

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 740 104**

51 Int. Cl.:

G10L 19/008 (2013.01)

H04S 3/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **01.02.2006 PCT/EP2006/000875**

87 Fecha y número de publicación internacional: **19.10.2006 WO06108462**

96 Fecha de presentación y número de la solicitud europea: **01.02.2006 E 06706552 (4)**

97 Fecha y número de publicación de la concesión europea: **08.05.2019 EP 1869667**

54 Título: **Codificación de audio jerárquica multicanal con información lateral compacta**

30 Prioridad:

21.12.2005 US 314711

15.04.2005 US 671544 P

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

05.02.2020

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (33.3%)**

Hansastraße 27c

80686 München, DE;

DOLBY INTERNATIONAL AB (33.3%) y

KONINKLIJKE PHILIPS N.V. (33.3%)

72 Inventor/es:

HOELZER, ANDREAS;

HERRE, JUERGEN;

ROEDEN, JONAS;

PURNHAGEN, HEIKO;

KJOERLING, KRISTOFER;

ENGDEGARD, JONAS;

BREEBAART, JEROEN;

SCHUIJERS, ERIK;

OOMEN, WERNER y

VILLEMOES, LARS

74 Agente/Representante:

SALVÀ FERRER, Joan

ES 2 740 104 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 jerárquica multicanal con información lateral compacta

5 Campo de la invención

[0001] La presente invención se refiere al procesamiento de audio de múltiples canales y, en particular, a la generación y el uso de información lateral paramétrica compacta para describir las propiedades espaciales de una señal de audio de múltiples canales.

10

Antecedentes de la invención y técnica anterior

[0002] En los últimos tiempos, la técnica de reproducción de audio multicanal se está volviendo cada vez más importante. Esto puede deberse al hecho de que las técnicas de compresión/codificación de audio, como la conocida técnica de mp3, han hecho posible la distribución de grabaciones de audio a través de Internet u otros canales de transmisión que tienen un ancho de banda limitado. La técnica de codificación de mp3 se ha vuelto tan famosa debido al hecho de que permite la distribución de todos los registros en un formato estéreo, es decir, una representación digital de la grabación de audio que incluye un primer canal estéreo o izquierdo y un segundo canal estéreo o derecho.

15

[0003] Sin embargo, existen deficiencias básicas de los sistemas de sonido de dos canales convencionales. Por ello, ha sido desarrollada la técnica envolvente. Un formato recomendado de presentación envolvente multicanal incluye, además de dos canales estéreo L y R, un canal central adicional C y dos canales envolventes Ls, Rs. Este formato de sonido de referencia también se conoce como tres/dos estéreo, lo que significa tres canales frontales y dos canales envolventes. En un entorno de reproducción, se necesitan al menos cinco altavoces en cinco ubicaciones apropiadas para obtener un punto agradable óptimo a una cierta distancia de los cinco altavoces bien colocados.

20

25

[0004] Las técnicas recientes para la codificación paramétrica de señales de audio multicanal (estéreo paramétrico (PS), "codificación de audio espacial", "codificación de señal binaural" (BCC), etc.) representan una señal de audio multicanal por medio de una señal de mezcla descendente (podría ser monofónica o abarcar varios canales) y la información lateral paramétrica ("señales espaciales"), que caracteriza su etapa de sonido espacial percibida. Las distintas soluciones y técnicas serán revisadas en breve en los párrafos siguientes.

30

[0005] Una técnica relacionada, también conocida como estéreo paramétrico, se describe en J. Breebaart, S. van de Par, A. Kohlrausch, E. Schuijers, "High-Quality Parametric Spatial Audio Coding at Low Bitrates", AES 116th Convention, Berlín, Preprint 6072, mayo de 2004, y E. Schuijers, J. Breebaart, H. Purnhagen, J. Engdegard, "Low Complexity Parametric Stereo Coding", AES 116th Convention, Berlín, Preprint 6073, mayo de 2004.

35

[0006] Se conocen varias técnicas en la técnica para reducir la cantidad de datos requeridos para la transmisión de una señal de audio multicanal. Con este fin, se hace referencia a la fig. 11, que muestra un dispositivo estéreo conjunto 60. Este dispositivo puede ser un dispositivo de implementación, por ejemplo, estéreo de intensidad (IS) o codificación de señal binaural (BCC). Dicho dispositivo generalmente recibe, como entrada, al menos dos canales (CH1, CH2, ... CHn), y genera un solo canal portador y datos paramétricos. Los datos paramétricos se definen de manera tal que, en un decodificador, se puede calcular una aproximación de un canal original (CH1, CH2, ... CHn).

40

[0007] Normalmente, el canal portador incluirá muestras de subbanda, coeficientes espectrales, muestras en el dominio del tiempo, etc., que proporcionan una representación relativamente fina de la señal subyacente, mientras que los datos paramétricos no incluyen tales muestras de coeficientes espectrales, sino que incluyen parámetros de control para controlar un determinado algoritmo de reconstrucción, como la ponderación por multiplicación, el desplazamiento en el tiempo, el desplazamiento de frecuencia, el desplazamiento de fase, etc. Los datos paramétricos, por lo tanto, incluyen solo una representación relativamente tosca de la señal o del canal asociado. En términos numéricos, la cantidad de datos requeridos por un canal portador puede estar en el rango de 60 a 70 kbit/s en un esquema de codificación MPEG, mientras que la cantidad de datos requeridos por la información del lado paramétrico para un canal puede estar en el rango de aproximadamente 10 kbit/s para una señal de canales 5.1. Un ejemplo de datos paramétricos son los factores de escala conocidos, la información estéreo de intensidad o los parámetros de referencia binaurales que se describirán a continuación.

45

50

55

[0008] La técnica BCC se describe, por ejemplo, en el documento de la convención AES 5574, "Binaural Cue Coding applied to Stereo and Multi-Channel Audio Compression", C. Faller, F. Baumgarte, mayo de 2002, Múnich, en el documento IEEE WASPAA Paper "Efficient representation of spatial audio using perceptual parametrization", octubre de 2001, Mohonk, NY, y en los 2 documentos de ICASSP "Estimation of auditory spatial cues for binaural cue coding", y "Binaural cue coding: a novel and efficient representation of spatial audio", ambos escritos por C. Faller, y F. Baumgarte, Orlando, FL, mayo de 2002.

60

[0009] En la codificación BCC, una serie de canales de entrada de audio se convierten en una representación espectral utilizando una transformada basada en DFT (transformada discreta de Fourier) con ventanas superpuestas.

65

El espectro resultante se divide en particiones no superpuestas. Cada partición tiene un ancho de banda proporcional al ancho de banda rectangular equivalente (ERB). Las diferencias de nivel entre canales (ICLD) y las diferencias de tiempo entre canales (ICTD) se estiman para cada partición. Las diferencias de nivel entre canales ICLD y las diferencias de tiempo entre canales ICTD se dan normalmente para cada canal con respecto a un canal de referencia y, además, se cuantifican. Los parámetros transmitidos se calculan finalmente según las fórmulas prescritas (codificadas), que pueden depender de las particiones específicas de la señal a procesar.

[0010] En el lado del decodificador, el decodificador recibe una señal mono y el flujo de bits BCC. La señal mono se transforma en el dominio de la frecuencia y se introduce en un bloque de síntesis espacial, que también recibe valores decodificados de ICLD y ICTD. En el bloque de síntesis espacial, los valores de los parámetros BCC (ICLD y ICTD) se utilizan para realizar una operación de ponderación de la señal mono con el fin de sintetizar las señales multicanal, que, después de una conversión de frecuencia/tiempo, representan una reconstrucción de la señal de audio multicanal original. En el caso de BCC, el módulo estéreo conjunto 60 está operativo para emitir la información del lado del canal, de modo que los datos del canal paramétrico se cuantifican y codifican, lo que da como resultado parámetros ICLD o ICTD, donde uno de los canales originales se utiliza como canal de referencia mientras codifica información lateral del canal.

[0011] Normalmente, el canal portador se forma a partir de la suma de los canales originales participantes.

[0012] Por lo tanto, las técnicas anteriores proporcionan adicionalmente una representación mono adecuada para el equipo de reproducción que solo puede procesar el canal portador y no puede procesar los datos paramétricos para generar una o más aproximaciones de más de un canal de entrada.

[0013] La técnica de codificación de audio conocida como codificación de señal binaural (BCC) también está bien descrita en las publicaciones de solicitud de patente de Estados Unidos US 2003, 0219130 A1, 2003/0026441 A1 y 2003/0035553 A1. También se hace referencia adicional a "Binaural Cue Coding. Parte II: Schemes and Applications", C. Faller and F. Baumgarte, IEEE Trans. on Audio and Speech Proc., Vol. 11, N.º 6, noviembre de 2003 y a "Binaural cue coding applied to audio compression with flexible rendering", C. Faller and F. Baumgarte, AES 113th Convention, Los Angeles, octubre de 2002.

[0014] Si bien los parámetros de ICLD y ICTD representan los parámetros de localización de fuente de sonido más importantes, una representación espacial que utiliza estos parámetros solo limita la calidad máxima que se puede alcanzar. Para superar esta limitación y, por lo tanto, para permitir la codificación paramétrica de alta calidad, estéreo paramétrico (como se describe J. Breebaart, S. van de Par, A. Kohlrausch, E. Schuijers (2005) "Parametric coding of stereo audio", Eurasip J. Applied Signal Proc. 9, 1305-1322) aplica tres tipos de parámetros espaciales, denominados Diferencias de Intensidad Intercanal (IID), Diferencias de Fase Intercanal (IPD) y Coherencia Intercanal (IC). La extensión del conjunto de parámetros espaciales con parámetros de coherencia permite una parametrización de la "difusividad" espacial percibida o la "compacidad" espacial de la etapa de sonido.

[0015] Donde sigue, se elabora con más detalle un esquema BCC genérico típico para la codificación de audio multicanal con referencia a las figuras 12 a 14. La figura 9 muestra un esquema de codificación binaural genérico de este tipo para la codificación/transmisión de señales de audio multicanal. La señal de entrada de audio multicanal en una entrada 110 de un codificador BCC 112 se mezcla en un bloque 114 de mezcla descendente. En el presente ejemplo, la señal multicanal original en la entrada 110 es una señal envolvente de 5 canales que tiene un canal frontal izquierdo, un canal frontal derecho, un canal envolvente izquierdo, un canal envolvente derecho y un canal central. En una realización preferida de la presente invención, el bloque de mezcla descendente 114 produce una señal de suma mediante una simple adición de estos cinco canales en una señal mono. Otros esquemas de mezcla descendente son conocidos en la técnica, de manera que, utilizando una señal de entrada multicanal, se puede obtener una señal de mezcla descendente que tiene un solo canal. Este canal único se emite en una línea de señal de suma 115. Una información lateral obtenida por un bloque de análisis BCC 116 se emite en una línea de información lateral 117. En el bloque de análisis BCC, las diferencias de nivel entre canales (ICLD) y las diferencias de tiempo entre canales (ICTD) se calculan como se ha descrito anteriormente. El bloque de análisis BCC 116 se forma para calcular también los valores de correlación entre canales (valores ICC). La señal de suma y la información lateral se transmiten, preferentemente en forma cuantificada y codificada, a un decodificador BCC 120. El decodificador BCC descompone la señal de suma transmitida en una serie de subbandas y aplica escalas, retardos y otros procesamientos para generar las subbandas de las señales de audio multicanal de salida. Este procesamiento se realiza de manera tal que los parámetros ICLD, ICTD e ICC (señales) de una señal multicanal reconstruida en una salida 121 son similares a las señales respectivas para la señal multicanal original en la entrada 110 del codificador BCC 112. Para este fin, el decodificador BCC 120 incluye un bloque de síntesis BCC 122 y un bloque de procesamiento de información lateral 123.

[0016] A continuación, la construcción interna del bloque de síntesis BCC 122 se explica con referencia a la fig. 13. La señal de suma en la línea 115 se introduce en una unidad de conversión de tiempo/frecuencia o en el banco de filtros FB 125. En la salida del bloque 125, hay un número N de señales de subbanda o, en un caso extremo, un bloque de coeficientes espectrales, cuando el banco de filtros de audio 125 realiza una transformación 1:1, es decir,

una transformación que produce N Coeficientes espectrales de N muestras de dominio de tiempo (submuestreo crítico).

5 **[0017]** El bloque 122 de síntesis de BCC comprende además una etapa de retardo 126, una etapa de modificación de nivel 127, una etapa de procesamiento de correlación 128 y una etapa de banco de filtros inverso IFB 129. En la salida de la etapa 129, la señal de audio multicanal reconstruida que tiene, por ejemplo, cinco canales en el caso de un sistema de sonido envolvente de 5 canales, puede enviarse a un conjunto de altavoces 124 como se ilustra en la fig. 12.

10 **[0018]** Como se muestra en la fig. 13, la señal de entrada $s(n)$ se convierte en el dominio de frecuencia o dominio de banco de filtros por medio del elemento 125. La señal de salida por el elemento 125 se multiplica de manera que se obtienen varias versiones de la misma señal, como se ilustra en el nódulo de ramificación 130. El número de versiones de la señal original es igual al número de canales de salida en la señal de salida a reconstruir. Cuando, en general, cada versión de la señal original en el nódulo 130 se somete a un cierto retardo d_1, d_2, \dots, d_N . Los parámetros de retardo se calculan mediante el bloque de procesamiento de información lateral 123 en la fig. 12 y se derivan de las diferencias de tiempo entre canales determinadas por el bloque de análisis BCC 116.

15 **[0019]** Lo mismo es cierto para los parámetros de multiplicación a_1, a_2, \dots, a_N , que también se calculan mediante el bloque de procesamiento de información lateral 123 basado en las diferencias de nivel entre canales calculadas por el bloque de análisis BCC 116.

20 **[0020]** Los parámetros ICC calculados por el bloque de análisis BCC 116 se utilizan para controlar la funcionalidad del bloque 128 de manera que se obtengan ciertas correlaciones entre las señales manipuladas en el nivel y las retardadas en las salidas del bloque 128. Debe observarse aquí que el orden de las etapas 126, 127, 128 puede ser diferente del caso que se muestra en la fig. 13.

25 **[0021]** Se debe tener en cuenta que, en un procesamiento de una señal de audio en forma de trama, el análisis BCC también se realiza en forma de trama, es decir, que varía con el tiempo, y también en forma de frecuencia. Esto significa que, para cada banda espectral, los parámetros BCC se obtienen individualmente. Esto significa además que, en caso de que el banco de filtros de audio 125 descomponga la señal de entrada en, por ejemplo, señales de paso de 32 bandas, el bloque de análisis BCC obtiene un conjunto de parámetros BCC para cada una de las 32 bandas. Naturalmente, el bloque de síntesis BCC 122 de la fig. 12, que se muestra en detalle en la fig. 13, realiza una reconstrucción, que también se basa en las 32 bandas del ejemplo.

30 **[0022]** A continuación, se hace referencia a la fig. 14 que muestra una configuración para determinar ciertos parámetros de BCC. Normalmente, los parámetros ICLD, ICTD y ICC se pueden definir entre pares de canales arbitrarios. Un procedimiento, que se describirá aquí, consiste en los parámetros de ICLD y ICTD entre un canal de referencia y cada uno de los otros canales. Esto se ilustra en la fig. 14A.

35 **[0023]** Los parámetros ICC se pueden definir de distintas maneras. En general, se podrían estimar los parámetros ICC en el codificador entre todos los pares de canales posibles, como se indica en la fig. 14B. En este caso, un decodificador sintetizaría ICC de manera que sea aproximadamente el mismo que en la señal multicanal original entre todos los pares de canales posibles. Sin embargo, se propuso estimar solo los parámetros ICC entre los dos canales más fuertes a la vez. Este esquema se ilustra en la fig. 14C, donde se muestra un ejemplo, donde en una instancia de tiempo, se estima un parámetro ICC entre los canales 1 y 2, y, en otra instancia de tiempo, se calcula un parámetro ICC entre los canales 1 y 5. El decodificador a continuación sintetiza la correlación entre canales entre los canales más fuertes en el decodificador y aplica alguna regla heurística para calcular y sintetizar la coherencia entre canales para los pares de canales restantes.

40 **[0024]** Con respecto al cálculo de, por ejemplo, los parámetros de multiplicación a_1, \dots, a_N basados en los parámetros ICLD transmitidos, se hace referencia al documento de convención de AES 5574 citado anteriormente. Los parámetros ICLD representan una distribución de energía en una señal multicanal original. Sin pérdida de generalidad, en la fig. 14A se muestra que hay cuatro parámetros ICLD que muestran la diferencia de energía entre todos los demás canales y el canal frontal izquierdo. En el bloque de procesamiento de información lateral 123, los parámetros de multiplicación a_1, \dots, a_N se derivan de los parámetros ICLD, de modo que la energía total de todos los canales de salida reconstruidos es la misma que (o proporcional a) la energía de la señal de suma transmitida. Una forma sencilla de determinar estos parámetros es un procedimiento de 2 etapas, en el cual, en una primera etapa, el factor de multiplicación para el canal frontal izquierdo se establece en la unidad, mientras que los factores de multiplicación para los otros canales en la fig. 14A se determinan a partir de los valores de ICLD transmitidos. A continuación, en una segunda etapa, la energía de los cinco canales se calcula y se compara con la energía de la señal de suma transmitida. A continuación, todos los canales se reducen de escala utilizando un factor de reducción de escala que es igual para todos los canales, donde el factor de reducción de escala se selecciona de tal manera que la energía total de todos los canales de salida reconstruidos es, después de la reducción de escala, igual a la energía total de la señal de suma transmitida.

65

[0025] Naturalmente, también hay otros procedimientos para calcular los factores de multiplicación, que no se basan en el procedimiento de 2 etapas, sino que solo necesitan un procedimiento de 1 etapa.

[0026] Con respecto a los parámetros de retardo, debe observarse que los parámetros de retardo ICTD, que se transmiten desde un codificador BCC se pueden usar directamente, cuando el parámetro de retardo d_1 para el canal frontal izquierdo se pone a cero. No se debe realizar ningún cambio de escala aquí, ya que un retraso no altera la energía de la señal.

[0027] Como se ha descrito anteriormente con respecto a la fig. 14, la información del lado paramétrico, es decir, las diferencias de nivel entre canales (ICLD), las diferencias de tiempo entre canales (ICTD) o el parámetro de coherencia entre canales (ICC), se puede calcular y transmitir para cada uno de los cinco canales. Esto significa que uno, normalmente, transmite cuatro conjuntos de diferencias de nivel entre canales para una señal de cinco canales. Lo mismo es cierto para las diferencias de tiempo entre canales. Con respecto al parámetro de coherencia entre canales, también puede ser suficiente para transmitir solo, por ejemplo, dos conjuntos de estos parámetros.

[0028] Como se ha descrito anteriormente con respecto a la fig. 13, no hay un único parámetro de diferencia de nivel, parámetro de diferencia de tiempo o parámetro de coherencia para una trama o parte de tiempo de una señal. En su lugar, estos parámetros se determinan para varias bandas de frecuencia distintas, de modo que se obtiene una parametrización dependiente de la frecuencia. Dado que se prefieren utilizar, por ejemplo, 32 canales de frecuencia, es decir, un banco de filtros que tiene 32 bandas de frecuencia para el análisis de BCC y la síntesis de BCC, los parámetros pueden ocupar una gran cantidad de datos. Aunque, en comparación con otras transmisiones multicanal, la representación paramétrica da como resultado una velocidad de datos bastante baja, existe una necesidad continua de una reducción adicional de la velocidad de datos necesaria para representar una señal que tenga más de dos canales, como una señal de sonido envolvente multicanal.

[0029] La codificación de una señal de audio multicanal puede implementarse ventajosamente utilizando varios módulos existentes, que realizan una codificación estéreo paramétrica en un solo canal mono. La solicitud de patente internacional WO2004008805 A1 enseña cómo los codificadores estéreo paramétricos pueden ordenarse en una configuración jerárquica de tal manera que un número dado de canales de audio de entrada se mezclen posteriormente en un solo canal mono. La información del lado paramétrico, que describe las propiedades espaciales del monocanal de mezcla descendente, finalmente consiste en toda la información paramétrica producida posteriormente durante el procedimiento de mezcla descendente iterativa. Esto significa que si hay, por ejemplo, tres procedimientos de mezcla descendente estéreo a mono involucrados en la construcción de la señal mono final, el conjunto final de parámetros que construyen la representación paramétrica de la señal de audio multicanal consiste en los tres grupos de los parámetros derivados durante cada procedimiento de mezcla descendente de estéreo a mono.

[0030] En la fig. 15 se muestra un codificador de mezcla descendente jerárquico para explicar el procedimiento de la técnica anterior con más detalle. La fig. 15 muestra seis canales de audio originales 200a a 200f que se transforman en un solo canal de audio monofónico 202 más información lateral paramétrica. Por lo tanto, los seis canales de audio originales 200a a 200f deben transformarse del dominio del tiempo al dominio de la frecuencia, que se realiza mediante la transformación de las unidades 204, lo que transforma los canales de audio 200a a 200f en los canales correspondientes 206a a 206f en el dominio de la frecuencia. Después de la técnica jerárquica, los canales 206a a 206f se mezclan por pares en tres canales monofónicos L, R y C (208a, 208b y 208c, respectivamente). Durante la mezcla descendente de los tres pares de canales, se deriva un conjunto de parámetros para cada par de canales, que describe las propiedades espaciales de la señal estereofónica original y se mezcla en una señal monofónica. Por lo tanto, en esta primera etapa de mezcla descendente, se generan tres conjuntos de parámetros 210a a 210c para preservar la información espacial de las señales 206a a 206f.

[0031] En el siguiente paso de la mezcla descendente jerárquica, los canales 208a y 208b se mezclan en un canal 212 (LR), lo que genera un conjunto de parámetros 210d (conjunto de parámetros 4). Para finalmente derivar un solo canal monofónico, es necesario realizar una mezcla descendente de los canales 208c y 212, lo que da como resultado el canal 214 (M). Esto genera un quinto conjunto de parámetros 210e (conjunto de parámetros 5). Finalmente, la señal de audio monofónica de mezcla descendente 214 se transforma inversamente en el dominio del tiempo para derivar una señal de audio 202 que puede ser reproducida por un equipo estándar.

[0032] Como se describió anteriormente, una representación paramétrica de la señal de audio de mezcla descendente 202 según la técnica anterior consiste en todos los conjuntos de parámetros 210a a 210e, lo que significa que si uno quiere reconstruir la señal de audio multicanal original (canales 200a a 200f) a partir de la señal de audio monofónica 202, todos los conjuntos de parámetros 210a a 210e se requieren como información lateral de la señal de mezcla monofónica 202.

[0033] La solicitud de patente de EE. UU. 11/032.689 (denominada aquí como "combinación de referencia de la técnica anterior") describe un procedimiento para combinar varios valores de referencia en un único transmitido para guardar información complementaria en un esquema de codificación no jerárquica. Para hacerlo, todos los canales se mezclan primero y los códigos de referencia se combinan posteriormente para formar valores de referencia

transmitidos (también podría ser un solo valor), la combinación depende de una función matemática predefinida, donde los parámetros espaciales, que son derivados directamente de las señales de entrada, se introducen como variables.

[0034] Las técnicas de vanguardia para la codificación paramétrica de dos (“estéreo”) o más (“multicanal”) canales de entrada de audio derivan los parámetros espaciales directamente de las señales de entrada. Ejemplos de dichos parámetros son las diferencias de nivel entre canales (ICLD) o las diferencias de intensidad entre canales (IID), el retardo de tiempo entre canales (ICTD) o las diferencias de fase entre canales (IPD), y la correlación/coherencia entre canales (ICC), donde cada uno de ellos se transmite de manera selectiva en frecuencia, es decir, por banda de frecuencia. La aplicación de la combinación de claves de la técnica anterior enseña que varios valores de claves pueden combinarse con un único valor que se transmite desde el codificador al lado del decodificador. El procedimiento de decodificación utiliza el valor único transmitido en lugar de los valores de referencia originalmente transmitidos individualmente para reconstruir la señal de salida multicanal. En una realización preferida, este esquema se ha aplicado a los parámetros ICC. Se ha demostrado que esto conduce a una reducción considerable en el tamaño de la información del lado de la señal mientras se preserva la calidad espacial de la gran mayoría de las señales. Sin embargo, no está claro cómo se puede explotar esto en un esquema de codificación jerárquica.

[0035] La solicitud de patente sobre combinación de referencia de la técnica anterior ha detallado el principio de la invención mediante un ejemplo para un sistema basado en dos canales de mezcla descendente transmitidos. En el procedimiento propuesto, con referencia a la figura 15, los valores ICC de los pares de canales Lf/Lr y Rf/Rr se combinan en un solo parámetro ICC transmitido. Los dos valores ICC combinados se obtuvieron durante la mezcla descendente de un canal frontal izquierdo Lf y un canal trasero izquierdo Lr en el canal L y durante la mezcla descendente de un canal frontal derecho Rf y un canal trasero derecho Rr en el canal R. Por lo tanto, los dos valores ICC combinados que finalmente se combinan en el único parámetro ICC transmitido, ambos contienen información sobre la correlación frontal/trasero de los canales originales y una combinación de estos dos valores ICC generalmente conservará la mayor parte de esta información. Si uno tuviera que mezclar más a fondo los canales L y R en un solo canal mono, obtendría un tercer valor ICC, con información sobre la correlación izquierda/derecha de los canales L y R de la mezcla descendente. Según la combinación de referencia de la técnica anterior, uno ahora tendría que combinar los tres valores ICC aplicando una función dada, lo que transformaría los tres valores ICC en un parámetro ICC transmitido.

[0036] Uno tiene el problema, entonces, de que la información delantera/trasera se mezcla con la información izquierda/derecha, lo que obviamente es desventajoso para la reproducción de la señal de audio multicanal original. En la solicitud de EE. UU. 11/032.689, esto se evita transmitiendo dos canales de mezcla descendente, los canales L y R, que contienen la información izquierda/derecha, y además transmiten un solo valor ICC, lo que mantiene la información delantera/trasera. Esto preserva las propiedades espaciales de los canales originales al costo de una velocidad de datos sustancialmente mayor, como resultado del canal de mezcla descendente adicional completo que se transmitirá.

[0037] Muchas otras publicaciones y solicitudes de patentes se refieren al tema de la codificación de múltiples canales utilizando representaciones paramétricas de los canales. Por ejemplo, el documento US 2003/0219130 A1 se relaciona con la codificación y síntesis de audio basadas en la coherencia, donde se sintetiza una escena auditiva a partir de una señal de audio mono al modificar, para cada banda crítica, un parámetro de escena auditiva (por ejemplo, una diferencia de nivel entre canales), (ILD) y/o una diferencia de tiempo entre canales (ITD) para cada subbanda dentro de la banda crítica. La modificación se basa en una coherencia media estimada para la banda crítica. La modificación basada en la coherencia produce escenas auditivas con objetos cuyos anchos coinciden con mayor precisión con los anchos de los objetos en la escena auditiva de entrada original.

Resumen de la invención

[0038] El objetivo de la presente invención es proporcionar un concepto mejorado para generar y utilizar una representación paramétrica de una señal de audio multicanal con información lateral compacta en el contexto de un esquema de codificación jerárquica.

[0039] Este objetivo se logra mediante un codificador de la reivindicación 1, un decodificador de la reivindicación 14, un procedimiento para generar de la reivindicación 25, un procedimiento para procesar la reivindicación 26, datos de audio codificados de la reivindicación 27, un medio de almacenamiento legible por ordenador de la reivindicación 28, un receptor de reproductor de audio de la reivindicación 29, un transmisor o grabador de audio de la reivindicación 30, un procedimiento de recepción o reproducción de audio de la reivindicación 31, un procedimiento de transmisión o grabación de audio de la reivindicación 32, un sistema de transmisión de la reivindicación 33, un procedimiento de transmisión o recepción de la reivindicación 34, o un programa informático de la reivindicación 35.

[0040] La presente invención se basa en el hallazgo de que una representación paramétrica de una señal de audio multicanal describe las propiedades espaciales de la señal de audio utilizando una información lateral compacta, cuando la información de coherencia, que describe la coherencia entre un primer y un segundo canal, se deriva dentro

de un procedimiento de codificación jerárquica solo para pares de canales que incluyen un primer canal que tiene solo información de un lado izquierdo con respecto a una posición de escucha e incluye un segundo canal que tiene solo información de un lado derecho con respecto a una posición de escucha. Como en el procedimiento jerárquico los múltiples canales de audio de la señal de audio original se mezclan de manera iterativa y preferentemente en un canal monofónico, uno tiene la oportunidad de elegir los parámetros de información lateral relevantes durante el procedimiento de codificación para un paso que involucra solo pares de canales que llevan la información deseada necesaria para describir lo mejor posible las propiedades espaciales de la señal de audio original. Esto permite construir una representación paramétrica de la señal de audio original sobre la base de esos parámetros seleccionados o en una combinación de esos parámetros, lo que permite una reducción significativa del tamaño de la información lateral, que contiene la información espacial de la señal de mezcla descendente.

[0041] El concepto propuesto permite combinar valores de referencia para reducir la tasa de información lateral de una señal de audio de mezcla descendente incluso en el caso de que solo sea posible un único canal de transmisión (monofónico). El concepto inventivo permite incluso distintas topologías jerárquicas del codificador. Se aclara específicamente cómo se puede derivar un valor ICC único adecuado, que se puede aplicar en un decodificador de audio espacial utilizando el enfoque de codificación/decodificación jerárquica para reproducir fielmente la imagen de sonido original.

[0042] Una realización de la presente invención implementa una estructura de codificación jerárquica que combina el canal de audio frontal izquierdo y el canal trasero izquierdo de una señal de audio de canal 5.1 en un canal maestro izquierdo y que combina simultáneamente el canal frontal derecho y el canal trasero derecho en un canal principal derecho. Al combinar los canales de la izquierda y los canales de la derecha por separado, la importante información de coherencia izquierda/derecha se conserva principalmente y, según la invención, se deriva en la segunda etapa de codificación, donde los canales maestro izquierdo y derecho están mezclados en un estéreo canal maestro. Durante este procedimiento de mezcla descendente, se deriva el parámetro ICC para todo el sistema, ya que este parámetro ICC será el parámetro ICC que se asemeja con mayor precisión a la coherencia izquierda/derecha. Dentro de esta realización de la presente invención, se obtiene un parámetro ICC, que describe la coherencia izquierda/derecha más importante de los seis canales de audio simplemente organizando los pasos de codificación jerárquica de una manera apropiada y no aplicando alguna función artificial a un conjunto de parámetros ICC, que describen pares arbitrarios de canales, como es el caso en las técnicas de la técnica anterior.

[0043] En una modificación de la realización descrita de la presente invención, el canal central y el canal de baja frecuencia de la señal de audio 5.1 se mezclan en un canal maestro central, este canal contiene principalmente información sobre el canal central, ya que el canal de baja frecuencia solo contiene señales con una frecuencia tan baja que los seres humanos difícilmente pueden localizar el origen de las señales. Puede ser ventajoso dirigir adicionalmente el valor ICC, derivado como se describe anteriormente, por los parámetros que describen el canal maestro central. Esto se puede hacer, por ejemplo, ponderando el valor ICC con información de energía, la información de energía que indica la cantidad de energía que se transmite a través del canal maestro central con respecto al canal maestro estéreo.

[0044] En una realización adicional de la presente invención, el procedimiento de codificación jerárquica se realiza de manera tal que, en un primer paso, los canales frontal izquierdo y frontal derecho de una señal de audio 5.1 se mezclan en un canal principal frontal, mientras que los canales traseros izquierdo y los canales traseros derechos se mezclan en un canal maestro trasero. Por lo tanto, en cada uno de los procedimientos de mezcla descendente se genera un valor ICC, que contiene información sobre la importante coherencia izquierda/derecha. El parámetro ICC combinado y transmitido a continuación se deriva de una combinación de los dos valores ICC separados, una forma ventajosa de derivar el parámetro ICC transmitido es construir la suma ponderada de los valores ICC, utilizando los parámetros de nivel de los canales como ponderaciones.

[0045] En una modificación de la invención, el canal central y el canal de baja frecuencia se mezclan en un canal maestro central y a continuación el canal maestro central y el canal maestro frontal se mezclan en un canal maestro estéreo. En el último procedimiento de mezcla descendente, se recibe una correlación entre el centro y los canales estéreo, que se utiliza para dirigir o modificar un parámetro ICC transmitido, donde también se tiene en cuenta la contribución del centro a la señal de audio frontal. Una ventaja importante del sistema descrito anteriormente es que se puede construir la información de coherencia de modo que los canales que contribuyen más a la señal de audio definen principalmente el valor ICC transmitido. Normalmente, estos serán los canales frontales, pero, por ejemplo, en una representación multicanal de un concierto de música, la señal de la audiencia aplaudida podría enfatizarse utilizando principalmente el valor ICC de los canales traseros. Otra ventaja es que la ponderación entre los canales frontal y trasero puede variar dinámicamente según las propiedades espaciales de la señal de audio multicanal.

[0046] En una realización de la presente invención, un decodificador jerárquico de la invención está operativo para recibir menos parámetros ICC de los requeridos por el número de etapas de decodificación existentes. El decodificador está operativo para derivar los parámetros ICC necesarios para cada paso de decodificación a partir de los parámetros ICC recibidos.

[0047] Esto podría hacerse derivando los parámetros ICC adicionales usando una regla derivada que se basa en los parámetros ICC recibidos y los valores ICLD recibidos o usando valores predefinidos en su lugar.

[0048] En una realización preferida, sin embargo, el decodificador está operativo para usar un único parámetro ICC transmitido para cada paso de decodificación individual. Esto es ventajoso ya que la correlación más importante, la correlación izquierda/derecha, se conserva en un parámetro ICC transmitido dentro del concepto inventivo. Como este es el caso, un oyente experimentará una reproducción de la señal que se asemeja muy bien a la señal original. Debe recordarse que el parámetro ICC define la amplitud perceptiva de una señal reconstruida. Si el decodificador modificara un parámetro ICC transmitido después de la transmisión, los parámetros ICC que describen la amplitud perceptiva de la señal reconstruida pueden ser bastante distintos para la correlación izquierda/derecha y la frontal/posterior dentro de la reproducción jerárquica. Esto sería lo más desventajoso, ya que un oyente que mueva o gire su cabeza experimentará una señal que se hará perceptualmente más ancha o más estrecha, lo que por supuesto es lo más molesto. Esto se puede evitar distribuyendo un único parámetro ICC recibido a las unidades de decodificación de un decodificador jerárquico.

[0049] En otra realización preferida, un decodificador de la invención está operativo para recibir un conjunto completo de valores ICC o, alternativamente, un solo valor ICC, donde el decodificador reconoce la estrategia de decodificación que se aplica al recibir una indicación de estrategia dentro del flujo de bits. Tal decodificador compatible con versiones anteriores también está operativo en entornos de la técnica anterior, donde decodifica señales de la técnica anterior que transmiten un conjunto completo de datos ICC.

Breve descripción de los dibujos

[0050] Las realizaciones preferidas de la presente invención se describen posteriormente haciendo referencia a los dibujos adjuntos, donde:

la figura 1 muestra un diagrama de bloques de una realización del codificador de audio jerárquico de la invención;

la figura 2 muestra una realización de un codificador de audio de la invención;

la figura 2a muestra un posible esquema de dirección de los parámetros IIC de un codificador de audio de la invención;

la figura 3a, b muestra representaciones gráficas de información de canal lateral;

la figura 4 muestra una segunda realización de un codificador de audio de la invención;

la figura 5 muestra un diagrama de bloques de una realización preferida de un decodificador de audio de la invención;

la figura 6 muestra una realización de un decodificador de audio de la invención;

la figura 7 muestra otra realización de un decodificador de audio de la invención;

la figura 8 muestra un transmisor o grabador de audio de la invención;

la figura 9 muestra un receptor o reproductor de audio de la invención;

la figura 10 muestra un sistema de transmisión de la invención;

la figura 11 muestra un codificador estéreo conjunto de la técnica anterior;

la figura 12 muestra una representación en diagrama de bloques de una cadena de codificador/decodificador BCC de la técnica anterior;

la figura 13 muestra un diagrama de bloques de una implementación de la técnica anterior de un bloque de síntesis de BCC;

la figura 14 muestra una representación de un esquema para determinar los parámetros de BCC; y

la figura 15 muestra un codificador jerárquico de la técnica anterior.

Descripción detallada de las realizaciones preferidas

[0051] La figura 1 muestra un diagrama de bloques de un codificador de la invención para generar una representación paramétrica de una señal de audio. La figura 1 muestra un generador 220 para combinar posteriormente los canales de audio y generar parámetros espaciales que describen las propiedades espaciales de

pares de canales que se combinan en un solo canal. La figura 1 muestra además un proveedor 222 para proporcionar una representación paramétrica de una señal de audio multicanal seleccionando información de diferencia de nivel entre pares de canales y determinando una medida de coherencia izquierda/derecha utilizando información de coherencia generada por el generador 220.

5

[0052] Para demostrar el principio del concepto inventivo de codificación de audio multicanal jerárquica, la figura 1 muestra un caso, donde cuatro canales de audio originales 224a a 224d se combinan de forma iterativa y dan como resultado un solo canal 226. Los canales de audio originales 224a y 224b representan los canales frontal izquierdo y trasero izquierdo de una señal de audio original de cuatro canales, los canales 224c y 224d representan los canales frontal derecho y trasero derecho, respectivamente. Sin pérdida de generalidad, solo dos de los diversos parámetros espaciales se muestran en la figura 1 (ICLD e ICC). Según la invención, el generador 220 combina los canales de audio 224a a 224d de tal manera que durante el procedimiento de combinación se puede derivar un parámetro ICC que lleva la información importante de coherencia izquierda/derecha.

10

[0053] En un primer paso, los canales que contienen solo información del lado izquierdo 224a y 224b se combinan en un canal maestro izquierdo 228a (L) y los dos canales que contienen solo información del lado derecho 224c y 224d se combinan en un canal maestro derecho 228b (R). Durante esta combinación, el generador genera dos parámetros ICLD 230a y 230b, ambos parámetros espaciales que contienen información sobre la diferencia de nivel de dos canales originales que se combinan en un solo canal. El generador también genera dos parámetros ICC 232a y 232b, que describen la correlación entre los dos canales que se combinan en un solo canal. Los parámetros ICLD e ICC 230a, 230b, 232a y 232b se transfieren al proveedor 222.

15

[0054] En la siguiente etapa del procedimiento de generación jerárquica, el canal maestro izquierdo 228a se combina con el canal maestro derecho 228b en el canal de audio resultante 226, donde el generador proporciona un parámetro ICLD 234 y un parámetro ICC 236, ambos transmitiéndose al proveedor 222. Es importante tener en cuenta que el parámetro ICC 236 generado en este paso de combinación representa principalmente la importante información de coherencia izquierda/derecha de la señal de audio de cuatro canales original representada por los canales de audio 224a a 224d.

25

[0055] Por lo tanto, el proveedor 222 construye una representación paramétrica 238 a partir de los parámetros espaciales disponibles 230a,b, 232a,b, 234 y 236, de modo que la representación paramétrica comprende los parámetros 230a, 230b, 234 y 236.

30

[0056] La figura 2 muestra una realización preferida de un codificador de audio de la invención que codifica una señal multicanal 5.1 en una única señal monofónica.

35

[0057] La figura 2 muestra tres unidades de transformación 240a a 240c, cinco mezcladores descendentes 2 a 1 242a a 242e, una unidad de combinación de parámetros 244 y una unidad de transformación inversa 246. La señal de audio de canal 5.1 original viene dada por el canal frontal izquierdo 248a, el canal trasero izquierdo 248b, el canal frontal derecho 248c, el canal trasero derecho 248d, el canal central 248e y el canal de baja frecuencia 248f. Es importante tener en cuenta que los canales originales están agrupados de tal manera que los canales que contienen solo información del lado izquierdo 248a y 248b forman un par de canales, los canales que contienen solo información del lado derecho 248c y 248d forman otro par de canales y que el canal central 248e y 248f forman un tercer par de canales.

40

45

[0058] Las unidades de transformación 240a a 240c convierten los canales 248a a 248f del dominio de tiempo en su representación espectral 250a a 250f en el dominio de subbanda de frecuencia. En la primera etapa de codificación jerárquica 252, los canales izquierdos 250a y 250b están codificados en un canal maestro izquierdo 254a, los canales derechos 250c y 250d están codificados en un canal maestro derecho 254b y el canal central 250e y el canal de baja frecuencia 250f están codificados en un canal maestro central 256. Durante esta primera etapa de codificación jerárquica 252, los tres codificadores 2 a 1 242a a 242c involucrados generan los canales de mezcla descendente 254a, 254b y 256, y además los conjuntos de parámetros espaciales importantes 260a, 260b y 260c, donde el conjunto de parámetros 260a (conjunto de parámetros 1) describe la información espacial entre los canales 250a y 250b, el conjunto de parámetros 260b (conjunto de parámetros 2) describe la relación espacial entre los canales 250c y 250d y el conjunto de parámetros 260c (conjunto de parámetros 3) describe la relación espacial entre los canales 250e y 250f.

50

55

[0059] En un segundo paso jerárquico 262, el canal maestro izquierdo 254a y el canal maestro derecho 254b se mezclan en un canal maestro estéreo 264 y generan un conjunto de parámetros espaciales 266 (conjunto de parámetros 4), donde el parámetro ICC de este conjunto de parámetros 266 contiene la información importante de correlación izquierda/derecha. Para construir un valor ICC combinado a partir del conjunto de parámetros 266, el conjunto de parámetros 266 se puede transferir a la unidad de combinación de parámetros 244 a través de una conexión de datos 268. En la tercera etapa de codificación jerárquica 272, el canal maestro estéreo 264 se combina con el canal maestro central 256 para formar un canal de resultado monofónico 274. El conjunto de parámetros 276, que se deriva durante este procedimiento de mezcla descendente, puede transferirse a través de una conexión de

60

65

datos 278 a la unidad de combinación de parámetros 244. Finalmente, el canal de resultado 274 se transforma en el dominio de tiempo mediante la unidad de transformación inversa 246, para construir la señal de audio de mezcla descendente monofónica 280, que es la representación monofónica final de la señal de canal 5.1 original representada por los canales de audio 248a a 248f.

5

[0060] Para reconstruir la señal de audio de canal 5.1 original del canal de audio de mezcla descendente monofónica 280, se necesita además la representación paramétrica de la señal de audio de canal 5.1. Mediante la estructura de árbol que se muestra en la figura 2, se puede ver que los canales frontales y traseros izquierdos se combinan en una señal L 254a. De manera similar, los canales frontales y traseros derechos se combinan en una
10 señal R 254b. Posteriormente, se lleva a cabo la combinación de las señales L y R, que entrega el conjunto de parámetros número 4 (266). En el caso de esta estructura jerárquica, una forma simple de obtener un valor ICC combinado es elegir el valor ICC del conjunto de parámetros número 4 y tomarlo como un valor ICC combinado, que a continuación se incorpora a la representación paramétrica de la señal de canal 5.1 mediante la unidad de combinación de parámetros 244. Los procedimientos más sofisticados también pueden tener en cuenta la influencia
15 del canal central (por ejemplo, mediante el uso de parámetros del conjunto de parámetros número 5), como se muestra en la figura 2a.

[0061] Como ejemplo, la relación de energía $E(LR)/E(C)$ de la energía contenida en el canal LR (264) y en el canal C (256) del conjunto de parámetros número 5 se puede utilizar para dirigir el ICC de valor. En caso de que la
20 mayor parte de la energía provenga de la ruta LR, el valor ICC transmitido debe estar cerca del valor ICC (LR) del conjunto de parámetros número 4. En el caso de que la mayor parte de la energía provenga de la ruta C 256, el valor ICC transmitido debería ser posteriormente cercano a 1, como se indica en la figura 2a. La figura muestra dos formas posibles de implementar esta dirección del parámetro ICC, ya sea cambiando entre dos valores extremos cuando la relación de energía cruza un umbral dado 286 (función de dirección 288a) o mediante una transición suave entre los
25 valores extremos (función de dirección 288b).

[0062] Las figuras 3a y 3b muestran una comparación de una posible representación paramétrica de un canal de audio 5.1 entregado desde una estructura de codificador jerárquico mediante una técnica de la técnica anterior (figura 3a) y mediante el concepto de la invención para la codificación de audio (figura 3b).
30

[0063] La figura 3a muestra una representación paramétrica de un marco de tiempo único y un intervalo de frecuencia discreto, tal como sería proporcionado por la técnica de la técnica anterior. Cada uno de los codificadores 2 a 1 242a a 242e de la figura 2 entrega un par de parámetros ICLD e ICC, el origen de los pares de parámetros se indica dentro de la figura 3a. Después de la técnica de la técnica anterior, todos los conjuntos de parámetros
35 proporcionados por los codificadores 2 a 1 242a a 242e deben transmitirse junto con la señal de audio monofónica de mezcla descendente 280 como información lateral para reconstruir una señal de audio de canal 5.1.

[0064] La figura 3b muestra los parámetros derivados según el concepto inventivo. Cada uno de los codificadores 2 a 1 242a a 242e contribuye solo con un parámetro directamente, el parámetro ICLD. El parámetro ICC transmitido único ICC_c se deriva de la unidad de combinación de parámetros 244, y no es proporcionado directamente por los codificadores 2 a 1 242a a 242e. Como se ve claramente en las figuras 3a y 3b, el concepto de la invención para un codificador jerárquico puede reducir la cantidad de datos de información lateral de manera importante en comparación con las técnicas de la técnica anterior.
40

[0065] La figura 4 muestra otra realización preferida de la presente invención, que permite codificar una señal de audio de canal 5.1 en una señal de audio monofónica en un procedimiento de codificación jerárquica y suministrar información lateral compacta. Como la estructura de hardware principal es igual a la descrita en la figura 2, los mismos elementos en las dos figuras están etiquetados con los mismos números. La diferencia se debe a la diferente agrupación de los canales de entrada 248a a 248f y, por lo tanto, el orden donde los canales individuales se mezclan
50 en el canal monofónico 274 difiere del orden de mezcla descendente en la figura 2. Por lo tanto, solo los aspectos que difieren de la descripción de la figura 2, que son vitales para la comprensión de la realización de la presente invención mostrada en la figura 4, se describen a continuación.

[0066] El canal frontal izquierdo 248a y el canal frontal derecho 248c se agrupan para formar un par de canales, el canal central 248e y el canal de baja frecuencia 248f forman otro par de canales de entrada y el tercer par de canales de entrada de la señal de audio 5.1 está formado por el canal trasero izquierdo 248b y el canal derecho trasero 248d.
55

[0067] En una primera etapa de codificación jerárquica 252, el canal frontal izquierdo 250a y el canal frontal derecho 250c se mezclan en un canal maestro frontal 290 (F), el canal central 250e y el canal de baja frecuencia 250f se mezclan en un canal maestro central 292 (C) y el canal trasero izquierdo 250b y el canal trasero derecho 250d se mezclan en un canal maestro trasero 294 (S). Un conjunto de parámetros 300a (conjunto de parámetros 1) describe el canal maestro frontal 290, un conjunto de parámetros 300b (conjunto de parámetros 2) describe el canal maestro central 292, y un conjunto de parámetros 300c (conjunto de parámetros 3) describe el canal maestro trasero 294.
60

[0068] Es importante tener en cuenta que el conjunto de parámetros 300a, así como el conjunto de parámetros
65

300c, contienen información que describe la importante correlación izquierda/derecha entre los canales originales 248a a 248f. Por lo tanto, el conjunto de parámetros 300a y el conjunto de parámetros 300c se ponen a disposición de la unidad de combinación de parámetros 244 a través de los enlaces de datos 302a y 302b.

5 **[0069]** En una segunda etapa de codificación 262, el canal maestro frontal 290 y el canal maestro central 292 se mezclan en un canal frontal puro 304, lo que genera un conjunto de parámetros 300d (conjunto de parámetros 4). Este conjunto de parámetros 300d también está disponible para la unidad de combinación de parámetros 244 a través de un enlace de datos 306.

10 **[0070]** En una tercera etapa de codificación jerárquica 272, el canal frontal puro 304 se mezcla en sentido descendente con el canal maestro trasero 294 en el canal de resultados 274 (M), que a continuación se transforma en el dominio del tiempo por la unidad de transformación inversa 246 para formar el canal de audio de mezcla descendente monofónico final 280. El conjunto de parámetros 300e (conjunto de parámetros 5), originado por la mezcla descendente del canal frontal puro 304 y el canal maestro trasero 294, también está disponible para la unidad
15 de combinación de parámetros 244 a través de un enlace de datos 310.

[0071] La estructura de árbol en la figura 4 primero realiza una combinación de los canales izquierdo y derecho por separado para los frontales y traseros. Por lo tanto, la correlación/coherencia izquierda/derecha básica está presente en los conjuntos de parámetros 1 y 3 (300a, 300c). La unidad de combinación de parámetros 244 podría
20 construir un valor ICC combinado construyendo el promedio ponderado entre los valores ICC de los conjuntos de parámetros 1 y 3. Esto significa que se dará más peso a los pares de canales más fuertes (Lf/Rf frente a Lr/Rr). Uno puede lograr lo mismo derivando un parámetro ICC ICC_c combinado que construye la suma ponderada:

$$ICC_c = (A \cdot ICC_1 + B \cdot ICC_2) / (A + B)$$

25 donde A denota la energía dentro del par de canales correspondientes a ICC_1 y B denota la energía dentro del par de canales correspondientes a ICC_2 .

[0072] En una realización alternativa, los procedimientos más sofisticados también pueden tener en cuenta la
30 influencia del canal central (por ejemplo, teniendo en cuenta los parámetros del conjunto de parámetros número 4).

[0073] La figura 5 muestra un decodificador de la invención, para procesar información lateral compacta recibida, que es una representación paramétrica de una señal de audio original de cuatro canales. La figura 5 comprende un receptor 310 para proporcionar una representación paramétrica compacta de la señal de audio de
35 cuatro canales y un procesador 312 para procesar la representación paramétrica compacta tal que se suministra una representación paramétrica completa de la señal de audio de cuatro canales, que permite reconstruir la señal de audio de cuatro canales a partir de una señal de audio monofónica recibida.

[0074] El receptor 310 recibe los parámetros espaciales ICLD (B) 314, ICLD (F) 316, ICLD (R) 318 y ICC 320.
40 La representación paramétrica proporcionada, que consta de los parámetros 314 a 320, describe las propiedades espaciales de los canales de audio originales 324a a 324d.

[0075] Como primera etapa de mezcla ascendente, el procesador 312 suministra los parámetros espaciales que describen un primer par de canales 326a, que es una combinación de dos canales 324a y 324b (Rf y Lf), y un
45 segundo par de canales 326b, que es una combinación de dos canales 324c y 324d (Rr y Lr). Para ello, se requiere la diferencia de nivel 314 de los pares de canales. Dado que ambos pares de canales 326a y 326b contienen un canal izquierdo y un canal derecho, la diferencia entre los pares de canales describe principalmente una correlación delantera/trasera. Por lo tanto, el parámetro ICC 320, que contiene principalmente información sobre la coherencia izquierda/derecha, es proporcionado por el procesador 312 de manera que la información de coherencia
50 izquierda/derecha se utiliza preferentemente para suministrar los parámetros ICC individuales para los pares de canales 326a y 326b.

[0076] En el siguiente paso, el procesador 312 suministra los parámetros espaciales apropiados para poder reconstruir los canales de audio individuales 324a y 324b desde el canal 326a, y los canales 324c y 324d desde el
55 canal 326b. Para hacerlo, el procesador 312 suministra las diferencias de nivel 316 y 318, y el procesador 312 tiene que proporcionar valores ICC apropiados para los dos pares de canales, ya que cada uno de los pares de canales 326a y 326b contiene información importante de coherencia izquierda/derecha.

[0077] En un ejemplo, el procesador 312 podría simplemente proporcionar el valor ICC recibido combinado 320
60 a los pares de canales de mezcla ascendente 326a y 326b. Alternativamente, el valor ICC recibido combinado 320 podría ponderarse para obtener valores ICC individuales para los dos pares de canales, las ponderaciones se basan, por ejemplo, en la diferencia de nivel 314 de los dos pares de canales.

[0078] En una realización preferida de la presente invención, el procesador proporciona el parámetro ICC 320
65 recibido por cada etapa de mezcla ascendente individual para evitar la introducción de artefactos adicionales durante

la reproducción de los canales 324a a 324d.

[0079] La figura 6 muestra una realización preferida de un decodificador que incorpora un procedimiento de decodificación jerárquica según la presente invención, para decodificar una señal de audio monofónica a una señal de audio multicanal 5.1. Hace uso de una representación paramétrica compacta de una señal de audio 5.1 original.

[0080] La figura 6 muestra una unidad de transformación 350, una unidad de procesamiento de parámetros 352, cinco decodificadores 1 a 2 354a a 354e y tres unidades de transformación inversa 356a a 356c.

10 **[0081]** Debe observarse que la realización de un decodificador de la invención según la figura 6 es la contrapartida del codificador descrito en la figura 2 y está diseñado para recibir un canal de audio de mezcla descendente monofónica 358, que finalmente se mezclará en una señal de audio 5.1 que consta de los canales de audio 360a (lf), 360b (lr), 360c (rf), 360d (rr), 360e (co) y 360f (lfe). El canal de mezcla descendente 358 (m) se recibe y se transforma desde el dominio de tiempo al dominio de frecuencia en su representación de frecuencia 362 utilizando la unidad de transformación 350. La unidad de procesamiento de parámetros 352 recibe un conjunto combinado y compacto de parámetros espaciales 364 en paralelo con el canal de mezcla descendente 358.

20 **[0082]** En un primer paso 363 del procedimiento de decodificación jerárquica, el canal de mezcla descendente monofónico 362 se mezcla en un canal maestro estéreo 364 (LR) y un canal maestro central 366 (C).

[0083] En un segundo paso 368 del procedimiento de decodificación jerárquica, el canal maestro estéreo 364 se mezcla en un canal maestro izquierdo 370 (L) y un canal maestro derecho 372 (R).

25 **[0084]** En un tercer paso del procedimiento de decodificación, el canal maestro izquierdo 370 se mezcla de forma ascendente en un canal frontal izquierdo 374a y un canal trasero izquierdo 374b, el canal maestro derecho 372 se mezcla de forma ascendente en un canal frontal derecho 374c y el canal trasero derecho 374d, y el canal maestro central 366 se mezclan en un canal central 374e y un canal de baja frecuencia 374f.

30 **[0085]** Finalmente, los seis canales de audio únicos 374a a 374f son transformados por las unidades de transformación inversa 356a a 356c en su representación en el dominio del tiempo y, por lo tanto, construyen la señal de audio 5.1 reconstruida, que tiene seis canales de audio 360a a 360f. Para conservar la propiedad espacial original de la señal de audio 5.1, la unidad de procesamiento de parámetros 352, especialmente la forma en que la unidad de procesamiento de parámetros proporciona los conjuntos de parámetros individuales 380a a 380e, es vital, especialmente la forma en que la unidad de procesamiento de parámetros 352 deriva los conjuntos de parámetros individuales 380a a 380e.

40 **[0086]** El parámetro ICC combinado recibido describe la importante coherencia izquierda/derecha de la señal de audio de seis canales original. Por lo tanto, la unidad de procesamiento de parámetros 352 construye el valor ICC del conjunto de parámetros 4 (380d) de manera que se asemeja a la información de correlación izquierda/derecha del valor espacial recibido originalmente y se transmite dentro del conjunto de parámetros 364. En la implementación más simple posible, la unidad de procesamiento de parámetros 352 simplemente utiliza el parámetro ICC combinado recibido.

45 **[0087]** Otra realización preferida de un decodificador según la presente invención se muestra en la figura 7, donde el decodificador en la figura 7 es la contrapartida del codificador de la figura 4.

50 **[0088]** Como el codificador en la figura 7 comprende los mismos bloques funcionales que el decodificador en la figura 6, la siguiente explicación se limita a los pasos en los cuales el procedimiento de decodificación jerárquica difiere del de la figura 6. Esto se debe principalmente al hecho de que la señal monofónica 362 se mezcla de forma ascendente en un orden distinto y una combinación de canales distinta, ya que la señal de audio 5.1 original se había mezclado de manera distinta a la recibida en la figura 6.

55 **[0089]** En la primera etapa 363 del procedimiento de decodificación jerárquica, la señal monofónica 362 se mezcla de forma ascendente en un canal maestro trasero 400 (S) y un canal frontal puro 402 (CF).

[0090] En un segundo paso 368, el canal frontal puro 402 se mezcla de forma ascendente en un canal maestro frontal 404 y un canal maestro central 406.

60 **[0091]** En una tercera etapa de decodificación 372, el canal maestro frontal se mezcla de forma ascendente en un canal frontal izquierdo 374a y un canal frontal derecho 374c, el canal maestro central 406 se mezcla de forma ascendente en un canal central 374e, y un canal de baja frecuencia 374f y el canal maestro trasero 400 se mezclan de forma ascendente en un canal trasero izquierdo 374b y un canal trasero derecho 374d. Finalmente, los seis canales de audio 374a a 374f se transforman desde el dominio de la frecuencia a sus representaciones de dominio de tiempo 360a a 360f y así construyen la señal de audio 5.1 reconstruida.

65

- [0092]** Para preservar las propiedades espaciales de la señal 5.1 original, que ha sido codificada como información lateral por el codificador, la unidad de procesamiento de parámetros 352 suministra los conjuntos de parámetros 410a a 410e para los decodificadores 1 a 2 354a a 354e. Como la información importante de correlación izquierda/derecha es necesaria en el tercer procedimiento de mezcla ascendente 372 para construir los canales Lf, Rf, Lr y Rr, la unidad de procesamiento de parámetros 352 puede suministrar un valor ICC apropiado en los conjuntos de parámetros 410a y 410c en la implementación más sencilla simplemente al tomar el parámetro ICC transmitido para construir los conjuntos de parámetros 410a y 410c. En una posible alternativa, el parámetro ICC recibido podría transformarse en parámetros individuales para los conjuntos de parámetros 410a y 410c al aplicar una función de ponderación adecuada al parámetro ICC recibido, donde su peso, por ejemplo, es dependiente de la energía transmitida en el canal maestro frontal 404 y en el canal maestro trasero 400. En una implementación aún más sofisticada, la unidad de procesamiento de parámetros 352 también podría tener en cuenta la información del canal central para suministrar un valor ICC individual para el conjunto de parámetros 5 y el conjunto de parámetros 4 (410a, 410b).
- 15 **[0093]** La figura 8 muestra un transmisor o grabador de audio de la invención 500 que tiene un codificador 220, una interfaz de entrada 502 y una interfaz de salida 504.
- [0094]** Se puede suministrar una señal de audio en la interfaz de entrada 502 del transmisor/grabador 500. La señal de audio se codifica utilizando un codificador 220 de la invención dentro del transmisor/grabador y la representación codificada se emite en la interfaz de salida 504 del transmisor/grabador 500. La representación codificada puede ser transmitida o almacenada en un medio de almacenamiento.
- 20 **[0095]** La figura 9 muestra un receptor o reproductor de audio de la invención 520, que tiene un decodificador de la invención 312, una entrada de flujo de bits 522 y una salida de audio 524.
- 25 **[0096]** Se puede ingresar un flujo de bits en la entrada 522 del receptor/reproductor de audio 520 de la invención. El flujo de bits se decodifica a continuación mediante el decodificador 312 y la señal decodificada se emite o reproduce en la salida 524 del receptor/reproductor de audio 520 de la invención.
- 30 **[0097]** La figura 10 muestra un sistema de transmisión que comprende un transmisor 500 de la invención y un receptor 520 de la invención.
- [0098]** La entrada de señal de audio en la interfaz de entrada 502 del transmisor 500 se codifica y se transfiere desde la salida 504 del transmisor 500 a la entrada 522 del receptor 520. El receptor decodifica la señal de audio y reproduce o emite la señal de audio en su salida 524.
- 35 **[0099]** Los ejemplos expuestos de los decodificadores de la invención mezclan de forma descendente una señal de audio multicanal en una señal de audio monofónica. Por supuesto, es posible alternativamente mezclar una señal multicanal en una señal estereofónica, lo que por ejemplo significaría que para las realizaciones discutidas en las figuras 2 y 4, ese paso en el procedimiento de codificación jerárquica podría ser pasado por alto. Todos los demás números de canales resultantes también son posibles.
- 40 **[0100]** El procedimiento propuesto para codificar o decodificar jerárquicamente la información de audio multicanal que proporciona/utiliza una representación paramétrica compacta de las propiedades espaciales de la señal de audio se describe principalmente al reducir la información lateral al combinar múltiples valores ICC en un solo valor ICC transmitido. Cabe señalar aquí que la invención descrita no se limita de ninguna manera al uso de un solo valor ICC combinado. En su lugar, por ejemplo, se pueden generar dos valores combinados, uno que describe la importante correlación izquierda/derecha y el otro que describe una correlación frontal/posterior.
- 45 **[0101]** Esto puede implementarse ventajosamente, por ejemplo, en la realización de la presente invención mostrada en la figura 2, donde, por un lado, un canal frontal izquierdo 250a y un canal trasero izquierdo 250b se combinan en un canal maestro izquierdo 254a, y donde un canal frontal derecho 250c y un canal trasero derecho 250d se combinan en un canal maestro trasero 254b. Por lo tanto, estos dos pasos de codificación brindan información sobre la correlación de la parte frontal anterior de la señal de audio original, que puede procesarse fácilmente para proporcionar un valor ICC adicional y mantiene la información de correlación frontal/posterior.
- 50 **[0102]** Además, en una modificación preferida de la presente invención, es ventajoso tener procedimientos de codificación/decodificación, que puedan tanto usar los parámetros transmitidos individualmente de la técnica anterior como, dependiendo de la información lateral de señalización que se envía desde el codificador para decodificar, también utilizar parámetros transmitidos combinados. Dicho sistema puede lograr ventajosamente una mayor precisión de representación (utilizando parámetros transmitidos individualmente) y, alternativamente, una baja velocidad de bits de información lateral (utilizando parámetros combinados).
- 60 **[0103]** Normalmente, la elección de este ajuste es realizada por el usuario según los requisitos de la aplicación, como la cantidad de información complementaria que puede acomodar el sistema de transmisión utilizado. Esto

permite utilizar la misma arquitectura unificada de codificador/decodificador mientras se puede operar dentro de una amplia gama de compensaciones de precisión/velocidad de bits de información lateral. Esta es una capacidad importante para cubrir una amplia gama de aplicaciones posibles con distintos requisitos y capacidad de transmisión.

5 **[0104]** En otra modificación de una realización tan ventajosa, la elección del modo de funcionamiento también podría hacerse automáticamente por el codificador, que analiza, por ejemplo, la desviación de los valores decodificados del resultado ideal en caso de que se usara el modo de transmisión combinado. Si no se encuentra una desviación importante, se emplea la transmisión combinada de parámetros. Un decodificador incluso podría decidir por sí mismo, según un análisis de la información complementaria proporcionada, qué modo es el apropiado para
10 usar. Por ejemplo, si solo se proporcionara un parámetro espacial, el decodificador cambiaría automáticamente al modo de decodificación utilizando parámetros transmitidos combinados.

[0105] En otra modificación ventajosa de la presente invención, el codificador/decodificador cambia automáticamente del modo que utiliza parámetros transmitidos combinados al modo que utiliza parámetros
15 transmitidos individualmente, para asegurar el mejor compromiso posible entre una calidad de reproducción de audio y una velocidad de bits de información lateral baja deseada.

[0106] Como se puede ver en las realizaciones preferidas descritas de los codificadores/decodificadores en las figuras 2, 4, 6 y 7, estas unidades hacen uso de los mismos bloques funcionales. Por lo tanto, otra realización preferida
20 construye un codificador y un decodificador mediante el mismo hardware dentro de una carcasa.

[0107] En una realización alternativa de la presente invención, es posible cambiar dinámicamente entre los distintos esquemas de codificación agrupando distintos canales como pares de canales, lo que hace posible usar dinámicamente el esquema de codificación que proporciona la mejor calidad de audio posible para la señal de audio
25 multicanal dada.

[0108] No es necesario transmitir el canal de mezcla descendente monofónico junto con la representación paramétrica de una señal de audio multicanal. También es posible transmitir la representación paramétrica sola, para permitir que un oyente, que ya posea una mezcla descendente monofónica de la señal de audio multicanal, por ejemplo
30 como un registro, reproduzca una señal multicanal mediante su equipo multicanal existente y una información lateral paramétrica.

[0109] Para resumir, la presente invención permite determinar estos parámetros combinados ventajosamente a partir de parámetros conocidos de la técnica anterior. Al aplicar el concepto inventivo de combinar parámetros en
35 una estructura codificador/decodificador jerárquica, se puede mezclar una señal de audio multicanal en una representación paramétrica basada en mono, por lo que se obtiene una parametrización precisa de la señal original a una velocidad de información lateral baja (= reducción de la velocidad de bits).

[0110] Un objetivo de la presente invención es que el codificador combine ciertos parámetros con el objetivo
40 de reducir el número de parámetros que deben transmitirse. A continuación, el decodificador deriva los parámetros que faltan de los parámetros que se han transmitido, en lugar de usar valores de parámetros predeterminados, como es el caso en los sistemas de la técnica anterior, por ejemplo, el que se muestra en la figura 15.

[0111] Esta ventaja se vuelve evidente al revisar nuevamente la realización de un codificador de audio
45 multicanal paramétrico jerárquico mediante técnicas de la técnica anterior, un ejemplo que se muestra en la figura 15. Allí, las señales de entrada (Lf, Rf, Lr, Rr, C y LFE, correspondientes a los canales de mejora frontal izquierdo, frontal derecho, trasero izquierdo, trasero derecho, central y de baja frecuencia, respectivamente) se segmentan y se transforman en el dominio de la frecuencia para obtener los mosaicos de tiempo/frecuencia requeridos. Las señales resultantes se combinan posteriormente de forma pareada. Por ejemplo, las señales Lf y Lr se combinan para formar
50 la señal "L". Se genera un conjunto de parámetros espaciales correspondientes (1) para modelar las propiedades espaciales entre las señales Lf y Lr (es decir, que consiste en uno o más de IID, ICC, IPD). En la realización según la técnica anterior mostrada en la figura 15, este procedimiento se repite hasta que se obtiene un solo canal de salida (M), el canal de salida está acompañado por cinco conjuntos de parámetros. La aplicación de técnicas de codificación jerárquica de la técnica anterior implicaría entonces la transmisión de todos los conjuntos de parámetros.

[0112] Debe observarse, sin embargo, que no todos los conjuntos de parámetros deben contener valores para
55 todos los posibles parámetros espaciales. Por ejemplo, el conjunto de parámetros 1 en la figura 15 puede consistir en parámetros IID e ICC, mientras que el conjunto de parámetros 3 puede consistir solo en parámetros IDD. Si ciertos parámetros no se transmiten para conjuntos específicos, el decodificador jerárquico de la técnica anterior aplicará un valor predeterminado para estos parámetros (por ejemplo, ICC = + 1, IPD = 0, etc.). Por lo tanto, cada conjunto de
60 parámetros representa solo una combinación de señal específica y no describe las propiedades espaciales de los pares de canales restantes.

[0113] Esta pérdida de conocimiento sobre las propiedades espaciales de las señales, cuyos parámetros no
se están transmitiendo, se puede evitar mediante el concepto de la invención, donde el codificador combina parámetros

específicos de manera que se conservan las propiedades espaciales más importantes de la señal original.

- 5 **[0114]** Cuando, por ejemplo, los parámetros ICC se combinan en un solo valor, los parámetros combinados pueden usarse en el decodificador como un sustituto de todos los parámetros individuales (o el parámetro individual usado en el decodificador puede derivarse de los transmitidos). Es una característica importante que el procedimiento de combinación de parámetros del codificador se lleve a cabo de manera tal que la imagen de sonido de la señal multicanal original se conserve lo más cerca posible después de la reconstrucción por parte del decodificador. Al transmitir parámetros ICC, esto significa que debe mantenerse el ancho (descorrelación) del campo de sonido original.
- 10 **[0115]** Debe tenerse en cuenta que el valor ICC más importante se encuentra entre el eje izquierdo/derecho, ya que el oyente generalmente está mirando hacia adelante en la configuración de escucha. Esto se puede tener en cuenta de manera ventajosa para construir la estructura de codificación jerárquica de modo que se pueda obtener una representación paramétrica adecuada de la señal de audio durante el procedimiento de codificación iterativo, donde el valor ICC combinado resultante representa principalmente la descorrelación izquierda/derecha. Esto se explicará con más detalle más adelante cuando se analicen las realizaciones preferidas de la presente invención.

- 15 **[0116]** El esquema de codificación/decodificación de la invención permite reducir el número de parámetros transmitidos desde un codificador a un decodificador mediante una estructura jerárquica de un sistema de audio espacial por medio de las dos medidas siguientes:
- 20
- la combinación de los parámetros del codificador individual para formar un parámetro combinado, que se transmite al decodificador en lugar de los individuales. La combinación de los parámetros se lleva a cabo de manera que la imagen de sonido de la señal (incluida la correlación/coherencia L/R) se conserve en la medida de lo posible.
- 25
- el parámetro combinado transmitido se utiliza en el decodificador en lugar de varios parámetros individuales transmitidos (o los parámetros realmente utilizados se derivan del combinado).

- [0117]** Según ciertos requisitos de implementación de los procedimientos de la invención, los procedimientos de la invención pueden ser implementados en hardware o en software. La implementación puede ser realizada mediante el uso de un medio de almacenamiento digital, en particular, un disco, un DVD o un CD que tiene señales de control legibles electrónicamente almacenadas en el mismo, que cooperan con sistemas informáticos programables, de tal manera que se realicen los procedimientos de la invención. En general, la presente invención es por lo tanto un producto de programa informático con un código de programa almacenado en un soporte legible por máquina, el código de programa se opera para realizar los procedimientos de la invención cuando el producto de programa informático se ejecuta en un ordenador. En otras palabras, los procedimientos de la invención son por lo tanto un programa informático que tiene un código de programa para realizar al menos uno de los procedimientos de la invención cuando el programa informático se ejecuta en un ordenador.
- 30
- 35

REIVINDICACIONES

1. Codificador para generar una representación paramétrica (238) de una señal de audio que tiene al menos dos canales izquierdos originales (224a, 224b) en un lado izquierdo y dos canales derechos originales (224c, 224d) en un lado derecho con respecto a una posición de escucha, que comprende:
- un generador (220) para generar información paramétrica, el generador (220) está operativo para procesar por separado varios pares de canales para derivar una información de nivel (230a, 230b, 234) para pares de canales procesados, y para derivar información de coherencia (236) para un par de canales que incluye un primer canal (228a; 224a, 224b) que solo tiene información del lado izquierdo y un segundo canal (228b; 224c, 224d) que solo tiene información del lado derecho; y
- un proveedor (222) para proporcionar la representación paramétrica (238) seleccionando la información de nivel (230a, 230b, 234) para pares de canales y determinando una medida de coherencia izquierda/derecha mediante la información de coherencia (236) e introduciendo la medida de coherencia izquierda/derecha en un flujo de datos de salida como la única información de coherencia de la señal de audio dentro de la representación paramétrica (238).
2. Codificador según la reivindicación 1, donde el generador (220) está operativo
- para procesar (242a) un canal frontal izquierdo lf y un canal trasero izquierdo lr para obtener una información de nivel lf/lr, donde una combinación del canal frontal izquierdo lf y el canal trasero izquierdo lr forma un canal maestro izquierdo LM, y para procesar (242b) un canal frontal derecho rf y un canal trasero derecho rr para obtener una información de nivel rf/rr, donde una combinación del canal frontal derecho rf y el canal trasero derecho rr forma un canal maestro derecho RM; y
- para procesar (242d) el canal maestro izquierdo LM y el canal maestro derecho RM para derivar una información de nivel lm/rm y obtener la información de coherencia (236), donde una combinación del canal maestro izquierdo LM y el canal maestro derecho RM forma un canal maestro estéreo SM.
3. Codificador según la reivindicación 2, donde el generador (220) está operativo para procesar (242c) un canal central ce y un canal de baja frecuencia para obtener una información de nivel ce/lo, donde una combinación del canal central ce y el canal de baja frecuencia forma un canal maestro central CM.
4. Codificador según la reivindicación 3, en el cual el generador (220) está operativo para procesar (242e) el canal maestro estéreo SM y el canal maestro central CM para obtener una información de nivel sm/cm, donde una combinación del canal maestro estéreo SM y el canal maestro central CM forma un canal de mezcla descendente (274); y donde el proveedor (222) está operativo para determinar la medida de coherencia izquierda/derecha mediante la información de coherencia (236) y la información de nivel sm/cm.
5. Codificador según la reivindicación 4, donde el proveedor (222) está operativo para calcular (244) la medida de coherencia izquierda/derecha según la información del nivel sm/cm tal que, en un caso, donde la información del nivel sm/cm indica que hay más energía en el canal maestro estéreo SM que en el canal maestro central CM, la medida de coherencia izquierda/derecha está más cerca de la información de coherencia en comparación con una situación donde la información del nivel de sm/cm indica que hay más energía en el canal maestro central CM, en cuyo caso la medida de coherencia izquierda/derecha está más cerca de la unidad.
6. Codificador según la reivindicación 4, donde el proveedor (222) está operativo para calcular (244) la medida de coherencia izquierda/derecha según la información del nivel sm/cm, de manera que, en un caso, donde la información del nivel sm/cm indica que una relación de la energía en el canal maestro estéreo SM y la energía en el canal maestro central CM excede un valor predefinido, la medida de coherencia izquierda/derecha se ajusta a la información de coherencia en comparación con una situación, donde la información del nivel sm/cm indica que la relación de la energía en el canal maestro estéreo SM con la energía en el canal maestro central CM permanece por debajo o es igual que el valor predefinido, donde la medida de coherencia izquierda/derecha se establece en la unidad.
7. Codificador según la reivindicación 1, en el cual el generador (220) está operativo para procesar (242a) un canal izquierdo frontal lf y un canal derecho frontal rf para obtener una información del nivel lf/rf y una primera información de coherencia (302a), donde una combinación del canal frontal izquierdo lf y el canal frontal derecho rf forma un canal maestro frontal FM, y para procesar (242c) un canal trasero izquierdo lr y un canal trasero derecho rr para obtener información del nivel lr/rr y obtener una segunda información de coherencia (302b), donde una combinación del canal trasero izquierdo lr y el canal trasero derecho rr forma un canal maestro trasero RM, y donde el proveedor (222) está operativo para determinar la medida de coherencia izquierda/derecha que combina la primera información de coherencia (302a) y la segunda información de coherencia (302b).
8. Codificador según la reivindicación 7, donde el proveedor (222) está operativo para determinar la medida de coherencia izquierda/derecha según una suma ponderada de la primera y la segunda información de coherencia

(302a, 302b), mediante información del nivel del canal maestro frontal FM e información del nivel del canal maestro trasero RM como ponderaciones.

9. Codificador según la reivindicación 7, donde el generador (220) está operativo para procesar (242c) un canal central ce y un canal de baja frecuencia para obtener una información del nivel ce/lo, donde una combinación del canal central ce y el canal de baja frecuencia forma un canal maestro central CM.
10. Codificador según la reivindicación 9, en el cual el generador (220) está operativo para procesar (242d) el canal maestro frontal FM y el canal maestro central CM para obtener una información del nivel fm/cm, donde una combinación del canal maestro frontal FM y el canal maestro central CM forma un canal frontal puro PF; y donde el proveedor (222) está operativo para determinar la medida de coherencia izquierda/derecha al combinar la primera y la segunda información de coherencia (302a, 302b) además de usar la información del nivel fm/cm.
11. Codificador según la reivindicación 11, donde el generador (220) está operativo para procesar (242e) el canal frontal puro PF y el canal maestro trasero RM para obtener una información del nivel pf/rm, donde una combinación del canal frontal puro PF y el canal maestro trasero RM forma un canal de mezcla descendente.
12. Codificador según la reivindicación 1, donde el generador (220) está operativo para procesar los pares de canales en marcos de tiempo discretos de una longitud dada.
13. Codificador según la reivindicación 1, donde el generador (220) está operativo para procesar los pares de canales en intervalos de frecuencia discreta de un ancho de banda determinado.
14. El decodificador para procesar una representación paramétrica (314, 316, 318, 320) de una señal de audio original, donde la señal de audio original tiene al menos dos canales originales a la izquierda (324b, 324d; 360a, 360b) en un lado izquierdo y al menos dos canales derechos originales (324a, 324d; 360c, 360d) en un lado derecho con respecto a una posición de escucha, que comprende:
- un receptor (310) para proporcionar la representación paramétrica (314, 316, 318, 320) de la señal de audio, donde el receptor (310) está operativo para proporcionar información de nivel (314, 316, 318) para pares de canales y para proporcionar una medida de coherencia izquierda/derecha (320) para un par de canales que incluye un canal izquierdo y un canal derecho como la única información de coherencia de la señal de audio original dentro de la representación paramétrica (314, 316, 318, 320), la medida de coherencia izquierda/derecha (320) que representa una información de coherencia entre al menos un par de canales que incluye un primer canal que solo tiene información del lado izquierdo y un segundo canal que solo tiene información del lado derecho; y un procesador (312) para suministrar información paramétrica para pares de canales, el procesador (312) está operativo para seleccionar información de nivel (314, 316, 318) de la representación paramétrica (314, 316, 318, 320) y para derivar la información de coherencia para al menos un par de canales que utiliza la medida de coherencia izquierda/derecha (320), el al menos un par de canales que incluye un primer canal (324b, 324d) que solo tiene información del lado izquierdo y un segundo canal (324a, 324c) que solo tiene información del lado derecho.
15. Decodificador según la reivindicación 14, en el cual el receptor (310) está operativo para proporcionar una información del nivel lf/lr (380a) para un par de canales de un canal frontal izquierdo original lf y un canal trasero izquierdo original lr, donde una combinación del canal frontal izquierdo original lf y el canal trasero izquierdo original lr forma un canal maestro izquierdo LM; proporcionar una información del nivel rf/rr (380b) para un par de canales de un canal frontal derecho original rf y un canal trasero derecho original rr, donde una combinación del canal frontal derecho original rf y el canal trasero derecho original rr forma un canal maestro derecho RM; proporcionar una información del nivel lm/rm (380d) para un par de canales del canal maestro izquierdo LM y el canal maestro derecho RM, donde una combinación del canal maestro izquierdo LM y el canal maestro derecho RM forma un canal maestro estéreo SM; y donde el procesador (312) está operativo para proporcionar (352) información de coherencia para el canal maestro izquierdo LM y el canal maestro derecho RM mediante la medida de coherencia izquierda/derecha (320); el decodificador comprende además un mezclador ascendente, donde mezclador ascendente tiene:
- un primer mezclador ascendente 1 a 2 (354b) para la generación del canal maestro izquierdo LM y el canal maestro derecho RM desde el canal maestro estéreo SM mediante la información del nivel lm/rm (380d) y la medida de coherencia izquierda/derecha (320); un segundo mezclador ascendente 1 a 2 (354e) para la generación del canal frontal izquierdo original lf (360a) y el canal trasero izquierdo original lr (360b) del canal maestro izquierdo LM mediante la información del nivel lf/lr (380a) y una información de coherencia predefinida; y

un tercer mezclador ascendente 1 a 2 (354d) para la generación del canal frontal derecho original rf (360c) y el canal trasero derecho original rr (360d) del canal maestro derecho RM mediante la información del nivel rf/rr (380b) y una información de coherencia predefinida.

5 16. Decodificador según la reivindicación 15,
donde el receptor (310) está operativo para proporcionar una información de nivel ce/lo (380c) para un par de canales de un canal central original ce y de un canal de baja frecuencia original lo, donde una combinación del canal central original ce y del canal de baja frecuencia original lo forma un canal maestro central CM; y
donde el mezclador ascendente comprende además un cuarto mezclador ascendente de 1 a 2 (354c) para la
10 generación del canal central original ce (360e) y el canal de baja frecuencia original lo (360f) del canal maestro central CM mediante la información del nivel ce/lo (380c) y una información de coherencia predefinida.

17. Decodificador según la reivindicación 16,
en el cual el receptor (310) está operativo para proporcionar una información del nivel sm/cm (380e) para un par de
15 canales del canal maestro estéreo SM y del canal maestro central CM, donde una combinación del canal maestro estéreo SM y el canal maestro estéreo SM forma un canal de mezcla descendente; y
donde el mezclador ascendente comprende además un quinto mezclador ascendente 1 a 2 (354a) para la generación del canal maestro estéreo SM y el canal maestro central CM del canal de mezcla descendente mediante la información del nivel sm/cm (380e) y una información de coherencia predefinida.

20 18. Decodificador según la reivindicación 14, donde el receptor (310) está operativo para
proporcionar una información del nivel lf/rf (410a) para un par de canales de un canal frontal izquierdo original lf y de un canal frontal derecho original rf, donde una combinación del canal frontal izquierdo original lf y el canal frontal
25 derecho original rf forma un canal maestro frontal FM;
proporcionar una información del nivel lr/rr (410c) para un par de canales de un canal izquierdo trasero original lr y un canal trasero derecho original rr, donde una combinación del canal izquierdo trasero original lr y el canal derecho trasero original rr forma un canal maestro trasero RM; y

30 donde el procesador (312) está operativo para suministrar una primera información de coherencia para el canal original izquierdo frontal y el canal original derecho frontal y para proporcionar una segunda información de coherencia para el canal izquierdo trasero original lr y el canal derecho trasero original rr mediante la medida de coherencia izquierda/derecha (320);
el decodificador comprende además un mezclador ascendente, donde mezclador ascendente tiene:

35 un primer mezclador ascendente 1 a 2 (354e) para la generación del canal frontal izquierdo original lf (360a) y el canal frontal derecho original rf (360c) desde el canal maestro frontal FM mediante la información del nivel lf/rf (410a) y la medida de coherencia izquierda/derecha (320);
un segundo mezclador ascendente 1 a 2 (354c) para la generación del canal trasero izquierdo original lr (360b) y del
40 canal trasero derecho original rr (360d) del canal maestro trasero RM mediante la información del nivel lr/rr (410c) y la medida de coherencia izquierda/derecha (320).

19. Decodificador según la reivindicación 18,
en el cual el receptor (310) está operativo para proporcionar una información del nivel ce/lo (410b) para un par de
45 canales de un canal central original ce y de un canal de baja frecuencia original lo, donde una combinación del canal central original ce y del canal original de baja frecuencia forma un canal maestro central CM; y
donde el mezclador ascendente comprende además un tercer mezclador ascendente 1 a 2 (354d) para la generación del canal central original ce (360e) y el canal de baja frecuencia original lo (360f) del canal maestro central CM mediante la información del nivel ce/lo (410b) y una información de coherencia predefinida.

50 20. Decodificador según la reivindicación 19,
en el cual el receptor (310) está operativo para proporcionar una información del nivel fm/cm (410b) para un par de canales del canal maestro frontal FM y el canal maestro central CM, donde una combinación del canal maestro frontal FM y el canal maestro central CM forma un canal frontal puro PF; y
55 donde el mezclador ascendente comprende además un cuarto mezclador ascendente 1 a 2 (354b) para la generación del canal maestro frontal FM y el canal maestro central CM del canal frontal puro PF mediante la información del nivel fm/cm (410b) y una información de coherencia predefinida.

21. Decodificador según la reivindicación 20,
60 en el cual el receptor (310) está operativo para proporcionar una información del nivel pf/rm (410a) para un par de canales del canal frontal puro PF y el canal maestro trasero RM, donde una combinación del canal frontal puro PF y el canal maestro trasero RM forma un canal de mezcla descendente; y
donde el mezclador ascendente comprende además un quinto mezclador ascendente 1 a 2 (354a) para la generación del canal frontal puro PF y el canal maestro trasero RM del canal de mezcla descendente mediante la información del
65 nivel pf/rm (410a) y una información de coherencia predefinida.

22. Decodificador según la reivindicación 14, donde el procesador (312) está operativo para derivar medidas de coherencia para todos los pares de canales distribuyendo la medida de coherencia izquierda/derecha (320) recibida como las medidas de coherencia.

5

23. Decodificador según la reivindicación 14, donde el receptor (310) está operativo para

operar en un primer modo, proporcionando información de nivel (314, 316, 318) para pares de canales y proporcionando una medida de coherencia izquierda/derecha (320) para un par de canales que incluye un canal izquierdo y un canal derecho como la única información de coherencia de la señal de audio dentro de la representación paramétrica (314, 316, 318, 320), la medida de coherencia izquierda/derecha (320) representa una información de coherencia entre al menos un par de canales que incluye un primer canal que solo tiene información del lado izquierdo y un segundo canal que solo tiene información del lado derecho con respecto a una posición de escucha; o para operar en un segundo modo, proporcionando la información de nivel (314, 316, 318) para pares de canales y la información de coherencia para los mismos pares de canales; y

15

donde el procesador (312) está operativo para suministrar información paramétrica para pares de canales

en el primer modo, el procesador (312) está operativo para seleccionar la información de nivel (314, 316, 318) de la representación paramétrica (314, 316, 318, 320) y para derivar la información de coherencia para al menos un par de canales mediante la medida de coherencia izquierda/derecha (320), el al menos un par de canales que incluye un primer canal que solo tiene información del lado izquierdo y un segundo canal que solo tiene información del lado derecho; o

20

en el segundo modo, el procesador (312) está operativo para seleccionar la información de nivel (314, 316, 318) de la representación paramétrica (314, 316, 318, 320) y para seleccionar la información de coherencia de la representación paramétrica (314, 316, 318, 320).

25

24. Decodificador según la reivindicación 23, donde el receptor (310) comprende además un receptor de modo para seleccionar un modo de operación mediante información de modo recibida, la información de modo que indica el primer o el segundo modo a usar.

30

25. Un procedimiento para generar una representación paramétrica (238) de una señal de audio que tiene al menos dos canales izquierdos originales (224a, 224b) y al menos dos canales derechos originales (224c, 224c) con respecto a una posición de escucha, el procedimiento que comprende:

35

generar información paramétrica procesando por separado varios pares de canales para obtener una información de nivel (230a, 230b, 234) para pares de canales procesados y derivando información de coherencia (236) para un par de canales que incluye un primer canal (228a; 224a, 224b) que solo tiene información del lado izquierdo y un segundo canal (228b; 224c, 224d) que solo tiene información del lado derecho, y

40

proporcionar la representación paramétrica (238) seleccionando información de nivel (230a, 230b, 234) para pares de canales y determinando una medida de coherencia izquierda/derecha utilizando la información de coherencia (236) e introduciendo la medida de coherencia izquierda/derecha en un flujo de datos de salida como la única información de coherencia (236) de la señal de audio dentro de la representación paramétrica (238).

45

26. Un procedimiento para procesar una representación paramétrica (314, 316, 318, 320) de una señal de audio original, donde la señal de audio original tiene al menos dos canales izquierdos originales (324b, 324d; 360a, 360b) en el lado izquierdo y al menos dos canales derechos originales (324a, 324d; 360c, 360d) en el lado derecho con respecto a una posición de escucha, el procedimiento comprende:

50

proporcionar la representación paramétrica (314, 316, 318, 320) de la señal de audio, proporcionando información de nivel (314, 316, 318) para pares de canales y proporcionando una medida de coherencia izquierda/derecha (320) para un par de canales que incluye un canal izquierdo y un canal derecho como la única información de coherencia de la señal de audio original dentro de la representación paramétrica (314, 316, 318, 320), la medida de coherencia izquierda/derecha (320) que representa una información de coherencia entre al menos un par de canales que incluye un primer canal que solo tiene información del lado izquierdo y un segundo canal que solo tiene información del lado derecho; y

55

suministrar información paramétrica para pares de canales seleccionando información de nivel (314, 316, 318) de la representación paramétrica (314, 316, 318, 320) y derivando información de coherencia para al menos un par de canales mediante la medida de coherencia izquierda/derecha (320), el al menos un par de canales que incluye un primer canal (324b, 324d) que solo tiene información del lado izquierdo y un segundo canal (324a, 324c) que solo tiene información del lado derecho.

60

27. Datos de audio codificados con información lateral, la información lateral que comprende una representación paramétrica (314, 316, 318, 320) de una señal de audio que tiene al menos dos canales izquierdos originales (324b, 324d; 360a, 360b) en un lado izquierdo y dos canales derechos originales (324a, 324d; 360c, 360d)

65

en un lado derecho con respecto a una posición de escucha, donde la representación paramétrica (314, 316, 318, 320) comprende diferencias de nivel (314, 316, 318) para pares de canales y una medida de coherencia izquierda/derecha (320) derivada de la información de coherencia de un par de canales que incluye un primer canal que solo tiene información del lado izquierdo y un segundo canal que solo tiene información del lado derecho, como la única información de coherencia de la señal de audio original dentro de la representación paramétrica.

28. Medio de almacenamiento legible por ordenador, que ha almacenado en él datos de audio codificados según la reivindicación 27.
- 10 29. Receptor (520) o reproductor de audio que tiene un decodificador (312) según la reivindicación 14.
30. Transmisor (500) o grabador de audio que tiene un codificador (220) según la reivindicación 1.
31. Procedimiento de recepción o reproducción de audio, donde el procedimiento tiene un procedimiento para procesar según la reivindicación 26.
- 15 32. Procedimiento de transmisión o grabación de audio, donde el procedimiento tiene un procedimiento para generar según la reivindicación 25.
- 20 33. Sistema de transmisión que tiene un transmisor (500) y un receptor (520),
donde el transmisor (500) tiene un codificador (220) según la reivindicación 1 y donde el receptor (520) tiene un decodificador (312) según la reivindicación 14.
- 25 34. Procedimiento de transmisión y recepción,
donde el procedimiento de transmisión tiene un procedimiento para generar según la reivindicación 25 y donde el procedimiento de recepción tiene un procedimiento para procesar según la reivindicación 26.
- 30 35. Ordenador que tiene un código de programa para ejecutar, cuando se ejecuta en un ordenador, cualquiera de los procedimientos de las reivindicaciones 25, 26, 31, 32 o 34.

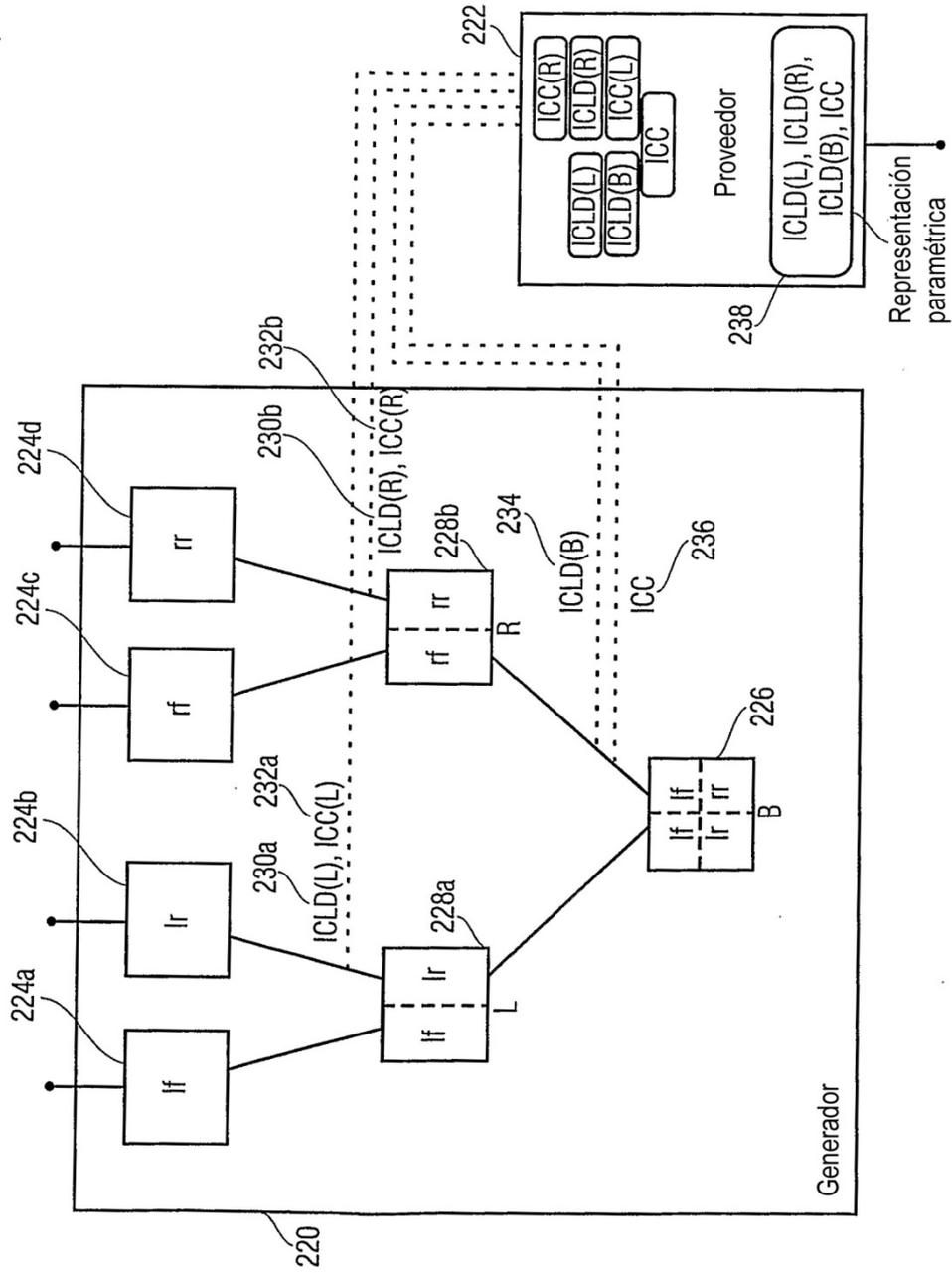


Fig. 1

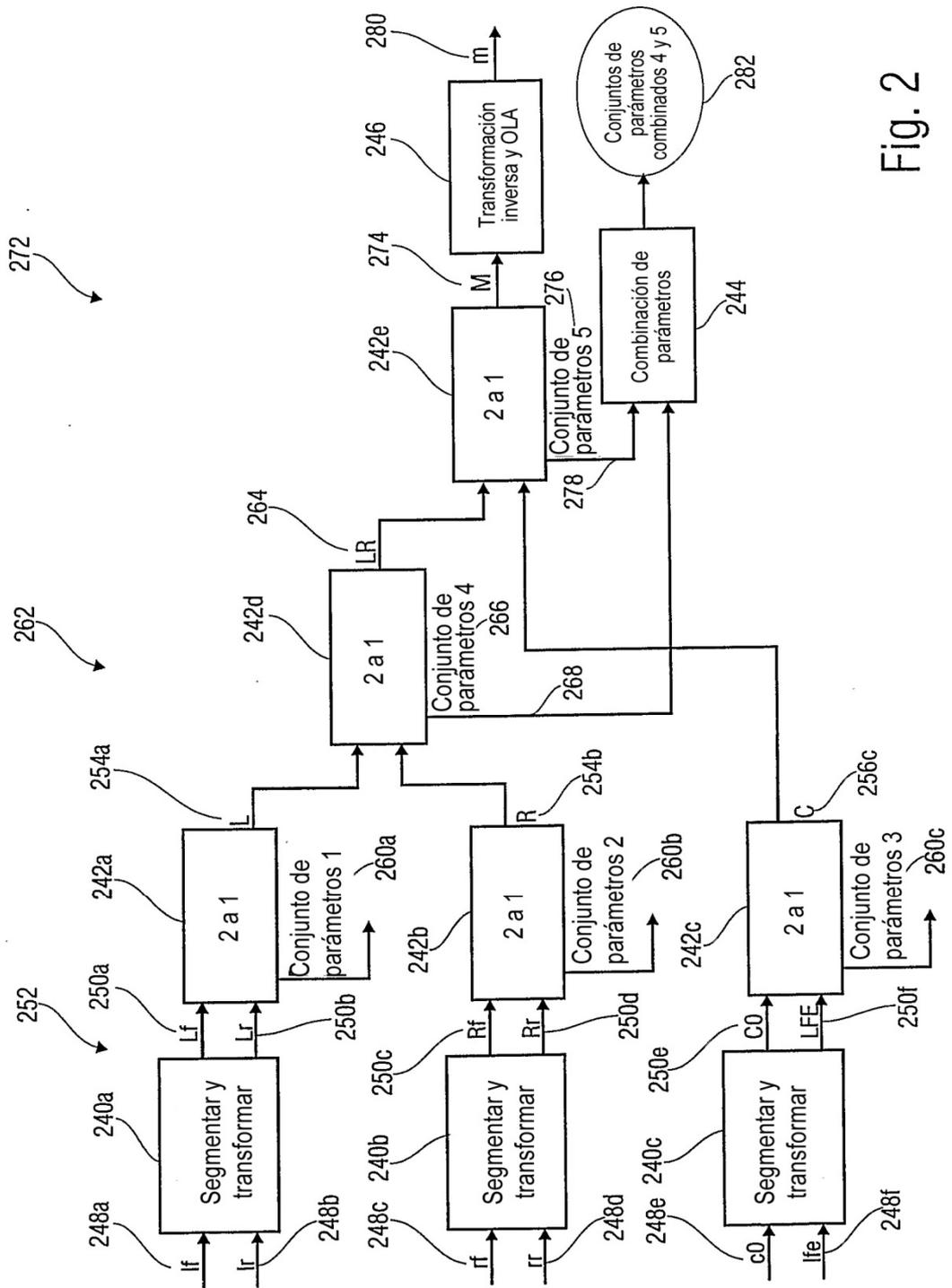


Fig. 2

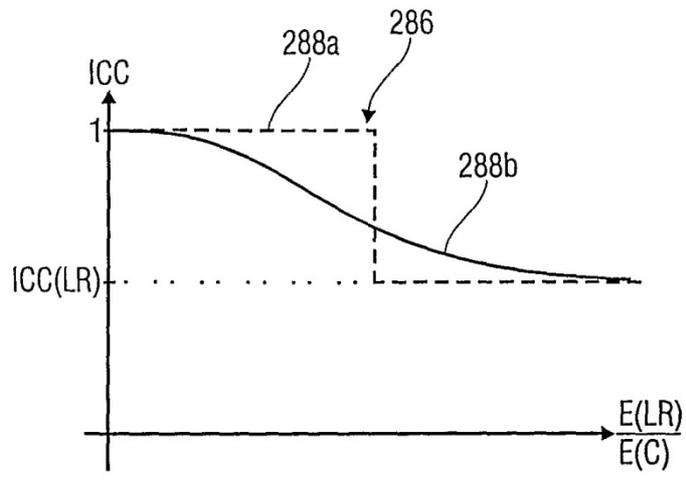


Fig. 2a

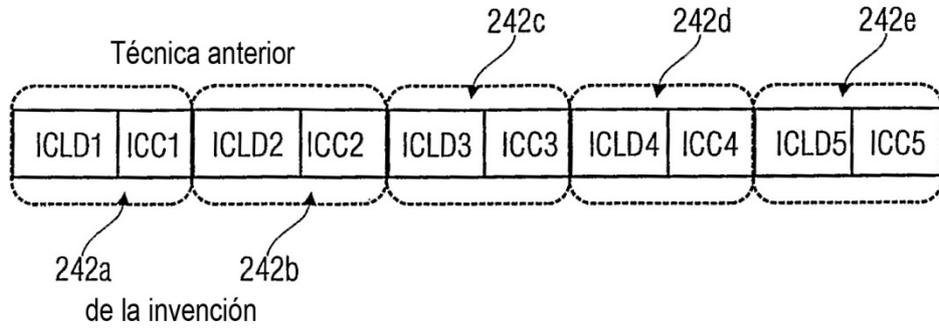


Fig. 3a

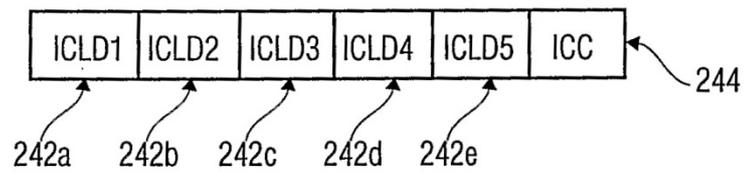


Fig. 3b

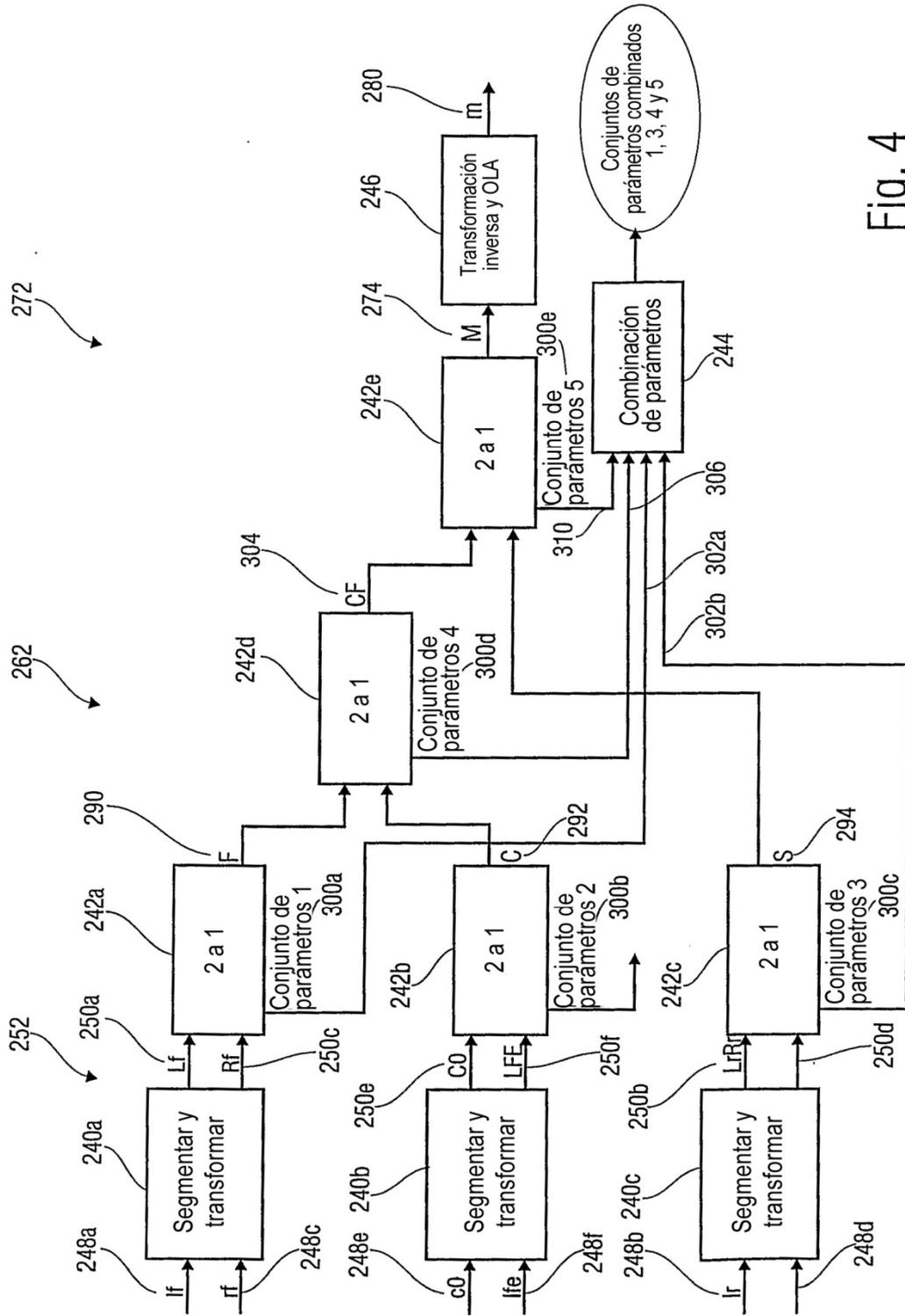


Fig. 4

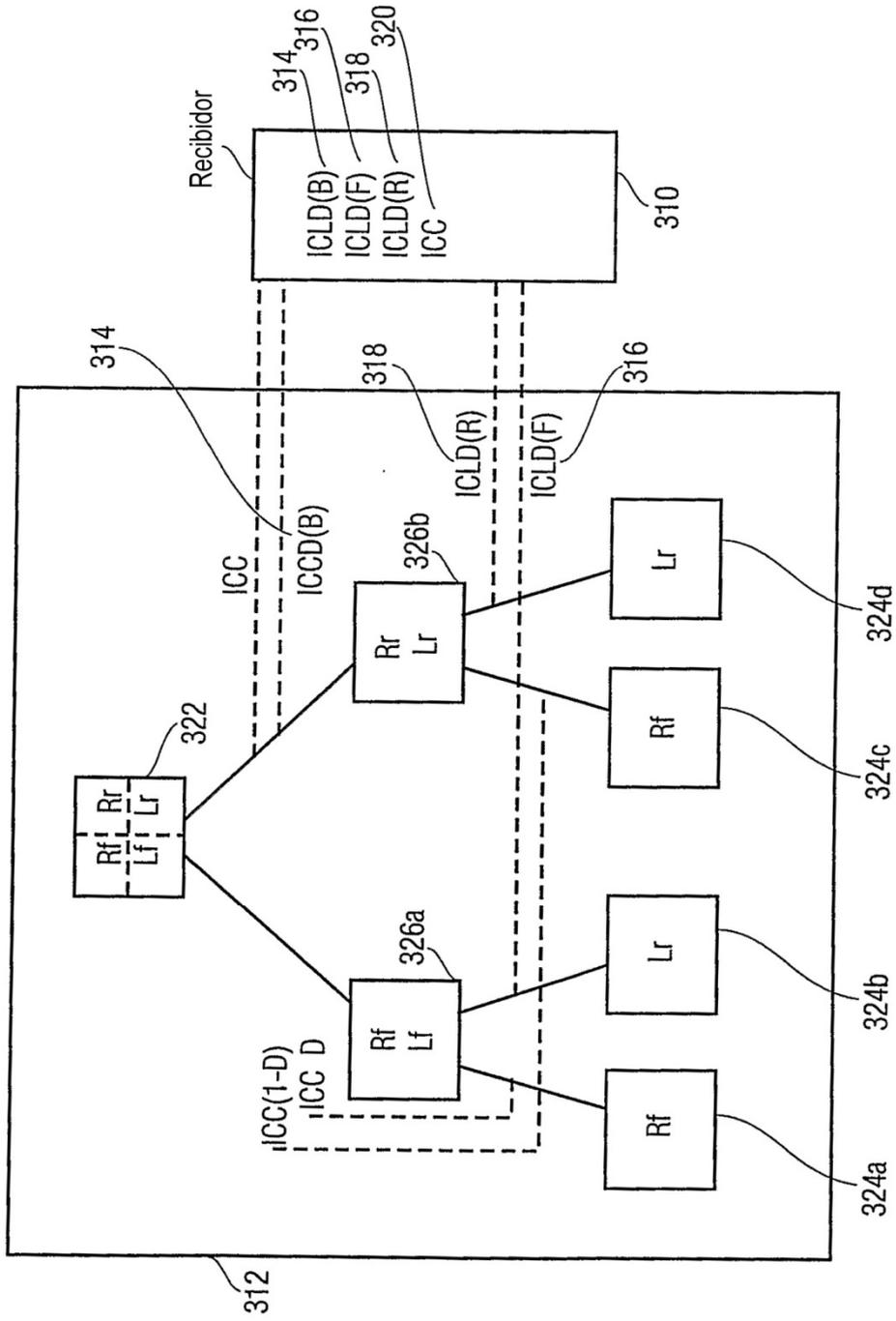


Fig. 5

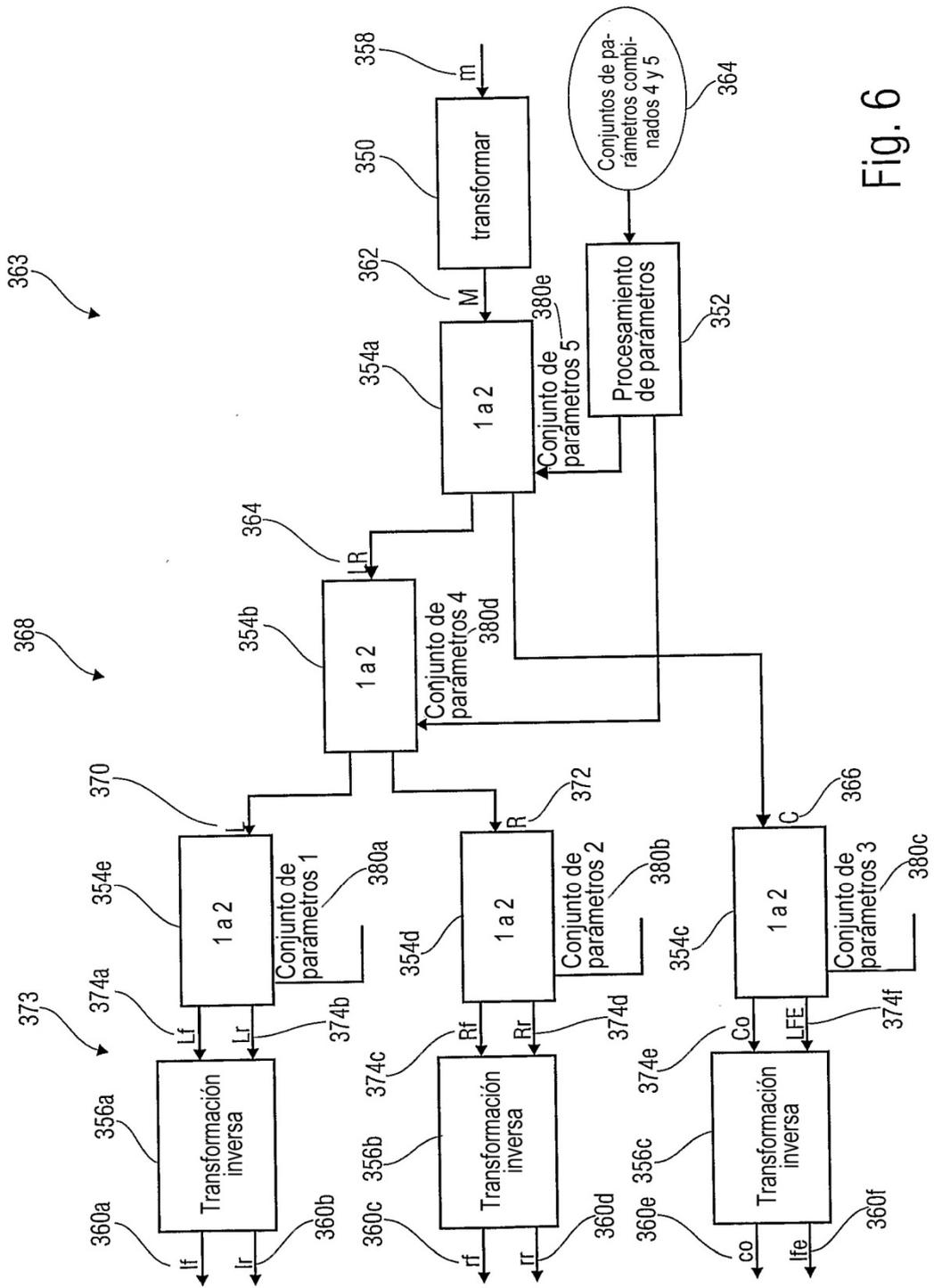


Fig. 6

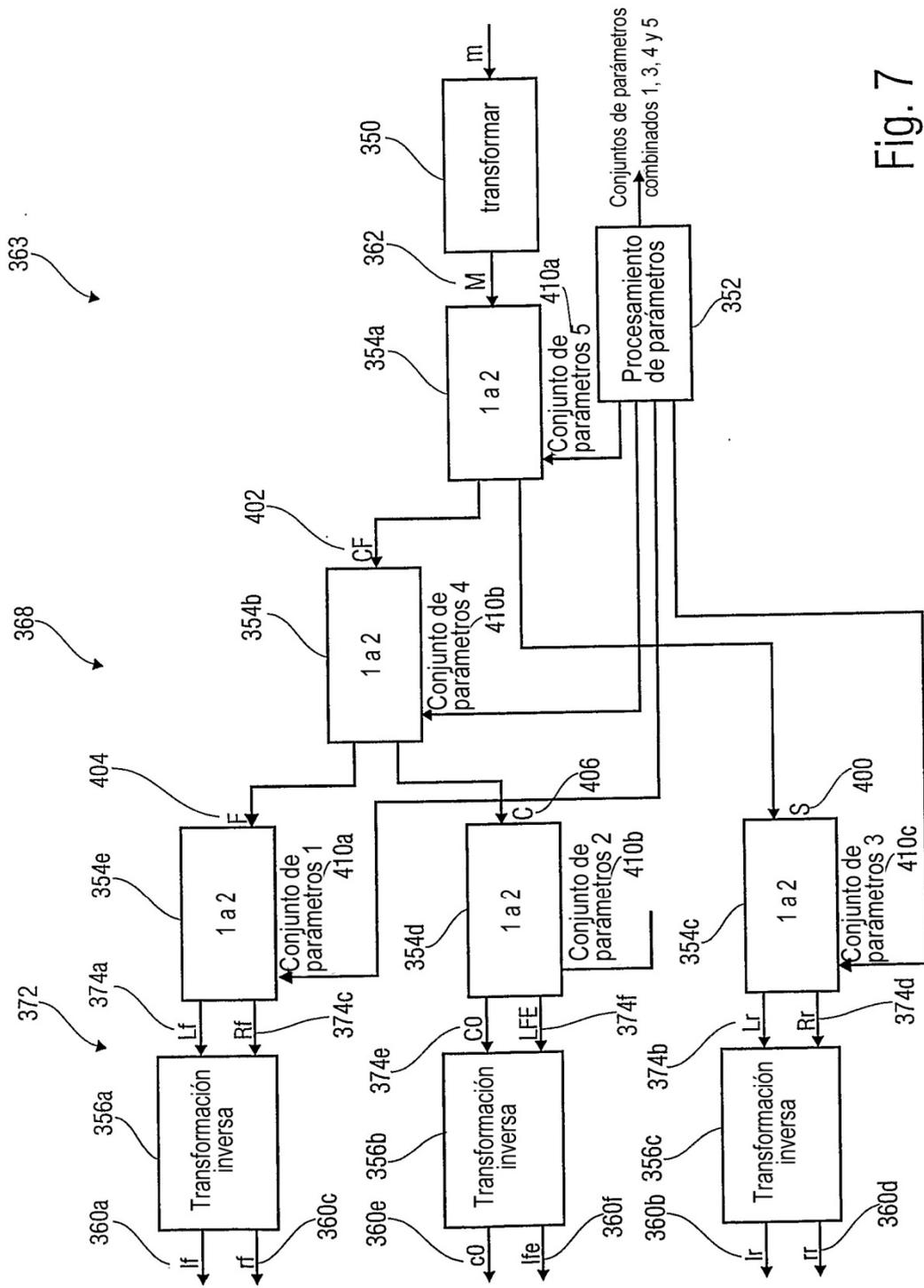


Fig. 7

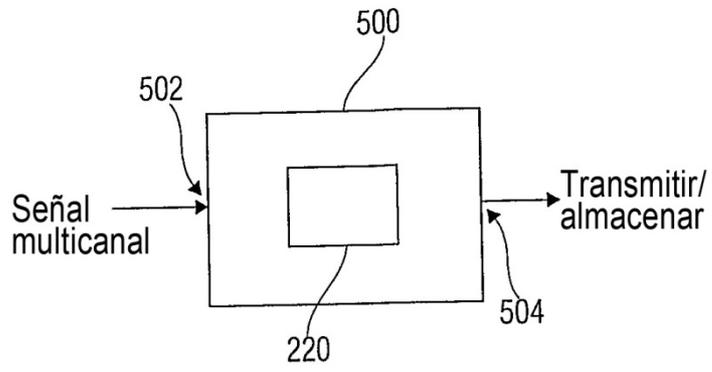


Fig. 8

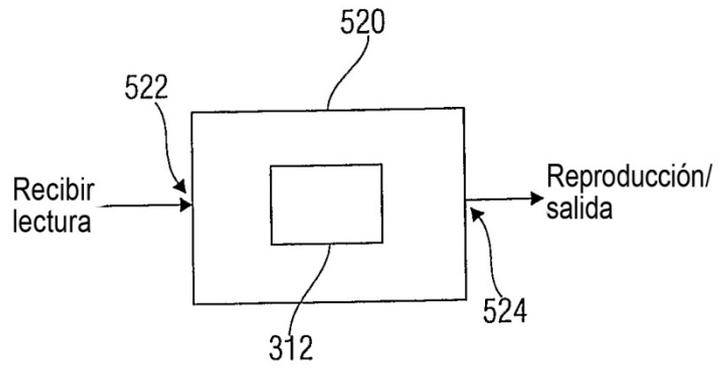


Fig. 9

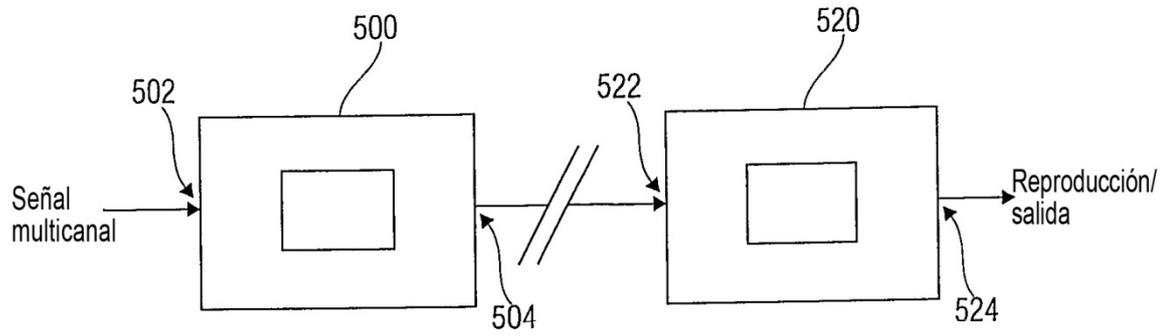


Fig. 10

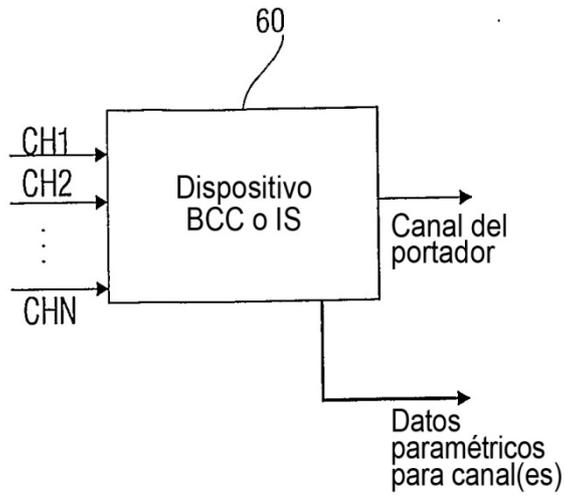


Fig. 11
(Técnica anterior)

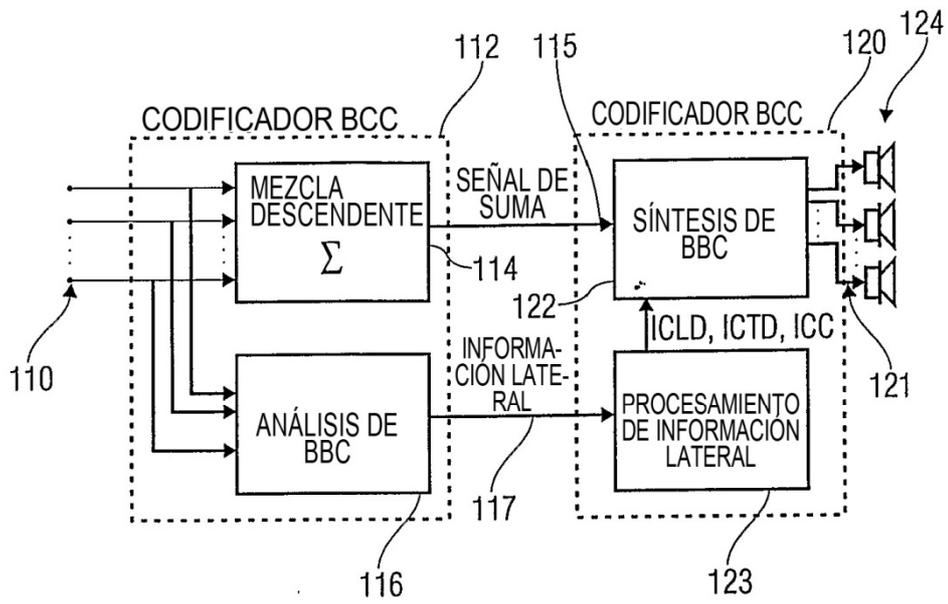


Fig. 12
(Técnica anterior)

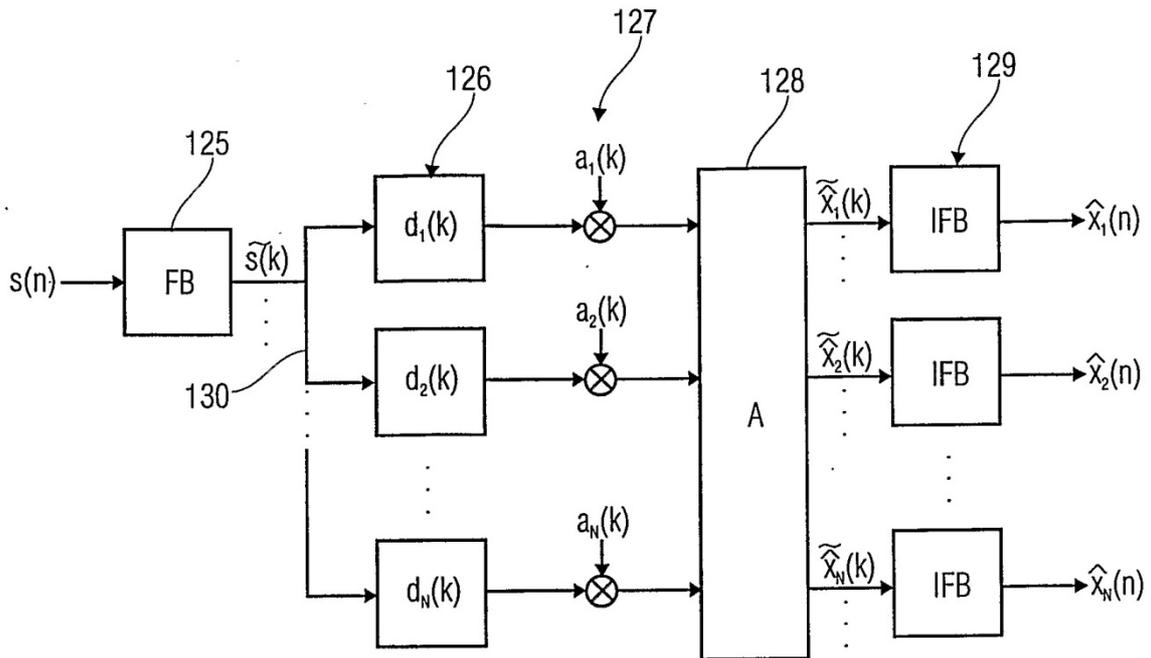


Fig. 13
(Técnica anterior)

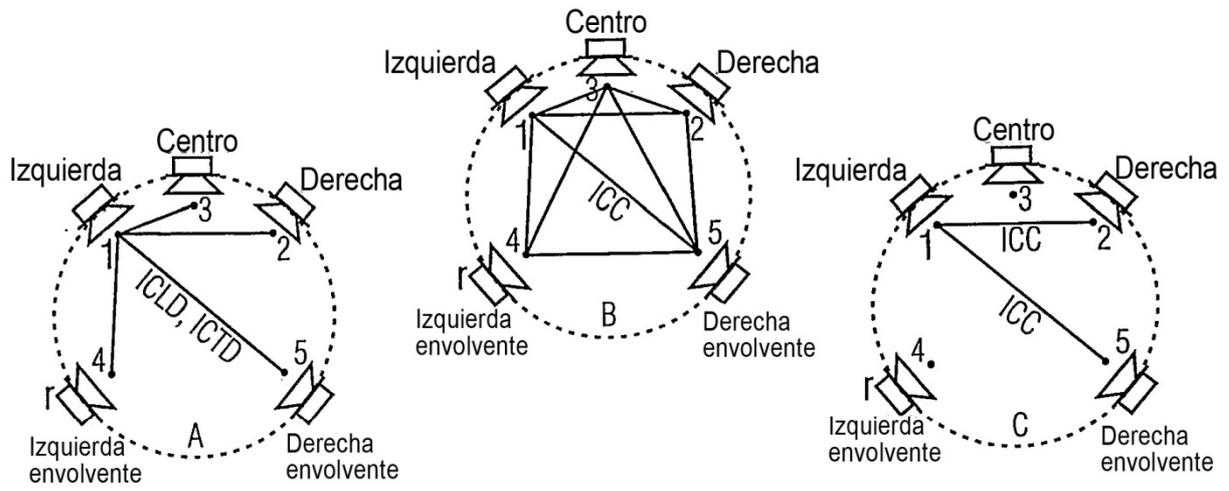


Fig. 14
(Técnica anterior)

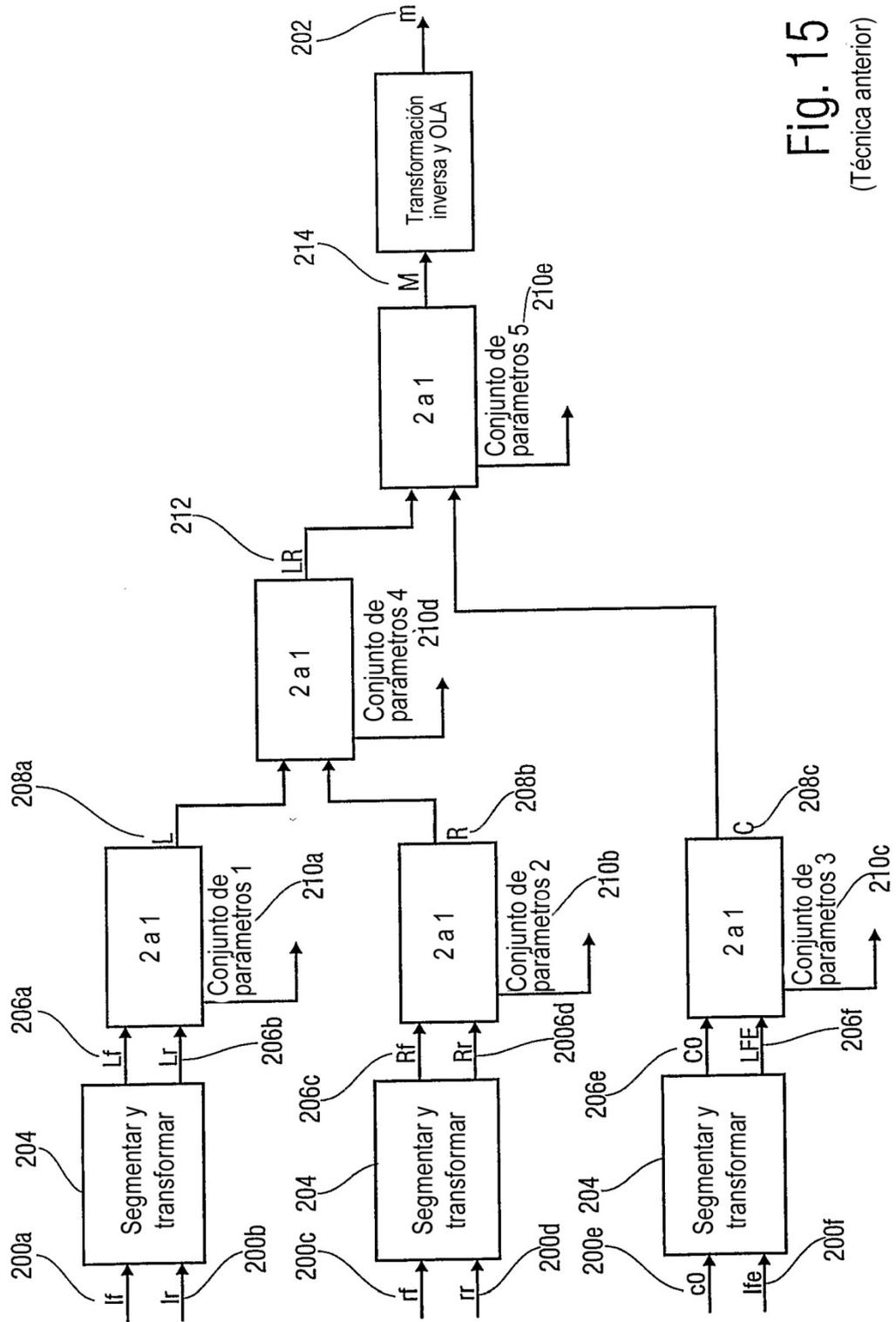


Fig. 15
(Técnica anterior)