(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 387 248**

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

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Número de solicitud europea: **05740130 .9**
(96) Fecha de presentación: **12.05.2005**
(97) Número de publicación de la solicitud: **1774515**
(97) Fecha de publicación de la solicitud: **18.04.2007**

(54) Título: **Aparato y procedimiento para generar una señal de salida multi-canal**

(30) Prioridad:
09.07.2004 US 586578 P
07.09.2004 US 935061

(45) Fecha de publicación de la mención BOPI:
19.09.2012

(45) Fecha de la publicación del folleto de la patente:
19.09.2012

(73) Titular/es:
Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.
Hansastrasse 27c
80686 München, DE y
AGERE SYSTEM INC.

(72) Inventor/es:
HERRE, Jürgen;
FALLER, Christof;
DISCH, Sascha y
HILPERT, Johannes

(74) Agente/Representante:
Ponti Sales, Adelaida

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).

ES 2 387 248 T3

DESCRIPCIÓN

Aparato y procedimiento para generar una señal de salida multi-canal.

Sector de la invención

[0001] La presente invención se refiere a la descodificación multi-canal y, en particular, a la descodificación multi-canal, en la que por ejemplo están presentes al menos dos canales de transmisión, que es estéreo compatible.

[0002] En los últimos tiempos, la técnica de reproducción de audio multi-canal se está convirtiendo cada vez más importante. Esto puede ser debido al hecho de que las técnicas de compresión/ codificación de audio tales como la bien conocida técnica mp3 han hecho posible distribuir los registros de audio a través de Internet u otros canales de transmisión con un ancho de banda limitado. La técnica de codificación mp3 se ha hecho tan famosa por el 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 canal estéreo primero o izquierdo y un segundo canal estéreo o derecho.

[0003] Sin embargo, hay carencias básicas en los sistemas de sonido de dos canales convencionales. Por lo tanto, se ha desarrollado la técnica envolvente. Una representación envolvente multi-canal recomendada incluye, además de los dos canales estéreo L y R, un canal central adicional C y dos canales de sonido envolvente Ls, Rs. Este formato de sonido de referencia también se conoce como estéreo tres / dos, lo que significa tres canales frontales y dos canales de sonido envolvente. Generalmente, se requieren cinco canales de transmisión. En un entorno de reproducción, son necesarios por lo menos cinco altavoces en los respectivos cinco lugares diferentes para obtener un punto óptimo agradable a una cierta distancia de los cinco altavoces bien colocados.

[0004] Son conocidas varias técnicas en la técnica para reducir la cantidad de datos necesarios para la transmisión de una señal de audio de canales múltiples. Estas técnicas se denominan técnicas estéreo conjuntas. Para este fin, se hace referencia a la figura 10, que muestra un dispositivo estéreo conjunto 60. Este dispositivo puede ser un dispositivo que implementa por ejemplo intensidad estéreo (IS) o codificación binaural (BCC). Este dispositivo recibe generalmente - como entrada- al menos dos canales (CH1, CH2,...CHn), y tiene como salida un canal portador único y datos paramétricos. Los datos paramétricos se definen se modo que, en un descodificador, se puede calcular una aproximación de un canal original (CH1, CH2,...

[0005] Normalmente, el canal portador incluirá muestras de sub-banda, coeficientes espectrales, muestras en el dominio del tiempo etc, que proporcionan una representación comparativamente fina de la señal subyacente, mientras que los datos paramétricos no incluyen estas muestras de coeficientes espectrales sino que incluyen los parámetros de control para controlar un algoritmo de reconstrucción determinado, tal como la ponderación por multiplicación, desfase de tiempo, cambio de frecuencia, ... Los datos paramétricos, por lo tanto, incluyen solamente una representación relativamente basta de la señal o del canal asociado. Expresado en números, la cantidad de datos requeridos por un canal portador estará en el intervalo de 60-70 kbit / s, mientras que la cantidad de datos requeridos por el lado de información paramétrico para un canal estará en el intervalo de 1,5 - 2, 5 kbit / s. Un ejemplo de datos paramétricos son los bien conocidos factores de escala, la información de intensidad estéreo o los parámetros binaurales tal como se describen a continuación.

[0006] El documento de HERRE J., FALLER C.: "MP3 Surround: Efficient y Compatible Coding de Multi-Channel Audio" AES CONVENTION, 8 May 2004 (2004-05-08), páginas 1-14, describe: un aparato para generar una señal de salida multi-canal que tiene K canales de salida, correspondiendo la señal de salida multi-canal a una señal de entrada multi-canal que tiene C canales de entrada, empleando E canales de transmisión, los E canales de transmisión representando un resultado de una operación de mezclado hacia abajo que tiene C canales de entrada como entrada, y utilizar información paramétrica relacionada con los canales de entrada, donde E es 2, C es > E, y K es > 1 y ≤C.

[0007] La codificación de intensidad estéreo se describe en la pre-impresión AES 3799, "Intensity Stereo Coding", J. Herre, K. H. Brandenburg, D. Lederer, Febrero 1994, Amsterdam. En general, el concepto de intensidad estéreo se basa en una transformada de eje principal que debe ser aplicada a los datos de ambos canales de audio estereofónicos. Si la mayoría de los puntos de datos se concentran alrededor del primer eje principal, puede lograrse una ganancia de codificación mediante la rotación de ambas señales en un ángulo determinado antes de la codificación. Esto es, sin embargo, no siempre es cierto para las técnicas de producción estereofónicas reales. Por lo tanto, esta técnica se ha modificado mediante la exclusión de la segunda componente ortogonal de la transmisión en el flujo de bits. Así, las señales reconstruidas para los canales izquierdo y derecho consisten en versiones de la misma señal transmitida ponderadas o escaladas de manera diferente. Sin embargo, las señales reconstruidas difieren en su amplitud, pero son idénticas en cuanto a su información de fase. Las envolventes energía-tiempo de ambos canales de audio originales, sin embargo, se conservan por medio de la operación de escalado selectivo, que normalmente opera de una manera selectiva en frecuencia. Esto se ajusta a la percepción humana del sonido en las frecuencias altas, donde las señales espaciales dominantes están determinadas por las envolventes de energía.

[0008] Además, en implementaciones prácticas, la señal transmitida, es decir, el canal portador se genera a partir de la señal suma de los canales izquierdo y derecho en lugar de girar ambos componentes. Además, este procesamiento, es decir, la generación de parámetros de intensidad estéreo para realizar la operación de escalado, se realiza en frecuencia selectiva, es decir, de forma independiente para cada banda de factor de escala, es decir, partición de frecuencia de codificación. Preferentemente, los dos canales se combinan para formar un combinado o canal "portador", y, sumado al canal combinado, la información de intensidad estéreo se determina como dependiente de la energía del primer canal, la energía del segundo canal o la energía del combinado o canal.

[0009] La técnica de BCC se describe en el documento de la convención AES 5574 5574, "Binaural cue coding applied to stereo y multi-channel audio compression", C. Faller, F. Baumgarte, mayo de 2002, Munich. En la codificación de BCC, un número de canales de entrada de audio se convierten en una representación espectral, utilizando una transformada basada en DFT con ventanas solapadas. El espectro uniforme resultante se divide en particiones no superpuestas cada una con un índice. 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 () se calculan para cada partición para cada marco k. El ICLD y el ICTD se cuantifican y se codifican resultando en un flujo de bits BCC. Las diferencias de nivel entre canales y diferencias de tiempo entre canales se dan para cada canal con respecto a un canal de referencia. A continuación, los parámetros se calculan de acuerdo con las fórmulas prescritas, que dependen de las determinadas particiones de la señal a procesar.

[0010] En un lado decodificador, el descodificador recibe una señal mono y el flujo de bits BCC. La señal mono se transforma en el dominio de la frecuencia y entra en un bloque de síntesis espacial, que también recibe los valores decodificados ICLD y ICTD. En el bloque de síntesis espacial, se utilizan los valores de los parámetros BCC (ICLD y) para realizar una operación de ponderación de la señal mono con el fin de sintetizar las señales del canal múltiple, que, después de una conversión de frecuencia / tiempo, representan una reconstrucción de la señal audio multi-canal original.

[0011] En el caso de BCC, el módulo estéreo conjunto 60 es operativo para dar como salida información de lado de canal de tal manera que los datos de canal paramétricos se cuantifican y codifican en parámetros ICLD o TICD, donde uno de los canales originales se utiliza como canal de referencia para la codificación de la información de lado de canal.

[0012] Normalmente, el canal portador está formado por la suma de los canales participantes originales.

[0013] Naturalmente, las técnicas anteriores sólo proporcionan una representación mono para un decodificador, que sólo puede procesar el canal portador, pero no es capaz de procesar los datos paramétricos para generar una o más aproximaciones de más de un canal de entrada.

[0014] La técnica de codificación de audio conocida como codificación de señal binaural (BCC) también está bien descrita en las solicitudes de patente americana publicadas US 2003, 0219130 A1, 2003/0026441 A1 y 2003/0035553 A1. Referencia adicional también se hace a "Binaural Cue Coding. Part II: Schemes y Applications", C. Faller y F. Baumgarte, IEEE Trans. On Audio y Speech Proc., Vol. 11, No. 6, Nov. 2993. Las publicaciones de solicitud de patentes americanas citadas y las dos publicaciones técnicas citadas sobre la técnica BCC escritas por Faller y Baumgarte se incorporan aquí por referencia en su totalidad.

[0015] En lo que sigue, se elabora un típico esquema genérico BCC para codificación de audio multi-canal con más detalle con referencia a las figuras 11 a 13. La figura 11 muestra un esquema de codificación de señal binaural genérico para la la codificación / transmisión de señales de audio multi-canal. La señal de entrada de audio multi-canal en una entrada 110 de un codificador 112 se mezcla en un bloque de mezcla hacia abajo 114. En el presente ejemplo, la señal en la entrada multi-canal original 110 es una señal de sonido envolvente de 5 canales que tiene un canal delantero izquierdo, un canal delantero derecho, un canal izquierdo envolvente, un canal derecho envolvente y un canal central. Por ejemplo, el bloque de mezclado hacia abajo 114 produce una señal de suma por una simple adición de estos cinco canales en una señal mono. Otros esquemas downmix (mezcla hacia abajo) son conocidos en la técnica de tal manera que, utilizando una señal de entrada de múltiples canales, se puede obtener una señal downmix que tiene un solo canal. Este canal único se da como salida en una línea de señal de suma 115. Una información de lado obtenida por un bloque de análisis BCC 116 se da como salida a una línea de información de lado 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 Recientemente, el bloque de análisis BCC 116 se ha mejorado para calcular también valores de correlación entre canales (valores ICC). La señal de suma y la información de lado se transmiten, preferentemente de forma cuantificada y codificada, a un decodificador BCC 120. El decodificador BCC descompone la señal de suma transmitida en un número de sub-bandas y aplica escalados, retardos y otras operaciones para generar las sub-bandas de las señales de audio multi-canal de salida. Este procesamiento se lleva a cabo de tal manera que los parámetros ICLD, ICTD y ICC (señales) de una señal multi-canal reconstruida en una salida 121 son similares a las señales respectivas para la señal multi-canal original en la

entrada 110 en el codificador BCC 112. Para este fin, el descodificador BCC 120 incluye un bloque de síntesis BCC 122 y un bloque de procesamiento de información de lado 123.

[0016] A continuación, se explica la construcción interna del bloque de síntesis BCC 122 con referencia a la figura 12. La señal de suma en la línea 115 se introduce en una unidad de conversión tiempo / de frecuencia o banco de filtros FB 125. En el bloque de salida 125, existe un número N de señales de sub banda o, en un caso extremo, un bloque de coeficientes espectrales, cuando el banco de filtros de audio 125 realiza una transformada 1: 1, es decir, una transformada que produce N coeficientes espectrales de N muestras en el dominio del tiempo.

[0017] El bloque de síntesis BCC 122 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 inversos IFB 129. En la salida de la etapa 129, la señal de audio multi-canal reconstruida que tiene por ejemplo cinco canales en el caso de un sistema de sonido envolvente de 5-canales, puede ser la salida a un conjunto de altavoces 124, tal como se ilustra en la figura 11.

[0018] Tal como se muestra en la figura 12, las señales de entrada (n) se convierten en el dominio de la frecuencia o el dominio del banco de filtros por medio del elemento 125. La salida de señal por el elemento 125 se multiplica de tal manera que se obtienen varias versiones de la misma señal tal como se ilustra por el nodo de multiplicación 130. El número de versiones de la señal original es igual al número de canales de salida en la señal de salida, que hay que reconstruir cuando, en general, cada versión de la señal original en el nodo 130 es sometida a un cierto retardo d1, d2, ..., di, ..., dN. Los parámetros de retardo se calculan mediante el bloque de procesamiento de información de lado 123 en la figura 11 y se derivan de las diferencias de tiempo entre canales determinadas por el bloque de análisis BCC 116.

[0019] Lo mismo es cierto para los parámetros de multiplicación a1, a2, ..., ai, ..., aN, que también se calculan mediante el bloque de procesamiento de información de lado 123 a partir de las diferencias de nivel entre canales tal como fueron calculadas por el bloque de análisis BCC 116.

[0020] Los parámetros ICC calculados por el bloque de análisis BCC 116 se utilizan para controlar la funcionalidad del bloque 128 de tal manera que se obtienen ciertas correlaciones entre las señales retardadas y manipuladas de nivel en las salidas del bloque 128. Es de notar aquí que el orden entre las etapas 126, 127, 128 puede ser diferente del caso que se muestra en la figura 12.

[0021] Es de notar aquí que, en procesamiento a lo largo de un marco de una señal de audio, el análisis BCC se realiza a lo largo de un marco, es decir, variable en el tiempo, y también a lo largo de frecuencia. Esto significa que, para cada banda espectral, se obtienen los parámetros BCC. Esto significa que, en el caso de que el banco de filtros de audio 125 descomponga la señal de entrada en, por ejemplo, 32 señales de paso de banda, 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 figura 11, que se muestra en detalle en la figura 12, realiza una reconstrucción que también se basa en las 32 bandas en el ejemplo.

[0022] En lo que sigue, se hace referencia a la figura 13 que muestra una configuración para determinar ciertos parámetros BCC. Normalmente, los parámetros ICLD, ICTD e ICC se pueden definir entre pares de canales. Sin embargo, se prefiere determinar los parámetros ICLD e ICTD entre un canal de referencia y cada uno de los otros canales. Esto se ilustra en la figura 13A.

[0023] Los parámetros ICC se pueden definir de diferentes maneras. Más en general, se podrían estimar los parámetros ICC en el codificador de entre todos los pares de canales posibles, tal como se indica en la figura 13B. En este caso, un decodificador sintetizaría los ICC de manera que fueran aproximadamente los mismos que en la señal multi-canal original de entre todos los pares de canales posibles. Sin embargo, se propuso estimar los parámetros ICC sólo entre los más fuertes de los dos canales en cada momento. Este esquema se ilustra en la figura 13C, donde se muestra un ejemplo, en el que en un instante de tiempo se estima un parámetro ICC entre los canales 1 y 2, y, en otro ejemplo de tiempo, se calcula un parámetro ICC entre los canales 1 y 5. El decodificador sintetiza a continuación la correlación entre canales entre los más fuertes de los canales en el decodificador y se aplica una regla heurística para el cálculo y la síntesis de la coherencia entre canales para los pares de canales restantes.

[0024] En cuanto al cálculo de, por ejemplo, los parámetros de multiplicación a1, aN basado en los parámetros de transmisión ICLD, se hace referencia al documento de la convención AES 5574 citado más arriba. Los parámetros ICLD representan una distribución de energía en una señal multi-canal original. Sin pérdida de generalidad, se muestra en la figura 13 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 a1, ..., aN se derivan de los parámetros ICLD tales 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

simple para determinar estos parámetros es un proceso de dos etapas, en el que, en una primera etapa, se establece el factor de multiplicación para el canal frontal izquierdo a la unidad, mientras que los factores de multiplicación para los otros canales de la figura 13 se establecen en los valores ICLD transmitidos. Luego, en una segunda etapa, se calcula la energía de los cinco canales y se compara con la energía de la señal de suma transmitida. A continuación, todos los canales se escalan a la baja 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 de transmitida.

[0025] Naturalmente, hay otros métodos para el cálculo de los factores de multiplicación, que no se basan en el proceso de 2 etapas, pero que sólo necesitan un proceso de una etapa.

[0026] En cuanto a los parámetros de retardo, es de señalar que el retardo de parámetros ICTD, que se transmiten desde un codificador BCC pueden ser utilizados directamente, cuando el parámetro de retardo d1 para el canal frontal izquierdo se establece en cero. Aquí no es necesario reescalado alguno, ya que un retardo no altera la energía de la señal.

[0027] En cuanto a la medida de coherencia inter-canal ICC transmitida desde el codificador BCC al decodificador BCC, es de notar aquí que puede hacerse una manipulación de coherencia mediante la modificación de los factores de multiplicación a1, ..., an como por ejemplo multiplicando los factores de ponderación de todas las sub-bandas con números aleatorios con un intervalo de [20log10(-6) y 20log10 (6)]. La secuencia pseudo-aleatoria se elige preferentemente de manera que la variación sea aproximadamente constante para todas las bandas críticas, y la media sea cero dentro de cada banda crítica. La misma secuencia se aplica a los coeficientes espectrales para cada marco diferente. Así, la anchura de la imagen auditiva se controla mediante la modificación de la varianza de la secuencia pseudo-aleatoria. La modificación de varianza se puede realizar en bandas individuales que son de anchura de banda crítica. Esto permite la existencia simultánea de varios objetos en una escena auditiva, teniendo cada objeto una anchura de imagen diferente. Una distribución de amplitud adecuada para la secuencia pseudo-aleatoria es una distribución uniforme en una escala logarítmica tal como se describe en la publicación de solicitud de patente americana 2003/0219130 A1. Sin embargo, todo el procesamiento de síntesis BCC se relaciona con un canal de entrada único transmitido como la suma de señales del codificador BCC al decodificador BCC tal como se muestra en la figura 11.

[0028] Para transmitir los cinco canales en una forma compatible, es decir, en un formato de cadena de bits, que también es comprensible para un descodificador estéreo normal, la técnica llamada de matrizado se ha utilizado tal como se describe en "", G. Theile y G. Stoll, pre-impresión de AES 3403, Octubre de 1992, San Francisco. Los cinco canales de entrada L, R, C, Ls, y Rs se introducen en un dispositivo de matrizado realizando una operación de matrizado para calcular los canales estéreo básicos o compatibles Lo, Ro, a partir de los cinco canales de entrada. En particular, estos canales estéreo básicos Lo/Ro se calculan tal como se establece a continuación:

[0029] X e Y son constantes. Los otros tres canales C, Ls, Rs se transmiten tal como están en una capa de extensión, además de una capa básica estéreo, que incluye una versión codificada de las señales estéreo básicas Lo/Ro. Con respecto a la corriente de bits, esta capa estéreo básica Lo/Ro incluye un encabezamiento, información tal como factores de escala y muestras de sub-banda. La capa de extensión multi-canal, es decir, el canal central y los dos canales de sonido envolvente se incluyen en el ámbito de extensión multi-canal, que también se denomina campo de datos auxiliares.

[0030] En el lado decodificador, se realiza una operación de matrizado inversa con el fin de formar reconstrucciones de los canales izquierdo y derecho en la representación de cinco canales utilizando los canales de base estéreo Lo, Ro y los tres canales adicionales. Adicionalmente, los tres canales adicionales son decodificados de la información auxiliar con el fin de obtener una representación decodificada de cinco canales o envolvente de la señal de audio multi-canal original.

[0031] Otro enfoque para la codificación multi-canal se describe en la publicación "Improved MPEG-2 audio multichannel encoding", B. Grill, J. Herre, K. H. Brandenburg, E. Eberlein, J. Koller, J. Mueller, AES preprint 3865, February 1994, Amsterdam, en el que, con el fin de alcanzar la compatibilidad hacia atrás, se consideran unos modos compatibles hacia atrás. Para este fin, se utiliza una matriz de compatibilidad para obtener dos así llamados canales downmix Lc, Rc a partir de los cinco canales de entrada originales. Además, es posible seleccionar dinámicamente los tres canales de transmisión auxiliares como datos auxiliares.

[0032] Con el fin de explotar la irrelevancia estéreo, se aplica una técnica estéreo conjunta a grupos de canales, por ejemplo, a los tres canales frontales, es decir, al canal izquierdo, el canal derecho y el canal central. Para este fin, estos tres canales se combinan para obtener un canal combinado. Este canal combinado se cuantifica y se empaqueta en el flujo de bits. A continuación, este canal combinado junto con la información estéreo conjunta correspondiente se introduce en un módulo de decodificación estéreo conjunto para obtener conjuntos de canales estéreo decodificados, es decir, un canal izquierdo decodificado estéreo conjunto, un canal derecho decodificado estéreo conjunto y un canal central decodificado estéreo conjunto. Estos canales decodificados estéreo conjuntos, junto con el canal izquierdo de sonido envolvente y el canal derecho de sonido envolvente se entran en un bloque de matriz de compatibilidad para formar los canales downmix primero y segundo Lc, Rc. Entonces, las versiones cuantificadas de los dos canales downmix y una versión cuantificada del canal combinado se empaquetan en el flujo de bits, junto con parámetros de codificación estéreo conjuntos.

[0033] Por lo tanto, utilizando la codificación estéreo de intensidad, se transmite un grupo de señales de canales originales independientes dentro de una sola porción de datos "portadores". Entonces el decodificador reconstruye las señales involucradas como datos idénticos, que son reescalados de acuerdo con sus envolventes energía/ tiempo originales. En consecuencia, una combinación lineal de los canales transmitidos dará lugar a resultados, que son bastante diferentes del downmix original. Esto se aplica a cualquier tipo de codificación estéreo conjunta basada en el concepto de intensidad estéreo. Para un sistema de codificación que proporcione canales downmix compatibles, hay una consecuencia directa: La reconstrucción por supresión de matriz, tal como se describe en la publicación anterior, sufre de artefactos causados por la reconstrucción imperfecta. El uso de un así llamado esquema de predistorsión estéreo conjunto, en el que se realiza una codificación estéreo conjunta de los canales izquierdo, derecho y central antes de la matrización en el codificador, alivia este problema. De esta manera, el esquema de supresión de matriz para la reconstrucción introduce menos artefactos, ya que, en el lado codificador, se han utilizado las señales decodificadas estéreo conjuntas para generar los canales downmix. Así, el proceso de reconstrucción imperfecta se desplaza en los canales downmix compatibles Lc y Rc, donde es mucho más probable que se enmascare por la propia señal de audio.

[0034] Aunque este sistema ha dado lugar a un menor número de artefactos por la supresión de matriz en el lado decodificador, tiene no obstante algunos inconvenientes. Un inconveniente es que los canales downmix compatibles estéreo Lc y Rc no se derivan de los canales originales, sino de las versiones codificada / decodificada de intensidad estéreo de los canales originales. Por lo tanto, las pérdidas de datos debidas al sistema de codificación de intensidad estéreo se incluyen en los canales downmix compatibles. Por lo tanto, un decodificador solamente estéreo, que sólo decodifica los canales compatibles en lugar de los canales codificados de intensidad estéreo mejorada, proporciona una señal de salida, que está afectada por las pérdidas de datos inducidas de intensidad estéreo.

[0035] Además, debe transmitirse un canal completo adicional, además de los dos canales downmix. Este canal es el canal combinado, que se forma por medio de codificación estéreo conjunta de los canales izquierdo, derecho y central. Además, la información de intensidad estéreo para reconstruir los canales originales, L, R, C a partir del canal combinado también tiene que ser transmitida al decodificador. En el decodificador, se realiza un matrizado inverso, es decir, una operación de supresión de matriz para obtener los canales de sonido envolvente a partir de los dos canales downmix. Además, los canales izquierdo, derecho y central originales se aproximan por decodificación estéreo conjunta a través del canal combinado transmitido y los parámetros estéreo conjuntos transmitidos. Es de señalar que los canales izquierdo, derecho y central originales se derivan por decodificación estéreo conjunta del canal combinado.

[0036] Una mejora del esquema BCC mostrado en la figura 11 es un esquema BCC con al menos dos canales de transmisión de audio de modo que se obtiene un procesamiento estéreo compatible. En el codificador, C canales de entrada se mezclan hacia abajo a E canales de transmisión de audio. Las señales ICTD, ICLD e ICC señales entre ciertos pares de canales de entrada se calcula en función de la frecuencia y el tiempo. Las señales estimadas se transmiten al decodificador como información lateral. Un esquema BCC con C canales de entrada y E canales transmitidos se denota como C-2-E BCC.

[0037] En términos generales, el procesamiento BCC es un procesamiento posterior de frecuencia selectiva y variable en el tiempo de los canales transmitidos. En lo que sigue, con el entendimiento implícito de esto, no se introducirá un índice de banda de frecuencia. En cambio, se asumen las variables xn, Sn, yn, an, etc. como vectores de dimensión (1,f), donde f indica el número de bandas de frecuencia.

[0038] El así llamado esquema regular BCC se describe en C. Faller y F. Baumgarte, "Binaural Cue Coding applied to stereo and multi-channel audio compression," en preimpresión 112th Conv. Aud. Engl. Soc., May 2002, F. Baumgarte and C. Faller, "Binaural Cue Coding - Part I: Psychoacoustic fundamentals and design principles," IEEE Trans. On Speech and Audio Proc., vol. 11, no. 6, Nov. 2003, y C. Faller y F. Baumgarte, "Binaural Cue Coding - Part II; Schemes and applications," IEEE Trans. On Speech y Audio Proc., vol. 11, no. 6, Nov. 2003. Aquí, solo se dispone de un canal de transmisión de audio tal como se muestra en la figura 11, que es una extensión compatible

con versiones anteriores de los sistemas mono existentes para la reproducción audio estéreo o multicanal. Puesto que el canal de transmisión de audio único es una señal mono válida, es adecuado para la reproducción de los receptores de la herencia.

[0039] Sin embargo, la mayor parte de la infraestructura de radiodifusión de audio instalada (radio analógica y digital, televisión, etc) y los sistemas de almacenamiento de audio (discos de vinilo, casete compacto, disco compacto, vídeo VHS, MP3 de almacenamiento de sonido, etc) están basados en estéreo de dos canales. Por otro lado, "los sistemas de cine en casa", según la norma 5.1 (Rec. ITU-R BS.775, Multi-Channel Stereophonic Sound sytem with or without Accompanying Picture, ITU, 1993, http://www.itu.org) son cada vez más populares. Por lo tanto, con dos canales de transmisión BCC (C-to-2 BCC), tal como se describe en J. Herre, C. Faller, C. Ertel, J. Hilpert, A. Hoelzer, y C. Spenger, "MP3 Surround: Efficient y compatible coding de multi-channel audio," en la pre-impresión 116th Conv. Aud. Eng. Soc., Mayo 2004, es particularmente interesante para ampliar los sistemas estéreo existentes para sonido envolvente multicanal. A este respecto, también se hace referencia a la solicitud de patente americana " Apparatus and method for constructing a multi-channel output signal or for generating a downmix signal ", Número de serie US 10/762,100, presentada el 20 de enero de 2004.

[0040] En el dominio analógico, los algoritmos de matrizado tales como "Dolby Surround", "Dolby Pro Logic", y "Dolby Pro Logic II" (J. Hull, "Surround sound past, present, and future," Techn. Rep., Dolby Laboratories, 1999, www.dolby.com/tech/; R. Dressler, "Dolby Surround Prologic II Decoder - Principles of operation," Techn Rep., Dolby Laboratories, 2000, www.dolby.com/tech/) ya son populares desde hace años. Estos algoritmos aplican el "matrizado" para la asignación de los 5.1 canales de audio a un par de canales estéreo compatibles. Sin embargo, los algoritmos de matrizado sólo proporcionan una flexibilidad significativamente reducida y calidad en comparación con canales de audio discretos tal como se describe en J. Herre, C. Faller, C. Ertel, J. Hilpert, A. Hoelzer, y C. Spenger, "MP3 Surround: Efficient and compatible coding of multi-channel audio," en Preimpresión 116th Conv. Aud. Eng. Soc., May 2004. Si ya se consideran las limitaciones de los algoritmos de matrizado a la hora de mezclar señales de audio para sonido envolvente 5.1, algunos de los efectos de esta imperfección se pueden reducir tal como se indica en J. Hilson, "Mixing with Dolby Pro Logic II Technology," Tech. Rep., Dolby Laboratories, 2004, www.dolby.com/tech/PLII.Mixing.JimHilson.html.

[0041] El C-to-2 BCC puede ser visto como un esquema con una funcionalidad similar a un algoritmo de matrizado con información de ayuda de lado adicional. Es, sin embargo, más general en su naturaleza, ya que admite la asignación de cualquier número de canales originales a cualquier número de canales transmitidos. El C-to-E BCC está pensado para el dominio digital y por lo general su información de lado de baja velocidad de bits puede ser incluida en la transmisión de datos existentes de una manera compatible con versiones anteriores. Esto significa que los receptores existentes ignorarán la información de lado adicional y reproducirán directamente los 2 canales transmitidos tal como destaca en J. Herre, C. Faller, C. Ertel, J. Hilpert, A. Hoelzer, y C. Spenger, "MP3 Surround: Efficient and compatible coding of multi-channel audio," en Preimpresión 116th Conv. Aud. Eng. Soc., Mayo 2004. El eterno objetivo es lograr una calidad de audio similar a una transmisión discreta de todos los canales de audio originales, es decir, una calidad mucho mejor de lo que puede esperarse de un algoritmo matrizado convencional.

[0042] En lo que sigue, se hace referencia a la figura 6a a fin de ilustrar el funcionamiento del codificador downmix convencional para generar dos canales de transmisión a partir de cinco canales de entrada, que son un canal izquierdo L o x1, un canal derecho R o x2, un canal central C o x3, un canal izquierdo de sonido envolvente SL o x4 y un derecho de sonido envolvente de canal SR o x5. La situación downmix se muestra esquemáticamente en la figura 6a. Se hace evidente que el primer canal de transmisión y1 se forma con un canal de la izquierda x1, un canal central x3 y el canal izquierdo de sonido envolvente x4. Además, la figura 6 bis establece claramente que el canal de transmisión derecho y2 se forma utilizando el canal derecho x2, el canal central x3 y el canal derecho de sonido envolvente x5.

[0043] La regla generalmente preferida de downmixing o matriz downmixing se muestra en la figura 6c. Se hace evidente que el canal central x3 está ponderado con un factor de ponderación de 1 / [raíz cuadrada de] 2, lo que significa que la primera mitad de la energía del canal central x3 está puesta en el canal de transmisión izquierdo o primer canal de transmisión Lt, mientras que la segunda mitad de la energía en el canal central es introducida en el segundo canal de transmisión o canal derecho de transmisión Tr. Por lo tanto, el downmix asigna los canales de entrada a los canales transmitidos. El downmix está convenientemente descrito por una matriz (m, n), que asigna n muestras de entrada a m muestras de salida. Las entradas de esta matriz son los pesos aplicados a los canales correspondientes antes de sumar para formar el canal de salida correspondiente.

[0044] Existen diferentes métodos de downmix que se pueden encontrar en las recomendaciones de la ITU (Rec. ITU-R BS. 775, Multi-Channel Stereophonic Sound system with or without Accompanying Picture, ITU, 1993, http://www.itu.org). Además, se hace referencia a J. Herre, C. Faller, C. Ertel, J. Hilpert, A. Hoelzer, y C. Spenger, "MP3 Surround: Efficient and compatible coding of multi-channel audio," en Preimpresión 116th Conv. Aud. Eng. Soc., Mayo 2004, Sección 4.2 con respecto a los métodos downmix diferentes. El downmix se puede realizar ya sea

en tiempo o en el dominio de la frecuencia. Puede ser que sea variable en el tiempo de una forma de adaptación de la señal o dependiente de la frecuencia (banda).

La asignación de canal se muestra por la matriz a la derecha de la figura 6a y se da como sigue:

$$IN_5 = \begin{bmatrix} left \\ right \\ center \\ rear - left \\ rear - right \end{bmatrix}$$

[0045] Por lo tanto, para el caso importante de 5-to-2 BCC, se calcula un canal transmitido desde la derecha, posterior derecha y central, y el otro canal transmitido desde la izquierda a derecha, posterior izquierda y central, que corresponde a una matriz downmix que también se muestra en la figura 6c., por ejemplo

$$D_{52} = \begin{bmatrix} 1 & 0 & \frac{1}{\sqrt{2}} & 1 & 0 \\ 0 & 1 & \frac{1}{\sqrt{2}} & 0 & 1 \end{bmatrix}$$

[0046] En esta matriz downmix, los factores de ponderación se pueden elegir de tal manera que la suma de los cuadrados de los valores en cada columna sea uno, de tal manera que la potencia de cada señal de entrada contribuya igualmente a las señales mezcladas hacia abajo. Por supuesto podrían utilizarse también otros esquemas downmix.

[0047] En particular, se hace referencia a la figura 6b o 7b, que muestra una implementación específica de un esquema downmix del codificador. Se muestra el procesamiento de una sub-banda. En cada sub-banda, los factores de escalado E1 y E2 están controlados para "ecualizar" el volumen de los componentes de la señal en la señal mezclada hacia abajo. En este caso, el downmix se realiza en dominio de frecuencia, designando la variable N (Fig. 7b) un índice de tiempo de sub-banda en el dominio de la frecuencia y siendo k el índice del bloque de señal en el dominio del tiempo transformada. En particular, se destaca el dispositivo de ponderación para la ponderación del canal central antes de introducir la versión ponderada del canal central en el canal izquierdo transmitido de los canales y el canal derecho transmitido por los respectivos dispositivos de suma.

[0048] La operación Upmix correspondiente en el decodificador se muestra con respecto a las figuras 7a, 7b y 7c. En el decodificador debe calcularse un Upmix, que asigna el canal transmitido a los canales de salida. El Upmix está convenientemente descrito por una matriz (i, j) (i filas, columnas j), que asigna i muestras transmitidas a j muestras de salida. Una vez más, las entradas de esta matriz son los pesos aplicados a los canales correspondientes antes de sumar para formar el canal de salida correspondiente. El Upmix se puede realizar ya sea en el dominio del tiempo o de la frecuencia. Adicionalmente, podría ser variable en el tiempo de una manera adaptativa a la señal dependiente de la frecuencia (banda). A diferencia de la matriz downmix, los valores absolutos de las entradas de la matriz no representan los pesos finales de los canales de salida, puesto que estos canales upmixed se modificadan aún más en caso de tratamiento BCC. En particular, la modificación se lleva a cabo utilizando la información proporcionada por las señales espaciales tales como ICLD, etc. Aquí, en este ejemplo, todas las entradas están a 0 o 1.

[0049] La figura 7a muestra la situación upmixing para un sistema de sonido envolvente de 5 altavoces. Al lado de cada altavoz, se muestra el canal de base que se utiliza para la síntesis BCC. En particular, con respecto al canal de salida de sonido envolvente izquierdo, se utiliza un primer canal transmitido Y1. Lo mismo es cierto para el canal izquierdo. Este canal se usa como un canal base, también llamado el "canal izquierdo transmitido".

[0050] En cuanto a canal de salida derecho y al canal de salida de sonido envolvente derecho, también utilizan el mismo canal, es decir, el canal segundo o derecho transmitido y2. En lo que respecta al canal central, es de notar aquí que el canal base para la síntesis BCC del canal central está formado de acuerdo con la matriz upmixing que se muestra en la figura 7c, es decir, mediante la adición de los dos canales transmitidos.

[0051] El proceso de generación de la señal de salida de 5 canales, dados los dos canales transmitidos se muestra en la figura 7b. Aquí, el Upmix se realiza en dominio de la frecuencia donde la variable n indica un índice de tiempo de sub-banda en el dominio de la frecuencia, y k el índice del bloque de señal de dominio en el tiempo transformada. Es de notar aquí que la síntesis ICTD e ICC ICC se aplica entre los pares de canales para los cuales se utiliza el mismo canal base, es decir, entre el izquierdo e izquierda posterior, derecho y derecho posterior, respectivamente. Los dos bloques indicados A en la figura 7b incluyen esquemas para la síntesis ICC de dos canales.

[0052] La información de lado estimada en el codificador, que es necesaria para calcular todos los parámetros para la señal de decodificador de síntesis de salida incluye las siguientes señales: ΔL12, ΔL13, ΔL14, ΔL15, τ14, τ25, c14, y c25 (ΔLij es la diferencia de nivel los canales i y j, τij es la diferencia de tiempo entre el canal i y j, y cij es un coeficiente de correlación entre el canal i y j). Es de notar aquí que pueden utilizarse también otras diferencias de nivel. Existe el requisito de que haya suficiente información disponible en el decodificador para la computación por ejemplo, de los factores de escala, los retardos etc para la síntesis BCC.

[0053] En lo que sigue, se hace referencia a la figura 7d con el fin de ilustrar aún más la modificación de nivel para cada canal, es decir, el cálculo de ai y la posterior normalización en general, que no se muestra en la figura 7b. Preferentemente, diferencias de nivel entre canales ΔLi se transmiten como información lateral, es decir como ICLD. Aplicado a la señal del canal, se tiene que usar la relación exponencial entre el canal de referencia Fref y y un canal que se calcula, es decir Fi. Esto se muestra en la parte superior de la figura 7d.

[0054] Lo que no se muestra en la figura 7b es la normalización global posterior o final, que puede tener lugar antes de los bloques de correlación A o después de los bloques de correlación A. Cuando los bloques de correlación afectan a la energía de los canales ponderados por ai, la normalización global debería tener lugar después de los bloques de correlación A. Para asegurarse de que la energía de todos los canales de salida es igual a la energía de todos los canales transmitidos, el canal de referencia se escala tal como se muestra en la figura 7d. Preferentemente, el canal de referencia es la raíz de la suma de los cuadrados de los canales trasmitidos.

[0055] En lo que sigue, se describen los problemas asociados con estos sistemas downmix / upmixing. Cuando se considera el esquema de BCC 5-a-2 tal como se ilustra en la figura 6 y la figura 7, se aprecia lo siguiente.

[0056] El canal central original se introduce en los dos canales transmitidos y, por consiguiente, también en los canales de salida reconstruidos izquierdo y derecho.

[0057] Además, en este esquema, la contribución central común tiene la misma amplitud en ambos canales de salida reconstruidos.

[0058] Además, la señal central original se sustituye durante la decodificación por una señal central, que se deriva de los canales izquierdo y derecho transmitidos y, por tanto, no puede ser independiente de (es decir, no correlacionado con) los canales izquierdo y derecho reconstruidos.

[0059] Este efecto tiene consecuencias desfavorables sobre la calidad del sonido percibido para señales con una imagen de sonido muy amplia que se caracteriza por un alto grado de decorrelación (es decir, baja coherencia) entre todos los canales de audio. Un ejemplo de estas señales es el sonido de un público aplaudiendo, al utilizar diferentes micrófonos con un espaciado suficientemente amplio para la generación de las señales multi-canal originales. Para tales señales, la imagen de sonido del sonido decodificado se estrecha y se reduce su amplitud natural.

Resumen de la invención

[0060] Es objeto de la presente invención proporcionar un concepto de reconstrucción multi-canal de mayor calidad que se traduce en una señal de salida multi-canal que tiene una percepción del sonido mejorada.

[0061] Según un primer aspecto de esta invención, este objeto se logra con un aparato para generar una señal de salida multi-canal que tiene K canales de salida, correspondiendo la señal de salida multi-canal a una señal de entrada multi-canal que tiene C canales de entrada, empleando E canales de transmisión, representando los E canales de transmisión un resultado de una operación de mezclado hacia abajo que tiene C canales de entrada como entrada, y utilizar información paramétrica lateral relacionada con los canales de entrada, donde E es ≥2, C es > E, y K es > 1 y ≤C, y donde la operación de mezclado hacia abajo es eficaz para introducir un primer canal de entrada en un primer canal de transmisión y en un segundo canal de transmisión, y para introducir adicionalmente un segundo canal de entrada en el primer canal de transmisión, que comprende: un calculador de canal de anulación para calcular un canal de anulación empleando información relacionada con el primer canal de entrada incluido en el primer canal de transmisión, el segundo canal de transmisión o la información paramétrica de lado; un combinador para combinar el canal de anulación y el primer canal de transmisión o una versión procesada de estos para obtener un segundo canal de base, en el que una influencia del primer canal de entrada es reducida comparado con la influencia del primer canal de entrada en el primer canal de transmisión; y un canal reconstructor para reconstruir un segundo canal de salida correspondiente al segundo canal de entrada empleando el segundo canal de base e información paramétrica lateral relacionada con el segundo canal de entrada, y para reconstruir un primer canal de salida correspondiente al primer canal de entrada empleando un primer canal de base que sea diferente del segundo canal de base en el hecho de que la influencia del primer canal es mayor comparado con el segundo canal de base, e información paramétrica lateral relacionada con el primer canal de entrada.

[0062] Según un segundo aspecto de la presente invención, este objeto se logra con un procedimiento para generar una señal de salida multi-canal que tiene K canales de salida, correspondiendo la señal de salida multi-canal a una señal de entrada multi-canal que tiene C canales de entrada, empleando E canales de transmisión, representando los E canales de transmisión un resultado de una operación de mezclado hacia abajo que tiene C canales de entrada como entrada, y utilizar información paramétrica lateral relacionada con los canales de entrada, donde E es ≥2, C es > E, y K es > 1 y ≤C, y donde la operación de mezclado hacia abajo es eficaz para introducir un primer canal de entrada en un primer canal de transmisión y en un segundo canal de transmisión, y para introducir adicionalmente un segundo canal de entrada en el primer canal de transmisión, que comprende: calcular un canal de anulación empleando información relacionada con el primer canal de entrada incluido en el primer canal de transmisión, el segundo canal de transmisión o la información paramétrica lateral; combinar el canal de anulación y el primer canal de transmisión o una versión procesada de estos para obtener un segundo canal de base, en el que una influencia del primer canal de entrada es reducida comparado con la influencia del primer canal de entrada en el primer canal de transmisión; y reconstruir un segundo canal de salida correspondiente al segundo canal de entrada empleando el segundo canal de base e información paramétrica lateral relacionada con el segundo canal de entrada, y un primer canal de salida correspondiente al primer canal de entrada empleando un primer canal de base que sea diferente del segundo canal de base en el hecho de que la influencia del primer canal es mayor comparado con el segundo canal de base, e información paramétrica lateral relacionada con el primer canal de entrada.

[0063] Según un tercer aspecto de la presente invención, este objeto se logra con un programa de ordenador que tiene un código de programa para realizar el procedimiento para generar una señal de salida multi-canal, cuando el programa se ejecuta en un ordenador.

[0064] Es de notar aquí, que preferentemente, K es igual a C. No obstante, también se podrían reconstruir menos canales de salida, tales como tres canales de salida L, R, C y no reconstruir Ls y Rs. En este caso, los K (= 3) canales de salida corresponden a tres de los C (= 5) canales de entrada L, R, C originales.

[0065] La presente invención se basa en el hallazgo de que, para mejorar la calidad de sonido de la señal de salida de varios canales, se calcula un determinado canal de base mediante la combinación de un canal transmitido y un canal de cancelación, que se calcula en el extremo receptor o decodificador. El canal de cancelación se calcula de manera que el canal de base modificado obtenido mediante la combinación del canal de cancelación y el canal transmitido tiene una influencia reducida del canal central, es decir, el canal que se introduce en ambos canales transmitidos. Dicho en otras palabras, la influencia del canal central, es decir, el canal que se introduce en ambos canales transmitidos, lo que inevitablemente se produce cuando se llevan a cabo las operaciones upmixing downmix y posterior, se reduce en comparación con una situación en la que no se calcula dicho canal de cancelación ni se combina con un canal transmitido.

[0066] En contraste con la técnica anterior, por ejemplo el canal izquierdo transmitido no se utiliza simplemente como canal de base para la reconstrucción el canal izquierdo o de sonido envolvente izquierdo. En contraste con ello, el canal izquierdo transmitido se modifica mediante la combinación con el canal de cancelación de modo que se reduce la influencia del canal central de entrada original en el canal de base para la reconstrucción del canal de salida izquierdo o derecho o incluso se cancela por completo.

[0067] Según la invención, el canal de cancelación se calcula en el decodificador utilizando la información sobre el canal central original, que ya está presente en el decodificador o generador de salida multi-canal. La información sobre el canal central está incluida en el canal izquierdo transmitido, el canal derecho transmitido y la información paramétrica lateral tal como diferencias de nivel, diferencias de tiempo o parámetros de correlación para el canal central. Dependiendo de ciertas realizaciones, toda esta información puede utilizarse para obtener un canal central de cancelación de alta calidad. En otras realizaciones de nivel más bajo, sin embargo, sólo una parte de esta información se utiliza en el canal de entrada central. Esta información puede ser el canal transmitido izquierdo, el canal transmitido derecho o la información paramétrica lateral. Adicionalmente, también se puede utilizar la información estimada en el codificador y transmitida al decodificador.

[0068] Así, en un entorno 5-a-2, el canal izquierdo transmitido o el canal derecho transmitido no se utilizan directamente para la reconstrucción izquierda y derecha, pero se modifican al ser combinados con el canal de cancelación para obtener un canal de base modificado, que es diferente del canal transmitido correspondiente. Preferentemente, un factor de ponderación adicional, que dependerá de la operación downmix realizada en un codificador para generar los canales transmitidos también se incluye en el cálculo del canal de cancelación. En un entorno 5-a-2, se calculan al menos dos canales de cancelación de modo que cada canal transmitido se puede combinar con un canal de cancelación designado para obtener canales de base modificados para la reconstrucción los canales de salida izquierdo e izquierdo envolvente, derecho y derecho envolvente respectivamente.

[0069] La presente invención se puede incorporar en una serie de sistemas o aplicaciones, incluyendo, por ejemplo, consolas de videojuego digitales, reproductores de audio digital, computadoras, receptores de satélite, receptores de cable, receptores de radiodifusión terrestre y sistemas de entretenimiento domésticos.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0070] Unas realizaciones preferidas de la presente invención se describen a continuación con referencia a las figuras adjuntas, en las cuales:

La figura 1 es un diagrama de bloques de un codificador multicanal que produce canales transmitidos y la información de lado paramétrica en los canales de entrada;

La figura 2 es un diagrama esquemático de bloques del aparato preferido para generar una señal de salida multi-canal de acuerdo con la presente invención;

La figura 3 es un diagrama esquemático del aparato de la invención de acuerdo con una primera realización de la presente invención;

La figura 4 es una implementación del circuito de la realización preferida de la figura 3;

La figura 5a es un diagrama de bloques del aparato de la invención de acuerdo con una segunda realización de la presente invención;

La figura 5b es una representación matemática de upmixing dinámico tal como se muestra en la figura 5a;

La figura 6a es un diagrama general para ilustrar la operación downmixing;

La figura 6b es un diagrama de circuito para implementar la operación downmix de la figura 6a;

La figura 6c es una representación matemática de la operación downmix;

La figura 7a es un diagrama esquemático para indicar los canales de base utilizados para upmixing en un entorno estéreo compatible;

La figura 7b es un diagrama de circuito para la aplicación de una reconstrucción multi-canal en un entorno estéreo compatible;

La figura 7c es una presentación matemática de la matriz upmixing utilizada en la figura 7b;

La figura 7d es una ilustración matemática de la modificación de nivel para cada canal y la normalización global posterior; La figura 8 ilustra un codificador; La figura 9 ilustra un descodificador; La figura 10 ilustra un codificador estéreo conjunto del estado de la técnica.

La figura 11 es una representación de diagrama de bloques de un sistema codificador / decodificador BCC del estado de la técnica;

La figura 12 es un diagrama de bloques de una implementación según el estado de la técnica de un bloque de síntesis BCC de la figura 11, y

La figura 13 es una representación de un esquema bien conocido para la determinación de los parámetros ICLD, ICTD e ICC.

[0071] Antes de dar una descripción detallada de realizaciones preferidas se presentará el problema que subyace a la invención y la solución al problema en términos generales. La técnica de la invención para mejorar la anchura auditivo imagen espacial para canales de salida reconstruidos es aplicable a todos los casos en los que un canal de entrada se mezcla en más de uno de los canales transmitidos en un sistema multicanal paramétrico C-a-E. La realización preferida es la aplicación de la invención a un sistema de codificación de señal binaural (BCC). Por simplicidad de discusión, pero sin pérdida de generalidad, la técnica de la invención se ha descrito para el caso específico de un esquema BCC para la codificación / decodificación de señales envolventes 5.1 rodea de forma compatible hacia atrás.

[0072] El problema antes mencionado de la reducción de anchura de la imagen auditiva se produce sobre todo para señales de audio que contienen transitorios rápidos independientes que se repiten desde distintas direcciones, tal como una señal de aplausos de una audiencia en cualquier tipo de grabación en vivo. Mientras que la reducción de anchura de la imagen puede, en principio, ser abordada mediante una resolución de tiempo mayor para la síntesis ICLD, esto daría lugar a una mayor tasa de información de lado y también requiere un cambio en el tamaño de la ventana del banco de filtros de análisis / síntesis utilizado. Es de notar aquí que esta posibilidad adicionalmente resulta en efectos negativos sobre componentes tonales, ya que un aumento de la resolución de tiempo conlleva automáticamente una disminución de la resolución de frecuencia.

[0073] En su lugar, la invención es un concepto simple que no tiene estas desventajas y su objetivo es reducir la influencia del componente del canal central en los canales laterales.

[0074] Tal como se ha discutido en relación con las figuras 7a-7d, los canales de base para los cinco canales de salida reconstruidos 5-a-2 BCC son:

$$\tilde{s}_1(k) = \tilde{y}_1(k) = \tilde{x}_1(k) + \tilde{x}_3(k)/\sqrt{2} + \tilde{x}_4(k)$$

$$\tilde{s}_2(k) = \tilde{y}_2(k) = \tilde{x}_2(k) + \tilde{x}_3(k)/\sqrt{2} + \tilde{x}_5(k)$$

$$\tilde{s}_3(k) = \tilde{y}_1(k) + \tilde{y}_2(k) = \tilde{x}_1(k) + \tilde{x}_2(k) + \sqrt{2}\tilde{x}_3(k) + \tilde{x}_4(k) + \tilde{x}_5(k)$$

$$\tilde{s}_4(k) = \tilde{s}_1(k)$$

$$\tilde{s}_5(k) = \tilde{s}_2(k)$$

[0075] Es de señalar que la componente de la señal del canal central original x3 aparece amplificado 3 dB en la sub-banda de canal de base central s3 (factor $1/\sqrt{2}$) y atenuada 3 dB en las sub-bandas canal de base restantes (canal lateral).

[0076] Con el fin de atenuar aún más la influencia de la componente de señal de canal central en las sub-bandas de canal de base lateral de acuerdo con esta invención, se aplica la idea general siguiente, tal como se ilustra en la figura 2.

[0077] Una estimación de la señal de canal central decodificada final se calcula preferentemente escalándola al nivel objetivo deseado tal como se describe en la información de nivel correspondiente, tal como un valor ICLD en entornos BCC. Preferentemente, esta señal central descodificada se calcula en el dominio espectral con el fin de ahorrar cálculo, es decir, sin procesamiento de bancos de filtro de síntesis.

[0078] Además, esta señal decodificada central o señal reconstruida central, que corresponde al canal de cancelación, se puede ponderar y luego se combina para ambas señales de canal de base de los otros canales de salida. Esta combinación es preferentemente una resta. No obstante, cuando los factores de ponderación tienen un signo diferente, entonces una adición también resulta en la reducción de la influencia del canal central en el canal de base utilizado para la reconstrucción del canal de salida izquierdo o derecho. Este proceso resulta en la formación de un canal de base modificado para la reconstrucción del izquierdo o sonido envolvente izquierdo y derecho o envolvente derecho. Preferentemente un factor de ponderación de -3 dB es el preferido, aunque también es posible cualquier otro valor.

[0079] En lugar de las señales de canal de base transmitido originales tal como se utilizan en la figura 7b, se utilizan unas señales del canal de base modificada las para el cálculo del canal de salida decodificado de los otros canales de salida, es decir, los canales que no sean el canal central.

[0080] A continuación, se discutirá un diagrama de bloques del concepto inventivo con referencia a la figura 2. La figura 2 muestra un aparato para generar una señal de salida multi-canal que tiene K canales de salida, correspondiendo la señal de salida multi-canal a una señal de entrada multi-canal que tiene C canales de entrada, empleando E canales de transmisión, los E canales de transmisión representando un resultado de una operación de mezclado hacia abajo que tiene los C canales de entrada como entrada, y utilizar información paramétrica lateral on los canales de entrada, donde C es $\geq$2, C es > E, y K es > 1 y $\leq$C. Adiciona lmente, la operación de mezclado hacia abajo es eficaz para introducir un primer canal de entrada en un primer canal de transmisión y en un segundo canal de transmisión. El aparato de la invención incluye el calculador de canal de anulación 20 para calcular al menos un canal de anulación 21, que entra en un combinador 22, que recibe, en una segunda entrada 23, el primer canal de transmisión directamente o una versión procesada del primer canal de transmisión. El procesamiento del primer canal de transmisión para obtener la versión procesada del primer canal de transmisión se realiza mediante un procesador 24, que puede estar presente en algunas realizaciones, pero que en general es opcional. El combinador funciona para obtener un segundo canal de base 25 para ser la entrada en un canal reconstructor 26.

[0081] El canal reconstructor utiliza el segundo canal de base 25 y la información paramétrica lateral en el canal de entrada izquierdo original, que son entrados en el canal reconstructor 26 en otra entrada 27, para generar el segundo canal de salida. En la salida del canal reconstructor, se obtiene un segundo canal de salida 28, que puede ser el canal de salida izquierdo reconstruido, que se genera, comparado con el escenario de la figura 7b, por un canal de base, que tiene una influencia pequeña o incluso una influencia totalmente cancelada del canal de entrada central original en comparación con la situación de la figura 7b.

[0082] Mientras que el canal de salida izquierdo generado tal como se muestra en la figura 7b incluye una cierta influencia tal como se ha descrito anteriormente, esta cierta influencia se reduce en el segundo canal de base tal como se ha generado en la figura 2 debido a la combinación del canal de anulación y el primer canal de transmisión o el primer canal de transmisión procesado.

[0083] Tal como se muestra en la figura 2,el calculador de canal de anulación 20 calcula el canal de anulación empleando información sobre el canal central original disponible como un decodificador, es decir información para generar la señal de salida multi-canal. Esta información incluye información paramétrica lateral en el primer canal de entrada 30, o incluye el primer canal de transmisión 31, que también incluye alguna información sobre el canal central a causa de la operación downmix, o incluye el segundo canal de transmisión 32, que también incluye información sobre el canal central, debido a la operación downmix. Preferentemente, toda esta información se utiliza para la reconstrucción óptima del canal central para obtener el canal de anulación 21.

[0084] Tal realización óptima posteriormente se describirá con respecto a la figura 3 y la figura 4. En contraste con la figura 2, la figura 3 muestra el dispositivo de 2-pliegue de la figura 2, es decir, un dispositivo para la cancelación de la influencia del canal central en el canal de base izquierdo S1, así como el canal de base derecha s2.

El canal de cancelación calculadora 20 de la figura 2 incluye un dispositivo de reconstrucción de canal central 20a y un dispositivo de ponderación 20b para obtener el canal de cancelación 21 en la salida del dispositivo de ponderación. El combinador 22 en la figura 2 es un simple restador que es operativo para sustraer el canal de anulación 21 del primer canal de transmisión 21 para obtener -en términos de la figura 2- el segundo canal de base 25 para reconstruir el segundo canal de salida (tal como el canal de salida izquierdo) y,, opcionalmente, también el canal izquierdo de salida de sonido envolvente. El canal central reconstruido x3 (k) se puede obtener en la salida del dispositivo de reconstrucción del canal central 20a.

[0085] La figura 4 indica una forma de realización preferida implementada como un diagrama de circuito, que utiliza la técnica que se ha discutido con respecto a la figura 3. Además, la figura 4 muestra el procesamiento de frecuencia selectiva que es óptimo para ser integrado en un sencillo dispositivo de reconstrucción de frecuencia selectiva BCC.

[0086] La reconstrucción del canal central 26 se lleva a cabo mediante la suma de los dos canales transmitidos en un sumador 40. Entonces, se utiliza la información paramétrica de lado de las diferencias de nivel de canal, o el factor a3derivado de la diferencia de nivel inter-canal tal como se discute en la figura 7d para generar una versión modificada del primer canal de base (en términos de la figura 2), que se introduce en el reconstructor de canal 26 en la entrada del primer canal de base 29 en la figura 2. El canal central reconstruido en la salida del multiplicador 41 puede utilizarse para la reconstrucción del canal de salida central (después de la normalización general que se describe en la figura 7d).

[0087] Para reconocer la influencia del canal central en la reconstrucción de los canales de base izquierdo y derecho, se aplica un factor de ponderación de $1/\sqrt{2}$ que se ilustra por medio de un multiplicador 42 en la figura 4. A continuación, el canal central reconstruido y ponderado de nuevo se alimenta de nuevo a los sumadores 43a y 43b, que corresponden al combinador 22 en la figura 2.

[0088] De este modo, el segundo canal de base s1 o s4 (o s2 y s5) es diferente del canal de transmisión y1 en que la influencia del canal central se reduce en comparación con el caso en la figura 7b.

[0089] Los sub-bandas del canal de base resultantes se dan en términos matemáticos tal como sigue:

$$\tilde{s}_1(k) = \tilde{y}_1(k) - a_3(k)(\tilde{y}_1(k) + \tilde{y}_2(k)) / \sqrt{2}$$

$$\tilde{s}_2(k) = \tilde{y}_2(k) - a_3(k)(\tilde{y}_1(k) + \tilde{y}_2(k)) / \sqrt{2}$$

$$\tilde{s}_3(k) = \tilde{y}_1(k) + \tilde{y}_2(k)$$

$$\tilde{s}_4(k) = \tilde{s}_1(k)$$

$$\tilde{s}_5(k) = \tilde{s}_2(k)$$

[0090] Así, el dispositivo de la figura 4 prevé una sustracción de una estimación de sub-banda del canal central a partir de los canales de base para los canales laterales con el fin de mejorar la independencia entre los canales y, por lo tanto, para proporcionar una mejor anchura espacial de la señal multi-canal de salida reconstruida.

[0091] De acuerdo con otra realización de la presente invención, que se describirá posteriormente con respecto a la figura 5a y la figura 5b, se determina un canal de cancelación diferente del canal cancelación calculado en la figura 3. En contraste con la realización de las figuras 3/ 4, el canal de anulación 21 para calcular el segundo canal de base s1(k) no se deriva del primer canal de transmisión así como el segundo canal de transmisión si no que se deriva solamente del segundo canal de transmisión y2(k) empleando un determinado factor de ponderación xlr, que se ilustra por el dispositivo de multiplicación 51 en la figura 5a. Así, el canal de cancelación 21en la figura 5a es diferente del canal de cancelación de la figura 3, y también contribuye a una reducción de la influencia del canal central en el canal de base s1(k) utilizado para reconstruir el segundo canal de salida, es decir canal de salida izquierdo x1(k).

[0092] En la realización de la figura 5a, también se muestra una realización preferida del procesador 24. En particular, el procesador 24 se implementa como otro dispositivo de multiplicación 52, que aplica una multiplicación por un factor de multiplicación (1-xlr). Preferentemente, tal como se muestra en la figura 1a, el factor de multiplicación aplicado por el procesador 24 al primer canal de transmisión depende del factor de multiplicación 51, que se utiliza para multiplicar el segundo canal transmitido para obtener el canal de cancelación 21. Finalmente, la versión procesada del primer canal transmitido en una entrada 23 al combinador 22 se utiliza para combinar, lo cual consiste en restar el canal de cancelación 21 de la versión procesada del primer canal transmitido. Todo esto resulta de nuevo en el canal de base 25, que tiene una influencia reducida o totalmente cancelada del canal de entrada central original.

[0093] Tal como se muestra en la figura 5a se repite, el mismo procedimiento para obtener el tercer canal de base s2(k) en una entrada en el dispositivo de reconstrucción envolvente derecho / derecho. Sin embargo, tal como se muestra en la figura 5a, el tercer canal de base s2(k se obtiene mediante la combinación de la versión procesada de el segundo canal de transmisión y(k) y otro canal de anulación 53, que se deriva del primer canal de transmisión y1(k) a través de la multiplicación en un dispositivo de multiplicación 54, que tiene un factor de multiplicación xrl, que puede ser idéntico a xlr para un dispositivo 51, para un dispositivo 51, pero que también puede ser diferente de este valor. El procesador para procesar el segundo canal transmitido, tal como se indica en la figura 5a, es un dispositivo de multiplicación 55. El combinador para combinar el segundo canal de anulación 53 y la versión procesada del segundo canal de transmisión y2(k) se ilustrado por el número de referencia 56 en la figura 5a. El calculador de canal de anulación de la figura 2 también incluye un dispositivo para calcular los coeficientes de cancelación, que está indicado por el número de referencia 57 en la figura 5a. El dispositivo 57 es operativo para obtener información de lado paramétrica en el canal central original o de entrada, tales como la diferencia de nivel entre canales, etc Lo mismo es cierto para el dispositivo 20a en la figura 3, donde el dispositivo de reconstrucción de canal central 20a también incluye una entrada para recibir información de lado paramétrica tal como los valores de nivel o diferencias de nivel entre canales, etc.

[0094] La siguiente ecuación

$$\tilde{s}_1(k) = \tilde{y}_1(k) - a_3(k)(\tilde{y}_1(k) + \tilde{y}_2(k)) / \sqrt{2} = \left(1 - \frac{a_3}{\sqrt{2}}\right)\tilde{y}_1(k) - \frac{a_3}{\sqrt{2}}\tilde{y}_2(k)$$

$$\tilde{s}_2(k) = \tilde{y}_2(k) - a_3(k)(\tilde{y}_1(k) + \tilde{y}_2(k)) / \sqrt{2} = \left(1 - \frac{a_3}{\sqrt{2}}\right)\tilde{y}_2(k) - \frac{a_3}{\sqrt{2}}\tilde{y}_1(k)$$

$$x_{lr} = x_{rl} = \frac{a_3}{\sqrt{2}}$$

muestra la descripción matemática de la realización de la figura 5a e ilustra, en el lado derecho del mismo, el procesamiento de cancelación en el calculador de canal de anulación por un lado y los procesadores (21, 24 en la figura 2) por otro lado. En esta realización específica, que se ilustra aquí, los factores xlr y xrl son idénticos entre sí.

[0095] La realización anterior deja claro que la invención incluye una composición de los canales de base de reconstrucción como combinación lineal de adaptación a la señal de los canales transmitidos izquierdo y derecho. Esta topología se ilustra en la figura 5a.

[0096] Cuando se observa desde un ángulo diferente, el dispositivo de la invención también puede entenderse como un procedimiento de upmixing dinámico, en el que se utiliza una matriz upmixing diferente para cada sub-banda y cada instante de tiempo k. Esta matriz upmixing dinámica se ilustra en la figura 5b. Es de señalar que, para cada sub-banda, es decir, para cada salida del dispositivo de banco de filtros en la figura 4, existe esta matriz upmixing U. En cuanto a la manera dependiente del tiempo, es de señalar que la figura 5b incluye el índice de tiempo k. Cuando se tiene información de nivel para cada índice de tiempo, la matriz upmixing cambiaría para cada instante de tiempo a cada instante de tiempo. Sin embargo, cuando se utiliza la misma información de nivel a3 para un bloque completo de valores transformado en una representación en frecuencia por el banco de filtros de entrada FB, entonces estará presente un valor a3 para un bloque completo de por ejemplo 1024 o 2048 valores de muestreo. En este caso, la matriz upmixing cambiaría en el sentido del tiempo de bloque a bloque y no de valor a valor. Sin embargo, existen técnicas para suavizar los valores de nivel paramétricos de manera que se puedan obtener diferentes factores de modificación de amplitud a3 durante el upmixing en una banda de frecuencia determinada.

[0097] Dicho de forma general, también se podrían utilizar diferentes factores para el cálculo de las sub-bandas de canal de salida central y de los factores para "upmixing dinámico", lo que resulta en un factor a3, que es una versión escalada de a3 tal como se ha calculado anteriormente.

[0098] En una realización preferida, la fuerza de ponderación de la cancelación de componente central está controlada de forma adaptativa por medio de una transmisión explícita de información de lado del codificador al decodificador. En este caso, el calculador de canal de anulación 20 que se muestra en la figura 2 incluirá una entrada de control adicional, que recibe una señal de control explícita que podría ser calculada para indicar una interdependencia directa entre los canales izquierdo y central o derecho y central. En este sentido, esta señal de control sería diferente de las diferencias de nivel para el canal central y el canal izquierdo, debido a que estas diferencias de nivel están relacionadas con un tipo de canal de referencia virtual, que podría ser la suma de la energía en el primer canal de transmisión y la suma de la energía en el segundo canal de transmisión tal como se ilustra en la parte superior de la figura 7d.

[0099] Por ejemplo, este parámetro de control podría indicar que el canal central está por debajo de un umbral y se aproxima a cero, mientras que hay una señal en el canal izquierdo o derecho, que está por encima del umbral. En este caso, una reacción adecuada del calculador de canal de anulación a una señal de control correspondiente sería apagar la cancelación del canal y aplicar un esquema upmixing normal tal como se muestra en la figura 7b para evitar la "sobre-cancelación" del canal central, que no está presente en la entrada. En este sentido, esta sería una especie extrema de controlar la fuerza de ponderación tal como se describió anteriormente.

[0100] Preferentemente, como se evidencia en la figura 4, no se realiza operación de tratamiento de retardo de tiempo para el cálculo del canal central de reconstrucción. Esto es ventajoso en que la realimentación funciona sin tener que tomar en consideración los retrasos de tiempo. Sin embargo, esto puede ser obtenido sin pérdida de calidad, cuando el canal central original se utiliza como canal de referencia para calcular las diferencias de tiempo di. Lo mismo es cierto para cualquier medida de correlación. Se prefiere no realizar ningún procesamiento de correlación para reconstruir el canal central.

Dependiendo del tipo de cálculo de correlación, esto puede hacerse sin pérdida de calidad, cuando el canal central original se utiliza como referencia para todos los parámetros de correlación.

[0101] Es de señalar que la invención no depende de un determinado esquema downmix. Esto significa que uno puede utilizar un downmix automático o manual realizado por un ingeniero de sonido. Incluso se puede utilizar la información paramétrica generada automáticamente junto con canales downmix generados manualmente.

[0102] Según el entorno de aplicación, los métodos de la invención para la construcción o la generación puede ser implementado en hardware o software. La implementación puede ser un medio de almacenamiento digital, tal como un disco o un CD que tiene señales de control electrónicamente legibles, que pueden cooperar con un sistema de ordenador programable de tal manera que lleven a cabo los procedimientos de la invención. Por lo tanto, dicho de forma general, la invención, también se refiere a un producto de programa de ordenador que tiene un código de programa almacenado en un soporte legible por máquina, estando el código de programa adaptado para realizar los procedimientos de la invención, cuando el producto de programa de ordenador se ejecuta en un ordenador. En otras palabras, la invención, por tanto, también se refiere a un programa de ordenador que tiene un código de programa para llevar a cabo los procedimientos, cuando el programa de ordenador se ejecuta en un ordenador.

[0103] La presente invención puede ser utilizada en conjunción con o incorporados en una variedad de diferentes aplicaciones o sistemas, incluidos los sistemas de televisión o distribución electrónica de música, radiodifusión, transmisión por corrientes de datos y / o recepción. Estos incluyen sistemas de decodificación / codificación de transmisiones a través de, por ejemplo, vía terrestre, satélite, cable, Internet, intranets, o medios físicos (por ejemplo, discos compactos, discos versátiles digitales, chips semiconductores, discos duros, tarjetas de memoria y similares).

La presente invención también se puede emplear en los juegos y sistemas de juego que incluyen, por ejemplo, los productos de software interactivos destinados a interactuar con el usuario en juegos de entretenimiento (de acción, juegos de rol, estrategia, aventura, simulación, carreras, deportes, arcade, juegos de cartas y de mesa) y / o educación que pueda ser publicada por varios equipos, plataformas o medios de comunicación. Además, la presente invención puede incorporarse en los reproductores de audio o sistemas CD-ROM/DVD. La presente invención también se puede incorporar en las aplicaciones de software para PC que incorporan decodificación digital (por ejemplo jugador, decodificador) y aplicaciones de software que incorporan las capacidades de codificación digital (por ejemplo de codificador, desgarrador, Recodificador, y máquina de reproducción de discos).

REIVINDICACIONES

1. Aparato para generar una señal de salida multi-canal que tiene K canales de salida, correspondiendo la señal de salida multi-canal a una señal de entrada multi-canal que tiene C canales de entrada, empleando E canales de transmisión, representando los E canales de transmisión un resultado de una operación de mezclado hacia abajo que tiene C canales de entrada como entrada, y utilizar información paramétrica relacionada con los canales de entrada, donde E es≥2, C es > E, y K es > 1 y ≤C, y donde la operación de mezclado hacia abajo es eficaz para introducir un primer canal de entrada en un primer canal de transmisión y en un segundo canal de transmisión, y para introducir adicionalmente un segundo canal de entrada en el primer canal de transmisión, que comprende: un calculador de canal de anulación (20) para calcular un canal de anulación (21) empleando información relacionada con el primer canal de entrada incluido en el primer canal de transmisión, el segundo canal de transmisión o la información paramétrica; un combinador (23) para combinar el canal de anulación (21) y el primer canal de transmisión (23) o una versión procesada de estos para obtener un segundo canal de base (25), en el que una influencia del primer canal de entrada es reducida comparado con la influencia del primer canal de entrada en el primer canal de transmisión; y un canal reconstructor (26) para reconstruir un segundo canal de salida correspondiente al segundo canal de entrada empleando el segundo canal de base e información paramétrica relacionada con el segundo canal de entrada, y para reconstruir un primer canal de salida correspondiente al primer canal de entrada empleando un primer canal de base que sea diferente del segundo canal de base en el hecho de que la influencia del primer canal es mayor comparado con el segundo canal de base, e información paramétrica relacionada con el primer canal de entrada.

2. Aparato según la reivindicación 1, en el que el combinador (22) es operativo para sustraer el canal de anulación del primer canal de transmisión o la versión procesada de estos.

3. Aparato según la reivindicación 1 o la reivindicación 2, en el que el calculador de canal de anulación (20) es operativo para calcular una estimación para el primer canal de entrada empleando el primer canal de transmisión y el segundo canal de transmisión para obtener el canal de anulación (21).

4. Aparato según cualquiera de las reivindicaciones 1 - 3, en el que la información paramétrica incluye un parámetro de diferencia entre el primer canal de entrada y un canal de referencia, y en el que el calculador de canal de anulación (20) es operativo para calcular una suma del primer canal de transmisión y el segundo canal de transmisión y para ponderar la suma empleando el parámetro de diferencia.

5. Aparato según cualquiera de las reivindicaciones 1 - 4, en el que la operación de mezclado hacia abajo es tal que el primer canal de entrada es introducido en el primer canal de transmisión tras ser escalado por un factor de mezclado hacia abajo, y en el que el calculador de canal de anulación (20) es operativo para escalar la suma de los canales de transmisión primero y segundo empleando un factor de escala, que depende del factor de mezclado hacia abajo.

6. Aparato según la reivindicación 5, en el que el factor de ponderación es igual al factor de mezclado hacia abajo.

7. Aparato según cualquiera de las reivindicaciones 1 - 6, en el que el calculador de canal de anulación (20) es operativo para determinar una suma de los canales de transmisión primero y segundo para obtener el primer canal de base.

8. Aparato según cualquiera de las reivindicaciones 1 - 7, que comprende además un procesador (24) que es operativo para procesar el primer canal de transmisión por ponderación empleando un primer factor de ponderación, y en el que el calculador de canal de anulación (20) es operativo para ponderar el segundo canal de transmisión empleando un segundo factor de ponderación.

9. Aparato según la reivindicación 8, en el que la información paramétrica incluye el parámetro de diferencia entre el primer canal de entrada y un canal de referencia, y en el que el calculador de canal de anulación (20) es operativo para determinar el segundo factor de ponderación basado en un parámetro de diferencia.

10. Aparato según la reivindicación 8 o 9, en el que el primer factor de ponderación es igual a (1-h), donde h es un valor real, y en el que el segundo factor de ponderación es igual a h.

11. Aparato según la reivindicación 10, en el que la información paramétrica incluye un valor de diferencia de nivel, y donde h se deriva de el valor de diferencia de nivel paramétrico.

12. Aparato según la reivindicación 11, en el que h es igual a un valor derivado de la diferencia de nivel dividida por un factor que depende de la operación de mezclado hacia abajo.

13. Aparato según la reivindicación 10, en el que la información paramétrica incluye la diferencia de nivel entre el primer canal y el canal de referencia, y en el que h es igual a 1√2 x 10L/20, donde L es la diferencia de nivel.

14. Aparato según cualquiera de las reivindicaciones 1 - 13, en el que la información paramétrica también incluye una señal de control que depende de la relación entre el primer canal de entrada y el segundo canal de entrada, y en el que el calculador de canal de anulación (20) está controlado por la señal de control para aumentar o disminuir activamente una energía del canal de anulación o incluso deshabilitar el cálculo del canal de anulación del todo.

15. Aparato según cualquiera de las reivindicaciones 1 - 14, en el que la operación de mezclado hacia abajo también es operativa para introducir un tercer canal de entrada en el segundo canal de transmisión, comprendiendo el aparato además otro combinador para combinar el canal de anulación y el segundo canal de transmisión o una versión procesada de estos para obtener un tercer canal de base, en el que una influencia del primer canal de entrada es reducida comparado con la influencia del primer canal de entrada on el segundo canal de transmisión; y un canal reconstructor para reconstruir el tercer canal de salida correspondiente al tercer canal de entrada empleando el tercer canal de base e información paramétrica relacionada con el tercer canal de entrada.

16. Aparato según cualquiera de las reivindicaciones 1 - 15, en el que la información paramétrica incluye diferencias de nivel entre canales, diferencias de tiempo entre canales, diferencias de fase entre canales o valores de correlación entre canales, y en el que el canal reconstructor (26) es operativo para aplicar cualquiera de los parámetros del grupo mencionado en un canal de base para obtener un canal de salida en bruto.

17. Aparato según la reivindicación 16, en el que el canal reconstructor (26) es operativo para escalar the canal de salida en bruto de modo que toda la energía en el canal de salida reconstruido final es igual a toda la energía de los E canales de transmisión.

18. Aparato según cualquiera de las reivindicaciones 1 - 17, en el que la información paramétrica se dan por bandas, y en el que el calculador de canal de anulación (20), el combinador (22) y el canal reconstructor (26) son operativos para procesar la pluralidad de bandas empleando información paramétrica determinada por bandas, y en el que el aparato también comprende una unidad de conversión tiempo/ frecuencia (IFB) para convertir los canales de transmisión en una representación en frecuencia que tiene bandas de frecuencia, y una unidad de conversión frecuencia/ tiempo para convertir bandas de frecuencia reconstruidas en el dominio del tiempo.

19. El aparato según cualquiera de las reivindicaciones 1 - 18 que comprende además: un sistema que se selecciona de entre el grupo que consiste en un reproductor de video digital, un reproductor de audio digital, un ordenador, un receptor satélite, un receptor cable , un receptor de difusión terrestre, y un sistema de entretenimiento doméstico; y donde el sistema comprende el calculador de canal, el combinador, y el canal reconstructor.

20. Procedimiento para generar una señal de salida multi-canal que tiene K canales de salida, correspondiendo la señal de salida multi-canal a una señal de entrada multi-canal que tiene C canales de entrada, empleando E canales de transmisión, representando los E canales de transmisión un resultado de una operación de mezclado hacia abajo que tiene C canales de entrada como entrada, y utilizar información paramétrica relacionada con los canales de entrada, donde E es≥2, C es > E, y K es > 1 y ≤C, y donde la operación de mezclado hacia abajo es eficaz para introducir un primer canal de entrada en un primer canal de transmisión y en un segundo canal de transmisión, y para introducir adicionalmente un segundo canal de entrada en el primer canal de transmisión, que comprende: calcular (20) un canal de anulación empleando información relacionada con el primer canal de entrada incluido en el primer canal de transmisión, el segundo canal de transmisión o la información paramétrica; combinar (22) el canal de anulación y el primer canal de transmisión o una versión procesada de estos para obtener un segundo canal de base, en el que una influencia del primer canal de entrada es reducida comparado con la influencia del primer canal de entrada en el primer canal de transmisión; y reconstruir (26) un segundo canal de salida correspondiente al segundo canal de entrada empleando el segundo canal de base e información paramétrica relacionada con el segundo canal de entrada, y un primer canal de salida correspondiente al primer canal de entrada empleando un primer canal de base que sea diferente del segundo canal de base en el hecho de que la influencia del primer canal es mayor comparado con el segundo canal de base, e información paramétrica relacionada con el primer canal de entrada.

21. Programa de ordenador que tiene un código de programa para implementar, cuando se ejecuta en un ordenador, un procedimiento para generar una señal de salida multi-canal que tiene K canales de salida, correspondiendo la señal de salida multi-canal a una señal de entrada multicanal que tiene C canales de entrada, empleando E canales de transmisión, representando los E canales de transmisión un resultado de una operación de mezclado hacia abajo que tiene C canales de entrada como entrada, y utilizar información paramétrica relacionada con los canales de entrada, donde E es≥2, C es > E, y K es > 1 y ≤C, y dond e la operación de mezclado hacia abajo es eficaz para introducir un primer canal de entrada en un primer canal de transmisión y en un segundo canal de transmisión, y para introducir adicionalmente un segundo canal de entrada en el primer canal de transmisión, comprendiendo el procedimiento: calcular (20) un canal de anulación empleando información relacionada con el primer canal de entrada incluido en el primer canal de transmisión, el segundo canal de transmisión o la información paramétrica; combinar (22) el canal de anulación y el primer canal de transmisión o una versión procesada de estos para obtener un segundo canal de base, en el que una influencia del primer canal de entrada es reducida comparado con la

influencia del primer canal de entrada en el primer canal de transmisión; y reconstruir (26) un segundo canal de salida correspondiente al segundo canal de entrada empleando el segundo canal de base e información paramétrica relacionada con el segundo canal de entrada, y un primer canal de salida correspondiente al primer canal de entrada empleando un primer canal de base que sea diferente del segundo canal de base en el hecho de que la influencia del primer canal es mayor comparado con el segundo canal de base, e información paramétrica relacionada con el primer canal de entrada.

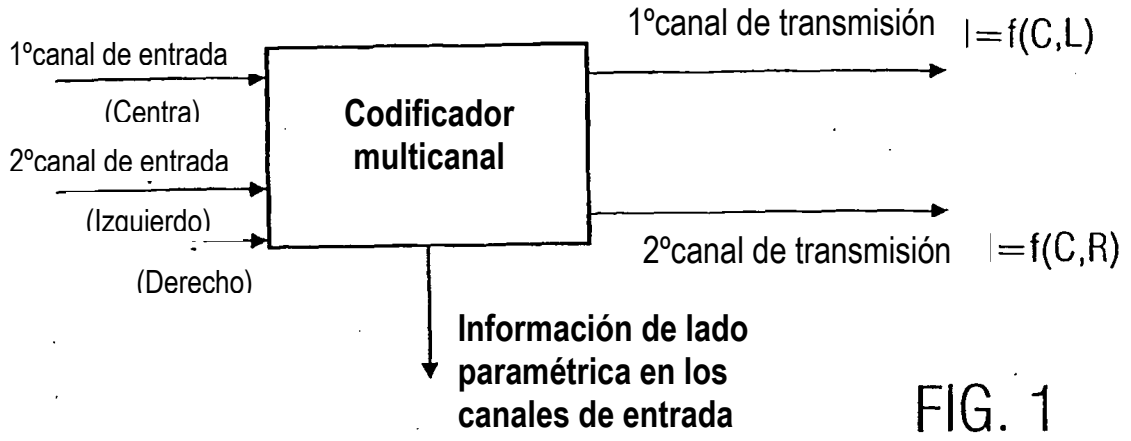

1ºcanal de entrada
(Centra)
2ºcanal de entrada
(Izquierdo)
(Derecho)
Codificador multicanal
1ºcanal de transmisión
I=f(C,L)
2ºcanal de transmisión
I=f(C,R)
Información de lado paramétrica en los canales de entrada

FIG. 1

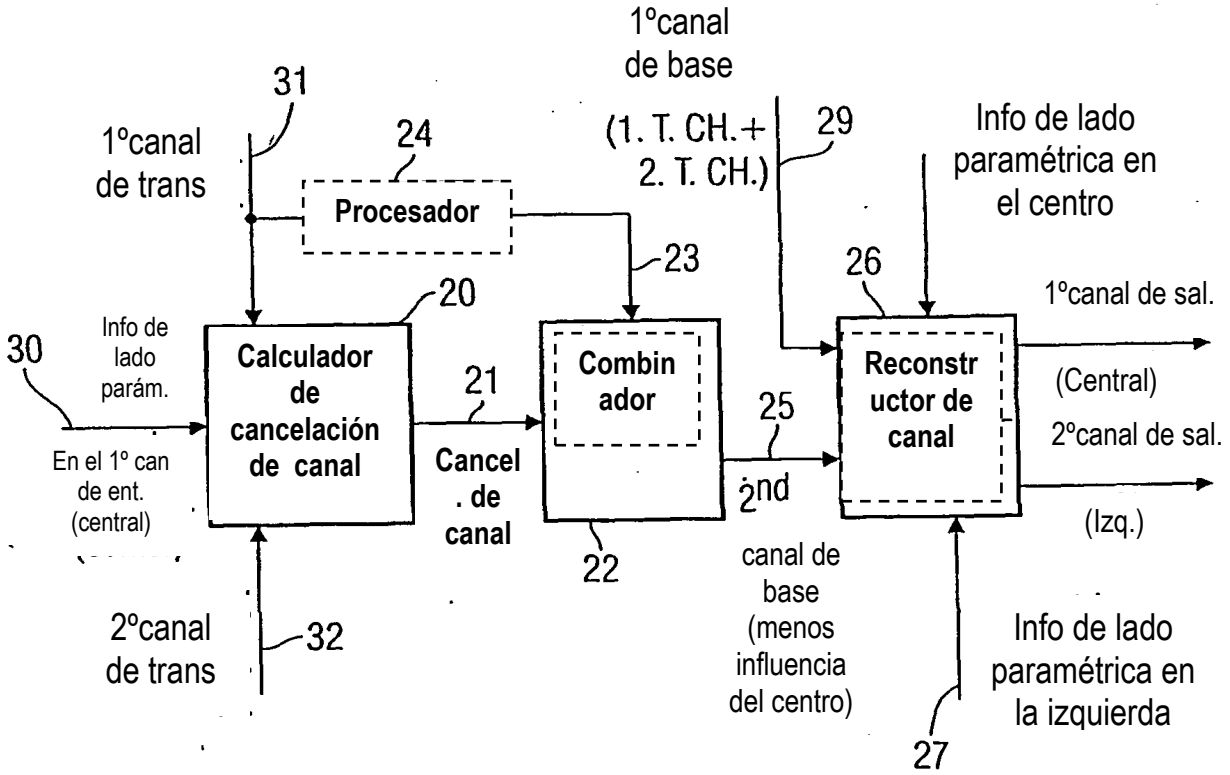

1ºcanal de trans
31
24
Procesador
1ºcanal de base
(1. T. CH.+ 2. T. CH.)
29
Info de lado paramétrica en el centro
23
26
20
30
Info de lado parám.
Calculador de cancelación de canal
21
Combin ador
25
Reconstr uctor de canal
1ºcanal de sal.
(Central)
2ºcanal de sal.
(Izq.)
En el 1º can de ent. (central)
Cancel . de canal
22
2nd
canal de base (menos influencia del centro)
2ºcanal de trans
32
Info de lado paramétrica en la izquierda
27

FIG. 2

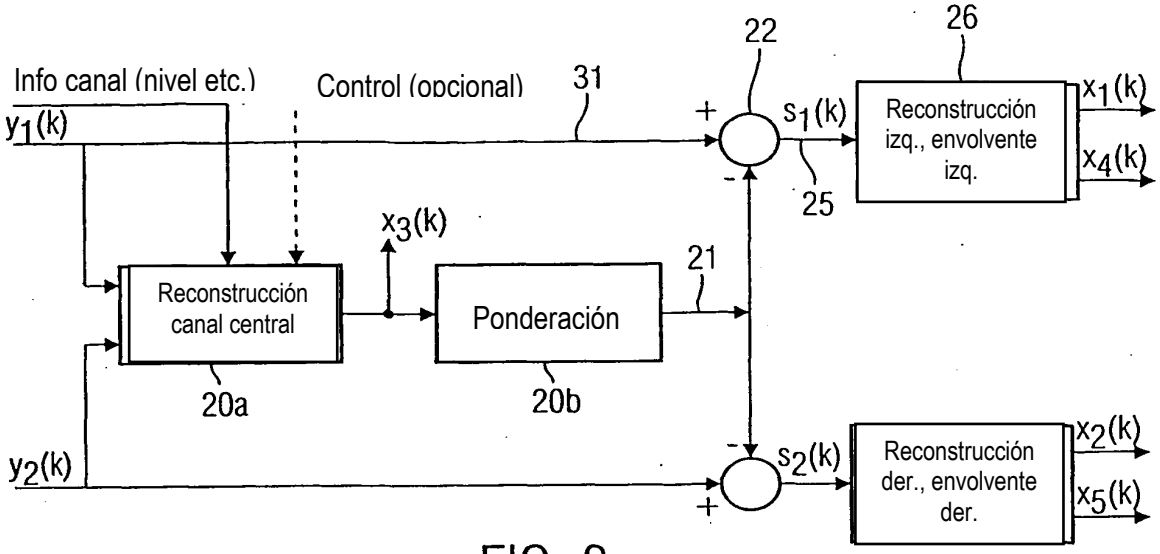

Info canal (nivel etc.)
Control (opcional)
31
22
26
$y_1(k)$
$s_1(k)$
Reconstrucción izq., envolvente izq.
$x_1(k)$
$x_4(k)$
25
$x_3(k)$
21
Reconstrucción canal central
Ponderación
20a
20b
$y_2(k)$
$s_2(k)$
Reconstrucción der., envolvente der.
$x_2(k)$
$x_5(k)$

FIG. 3

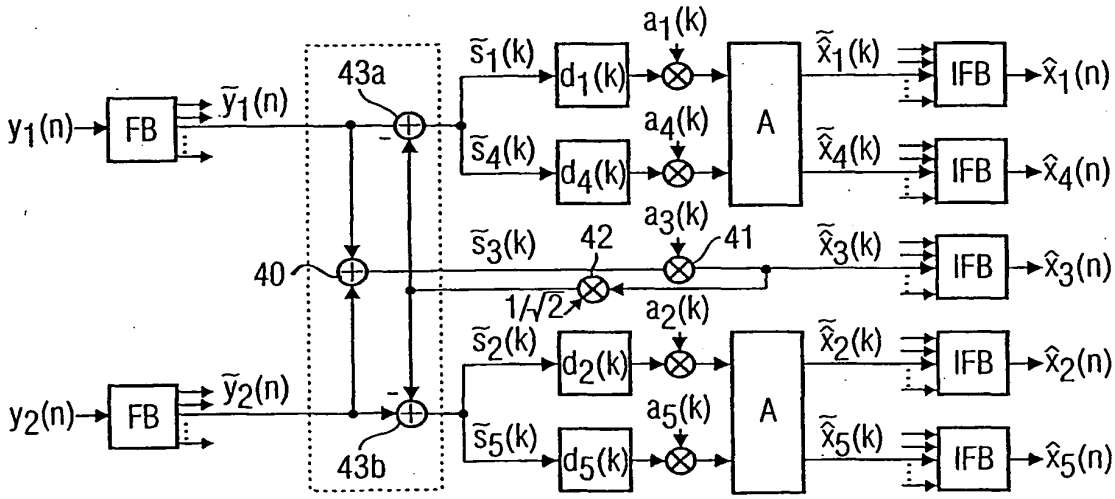

$y_1(n)$
FB
$\tilde{y}_1(n)$
43a
$\tilde{s}_1(k)$
$d_1(k)$
$a_1(k)$
$\tilde{s}_4(k)$
$d_4(k)$
$a_4(k)$
A
$\tilde{\hat{x}}_1(k)$
$\tilde{\hat{x}}_4(k)$
IFB
$\hat{x}_1(n)$
$\hat{x}_4(n)$
40
$\tilde{s}_3(k)$
42
$a_3(k)$
41
$1/\sqrt{2}$
$\tilde{\hat{x}}_3(k)$
$\hat{x}_3(n)$
$\tilde{s}_2(k)$
$d_2(k)$
$a_2(k)$
$\tilde{\hat{x}}_2(k)$
$\hat{x}_2(n)$
$y_2(n)$
$\tilde{y}_2(n)$
$\tilde{s}_5(k)$
$d_5(k)$
$a_5(k)$
$\tilde{\hat{x}}_5(k)$
$\hat{x}_5(n)$
43b

FIG. 4

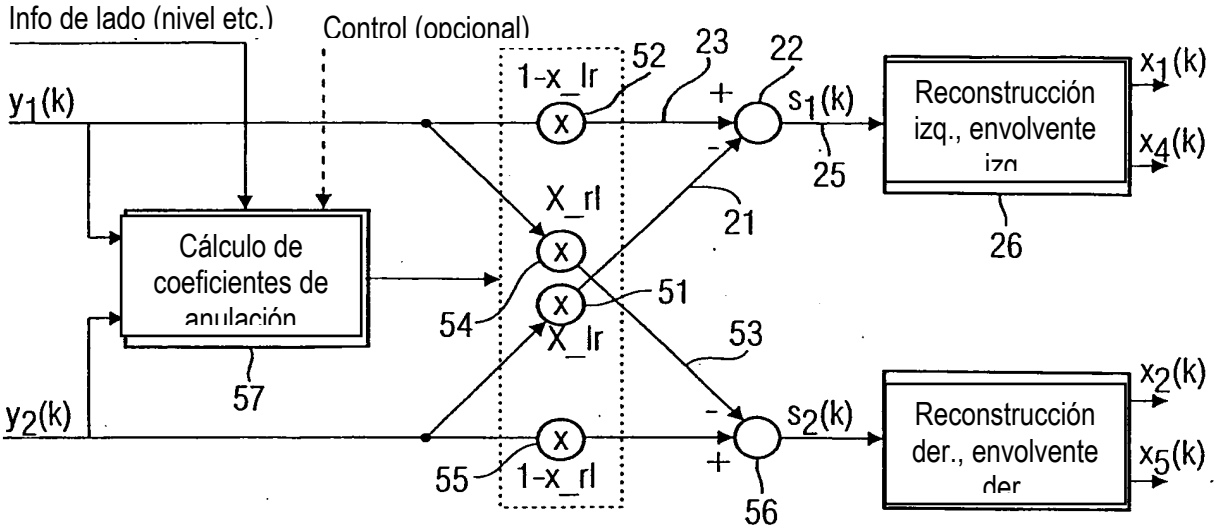
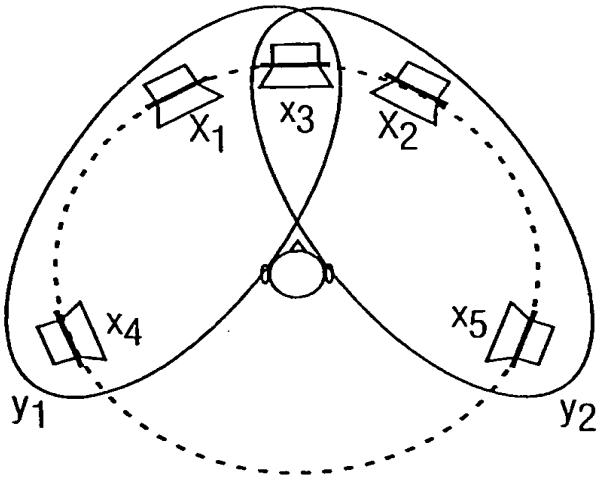


FIG. 5a

$$U_{25}=\begin{bmatrix} 1-a_3(k)/\sqrt{2} & -a_3(k)/\sqrt{2} \\ -a_3(k)/\sqrt{2} & 1-a_3(k)/\sqrt{2} \\ 1 & 1 \\ 1-a_3(k)/\sqrt{2} & -a_3(k)/\sqrt{2} \\ -a_3(k)/\sqrt{2} & 1-a_3(k)/\sqrt{2} \end{bmatrix}$$

FIG. 5b



$$IN_5 = \begin{bmatrix} \text{Izquierda} \\ \text{Derecha} \\ \text{Central} \\ \text{Posterior – izq} \\ \text{Posterior - der} \end{bmatrix}$$

FIG. 6a

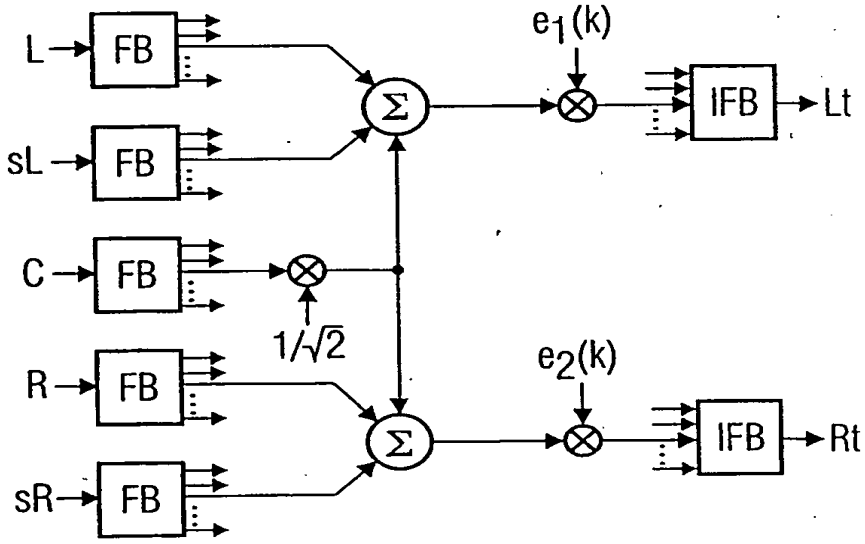


FIG. 6b

$$D_{52} = \begin{bmatrix} 1 & 0 & 1/\sqrt{2} & 1 & 0 \\ 0 & 1 & 1/\sqrt{2} & 0 & 1 \end{bmatrix}$$

FIG. 6c

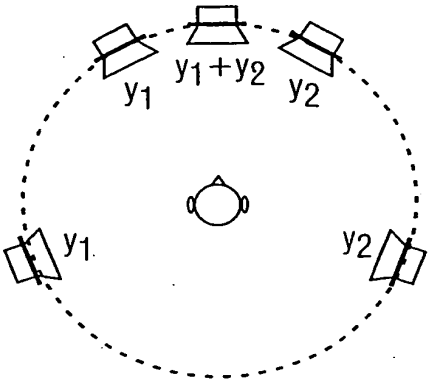


FIG. 7a

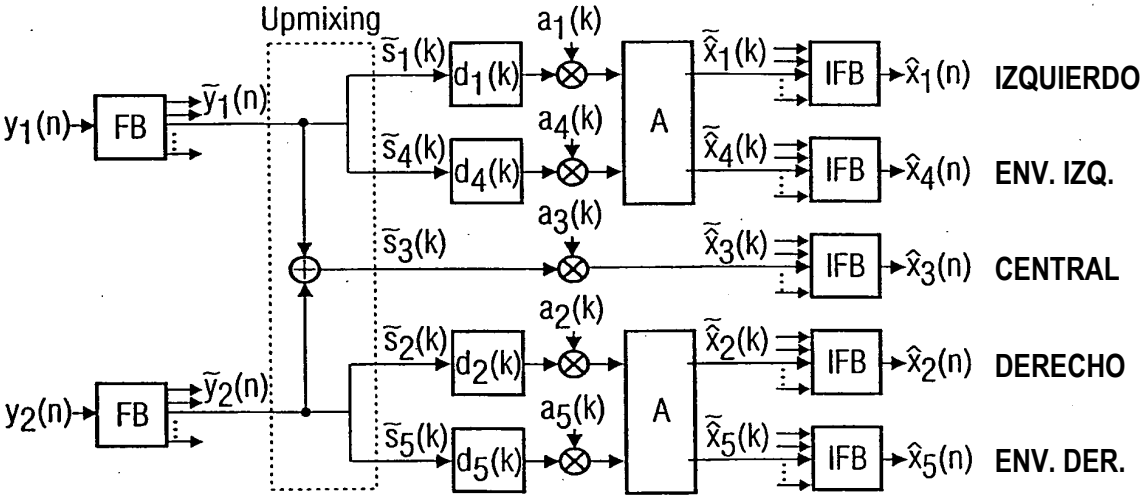


FIG. 7b

$$U_{25}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

FIG. 7c

- **Modificación de nivel para cada canal**

$$F_i = 10^{\Delta Li/20} \cdot F_{ref} \; ; \; F_{ref} = y_1 + y_2$$

$$a_i = 10^{\Delta Li/20}$$

- **Normalización global (energía de todos los canales de salida es igual a la energía de todos los canales transmitidos)**

$$F_{ref} = \frac{1}{\sqrt{\sum_i 10^{\Delta Li/10}}}$$

F_i: **Uno de los canales de salida de índice i**

ΔLi: **Diferencia de nivel de canal entre el canal de entrada y un canal de referencia**

F_{ref}: **Canal de referencia (tal como y_1+y_2)**

i: **Índice de canal, i=1...C**

C: **Número total de canales de entrada**

FIG. 7d

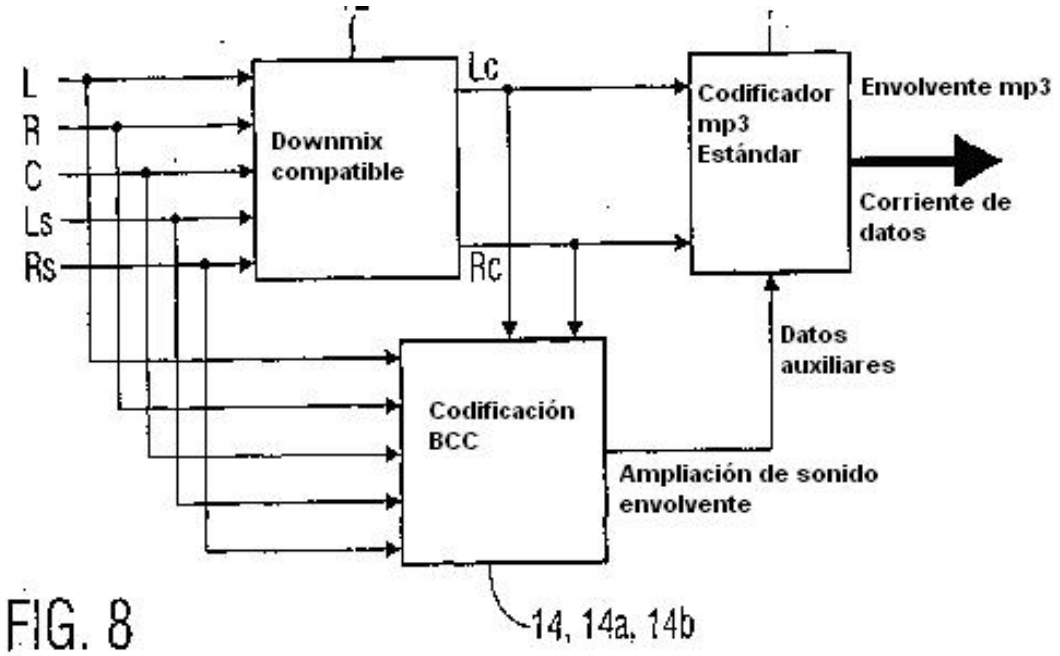

L
R
C
Ls
Rs
Downmix compatible
Lc
Rc
Codificador mp3 Estándar
Envolvente mp3
Corriente de datos
Datos auxiliares
Codificación BCC
Ampliación de sonido envolvente
14, 14a, 14b

FIG. 8

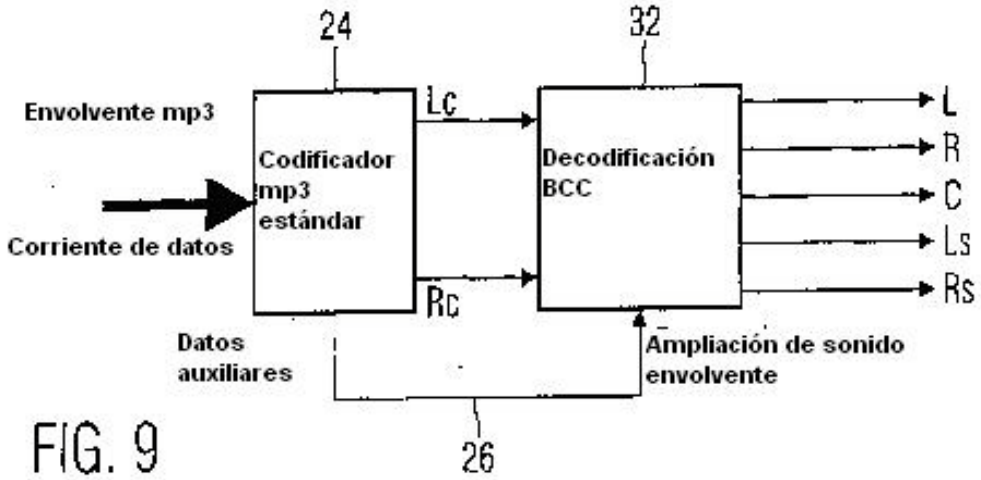
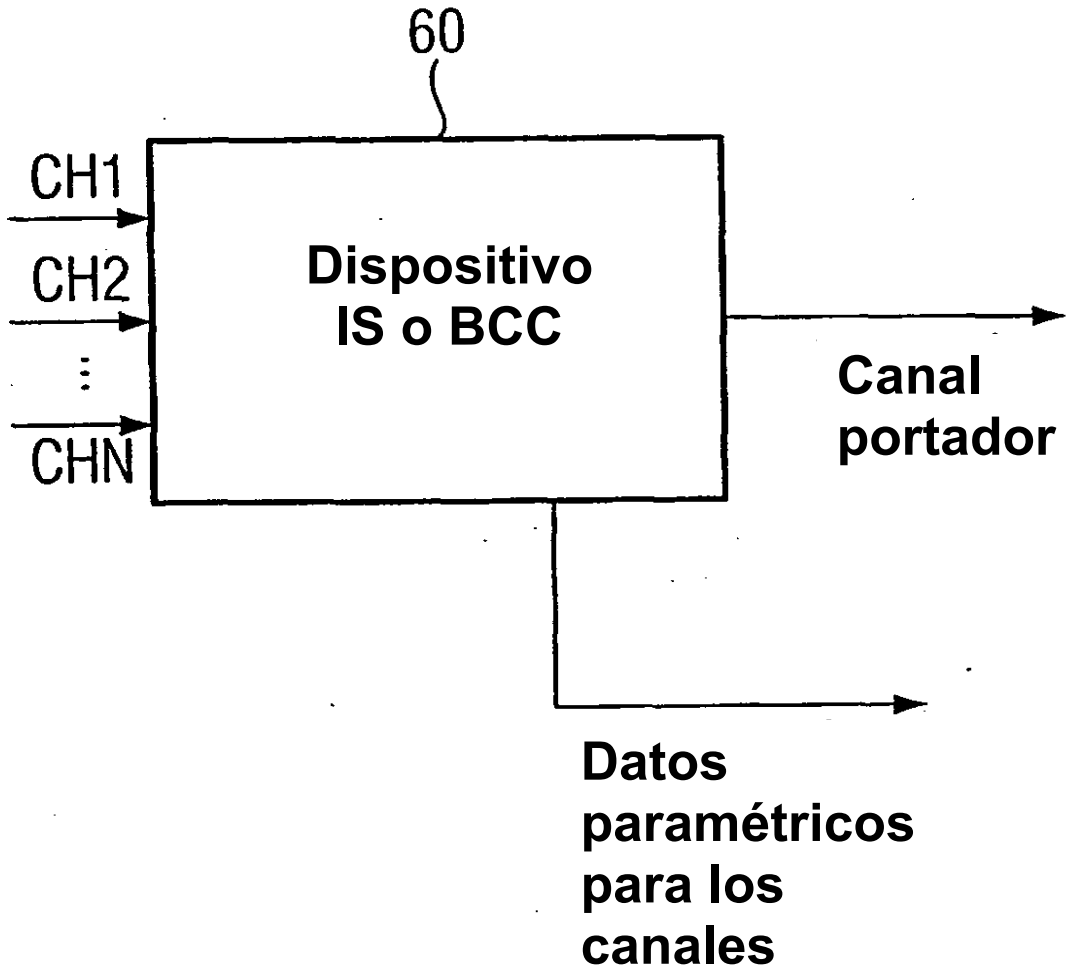

24
32
Envolvente mp3
Corriente de datos
Codificador mp3 estándar
Lc
Rc
Decodificación BCC
L
R
C
Ls
Rs
Datos auxiliares
Ampliación de sonido envolvente
26

FIG. 9

FIG. 10 **(Estado de la técnica)**



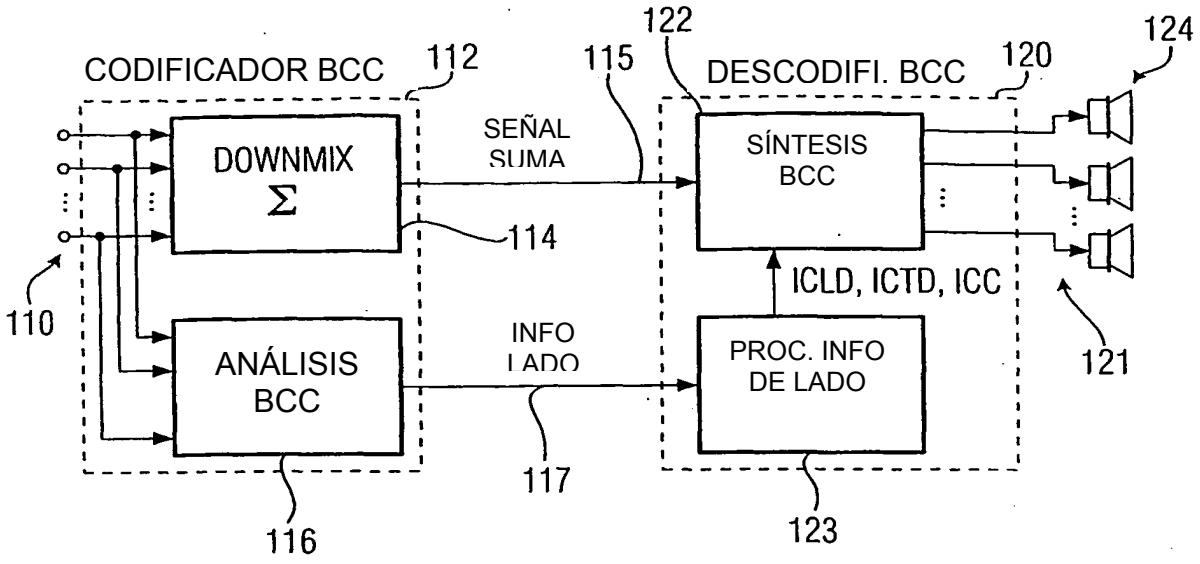


FIG. 11 **(ESTADO DE LA TÉCNICA)**

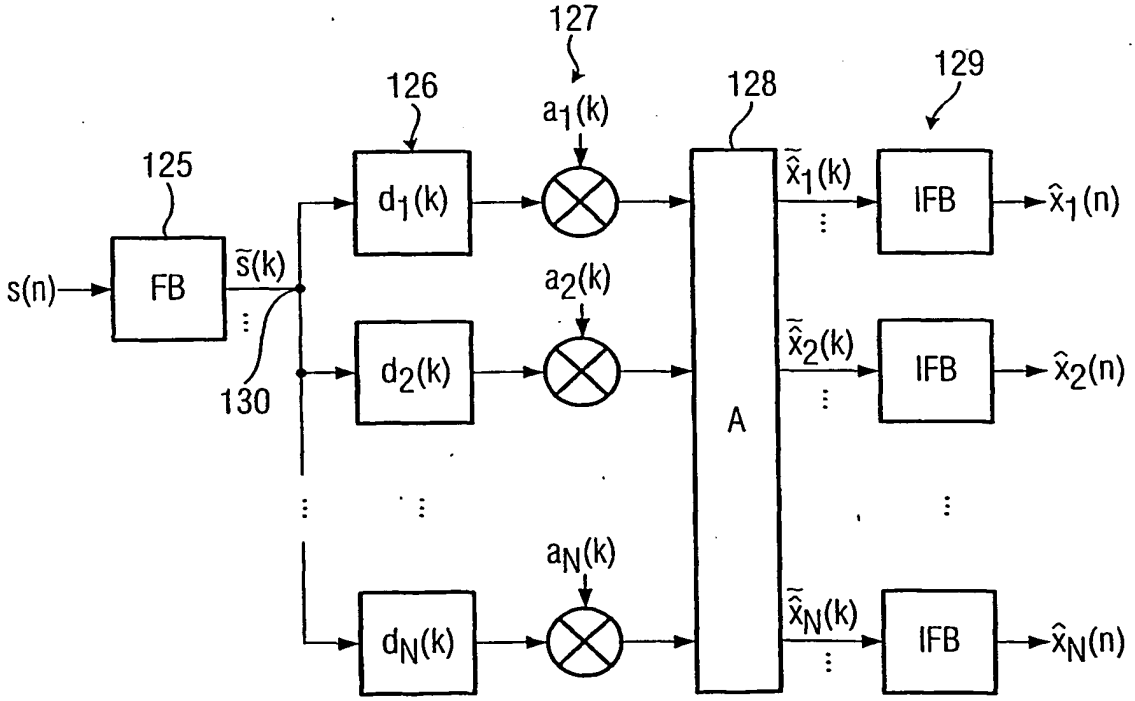


FIG. 12 **(ESTADO DE LA TÉCNICA)**

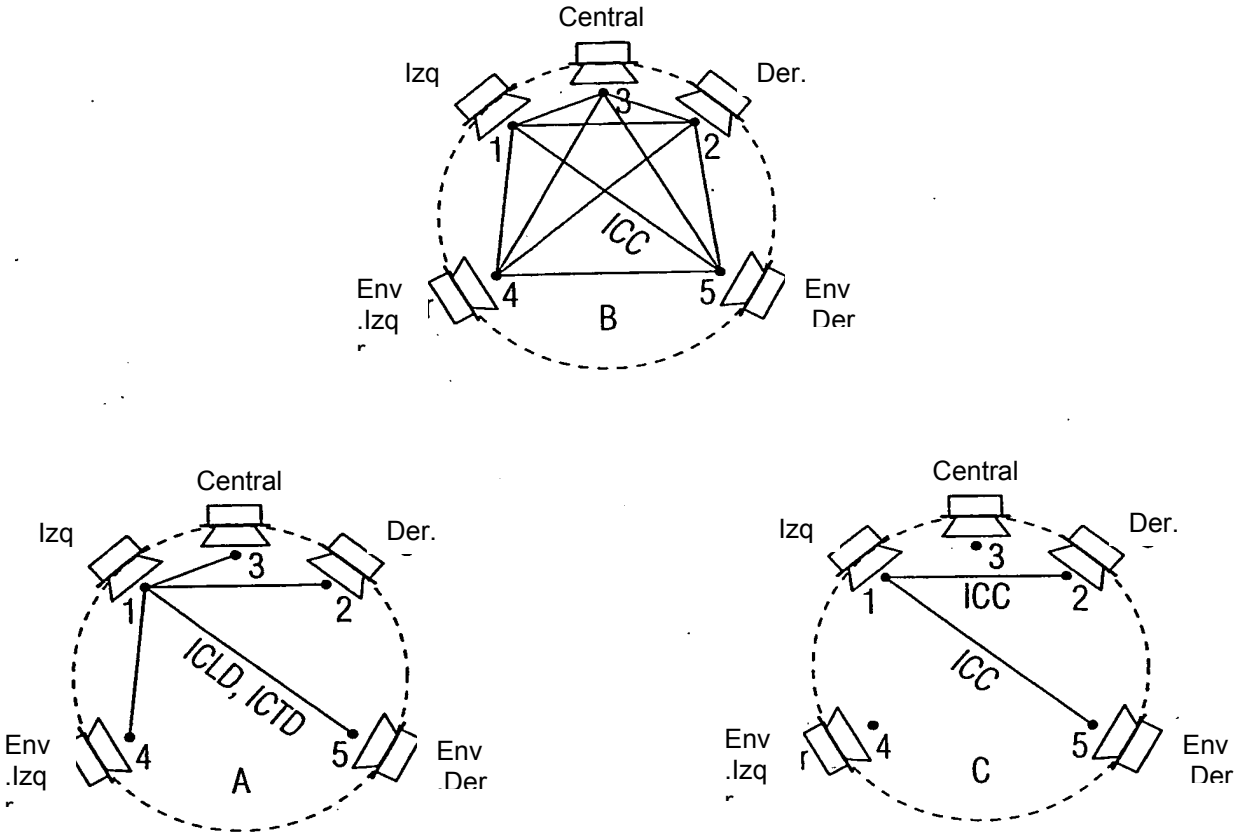


FIG. 13 **(ESTADO DE LA TÉCNICA)**