



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 306 235**

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

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **05798859 .4**

86 Fecha de presentación : **31.10.2005**

87 Número de publicación de la solicitud: **1784819**

87 Fecha de publicación de la solicitud: **16.05.2007**

54 Título: **Codificación de audio multicanal compatible estéreo.**

30 Prioridad: **02.11.2004 SE 0402650**

45 Fecha de publicación de la mención BOPI:
01.11.2008

45 Fecha de la publicación del folleto de la patente:
01.11.2008

73 Titular/es: **Coding Technologies AB.**
Döbelnsgatan 64
113 52 Stockholm, SE
Koninklijke Philips Electronics N.V.

72 Inventor/es: **Purnhagen, Heiko;**
Breebaart, Jeroen;
Schuijers, Erik;
Villemoes, Lars;
Röden, Jonas y
Engdegard, Jonas

74 Agente: **Arizti Acha, Mónica**

ES 2 306 235 T3

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

DESCRIPCIÓN

Codificación de audio multicanal compatible estéreo.

5 La presente invención se refiere a codificación de audio multicanal y en particular a un concepto de generación y utilización de una representación paramétrica de una señal de audio multicanal que es totalmente compatible hacia atrás para entornos de reproducción de estéreo paramétrico.

10 La presente invención se refiere a codificación de representaciones multicanal de señales de audio utilizando parámetros de audio espacial de una manera que es compatible con codificación de señales estéreo de 2 canales utilizando parámetros de estéreo paramétrico. La presente invención enseña nuevos métodos para codificación eficaz tanto de parámetros de audio espacial como de parámetros de estéreo paramétrico y para embeber los parámetros codificados en un flujo de bits de una manera compatible hacia atrás. En particular su objetivo es minimizar la tasa de bits global para los parámetros de estéreo paramétrico y audio espacial en el flujo de bits compatible hacia atrás sin comprometer la calidad del estéreo descodificado o señal de audio multicanal. Cuando una calidad ligeramente comprometida de la señal estéreo descodificada es aceptable, puede reducirse aún más la tasa de bits global.

20 Últimamente, las técnicas de reproducción de audio multicanal están volviéndose cada vez más importantes. Teniendo como objetivo una transmisión eficaz de señales de audio multicanal que presentan 5 o más canales de audio separados, se han desarrollado varias maneras de comprimir una señal multicanal o estéreo. Enfoques recientes para la codificación paramétrica de señales de audio multicanal (estéreo paramétrico (PS, *parametric stereo*), codificación de indicación binaural (BCC, *Binaural Cue Coding*) etc.) representan una señal de audio multicanal por medio de una señal de downmix (podría ser monofónica o comprender varios canales) e información lateral paramétrica, también denominada “indicaciones espaciales”, caracterizando su fase de sonido espacial percibida.

25 Un dispositivo de codificación multicanal recibe en general, como entrada, al menos dos canales, y emite uno o más canales portadores y datos paramétricos. Los datos paramétricos se obtienen de tal manera que, en un descodificador, puede calcularse una aproximación de la señal multicanal original. Normalmente, el(los) canal (canales) portador(es) incluirá(n) muestras de subbanda, coeficientes espectrales, muestras de dominio de tiempo, etc., que proporcionan una representación comparativamente precisa de la señal subyacentes, mientras que los datos paramétricos no incluyen tales muestras de coeficientes espectrales sino que incluyen parámetros de control para controlar un cierto algoritmo de reconstrucción en su lugar. Tal reconstrucción podría comprender ponderación por multiplicación, desplazamiento de tiempo, desplazamiento de frecuencia, desplazamiento de fase, etc. Por tanto, los datos paramétricos, incluyen sólo una representación comparativamente imprecisa de la señal o el canal asociado.

30 La técnica de codificación de indicación binaural (BCC) se describe en varias publicaciones, como en “*Binaural Cue Coding applied to Stereo and Multi-Channel Audio Compression*”, C. Faller, F. Baumgarte, artículo de convención AES 5574, mayo de 2002, Munich, en las 2 publicaciones de ICASSP “*Estimation of auditory spatial cues for binaural cue coding*”, y “*Binaural Cue Coding: a normal an efficient representation of spatial audio*”, ambas escritas por C. Faller y F. Baumgarte, Orlando, FL, mayo de 2002.

35 En la codificación BCC, varios canales de entrada de audio se convierten en una representación espectral utilizando una transformada basada en DFT (transformada discreta de Fourier, *Discrete Fourier Transform*) con ventanas superpuestas. El espectro uniforme resultante se divide entonces en particiones no superpuestas. Cada partición presenta un ancho de banda proporcional al ancho de banda rectangular equivalente (ERB, *equivalent rectangular bandwidth*). Entonces, se estiman parámetros espaciales denominados ICLD (diferencia de nivel entre canales, *Inter-Channel Level Difference*) e ICTD (diferencia de tiempo entre canales, *Inter-Channel Time Difference*) para cada partición. El parámetro ICLD describe una diferencia de nivel entre dos canales y el parámetro ICTD describe la diferencia de tiempo (desfase) entre dos señales de diferentes canales. Las diferencias de nivel y las diferencias de tiempo se dan normalmente para cada canal con respecto a un canal de referencia. Tras la obtención de estos parámetros, los parámetros se cuantifican y finalmente se codifican para su transmisión.

40 Aunque los parámetros ICLD e ICTD representan los parámetros de localización de fuente de sonido más importantes, puede mejorarse una representación espacial utilizando estos parámetros introduciendo parámetros adicionales.

55 Una técnica relacionada, denominada “estéreo paramétrico” describe la codificación paramétrica de una señal estéreo de dos canales basándose en una señal mono transmitida más información lateral de parámetros. Se introducen tres tipos de parámetros espaciales, denominados diferencia de intensidad entre canales (IID, *inter-channel intensity difference*), diferencias de fase entre canales (IPD, *inter-channel phase differences*) y coherencia entre canales (IC, *inter-channel coherence*). La extensión del conjunto de parámetros espaciales con un parámetro de coherencia (parámetro de correlación) permite una parametrización de la “difusividad” espacial o “compacidad” espacial percibidas de la fase de sonido. El estéreo paramétrico se describe con más detalle en: “*Parametric Coding of stereo audio*”, J. Breebaart, S. van de Par, A. Kohlrausch, E. Schuijers (2005) *Eurasip, J. Applied Signal Proc.* 9, páginas 1305 a 1322”, en “*High-Quality Parametric Spatial Audio Coding at Low Bitrates*”, J. Breebaart, S. van de Par, A. Kohlrausch, E. Schuijers, 116ª convención AES, preimpresión 6072, Berlín, mayo de 2004, y en “*Low Complexity Parametric Stereo Coding*”, E. Schuijers, J. Breebaart, H. Purnhagen, J. Engdegard, 116ª convención AES, preimpresión 6073, Berlín, mayo de 2004.

Tal como se mencionó anteriormente, últimamente se han desarrollado sistemas para codificación de estéreo paramétrico así como para codificación de audio espacial. Como en estéreo paramétrico se representa una señal de audio estéreo de dos canales por medio de una señal de audio de downmix mono e información lateral adicional que lleva parámetros de estéreo (véase PCT/SE02/01372 “*Efficient and scalable Parametric Stereo Coding for Low Bitrate Audio Coding Applications*”), un descodificador de estéreo legado reconstruye una señal estéreo de dos canales a partir de la señal mono y la información lateral.

En esquemas de codificación de audio espacial, se representa una señal de audio envolvente multicanal por medio de una señal de audio de downmix estéreo o mono e información lateral adicional que lleva parámetros de audio espacial. Un ejemplo ampliamente conocido es la configuración de canales 5.1 utilizada para sistemas de entretenimiento para el hogar.

Un descodificador de audio espacial legado reconstruye la señal multicanal 5.1 basándose en la señal mono o estéreo y los parámetros espaciales de audio adicionales.

Normalmente las señales de downmix empleadas en sistemas de codificación de audio espacial o estéreo paramétrico se codifican adicionalmente, utilizando técnicas de codificación de audio perceptiva de baja tasa de bits (tal como MPEG AAC) para reducir adicionalmente el ancho de banda de transmisión requerido para la transmisión de los diferentes tipos de señales. Además, la señal de downmix se combina normalmente con la información lateral de audio espacial o de estéreo paramétrico en un flujo de bits de una manera que garantiza la compatibilidad hacia atrás con descodificadores legados, es decir, con descodificadores que no están operativos para procesar los parámetros de audio espacial o estéreo paramétrico. De esta manera, un descodificador de audio legado sólo reconstruye la señal de downmix estéreo o mono transmitida. Cuando se utiliza un descodificador que implementa estéreo paramétrico o codificación de audio espacial, el descodificador también recuperará la información lateral embebida en el flujo de bits y reconstruirá la señal completa envolvente de canales 5.1 o estéreo de dos canales.

Cuando se utiliza codificación de audio espacial basándose en una señal de downmix mono es deseable además aumentar la compatibilidad hacia atrás proporcionando una señal de tal manera que no sólo un descodificador de audio perceptivo legado pueda obtener la señal de downmix mono, sino que adicionalmente sea posible para un descodificador de estéreo paramétrico que no soporta descodificación de audio espacial una descodificación de estéreo paramétrico de tal flujo de bits. Para conseguir este objetivo, es necesario incluir ambas informaciones, la información lateral de estéreo paramétrico y la información lateral de audio espacial en el flujo de bits. Este enfoque obvio lleva a una cantidad indeseablemente alta de información lateral dentro del flujo de bits. Esto significaría para un escenario en el que tiene que mantenerse una tasa de bits máxima total para transmitir la señal mono y la información lateral, que un aumento en la información lateral llevaría a menos tasa de datos disponible para el downmix mono codificado de manera perceptiva, lo que obviamente reduce la calidad de audio de la señal de downmix mono descodificada.

Otro enfoque de la técnica anterior de incluir simultáneamente tanto los parámetros de audio espacial y de estéreo paramétrico como la información lateral, requiere un conjunto de parámetros de audio espacial que están estructurados de tal manera que un subconjunto de estos parámetros permite reconstruir una señal estéreo de dos canales a partir de la señal de downmix mono. Este subconjunto se embebe como información lateral paramétrica dentro del flujo de bits de una manera compatible con flujos de bits de estéreo paramétrico, mientras que los parámetros de audio espacial restantes que no pertenecen al subconjunto se embeben en información lateral de audio espacial en el flujo de bits compatible con codificadores de audio espacial. En el lado del descodificador, un descodificador que implementa sólo estéreo paramétrico reconstruirá una señal estéreo de dos canales basándose en el subconjunto de parámetros que están embebidos como información lateral de estéreo paramétrico. Por otro lado, un descodificador que implementa audio espacial recuperará el subconjunto de estéreo paramétrico y los parámetros de audio espacial restantes. Con este conjunto completo de parámetros espaciales, puede reconstruirse la señal multicanal.

Sin embargo, este enfoque presenta la desventaja de que compromete la calidad de audio de o bien la reconstrucción de estéreo paramétrico compatible hacia atrás o bien la reconstrucción multicanal. Esto es evidente, puesto que en el primer caso, el subconjunto de parámetros que también se utilizan como parámetros de audio espacial describe la interrelación entre dos canales de una señal 5.1. La elección más natural sería el canal izquierdo frontal (l) y el derecho frontal (r), lo que, sin embargo, puede diferir sustancialmente de los valores correctos para la relación de los canales izquierdo (l0) y derecho (r0) de un downmix estéreo. En el segundo caso, los valores correctos de un downmix estéreo forman dicho subconjunto, lo que significa que se utilizan para describir una interrelación entre el canal izquierdo frontal y el derecho frontal de una señal envolvente multicanal. Esto, sin embargo, puede llevar a una imperfección significativa de la reconstrucción de audio espacial debido a la cuantificación de los parámetros, que se requiere con el fin de embeberlos en el flujo de bits de una manera compatible multicanal.

Es el objeto de la presente invención proporcionar un concepto para crear y utilizar una representación paramétrica de una señal de audio multicanal que permita una representación más eficaz que apenas comprometa la calidad de una reconstrucción de estéreo paramétrico o la calidad de una reconstrucción de audio espacial.

Este objeto se consigue mediante un descodificador de audio multicanal según la reivindicación 1 o un codificador de audio según la reivindicación 11.

La presente invención se basa en el descubrimiento de que puede obtenerse eficazmente una representación paramétrica de una señal de audio multicanal que presenta parámetros adecuados para utilizarse junto con una señal de downmix monofónico de una manera compatible hacia atrás cuando se utiliza un combinador de parámetros para generar la representación paramétrica combinando un conjunto de parámetros espaciales y un parámetro de estéreo dando como resultado una representación paramétrica que presenta un parámetro de estéreo utilizable por el descodificador y una información sobre el conjunto de parámetros espaciales que representa, junto con el parámetro de estéreo utilizable por el descodificador, el conjunto de parámetros espaciales.

Utilizando una interrelación entre los parámetros espaciales y los parámetros de estéreo que están describiendo un downmix estéreo de la misma señal multicanal de audio descrita también por los parámetros espaciales, se puede predecir de manera ventajosa un subconjunto de los parámetros espaciales basándose en los parámetros de estéreo paramétrico.

Puesto que la señal estéreo de dos canales descrita por los parámetros de estéreo representa alguna forma de una downmix estéreo de la señal multicanal 5.1, hay dependencias entre los parámetros de estéreo del sistema de estéreo paramétrico y los parámetros espaciales del sistema de codificación de audio espacial, tal como se mencionó anteriormente. La presente invención utiliza estos parámetros de estéreo en combinación con un subconjunto de los parámetros de audio espacial para predecir los valores de los parámetros de audio espacial restantes no incluidos en dicho subconjunto. Entonces, sólo es necesario transmitir la diferencia entre los valores predicho y real de los parámetros de audio espacial que no están en el subconjunto. La entropía de esta diferencia (es decir el error de predicción) es normalmente inferior a la entropía del propio parámetro real. Esto puede utilizarse por un sistema que emplea la presente invención y algún tipo de codificación de entropía posterior. Tal sistema requiere menos tasa de bits de información lateral para los parámetros de estéreo paramétrico y audio espacial que un sistema que simplemente embebería todos los parámetros de manera independiente. Ha de observarse que al mismo tiempo, tal sistema que emplea la presente invención no compromete ni la calidad de la reconstrucción de estéreo paramétrico ni la calidad de la reconstrucción de audio espacial.

Como el objetivo es proporcionar una representación paramétrica que sea compatible hacia atrás con descodificaciones de estéreo paramétrico; se prefiere que los parámetros correctos que representan el downmix estéreo se utilicen para no comprometer la calidad de la señal estéreo de dos canales reconstruida a partir de un descodificador de estéreo paramétrico. Sin embargo, en una realización alternativa de la presente invención, se emplea una pequeña modificación de los parámetros de estéreo paramétrico en el codificador, basándose en los parámetros espaciales estimados, con el fin de mejorar el rendimiento de la predicción de parámetros para los parámetros de audio espacial. Está claro que esta modificación de los parámetros de estéreo paramétrico (PS) lleva a calidad ligeramente reducida de la señal estéreo reconstruida por un descodificador que sólo implementa descodificación de estéreo paramétrico. Mediante esta realización de la presente invención, la calidad de la señal de audio espacial reconstruida permanece no afectada por la modificación de parámetros PS, mientras que se reduce la tasa de bits global requerida para la información lateral espacial y de PS embebida en un flujo de bits compatible.

En una realización preferida de la presente invención, se utiliza un codificador para obtener una representación paramétrica de una señal de audio multicanal que genera un flujo de bits, en la que se embeben parámetros de audio espacial así como parámetros de estéreo paramétrico de un downmix estéreo de la señal multicanal de una manera compatible completamente hacia atrás. Es decir, un descodificador de estéreo paramétrico que puede procesar sólo parámetros de estéreo paramétrico, podrá reconstruir una señal estéreo de alta calidad utilizando los parámetros de estéreo paramétrico. Además, el codificador de la invención sustituye algunos de los parámetros espaciales por una representación diferencial de los parámetros espaciales reales y una predicción del parámetro espacial, mientras que la predicción del parámetro espacial se basa en los parámetros de estéreo y en un conjunto de los parámetros de audio espacial no sustituidos. Puesto que tanto la representación de parámetros de audio espacial como los parámetros de representación de estéreo paramétrico describen diferencias de nivel y correlación entre pares de canales, existe una interrelación entre los parámetros de audio espacial y los parámetros de estéreo, puesto que ambos se obtienen a partir de la misma base de datos, es decir la señal multicanal. Por lo tanto, utilizando la diferencia entre la predicción y el valor real para la transmisión, puede ahorrarse tasa de bits, puesto que las diferencias normalmente presentan una entropía que es mucho menor que la entropía del parámetro de audio espacial subyacente. Cuando la predicción es perfecta, la diferencia de la predicción y el valor real es obviamente cero, lo que significa que como representación de los parámetros espaciales sustituidos sólo tienen que transmitirse o almacenarse valores cero dentro de la representación paramétrica, que es lo más ventajoso cuando se realizan etapas de codificación de entropía adicionales sobre la representación, tal como es normalmente el caso.

Utilizando el concepto descrito anteriormente, un codificación o descodificador de la invención presenta la ventaja obvia de que a pesar de la transmisión compatible hacia atrás de parámetros de audio espacial y estéreo paramétrico sin pérdida de precisión, puede disminuirse la tasa de bits en comparación con un escenario en el que los parámetros de audio espacial y los parámetros estéreo paramétrico se transmiten simplemente de manera independiente dentro de un flujo de bits.

En una realización adicional de la presente invención, se aplica un pequeño cambio a los parámetros de estéreo paramétrico antes de la predicción de los parámetros espaciales y la transmisión de los parámetros espaciales alterados. Esto presenta la gran ventaja de que puede mejorarse la estabilidad de la predicción mediante un pequeño cambio de los parámetros de estéreo paramétrico y, por lo tanto, puede disminuirse adicionalmente la tasa de bits global. El

coste es una pequeña degradación en la calidad de un upmix estéreo reconstruido utilizando los parámetros de estéreo modificados, puesto que los parámetros de estéreo paramétrico realmente óptimos se cambian dentro del proceso de codificación.

5 En una realización adicional de la presente invención, un codificador de audio de la invención comprende un elemento de downmix espacial para generar una señal monofónica a partir de una señal multicanal introducida en el codificador. La señal monofónica se comprime adicionalmente mediante un codificador de audio, utilizando por ejemplo compresión de audio perceptiva, para disminuir adicionalmente la tasa de transmisión que la señal de downmix monofónico utiliza durante la transmisión. Un generador de flujo de bits genera finalmente un flujo de bits para
10 combinar la señal mono, los parámetros de audio espacial y los parámetros de estéreo paramétrico en un único flujo de bits compatible estéreo paramétrico.

En una realización adicional de la presente invención, un codificador o decodificador paramétrico comprende una unidad de control, permitiendo una disminución adicional de la tasa de bits requerida. Esto se consigue comparando
15 la tasa de bits que necesita la representación diferencial de los parámetros espaciales generada utilizando la diferencia del parámetro espacial real y una predicción del mismo con la tasa de bits necesaria para codificar directamente los parámetros espaciales. La codificación se realiza por medio de un procedimiento de codificación de dos etapas, que comprende en primer lugar, la codificación diferencial del tiempo y/o la frecuencia de cada parámetro individualmente, y una codificación de entropía posterior (utilizando por ejemplo un codificador Huffman, un codificador aritmético o
20 un codificador de segmentos en blanco (*run-length*)). Este proceso aprovecha la predictabilidad (o redundancia) para cada parámetro basándose en su propia historia (comparado con la predicción a través de conjuntos de parámetros tal como se describió anteriormente). En los casos en los que la codificación predictiva diferencial da como resultado una tasa de bits superior, puede ahorrarse tasa de bits adicional transmitiendo directamente los parámetros espaciales para tramas de tiempo dadas. La decisión, cuya estrategia se eligió, puede o bien transmitirse dentro del flujo de bits que
25 va a procesarse en el lado del decodificador o bien el decodificador puede decidir sin notificación qué estrategia se había utilizado originariamente aplicando algoritmos de detección apropiados.

Tal como ya se mencionó, una señal generada según la presente invención presenta la gran ventaja de ser compatible hacia atrás para un decodificador de estéreo paramétrico y además albergar la información requerida para la
30 reproducción de una señal (envolvente) espacial completa cuando se transmite a un decodificador de la invención.

Por lo tanto, un decodificador de la invención que recibe los parámetros de estéreo paramétrico y los parámetros de audio espacial puede reconstruir un conjunto completo de parámetros espaciales aplicando la misma predicción y transformación inversa de los parámetros de audio espacial transmitidos de manera diferencial para obtener el conjunto
35 completo de parámetros de audio espacial que representan la propiedad espacial de una señal multicanal a partir de un flujo de bits de la invención.

Dicho de otro modo, la regla de combinación utilizada para combinar los parámetros de estéreo paramétrico y los parámetros de audio espacial recibidos para reconstruir un conjunto completo de parámetros espaciales es la inversa de la regla aplicada en un lado de codificador. En el caso de codificación diferencial tal como se mencionó anteriormente, esto significaría que en primer lugar se calcula la predicción del parámetro deseado utilizando uno o más de los
40 parámetros de estéreo paramétrico y uno o más de los parámetros de audio espacial recibidos. Entonces se calcula la suma entre el valor predicho y el valor transmitido, siendo esta suma el parámetro deseado del conjunto completo de parámetros espaciales.

En una realización adicional de la presente invención, un decodificador de la invención puede reconstruir también una representación estéreo de la señal multicanal utilizando los parámetros de estéreo paramétrico de alta calidad. Esto presenta la gran ventaja de que puede configurarse un decodificador de la invención según las necesidades, es decir, cuando sólo esté disponible un entorno sólo de reproducción estéreo, puede reproducirse una señal estéreo de
50 alta calidad mediante un decodificador de la invención, mientras que, cuando esté a mano un entorno de reproducción multicanal, la representación multicanal de la señal puede reproducirse para permitir la escucha agradable para sonido envolvente.

En una realización adicional de la presente invención, un codificador de la invención está comprendido dentro de un transmisor o grabador de audio, permitiendo el almacenamiento o la transmisión con ahorro de tasa de bits de
55 una señal de audio, que puede reproducirse con excelente calidad o bien como una señal estéreo o bien como señal envolvente completa.

En una realización adicional de la presente invención, un decodificador de la invención está comprendido dentro de un receptor o reproductor de audio, permitiendo recibir o reproducir señales utilizando diferentes configuraciones de altavoces, en la que la señal de audio puede reproducirse en la representación que mejor se ajusta al entorno de
60 reproducción existente.

En resumen, la presente invención comprende las siguientes propiedades ventajosas:

65 - codificación compatible de señales de audio multicanal, que incluye,

en el lado del codificador, someter a downmix la señal multicanal para dar una representación de un canal,

ES 2 306 235 T3

en el lado del codificador dada dicha señal multicanal, definición de los parámetros que representan la señal multicanal,

5 en el lado del codificador dada dicha señal multicanal, definición de los parámetros que representan un downmix estéreo de la señal multicanal,

en el lado del codificador, embeber ambos conjuntos de parámetros de una manera eficaz en tasa de bits y compatible hacia atrás en un flujo de bits,

10 en el lado del decodificador, extraer los parámetros embebidos de un flujo de bits,

en el lado del decodificador, reconstruir los parámetros que representan una señal multicanal a partir de los parámetros extraídos del flujo de bits,

15 en el lado del decodificador, reconstruir las señales de salida multicanal dados los parámetros reconstruidos a partir de los datos del flujo de bits, y dicha señal sometida a downmix;

- embeber los parámetros que representan un downmix estéreo en el flujo de bits, de tal manera que puede descodificarse mediante un método de descodificación (legado) que sólo soporta descodificación de estéreo paramétrico;

20 - dividir el conjunto de parámetros que representan la señal multicanal en un primer subconjunto y un segundo subconjunto;

25 - predecir los valores en dicho primer subconjunto de parámetros basándose en dicho segundo subconjunto de parámetros y basándose en los parámetros que representan un downmix estéreo de la señal multicanal;

- un mecanismo de control que selecciona automáticamente si el primer subconjunto de parámetros se codifica directamente o si sólo se codifican las diferencias con respecto a los valores de parámetros predichos;

30 - modificación de los parámetros que representan un downmix estéreo, en la que tanto los parámetros originales que representan la señal multicanal como los parámetros originales que representan el downmix estéreo se utilizan como base para obtener los parámetros modificados;

35 - una tabla de consulta que se utiliza para encontrar dichos valores de parámetros predichos;

- una función polinómica que se utiliza para encontrar dichos valores de parámetros predichos;

40 - una función matemática obtenida a partir del método empleado para generar el downmix estéreo que se utiliza para encontrar dichos valores de parámetros predichos.

A continuación se describen realizaciones preferidas de la presente invención haciendo referencia a los dibujos adjuntos, en los que

45 la figura 1 muestra un diagrama de bloques de un codificador de la invención;

la figura 2 muestra un flujo de bits generado según la presente invención;

la figura 3 muestra una realización adicional de un codificador de la invención;

50 la figura 4 muestra detalles del codificador de la invención de la figura 3;

la figura 5 muestra un decodificador de la invención;

55 la figura 6 muestra una realización preferida de un decodificador multicanal de la invención;

la figura 7 muestra detalles del decodificador multicanal de la figura 6;

60 la figura 8 ilustra la compatibilidad hacia atrás de una señal de la invención;

la figura 9 muestra un transmisor o grabador de audio que presenta un codificador de la invención;

la figura 10 muestra un receptor o reproductor de audio que presenta un decodificador multicanal de la invención;

y

65 la figura 11 muestra un sistema de transmisión.

Las realizaciones descritas a continuación son meramente ilustrativas para los principios de la presente invención para codificación compatible de estéreo paramétrico mejorada de audio espacial. Se entiende que para otros expertos en la técnica serán evidentes modificaciones y variaciones de las disposiciones y los detalles descritos en el presente documento. Por lo tanto, se pretende quedar limitado sólo por el alcance de las reivindicaciones de patente siguientes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en el presente documento.

La figura 1 muestra un codificador 10 de la invención para obtener una representación 12 paramétrica de una señal de audio multicanal. El codificador 10 comprende un calculador 14 de parámetros espaciales, un calculador 16 de parámetros de estéreo y un combinador 18 de parámetros.

El calculador 14 de parámetros espaciales calcula un conjunto de parámetros 20 espaciales que describen las propiedades espaciales de una señal multicanal. El calculador 16 de parámetros de estéreo calcula parámetros 22 de estéreo que describen propiedades espaciales de un downmix estéreo de la señal multicanal. El conjunto de parámetros 20 espaciales y los parámetros 22 de estéreo se transfieren al combinador 18 de parámetros para obtener la representación 12 paramétrica, que comprende un parámetro 24 de estéreo utilizable por el descodificador y una información sobre el conjunto de parámetros 26 espaciales.

La figura 2 muestra un ejemplo para un flujo de bits compatible hacia atrás que es la representación paramétrica de una señal de audio multicanal según se produce mediante un codificador de la invención según la figura 1. El flujo de bits comprende una sección 30 de parámetros de estéreo y una sección 32 de parámetros espaciales. La sección 30 de parámetros de estéreo presenta una cabecera 34 de estéreo al comienzo de la sección 30 de parámetros de estéreo, seguida por dos parámetros 36a y 36b de estéreo utilizables por el descodificador, que se utilizarían por un descodificador de estéreo paramétrico para reconstruir la señal estéreo. Un descodificador que pueda procesar parámetros de estéreo paramétrico sólo identificaría los parámetros 36a y 36b de estéreo paramétrico por la información comprendida en la cabecera 34 de estéreo.

La sección 32 de audio espacial comienza con una cabecera 38 espacial y comprende cuatro parámetros 40a a 40d de audio espacial. Un descodificador multicanal según la presente invención utilizaría los parámetros 40a a 40d espaciales identificándolos con la ayuda de la cabecera 38 espacial así como los parámetros 36a y 36b de estéreo tal como se identifican por la cabecera 34 de estéreo. Tal como se indica en la figura 2, el parámetro 40a espacial consume menos tasa de bits que los parámetros 40b a 40d espaciales. En el ejemplo mostrado en la figura 2, el parámetro 40a espacial se representa por la diferencia del parámetro espacial original subyacente, y un parámetro espacial predicho obtenido utilizando uno o más de los parámetros 36a o 36b de estéreo y uno o más de los parámetros 40b a 40d de audio espacial. Por lo tanto, un descodificador multicanal de la invención necesitaría utilizar tanto los parámetros 36a y 36b de estéreo como los parámetros 40b a 40d espaciales para reconstruir el parámetro espacial que subyace a la información sobre el parámetro 40a espacial que se transmite en el flujo de bits.

La figura 3 muestra una realización preferida de un codificador 52 de la invención para obtener una representación paramétrica de una señal 50 de audio multicanal, que presenta tres canales, un canal izquierdo l, un canal derecho r y un canal central c.

El codificador 52 de la invención comprende un elemento 54 de downmix espacial, un estimador 56 de parámetros espaciales, un elemento 58 de downmix estéreo, un estimador 60 de parámetros de estéreo paramétrico, un codificador 62 de audio, un combinador 64 de parámetros (bloque de codificación de unión) y un calculador 66 de flujo de bits (multiplexor).

El elemento 54 de downmix espacial, el estimador 56 de parámetros espaciales y el elemento 58 de downmix estéreo reciben como una entrada la señal 50 multicanal. El elemento 54 de downmix espacial crea una señal 68 de downmix monofónico a partir de la señal 50 multicanal, el estimador 56 de parámetros espaciales obtiene parámetros 70 espaciales que describen propiedades espaciales de la señal multicanal, y el elemento 58 de downmix estéreo crea una señal 72 de downmix estéreo a partir de la señal 50 multicanal.

La señal 72 de downmix estéreo se introduce al estimador 60 de parámetros de estéreo paramétrico, que obtiene parámetros 74 de estéreo a partir de la señal de downmix estéreo que describen propiedades espaciales de la señal 72 de downmix estéreo. La señal 68 de downmix monofónico se introduce en el codificador 62 de audio que obtiene un flujo 76 de bits de audio que representa la señal 68 de downmix monofónico por medio de codificación, utilizando por ejemplo técnicas de codificación de audio perceptiva. El combinador 64 de parámetros recibe como una entrada los parámetros 70 espaciales así como los parámetros 74 de estéreo paramétrico y obtiene como una salida parámetros 78 de estéreo utilizables por el descodificador (información lateral de estéreo paramétrico) e información sobre los parámetros 80 espaciales (información lateral espacial) sustituyendo conjuntos de parámetros espaciales por la diferencia de una predicción de los parámetros espaciales y los propios parámetros espaciales. Esto se describirá con más detalle mediante la siguiente figura.

El calculador 66 de flujo de bits recibe finalmente como una entrada el flujo 76 de bits de audio, la información sobre el conjunto de parámetros 80 espaciales y los parámetros 78 de estéreo utilizables por el descodificador y combina dicha entrada en un flujo 82 de bits compatible estéreo paramétrico que podría comprender por ejemplo segmentos de parámetros tal como se detalla en la figura 2.

El calculador 66 de flujo de bits puede ser un simple multiplexor. Sin embargo, también pueden implementarse otros medios para combinar las tres entradas en un flujo de bits compatible para obtener un flujo de bits según la presente invención.

Dicho de otro modo, la figura 3 ilustra un codificador que toma una señal de audio multicanal, que comprende los canales l, r, y c, como entrada y genera un flujo de bits compatible que permite la descodificación mediante un descodificador espacial así como descodificación compatible hacia atrás mediante un descodificador PS. El downmix espacial toma la señal multicanal l, r, c y genera una señal de downmix mono m. Esta señal puede codificarse entonces mediante un codificador de audio perceptivo opcional para producir un flujo de bits de audio compacto que representa la señal mono. La estimación de parámetros espaciales toma la señal multicanal l, r, c como entrada y genera un conjunto de parámetros espaciales cuantificados. Estos parámetros pueden estar en función del tiempo y de la frecuencia. El downmix a estéreo produce un downmix l0, r0 estéreo de 2 canales de la señal multicanal, por ejemplo utilizando las ecuaciones de downmix ITU-R o enfoques alternativos. La estimación de parámetros de estéreo paramétrico (PS) toma este downmix estéreo como entrada y genera un conjunto de parámetros PS cuantificados, que pueden estar en función del tiempo y de la frecuencia. El bloque de codificación de unión toma tanto el parámetro espacial como el parámetro PS como entrada y produce la información lateral de estéreo paramétrico (información lateral PS) y la información lateral espacial. Finalmente, un multiplexor toma el flujo de bits de audio y los flujos de bits de información lateral tanto espacial como PS como entrada y embebe la información lateral de tal manera en el flujo de bits que es posible la descodificación compatible hacia atrás mediante un descodificador legado (que implementa sólo PS).

La figura 4 detalla el combinador 64 de parámetros mostrado en la figura 3. El combinador 64 de parámetros presenta un divisor 90 de parámetros, un modificador 92 de parámetros de estéreo paramétrico, un predictor 94 de parámetros espaciales, un combinador 96, una unidad 98 de control, un ensamblador 100 de parámetros espaciales y un primer codificador 102 diferencial, un segundo codificador 104 diferencial, un tercer codificador 106a diferencial y un cuarto codificador 106b diferencial.

El combinador 64 de parámetros recibe como entrada parámetros 70 espaciales y los parámetros 74 de estéreo paramétrico. Los parámetros 74 de estéreo paramétrico se introducen en el modificador 92 de parámetros de estéreo paramétrico en una primera entrada del mismo, y los parámetros 70 espaciales se introducen en el modificador 92 de parámetros de estéreo paramétrico en una segunda entrada.

Los parámetros 70 espaciales se introducen además en el divisor 90 de parámetros. El modificador 92 de parámetros de estéreo paramétrico es un dispositivo opcional, que puede utilizarse para obtener parámetros 110 de estéreo utilizables por el descodificador modificando los parámetros 74 de estéreo paramétrico utilizando información de los parámetros 70 espaciales.

El divisor 90 de parámetros divide los parámetros 70 espaciales en un primer subconjunto 112 de los parámetros espaciales y en un segundo subconjunto 114 de los parámetros espaciales, donde el primer subconjunto 112 es el subconjunto de los parámetros espaciales que pueden sustituirse por una predicción diferencial dentro la representación paramétrica final de la señal multicanal.

Como la predicción de los parámetros dentro del primer subconjunto se realiza utilizando los parámetros 110 de estéreo utilizables por el descodificador y el segundo subconjunto 114 de los parámetros espaciales, tanto los parámetros 110 utilizables por el descodificador como el segundo subconjunto 114 de parámetros espaciales se introducen en el predictor 94 de parámetros espaciales. El predictor 94 de parámetros espaciales obtiene parámetros 116 predichos utilizando los parámetros 110 de estéreo paramétrico utilizables por el descodificador y el segundo subconjunto 114 de los parámetros espaciales. Los parámetros 116 predichos son una predicción de los parámetros del primer subconjunto 112 y han de compararse con los parámetros del primer subconjunto 112.

Por lo tanto, se calcula la diferencia de los parámetros 116 predichos y el primer subconjunto 112 de parámetros para cada parámetro mediante el combinador 96, que es tal que obtiene parámetros 118 de diferencia. El primer subconjunto 112 de parámetros se introduce en el tercer codificador 106a diferencial que codifica de manera diferencial el primer subconjunto de parámetros aplicando codificación diferencial o bien en tiempo o bien en frecuencia. Los parámetros 118 diferenciales se introducen en el cuarto codificador 106b diferencial.

Según la realización preferida de la presente invención mostrada en la figura 4, la representación codificada de manera diferencial del primer subconjunto 112 se compara con la representación codificada de manera diferencial de los parámetros 118 diferenciales mediante la unidad 98 de control para estimar qué representación requiere más bits dentro de un flujo de bits. La unidad 98 de control controla un conmutador 120, para suministrar esa representación del primer subconjunto 112 al ensamblador 100 de parámetros espaciales que requiere menos bits, mientras que la información cuya representación se utilizó se transfiere adicionalmente desde la unidad 98 de control al ensamblador 100 de parámetros espaciales.

El segundo subconjunto 114 de los parámetros espaciales también se codifica de manera diferencial mediante el segundo codificador 104 diferencial, y la representación codificada de manera diferencial del segundo subconjunto 114 se introduce en el ensamblador 100 de parámetros espaciales, que es tal que tiene la información completa sobre los parámetros 70 espaciales. El ensamblador 100 de parámetros espaciales obtiene finalmente la información sobre los parámetros 80 espaciales reensamblando las representaciones del primer subconjunto 112 y el segundo subconjunto

ES 2 306 235 T3

114 en la información sobre el conjunto de parámetros 80 espaciales que alberga la información completa sobre los parámetros 70 espaciales.

La información final sobre el conjunto de parámetros 80 espaciales, por lo tanto, comprende un segundo subconjunto de parámetros espaciales que no se modifican a pesar de una codificación diferencial del mismo y una representación del primer subconjunto de parámetros espaciales, que puede o bien ser directamente la representación codificada de manera diferencial del primer subconjunto 112 o bien una representación codificada de manera diferencial de parámetros 118 diferenciales, dependiendo de qué representación requiera menos tasas de bits.

Los parámetros 78 de estéreo paramétrico utilizables por el descodificador que se obtienen mediante un combinador 64 de parámetros de la invención, se obtienen mediante el primer codificador 102 diferencial. El primer codificador 102 diferencial recibe como una entrada los parámetros 110 de estéreo paramétrico modificados y obtiene los parámetros 78 de estéreo paramétrico utilizables por el descodificador codificando de manera diferencial los parámetros 110 de estéreo paramétrico modificados.

Dicho de otro modo, la figura 4 ilustra el bloque de codificación de unión que toma tanto el parámetro espacial como el parámetro PS como entrada y genera tanto la información lateral espacial como la información lateral PS. Un bloque de modificación de parámetros PS opcional toma tanto el parámetro espacial como el parámetro PS como entrada y genera un parámetro PS modificado. Esto permite conseguir mejor predicción del parámetro espacial a expensas de comprometer la calidad de la señal estéreo de 2 canales reconstruida a partir del parámetro PS modificado. Si no se emplea el bloque de modificación de parámetros PS, el parámetro PS entrante sirve directamente como entrada al bloque de predicción de parámetros espaciales y a la codificación PS. El conjunto de parámetros PS (modificados) puede codificarse utilizando codificación de diferencial de tiempo (dt) o diferencial de frecuencia (df), es decir, codificación de diferencias de parámetros posteriores en la dirección del tiempo o la frecuencia respectivamente, y codificación Huffman, es decir, codificación de entropía sin pérdidas, con el fin de minimizar el número de bits requeridos para representar el conjunto de parámetros. El bloque de división de parámetros separa el conjunto de parámetros espaciales en un segundo subconjunto que se codifica directamente y un primer subconjunto complementario que contiene todos los parámetros restantes y que puede codificarse utilizando predicción de parámetros. El bloque de predicción de parámetros espaciales toma el segundo subconjunto del parámetro espacial y el parámetro PS (modificado) como entrada y calcula valores predichos para el primer subconjunto del parámetro espacial. Estos valores predichos se restan entonces de los valores reales de los parámetros espaciales en el primer subconjunto, dando como resultado un conjunto de valores de error de predicción.

El segundo subconjunto de parámetros puede codificarse utilizando codificación de diferencial de tiempo o frecuencia y codificación Huffman con el fin de minimizar el número de bits requeridos para representar el subconjunto de parámetros. El primer subconjunto de parámetros puede codificarse utilizando codificación de diferencial de tiempo o frecuencia y codificación Huffman con el fin de minimizar el número de bits requeridos para representar el subconjunto de parámetros. Los valores de error de predicción para el primer subconjunto de parámetros pueden codificarse utilizando codificación de diferencial de tiempo o frecuencia y codificación Huffman con el fin de minimizar el número de bits requeridos para representar el subconjunto de parámetros. Un bloque de control selecciona si el primer subconjunto de parámetros debería codificarse directamente o si el error de predicción debería codificarse con el fin de minimizar el número de bits requeridos para representar el primer subconjunto de parámetros. Esta selección puede realizarse individualmente para cada parámetro en el subconjunto. La decisión de selección real puede o bien transmitirse como información lateral en el flujo de bits o bien puede basarse en reglas que son parte de la predicción de parámetros espaciales. En este último caso, esta decisión no tiene que transmitirse como información lateral. Finalmente, un multiplexor combina todos los datos codificados para formar la información lateral espacial.

Para utilizar el concepto inventivo de codificación o descodificación, son posibles diferentes implementaciones de la predicción de los parámetros. En general, se tiene la posibilidad de utilizar una tabla de consulta diseñada de manera apropiada para obtener una predicción del primer subconjunto de los parámetros espaciales a partir de los parámetros de estéreo y el segundo subconjunto de los parámetros espaciales o se podría aplicar como alternativa una función analítica para obtener los parámetros predichos basándose en el conocimiento de procesos de downmix específicos y la manera en que se obtienen los parámetros espaciales y los parámetros de estéreo. Los siguientes párrafos dan una visión general de algunos ejemplos específicos para conseguir una predicción apropiada.

Esta visión general se basa en una señal multicanal que presenta tres canales,

- l: izquierdo,
- c: central,
- r: derecho,

que ha de considerarse sólo como un ejemplo. Los principios presentados obviamente se aplican de manera correspondiente también a otras configuraciones de canales. Por ejemplo, en caso de una configuración de canales 5.1, el canal izquierdo frontal y el envolvente izquierdo pueden combinarse utilizando un módulo de estéreo paramétrico para formar la señal izquierda (l), el canal derecho frontal y el envolvente derecho pueden combinarse utilizando un módulo de estéreo paramétrico para formar la señal derecha (r), y el canal central y el de mejora de bajas frecuencias pueden combinarse utilizando un módulo de estéreo paramétrico para formar la señal central (c).

ES 2 306 235 T3

La siguiente descripción trata el bloque de predicción de parámetros espaciales con más detalle. Los 2 canales de la señal de downmix estéreo se denotan:

l_0 : downmix izquierdo,
 r_0 : downmix derecho,
y el downmix mono se denota
 m : downmix mono.

El bloque de predicción emite valores predichos $\hat{s}_1, \dots, \hat{s}_K$ de los primeros K parámetros espaciales cuantificados $s_1, \dots, s_K$ (es decir, un primer subconjunto de los parámetros espaciales), dados los parámetros p_1, p_2 PS modificados o no modificados cuantificados y un segundo subconjunto $s_{K+1}, s_{K+2}, \dots, s_N$ de los parámetros espaciales cuantificados restantes.

En el sentido más general, consiste en una función tabulada (tabla de consulta)

$$(\hat{s}_1, \dots, \hat{s}_K) = F(p_1, p_2, s_{K+1}, s_{K+2}, \dots, s_N) \quad (1)$$

La señal de diferencia es entonces igual al error de predicción

$$(d_1, \dots, d_K) = (s_1 - \hat{s}_1, \dots, s_K - \hat{s}_K) \quad (2)$$

Un primer método de diseño es permitir dejar que F sea una función tabulada o un polinomio multivariable elegido para minimizar el error de predicción en el sentido de mínimos cuadrados a lo largo de una gran base de datos de parámetros. Como alternativa, F puede elegirse para minimizar la tasa de bits resultante requerida para representar el primer subconjunto de parámetros espaciales, donde se utiliza una gran base de datos de parámetros como datos de entrenamiento para encontrar la F óptima en este sentido. Antes del uso en la unidad de predicción, una operación de redondeo o cuantificación puede seguir a tal función tabulada o polinomio con el fin de producir resultados enteros.

Un caso especial importante de esto es el uso de una predicción lineal si F es un polinomio de grado uno.

Una segunda clase de diseños de predictor son los que tienen en cuenta la estructura de parámetros reales utilizada. En la realización preferida de la invención, $K=2$ y $N=4$, y los parámetros transmiten información según:

p_1 : iid_l0_r0 diferencia de intensidad entre canales (IID) entre los canales l_0 y r_0 ;
 p_2 : icc_l0_r0 coherencia entre canales o correlación cruzada (ICC) entre los canales l_0 y r_0 ;
 s_1 : iid_l_r diferencia de intensidad entre canales (IID) entre los canales l y r ;
 s_2 : icc_l_r coherencia entre canales o correlación cruzada (ICC) entre los canales l y r ;
 s_3 : iid_lr_c diferencia de intensidad entre canales (IID) entre los canales $l+r$ y c ;
 s_4 : icc_lr_c coherencia entre canales o correlación cruzada (ICC) entre los canales $l+r$ y c .

El primer ejemplo de un diseño de este tipo es un caso especial del diseño de predictor lineal anterior y consiste simplemente en establecer

$$\hat{s}_1 = p_1, \quad \hat{s}_2 = p_2. \quad (3)$$

Este predictor sencillo tiene la ventaja de que da como resultado un error de predicción más estable (en lugar de un error de predicción mínimo) que es bastante adecuado para la codificación de diferencial de tiempo o diferencial de frecuencia de dicho error de predicción. Esto se cumple para todos los predictores como polinomios mencionados anteriormente.

El segundo ejemplo se basa en la suposición de que el downmix estéreo se produce mediante

$$l_0 = l + q \cdot c, \quad r_0 = r + q \cdot c, \quad (4)$$

con una ganancia de canal central conocida q , (normalmente 1 ó $1/\sqrt{2}$). Todas las señales l , r , c normalmente son vectores de longitud finita que resultan de un intervalo de tiempo y frecuencia de muestras de subbanda de un análisis de banco de filtros modulado complejo de señales de tiempo. Para vectores complejos x , y , el producto interior complejo y la norma cuadrada se definen mediante

$$\left\{ \begin{array}{l} \langle x, y \rangle = \sum_n x(n) y^*(n), \\ X = \|x\|^2 = \langle x, x \rangle = \sum_n |x(n)|^2, \\ Y = \|y\|^2 = \langle y, y \rangle = \sum_n |y(n)|^2, \end{array} \right. \quad (5)$$

donde el asterisco denota conjugación compleja. Se asume entonces que las versiones lineales y no cuantificadas de los parámetros IID se obtienen mediante

$$P_1 = \sqrt{\frac{L_0}{R_0}}, \quad S_1 = \sqrt{\frac{L}{R}}, \quad S_3 = \sqrt{\frac{L+R}{C}}. \quad (6)$$

Para los parámetros ICC, en el caso de correlación cruzada, las fórmulas son

$$P_2 = \frac{\text{Re}\langle l_0, r_0 \rangle}{\sqrt{L_0 \cdot R_0}}, \quad S_2 = \frac{\text{Re}\langle l, r \rangle}{\sqrt{L \cdot R}}, \quad S_4 = \frac{\text{Re}\langle l+r, c \rangle}{\|l+r\| \cdot \|c\|}. \quad (7)$$

En el caso de coherencia, las operaciones de valor real se sustituyen con operaciones de valor absoluto (magnitud compleja) en las fórmulas (7).

Suponiendo para mayor simplicidad que $\langle l, c \rangle = \langle r, c \rangle = 0$, se deduce que $L_0 = L + q^2 C$ y $R_0 = R + q^2 C$, lo que puede insertarse en la primera fórmula de (6). Resolviendo dos ecuaciones con dos incógnitas, se obtienen entonces las siguientes estimaciones de $X = L/C$ e $Y = R/C$ a partir de P_1 y S_3 ,

$$\hat{X} = \frac{P_1^2 S_3^2 + q^2 (P_1^2 - 1)}{P_1^2 + 1}, \quad \hat{Y} = \frac{S_3^2 - q^2 (P_1^2 - 1)}{P_1^2 + 1} \quad (8)$$

Cuando ambos valores en la fórmula (8) son positivos, la estimación de S_1 se forma como $\hat{S}_1 = \sqrt{\hat{X}/\hat{Y}}$. En este caso, los valores de parámetros lineales requeridos se obtienen descuantificando los parámetros enteros dados y se obtiene entonces la estimación $\hat{s}_1$ mediante la cuantificación de $\hat{S}_1$.

Cuando es aceptable una calidad ligeramente comprometida de la señal estéreo descodificada, puede reducirse adicionalmente la tasa de bits global empleando modificación de los parámetros de estéreo paramétrico. El objetivo de esta modificación es conseguir una predicción más estable del primer subconjunto de parámetros espaciales y error de predicción reducido. Puede considerarse como un medio para estabilizar los cálculos anteriores. El caso más extremo de una modificación de parámetros de este tipo sería utilizar $p_1' = s_1$, $p_2' = s_2$ donde p_1' , p_2' denotan los parámetros de estéreo paramétrico modificados. Puesto que esta operación de modificación de parámetros se lleva a cabo sólo en el lado del codificador, no es necesario tener cuidado especial en el lado del descodificador.

Un enfoque más general incorpora la potencia completa y la información de estructura de correlación disponible en P_1, P_2, S_3, S_4 a través de las fórmulas (6) y (7) para obtener estimaciones de S_1 y S_2 . Por la invarianza de escala de los parámetros, no hay pérdida de generalidad al suponer para fines de cálculo que $C=1$. Entonces con las definiciones

$$a = \text{Re}\langle l, c \rangle, \quad b = \text{Re}\langle r, c \rangle, \quad \rho = \text{Re}\langle l, r \rangle, \quad (9)$$

surge el siguiente sistema de ecuaciones:

$$\left\{ \begin{array}{l} L + q^2 + 2qa = P_1^2(R + q^2 + 2qb) \\ \rho + q^2 + 2q(a+b) = P_2(L + q^2 + 2qa)^{1/2}(R + q^2 + 2qb)^{1/2} \\ L + R = S_3^2 \\ a + b = S_4(L + R + 2\rho)^{1/2}. \end{array} \right. \quad (10)$$

Las incógnitas de interés para la estimación de L, R, p y a, b son incógnitas adicionales. Este sistema de ecuaciones (indeterminado) puede utilizarse como guía para una multitud de fórmulas de predicción, dependiendo de la selección de restricciones sobre el par a, b . Por ejemplo, la primera y la tercera ecuación de (10) implican

$$\left\{ \begin{array}{l} (1 + P_1^2)L = q^2(P_1^2 - 1) + 2q(P_1^2b - a) + P_1^2S_3^2 \\ (1 + P_1^2)R = S_3^2 - q^2(P_1^2 - 1) - 2q(P_1^2b - a) \end{array} \right. \quad (11)$$

por lo que los cálculos que llevan a las fórmulas (8) corresponden al caso en el que $P_1^2b=a$. De manera más general, un parámetros heurístico γ defines una restricción sobre el par a, b a través de $\gamma=P_1^2b-a$.

Se resalta de nuevo que los esquemas de predicción anteriores son sólo ejemplos para posibles esquemas de predicción que pueden implementarse tanto en un lado de codificador como en un lado de decodificador.

La figura 5 muestra un decodificador 200 de audio multicanal de la invención para procesar una representación 202 paramétrica.

La representación 202 paramétrica comprende información sobre un conjunto de parámetros 204 espaciales que describen las propiedades espaciales de una señal multicanal y parámetros 206 de estéreo utilizables por el decodificador que describen propiedades espaciales de un downmix estéreo de la señal multicanal. El decodificador 200 de audio multicanal de la invención presenta un reconstructor 208 de parámetros para combinar los parámetros 206 de estéreo utilizables por el decodificador y la información sobre el conjunto de parámetros espaciales para obtener parámetros 210 espaciales.

La figura 6 muestra una realización de un decodificador 220 de audio multicanal según la presente invención. El decodificador 220 de audio multicanal presenta un elemento 222 de descomposición (demultiplexor) de flujo de bits, un decodificador 224 de audio, un reconstructor 226 de parámetros (decodificador de unión) y un elemento 228 de upmix.

El elemento 222 de descomposición de flujo de bits recibe un flujo 230 de bits compatible hacia atrás que comprende un flujo 231 de bits de audio, información sobre un conjunto de parámetros 232 espaciales y parámetros 234 de estéreo utilizables por el decodificador (información lateral PS). El elemento de descomposición de flujo de bits descompone o demultiplexa el flujo 230 de bits compatible hacia atrás para obtener el flujo 231 de bits de audio, la información sobre el conjunto de parámetros 232 espaciales y los parámetros 234 de estéreo utilizables por el decodificador. El decodificador 224 de audio recibe el flujo 231 de bits de audio como entrada y obtiene una señal 236 de downmix monofónico a partir del flujo 231 de bits de audio.

El reconstructor 226 de parámetros recibe la información sobre el conjunto de parámetros 232 espaciales y los parámetros 234 de estéreo utilizables por el decodificador como una entrada. El reconstructor 226 de parámetros combina la información sobre el conjunto de parámetros espaciales y los parámetros de estéreo utilizables por el decodificador para obtener un conjunto de parámetros 238 espaciales que sirven como una entrada al elemento 228 de upmix, que recibe además la señal 236 de downmix monofónico como segunda entrada. Basándose en los parámetros 238 espaciales y en la señal 236 de downmix monofónico, el elemento 228 de upmix obtiene una reconstrucción de una señal 240 multicanal en su salida.

ES 2 306 235 T3

Por lo tanto, la figura 6 ilustra un descodificador de audio espacial que toma un flujo de bits compatible como entrada y genera la señal de audio multicanal, que comprende los canales l, r, y c. En primer lugar, un demultiplexor toma el flujo de bits compatible como entrada y lo descompone en un flujo de bits de audio y la información lateral tanto espacial como PS. Si se aplicó codificación de audio perceptiva a la señal mono, un descodificador de audio correspondiente toma el flujo de bits de audio como entrada y genera la señal de audio mono m descodificada, sometida a distorsión ya que se introduce mediante el códec de audio perceptivo. El bloque de descodificación de unión toma la información lateral tanto espacial como PS como entrada y reconstruye los parámetros espaciales. Finalmente, la reconstrucción espacial toma la señal mono m descodificada y los parámetros espaciales como entrada y reconstruye la señal de audio multicanal.

La figura 7 da una descripción detallada del reconstructor 226 de parámetros utilizado por el descodificador 220 de audio multicanal. El reconstructor 226 de parámetros comprende un desensamblador 250 de parámetros espaciales, una unidad 252 de control, un predictor 254 de parámetros espaciales, un ensamblador 256 de parámetros espaciales y un primer descodificador 258 diferencial, a segundo descodificador 260 diferencial, un tercer codificador 262a diferencial, y un cuarto descodificador 262b diferencial.

El desensamblador 250 de parámetros espaciales recibe la información sobre el conjunto de parámetros 232 espaciales como una entrada y obtiene un primer subconjunto 266 y un segundo subconjunto 268 a partir de la información sobre el conjunto de parámetros 232 de audio espacial. El primer subconjunto 266 comprende los parámetros que están representados posiblemente por una representación diferencial predictiva realizada sobre el lado del codificador, y el segundo subconjunto 268 comprende un subconjunto de la información sobre el conjunto de parámetros espaciales que se transmiten no modificados dentro del flujo de bits.

Además, la unidad 252 de control recibe opcionalmente información de control desde el desensamblador de parámetros espaciales, indicando si se ha utilizado una representación diferencial predictiva durante la codificación o no. Esta información es opcional en el sentido de que la unidad 252 de control podría obtener como alternativa, utilizando algoritmos apropiados, si tal predicción se ha realizado o no sin tener acceso a un parámetro de indicación.

El segundo subconjunto 268 de parámetros se introduce en el segundo descodificador 260 diferencial, que descodifica de manera diferencial el segundo subconjunto para obtener un segundo subconjunto de parámetros 270 espaciales.

El primer descodificador 258 diferencial recibe como una entrada los parámetros 234 de estéreo utilizables por el descodificador, para obtener parámetros 272 de estéreo paramétrico a partir de la representación codificada. El predictor 254 de parámetros espaciales opera de la misma manera que su parte complementaria en el lado del codificador, por lo tanto recibe como una primera entrada los parámetros 272 de estéreo paramétrico y como una segunda entrada el segundo subconjunto de parámetros 270 espaciales para obtener parámetros 274 predichos.

La unidad 252 de control controla dos posibles trayectorias de datos diferentes para el primer subconjunto de la información sobre el conjunto de parámetros espaciales. Cuando la unidad 252 de control indica que el primer subconjunto de la información del conjunto de parámetros espaciales se ha transmitido utilizando codificación diferencial predictiva, la unidad 252 de control conduce los conmutadores 278a y 278b de tal manera que el primer subconjunto 266 se introduce en el tercer descodificador 262a diferencial para obtener un primer subconjunto del conjunto de parámetros 280 espaciales sin aplicar predicción inversa. El primer subconjunto de parámetros 280 espaciales se introduce entonces en el ensamblador 256 de parámetros espaciales en una segunda entrada del mismo.

Sin embargo, si la unidad 252 de control indica parámetros predichos de manera diferencial, el primer subconjunto 266 de la información del conjunto de parámetros espaciales se introduce en el cuarto descodificador 262b diferencial para obtener una representación predicha de manera diferencial del primer subconjunto 266 en una salida 282 del descodificador diferencial. Entonces se calcula la suma de la representación diferencial y los parámetros 274 predichos mediante un sumador 284, invirtiendo por tanto la operación de predicción diferencial realizada sobre un lado de codificador. Como resultado, el primer conjunto de parámetros 280 espaciales está disponible en la segunda entrada del ensamblador 256 de parámetros espaciales. El ensamblador 256 de parámetros espaciales combina el primer conjunto de parámetros 280 espaciales y el segundo conjunto de parámetros 270 espaciales para proporcionar un conjunto completo de parámetros 290 espaciales en su salida, que es la base de una reconstrucción multicanal de una señal codificada.

En resumen, la figura 7 ilustra el bloque de descodificación de unión que toma tanto la información lateral espacial como la información lateral PS como entrada y reconstruye el parámetro espacial. Un demultiplexor separa la información lateral espacial en un segundo subconjunto codificado de parámetro espacial y un primer subconjunto codificado de parámetro espacial e información de control. El bloque de descodificación toma el segundo subconjunto codificado de parámetro espacial como entrada y reconstruye este subconjunto de parámetros. Esto incluye descodificación Huffman y descodificación de diferencial de tiempo (dt) o diferencial de frecuencia en caso de que se emplease tal codificación en el codificador. El bloque de descodificación toma la información lateral PS como entrada y reconstruye el parámetro PS (modificado). El bloque de predicción de parámetros espaciales toma el segundo subconjunto del parámetro espacial y el parámetro PS (modificado) como entrada y calcula valores predichos para el primer subconjunto del parámetro espacial de la misma manera que lo hace su parte complementaria en el codificador. El bloque de control determina qué decisión de selección tomó su parte complementaria, el bloque de control en el codificador. Dependiendo de esta selección, el primer subconjunto codificado de parámetro espacial o bien se descodifica directa-

mente o bien se descodifica teniendo en cuenta la predicción. En ambos casos, esto incluye descodificación Huffman y descodificación de diferencial de tiempo o frecuencia en caso de que se emplease tal codificación en el codificador. En caso de que el bloque de control determinase que no se utilizó ninguna predicción, la salida del bloque de descodificación se toma como el primer subconjunto reconstruido de parámetro espacial. Si no, la salida del bloque de descodificación contiene los valores de error de predicción que se añaden entonces a los valores de parámetros predichos según se generan mediante la predicción de parámetros espaciales con el fin de obtener los valores originales del primer subconjunto de parámetros espaciales. Finalmente, el primer y segundo subconjuntos reconstruidos de parámetros espaciales se funden para formar el conjunto completo de parámetros espaciales.

La figura 8 ilustra cómo se procesa un flujo de bits de la invención compatible mediante un descodificador de estéreo paramétrico legado para obtener un upmix estéreo de una señal para resaltar la gran ventaja de la completa compatibilidad hacia atrás del concepto inventivo.

Un descodificador 300 de estéreo paramétrico recibe un flujo 302 de bits compatible como entrada. El descodificador 300 de estéreo paramétrico comprende un demultiplexor 304, un descodificador 306 de audio, un descodificador 308 diferencial y un elemento 310 de upmix. El demultiplexor 304 obtiene un flujo 312 de bits de audio y parámetros 314 de estéreo paramétrico utilizables por el descodificador a partir del flujo 302 de bits compatible.

Como el descodificador 300 de estéreo paramétrico no puede operar sobre parámetros de audio espacial, el demultiplexor 304 simplemente ignora los parámetros de audio espacial comprendidos dentro del flujo 302 de bits compatible, por ejemplo saltando campos de cabecera y secciones de datos asociadas dentro del flujo de bits no conocidas para el descodificador. El flujo 312 de bits de audio se introduce en el descodificador 306 de audio que obtiene una señal 316 de downmix monofónico mientras que los parámetros 314 de estéreo utilizables por el descodificador se descodifican de manera diferencial mediante el descodificador 308 diferencial para obtener parámetros 318 de estéreo paramétrico. La señal 316 de downmix monofónico y los parámetros 318 de estéreo paramétrico se introducen en el elemento 310 de upmix, que obtiene una señal 320 de upmix estéreo utilizando la señal 316 de downmix monofónico y los parámetros 318 de estéreo paramétrico.

Dicho de otro modo, la figura 8 ilustra un descodificador de estéreo paramétrico (PS) que toma un flujo de bits compatible como entrada y genera una señal de audio estéreo de 2 canales, que comprende los canales l0 y r0. En primer lugar, un demultiplexor toma el flujo de bits compatible como entrada y lo descompone en un flujo de bits de audio y la información lateral PS. Puesto que la información lateral espacial estaba embebida en el flujo de bits compatible de una manera compatible hacia atrás, esto no afecta al demultiplexor. Si se aplicó codificación de audio perceptiva a la señal mono, un descodificador de audio correspondiente toma el flujo de bits de audio como entrada y genera la señal de audio mono m descodificada, sometida a distorsión puesto que se introduce mediante el códec de audio perceptivo. El bloque de descodificación PS toma la información lateral PS como entrada y reconstruye el parámetro PS. Esto incluye descodificación Huffman y descodificación de diferencial de tiempo (dt) o diferencial de frecuencia (df) en caso de que se emplease tal codificación en el codificador. Finalmente, la reconstrucción PS toma la señal mono m descodificada y los parámetros PS como entrada y reconstruye la señal estéreo de 2 canales.

La figura 9 muestra un transmisor o grabador 330 de audio de la invención que presenta un codificador 10 de audio, una interfaz 332 de entrada y una interfaz 334 de salida.

Puede suministrarse una señal de audio en la interfaz 332 de entrada del transmisor/grabador 330. La señal de audio se codifica mediante un codificador 10 de la invención dentro del transmisor/grabador y la representación codificada se emite en la interfaz 334 de salida del transmisor/grabador 330. La representación codificada puede transmitirse o almacenarse entonces sobre un medio de almacenamiento.

La figura 10 muestra un receptor o reproductor 340 de audio de la invención, que presenta un descodificador 180 de audio de la invención, una entrada 342 de flujo de bits y una salida 344 de audio.

Puede introducirse un flujo de bits en la entrada 342 del receptor/reproductor 340 de audio de la invención. El flujo de bits se descodifica entonces mediante el descodificador 180 y la señal descodificada se emite o reproduce en la salida 344 del receptor/reproductor 340 de audio de la invención.

La figura 11 muestra un sistema de transmisión que comprende un transmisor 330 de la invención y un receptor 340 de la invención.

La entrada de señal de audio en la interfaz 332 de entrada del transmisor 330 se codifica y transfiere desde la salida 334 del transmisor 330 a la entrada 342 del receptor 340. El receptor descodifica la señal de audio y reproduce o emite la señal de audio sobre su salida 344.

Resumiendo el concepto inventivo, se puede decir que la presente invención se refiere a la codificación de representaciones multicanal de señales de audio que utiliza parámetros de audio espacial de una manera que es compatible con la codificación de señales estéreo de 2 canales que utiliza parámetros de estéreo paramétrico. La presente invención enseña nuevos métodos para la codificación eficaz tanto de parámetros de audio espacial como de parámetros de estéreo paramétrico y para embeber los parámetros codificados en un flujo de bits de una manera compatible hacia atrás. En particular, su objetivo es minimizar la tasa de bits global para los parámetros de estéreo paramétrico y

audio espacial en un flujo de bits compatible hacia atrás sin comprometer la calidad de la señal multicanal de audio o estéreo descodificada. Sin embargo, cuando es aceptable una calidad ligeramente comprometida de la señal estéreo descodificada, la tasa de bits global puede reducirse adicionalmente.

5 Aunque los flujos de bits que describen la compatibilidad hacia atrás de la señal de la invención y la generación de los mismos no comprenden parámetros que describen la señal de downmix monofónico, por supuesto tales parámetros puede incorporarse fácilmente en el flujo de bits mostrado.

10 Pueden predecirse números arbitrarios de los parámetros de audio espacial utilizando parámetros de estéreo paramétrico si se puede obtener una regla apropiada para predecir los parámetros. Por lo tanto, las reglas de predicción detalladas dadas anteriormente deben entenderse sólo como ejemplos. Está claro que otras reglas de predicción pueden llevar al mismo efecto de ahorro de bits y, por lo tanto, la presente invención no está limitada de ninguna manera a utilizar una de las reglas de predicción descritas anteriormente.

15 Aunque en los ejemplos de los codificadores de la invención dados existe un elemento 58 de downmix de estéreo paramétrico que obtiene un downmix estéreo de una señal multicanal, en implementaciones prácticas, puede omitirse el elemento de downmix estéreo, si se conoce la regla de downmix y, por lo tanto, cuando los parámetros de estéreo paramétrico pueden obtenerse directamente a partir de la señal multicanal.

20 En las implementaciones dadas, la señal de downmix monofónico se codifica adicionalmente mediante un codificador de audio o se descodifica en un lado de descodificador. La codificación y descodificación es opcional, es decir omitir una compresión adicional de la señal de downmix monofónico también proporcionará codificadores y descodificadores de la invención que incorporan el concepto inventivo.

25 La unidad de control dentro de los codificadores y descodificadores de la invención puede omitirse y se puede tomar una decisión general de representar subconjuntos de parámetros espaciales mediante parámetros predichos diferenciales con el provecho de ahorrar la unidad de control y al coste de aceptar una tasa de bits ligeramente superior para los casos extraños, cuando la representación predicha diferencial no ahorra tasa de bits de transmisión.

30 Aunque, dentro de los ejemplos dados, los codificadores adicionales aplicados en las trayectorias de señal se denominan sólo codificadores diferenciales o descodificadores diferenciales, se entiende que también puede utilizarse cualquier otro codificador o descodificador apropiado adecuado para comprimir los parámetros, especialmente una combinación de un descodificador o codificador diferencial y un descodificador o codificador Huffman. Tal combinación se utiliza de una manera que, en primer lugar, los parámetros se codifican de manera diferencial y luego
35 los parámetros codificados de manera diferencial se codifican según Huffman, lo que finalmente da como resultado una representación paramétrica que utiliza tasas de bits inferiores, puesto que la representación predicha de manera diferencial en general presenta entropía inferior que los parámetros espaciales subyacentes a los mismos.

40 Resumiendo las ideas de la invención, la presente invención enseña lo siguiente:

En un primer aspecto, un método para la codificación compatible de señales de audio multicanal, caracterizado por: en el lado del codificador, someter a downmix la señal multicanal para dar una representación de un canal; en el lado del codificador dada dicha señal multicanal, define parámetros que representan la señal multicanal; en el
45 lado del codificador dada dicha señal multicanal, definir parámetros que representan un downmix estéreo de la señal multicanal; en el lado del codificador, embeber ambos conjuntos de parámetros de una manera compatible hacia atrás y eficaz en tasa de bits en un flujo de bits; en el lado del descodificador, extraer los parámetros embebidos de un flujo de bits; en el lado del descodificador, reconstruir parámetros que representan una señal multicanal a partir de los parámetros extraídos del flujo de bits; en el lado del descodificador, reconstruir las señales de salida multicanal dados
50 los parámetros a partir de los datos del flujo de bits, y dicha señal sometida a downmix.

Como un segundo aspecto, un método según el primer aspecto, caracterizado por embeber los parámetros que representan un downmix estéreo en el flujo de bits, de tal manera que pueden descodificarse mediante un método de descodificación (legado) que sólo soporta descodificación de estéreo paramétrico.

55 Como un tercer aspecto, un método según el primer aspecto, caracterizado por dividir el conjunto de parámetros que representan la señal multicanal en un primer subconjunto y un segundo subconjunto.

Como un cuarto aspecto, un método según el tercer aspecto, caracterizado por una predicción de los valores en dicho primer subconjunto de parámetros basándose en dicho segundo subconjunto de parámetros y basándose en los parámetros que representan un downmix estéreo de la señal multicanal.

60 Como un quinto aspecto, un método según el cuarto aspecto, caracterizado por un método de control que selecciona automáticamente si el primer subconjunto de parámetros se codifica directamente o si sólo se codifican las diferencias con respecto a los valores de parámetros predichos.

65 Como un sexto aspecto, un método según el tercer aspecto, caracterizado por la modificación de los parámetros que presentan un downmix estéreo, en el que se utilizan tanto los parámetros originales que representan la señal

ES 2 306 235 T3

multicanal como los parámetros originales que representan el downmix estéreo como base para obtener los parámetros modificados.

5 Como un séptimo aspecto, un método según el cuarto aspecto, caracterizado por una tabla de consulta que se utiliza para encontrar dichos valores de parámetros predichos.

Como un octavo aspecto, un método según el cuarto aspecto, en el que en el cuarto aspecto se utiliza una función polinómica para encontrar dichos valores de parámetros predichos.

10 Como un noveno aspecto, un método según el cuarto aspecto, caracterizado por una función matemática obtenida a partir de método empleado para generar el downmix estéreo que se utiliza para encontrar dichos valores de parámetros predichos.

15 Como un décimo aspecto, un aparato para codificar una representación de una señal de audio multicanal, caracterizado por: medios para someter a downmix la señal multicanal para dar una representación de un canal; medios para definir parámetros que representan la señal multicanal; medios para definir parámetros que representan un downmix estéreo de la señal multicanal; medios para embeber ambos conjuntos de parámetros de una manera compatible hacia atrás y eficaz en tasa de bits en un flujo de bits.

20 Como un undécimo aspecto, un aparato para reconstruir una señal multicanal basándose en una señal sometida a downmix y conjuntos de parámetros correspondientes, caracterizado por: medios para extraer los conjuntos de parámetros embebidos en un flujo de bits; medios para reconstruir parámetros que representan una señal multicanal a partir de los parámetros extraídos del flujo de bits; medios para reconstruir la señal de salida multicanal dado el conjunto de parámetros reconstruido a partir de los datos del flujo de bits, y dicha señal sometida a downmix.

25 Dependiendo de ciertos requisitos de implementación de los métodos de la invención, los métodos de la invención pueden implementarse en hardware o en software. La implementación puede realizarse utilizando un medio de almacenamiento digital, en particular un disco, DVD o un CD que presente señales de control legibles electrónicamente almacenadas sobre el mismo, que actúe conjuntamente con un sistema informático programable de tal manera que se realicen los métodos de la invención. Por lo tanto, en general, la presente invención es un producto de programa informático con un código de programa almacenado sobre un portador legible por máquina, estando operativo el código de programa para realizar los métodos de la invención cuando el producto de programa informático se ejecuta en un ordenador. Dicho de otro modo, los métodos de la invención son, por lo tanto, un programa informático que presenta un código de programa para realizar al menos uno de los métodos de la invención cuando el programa informático se ejecuta en un ordenador.

30 Aunque lo anterior se ha mostrado y descrito en particular con referencia a realizaciones particulares de la misma, los expertos en la técnica entenderán que pueden realizarse otros cambios diversos en la forma y los detalles sin apartarse del alcance de la misma. Ha de entenderse que pueden realizarse diversos cambios a la hora de adaptarse a diferentes realizaciones sin apartarse del alcance de la invención que se define en las reivindicaciones.

REIVINDICACIONES

1. Descodificador (200) de audio multicanal para procesar una representación (202) paramétrica, en el que la representación (202) paramétrica comprende un parámetro (206) de estéreo que describe propiedades espaciales de un downmix estéreo de la señal multicanal e información sobre uno o más parámetros (204) espaciales que describen propiedades espaciales de una señal multicanal, dependiendo la información sobre uno o más parámetros espaciales del parámetro de estéreo, en el que la información sobre el uno o más parámetros (204) espaciales y el parámetro (206) de estéreo, cuando se combinan utilizando una regla de combinación, da como resultado uno o más parámetros (210) espaciales, comprendiendo el decodificador:

un reconstructor (208) de parámetros para combinar el parámetro (206) de estéreo y la información sobre el uno o más parámetros (204) espaciales utilizando la regla de combinación para obtener el uno o más parámetros (210) espaciales.

2. Descodificador (200) de audio multicanal según la reivindicación 1, en el que la regla de combinación es de tal manera que la combinación comprende una sustitución de un primer subconjunto de parámetros de la información sobre el uno o más parámetros (204) espaciales por parámetros de sustitución obtenidos combinando el parámetro (206) de estéreo y el primer subconjunto de los parámetros.

3. Descodificador (200) de audio multicanal según la reivindicación 2, en el que la regla de combinación es de tal manera que se obtiene un parámetro de sustitución mediante una combinación lineal del parámetro correspondiente del primer subconjunto de parámetros y de una predicción del mismo parámetro, en el que la predicción se obtiene utilizando parámetros de un segundo subconjunto de las informaciones sobre el uno o más parámetros (204) espaciales y el parámetro (206) de estéreo, combinándolos utilizando una regla de predicción.

4. Descodificador (200) de audio multicanal según la reivindicación 3, en el que la regla de predicción es de tal manera que la predicción se obtiene utilizando el parámetro (206) de estéreo.

5. Descodificador (200) de audio multicanal según la reivindicación 4, en el que la regla de predicción es de tal manera que el parámetro (206) de estéreo se utiliza como la predicción del parámetro espacial.

6. Descodificador (200) de audio multicanal según una de las reivindicaciones anteriores, en el que el parámetro (206) de estéreo comprende un primer parámetros P_1 que describe una diferencia de intensidad entre los canales del downmix estéreo y un segundo parámetro P_2 que describe una correlación entre los canales del downmix estéreo;

en el que el segundo subconjunto de parámetros comprende un parámetro S_3 que describe una diferencia de intensidad entre una suma de un canal izquierdo y un canal derecho de la señal multicanal y un canal central de la señal multicanal; y

en el que la regla de predicción es de tal manera que un parámetro S_1 del primer subconjunto de parámetros, describiendo el parámetro una diferencia de intensidad entre el canal izquierdo y el canal derecho de la señal multicanal, se predice mediante un parámetro $\hat{S}_1$ de predicción según las siguientes fórmulas:

$$\hat{S}_1 = \sqrt{\hat{X} / \hat{Y}},$$

donde

$$\hat{X} = \frac{P_1^2 S_3^2 + q^2 (P_1^2 - 1)}{P_1^2 + 1}, \quad \hat{Y} = \frac{S_3^2 - q^2 (P_1^2 - 1)}{P_1^2 + 1}$$

7. Descodificador (200; 220) de audio multicanal según una de las reivindicaciones anteriores, en el que el reconstructor (226) de parámetros comprende además una unidad (252) de decisión para decidir si el primer subconjunto de parámetros se sustituye por parámetros de sustitución o no.

8. Descodificador (200; 220) de audio multicanal según una de las reivindicaciones anteriores, que comprende además un elemento (222) de descomposición de flujo de bits para descomponer una representación del parámetro (234) de estéreo y representación de la información sobre el uno o más parámetros (232) espaciales de un flujo (230) de bits, en el que el flujo (230) de bits es compatible hacia atrás para poder procesarse por dispositivos de estéreo paramétrico legados.

9. Descodificador (200; 220) de audio multicanal según la reivindicación 8, que comprende además un decodificador de entropía y un decodificador (258, 260, 262a, 262b) diferencial para obtener el parámetro (272) de estéreo y

ES 2 306 235 T3

la información sobre el uno o más parámetros (266, 268) espaciales a partir de la representación del parámetro (234) de estéreo y de la representación de la información sobre el uno o más parámetros (232) espaciales.

10. Descodificador (220) de audio multicanal según la reivindicación 8 ó 9, en el que el elemento (222) de descomposición de flujo de bits está operativo además para descomponer una señal (236) de downmix monofónico del flujo de bits, siendo la señal (236) de downmix monofónica un downmix monofónico de la señal multicanal; y

que comprende además un elemento (228) de upmix para obtener una reconstrucción de la señal (240) multicanal utilizando la señal (236) de downmix y el uno o más parámetros (238).

11. Descodificador (220) de audio multicanal según la reivindicación 10, que comprende además un decodificador (224) de audio para obtener la señal (236) de downmix monofónico a partir de una representación codificada de la señal (231) de downmix monofónico descompuesta del flujo (230) de bits.

12. Codificador (10) para obtener una representación (12) paramétrica de una señal de audio multicanal, presentando la representación (12) paramétrica parámetros adecuados para utilizarse junto con una señal de downmix monofónico, comprendiendo el codificador (10):

un calculador (14) de parámetros espaciales para calcular uno o más parámetros (20) espaciales que describen propiedades espaciales de la señal multicanal;

un calculador (16) de parámetros de estéreo para calcular un parámetro (22) de estéreo que describe propiedades espaciales de una señal de downmix estéreo obtenida a partir de la señal multicanal; y

un combinador (18) de parámetros para generar la representación (12) paramétrica combinando el uno o más parámetros (20) espaciales y los parámetros (22) de estéreo utilizando una regla de combinación, en el que el combinador (18) de parámetros está operativo para utilizar una regla de combinación que da como resultado un parámetro (24) de estéreo utilizable por el decodificador y una información sobre el uno o más parámetros (26) espaciales, dependiendo la información sobre uno o más parámetros espaciales del parámetro de estéreo, que representa, junto con el parámetro (24) de estéreo utilizable por el decodificador, el uno o más parámetros (20) espaciales.

13. Codificador (10; 52) según la reivindicación 12, en el que el calculador (16) de parámetros de estéreo comprende además un elemento (58) de downmix estéreo para obtener la señal (72) de downmix estéreo a partir de la señal (50) multicanal.

14. Codificador (10; 52) según las reivindicaciones 12 ó 13, que comprende además un elemento (54) de downmix espacial para obtener la señal (68) de downmix monofónico a partir de la señal (50) multicanal.

15. Codificador (10; 52) según una de las reivindicaciones 12 a 14, que comprende además un calculador (66) de flujo de bits para obtener un flujo (82) de bits que comprende la representación paramétrica y el downmix monofónico de una manera que es compatible hacia atrás con decodificadores de estéreo paramétrico legados.

16. Codificador según la reivindicación 14 ó 15, en el que el elemento (54) de downmix espacial comprende además un codificador (62) de audio para la compresión de la señal (68) de downmix monofónico utilizando una regla de compresión.

17. Método para procesar una representación paramétrica, en el que la representación paramétrica comprende un parámetro (206) de estéreo que describe propiedades espaciales de un downmix estéreo de la señal multicanal e información sobre uno o más parámetros espaciales que describen propiedades espaciales de una señal multicanal, dependiendo la información sobre uno o más parámetros espaciales del parámetro de estéreo, en el que la información sobre el uno o más parámetros espaciales y los parámetros de estéreo, cuando se combinan utilizando una regla de combinación, da como resultado el uno o más parámetros espaciales, comprendiendo el método:

combinar el parámetro de estéreo y la información sobre el uno o más parámetros espaciales utilizando la regla de combinación para obtener el uno o más parámetros espaciales.

18. Método para obtener una representación paramétrica de una señal de audio multicanal, presentando la representación paramétrica parámetros adecuados para utilizarse junto con una señal de downmix monofónico, comprendiendo el método:

calcular uno o más parámetros espaciales que describen propiedades espaciales de la señal multicanal;

calcular un parámetro de estéreo que describe propiedades espaciales de una señal de downmix estéreo obtenida a partir de la señal multicanal; y

generar la representación paramétrica combinando el uno o más parámetros espaciales y el parámetro de estéreo utilizando una regla de combinación, en el que utilizar la regla de combinación da como resultado un parámetro de estéreo utilizable por el decodificador e información sobre el uno o más parámetros espaciales, dependiendo la

información sobre uno o más parámetros espaciales del parámetro de estéreo, que representa, junto con el parámetro de estéreo utilizable por el decodificador, el uno o más parámetros espaciales.

19. Representación paramétrica de una señal de audio multicanal, presentando la representación paramétrica parámetros adecuados para utilizarse junto con una señal de downmix monofónico, en el que la representación paramétrica presenta un parámetro de estéreo utilizable por el decodificador que describe propiedades espaciales de un downmix estéreo de la señal multicanal e información sobre uno o más parámetros espaciales generados combinando uno o más parámetros espaciales que describen propiedades espaciales de la señal de audio multicanal y el parámetro de estéreo, dependiendo la información sobre uno o más parámetros espaciales del parámetro de estéreo, de tal manera que la información sobre el uno o más parámetros espaciales representa, junto con el parámetro de estéreo utilizable por el decodificador, el uno o más parámetros espaciales.

20. Medio de almacenamiento legible por ordenador que presenta almacenado sobre el mismo una representación paramétrica de una señal de audio multicanal según la reivindicación 19.

21. Transmisor o grabador (330) de audio que presenta un codificador para obtener una representación paramétrica de una señal de audio multicanal según la reivindicación 12.

22. Receptor o reproductor (340) de audio que presenta un decodificador (200) de audio multicanal según la reivindicación 1.

23. Método de transmisión o grabación de audio, presentando el método un método para obtener una representación paramétrica de una señal de audio multicanal según la reivindicación 18.

24. Método de recepción o reproducción de audio, presentando el método un método para procesar una representación paramétrica según la reivindicación 17.

25. Sistema de transmisión que presenta un transmisor (330) y un receptor (340);

presentando el transmisor (330) un codificador para obtener una representación paramétrica de una señal de audio multicanal según la reivindicación 12; y

presentando el receptor (340) un decodificador de audio multicanal según la reivindicación 1.

26. Método de transmisión y recepción, incluyendo el método un método de transmisión que presenta un método para obtener una representación paramétrica de una señal de audio multicanal según la reivindicación 18 y un método de recepción, que presenta un método para procesar una representación paramétrica según la reivindicación 17.

27. Programa informático para realizar, cuando se ejecuta en un ordenador, un método según cualquiera de las reivindicaciones 17, 18, 23, 24 ó 26 de método.

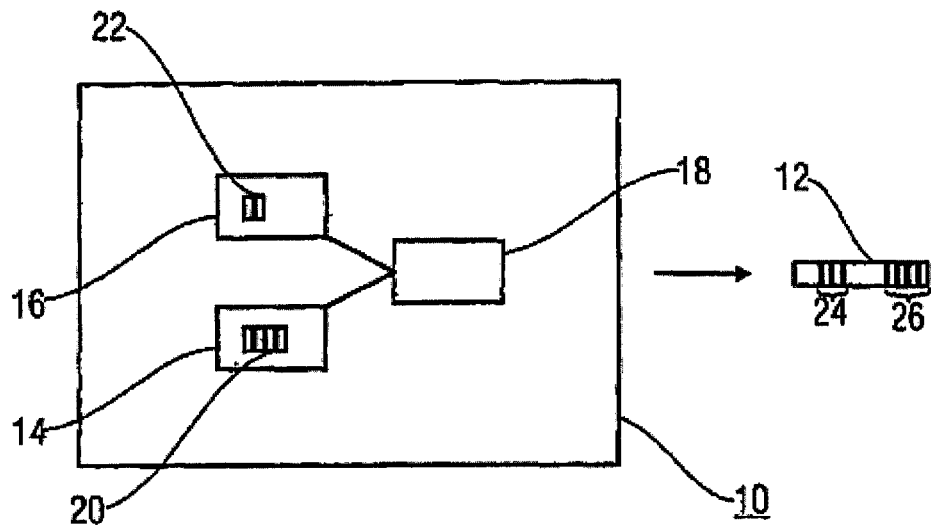


Fig. 1

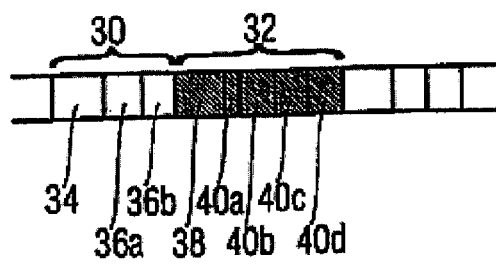


Fig. 2

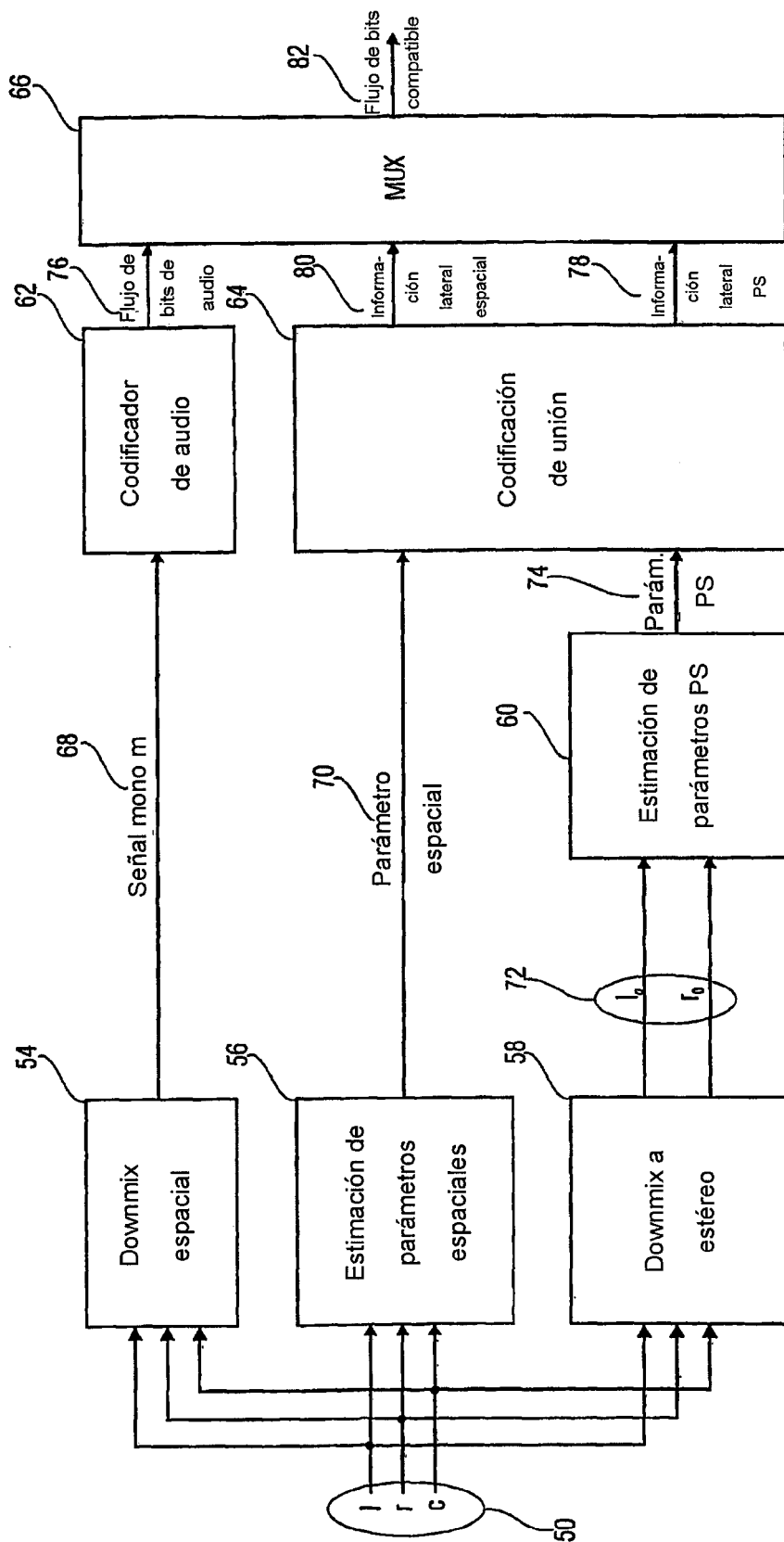


Fig. 3

52

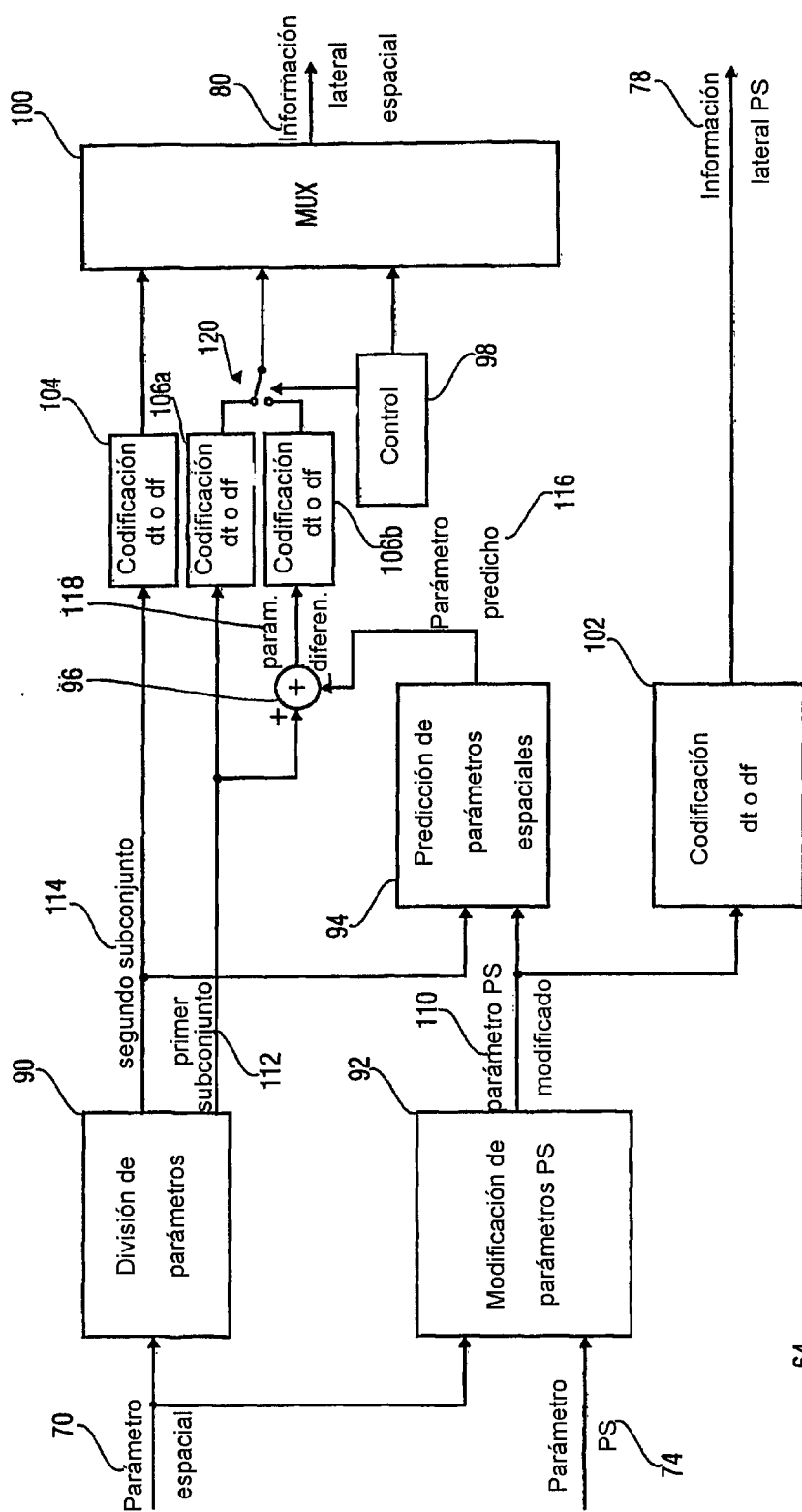


Fig. 4

64

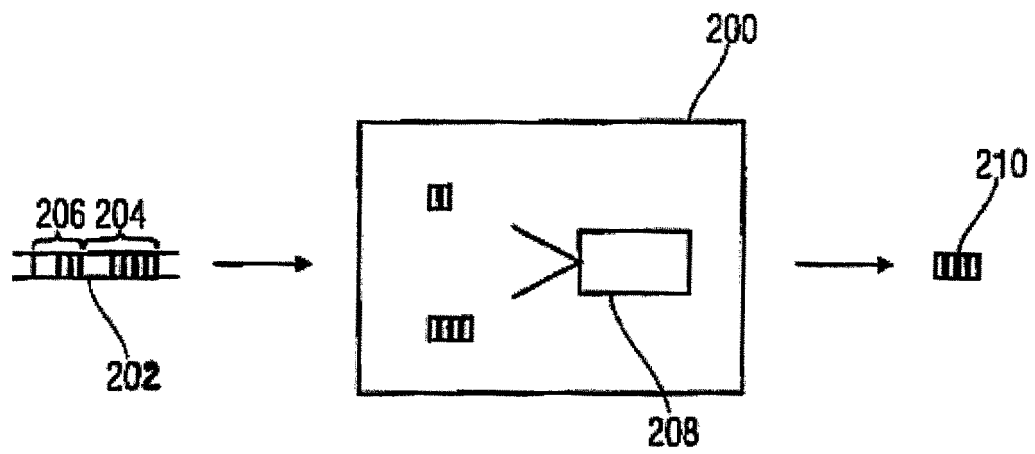


Fig. 5

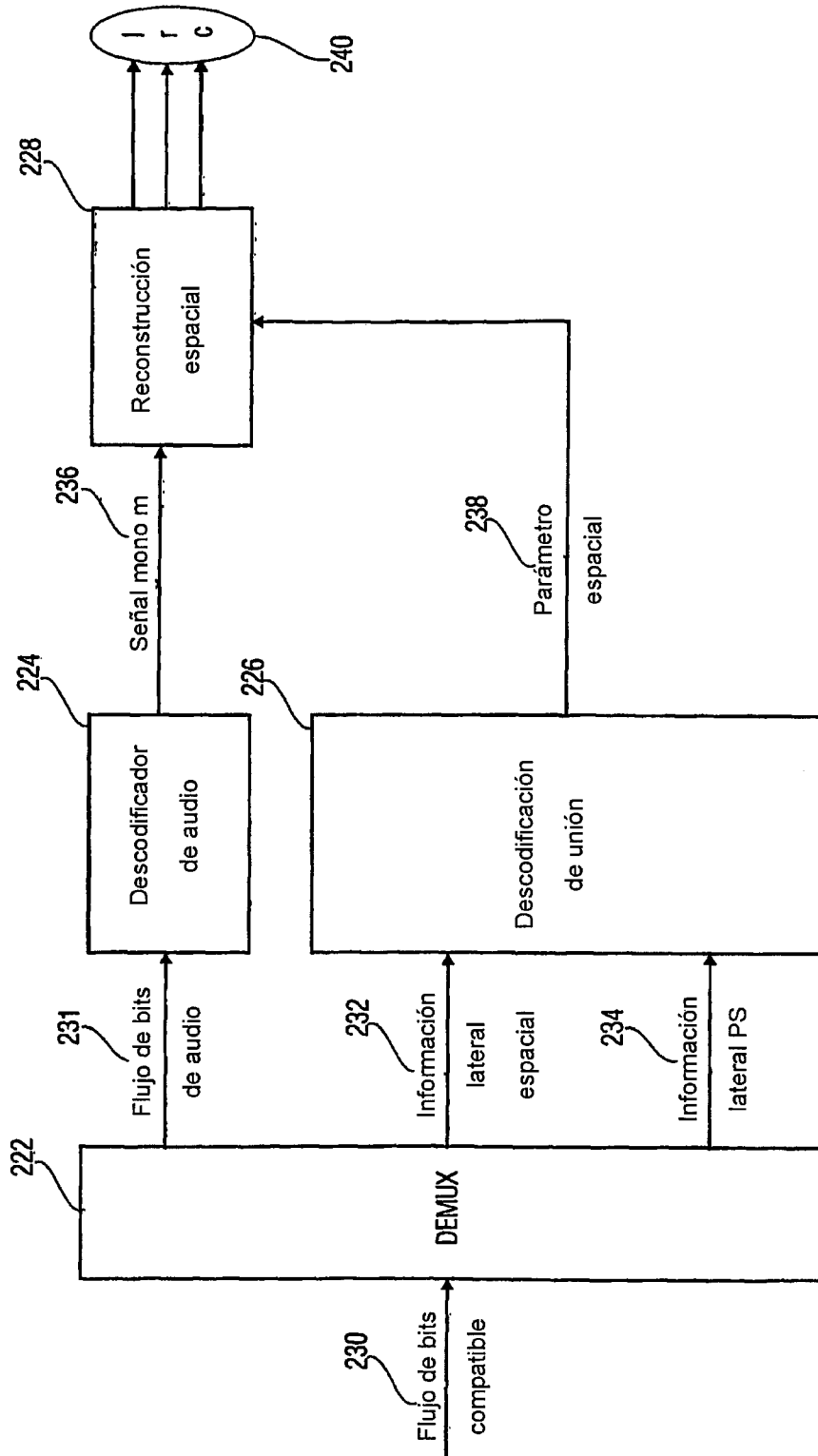


Fig. 6

220

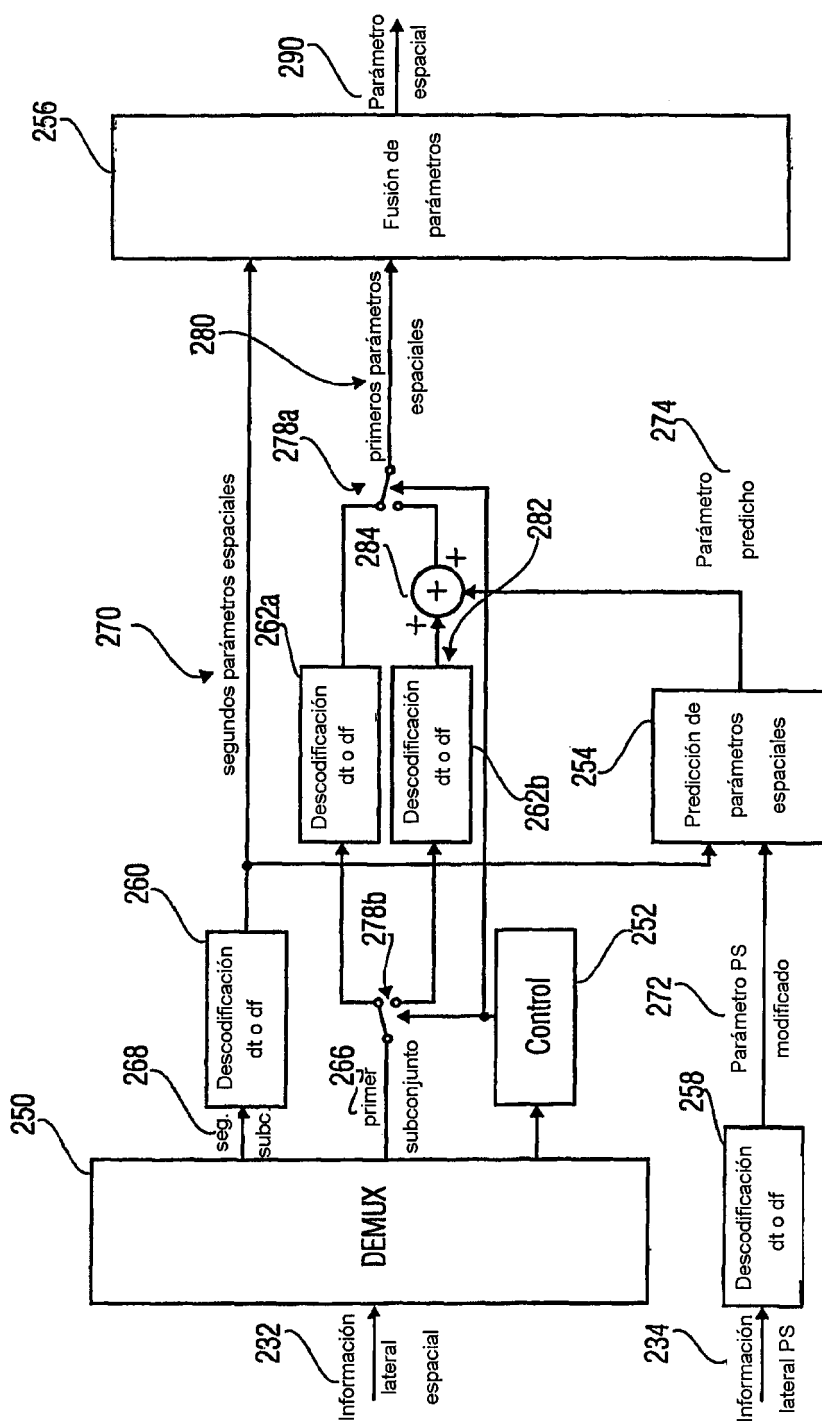


Fig. 7

226

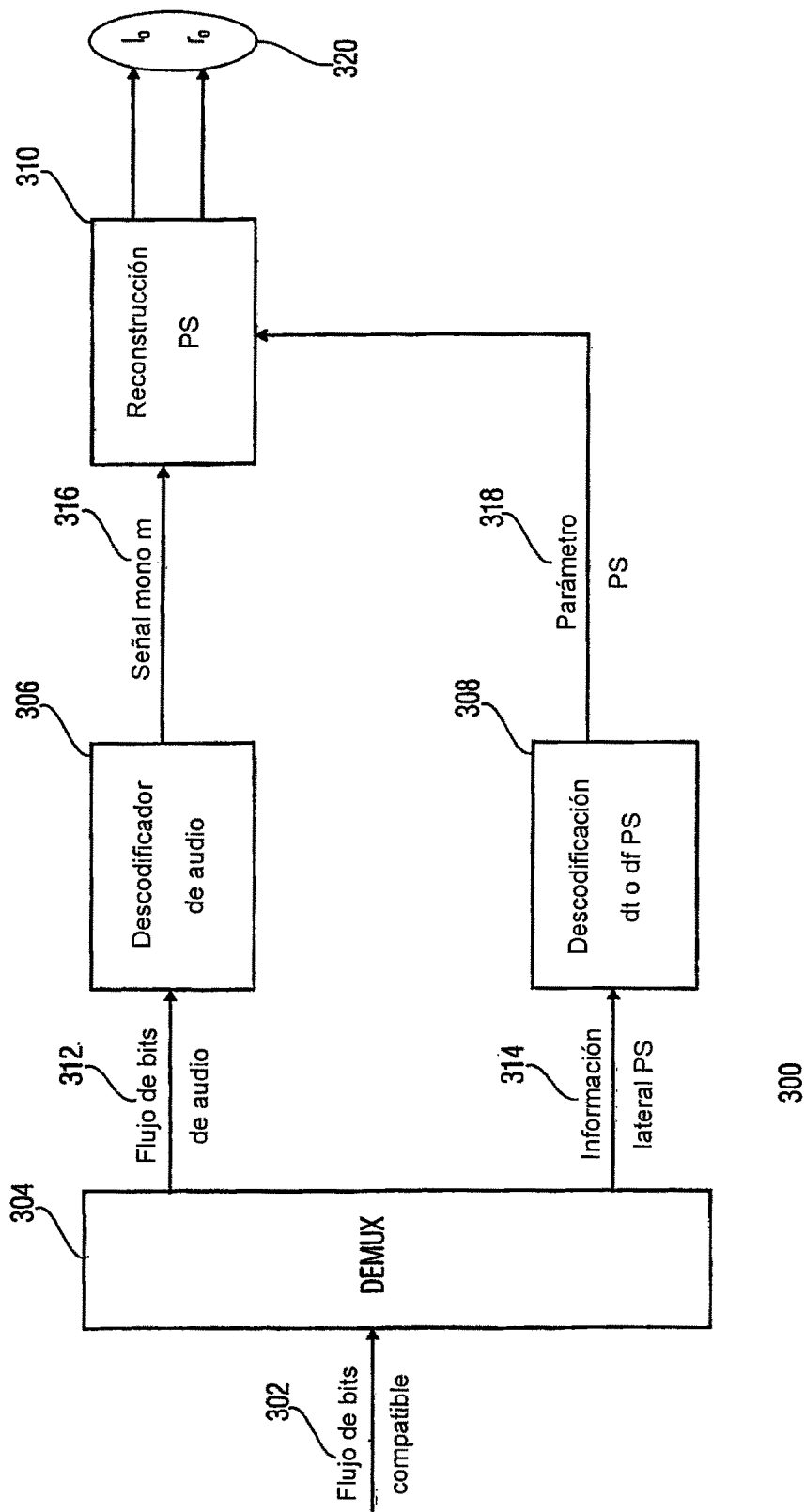


Fig. 8

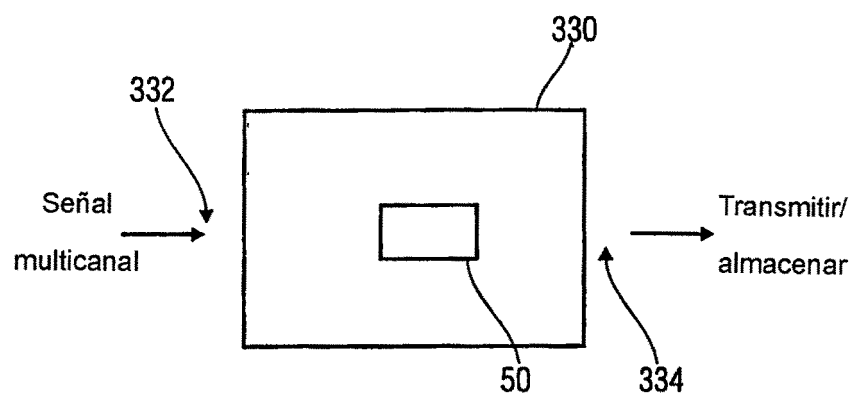


Fig. 9

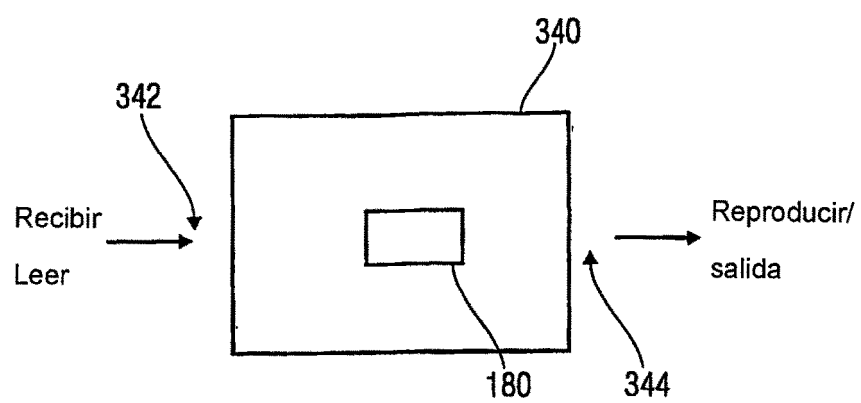


Fig. 10

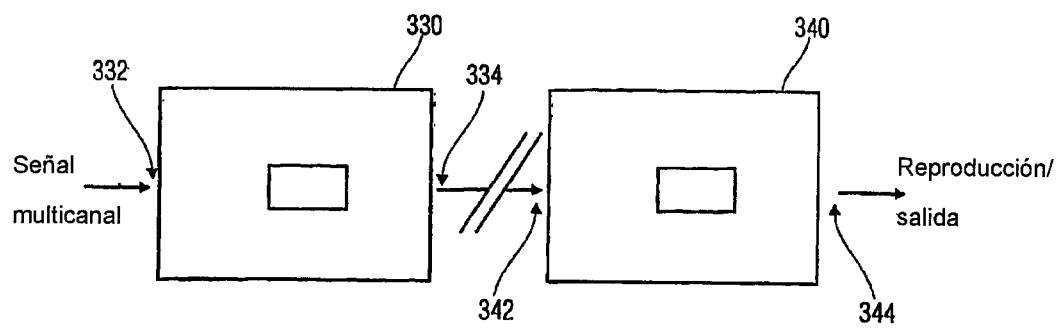


Fig. 11