(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 717 606**

(51) Int. CI.:

G10L 19/24 (2013.01)
G10L 19/02 (2013.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **16.06.2006 PCT/IB2006/003986**

(87) Fecha y número de publicación internacional: **05.07.2007 WO07074401**

(96) Fecha de presentación y número de la solicitud europea: **16.06.2006 E 06848793 (3)**

(97) Fecha y número de publicación de la concesión europea: **09.01.2019 EP 1891740**

(54) Título: **Codificación y decodificación de audio escalable usando un banco de filtros jerárquico**

(30) Prioridad:

17.06.2005 US 691558 P
12.06.2006 US 452001

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
24.06.2019

(73) Titular/es:

DTS(BVI) LIMITED (100.0%)
Beaufort House 438
Road Town Tortola, VG

(72) Inventor/es:

SHMUNK, DMITRY V. y
BEATON, RICHARD J.

(74) Agente/Representante:

CARPINTERO LÓPEZ, Mario

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

DESCRIPCIÓN

Codificación y decodificación de audio escalable usando un banco de filtros jerárquico

Antecedentes de la invención

Campo de la invención

La presente invención se refiere a la codificación escalable de una señal de audio y más específicamente a procedimientos de realización de este escalado de tasa de datos de una manera eficiente para señales de audio multicanal que incluyen filtrado jerárquico, codificación conjunta de componentes tonales y codificación conjunta de canal de componentes de dominio tiempo en la señal residual.

Descripción de la técnica relacionada

El principal objetivo de un algoritmo de compresión de audio es crear una representación sónicamente aceptable de una señal de audio de entrada usando tan pocos bits digitales como sea posible. Esto permite que una versión de tasa de datos baja de la señal de audio de entrada se distribuya a través de canales de transmisión de ancho de banda limitado, tal como internet, y reduce la cantidad de almacenamiento necesario para almacenar la señal de audio de entrada para reproducción futura. Para esas aplicaciones en las que la capacidad de datos del canal de transmisión es fija, y no varía con el paso del tiempo, o la cantidad, en términos de minutos, de audio que necesita almacenarse se conoce por adelantado y no aumenta, procedimientos de compresión de audio tradicionales fijan la tasa de datos y por lo tanto el nivel de calidad de audio en el momento de codificación de compresión. Ninguna reducción adicional en tasa de datos puede efectuarse sin o bien grabar la señal original a una tasa de datos menor o bien descomprimir la señal de audio comprimida y a continuación comprimir de nuevo esta señal descomprimida en una tasa de datos menor. Estos procedimientos no son "escalables" para abordar problemas de capacidad de canal variable, almacenamiento de contenido adicional en una memoria fija o extracción de flujos de datos a tasas de datos variables para diferentes aplicaciones.

Una técnica usada para crear un flujo de bits con características escalables, y evitar las limitaciones anteriormente descritas, codifica la señal de audio de entrada como un flujo de datos de tasa de datos alta compuesta de subconjuntos de flujo de datos de tasa de datos baja. Estos flujos de datos de tasa de datos baja codificados pueden extraerse de la señal codificada y combinarse para proporcionar un flujo de datos de salida cuya tasa de datos es ajustable en un amplio intervalo de tasas de datos. Un enfoque para implementar este concepto es decodificar primero datos a tasa de datos más baja soportada, a continuación codificar un error entre la señal original y una versión decodificada de este flujo de datos de la menor tasa de datos. Este error codificado se almacena y también combina con el flujo de datos de la menor tasa de datos soportada para crear un flujo de datos de la segunda menor tasa de datos. El error entre la señal original y una versión decodificada de esta señal de la segunda menor tasa de datos se codifica, almacena y añade al flujo de datos de la segunda menor tasa de datos para formar un flujo de datos de la tercera menor tasa de datos y así sucesivamente. Este procedimiento se repite hasta que la suma de las tasas de datos asociadas con flujos de datos de cada una de las señales de error así obtenidas y la tasa de datos del flujo de datos de la menor tasa de datos soportada es igual al flujo de datos de la mayor tasa de datos a soportar. El flujo de datos de tasa de datos alta escalable final se compone del flujo de datos de la menor tasa de datos y cada uno de los flujos de datos de error codificado.

Una segunda técnica, normalmente usada para soportar un número pequeño de diferentes tasas de datos entre menores y mayores tasas de datos ampliamente espaciadas, emplea el uso de más de un algoritmo de compresión para crear un flujo de bits escalable "por capas". El aparato que realiza la operación de cambio de escala en un flujo de bits codificado de esta manera elige, dependiendo de requisitos de tasa de datos de salida, cuál de los múltiples flujos de datos transportados en el flujo de datos por capas usar como la salida de audio codificada. Para mejorar la eficiencia de codificación y proporcionar un intervalo más amplio de tasas de datos escaldas, datos transportados en los flujos de datos de menor tasa pueden usarse por flujos de datos de mayor tasa para formar flujos de datos de mayor calidad y mayor tasa adicionales.

El documento DAUDET L ET AL: "Hybrid representations for audiophonic signal encoding", SIGNAL PROCESSING, ELSEVIER SCIENCE PUBLISHERS, vol. 82, n.º 11, 1 de noviembre 2002, páginas 1595-1617, XP004381254, ISSN: 0165-1684, DOI: 10.1016/S0165-1684(02)00304-3, desvela un procedimiento en el que componentes transitorios, tonales y estocásticos de una señal de audio se estiman y codifican usando una estrategia que incluye codificación por transformación asociada con espacios de representación excesivos. Entre espacios de representación, se usan una unión de coseno local y ondículas diádicas mostrando buenas propiedades de separación para características estructurales de señales de audio, su parte tonal y parte transitoria, respectivamente. La separación de estas dos capas se mejora mediante el uso de representaciones estructuradas. El enfoque descrito no se basa en ninguna segmentación anterior de la señal de audio.

El documento US 2002/0176353 A1 describe un procedimiento y sistema de codificación y decodificación de una señal de entrada en relación con los aspectos perceptualmente más relevantes de la señal de entrada. Más particularmente, se aplica una transformación bidimensional a la señal de entrada para producir una matriz de magnitud y una matriz de fase que puede cuantificarse inversamente mediante un decodificador. Una primera

columna de coeficientes de la matriz de magnitud representa una función de densidad espectral media de la señal de entrada. Aspectos relevantes de la función de densidad espectral media se codifican en un comienzo de un paquete de datos para un uso posterior por un decodificador para recrear la señal de entrada, a base de una codificación de las matrices de magnitud y fase con el resto del paquete de datos.

Sumario de la invención

La invención proporciona un procedimiento de reconstrucción de una señal de audio de salida de dominio de tiempo a partir de un flujo de datos codificado con las características de la reivindicación 1, un decodificador de reconstrucción de una señal de audio de salida de dominio de tiempo a partir de un flujo de datos codificado con las características de la reivindicación 9, un procedimiento de codificación de una señal de audio de entrada con las características de la reivindicación 11 y un codificador de flujo de bits escalable de codificación de una señal de audio de entrada y formación de un flujo de bits escalable con las características de la reivindicación 31. En este contexto, se observa que la invención se expone en las reivindicaciones independientes anteriormente mencionadas, y que todas las ocurrencias siguientes de la palabra "realización o realizaciones", si se refieren a combinaciones de características diferentes de las definidas por las reivindicaciones independientes, se refieren a ejemplos que se presentaron originalmente pero que no representan realizaciones de la invención reivindicada actualmente, en el que estos ejemplos se muestran aún únicamente para fines de ilustración.

Por consiguiente, la presente invención proporciona un procedimiento de codificación de señales de entrada de audio para formar un flujo de bits maestro que puede escalarse para formar un flujo de bits escalado que tiene una tasa de datos prescita arbitrariamente y de decodificación del flujo de bits escalado para reconstruir las señales de audio.

Esto se logra generalmente comprimiendo las señales de entrada de audio y disponiendo las mismas para formar un flujo de bits maestro. El flujo de bits maestro incluye componentes cuantificados que se clasifican sobre la base de su contribución relativa a calidad de señal decodificada. La señal de entrada se comprime adecuadamente separando la misma en una pluralidad de componentes tonales y residuales, y clasificando y a continuación cuantificando los componentes. La separación se realiza adecuadamente usando un banco de filtros jerárquico. Los componentes se clasifican y cuantifican adecuadamente con referencia a la misma función de enmascaramiento o diferente criterio psicoacústico. Los componentes pueden ordenarse a continuación a base de su clasificación para facilitar escalado eficiente. El flujo de bits maestro se escala eliminando un número suficiente de los componentes de baja clasificación para formar el flujo de bits escalado que tiene una tasa de datos escalada menor que o aproximadamente igual a una tasa de datos deseada. El flujo de bits escalado incluye información que indica la posición de los componentes en el espectro de frecuencia. Un flujo de bits escalado se decodifica adecuadamente usando un banco de filtros jerárquico inverso disponiendo los componentes cuantificados a base de la formación de posición, ignorando los componentes faltantes y decodificando los componentes dispuestos para producir un flujo de datos de salida.

En una realización, el codificador usa un banco de filtros jerárquico para descomponer la señal de entrada en una representación de tiempo/frecuencia de múltiples resoluciones. El codificador extrae componentes tonales en cada iteración del HFB a diferentes resoluciones de frecuencia, elimina esos componentes tonales de la señal de entrada para pasar una señal residual a la siguiente iteración del HFB y a continuación extrae componentes residuales de la señal residual final. Los componentes tonales se agrupan en al menos un subdominio de frecuencia por resolución de frecuencia y clasifican de acuerdo con su importancia psicoacústica a la calidad de la señal codificada. Los componentes residuales incluyen componentes de muestra de tiempo (por ejemplo una cuadrícula G) y componentes de factor de escala (por ejemplo cuadrículas G0, G1) que modifican los componentes de muestra de tiempo. Los componentes de muestra de tiempo se agrupan en al menos un subdominio de muestra de tiempo y clasifican de acuerdo con su contribución a la calidad de la señal decodificada.

En el decodificador, el banco de filtros jerárquico inverso puede usarse para extraer tanto los componentes tonales como los componentes residuales dentro de una estructura de banco de filtros eficiente. Todos los componentes se cuantifican inversamente y la señal residual se reconstruye aplicando los factores de escala a las muestras de tiempo. Las muestras de frecuencia se reconstruyen y añaden a las muestras de tiempo reconstruidas para producir la señal de audio de salida. Obsérvese que el banco de filtros jerárquico inverso puede usarse en el decodificador independientemente de si el banco de filtros jerárquico se usó durante el procedimiento de codificación.

En una realización ilustrativa, los componentes tonales seleccionados en una señal de audio multicanal se codifican usando codificación diferencial. Para cada componente tonal, se selecciona un canal como el canal primario. El número de canal del canal primario y su amplitud y fase se almacenan en el flujo de bits. Una máscara de bits se almacena que indica cuál de los otros canales incluyen el componente tonal indicado y, por lo tanto, deberían codificarse como canales secundarios. La diferencia entre las amplitudes primarias y secundarias y fases se codifican por entropía y almacenan para cada canal secundario en el que está presente el componente tonal.

En una realización ilustrativa, los componentes de muestra de tiempo y de factor de escala que componen la señal residual se codifican usando codificación conjunta de canal (JCC) extendida a audio multicanal. Un procedimiento de agrupamiento de canal primero determina cuál de los múltiples canales puede codificarse conjuntamente y todos los

canales se forman en grupos con el último grupo estando posiblemente incompleto.

Objetos, características y ventajas adicionales de la presente invención se incluyen en la siguiente descripción de realizaciones ilustrativas, cuya descripción debería leerse con los dibujos adjuntos. Aunque estas realizaciones ilustrativas pertenecen a datos de audio, se entenderá que video, multimedia y otros tipos de datos también pueden procesarse de maneras similares.

Breve descripción de los dibujos

La Figura 1 es una ilustración de diagrama de bloques de un codificador de flujo de bits escalable usando una topología de codificación residual de acuerdo con la presente invención;
Las Figuras 2a y 2b son representaciones de dominio de frecuencia y tiempo de una ventana de Shmunk para uso con el banco de filtros jerárquico;
La Figura 3 es una ilustración de un banco de filtros jerárquico para la provisión de una representación de tiempo/frecuencia de múltiples resoluciones de una señal de entrada desde la que pueden extraerse tanto componentes tonales como residuales con la presente invención;
La Figura 4 es un diagrama de flujo de las etapas asociadas con el banco de filtros jerárquico;
Las Figuras 5a a 5c ilustran una ventaja de formación de ventana de 'solapar-añadir';
La Figura 6 es un gráfico de la respuesta de frecuencia de banco de filtros jerárquico;
La Figura 7 es un diagrama de bloques de una implementación ilustrativa de un banco de filtros de análisis jerárquico para su uso en el codificador;
Las Figuras 8a y 8b son un diagrama de bloques simplificado de un banco de filtros jerárquico de 3 etapas y un diagrama de bloques más detallado de una única etapa;
La Figura 9 es una máscara de bits para la extensión de codificación diferencial de componentes tonales a audio multicanal;
La Figura 10 representa la realización detallada del codificador residual usado en una realización del codificador de la presente invención;
La Figura 11 es un diagrama de bloques para la codificación conjunta de canal para audio multicanal;
La Figura 12 representa esquemáticamente una trama escalable de datos producidos por el codificador de flujo de bits escalable de la presente invención;
La Figura 13 muestra el diagrama de bloques detallado de una implementación del decodificador usado en la presente invención;
La Figura 14 es una ilustración de un banco de filtros jerárquico inverso para la reconstrucción de datos de serie de tiempo a partir de tanto componentes de muestra de tiempo como frecuencia de acuerdo con la presente invención;
La Figura 15 es un diagrama de bloques de una implementación ilustrativa de un banco de filtros jerárquico inverso;
La Figura 16 es un diagrama de bloques de la combinación de componentes tonales y residuales usando un banco de filtros jerárquico inverso en el decodificador;
Las Figuras17a y 17b son un diagrama de bloques simplificado de un banco de filtros jerárquico inverso de 3 etapas y un diagrama de bloques más detallado de una única etapa;
La Figura 18 es un diagrama de bloques detallado del decodificador residual;
La Figura 19 es una tabla de correlación de G1;
La Figura 20 es una tabla de coeficientes de corrección de síntesis de función base; y
Las Figuras 21 y 22 son diagramas de bloque funcionales del codificador y decodificador, respectivamente, que ilustran una aplicación de la representación de tiempo/frecuencia de múltiples resoluciones del banco de filtros jerárquico en un codificador/decodificador de audio.

Descripción de realizaciones ilustrativas

La presente invención proporciona un procedimiento de compresión y codificación de señales de entrada de audio para formar un flujo de bits maestro que puede escalarse para formar un flujo de bits escalado que tiene una tasa de datos prescrita arbitrariamente y de decodificación del flujo de bits escalado para reconstruir las señales de audio. Un banco de filtros jerárquico (HFB) proporciona una representación de tiempo/frecuencia de múltiples resoluciones de la señal de entrada de la que el codificador puede extraer de forma eficiente tanto los componentes tonales como residuales. Para audio multicanal, se implementa codificación conjunta de componentes tonales y codificación conjunta de canal de componentes residuales en la señal residual. Los componentes se clasifican sobre la base de su contribución relativa a calidad de señal decodificada y cuantificada con referencia a una función de enmascaramiento. El flujo de bits maestro se escala eliminando un número suficiente de los componentes de baja clasificación para formar el flujo de bits escalado que tiene una tasa de datos escalada menor que o aproximadamente igual a una tasa de datos deseada. El flujo de bits escalado se decodifica adecuadamente usando un banco de filtros jerárquico inverso disponiendo los componentes cuantificados a base de posición información, ignorando los componentes faltantes y decodificando los componentes dispuestos para producir un flujo de datos de salida. En una posible aplicación, el flujo de bits maestro se almacena y a continuación se escala hacia abajo a una tasa de datos deseada para la grabación en otros medios o para la transmisión a través de un canal de banda limitada. En otra aplicación, en la que se almacenan múltiples flujos de bits escalados en medios, la tasa de datos de cada flujo se controla independiente y dinámicamente para maximizar la calidad percibida mientras se satisface una

restricción de tasa de datos agregada en todos los flujos de bits.

Como se usa en el presente documento los términos "Dominio", "subdominio" y "componente" describen la jerarquía de elementos escalables en el flujo de bits. Ejemplos incluirán:

Dominio	Subdominio	Componente
Tonal	Transformada de resolución de 1024 puntos (4 subtramas)	Componente tonal (fase/amplitud/posición)
Cuadrículas de factor de escala residual	Cuadrícula 1	Factor de escala dentro de cuadrícula 1
Subbandas residuales	Conjunto de todas las muestras de tiempo en subbanda 3	Cada muestra de tiempo en subbanda 3

Codificador de flujo de bits escalable con una topología de codificación residual

Como se muestra en la Figura 1, en una realización ilustrativa un codificador de flujo de bits escalable usa una topología de codificación residual para escalar el flujo de bits a una tasa de datos arbitraria eliminando selectivamente los componentes con la menor clasificación del núcleo (componentes tonales) y/o los componentes residuales (muestra de tiempo y factor de escala). El codificador usa un banco de filtros jerárquico para descomponer de forma eficiente la señal de entrada en una representación de tiempo/frecuencia de múltiples resoluciones de la que el codificador puede extraer de forma eficiente los componentes tonales y residuales. El banco de filtros jerárquico (HFB) descrito en el presente documento para la provisión de la representación de tiempo/frecuencia de múltiples resoluciones puede usarse en muchas otras aplicaciones en la que se desea una representación de este tipo de una señal de entrada. Una descripción general del banco de filtros jerárquico y su configuración para su uso en el codificador de audio se describen a continuación así como el HFB modificado usado por el codificador de audio particular.

La señal 100 de entrada se aplica tanto al calculador 101 de enmascaramiento como el extractor 102 de tonos de múltiples órdenes. El calculador 101 de enmascaramiento analiza la señal 100 de entrada e identifica un nivel de enmascaramiento como una función de frecuencia debajo de la cual frecuencias presentes en la señal 101 de entrada no son audibles para el oído humano. El extractor 102 de tonos de múltiples órdenes identifica frecuencias presentes en la señal 101 de entrada usando, por ejemplo, múltiples FFT superpuestas o como se muestra un banco de filtros jerárquico a base de MDCT, que cumplen el criterio psicoacústico que se han definido para tonos, selecciona tonos de acuerdo con este criterio, cuantifica la amplitud, frecuencia, fase y posición componentes de estos tonos seleccionados, y sitúa estos tonos en una lista de tonos. En cada iteración o nivel, los tonos seleccionados se eliminan de la señal de entrada para pasar una señal residual hacia delante. Una vez completado, se extraen de la señal de entrada todas las otras frecuencias que no cumplen el criterio para tonos y emiten desde el extractor 102 de tonos de múltiples órdenes, específicamente la última etapa del banco de filtros jerárquico MDCT(256), en el dominio del tiempo en la línea 111 como la señal residual final.

El extractor 102 de tonos de múltiples órdenes usa, por ejemplo, cinco órdenes de transformadas superpuestas, comenzando desde la mayor y trabajando hasta la menor, para detectar tonos a través del uso de una función base. Se usan transformadas de tamaño: 8192, 4096, 2048, 1024 y 512 respectivamente, para una señal de audio cuya tasa de muestreo es 44100 Hz. Podrían elegirse otros tamaños de transformada. La Figura 7 muestra gráficamente cómo las transformadas se superponen entre sí. La función base se define mediante las ecuaciones:

$$F(t;A,l,f,\varphi) = A \cdot \frac{1-\cos(\frac{2\pi}{l}\cdot t)}{2} \cdot \text{sen}(\frac{2\pi}{l}\cdot f \cdot t + \varphi) \qquad t \in [0,l]$$

$$F(t;A,l,f,\varphi) = 0; \qquad t \notin [0,l]$$

en las que:

$$A_i = Amplitud = (\mathrm{Re}_i \cdot \mathrm{Re}_i + \mathrm{Im}_i \cdot \mathrm{Im}_i) - (\mathrm{Re}_{i+1} \cdot \mathrm{Re}_{i+1} + \mathrm{Im}_{i+1} \cdot \mathrm{Im}_{i+1})$$

t = tiempo $(t \in N$ siendo un valor entero positivo)

l = tamaño de transformada como una potencia de 2 $(l \in 512,1024,\ldots,8192)$

φ = fase

f = frecuencia $\left(f \in \left[1, \frac{l}{2}\right]\right)$

Tonos detectados en cada tamaño de transformada se decodifican localmente usando el mismo procedimiento de decodificación que el usado por el decodificador de la presente invención, que se describirá más adelante. Estos tonos decodificados localmente se invierten de fase y combinan con la señal de entrada original a través de suma de dominio de tiempo para formar la señal residual que se pasa a la siguiente iteración o nivel del HFB.

El nivel de enmascaramiento del calculador 101 de enmascaramiento y la lista de tonos del extractor 102 de tonos de múltiples órdenes son entradas al selector 103 de tono. El selector 103 de tono primero clasifica la lista de tonos proporcionada al mismo desde el extractor 102 de tonos de múltiples órdenes mediante potencia relativa sobre el nivel de enmascaramiento proporcionado por el calculador 101 de enmascaramiento. A continuación usa un procedimiento iterativo para determinar qué componentes tonales encajarán en una trama de datos codificados en el flujo de bits maestro. La cantidad de espacio disponible en una trama para componentes tonales depende de la tasa de datos predeterminada, antes de escalado, del flujo de datos maestro codificado. Si toda la trama se asigna para componentes tonales entonces no se realiza ninguna codificación residual. En general, alguna porción de la tasa de datos disponible se asigna para los componentes tonales con el resto (menos sobrecarga) reservado para los componentes residuales.

Grupos de canales se seleccionada adecuadamente para señales multicanal y canales primarios/secundarios identificados dentro de cada grupo de canales de acuerdo con una métrica tal como contribución a calidad perceptual. Los componentes tonales seleccionados se almacenan preferentemente usando codificación diferencial. Para audio estéreo, el campo de dos bits indica los canales primarios y secundarios. La amplitud/fase y amplitud/fase diferencial se almacenan para los canales primarios y secundarios, respectivamente. Para audio multicanal el canal primario se almacena con su amplitud y fase y se almacena una máscara de bits (Véase la Figura 9) para todos los canales secundarios con amplitud/fase diferencial para los canales secundarios incluidos. La máscara de bits indica qué otros canales se codifican conjuntamente con el canal primario y se almacena en el flujo de bits para cada componente tonal en el canal primario.

Durante este procedimiento iterativo, algunos o todos los componentes tonales que se determinan que no encajan en una trama pueden convertirse de vuelta en el dominio del tiempo y combinarse con la señal 111 residual. Si, por ejemplo, la tasa de datos es lo suficientemente alta, entonces habitualmente todos los componentes tonales deseleccionados se recombinan. Si, sin embargo, la tasa de datos es menor, los componentes tonales relativamente fuertes 'deseleccionados' se excluyen adecuadamente de la residual. Se ha encontrado que esto mejora la calidad perceptual en tasas de datos menores. Los componentes tonales deseleccionados representados por la señal 110, se decodifican localmente a través del decodificador 104 local para volver a convertirlos en el dominio del tiempo en la línea 114 y combinarse con la señal 111 residual desde el extractor 102 de tonos de múltiples órdenes en el combinador 105 para formar una señal 113 residual combinada. Obsérvese que las señales que aparecen en 114 y 111 son ambas señales de dominio de tiempo de modo que este procedimiento de combinación puede afectarse fácilmente. La señal 113 residual combinada se procesa adicionalmente por el codificador 107 residual.

La primera acción realizada por el codificador 107 residual es procesar la señal 113 residual combinada a través de un banco de filtros que subdivide la señal en subbandas de frecuencia de dominio de tiempo críticamente muestreadas. En una realización preferida, cuando el banco de filtros jerárquico se usa para extraer los componentes tonales, estos componentes de muestra de tiempo pueden leerse directamente del banco de filtros jerárquico eliminando de este modo la necesidad de un segundo banco de filtros dedicado al procesamiento de señal residual. En este caso, como se muestra en la Figura 21, el combinador 104 opera en la salida de la última etapa del banco de filtros jerárquico (MDCT(256)) para combinar los componentes 114 tonales 'deseleccionados' y decodificados con la señal 111 residual antes de calcular la IMDCT 2106, que produce las muestras de tiempo de subbanda (véase también la Figura 7 etapas 3906, 3908 y 3910). Adicionalmente a continuación se realiza la descomposición, cuantificación y disposición de estas subbandas en orden psicoacústicamente relevante. Los componentes residuales (muestras de tiempo y factores de escala) se codifican adecuadamente usando codificación conjunta de canal en el que las muestras de tiempo se representan mediante una cuadrícula G y los factores de escala mediante las cuadrículas G0 y G1 (Véase la Figura 11). La codificación conjunta de la señal residual usa

cuadrículas parciales, aplicadas a grupos de canales, que representan la relación de energías de señales entre canal primario y grupos de canales secundarios. Los grupos se seleccionan (dinámica o estáticamente) a través de correlaciones cruzadas u otras métricas. Pueden combinarse más de un canal y usarse como un canal primario (por ejemplo L+R primario, C secundario). El uso de cuadrículas de factor de escala parciales, G0, G1 en dimensiones de tiempo/frecuencia es novedoso como se aplica a estos grupos multicanal y puede asociarse más de un canal secundario con un canal primario dado. Los elementos de cuadrícula individuales y muestras de tiempo se clasifican por frecuencia clasificando más alto las frecuencias más bajas. Las cuadrículas se clasifican de acuerdo con la tasa de bits. La información de canal secundario se clasifica con menor prioridad que la información de canal primario.

El generador 108 de cadena de código toma una entrada del selector 103 de tono, en la línea 120, y codificador 107 residual en la línea 122, y codifica valores desde estas dos entradas usando codificación por entropía bien conocida en la técnica en el flujo 124 de datos. El formateador 109 de flujo de bits garantiza que aparezcan elementos psicoacústicos desde el selector 103 de tono y codificador 107 residual, después de codificarse a través del generador 108 de cadena de código, en la posición correcta en el flujo 126 de bits maestro. Las 'clasificaciones' se incluyen implícitamente en el flujo de bits maestro mediante la reordenación de los diferentes componentes.

Un escalador 115 elimina un número suficiente de los componentes codificados de menor clasificación desde cada trama del flujo **126** de bits maestro producido por el codificador para formar un flujo 116 de bits escalado que tiene una tasa de datos menor que o aproximadamente igual a una tasa de datos deseada.

Banco de filtros jerárquico

El extractor 102 de tonos de múltiples órdenes usa preferentemente un banco de filtros jerárquico 'modificado' para proporcionar una resolución de tiempo/frecuencia de múltiples resoluciones de la que pueden extraerse de forma eficiente tanto los componentes tonales como los componentes residuales. El HFB descompone la señal de entrada en coeficientes de transformada en resoluciones de frecuencia sucesivamente menores y de vuelta en muestras de subbanda de dominio de tiempo en resolución de escala de tiempo sucesivamente más fina en cada iteración sucesiva. Los componentes tonales generados por el banco de filtros jerárquico son exactamente los mismos que los generador por múltiples FFT superpuestas, sin embargo la carga de cálculo es mucho menor. El banco de filtros jerárquico aborda el problema de modelización de resolución de tiempo/frecuencia desigual del sistema auditivo humano analizando simultáneamente la señal de entrada en diferentes resoluciones de tiempo/frecuencia en paralelo para conseguir una descomposición de tiempo/frecuencia casi arbitraria. El banco de filtros jerárquico hace uso de una etapa de formación de ventana y superposición-adición en la transformada interior no encontrada en descomposiciones conocidas. Esta etapa y el diseño novedoso de la función de ventana permiten que esta estructura se itere en un árbol arbitrario para conseguir la descomposición deseada, y podría hacerse de una manera adaptativa a la señal.

Como se muestra en la Figura 21, un codificador 2100 de un solo canal extrae componentes tonales de los coeficientes de transformada en cada iteración 2101a,...2101e, cuantifica y almacena los componentes tonales extraídos en una lista 2106 de tonos. A continuación se analiza la codificación conjunta de los tonos y señales residuales para señales multicanal. En cada iteración la señal de entrada de dominio de tiempo (señal residual) se forma en ventana 2107 y se aplica 2108 una MDCT de N puntos para producir coeficientes de transformada. Los tonos se extraen 2109 de los coeficientes de transformada, cuantifican 2110 y añaden a la lista de tonos. Los componentes tonales seleccionados se decodifican localmente 2111 y restan 2112 de los coeficientes de transformada antes de realizar la transformada 2113 inversa para generar las muestras de subbanda de dominio de tiempo que forman la señal 2114 residual para la siguiente iteración del HFB. Una transformada 2115 inversa final con resolución de frecuencia relativamente menor que la iteración final del HFB se realiza en la residual 113 combinada y forma en ventana 2116 para extraer los componentes 2117 residuales de G. Como se describe anteriormente, cualquier tono 'deseleccionado' se decodifica 104 localmente y combina 105 con la señal 111 residual antes del cálculo de la transformada inversa final. Los componentes residuales incluyen componentes de muestra de tiempo (Cuadrícula G) y componentes de factor de escala (Cuadrícula G0, G1) que se extraen de la cuadrícula G en 2118 y 2119. La cuadrícula G se recalcula 2120 y la cuadrícula G y G1 se cuantifican 2121, 2122. A continuación se describe el cálculo de las cuadrículas G, G1 y G0. Los tonos cuantificados en la lista de tonos, cuadrícula G y cuadrícula de factor de escala G1 se codifican todos y sitúan en el flujo de bits maestro. La eliminación de los tonos seleccionados de la señal de entrada en cada iteración y el cálculo de la transformada inversa final son las modificaciones impuestas en el HFB por el codificador de audio.

Un reto fundamental en la codificación de audio es la modelización de la resolución de tiempo/frecuencia de la percepción humana. Señales transitorias, tal como un aplauso, requieren una alta resolución en el dominio del tiempo, mientras que señales armónicas, tal como un claxon, requieren resolución alta en el dominio de la frecuencia para representarse de forma precisa mediante un flujo de datos codificado. Pero es un principio bien conocido que resolución de tiempo y frecuencia son inversas entre sí y ninguna transformada sola puede convertir simultáneamente precisión alta en ambos dominios. El diseño de un códec de audio efectivo requiere equilibrar esta compensación entre resolución de tiempo y frecuencia.

Soluciones conocidas a este problema utilizan conmutación de ventana, adaptando el tamaño de transformada a la naturaleza transitoria de la señal de entrada (véase K. Brandenburg y col., "The ISO-MPEG-Audio Codec: A Generic

Standard for Coding of High Quality Digital Audio", Journal of Audio Engineering Society, Vol. 42, N.º 10, Octubre, 1994). Esta adaptación del tamaño de ventana de análisis introduce complejidad adicional y requiere una detección de eventos transitorios en la señal de entrada. Para gestionar la complejidad algorítmica, los procedimientos de conmutación de ventana de la técnica anterior habitualmente limitan el número de diferentes tamaños de ventana a dos. El banco de filtros jerárquico analizado en el presente documento evita este ajuste basto a las características de señal/auditivas representando/procesando la señal de entrada mediante un banco de filtros que proporciona múltiples resoluciones de tiempo/frecuencia en paralelo.

Existen muchos bancos de filtros, conocidos como bancos de filtros híbridos, que descomponen la señal de entrada en una representación de tiempo/frecuencia dada. Por ejemplo, el algoritmo de Capa 3 de MPEG descrito en ISO/IEC 11172-3 utiliza un banco de filtros de espejo de pseudocuadratura seguido por una transformada MDCT en cada subbanda para proporcionar la resolución de frecuencia deseada. En nuestro banco de filtros jerárquico utilizamos una transformada, tal como una MDCT, seguida por la transformada inversa (por ejemplo IMDCT) en grupos de líneas espectrales para realizar una transformación de tiempo/frecuencia flexible de la señal de entrada.

A diferencia de los bancos de filtros híbridos, el banco de filtros jerárquico usa resultados de dos transformadas exteriores superpuestas y consecutivas para calcular transformadas interiores 'superpuestas'. Con el banco de filtros jerárquico es posible agregar más de una transformadas encima de la primera transformada. Esto también es posible con bancos de filtros de la técnica anterior (por ejemplo bancos de filtros de tipo árbol), pero no es práctico debido a la rápida degradación de separación de dominio de frecuencia con el aumento de número de niveles. El banco de filtros jerárquico evita esta degradación de dominio de frecuencia a costa de algo de degradación de dominio de tiempo. Esta degradación de dominio de tiempo puede controlarse, sin embargo, a través de la correcta selección de forma o formas de ventana. Con la selección de la ventana de análisis correcta, los coeficientes de la transformada interior también pueden hacerse invariables a desplazamiento de tiempo iguales al tamaño de la transformada interior (no al tamaño de la transformada más exterior como en enfoques convencionales).

Una ventana adecuada W(x) denominada en el presente documento como la "Ventana de Shmunk", para uso con el banco de filtros jerárquico se define mediante:

$$W^2(x) = \frac{128 - 150\cos\left(\frac{2\pi x}{L}\right) + 25\cos\left(\frac{6\pi x}{L}\right) - 3\cos\left(\frac{10\pi x}{L}\right)}{256}$$

En la que x es el índice de muestra de dominio del tiempo (0 < x <= L), y L es la longitud de la ventana en muestras.

La respuesta de frecuencia 2603 de la ventana de Shmunk en comparación con la ventana 2602 derivada de Kaiser-Bessel usada comúnmente se muestra en la Figura 2a. Puede observarse que las dos ventanas son similares en forma pero la atenuación de lóbulos laterales es mayor con la ventana propuesta. La respuesta 2604 de dominio de tiempo de la ventana de Shmunk se muestra en la Figura 2b.

En las Figuras 3 y 4 se ilustra un banco de filtros jerárquico de aplicabilidad general para la provisión de una descomposición de tiempo/frecuencia. El HFB tendría que modificarse como se ha descrito anteriormente para su uso en el códec de audio. En la Figura 3, el número en cada línea discontinua representa el número de contenedores de frecuencia igualmente espaciados en cada nivel (aunque no se calculan todos estos contenedores). Las flechas hacia abajo representan una transformada de MDCT de N puntos que resulta en N/2 subbandas. Las flechas hacia arriba representan una IMDCT que toma N/8 subbandas y transforma las mismas a N/4 muestras de tiempo dentro de una subbanda. Cada cuadrado representa una subbanda. Cada rectángulo representa N/2 subbandas. El banco de filtros jerárquico realiza las siguientes etapas:

(a) Como se muestra en la Figura 5a, las muestras 2702 de señal de entrada se almacenan en memoria intermedia en tramas de N muestras 2704, y cada trama se multiplica por una función 2706 de ventana de N muestras (Figura 5b) para producir N muestras 2708 formadas en ventanas (Figura 5c) (etapa 2900);
(b) Como se muestra en la Figura 3, una transformada de N puntos (representada por la flecha 2802 hacia abajo en la Figura 3) se aplica a las muestras 2708 formadas en ventanas para producir N/2 coeficientes 2804 de transformada (etapa 2902);
(c) Opcionalmente se aplica reducción de timbre a uno o más de los coeficientes 2804 de transformada aplicando una combinación lineal de uno o más coeficientes de transformada adyacentes (etapa 2904);
(d) Los N/2 coeficientes 2804 de transformada se dividen en P grupos de M_i coeficientes, de tal forma que la suma de los M_i coeficientes es N/2 ($\sum_{i=1}^{P} M_i = N/2$);
(e) Para cada de P grupos, una transformada inversa de (2* M_i) puntos (representada por la fecha 2806 hacia arriba en la Figura 3) se aplica a los coeficientes de transformada para producir (2* M_i) muestras de subbanda de cada grupo (etapa 2906);
(d) En cada subbanda, las (2* M_i) muestras de subbanda se multiplican por una función 2706 de ventana de (2* M_i) puntos (etapa 2908);

(e) En cada subbanda, las M_i muestras anteriores se superponen y añaden a correspondientes valores actuales para producir M_i nuevas muestras para cada subbanda (etapa 2910);
(f) N se establece igual al anterior M_i y selecciona nuevos valores para P y M_i, y
(g) Las etapas anteriores se repiten (etapa 2912) en una o más de las subbandas de M_i nuevas muestras usando los tamaños de transformada sucesivamente más pequeños para N hasta que se consigue la resolución de tiempo/transformada deseada (etapa 2914). Obsérvese, las etapas pueden iterarse en todas las subbandas, únicamente la menores subbandas o cualquier combinación deseada de las mismas. Si las etapas se iteran en todas las subbandas el HFB es uniforme, de lo contrario no es uniforme.

La gráfica de la respuesta 3300 de frecuencia de una implementación del banco de filtros de la Figura 3 y descrita anteriormente se muestra en la Figura 6 en la que N =128, M_i =16 y P= 4, y las etapas se iteran en las dos menores subbandas en cada etapa.

Las aplicaciones potenciales para este banco de filtros jerárquico van más allá de audio, para procesamiento de video y otros tipos de señales (por ejemplo sísmicas, médicas, otras señales de serie de tiempo). La codificación de y compresión de video tienen requisitos similares para la descomposición de tiempo/frecuencia, y la naturaleza arbitraría de la descomposición proporcionada por el banco de filtros jerárquico puede tener ventajas significativas sobre las técnicas del estado de la técnica actuales a base de Transformada de Coseno Discreta y descomposición de ondícula. El banco de filtros también puede aplicarse en el análisis y procesamiento de mediciones sísmicas o mecánicas, procesamiento de señales biomédicas, análisis y procesamiento de señales naturales o fisiológicas, voz y otras señales de serie de tiempo. Puede extraerse información de dominio de frecuencia de los coeficientes de transformada producidos en cada iteración en resoluciones de frecuencia sucesivamente menores. Asimismo puede extraerse información de dominio de tiempo de las muestras de subbanda de dominio de tiempo producidas en cada iteración en escalas de tiempo sucesivamente más finas.

Banco de filtros jerárquico: subbandas separadas uniformemente

La Figura 7 muestra un diagrama de bloques de una realización ilustrativa del banco de filtros jerárquico 3900, que implementa un banco de filtros de subbandas separadas uniformemente. Para un banco de filtros uniforme M_i =M =N/(2*P). La descomposición de la señal de entrada en señales 3914 de subbanda se describe como se indica a continuación:

1. Muestras 3902 de tiempo de entrada se forman en ventana en N puntos, 50 % tramas 3904 superpuestas.
2. Se realiza una MDCT de N puntos 3906 en cada trama.
3. Los coeficientes de MDCT resultantes se agrupan en P grupos 3908 de M coeficientes en cada grupo.
4. Se realiza una IMDCT de (2*M) puntos 3910 en cada grupo para formar (2*M) muestras 3911 de subbanda de tiempo.
5. Las muestras 3911 de tiempo resultantes se forman en ventana en (2*M) puntos, 50 % tramas superpuestas y añadidas superpuestas (OLA) 3912 para formar M muestras de tiempo en cada subbanda 3914.

En una implementación ilustrativa, N=256, P=32 y M=4. Obsérvese que diferentes tamaños de transformada y agrupamientos de subbanda representados por diferentes elecciones para N, P y M también pueden conseguirse para conseguir una descomposición de tiempo/frecuencia deseada.

Banco de filtros jerárquico: subbandas no separadas uniformemente

En las Figuras 8a y 8b se muestra otra realización de un banco 3000 de filtros jerárquico. En esta realización, algunas de las etapas de banco de filtros están incompletas para producir una transformada con tres diferentes intervalos de frecuencia con los coeficientes de transformada que representan una diferente resolución de frecuencia en cada intervalo. La señal de dominio de tiempo se descompone en estos coeficientes de transformada usando una serie de bancos de filtros de un solo elemento en cascada. El elemento de banco de filtros detallado puede iterarse un número de veces para producir una descomposición de tiempo/frecuencia deseada. Obsérvese que los números para tamaños de memoria intermedia, tamaños de transformada y tamaños de ventana y el uso de la MDCT/IMDCT para la transformada son únicamente para una realización ilustrativa y no limitan el alcance de la presente invención. También pueden usarse otros tamaños de memoria intermedia, ventana y transformada y otros tipos de transformada. En general, las M_i difieren entre sí pero satisfacen la restricción de que la suma de las M_i es igual a N/2.

Como se muestra en la Figura 8b, un único elemento de banco de filtros almacena en memoria intermedia 3022 muestras 3020 de entrada para formar memorias intermedias de 256 muestras 3024, que se forman en ventana 3026 multiplicando las muestras por una función de ventaja de 256 muestras. Las muestras formadas en ventanas 3028 se transforman a través de una MDCT 3030 de 256 puntos para formar 128 coeficientes de transformada 3032. De estos 128 coeficientes, se seleccionan 3034 los 96 coeficientes de frecuencia más alta para salida 3037 y no se procesan adicionalmente. Los 32 coeficientes de frecuencia más baja se transforman inversamente 3042 para producir 64 muestras de dominio de tiempo, que a continuación se forman en ventana 3044 en muestras 3046 y añaden superpuestas 3048 con la trama de salida anterior para producir 32 muestras de salida 3050.

En el ejemplo mostrado en la Figura 8a, el banco de filtros se compone de un elemento de banco de filtros 3004

iterado una vez con un tamaño de memoria intermedia de entrada de 256 muestras seguido por un elemento de banco de filtros 3010 también iterado con un tamaño de memoria intermedia de entrada de 256 muestras. La última etapa 3016 representa un único elemento de banco de filtros abreviado y se compone de las etapas de almacenamiento en memoria intermedia 3022, formación de ventana 3026 y MDCT 3030 únicamente para emitir 128 coeficientes de dominio de frecuencia que representan el intervalo de frecuencias más bajas de 0-1378 Hz.

Por lo tanto, suponiendo una entrada 3002 con una tasa de muestras de 44100 Hz, el banco de filtros mostrado produce 96 coeficientes que representan el intervalo de frecuencia 5513 a 22050 Hz en "Salida1" 3008, 96 coeficientes que representan el intervalo de frecuencia 1379 a 5512 Hz en "Salida2" 3014, y 128 coeficientes que representan el intervalo de frecuencia 0 a 1378 Hz en "Salida3" 3018,

Debería observarse que el uso de MDCT/IMDCT para la transformada/transformada inversa de frecuencia es ilustrativo y pueden aplicarse otras transformaciones de tiempo/frecuencia como parte de la presente invención. Son posibles otros valores para los tamaños de transformada y son posibles otras descomposiciones con este enfoque, expandiendo selectivamente cualquier rama en la jerarquía descrita anteriormente.

Codificación conjunta multicanal de componentes tonales y residuales

El selector 103 de tono en la Figura 1 toma como entrada datos desde el calculador 101 de máscara y la lista de tonos desde el extractor 102 de tonos de múltiples órdenes. El selector 103 de tono primero clasifica la lista de tonos mediante potencia relativa sobre el nivel de enmascaramiento desde el calculador 101 de máscara, formando una ordenación por importancia psicoacústica. La fórmula empleada se proporciona mediante:

$$P_k = A_k \cdot \frac{\sum_{i=0}^{l-1}\left(1-\cos\left(\frac{\pi(2i+1)}{l}\right)\right)}{\sqrt{M_{i,k}}}$$

en la que:

A_k = amplitud de línea espectral
Mi,k = nivel de enmascaramiento para línea espectral de k en la subtrama de máscara de i
l = longitud de función base en términos de subtramas de máscara

La suma se realiza sobre las subtramas en las que el componente espectral tiene un valor distinto de cero.

El selector 103 de tonos usa a continuación un procedimiento iterativo para determinar qué componentes tonales de la lista de tonos clasificada para la misma trama encajarán en el flujo de bits. En señales de audio estéreo o multicanal, en las que la amplitud de un tono es aproximadamente la misma en más de un canal, únicamente toda la amplitud y fase se almacenan en el canal primario; siendo el canal primario el canal con la mayor amplitud para el componente tonal. Otros canales que tienen características tonales similares almacenan la diferencia del canal primario.

Los datos para cada tamaño de transformada incluyen un número de subtramas, cubriendo el tamaño de transformada más pequeño 2 subtramas; el segundo 4 subtramas; el tercero 8 subtramas; el cuarto 16 subtramas; y el quinto 32 subtramas. Existen 16 subtramas a 1 trama. Datos de tono se agrupan por tamaño de la transformada en la que se encontró la información de tono. Para cada tamaño de transformada, los siguientes datos de componente tonal se cuantifican, codifican por entropía y sitúan en el flujo de bits: posición de subtrama codificada por entropía, posición espectral codificada por entropía, amplitud cuantificada codificada por entropía y fase cuantificada.

En el caso de audio multicanal, para cada componente tonal, se selecciona un canal como el canal primario. La determinación de qué canal debería ser el canal primario puede fijarse o puede hacerse a base de la señal características o criterios perceptuales. El número de canal del canal primario y su amplitud y fase se almacenan en el flujo de bits. Como se muestra en la Figura 9, se almacena una máscara 3602 de bits indicando cuál de los otros canales incluyen el componente tonal indicado y, por lo tanto, deberían codificarse como canales secundarios. La diferencia entre las amplitudes y fases primarias y secundarias se codifican por entropía y almacenan para cada canal secundario en el que está presente el componente tonal. Este ejemplo particular supone que existen 7 canales, y el canal principal es el canal 3. La máscara 3602 de bits indica la presencia del componente tonal en los canales secundarios 1, 4 y 5. No se usa ningún bit para el canal primario.

La salida 4211 del extractor 102 de tonos de múltiples órdenes se compone de tramas de coeficientes de MDCT en una o más resoluciones. El selector 103 de tono determina qué componentes tonales pueden retenerse para inserción en la trama de salida de flujo de bits por el generador 108 de cadena de código, a base de su relevancia a la calidad de señal decodificada. Esos componentes tonales determinados que no encajan en la trama son la salida 110 al decodificador 104 local. El decodificador 104 local toma la salida 110 del selector 103 de tono y sintetiza todos los componentes tonales añadiendo cada componente tonal escalado con coeficientes 2000 de síntesis de

una tabla de consulta (Figura 20) para producir tramas de coeficientes de MDCT (Véase la Figura 16). Estos coeficientes se añaden a la salida 111 del extractor 102 de tonos de múltiples órdenes en el combinador 105 para producir una señal 113 residual en la resolución de MDCT de la última iteración del banco de filtros jerárquico.

Como se muestra en la Figura 10, la señal 113 residual para cada canal se pasa al codificador 107 residual como los coeficientes de MDCT 3908 del banco de filtros jerárquico 3900, antes de las etapas de formación de ventana y superposición-adición 3904 e IMDCT 3910 mostradas en la Figura 7. Las etapas posteriores de IMDCT 3910, formación de ventana y superposición-adición 3912 se realizan para producir 32 subbandas de frecuencia críticamente muestreadas espaciadas igualmente 3914 en el dominio del tiempo para cada canal. Las 32 subbandas, que componen los componentes de muestra de tiempo, se denominan cuadrícula G. Obsérvese que podrían usarse otras realizaciones del banco de filtros jerárquico en un codificador para implementar diferentes descomposiciones de tiempo/frecuencia que las descritas anteriormente y podrían usarse otras transformadas para extraer componentes tonales. Si no se usa un banco de filtros jerárquico para extraer componentes tonales, puede usarse otra forma de banco de filtros para extraer las subbandas pero en una carga de cálculo mayor.

Para audio estéreo o multicanal, se hacen varios cálculos en el bloque 501 de selección de canal para determinar el canal primario y secundario para la codificación de componentes tonales, así como el procedimiento de codificación de componentes tonales (por ejemplo, Izquierda-Derecha o Medio-Lateral). Como se muestra en la Figura 11, un procedimiento 3702 de agrupamiento de canal primero determina cuál de los múltiples canales puede codificarse conjuntamente y todos los canales se forman en grupos con el último grupo estando posiblemente incompleto. Los agrupamientos se determinan por criterios perceptuales de un oyente y eficiencia de codificación, y pueden construirse grupos de canales de combinaciones de más de dos canales (por ejemplo, una señal de 5 canales compuesta de canales L, R, Ls, Rs y C puede agruparse como {L,R}, {Ls, Rs}, {L+R, C}. Los grupos de canales se ordenan a continuación como canales primarios y secundarios. E una realización multicanal ilustrativa, la selección del canal primario se hace a base de la potencia relativa de los canales sobre la trama. Las siguientes ecuaciones definen las potencias relativas:

$$P_l = \sum_{i=0}^{15} L_i^2 \qquad P_r = \sum_{i=0}^{15} R_i^2 \qquad P_m = \sum_{i=0}^{15} (L_i + R_i)^2 \qquad P_s = \sum_{i=0}^{15} (L_i - R_i)^2$$

El modo de agrupamiento también se determina como se muestra en la etapa 3704 de la Figura 11. Los componentes tonales pueden codificarse como representación de Izquierda-Derecha o Medio-Lateral, o la salida de esta etapa puede resultar en un único canal primario únicamente como se muestra mediante las líneas discontinuas. En representación de Izquierda-Derecha, el canal con la mayor potencia para la subbanda se considera el primario y se establece un único bit en el flujo de bits 3706 para la subbanda si el canal derecho es el canal de mayor potencia. Se usa codificación Medio-Lateral para una subbanda si se cumple la siguiente condición para la subbanda:

$$P_m > 2 \cdot P_s$$

Para señales multicanal, lo anterior se realiza para cada grupo de canales.

Para una señal estéreo, el cálculo 502 de cuadrícula proporciona una cuadrícula de panoramización estéreo en la que la panoramización estéreo puede reconstruirse difícilmente y aplicarse a la señal residual. La cuadrícula estéreo es 4 subbandas por 4 intervalos de tiempo, cada subbanda en la cuadrícula estéreo cubre 4 subbandas y 32 muestras de la salida del banco 500 de filtros, comenzando con bandas de frecuencia por encima de 3 kHz. Podrían elegirse otros tamaños de cuadrícula, subbandas de frecuencia cubiertas y divisiones de tiempo. Valores en las celdas de la cuadrícula estéreo son la relación de la potencia del canal dado a la del canal primario, para el intervalo de valores cubiertos por la celda. La relación se cuantifica a continuación a la misma tabla que la usada para codificar componentes tonales. Para señales multicanal, la cuadrícula estéreo anterior se calcula para cada grupo de canales.

Para señales multicanal, el cálculo 502 de cuadrícula proporciona múltiples cuadrículas de factor de escala, una por cada grupo de canales, que se insertan en el flujo de bits en orden de su importancia psicoacústica en el dominio espacial. Le calcula la relación de la potencia del canal dado al canal primario para cada grupo de 4 subbandas por 32 muestras. Esta relación se cuantifica a continuación y este valor cuantificado más signo de logaritmo de la relación de potencia se inserta en el flujo de bits.

El cálculo 503 de cuadrícula de factor de escala calcula la cuadrícula G1, que se sitúa en el flujo de bits. Se describe ahora el procedimiento de cálculo de G1. G0 se obtiene primero a partir de G. G0 contiene todas las 32 subbandas pero únicamente la mitad de resolución de tiempo de G. Los contenidos de las celdas en G0 son valores cuantificados del máximo de dos valores vecinos de una subbanda dada de G. Cuantificación (denominada en las siguientes ecuaciones como *Cuantificar*) se realiza usando la misma tabla de cuantificación logarítmica modificada como se usó para codificar los componentes tonales en el extractor 102 de tonos de múltiples órdenes. Cada celda en G0 se determina por lo tanto mediante:

$$G0_{m,n} = \textit{Cuantificar}\,(\textit{Máximo}(G_{m,2n}, G_{m,2n+1})) \qquad n \in [0 \ldots 63]$$

en la que:

m es el número de subbandas
n es el número de columnas de G0

G1 se obtiene a partir de G0. G1 tiene 11 subbandas superpuestas y 1/8 de la resolución de tiempo de G0, formando una cuadrícula de dimensión de 11 x 8. Cada celda en G1 se cuantifica usando la misma tabla que la usada para componentes tonales y encontrada usando la siguiente fórmula: $G1_{m,n} = Cuantificar\left(\sum_{l=0}^{31}\left(W_l \cdot \sqrt{\sum_{i=8n}^{8n+7} G_{l,i}^2}\right)\right)$

en la que: W_l es un valor de ponderación obtenido de la Tabla 1 en la Figura 19.

G0 se recalcula a partir de G1 en decodificador 506 de cuadrícula local. En el bloque 507 de cuantificación de muestra de tiempo, se extraen muestras de salida de tiempo ("componentes de muestra de tiempo") del banco de filtros jerárquico (Cuadrícula G), que pasan a través del bloque 504 de selección de nivel de cuantificación, escaladas dividiendo los componentes de muestra de tiempo por los respectivos valores en la recalculada G0 del decodificador 506 de cuadrícula local y cuantificados al número de niveles de cuantificación, como una función de subbanda, determinados por el bloque 504 de selección de nivel de cuantificación. Estas muestras de tiempo cuantificadas se sitúan a continuación en el flujo de datos codificado junto con la cuadrícula G1 cuantificada. En todos los casos, se usa un modelo que refleja la importancia psicoacústica de estos componentes para determinar la prioridad para la operación de almacenamiento de flujo de bits.

En una etapa de mejora adicional para mejorar la ganancia de codificación para algunas señales, pueden procesarse adicionalmente cuadrículas que incluyen G, G1 y cuadrículas parciales aplicando una Transformada de Coseno Discreta bidimensional (DCT) antes de cuantificación y codificación. La correspondiente DCT inversa se aplica en el decodificador tras la cuantificación inversa para reconstruir las cuadrículas originales.

Flujo de bits escalable y mecanismo de escalado

Típicamente, cada trama del flujo de bits maestro incluirá (a) una pluralidad de componentes tonales cuantificados que representan contenidos de dominio de frecuencia a diferentes resoluciones de frecuencia de la señal de entrada, b) componentes de muestra de tiempo residuales cuantificados que representan la residual de dominio de tiempo formada a partir de la diferencia entre los componentes tonales reconstruidos y la señal de entrada, y c) cuadrículas de factor de escala que representan las energías de señal de la señal residual, que se extienden un intervalo de frecuencia de la señal de entrada. Para una señal multicanal cada trama también puede contener d) cuadrículas parciales que representan las relaciones de energía de señal de los canales de señal residuales dentro de grupos de canales y e) una máscara de bits para cada primario especificando la codificación conjunta de canales secundarios para componentes tonales. Normalmente una porción de la tasa de datos disponible en cada trama se asigna desde los componentes tonales (a) y se asigna una porción para los componentes residuales (b, c). Sin embargo, en algunos casos toda la tasa disponible puede asignarse para codificar los componentes tonales. Como alternativa, toda la tasa disponible puede asignarse para codificar los componentes residuales. En casos extremos, únicamente las cuadrículas de factor de escala pueden codificarse, en cuyo caso el decodificador usa una señal de ruido para reconstruir una señal de salida. En la mayoría de aplicaciones reales, el flujo de bits escalado incluirá al menos algunas tramas que contienen componentes tonales y algunas tramas que incluyen cuadrículas de factor de escala.

La estructura y orden de componentes situados en el flujo de bits maestro, según se define mediante la presente invención, proporciona escalabilidad de flujo de datos de intervalo de bits amplia y granularidad fina. Es esta estructura y orden la que permite que el flujo de bits se escale suavemente por mecanismos externos. La Figura 12 representa la estructura y orden de componentes a base del códec de descompresión de audio de la Figura 1 que descompone el flujo de datos original en un conjunto particular de componentes psicoacústicamente relevantes. El flujo de bits escalable usado en este ejemplo se compone de un número de estructuras de datos de Formato de Archivo de Intercambio de Recursos, o RIFF, llamadas "fragmentos", aunque pueden usarse otras estructuras de datos. Este formato de archivo que se conoce bien por los expertos en la materia, permite la identificación del tipo de datos transportados por un fragmento así como la cantidad de datos transportada por un fragmento. Obsérvese que puede usarse cualquier formato de flujo de datos que transporta información con respecto a la cantidad y tipo de datos transportados en sus estructuras de datos de flujo de datos para practicar la presente invención.

La Figura 12 muestra el diseño de un fragmento 900 de trama de tasa de datos escalable, junto con subfragmentos 902, 903, 904, 905, 906, 906, 907, 908, 909, 910 y 912, que comprenden los datos psicoacústicos que se transportan dentro del fragmento 900 de trama. Aunque la Figura 12 únicamente representa ID de fragmento y longitud de fragmento para el fragmento de trama, ID de subfragmento y datos de longitud de subfragmento se incluyen dentro de cada subfragmento. La Figura 12 muestra el orden de subfragmentos en una trama del flujo de bits escalable. Estos subfragmentos contienen los componentes psicoacústicos producidos por el codificador de flujo de bits escalable, con un subfragmento único usado para cada subdominio del flujo de datos codificado. Además de los subfragmentos que se disponen en importancia psicoacústica, ya sea por una decisión o cálculo a priori, los componentes dentro de los subfragmentos también se disponen en importancia psicoacústica. El fragmento 911

nulo, que es el último fragmento en la trama, se usa para empaquetar fragmentos en el caso en el que se requiera que la trama tenga un tamaño constante o específico. Por lo tanto el fragmento 911 no tiene relevancia psicoacústica y es el fragmento psicoacústico menos importante. Fragmento 910 de muestras de tiempo 2 aparece en el lado derecho de la figura y el fragmento psicoacústico más importante, fragmento 902 de cuadrícula 1 aparece en el lado izquierdo de la figura. Operando para eliminar primero datos del fragmento psicoacústicamente menos relevante en el final del flujo de bits, el fragmento 910, y trabajando hacia la eliminación de componentes psicoacústicamente cada vez más relevantes hacia el comienzo del flujo de bits, el fragmento 902, se mantiene la mayor calidad posible para cada reducción sucesiva en tasa de datos. Debería observarse que la mayor tasa de datos, junto con la mayor calidad de audio, capaz de soportarse por el flujo de bits, se define en tiempo de codificación. Sin embargo, la menor tasa de datos después de escalado se define por el nivel de calidad de audio que es aceptable para su uso mediante una aplicación o mediante la restricción de tasa situada en el canal o medios.

Cada componente psicoacústico eliminado no utiliza el mismo número de bits. La resolución de escalado para la implementación actual de la presente invención oscila desde 1 bit para componentes de la menor importancia psicoacústica hasta 32 bits para esos componentes de la mayor importancia psicoacústica. El mecanismo de escalado del flujo de bits no necesita eliminar fragmentos enteros cada vez. Como se ha mencionado anteriormente, componentes dentro de cada fragmento se disponen de modo que los datos más psicoacústicamente importantes se sitúan en el comienzo del fragmento. Por esta razón, pueden eliminarse componentes del final del fragmento, un componente cada vez, mediante un mecanismo de escalado mientras mantiene la mejor calidad de audio posible con cada componente eliminado. En una realización de la presente invención, se eliminan componentes enteros mediante el mecanismo de escalado, mientras en otras realizaciones, pueden eliminarse algunos o todos los componentes. El mecanismo de escalado elimina componentes dentro de un fragmento según se requiera, actualizando el campo de longitud de fragmento del fragmento particular del que se eliminaron los componentes, la longitud 915 de fragmento de trama y la suma 901de control de trama. Como se observará a partir de la descripción detallada de las realizaciones ilustrativas de la presente invención, con longitud de fragmento actualizada para cada fragmento escalado, así como longitud de fragmento de trama actualizada e información de suma de control de trama disponible al decodificador, el decodificador puede procesar apropiadamente el flujo de bits escalado, y automáticamente producir una señal de salida de audio de tasa de muestra fija para distribución al DAC, incluso aunque hay fragmentos dentro del flujo de bits que son componentes faltantes, así como fragmentos que faltan completamente del flujo de bits.

Decodificador de flujo de bits escalable para una topología de codificación residual

La Figura 13 muestra el diagrama de bloques para el decodificador. El analizador 600 de flujo de bits lee información secundaria inicial que consiste en: la tasa de muestras en hercios de la señal codificada antes de codificación, el número de canales de audio, la tasa de datos original del flujo y la tasa de datos codificada. Esta información secundaria inicial permite que la misma reconstruya la tasa de datos completa de la señal original. Adicionalmente el analizador 600 de flujo de bits analiza componentes en el flujo 599 de datos y pasan al elemento de decodificación apropiado: decodificador 601 de tono o decodificador 602 residual. Componentes decodificados a través del decodificador 601 de tono se procesan a través de la transformada 604 de frecuencia inversa que convierte la señal de vuelta al dominio del tiempo. El bloque 608 de superponer-añadir añade los valores de la última mitad de la trama anteriormente decodificada a los valores de la primera mitad de la trama recién decodificada que es la salida de la transformada 604 de frecuencia inversa. Componentes que el analizador 600 de flujo de bits determina que sean parte del procedimiento de decodificación residual se procesan a través del decodificador 602 residual. La salida del decodificador 602 residual, que contiene 32 frecuencia subbandas representadas en el dominio del tiempo, se procesa a través del banco 605 de filtros inverso. El banco 605 de filtros inverso recombina las 32 subbandas en una señal a combinar con la salida del añadir-superponer 608 en el combinador 607. La salida del combinador 607 es la señal de salida decodificada 614.

Para reducir la carga de cálculo, la transformada 604 de frecuencia inversa y el banco 605 de filtros inverso que convierten las señales de vuelta al dominio del tiempo pueden implementarse con un banco de filtros jerárquico inverso, que integra estas operaciones con el combinador 607 para formar la señal 614 de salida de dominio de tiempo decodificada. El uso del banco de filtros jerárquico en el decodificador es novedoso en la forma en que se combinan los componentes tonales con la residual en el banco de filtros jerárquico en el decodificador. Las señales residuales se transforman hacia delante usando MDCT en cada subbanda, y a continuación los componentes tonales se reconstruyen y combinan antes de la última etapa de IMDCT. El enfoque de múltiples resoluciones podría generalizarse para otras aplicaciones (por ejemplo múltiples niveles, diferentes descomposiciones aún se cubrirían mediante este aspecto de la invención).

Banco de filtros jerárquico inverso

Para reducir la complejidad del decodificador, puede usarse el banco de filtros jerárquico para combinar las etapas de la transformada 604 de frecuencia inversa, banco 605 de filtros inverso, añadir-superponer 608 y combinador 607. Como se muestra en la Figura 15, la salida del decodificador 602 residual se pasa a la primera etapa del banco 4000 de filtros jerárquico inverso mientras la salida del decodificador 601 de tono se añade a las muestras residuales en la etapa de mayor resolución de frecuencia antes de la transformada 4010 inversa final. Las muestras

transformadas inversas resultantes se añaden superpuestas a continuación para producir las muestras 4016 de salida lineales.

En la Figura 22 se muestra la operación general del decodificador para un único canal que usa el HFB 2400. Las etapas adicionales para decodificación multicanal de los tonos y señales residuales se muestran en las Figuras 10, 11 y 18. Las cuadrículas cuantificadas G1 y G' se leen del flujo 599 de bits mediante el analizador 600 de flujo de datos. El decodificador 602 residual cuantifica inversamente (Q^{-1}) 2401, 2402 las cuadrículas G' 2403 y G1 2404 y reconstruye la cuadrícula G0 2405 a partir de la cuadrícula G1. La cuadrícula G0 se aplica a la cuadrícula G' multiplicando 2406 elementos correspondientes en cada cuadrícula para formar la cuadrícula G escalada, que consiste en muestras 4002 de subbanda de tiempo que se introducen en la siguiente etapa en el banco de filtros jerárquico 2401. Para una señal multicanal, la cuadrícula 508 parcial se usará para decodificar los canales secundarios.

Los componentes tonales (T5) 2407 a la menor resolución de frecuencia (P=16, M=256) se leen del flujo de bits por el analizador 600 de flujo de datos. El decodificador 601 de tono cuantifica inversamente 2408 y sintetiza 2409 el componente tonal para producir P grupos de M coeficientes de dominio de frecuencia.

Las muestras 4002 de tiempo de cuadrícula G se forman en ventana y añaden superpuestas 2410 como se muestra en la Figura 15, se transforman hacia delante mediante P MDCT de (2*M) puntos 2411 para formar P grupos de M coeficientes de dominio de frecuencia que se combinan a continuación 2412 con los P grupos de M coeficientes de dominio de frecuencia sintetizados a partir de los componentes tonales como se muestra en la Figura 16. Los coeficientes de dominio de frecuencia combinados se concatenan a continuación y transforman inversamente mediante una IMDCT de longitud N 2413, forman en ventanas y añaden superpuestas 2414 para producir N muestras de salida 2415 que se introducen en la siguiente etapa del banco de filtros jerárquico.

Los siguientes componentes tonales (T4) de menor resolución de frecuencia se leen del flujo de bits, y combinan con la salida de la etapa previa del banco de filtros jerárquico como se ha descrito anteriormente, y a continuación esta iteración continua P=8, 4, 2, 1 y M=512, 1024, 2048 y 4096 hasta que todos los componentes de frecuencia se han leído del flujo de bits, combinado y reconstruido.

En la etapa final del decodificador, la transformada inversa produce N muestras de ancho de banda completo de tiempo que se emiten como la salida 614decodificada. Los valores anteriores de P, M y N son únicamente para una realización ilustrativa y no limitan el alcance de la presente invención. También pueden usarse otros tamaños de memoria intermedia, ventana y transformada y otros tipos de transformada.

Como se describe, el decodificador anticipa la recepción de una trama que incluye componentes tonales, componentes de muestra de tiempo y cuadrículas de factor de escala. Sin embargo, si falta uno o más de estos del flujo de bits escalado el decodificador reconstruye sin problemas la salida decodificada. Por ejemplo, si la trama incluye únicamente componentes tonales entonces las muestras de tiempo en 4002 son cero y no se combina 2403 ninguna residual con los componentes tonales sintetizados en la primera etapa del HFB inverso. Si falta uno o más de los componentes tonales T5, ... T1, entonces se combina 2403 un valor cero en esa iteración. Si la trama incluye únicamente las cuadrículas de factor de escala, entonces el decodificador sustituye una señal de ruido para la cuadrícula G para decodificar la señal de salida. Como resultado, el decodificador puede reconstruir sin problemas la señal de salida decodificada ya que la composición de cada trama del flujo de bits escalado puede cambiar debido al contenido de la señal, cambiando restricciones de tasa de datos, etc.

La Figura 16 muestra en más detalle cómo se combinan los componentes tonales dentro del banco de filtros jerárquico inverso de la Figura 15. En este caso, las señales 4004 residuales de subbanda se forman en ventana y añaden superpuestas 4006, transforman hacia delante 4008 y los coeficientes resultantes de todas las subbandas se agrupan para formar la única trama de coeficientes 4010. Cada coeficiente tonal se combina a continuación con la trama de coeficientes residuales multiplicando 4106 la envolvente 4102 de amplitud de componente tonal por un grupo de coeficientes de síntesis 4104 (normalmente proporcionados por una tabla de consulta) y añadiendo los resultados a los coeficientes centrados alrededor de la frecuencia 4106 de componente tonal dada. La adición de estos coeficientes de síntesis tonales se realiza en las líneas espectrales de la misma región de frecuencias por toda la longitud de componente tonal. Después de que se añaden todos los componentes tonales de esta manera, se realiza la IMDCT 4012 final y los resultados se forman en ventana y añaden superpuestos 4014 con la trama anterior para producir las muestras 4016 de tiempo de salida.

En la Figura 14 se muestra la forma general del banco 2850 de filtros jerárquico inverso que es compatible con el banco de filtros jerárquico mostrado en la Figura 3. Cada trama de entrada contiene M_i muestras de tiempo en cada de P subbandas, de tal forma que la suma de los M_i coeficientes es N/2:

$$\sum_{i=1}^{P} M_i = N/2;$$

En la Figura 14, las flechas hacia arriba representan una transformada de IMDCT de N puntos que toma N/2 coeficientes de MDCT y transforma las mismas a N muestras de dominio tiempo. Las flechas hacia abajo

representan una MDCT que toma N/4 muestras dentro de una subbanda y transforma las mismas a N/8 coeficientes de MDCT. Cada cuadrado representa una subbanda. Cada rectángulo representa N/2 coeficientes de MDCT. Las siguientes etapas se muestran en la Figura 14:

(a) En cada subbanda, las M_i muestras anteriores se almacenan en memoria y concatenan con las M_i muestras actuales para producir (2* M_i) nuevas muestras para cada subbanda 2828;
(b) En cada subbanda, las (2* M_i) muestras de subbanda se multiplican por una función 2706 de ventana de (2* M_i) puntos (Figura 5a-5c);
(c) Una transformada de (2* M_i) puntos (representada por la flecha 2826 hacia abajo) se aplica para producir M_i coeficientes de transformada para cada subbanda;
(d) Los M_i coeficientes de transformada para cada subbanda se concatenan para formar un único grupo 2824 de N/2 coeficientes;
(e) Una transformada inversa de N puntos (representada por la fecha 2822 hacia arriba) se aplica a los coeficientes concatenados para producir N muestras;
(f) Cada trama de N muestras 2704 se multiplica por una función de ventana de N muestras 2706 para producir N muestras 2708 formadas en ventanas;
(g) Las muestras 2708 formadas en ventanas resultantes se añaden superpuestas para producir N/2 nuevas muestras de salida en el nivel de subbanda dado;
(h) Las etapas anteriores se repiten en el nivel actual y todos los niveles posteriores hasta que se hayan procesado todas las subbandas y las muestras 2840 de tiempo originales se reconstruyen.

Banco de filtros jerárquico inverso : subbandas separadas uniformemente

La Figura 15 muestra un diagrama de bloques de una realización ilustrativa de un banco 4000 de filtros jerárquico inverso compatible con el banco de filtros hacia delante mostrado en la Figura 7. La síntesis de la señal de salida decodificada 4016 se describe en más detalle como se indica a continuación:

1. Cada trama 4002 de entrada contiene M muestras de tiempo en cada de P subbandas.
2. Almacenar en memoria intermedia cada subbanda 4004, desplazar en M nuevas muestras, aplicar ventana de (2*M) puntos, 50 % superponer-añadir (OLA) 4006 para producir M nuevas muestras de subbanda.
3. Una MDCT de (2*M) puntos 4008 realizada dentro de cada subbanda para formar M coeficientes de MDCT en cada una de P subbandas.
4. Los coeficientes de MDCT resultantes se agrupan para formar una única trama 4010 de (N/2) coeficientes de MDCT.
5. Una IMDCT 4012 de N puntos realizada en cada trama
6. La salida de IMDCT se forma en ventana en N puntos, 50 % tramas superpuestas y añadidas superpuestas 4014 para formar N/2 nuevas muestras 4016 de salida.

En una implementación ilustrativa, N=256, P=32 y M=4. Obsérvese que también pueden conseguirse diferentes tamaños de transformada y agrupamientos de subbanda representados por diferentes elecciones para N, P y M para conseguir una descomposición de tiempo/frecuencia deseada.

Banco de filtros jerárquico inverso: subbandas espaciadas no uniformemente

Otra realización del banco de filtros jerárquico inverso se muestra en la Figura 17a-b, que es compatible con el banco de filtros mostrado en la Figura 8a-b. En esta realización, algunos de los elementos de banco de filtros detallados están incompletos para producir una transformada con tres diferentes intervalos de frecuencia con los coeficientes de transformada que representan una diferente resolución de frecuencia en cada intervalo. La reconstrucción de la señal de dominio de tiempo a partir de estos coeficientes de transformada se describe como se indica a continuación:
En este caso, el primer elemento 3110 de síntesis omite las etapas de almacenamiento en memoria intermedia 3122, formación de ventana 3124 y la MDCT 3126 del elemento detallado mostrado en la Figura 17b. En su lugar, la salida 3102 forma un único conjunto de coeficientes que se transforman inversamente 3130 para producir 256 muestras de tiempo, que se forman en ventana 3132 y añaden superpuestas 3134 con la trama anterior para producir la salida 3136 de 128 nuevas muestras de tiempo para esta etapa.

La salida del primer elemento 3110 y 96 coeficientes 3106 se introducen al segundo elemento 3112 y combinan como se muestra en la Figura 17b para producir 128 muestras de tiempo para introducir al tercer elemento 3114 del banco de filtros. El segundo elemento 3112 y tercer elemento 3114 en la Figura 17a implementan todo el elemento detallado de la Figura 17b, en cascada para producir 128 nuevas muestras de tiempo emitidas desde el banco de filtros 3116. Obsérvese que la memoria intermedia y tamaños de transformada se proporcionan como ejemplos únicamente y pueden usarse otros tamaños. En particular obsérvese que el almacenamiento en memoria intermedia 3122 en la entrada al elemento detallado puede cambiar para acomodar diferentes tamaños de entrada dependiendo de donde se usa en la jerarquía del banco de filtros general.

Se describirán ahora detalles adicionales con respecto a los bloques de decodificador.

Analizador 600 de flujo de datos

El analizador 600 de flujo de bits lee información de fragmento de IFF del flujo de bits y pasa elementos de esa información en el decodificador apropiado, decodificador 601 de tono o decodificador 602 residual. Es posible que el flujo de bits pueda haber escalado antes de alcanzar el decodificador. Dependiendo del procedimiento de escalado empleado, elementos de datos psicoacústicos al final de un fragmento pueden no ser válidos debido a bits faltantes. El decodificador 601 de tono y decodificador 602 residual ignoran apropiadamente los datos que se encuentran como no válidos en el final de un fragmento. Una alternativa al decodificador 601 de tono y decodificador 602 residual que ignora elementos de datos psicoacústicos completos, cuando faltan bits del elemento, es que estos decodificadores recuperen tanto del elemento como sea posible leyendo en los bits que salen y rellenen los restantes bits faltantes con ceros, patrones aleatorios o patrones a base de elementos de datos psicoacústicos anteriores. Aunque es más intensivo en cálculo, se prefiere el uso de datos a base de elementos de datos psicoacústicos anteriores porque el audio decodificado resultante puede coincidir más estrechamente con la señal de audio original.

Decodificador 601 de tono

Información de tono encontrada por el analizador 600 de flujo de bits se procesa a través del decodificador 601 de tono. La resíntesis de componentes tonales se realiza usando el banco de filtros jerárquico como se ha descrito anteriormente. Como alternativa, puede usarse una Transformada rápida de Fourier inversa cuyo tamaño es el mismo tamaño que el tamaño de transformada más pequeño que se usó para extraer los componentes tonales en el codificador.

Las siguientes etapas se realizan para decodificación tonal:

a) Inicializar el dominio de la frecuencia subtrama con valores cero.
b) Resintetizar la porción requerida de componentes tonales a partir del tamaño de transformada más pequeño en el dominio de la frecuencia subtrama.
c) Resintetizar y añadir en las posiciones requeridas, componentes tonales de los otros cuatro tamaños de transformada en la misma subtrama. La resísntesis de estos otros cuatro tamaños de transformada pueden producirse en cualquier orden.

El decodificador 601 de tono decodifica los siguientes valores para cada agrupamiento de tamaño de transformada: amplitud cuantificada, fase cuantificada, distancia espectral desde el componente tonal anterior para el agrupamiento y la posición del componente dentro de toda la trama. Para señales multicanal, la información secundaria se almacena como diferencias del canal primario valores y necesita restaurar a valores absolutos añadiendo los valores obtenidos del flujo de bits al valor obtenido para el canal primario. Para señales multicanal, 'presencia' por canal del componente tonal también se proporciona por la máscara 3602 de bits que se decodifica del flujo de bits. Se hace procesamiento adicional en canales secundarios independientemente del canal primario. Si el decodificador 601 de tono no es capaz de adquirir completamente los elementos necesarios para reconstruir un tono a partir del fragmento, ese elemento tonal se descarta. La amplitud cuantificada se decuantifica usando la inversa de la tabla usada para cuantificar el valor en el codificador. La fase cuantificada se decuantifica usando la inversa de la cuantificación lineal usada para cuantificar la fase en el codificador. La posición espectral de frecuencia absoluta se determina añadiendo el valor de diferencia obtenida del flujo de bits al valor decodificado anteriormente. Definir *Amplitud* para ser la amplitud decuantificada, *Fase* para ser la fase decuantificada, y *Frec* para ser la posición de frecuencia absoluta, el siguiente pseudocódigo describe la resíntesis de componentes tonales del tamaño de transformada más pequeño:

```
Re[Frec ]      += Amplitud * sin( 2 * Pi * Fase / 8 );
Im[Frec ]      += Amplitud * cos( 2 * Pi * Fase / 8 );
Re[Frec + 1]   += Amplitud * sin( 2 * Pi * Fase / 8 );
Im[Frec + 1]   += Amplitud * cos( 2 * Pi * Fase / 8);
```

Resíntesis de funciones base más largas se extienden sobre más subtramas por lo tanto la amplitud y fase valores necesarios para actualizarse de acuerdo con la frecuencia y longitud de la función base. El siguiente pseudo código describe cómo se hace esto:

```
xFrec = Frec >> (Grupo – 1);
CurrentPhase = Fase – 2 * (2 * xFrec + 1);
for (i = 0; i <longitud; i = i + 1)
{
        CurrentPhase += 2 * (2 * Frec + 1) / longitud;
        CurrentAmplitude = Amplitud * Envolvente[Grupo][i];

        Re[i][xFrec ] += CurrentAmplitude * sin( 2 * Pi * CurrentPhase / 8 );
        Im[i][xFrec ] += CurrentAmplitude * cos( 2 * Pi * CurrentPhase / 8 );
        Re[i][xFrec+1] += CurrentAmplitude * sin( 2 * Pi * CurrentPhase / 8 );
        Im[i][xFrec+1] += CurrentAmplitude * cos( 2 * Pi * CurrentPhase / 8 );
        }
```

en la que:

Amplitud, *Frec* y *Fase* son los mismos como se define anteriormente.
Grupo es un número que representan la función base tamaño de transformada, 1 para la transformada más pequeña y 5 para la más larga.
longitud es las subtramas para el *Grupo* y se proporciona mediante:
longitud = 2 ^ (Grupo - 1).
>> es el operador derecho de desplazamiento.
CurrentAmplitud y *CurrentFase* se almacenan para la siguiente subtrama. *Envolvente[Grupo][i]* es la envolvente con forma triangular de longitud apropiada (*longitud*) para cada grupo, teniendo valor cero en cada extremo y teniendo un valor de 1 en el medio.

Resíntesis de menores frecuencias en los tres mayores tamaños de transformada a través del procedimiento descrito anteriormente, provoca distorsión audible en la salida audio, por lo tanto la siguiente corrección empíricamente basada se aplica a líneas espectrales menos de 60 en grupos 3, 4 y 5:

xFrec = Frec >> (Grupo -1);

en la que:

```
CurrentPhase = Fase – 2 * (2 * xFreq + 1);

f_dlt = Frec – (xFrec << (Grupo – 1));

for (i = 0; i < longitud; i = i + 1)
{
        CurrentPhase += 2 * (2 * Frec + 1) / longitud;
        CurrentAmplitude = Amplitud * Envolvente[Grupo][i];

        Re_Amp = CurrentAmplitude * sin( 2 * Pi * CurrentPhase / 8);
        Im_Amp = CurrentAmplitude * cos( 2 * Pi * CurrentPhase / 8);

        a0 = Re_Amp * CorrCf[f_dlt][0];
        b0 = Im_Amp * CorrCf[f_dlt][0];
        a1 = Re_Amp * CorrCf[f_dlt][1];
        b1 = Im_Amp * CorrCf[f_dlt][1];
        a2 = Re_Amp * CorrCf[f_dlt][2];
        b2 = Im_Amp * CorrCf[f_dlt][2];
        a3 = Re_Amp * CorrCf[f_dlt][3];
        b3 = Im_Amp * CorrCf[f_dlt][3];
        a4 = Re_Amp * CorrCf[f_dlt][4];
        b4 = Im_Amp * CorrCf[f_dlt][4];
```

```
Re[i][abs(xFrec – 2)] -=    a4;
Im[i][abs(xFrec – 2)] -=    b4;
Re[i][abs(xFrec – 1)] +=    (a3-a0);
Im[i][abs(xFrec – 1)] +=    (b3-b0);

        Re[i][xFrec ]       +=    Re_Amp – a2 - a3;
        Im[i][xFrec ]       +=    Im_Amp – b2 - b3;
        Re[i][xFrec + 1]    +=    a1 + a4 - Re_Amp;
        Im[i][xFrec + 1]    +=    b1 + b4 - Im_Amp;
        Re[i][xFrec + 2]    +=    a0 - a1;
        Re[i][xFrec + 3]    +=    a2;
        Im[i][xFrec + 3]    +=    a2;
}
```

Amplitud, *Frec*, *Fase*, *Envolvente[Grupo][i]*, *Grupo*, y *Longitud* son todos como se definen anteriormente.
CorrCf se proporciona mediante la Tabla 2 (Figura 20).
abs(val) es una función que devuelve el valor absoluto de val

Ya que el flujo de bits no contiene ninguna información como el número de componentes tonales codificados, el decodificador solo lee datos de tono para cada tamaño de transformada hasta que se agota de datos para ese tamaño. Por lo tanto, los componentes tonales eliminados del flujo de bits mediante medios externos, no tienen ningún efecto sobre la capacidad del decodificador para tratar datos aún contenidos en el flujo de bits. Eliminar elementos del flujo de bits solo degrada calidad de audio mediante la cantidad del componente de datos eliminado. También pueden eliminarse fragmentos tonales, en cuyo caso el decodificador no realiza ningún trabajo de reconstrucción de componentes tonales para ese tamaño de transformada.

Transformada 604 de frecuencia inversa

La transformada 604 de frecuencia inversa es la inversa de la transformada usada para crear el dominio de la representación de frecuencia en el codificador. La realización actual emplea el banco de filtros jerárquico inverso descrito anteriormente. Como alternativa, una Transformada rápida de Fourier inversa que es la inversa de la FFT más pequeña usada para extraer tonos por el codificador proporcionado FFT superpuestas se usó en tiempo de codificación.

Decodificador 602 residual

En la Figura 18 se muestra un diagrama de bloques detallado del decodificador 602 residual. El analizador 600 de flujo de datos pasa elementos de G1 desde el flujo de bits al decodificador 702 de cuadrícula en línea 610. El decodificador 702 de cuadrícula decodifica G1 para recrear G0 que es 32 subbandas de frecuencia por 64 intervalos de tiempo. El flujo de bits contiene valores de G1 cuantificados y las distancias entre esos valores. Valores de G1 del flujo de bits se decuantifican usando la misma tabla de decuantificación como se usa para de cuantificar amplitudes de componente tonal. Interpolación lineal entre los valores del flujo de bits conduce a 8 amplitudes de G1 finales para cada subbanda de G1. Subbandas 0 y 1 de G1 se inicializan a cero, reemplazándose los valores cero cuando se encuentra información de subbanda para estas dos subbandas en el flujo de bits. Estas amplitudes se ponderan a continuación en la cuadrícula G0 recreada usando las ponderaciones 1900 de correlación obtenidos de la Tabla 1 en la Figura 19. Una fórmula general G0 se proporciona mediante:

$$G0_{m,n} = \sum_{k=0}^{10} \left(W_{m,k} \cdot G1_{k,\lfloor n/8 \rfloor} \right)$$

en la que:

m es el número de subbandas
W es la entrada de la Tabla 1
n es el número de columnas de G0
k se extiende a través de 11 subbandas de G1

Decuantificador 700

Muestras de tiempo encontradas por el analizador 600 de flujo de datos se decuantifican en el decuantificador 700. El decuantificador 700 decuantifica muestras de tiempo del flujo de bits usando el procedimiento inverso del codificador. Muestras de tiempo de subbanda cero se decuantifican a 16 niveles, subbandas 1 y 2 a 8 niveles, subbandas 11 a 25 a tres niveles y subbandas 26 a 31 a 2 niveles. Cualquier muestra de tiempo faltante o no válida se sustituyen con una secuencia pseudoaleatoria de valores en el intervalo de -1 a 1 que tiene una distribución de energía espectral de ruido blanco. Esto mejora la calidad de audio de flujo de bits escalado ya que una secuencia de valores de este tipo tiene características que se parecen más estrechamente a la señal original que la sustitución con valores cero.

Demultiplexor 701 de canal

Información de canal secundario en el flujo de bits se almacena como la diferencia del canal primario para algunas subbandas, dependiendo de banderas establecidas en el flujo de bits. Para estas subbandas, el demultiplexor 701 de canal restaura valores en el canal secundario a partir de los valores en el canal primario y valores de diferencia en el flujo de bits. Si a la información de canal secundario le falta el flujo de bits, información de canal secundario puede recuperarse difícilmente del canal primario duplicando la información de canal primario en canales secundarios y usando la cuadrícula estéreo, a analizar posteriormente.

Reconstrucción 706 de canal

Reconstrucción 706 estéreo se aplica a canales secundarios cuando no se encuentra ninguna información de canal secundario (muestras de tiempo) en el flujo de bits. La cuadrícula estéreo, reconstruida por el decodificador 702 de cuadrícula, se aplica a las muestras de tiempo secundarias, recuperadas duplicando la información de muestra de tiempo de canal primario, para mantener la relación de potencia estéreo original entre canales.

Reconstrucción multicanal

Reconstrucción 706 multicanal se aplica a canales secundarios cuando ninguna información secundaria (ya sean muestras de tiempo o cuadrículas) para los canales secundarios está presente en el flujo de bits. El procedimiento es similar a reconstrucción 706 estéreo, excepto que la cuadrícula parcial reconstruida por el decodificador 702 de cuadrícula se aplica a las muestras de tiempo del canal secundario dentro de cada grupo de canales, recuperadas duplicando información de muestra de tiempo de canal primario para mantener el nivel de potencia apropiado en el canal secundario. La cuadrícula parcial se aplica individualmente a cada canal secundario en el grupo de canales reconstruido tras escalado por otra cuadrícula o cuadrículas de factor de escala que incluyen la cuadrícula G0 en la etapa 703 de escalado multiplicando muestras de tiempo de la cuadrícula G por correspondientes elementos de la cuadrícula parcial para cada canal secundario. La cuadrícula G0, cuadrículas parciales pueden aplicarse en cualquier orden de acuerdo con la presente invención.

Mientras varias realizaciones ilustrativas de la invención se han mostrado y descrito, numerosas variaciones y realizaciones alternativas se ocurrirán a los expertos en la materia. Tales variaciones y realizaciones alternativas se contemplan y pueden hacerse sin alejarse del ámbito de la invención como se define en las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un procedimiento de reconstrucción de una señal de audio de salida de dominio de tiempo a partir de un flujo de bits codificado, que comprende:

recibir un flujo **(599)** de bits escalado que tiene una tasa de datos predeterminada dentro de un intervalo dado como una secuencia de tramas, conteniendo cada trama al menos uno de los siguientes y conteniendo al menos algunas de las tramas todos los siguientes (a) una pluralidad de componentes **(2407)** tonales cuantificados que representan contenidos de dominio de frecuencia a diferentes resoluciones de frecuencia de una señal de entrada, b) componentes **(2403)** de muestra de tiempo residuales cuantificados que representan un residual de dominio de tiempo formado a partir de la diferencia entre componentes tonales reconstruidos y una señal de entrada y c) cuadrículas de factor de escala **(2404)** que representan energías de señal de una señal residual formada a partir de la diferencia entre componentes tonales reconstruidos y una señal de entrada, en el que las cuadrículas **(2404)** de factor de escala se extienden al menos parcialmente en un intervalo de frecuencia de la señal de entrada;
recibir información **(599)** para cada trama acerca de la posición de los componentes cuantificados y/o cuadrículas dentro del intervalo de frecuencia;
analizar las tramas del flujo de bits escalado en los componentes y cuadrículas **(600)**;
decodificar cualquier componente tonal para formar coeficientes **(2408)** de transformada;
decodificar cualquier componente de muestra de tiempo y cualquier cuadrícula **(2401-2405)**;
multiplicar los componentes de muestra de tiempo por elementos de cuadrícula para formar muestras **(2406)** de dominio tiempo; y
aplicar un banco **(2400)** de filtros jerárquico inverso a los coeficientes **(2407)** de transformada y muestras **(4002)** de dominio tiempo para reconstruir una señal **(614)** de audio de salida de dominio de tiempo.

2. El procedimiento de la reivindicación 1, en el que las muestras de dominio tiempo se forman mediante,
el análisis del flujo de bits en una cuadrícula **(2404)** de factor de escala G1 y los componentes **(2403)** de muestra de tiempo;
la decodificación y cuantificación inversa de la cuadrícula de factor de escala de cuadrícula G1 para producir una cuadrícula **(2405)** de factor de escala G0; y
la decodificación y cuantificación inversa de los componentes de muestra de tiempo, multiplicando esos valores de muestra de tiempo por valores **(2406)** de cuadrícula de factor de escala G0 para producir muestra **(4002)** de tiempo reconstruida.

3. El procedimiento de la reivindicación 2, en el que la señal es una señal multicanal en la que los canales residuales se han agrupado y codificado, conteniendo también cada una de dichas tramas d) cuadrículas parciales que representan las relaciones de energía de señal de los canales de señal residuales dentro de grupos de canales comprendiendo además:

analizar el flujo de bits en las cuadrículas **(508)** parciales;
decodificar y cuantificar **(2401)** inversamente las cuadrículas parciales; y
multiplicar las muestras de tiempo reconstruidas por la cuadrícula **(508)** parcial aplicada a cada canal secundario en un grupo de canales para producir las muestras de dominio tiempo reconstruidas.

4. El procedimiento de la reivindicación 1, en el que la señal de entrada es multicanal en la que grupos de componentes tonales que contienen un canal primario y uno o más canales secundarios, conteniendo también cada una de dichas tramas e) una máscara de bits asociada con el canal primario en cada grupo en el que cada bit identifica la presencia de un canal secundario que se ha codificado conjuntamente con el canal primario,
analizar el flujo de bits en las máscaras **(3602)** de bits;
decodificar los componentes tonales para el canal primario en cada grupo **(601)**;
decodificar los componentes tonales conjuntamente codificados en cada grupo **(601)**;
para cada grupo, usar la máscara de bits para reconstruir los componentes tonales para cada uno de dichos canales secundarios a partir de los componente tonales de canal primario y los componentes **(601)** tonales conjuntamente codificados.

5. El procedimiento de la reivindicación 4, en el que los componentes tonales de canal secundario se decodifican decodificando la información de diferencia entre las frecuencias primarias y secundarias, decodificándose por entropía y almacenándose amplitudes y fases para cada canal secundario en el que el componente tonal está presente.

6. El procedimiento de la reivindicación 1, en el que el banco **(2400)** de filtros jerárquico inverso reconstruye la señal **(614)** de audio de salida transformando las muestras **(4002)** de dominio tiempo en coeficientes **(2411)** de transformada residuales, combinando **(2412)** los mismos con los coeficientes **(2409)** de transformada para un conjunto de componentes **(2407)** tonales a una resolución de frecuencia baja y transformado inversamente **(2413)** los coeficientes de transformada combinados para formar una señal **(2415)** de audio de salida parcialmente reconstruida, y repitiendo las etapas en esta señal de audio de salida parcialmente reconstruida con los coeficientes de transformada para otro conjunto de componentes tonales a la siguiente resolución de frecuencia más alta hasta

que se reconstruya la señal **(614)** de audio de salida.

7. El procedimiento de la reivindicación 6, en el que las muestras de dominio tiempo se representan como subbandas, reconstruyendo dicho banco de filtros jerárquico inverso la señal de audio de salida de dominio de tiempo mediante:

a) la formación en ventanas de la señal o señales en cada una de las subbandas de dominio de tiempo de la trama de entrada para formar subbandas **(2410)** de tiempo de dominio formadas en ventana;
b) la aplicación de una transformada de dominio de tiempo a frecuencia a cada una de las subbandas de tiempo de dominio formadas en ventana para formar coeficientes **(2411)** de transformada;
c) la concatenación de los coeficientes de transformada resultantes para formar conjunto o conjuntos más grandes de los coeficientes **(2411)** de transformada residuales;
d) sintetizar los coeficientes de transformada del conjunto de componentes **(2409)** tonales;
e) la combinación de los coeficientes de transformada reconstruidos a partir de los componentes tonales y de dominio de tiempo en un único conjunto de coeficientes **(2412)** de transformada combinados;
f) la aplicación de una transformada inversa a los coeficientes **(2413)** de transformada combinados, formando en ventanas y añadiendo solapamiento **(2414)** con la trama anterior para reconstruir una señal de dominio de tiempo parcialmente reconstruida **(2415)**; y
g) la aplicación de iteraciones sucesivas de las etapas (a) a (f) en la señal o señales de dominio de tiempo parcialmente reconstruida(s) usando el siguiente conjunto de componentes **(2407)** tonales hasta que se reconstruya la señal **(614)** de audio de salida de dominio de tiempo.

8. El procedimiento de la reivindicación 6, en el que cada trama de entrada contiene M_i muestras de tiempo en cada una de las P subbandas, realizando dicho banco de filtros jerárquico inverso las siguientes etapas:

a) en cada subbanda i, almacenar en memoria intermedia y concatenar las M_i muestras anteriores con las M_i muestras actuales para producir 2* M_i nuevas muestras **(4004)**;
b) en cada subbanda i, multiplicar las 2* M_i muestras de subbanda por una función de ventana de 2* M_i puntos **(4006)**;
c) aplicar una transformada de (2* M_i) puntos a las muestras de subbanda para producir M_i coeficientes de transformada para cada subbanda i **(4008)**;
d) concatenar los M_i coeficientes de transformada para cada subbanda i para formar un único conjunto de N/2 coeficientes **(4010)**;
e) sintetizar los coeficientes de transformada tonales a partir del conjunto decodificado e inversamente cuantificado de componentes tonales y combinar los mismos con los coeficientes concatenados de la etapa anterior para formar un único conjunto de coeficientes **(2407, 2408, 2409, 2412)** concatenados combinados;
f) aplicar una transformada inversa de N puntos a los coeficientes concatenados combinados para producir N muestras **(4012)**;
g) multiplicar cada trama de N muestras por una función de ventana de N muestras para producir N muestras **(4014)** formadas en ventanas;
h) añadir por superposición las muestras **(4014)** formadas en ventanas resultantes para producir N/2 nuevas muestras de salida en el nivel de subbanda dado como la señal **(4016)** de audio de salida parcialmente reconstruida; y
i) repetir las etapas (a)-(h) en las N/2 nuevas muestras de salida usando el siguiente conjunto de componentes **(2407)** tonales hasta que se hayan procesado todas las subbandas y las N muestras de tiempo originales sean reconstruidas como la señal (614) de audio de salida.

9. Un decodificador para reconstrucción de una señal de audio de salida de dominio de tiempo a partir de un flujo de bits codificado, que comprende:

un analizador **(600)** de flujo de bits para analizar cada trama de un flujo de bits escalado en sus componentes de audio, conteniendo cada trama al menos uno de los siguientes y conteniendo al menos algunas de las tramas todos los siguientes (a) una pluralidad de componentes tonales cuantificados que representan contenidos de dominio de frecuencia a diferentes resoluciones de frecuencia de una señal de entrada, b) componentes de muestra de tiempo residuales cuantificados que representan un residual de dominio de tiempo formado a partir de la diferencia entre componentes tonales reconstruidos y una señal de entrada y c) cuadrículas de factor de escala que representan energías de señal de una señal residual formada a partir de la diferencia entre los componentes tonales reconstruidos y una señal de entrada;
un decodificador **(602)** residual para codificar cualquier componente de muestra de tiempo y cualquier cuadrícula para reconstruir muestras de tiempo;
un decodificador **(601)** tonal para decodificar cualquier componente tonal para formar coeficientes de transformada; y
un banco **(2400)** de filtros jerárquico inverso que reconstruye la señal de salida transformando las muestras de tiempo en coeficientes de transformada residuales, combinando los mismos con los coeficientes de transformada para un conjunto de los componentes tonales a una resolución de frecuencia baja y transformado inversamente los coeficientes de transformada combinados para formar una señal de salida parcialmente reconstruida, y repitiendo las etapas en esta señal de salida parcialmente reconstruida con los coeficientes de transformada para

otro conjunto de componentes tonales a la siguiente resolución de frecuencia más alta hasta que la señal de audio de salida se reconstruya.

10. El decodificador de la reivindicación 9, en el que cada trama de entrada contiene M_i muestras de tiempo en cada una de las P subbandas, realizando dicho banco de filtros jerárquico inverso las siguientes etapas:

a) en cada subbanda i, almacenar en memoria intermedia y concatenar las M_i muestras anteriores con las M_i muestras actuales para producir 2* M_i nuevas muestras **(4004)**;
b) en cada subbanda i, multiplicar las 2* M_i muestras de subbanda por una función **(4006)** de ventana de 2* M_i puntos;
c) aplicar una transformada de (2* M_i) puntos a las muestras de subbanda para producir M_i coeficientes de transformada residuales para cada subbanda i **(4008)**;
d) concatenar los M_i coeficientes de transformada residuales para cada subbanda i para formar un único conjunto de N/2 coeficientes **(4010)**;
e) sintetizar los coeficientes de transformada tonales a partir del conjunto decodificado e inversamente cuantificado de componentes tonales y combinar los mismos con los coeficientes de transformada residuales concatenados para formar un único conjunto de coeficientes **(2407, 2408, 2409, 2412)** concatenados combinados;
f) aplicar una transformada inversa de N puntos a los coeficientes concatenados combinados para producir N muestras **(4012)**;
g) multiplicar cada trama de N muestras por una función de ventana de N muestras para producir N muestras **(4014)** formadas en ventanas;
h) añadir por superposición las muestras **(4014)** formadas en ventanas resultantes para producir N/2 nuevas muestras de salida en el nivel de subbanda dado como la señal **(4016)** de salida parcialmente reconstruida; y
i) repetir las etapas (a)-(h) en las N/2 nuevas muestras de salida usando el siguiente conjunto de componentes **(2407)** tonales hasta que se hayan procesado todas las subbandas y las N muestras de tiempo originales sean reconstruidas como la señal **(614)** de salida.

11. Un procedimiento de codificación de una señal de audio de entrada para formar un flujo **(116)** de bits escalable, que comprende:

usar un banco **(2101a,... 2101e)** de filtros jerárquico (HFB) para descomponer una señal **(100)** de audio de entrada en una representación de tiempo/frecuencia de múltiples resoluciones;
extraer componentes tonales en cada iteración del HFB en múltiples resoluciones de frecuencia de la representación **(2109)** de tiempo/frecuencia;
extraer componentes residuales **(2117, 2118, 2119)** de la representación de tiempo/frecuencia eliminando componentes tonales de la señal de entrada para pasar una señal residual a la siguiente iteración del HFB y extrayendo los componentes residuales **(2117, 2118, 2119**) de la señal residual final;
clasificando los componentes en base a su contribución relativa a la calidad **(103, 107, 109)** de señal decodificada;
cuantificar y codificar los componentes **(102, 107, 108)**;
formar un flujo **(126)** de bits maestro que incluye los componentes **(109)** cuantificados clasificados, y
escalar el flujo **(126)** de bits maestro eliminando un número suficiente de componentes **(115)** codificados de menor clasificación para formar el flujo **(116)** de bits escalado que tiene una tasa de datos menor que o aproximadamente igual a una tasa de datos deseada,

en el que los componentes se clasifican y cuantifican con referencia a la misma función de enmascaramiento o diferentes criterios psicoacústicos, y
en el que el flujo de bits escalado incluye información que indica la posición de los componentes en el espectro de frecuencia.

12. El procedimiento de la reivindicación 11, en el que los componentes se clasifican agrupando primero los componentes tonales en al menos un subdominio **(903, 904, 905, 906, 907)** de frecuencia a diferentes resoluciones de frecuencia y agrupando los componentes residuales en al menos un subdominio **(908, 909, 910)** residual en diferentes escalas de tiempo y/o resoluciones de frecuencia, clasificando los subdominios en base a su contribución relativa a la calidad de señal decodificada y clasificando los componentes dentro de cada subdominio en base a su contribución relativa a la calidad de señal decodificada.

13. El procedimiento de la reivindicación 12, que comprende adicionalmente:

formar el flujo **(126)** de bits maestro en el que los subdominios y componentes dentro de cada subdominio se ordenan en base a su clasificación **(109)**, eliminándose dichos componentes de baja clasificación comenzando con el componente de clasificación más baja en el subdominio de clasificación más baja y eliminando componentes en orden hasta que se consiga **(115)** la tasa de datos deseada .

14. El procedimiento de la reivindicación 11, en el que el flujo **(116)** de bits escalado se graba en o se transmite a través de un canal que tiene la tasa de datos deseada como una restricción.

15. El procedimiento de la reivindicación 14, en el que el flujo **(116)** de bits escalado es uno de los múltiples flujos de bits escalados y la tasa de datos de cada flujo de bits individual se controla independientemente, con la restricción de que la suma de las tasas de datos individuales no debe exceder una tasa de datos total máxima, controlándose dinámicamente cada una de dichas tasas de datos en tiempo de acuerdo con la calidad de señal decodificada a través de todos los flujos de bits.

16. El procedimiento de la reivindicación 11, por el que componentes tonales que se eliminan para formar el flujo de bits escalado también se eliminan **(2112)** de la señal **(2114)** residual.

17. El procedimiento de la reivindicación 11, en el que los componentes residuales incluyen componentes **(2117)** de muestra de tiempo y componentes **(2118, 2119)** de factor de escala que modifican los componentes de muestra de tiempo en diferentes escalas de tiempo y/o resoluciones de frecuencia.

18. El procedimiento de la reivindicación 17, en el que los componentes de muestra de tiempo se representan mediante una cuadrícula G **(2117)** y los componentes de factor de escala comprenden una serie de una o más cuadrículas G0, G1 **(2118, 2119)** en múltiples escalas de tiempo y resoluciones de frecuencia que se aplican a los componentes de muestra de tiempo dividiendo la cuadrícula G por elementos de cuadrícula de G0, G1 en el plano de tiempo/frecuencia, teniendo cada cuadrícula G0, G1 un número diferente de factores de escala en tiempo y/o frecuencia.

19. El procedimiento de la reivindicación 17, en el que los factores de escala se codifican **(107)** aplicando una transformación bidimensional a los componentes de factor de escala y cuantificando los coeficientes de transformada.

20. El procedimiento de la reivindicación 19, en el que la transformada es una Transformada de Coseno Discreta bidimensional.

21. Un procedimiento de la reivindicación 11, en el que el HFB descompone la señal de audio de entrada en coeficientes de transformada en niveles de resolución de frecuencia sucesivamente más bajos en iteraciones sucesivas, en el que dichos componentes tonales y residuales se extraen mediante:

la extracción de componentes **(2109)** tonales de los coeficientes de transformada en cada iteración, cuantificando **(2110)** y almacenando los componentes tonales extraídos en una lista **(2106)** de tonos;
la eliminación de los componentes **(2111, 2112)** tonales de la señal de audio de entrada para pasar una señal **(2114)** residual a la siguiente iteración del HFB; y
la aplicación de una transformada **(2115)** inversa final con resolución de frecuencia relativamente menor que la iteración final del HFB a la señal **(113)** residual para extraer los componentes **(2117)** residuales.

22. El procedimiento de la reivindicación 21, que comprende adicionalmente:

eliminar algunos de los componentes **(114)** tonales de la lista de tonos después de la iteración final; y
decodificar y cuantificar **(104)** inversamente localmente los componentes **(114)** tonales cuantificados eliminados, y combinar **(105)** los mismos con la señal **(111)** residual en la iteración final.

23. El procedimiento de la reivindicación 22, en el que al menos algunos de los componentes tonales relativamente fuertes eliminados de la lista no se decodifican localmente ni se recombinan.

24. El procedimiento de la reivindicación 21, en el que los componentes tonales en cada resolución de frecuencia se extraen **(2109)** mediante:

la identificación de los componentes tonales deseados a través de aplicación de un modelo perceptual;
la selección del más perceptualmente significativo de los coeficientes de transformada;
el almacenamiento de parámetros de cada coeficiente de transformada seleccionado como el componente tonal, incluyendo dichos parámetros la amplitud, frecuencia, fase y posición en la trama del correspondiente coeficiente de transformada; y
la cuantificación y codificación **(2110)** de los parámetros para cada componente tonal en la lista de tonos para inserción en el flujo de bits.

25. El procedimiento de la reivindicación 21, en el que los componentes residuales incluyen componentes de muestra de tiempo representados como una cuadrícula G **(2117)**, la extracción de los componentes residuales comprende además:

construir una o más cuadrículas **(2118, 2119)** de factor de escala de diferentes resoluciones de tiempo/frecuencia, cuyos elementos representan valores de señal máximos o energías de señales en una región de tiempo/frecuencia;
dividir los elementos de cuadrícula G de muestra de tiempo mediante correspondientes elementos de las cuadrículas de factor de escala para producir una cuadrícula G **(2120)** de muestra de tiempo escalada; y
cuantificar y codificar la cuadrícula G **(2122)** de muestra de tiempo escalada y cuadrículas **(2121)** de factor de

escala para inserción en el flujo de bits codificado.

26. El procedimiento de la reivindicación 11, en el que la señal de audio de entrada se descompone y los componentes tonales y residuales se extraen mediante,

(a) el almacenamiento en memoria intermedia de muestras de la señal de audio de entrada en tramas de N muestras **(2900)**;
(b) la multiplicación de las N muestras en cada trama por una función de ventana de N muestras **(2900)**;
(c) la aplicación de una transformada de N puntos para producir N/2 coeficientes de transformada originales **(2902)**;
(d) la extracción de componentes tonales de los N/2 coeficientes **(2109)** de transformada originales, cuantificando **(2110)** y almacenando los componentes tonales extraídos en una lista **(2106)** de tonos;
(e) la resta de los componentes tonales cuantificando inversamente **(2111)** y restando los coeficientes de transformada tonales resultantes de los coeficientes de transformada originales **(2112)** para proporcionar N/2 coeficientes de transformada residuales;
(f) la división de los N/2 coeficientes de transformada residuales en P grupos de M_i coeficientes **(2906)**, de tal forma que la suma de los M_i coeficientes es N/2 ($\sum_{i=1}^{P} M_i = N/2$;)
(g) para cada uno de los P grupos, la aplicación de una transformada inversa de (2* M_i) puntos a los coeficientes de transformada residuales para producir (2* M_i) muestras de subbanda desde cada grupo **(2906)**;
(h) en cada subbanda, la multiplicación de las 2* M_i muestras de subbanda por una función de ventana de 2* M_i puntos **(2908)**;
(i) en cada subbanda, la superposición con M_i muestras anteriores y la adición de los valores correspondientes para producir M_i nuevas muestras para cada subbanda **(2910)**;
(j) la repetición de las etapas (a)-(i) en una o más de las subbandas de M_i nuevas muestras usando tamaños N de transformada sucesivamente más pequeños **(2912)** hasta que la resolución de tiempo/transformada deseada se logre **(29014)**; y
(k) la aplicación de una transformada **(2115)** inversa final con resolución N de frecuencia relativamente menor a las M_i nuevas muestras para cada salida de subbanda en la iteración final para producir subbandas de muestras de tiempo en una cuadrícula G de subbandas y múltiples muestras de tiempo en cada subbanda.

27. El procedimiento de la reivindicación 11, en el que la señal de audio de entrada es una señal de audio de entrada multicanal, codificándose juntos cada uno de dichos componentes tonales formando grupos de dichos canales y para cada uno de dichos grupos,
seleccionar un canal primario y al menos un canal secundario, que se identifican a través de una máscara **(3602)** de bits, identificando cada bit la presencia de un canal secundario,
cuantificar y codificar el canal primario **(102, 108)**; y
cuantificar y codificar la diferencia entre el canal primario y cada canal secundario **(102, 108)**.

28. El procedimiento de la reivindicación 27, en el que se selecciona un modo de canal conjunto de codificación de cada grupo de canales en base a una métrica que indica qué modo proporciona la menor distorsión percibida para la tasa de datos deseada en la señal de salida decodificada.

29. El procedimiento de la reivindicación 11, en el que la señal de audio de entrada es una señal multicanal, comprendiendo adicionalmente:

restar los componentes tonales extraídos de la señal de audio de entrada para cada canal para formar señales **(2109a,...2109e)** residuales;
formar los canales de la señal residual en grupos determinados por criterios perceptuales y eficiencia de codificación **(3702)**;
determinar canales primarios y secundarios para cada grupo de señal residual **(3704)**;
calcular una cuadrícula **(508)** parcial para codificar información espacial relativa entre cada canal primario/secundario que se emparejan en cada grupo **(502)** de señales residuales;
cuantificar y codificar componentes residuales para el canal primario en cada grupo como respectivas cuadrículas G **(2110a)**;
cuantificar y codificar la cuadrícula parcial para reducir la tasa **(2110a)** de datos requerida; e
insertar la cuadrícula parcial codificada y la cuadrícula G para cada grupo en el flujo de bits escalado **(3706)**.

30. El procedimiento de la reivindicación 29, en el que los canales secundarios se construyen a partir de combinaciones lineales de uno o más canales **(3704)**.

31. Un codificador de flujo de bits escalable para codificar una señal de audio de entrada y formar un flujo de bits escalable, que comprende:

un banco **(2100)** de filtros jerárquicos (HFB) que descompone la señal de audio de entrada en coeficientes **(2108)** de transformada en niveles de resolución de frecuencia sucesivamente más bajos y de vuelta en muestras **(2114)** de subbanda de dominio de tiempo en escalas de tiempo sucesivamente más finas en

iteraciones sucesivas;

un codificador **(102)** de tono que (a) extrae componentes **(2109)** tonales de los coeficientes de transformada en cada iteración, cuantifica **(2110)** y almacena los mismos en una lista **(2106)** de tonos, (b) elimina los componentes **(2111, 2112)** tonales de la señal de audio de entrada para pasar una señal **(2114b)** residual a la siguiente iteración del HFB y (c) clasifica todos los componentes tonales extraídos en base a su contribución relativa a calidad de señal decodificada;
un codificador **(107)** residual que aplica una transformada **(2115)** inversa final con resolución de frecuencia relativamente menor que la iteración final del HFB **(2101e)** a la señal **(113)** residual final para extraer los componentes residuales **(2117, 2118, 2119)** y clasifica los componentes residuales en base a su contribución relativa a calidad de señal decodificada;
en el que los componentes tonales extraídos y componentes residuales se clasifican y cuantifican con referencia a la misma función de enmascaramiento o diferentes criterios psicoacústicos;
un formateador **(109)** de flujo de bits que ensambla los componentes tonales y residuales en una base de trama por trama para formar un flujo **(126)** de bits maestro; y
un escalador **(115)** que elimina un número suficiente de los componentes codificados de menor clasificación de cada trama del flujo de bits maestro para formar un flujo **(116)** de bits escalado que tiene una tasa de datos menor que o aproximadamente igual a una tasa de datos deseada,

en el que el flujo de bits escalado incluye información que indica la posición de los componentes en el espectro de frecuencia.

32. El codificador de la reivindicación 31, en el que el codificador de tono agrupa los componentes tonales en subdominios de frecuencia a diferentes resoluciones **(903, 904, 905, 906, 907)** de frecuencia y clasifica los componentes con cada subdominio, el codificador residual agrupa los componentes residuales en subdominios residuales en diferentes escalas de tiempo y/o resoluciones **(908, 909, 910)** de frecuencia y clasifica los componentes con cada subdominio, y dicho formateador de flujo de bits clasifica los subdominios en base a su contribución relativa a calidad de señal decodificada.

33. El codificador de la reivindicación 32, en el que el formateador de flujo de bits ordena los subdominios y los componentes dentro de cada subdominio en base a su clasificación, eliminando dicho escalador **(115)** dichos componentes de baja clasificación comenzando con el componente de clasificación más baja en el subdominio de clasificación más baja y eliminando componentes en orden hasta que se consiga la tasa de datos deseada.

34. El codificador de la reivindicación 31, en el que la señal de audio de entrada es una señal de audio de entrada multicanal, codificando conjuntamente dicho codificador de tono cada uno de dichos componentes tonales formando grupos de dichos canales y para cada uno de dichos grupos,
seleccionando un canal primario y al menos un canal secundario, que se identifican a través de una máscara **(3602)** de bits, identificando cada bit la presencia de un canal secundario;
cuantificando y codificando el canal primario **(102, 108)**; y
cuantificando y codificando la diferencia entre el canal primario y cada canal secundario **(102, 108)**.

35. El codificador de la reivindicación 31, en el que la señal de entrada es una señal de audio multicanal, dicho codificador residual,
formando los canales de la señal residual en grupos determinados por criterios perceptuales y eficiencia de codificación **(3702)**;
determinando canales primarios y secundarios para cada uno de dichos grupos de señales residuales **(3704)**;
calculando una cuadrícula **(508)** parcial para codificar información espacial relativa entre cada canal primario/secundario que se emparejan en cada grupo **(502)** de señales residuales;
cuantificando y codificando componentes residuales para el canal primario en cada grupo como respectivas cuadrículas G **(2110a)**;
cuantificando y codificando la cuadrícula parcial para reducir la tasa **(2110a)** de datos requerida; e
insertando la cuadrícula parcial codificada y la cuadrícula G para cada grupo en el flujo de bits escalado **(3706)**.

36. El codificador de la reivindicación 31, en el que el codificador residual extrae componentes de muestra de tiempo representados por una cuadrícula G **(2117)** y una serie de una o más cuadrículas G0, G1 **(2118, 2119)** de factor de escala en múltiples resoluciones de tiempo y frecuencia que se aplican a los componentes de muestra de tiempo dividiendo la cuadrícula G por elementos de cuadrícula de G0, G1 en el plano **(2120)** de tiempo/frecuencia, teniendo cada cuadrícula G0, G1 un número diferente de factores de escala en tiempo y/o frecuencia.

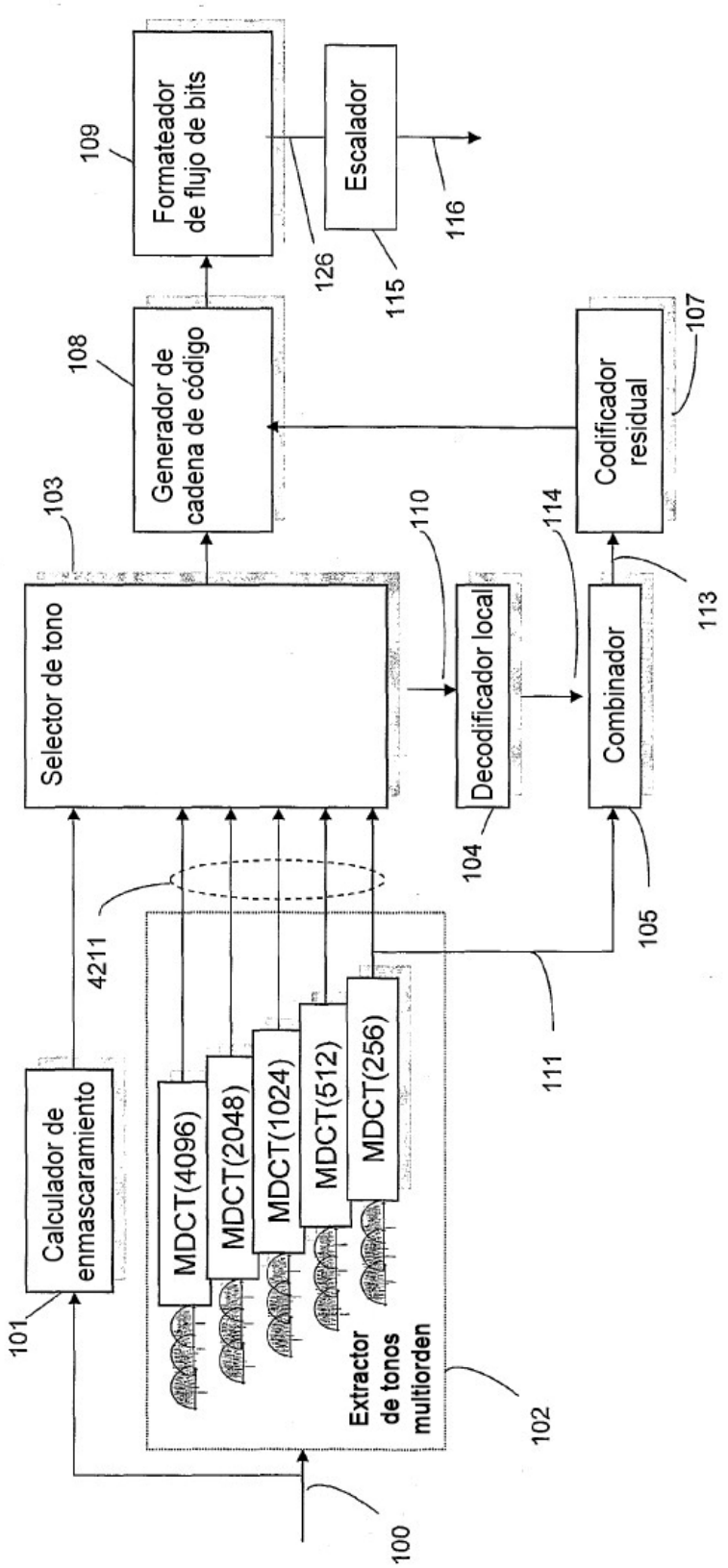

Formateador de flujo de bits
109
Escalador
116
126
115
Generador de cadena de código
108
Codificador residual
107
103
110
114
Selector de tono
Decodificador local
Combinador
113
104
105
4211
111
Calculador de enmascaramiento
MDCT(4096)
MDCT(2048)
MDCT(1024)
MDCT(512)
MDCT(256)
101
Extractor de tonos multiorden
102
100

Fig. 1

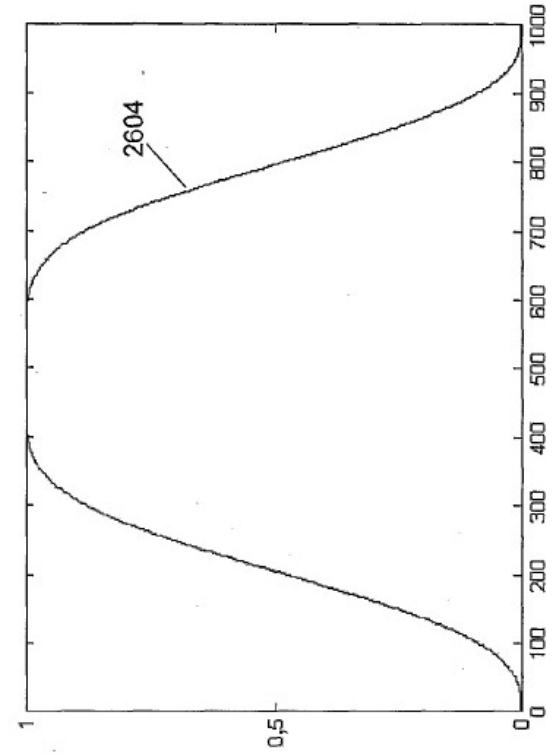

2604
1
0,5
0
0
100
200
300
400
500
600
700
800
900
1000

Fig. 2b

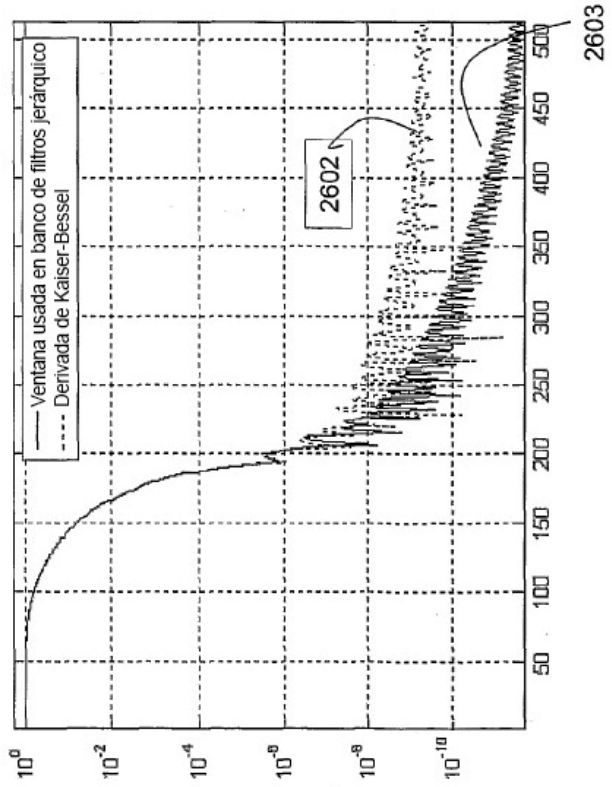

Ventana usada en banco de filtros jerárquico
Derivada de Kaiser-Bessel
2602
2603
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
50
100
150
200
250
300
350
400
450
500

Fig. 2a

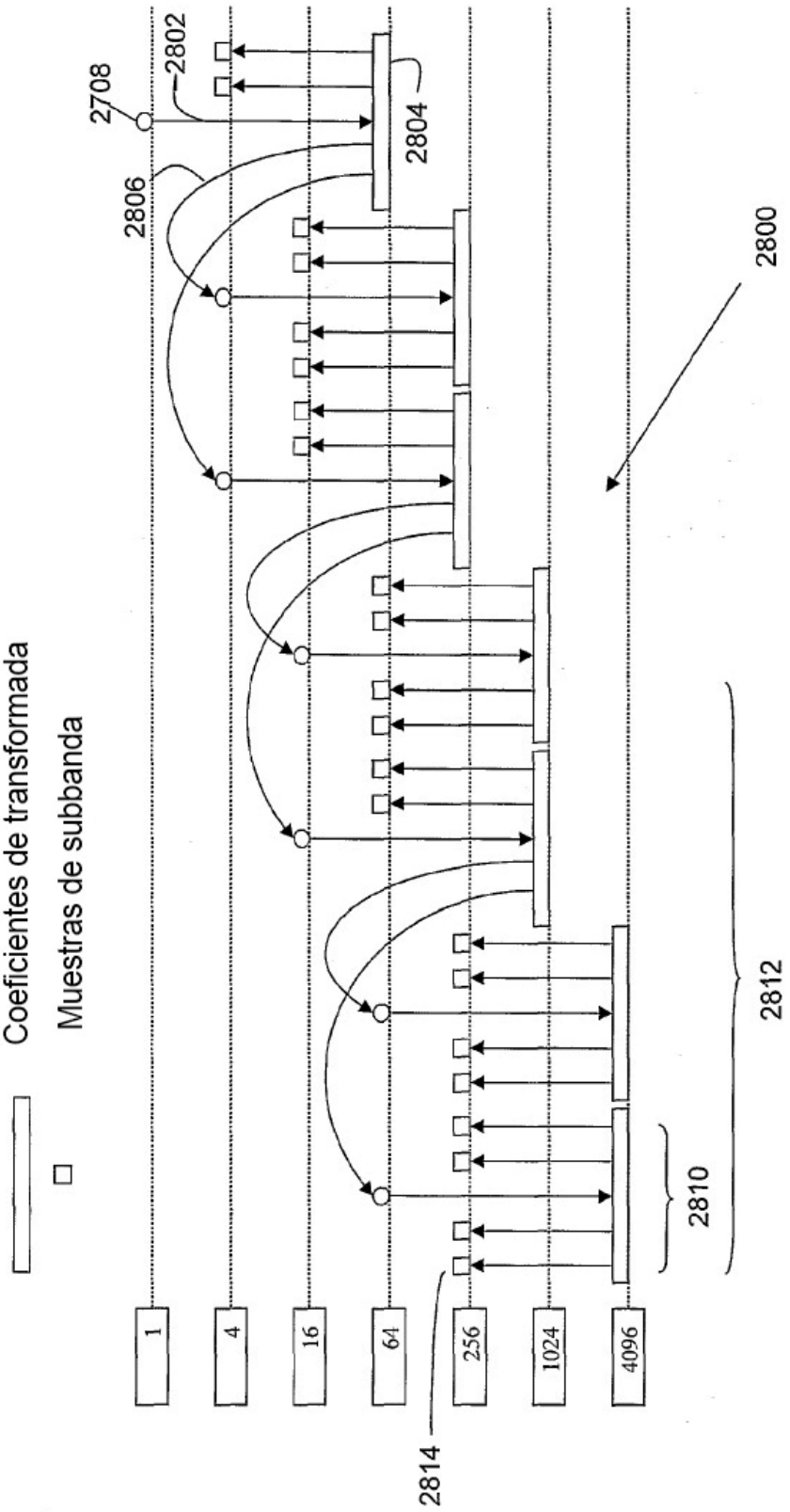

Coeficientes de transformada
Muestras de subbanda
2708
2802
2804
2806
2800
2812
2810
2814
1
4
16
64
256
1024
4096

Fig. 3

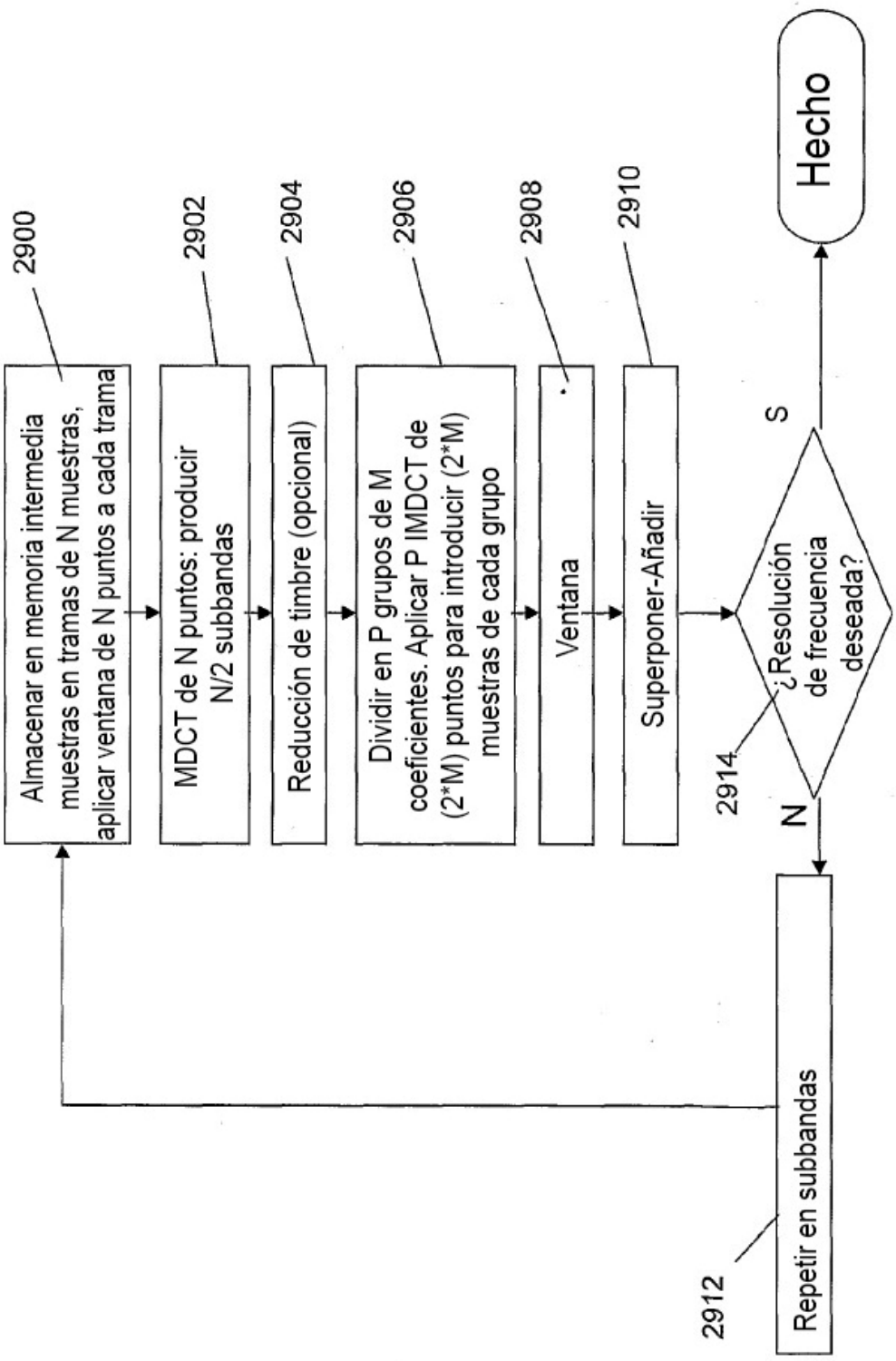

2900
Almacenar en memoria intermedia muestras en tramas de N muestras, aplicar ventana de N puntos a cada trama
2902
MDCT de N puntos: producir N/2 subbandas
2904
Reducción de timbre (opcional)
2906
Dividir en P grupos de M coeficientes. Aplicar P IMDCT de (2*M) puntos para introducir (2*M) muestras de cada grupo
2908
Ventana
2910
Superponer-Añadir
2914
¿Resolución de frecuencia deseada?
S
Hecho
N
2912
Repetir en subbandas

Fig. 4

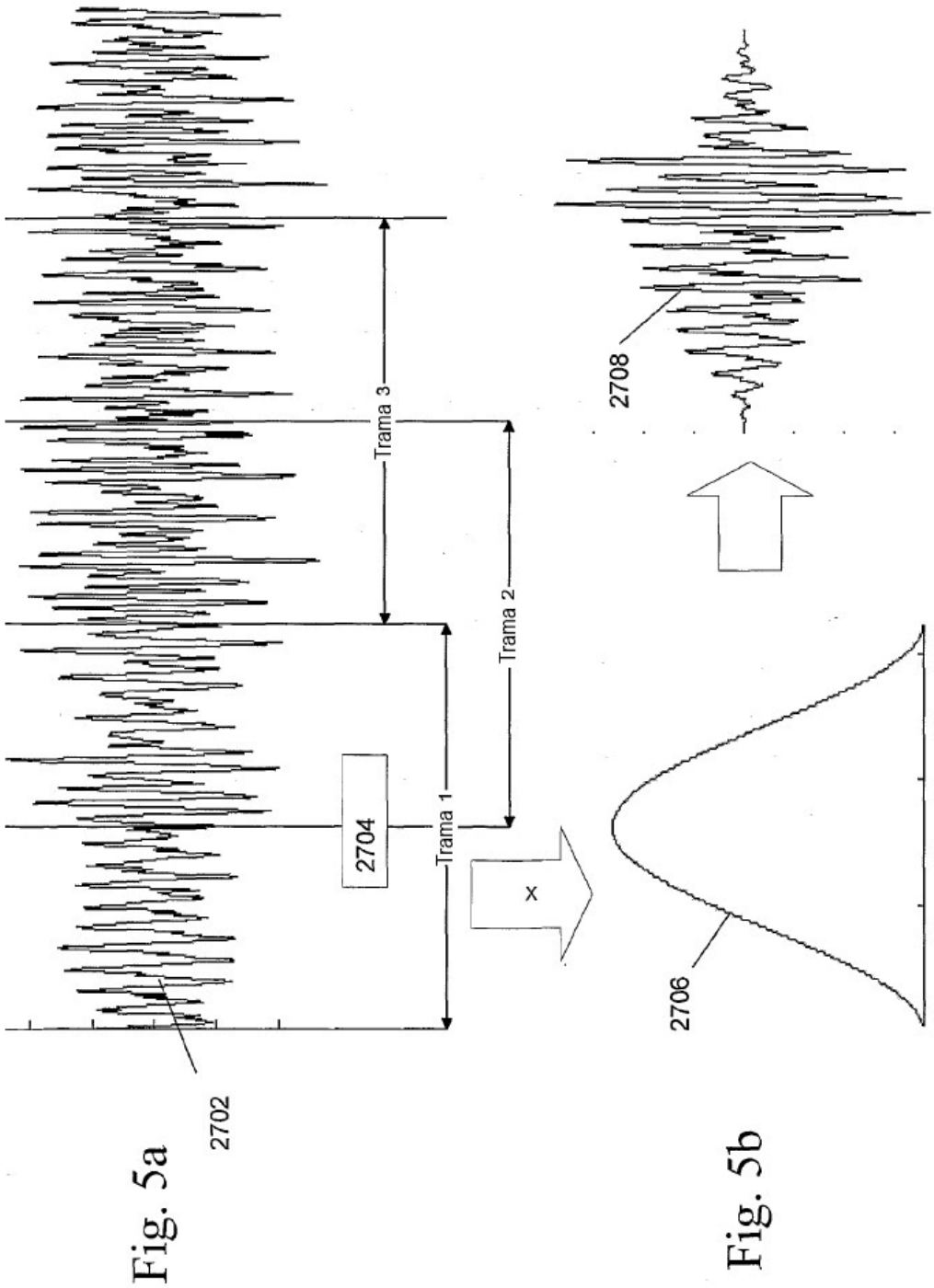


Fig. 5a

Fig. 5b

Fig. 5c

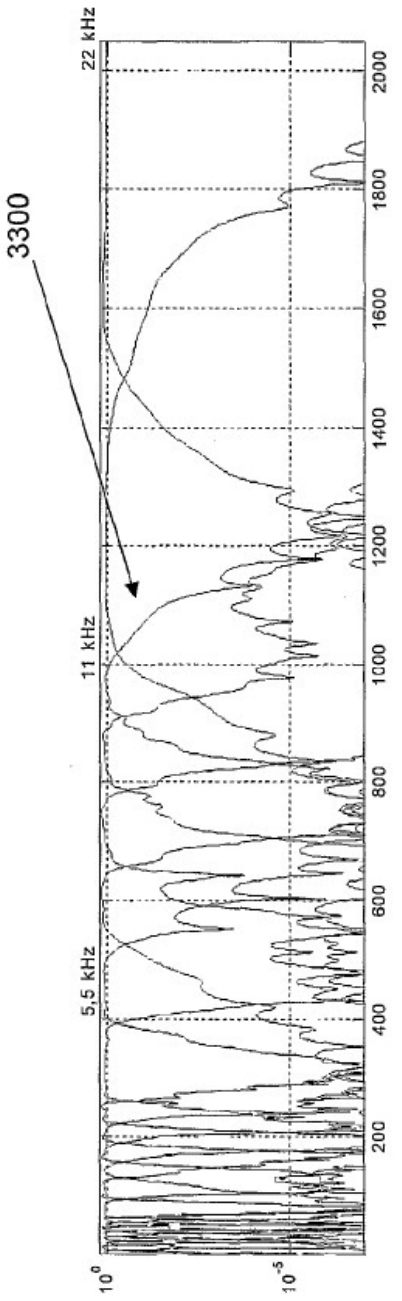

3300
22 kHz
11 kHz
5,5 kHz
2000
1800
1600
1400
1200
1000
800
600
400
200
10^{0}
10^{-5}

Fig. 6

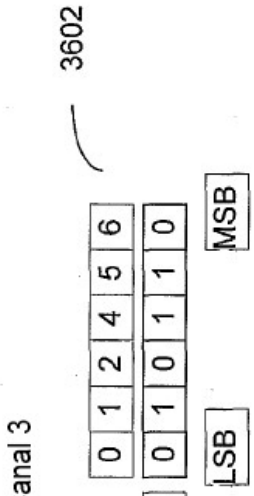

3602
Primario: Canal 3
Canal
Máscara de bits
0 1 2 4 5 6
0 1 0 1 1 0
LSB
MSB

Fig. 9

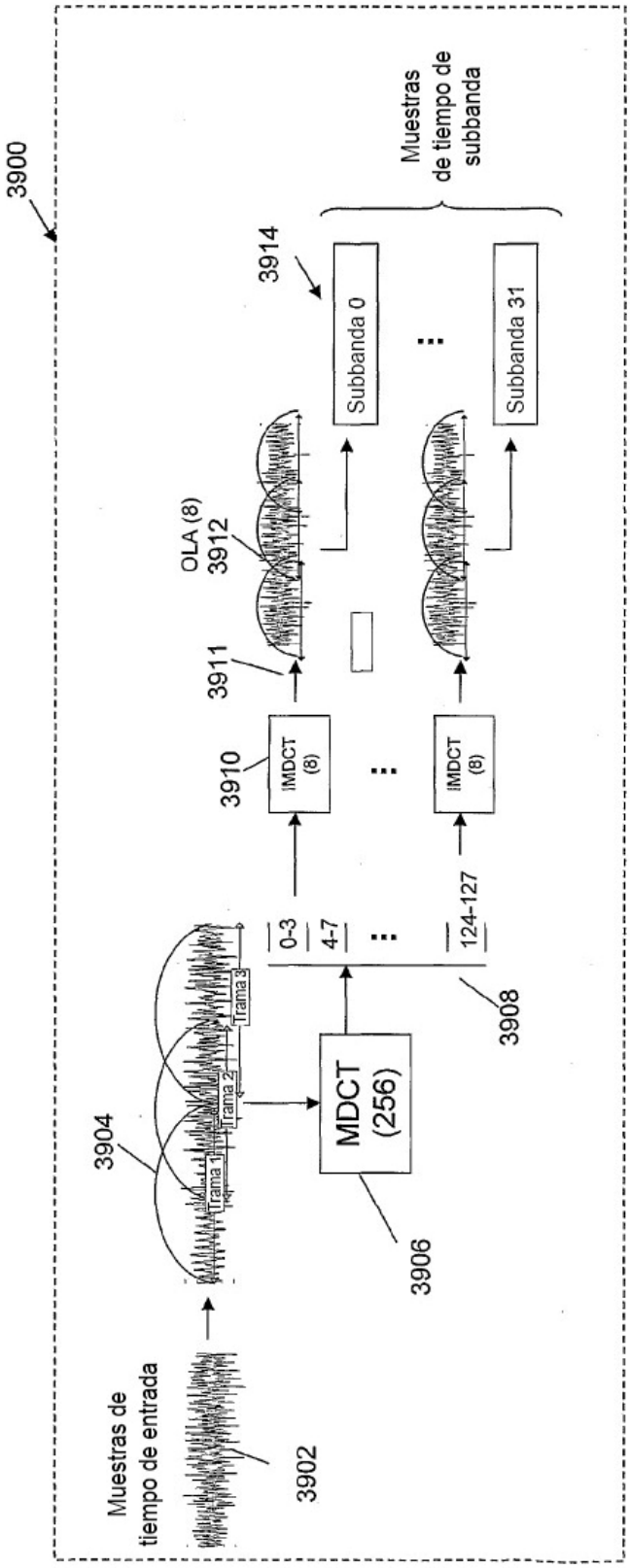

3900
Muestras de tiempo de entrada
3902
3904
Trama 1
Trama 2
Trama 3
MDCT (256)
3906
0-3
4-7
124-127
3908
IMDCT (8)
3910
3911
OLA (8)
3912
3914
Subbanda 0
Subbanda 31
Muestras de tiempo de subbanda

Fig. 7

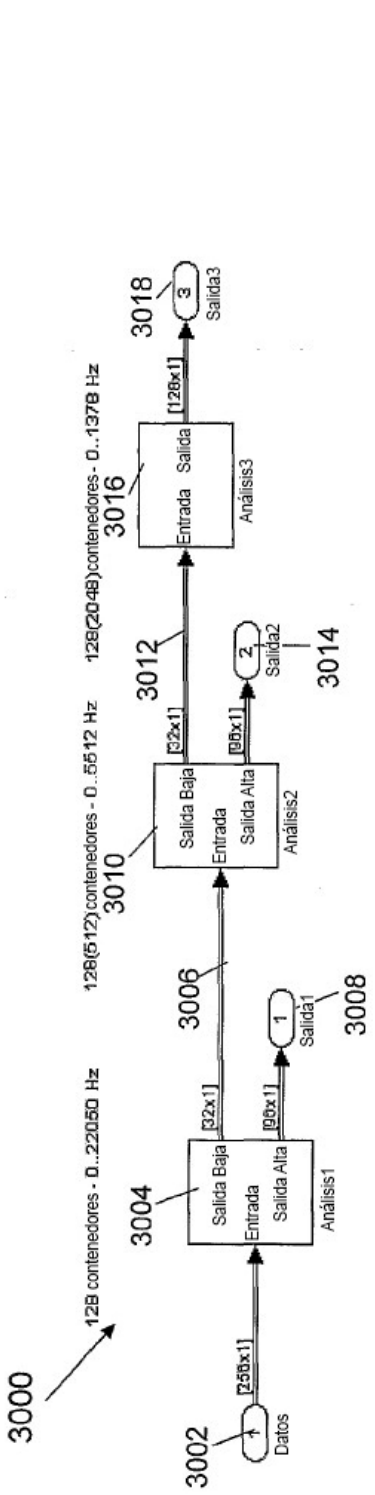

3000
128 contenedores - 0..22050 Hz
3002
Datos
[256x1]
3004
Salida Baja
Entrada
Salida Alta
Análisis1
[32x1]
[96x1]
3006
1
Salida1
3008
128(512)contenedores - 0..5512 Hz
3010
Salida Baja
Entrada
Salida Alta
Análisis2
[32x1]
[96x1]
3012
2
Salida2
3014
128(2048)contenedores - 0..11378 Hz
3016
Entrada
Salida
Análisis3
[128x1]
3018
3
Salida3

Fig. 8a

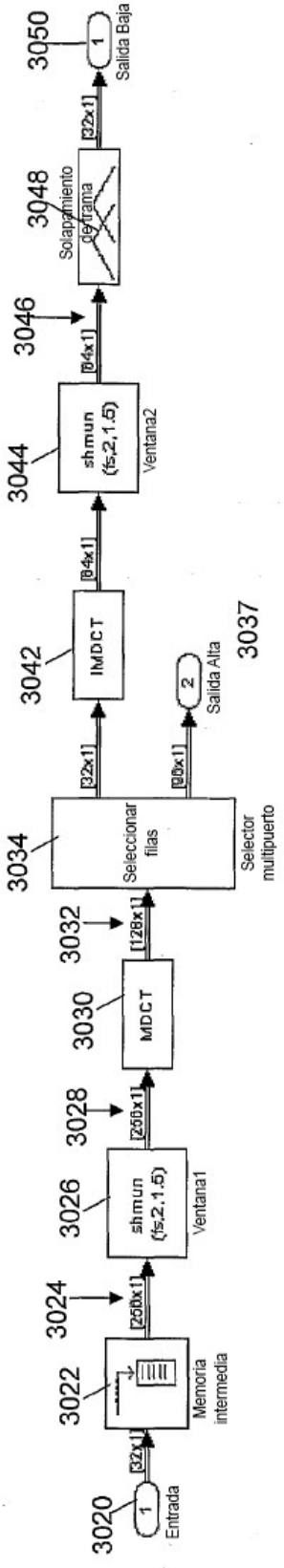

3020
1
Entrada
[32x1]
3022
Memoria intermedia
[256x1]
3024
3026
shmun (fs,2,1.5)
Ventana1
[256x1]
3028
3030
MDCT
[128x1]
3032
3034
Seleccionar filas
Selector multipuerto
[32x1]
[96x1]
2
Salida Alta
3037
3042
IMDCT
[64x1]
3044
shmun (fs,2,1.5)
Ventana2
[64x1]
3046
3048
Solapamiento diorama
[32x1]
3050
1
Salida Baja

Fig. 8b

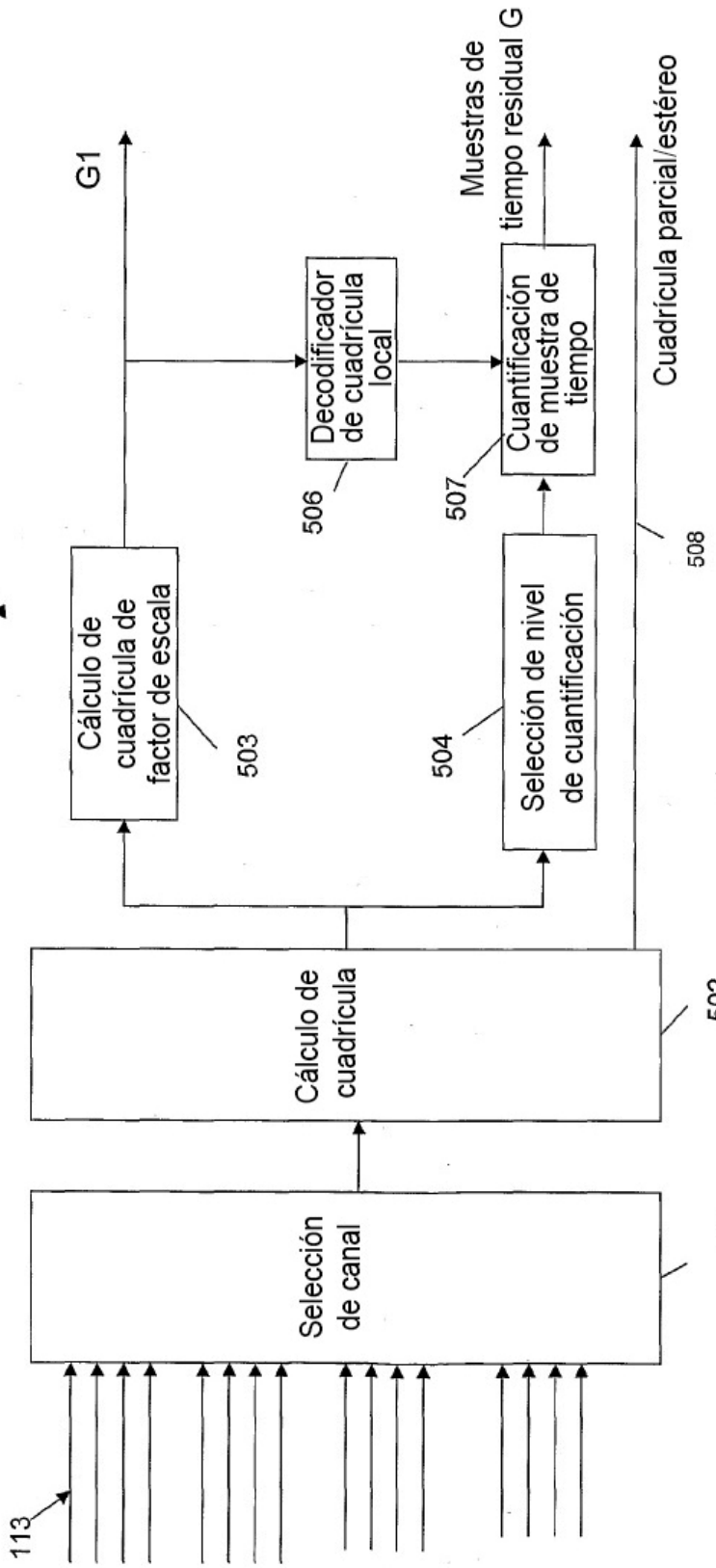

107
113
Selección de canal
501
Cálculo de cuadrícula
502
Cálculo de cuadrícula de factor de escala
503
Selección de nivel de cuantificación
504
Decodificador de cuadrícula local
506
Cuantificación de muestra de tiempo
507
Cuadrícula parcial/estéreo
508
G1
Muestras de tiempo residual G

Fig. 10

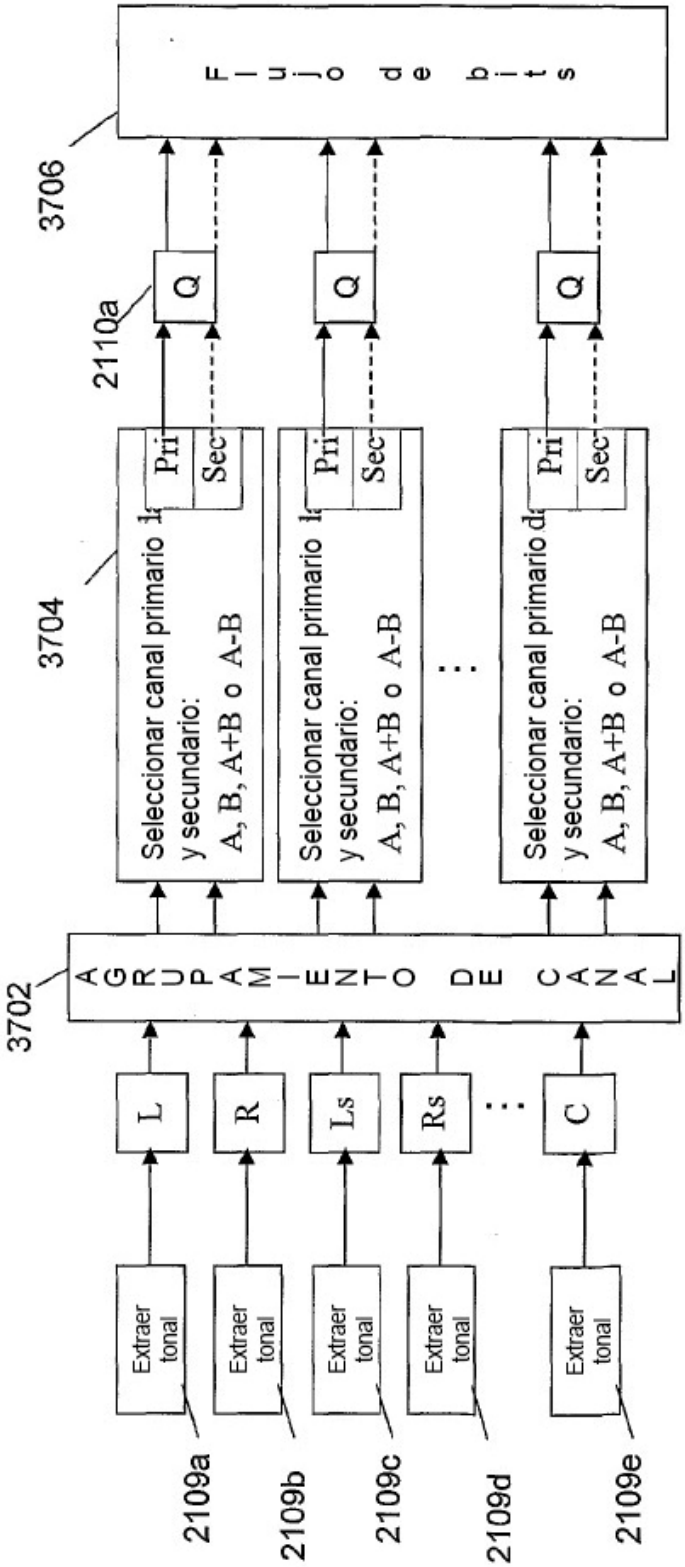

3706
Flujo de bits
2110a
Q
3704
Seleccionar canal primario
y secundario:
A, B, A+B o A-B
Pri
Sec
3702
AGRUPAMIENTO DE CANAL
L
R
Ls
Rs
C
Extraer tonal
2109a
2109b
2109c
2109d
2109e

Fig. 11

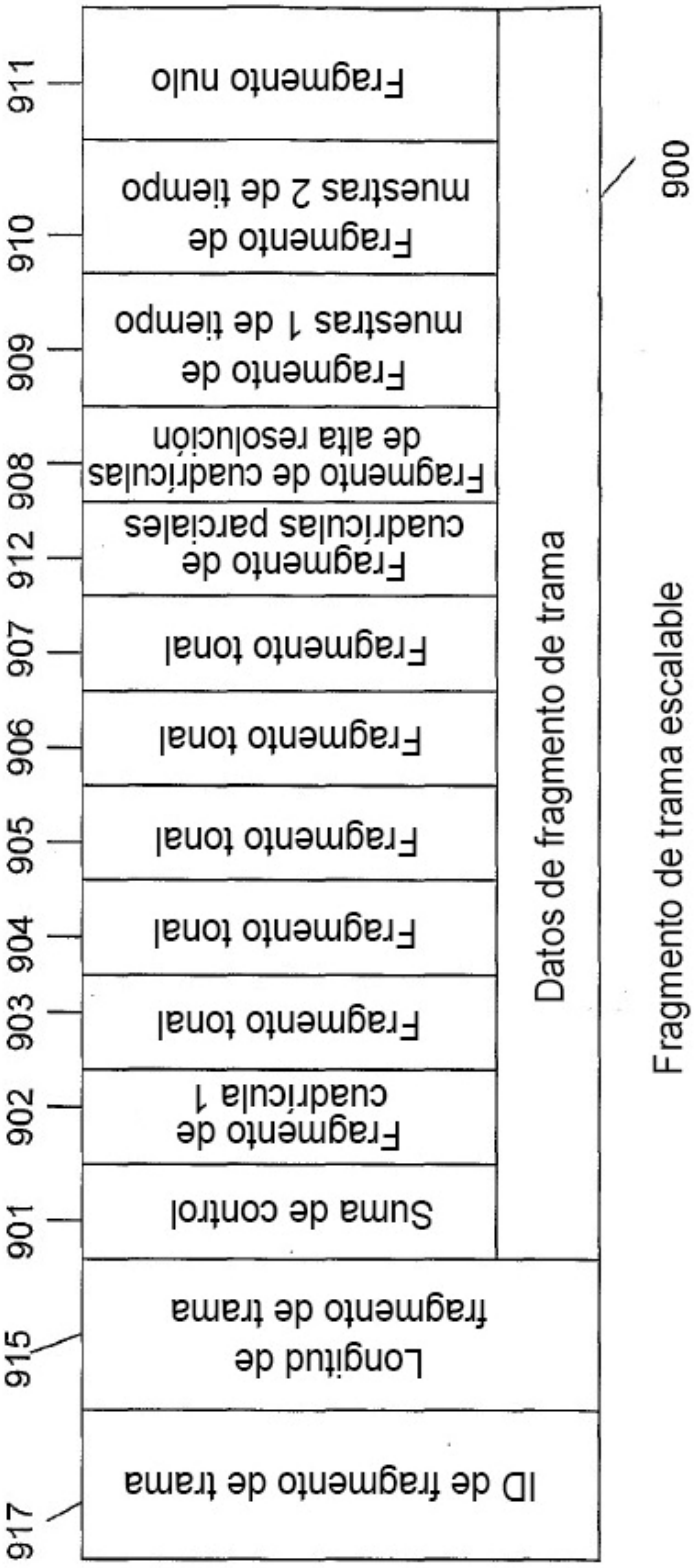

ID de fragmento de trama
Longitud de fragmento de trama
Suma de control
Fragmento de cuadrícula 1
Fragmento tonal
Fragmento tonal
Fragmento tonal
Fragmento tonal
Fragmento tonal
Fragmento tonal
Fragmento de cuadrículas parciales
Fragmento de cuadrículas de alta resolución
Fragmento de muestras 1 de tiempo
Fragmento de muestras 2 de tiempo
Fragmento nulo
Datos de fragmento de trama
Fragmento de trama escalable
900
911
910
909
908
912
907
906
905
904
903
902
901
915
917

Fig. 12

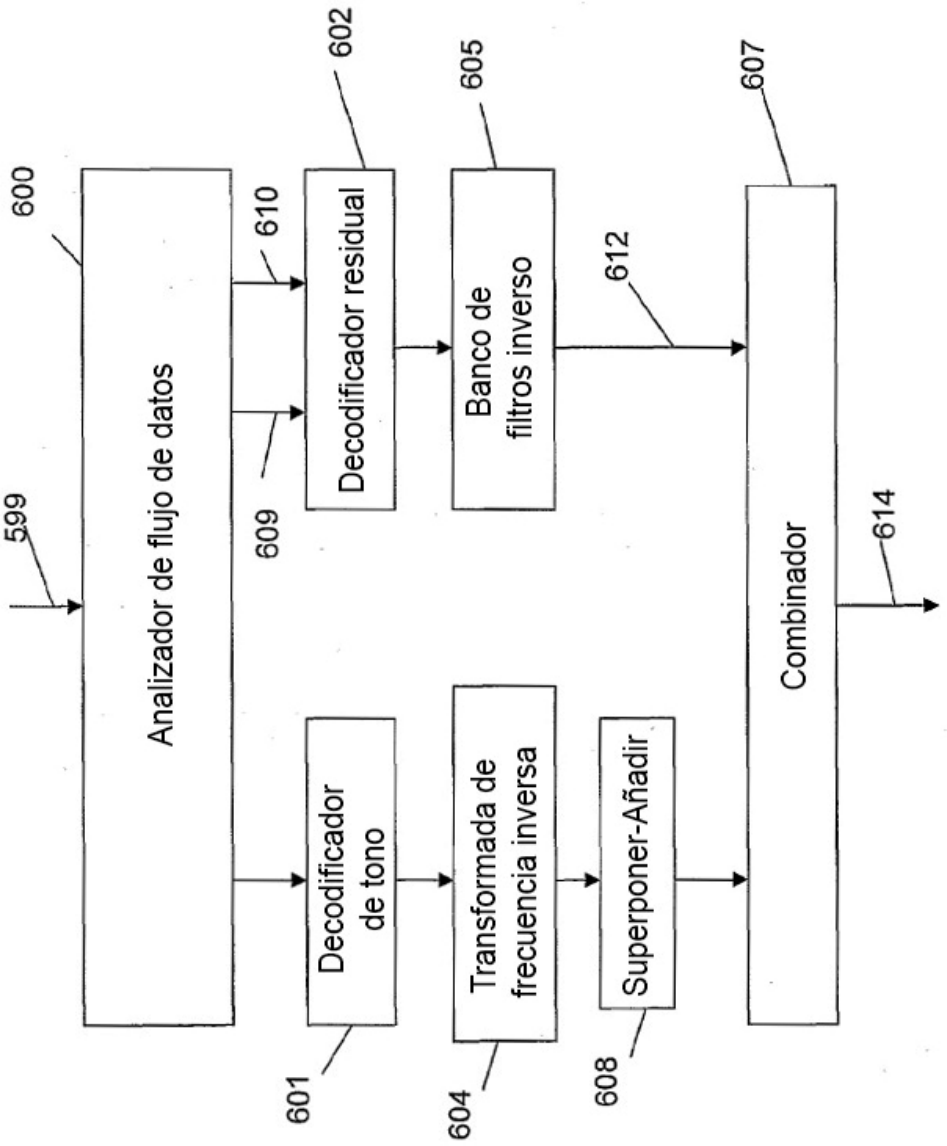

600
599
Analizador de flujo de datos
610
609
602
Decodificador residual
605
Banco de filtros inverso
612
607
Combinador
614
601
Decodificador de tono
604
Transformada de frecuencia inversa
608
Superponer-Añadir

Fig. 13

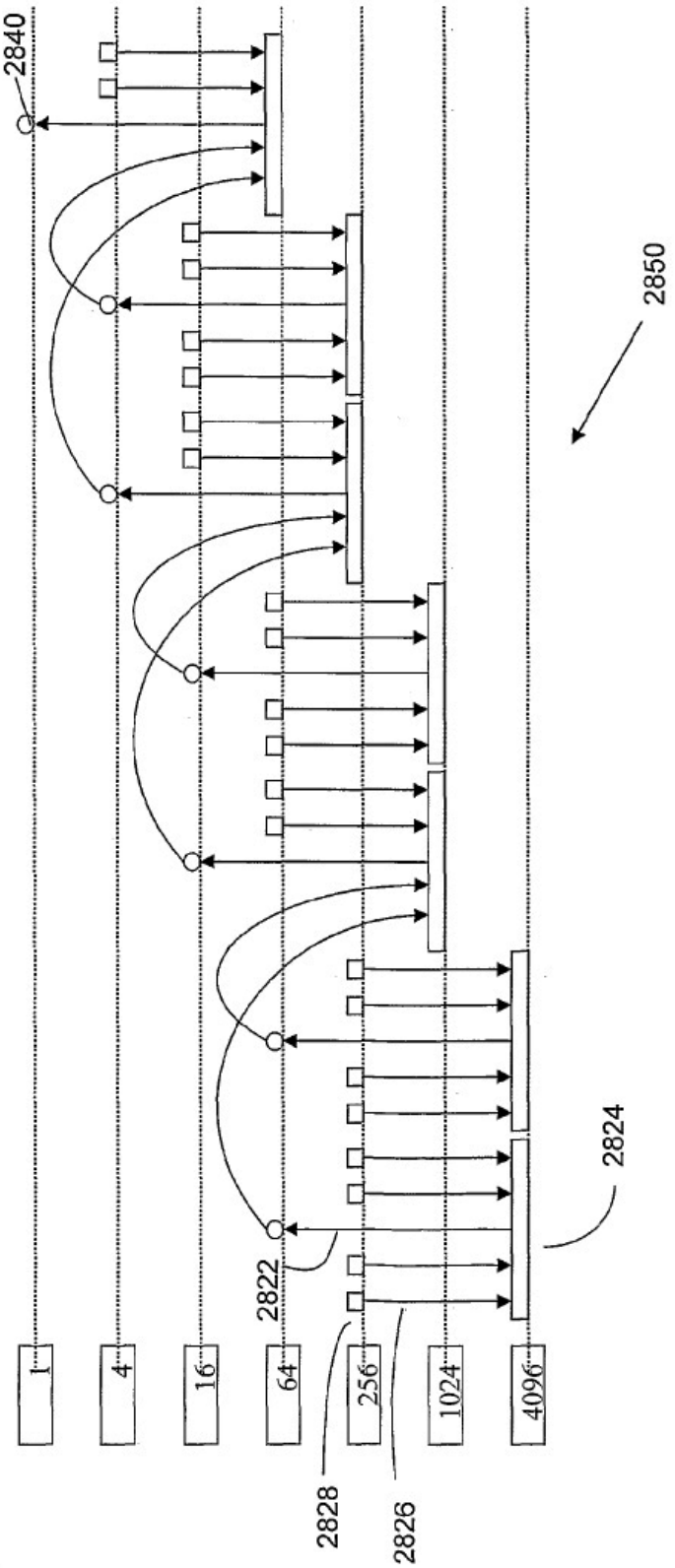


Fig. 14

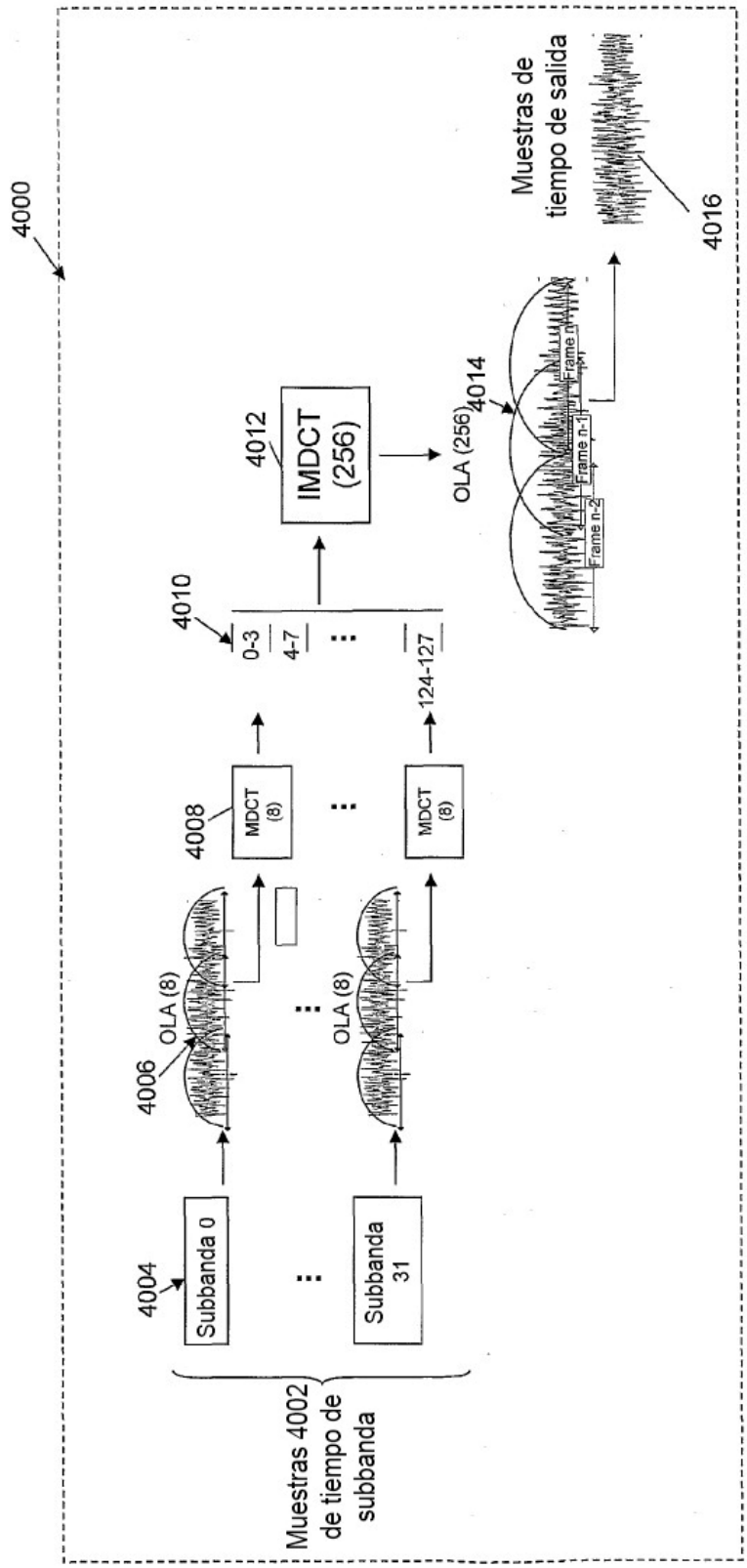

4000
Muestras de tiempo de salida
4016
4014
OLA (256)
4012
IMDCT (256)
4010
0-3
4-7
124-127
4008
MDCT (8)
MDCT (8)
4006
OLA (8)
OLA (8)
4004
Subbanda 0
Subbanda 31
Muestras 4002 de tiempo de subbanda
Frame n
Frame n-1
Frame n-2

Fig. 15

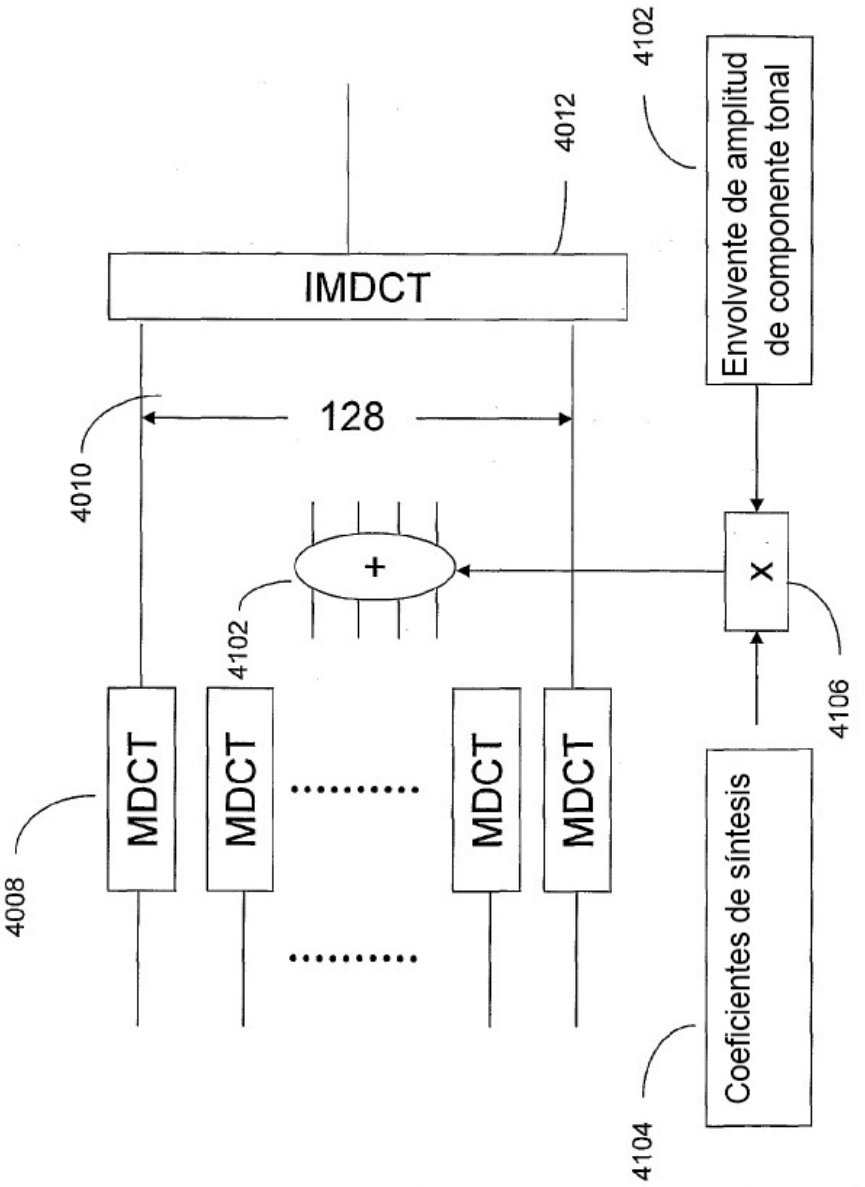

4102
Envolvente de amplitud de componente tonal
4012
IMDCT
128
4010
+
x
4102
4106
MDCT
MDCT
MDCT
MDCT
4008
Coeficientes de síntesis
4104

Fig. 16

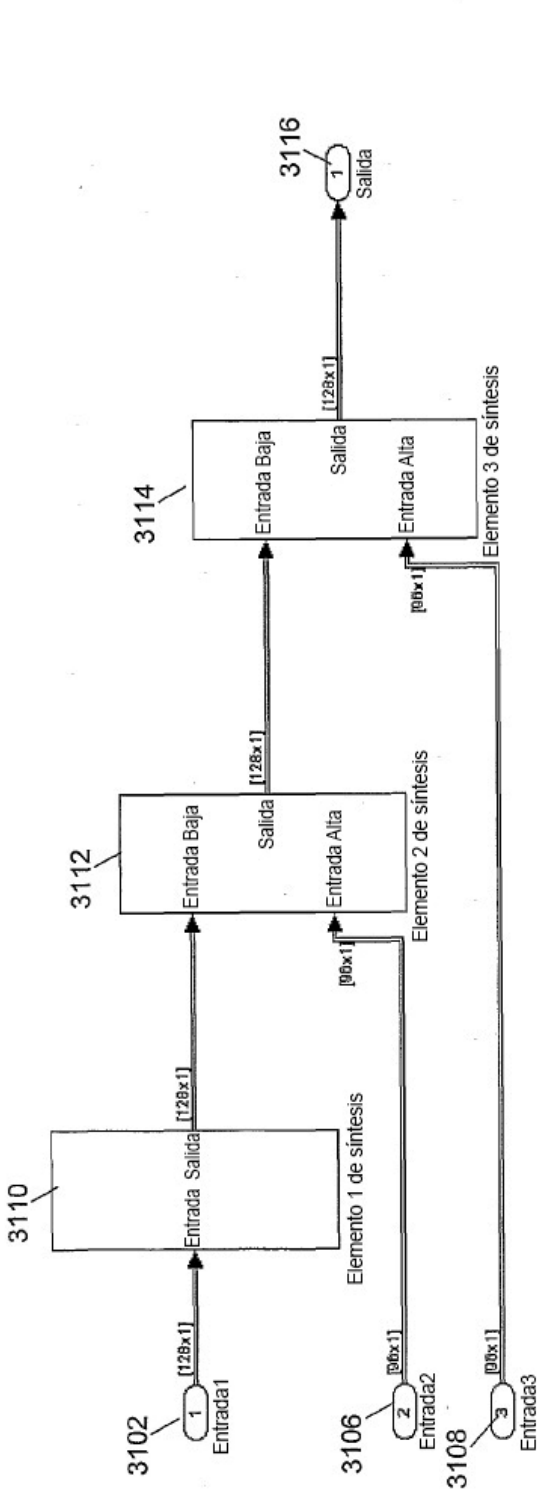


Fig. 17a

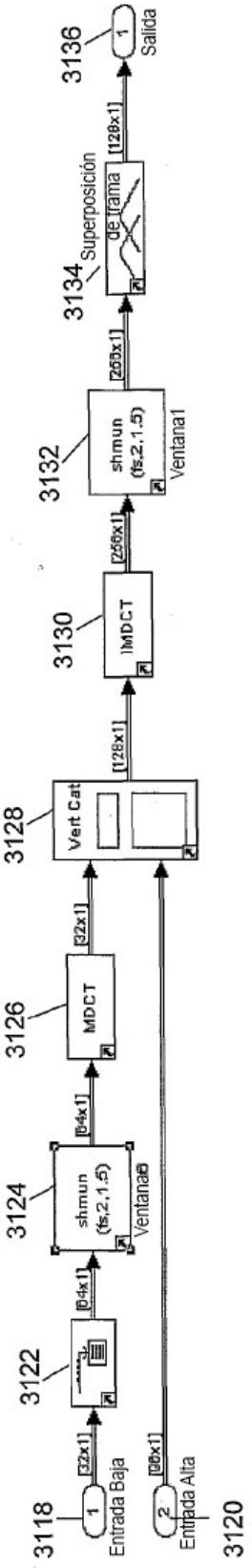


Fig. 17b

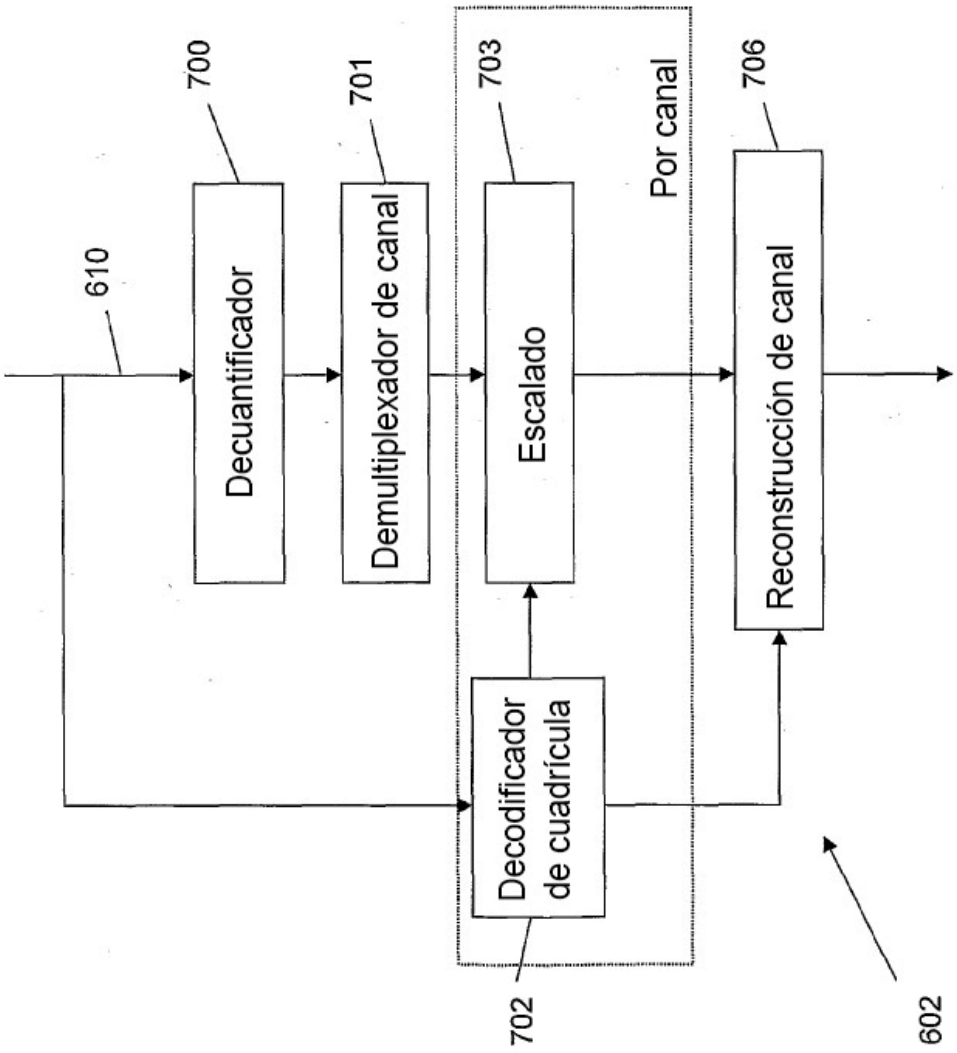
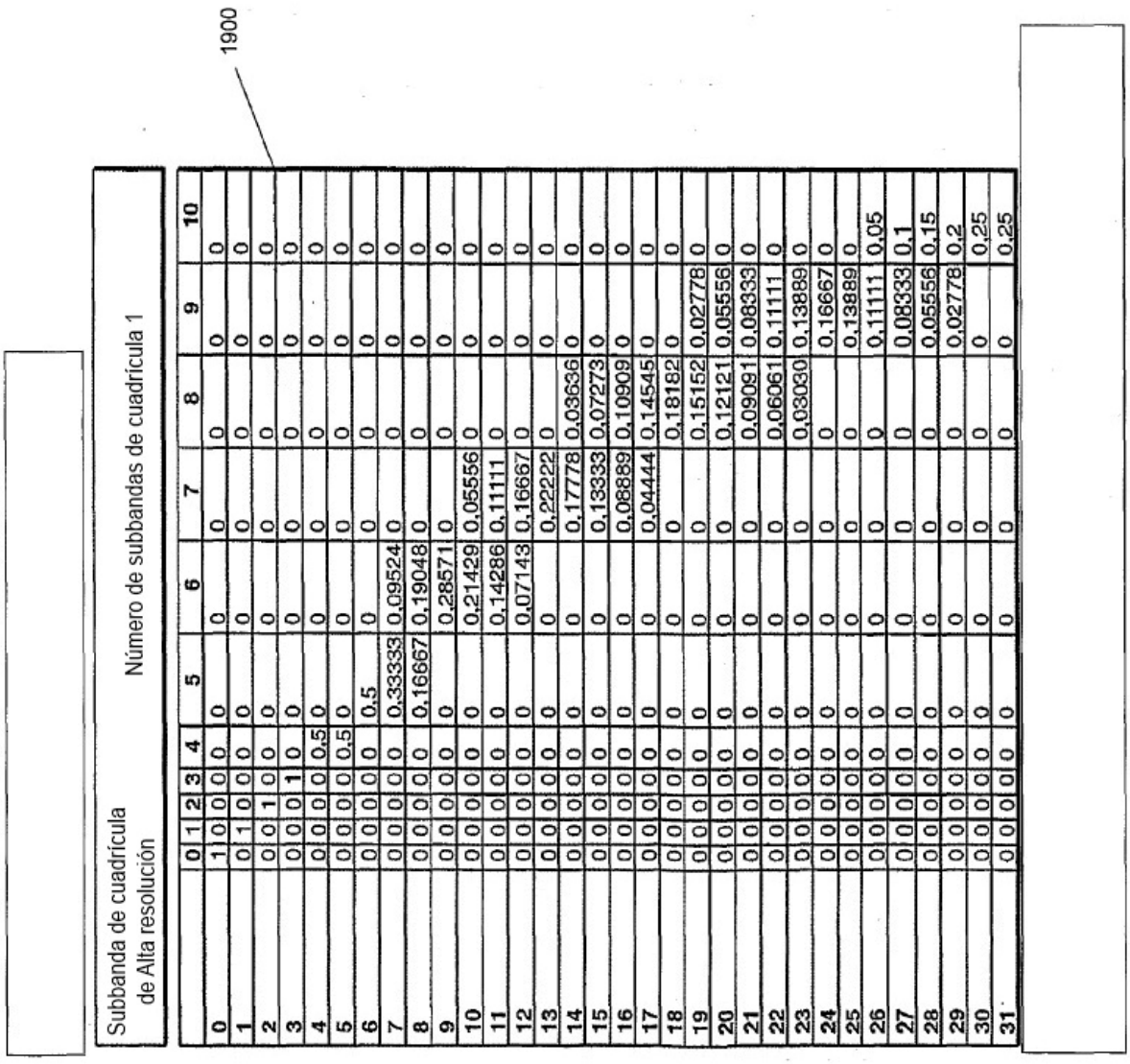

610
700
Decuantificador
701
Demultiplexador de canal
703
Escalado
Por canal
706
Reconstrucción de canal
Decodificador de cuadrícula
702
602

Fig. 18



	0	1	2	3	4	5	6	7	8	9	10
0	1	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0	0	0
2	0	0	1	0	0	0	0	0	0	0	0
3	0	0	0	1	0	0	0	0	0	0	0
4	0	0	0	0	0,5	0	0	0	0	0	0
5	0	0	0	0	0,5	0	0	0	0	0	0
6	0	0	0	0	0	0,5	0	0	0	0	0
7	0	0	0	0	0	0,33333	0,09524	0	0	0	0
8	0	0	0	0	0	0,16667	0,19048	0	0	0	0
9	0	0	0	0	0	0	0,28571	0	0	0	0
10	0	0	0	0	0	0	0,21429	0,05556	0	0	0
11	0	0	0	0	0	0	0,14286	0,11111	0	0	0
12	0	0	0	0	0	0	0,07143	0,16667	0	0	0
13	0	0	0	0	0	0	0	0,22222	0	0	0
14	0	0	0	0	0	0	0	0,17778	0,03636	0	0
15	0	0	0	0	0	0	0	0,13333	0,07273	0	0
16	0	0	0	0	0	0	0	0,08889	0,10909	0	0
17	0	0	0	0	0	0	0	0,04444	0,14545	0	0
18	0	0	0	0	0	0	0	0	0,18182	0	0
19	0	0	0	0	0	0	0	0	0,15152	0,02778	0
20	0	0	0	0	0	0	0	0	0,12121	0,05556	0
21	0	0	0	0	0	0	0	0	0,09091	0,08333	0
22	0	0	0	0	0	0	0	0	0,06061	0,11111	0
23	0	0	0	0	0	0	0	0	0,03030	0,13889	0
24	0	0	0	0	0	0	0	0	0	0,16667	0
25	0	0	0	0	0	0	0	0	0	0,13889	0
26	0	0	0	0	0	0	0	0	0	0,11111	0,05
27	0	0	0	0	0	0	0	0	0	0,08333	0,1
28	0	0	0	0	0	0	0	0	0	0,05556	0,15
29	0	0	0	0	0	0	0	0	0	0,02778	0,2
30	0	0	0	0	0	0	0	0	0	0	0,25
31	0	0	0	0	0	0	0	0	0	0	0,25

Fig. 19

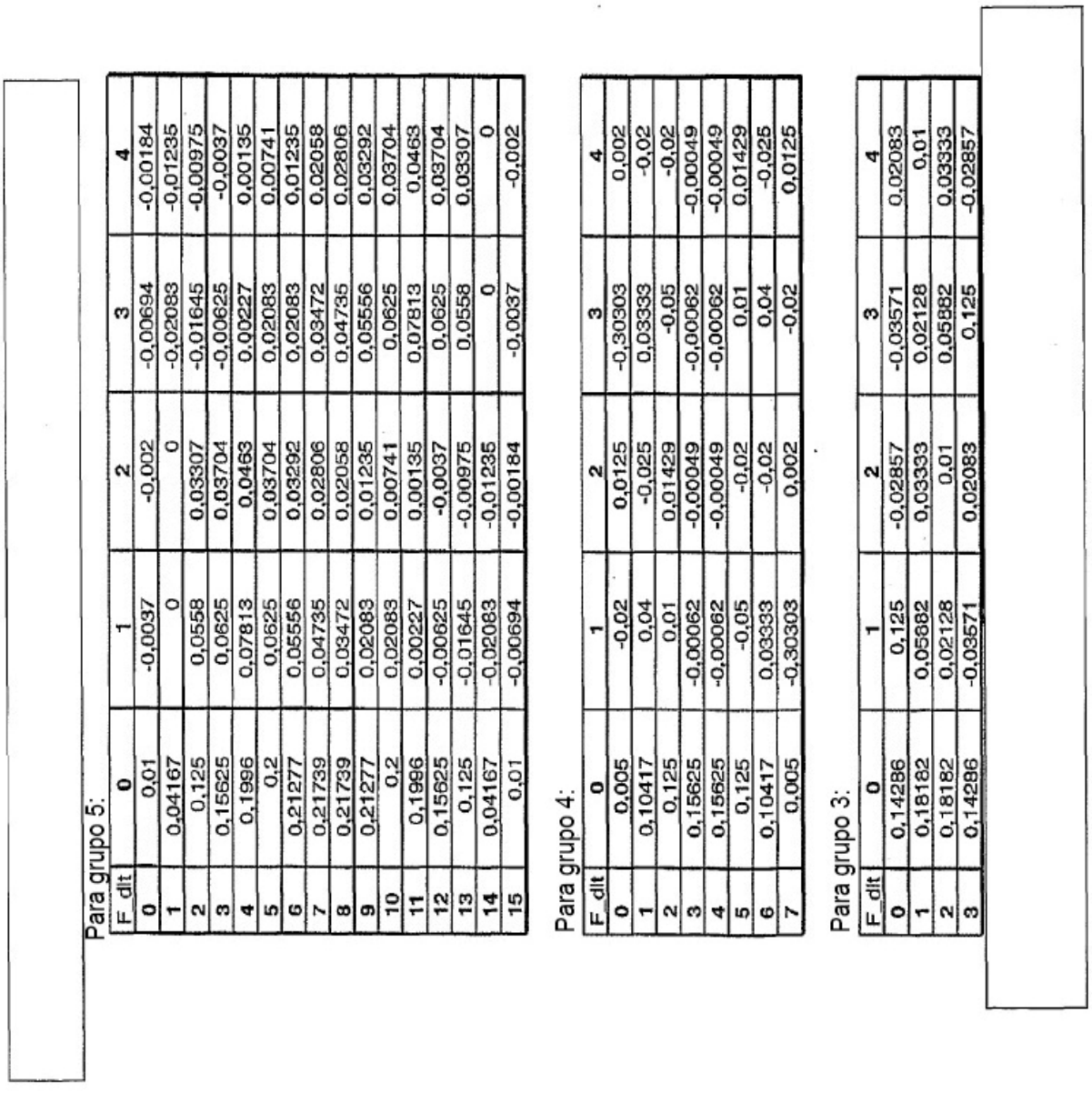

2000
Para grupo 5:
F_dlt | 0 | 1 | 2 | 3 | 4
0 | 0,01 | -0,0037 | -0,002 | -0,00694 | -0,00184
1 | 0,04167 | 0 | 0 | -0,02083 | -0,01235
2 | 0,125 | 0,0558 | 0,03307 | -0,01645 | -0,00975
3 | 0,15625 | 0,0625 | 0,03704 | -0,00625 | -0,0037
4 | 0,1996 | 0,07813 | 0,0463 | 0,00227 | 0,00135
5 | 0,2 | 0,0625 | 0,03704 | 0,02083 | 0,00741
6 | 0,21277 | 0,05556 | 0,03292 | 0,02083 | 0,01235
7 | 0,21739 | 0,04735 | 0,02806 | 0,03472 | 0,02058
8 | 0,21739 | 0,03472 | 0,02058 | 0,04735 | 0,02806
9 | 0,21277 | 0,02083 | 0,01235 | 0,05556 | 0,03292
10 | 0,2 | 0,02083 | 0,00741 | 0,0625 | 0,03704
11 | 0,1996 | 0,00227 | 0,00135 | 0,07813 | 0,0463
12 | 0,15625 | -0,00625 | -0,0037 | 0,0625 | 0,03704
13 | 0,125 | -0,01645 | -0,00975 | 0,0558 | 0,03307
14 | 0,04167 | -0,02083 | -0,01235 | 0 | 0
15 | 0,01 | -0,00694 | -0,00184 | -0,0037 | -0,002
Para grupo 4:
F_dlt | 0 | 1 | 2 | 3 | 4
0 | 0,005 | -0,02 | 0,0125 | -0,30303 | 0,002
1 | 0,10417 | 0,04 | -0,025 | 0,03333 | -0,02
2 | 0,125 | 0,01 | 0,01429 | -0,05 | -0,02
3 | 0,15625 | -0,00062 | -0,00049 | -0,00062 | -0,00049
4 | 0,15625 | -0,00062 | -0,00049 | -0,00062 | -0,00049
5 | 0,125 | -0,05 | -0,02 | 0,01 | 0,01429
6 | 0,10417 | 0,03333 | -0,02 | 0,04 | -0,025
7 | 0,005 | -0,30303 | 0,002 | -0,02 | 0,0125
Para grupo 3:
F_dlt | 0 | 1 | 2 | 3 | 4
0 | 0,14286 | 0,125 | -0,02857 | -0,03571 | 0,02083
1 | 0,18182 | 0,05682 | 0,03333 | 0,02128 | 0,01
2 | 0,18182 | 0,02128 | 0,01 | 0,05682 | 0,03333
3 | 0,14286 | -0,03571 | 0,02083 | 0,125 | -0,02857

Fig. 20

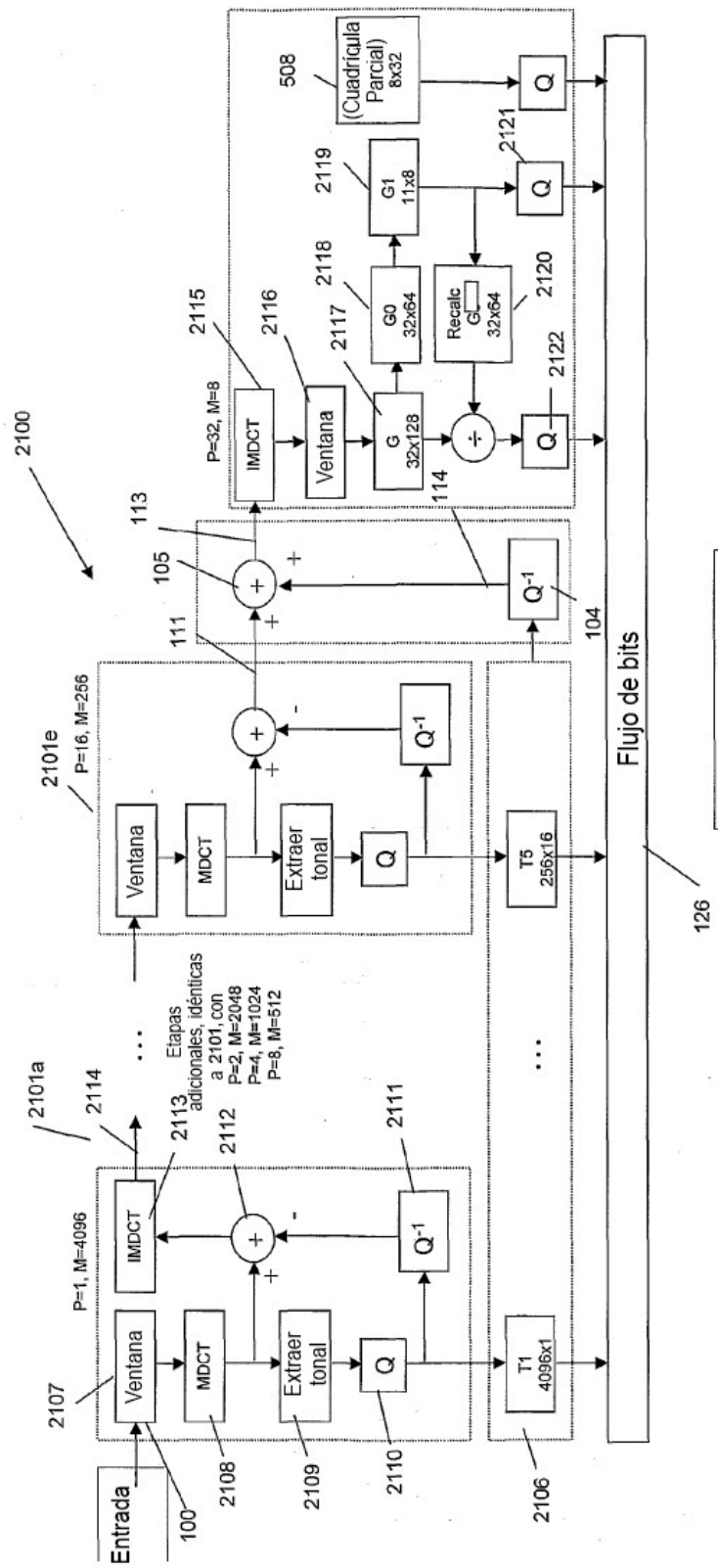

2100
Entrada
100
2107
Ventana
MDCT
Extraer tonal
Q
2108
2109
2110
2111
2112
2113
2114
2101a
P=1, M=4096
IMDCT
Q-1
T1 4096x1
2106
Etapas adicionales, idénticas a 2101, con P=2, M=2048 P=4, M=1024 P=8, M=512
2101e
P=16, M=256
Ventana
MDCT
Extraer tonal
Q
Q-1
T5 256x16
111
105
113
114
104
Q-1
P=32, M=8
IMDCT
Ventana
G 32x128
G0 32x64
G1 11x8
Recalc G 32x64
(Cuadrícula Parcial) 8x32
508
2115
2116
2117
2118
2119
2120
2121
2122
Q
Flujo de bits
126

Fig. 21

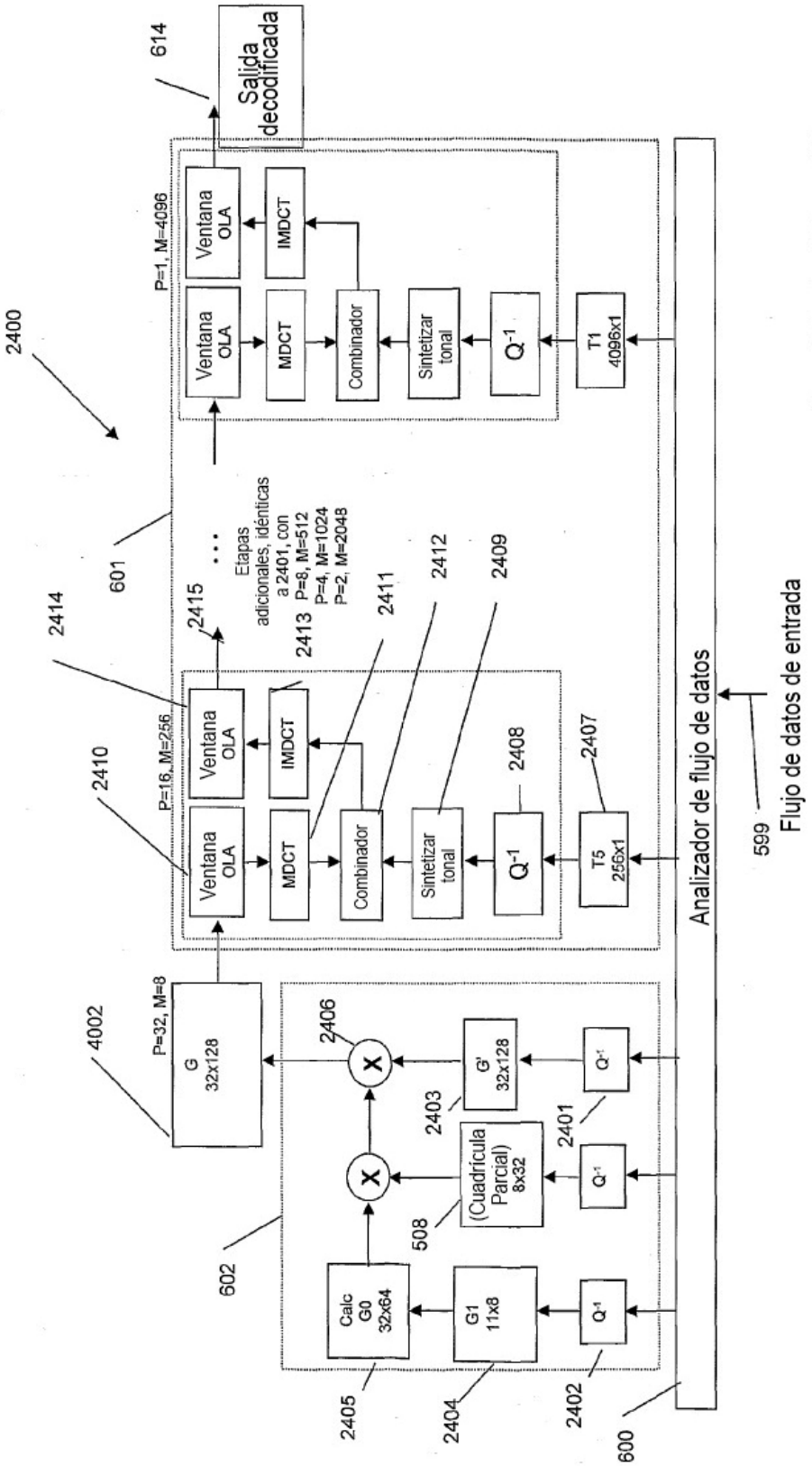

2400
614
Salida decodificada
P=1, M=4096
Ventana OLA
IMDCT
MDCT
Combinador
Sintetizar tonal
Q-1
T1 4096x1
601
2415
Etapas adicionales, idénticas a 2401, con P=8, M=512 P=4, M=1024 P=2, M=2048
2413
2414
2411
2412
2409
P=16, M=256
2410
2408
2407
T5 256x1
Analizador de flujo de datos
599
Flujo de datos de entrada
4002
P=32, M=8
G 32x128
2406
G' 32x128
2403
2401
(Cuadrícula Parcial) 8x32
508
602
Calc G0 32x64
G1 11x8
2405
2404
2402
600

Fig. 22