

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 995 111**

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 23214134 (1)**

97 Fecha y número de publicación de la concesión europea: **31.07.2024** **EP 4307126**

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

**06.02.2025**

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:

**ARIZTI ACHA, Monica**

ES 2 995 111 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 la codificación de audio multicanal paramétrica y la codificación multicanal envolvente matricial

5

### **Campo de la invención**

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

10

### **Antecedentes de la invención y técnica anterior**

La codificación paramétrica de señales de audio multicanal es un tema de investigación en curso. En general, pueden distinguirse dos enfoques para codificar señales de audio multicanal. El Grupo de Expertos en Imágenes en Movimiento (*Moving Pictures Experts Group*, MPEG), un subgrupo de la Organización Internacional de Normalización (ISO), está trabajando actualmente en la normalización de la tecnología para la reconstrucción de contenido de audio multicanal a partir de señales de mezcla descendente estéreo o incluso mono añadiendo sólo una pequeña cantidad de información auxiliar a las señales de mezcla descendente.

20

En paralelo, están desarrollándose métodos de mezcla ascendente estéreo a multicanal que no necesitan ninguna información secundaria adicional que no esté ya (implícitamente) contenida en la señal de mezcla descendente con el fin de reconstruir la imagen espacial de la señal de audio multicanal original.

25

Los métodos existentes para la transmisión multicanal compatible con estéreo sin información secundaria adicional que adquirieron relevancia práctica pueden caracterizarse principalmente como métodos envolventes matriciales, tales como Dolby Pro Logic (Dolby Pro Logic II) y Logic-7, tal como se describe con 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](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 AES, Los Angeles, EE. UU., 1996, preimpresión 4402. El principio común de estos métodos es que hacen uso de formas dedicadas de mezclado descendente multicanal o estéreo donde el codificador aplica desplazamientos de fase a los canales envolventes antes de mezclarlos junto con 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:

30

35

$$\begin{bmatrix} Lf \\ Rt \end{bmatrix} = \begin{bmatrix} 1 & 0 & q & a \cdot j & b \cdot j \\ 0 & 1 & q & -b \cdot j & -a \cdot j \end{bmatrix} \begin{bmatrix} Lf \\ Rf \\ C \\ Ls \\ Rs \end{bmatrix} \quad (1)$$

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 90 grados (j) y escalada en un factor a, y la señal envolvente derecha (Rs) que también está girada en fase 90 grados y escalada en 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 fundamento de los diferentes signos 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 antifase en el par de mezcla descendente (Lt, Rt). Esta propiedad ayuda al decodificador a discriminar entre canales delanteros y traseros del par de señales de mezcla descendente. Por tanto, la matriz de mezcla descendente permite una reconstrucción parcial de una señal de salida multicanal fuera de la mezcla descendente estéreo dentro del decodificador aplicando una operación de desmatrizado. Sin embargo, la proximidad con la que la señal multicanal recreada se asemeja a la señal de entrada del codificador original depende de las propiedades específicas del contenido de audio multicanal.

40

45

50

Un ejemplo de un método de codificación que añade información auxiliar, también denominada información secundaria, es la codificación de audio envolvente MPEG. Esta manera eficiente de 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 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.

55

En la figura 6 se muestra una vista general esquemática de un codificador usado en la codificación de audio espacial.

El codificador divide las señales entrantes 10 (entrada 1,... entrada N) en mosaicos de tiempo-frecuencia separados por medio de filtros espejo en cuadratura 12 (QMF). Los grupos de los mosaicos de frecuencia resultantes (bandas) se denominan "bandas de parámetros". Para cada banda de parámetros, se determinan varios parámetros espaciales 14 mediante un estimador de parámetros 16 que describe las propiedades de la imagen espacial, por ejemplo, diferencias de nivel entre pares de canales (CLD), correlación cruzada entre pares de canales (ICC) o información sobre envolturas de señal (CPC). Estos parámetros se cuantifican, codifican y compilan posteriormente de manera conjunta en un flujo de bits de datos espaciales. Dependiendo del modo de funcionamiento, este flujo de bits puede cubrir una amplia gama de velocidades de bits, partiendo de unos pocos kBit/s para audio multicanal de buena calidad hasta decenas de kBit/s para calidad casi transparente.

Además de la extracción de parámetros, el codificador genera también una mezcla descendente mono o estéreo a partir de la señal de entrada multicanal. Además, en 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 envolventes matriciales. Finalmente, la mezcla descendente estéreo se transfiere al dominio de tiempo por medio de bancos de síntesis de QMF 18. La mezcla descendente resultante puede transmitirse a un decodificador, acompañada por los parámetros espaciales o el flujo de bits de parámetros espaciales 14. Preferiblemente, la mezcla descendente también se codifica antes de la transmisión (usando un codificador central mono o estéreo convencional), mientras que los flujos de bits del codificador central y los parámetros espaciales pueden combinarse (multiplexarse) adicionalmente para formar un único flujo de bits de salida.

Un decodificador, como se representa en líneas generales en la figura 7, en principio realiza el proceso inverso del codificador. Un flujo de entrada se divide en un flujo de bits de codificador central y un flujo de bits de parámetros. Esto no se muestra en la figura 7. Posteriormente, la mezcla descendente decodificada 20 se procesa por un banco de análisis de QMF 22 para derivar bandas de parámetros que son las mismas que las aplicadas en el codificador. Una fase 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 de QMF se transfieren al dominio de tiempo por medio de un banco de síntesis de QMF 27 que deriva las señales de salida multicanal finales 28.

La figura 8 muestra un ejemplo simple de un análisis de QMF, tal 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 filtros 32. El banco de filtros 32 deriva tres muestras de salida 34a, 34b y 34c que tienen cuatro valores de muestra cada una. En un caso ideal, el banco de filtros 32 deriva las muestras de salida 34a a 34c de tal manera que las muestras dentro de las señales de salida sólo comprenden información sobre intervalos de frecuencia discretos de la señal de audio subyacente 30. En el caso mostrado en la figura 8, la muestra 34a tiene información sobre el intervalo de frecuencia que oscila entre  $f_0$  y  $f_1$ , la muestra 34b tiene información del intervalo de frecuencia  $[f_1, f_2]$  y la muestra 34c tiene información sobre el intervalo de frecuencia  $[f_2, f_3]$ . Aunque los intervalos de frecuencia en la figura 8 no se superponen, en un caso más general, los intervalos de frecuencia de las muestras de salida que salen de un banco de filtros pueden tener muy bien una superposición de frecuencias.

Un codificador de la técnica anterior puede, como ya se ha descrito anteriormente, suministrar o bien una mezcla descendente de estilo ITU o bien una mezcla descendente compatible envolvente matricial, cuando se desea una mezcla descendente de dos canales. En el caso de una mezcla descendente compatible envolvente matricial (usando, por ejemplo, el enfoque de matrizado proporcionado en la ecuación 1), una posibilidad sería que el codificador genere directamente una mezcla descendente compatible envolvente matricial.

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

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

Aunque ambos enfoques son adecuados para transmitir una señal multicanal, existen inconvenientes específicos de los sistemas del estado de la técnica. Los métodos de envolvente matricial son muy eficientes (ya que no se requieren parámetros adicionales) a costa de una calidad de reconstrucción multicanal muy limitada.

Por otro lado, los enfoques multicanal paramétricos requieren una velocidad de bits más alta debido a la información secundaria, lo que se convierte en un problema cuando se establece un límite como velocidad de bits máxima

aceptable para la representación paramétrica. Cuando los parámetros codificados requieren una cantidad comparativamente alta de velocidad de bits, la única manera posible de permanecer dentro de tal límite de velocidad de bits es disminuir la calidad de un canal de mezcla descendente codificado aumentando la compresión del canal. Por tanto, el resultado es una pérdida general en la calidad de audio, que puede ser inaceptablemente alta. En otras palabras, para enfoques multicanal paramétricos, hay a menudo un límite estricto de la velocidad de bits mínima que se requiere para la capa de parámetros espaciales, que en algunos casos puede ser inaceptablemente alta.

Aunque la compatibilidad hacia atrás principal entre los métodos envolventes matriciales y los métodos de audio espacial puede lograrse mediante un codificador de la técnica anterior como se ilustra en la figura 9, no puede ahorrarse ninguna velocidad de bits adicional con este enfoque cuando solo se requiere una decodificación basada en matriz. Incluso entonces, tiene que transmitirse el conjunto completo de parámetros espaciales, desperdiciando ancho de banda de transmisión.

Mientras que la velocidad 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 de audio suministrada por los métodos sin transmisión de información secundaria podría no ser suficiente.

La solicitud de patente estadounidense 2005157883 muestra un aparato para construir una señal de audio multicanal usando 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 la información secundaria paramétrica que describe interrelaciones entre canales de la señal original multicanal.

La solicitud de patente estadounidense 2005/157883 A1 da a conocer un aparato para construir una señal de audio multicanal usando 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 la información secundaria paramétrica que describe interrelaciones entre canales de la señal multicanal original.

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 indicación binaural (BCC). En la codificación BCC, los componentes espectrales de las señales de entrada se mezclan de manera descendente y se generan parámetros de BCC (por ejemplo, nivel intercanal y/o diferencias de tiempo). En una implementación estéreo, después de convertir los canales izquierdo y derecho al dominio de frecuencia, los pares de componentes espectrales de canal izquierdo y derecho se mezclan de manera descendente a mono. Los componentes mono se convierten entonces 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 pueden codificarse entonces usando técnicas de codificación convencionales. Para la reproducción, el flujo de bits codificado se decodifica usando técnicas de decodificación convencionales. Las técnicas de síntesis de BCC pueden aplicar entonces los parámetros de BCC para sintetizar una escena auditiva basándose en los componentes mono así como en los componentes estéreo no mezclados.

## **Sumario de la invención**

Un objeto de la presente invención es proporcionar un concepto para una codificación más eficiente de señales de audio multicanal al tiempo que es compatible hacia atrás con soluciones de codificación basadas en matriz.

Este objeto se logra mediante un decodificador de audio multicanal según la reivindicación 1, un método para el procesamiento según la reivindicación 12 o un programa informático según la reivindicación 13.

La presente invención se basa en el hallazgo de que una señal de audio multicanal puede representarse eficientemente mediante una representación paramétrica, cuando se usa una primera regla de derivación 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 se incluyen segundos datos paramétricos o no se incluyen datos paramétricos 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.

Por tanto, una primera porción de la señal multicanal está representada por primeros parámetros que permiten una reconstrucción de la señal multicanal con mayor calidad y una segunda porción puede estar representada por segundos parámetros que permiten una reconstrucción con una calidad ligeramente menor. La velocidad de bits consumida por los primeros datos paramétricos es, en consecuencia, mayor que la velocidad de bits consumida por los segundos datos paramétricos cuando ambos datos paramétricos van a describir la misma porción de una señal multicanal. En otras palabras, los primeros parámetros requieren más velocidad 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 mejorando gradualmente el sonido de la señal de mezcla ascendente al tiempo que se eleva la velocidad de bits consumida por la información secundaria partiendo de 0 hasta las velocidades de bits de los métodos paramétricos. Es decir, la presente invención tiene como objetivo puentear el espacio en las velocidades de bits y la calidad perceptual entre métodos completamente paramétricos y métodos envolventes matriciales. Más específicamente, proporciona un método de elección de manera flexible de un "punto operativo" en algún lugar entre la reconstrucción envolvente matricial (sin información secundaria, calidad de audio limitada) y la reconstrucción completamente paramétrica (velocidad de información secundaria completa requerida, buena calidad). Este punto operativo puede elegirse dinámicamente (es decir, variando en el tiempo) y en respuesta a la velocidad de información secundaria permisible, tal como lo dicte la aplicación individual.

Al elegir dinámicamente el tamaño de la primera porción de la señal de audio multicanal que es la parte de la señal de audio multicanal que está representada por los parámetros de audio espaciales, la velocidad 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 velocidad de bits comparativamente alta en beneficio de una buena calidad perceptual. Puesto que para la segunda porción de la señal de audio multicanal se elige una regla de derivación de parámetros que da como resultado parámetros que consumen menos velocidad de bits, la velocidad de bits total resultante puede disminuirse aumentando el tamaño de la segunda porción de la señal multicanal. En una realización preferida de la presente invención, no se transmiten datos paramétricos en absoluto para la segunda porción de la señal multicanal, que es, por supuesto, la mayor parte del ahorro de bits. Por tanto, desplazando dinámicamente el tamaño de la primera porción con respecto al tamaño de la segunda porción, la velocidad de bits (o la calidad perceptual) puede ajustarse dinámicamente a las necesidades.

Se deriva una señal de mezcla descendente de un modo compatible con la matriz. Por 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 una mayor calidad. Al mismo tiempo, se disminuye la velocidad de bits global basándose en una reproducción basada en matriz para partes de la señal menos vitales para la calidad de una señal reproducida.

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

Preprocesamiento de las señales de entrada (señales de mezcla descendente recibidas en el dominio de QMF) mediante la aplicación de una matriz predecorrelacionadora;

decorrelación de parte de las señales preprocesadas; y mezclado de las señales así derivadas (señales preprocesadas y señales decorrelacionadas) dentro de una matriz de mezcla, siendo la salida del mezclado los canales de la señal de mezcla ascendente.

Tanto la matriz predecorrelacionadora así como la matriz mixta son matrices bidimensionales con las dimensiones "número de franjas de tiempo" por un lado y el "número de bandas de parámetros" por otro lado. Dentro de un proceso de decodificación, los elementos de estas matrices se rellenan con valores que se derivan de los parámetros leídos del flujo de bits espacial, es decir, mediante los primeros datos paramétricos. Cuando los primeros datos paramétricos se reciben solo para una primera porción de la señal multicanal, solo esa porción de una reconstrucción de una señal multicanal puede derivarse usando los primeros datos paramétricos enviados. Los elementos de matriz para derivar la segunda parte de la reconstrucción de la señal multicanal se derivan, según la presente invención, usando esquemas de codificación compatibles con matriz. Estos elementos de matriz, por tanto, o bien pueden derivarse basándose solo en el conocimiento logrado a partir de la señal de mezcla descendente o bien pueden reemplazarse por valores predefinidos.

En una realización preferida, un decodificador de audio multicanal según la presente invención reconoce mediante la cantidad de los primeros datos paramétricos transmitidos qué parte de la matriz o qué parte de la señal de audio multicanal va a procesarse mediante la regla dependiendo de los parámetros espaciales y qué parte va a procesarse mediante la solución basada en matriz.

En otro ejemplo, un codificador de audio crea información de ventana, que indica qué partes de una señal multicanal está procesando la solución basada en matriz o el enfoque compatible con audio espacial. La información de ventana se incluye en la representación paramétrica de una señal multicanal.

Un decodificador de la invención, por 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.

El concepto de la invención se aplica en el dominio de QMF durante el procesamiento de señales, es decir, en un dominio donde las señales están representadas por múltiples representaciones, conteniendo cada representación información sobre una cierta banda de frecuencia.

5 En una realización adicional preferida de la presente invención, el método libre de información secundaria (enfoque basado en matriz) se aplica solo a las partes de frecuencia más alta mientras que se aplica información paramétrica (explícita) (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 de la audición humana de notar pequeñas desviaciones de dos señales similares (por ejemplo, desviaciones de fase) mucho más fácilmente para frecuencias bajas que para frecuencias altas.

10 Un gran beneficio de la presente invención es que se consigue una compatibilidad hacia atrás de un esquema de codificación y decodificación de audio espacial con soluciones basadas en matriz sin tener que introducir hardware o software cuando se eligen apropiadamente las reglas de codificación y decodificación de los codificadores de audio espacial.

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 según la presente invención es además extremadamente flexible, ya que permite un ajuste sin interrupciones de la velocidad de bits o la calidad, es decir, una transición suave entre la codificación basada en matriz completa y la codificación de audio espacial completa de una señal dada. Es decir, el esquema de codificación aplicado puede ajustarse a las necesidades reales, o bien con respecto a la velocidad de bits requerida o bien con respecto a la calidad deseada.

## 25 **Breve descripción de los dibujos**

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

30 la figura 1 muestra un codificador;

la figura 2 muestra un ejemplo de un flujo de bits de parámetros creado por el concepto a modo de ejemplo;

35 la figura 2a muestra un transcodificador;

la figura 3 muestra un decodificador de la invención;

la figura 4 muestra un ejemplo de un decodificador de audio espacial que implementa el concepto de la invención;

40 la figura 5 ilustra el uso de los diferentes esquemas de codificación en un lado del 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 filtros; y

la figura 9 muestra un ejemplo adicional de un codificador de la técnica anterior.

## 50 **Descripción detallada de las realizaciones preferidas**

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

55 Se introduce una señal de audio multicanal 106 en el codificador 100, donde 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 donde, para una segunda porción 110 de la señal multicanal 106, se incluyen segundos datos paramétricos que requieren menos información que los primeros datos paramétricos o no se incluyen datos paramétricos 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 que dan como resultado diferentes conjuntos de parámetros que luego se transfieren a la interfaz de salida 104 que combina los diferentes conjuntos de parámetros en la representación paramétrica 112. Un caso especial y preferido es que, para la segunda porción 110, no se incluyan parámetros en la representación paramétrica (y por tanto no se deriven por el generador de parámetros 102) puesto que, en un lado del decodificador, el decodificador deriva los parámetros de decodificación requeridos mediante algunas reglas heurísticas.

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

Además, la interfaz de salida 104 podría añadir una señal de ventana adicional a la representación paramétrica 112 que señalará 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 porciones para decidir qué parte de la señal multicanal 106 se usa como primera porción 108 y qué parte se usa como segunda porción 110, basándose la decisión en un criterio de calidad.

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

Una ventaja principal es que la velocidad de bits consumida por la representación paramétrica puede por tanto variarse en el tiempo, garantizando que se cumple el criterio de calidad en cualquier momento durante la codificación mientras que se permite una reducción global de la velocidad de bits requerida en comparación con los métodos de la técnica anterior.

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

Tal como se mencionó anteriormente, el procesamiento de las señales de audio se realiza por bloques, es decir, se procesa en una etapa un número de muestras posteriores de la señal multicanal en el dominio de tiempo, construyendo una denominada trama. La figura 2 muestra un flujo de bits de parámetros, es decir, una representación paramétrica de dos tramas consecutivas. El flujo de bits de parámetros 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 tomó la decisión de que la primera porción 108, que está representada por datos paramétricos, tiene que ser grande en comparación con la segunda porción, lo que puede ser el caso, por ejemplo, si la escena de audio a codificar es bastante compleja. El flujo de bits de parámetros 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. Tal como puede observarse en la figura 2, se incluyen 28 parámetros espaciales ICC e ICLD 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 de QMF de la señal multicanal.

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

La figura 2a muestra un transcodificador 150. El transcodificador recibe como entrada un flujo 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 un flujo de bits 158 que contiene 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 disminuye solo ligeramente. El número de parámetros 162 que describen la segunda trama correspondiente a los parámetros de entrada 156 disminuye enormemente, lo que reduce la cantidad de velocidad de bits que necesita la representación paramétrica resultante significativamente. Tal transcodificador 150 puede usarse por tanto para el procesamiento posterior de un flujo de bits ya existente de parámetros espaciales para derivar una representación paramétrica de la invención que requiere menos velocidad de bits durante la transmisión o menos espacio de almacenamiento cuando se almacena en un medio legible por ordenador. Debe indicarse en este caso que, por supuesto, es también posible implementar un transcodificador para transcodificar en la otra dirección, es decir, usando la representación paramétrica para generar parámetros espaciales.

El transcodificador 150 puede implementarse de diversos modos 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 la velocidad de bits posible sin alterar la calidad perceptual más allá de un límite aceptable.

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

El procesador está recibiendo como 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 que 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 donde 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 donde la segunda regla de derivación usa los segundos datos paramétricos o ningún dato paramétrico y en donde la primera regla de derivación depende de los primeros datos paramétricos 206.

La señal intermedia 210 derivada por el procesador 202 se construye 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 van a 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, el procesador 202 puede señalar las fracciones de la porción de alta calidad 212 y de la segunda porción 214 dentro de la señal de mezcla descendente 204 mediante alguna información de ventana adicional que se deriva de 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 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 de la invención de compatibilidad de matriz en un decodificador de audio espacial. El decodificador de audio multicanal 600 comprende un predecorrelacionador 601, un decorrelacionador 602 y una matriz de mezcla 603.

El decodificador de audio multicanal 600 es un dispositivo flexible que permite funcionar en diferentes modos dependiendo de la configuración de señales de entrada 605 introducidas en el predecorrelacionador 601. Generalmente, el predecorrelacionador 601 deriva señales intermedias 607 que sirven como entrada para el decorrelacionador 602 y que se transmiten parcialmente sin alteraciones 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, se suministra una señal de mezcla descendente y una señal residual opcional al predecorrelacionador 601, que deriva cuatro señales intermedias (e1 a e4) que se usan como 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 donde se suministra una señal residual adicional como entrada, el decorrelacionador 602 que está funcionando generalmente en un dominio de sub-banda, puede ser operativo para simplemente reenviar la señal residual en lugar de derivar una señal decorrelacionada. Esto puede realizarse de una manera selectiva de frecuencia para ciertas bandas de frecuencia solo.

En la configuración de 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 predecorrelacionadora 601 deriva un canal izquierdo, uno derecho y uno central y, además, dos canales intermedios (e1, e2). Por tanto, las señales de entrada a la matriz de mezcla 603 están formadas por el canal izquierdo, el canal derecho, el canal central y dos señales decorrelacionadas (d1 y d2).

En una modificación adicional, la matriz predecorrelacionadora puede derivar una señal intermedia adicional (e5) que se usa como 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 izquierdo y derecho.

El decodificador de audio de la invención 600 implementa el concepto de la invención en la configuración de 2 a 5. La representación paramétrica transmitida se usa en la matriz de predecorrelación 601 y en la matriz de mezcla 603. Allí,

puede implementarse el concepto de la invención de diferentes modos tal como se muestra en más detalle en la figura 5.

5 La figura 5 muestra el predecorrelacionador, implementado como matriz predecorrelacionadora 601 y matriz de mezcla 603 en un esquema de principios, en donde los otros componentes del decodificador de audio multicanal 600 se omiten.

10 La matriz usada para realizar la predecorrelación y el mezclado tiene columnas que representan franjas de tiempo, es decir, las muestras de tiempo individuales de una señal, y filas que representan las diferentes bandas de parámetros, es decir, cada fila está asociada con una banda de parámetros de una señal de audio.

15 Según el concepto de la presente invención, los elementos de matriz de las matrices 601 y 603 se derivan solo parcialmente de los datos paramétricos transmitidos, en donde los elementos de matriz restantes los deriva el decodificador, basándose en, por ejemplo, el conocimiento de la señal de mezcla descendente. La figura 5 muestra un ejemplo donde, por debajo de una línea límite de frecuencia dada 622, los elementos de la matriz predecorrelacionadora 601 y la matriz de mezcla 603 se derivan de parámetros 620 que se leen a partir del flujo de bits, es decir, basándose en información transmitida desde el decodificador. 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 solo.

20 La frecuencia límite (o en general: la cantidad de elementos de matriz derivados de datos transmitidos) puede adaptarse libremente según las restricciones de calidad y/o velocidad de bits que tienen que cumplirse para el escenario de aplicación particular.

25 Se prefiere para el método de codificación novedoso explicado de manera resumida en el presente documento que pueda realizarse un proceso de mezcla ascendente libre de información secundaria con la misma estructura que se ha explicado de manera resumida en el Modelo de Referencia de Codificación de Audio Espacial MPEG 0. Esta invención puede consistir en describir un método para la mezcla ascendente libre de información secundaria, pero preferiblemente proporciona un método para una combinación sin interrupciones y ventajosa de tales conceptos con métodos para la 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 preferiblemente de datos transmitidos en un flujo de bits, sino por diferentes medios sin la ayuda de información secundaria, por ejemplo aplicando reglas heurísticas basadas solo en el conocimiento logrado a partir de la señal de mezcla descendente.

40 De este modo, es posible lograr un escalado gradual entre ambas técnicas, en cuanto a tasa de bits así como en cuanto a calidad de sonido, adquiriendo solo partes de las matrices basándose en los parámetros transmitidos y aplicando las reglas del método sin información secundaria para rellenar las partes restantes. En términos conceptuales, esto corresponde a transmitir para ciertas partes de las matrices los parámetros espaciales y para otras partes generarlos en el decodificador.

45 La determinación de las partes de matrices que van a derivarse del uno u otro método puede realizarse de muchos modos diferentes, tales como

- derivar las partes de las matrices por debajo de una línea límite horizontal dada mediante un método y por encima de esta línea límite mediante el otro método;
- 50 • derivar las partes de las matrices a la izquierda de una línea límite vertical dada mediante un método y a la derecha de esta línea límite mediante el otro método;
- determinar mosaicos de tiempo-frecuencia arbitrarios dentro de ambas matrices cuyos elementos se derivan mediante un método y derivar los elementos de los mosaicos de tiempo-frecuencia restantes por medio del otro método.

55 Se ha detallado en los párrafos anteriores que es ventajoso describir todas las partes de frecuencia de una señal multicanal hasta una cierta frecuencia límite mediante parámetros espaciales, mientras que las partes de frecuencia restantes de la señal multicanal no están representadas por 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 supuesto, la presente invención no se limita de ninguna manera 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. Este puede ser especialmente el caso cuando, en la región de frecuencia más baja, sólo está contenida poca energía en la señal, ya que la mayor parte de la energía está contenida

en un dominio de alta frecuencia de la señal de audio. Debido a los efectos de enmascaramiento, la parte de baja frecuencia estará dominada en su mayor parte por las partes de alta frecuencia y entonces puede ser ventajoso proporcionar la posibilidad de una reproducción de alta calidad de la parte de alta frecuencia de la señal.

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

## REIVINDICACIONES

1. Decodificador de audio multicanal (200) para procesar una señal de audio de mezcla descendente (204) derivada de un modo compatible con matriz y para procesar primeros datos paramétricos (206) que describen una primera porción de una señal multicanal, en donde, para una segunda porción de la señal multicanal, no se procesan datos paramétricos o se procesan 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) 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,
  - usar 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 donde la primera regla de derivación depende de los primeros datos paramétricos (206) y comprende un primer conjunto de elementos de matriz de una matriz predecorrelacionadora (M1) o una matriz de mezcla (M2); y
  - usar una segunda regla de derivación para derivar una segunda porción de la señal intermedia (210), no usando la segunda regla de derivación datos paramétricos o usando los segundos datos paramétricos (208) y comprendiendo un segundo conjunto de elementos de matriz de la matriz predecorrelacionadora (M1) o la matriz de mezcla (M2),
  - en donde la primera porción es una porción de tiempo o una porción de frecuencia.
2. Decodificador de audio multicanal según la reivindicación 1, en donde el procesador (202) está configurado para recibir, como 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) según 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 donde solo está contenida información sobre una banda de frecuencia predeterminada de la señal de audio multicanal dentro de la descripción.
4. Decodificador de audio multicanal (200) según 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 entre un límite de baja frecuencia y un límite de alta frecuencia.
5. Decodificador de audio multicanal (200) según 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) según 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 ventana.
7. Decodificador de audio multicanal (200) según la reivindicación 6, en el que la regla de estimación de ventana incluye contar la cantidad de datos paramétricos enviados 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) según la reivindicación 1, en el que el procesador (202) es operativo para calcular la segunda regla de derivación a partir de la información sobre la señal de audio de mezcla descendente (204) o a partir de los segundos datos paramétricos (208).
9. Decodificador de audio multicanal (200) según la reivindicación 1, en el que el procesador (202) es operativo para usar una regla predefinida como segunda regla de derivación.
10. Decodificador de audio multicanal (200) según 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.
11. Decodificador de audio multicanal (200) según 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 mezclado ascendente para calcular una primera porción de la reconstrucción de la señal multicanal correspondiente a la primera porción de la señal multicanal, en donde la primera regla de mezcla ascendente depende de los primeros datos paramétricos (206); y
- 10   12.       usar una segunda regla de mezclado ascendente para calcular una segunda porción de la reconstrucción de la señal multicanal, no usando la segunda regla de mezclado ascendente de datos paramétricos o usando los segundos datos paramétricos (208).
- 15           Método para procesar una señal de audio de mezcla descendente (204) derivada de un modo compatible con matriz y para procesar primeros datos paramétricos (206) que describen una primera porción de una señal multicanal, en donde, para una segunda porción de la señal multicanal, no se procesan datos paramétricos o se procesan 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) 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) y que comprende un primer conjunto de elementos de matriz de una matriz predecorrelacionadora (M1) o una matriz de mezcla (M2), correspondiendo la primera porción de la señal intermedia (210) a la primera porción de la señal de audio multicanal, estando la señal de audio de mezcla descendente (204) en una representación de QMF, banco de filtros 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 y comprendiendo un segundo conjunto de elementos de matriz de la matriz predecorrelacionadora (M1) o la matriz de mezcla (M2),
- 30           en donde la primera porción es una porción de tiempo o una porción de frecuencia.
13.       Programa informático que tiene un código de programa para realizar, cuando se ejecuta en un ordenador, un método según la reivindicación 12.

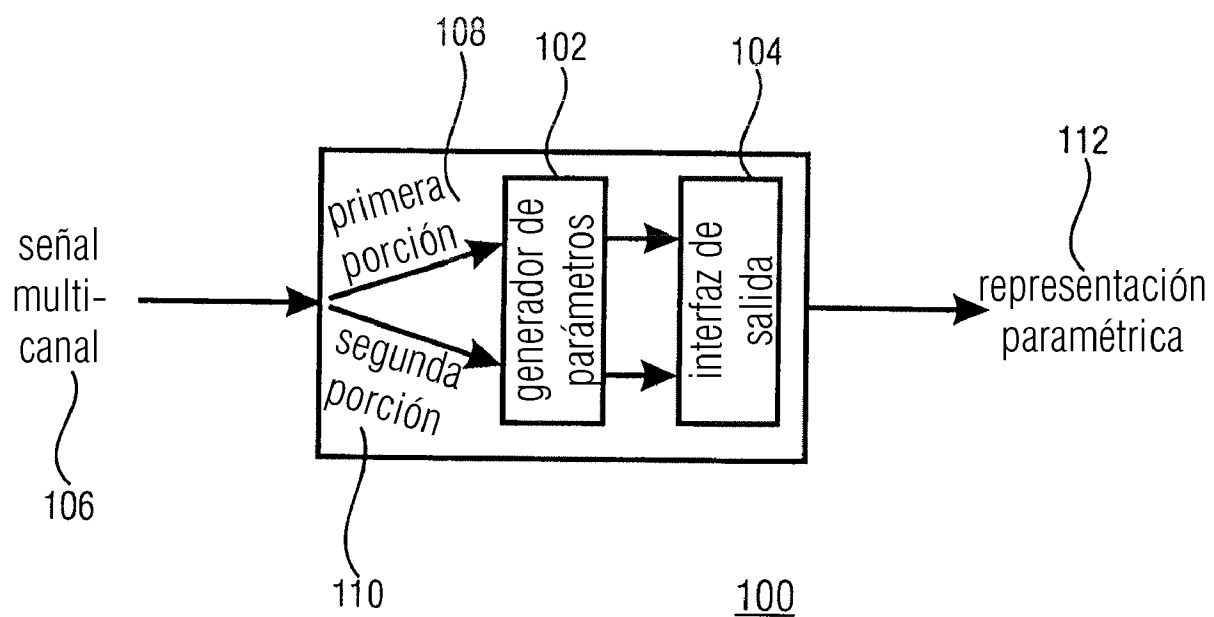


FIGURA 1

## Flujo de bits de parámetros

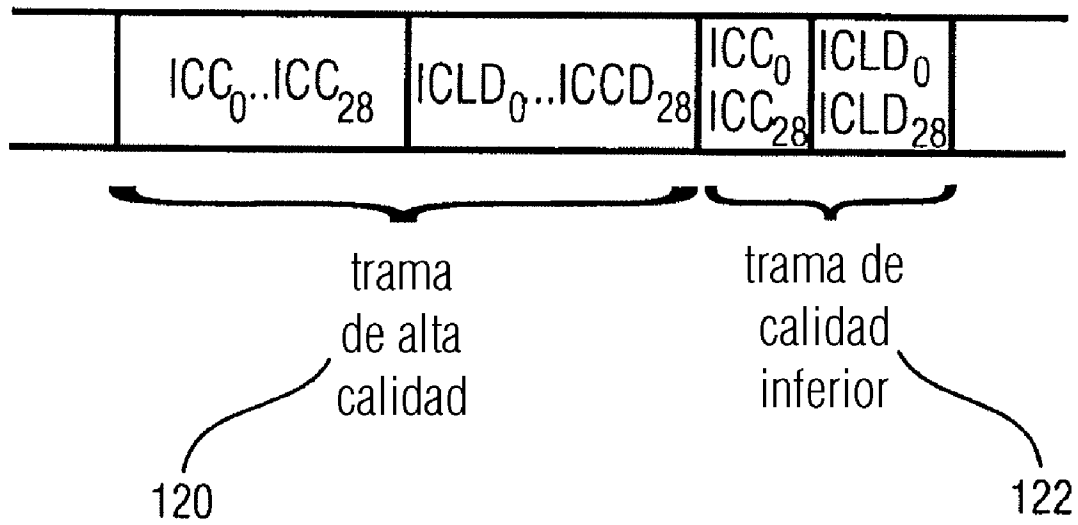


FIGURA 2

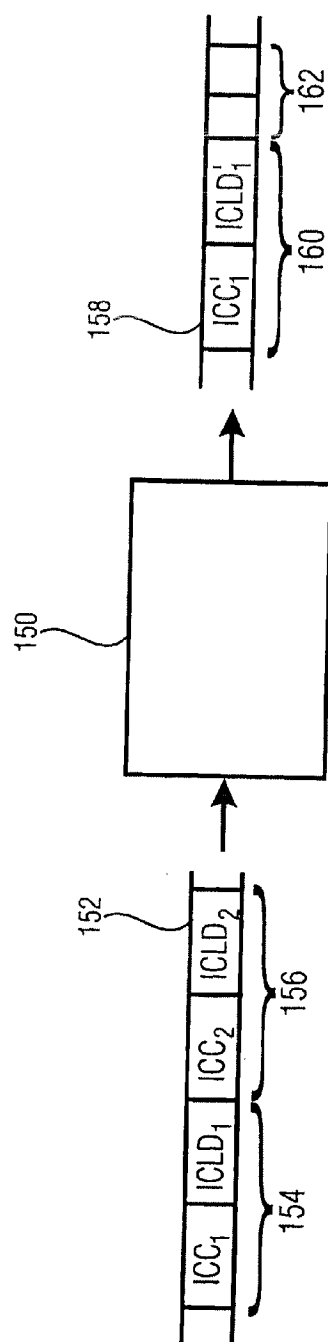


FIGURA 2A

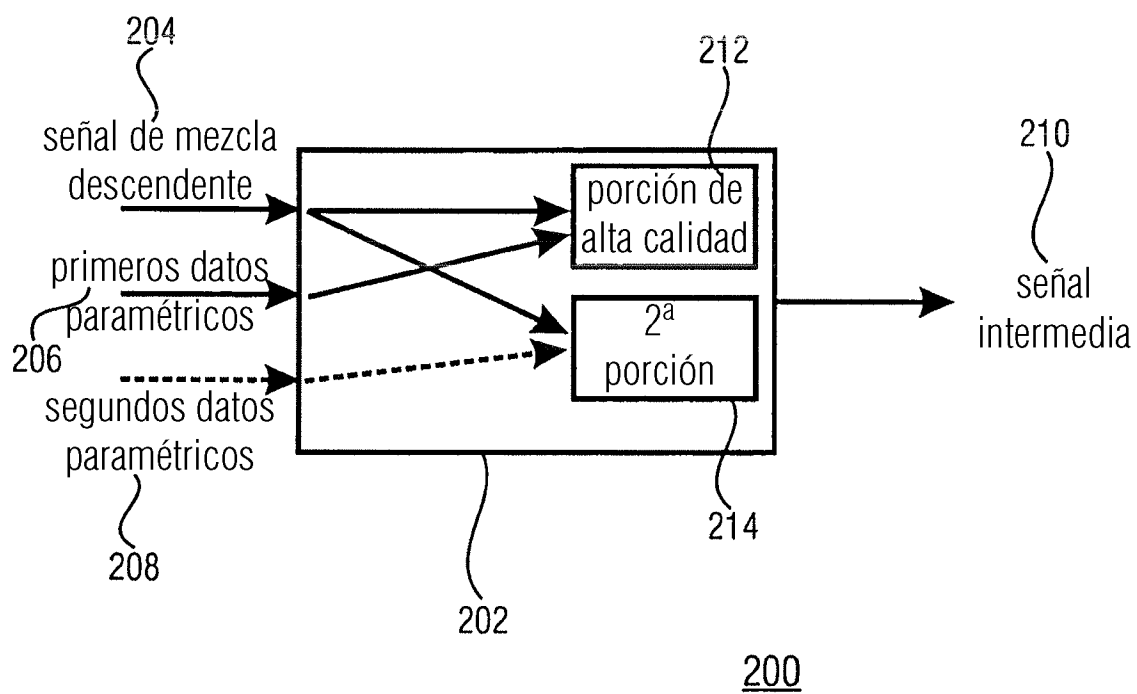


FIGURA 3

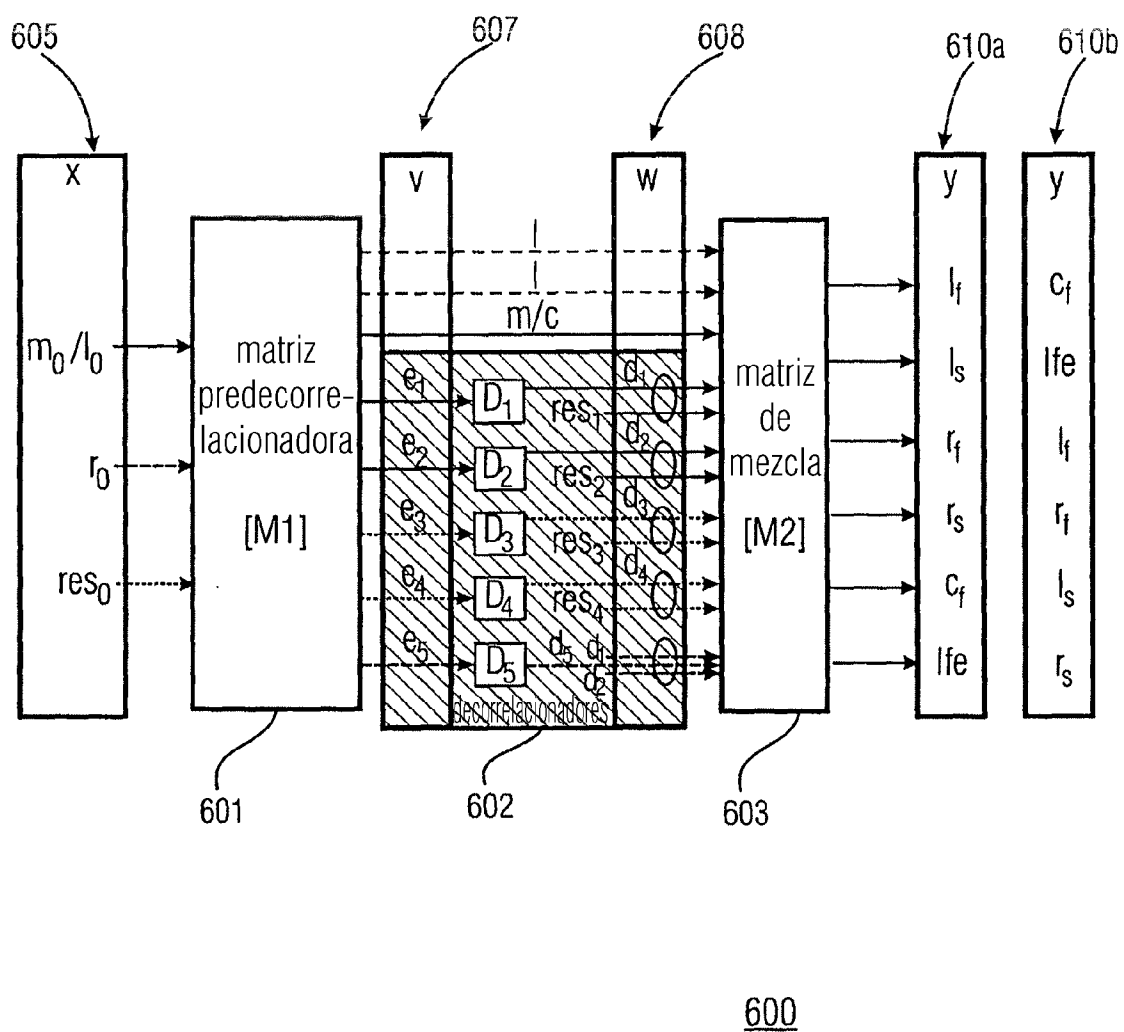


FIGURA 4

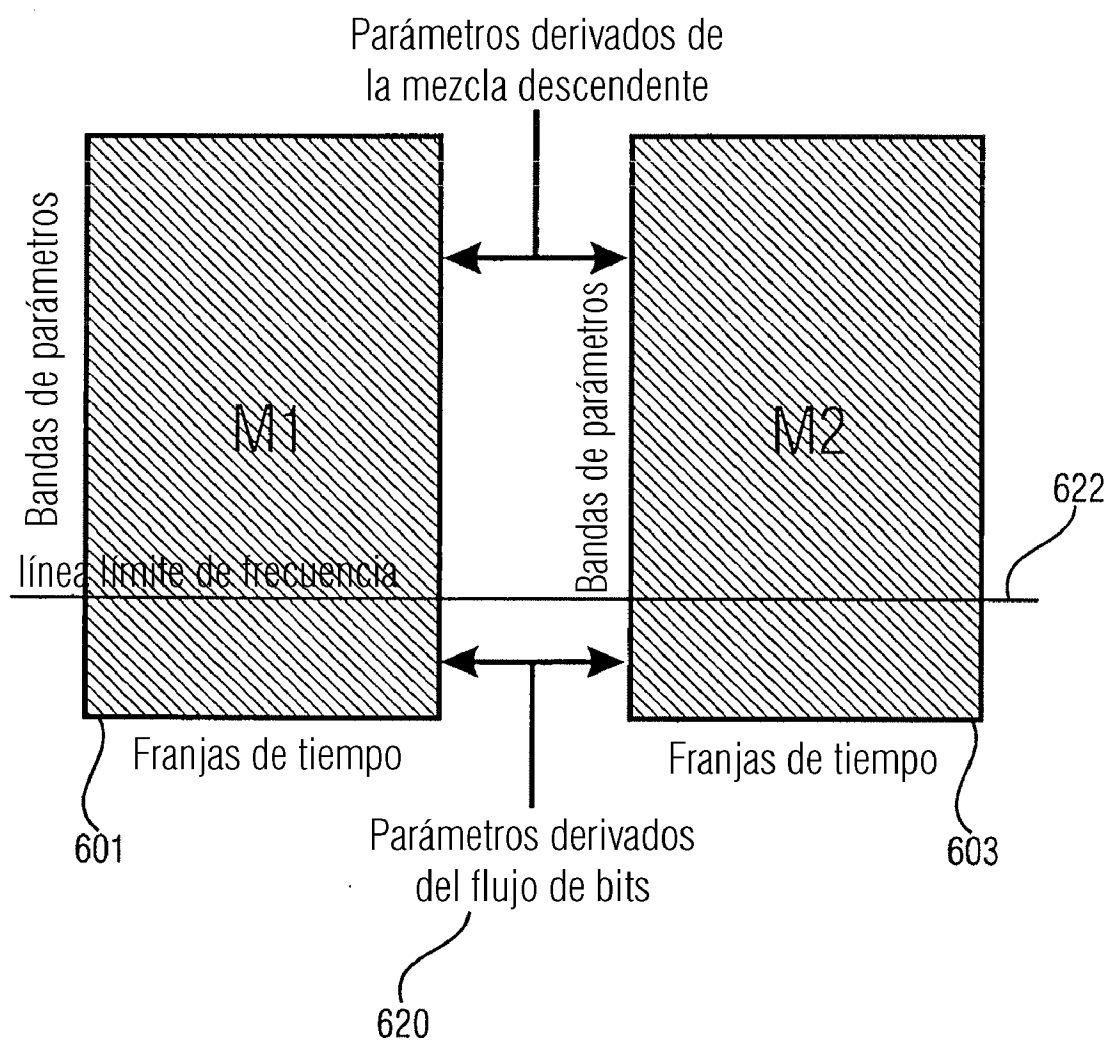


FIGURA 5

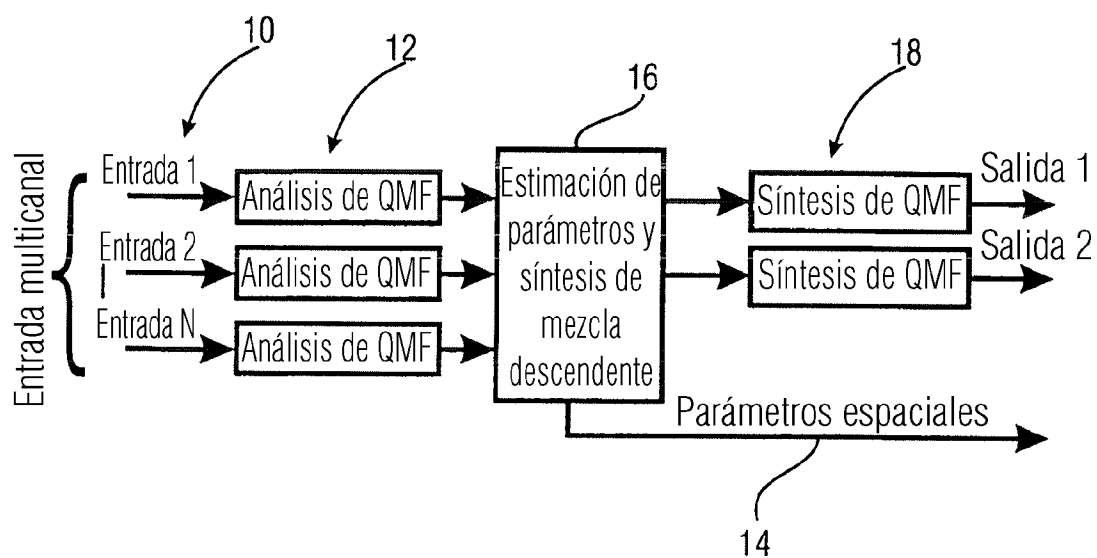


FIGURA 6

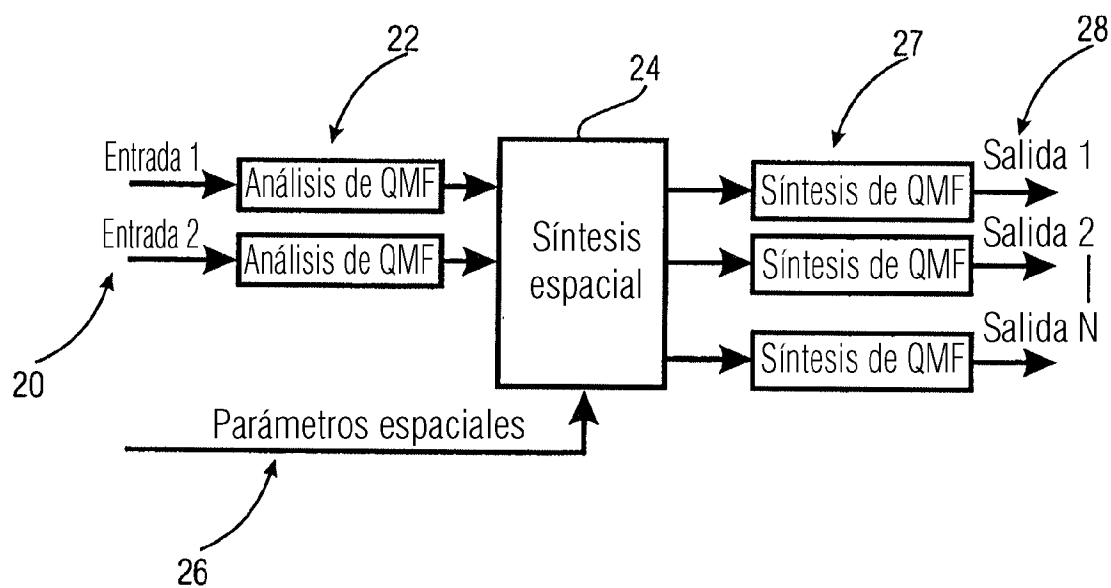


FIGURA 7

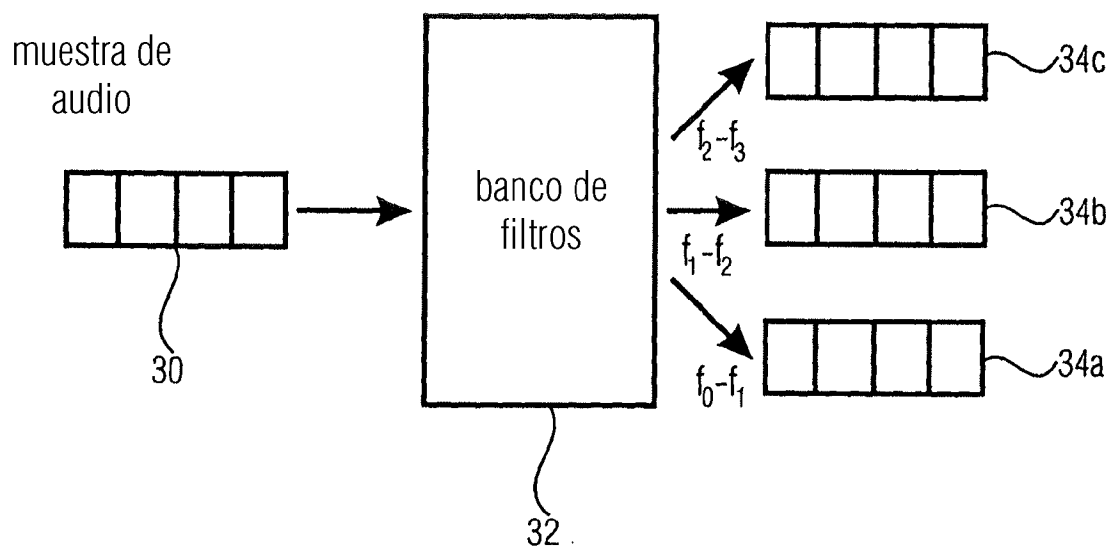


FIGURA 8

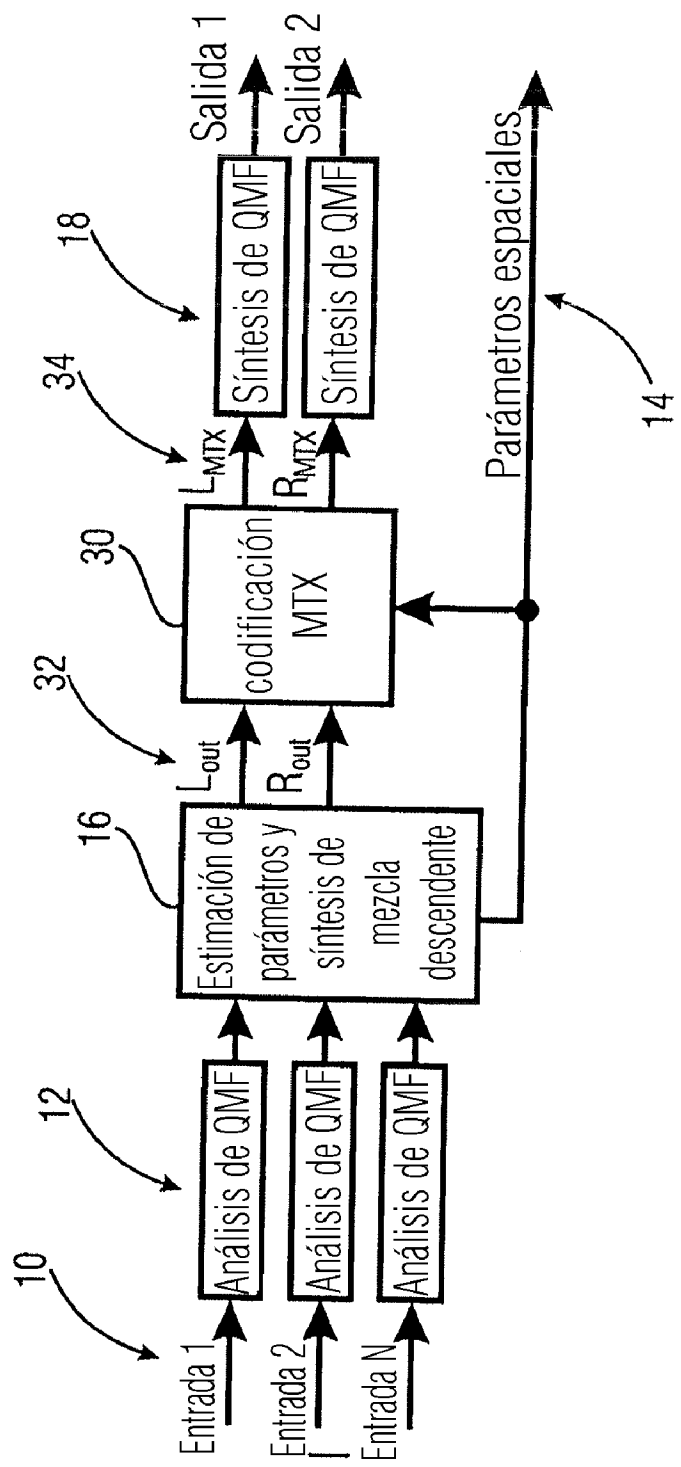


FIGURA 9