

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 984 501**

51 Int. Cl.:

**G10L 19/02** (2013.01)

**G10L 19/038** (2013.01)

**G10L 19/002** (2013.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **05.11.2018** **PCT/EP2018/080137**

87 Fecha y número de publicación internacional: **16.05.2019** **WO19091904**

96 Fecha de presentación y número de la solicitud europea: **05.11.2018** **E 18793692 (7)**

97 Fecha y número de publicación de la concesión europea: **24.04.2024** **EP 3707709**

54 Título: **Aparato y método para codificar y decodificar una señal de audio usando submuestreo o interpolación de parámetros de escala**

30 Prioridad:

**10.11.2017 WO PCT/EP2017/078921**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**29.10.2024**

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR  
FÖRDERUNG DER ANGEWANDTEN FORSCHUNG  
E.V. (100.0%)  
Hansastr. 27c  
80686 München, DE**

72 Inventor/es:

**RAVELLI, EMMANUEL;  
SCHNELL, MARKUS;  
BENNDORF, CONRAD;  
LUTZKY, MANFRED y  
DIETZ, MARTIN**

74 Agente/Representante:

**ARIZTI ACHA, Monica**

ES 2 984 501 T3

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

## DESCRIPCIÓN

Aparato y método para codificar y decodificar una señal de audio usando submuestreo o interpolación de parámetros de escala

5

La presente invención está relacionada con el procesamiento de audio y, en particular, con el procesamiento de audio que opera en un dominio espectral usando parámetros de escala para bandas espectrales.

### **Técnica anterior 1: Codificación de audio avanzada (AAC, por sus siglas en inglés: *Advanced Audio Coding*)**

10

En uno de los códec de audio perceptuales del estado de la técnica más ampliamente utilizado, codificación de audio avanzada (AAC) [1-2], la conformación del ruido espectral se realiza con la ayuda de los así llamados factores de escala.

15

En este enfoque, el espectro MDCT (por sus siglas en inglés: *Modified Discrete Cosine Transform* – Transformada de coseno discreta modificada) se divide en un número de bandas de factores de escala no uniformes. Por ejemplo, a 48kHz, la MDCT tiene 1.024 coeficientes y se divide en 49 bandas de factores de escala. En cada banda, se usa un factor de escala para escalar los coeficientes MDCT de esa banda. Luego se emplea un cuantizador escalar con un tamaño de etapa constante para cuantizar los coeficientes MDCT escalados. Del lado del decodificador, se realiza el escalado inverso en cada banda, conformando el ruido de cuantización introducido por el cuantizador escalar.

20

Los 49 factores de escala se codifican en la corriente de bits como información lateral. Esto usualmente requiere una cantidad considerablemente alta de bits para codificar los factores de escala, debido al número relativamente alto de factores de escala y la alta precisión requerida. Esto puede volverse un problema con baja tasa de bits y/o con bajo retardo.

25

### **Técnica anterior 2: TCX (por sus siglas en inglés: *Transform Coded Excitation* - Excitación de codificación por transformada) basada en MDCT**

30

En el TCX basado en MDCT, un códec de audio basado en transformadas utilizado en los estándares MPEG-D USAC [3] y 3GPP EVS [4], la conformación del ruido espectral se realiza con la ayuda de un filtro perceptual basado en LPC, el mismo filtro perceptual que se utiliza en los recientes códec de voz basados en ACELP (por ejemplo, AMR-WB).

35

En este enfoque, se estima en primer lugar un conjunto de 16 LPC en una señal de entrada preenfatisada. Luego, los LPC son ponderados y cuantizados. La respuesta en frecuencia de los LPC ponderados y cuantizados luego se computa en 64 bandas uniformemente espaciadas. Los coeficientes MDCT luego se escalan en cada banda usando la respuesta en frecuencia computada. Los coeficientes MDCT escalados luego se cuantizan usando un cuantizador escalar con un tamaño de etapa controlada por una ganancia global. En el decodificador, se realiza el escalado inverso en cada una de las 64 bandas, conformando el ruido de cuantización introducido por el cuantizador escalar.

40

Este enfoque tiene una clara ventaja sobre el enfoque AAC: requiere la codificación de solo 16 (LPC) + 1 (ganancia global) parámetros como información lateral (contrariamente a los 49 parámetros en AAC). Asimismo, los 16 LPC se pueden codificar eficientemente con un pequeño número de bits mediante el empleo de una representación LSF (por sus siglas en inglés: *Line Spectral Frequencies* - Frecuencias espectrales lineales) y un cuantizador vectorial. En consecuencia, el enfoque de la técnica anterior 2 requiere menos bits de información lateral que el enfoque de la técnica anterior 1, lo cual puede ser una diferencia significativa con baja tasa de bits y/o con bajo retardo.

45

No obstante, este enfoque también tiene algunos inconvenientes. El primer inconveniente es que la escala de frecuencia del conformación del ruido está limitada a ser lineal (es decir, usando bandas uniformemente espaciadas) por los LPC están estimados en el dominio del tiempo. Esto es una desventaja ya que el oído humano es más sensible en las bajas frecuencias que en las altas frecuencias. El segundo inconveniente es la alta complejidad requerida por este enfoque. La estimación de LPC (autocorrelación, Levinson-Durbin), la cuantización de LPC (conversión LPC<->LSF, cuantización vectorial) y el cálculo de la respuesta en frecuencia de los LPC son todas operaciones muy costosas. El tercer inconveniente es que este enfoque no es muy flexible debido a que el filtro perceptual basado en LPC no se puede modificar fácilmente y esto impide algunos ajustes específicos que se requerirían para ítems de audio críticos.

50

55

### **Técnica anterior 3: TCX basado en MDCT mejorado**

60

Algunos trabajos recientes han superado el primer inconveniente y parcialmente el segundo inconveniente de la técnica anterior 2. Dichos trabajos se publicaron en los documentos US 9595262 B2, EP2676266 B1. En este nuevo enfoque, la autocorrelación (para estimar los LPC) ya no se realiza en el dominio del tiempo sino que en cambio se calcula en el dominio MDCT usando una transformada inversa de las energías del coeficiente MDCT. Esto permite usar una escala de frecuencias no uniformes al simplemente agrupar los coeficientes MDCT en 64 bandas no

uniformes y computar la energía de cada banda. Se reduce así la complejidad requerida para computar la autocorrelación.

5 No obstante, la mayor parte del segundo inconveniente y del tercer inconveniente persisten, inclusive con el nuevo enfoque.

10 El documento US 4 972 484 A divulga que, en la transmisión de señales de audio, la señal de audio se representa digitalmente mediante el uso de filtrado de espejo en cuadratura en forma de una pluralidad de señales de sub-bandas espectrales. La cuantización de los valores de muestra en las sub-bandas, por ejemplo 24 sub-bandas, se controla en la medida en que los niveles de ruido de cuantización de las señales individuales de las sub-bandas están aproximadamente a la misma diferencia de nivel del umbral de enmascaramiento del sistema auditivo humano resultante de las señales individuales de las sub-bandas. Las diferencias de los niveles de ruido de cuantificación de las señales de sub-banda con respecto al umbral de enmascaramiento resultante se establecen por la diferencia entre el flujo total de información necesario para la codificación y el flujo total de información disponible para la codificación.

15 El flujo de información total disponible se fija y puede fluctuar en función de la señal

Es un objeto de la presente invención proporcionar un concepto mejorado para procesar una señal de audio.

20 Este objeto se logra mediante un aparato para codificar una señal de audio de la reivindicación 1, un método para codificar una señal de audio de la reivindicación 24, un aparato para decodificar una señal de audio codificada de la reivindicación 25, un método para decodificar una señal de audio codificada de la reivindicación 40 o un programa de computación de la reivindicación 41.

25 Un aparato para codificar una señal de audio comprende un conversor para convertir la señal de audio en una representación espectral. Asimismo, se proporciona un calculador de parámetros de escala para calcular un primer conjunto de parámetros de escala de la representación espectral. Además, a fin de mantener la tasa de bits tan baja como sea posible, el primer conjunto de parámetros de escala es submuestreado para obtener un segundo conjunto de parámetros de escala, donde un segundo número de parámetros de escala en el segundo conjunto de parámetros de escala es menor que un primer número de parámetros de escala en el primer conjunto de parámetros de escala.

30 Asimismo, se proporciona un codificador de parámetros de escala para generar una representación codificada del segundo conjunto de parámetros de escala además de un procesador espectral para procesar la representación espectral usando un tercer conjunto de parámetros de escala, el tercer conjunto de parámetros de escala tiene un tercer número de parámetros de escala que es mayor que el segundo número de parámetros de escala. Particularmente, el procesador espectral está configurado para usar el primer conjunto de parámetros de escala o para derivar el tercer conjunto de parámetros de escala del segundo conjunto de parámetros de escala o de la representación codificada del segundo conjunto de parámetros de escala usando una operación de interpolación para obtener una representación codificada de la representación espectral. Asimismo, se proporciona una interfaz de salida para generar una señal de salida codificada que comprende información sobre la representación codificada de la representación espectral y también comprende información sobre la representación codificada del segundo conjunto de parámetros de escala.

40

La presente invención está basada en el hallazgo de que se puede obtener una baja tasa de bits sin pérdida sustancial de calidad al escalar, en el lado del codificador, con un número mayor de factores de escala y mediante el submuestreo de los parámetros de escala en el lado del codificador en un segundo conjunto de parámetros de escala o factores de escala, donde los parámetros de escala en el segundo conjunto que luego es codificado y transmitido o almacenado mediante una interfaz de salida es menor que el primer número de parámetros de escala. Así se obtiene, por un lado, un fino escalamiento y una baja tasa de bits por el otro, en el lado del codificador.

45

En el lado del decodificador, el pequeño número de factores de escala transmitido es decodificado por un decodificador de factores de escala para obtener un primer conjunto de factores de escala donde el número de factores de escala o parámetros de escala en el primer conjunto es mayor que el número de factores de escala o los parámetros de escala del segundo conjunto, y luego, una vez más, se realiza un fino escalamiento usando el número más alto de parámetros de escala en el lado del decodificador dentro de un procesador espectral para obtener una representación espectral finamente escalada.

50

55 Así se obtiene, una baja tasa de bits, por un lado y, sin embargo, un procesamiento espectral de alta calidad del espectro de la señal de audio, por el otro.

La conformación del ruido espectral como el que se realiza en realizaciones preferidas se implementa usando solo una tasa de bits muy baja. Así, esta conformación del ruido espectral puede ser una herramienta esencial inclusive en un códec de audio basado en la transformación con baja tasa de bits. La conformación del ruido espectral conforma el ruido de cuantización en el dominio de la frecuencia de forma tal que el ruido de cuantización es mínimamente percibido por el oído humano, y por lo tanto, se puede maximizar la calidad perceptual de la señal de salida decodificada.

60

- Las realizaciones preferidas se remiten a parámetros espectrales calculados a partir de mediciones relacionadas con la amplitud, tales como las energías de una representación espectral. Particularmente, las energías en cuanto a las bandas o, en general, las mediciones relacionadas con la amplitud en cuanto a la energía se calculan como la base de los parámetros de escala, donde los ancho de banda usados para calcular las mediciones relacionadas con la amplitud en cuanto a la energía aumentan de bandas bajas a altas a fin de aproximarse a la característica de la audición humana tanto como sea posible. Preferentemente, la división de la representación espectral en bandas se realiza de acuerdo con la ampliamente conocida escala Bark.
- En realizaciones adicionales, se calculan los parámetros de escala en el dominio lineal y se calculan particularmente para el primer conjunto de parámetros de escala con el número alto de parámetros de escala, y este número alto de parámetros de escala se convierte en un dominio del tipo logarítmico. Un dominio del tipo logarítmico es generalmente un dominio en el cual los valores pequeños se expanden y los valores altos se comprimen. Luego la operación de submuestreo o diezmado de los parámetros de escala se realiza en el dominio del tipo logarítmico que puede ser un dominio logarítmico con base 10, o un dominio logarítmico con base 2, donde se prefiere este último a los fines de la implementación. El segundo conjunto de factores de escala luego se calcula en el dominio del tipo logarítmico, preferentemente, se realiza una cuantización vectorial del segundo conjunto de factores de escala, donde los factores de escala están en el dominio del tipo logarítmico. Por lo tanto, el resultado de la cuantización vectorial indica los parámetros de escala del dominio del tipo logarítmico. El segundo conjunto de factores de escala o parámetros de escala tiene, por ejemplo, un número de factores de escala igual a la mitad del número de factores de escala del primer conjunto, o incluso un tercio o, aún más preferiblemente, un cuarto. Luego, el pequeño número de parámetros de escala cuantizados en el segundo conjunto de parámetros de escala pasa a la corriente de bits y luego se transmite desde el lado del codificador al lado del decodificador o se almacena como una señal de audio codificada junto con un espectro cuantizado que también se ha procesado usando estos parámetros, donde este procesamiento además supone cuantización usando una ganancia global. Preferentemente, sin embargo, el codificador deriva de estos segundos factores de escala del dominio del tipo logarítmico una vez más un conjunto de factores de escala del dominio lineal, que es el tercer conjunto de factores de escala, y el número de factores de escala en el tercer conjunto de factores de escala es mayor que el número en el segundo y es preferentemente inclusive igual al primer número de factores de escala en el primer conjunto de los primeros factores de escala. Por lo tanto, en el lado del codificador, estos factores de escala interpolados se usan para procesar la representación espectral, donde la representación espectral procesada es finalmente cuantizada, y codificada entrópicamente de cualquier forma, tal como mediante la codificación de Huffman, codificación aritmética o codificación basada en cuantización vectorial, etc.
- En el decodificador que recibe una señal codificada que tiene un número bajo de parámetros espectrales junto con la representación codificada de la representación espectral, el bajo número de parámetros de escala se interpola a un alto número de parámetros de escala, es decir, para obtener un primer conjunto de parámetros de escala donde el número de parámetros de escala de los factores de escala del segundo conjunto de factores de escala o parámetros de escala es menor que el número de parámetros de escala del primer conjunto, es decir, el conjunto calculado mediante el decodificador de factores/parámetros de escala. Luego, un procesador espectral ubicado dentro del aparato para decodificar una señal de audio codificada procesa la representación espectral decodificada usando este primer conjunto de parámetros de escala para obtener una representación espectral escalada. Luego, un conversor para convertir la representación espectral escalada opera para finalmente obtener una señal de audio decodificada que preferentemente está en el dominio del tiempo.
- Realizaciones ulteriores dan por resultado ventajas adicionales tal como se establece a continuación. En realizaciones preferidas, la conformación del ruido espectral se realiza con la ayuda de 16 parámetros de escalamiento similares a los factores de escala usados en la técnica anterior 1. Estos parámetros se obtienen en el codificador al computar primeramente la energía del espectro MDCT en 64 bandas no uniformes (similares a las 64 bandas no uniformes de la técnica anterior 3), luego al aplicar algún procesamiento a las 64 energías (suavizado, preénfasis, ruido de fondo, conversión logarítmica), luego mediante submuestreo de las 64 energías procesadas por un factor de 4 para obtener 16 parámetros que finalmente son normalizados y escalados. Estos 16 parámetros luego se cuantizan usando cuantización vectorial (usando cuantización vectorial similar a la empleada en la técnica anterior 2/3). Los parámetros cuantizados luego son interpolados para obtener 64 parámetros de escala interpolados. Estos 64 parámetros de escala luego se emplean para modelar directamente el espectro MDCT en las 64 bandas no uniformes. De modo similar a la técnica anterior 2 y 3, los coeficientes MDCT escalados luego se cuantizan usando un cuantizador escalar con un tamaño de etapa controlada por una ganancia global. En el decodificador, se realiza el escalado inverso en cada una de las 64 bandas, conformando el ruido de cuantización introducido por el cuantizador escalar.
- Como en la técnica anterior 2/3, la realización preferida emplea solo 16+1 parámetros como información lateral y los parámetros se pueden codificar eficazmente con un número bajo de bits usando cuantización vectorial. En consecuencia, la realización preferida tiene las mismas ventajas que la técnica anterior 2/3: requiere menos bits de información lateral que el enfoque de la técnica anterior 1, lo cual puede ser una diferencia significativa con baja tasa de bits y/o con bajo retardo.

Como en la técnica anterior 3, la realización preferida usa escalamiento en frecuencia no lineal y, por lo tanto, no tiene el primer inconveniente de la técnica anterior 2.

Contrariamente a la técnica anterior 2/3, la realización preferida no emplea ninguna de las funciones relacionadas con LPC que tienen alta complejidad. Comparativamente, las funciones de procesamiento requeridas (suavizado, preénfasis, ruido de fondo, conversión logarítmica, normalización, escalamiento, interpolación) conllevan una complejidad mucho menor. Solo la cuantización vectorial aún tiene una complejidad relativamente alta. Pero se pueden usar algunas técnicas de cuantización vectorial de baja complejidad con pérdida pequeña en el desempeño (enfoques por divisiones múltiples/etapas múltiples). La realización preferida, por lo tanto, no tiene el segundo inconveniente de la técnica anterior 2/3 en cuanto a la complejidad.

Contrariamente a la técnica anterior 2/3, la realización preferida no se remite a un filtro perceptual basado en LPC. Esta emplea 16 parámetros de escalamiento que se pueden computar muy libremente. La realización preferida es más flexible que la técnica anterior 2/3 y por lo tanto no tiene el tercer inconveniente de la técnica anterior 2/3.

En conclusión, la realización preferida tiene todas las ventajas de la técnica anterior 2/3 y ninguno de sus inconvenientes.

Las realizaciones preferidas de la presente invención se describen a continuación con mayor detalle con respecto a las figuras adjuntas, en las cuales:

La Fig. 1 es un diagrama de bloques de un aparato para codificar una señal de audio;

La Fig. 2 es una representación esquemática de una implementación preferida del calculador del factor de escala de la Fig. 1;

La Fig. 3 es una representación esquemática de una implementación preferida del submuestreador de la Fig. 1;

La Fig. 4 es una representación esquemática del codificador de factores de escala de la Fig. 4;

La Fig. 5 es una ilustración esquemática del procesador espectral de la Fig. 1;

La Fig. 6 ilustra una representación general de un codificador por una parte, y un decodificador por otra parte, implementando la conformación del ruido espectral (SNS, por sus siglas en inglés: *Spectral Noise Shaping*);

La Fig. 7 ilustra una representación más detallada del lado del codificador por una parte, y el lado del decodificador por otra parte, en donde la conformación del ruido temporal (TNS, por sus siglas en inglés: *Temporal Noise Shaping*) se implementa junto con la conformación del ruido espectral (SNS);

La Fig. 8 ilustra un diagrama de bloques de un aparato para decodificar una señal de audio codificada;

La Fig. 9 ilustra una ilustración esquemática que ilustra los detalles del decodificador de factores de escala, el procesador espectral y el decodificador espectral de la Fig. 8;

La Fig. 10 ilustra una subdivisión del espectro en 64 bandas;

La Fig. 11 ilustra una ilustración esquemática de la operación de submuestreo por una parte, y la operación de interpolación, por otra parte;

La Fig. 12a ilustra una señal de audio en el dominio del tiempo con tramas superpuestas;

La Fig. 12b ilustra una implementación del convertidos de la Fig. 1; y

La Fig. 12c ilustra una ilustración esquemática del conversor de la Fig. 8.

La Fig. 1 ilustra un aparato para codificar una señal de audio 160. La señal de audio 160 preferentemente está disponible en el dominio del tiempo, aunque otras representaciones de la señal de audio tales como un dominio de predicción o cualquier otro dominio principalmente serían también de utilidad. El aparato comprende un conversor 100, un calculador de factores de escala 110, un procesador espectral 120, un submuestreador 130, un codificador de factores de escala 140, y una interfaz de salida 150. El conversor 100 está configurado para convertir la señal de audio 160 en una representación espectral. El calculador de factores de escala 110 está configurado para calcular un primer conjunto de parámetros de escala o factores de escala de la representación espectral.

En toda la memoria descriptiva, la expresión "factor de escala" o "parámetro de escala" se usa para hacer referencia

al mismo parámetro o valor, es decir, un parámetro o valor que, luego de algún procesamiento, se emplea para ponderar alguna clase de valores espectrales. Esta ponderación, cuando se realiza en el dominio lineal es realmente una operación de multiplicación por un factor de escalamiento. No obstante, cuando la ponderación se realiza en un dominio logarítmico, la operación de ponderación por un factor de escala se realiza mediante una operación real de suma o resta. Por lo tanto, en los términos de la presente solicitud, el escalamiento no solo significa multiplicación o división sino que también significa, según el dominio, suma o resta, o generalmente significa cada operación, por la cual el valor espectral, por ejemplo, es ponderado o modificado usando el factor de escala o el parámetro de escala.

El submuestreador 130 está configurado para submuestrear el primer conjunto de parámetros de escala para obtener un segundo conjunto de parámetros de escala, donde un segundo número de los parámetros de escala en el segundo conjunto de parámetros de escala es menor que un primer número de parámetros de escala en el primer conjunto de parámetros de escala. Esto también se señala en la caja en la Fig. 1 indicando que el segundo número es menor que el primer número. Como se ilustra en la Fig. 1, el codificador de factores de escala está configurado para generar una representación codificada del segundo conjunto de factores de escala, y esta representación codificada se envía a la interfaz de salida 150. Debido al hecho de que el segundo conjunto de factores de escala tiene un número menor de factores de escala que el primer conjunto de factores de escala, la tasa de bits para transmitir o almacenar la representación codificada del segundo conjunto de factores de escala es menor comparada con una situación en la cual no se hubiera realizado el submuestreo de los factores de escala realizado en el submuestreador 130.

Asimismo, el procesador espectral 120 está configurado para procesar la salida de la representación espectral mediante el conversor 100 en la Fig. 1 usando un tercer conjunto de parámetros de escala, el tercer conjunto de parámetros de escala o factores de escala tiene un tercer número de factores de escala que es mayor que el segundo número de factores de escala, donde el procesador espectral 120 está configurado para usar, a los fines del procesamiento espectral, el primer conjunto de factores de escala como ya está disponible del bloque 110 mediante la línea 171. Alternativamente, el procesador espectral 120 está configurado para usar el segundo conjunto de factores de escala como salida por el submuestreador 130 para el cálculo del tercer conjunto de factores de escala como se ilustra con la línea 172. En una implementación adicional, el procesador espectral 120 usa la salida de representación codificada por el codificador de factores/parámetros de escala 140 con el propósito de calcular el tercer conjunto de factores de escala como se ilustra con la línea 173 en la Fig. 1. Preferentemente, el procesador espectral 120 no usa el primer conjunto de factores de escala, sino que usa el segundo conjunto de factores de escala calculado por el submuestreador o incluso más preferentemente usa la representación codificada o, generalmente, el segundo conjunto de factores de escala cuantizados y, luego, realiza una operación de interpolación para interpolar el segundo conjunto de parámetros espectrales cuantizados para obtener el tercer conjunto de parámetros de escala que tiene un número más alto de parámetros de escala debido a la operación de interpolación.

Por lo tanto, la representación codificada del segundo conjunto de factores de escala que sale del bloque 140 comprende un índice del libro de códigos para un libro de códigos de parámetros de escala usado preferentemente o un conjunto de índices correspondiente del libro de códigos. En otras realizaciones, la representación codificada comprende los parámetros de escala cuantizados de los factores de escala cuantizados que se obtienen, cuando el índice del libro de códigos o el conjunto de índices del libro de códigos, o generalmente, la representación codificada es ingresada en el lado del decodificador de un decodificador vectorial o en cualquier otro decodificador.

Preferentemente, el procesador espectral 120 usa el mismo conjunto de factores de escala que también está disponible en el lado del decodificador, es decir, usa el segundo conjunto de parámetros de escala cuantizados junto con una operación de interpolación para finalmente obtener el tercer conjunto de factores de escala.

En una realización preferida, el tercer número de factores de escala en el tercer conjunto de factores de escala es igual al primer número de factores de escala. No obstante, también resulta útil un número más pequeño de factores de escala. Por ejemplo, se podrían derivar 64 factores de escala en el bloque 110, y luego se podrían submuestrear los 64 factores de escala a 16 factores de escala para transmisión. Luego, se podría realizar una interpolación no necesariamente a 64 factores de escala, sino a 32 factores de escala en el procesador espectral 120. Alternativamente, se podría realizar una interpolación a un número todavía más alto tal como más de 64 factores de escala según sea el caso, siempre y cuando el número de factores de escala transmitidos en la señal de salida codificada 170 sea menor que el número de factores de escala calculados en el bloque 110 o calculados y usados en el bloque 120 de la Fig. 1.

Preferentemente, el calculador de factores de escala 110 está configurado para realizar diversas operaciones que se ilustran en la Fig. 2. Estas operaciones se refieren al cálculo 111 de una medición relacionada con la amplitud por banda. Una medición preferida relacionada con la amplitud por banda es la energía por banda, pero también se pueden usar otras mediciones relacionadas con la amplitud, por ejemplo, la sumatoria de las magnitudes de las amplitudes por banda o la sumatoria del cuadrado de las amplitudes que corresponden a la energía. No obstante, además de la potencia de 2 usada para calcular la energía por banda, también podrían usarse otras potencias tales como una potencia de 3 que reflejaría la intensidad subjetiva de la señal, e inclusive se pueden usar potencias distintas de números enteros tales como potencias de 1,5 o 2,5 como así también a fin de calcular las mediciones relacionadas con la amplitud por banda. Incluso se pueden usar potencias menores que 1,0 siempre y cuando se asegure que los

valores procesados por tales potencias están valuados positivamente.

5 Una operación adicional realizada por el calculador de factores de escala puede ser un suavizado entre bandas 112. Este suavizado entre bandas se usa preferentemente para suavizar las posibles inestabilidades que pueden aparecer en el vector de mediciones relacionadas con la amplitud obtenidas en la etapa 111. Si no se realizara este suavizado, estas inestabilidades se amplificarían cuando se conviertan posteriormente a un dominio logarítmico como se ilustra en 115, especialmente en valores espectrales donde la energía está próxima a 0. No obstante, en otras realizaciones, el suavizado entre bandas no se lleva a cabo.

10 Una operación adicional preferida realizada por el calculador de factores de escala 110 es la operación de preénfasis 113. Esta operación de preénfasis tiene un propósito similar al de una operación de preénfasis usada en un filtro perceptual sobre la base de los LPC del procesamiento TCX basada en MDCT según se expuso anteriormente con respecto a la técnica anterior. Este procedimiento aumenta la amplitud del espectro modelado en las frecuencias bajas que resulta en una reducción del ruido de cuantización en las frecuencias bajas.

15 No obstante, dependiendo de la implementación, la operación de preénfasis -así como las otras operaciones específicas- no se tiene que llevar necesariamente a cabo.

20 Una operación de procesamiento opcional adicional es el procesamiento de suma de ruido de fondo 114. Este procedimiento mejora la calidad de las señales que contienen dinámicas espectrales muy altas tales como, por ejemplo, el Glockenspiel, al limitar la amplificación de la amplitud del espectro modelado en los valles, lo cual tiene el efecto indirecto de reducir el ruido de cuantización en los picos, con el costo de un incremento del ruido de cuantización en los valles, donde de todas maneras el ruido de cuantización no es perceptible debido a las propiedades de enmascaramiento del oído humano tal como el umbral de audición absoluto, el pre-enmascaramiento, el post-  
25 enmascaramiento o el umbral de enmascaramiento general que indica, generalmente, que un tono de baja intensidad relativamente próximo en frecuencia a un tono de alta intensidad no es perceptible en absoluto, es decir, está completamente enmascarado o solo es apenas percibido por el mecanismo humano de audición, de modo que esta contribución espectral se puede cuantizar groseramente.

30 Sin embargo, la operación de suma de ruido de fondo 114, no tiene que llevarse a cabo necesariamente.

Asimismo, el bloque 115 indica una conversión en el dominio de tipo logarítmico. Preferentemente, una transformación de una salida de uno de los bloques 111, 112, 113, 114 en la Fig. 2 es realizada en un dominio de tipo logarítmico. Un dominio del tipo logarítmico es un dominio, en el cual los valores cercanos a cero se expanden y los valores altos se comprimen. Preferentemente, el dominio logarítmico es un dominio con base de 2, pero también se pueden usar otros dominios logarítmicos. No obstante, un dominio logarítmico con la base de 2 es mejor para una implementación en un procesador de señales de punto fijo.

40 La salida del calculador de factores de escala 110 es un primer conjunto de factores de escala.

Como se ilustra en la Fig. 2, cada uno de los bloques 112 a 115 se puede puentear, es decir, la salida del bloque 111, por ejemplo, podría ya ser el primer conjunto de factores de escala. No obstante, se prefieren todas las operaciones de procesamiento y, particularmente, la conversión en el dominio del tipo logarítmico. Por lo tanto, aún se podría implementar el calculador de factores de escala llevando a cabo solamente las etapas 111 y 115 sin los procedimientos en las etapas 112 a 114, por ejemplo.

45 Por lo tanto, el calculador de factores de escala está configurado para realizar uno o dos o más de los procedimientos ilustrados en la Fig. 2 como se indica con las líneas de entrada/salida que conectan diversos bloques.

50 La Fig. 3 ilustra una implementación preferida del submuestreador 130 de la Fig. 1. Preferentemente, se realiza un filtrado de paso bajo o, generalmente, un filtrado con una cierta ventana  $w(k)$  en la etapa 131, y luego, se realiza una operación de submuestreo/diezrado del resultado del filtrado. Debido al hecho de que el filtrado paso bajo 131 y en realizaciones preferidas la operación de submuestreo/diezrado 132 son operaciones aritméticas, el filtrado 131 y el submuestreo 132 se pueden realizar dentro de una sola operación como se señalará más adelante. Preferentemente, la operación de submuestreo/diezrado se realiza de forma tal que se realiza una superposición entre los grupos  
55 individuales de los parámetros de escala del primer conjunto de parámetros de escala. Preferentemente, se realiza una superposición de un factor de escala en la operación de filtrado entre dos parámetros diezrados calculados. Por lo tanto, la etapa 131 realiza un filtrado de paso bajo en el vector de los parámetros de escala antes del diezrado. Este filtro paso bajo tiene un efecto similar al de la función de dispersión usada en los modelos psicoacústicos. Este filtro reduce el ruido de cuantización en los picos, con el costo de un aumento del ruido de cuantización alrededor de los picos donde este de todas maneras está enmascarado perceptualmente al menos en un mayor grado con respecto al ruido de cuantización en los picos.

Asimismo, el submuestreador además realiza una remoción del valor de la media 133 y una etapa de escalamiento

adicional 134. No obstante, la operación de filtrado de paso bajo 131, la etapa de remoción del valor de la media 133 y la etapa de escalamiento 134 solo son etapas opcionales. Por lo tanto, el submuestreador ilustrado en la Fig. 3 o ilustrado en la Fig. 1 se pueden implementar para realizar solamente la etapa 132 o para realizar las dos etapas ilustradas en la Fig. 3 tal como la etapa 132 y una de las etapas 131, 133 y 134. Alternativamente, el submuestreador puede realizar las cuatro etapas o solo tres etapas de las cuatro etapas ilustradas en la Fig. 3 siempre que la operación de submuestreo/diezmando 132 se lleve a cabo.

Como se señala en la Fig. 3, las operaciones de audio en la Fig. 3 realizadas por el submuestreador se realizan en el dominio del tipo logarítmico a fin de obtener mejores resultados.

La Fig. 4 ilustra una implementación preferida del codificador de factores de escala 140. El codificador de factores de escala 140 recibe el segundo conjunto de factores de escala preferentemente en el dominio del tipo logarítmico y realiza una cuantización vectorial como se ilustra en el bloque 141 para finalmente producir uno o varios índices por trama. Estos uno o varios índices por trama se pueden enviar a la interfaz de salida y escribirse en la corriente de bits, es decir, introducirse en la señal de audio codificada de salida 170 por medio de cualquier procedimiento de interfaz de salida disponible. Preferentemente, el cuantizador vectorial 141 adicionalmente produce el segundo conjunto de factores de escala del dominio del tipo logarítmico cuantizado.

Por lo tanto, esta información puede salir directamente del bloque 141 como se indica con la flecha 144. No obstante, alternativamente, el libro de códigos del decodificador 142 también está disponible separadamente en el codificador. Este libro de códigos del decodificador recibe el único o los varios índices por trama y deriva, de estos uno o varios índices por trama el segundo conjunto de factores de escala del dominio del tipo logarítmico cuantizados preferentemente como se indica con la línea 145. En las implementaciones típicas, el libro de códigos del decodificador 142 estará integrado dentro del cuantizador vectorial 141. Preferentemente, el cuantizador vectorial 141 es un cuantizador vectorial de etapas múltiples o por niveles o una combinación de cuantizador vectorial de etapas múltiples/por niveles como se usa, por ejemplo, en cualquiera de los procedimientos de la técnica anterior indicados.

Por lo tanto, se asegura que el segundo conjunto de factores de escala es el mismo segundo conjunto de factores de escala cuantizados que está también disponible del lado del decodificador, es decir, en el decodificador que solo recibe la señal de audio codificada que tiene el único o los varios índices por trama producidos por el bloque 141 mediante la línea 146.

La Fig. 5 ilustra una implementación preferida del procesador espectral. El procesador espectral 120 incluido dentro del codificador la Fig. 1 comprende un interpolador 121 que recibe el segundo conjunto de parámetros de escala cuantizados y que produce el tercer conjunto de parámetros de escala donde el tercer número es mayor que el segundo número y preferentemente igual al primer número. Asimismo, el procesador espectral comprende un conversor de dominio lineal 120. Luego, se realiza una conformación espectral en el bloque 123, usando los parámetros de escala lineales por una parte y la representación espectral por otra parte que se obtiene mediante el conversor 100. Preferentemente, se realiza una operación subsiguiente de conformación de ruido temporal, es decir, una predicción sobre la frecuencia a fin de obtener valores residuales espectrales en la salida del bloque 124, mientras que la información lateral TNS es enviada a la interfaz de salida como se indica con la flecha 129.

Finalmente, el procesador espectral 125 tiene un cuantizador/codificador escalar que está configurado para recibir una sola ganancia global para toda la representación espectral, es decir, para una trama completa. Preferentemente, la ganancia global se deriva según ciertas consideraciones sobre la tasa de bits. Por lo tanto, la ganancia global se establece de forma tal que la representación codificada de la representación espectral generada por el bloque 125 satisface ciertos requerimientos tales como un requerimiento de tasa de bits, un requerimiento de calidad, o ambos. La ganancia global se puede calcular iterativamente o se puede calcular en una medición pre-alimentada según sea el caso. Generalmente, la ganancia global se usa junto con un cuantizador y una ganancia global elevada generalmente resulta en una cuantización más burda donde una ganancia global menor resulta en una cuantización más fina. Por lo tanto, en otras palabras, una ganancia global mayor resulta en un tamaño de etapa de cuantización mayor mientras que una ganancia global menor resulta en un tamaño de etapa de cuantización menor cuando se obtiene un cuantizador fijo. No obstante, se pueden emplear otros cuantizadores junto con la funcionalidad de ganancia global tal como un cuantizador que tiene alguna clase de funcionalidad de compresión para valores altos, es decir, alguna clase de funcionalidad de compresión no lineal, de modo que, por ejemplo, los valores más altos están más comprimidos que los valores más bajos. La dependencia señalada entre la ganancia global y la aproximación de la cuantización es válida, cuando la ganancia global se multiplica por los valores antes de la cuantización en el dominio lineal correspondiendo a una suma en el dominio logarítmico. No obstante, si la ganancia global se aplica mediante una división en el dominio lineal, o mediante una resta en el dominio logarítmico, la dependencia resulta a la inversa. Lo mismo es verdadero, cuando la "ganancia global" representa un valor inverso.

A continuación se brindan implementaciones preferidas de los procedimientos individuales descriptos con respecto a la Fig. 1 a la Fig. 5.



## Descripción detallada etapa a etapa de las realizaciones preferidas

### **CODIFICADOR:**

#### 5 **Etapla 1: Energía por banda (111)**

Las energías por banda  $E_B(n)$  se calculan como sigue:

$$E_B(b) = \sum_{k=Ind(b)}^{Ind(b+1)-1} \frac{X(k)^2}{Ind(b+1) - Ind(b)} \quad \text{para } b = 0 \dots N_B - 1$$

10

siendo  $X(k)$  los coeficientes MDCT,  $N_B = 64$  el número de bandas e  $Ind(n)$  los índices de banda. Las bandas no son uniformes y siguen la escala Bark perceptualmente pertinente (menor en frecuencias bajas, mayor en frecuencias altas).

#### 15 **Etapla 2: Suavizado (112)**

La energía por banda  $E_B(b)$  se suaviza empleando

$$E_S(b) = \begin{cases} 0,75 \cdot E_B(0) + 0,25 \cdot E_B(1) & , \text{ si } b = 0 \\ 0,25 \cdot E_B(62) + 0,75 \cdot E_B(63) & , \text{ si } b = 63 \\ 0,25 \cdot E_B(b-1) + 0,5 \cdot E_B(b) + 0,25 \cdot E_B(b+1) & , \text{ si no} \end{cases}$$

20

Observación: esta etapa se usa principalmente para suavizar las posibles inestabilidades que pueden aparecer en el vector  $E_B(b)$ . Si no se suavizan, estas inestabilidades se amplifican cuando se convierten al dominio logarítmico (ver etapa 5), especialmente en los valles donde la energía está próxima a 0.

#### 25 **Etapla 3: Preénfasis (113)**

La energía por banda suavizada  $E_S(b)$  luego se preenfatisa usando

$$E_P(b) = E_S(b) \cdot 10^{\frac{b \cdot g_{tilt}}{10 \cdot 63}} \quad \text{para } b = 0..63$$

30

con  $g_{tilt}$  controla la inclinación de preénfasis y depende de la frecuencia de muestreo. Es, por ejemplo, 18 a 16kHz y 30 a 48kHz. El preénfasis empleado en esta etapa tiene la misma finalidad que el preénfasis empleado en el filtro perceptual basado en LPC de la técnica anterior 2, aumenta la amplitud del espectro modelado en las bajas frecuencias, resultando en la reducción del ruido de cuantización en las frecuencias bajas.

35

#### **Etapla 4: Ruido de fondo (114)**

Se añade un ruido de fondo a -40dB a  $E_P(b)$  usando

$$40 \quad E_P(b) = \max(E_P(b), \text{ruidodeFondo}) \quad \text{para } b = 0..63$$

calculándose el ruido de fondo mediante

$$\text{ruidodeFondo} = \max\left(\frac{\sum_{b=0}^{63} E_P(b)}{64} \cdot 10^{-\frac{40}{10}}, 2^{-32}\right)$$

45

Esta etapa mejora la calidad de las señales que contienen una dinámica espectral muy elevada tales como, por ejemplo, el glockenspiel, limitando la amplificación de la amplitud del espectro modelado en los valles, que tiene el efecto indirecto de reducir el ruido de cuantización en los picos, con el costo de un aumento del ruido de cuantización en los valles donde de todas maneras no es perceptible.

50

#### **Etapla 5: Logaritmo (115)**

Luego se realiza una transformación en el dominio logarítmico usando

$$E_L(b) = \frac{\log_2(E_P(b))}{2} \quad \text{para } b = 0..63$$

**Etapla 6: Submuestreo (131, 132)**

El vector  $E_L(b)$  luego es submuestreado por un factor de 4 usando

$$E_4(b) = \begin{cases} w(0)E_L(0) + \sum_{k=1}^5 w(k)E_L(4b+k-1) & , \text{si } b = 0 \\ \sum_{k=0}^4 w(k)E_L(4b+k-1) + w(5)E_L(63) & , \text{si } b = 15 \\ \sum_{k=0}^5 w(k)E_L(4b+k-1) & , \text{si no} \end{cases}$$

Con

$$w(k) = \left\{ \frac{1}{12}, \frac{2}{12}, \frac{3}{12}, \frac{3}{12}, \frac{2}{12}, \frac{1}{12} \right\}$$

Esta etapa aplica un filtro de paso bajo ( $w(k)$ ) en el vector  $E_L(b)$  antes del diezmado. Este filtro paso bajo tiene un efecto similar al de la función de dispersión usada en los modelos psicoacústicos: reduce el ruido de cuantización en los picos, con el costo de un aumento del ruido de cuantización alrededor de los picos donde de todas maneras está enmascarado perceptualmente.

**Etapla 7: Remoción de la media y escalamiento (133, 134)**

Los factores de escala finales se obtienen después de la remoción de la media y el escalamiento por un factor de 0,85

$$scf(n) = 0.85 \left( E_4(n) - \frac{\sum_{b=0}^{15} E_4(b)}{16} \right) \quad \text{para } n = 0..15$$

Dado que el códec tiene una ganancia global adicional, la media se puede remover sin pérdida de información alguna. Remover la media también permite una cuantización vectorial más eficiente. El escalamiento de 0,85 comprime ligeramente la amplitud de la curva de conformación del ruido. Este tiene un efecto perceptual similar al de la función de dispersión mencionada en la Etapa 6: ruido de cuantización reducido en los picos y ruido de cuantización aumentado en los valles.

**Etapla 8: Cuantización (141, 142)**

Los factores de escala se cuantizan usando cuantización vectorial, produciendo índices que luego se empaquen en la corriente de bits y se envían al decodificador, y factores de escala cuantizados  $scfQ(n)$ .

**Etapla 9: Interpolación (121, 122)**

Los factores de escala cuantizados  $scfQ(n)$  se interpolan usando

$$\begin{aligned}
scfQint(0) &= scfQ(0) \\
scfQint(1) &= scfQ(0) \\
scfQint(4n+2) &= scfQ(n) + \frac{1}{8}(scfQ(n+1) - scfQ(n)) \quad \text{para } n = 0..14 \\
scfQint(4n+3) &= scfQ(n) + \frac{3}{8}(scfQ(n+1) - scfQ(n)) \quad \text{para } n = 0..14 \\
scfQint(4n+4) &= scfQ(n) + \frac{5}{8}(scfQ(n+1) - scfQ(n)) \quad \text{para } n = 0..14 \\
scfQint(4n+5) &= scfQ(n) + \frac{7}{8}(scfQ(n+1) - scfQ(n)) \quad \text{para } n = 0..14 \\
scfQint(62) &= scfQ(15) + \frac{1}{8}(scfQ(15) - scfQ(14)) \\
scfQint(63) &= scfQ(15) + \frac{3}{8}(scfQ(15) - scfQ(14))
\end{aligned}$$

y se transforman nuevamente al dominio lineal usando

$$g_{SNS}(b) = 2^{scfQint(b)} \quad \text{para } b = 0..63$$

Se usa la interpolación para obtener una curva de conformación de ruido suavizada y por lo tanto evitar cualquier salto grande de amplitud entre bandas adyacentes.

#### 10 Etapa 10: Conformación espectral (123)

Los factores de escala SNS  $g_{SNS}(b)$  se aplican a las líneas de frecuencia MDCT para cada banda separadamente a fin de generar el espectro modelado  $X_s(k)$

$$X_s(k) = \frac{X(k)}{g_{SNS}(b)} \quad \text{para } k = Ind(b)..Ind(b+1)-1, \quad \text{para } b = 0..63$$

La Fig. 8 ilustra una implementación preferida de un aparato para decodificar una señal de audio codificada 250 que comprende información sobre una representación espectral codificada e información sobre una representación codificada de un segundo conjunto de parámetros de escala. El decodificador comprende una interfaz de entrada 200, un decodificador espectral 210, un decodificador de factores de escala/parámetros 220, un procesador espectral 230 y un conversor 240. La interfaz de entrada 200 está configurada para recibir la señal de audio codificada 250 y para extraer la representación espectral codificada que es enviada al decodificador espectral 210 y para extraer la representación codificada del segundo conjunto de factores de escala que es enviada al decodificador de factores de escala 220. Asimismo, el decodificador espectral 210 está configurado para decodificar la representación espectral codificada para obtener una representación espectral decodificada que es enviada al procesador espectral. El decodificador de factores de escala 220 está configurado para decodificar el segundo conjunto de parámetros de escala codificados para obtener un primer conjunto de parámetros de escala enviado al procesador espectral 230. El primer conjunto de factores de escala tiene un número de factores de escala o parámetros de escala que es mayor que el número de factores de escala o parámetros de escala en el segundo conjunto. El procesador espectral 230 está configurado para procesar la representación espectral decodificada usando el primer conjunto de parámetros de escala para obtener una representación espectral escalada. La representación espectral escalada luego es convertida por el conversor 240 para obtener finalmente la señal de audio decodificada 260.

Preferentemente, el decodificador de factores de escala 220 está configurado para operar sustancialmente de la misma manera expuesta con respecto al procesador espectral 120 de la Fig. 1 en relación con el cálculo del tercer conjunto de factores de escala o parámetros de escala según se expuso en conexión con los bloques 141 o 142 y, en particular, con respecto a los bloques 121, 122 de la Fig. 5. En particular, el decodificador de factores de escala está configurado para realizar sustancialmente el mismo procedimiento para la interpolación y la transformación nuevamente en el dominio lineal según se expuso anteriormente con respecto a la etapa 9. Por lo tanto, según se ilustra en la Fig. 9, el decodificador de factores de escala 220 está configurado para aplicar un libro de códigos decodificador 221 al único o a los varios índices por trama que representan la representación del parámetro de escala codificado. Luego, se realiza una interpolación en el bloque 222 que es sustancialmente la misma interpolación mencionada con respecto al bloque 121 en la Fig. 5. Luego, se emplea un conversor de dominio lineal 223 que es sustancialmente el mismo conversor de dominio lineal 122 mencionado con respecto a la Fig. 5. No obstante, en otras implementaciones, los bloques 221, 222, 223 pueden operar de manera diferente de la expuesta con respecto a los

bloques correspondientes en el lado del codificador.

Asimismo, el decodificador espectral 210 ilustrado en la Fig. 8 comprende un bloque decuantizador/decodificador que recibe, como una entrada, el espectro codificado y que produce un espectro decuantizado que es decuantizado preferentemente usando la ganancia global que se transmite adicionalmente desde el lado del codificador al lado del decodificador dentro de la señal de audio codificada en una forma codificada. El decuantizador/decodificador 210 puede, por ejemplo, comprender una funcionalidad decodificadora aritmética o de Huffman que recibe, como entrada, alguna clase de cógidos y que produce índices de cuantización que representan valores espectrales. Luego, estos índices de cuantización ingresan a un decuantizador junto con la ganancia global y la salida son valores espectrales decuantizados que luego pueden ser sometidos a un procesamiento TNS tal como una predicción inversa sobre la frecuencia en un bloque de procesamiento decodificador del TNS 211 que, no obstante, es opcional. Particularmente, el bloque de procesamiento decodificador del TNS además recibe la información lateral TNS que ha sido generada en el bloque 124 de la Fig. 5 como se indica con la línea 129. La salida de la etapa de procesamiento del decodificador TNS 211 ingresa al bloque de conformación espectral 212, donde el primer conjunto de factores de escala calculado por el decodificador de factores de escala se aplica a la representación espectral decodificada que puede o no estar procesada por TNS según sea el caso, y la salida es la representación espectral escalada que luego ingresa al conversor 240 de la Fig. 8.

A continuación se exponen procedimientos adicionales de realizaciones preferidas del decodificador.

## **DECODIFICADOR:**

### **Etapa 1: Cuantización (221)**

Los índices del cuantizador vectorial producidos en el codificador paso 8 son leídos de la corriente de bits y usados para decodificar los factores de escala cuantizados  $scfQ(n)$ .

### **Etapa 2: Interpolación (222, 223)**

Igual al codificador de Etapa 9.

### **Etapa 3: Conformación espectral (212)**

Los factores de escala SNS  $g_{SNS}(b)$  se aplican a las líneas de frecuencia MDCT cuantizadas para cada banda separadamente a fin de generar el espectro decodificado  $\hat{X}(k)$  según lo delineado por el código siguiente.

$$\hat{X}(k) = \widehat{X}_S(k) \cdot g_{SNS}(b) \text{ para } k = \text{Ind}(b)..\text{Ind}(b+1) - 1, \text{ para } b = 0..63$$

La Fig. 6 y la Fig. 7 ilustran una configuración general de codificador/decodificador donde la Fig. 6 representa una implementación sin procesamiento TNS, mientras que la Fig. 7 ilustra una implementación que comprende procesamiento TNS. Las funcionalidades similares ilustradas en la Fig. 6 y la Fig. 7 corresponden a funcionalidades similares en las otras figuras cuando se indican números de referencia idénticos. Particularmente, según se ilustra en la Fig. 6, la señal de entrada 160 ingresa en una etapa de transformación 110 y, a continuación, se realiza el procesamiento espectral 120. Particularmente, el procesamiento espectral se refleja mediante un codificador SNS indicado por los números de referencia 123, 110, 130, 140 que indican que el bloque codificador SNS implementa las funcionalidades indicadas por estos números de referencia. Posteriormente al bloque codificador SNS, se realiza una operación de codificación de cuantización 125, y la señal codificada se introduce en el flujo de bits como se indica en 180 en la Fig. 6. La corriente de bits 180 luego se produce en el lado del decodificador y posteriormente a una cuantización inversa y decodificación ilustrada por el número de referencia 210, se realiza la operación del decodificador SNS ilustrada por los bloques 210, 220, 230 de la Fig. 8 de modo que, al final, posteriormente a una transformada inversa 240, se obtiene la señal de salida decodificada 260.

La Fig. 7 ilustra una representación similar a la de la Fig. 6, pero se indica que, preferentemente, el procesamiento TNS se realiza posteriormente al procesamiento SNS en el lado del codificador y, de manera correspondiente, el procesamiento TNS 211 se realiza antes del procesamiento SNS 212 con respecto a la secuencia de procesamiento en el lado del decodificador.

Preferentemente, se emplea la herramienta adicional TNS entre la conformación del ruido espectral (SNS) y la cuantización/codificación (ver el diagrama de bloques a continuación). El TNS (conformación del ruido temporal) también conforma el ruido de cuantización pero realiza una conformación en el dominio del tiempo (en oposición a la conformación en el dominio de la frecuencia de SNS). TNS es útil para señales que contienen ataques agudos y para señales de voz.

TNS usualmente se aplica (por ejemplo, en AAC) entre la transformación y SNS. Sin embargo, preferentemente, se

prefiere aplicar TNS en el espectro modelado. Esto evita algunos defectos que fueron producidos por el decodificador TNS al operar el códec con bajas tasas de bits.

5 La Fig. 10 ilustra una subdivisión preferida de los coeficientes espectrales o las líneas espectrales obtenidos por el bloque 100 en el lado del codificador en bandas. Particularmente, se indica que las bandas bajas tienen un número más pequeño de líneas espectrales que las bandas altas.

Particularmente, el eje X en la Fig. 10 corresponde al índice de bandas e ilustra la realización preferida de 64 bandas y el eje Y corresponde al índice de las líneas espectrales que ilustra 320 coeficientes espectrales en una trama. 10 Particularmente, la Fig. 10 ilustra de modo ejemplar la situación del caso de la súper banda ancha (SWB, por sus siglas en inglés: *Super Wide Band*) donde existe una frecuencia de muestreo de 32 kHz.

Para el caso de la súper banda ancha, la situación con respecto a las bandas individuales es tal que una trama resulta en 160 líneas espectrales y la frecuencia de muestreo es 16 kHz de modo que, para ambos casos, una trama tiene 15 una longitud en el tiempo de 10 milisegundos.

La Fig. 11 ilustra más detalles del submuestreo preferido realizado en el submuestreador 130 de la Fig. 1 o el sobremuestreo o la interpolación correspondiente realizados en el decodificador de factores de escala 220 de la Fig. 8 o como se ilustra en el bloque 222 de la Fig. 9. 20

A lo largo del eje X, se proporciona el índice para las bandas 0 a 63. Particularmente, existen 64 bandas que van desde 0 hasta 63.

Los 16 puntos de submuestreo correspondientes a  $scfQ(i)$  se ilustran como líneas verticales 1100. Particularmente, la 25 Fig. 11 ilustra cómo se realiza un determinado agrupamiento de parámetros de escala para finalmente obtener el punto submuestreado 1100. Por ejemplo, el primer bloque de cuatro bandas consiste en (0, 1, 2, 3) y el punto medio de este primer bloque está en 1,5 indicado por el ítem 1100 en el índice 1,5 en el eje X.

De manera correspondiente, el segundo bloque de cuatro bandas es (4, 5, 6, 7), y el punto medio del segundo bloque es 5,5. 30

Las ventanas 1110 corresponden a las ventanas  $w(k)$  señaladas con respecto al submuestreo de la etapa 6 descrito anteriormente. Se puede observar que estas ventanas están centradas en los puntos submuestreados y existe el solapamiento de un bloque a cada lado como se señaló anteriormente. 35

La etapa de interpolación 222 de la Fig. 9 recupera las 64 bandas de los 16 puntos submuestreados. Esto se observa en la Fig. 11 al computar la posición de cualquiera de las líneas 1120 como una función de los dos puntos submuestreados indicados en 1100 alrededor de una línea 1120 determinada. El ejemplo a continuación lo muestra.

40 La posición de la segunda banda se calcula como una función de las dos líneas verticales alrededor de ella (1,5 y 5,5):  $2=1,5+1/8 \times (5,5-1,5)$ .

De manera correspondiente, la posición de la tercera banda como una función de las dos líneas verticales 1100 alrededor de ella (1,5 y 5,5):  $3=1,5+3/8 \times (5,5-1,5)$ . 45

Se realiza un procedimiento específico para las primeras dos bandas y las últimas dos bandas. Para estas bandas, no se puede realizar una interpolación, debido a que no existirían líneas verticales o valores correspondientes a las líneas verticales 1100 fuera del rango que va de 0 a 63. Por lo tanto, a fin de solucionar este problema, se realiza una extrapolación como se describió con respecto a la etapa 9: interpolación como se delineó antes para las dos bandas 50 0, 1 por una parte y 62 y 63 por la otra parte.

A continuación se expone sobre una implementación preferida del conversor 100 de la Fig. 1 por una parte y el conversor 240 de la Fig. 8 por la otra parte.

55 Particularmente, la Fig. 12a ilustra un esquema para indicar el entramado realizado en el lado del codificador dentro del conversor 100. La Fig. 12b ilustra una implementación preferida del conversor 100 de la Fig. 1 en el lado del codificador y la Fig. 12c ilustra una implementación preferida del conversor 240 del lado del decodificador.

El conversor 100 del lado del codificador está implementado preferentemente para realizar un entramado con tramas solapadas tal como un solapamiento de 50% de modo que la trama 2 se solapa con la trama 1 y la trama 3 se solapa con la trama 2 y la trama 4. No obstante, también se pueden realizar otros solapamientos o un procesamiento sin solapamiento, pero se prefiere realizar un solapamiento de 50% junto con un algoritmo MDCT. A tal fin, el conversor 100 comprende una ventana de análisis 101 y un conversor espectral conectado posteriormente 102 para realizar un procesamiento FFT (por sus siglas en inglés: *Fast Fourier Transform* - Transformada rápida de Fourier), un 60

procesamiento MDCT o cualquier otra clase de conversión de tiempo a espectro para obtener una secuencia de tramas correspondiente a una secuencia de representaciones espectrales como entrada en la Fig. 1 a los bloques posteriores al conversor 100.

- 5 De forma correspondiente, una o más representaciones espectrales ingresan en el conversor 240 de la Fig. 8. Particularmente, el conversor comprende un conversor a tiempo 241 que implementa una operación FFT inversa, una operación MDCT inversa o una operación correspondiente de conversión de espectro a tiempo. La salida se inserta en una ventana de síntesis 242 y la salida de la ventana de síntesis 242 ingresa en un procesador solapamiento-suma 243 para realizar una operación de solapamiento-suma para obtener finalmente la señal de audio decodificada.
- 10 Particularmente, el procesamiento solapamiento-suma en el bloque 243, por ejemplo, realiza una suma muestra por muestra entre las muestras correspondientes de la segunda mitad de, por ejemplo, la trama 3 y la primera mitad de la trama 4 de modo que los valores de muestreo de audio para obtener el solapamiento entre la trama 3 y la trama 4 como se indica en el ítem 1200 en la Fig. 12a. Se realizan operaciones similares de solapamiento-suma en una forma muestra por muestra para obtener los valores de muestreo de audio remanentes de la señal de salida de audio decodificada.
- 15

Una señal de audio codificada con la invención se puede almacenar en un medio de almacenamiento digital o en un medio de almacenamiento no transitorio o se puede transmitir por un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable tal como Internet.

- 20 Aunque se han descrito algunos aspectos en el contexto de un aparato, está claro que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a una etapa de método o una característica de una etapa de método. De manera análoga, los aspectos descritos en el contexto de una etapa de método también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente.
- 25

- Según ciertos requerimientos de implementación, las realizaciones de la invención se pueden implementar en hardware o en software. La implementación se puede realizar usando un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un CD, una memoria ROM (por sus siglas en inglés: *Read Only Memory* - Memoria de solo lectura), PROM (por sus siglas en inglés: *Programmable ROM* - Memoria de solo lectura programable), EPROM (por sus siglas en inglés: *Erasable PROM* - Memoria de solo lectura programable borrrable), EEPROM (por sus siglas en inglés: *Electrically EPROM* - Memoria de solo lectura programable borrrable electrónicamente) o una tarjeta de memoria FLASH, que tiene señales de control legibles electrónicamente almacenadas en él, que cooperan (o son capaces de cooperar) con un sistema de computación programable de forma tal que se lleva a cabo el método respectivo.
- 30
- 35

- Algunas realizaciones de acuerdo con la invención comprenden un portador de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema de computación programable, de forma tal que se lleva a cabo uno de los métodos descriptos en la presente.
- 40

- Generalmente, las realizaciones de la presente invención pueden implementarse como un producto de programa de computación con un código de programa, el código de programa es operativo para llevar a cabo uno de los métodos cuando el producto de programa de computación se ejecuta en una computadora. El código de programa puede, por ejemplo, estar almacenado en un portador legible por una máquina.
- 45

- Otras realizaciones comprenden el programa de computación para llevar a cabo uno de los métodos descriptos en la presente, almacenados en un portador legible por una máquina o en un medio de almacenamiento no transitorio.

- En otras palabras, una realización del método de la invención es, por lo tanto, un programa de computación que tiene un código de programa para llevar a cabo uno de los métodos descriptos en la presente, cuando se ejecuta el programa en una computadora.
- 50

- Una realización adicional del método de la invención es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible por computadora) que comprende, grabado en él, el programa de computación para llevar a cabo uno de los métodos descriptos en la presente.
- 55

- Una realización adicional del método de la invención es, por lo tanto, una corriente de datos o una secuencia de señales que representa el programa de computación para llevar a cabo uno de los métodos descriptos en la presente. La corriente de datos o la secuencia de señales puede, por ejemplo, estar configurada para ser transferida mediante una conexión de comunicación de datos, por ejemplo, mediante Internet.
- 60

Una realización adicional comprende un medio de procesamiento, por ejemplo, una computadora, o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los métodos descriptos en la presente.

Una realización adicional comprende una computadora que tiene instalado en ella el programa de computación para llevar a cabo uno de los métodos descritos en la presente.

- 5 En algunas realizaciones, se puede usar un dispositivo lógico programable (por ejemplo, una matriz de puertas programables por campo) para llevar a cabo algunas o todas las funcionalidades de los métodos descritos en la presente. En algunas realizaciones, una matriz de puertas programables por campo puede cooperar con un microprocesador para llevar a cabo uno de los métodos descritos en la presente. Generalmente, los métodos preferentemente se llevan a cabo mediante cualquier aparato de hardware.
- 10 Las realizaciones descritas anteriormente son meramente ilustrativas de los principios de la presente invención. Se entiende que modificaciones y variaciones de las disposiciones y los detalles descritos en la presente se tomarán evidentes a otros expertos en la técnica. Por lo tanto, la intención es limitar la invención solo al alcance de las reivindicaciones de la patente a otorgarse y no a los detalles específicos presentados a modo de descripción y explicación de las realizaciones de la presente.

15

#### **Bibliografía**

[1] ISO/IEC 14496-3:2001; *Information technology – Coding of audio-visual objects – Part 3: Audio*.

- 20 [2] 3GPP TS 26.403; *General audio codec audio processing functions; Enhanced aacPlus general audio codec; Encoder specification; Advanced Audio Coding (AAC) part*.

[3] ISO/IEC 23003-3; *Information technology – MPEG audio technologies – Part 3: Unified speech and audio coding*.

- 25 [4] 3GPP TS 26.445; *Codec for Enhanced Voice Services (EVS); Detailed algorithmic description*.

## REIVINDICACIONES

1. Un aparato para codificar una señal de audio (160), que comprende:
  - 5 un conversor (100) para convertir la señal de audio en una representación espectral;  
un calculador de parámetros de escala (110) para calcular un primer conjunto de parámetros de escala de la representación espectral;
  - 10 un submuestreador (130) para submuestrear el primer conjunto de parámetros de escala para obtener un segundo conjunto de parámetros de escala, donde un segundo número de los parámetros de escala en el segundo conjunto de parámetros de escala es menor que un primer número de parámetros de escala en el primer conjunto de parámetros de escala;
  - 15 un codificador de parámetros de escala (140) para generar una representación codificada del segundo conjunto de parámetros de escala;  
un procesador espectral (120) para procesar la representación espectral usando un tercer conjunto de parámetros de escala, el tercer conjunto de parámetros de escala que tiene un tercer número de parámetros de escala que es mayor que el segundo número de parámetros de escala, donde el procesador espectral (120) está configurado para usar el primer conjunto de parámetros de escala o para derivar el tercer conjunto de parámetros de escala del segundo conjunto de parámetros de escala o de la representación codificada del segundo conjunto de parámetros de escala usando una operación de interpolación; y
  - 20 una interfaz de salida (150) para generar una señal de salida codificada (170) que comprende información sobre la representación codificada de la representación espectral e información sobre la representación codificada del segundo conjunto de parámetros de escala,  
donde el calculador de parámetros de escala (110) está configurado para calcular, para cada banda de una pluralidad de bandas de la representación espectral, una medida relacionada con la amplitud en un dominio lineal para obtener un primer conjunto de medidas de dominio lineal; y para transformar el primer conjunto de medidas de dominio lineal en un dominio logarítmico para obtener un primer conjunto de medidas de dominio logarítmico; y
  - 25 donde el submuestreador (130) está configurado para reducir el muestreo del primer conjunto de parámetros de escala en el dominio logarítmico para obtener el segundo conjunto de parámetros de escala en el dominio logarítmico.
2. El aparato de la reivindicación 1,
  - 30 donde el procesador espectral (120) está configurado para utilizar el primer conjunto de parámetros de escala en el dominio lineal para procesar la representación espectral o para interpolar el segundo conjunto de parámetros de escala en el dominio logarítmico para obtener parámetros de escala de dominio logarítmico interpolados y transformar los parámetros de escala de dominio logarítmico en el dominio lineal para obtener el tercer conjunto de parámetros de escala.
3. Aparato según una de las reivindicaciones anteriores,
  - 35 donde el calculador de parámetros de escala (110) está configurado para calcular el primer conjunto de parámetros de escala para bandas no uniformes, y  
donde el reductor de muestreo (130) está configurado para reducir el muestreo del primer conjunto de parámetros de escala para obtener un primer parámetro de escala del segundo conjunto combinando un primer grupo que tiene un primer número predefinido de parámetros de escala de frecuencia adyacente del primer conjunto, y en el que el divisor de frecuencias (130) está configurado para reducir el primer conjunto de parámetros de escala para obtener un segundo parámetro de escala del segundo conjunto combinando un segundo grupo que tiene un segundo número predefinido de parámetros de escala de frecuencia adyacente del primer conjunto, en el que el segundo número predefinido es igual al primer número predefinido, y en el que el segundo grupo tiene miembros que son diferentes de los miembros del primer grupo.
4. Aparato de la reivindicación 3, en donde el primer grupo de parámetros de escala de frecuencia adyacente del primer conjunto y el segundo grupo de parámetros de escala de frecuencia adyacente del primer conjunto tienen al menos un parámetro de escala del primer conjunto en común, de modo que el primer grupo y el



segundo grupo se solapan entre sí.

5. Aparato de una de las reivindicaciones precedentes, en donde el submuestreador (130) está configurado para utilizar una operación media entre un grupo de primeros parámetros de escala, teniendo el grupo dos o más miembros.
6. Aparato según la reivindicación 5, en donde la operación de promedio es una operación de promedio ponderado configurada para ponderar un parámetro de escala en el medio del grupo más fuerte que un parámetro de escala en un borde del grupo.
7. Aparato según una de las reivindicaciones anteriores, en donde el submuestreador (130) está configurado para realizar una eliminación del valor medio (133) de modo que el segundo conjunto de parámetros de escala esté libre de media, o en donde el submuestreador (130) está configurado para realizar una operación de escala (134) utilizando un parámetro de escala inferior a 1,0 y superior a 0,0 en el dominio logarítmico, o en donde el codificador de parámetros de escala (140) está configurado para cuantificar y codificar el segundo conjunto utilizando un cuantificador vectorial (141), en donde la representación codificada comprende uno o más índices (146) para uno o más libros de códigos del cuantificador vectorial, o en donde el codificador de parámetros de escala (140) está configurado para proporcionar un segundo conjunto de parámetros de escala cuantizados asociados con la representación codificada (142), y en el que el procesador espectral (120) está configurado para derivar el segundo conjunto de parámetros de escala a partir del segundo conjunto de parámetros de escala cuantizados (145), o en donde el procesador espectral (120) está configurado para determinar este tercer conjunto de parámetros de escala de modo que el tercer número sea igual al primer número, o en donde el procesador espectral (120) está configurado para determinar un parámetro de escala interpolado (121) basado en un parámetro de escala cuantizado y una diferencia entre el parámetro de escala cuantizado y el siguiente parámetro de escala cuantizado en una secuencia ascendente de parámetros de escala cuantizados con respecto a la frecuencia.
8. Aparato según la reivindicación 7, en donde el procesador espectral (120) está configurado para determinar, a partir del parámetro de escala cuantizado y de la diferencia, al menos dos parámetros de escala interpolados, en donde para cada uno de los dos parámetros de escala interpolados se utiliza un factor de ponderación diferente.
9. Aparato según una de las reivindicaciones anteriores, en donde el procesador espectral (120) está configurado para realizar la operación de interpolación (121) en el dominio logarítmico, y convertir (122) los parámetros de escala interpolados en el dominio lineal para obtener el tercer conjunto de parámetros de escala, o en donde el calculador de parámetros de escala (110) está configurado para calcular una medida relacionada con la amplitud para cada banda a fin de obtener un conjunto de medidas relacionadas con la amplitud (111), y para suavizar (112) las medidas relacionadas con la amplitud a fin de obtener un conjunto de medidas relacionadas con la amplitud suavizadas como el primer conjunto de parámetros de escala, o en donde el calculador de parámetros de escala (110) está configurado para calcular una medida relacionada con la amplitud para cada banda a fin de obtener un conjunto de medidas relacionadas con la amplitud, y para realizar (113) una operación de preénfasis al conjunto de medidas relacionadas con la amplitud, en donde la operación de preénfasis es tal que las amplitudes de baja frecuencia se enfatizan con respecto a las amplitudes de alta frecuencia, o en donde el calculador de parámetros de escala (110) está configurado para calcular una medida relacionada con la amplitud para cada banda a fin de obtener un conjunto de medidas relacionadas con la amplitud, y para realizar una operación de adición de piso de ruido (114), en donde se calcula un piso de ruido a partir de una medida relacionada con la amplitud derivada como valor medio de dos o más bandas de frecuencia de la representación espectral, o en el que el calculador de parámetros de escala (110) está configurado para

5 realizar al menos una de un grupo de operaciones, el grupo de operaciones comprende calcular (111) medidas relacionadas con la amplitud para una pluralidad de bandas, realizar (112) una operación de suavizado, realizar (113) una operación de preénfasis, realizar (114) una operación de adición de piso de ruido, y realizar una operación de conversión de dominio logarítmico (115) para obtener el primer conjunto de parámetros de escala, o

10 en donde el procesador espectral (120) está configurado para ponderar (123) valores espectrales en la representación espectral utilizando el tercer conjunto de parámetros de escala para obtener una representación espectral ponderada y aplicar una operación de conformación de ruido temporal (TNS) (124) sobre la representación espectral ponderada, y en donde el procesador espectral (120) está configurado para cuantizar (125) y codificar un resultado de la operación de conformación de ruido temporal (124) para obtener la representación codificada de la representación espectral, o

15 en donde el convertidor (100) comprende una ventana de análisis (101) para generar una secuencia de bloques de muestras de audio con ventanas, y un convertidor tiempo-espectro (102) para convertir los bloques de muestras de audio con ventanas en una secuencia de representaciones espectrales, siendo una representación espectral una trama espectral, o

20 en donde el convertidor (100) está configurado para aplicar una operación MDCT (transformada discreta del coseno modificada) para obtener un espectro MDCT a partir de un bloque de muestras en el dominio del tiempo, o

25 en donde el calculador de parámetros de escala (110) está configurado para calcular, para cada banda, una energía de la banda, comprendiendo el cálculo elevar al cuadrado las líneas espectrales, sumar las líneas espectrales elevadas al cuadrado y dividir las líneas espectrales elevadas al cuadrado por un número de líneas en la banda, o

30 en donde el procesador espectral (120) está configurado para ponderar (123) valores espectrales de la representación espectral o para ponderar (123) valores espectrales derivados de la representación espectral de acuerdo con un esquema de bandas, siendo el esquema de bandas idéntico al esquema de bandas utilizado en el cálculo del primer conjunto de parámetros de escala por el calculador de parámetros de escala (110), o

35 en donde un número de bandas es 64, el primer número es 64, el segundo número es 16, y el tercer número es 64, o

40 en donde el procesador espectral (120) está configurado para calcular una ganancia global para todas las bandas y para cuantizar (125) los valores espectrales subsiguientes a un escalado (123) que implica el tercer número de parámetros de escala utilizando un cuantificador escalar, donde el procesador espectral (120) está configurado para controlar un tamaño de etapa del cuantificador escalar (125) dependiente de la ganancia global.

10. Un método para codificar una señal de audio (160), que comprende:

45 convertir (100) la señal de audio (160) en una representación espectral;

calcular (110) un primer conjunto de parámetros de escala de la representación espectral:

50 submuestrear (130) el primer conjunto de parámetros de escala para obtener un segundo conjunto de parámetros de escala, donde un segundo número de los parámetros de escala en el segundo conjunto de parámetros de escala es menor que un primer número de parámetros de escala en el primer conjunto de parámetros de escala;

55 generar (140) una representación codificada del segundo conjunto de parámetros de escala;

60 procesar (120) la representación espectral usando un tercer conjunto de parámetros de escala, el tercer conjunto de parámetros de escala que tiene un tercer número de parámetros de escala que es mayor que el segundo número de parámetros de escala, en donde el proceso (120) usa el primer conjunto de parámetros de escala o deriva el tercer conjunto de parámetros de escala del segundo conjunto de parámetros de escala o de la representación codificada del segundo conjunto de parámetros de escala usando una operación de interpolación; y

generar (150) una señal de salida codificada (170) que comprende información sobre la representación codificada de la representación espectral e información sobre la representación codificada del segundo

- conjunto de parámetros de escala,
- 5 en donde el cálculo (110) de un primer conjunto de parámetros de escala comprende calcular, para cada banda de una pluralidad de bandas de la representación espectral, una medida relacionada con la amplitud en un dominio lineal para obtener un primer conjunto de medidas de dominio lineal; y transformar el primer conjunto de medidas de dominio lineal en un dominio logarítmico para obtener el primer conjunto de medidas de dominio logarítmico; y
- 10 en donde el submuestreo (130) comprende submuestreo para reducir el muestreo del primer conjunto de parámetros de escala en el dominio logarítmico para obtener el segundo conjunto de parámetros de cálculo de escala en el dominio logarítmico.
11. Aparato para decodificar una señal de audio codificada (250) que comprende información sobre una representación espectral codificada e información sobre una representación codificada de un segundo conjunto de parámetros de escala, que comprende:
- 15 una interfaz de entrada (200) para recibir la señal de audio codificada (250) y extraer la representación espectral codificada y la representación codificada del segundo conjunto de parámetros de escala;
- 20 un decodificador de espectro (210) para decodificar la representación espectral codificada y obtener una representación espectral decodificada;
- 25 un decodificador de parámetros de escala (220) para decodificar el segundo conjunto codificado de parámetros de escala para obtener un primer conjunto de parámetros de escala, en donde un número de parámetros de escala del segundo conjunto es menor que un número de parámetros de escala del primer conjunto;
- 30 un procesador espectral (230) para procesar la representación espectral decodificada usando el primer conjunto de parámetros de escala para obtener una representación espectral escalada; y
- 35 un convertidor (240) para convertir la representación espectral escalada con el fin de obtener una señal de audio descodificada (260),
- en donde el decodificador de parámetros de escala (220) está configurado para interpolar (222) el segundo conjunto de parámetros de escala en el dominio logarítmico para obtener parámetros de escala interpolados en el dominio logarítmico.
12. Aparato según la reivindicación 11,
- 40 en donde el decodificador de parámetros de escala (220) está configurado para descodificar la representación espectral codificada utilizando un decuantificador vectorial (210) que proporciona, para uno o más índices de cuantización, el segundo conjunto de parámetros de escala descodificados, y en el que el decodificador de parámetros de escala (220) está configurado para interpolar (222) el segundo conjunto de parámetros de escala decodificados para obtener el primer conjunto de parámetros de escala, o
- 45 en donde el decodificador de parámetros de escala (220) está configurado para determinar un parámetro de escala interpolado basado en un parámetro de escala cuantizado y una diferencia entre el parámetro de escala cuantizado y el siguiente parámetro de escala cuantizado en una secuencia ascendente de parámetros de escala cuantizados con respecto a la frecuencia.
- 50 13. Aparato según la reivindicación 12,
- 55 en donde el decodificador de parámetros de escala (220) está configurado para determinar, a partir del parámetro de escala cuantizado y de la diferencia, al menos dos parámetros de escala interpolados, en donde para la generación de cada uno de los dos parámetros de escala interpolados se usa un factor de ponderación diferente.
14. Aparato según la reivindicación 13,
- 60 en donde el decodificador de parámetros de escala (220) está configurado para utilizar los factores de ponderación, en donde los factores de ponderación aumentan con frecuencias crecientes asociadas a los parámetros de escala interpolados, o
- en donde el decodificador de parámetros de escala (220) está configurado para llevar a cabo una operación

de interpolación (222) en el dominio logarítmico, y para convertir (223) los parámetros de escala interpolados en el dominio lineal para obtener el primer conjunto de parámetros de escala, en donde el dominio logarítmico es un dominio logarítmico con una base de 10 o con una base de 2, o

5 en donde el procesador espectral (230) está configurado para aplicar (211) una operación de decodificación de conformación de ruido temporal (TNS) a la representación espectral decodificada para obtener una representación espectral decodificada TNS, y para ponderar (212) la representación espectral decodificada TNS utilizando el primer conjunto de parámetros de escala, o

10 en donde el decodificador de parámetros de escala (220) está configurado para interpolar parámetros de escala cuantificados de modo que los parámetros de escala cuantificados interpolados tengan valores que se encuentren en un intervalo de  $\pm 20\%$  de los valores obtenidos utilizando las siguientes ecuaciones:

$$\begin{aligned} scfQint(0) &= scfQ(0) \\ scfQint(1) &= scfQ(0) \\ scfQint(4n+2) &= scfQ(n) + \frac{1}{8}(scfQ(n+1) - scfQ(n)) \quad \text{para } n = 0..14 \\ scfQint(4n+3) &= scfQ(n) + \frac{3}{8}(scfQ(n+1) - scfQ(n)) \quad \text{para } n = 0..14 \\ scfQint(4n+4) &= scfQ(n) + \frac{5}{8}(scfQ(n+1) - scfQ(n)) \quad \text{para } n = 0..14 \\ scfQint(4n+5) &= scfQ(n) + \frac{7}{8}(scfQ(n+1) - scfQ(n)) \quad \text{para } n = 0..14 \\ scfQint(62) &= scfQ(15) + \frac{1}{8}(scfQ(15) - scfQ(14)) \\ scfQint(63) &= scfQ(15) + \frac{3}{8}(scfQ(15) - scfQ(14)) \end{aligned}$$

15 en donde  $scfQ(n)$  es el parámetro de escala cuantizado para un índice  $n$ , y en donde  $scfQint(k)$  es el parámetro de escala interpolado para un índice  $k$ , o

20 en donde el decodificador de parámetros de escala (220) está configurado para realizar una interpolación (222) para obtener parámetros de escala dentro, con respecto a la frecuencia, del primer conjunto de parámetros de escala y para realizar una operación de extrapolación para obtener parámetros de escala en los bordes, con respecto a la frecuencia, del primer conjunto de parámetros de escala.

25 15. Aparato según la reivindicación 14,

en donde el decodificador de parámetros de escala (220) está configurado para determinar al menos un primer parámetro de escala y un último parámetro de escala del primer conjunto de parámetros de escala con respecto a bandas de frecuencia ascendentes mediante una operación de extrapolación.

30 16. Aparato según una de las reivindicaciones 10 a 15,

35 en donde el decodificador de parámetros de escala (220) está configurado para realizar una interpolación (222) y una transformación posterior del dominio logarítmico al dominio lineal, en el que el dominio logarítmico es un dominio  $\log 2$  y en el que los valores del dominio lineal se calculan utilizando una exponenciación con una base de dos, o

40 en donde la señal de audio codificada (250) comprende información sobre una ganancia global para la representación espectral codificada, en la que el decodificador espectral (210) está configurado para descuantificar (210) la representación espectral codificada utilizando la ganancia global, y en donde el procesador espectral (230) está configurado para procesar la representación espectral descuantificada o los valores derivados de la representación espectral descuantificada ponderando cada valor espectral descuantificado o cada valor derivado de la representación espectral descuantificada de una banda utilizando el mismo parámetro de escala del primer conjunto de parámetros de escala para la banda, o

45 en donde el convertidor (240) está configurado para convertir (241) representaciones espectrales a escala subsiguiente en el tiempo; para realizar una ventana de síntesis (242) de las representaciones espectrales a escala subsiguiente en el tiempo convertidas, y para superponer y sumar (243) las representaciones

convertidas en ventana para obtener una señal de audio descodificada (260), o

5 en donde el convertidor (240) comprende un convertidor de transformada discreta de coseno inversa modificada (MDCT), o

en donde el procesador espectral (230) está configurado para multiplicar los valores espectrales por los parámetros de escala correspondientes del primer conjunto de parámetros de escala, o

10 en donde un segundo número de parámetros de escala en el segundo conjunto de parámetros de escala es 16 y el primer número es 64, o

15 en donde cada parámetro de escala del primer conjunto está asociado a una banda, en donde las bandas correspondientes a frecuencias más altas son más amplias que las bandas asociadas a frecuencias más bajas, de modo que un parámetro de escala del primer conjunto de parámetros de escala asociado a una banda de frecuencias altas se utiliza para ponderar un mayor número de valores espectrales en comparación con un parámetro de escala asociado a una banda de frecuencias más bajas, en el que el parámetro de escala asociado a la banda de frecuencias más bajas se utiliza para ponderar un menor número de valores espectrales en la banda de frecuencias bajas.

20 17. Método para decodificar una señal de audio codificada (250) que comprende información sobre una representación espectral codificada e información sobre una representación codificada de un segundo conjunto de parámetros de escala, que comprende:

25 recibir (200) la señal de audio codificada (250) y extraer la representación espectral codificada y la representación codificada del segundo conjunto de parámetros de escala;

decodificar (210) la representación espectral codificada y obtener una representación espectral decodificada;

30 decodificar (220) el segundo conjunto codificado de parámetros de escala para obtener un primer conjunto de parámetros de escala, en donde un número de parámetros de escala del segundo conjunto es menor que un número de parámetros de escala del primer conjunto;

35 procesar (230) la representación espectral decodificada usando el primer conjunto de parámetros de escala para obtener una representación espectral escalada; y

convertir (240) la representación espectral escalada con el fin de obtener una señal de audio descodificada (260),

40 en donde la decodificación (220) del segundo conjunto codificado de parámetros de escala comprende interpolar (222) el segundo conjunto de parámetros de escala en un dominio logarítmico para obtener parámetros de escala de dominio logarítmico interpolados.

18. Programa informático para realizar, cuando se ejecuta en un ordenador o un procesador, el método según la reivindicación 10 o el método según la reivindicación 17.

45

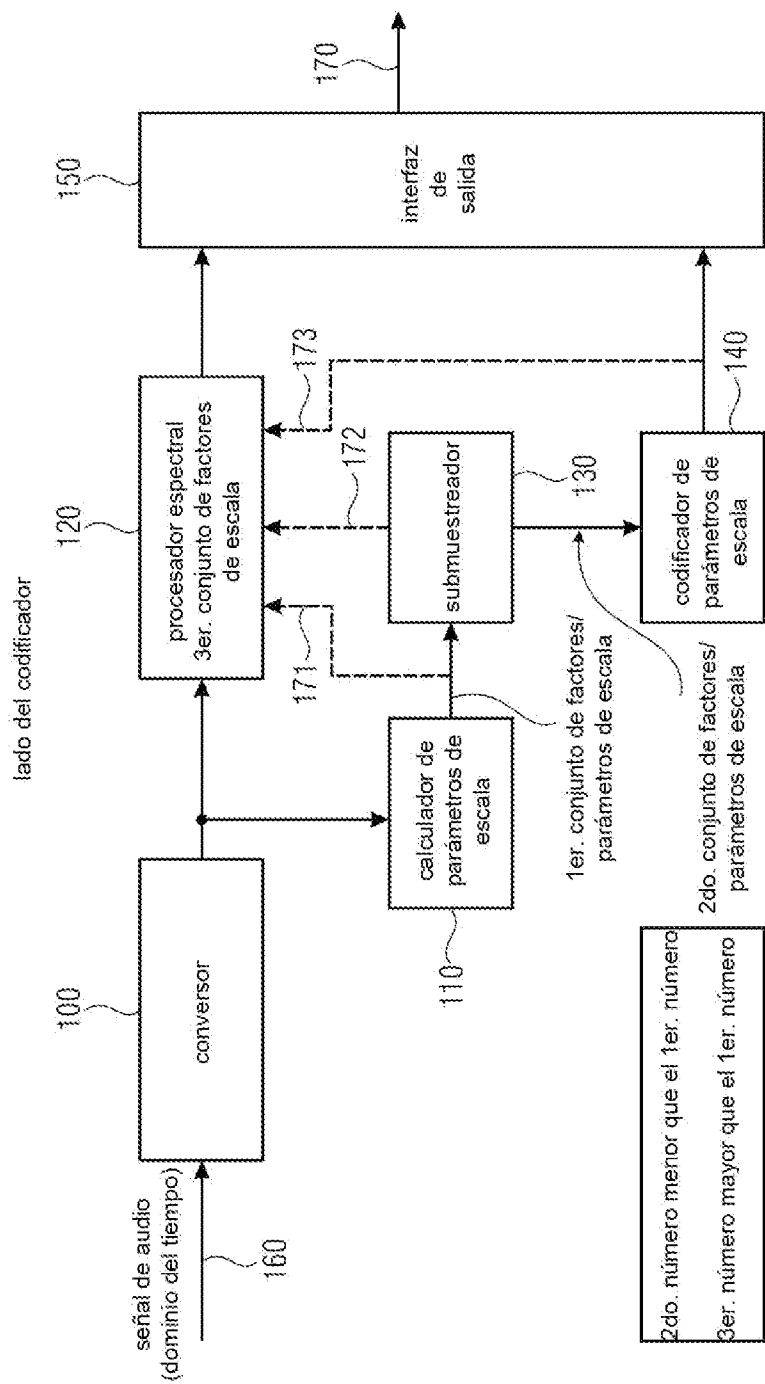
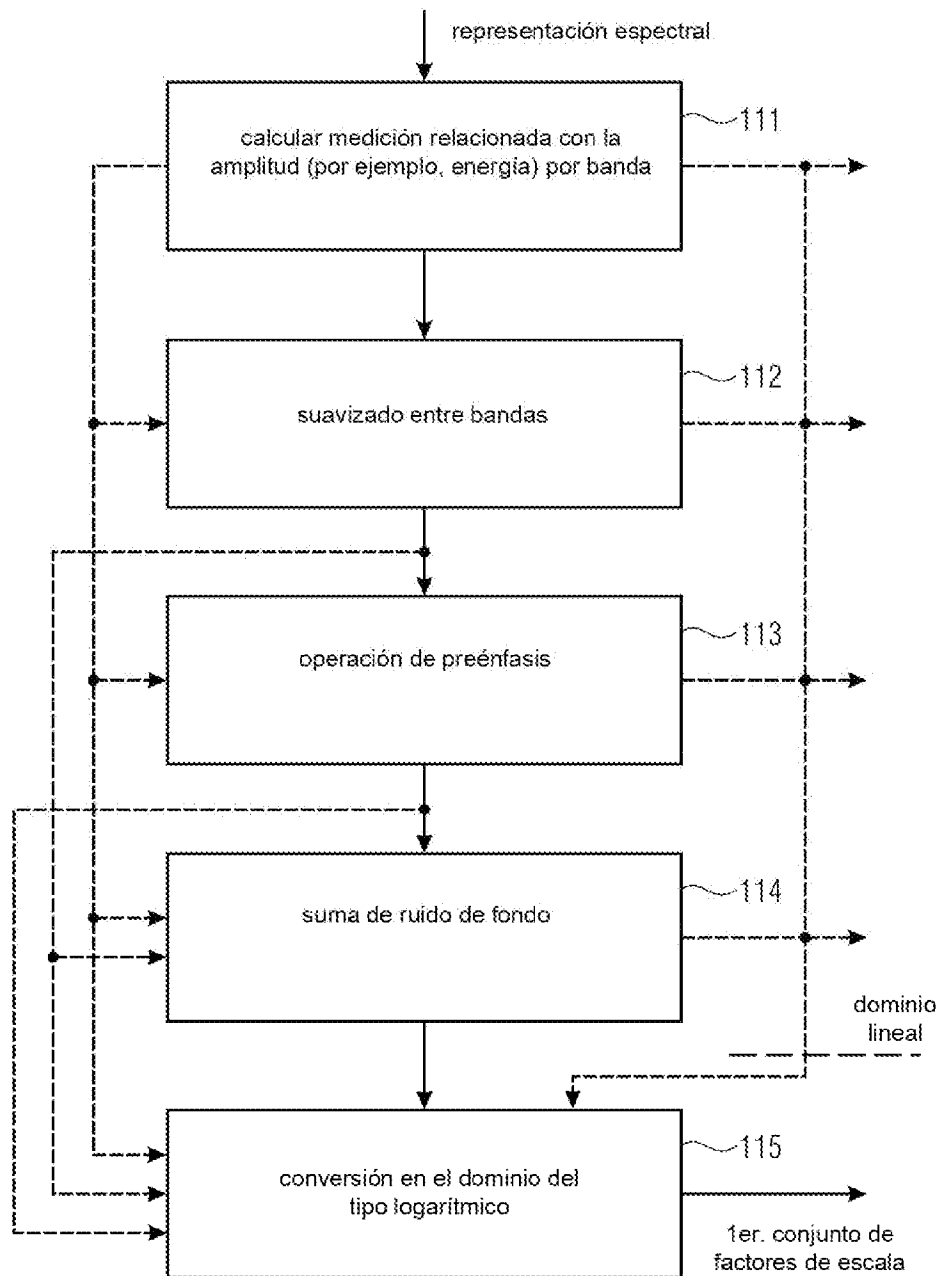


Fig. 1



**Fig. 2**  
 (    CALCULADOR DE FACTORES  
      DE ESCALA    )

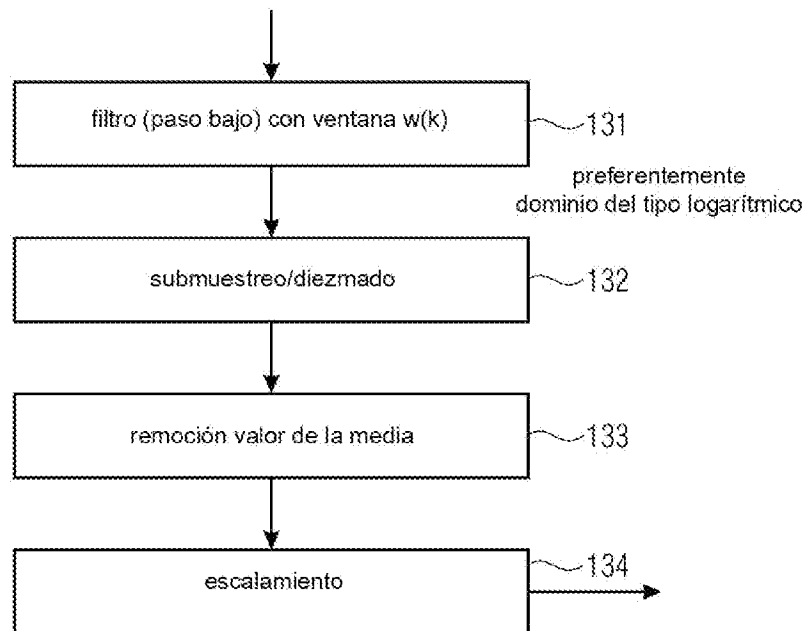


Fig. 3  
( SUBMUESTREADOR )

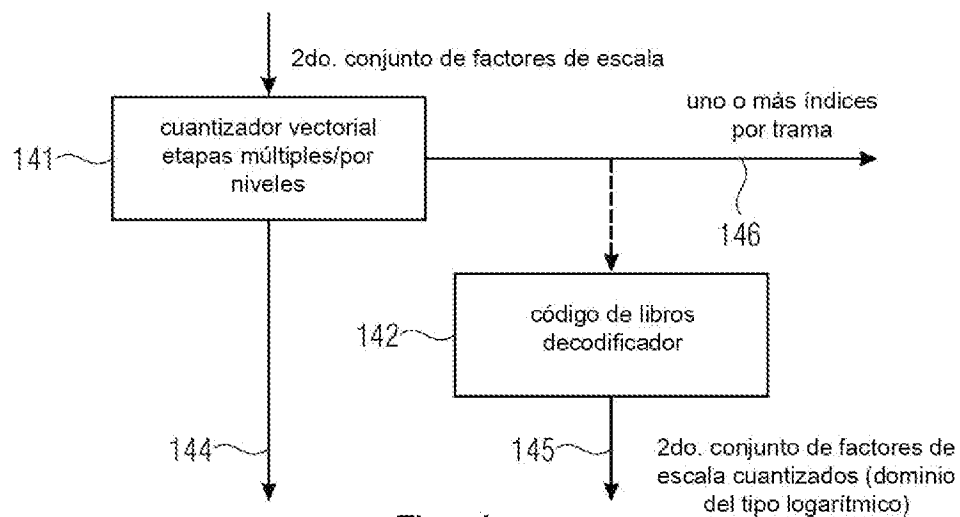
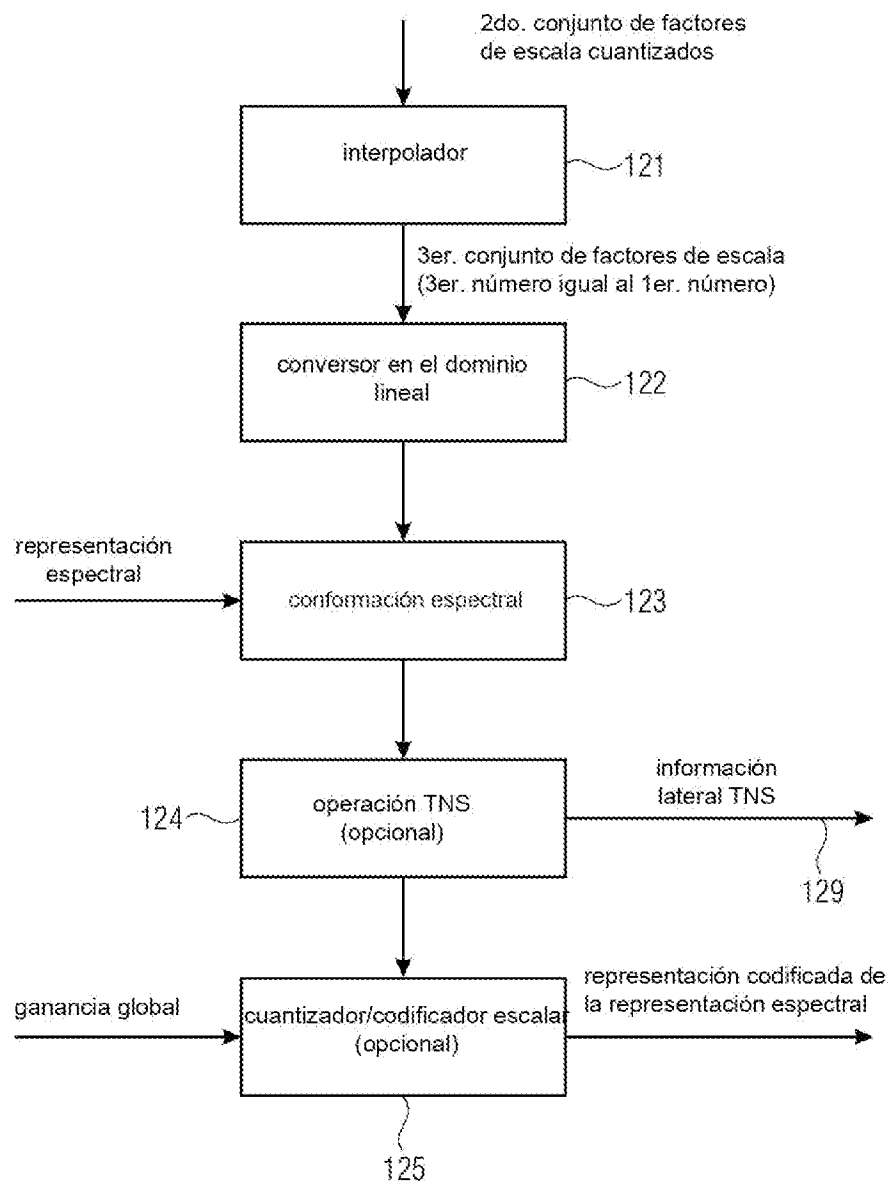


Fig. 4  
( CODIFICADOR DE FACTORES DE ESCALA )





**Fig. 5**  
( PROCESADOR ESPECTRAL )

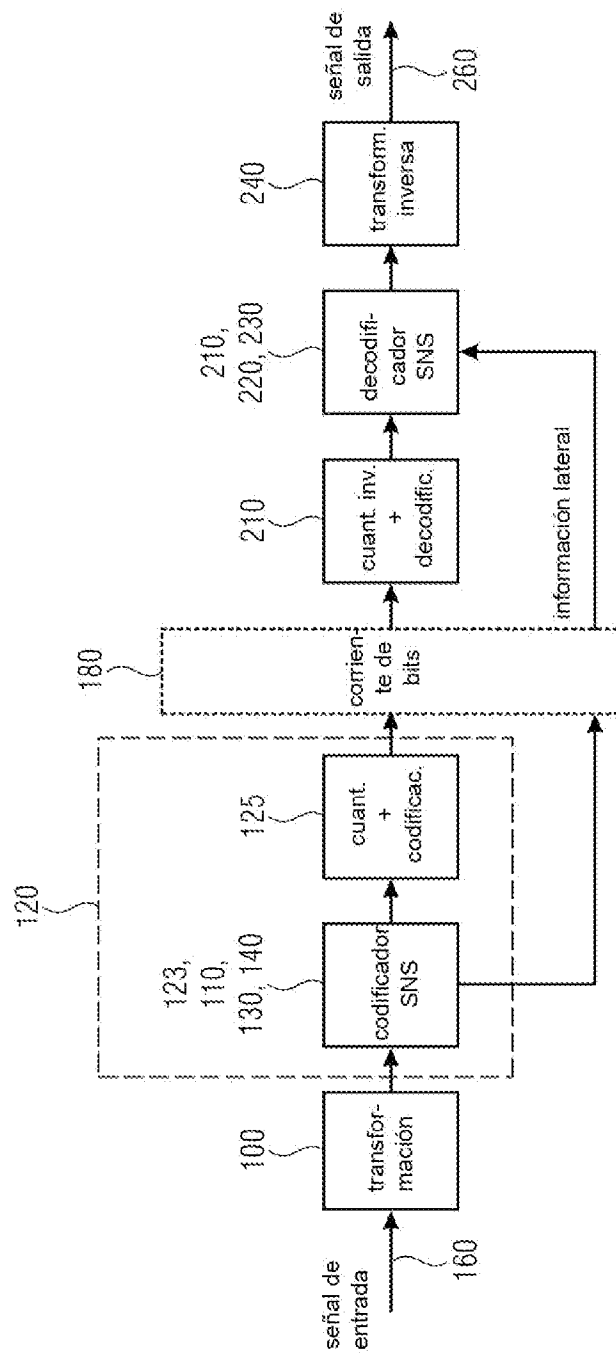


Fig. 6

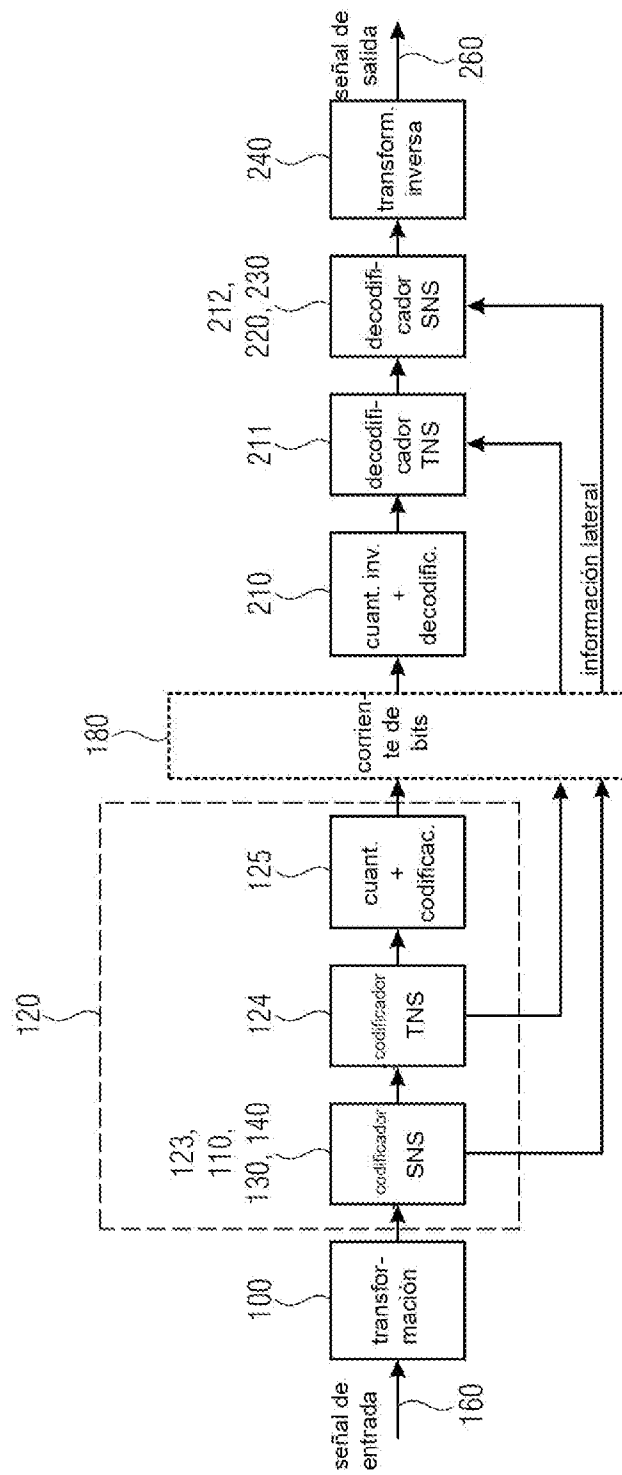


Fig. 7

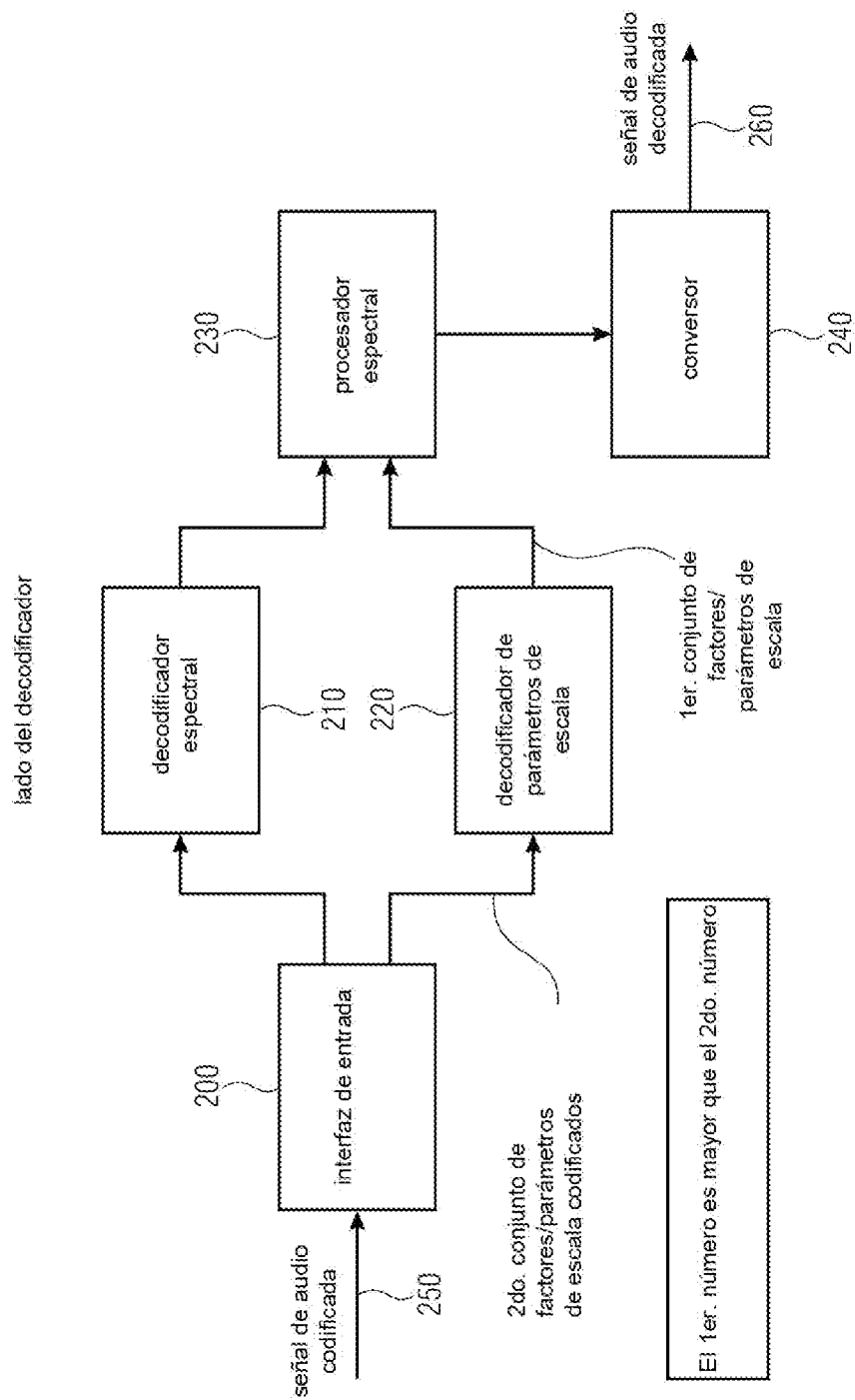


Fig. 8

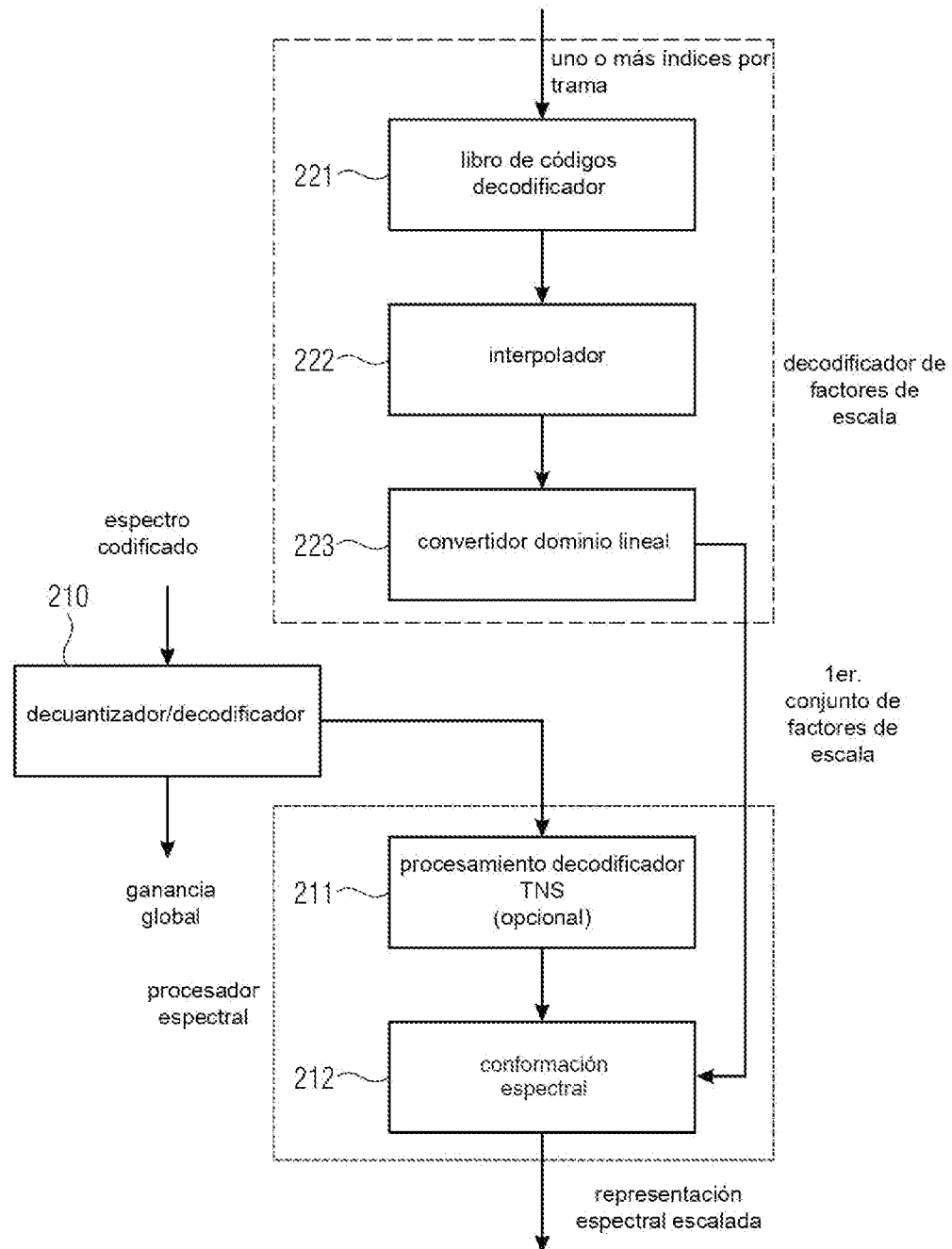


Fig. 9

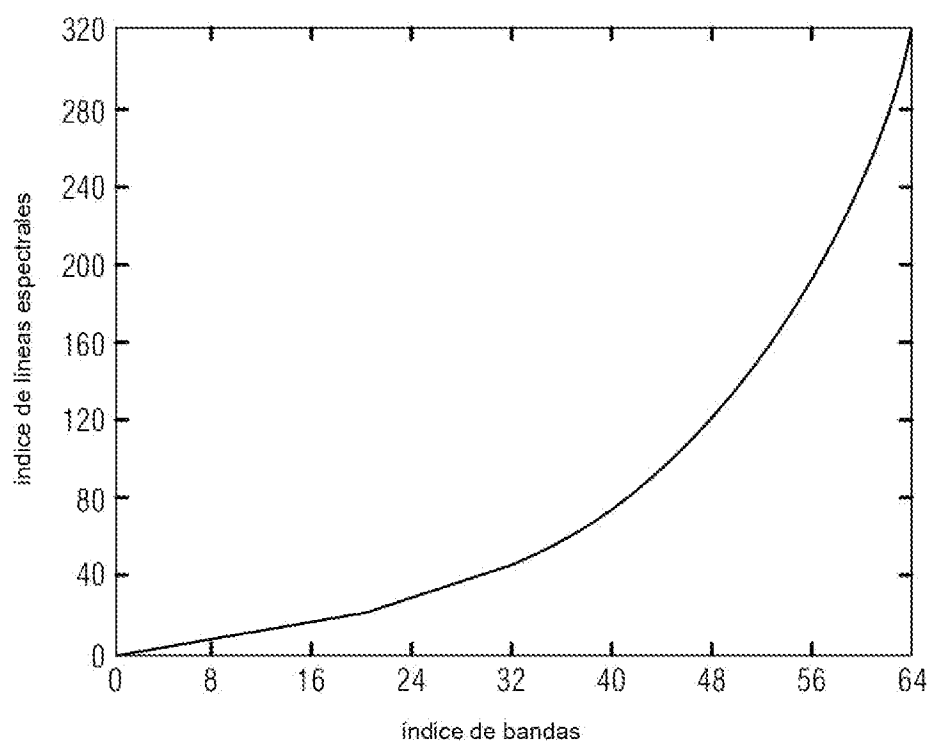


Fig. 10

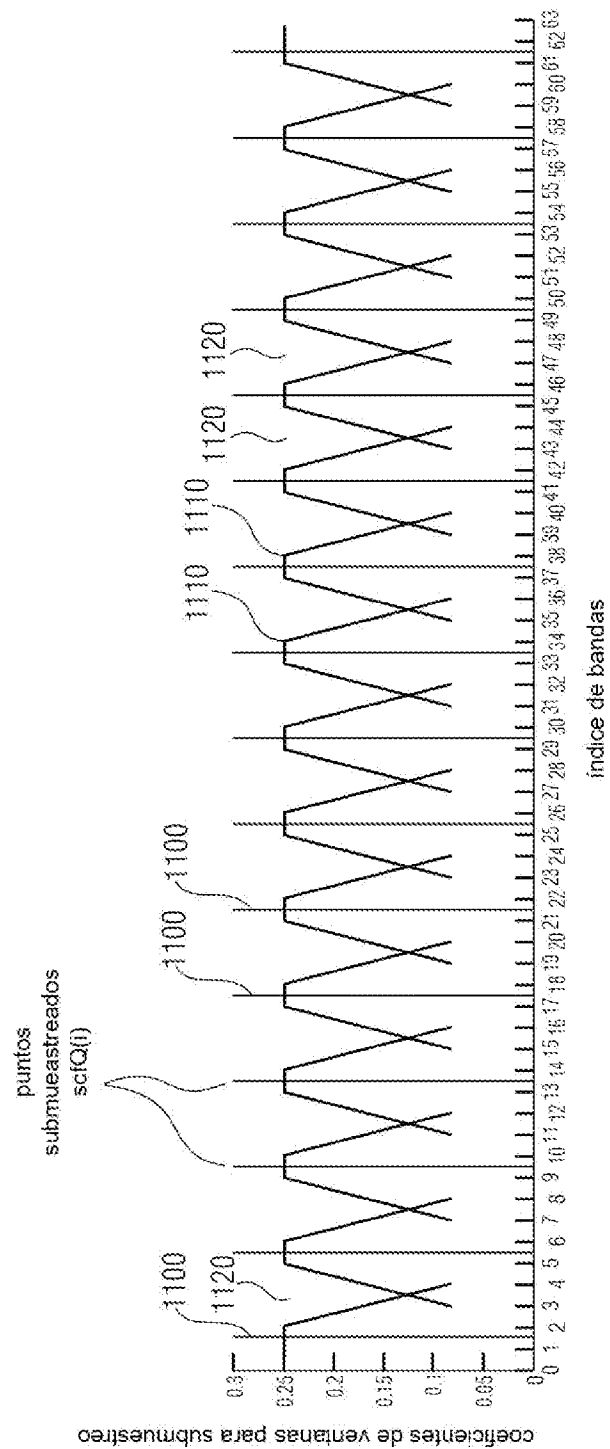


Fig. 11

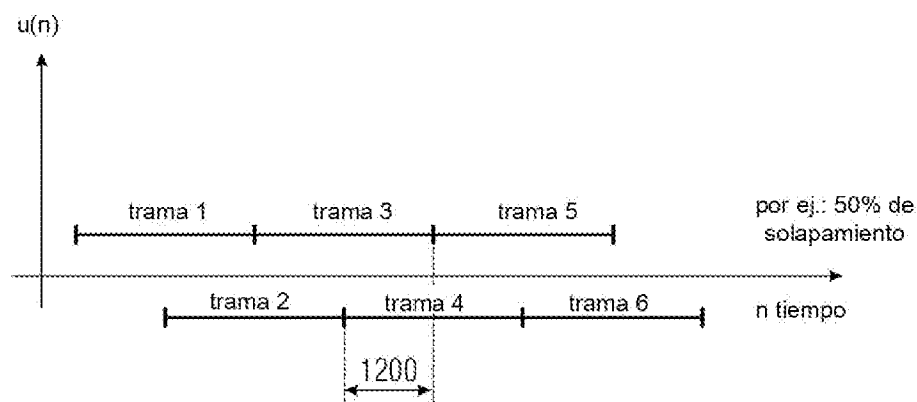


Fig. 12a

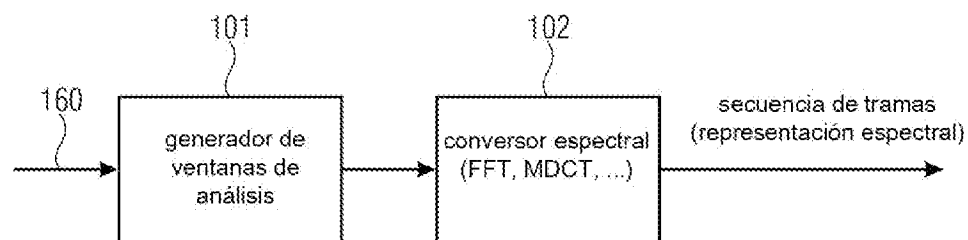


Fig. 12b

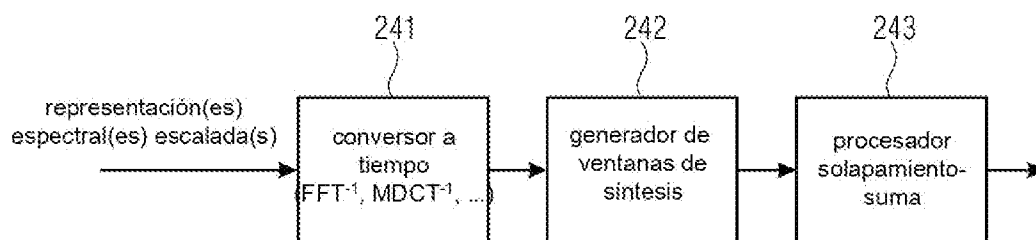


Fig. 12c