

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 002 034**

51 Int. Cl.:

G10L 19/008 (2013.01)

G06F 12/0815 (2006.01)

H04S 3/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **21.06.2006** **E 23180543 (3)**

97 Fecha y número de publicación de la concesión europea: **27.11.2024** **EP 4235440**

54 Título: **Concepto para puentear el espacio entre codificación paramétrica de audio multicanal y codificación multicanal envolvente matricial**

30 Prioridad:

19.07.2005 US 701001 P

29.12.2005 US 323965

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
06.03.2025

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (50.00%)**

Hansastraße 27c

80686 München, DE y

**DOLBY LABORATORIES LICENSING
CORPORATION (50.00%)**

72 Inventor/es:

HILPERT, JOHANNES;

FALLER, CHRISTOF;

LINZMEIER, KARSTEN y

SPERSCHNEIDER, RALPH

74 Agente/Representante:

RUÍZ GALLEGOS, Natalia

ES 3 002 034 T3

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

DESCRIPCIÓN

Concepto para puentear el espacio entre codificación paramétrica de audio multicanal y codificación multicanal envolvente matricial

Campo de la invención

La presente invención se refiere a codificación y transmisión de audio multicanal y, en particular, a técnicas para codificar y decodificar audio multicanal en una manera que es completamente compatible de manera inversa con dispositivos y formatos estéreo, permitiendo una codificación eficiente de audio multicanal.

Antecedentes de la invención y técnica anterior

La codificación paramétrica de señales de audio multicanal es un tema de investigación actual. Generalmente pueden distinguirse dos procedimientos para codificar señales de audio multicanal. El Grupo de Expertos en Imágenes en Movimiento (MPEG), un subgrupo de la Organización Internacional de Normalización (ISO), está actualmente trabajando en la estandarización de la tecnología para la reconstrucción de contenido de audio multicanal de señales estéreo o incluso mono de mezcla descendente añadiendo únicamente una cantidad pequeña de información auxiliar a las señales de mezcla descendente.

En paralelo al estéreo, se están desarrollando métodos de mezcla ascendente multicanal que no necesitan ninguna información secundaria adicional que no se encuentra ya contenida (implícitamente) en la señal de mezcla descendente para reconstruir la imagen espacial de la señal de audio multicanal original.

Los métodos existentes para transmisión multicanal compatible con estéreo sin información secundaria adicional que obtiene relevancia práctica pueden caracterizarse mayormente como métodos de envolvente matricial, tal como Dolby Pro Logic (Dolby Pro Logic II) y Logic-7, como se describe en más detalle en "Dolby Surround Pro Logic II Decoder - Principles of Operation", http://www.dolby.com/assets/pdf/tech_library/209_Dolby_Surround_Pro_Logic_II_Decoder_Principles_of_Operation.pdf y en "Multichannel Matrix Surround Decoders for Two-Eared Listeners", Griesinger, D., 101ª Convención de AES, Los Ángeles, Estados Unidos, 1996, Publicación 4402. El principio común de estos métodos es que usan maneras especializadas de mezcla descendente estéreo o multicanal en el que el codificador aplica desplazamientos de fase a los canales envolventes antes de mezclarlos con los canales, frontal y central, para formar una señal de mezcla descendente estéreo. La generación de la señal de mezcla descendente (Lt, Rt) se representa en la siguiente ecuación:

$$\begin{bmatrix} L_t \\ R_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & q & a \cdot j & b \cdot j \\ 0 & 1 & q & -b \cdot j & -a \cdot j \end{bmatrix} \begin{bmatrix} L_f \\ R_f \\ C \\ L_s \\ R_s \end{bmatrix} \quad (1)$$

La señal de mezcla descendente izquierda (Lt) consiste en la señal frontal izquierda (Lf), la señal central (C) multiplicada por un factor q, la señal envolvente izquierda (Ls) girada en fase por 90 grados, ('j') y escalada por un factor a, y la señal envolvente derecha (Rs) que también se gira en fase por 90 grados y escalada por un factor b. La señal de mezcla descendente derecha (Rt) se genera de manera similar. Los factores de mezcla descendente típicos son 0,707 para q y a, y 0,408 para b. El razonamiento para las diferentes señales de los canales envolventes para la señal de mezcla descendente derecha (Rt) y la señal de mezcla descendente izquierda (Lt) es que es ventajoso mezclar los canales envolventes en anti-fase en el par de mezcla descendente (Lt, Rt). Esta propiedad ayuda al decodificador a discriminar entre los canales frontal y posterior del par de señal de mezcla descendente. Por lo tanto, la matriz de mezcla descendente permite una reconstrucción parcial de una señal de salida multicanal que sale de la mezcla descendente estéreo dentro del decodificador aplicando una operación de de-matrización. Cuánto se parece la señal multicanal recreada a la señal de entrada del codificador original, sin embargo, depende de las propiedades específicas del contenido de audio multicanal.

Un ejemplo de un método de codificación para agregar información auxiliar, también llamada información secundaria, es la codificación de audio envolvente MPEG. Esta manera eficiente para codificación de audio multicanal paramétrica se describe, por ejemplo, en "The Reference Model Architecture for MPEG Spatial Audio Coding", Herre, J., Purnhagen, H., Breebaart, J., Faller, C., Disch, S., Kjoerling, K., Schuijers, E., Hilpert, J., Myburg, F., Proc. 118ª Convención de AES, Barcelona, España, 2005 y en "Text of Working Draft for Spatial Audio Coding

(SAC)", ISO/IEC JTC1/SC29/WG11 (MPEG), Documento N7136, Busan, Corea, 2005.

Una revisión esquemática de un codificador usado en codificación de audio espacial se muestra en la figura 6. El codificador divide las señales entrantes 10 (entrada 1, ... entrada N) en cintas de frecuencia de tiempo separadas por medio de filtros espejo en cuadratura 12 (QMF). Los grupos de las cintas (bandas) de frecuencia resultantes se denominan "bandas de parámetro". Para cada banda de parámetro, se determina un número de parámetros espaciales 14 mediante un estimador de parámetro 16 que describe las propiedades de la imagen espacial, por ejemplo, diferencias de nivel entre pares de canales (CLD), correlación transversal entre pares de canales (ICC) o información en envolturas de señal (CPC). Estos parámetros se cuantifican posteriormente, codifican y compilan de manera conjunta en una secuencia de bits de datos espaciales. Dependiendo del modo de operación, esta secuencia de bits puede cubrir un amplio intervalo de tasas de bits, iniciando desde unos pocos kBit/s para audio multicanal de buena calidad hasta décimas de kBit/s para calidad casi transparente.

Además de la extracción de los parámetros, el codificador también genera una mezcla descendente mono o estéreo de la señal de entrada multicanal. Además, en el caso de una mezcla descendente estéreo, el usuario tiene la elección de una mezcla descendente estéreo convencional (estilo ITU) o de una mezcla descendente que es compatible con sistemas de envolvente matricial. Finalmente, la mezcla descendente estéreo se transfiere al dominio de tiempo por medio de bancos de síntesis QMF 18. La mezcla descendente resultante puede transmitirse a un decodificador, acompañado por los parámetros espaciales o la secuencia de bits de parámetro espacial 14. Preferentemente, la mezcla descendente se codifica también antes de la transmisión (usando un codificador de núcleo mono o estéreo convencional), mientras las secuencias de bits del codificador de núcleo y los parámetros espaciales adicionalmente pueden combinarse (multiplexarse) para formar una secuencia de bits de salida única.

Un decodificador, como se representa en la figura 7, en principio realiza el proceso inverso del codificador. Un flujo de entrada se divide en una secuencia de bits del codificador núcleo y una secuencia de bits de parámetro. Esto no se muestra en la figura 7. Posteriormente, la mezcla descendente decodificada 20 se procesa por un banco de análisis QMF 22 para derivar las bandas de parámetro que son las mismas que aquellas aplicadas en el codificador. Una etapa de síntesis espacial 24 reconstruye la señal multicanal por medio de datos de control 26 (es decir, los parámetros espaciales transmitidos). Finalmente, las señales de dominio QMF se transfieren al dominio de tiempo por medio de un banco de síntesis QMF 27 que deriva las señales de salida multicanal finales 28.

La figura 8 muestra un ejemplo simple de un análisis QMF, como se realiza dentro del codificador de la técnica anterior en la figura 6 y el decodificador de la técnica anterior en la figura 7. Una muestra de audio 30, muestreada en el dominio de tiempo y que tiene cuatro valores de muestra se introduce en un banco de filtro 32. El banco de filtro 32 deriva tres muestras de salida 34a, 34b y 34c teniendo cuatro valores de muestra cada una. En un caso ideal, el banco de filtro 32 deriva las muestras de salida 34a a 34c de manera que las muestras dentro de las señales de salida comprenden únicamente información acerca de los intervalos de frecuencia discretos de la señal de audio subyacente 30. En el caso mostrado en la figura 8, la muestra 34a tiene información acerca del intervalo de frecuencia que va desde f_0 a f_1 , la muestra 34b tiene información del intervalo de frecuencia $[f_1, f_2]$ y la muestra 34c tiene información acerca del intervalo de frecuencia $[f_2, f_3]$. Aunque los intervalos de frecuencia en la figura 8 no se solapan, en un caso más general los intervalos de frecuencia de las muestras de salida que salen de un banco de filtro pueden tener perfectamente un solapamiento de frecuencia.

Un codificador de la técnica anterior puede, como ya se describe anteriormente, suministrar o bien una mezcla descendente estilo ITU o bien una mezcla descendente compatible con envolvente matricial, cuando se desea una mezcla descendente de dos canales. En el caso de una mezcla descendente compatible con envolvente matricial (usando por ejemplo el procedimiento de matriz dado en la ecuación 1), una posibilidad sería que el codificador generase una mezcla descendente compatible con envolvente matricial directamente.

La figura 9 muestra un enfoque alternativo para generar una mezcla descendente compatible con envolvente matricial usando una unidad de procesamiento posterior de mezcla descendente 30 funcionando en una mezcla descendente estéreo regular 32. El procesador de envolvente matricial 30 (codificador MTX) modifica la mezcla descendente de estéreo regular 32 para hacerlo compatible con envolvente matricial guiado por los parámetros espaciales 14 extraídos por la etapa de extracción de parámetro 16. Para transmisión, se transfiere una mezcla descendente compatible con envolvente matricial 34 al dominio de tiempo por una síntesis QMF usando el banco de síntesis QMF 18.

Derivar la señal compatible con envolvente matricial mediante el procesamiento posterior de una mezcla descendente estéreo regular tiene la ventaja de que el procesamiento de compatibilidad con envolvente matricial puede invertirse completamente en un lado de decodificador si los parámetros espaciales están disponibles.

Aunque ambos enfoques son adecuados para transmitir una señal multicanal, existen desventajas específicas de los sistemas del estado de la técnica. Los métodos de envolvente matricial son muy eficientes (ya que no se requiere ningún parámetro adicional) al precio de una calidad de reconstrucción multicanal muy limitada.

Los procedimientos multicanal paramétricos por el otro lado requieren una tasa de transmisión de bits más alta debido a la información secundaria, que se vuelve un problema cuando se fija un límite como una tasa de transmisión de bits aceptable máxima para la representación paramétrica. Cuando los parámetros codificados requieren una cantidad comparativamente alta de tasa de bits, la única manera posible de permanecer dentro de tal límite de tasa de transmisión de bits es disminuir la calidad de un canal de mezcla descendente codificado al incrementar la compresión del canal. Por lo tanto, el resultado es una pérdida general en calidad de audio, que puede ser inaceptablemente alta. En otras palabras, para enfoques multicanal paramétricos, con frecuencia existe un límite estricto de la tasa de transmisión de bits mínima que se requiere para la capa de parámetro espacial que puede, en algunos casos, ser inaceptablemente alta.

Aunque la compatibilidad inversa principal entre los métodos de envolvente matricial y métodos de audio espacial pueden lograrse mediante un codificador de la técnica anterior como se ilustra en la figura 9, ninguna tasa de transmisión de bits adicional puede ahorrarse con este enfoque cuando únicamente se requiere la decodificación basada en matriz. Aún después el conjunto completo de parámetros espaciales tiene que transmitirse, desperdiciando el ancho de banda de transmisión.

Mientras que la tasa de bits que tiene que gastarse cuando se aplica el método paramétrico puede ser demasiado alta en caso de ciertos escenarios de aplicación, la calidad del audio suministrado por los métodos sin transmisión de información secundaria puede no ser suficiente.

La solicitud de patente de Estados Unidos 2005157883 muestra un aparato para construir una señal de audio multicanal utilizando una señal de entrada e información paramétrica lateral, la señal de entrada incluye el primer canal de entrada y el segundo canal de entrada derivados de una señal multicanal original, y la información paramétrica lateral describe las interrelaciones entre los canales de la señal multicanal original.

La solicitud de patente de Estados Unidos 2005/157883 A1 da a conocer un aparato para construir una señal de audio multicanal que usa una señal de entrada e información secundaria paramétrica, incluyendo la señal de entrada el primer canal de entrada y el segundo canal de entrada derivados de una señal multicanal original, y describiendo la información secundaria paramétrica las interrelaciones entre canales de la señal original multicanal.

El documento EP 1376538 A1 da a conocer un procedimiento, en el que una parte del espectro de dos o más señales de entrada se codifica usando técnicas de codificación convencionales, mientras que el resto del espectro se codifica usando codificación de referencia binaural (BCC). En la codificación BCC, componentes espectrales de las señales de entrada se mezclan de manera descendente y se generan parámetros BCC (por ejemplo, nivel entre canales y/o diferencias de tiempo). En una implementación estéreo, después de convertir los canales izquierdo y derecho al dominio de frecuencia, pares de componentes espectrales del canal izquierdo y derecho se mezclan de manera descendente a mono. Después, los componentes mono se convierten de nuevo al dominio de tiempo, junto con aquellos componentes espectrales de canal izquierdo y derecho que no se mezclaron de manera descendente, para formar señales estéreo híbridas, que después se pueden codificar usando técnicas de codificación convencionales. Para reproducción, la tasa de transmisión de bits se decodifica usando técnicas de decodificación convencionales. Después, las técnicas de síntesis BCC pueden aplicar los parámetros BCC para sintetizar una escena auditiva basada en los componentes mono así como en los componentes estéreo sin mezclar.

Sumario de la invención

Es el objeto de la presente invención proporcionar un concepto de una codificación más eficiente de señales de audio multicanal mientras es compatible de manera inversa con las soluciones de codificación basada en matriz.

Este objeto se logra mediante un decodificador de audio multicanal de la reivindicación 1, un método para el procesamiento de la reivindicación 13, una representación paramétrica de la reivindicación 14, o un programa informático de la reivindicación 15.

La presente invención se basa en el descubrimiento de que una señal de audio multicanal puede representarse de manera eficiente por una representación paramétrica, cuando una primera regla de derivación se usa para derivar primeros datos paramétricos de la representación paramétrica que describe una primera porción de la señal multicanal, y cuando para una segunda porción de la señal multicanal no se incluyen segundos datos paramétricos o ningún dato paramétrico en la representación paramétrica, mientras que los segundos datos paramétricos requieren menos unidades de información que los primeros datos paramétricos cuando se describe una porción idéntica de la señal multicanal.

De esta manera, una primera porción de la señal multicanal se representa por los primeros parámetros que permiten una reconstrucción de la señal multicanal con calidad más alta y una segunda porción puede representarse por

segundos parámetros que permiten una reconstrucción con calidad ligeramente inferior. La tasa de transmisión de bits consumida por los primeros datos paramétricos es consecuentemente más alta que la tasa de transmisión de bits consumida por los segundos datos paramétricos cuando ambos datos paramétricos son para describir la misma porción de una señal multicanal. En otras palabras, los primeros parámetros requieren más tasa de transmisión de bits por porción de señal que los segundos parámetros.

El propósito de la invención es puentear el espacio entre ambos mundos de la técnica anterior al mejorar gradualmente el sonido de la señal de mezcla ascendente mientras eleva la tasa de transmisión de bits consumida por la información secundaria empezando desde 0 hasta la tasa de bits de los métodos paramétricos. Es decir, la presente invención ayuda a puentear el espacio en tasas de transmisión de bits y la calidad perceptual entre métodos completamente paramétricos y métodos de envolvente matricial. Más específicamente, proporciona un método para elegir de manera flexible un "punto de operación" en algún lugar entre envolvente matricial (sin información secundaria, calidad de audio limitada) y reconstrucción completamente paramétrica (tasa de información secundaria completa requerida, buena calidad). Este punto de operación puede elegirse dinámicamente (es decir, variando en tiempo) y en respuesta a la tasa de información secundaria permisible, según se dicte por la aplicación individual.

Al elegir de manera dinámica el tamaño de la primera porción de la señal de audio multicanal que es la parte de la señal de audio multicanal que se representa por los parámetros de audio espaciales, la tasa de transmisión de bits demandada puede variarse dentro de un intervalo amplio. La representación de las partes principales de una señal multicanal por los parámetros de audio espaciales consumirá una tasa de transmisión de bits comparativamente alta en el beneficio de una buena calidad perceptual. Ya que para la segunda porción de la señal de audio multicanal se elige una regla para derivar un parámetro, lo que resulta en parámetros que consumen menos tasa de bits, la tasa de transmisión de bits total resultante puede disminuirse al incrementar el tamaño de la segunda porción de la señal multicanal. En una realización preferida de la presente invención, no se transmite ningún dato paramétrico en absoluto para la segunda porción de la señal multicanal, que por supuesto ahorra más bits. Por lo tanto, al cambiar de manera dinámica el tamaño de la primera porción con respecto al tamaño de la segunda porción, la tasa de transmisión de bits (o la calidad perceptual) puede ajustarse de manera dinámica a las necesidades.

En la presente invención, una señal de mezcla descendente se deriva en una manera compatible con la matriz. Por lo tanto, la primera porción de la señal de audio multicanal puede reproducirse con alta calidad perceptual usando los parámetros de audio espaciales y la segunda porción de la señal multicanal puede reproducirse usando soluciones basadas en matriz. Esto permite una reproducción de alta calidad de partes de las señales que requieren calidad más alta. Al mismo tiempo, la tasa de transmisión de bits total se disminuye al depender de la reproducción basada en matriz para las partes de señal menos vitales para la calidad de una señal reproducida.

En la presente invención, el concepto inventivo se aplica en el lado de decodificador dentro de una representación QMF de una señal de mezcla descendente recibida. El proceso de mezclado ascendente puede subdividirse principalmente en tres etapas:

pre-procesamiento de las señales de entrada (señales de mezcla descendente recibidas en el dominio QMF) mediante la aplicación de una matriz de pre-decorrelacionador;

decorrelación de parte de las señales pre-procesadas; y

mezcla de las señales así derivadas (señales pre-procesadas y señales decorrelacionadas) dentro de una matriz de mezcla, siendo la salida de la mezcla los canales de la señal de mezcla ascendente.

Tanto la matriz de pre-decorrelacionador así como la matriz mezclada son matrices bidimensionales con las dimensiones "número de espacios de tiempo" por un lado y "número de bandas de parámetro" por el otro lado. Dentro de un proceso de decodificación, los elementos de estas matrices se rellenan con valores que se derivan de los parámetros leídos de la secuencia de bits espacial, es decir, por los primeros datos paramétricos. Cuando los primeros datos paramétricos únicamente se reciben para una primera porción de la señal multicanal, únicamente esa porción de una reconstrucción de una señal multicanal puede derivarse usando los primeros datos paramétricos presentados. Los elementos de matriz para derivar la segunda parte de la reconstrucción de la señal multicanal, de acuerdo con la presente invención, se derivan usando esquemas de codificación compatibles con la matriz. Estos elementos de matriz, por lo tanto, pueden o bien derivarse basándose en únicamente el conocimiento logrado de la señal de mezcla descendente o bien reemplazarse por los valores predefinidos.

En una realización preferida, un decodificador de audio multicanal de acuerdo con la presente invención reconoce por la cantidad de los primeros datos paramétricos transmitidos, qué parte de la matriz o qué parte de la señal de audio multicanal tiene que procesarse por la regla que depende de los parámetros espaciales y qué parte tiene que procesarse por la solución basada en matriz.

En otro ejemplo, un codificador de audio crea información de ventana, indicando qué partes de una señal multicanal tienen que procesarse por la solución basada en matriz o por el procedimiento compatible con audio espacial. La información de ventana se incluye en la representación paramétrica de una señal multicanal.

- 5 Un decodificador inventivo, por lo tanto, es capaz de recibir y procesar la información de ventana creada para aplicar las reglas de mezclado ascendente apropiadas en las porciones de la señal de audio multicanal indicadas por la información de ventana.

- 10 En una realización preferida de la presente invención, el concepto inventivo se aplica en el dominio QMF durante el procesamiento de señal, es decir, en un dominio en el que las señales se representan mediante múltiples representaciones, manteniendo cada representación información acerca de una cierta banda de frecuencia.

- 15 En una realización preferida adicional de la presente invención, el método libre de información secundaria (procedimiento basado en matriz) se aplica únicamente a las partes de frecuencia más alta mientras aplica (explícitamente) información paramétrica (es decir, la primera regla de codificación y decodificación) para una reproducción apropiada de las partes de baja frecuencia. Esto es ventajoso debido a la propiedad del oído humano para notar desviaciones pequeñas de dos señales similares (por ejemplo, desviaciones de fase) mucho más fácil para frecuencias bajas que para frecuencias altas.

- 20 Un gran beneficio de la presente invención es que se consigue una compatibilidad inversa de un esquema de codificación y decodificación de audio espacial con soluciones basadas en matriz sin tener que introducir hardware o software adicional cuando las reglas de codificación y decodificación de los codificadores de audio espaciales se eligen de manera apropiada.

- 25 Además, la compatibilidad se consigue sin tener que transmitir datos adicionales, como es el caso en otros intentos de la técnica anterior. El esquema de codificación de acuerdo con la presente invención es además extremadamente flexible, ya que permite un ajuste sin interrupciones de la tasa de transmisión de bits o la calidad, es decir, una transmisión suave entre codificación basada en matriz completa a codificación de audio espacial completa de una señal dada. Es decir, el esquema de codificación aplicado puede ajustarse a las necesidades reales, ya sea con respecto a la tasa de transmisión de bits requerida o con respecto a la calidad deseada.

Breve descripción de los dibujos

- 35 Las realizaciones preferidas de la presente invención se describen posteriormente haciendo referencia a los dibujos adjuntos, en los que:

- la figura 1 muestra un codificador;
la figura 2 muestra un ejemplo de una secuencia de bits del parámetro creada por el concepto a modo de ejemplo;
40 la figura 2a muestra un transcodificador;
la figura 3 muestra un decodificador inventivo;
la figura 4 muestra un ejemplo de un decodificador de audio espacial que implementa el concepto inventivo;
la figura 5 ilustra el uso de los diferentes esquemas de codificación en un lado de decodificador;
la figura 6 muestra un codificador de la técnica anterior;
45 la figura 7 muestra un decodificador de la técnica anterior;
la figura 8 muestra un diagrama de bloques de un banco de filtro; y
la figura 9 muestra un ejemplo adicional de un codificador de la técnica anterior.

Descripción detallada de las realizaciones preferidas

- 50 La figura 1 muestra un codificador multicanal. El codificador multicanal 100 tiene un generador de parámetros 102 y una interfaz de salida 104.

- 55 Una señal de audio multicanal 106 se introduce en el codificador 100, en el que se procesan una primera porción 108 y una segunda porción 110 de la señal multicanal 106. El generador de parámetros 102 recibe la primera porción 108 y la segunda porción 110 y deriva parámetros espaciales que describen propiedades espaciales de la señal multicanal 106.

- 60 Los parámetros espaciales se transfieren a la interfaz de salida 104 que deriva una representación paramétrica 112 de la señal multicanal 106 de manera que la representación paramétrica 112 incluye primeros datos paramétricos para una primera porción 108 de la señal multicanal y en el que para una segunda porción 110 de la señal multicanal 106 segundos datos paramétricos que requieren menos información que los primeros datos paramétricos o ningún dato paramétrico se incluyen en la representación paramétrica 112.

Son posibles varias variaciones del codificador multicanal 100 para lograr el mismo objetivo. Por ejemplo, el generador de parámetros 102 puede aplicar dos reglas de derivación de parámetros diferentes en la primera porción 108 y en la segunda porción 110 lo que resulta en diferentes conjuntos de parámetros que se transfieren entonces a la interfaz de salida 104 que combina los diferentes conjuntos de parámetros en la representación paramétrica 112.

5 Un caso especial y preferido es que para la segunda porción 110 no se incluya ningún parámetro en la representación paramétrica (y por lo tanto no se deriva mediante el generador de parámetros 102) ya que, en un lado de decodificador, el decodificador deriva los parámetros de decodificación requeridos mediante algunas reglas heurísticas.

10 Otra posibilidad es que el generador de parámetros 102 derive un conjunto completo de parámetros de audio espaciales tanto para la primera porción 108 como para la segunda porción 110. Por lo tanto, la interfaz de salida 104 tendría que procesar los parámetros espaciales de manera que los segundos datos paramétricos requieren menos bits que los primeros datos paramétricos.

15 Además, la interfaz de salida 104 podría agregar una señal de ventana adicional a la representación paramétrica 112 que señalaría a un decodificador, cómo se dividió la señal multicanal 106 en la primera porción 108 y en la segunda porción 110 durante la codificación. En una modificación de la realización preferida de un codificador multicanal 100, el codificador multicanal 100 puede tener adicionalmente un decodificador de porción para decidir qué parte de la señal multicanal 106 se usa como la primera porción 108 y qué parte se usa como la segunda porción 20 110, basándose la decisión en un criterio de calidad.

El criterio de calidad puede derivarse con respecto a una tasa de transmisión de bits total resultante de la representación paramétrica 112 o con respecto a los aspectos de calidad, teniendo en cuenta la calidad perceptual de una reproducción de la señal multicanal 106 basándose en la representación paramétrica 112.

25 Una ventaja principal es que la tasa de transmisión de bits consumida por la representación paramétrica puede de esta manera variarse en tiempo, asegurando que el criterio de calidad se satisface en cualquier momento durante la codificación mientras permite una reducción total de la tasa de transmisión de bits requerida en comparación con los métodos de la técnica anterior.

30 La figura 2 muestra un ejemplo de una representación paramétrica 112 creada por un codificador.

Como se menciona arriba, el procesamiento de las señales de audio se hace por bloques, es decir un número de muestras posteriores de la señal multicanal en el dominio de tiempo, que forma una denominada trama, se procesa en una etapa. La figura 2 muestra una secuencia de bits del parámetro, es decir, una representación paramétrica para dos tramas consecutivas. La secuencia de bits del parámetro tiene una representación de una trama de alta calidad 120 y una representación de una trama de calidad inferior 122. Durante la codificación de la trama de alta calidad 120, se toma la decisión de que la primera porción 108, que se representa mediante datos paramétricos, tiene que ser grande en comparación con la segunda porción, que puede, por ejemplo, ser el caso si la escena de audio por codificar es compleja. La secuencia de bits del parámetro de la figura 2 se crea además bajo la suposición de que se usa un codificador a modo de ejemplo que no deriva ningún dato paramétrico para la segunda porción 110 de la señal multicanal 106. Como puede observarse en la figura 2, 28 parámetros espaciales ICC e ICLD se incluyen en la representación paramétrica para describir la trama de alta calidad 120. Por ejemplo, los 28 parámetros espaciales describen las bandas de frecuencia inferior de una representación QMF de la señal multicanal.

45 La trama de calidad inferior 122 comprende únicamente 21 conjuntos de parámetros espaciales que tienen ICC e ICLD ya que se encontró que esto es suficiente para la calidad perceptual deseada.

50 La figura 2a muestra un transcodificador 150. El transcodificador recibe como una entrada una secuencia de bits de entrada 152 que tiene un conjunto completo de parámetros espaciales que describen una primera trama 154 y una segunda trama 156 de una señal de audio multicanal.

El transcodificador 150 genera una secuencia de bits 158 que mantiene una representación paramétrica que representa las propiedades espaciales de la señal de audio multicanal. En el ejemplo mostrado en la figura 2a, el transcodificador 150 deriva la representación paramétrica de manera que para la primera trama el número de parámetros 160 únicamente se disminuye ligeramente. El número de parámetros 162 que describe la segunda trama correspondiente a los parámetros de entrada 156 disminuye fuertemente, lo que reduce significativamente la cantidad de tasa de transmisión de bits necesaria para la representación paramétrica resultante. Tal transcodificador 150 puede, por lo tanto, usarse para posprocesar una secuencia de bits ya existente de parámetros espaciales para derivar una representación paramétrica inventiva que requiere menos tasa de transmisión de bits durante la transmisión o menos espacio de almacenamiento cuando se almacena en un medio legible por ordenador. Debe observarse aquí que por supuesto, también es posible implementar un transcodificador para transcodificar en la otra dirección, es decir usar la representación paramétrica para generar parámetros espaciales.

El transcodificador 150 puede implementarse de varias maneras diferentes, como, por ejemplo, reduciendo la cantidad de parámetros con una regla dada o recibiendo adicionalmente la señal de audio multicanal para analizar la reducción de tasa de transmisión de bits posible sin perturbar la calidad perceptual más allá de un límite aceptable.

- 5 La figura 3 muestra un decodificador inventivo de audio multicanal 200 que tiene un procesador 202.

El procesador recibe como una entrada una señal de mezcla descendente 204 derivada de una señal de audio multicanal, primeros datos paramétricos 206 que describen una primera porción de la señal multicanal y, para una segunda porción de la señal multicanal, segundos datos paramétricos opcionales 208 requieren menos bits que los primeros datos paramétricos 206. El procesador 202 deriva una señal intermedia 210 de la señal de mezcla descendente 204 usando una primera regla de derivación para derivar una porción de alta calidad 212 de la señal intermedia, en el que la porción de alta calidad 212 de la señal intermedia 212 corresponde a la primera porción de la señal de audio multicanal. El procesador 202 usa una segunda regla de derivación para una segunda porción 214 de la señal intermedia 210, en el que la segunda regla de derivación usa los segundos datos paramétricos o ningún dato paramétrico y en el que la primera regla de derivación depende de los primeros datos paramétricos 206.

La señal intermedia 210 derivada mediante el procesador 202 se forma a partir de una combinación de la porción de alta calidad 212 y de la segunda porción 214.

20 El decodificador de audio multicanal 200 puede derivar por sí mismo qué porciones de la señal de mezcla descendente 204 tienen que procesarse con los primeros datos paramétricos 206 aplicando algunas reglas apropiadas, por ejemplo, contando el número de parámetros espaciales incluidos en los primeros datos paramétricos 206. Alternativamente, pueden señalizarse al procesador 202 las fracciones de la porción de alta calidad 212 y de la segunda porción 214 dentro la señal de mezcla descendente 204 mediante alguna información de ventana adicional que se deriva en un lado de codificador y que se transmite adicionalmente al decodificador de audio multicanal 200.

En una realización preferida, se omiten los segundos datos paramétricos 208 y el procesador 202 deriva la segunda regla de derivación a partir de la información ya contenida en la señal de mezcla descendente 204.

30 La figura 4 muestra una realización adicional de la presente invención que combina la característica inventiva de compatibilidad de matriz en un decodificador de audio espacial. El decodificador de audio multicanal 600 comprende un pre-decorrelacionador 601, un decorrelacionador 602 y una matriz de mezcla 603.

El decodificador de audio multicanal 600 es un dispositivo flexible que permite operar en diferentes modos dependiendo de la configuración de señales de entrada 605 introducidas en el pre-decorrelacionador 601. Generalmente, el pre-decorrelacionador 601 deriva señales intermedias 607 que sirven como entrada para el decorrelacionador 602 y que se transmiten parcialmente sin alterar para formar, junto con las señales decorrelacionadas calculadas por el decorrelacionador 602, señales de entrada 608. Las señales de entrada 608 son las señales introducidas en la matriz de mezcla 603 que deriva configuraciones de canal de salida 610a o 610b, dependiendo de la configuración de canal de entrada 605.

En una configuración de 1 a 5, una señal de mezcla descendente y una señal residual opcional se suministran al pre-decorrelacionador 601, que deriva cuatro señales intermedias (e1 a e4) que se usan como una entrada del decorrelacionador, que deriva cuatro señales decorrelacionadas (d1 a d4) que forman los parámetros de entrada 608 junto con una señal directamente transmitida m derivada de la señal de entrada.

Puede observarse, que en el caso en el que se suministra una señal residual adicional como entrada, el decorrelacionador 602 que funciona generalmente en un dominio sub-banda, puede ser operativo para reenviar simplemente la señal residual en lugar de derivar una señal decorrelacionada. Esto también puede hacerse de una manera selectiva de frecuencia para únicamente ciertas bandas de frecuencia.

En la configuración 2 a 5, las señales de entrada 605 comprenden un canal izquierdo, un canal derecho y opcionalmente una señal residual. En esa configuración, la matriz de pre-decorrelacionador 601 deriva un canal izquierdo, uno derecho y uno central y además dos canales intermedios (e1, e2). Por lo tanto, las señales de entrada a la matriz de mezcla 603 se forman por el canal izquierdo, el canal derecho, el canal central y dos señales decorrelacionadas (d1 y d2).

En una modificación adicional, la matriz de pre-decorrelacionador puede derivar una señal intermedia adicional (e5) que se usa como una entrada para un decorrelacionador (D5) cuya salida es una combinación de la señal decorrelacionada (d5) derivada de la señal (e5) y las señales decorrelacionadas (d1 y d2). En este caso, puede garantizarse una decorrelación adicional entre el canal central y el canal derecho e izquierdo.

El decodificador de audio inventivo 600 implementa el concepto inventivo en la configuración 2 a 5. La representación paramétrica transmitida se usa en la matriz de pre-decorrelación 601 y en la matriz de mezcla 603.

Aquí, el concepto inventivo puede implementarse de diferentes maneras como se muestra en más detalle en la figura 5.

5 La figura 5 muestra el pre-decorrelacionador, implementado como matriz de pre-decorrelacionador 601 y la matriz de mezcla 603 en un boceto del principio, en el que se omiten los otros componentes del decodificador de audio multicanal 600.

10 La matriz usada para realizar la pre-decorrelación y la mezcla tiene columnas que representan espacios de tiempo, es decir, las muestras de tiempo individuales de una señal y filas que representan las diferentes bandas de parámetro, es decir, cada fila se asocia con una banda de parámetro de una señal de audio.

15 De acuerdo con el concepto de la presente invención, los elementos de matriz de las matrices 601 y 603 se derivan únicamente de manera parcial a partir de datos paramétricos transmitidos, en el que los elementos de matriz restantes se derivan mediante el decodificador, basándose en, por ejemplo, el conocimiento de la señal de mezcla descendente. La figura 5 muestra un ejemplo en el que por debajo de una línea límite de frecuencia dada 622 los elementos de la matriz de pre-decorrelacionador 601 y la matriz de mezcla 603 se derivan a partir de los parámetros 620 que se leen de la secuencia de bits, es decir, basándose en la información transmitida del codificador. Por encima de la línea límite de frecuencia 622 los elementos de matriz se derivan en el decodificador basándose en el conocimiento de la señal de mezcla descendente únicamente.

20 La frecuencia límite (o en general: la cantidad de elementos de matriz derivados a partir de datos transmitidos) puede adaptarse libremente de acuerdo con la calidad y/o limitaciones de la tasa de bits que tienen que satisfacerse para el escenario de aplicación particular.

25 Se prefiere para el nuevo método de codificación aquí esbozado que un proceso de mezcla ascendente libre de información secundaria pueda realizarse con la misma estructura que se ha esbozado en el Modelo de Referencia de Codificación de Audio Espacial MPEG 0. Esta invención puede consistir en describir un método para mezcla ascendente libre de información secundaria, pero preferentemente proporciona un método para combinación ventajosa y sin interrupciones de tales conceptos con métodos para mezcla ascendente asistida por información secundaria.

30 Al contrario que el Modelo de Referencia de Codificación de Audio Espacial MPEG 0, en el proceso de mezcla ascendente libre de información secundaria los elementos de las matrices M1 (601) y M2 (603) no se derivan preferentemente de datos transmitidos en una secuencia de bits sino mediante diferentes medios sin la ayuda de la información secundaria, por ejemplo, aplicando reglas heurísticas basándose en únicamente un conocimiento conseguido a partir de la señal de mezcla descendente.

40 De esta manera es posible lograr una escala gradual entre ambas técnicas - en términos de tasa de transmisión de bits, así como en términos de calidad de sonido - adquiriendo únicamente partes de las matrices basándose en los parámetros transmitidos y aplicando las reglas del método sin información secundaria para rellenar las partes restantes. Conceptualmente hablando, esto corresponde a transmitir los parámetros espaciales para ciertas partes de las matrices y generar los mismos para otras partes en el decodificador.

45 La determinación de las partes de las matrices que tienen que derivarse a partir de o bien un método o bien el otro puede hacerse de muchas diferentes maneras, tales como

- derivar las partes de las matrices por debajo de una línea límite horizontal dada por un método y por encima de esta línea límite por el otro método;
- derivar las partes de las matrices izquierda de una línea de límite vertical dada por un método y derecha de esta línea de límite por el método;
- determinar las cintas de frecuencia de tiempo arbitrarias dentro de ambas matrices, los elementos de las cuales se derivan por un método y derivar los elementos de las cintas de frecuencia de tiempo restantes por medio del otro método.

55 Se ha detallado en los párrafos anteriores es ventajoso describir todas las partes de frecuencia de una señal multicanal hasta una cierta frecuencia límite mediante parámetros espaciales mientras las partes de frecuencia restantes de la señal multicanal no se representan mediante parámetros espaciales. Esto tiene en cuenta las características del oído humano que tiene una mejor percepción de frecuencias más bajas que de frecuencias más altas. Por su puesto, la presente invención no se limita a esta división de la señal multicanal en una primera porción y una segunda porción ya que también puede ser ventajoso o apropiado describir partes de frecuencia más alta de la señal con mejor precisión. Esto puede ser especialmente el caso cuando en la región de frecuencia más baja se contiene únicamente poca energía en la señal ya que la mayoría de la energía se contiene en un dominio de alta frecuencia de la señal de audio. Debido a los efectos de enmascaramiento la parte de frecuencia más baja mayormente se dominará por las partes de alta frecuencia y puede ser ventajoso proporcionar la posibilidad de una

reproducción de alta calidad de la parte de frecuencia alta de la señal.

- Dependiendo de ciertos requerimientos de implementación de los métodos inventivos, los métodos inventivos pueden implementarse en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento digital, en particular un disco, DVD o un CD que tiene señales de control electrónicamente legibles almacenadas en el mismo, que cooperan con un sistema informático programable de manera que se realizan los métodos inventivos. Generalmente, la presente invención es, por lo tanto, un producto de programa informático con un código de programa almacenado en una portadora legible por máquina, siendo el código de programa operativo para realizar los métodos inventivos cuando el producto de programa informático se ejecuta en un ordenador. En otras palabras, los métodos inventivos son, por lo tanto, un programa informático que tiene un código de programa para realizar al menos uno de los métodos inventivos cuando el programa informático se ejecuta en un ordenador.

- Mientras que lo anterior se ha mostrado particularmente y descrito con referencia a realizaciones particulares, los expertos en la técnica entenderán que se pueden realizar diversos cambios diferentes en la forma y detalles sin alejarse del alcance de las mismas. Debe entenderse que se pueden realizar diversos cambios en la adaptación a diferentes realizaciones sin alejarse de los conceptos más amplio dados a conocer en el presente documento y entendidos a partir de las siguientes reivindicaciones.

REIVINDICACIONES

1. Decodificador de audio multicanal (200) para procesar una señal de audio de mezcla descendente (204) derivada de una manera compatible con matriz y para procesar primeros datos paramétricos (206) que describen una primera porción de una señal multicanal, en el que para una segunda porción de la señal multicanal no se procesan datos paramétricos o segundos datos paramétricos (208), requiriendo los segundos datos paramétricos (208) menos unidades de información que los primeros datos paramétricos (206) cuando se describe una porción idéntica de la señal multicanal, que comprende:
 - un procesador (202) para derivar una señal intermedia (210) a partir de la señal de audio de mezcla descendente (204), estando la señal de audio de mezcla descendente (204) en una representación de QMF, Banco de filtros de espejo en cuadratura, usando una primera regla de derivación para derivar una primera porción de la señal intermedia (210), correspondiendo la primera porción de la señal intermedia (210) a la primera porción de la señal de audio multicanal, en el que la primera regla de derivación depende de los primeros datos paramétricos (206); y usando una segunda regla de derivación para derivar una segunda porción de la señal intermedia (210), no usando la segunda regla de derivación ningún dato paramétrico o los segundos datos paramétricos (208), en el que la primera porción es una porción de tiempo o una porción de frecuencia.
2. Decodificador de audio multicanal de acuerdo con la reivindicación 1, en el que el procesador (202) se configura para recibir, como una entrada, la señal de audio de mezcla descendente (204), y los primeros datos paramétricos (206), y para construir la señal intermedia mediante una combinación de la primera porción de la señal intermedia (210) que es una porción de alta calidad y la segunda porción de la señal intermedia (210).
3. Decodificador de audio multicanal (200) de acuerdo con la reivindicación 1, que es operativo para procesar primeros datos paramétricos (206) que comprenden una descripción de una porción de tiempo de la señal de audio multicanal, en el que solo información en una banda de frecuencia predeterminada de la señal de audio multicanal está contenida dentro de la descripción.
4. Decodificador de audio multicanal (200) de acuerdo con la reivindicación 1, en el que el procesador (202) es operativo para usar la primera regla de derivación para derivar una porción de frecuencia de la señal intermedia (210) que oscila de un límite de baja frecuencia a un límite de alta frecuencia.
5. Decodificador de audio multicanal (200) de acuerdo con la reivindicación 4, que es operativo para procesar adicionalmente una información de ventana que señala al menos el límite de alta frecuencia.
6. Decodificador de audio multicanal (200) de acuerdo con la reivindicación 4, que es operativo para derivar al menos el límite de alta frecuencia de los primeros datos paramétricos (206) usando una regla de estimación de frecuencia.
7. Decodificador de audio multicanal (200) de acuerdo con la reivindicación 6, en el que la regla de estimación de ventana incluye contar la cantidad de datos paramétricos presentados para una porción de señal de la señal de audio de mezcla descendente (204) y comparar la cantidad contada de datos paramétricos con una cantidad nominal de datos paramétricos.
8. Decodificador de audio multicanal (200) de acuerdo con la reivindicación 1, en el que el procesador (202) es operativo para calcular la segunda regla de derivación de información sobre la señal de audio de mezcla descendente (204).
9. Decodificador de audio multicanal (200) de acuerdo con la reivindicación 1, en el que el procesador (202) es operativo para calcular la segunda regla de derivación a partir de los segundos datos paramétricos (208).
10. Decodificador de audio multicanal (200) de acuerdo con la reivindicación 1, en el que el procesador (202) es operativo para usar una regla predefinida como la segunda regla de derivación.
11. Decodificador de audio multicanal (200) de acuerdo con la reivindicación 1, que comprende además un decorrelacionador para derivar una señal decorrelacionada a partir de la señal intermedia (210) usando un filtro de decorrelación.
12. Decodificador de audio multicanal (200) de acuerdo con la reivindicación 1 que comprende además un mezclador ascendente para calcular una reconstrucción de la señal multicanal basándose en la señal intermedia

(210),

5 usar una primera regla de mezcla ascendente para calcular una primera porción de la reconstrucción de la señal multicanal correspondiente a la primera porción de la señal multicanal, en el que la primera regla de mezcla ascendente depende de los primeros datos paramétricos (206); y

10 usar una segunda regla de mezcla ascendente para calcular una segunda porción de la reconstrucción de la señal multicanal, no usando la segunda regla de mezcla ascendente ningún dato paramétrico o los segundos datos paramétricos (208).

13. Método para procesar una señal de audio de mezcla descendente (204) derivada de una manera compatible con matriz y para procesar primeros datos paramétricos (206) que describen una primera porción de una señal multicanal, en el que para una segunda porción de la señal multicanal no se procesan datos paramétricos o segundos datos paramétricos (208), requiriendo los segundos datos paramétricos (208) menos unidades de información que los primeros datos paramétricos (206) cuando se describe una porción idéntica de la señal multicanal, comprendiendo el método:

20 derivar una señal intermedia (210) a partir de la señal de audio de mezcla descendente (204) usando una primera regla de derivación que depende de los primeros datos paramétricos (206) para derivar una primera porción de la señal intermedia (210), correspondiendo la primera porción de la señal intermedia (210), a la primera porción de la señal de audio multicanal, la señal de audio de mezcla descendente (204) está en una representación de QMF, Banco de filtros de espejo en cuadratura; y

25 derivar una segunda porción de la señal intermedia (210) usando una segunda regla de derivación, usando la segunda regla de derivación los segundos datos paramétricos (208) o ningún dato paramétrico,

en el que la primera porción es una porción de tiempo o una porción de frecuencia.

30 14. Representación paramétrica que describe propiedades espaciales de una señal de audio multicanal, incluyendo la representación paramétrica primeros datos paramétricos para una primera porción de la señal multicanal y en la que para una segunda porción de la señal multicanal no se incluyen datos paramétricos o segundos datos paramétricos, requiriendo los seguidos datos paramétricos menos unidades de información que los primeros datos paramétricos para una porción idéntica de la señal multicanal, en la que la representación paramétrica comprende adicionalmente una señal de mezcla descendente derivada de la señal multicanal de una manera compatible con matriz, y en la que la primera porción es una porción de tiempo o una porción de frecuencia, y en el que los datos paramétricos comprenden uno o más parámetros de audio espacial elegidos de la siguiente lista de parámetros de audio espacial: ICC (correlación inter-canal); CLD (diferencia de nivel de canal); y CPC (coeficiente de predicción de canal).

40 15. Programa informático que tiene un código de programa para realizar, cuando se ejecuta en un ordenador, un método de la reivindicación 13.

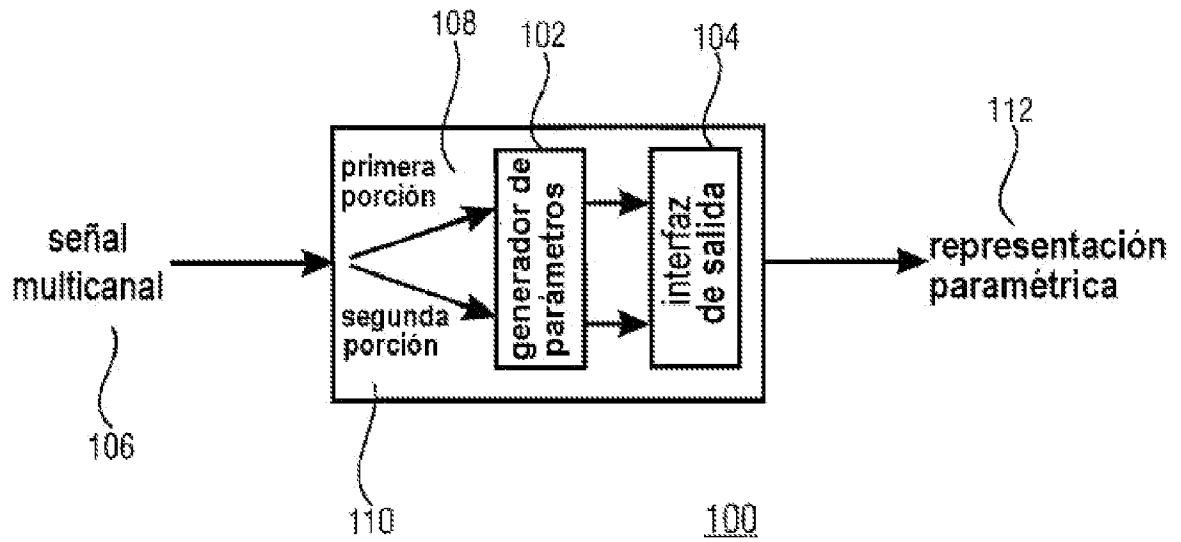


FIGURA 1

Secuencia de bits de parámetros

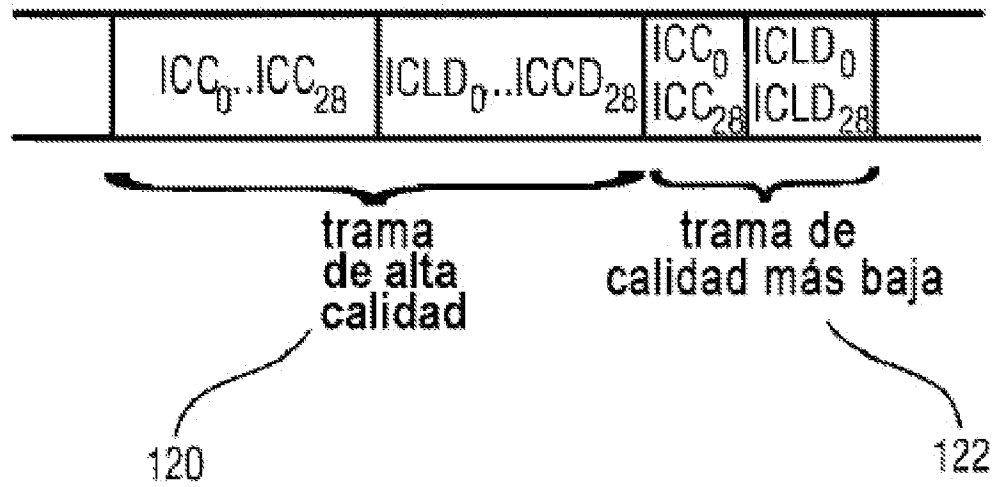


FIGURA 2

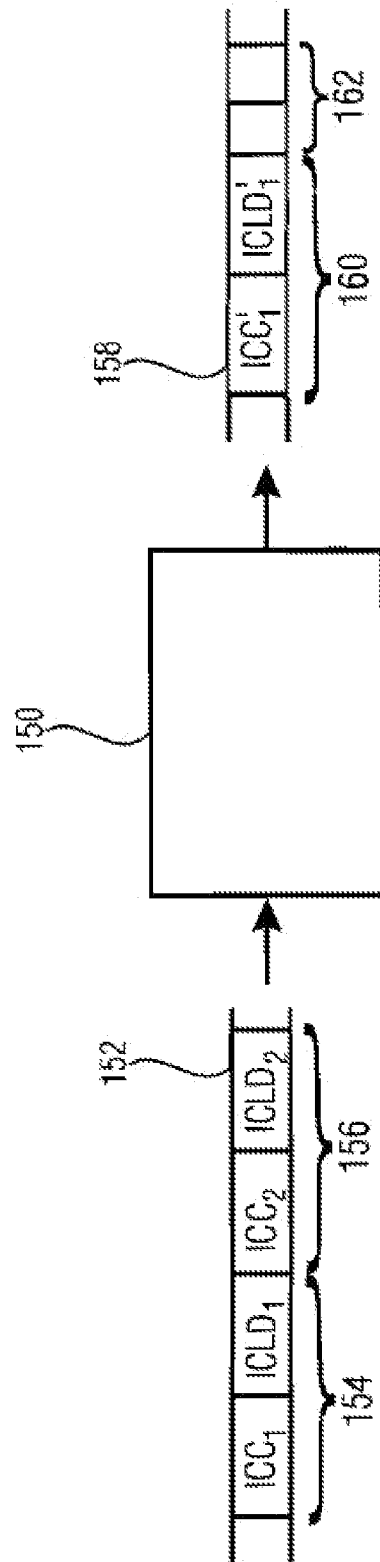


FIGURA 2A

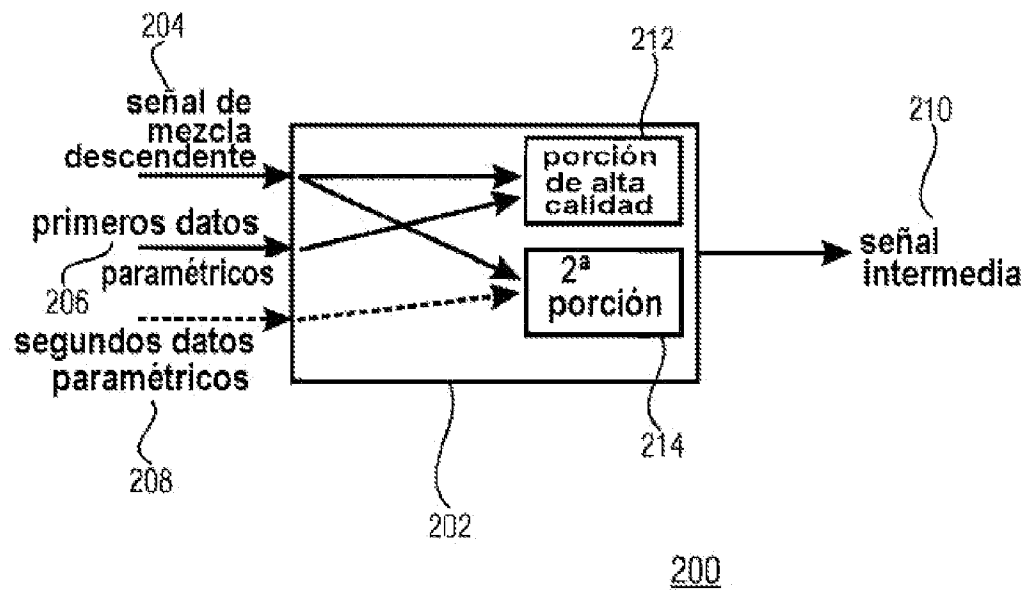


FIGURA 3

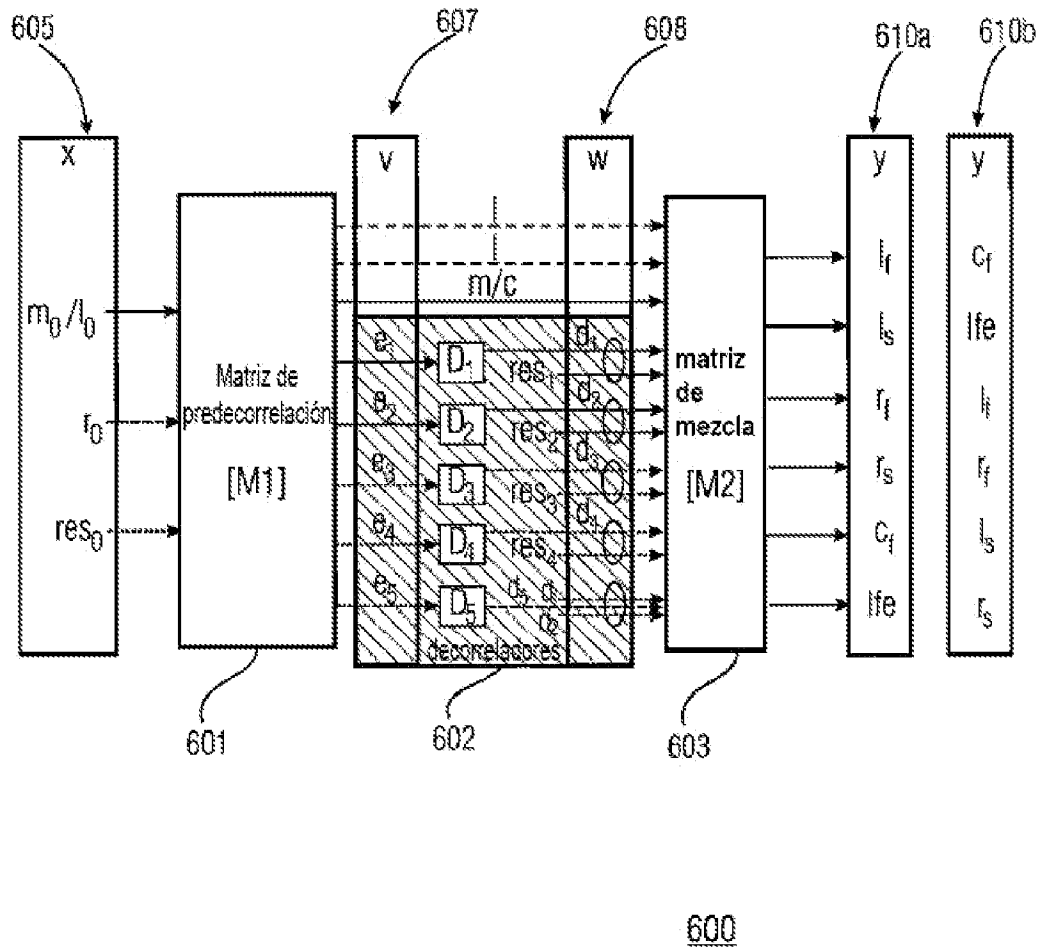


FIGURA 4

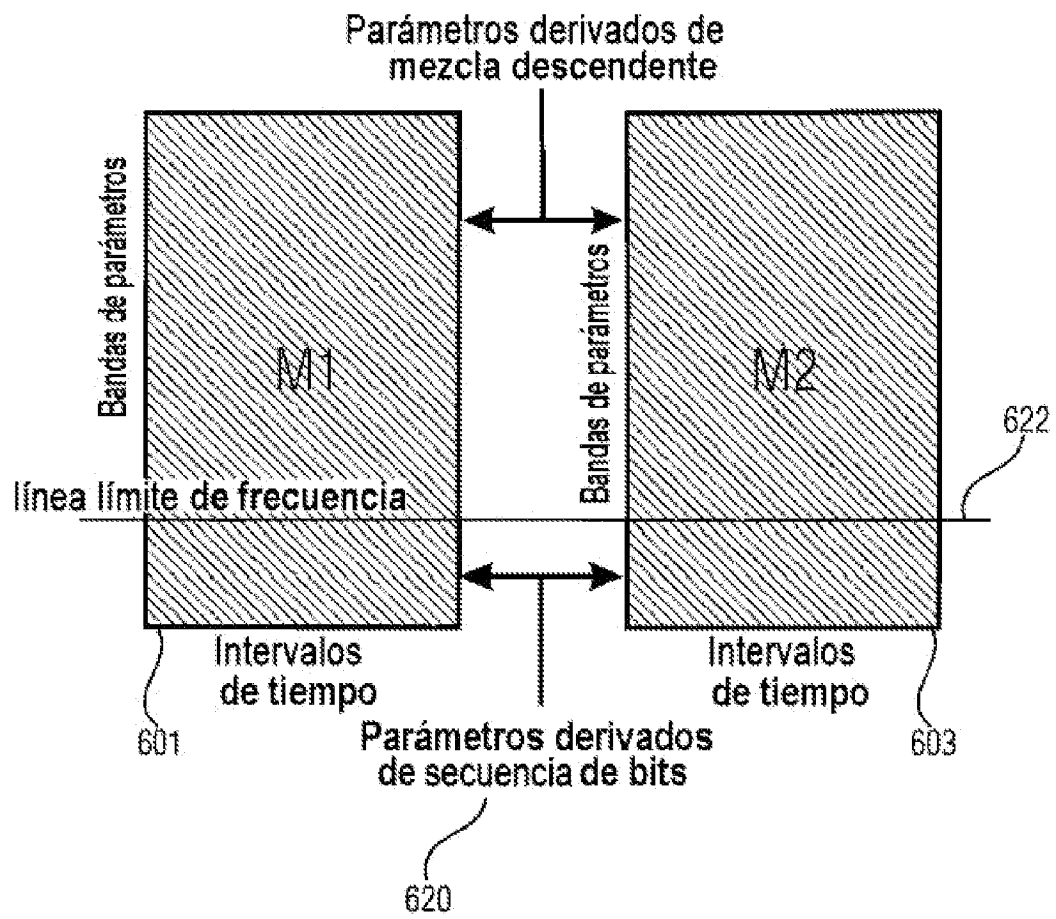


FIGURA 5

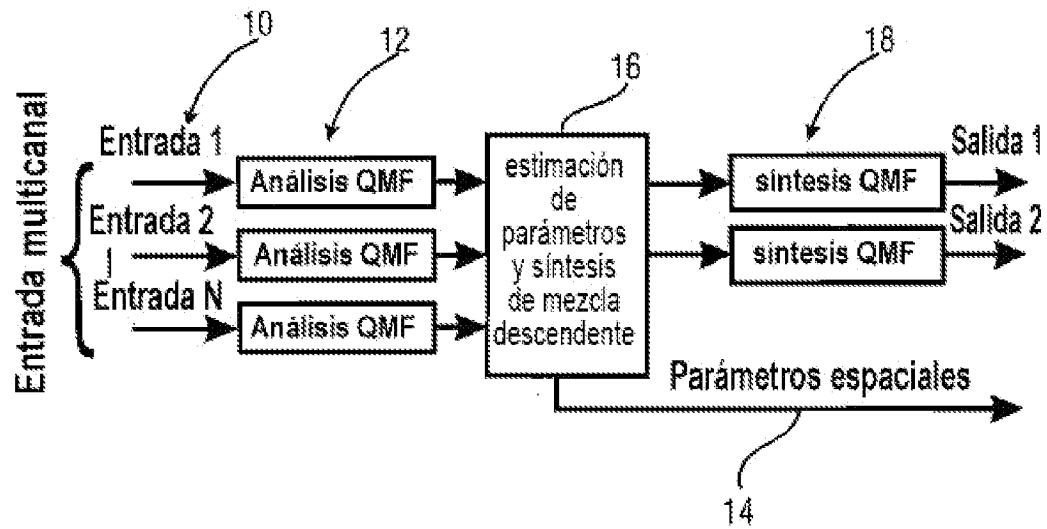


FIGURA 6

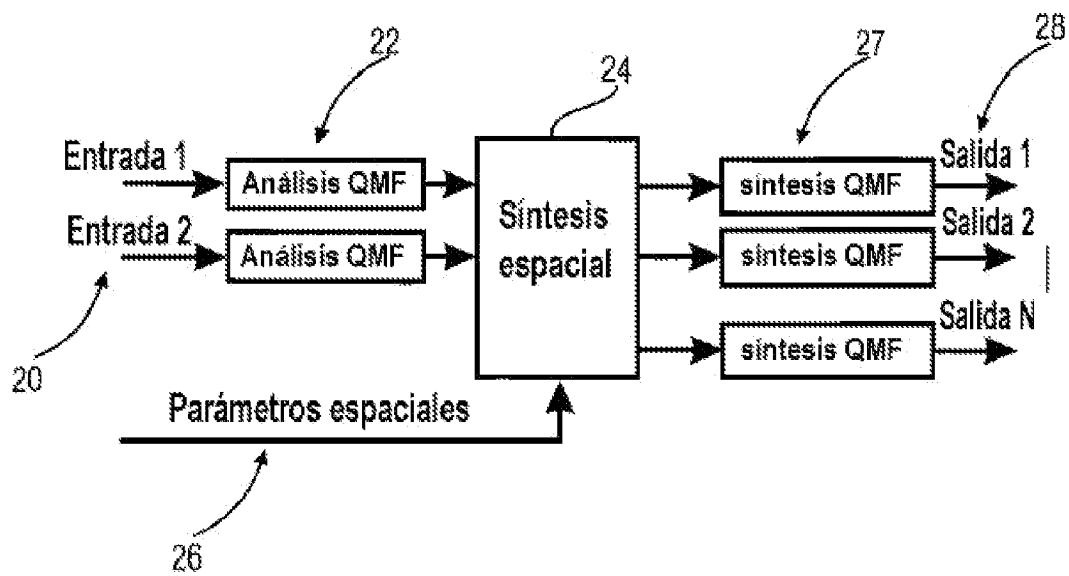


FIGURA 7

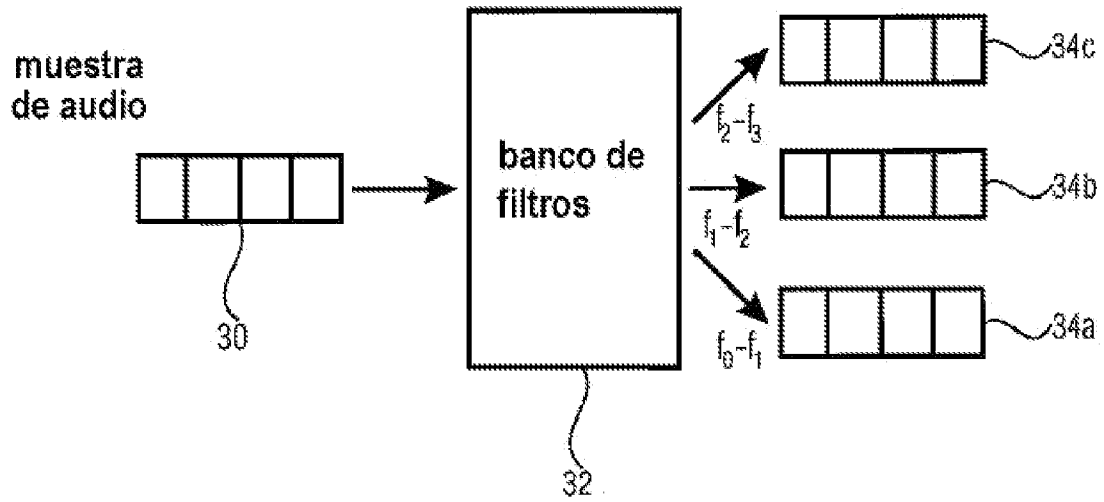


FIGURA 8

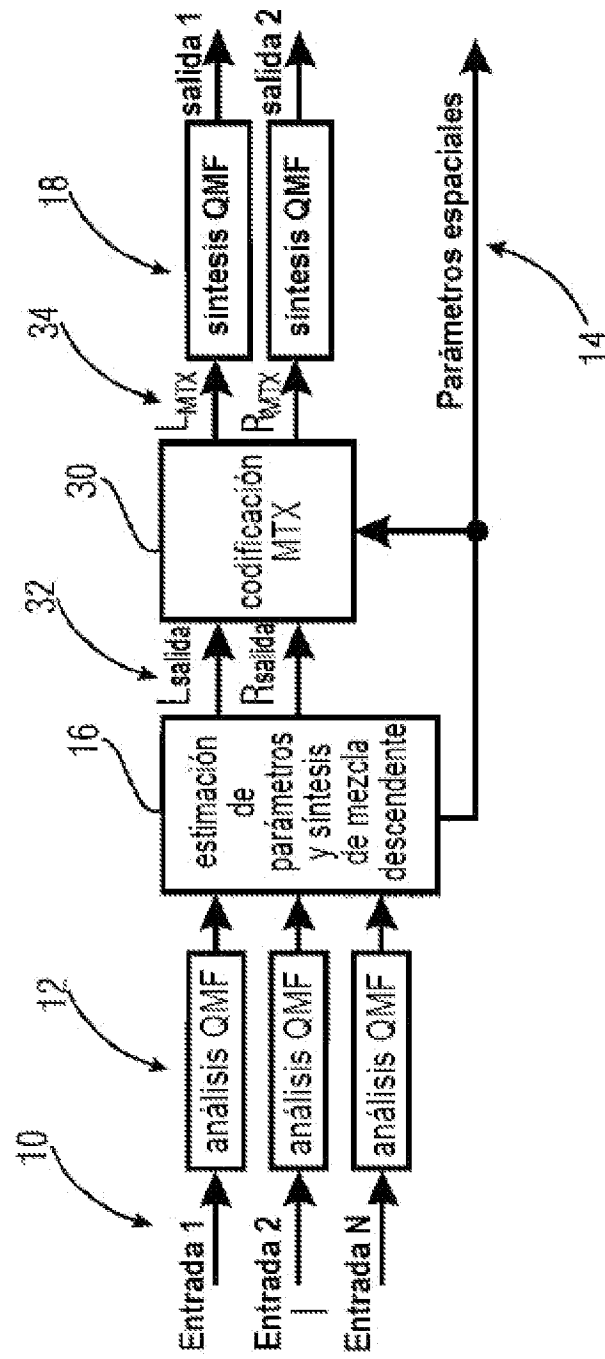


FIGURA 9