delta_R \underline{T}_{a,R} + (1 - \delta_R) \underline{T}_{e,R}]. \quad (22)$$

La matriz de transformación T para el procedimiento de mezclado II.a y el procedimiento III se obtiene como

$$T = \begin{bmatrix} T_L & T_R \end{bmatrix} = \begin{bmatrix} \delta_L T_{ce,L} + (1-\delta_L) T_{e,L} & \delta_R T_{ce,R} + (1-\delta_R) T_{e,R} \end{bmatrix}. \quad (23)$$

Ahora, considerando dos canales auxiliares correspondientes a dos canales de capa de mejora, la ecuación (1) anterior puede reescribirse como:

$$\begin{bmatrix} L_d & R_d \end{bmatrix} = \begin{bmatrix} L_a & R_a & L_{enh} & R_{enh} \end{bmatrix} T'. \quad (24)$$

donde L_a, R_a (como anteriormente) contienen las muestras de una losa de tiempo/frecuencia del canal izquierdo y derecho del downmix artístico respectivamente, L_d, R_d contienen las muestras de una losa de tiempo/frecuencia del canal izquierdo y derecho del downmix artístico modificado respectivamente y L_{enh}, R_{enh} contienen las muestras de una losa de tiempo/frecuencia de las señales de capa de mejora. La matriz de transformación T' 4x2 escribe por tanto la transformación del downmix artístico y las señales de capa de mejora al downmix artístico modificado. En relación con la ecuación (1), los únicos dos canales auxiliares utilizados son las señales de capa de mejora L_{enh}, R_{enh} .

En el sistema a modo de ejemplo específico, la segunda capa de mejora puede contener dos tipos diferentes de datos:

el primer tipo de datos comprende los parámetros contenidos en la matriz T de la ecuación (1). Estos parámetros son en el ejemplo calculado para todo el ancho de banda de señal y transforman el downmix estéreo artístico de manera que en algún sentido se parece al downmix espacial. Por tanto, este tipo de parámetros puede proporcionar un downmix artístico modificado que se parece más al downmix espacial original pero no permite (necesariamente) que un descodificador genere exactamente el downmix espacial. Para cada losa de tiempo/frecuencia sólo se requieren cuatro parámetros, concretamente se requieren los valores de T (T11, T12, T21 y T22). Estos parámetros pueden codificarse o bien de manera absoluta o bien de manera diferencial y el codificador 10 puede específicamente conmutar dinámicamente entre la codificación absoluta y la diferencial.

El segundo tipo de datos corresponde al downmix espacial real y es en el ejemplo específico una representación de una versión limitada en banda del downmix espacial. Específicamente, este tipo de datos representa una parte de baja frecuencia del downmix espacial (por ejemplo frecuencias por debajo de, digamos, 1,7 kHz). Esto hace posible reconstruir de manera muy precisa esta parte del downmix espacial en el descodificador en lugar de sólo generar una señal que tiene, por ejemplo, las mismas propiedades estadísticas (como con la matriz T). Este tipo de datos puede codificarse de manera absoluta o de manera relativa al downmix artístico. Específicamente, este tipo de datos puede codificarse de manera diferencial. Por ejemplo, la matriz de transformación T se aplica al downmix artístico (véase por ejemplo la ecuación (26)) y la diferencia de esa señal y el downmix espacial puede codificarse.

Por tanto, en algunas realizaciones los segundos datos de mejora se dividen en una primera y una segunda parte de datos de mejora donde la primera parte describe el downmix espacial de manera menos precisa que la segunda parte. Normalmente, la tasa de datos correspondiente de la primera parte de los segundos datos de mejora es inferior a la de la segunda parte. Los datos de mejora de la segunda parte de los segundos datos de mejora pueden referirse a sólo una parte del downmix y específicamente pueden sólo referirse a una parte de baja frecuencia.

En algunas realizaciones, el generador 123 puede estar dispuesto para seleccionar entre datos absolutos y relativos tanto para la primera parte como la segunda parte de los segundos datos de mejora o bien individualmente o bien en conjunto. En otras realizaciones, el generador 123 sólo puede seleccionar entre datos absolutos y relativos para una de las partes de datos. Específicamente, en las siguientes realizaciones se describirá el caso en el que la primera parte de los segundos datos de mejora comprende los parámetros de T mientras que la segunda parte comprende una representación de baja frecuencia del downmix espacial y la selección dinámica entre datos absolutos y relativos sólo se aplica a la segunda parte de los segundos datos de mejora.

Los datos relativos para la segunda parte de los segundos datos de mejora puede en estas realizaciones por ejemplo generarse como valores diferenciales con respecto al downmix artístico después de que se han aplicado los datos de mejora de la primera parte (es decir como valores diferenciales con respecto al downmix artístico modificado).

A continuación, se describen realizaciones en las que el generador 123 selecciona sólo entre datos relativos y absolutos para la segunda parte de los segundos datos de mejora.

En este ejemplo pueden derivarse datos de mejora absolutos para parte de la primera y la segunda parte de los segundos datos de mejora para las losas de tiempo/frecuencia asociadas estableciendo:

$$\begin{aligned} \underline{L}_{enh} &= \underline{L}_s, \\ \underline{R}_{enh} &= \underline{R}_s, \\ T' &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \end{aligned} \quad (25)$$

donde \underline{L}_s , \underline{R}_s contienen las muestras de una losa de tiempo/frecuencia del canal izquierdo y derecho del downmix estéreo espacial respectivamente. Por tanto, en el ejemplo específico, los datos de mejora absolutos simplemente corresponden a las muestras de losa de tiempo/frecuencia reales del downmix 102 espacial que pueden sustituir a las muestras de losa de tiempo/frecuencia correspondientes del downmix 103 artístico.

Además, para la parte de la primera y la segunda parte de los segundos datos de mejora, pueden derivarse específicamente datos de mejora relativos para las losas de tiempo/frecuencia asociadas como datos diferenciales estableciendo:

$$\begin{aligned} \underline{L}_{enh} &= \underline{L}_s - T_{11}\underline{L}_a - T_{21}\underline{R}_a, \\ \underline{R}_{enh} &= \underline{R}_s - T_{12}\underline{L}_a - T_{22}\underline{R}_a, \\ T' &= \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \\ 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned} \quad (26)$$

En este caso, los parámetros T_{11} , T_{12} , T_{21} y T_{22} constituyen la matriz T de la ecuación (2):

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}. \quad (27)$$

De esta manera, el generador 123 puede generar tanto datos de mejora absolutos como datos de mejora relativos para el downmix 103 artístico permitiendo que un descodificador genere un downmix artístico modificado que se parece más al downmix 102 espacial utilizado para generar los datos de mejora multicanal.

El generador 123 está dispuesto además para seleccionar entre los datos de mejora absolutos y los datos de mejora relativos. Esta selección se realiza en el ejemplo específico para bloques de señal individuales (por ejemplo segmentos individuales) y se basa en características de las señales dentro de estos bloques de señal. Específicamente, el generador 123 puede evaluar características de los datos de mejora absolutos y los datos de mejora relativos para un bloque de señal dado y puede decidir qué datos incluir en la capa de mejora para el bloque de señal dado. Además, el generador 123 puede incluir una indicación de qué datos se seleccionaron permitiendo de ese modo que el descodificador aplique los datos de mejora recibidos correctamente.

En algunas realizaciones, el generador 123 puede evaluar la codificación para determinar si los datos de mejora absolutos o los datos de mejora relativos pueden codificarse de la manera más eficaz (por ejemplo con el menor número de bits para una precisión dada). Un enfoque de fuerza bruta puede ser codificar realmente ambos tipos de datos de mejora y comparar el tamaño de los datos codificados. Sin embargo, este puede ser un enfoque complejo en algunas realizaciones, y en el codificador 10 a modo de ejemplo, el generador 123 evalúa la energía de señal de los datos de mejora absolutos con respecto a la energía de señal de los datos de mejora relativos y selecciona qué tipo de datos incluir basándose en una comparación entre los dos.

Específicamente, para codificadores de audio a menudo es beneficioso, en cuando a la tasa de bits, codificar una señal con la menor energía posible. En consecuencia, el generador 123 selecciona el tipo de datos de mejora que tiene la energía de señal más baja. En particular, los datos de mejora relativos se seleccionan cuando

$$\|\underline{L}_s - T_{11}\underline{L}_a - T_{21}\underline{R}_a\|^2 + \|\underline{R}_s - T_{12}\underline{L}_a - T_{22}\underline{R}_a\|^2 < \|\underline{L}_s\|^2 + \|\underline{R}_s\|^2 \quad (28)$$

y de otro modo se seleccionan los datos de mejora absolutos.

Un problema con la conmutación entre diferentes datos de mejora es que pueden producirse algunos artefactos notables. En el codificador 10 a modo de ejemplo, el generador 123 también comprende funcionalidad para conmutar gradualmente entre diferentes datos de mejora. Por tanto, en lugar de conmutar directamente de un tipo de datos de mejora en un bloque de señal a otro tipo en el siguiente bloque de señal, la conmutación se realiza de manera gradual de un conjunto de datos al otro.

Por tanto, durante un intervalo de tiempo (que puede tener una duración de menos o más que un bloque de señal), el generador 123 genera los datos de mejora como una combinación de los datos de mejora absolutos y los datos de mejora relativos. La combinación puede conseguirse por ejemplo mediante una interpolación entre los diferentes tipos de datos o puede utilizar una técnica de solapamiento y adición.

Como un ejemplo específico, en lugar de conmutar de manera brusca entre los diferentes tipos de datos de mejora:

$$\underline{L}_{enh} = \underline{L}_s - T_{11}\underline{L}_a - T_{21}\underline{R}_a, \underline{R}_{enh} = \underline{R}_s - T_{12}\underline{L}_a - T_{22}\underline{R}_a \text{ or } \underline{L}_{enh} = \underline{L}_s, \underline{R}_{enh} = \underline{R}_s$$

los datos de mejora que se transmiten pueden generarse como

$$\underline{L}_{enh} = \underline{L}_s - \alpha T_{11}\underline{L}_a - \alpha T_{21}\underline{R}_a, \underline{R}_{enh} = \underline{R}_s - \alpha T_{12}\underline{L}_a - \alpha T_{22}\underline{R}_a, \quad (29)$$

donde el valor de α para la k-ésima trama de datos puede determinarse como:

$$\alpha_k = \begin{cases} \max(0, \alpha_{k-1} - \delta), & \text{si la trama actual se codifica de manera absoluta,} \\ \min(1, \alpha_{k-1} + \delta), & \text{si la trama actual se codifica de manera diferencial,} \end{cases} \quad (30)$$

donde α_k denota el valor de α en la k-ésima trama y δ es la velocidad de adaptación. Un valor de $\delta = 0,33$ puede proporcionar una codificación libre de artefactos de manera fiable en muchos escenarios. Las señales \underline{L}_{enh} y \underline{R}_{enh} dadas en la ecuación (29) pueden obtenerse utilizando interpolación de parámetros o una técnica de solapamiento y adición y se codifican y añaden al flujo de bits. Además, la decisión con respecto a datos de mejora diferenciales o absolutos se incluye en el flujo de bits, haciendo posible de ese modo que un descodificador derive el mismo valor para α que se utiliza en el codificador.

Se apreciará que aunque la descripción se centra en la utilización de modos diferenciales y absolutos con codificación (intracanal) de cada uno de estos M canales individualmente, otras realizaciones pueden utilizar un enfoque de codificación diferente. Por ejemplo, para M=2, una siguiente etapa puede ser aplicar por ejemplo codificación M/S (codificación media/lateral, codificando por tanto la señal de la suma y la diferencia) cuando se realiza codificación (entre canales) de la señal estéreo. En muchas realizaciones esto puede ser ventajoso tanto en el modo diferencial como el absoluto de codificación (intracanal) de los canales individuales.

Los elementos de la matriz de transformación T' pueden ser de valor real o de valor complejo. Estos elementos pueden codificarse en parámetros de modificación según lo siguiente: aquellos elementos de la matriz de transformación T que son reales y positivos pueden cuantificarse logarítmicamente, como los parámetros IID utilizados en estéreo paramétrico MPEG4. Es posible establecer un límite superior para los valores de los parámetros para evitar la sobre amplificación de señales pequeñas. Este límite superior puede o bien ser fijo o bien una función de la correlación entre el canal izquierdo generado automáticamente y el canal izquierdo artístico y la correlación entre el canal derecho generado automáticamente y el canal derecho artístico. De los elementos de T' que son complejos, la magnitud puede cuantificarse utilizando parámetros IID, y la fase puede cuantificarse linealmente. Los elementos de T' que son reales y posiblemente negativos pueden codificarse tomando el logaritmo del valor absoluto de un elemento, mientras que se garantiza una distinción los valores negativos y positivos.

La figura 6 ilustra un ejemplo del generador 123 de la figura 5 con mayor detalle. En el ejemplo, el generador 123 comprende un procesador 145 de bloque de señal que recibe los downmix 102, 106 espacial y artístico en el

dominio de la frecuencia y divide las señales en bloques de señal. Cada bloque de señal puede corresponder a un intervalo de tiempo de una duración predeterminada. En algunas realizaciones, los bloques de señal pueden como alternativa o adicionalmente dividirse en el dominio de la frecuencia y por ejemplo los subcanales de transformada pueden agruparse juntos en diferentes bloques de señal.

El procesador 145 de bloque de señal está acoplado a un procesador 146 de datos de mejora absolutos que genera los datos de mejora absolutos para los bloques de señal individuales tal como se describió anteriormente. Además, el procesador 145 de bloque de señal está acoplado a un procesador 147 de datos de mejora relativos que genera los datos de mejora relativos para los bloques de señal individuales tal como se describió anteriormente. Los datos de mejora relativos y absolutos se determinan basándose en las características de señal dentro del bloque de señal y específicamente, los datos de mejora para un grupo de losas de tiempo/frecuencia dado pueden determinarse basándose sólo en ese grupo de losas de tiempo/frecuencia.

El procesador 146 de datos de mejora absolutos está acoplado a un primer procesador 148 de energía de señal que determina la energía de señal de los datos de mejora absolutos en cada bloque de señal tal como se describió anteriormente. De manera similar, el procesador 147 de datos de mejora relativos está acoplado a un segundo procesador 149 de energía de señal que determina la energía de señal de los datos de mejora relativos en cada bloque de señal tal como se describió anteriormente.

Los procesadores 148, 149 de energía de señal primero y segundo están acoplados a un procesador 150 de selección que para cada bloque de señal selecciona datos de mejora o bien absolutos o bien relativos dependiendo de qué tipo tiene la energía de señal más baja.

El procesador 150 de selección se proporciona a un procesador 151 de datos de mejora que está acoplado además al procesador 146 de datos de mejora y al procesador 147 de datos de mejora relativos. El procesador 151 de selección recibe una señal de control que indica qué tipo de datos de mejora se ha seleccionado y en consecuencia genera los datos de mejora como los datos de mejora seleccionados. Además, el procesador 151 de selección está dispuesto para realizar una conmutación gradual que incluye una interpolación entre los parámetros absolutos y relativos durante un intervalo de tiempo de conmutación.

El procesador 151 de selección está acoplado a un procesador 152 de codificación que codifica los datos de mejora según un protocolo dado. Además, el procesador 152 de codificación codifica datos que indican qué tipo de datos se selecciona en cada bloque de señal, por ejemplo ajustando un bit para cada bloque de señal para indicar el tipo de datos. Los datos codificados desde el procesador 152 de codificación se incluyen en el flujo de bits codificado generado por el codificador 10.

La figura 7 muestra un diagrama de bloques de otra realización de un descodificador de audio multicanal según algunas realizaciones de la invención que puede ser específicamente el descodificador 20 de audio de la figura 2.

El descodificador 20 comprende una primera unidad 210 y acoplada a la misma una segunda unidad 220. La primera unidad 210 recibe señales de downmix lo y ro y parámetros 105 de modificación como entradas. Las entradas pueden recibirse por ejemplo como un único flujo de bits desde el codificador 10 de la figura 1 ó 5. Las señales de downmix lo y ro pueden ser parte de un downmix 102 espacial o un downmix 103 artístico.

La primera unidad 210 comprende una unidad 211 de segmentación y transformación y una unidad 212 de modificación de downmix. Las señales de downmix lo y ro, respectivamente, se segmentan y las señales segmentadas se transforman al dominio de la frecuencia en la unidad 211 de segmentación y transformación. Las representaciones resultantes en el dominio de la frecuencia de las señales de downmix segmentadas se muestran como señales en el dominio de la frecuencia Lo y Ro, respectivamente. A continuación, las señales en el dominio de la frecuencia Lo y Ro se procesan en la unidad 212 de modificación de downmix. La función de esta unidad 212 de modificación de downmix es modificar el downmix de entrada de tal manera que se parezca al downmix 202 espacial, es decir reconstruir el downmix 202 espacial a partir del downmix 103 artístico y los parámetros 105 de modificación.

Si el downmix 102 espacial se recibe por el descodificador 20 la unidad 212 de modificación de downmix no tiene que modificar las señales de downmix Lo y Ro y estas señales de downmix Lo y Ro pueden simplemente pasarse a la segunda unidad 220 como señales de downmix Ld y Rd del downmix 202 espacial. Una señal 217 de control puede indicar si existe necesidad de modificación del downmix de entrada, es decir si el downmix de entrada es un downmix espacial o un downmix alternativo. La señal 217 de control puede generarse internamente en el descodificador 20, por ejemplo analizando el downmix de entrada y los parámetros asociados 105 que pueden describir propiedades de señal del downmix espacial deseado. Si el downmix de entrada coincide con las propiedades de señal deseadas la señal 217 de control puede ajustarse para indicar que no hay necesidad de modificación. Como alternativa, la señal 217 de control puede ajustarse manualmente o su ajuste puede recibirse como parte de la señal de audio multicanal codificada, por ejemplo en conjunto 105 de parámetros.

Si el codificador 20 recibe el downmix 103 artístico y la señal 217 de control indica que las señales de downmix recibidas Lo y Ro han de modificarse mediante la unidad 212 de modificación de downmix entonces el descodificador puede operar de dos maneras, dependiendo de la representación de los parámetros de modificación recibidos. Si los parámetros representan la transformación relativa del downmix artístico al downmix espacial (es decir si los parámetros

son datos de mejora relativos), las variables de transformación se obtienen directamente aplicando los parámetros de modificación al downmix artístico a la inversa de la operación realizada en el codificador. En diferentes realizaciones, esto puede aplicarse por ejemplo a la segunda parte de los segundos datos de mejora.

Por otro lado, si los parámetros transmitidos representan propiedades absolutas del downmix espacial, el descodificador puede sustituir directamente las muestras de downmix artístico por las muestras de downmix espacial. Por ejemplo, si la segunda parte de los segundos datos de mejora consiste simplemente en las muestras de losa de tiempo/frecuencia del downmix espacial, el descodificador puede sustituir directamente las muestras de losa de tiempo/frecuencia correspondientes del downmix artístico por estas. Se apreciará que también es posible que el descodificador calcule primero las propiedades correspondientes del downmix artístico transmitido realmente. Utilizando esta información (parámetros transmitidos y propiedades calculadas del downmix artístico transmitido), las variables de transformación se determinan entonces de modo que describen la transformada de (propiedades de) el downmix artístico transmitido a (propiedades de) el downmix espacial. Para ser más específicos, la matriz de transformación T puede determinarse utilizando o bien el procedimiento II.a o bien el II.b (ligeramente modificado) que se describieron anteriormente.

El procedimiento II.a puede utilizarse si se transmiten energías absolutas en la primera parte de los segundos datos de mejora. Los parámetros transmitidos (absolutos), E_{L_s} y E_{R_s} , representan la energía de la señal izquierda y derecha del downmix espacial respectivamente y vienen dados por

$$E_{L_s} = \sum_k \|L_s[k]\|^2, \quad E_{R_s} = \sum_k \|R_s[k]\|^2. \quad (31)$$

Las energías del downmix transmitido, E_{DL_s} y E_{DR_s} , se calculan en el descodificador. Utilizando estas variables se pueden calcular los parámetros α y β de (7), según lo siguiente

$$\alpha = \sqrt{\frac{E_{L_s}}{E_{DL_s}}}, \quad \beta = \sqrt{\frac{E_{R_s}}{E_{DR_s}}}. \quad (32)$$

La matriz de transformación T viene dada por

$$T = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}. \quad (33)$$

Específicamente, la unidad 212 de modificación de downmix comprende funcionalidad para extraer el downmix artístico y los parámetros 105 de modificación del flujo de bits recibido. El downmix artístico se divide en bloques de señal (correspondientes a los bloques de señal utilizados por el descodificador). Para cada bloque de señal la unidad 212 de modificación de downmix evalúa la indicación de datos recibida del flujo de bits para determinar si se proporcionan segundos datos de mejora relativos o absolutos para la primera y para la segunda parte para este bloque de señal. La unidad 212 de modificación de downmix aplica entonces la primera y la segunda parte de los segundos datos de mejora como datos de mejora absolutos o datos de mejora relativos en respuesta a los datos de indicación.

Se ha descubierto que puede conseguirse una baja complejidad pero con alto rendimiento cuando los elementos de la matriz de transformación T_{12} y T_{21} se ajustan a cero. A continuación, se describen algunas implementaciones específicas de la unidad 212 de modificación de downmix con esta restricción. Sin embargo, se apreciará que las implementaciones pueden ampliarse fácilmente al caso cuando T_{12} y/o T_{21} son diferentes a cero.

En el caso en el que no se transmiten datos de mejora de la segunda parte de los segundos datos de mejora para la señal de downmix artístico, la primera unidad 210 puede implementarse tal como se muestra en figura 8. Los canales de downmix estéreo en el dominio del tiempo, lo y ro, primero se segmentan y transforman al dominio de la frecuencia mediante una transformación QMF, dando como resultado las señales L_a y R_a , que representan una losa de tiempo/frecuencia del downmix estéreo artístico. A continuación, estas señales se transforman utilizando la matriz de transformación T, dando como resultado las señales $T_{11}L_a$ y $T_{22}R_a$.

Se apreciará que los datos de mejora pueden generarse y aplicarse en el dominio del tiempo y/o la frecuencia. Por tanto, es posible incluir los datos de mejora codificados en el dominio del tiempo (L_{enh} , R_{enh}) en el flujo de bits. Sin embargo, en algunas aplicaciones puede ser ventajoso incluir los datos de mejora codificados en el dominio de la frecuencia en lugar de los datos de mejora en el dominio del tiempo. Por ejemplo, en muchos codificadores los datos de mejora se generan en el dominio de la frecuencia para losas de tiempo/frecuencia y para generar la señal en el dominio del tiempo, se requiere una transformación de dominio de frecuencia a tiempo en el codificador. Además, para aplicar

tales datos de mejora, el descodificador convierte los datos del dominio del tiempo al dominio de la frecuencia. Las conversiones de dominio pueden reducirse por tanto incluyendo los datos de mejora en el dominio de la frecuencia.

En algunas realizaciones, pueden utilizarse diferentes conversiones de tiempo a frecuencia para generar el downmix artístico y los datos de mejora. Por ejemplo, la codificación del downmix artístico puede utilizar una transformada QMF mientras que los datos de mejora utilizan una transformada MDCT. En este caso, los datos de mejora puede incluirse en el dominio de la frecuencia (MDCT) y puede realizarse una transformada directamente entre los dos dominios de frecuencia mediante la unidad 212 de modificación de downmix tal como se ilustra en la figura 9.

En el ejemplo, la matriz de transformación T^* puede ser simplemente la matriz de transformación T de la ecuación (2). Sin embargo, para reducir los artefactos de conmutación T^* puede corresponder a la matriz de transformación T de la ecuación (2) pero modificada para una conmutación gradual. Específicamente, la matriz T^* puede incluir el factor α según se determina mediante la ecuación (30), donde la decisión con respecto a datos de mejora absolutos o relativos se recupera del flujo de bits. Este esquema se utiliza para esas bandas de bloques de señal/frecuencia en las que están presentes los datos de capa de mejora de la segunda parte de los segundos datos de mejora y de otro modo puede utilizarse el enfoque de la figura 8.

Si los datos de mejora (L_{enh} , R_{enh}) se proporcionan en el dominio del tiempo, puede utilizarse un enfoque similar al de la figura 9 tal como se ilustra en la figura 10. Sin embargo, en este caso la transformación de frecuencia a frecuencia transformación se sustituye por una transformación de tiempo a frecuencia que puede ser específicamente mediante una transformada de tiempo a QMF cuando se utilizan transformadas QMF para codificar el downmix artístico. Por tanto, en este ejemplo, los datos de mejora se aplican en el dominio de la frecuencia.

En muchas realizaciones, puede utilizarse una implementación de descodificador para los datos de mejora en el dominio del tiempo que sólo utilice una transformada de dominio de tiempo a frecuencia en la primera unidad 210.

Específicamente, pueden utilizarse los siguientes parámetros de datos de mejora diferenciales:

$$\begin{aligned} \underline{L}_{enh} &= \frac{T_{22}\underline{L}_s - T_{21}\underline{R}_s}{\det(T)} - \underline{L}_a, \\ \underline{R}_{enh} &= \frac{-T_{12}\underline{L}_s + T_{11}\underline{R}_s}{\det(T)} - \underline{R}_a, \\ T' &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \end{aligned} \quad (34)$$

siempre que la matriz T , dada por la ecuación (27), sea no singular (por tanto exista su inversa). Ahora la ecuación (1) puede cambiarse a:

$$[\underline{L}_d \ \underline{R}_d] = [\underline{L}_a \ \underline{R}_a \ \underline{L}_{enh} \ \underline{R}_{enh}] T'^T T. \quad (35)$$

La figura 11 ilustra una implementación eficaz de la unidad 212 de modificación de downmix para los datos de mejora en el dominio del tiempo basándose en las ecuaciones (34) y (35) que se proporcionan. Para mayor claridad, T_{12} y T_{21} de la matriz T se ajustan a cero. En comparación con la implementación de la figura 10, sólo se requiere una transformada de dominio de tiempo a QMF por la implementación de la figura 11.

Por tanto, tal como se describió anteriormente la unidad 212 de modificación de downmix genera una señal 202 que se parece mucho al downmix espacial utilizado para los datos de mejora multicanal. Esto puede utilizarse de manera efectiva por la segunda unidad 220 para ampliar la señal de audio de dos canales a una señal multicanal de sonido envolvente completa. Además, seleccionando de manera dinámica y flexible el tipo de datos de mejora más apropiado (relativos o absolutos) para cada bloque de señal, se consigue una codificación sustancialmente más eficaz y se consigue una codificación/descodificación multicanal con una calidad mejorada para relación de tasa de datos.

La segunda unidad 220 puede ser un descodificador multicanal 2 a 5.1 convencional que descodifica el downmix 202 espacial reconstruido y los datos 104 paramétricos asociados en una señal 203 de salida de canal 5.1. Tal como se describió anteriormente, los datos 104 paramétricos comprenden datos 141, 142, 143 y 144 paramétricos. La segunda unidad 220 realiza el procesamiento inverso de la primera unidad 110 en el codificador 10. La segunda unidad 220

comprende un elemento 221 de realización de upmix, que convierte el downmix 202 estéreo y los parámetros 144 asociados en tres señales de audio mono L, R y C. A continuación, cada una de las señales de audio mono L, R y C, respectivamente, se descorrelaciona en los descorrelacionadores 222, 225 y 228, respectivamente. Después de ello, una matriz 223 de mezclado transforma la señal de audio mono L, su homóloga descorrelacionada y parámetros 141 asociados en señales Lf y Lr. De manera similar, una matriz 226 de mezclado transforma la señal de audio mono R, su homóloga descorrelacionada y parámetros 142 asociados en señales Rf y Rr, y una matriz de 229 de mezclado transforma la señal de audio mono C, su homóloga descorrelacionada y parámetros 143 asociados en señales Co y LFE. Finalmente, se transforman los tres pares de señales en el dominio de la frecuencia segmentadas Lf y Lr, Rf y Rr, Co y LFE, respectivamente, al dominio del tiempo y se combinan mediante solapamiento-adición en transformadores 224, 227 y 230 inversos, respectivamente para obtener tres pares de señales de salida lf y lr, rf y rr, y co y lfe, respectivamente. Las señales de salida lf, lr, rf, rr, co y lfe forman la señal 203 de audio multicanal descodificada.

El codificador 10 de audio multicanal y el descodificador 20 de audio multicanal pueden implementarse por medio de hardware digital o por medio de software que se ejecuta mediante un procesador de señal digital o mediante un microprocesador de propósito general.

Se apreciará que la descripción anterior ha descrito para mayor claridad realizaciones de la invención con referencia a diferentes procesadores y unidades funcionales. Sin embargo, será evidente que puede utilizarse cualquier distribución adecuada de funcionalidad entre diferentes procesadores o unidades funcionales sin restarle valor a la invención. Por ejemplo, la funcionalidad ilustrada que va a realizarse mediante procesadores o controladores separados puede realizarse mediante el mismo procesador o controladores. Por tanto, las referencias a unidades funcionales específicas sólo deben verse como referencias a medios adecuados para proporcionar la funcionalidad descrita en lugar de indicativas de una organización o estructura lógica o física estricta.

La invención puede implementarse de cualquier forma adecuada incluyendo hardware, software, firmware o cualquier combinación de éstos. La invención puede implementarse de manera opcional al menos parcialmente como software informático que se ejecuta en uno o más procesadores de datos y/o procesadores de señal digital. Los elementos y componentes de una realización de la invención pueden implementarse de manera física, funcional y lógica de cualquier manera adecuada. De hecho la funcionalidad puede implementarse en una única unidad, en una pluralidad de unidades o como parte de otras unidades funcionales. Como tal, la invención puede implementarse en una única unidad o puede distribuirse física y funcionalmente entre diferentes unidades y procesadores.

Aunque la presente invención se ha descrito en conexión con algunas realizaciones, no está prevista para limitarse a la forma específica expuesta en el presente documento. En su lugar, el alcance de la presente invención está limitado sólo por las reivindicaciones adjuntas. Adicionalmente, aunque puede aparecer una propiedad descrita en conexión con realizaciones particulares, un experto en la técnica reconocería que diversas propiedades de las realizaciones descritas pueden combinarse según la invención. En las reivindicaciones, la expresión que comprende no excluye la presencia de otros elementos o etapas.

Además, aunque se enumeran de manera individual, pueden implementarse una pluralidad de medios, elementos o etapas de método mediante por ejemplo una única unidad o procesador. Adicionalmente, aunque pueden incluirse propiedades individuales en diferentes reivindicaciones, estas pueden posiblemente combinarse de manera ventajosa, y la inclusión en diferentes reivindicaciones no implica que no sea factible y/o ventajosa una combinación de propiedades. También la inclusión de una propiedad en una categoría de reivindicaciones no implica una limitación para esta categoría sino que indica que la propiedad es igualmente aplicable a otras categorías de reivindicaciones según sea apropiado. Además, el orden de las propiedades en las reivindicaciones no implica ningún orden específico en el que deban trabajarse las propiedades y en particular el orden de etapas individuales en una reivindicación de procedimiento no implica que las etapas deban realizarse en este orden. En su lugar, las etapas pueden realizarse en cualquier orden adecuado. Además, referencias singulares no excluyen una pluralidad. Por tanto, referencias a “uno”, “una”, “primero”, “segundo” etc. no excluyen una pluralidad. En las reivindicaciones se proporcionan símbolos de referencia meramente como un ejemplo clarificador y no deben interpretarse como limitativos del alcance de las reivindicaciones en modo alguno.

REIVINDICACIONES

1. Codificador (10) de audio multicanal para codificar una señal de audio de N canales, comprendiendo el codificador (10) de audio multicanal:
 - medios (110) para generar una primera señal de M canales para la señal de audio de N canales, siendo M menor que N;
 - medios (115, 116, 117, 118) para generar primeros datos de mejora para la primera señal de M canales con respecto a la señal de audio de N canales, permitiendo reconstruir la señal de audio de N canales a partir de la primera señal de M canales;
 - medios (121) para generar una segunda señal de M canales para la señal de audio de N canales;
 - medios (123) de mejora para generar segundos datos de mejora para la segunda señal de M canales con respecto a la primera señal de M canales, los segundos datos de mejora comprenden características de la primera señal de M canales que permiten reconstruir al menos parte de la primera señal de M canales;
 - medios para generar (120) una señal de salida codificada que comprende la segunda señal de M canales, los primeros datos de mejora y los segundos datos de mejora; y
 - en el que los medios (123) de mejora están dispuestos para seleccionar dinámicamente entre generar los segundos datos de mejora como datos de mejora absolutos o como datos de mejora relativos con respecto a la segunda señal de M canales.
2. Codificador (10) de audio multicanal según la reivindicación 1, en el que los medios (123) de mejora están dispuestos para seleccionar entre los datos de mejora absolutos y los datos de mejora relativos en respuesta a una característica de la señal de N canales.
3. Codificador (10) de audio multicanal según la reivindicación 1, en el que los medios (123) de mejora están dispuestos para seleccionar entre los datos de mejora absolutos y los datos de mejora relativos en respuesta a una característica relativa de los datos de mejora absolutos y los datos de mejora relativos.
4. Codificador (10) de audio multicanal según la reivindicación 1, en el que la característica relativa es una energía de señal de los datos de mejora absolutos con respecto a una energía de señal de los datos de mejora relativos.
5. Codificador (10) de audio multicanal según la reivindicación 1, en el que los medios (123) de mejora están dispuestos para dividir la segunda señal de M canales en bloques de señal y para seleccionar individualmente entre los datos de mejora absolutos y los datos de mejora relativos para cada bloque de señal.
6. Codificador (10) de audio multicanal según la reivindicación 5, en el que los medios (123) de mejora están dispuestos para seleccionar entre los datos de mejora absolutos y los datos de mejora relativos para un bloque de señal basándose sólo en características asociadas con el bloque de señal.
7. Codificador (10) de audio multicanal según la reivindicación 1, en el que los medios (123) de mejora están dispuestos para generar los datos de mejora como una combinación de los datos de mejora absolutos y los datos de mejora relativos durante un intervalo de tiempo de conmutación de una conmutación entre generar los datos de mejora como datos de mejora absolutos y como datos de mejora relativos.
8. Codificador (10) de audio multicanal según la reivindicación 7, en el que la combinación comprende una interpolación entre los datos de mejora absolutos y los datos de mejora relativos.
9. Codificador (10) de audio multicanal según la reivindicación 1, en el que los medios para generar (120) la señal de salida codificada están dispuestos para incluir datos que indican si se utilizan datos de mejora relativos o datos de mejora absolutos.
10. Codificador (10) de audio multicanal según la reivindicación 1, en el que los segundos datos de mejora comprenden una primera parte de datos de mejora y una segunda parte de datos de mejora, proporcionando la segunda parte una representación de calidad superior de la primera señal de M canales que la primera parte.
11. Codificador (10) de audio multicanal según la reivindicación 10, en el que los medios (123) de mejora están dispuestos para seleccionar dinámicamente sólo entre generar la segunda parte como datos de mejora absolutos o como datos de mejora relativos.
12. Codificador (10) de audio multicanal según la reivindicación 10, en el que los medios (123) de mejora están dispuestos para generar datos relativos de la segunda parte con respecto a una señal de referencia generada aplicando datos de mejora de la primera parte a la primera señal de M canales.

13. Decodificador (20) de audio multicanal para decodificar una señal de audio de N canales, comprendiendo el decodificador (20) de audio multicanal:

medios (210) para recibir una señal de audio codificada que comprende:

una primera señal de M canales para la señal de audio de N canales, siendo M menor que N,

primeros datos de mejora para la expansión multicanal, siendo los primeros datos de mejora con respecto a una segunda señal de M canales diferente de la primera señal de M canales;

segundos datos de mejora para la primera señal de M canales con respecto a la segunda señal de M canales, comprendiendo los segundos datos de mejora datos de mejora absolutos y datos de mejora relativos con respecto a la primera señal de M canales, y

datos de indicación indicativos de si los segundos datos de mejora para un bloque de señal son datos de mejora absolutos o datos de mejora relativos;

medios (212) de generación para generar una señal de expansión multicanal de M canales en respuesta a la primera señal de M canales y los segundos datos de mejora; y

medios (220) para generar una señal decodificada de N canales en respuesta a la señal de expansión multicanal de M canales y los primeros datos de mejora; y en el que los medios (212) de generación están dispuestos para seleccionar entre aplicar los segundos datos de mejora como datos de mejora absolutos o datos de mejora relativos en respuesta a los datos de indicación.

14. Decodificador (20) de audio multicanal según la reivindicación 13, en el que los medios (212) de generación están dispuestos para aplicar los segundos datos de mejora a la primera señal de M canales en el dominio del tiempo.

15. Decodificador (20) de audio multicanal según la reivindicación 13, en el que los medios (212) de generación están dispuestos para aplicar los segundos datos de mejora a la primera señal de M canales en el dominio de la frecuencia.

16. Decodificador de audio multicanal (10) según la reivindicación 13, en el que los segundos datos de mejora comprenden una primera parte de datos de mejora y una segunda parte de datos de mejora, proporcionando la segunda parte una representación de calidad superior de la primera señal de M canales que la primera parte.

17. Decodificador (20) de audio multicanal según la reivindicación 13, en el que los medios (212) de generación están dispuestos para seleccionar sólo entre aplicar segundos datos de mejora de la segunda parte como datos de mejora absolutos o datos de mejora relativos.

18. Decodificador (20) de audio multicanal según la reivindicación 13, en el que los medios (212) de generación están dispuestos para generar la expansión multicanal de M canales aplicando datos de mejora relativos de la segunda parte a una señal generada aplicando datos de mejora de la primera parte a la primera señal de M canales.

19. Procedimiento de codificación de una señal de audio de N canales, comprendiendo el procedimiento:

generar una primera señal de M canales para la señal de audio de N canales, siendo M menor que N;

generar primeros datos de mejora para la primera señal de M canales con respecto a la señal de audio de N canales, permitiendo reconstruir la señal de audio de N canales a partir de la primera señal de M canales;

generar una segunda señal de M canales para la señal de audio de N canales;

generar segundos datos de mejora para la segunda señal de M canales con respecto a la primera señal de M canales, los segundos datos de mejora comprenden características de la primera señal de M canales que permiten reconstruir al menos parte de la primera señal de M canales;

generar una señal de salida codificada que comprende la segunda señal de M canales, los primeros datos de mejora y los segundos datos de mejora; y

en el que la generación de los segundos datos de mejora comprende seleccionar dinámicamente entre generar los segundos datos de mejora como datos de mejora absolutos o como datos de mejora relativos con respecto a la segunda señal de M canales.

20. Procedimiento de decodificación de una señal de audio de N canales, comprendiendo el procedimiento: recibir una señal de audio codificada que comprende:

una primera señal de M canales para la señal de audio de N canales, siendo M menor que N,

ES 2 313 646 T3

primeros datos de mejora para la expansión multicanal, siendo los primeros datos de mejora con respecto a una segunda señal de M canales diferente de la primera señal de M canales; segundos datos de mejora para la primera señal de M canales con respecto a la segunda señal de M canales, comprendiendo los segundos datos de mejora datos de mejora absolutos y datos de mejora relativos con respecto a la primera señal de M canales, y

datos de indicación indicativos de si los segundos datos de mejora para un bloque de señal son datos de mejora absolutos o datos de mejora relativos;

generar una señal de expansión multicanal de M canales en respuesta a la primera señal de M canales y los segundos datos de mejora; y

generar una señal descodificada de N canales en respuesta a la señal de expansión multicanal de M canales y los primeros datos de mejora; y en el que la generación de la señal de expansión multicanal de M canales comprende seleccionar entre aplicar los segundos datos de mejora como datos de mejora absolutos o datos de mejora relativos en respuesta a los datos de indicación.

21. Señal de audio multicanal codificada para una señal de audio de N canales que comprende:

unos primeros datos de señal de M canales para la señal de audio de N canales, siendo M menor que N;

primeros datos de mejora para la expansión multicanal, siendo los primeros datos de mejora con respecto a una segunda señal de M canales diferente de la primera señal de M canales, permitiendo reconstruir la señal de audio de N canales a partir de la primera señal de M canales;

segundos datos de mejora para la primera señal de M canales con respecto a la segunda señal de M canales, los segundos datos de mejora comprenden características de la primera señal de M canales que permiten reconstruir al menos parte de la primera señal de M canales, comprendiendo los segundos datos de mejora datos de mejora absolutos y datos de mejora relativos con respecto a la primera señal de M canales; y

datos de indicación indicativos de si los segundos datos de mejora para un bloque de señal son datos de mejora absolutos o datos de mejora relativos.

22. Medio de almacenamiento que tiene almacenado en el mismo una señal según la reivindicación 21.

23. Transmisor (40) para transmitir una señal de audio multicanal codificada, comprendiendo el transmisor (40) un codificador (10) de audio multicanal según la reivindicación 1.

24. Receptor (50) para recibir una señal de audio multicanal, comprendiendo el receptor (50) un descodificador (20) de audio multicanal según la reivindicación 13.

25. Sistema (70) de transmisión que comprende un transmisor (40) para transmitir una señal de audio multicanal codificada a través de un canal (30) de transmisión a un receptor (50), comprendiendo el transmisor (40) un codificador (10) de audio multicanal según la reivindicación 1 y comprendiendo el receptor un descodificador (20) de audio multicanal según la reivindicación 13.

26. Procedimiento de transmisión de una señal de audio multicanal codificada, comprendiendo el procedimiento codificar una señal de audio de N canales según la reivindicación 19.

27. Procedimiento de recepción de una señal de audio multicanal codificada, comprendiendo el procedimiento descodificar la señal de audio multicanal codificada, comprendiendo la descodificación:

recibir la señal de audio multicanal codificada según la reivindicación 21;

generar una señal de expansión multicanal de M canales en respuesta a la primera señal de M canales y los segundos datos de mejora; y

generar una señal descodificada de N canales en respuesta a la señal de expansión multicanal de M canales y los primeros datos de mejora; y en el que la generación de la señal de expansión multicanal de M canales comprende seleccionar entre aplicar los segundos datos de mejora como datos de mejora absolutos o datos de mejora relativos en respuesta a los datos de indicación.

28. Procedimiento de transmisión y recepción de una señal de audio, comprendiendo el procedimiento un procedimiento de transmisión de una señal de audio multicanal codificada según la reivindicación 26;

transmitir la señal de salida codificada desde un transmisor hasta un receptor;

recibir, en el receptor, la señal de salida codificada;

ES 2 313 646 T3

descodificar la señal de salida codificada donde la descodificación comprende:

generar una señal de expansión multicanal de M canales en respuesta a la segunda señal de M canales y los segundos datos de mejora, comprendiendo la generación de la señal de expansión multicanal de M canales seleccionar entre aplicar los segundos datos de mejora como datos de mejora absolutos o datos de mejora relativos, y

generar una señal descodificada de N canales en respuesta a la señal de expansión multicanal de M canales y los primeros datos de mejora.

29. Producto de programa informático operativo para hacer que un procesador realice las etapas del procedimiento según una cualquiera de las reivindicaciones 19, 20, 26, 27 y 28.

30. Grabador (6) de audio multicanal que comprende un codificador (10) de audio multicanal según la reivindicación 1.

31. Reproductor (60) de audio multicanal que comprende un descodificador (20) de audio multicanal según la reivindicación 13.

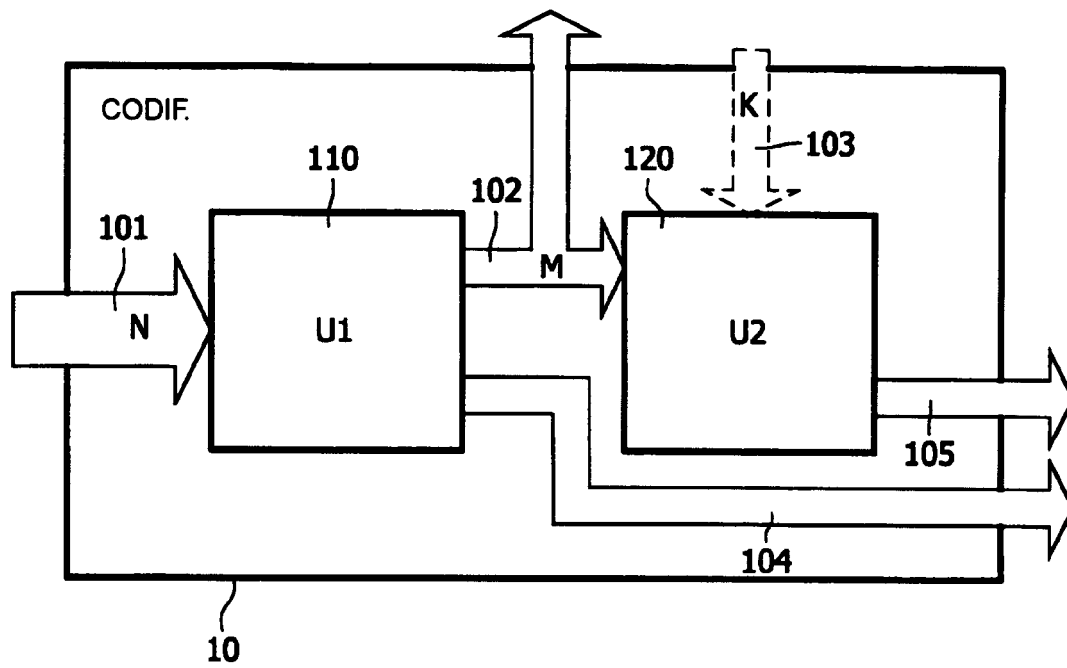


FIG. 1

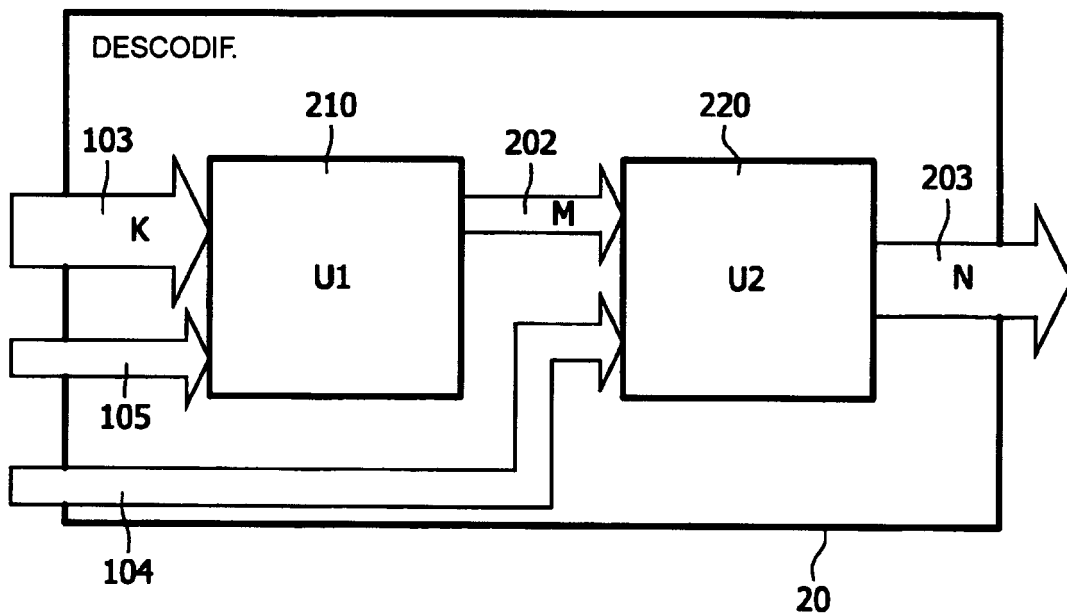


FIG. 2

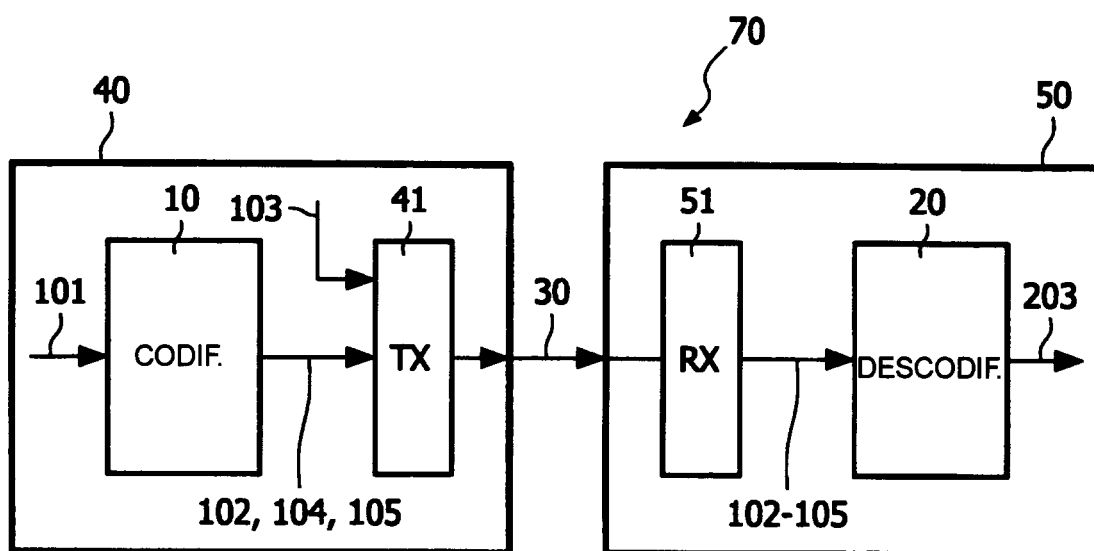


FIG. 3

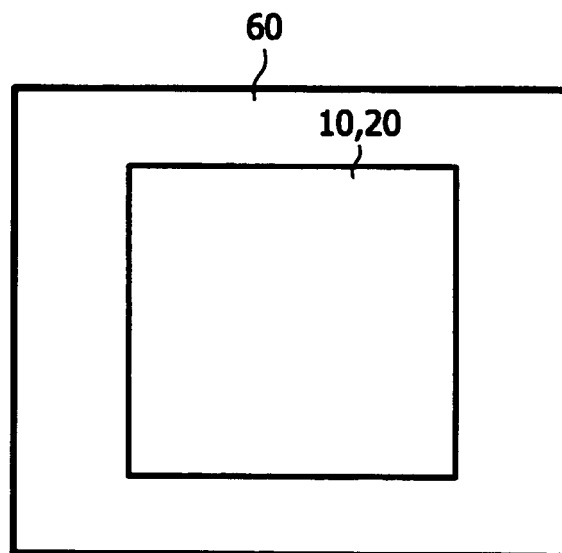


FIG. 4

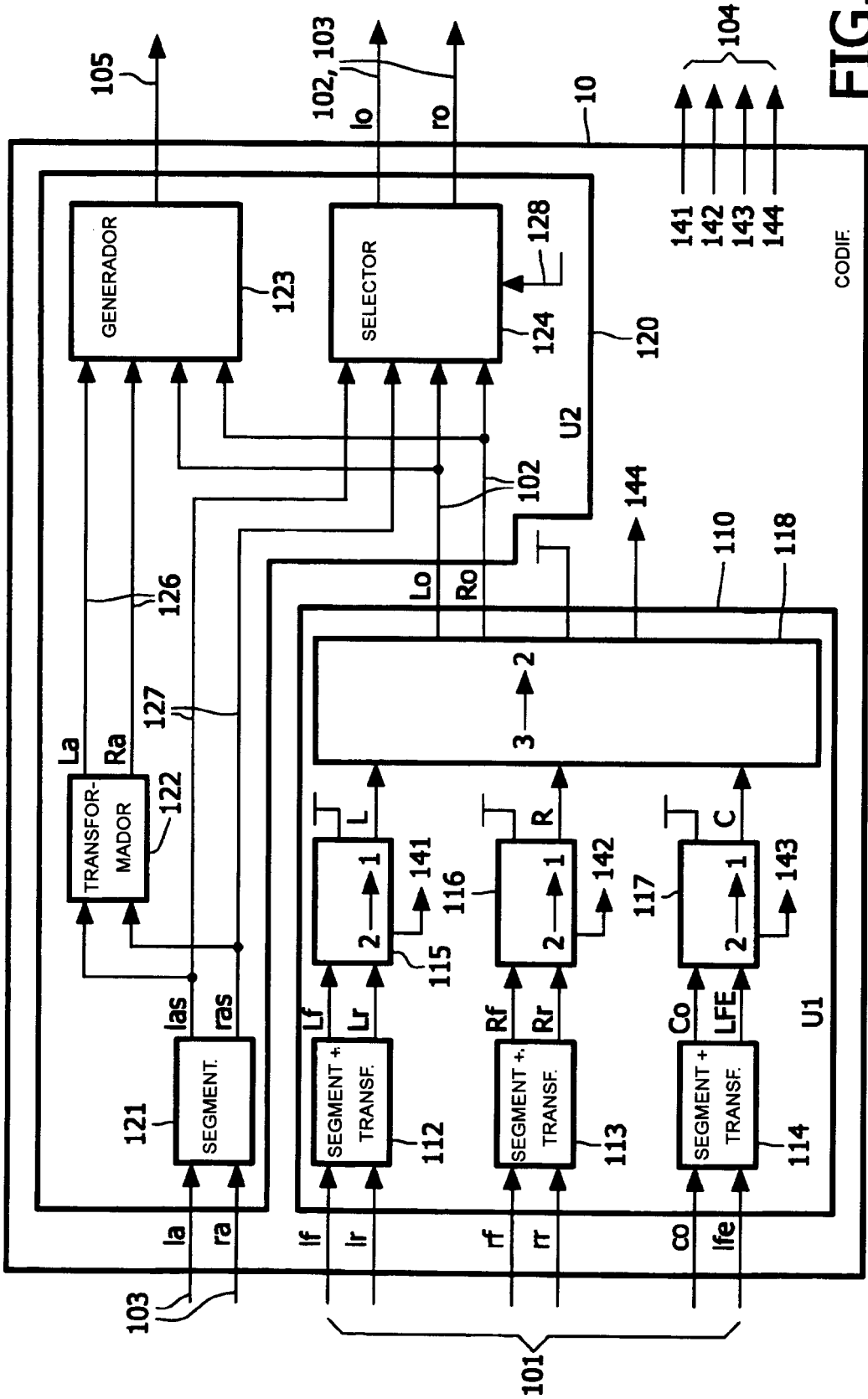


FIG. 5

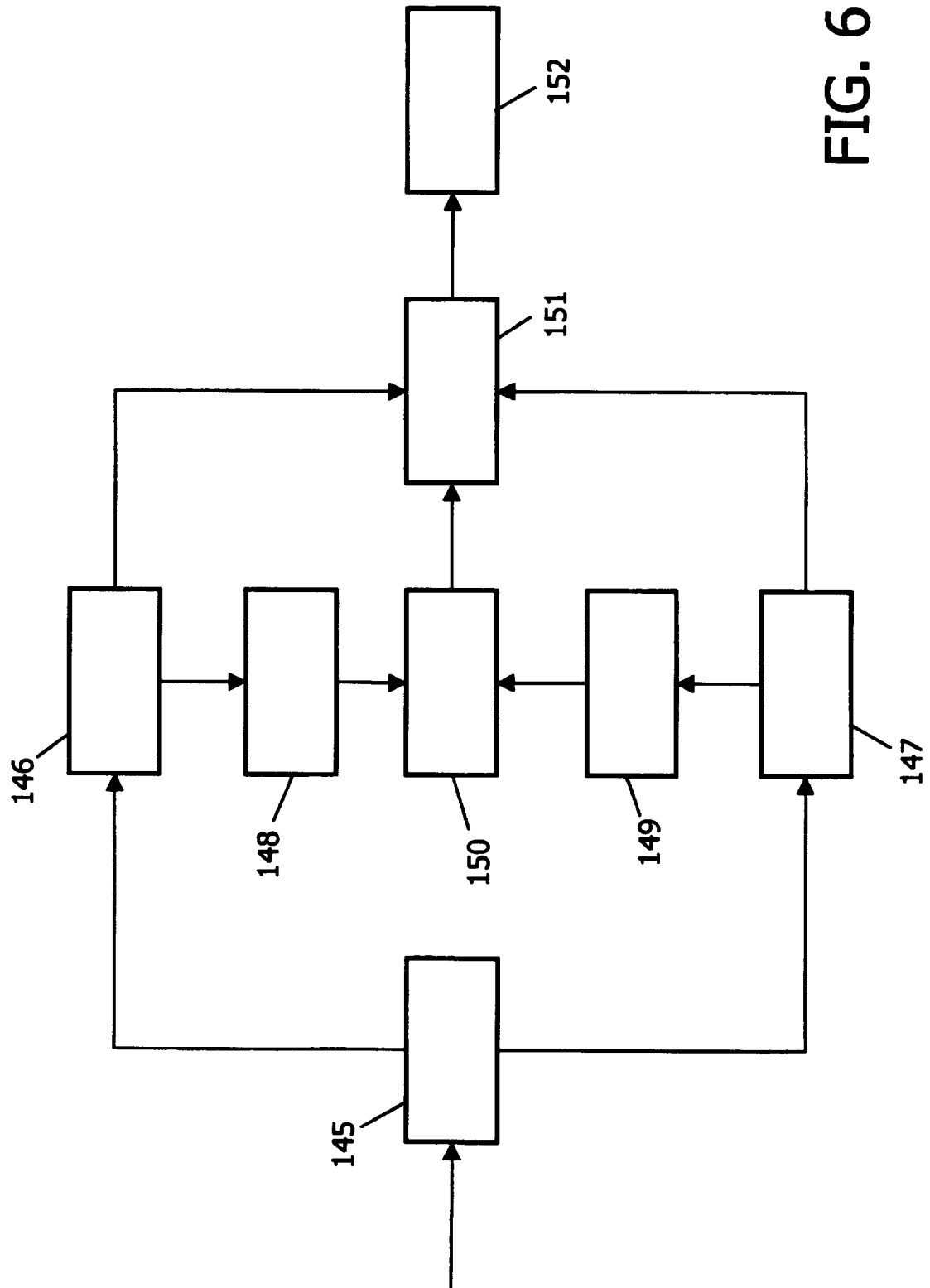


FIG. 6

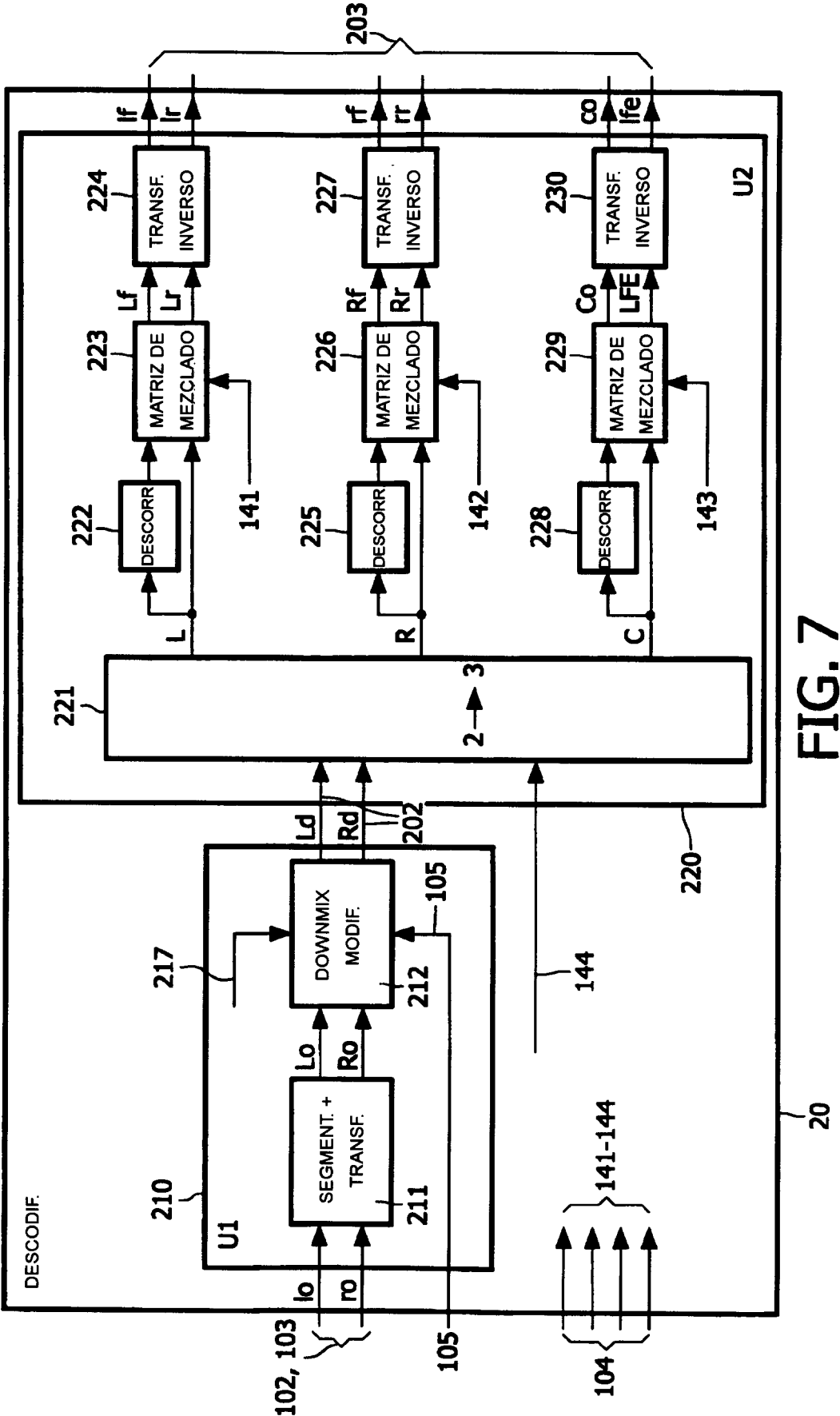


FIG. 7

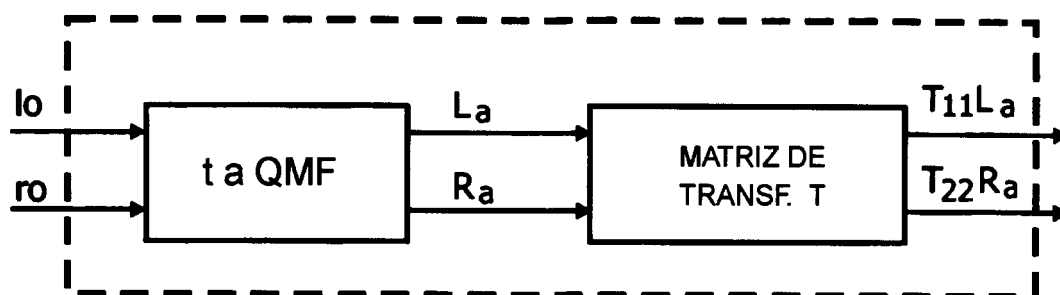


FIG. 8

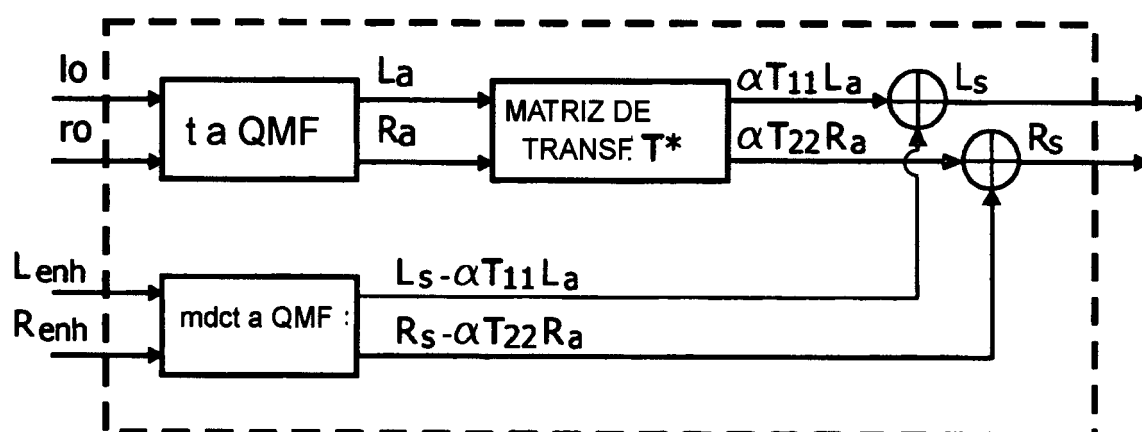


FIG. 9

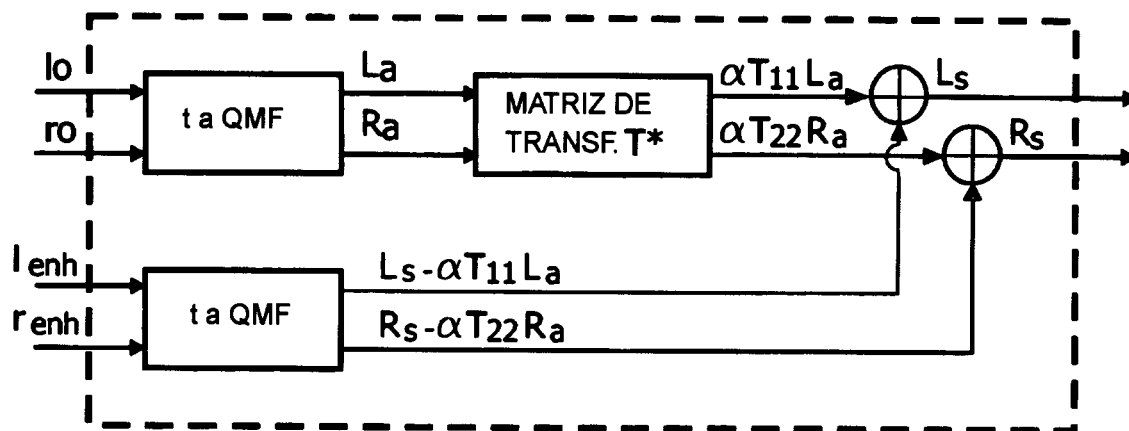


FIG. 10

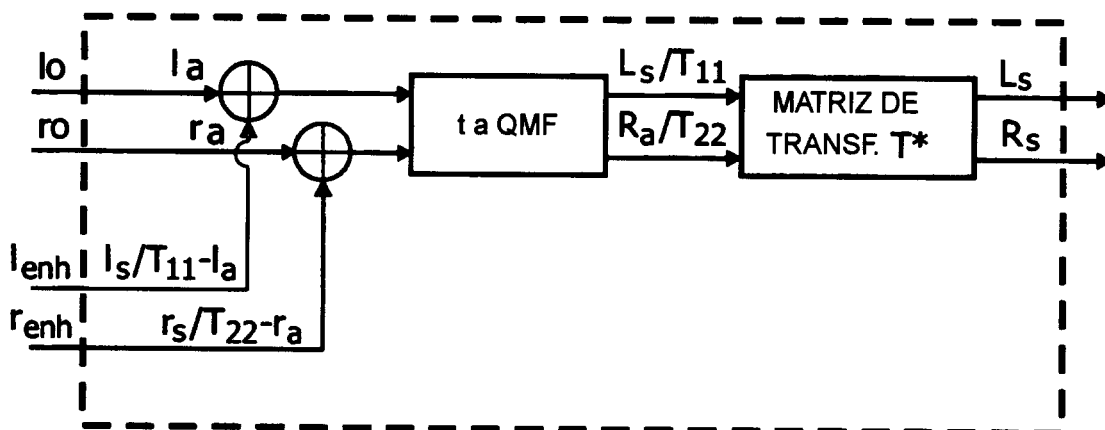


FIG. 11