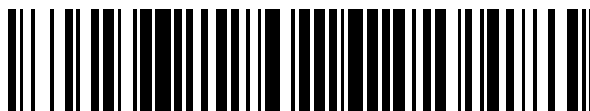


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 796 485**

51 Int. Cl.:

G10L 19/028 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.01.2014** E 18199319 (7)

97 Fecha y número de publicación de la concesión europea: **01.04.2020** EP 3451334

54 Título: **Concepto de llenado con ruido**

30 Prioridad:

29.01.2013 US 201361758209 P

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

27.11.2020

73 Titular/es:

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

72 Inventor/es:

**DISCH, SASCHA;
GAYER, MARC;
HELMRICH, CHRISTIAN;
MARKOVIC, GORAN y
LUIS VALERO, MARIA**

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 796 485 T3

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

DESCRIPCIÓN

Concepto de llenado con ruido

- 5 La presente solicitud se refiere a la codificación audio y, especialmente al llenado con ruido en conexión con la codificación de audio.

10 En la codificación por transformadas con frecuencia se reconoce (comparar [1], [2], [3]) que la cuantificación de partes de un espectro a cero lleva a una degradación perceptual. Esas partes cuantificadas a cero se denominan agujeros espectrales. Una solución a este problema presentada en [1], [2], [3] y [4] consiste en reemplazar las líneas espectrales cuantificadas a cero con ruido. En ocasiones, se evita la inserción de ruido por debajo de una determinada frecuencia. La frecuencia de partida para el llenado con ruido es fija, aunque difiere en la técnica anterior.

15 En ocasiones se utiliza FDNS (Modelado de Ruido en el Dominio de la Frecuencia) para modelar el espectro (incluyendo el ruido insertado) y para el control del ruido de cuantificación, como en USAC (comparar [4]). El FDNS se ejecuta empleando una respuesta de magnitud del filtro LPC. Los coeficientes de filtro LPC se calculan empleando una señal de entrada pre-enfatizada.

20 En [1] se señaló que la adición de ruido en la proximidad inmediata de un componente tonal lleva a una degradación y, en consecuencia, tal como en [5] solo se llenan con ruido largas sucesiones de ceros para evitar que los valores no cero queden ocultos por el ruido circundante inyectado.

25 En [3] se señala que existe un problema de un compromiso entre la granularidad del llenado con ruido y el tamaño de la información complementaria necesaria. En [1], [2], [3] y [5] se transmite un parámetro de llenado con ruido por espectro completo. Se modela espectralmente el ruido insertado utilizando LPC como en [2] o utilizando factores de escala como en [3]. En [3] se describe cómo se deben adaptar los factores de escala a un llenado con ruido con un nivel de llenado con ruido para la totalidad del espectro. En [3], se modifican los factores de escala correspondientes a las bandas que se cuantifican por completo a cero para evitar agujeros espectrales y para tener un correcto nivel de ruido.

30 Aunque las soluciones de [1] y [5] evitan la degradación de los componentes tonales puesto que sugieren no llenar los pequeños agujeros espectrales, subsiste la necesidad de mejorar aún más la calidad de la señal de audio codificada utilizando llenado con ruido, especialmente a muy bajas tasas de bits.

35 El documento US 2012/0046955 A1 se refiere a un esquema para inyectar ruido en elementos no codificados de un espectro que está controlado de acuerdo con una medida de una distribución de energía del espectro original entre las ubicaciones de los elementos no codificados.

40 El objetivo de la presente invención es dar a conocer un concepto para el llenado con ruido con características mejoradas.

Este objetivo se alcanza mediante la materia expuesta en las reivindicaciones independientes adjuntas a la presente, donde los aspectos ventajosos de la presente solicitud son el tema de las reivindicaciones dependientes.

45 Es un hallazgo básico de la presente solicitud que el llenado con ruido del espectro de una señal de audio puede tener una calidad mejorada con respecto al espectro llenado con ruido por lo que la reproducción de la señal de audio llenada con ruido es menos molesta, ejecutando el llenado con ruido de manera dependiente de una tonalidad de la señal de audio.

50 De acuerdo con una realización de la presente solicitud, se llena una porción espectral contigua de cero del espectro de la señal de audio con ruido modelado espectralmente utilizando una función que asume un máximo en el interior de la porción espectral contigua de cero, y que presenta bordes que caen en el exterior cuya pendiente absoluta depende negativamente de la tonalidad, es decir, que la pendiente decrece con el aumento de la tonalidad. Además o por otro lado, la función empleada para el llenado presume un máximo en la parte interna de la porción espectral contigua de cero y tiene bordes que caen en el exterior, cuyo ancho espectral depende positivamente de la tonalidad, es decir el ancho espectral se incrementa con el aumento de la tonalidad. Por añadidura, además o por otro lado, se puede utilizar una función constante o unimodal para llenar, una integral de la cual – normalizada a una integral de 1 – en los cuartos externos de la porción espectral contigua de cero depende negativamente de la tonalidad, es decir que la integral se reduce con el aumento de la tonalidad. Con todas estas medidas, el llenado con ruido tiende a ser menos perjudicial para las partes tonales de la señal de audio, aunque de todos modos es eficaz para las partes no tonales de la señal de audio en términos de reducción de los agujeros espectrales. En otras palabras, siempre que la señal de audio tenga un contenido de componentes tonales, el ruido introducido en el espectro de la señal de audio deja indemnes los picos tonales del espectro al mantener una distancia suficiente de los mismos, donde no obstante el carácter no tonal de las fases temporales de la señal de audio con el contenido de audio no tonal es satisfecho por

ES 2 796 485 T3

el llenado con ruido. De conformidad con una forma de realización de la presente solicitud, se identifican las porciones espectrales contiguas de ceros del espectro de la señal de audio y las porciones de ceros identificadas se llenan con ruido espectralmente modelado con funciones de manera tal que, por cada porción espectral contigua de cero se fija la respectiva función dependiendo de un ancho de la respectiva porción espectral contigua de cero y de la tonalidad de la señal de audio. Para facilitar la implementación, se puede determinar la dependencia mediante una búsqueda
5 de la señal de audio. Para facilitar la implementación, se pueden calcular las funciones analíticamente utilizando una fórmula matemática dependiendo del ancho de la porción espectral contigua de cero y de la tonalidad de la señal de audio. En todo caso, el esfuerzo para obtener la dependencia es relativamente ínfimo en comparación con las ventajas que se obtienen como resultado de la dependencia. En particular, la dependencia puede ser tal que se fije la respectiva
10 función dependiendo del ancho de la porción espectral contigua de cero de manera que la función quede limitada a la respectiva porción espectral contigua de cero, y dependiendo de la tonalidad de la señal de audio de manera que, para una tonalidad más alta de la señal de audio, la masa de una función se torne más compacta en la parte interna de la respectiva porción espectral contigua de cero y alejada de los bordes de la respectiva porción espectral contigua de cero.

De acuerdo con otra forma de realización, el ruido modelado espectralmente e introducido en las porciones espectrales contiguas de ceros por lo general es escalado utilizando un nivel espectralmente global de llenado con ruido. En particular, el ruido es escalado de tal manera que una integral del ruido en las porciones espectrales contiguas de ceros o una integral sobre las funciones de las porciones espectrales contiguas de ceros corresponda a, por ej., sea
20 igual a, un nivel global de llenado con ruido. Ventajosamente, un nivel global de llenado con ruido es codificado dentro de los códecs de audio existente de manera que no hay necesidad de suministrar sintaxis adicional para esos códecs de audio. En otras palabras, el nivel global de llenado con ruido puede ser explícitamente señalado en el flujo de datos en el cual se codifica la señal de audio con poco esfuerzo. En efecto, las funciones con las cuales el ruido de la porción espectral contigua de cero es modelado espectralmente pueden ser escaladas de tal manera que una integral sobre el ruido con el cual se llenan todas las porciones espectrales contiguas de ceros corresponde al nivel global de
25 llenado con ruido.

De conformidad con una forma de realización de la presente solicitud, la tonalidad se deriva de un parámetro de codificación utilizando el cual se codifica la señal de audio. Mediante esta medida, no es necesario transmitir ninguna información adicional dentro de un códec de audio existente. De acuerdo con formas de realización específicas, el parámetro de codificación es una bandera o ganancia de LTP (Predicción a Largo Plazo), una bandera o ganancia de habilitación de TNS (Modelado Temporal de Ruido) y/o una bandera de habilitación de reordenamiento del espectro.
30

De acuerdo con otra forma de realización, la ejecución del llenado con ruido se limita a una porción espectral de alta frecuencia, donde se fija una posición de partida de baja frecuencia de la porción espectral de alta frecuencia que corresponde a una señalización explícita en un flujo de datos y con la cual se codifica la señal de audio. Merced a esta medida, es factible una determinación adaptativa a la señal del límite inferior de la porción espectral de alta frecuencia en la cual se ejecuta el llenado con ruido. Merced a esta medida, a su vez, se puede aumentar la calidad del audio obtenida como resultado del llenado con ruido. A su vez, la información complementaria adicional necesaria causada por la señalización explícita es comparativamente pequeña.
35
40

De acuerdo con otra forma de realización de la presente solicitud, el aparato está configurado para ejecutar el llenado con ruido empleando un filtro espectral de paso bajo para contrarrestar la distorsión espectral causada por una preénfasis utilizada para codificar el espectro de la señal de audio. Merced a esta medida, la calidad del llenado con ruido se incrementa aún más, puesto que se reduce aún más la profundidad de los agujeros espectrales restantes. Hablando en términos más generales, se puede mejorar el llenado con ruido en los códecs de audio de transformada perceptual ejecutando, además del modelado espectral dependiente de la tonalidad del ruido dentro de los agujeros espectrales, el llenado con ruido con una inclinación espectral global, en lugar de en forma espectralmente plana. Por ejemplo, la inclinación espectralmente global puede tener pendiente negativa, es decir exhibir una reducción de bajas a altas frecuencias, para revertir por lo menos parcialmente la inclinación espectral causada al someter al espectro llenado con ruido a la función de ponderación espectral perceptual. También se puede concebir una pendiente positiva, por ej., en casos en que el espectro codificado exhibe un carácter de tipo paso alto. En particular, las funciones de ponderación perceptual espectral tienen por lo general a exhibir un aumento desde bajas a altas frecuencias. En consecuencia, el ruido introducido en el espectro de los codificadores de audio por transformada perceptual de manera espectralmente plana, terminaría en un piso de ruido inclinado en el espectro reconstruido en última instancia. Los inventores de la presente solicitud, sin embargo, notaron que esta inclinación en el espectro reconstruido en última instancia afecta a la calidad del audio, puesto que da lugar a agujeros espectrales restantes en las partes llenadas con ruido del espectro. En consecuencia, la inserción del ruido con una inclinación espectral global para que el nivel de ruido se reduzca desde las frecuencias bajas a las altas compensa, por lo menos en parte, dicha inclinación espectral causada por el modelado subsiguiente del espectro llenado con ruido usando la función de ponderación espectral perceptual, mejorando así la calidad del audio. Según las circunstancias, puede ser preferible una pendiente positiva, por ej., en ciertos espectros del tipo paso alto.
45
50
55
60

De conformidad con una forma de realización, la pendiente de la inclinación espectralmente global se varía en

ES 2 796 485 T3

- 5 respuesta a una señalización en el flujo de datos en el cual se codifica el espectro. La señalización puede señalar, por ejemplo, de manera explícita, el grado de empujado y se puede adaptar, del lado de la codificación, a la cantidad de inclinación espectral causada por la función de ponderación espectral perceptual. Por ejemplo, la cantidad de inclinación espectral causada por la función de ponderación espectral perceptual se puede originar en un preénfasis a la cual se somete a la señal de audio antes de aplicar el análisis LPC a la misma.
- 10 Se puede utilizar el llenado con ruido del lado de la codificación de audio y/o la decodificación de audio. Cuando se utiliza del lado de la codificación de audio, el espectro llenado con ruido se puede emplear para fines de análisis por síntesis.
- 15 De conformidad con una forma de realización, un codificador determina el nivel de escalado global de ruido tomando en cuenta la dependencia de la tonalidad.
- A continuación se describen las formas de realización preferidas de la presente solicitud en relación con las figuras, en las cuales:
- 20 Fig. 1 ilustra, en forma alineada en el tiempo, uno sobre otro, de arriba hacia abajo, un fragmento de tiempo extraído de una señal de audio, cuyo espectrograma utiliza una variación espectral indicada esquemáticamente "en escala de grises" de la energía espectral y la tonalidad de la señal de audio con fines ilustrativos;
- Fig. 2 ilustra un diagrama de bloques de un aparato de llenado con ruido de conformidad con una forma de realización;
- 25 Fig. 3 ilustra una vista esquemática de un espectro al que se ha de someter a llenado con ruido y una función empleada para el modelado espectral del ruido utilizado para llenar una porción espectral contigua de cero de este espectro de conformidad con una forma de realización;
- Fig. 4 ilustra una vista esquemática de un espectro al que se ha de someter a llenado con ruido y una función empleada para el modelado espectral del ruido utilizado para llenar una porción espectral contigua de cero de este espectro de acuerdo con otra forma de realización;
- 30 Fig. 5 ilustra una vista esquemática de un espectro al que se ha de someter a llenado con ruido y una función empleada para el modelado espectral del ruido utilizado para llenar una porción espectral contigua de cero de este espectro de acuerdo con una forma de realización adicional;
- 35 Fig. 6 ilustra un diagrama de bloques del dispositivo de llenado con ruido de la Fig. 2 de conformidad con una forma de realización;
- Fig. 7 ilustra esquemáticamente una relación posible entre la tonalidad de la señal de audio determinada por un lado y las funciones posibles disponibles para el modelado espectral de una porción espectral contigua de cero por el otro, de conformidad con una forma de realización;
- 40 Fig. 8 ilustra esquemáticamente un espectro que ha de ser llenado con ruido y además exhibe las funciones utilizadas para el modelado espectral del ruido para llenar las porciones espectrales contiguas de ceros del espectro a fin de ilustrar cómo se debe escalar el nivel de ruido de conformidad con una forma de realización;
- 45 Fig. 9 ilustra un diagrama de bloques de un codificador que se puede emplear en un códec de audio que adopta el concepto de llenado con ruido descrito con respecto a las Figs.
- 50 Fig. 10 ilustra esquemáticamente un espectro cuantificado para llenar con ruido, codificado por el codificador de la Fig. 9 junto con información complementaria transmitida, es decir factores de escala y nivel global de ruido, de conformidad con una forma de realización;
- Fig. 11 ilustra un diagrama de bloques de un decodificador que se ajusta al codificador de la Fig. 9 y que incluye un aparato de llenado con ruido de acuerdo con la Fig. 2;
- 55 Fig. 12 ilustra una vista esquemática de un espectrograma con datos de información complementaria asociados de acuerdo con una variante de una implementación del codificador y decodificador de las Figs. 9 y 11;
- 60 Fig. 13 ilustra un codificador de audio por transformada de predicción lineal que se puede incluir en un códec de audio utilizando el concepto de llenado con ruido de las Figs. 1 a 8 de conformidad con una forma de realización;
- Fig. 14 ilustra un diagrama de bloques de un decodificador que se adapta al codificador de la Fig. 13;

Fig. 15 ilustra ejemplos of fragmentos tomados de un espectro que ha de ser llenado con ruido;

Fig. 16 ilustra un ejemplo explícito de una función para modelar el ruido introducido en una determinada porción espectral contigua de cero del espectro que ha de ser llenado con ruido de conformidad con una forma de realización;

5

Figs. 17a–d ilustran diversos ejemplos de funciones para el modelado espectral del ruido introducido en porciones espectrales contiguas de ceros correspondientes a diferentes anchos de porciones de ceros y diferentes anchos de transición empleados para las diferentes tonalidades y

10

Fig. 18a ilustra un diagrama de bloques de un codificador de audio por transformada perceptual de conformidad con una forma de realización;

Fig. 18b ilustra un diagrama de bloques de un decodificador de audio por transformada perceptual i adaptado a la Fig. 18a;

15

Fig. 18c ilustra un diagrama esquemático que muestra una manera posible de obtener la inclinación espectralmente global introducida en el ruido introducido de conformidad con una forma de realización.

20

En toda la siguiente descripción de las figuras, se utilizan signos de referencia iguales para los elementos expuestos en estas figuras; se ha de interpretar que la descripción presentada con respecto a un elemento de una figura se puede transferir al elemento de otra figura que está indicado con el mismo signo de referencia. En virtud de esta medida, se evita en lo posible una descripción extensa y repetitiva, concentrando así la descripción de las diversas formas de realización en las diferencias entre sí en lugar de describir nuevamente y desde el comienzo todas las realizaciones una y otra vez.

25

La siguiente descripción se inicia con las formas de realización correspondientes a un aparato para ejecutar el llenado con ruido en el espectro de una señal de audio, en primer lugar. En segundo lugar se presentan diferentes formas de realización correspondientes a diversos códecs de audio, a los cuales se puede incorporar dicho llenado con ruido, junto con detalles específicos que se podrían aplicar en conexión con un respectivo códec de audio presentado. Cabe señalar que el llenado con ruido que se describe a continuación puede ser ejecutado, en todo caso, del lado de la decodificación. Sin embargo, el llenado con ruido que se describe a continuación también se ejecuta, en todo caso, del lado de la codificación tal como, por ejemplo, por razones de análisis por síntesis. A continuación también se describe un caso intermedio de acuerdo con el cual la forma modificada de llenado con ruido de acuerdo con las formas de realización descritas a continuación simplemente cambia parcialmente la manera de funcionar del codificador como por ejemplo, para determinar un nivel espectralmente global de llenado con ruido.

30

35

40

La Fig. 1 presenta, con fines ilustrativos, una señal de audio 10, es decir el curso temporal de sus muestras de audio, por ejemplo, el espectrograma alineado en el tiempo 12 de la señal de audio que ha sido derivado de la señal de audio 10, por lo menos entre otros modos, por medio de una transformación adecuada tal como una transformación solapada ilustrada en 14 a título de ejemplo respecto de dos ventanas consecutivas de transformada 16 y los espectros asociados 18 que, de esa manera, representan un corte de un espectrograma 12 en una instancia de tiempo que corresponde a un punto medio de la ventana de transformada asociada 16, por ejemplo. Más adelante se presentan ejemplos de espectrograma 12 y de cómo se obtiene el mismo. En todo caso, el espectrograma 12 ha sido objeto de algún tipo de cuantificación y, por consiguiente, tiene porciones de ceros donde los valores espectrales en los cuales se muestrea espectralmente el espectrograma 12 son contiguas de ceros. La transformada solapada 14 puede ser, por ejemplo, una transformada con muestreo crítico tal como una MDCT. Las ventanas de transformada 16 pueden tener un traslape de 50 % entre sí, aunque también son factibles otras formas de realización. Además, la resolución espectralmente con la cual se muestrea el espectrograma 12 en los valores espectrales puede variar en el tiempo. En otras palabras, la distancia temporal entre espectros consecutivos 18 del espectrograma 12 puede variar en el tiempo y lo mismo se aplica a la resolución espectral de cada espectro 18. En particular, la variación en tiempo en lo que se refiere a la distancia temporal entre espectros consecutivos 18, puede ser inversa a la variación de la resolución espectral de los espectros. La cuantificación utiliza, por ejemplo, un tamaño de paso de cuantificación adaptativa a la señal con variación espectral, que varía, por ejemplo, de acuerdo con una envolvente espectral de LPC de la señal de audio descrita por los coeficientes de LP señalizados en el flujo de datos en el cual se codifican los valores espectrales cuantificados del espectrograma 12 con los espectros 18 que se han de llenar con ruido, o de acuerdo con factores de escala determinados, a su vez, de acuerdo con un modelo psicoacústico, y señalizados en el flujo de datos.

45

50

55

60

Más allá de eso, en forma alineada en el tiempo, la Fig. 1 ilustra una característica de la señal de audio 10 y su variación temporal, es decir la tonalidad de la señal de audio. En términos generales, la "tonalidad" indica una medida que describe cuán condensada está la energía de la señal de audio en un determinado punto de tiempo en el espectro respectivo 18 asociado a ese punto de tiempo. Si la energía está muy dispersa, como por ejemplo en las fases temporales ruidosas de la señal de audio 10, la tonalidad es baja. Pero si la energía está sustancialmente concentrada en uno o más picos espectrales, en ese caso la tonalidad es alta.

ES 2 796 485 T3

La Fig. 2 ilustra un aparato configurado para ejecutar el llenado con ruido en el espectro de una señal de audio de conformidad con una forma de realización de la presente solicitud. Como se describe más adelante en forma más detallada, el aparato está configurado para ejecutar el llenado con ruido dependiendo de la tonalidad de la señal de audio.

5

El aparato de la Fig. 2 está indicado en general empleando el signo de referencia 30 y comprende un dispositivo de llenado con ruido 32 y un determinador de la tonalidad 34, que es opcional.

10 El llenado con ruido en sí es ejecutado por el dispositivo de llenado con ruido 32. El dispositivo de llenado con ruido 32 recibe el espectro al cual se ha de aplicar el llenado con ruido. Este espectro está ilustrado en la Fig. 2 en forma de espectro ralo 34. El espectro ralo 34 puede ser un espectro 18 tomado del espectrograma 12. Los espectros 18 ingresan al dispositivo de llenado con ruido 32 sucesivamente. El dispositivo de llenado con ruido 32 somete al espectro 34 a llenado con ruido y emite el "espectro llenado" 36. El dispositivo de llenado con ruido 32 ejecuta el llenado con ruido dependiendo de la tonalidad de la señal de audio, tal como la tonalidad 20 de la Fig. 1. Dependiendo de las circunstancias, la tonalidad puede no estar directamente disponible. Por ejemplo, los códecs de audio existentes no dan lugar a una señalización explícita de la tonalidad de la señal de audio en el flujo de datos, por lo que si el aparato 30 es instalado del lado de la decodificación, no sería factible reconstruir la tonalidad sin un alto grado de estimación falsa. Por ejemplo, el espectro 34 puede no ser una base óptima para el cálculo de la tonalidad debido a su poca densidad y/o debido a su cuantificación variante adaptativa a la señal.

15

20 En consecuencia, es la tarea del determinador de tonalidades 34 suministrar al dispositivo de llenado con ruido 32 una estimación de la tonalidad sobre la base de otro indicio de tonalidad 38 como se describe más adelante en forma más detallada. De acuerdo con las formas de realización que se describen más adelante, el indicio de tonalidad 38 puede estar disponible de todas maneras en los lados de la codificación y la decodificación, merced a un respectivo parámetro de codificación acarreado en el flujo de datos del códec de audio dentro del cual se utiliza, por ejemplo, el aparato 30.

25

La Fig. 3 ilustra un ejemplo correspondiente al espectro ralo 34, es decir un espectro cuantificado que consta de porciones contiguas 40 y 42 que consisten en sucesiones de valores espectrales espectralmente cercanos del espectro 34, que se cuantifican a cero. Por consiguiente, las porciones contiguas 40 y 42 están separadas o distanciadadas espectralmente entre sí por al menos una línea espectral no cuantificada a cero en el espectro 34.

30

La dependencia de la tonalidad del llenado con ruido descrito anteriormente en general con respecto a la Fig. 2 puede ser implementada de la siguiente manera. La Fig. 3 ilustra una porción temporal 44 que incluye una porción espectral contigua de cero 40, exagerada en 46. El dispositivo de llenado con ruido 32 está configurado para llenar esta porción espectral contigua de cero 40 de manera dependiente de la tonalidad de la señal de audio en el momento a que el espectro 34 pertenece. En particular, el dispositivo de llenado con ruido 32 llena la porción espectral contigua de cero con ruido modelado espectralmente utilizando una función que presume un máximo en el interior de la porción espectral contigua de cero, y que presenta bordes que caen en el exterior, cuya pendiente absoluta depende negativamente de la tonalidad. La Fig. 3 ilustra a manera de ejemplo dos funciones 48 correspondientes a dos tonalidades diferentes. Ambas funciones son "unimodales", es decir que presumen un máximo absoluto en la parte interna de la porción espectral contigua de cero 40 y tienen meramente un máximo local que puede ser una meseta o una única frecuencia espectral. En este caso, el máximo local es asumido por las funciones 48 y 50 en forma continua en todo un intervalo extendido 52, es decir una meseta, dispuesta en el centro de la porción de cero 40. El dominio de las funciones 48 y 50 es la porción de cero 40. El intervalo central 52 cubre solamente la porción central de la porción de cero 40 y está flanqueado por una porción de borde 54 en un extremo de frecuencia más elevada del intervalo 52, y una porción de borde de frecuencia más baja 56 en un extremo de menor frecuencia del intervalo 52. Dentro de la porción de borde 54, las funciones 48 y 52 tienen un borde de bajada 58, y dentro de la porción de borde 56, un borde de subida 60. Se puede atribuir una pendiente absoluta a cada borde 58 y 60, respectivamente, tal como la pendiente media dentro de la porción de borde 54 y 56, respectivamente. En otras palabras, la pendiente atribuida al borde de bajada 58 puede ser la pendiente media de la respectiva función 48 y 52, respectivamente, dentro de la porción de borde 54 y la pendiente atribuida al borde de subida 60 puede ser la pendiente media de la función 48 y 52, respectivamente, dentro de la porción de borde 56.

35

40

45

50

Como se puede apreciar, el valor absoluto de la pendiente de los bordes 58 y 60 es más elevada en el caso de la función 50 que en el de la función 48. El dispositivo de llenado con ruido 32 opta por llenar la porción de cero 40 con la función 50 para tonalidades más bajas que las tonalidades para las cuales el dispositivo de llenado con ruido 32 elige usar la función 48 para llenar la porción de cero 40. Merced a esta medida, el dispositivo de llenado con ruido 32 evita la aglomeración en la periferia inmediata de los picos espectrales potencialmente tonales del espectro 34, como por ejemplo, el pico 62. Cuanto menor es la pendiente absoluta de los bordes 58 y 60, más lejos está el ruido introducido en la porción de cero 40 de las porciones no cero del espectro 34 que rodean a la porción de cero 40.

55

60

El dispositivo de llenado con ruido 32 puede optar, por ejemplo, por seleccionar la función 48 en un caso en que la tonalidad de la señal de audio es τ_2 , y la función 50 en caso de que la tonalidad de la señal de audio sea τ_1 , aunque la descripción presentada más adelante ha de revelar que el dispositivo de llenado con ruido 32 puede discriminar entre

ES 2 796 485 T3

más de dos estados diferentes de la tonalidad de la señal de audio, es decir, puede admitir más de dos funciones diferentes 48, 50 para llenar una determinada porción espectral contigua de cero y elegir entre ellas dependiendo de la tonalidad mediante un mapeo sobreectivo de tonalidades a funciones.

5 Como nota menor, cabe señalar que la construcción de las funciones 48 y 50 de acuerdo con las cuales estas tienen una meseta en el intervalo interno 52, flanqueado por los bordes 58 y 60 para dar origen a funciones unimodales, es meramente un ejemplo. Por otro lado, se pueden utilizar funciones acampadas, por ejemplo, de acuerdo con una alternativa. El intervalo 52 se puede definir, por otro lado, como intervalo en el cual la función es superior a 95 % de su valor máximo.

10 La Fig. 4 ilustra una alternativa correspondiente a la variación de la función empleada para el modelado espectral del ruido con el cual el dispositivo de llenado con ruido 32 llena cierta porción espectral contigua de cero 40 y, en la tonalidad. De acuerdo con la Fig. 4, la variación se refiere al ancho espectral de las porciones de bordes 54 y 56 y los bordes de bajada externos 58 y 60, respectivamente. Como se ilustra en la Fig. 4, de acuerdo con el ejemplo de la Fig. 4, la pendiente de los bordes 58 y 60 pueden ser incluso independientes de la tonalidad, es decir no cambiar con esta. En particular, de acuerdo con el ejemplo de la Fig. 4, el dispositivo de llenado con ruido 32 establece la función empleando la cual se modela espectralmente el ruido para llenar la porción de cero 40 de tal manera que el ancho espectral de los bordes de bajada hacia el exterior 58 y 60 dependa positivamente de la tonalidad, es decir que para tonalidades más altas, se utiliza la función 48 por la cual el ancho espectral de los bordes de bajada hacia el exterior 58 y 60 es mayor, y en el caso de las tonalidades más bajas, se utiliza la función 50 respecto de la cual el ancho espectral de los bordes de bajada hacia el exterior 58 y 60 es menor.

La Fig. 4 ilustra otro ejemplo de variación de una función empleada por el dispositivo de llenado con ruido 32 para el modelado espectral del ruido con el cual se llena la porción espectral contigua de cero 40: en este caso, la característica de la función que varía con la tonalidad es la integral sobre los cuartos externos de la porción de cero 40. Cuanto más alta es la tonalidad, mayor es el intervalo. Antes de determinar el intervalo, se eculiza/normaliza el intervalo total de la función sobre la porción de cero completa 40, por ejemplo a 1.

Para explicar esto, véase la Fig. 5. La porción espectral contigua de cero 40 aparece dividida en cuatro cuartos de igual tamaño a, b, c, d, entre los cuales los cuartos a y d son cuartos externos. Como se puede apreciar, ambas funciones 50 y 48 tienen su centro de masa en la parte interna, en este caso, a título de ejemplo, en el punto medio de la porción de cero 40, aunque ambas se extienden desde los cuartos internos b, c hacia los cuartos externos a y d. La porción traslapada de las funciones 48 y 50, que se traslapa sobre los cuartos externos a y d, respectivamente, aparece simplemente sombreada.

En la Fig. 5, ambas funciones tienen la misma integral sobre toda la porción de cero 40, es decir sobre los cuatro cuartos a, b, c, d. La integral se normaliza, por ejemplo, a 1.

En esta situación, la integral de la función 50 sobre los cuartos a, d es mayor que la integral de la función 48 sobre los cuartos a, d y en consecuencia, el dispositivo de llenado con ruido 32 utiliza la función 50 para las tonalidades más altas y la función 48 para las tonalidades más bajas, es decir que la integral sobre los cuartos externos de las funciones normalizadas 50 y 48 depende negativamente de la tonalidad.

Con fines ilustrativos, en el caso de la Fig. 5 las dos funciones 48 y 50 han sido expuestas, a título de ejemplo, como funciones constantes o binarias. La función 50, por ejemplo, es una función que asume un valor constante en todo el dominio, es decir que la totalidad de la porción de cero 40, y la función 48 es una función binaria que es cero en los bordes externos de la porción de cero 40, y que asume un valor constante no cero entre las mismas. Debe resultar evidente que, en términos generales, las funciones 50 y 48 de acuerdo con el ejemplo de la Fig. 5 puede ser cualquier función constante o unimodal tal como las que corresponden a las expuestas en las Figs. 3 y 4. Para ser más precisos, por lo menos una puede ser unimodal y por lo menos una constante (por partes) y otras potenciales pueden ser unimodales o constantes.

Aunque el tipo de variación de las funciones 48 y 50 dependiendo de la tonalidad varía, todos los ejemplos de las Figs. 3 a 5 tienen en común que, para aumentar la tonalidad, se reduce el grado de dispersión en las proximidades inmediatas de picos tonales en el espectro 34 o se evita, por lo que la calidad del llenado con ruido se incrementa, puesto que el llenado con ruido no afecta negativamente a las fases tonales de la señal de audio y de todas maneras da lugar a una aproximación agradable de las fases no tonales de la señal de audio.

Hasta ahora la descripción de las Figs. 3 a 5 se enfocó en el llenado de una porción espectral contigua de cero. De acuerdo con una forma de realización de la Fig. 6, el aparato de la Fig. 2 está configurado para identificar porciones espectrales contiguas de ceros del espectro de la señal de audio y para aplicar el llenado con ruido a las porciones espectrales contiguas de ceros así identificadas. En particular, la Fig. 6 ilustra el dispositivo de llenado con ruido 32 de la Fig. 2 en más detalle mostrando que comprende un identificador de porciones de ceros 70 y un dispositivo de llenado de la porción de cero 72. El identificador de porciones de ceros busca en el espectro 34 las porciones

ES 2 796 485 T3

espectrales contiguas de ceros tales como 40 y 42 de la Fig. 3. Como ya se describiera anteriormente, las porciones espectrales contiguas de ceros se pueden definir como sucesiones de valores espectrales que han sido cuantificados a cero. El identificador de porciones de cero 70 puede estar configurado para confinar la identificación a una porción 5
espectral de alta frecuencia del espectro de la señal de audio a partir, es decir situado más arriba, de cierta frecuencia inicial. En consecuencia, el aparato puede estar configurado para confinar la ejecución del llenado con ruido a esa porción 10
espectral de alta frecuencia. La frecuencia inicial por encima de la cual el identificador de porciones de cero 70 ejecuta la identificación de porciones espectrales contiguas de cero, y por encima de la cual el aparato está configurado para confinar la ejecución del llenado con ruido, puede ser fija o puede variar. Por ejemplo, se puede emplear la señalización explícita en el flujo de datos de una señal de audio en el cual se codifica la señal de audio a través de su espectro para señalar la frecuencia inicial que se ha de utilizar.

El dispositivo de llenado de porciones de cero 72 está configurado para llenar las porciones espectrales contiguas de ceros identificadas por el identificador 70 con ruido modelado espectralmente de acuerdo con una función antes descrita con respecto a la Fig. 3, 4 o 5. En consecuencia, el dispositivo de llenado de porciones de cero 72 llena las 15
porciones espectrales contiguas de ceros identificadas por el identificador 70 con funciones establecidas dependiendo de un ancho de la respectiva porción espectral contigua de cero, como por ejemplo del número de valores espectrales que han sido cuantificados a cero de la sucesión de valores espectrales cuantificados a cero de la respectiva porción espectral contigua de cero y de la tonalidad de la señal de audio.

En particular, el llenado individual de cada porción espectral contigua de cero identificada por el identificador 70 puede ser ejecutado por el dispositivo de llenado 72 de la siguiente manera: la función se establece de conformidad con el ancho de la porción espectral contigua de cero por lo que la función queda confinada a la respectiva porción espectral contigua de cero, es decir que el dominio de la función coincide con el ancho de la porción espectral contigua de cero. El establecimiento de la función depende asimismo de la tonalidad de la señal de audio, es decir, de la manera antes reseñada con respecto a las Figs. 3 a 5, por lo que, si la tonalidad de la señal de audio aumenta, la masa de la función se torna más compacta en la parte interna de la respectiva porción contigua de cero y alejada de los bordes de la respectiva porción espectral contigua de cero. Usando esta función, se fija el estado preliminarmente llenado de la porción espectral contigua de cero de acuerdo con la cual se fija cada valor espectral en un valor aleatorio, pseudo-aleatorio o parcheado/copiado, se modela espectralmente, es decir multiplicando la función por los valores espectrales 20
preliminares. 25
30

Ya se ha reseñado que la dependencia del llenado con ruido puede discriminar entre más de solo dos tonalidades diferentes, como por ejemplo 3, 4 o incluso más de 4. La Fig. 7, por ejemplo, ilustra el dominio de las tonalidades posibles, es decir el intervalo de posibles valores entre tonalidades, determinados por el determinador 34 en el número de referencia 74. En 76, la Fig. 7 ilustra a título de ejemplo la serie de funciones posibles empleadas para el modelado espectral del ruido con el cual se pueden llenar las porciones espectrales contiguas de cero. La serie 76 ilustrada en la Fig. 7 es una serie de instanciaciones de funciones discretas que se distinguen mutuamente entre sí por el ancho espectral o la longitud y/o la forma del dominio, es decir la compactación y la distancia de los bordes externos. En 78, la Fig. 7 ilustra más detalladamente el dominio de los anchos posibles de la porción de cero. Aunque el intervalo 78 es un intervalo de valores discretos que oscilan entre un ancho mínimo y cierto ancho máximo, los valores de tonalidad emitidos por el determinador 34 para medir la tonalidad de la señal de audio pueden tener valores enteros o de algún otro tipo tales como valores de punto flotante. El mapeo del par de intervalos 74 y 78 con la serie de funciones posibles 76 se puede obtener mediante búsqueda en tablas o utilizando una función matemática. Por ejemplo, en el caso de una determinada porción espectral contigua de cero identificada por el identificador 70, el dispositivo de llenado de porciones de cero 72 puede utilizar el ancho de la respectiva porción espectral contigua de cero y la tonalidad actual determinada por el determinador 34 a fin de realizar una búsqueda en una tabla una función de la serie 76 definida, por ejemplo, como secuencia de valores de función, donde la longitud de la secuencia coincide con el ancho de la porción espectral contigua de cero. Por otro lado, el dispositivo de llenado de porciones de cero 72 busca parámetros de función e introduce estos parámetros de función en una función predeterminada a fin de derivar la función que se ha de utilizar para el modelado espectral del ruido que se ha de introducir en la respectiva porción espectral contigua de cero. En otra alternativa, el dispositivo de llenado de porciones de cero 72 puede insertar directamente el ancho de la respectiva porción espectral contigua de cero y la tonalidad actual en una fórmula matemática para arribar a los parámetros de función a fin de erigir la respectiva función de acuerdo con los parámetros de función computados matemáticamente. 35
40
45
50
55

Hasta ahora, la descripción de ciertas formas de realización de la presente solicitud se enfocó en la forma de la función utilizada para el modelado espectral del ruido con el cual se llenan ciertas porciones espectrales contiguas de cero. Sin embargo, es ventajoso controlar el nivel general de ruido agregado a un determinado espectro para ser llenado con ruido a fin de dar lugar a una reconstrucción agradable, o incluso controlar espectralmente el nivel de introducción de ruido. 60

La Fig. 8 ilustra un espectro que ha de ser llenado con ruido, donde las porciones no cuantificadas a cero y, en consecuencia, no sometidas a llenado con ruido, están indicadas con un cuadrículado, donde tres porciones espectrales contiguas de cero 90, 92 y 94 aparecen en un estado previo al llenado ilustrado por la porción de cero en

ES 2 796 485 T3

la cual se ha inscrito la función seleccionada para el modelado espectral del ruido introducido en estas porciones 90–94, usando una escala “de descuido” (don’t-care).

De acuerdo con una forma de realización, toda la serie de funciones disponibles 48, 50 para el modelado espectral del ruido que se ha de introducir en las porciones 90–94, tiene una escala predefinida conocida por el codificador y el decodificador. Se señala un factor de escala espectralmente global dentro del flujo de datos en el cual se codifica la señal de audio, es decir la parte no cuantificada del espectro. Este factor indica, por ejemplo, el RMS u otra medida correspondiente a un nivel de ruido, es decir valores de líneas espectrales aleatorios o pseudoaleatorios, con los cuales las porciones 90–94 son preestablecidas del lado de la decodificación para luego modelarlas espectralmente utilizando las funciones dependientes de la tonalidad 48, 50 tal como están. En cuanto a cómo se puede determinar el factor de escala de ruido total del lado del codificador, esto se describe más adelante. Digamos, por ejemplo, que A es la serie de índices i de líneas espectrales en que el espectro es cuantificado a cero y que pertenecen a cualquiera de las porciones 90–94, y digamos que N denota el factor de escala de ruido total. Los valores del espectro se indican como x_i . Además, “aleatorio(N)” ha de indicar una función que da un valor aleatorio de un nivel correspondiente al nivel “N” y restante(i) ha de ser una función que indica, por cualquier valor espectral cuantificado a cero en el índice i , el índice del valor cuantificado a cero en el extremo de baja frecuencia de la porción de cero a la cual i pertenece, y $F_i(j)$ donde $j=0$ a $J_i - 1$ debe indicar la función 48 o 50 que se asigna, dependiendo de la tonalidad, a la porción de cero 90–94 a partir del índice i , donde J_i indica el ancho de esa porción de cero. Luego, las porciones 90–94 se llenan de acuerdo con $x_i = F_{\text{restante}(i)}(i - \text{restante}(i)) \cdot \text{aleatorio}(N)$.

Además, el llenado con ruido de las porciones 90–94, puede ser controlado de tal manera que el nivel de ruido se reduzca de bajas a altas frecuencias. Esto se puede hacer mediante el modelado espectral del ruido con el cual se preestablecen las funciones, o el modelado espectral de la disposición de las funciones 48, 50 de acuerdo con una función de transferencia de filtro de paso bajo. Esto puede compensar una inclinación espectral causada al reescalar/descuantificar el espectro llenado debido, por ejemplo, a un preénfasis utilizado en la determinación del curso espectral del tamaño del paso de cuantificación. En consecuencia, se puede controlar lo abrupto de la reducción o la función de transferencia de filtro de paso bajo de acuerdo con un grado de preénfasis aplicado. Aplicando la nomenclatura antes utilizada, las porciones 90–94 pueden ser llenadas de acuerdo con $x_i = F_{\text{restante}(i)}(i - \text{restante}(i)) \cdot \text{aleatorio}(N) \cdot \text{LPF}(i)$ donde $\text{LPF}(i)$ denota la función de transferencia de filtro de baja frecuencia que puede ser lineal. Dependiendo de las circunstancias, la función LPF que corresponde a la función 15 puede tener una pendiente positiva y LPF se cambia para leer HPF de manera consiguiente.

En lugar de usar un escalado fijo de las funciones seleccionadas dependiendo de la tonalidad y el ancho de la porción de cero, la corrección de la inclinación espectral que se acaba de reseñar puede ser tenida en cuenta directamente utilizando la posición espectral de la respectiva de la porción contigua de cero también como índice para la búsqueda o de otro modo determinando 80 la función que se ha de utilizar para el modelado espectral del ruido con el cual se debe llenar la respectiva porción espectral contigua de cero. Por ejemplo, un valor medio de la función o su preescalado utilizado para el modelado espectral del ruido con que se ha de llenar una cierta porción de cero 90–94 puede depender de la posición espectral de la porción de cero 90–94 por lo que, en todo el ancho de banda del espectro, las funciones usadas para las porciones espectrales contiguas de cero 90–94 son preescaladas a fin de emular una función de transferencia de filtro de paso bajo para compensar cualquier función de transferencia de preénfasis de paso alto utilizada para derivar las porciones cuantificadas no cero del espectro.

Habiendo descrito las formas de realización atinentes a la ejecución del llenado con ruido, a continuación se presentan las formas de realización correspondientes a códecs de audio, a los que se puede incorporar ventajosamente el llenado con ruido antes reseñado. Las Figs. 9 y 10, por ejemplo, ilustran un par de codificador y decodificador, respectivamente, que juntos implementan un códec de audio perceptual basado en transformada del tipo que constituye la base de, por ejemplo, AAC (Codificación Avanzada de Audio). El codificador 100 expuesto en la Fig. 9 somete a la señal de audio original 102 a una transformada en un transformador 104. La transformación ejecutada por el transformador 104 es, por ejemplo, una transformada solapada que corresponde a una transformación 14 de la Fig. 1: descompone espectralmente la señal de audio original entrante 102 sometiendo a ventanas de transformada consecutivas, mutuamente solapadas de la señal de audio original a una secuencia de espectros 18 que, en conjunto, componen el espectrograma 12. Como se indicara anteriormente, el parche de la ventana entre transformada que define la resolución temporal del espectrograma 12 puede variar en tiempo, tal como puede variar la longitud temporal de las ventanas de transformada, lo que define la resolución espectral de cada espectro 18. El codificador 100 comprende asimismo un modelador perceptual 106 que deriva de la señal de audio original, sobre la base de la versión en el dominio del tiempo que entra al transformador 104 o de la versión espectralmente descompuesta emitida por el transformador 104, un umbral de enmascaramiento perceptual que define una curva espectral por debajo de la cual se puede ocultar el ruido de la cuantificación para que este no sea perceptible.

La representación por líneas espectrales de la señal de audio, es decir el espectrograma 12, y el umbral de enmascaramiento entran al cuantificador 108 que es responsable de la cuantificación de las muestras espectrales del espectrograma 12 utilizando un tamaño de paso de cuantificación que varía espectralmente que depende del umbral de enmascaramiento: cuanto más grande es el umbral de enmascaramiento, menor es el tamaño del paso de

- cuantificación. En particular, el cuantificador 108 informa al lado de la decodificación sobre la variación del tamaño del paso de cuantificación en forma de los denominados factores de escala que, mediante la relación recién presentada entre el tamaño del paso de cuantificación por un lado y el umbral de enmascaramiento perceptual por el otro, representan un tipo de representación del umbral de enmascaramiento perceptual en sí. Para hallar un buen
- 5 compromiso entre la cantidad de información complementaria que se ha de insumir para la transmisión de los factores de escala al lado de la decodificación, y la granularidad de la adaptación del ruido de la cuantificación al umbral de enmascaramiento perceptual, el cuantificador 108 fija/ varía los factores de escala en una resolución espectrot temporal que es menor, o más gruesa que, la resolución espectrot temporal a la cual los niveles espectrales cuantificados describen la representación por líneas espectrales del espectrograma de la señal de audio 12. Por ejemplo, el
- 10 cuantificador 108 subdivide cada espectro en bandas de factores de escala 110 tales como bandas bark y transmite un factor de escala por cada banda de factores de escala 110. En lo que respecta a la resolución temporal, esta también puede ser más baja en lo concerniente a la transmisión de los factores de escala, en comparación con los niveles espectrales de los valores espectrales del espectrograma 12.
- 15 Tanto los niveles espectrales de los valores espectrales del espectrograma 12, como también los factores de escala 112 se transmiten al lado de la decodificación. Sin embargo, para mejorar la calidad del audio, el codificador 100 transmite también, dentro del flujo de datos, un nivel global de ruido que señala al lado de la decodificación el nivel de ruido hasta el cual se tienen que llenar con ruido las porciones cuantificadas a cero de la representación 12 antes del reescalado o descuantificación del espectro mediante la aplicación de los factores de escala 112. Esto está
- 20 ilustrado en la Fig. 10. La Fig. 10 ilustra, utilizando un cuadrículado, el espectro aun no reescalado de la señal de audio, tal como 18 en la Fig. 9. Este tiene porciones espectrales contiguas de cero 40a, 40b, 40c y 40d. El nivel global de ruido 114 que también se puede transmitir en el flujo de datos correspondiente a cada espectro 18, indica al decodificador el nivel hasta el cual estas porciones de cero 40a a 40d deben ser llenadas con ruido antes de someter a este espectro llenado a un reescalado o recuantificación empleando los factores de escala 112.
- 25 Como ya se indicara anteriormente, el llenado con ruido al que se refiere el nivel global de ruido 114, puede ser objeto de una restricción por el hecho de que este tipo de llenado con ruido se refiere simplemente a las frecuencias superiores a cierta frecuencia inicial que se indica en la Fig. 10 únicamente con fines ilustrativos como f_{start} .
- 30 La Fig. 10 también ilustra otra característica específica que puede ser implementada en el codificador 100: ya que puede haber espectros 18 que comprenden bandas de factores de escala 110 donde todos los valores espectrales dentro de las respectivas bandas de factores de escala han sido cuantificados a cero, el factor de escala 112 asociado a esa banda de factores de escala es en realidad superfluo. En consecuencia, el cuantificador 100 usa este mismo factor de escala para llenar individualmente la banda de factores de escala con ruido además del ruido introducido en
- 35 la banda de factores de escala utilizando el nivel global de ruido 114, o, en otros términos, para escalar el ruido atribuido a la respectiva banda de factores de escala en respuesta al nivel global de ruido 114. Véase, por ejemplo, la Fig. 10. La Fig. 10 ilustra una subdivisión ejemplificativa del espectro 18 en bandas de factores de escala 110a a 110h.
- 40 La banda de factores de escala 110e es una banda de factores de escala cuyos valores espectrales han sido cuantificados a cero en su totalidad. En consecuencia, el factor de escala asociado 112 es "libre" y se utiliza para determinar 114 el nivel de ruido hasta el cual se llena por completo esta banda de factores de escala. Las otras bandas de factores de escala que comprenden valores espectrales cuantificados a niveles no cero, tienen factores de escala asociados a las mismas que se utilizan para reescalar los valores espectrales del espectro 18 que no han sido cuantificados a cero, incluyendo el ruido utilizado para llenar las porciones de cero 40a a 40d, escalado que se indica
- 45 usando la flecha 116, en forma representativa.
- El codificador 100 de la Fig. 9 ya puede tomar en cuenta que dentro del lado de la decodificación se ha de ejecutar el llenado con ruido usando el nivel global de ruido 114 empleando las formas de realización de llenado con ruido antes descritas, por ej., usando una dependencia de la tonalidad y/o imponiendo una inclinación espectralmente global al
- 50 ruido y/o variando la frecuencia inicial de llenado con ruido y así sucesivamente.
- En lo que respecta a la dependencia de la tonalidad, el codificador 100 puede determinar el nivel global de ruido 114, e insertar el mismo en el flujo de datos, asociando a las porciones de cero 40a a 40d la función para el modelado espectral del ruido para llenar la porción de cero respectiva. En particular, el codificador puede utilizar estas funciones
- 55 para ponderar los valores espectrales de la señal de audio original, es decir ponderados pero aun no cuantificados en estas porciones 40a a 40d para determinar el nivel global de ruido 114. De esa manera, el nivel global de ruido 114 determinado y transmitido en el flujo de datos, lleva a un llenado con ruido del lado de la decodificación que recupera en forma más aproximada el espectro de la señal de audio original.
- 60 El codificador 100 puede decidir, dependiendo del contenido de la señal de audio, sobre el uso de algunas opciones de codificación que, a su vez, se pueden utilizar como indicios de tonalidad tales como el indicio de tonalidad 38 expuesto en la Fig. 2 a fin de permitir que el lado de la decodificación establezca correctamente la función para el modelado espectral del ruido utilizado para llenar las porciones 40a a 40d. Por ejemplo, el codificador 100 puede utilizar la predicción temporal para predecir un espectro 18 a partir de un espectro anterior utilizando un denominado

ES 2 796 485 T3

- parámetro de ganancia por predicción a largo plazo. En otras palabras, la ganancia por predicción a largo plazo puede fijar el grado en el cual se utiliza o no dicha predicción temporal. En consecuencia, la ganancia por predicción a largo plazo o ganancia LTP, es un parámetro que se puede emplear como indicio de tonalidad ya que cuanto más alta sea la ganancia de LTP, más alta será probablemente la tonalidad de la señal de audio. Por consiguiente, un determinador de la tonalidad 34 de la Fig. 2, puede fijar por ejemplo, la tonalidad de acuerdo con una dependencia positiva monótona de la ganancia por LTP. En lugar, o además de una ganancia por LTP, el flujo de datos puede comprender una bandera de habilitación de LTP para señalar la activación/desactivación de la LTP, revelando así también, por ejemplo, un indicio con valor binario referente a la tonalidad.
- Además o por otro lado, el codificador 100 puede admitir el modelado temporal de ruido. En otras palabras, según el espectro 18, por ejemplo, el codificador 100 puede optar por someter al espectro 18 a modelado temporal de ruido indicando esta decisión al decodificador mediante una bandera de habilitación de modelado temporal de ruido. La bandera de habilitación de TNS indica si los niveles espectrales del espectro 18 forman la predicción residual de una predicción lineal espectral, es decir a lo largo de la dirección de frecuencia determinada, del espectro o si el espectro no ha sido sometido a predicción LP. Si se señala que la TNS está habilitada, el flujo de datos comprende además los coeficientes de predicción lineal para la predicción espectralmente lineal del espectro a fin de que el decodificador pueda recuperar el espectro usando estos coeficientes de predicción lineal mediante la aplicación de los mismos al espectro antes o después del reescalado o descuantificación. La bandera de habilitación de TNS también es un indicio de tonalidad: Si la bandera de habilitación de TNS señala que se ha de activar la TNS, por ej., en un transitorio, luego es muy improbable que la señal de audio sea tonal, ya que el espectro parece ser muy predecible por predicción lineal a lo largo del eje de frecuencia y, por ende, no estacionario. En consecuencia, se puede determinar la tonalidad sobre la base de la bandera de habilitación de TNS de manera tal que la tonalidad es más alta si la bandera de habilitación de TNS deshabilita la TNS, y es más baja si la bandera de habilitación de TNS señala la habilitación de TNS. En lugar, o además de una bandera de habilitación de TNS, existe la posibilidad de derivar de los coeficientes de filtro TNS una ganancia que indica un grado en el cual se puede emplear la TNS para predecir el espectro, revelando también de esa manera un indicio de un valor superior a dos concerniente a la tonalidad.

Otros parámetros de codificación también pueden ser codificados en el flujo de datos por el codificador 100. Por ejemplo, una bandera de habilitación de reordenamiento espectral puede señalar una opción de codificación de acuerdo con la cual se codifica el espectro 18 reordenando los niveles espectrales, es decir los valores espectrales cuantificados, espectralmente, transmitiendo además, dentro del flujo de datos, la receta de reordenamiento de manera que el decodificador pueda reordenar, o retransponer, los niveles espectrales a fin de recuperar el espectro 18. Si la bandera de habilitación de reordenamiento del espectro está habilitada, es decir si se aplica el reordenamiento del espectro, esto indica que es probable que la señal de audio sea tonal ya que el reordenamiento tiende a ser más efectivo en velocidad/ distorsión al comprimir el flujo de datos si hay muchos picos tonales en el espectro. En consecuencia, además o por otro lado, se puede utilizar la bandera de habilitación de reordenamiento del espectro como indicio tonal y se puede fijar la tonalidad utilizada para el llenado con ruido de un tamaño mayor que la bandera de habilitación de reordenamiento del espectro que se está habilitando y más baja si la bandera de habilitación de reordenamiento del espectro está deshabilitada.

Para completar, y también con referencia a la Fig. 2b, cabe señalar que el número de funciones diferentes para el modelado espectral de las porciones de cero 40a a 40d, es decir el número de tonalidades diferentes discriminadas para fijar la función para el modelado espectral, puede ser, por ejemplo, mayor que cuatro, o incluso mayor que ocho, por lo menos en el caso de anchos de porciones espectrales contiguas de cero superiores a un ancho mínimo predeterminado.

En lo que respecta al concepto de imponer una inclinación espectralmente global al ruido y tomando esto en cuenta al computar el parámetro de nivel de ruido del lado de la codificación, el codificador 100 puede determinar el nivel global de ruido 114, e insertarlo en el flujo de datos, mediante porciones de ponderación de los valores espectrales de la señal de audio aún no cuantificados, aunque cuantificados con la inversa de la función de ponderación perceptual, espectralmente ubicados conjuntamente en las porciones de cero 40a a 40d, donde una función se extiende espectralmente por lo menos a través de toda la porción de llenado con ruido del ancho de banda del espectro y que presenta una pendiente de signo opuesto a la función 15 usada del lado de la decodificación para el llenado con ruido, por ejemplo y midiendo el nivel sobre la base de los valores no cuantificados así ponderados.

La Fig. 11 ilustra un decodificador que se ajusta al codificador de la Fig. 9. El decodificador de la Fig. 11 está indicado en general empleando el signo de referencia 130 y comprende un dispositivo de llenado con ruido 30, un descuantificador 132 y un transformador inverso 134. El dispositivo de llenado con ruido 30 recibe la secuencia de espectros 18 dentro del espectrograma 12, es decir la representación por líneas espectrales que incluye los valores espectrales cuantificados, y, opcionalmente, indicios de tonalidad del flujo de datos tal como uno o varios de los parámetros de codificación antes descritos. El dispositivo de llenado con ruido 30 llena a continuación las porciones espectrales contiguas de cero 40a a 40d con ruido de acuerdo con lo descrito anteriormente, utilizando por ejemplo la dependencia de la tonalidad antes descrita y/o imponiendo una inclinación espectralmente global al ruido y utilizando el nivel global de ruido 114 para escalar el nivel de ruido de acuerdo con lo descrito anteriormente. Llenados de esa

ES 2 796 485 T3

manera, estos espectros llegan al descuantificador 132, que a su vez descuantifica o reescala el espectro llenado con ruido usando los factores de escala 112. El transformador inverso 134, a su vez, somete al espectro descuantificado a una transformación inversa a fin de recuperar la señal de audio. Como se describiera anteriormente, la transformación inversa 134 también puede comprender un proceso de solapamiento y suma para obtener la cancelación del solapamiento en el dominio del tiempo en caso de que la transformación usada por el transformador 104 sea una transformada solapada con muestreo crítico tal como MDCT, en cuyo caso la transformación inversa aplicada por el transformador inverso 134 sería una IMDCT (MDCT inversa).

Como ya se describiera con respecto a las Figs. 9 y 10, el descuantificador 132 aplica los factores de escala al espectro pre-llenado. En otras palabras, los valores espectrales contenidos en las bandas de factores de escala no cuantificados a cero son escalados empleando el factor de escala independientemente de los valores espectrales que representan un valor espectral no cero o un ruido que ha sido objeto de modelado espectral por el dispositivo de llenado con ruido 30 como se describiera anteriormente. Las bandas espectrales completamente cuantificadas a cero tienen factores de escala asociados a las mismas que son completamente libres para controlar el llenado con ruido y el dispositivo de llenado con ruido 30 puede emplear este factor de escala para escalar individualmente el ruido con el cual se ha llenado la banda de factores de escala por medio del llenado con ruido con el dispositivo de llenado con ruido 30 de las porciones espectrales contiguas de cero, o bien el dispositivo de llenado con ruido 30 puede utilizar el factor de escala para llenar adicionalmente, es decir sumar, ruido adicional en lo concerniente a las bandas espectrales cuantificadas a cero.

Cabe señalar que el ruido que el dispositivo de llenado con ruido 30 modela espectralmente en forma dependiente de la tonalidad antes descrita y/o somete a una inclinación espectralmente global de la manera antes descrita, puede originarse en una fuente de ruido pseudoaleatoria, o puede originarse en un dispositivo de llenado con ruido 30 sobre la base del parcheo o copia espectral de otras áreas del mismo espectro o espectros relacionados tales como el espectro alineado en tiempo de otro canal, o un espectro temporalmente precedente. Puede ser factible incluso el parcheo del mismo espectro, como por ejemplo copiando de áreas de frecuencia más baja del espectro 18 (copiado espectral). Independientemente de la manera en que el dispositivo de llenado con ruido 30 deriva el ruido, el dispositivo de llenado modela espectralmente el ruido para insertarlo en las porciones espectrales contiguas de cero 40a a 40d en la forma dependiente de la tonalidad antes descrita y/o somete al mismo a inclinación espectralmente global de la manera antes descrita.

Solo para completar más, se ilustra en la Fig. 12 que se pueden variar la forma de realización del codificador 100 y su correspondiente decodificador 130 de las Figs. 9 y 11 de manera que la yuxtaposición entre los factores de escala por un lado y el nivel de ruido específico de los factores de escala se implemente de manera diferente. De acuerdo con el ejemplo de la Fig. 12, el codificador transmite en el flujo de datos información sobre una envolvente de ruido, espectralmente muestreado en una resolución más gruesa que la resolución por líneas espectrales del espectrograma 12, tal como, por ejemplo, en la misma resolución espectral que los factores de escala 112, además de los factores de escala 112. Esta información de envolvente de ruido se indica empleando el signo de referencia 140 en la Fig. 12. Merced a esta medida, en el caso de las bandas de factores de escala no completamente cuantificados a cero existen dos valores: un factor de escala para reescalar o descuantificar los valores espectrales no cero dentro de esa banda de factores de escala respectiva, como así también un nivel de ruido 140 correspondiente a la banda de factores de escala individual que escala el nivel de ruido de los valores espectrales cuantificados a cero dentro de esa banda de factores de escala. Este concepto se denomina en ocasiones IGF (Llenado inteligente de Brechas).

Aun en este caso, el dispositivo de llenado con ruido 30 puede aplicar el llenado dependiente de la tonalidad de las porciones espectrales contiguas de cero 40a a 40d a título de ejemplo, como se ilustra en la Fig. 12.

De acuerdo con los ejemplos de códec de audio antes reseñados con respecto a las Figs. 9 a 12, se ha ejecutado el modelado espectral del ruido de la cuantificación mediante la transmisión de una información referente al umbral de enmascaramiento perceptual usando una representación espectral en forma de factores de escala. Las Figs. 13 y 14 ilustran un par de una forma de realización de un codificador y un decodificador correspondiente en el cual se pueden utilizar asimismo el concepto de llenado con ruido descrito en relación con las Figs. 1 a 8, pero donde el ruido de la cuantificación es modelado espectralmente de acuerdo con una descripción LP (Predicción Lineal) del espectro de la señal de audio. En el presente documento, el espectro que se ha de llenar con ruido está en el dominio ponderado, es decir, es cuantificado utilizando un tamaño de paso espectralmente constante en el dominio ponderado o el dominio perceptualmente ponderado.

La Fig. 13 ilustra un codificador 150 que comprende un transformador 152, un cuantificador 154, un preenfanzador 156, un analizador LPC 158 y un conversor de LPC a líneas espectrales 160. El preenfanzador 156 es opcional. El preenfanzador 156 somete a la señal de audio entrante 12 a preénfasis, es decir un filtrado de paso alto con una función plana de transferencia de filtro de paso alto utilizando, por ejemplo, un filtro FIR o IIR. Se puede utilizar un filtro de paso alto de primer orden, por ejemplo, para el preenfanzador 156 como, por ejemplo, $H(z) = 1 - az^{-1}$ donde a determina, por ejemplo, la cantidad o potencia de preénfasis en línea con la cual, de acuerdo con una de las formas

ES 2 796 485 T3

de realización, se varía la inclinación espectralmente global a la cual se somete al ruido para introducir en el espectro. Una configuración posible de α podría ser 0,68. La preénfasis causada por el preenfanzador 156 es para desplazar la energía de los valores espectrales cuantificados transmitidos por el codificador 150, de las altas a bajas frecuencias, tomando en cuenta de esa manera las leyes psicoacústicas de acuerdo con las cuales la percepción humana es mayor en la región de baja frecuencia que en la región de alta frecuencia. Sea o no preenfanzada la señal de audio, el analizador LPC 158 ejecuta un análisis de LPC de la señal de audio entrante 12 a fin de predecir linealmente la señal de audio α , para ser más precisos, estima su envolvente espectral. El analizador LPC 158 determina en unidades de tiempo, por ejemplo, de subtramas que consisten en un número de muestras de audio de la señal de audio 12, coeficientes de predicción lineal y transmite los mismos como se ilustra en 162 al lado de la decodificación dentro del flujo de datos. El analizador LPC 158 determina, por ejemplo, los coeficientes de predicción lineal usando autocorrelación en ventanas de análisis y usando, por ejemplo, un algoritmo de Levinson–Durbin.

Los coeficientes de predicción lineal se pueden transmitir en el flujo de datos en una versión cuantificada y/o transformada tal como en forma de pares de líneas espectrales o similar. En todo caso, el analizador LPC 158 envía al conversor de LPC a líneas espectrales 160 los coeficientes de predicción lineal también disponibles del lado de la decodificación a través del flujo de datos, y el conversor 160 convierte los coeficientes de predicción lineal en una curva espectral utilizada por el cuantificador 154 para variar/establecer el tamaño del paso de cuantificación. En particular, el transformador 152 somete a la señal de audio entrante 12 a una transformación, por ejemplo de manera igual a la que utiliza el transformador 104. Así, el transformador 152 emite una secuencia de espectros y el cuantificador 154 puede dividir, por ejemplo, cada espectro por la curva espectral obtenida del conversor 160 donde luego se utiliza un tamaño del paso de cuantificación espectralmente constante para todo el espectro. El espectrograma de una secuencia de espectros emitida por el cuantificador 154 está indicado en 164 en la Fig. 13 y comprende además algunas porciones espectrales contiguas de cero que se pueden llenar del lado de la decodificación. El codificador 150 puede transmitir un parámetro de nivel global de ruido en el flujo de datos.

La Fig. 14 ilustra un decodificador que se adapta al codificador de la Fig. 13. El decodificador de la Fig. 14 está indicado en general empleando el signo de referencia 170 y comprende un dispositivo de llenado con ruido 30, un conversor de LPC a líneas espectrales 172, un descuantificador 174 y un transformador inverso 176. El dispositivo de llenado con ruido 30 recibe los espectros cuantificados 164, ejecuta el llenado con ruido en las porciones espectrales contiguas de ceros como se describiera anteriormente, y envía el espectrograma llenado de esa manera al descuantificador 174. El descuantificador 174 recibe del conversor de LPC a líneas espectrales 172 una curva espectral que ha de ser utilizada por el descuantificador 174 para remodelar el espectro llenado o, en otras palabras, para descuantificarlo. Este proceso en ocasiones se denomina FDNS (Modelado de Ruido en el Dominio de la Frecuencia). El conversor de LPC a líneas espectrales 172 deriva la curva espectral sobre la base de la información de LPC 162 en el flujo de datos. El espectro descuantificado, o espectro remodelado emitido por el descuantificador 174 es sometido a una transformación inversa por el transformador inverso 176 para recuperar la señal de audio. Una vez más, la secuencia de los espectros remodelados puede ser sometida por el transformador inverso 176 a una transformación inversa seguida por un proceso de solapamiento y suma a fin de ejecutar la cancelación del solapamiento en el dominio del tiempo entre retransformadas consecutivas en caso de que una transformación del transformador 152 sea una transformada solapada con muestreo crítico tal como MDCT.

Por medio de las líneas de guiones de las Figs. 13 y 14 se indica que la preénfasis aplicada por el preenfanzador 156 puede variar en tiempo, donde la variación es señalizada en el flujo de datos. En ese caso, el dispositivo de llenado con ruido 30 puede tomar en cuenta la preénfasis al ejecutar el llenado con ruido como se describiera anteriormente con respecto a la Fig. 8. En particular, la preénfasis causa una inclinación espectral en el espectro cuantificado emitido por el cuantificador 154 por el hecho de que los valores espectrales cuantificados, es decir los niveles espectrales, tienden a decrecer desde las frecuencias más bajas a las frecuencias más altas, es decir que exhiben una inclinación espectral. Esta inclinación espectral puede ser compensada, o mejor emulada o adaptada, por el dispositivo de llenado con ruido 30 de la manera antes descrita. Si se señala en el flujo de datos, se puede utilizar el grado de preénfasis para ejecutar la inclinación adaptativa del ruido insertado de manera dependiente del grado de preénfasis. En otras palabras, el grado de preénfasis señalizado en el flujo de datos puede ser utilizado por el decodificador para fijar el grado de inclinación espectral impuesto al ruido introducido en el espectro por el dispositivo de llenado con ruido 30.

Hasta ahora se han descrito varias formas de realización, y en adelante se presentan ejemplos específicos de implementación. Los detalles expuestos con respecto a estos ejemplos deben ser considerados individualmente transferibles a las formas de realización antes expuestas para especificar las mismas aun más. Sin embargo, antes de eso se debe notar que todas las formas de realización antes descritas se pueden utilizar asimismo en la codificación de audio, como así también de voz. En general se refieren a la codificación por transformadas y utilizan un concepto de adaptación a la señal para reemplazar los ceros introducidos en el proceso de cuantificación con ruido espectralmente modelado utilizando una cantidad muy pequeña de información complementaria. En las formas de realización antes descritas, se aprovechó la observación de que en ocasiones también aparecen agujeros espectrales inmediatamente por debajo de la frecuencia inicial de ruido en caso de utilizarse esa frecuencia inicial, y que esos agujeros espectrales a veces son perceptualmente molestos. Las formas de realización expuestas que utilizan una señalización explícita de la frecuencia inicial permiten la remoción de los agujeros que causan degradación, aunque

permiten evitar insertar ruido a bajas frecuencias donde la inserción de ruido pudiera introducir distorsiones.

Más aún, algunas de las formas de realización antes reseñadas utilizan un llenado con ruido controlado con preénfasis para compensar la inclinación espectral causada por la preénfasis. Estas formas de realización toman en cuenta la observación de que si se calcula el filtro LPC en una señal de preénfasis, meramente aplicando una magnitud global o promedio o una energía promedio del ruido que se ha de insertar causaría que el modelado del ruido introduzca una inclinación espectral en el ruido insertado, ya que la FDNS del lado de la decodificación sometería al ruido insertado espectralmente plano a modelado espectral que de todas maneras muestra la inclinación espectral de la preénfasis. En consecuencia, estas últimas formas de realización ejecutaron un llenado con ruido de tal manera que se tome en cuenta y se compense la preénfasis.

Por consiguiente, en otras palabras, cada una de las Figs. 11 y 14 exhibía un decodificador de audio por transformada perceptual. Este comprende un dispositivo de llenado con ruido 30 configurado para ejecutar el llenado con ruido en un espectro 18 de una señal de audio. La ejecución se puede realizar dependiendo de la tonalidad como se describiera anteriormente. La ejecución se puede realizar llenando el espectro con ruido que exhibe una inclinación espectralmente global a fin de obtener un espectro llenado con ruido, como se describiera anteriormente. "Inclinación espectralmente global" ha de significar, por ejemplo, que la inclinación se manifiesta, por ejemplo, en una envolvente que envuelve el ruido a través de todas las porciones 40 que se han de llenar con ruido, que está inclinado, es decir que tiene una pendiente que no es cero. "Envolvente" se define, por ejemplo, como curva de regresión espectral tal como una función lineal u otro polinomio de orden dos o tres, por ejemplo, llevando a través de la máxima local del ruido introducido en las porciones 40 todas las cuales son auto-continuas, aunque espectralmente distanciadas. "Decrece de frecuencias bajas a altas" significa que esta inclinación tiene pendiente negativa y "aumenta de bajas a altas frecuencias" significa que esta inclinación tiene pendiente positiva. Ambos aspectos de la ejecución se pueden aplicar simultáneamente o solo uno de ellos.

Asimismo, el decodificador de audio por transformada perceptual comprende un modelador de ruido en el dominio de la frecuencia 6 en forma de descuantificador 132, 174, configurado para someter al espectro llenado con ruido a modelado espectral utilizando una función de ponderación espectral perceptual. En el caso de la Fig. 11, el modelador de ruido en el dominio de la frecuencia 132 está configurado para determinar la función de ponderación espectral perceptual a partir de la información de coeficientes de predicción lineal 162 señalizada en el flujo de datos en el cual se codifica el espectro. En el caso de la Fig. 14, el modelador de ruido en el dominio de la frecuencia 174 está configurado para determinar la función de ponderación espectral perceptual a partir de factores de escala 112 relacionados con las bandas de factores de escala 110, señalizados en el flujo de datos. Como se describiera en relación con la Fig. 8 y se ilustrara con respecto a la Fig. 11, el dispositivo de llenado con ruido 34 puede estar configurado para variar la pendiente de la inclinación espectralmente global en respuesta a una señalización explícita en el flujo de datos, o para deducirla de una porción del flujo de datos que señala la función de ponderación espectral perceptual como por ejemplo evaluando la envolvente espectral de LPC o los factores de escala, o para deducirla del espectro cuantificado y transmitido 18.

Además, el decodificador de audio por transformada perceptual comprende un transformador inverso 134, 176 configurado para transformar inversamente el espectro llenado con ruido, espectralmente modelado por el modelador de ruido en el dominio de la frecuencia, para obtener una transformada inversa y someter a la transformada inversa a un proceso de solapamiento y suma.

De manera correspondiente, tanto la Fig. 13 como la 9 exhibieron ejemplos de un codificador de audio por transformada perceptual configurado para ejecutar la ponderación de un espectro 1 y su cuantificación 2, ambas implementadas en los módulos del cuantificador 108, 154 expuestos en la Fig. 9 y 13. La ponderación del espectro 1 pondera espectralmente el espectro original de una señal de audio de acuerdo con la inversa de una función de ponderación espectral perceptual a fin de obtener un espectro perceptualmente ponderado y la cuantificación 2 cuantifica el espectro perceptualmente ponderado en forma espectralmente uniforme a fin de obtener un espectro cuantificado. El codificador de audio por transformada perceptual ejecuta además un cómputo del nivel de ruido 3 dentro de los módulos de cuantificación 108, 154, por ejemplo, computando un parámetro de nivel de ruido midiendo un nivel del espectro perceptualmente ponderado ubicado conjuntamente en las porciones de cero del espectro cuantificado de manera ponderada con una inclinación espectral global que aumenta de bajas a altas frecuencias. De acuerdo con la Fig. 13, el codificador de audio por transformada perceptual comprende un analizador LPC 158 configurado para determinar información de coeficientes de predicción lineal 162 que representa una envolvente espectral LPC del espectro original de la señal de audio, donde el ponderador espectral 154 está configurado para determinar la función de ponderación espectral perceptual a fin de seguir la envolvente espectral de LPC. Como se señalara, el analizador de LPC 158 puede estar configurado para determinar la información de coeficientes de predicción lineal 162 mediante la ejecución del análisis de LP en una versión de la señal de audio, sometida al filtro de preénfasis 156. Como se describiera anteriormente con respecto a la Fig. 13, el filtro de preénfasis 156 puede estar configurado para filtrar por paso alto la señal de audio con una cantidad variable de preénfasis para obtener la versión de la señal de audio sometida a un filtro de preénfasis, donde el cómputo del nivel de ruido puede estar configurado para fijar una cantidad de inclinación espectralmente global dependiendo de la cantidad de preénfasis. Se puede

ES 2 796 485 T3

utilizar la señalización explícita de la cantidad de inclinación espectralmente global o la cantidad de preénfasis en el flujo de datos. En el caso de la Fig. 9, el codificador de audio por transformada perceptual comprende una determinación de factores de escala, controlada mediante un modelo perceptual 106, que determina los factores de escala 112 relacionados con bandas de factores de escala 110 a fin de seguir un umbral de enmascaramiento. Esta determinación es implementada en el módulo de cuantificación 108, por ejemplo, que también actúa como ponderador espectral configurado para determinar la función de ponderación espectral perceptual para seguir los factores de escala.

Se toma ahora la alternativa recién aplicada y generalizando los términos para describir las Fig. 9 a 14 para describir las Figs. 18a y 18b.

La Fig. 18a ilustra un codificador de audio por transformada perceptual de conformidad con una forma de realización de la presente solicitud y la Fig. 18b ilustra un decodificador de audio por transformada perceptual adaptado al mismo para formar un códec de audio por transformada perceptual.

Como se ilustra en la Fig. 18a, el codificador de audio por transformada perceptual comprende un ponderador del espectro 1 configurado para ponderar espectralmente el espectro original de una señal de audio recibida por el ponderador del espectro 1 de acuerdo con la inversa de una función de ponderación espectral perceptual determinada por el ponderador del espectro 1 de manera predeterminada de la cual se exponen ejemplos más adelante. El ponderador espectral 1 obtiene, merced a esta medida, un espectro perceptualmente ponderado al que luego se somete a cuantificación en forma espectralmente uniforme, es decir de manera igual para las líneas espectrales, en un cuantificador 2 del codificador de audio por transformada perceptual. El resultado emitido por el cuantificador uniforme 2 es un espectro cuantificado 34 que finalmente es codificado en un flujo de datos emitido por el codificador de audio por transformada perceptual.

Para controlar el llenado con ruido que se ha de ejecutar del lado de la decodificación a fin de mejorar el espectro 34, en lo referente a la fijación del nivel de ruido, opcionalmente puede haber una calculadora de nivel de ruido 3 del codificador de audio por transformada perceptual presente para computar un parámetro de nivel de ruido midiendo un nivel del espectro perceptualmente ponderado 4 en las porciones 5 co-ubicadas en las porciones de cero 40 del espectro cuantificado 34. El parámetro de nivel de ruido así calculado puede ser codificado asimismo en el flujo de datos antes mencionado para que llegue al decodificador.

El decodificador de audio por transformada perceptual está expuesto en la Fig. 18b. El mismo comprende un aparato de llenado con ruido 30 configurado para ejecutar el llenado con ruido en el espectro entrante 34 de la señal de audio, codificada en el flujo de datos generado por el codificador de la Fig. 1a, llenando el espectro 34 con ruido que exhibe una inclinación espectralmente global de manera que el nivel de ruido se reduzca de bajas a altas frecuencias a fin de obtener un espectro llenado con ruido 36. Un modelador de ruido en el dominio de la frecuencia del decodificador de audio por transformada perceptual, indicado con el signo de referencia 6, está configurado para someter al espectro llenado con ruido a modelado espectral utilizando la función de ponderación espectral perceptual obtenida del lado de la codificación a través del flujo de datos de la manera descrita mediante los siguientes ejemplos específicos adicionales. Este espectro emitido por el modelador de ruido en el dominio de la frecuencia 6 puede ser reenviado a un transformador inverso 7 para reconstruir la señal de audio en el dominio del tiempo y del mismo modo, dentro del codificador de audio por transformada perceptual, un transformador 8 puede preceder al ponderador del espectro 1 para suministrar al ponderador del espectro 1 el espectro de la señal de audio. La significancia de llenar el espectro 34 con ruido 9 que exhibe una inclinación espectralmente global es la siguiente: cuando se somete al espectro llenado con ruido 36 a modelado espectral por el modelador de ruido en el dominio de la frecuencia 6, se somete al espectro 36 a una función de ponderación inclinada. Por ejemplo, el espectro se amplifica a las altas frecuencias en comparación con una ponderación de las bajas frecuencias. En otras palabras, el nivel del espectro 36 se eleva a frecuencias más altas en relación con las frecuencias más bajas. Esto causa una inclinación espectralmente global con pendiente positiva en las porciones en principio espectralmente planas del espectro 36. En consecuencia, en caso de introducir ruido 9 en el espectro 36 para llenar la porción de cero 40 del mismo, de manera espectralmente plana, luego el espectro emitido por FDNS 6 exhibiría en estas porciones 40 un piso de ruido que tiende a incrementarse, por ejemplo, de bajas a altas frecuencias. En otras palabras, al examinar la totalidad del espectro o por lo menos la porción del ancho de banda del espectro en la cual se ejecuta el llenado con ruido, se vería que el ruido dentro de las porciones 40 tiene tendencia a una función de regresión lineal con pendiente positiva o pendiente negativa. Sin embargo, cuando el aparato de llenado con ruido 30 llena el espectro 34 con ruido que exhibe una inclinación espectralmente global de pendiente positiva o negativa, indicada con α en la Fig. 1b, y que está inclinada en la dirección opuesta en comparación con la inclinación causada por el FDNS 9, la inclinación espectral causada por el FDNS 6 se compensa y el piso de ruido así introducido en el espectro reconstruido en última instancia a la salida del FDNS 6 es plano o por lo menos más plano, aumentando así la calidad del audio dejando menos agujeros profundos de ruido.

"Inclinación espectralmente global" ha de indicar que el ruido 9 llenado en el espectro 34 tiene un nivel que tiene a decrecer (o aumentar) de bajas a altas frecuencias. Por ejemplo, al trazar una línea de regresión lineal a través de un máximo local de ruido 9 insertado, por ejemplo, en porciones de cero espectrales contiguas mutuamente

espectralmente distanciadas 40, la línea de regresión lineal obtenida tiene la pendiente negativa (o positiva) a.

Aunque no es indispensable, la calculadora de nivel de ruido del codificador de audio por transformada perceptual puede dar cuenta de la manera inclinada del llenado con ruido del espectro 34 midiendo el nivel del espectro perceptualmente ponderado 4 en las porciones 5 de manera ponderada con una inclinación espectral global que tiene, por ejemplo, pendiente positiva en caso de que α sea negativo y pendiente negativa si α es positivo. No es necesario que la pendiente aplicada por la calculadora de nivel de ruido, que está indicada con β en la Fig. 18a, sea igual a la aplicada del lado de la decodificación en lo que respecta al valor absoluto de la misma, aunque de conformidad con una forma de realización este podría ser el caso. Al hacerlo así, la calculadora de nivel de ruido 3 puede adaptar el nivel del ruido 9 insertado del lado de la decodificación con más precisión al nivel de ruido que se aproxima mejor a la señal original y a través de todo el ancho de banda espectral. Más adelante se explicará que puede ser factible controlar una variación de una pendiente de la inclinación espectralmente global α mediante señalización explícita en el flujo de datos o mediante señalización implícita por el hecho de que, por ejemplo, el aparato de llenado con ruido 30 deduce lo abrupto de la pendiente, por ejemplo, de la función de ponderación espectral perceptual en sí o de una conmutación de largo de ventana de transformada. Mediante esta última deducción, por ejemplo, se puede adaptar la pendiente a la longitud de la ventana.

Hay diferentes maneras posibles por medio de las cuales el aparato de llenado con ruido 30 hace que el ruido 9 exhiba la inclinación espectralmente global. La Fig. 18c, por ejemplo, ilustra que el aparato de llenado con ruido 30 ejecuta una multiplicación por líneas espectrales 11 entre una señal de ruido intermediaria 13, que representa un estado intermedio en el proceso de llenado con ruido, y una función que aumenta (o decrece) en forma monótona 15, es decir una función que decrece (o aumenta) espectralmente en forma monótona a través de todo el espectro o por lo menos la porción en la que se ejecuta el llenado con ruido, para obtener el ruido 9. Como se ilustra en la Fig. 18c, la señal de ruido intermediaria 13 puede estar ya espectralmente modelada. Los detalles en este aspecto pertenecen a las formas de realización específicas reseñadas con más detalle más adelante, de acuerdo con las cuales también se ejecuta el llenado con ruido dependiendo de la tonalidad. Sin embargo, también se puede omitir el modelado espectral o se puede ejecutar después de la multiplicación 11. Se puede utilizar la señal de parámetro de nivel de ruido y el flujo de datos para establecer el nivel de la señal de ruido intermediaria 13, aunque por otro lado se puede generar la señal de ruido intermediaria utilizando un nivel estándar, aplicando el parámetro escalar de nivel de ruido para escalar el espectro después de la multiplicación 11. La función en decrecimiento monótono 15 puede ser una función lineal, como se ilustra en la Fig. 18c, una función lineal por fragmentos, una función polinómica o cualquier otra función.

Como se describe más adelante en forma más detallada, sería factible fijar adaptativamente la porción de todo el espectro dentro del cual el llenado con ruido es ejecutado por el aparato de llenado con ruido 30.

En conexión con las formas de realización que se describen más detalladamente más adelante, de acuerdo con las cuales se llenan las porciones espectrales contiguas de cero del espectro 34, es decir los agujeros espectrales, de manera específica no plana y dependiente de la tonalidad, se debe explicar que también hay alternativas para la multiplicación 11 ilustrada en la Fig. 18c para provocar la inclinación espectralmente global descrita hasta ahora.

Todas las formas de realización antes descritas tienen en común que se evitan los agujeros espectrales y también se evita el ocultamiento de las líneas tonales cuantificadas no cero. De la manera antes descrita, se puede preservar la energía en las partes ruidosas de una señal y se evita la adición de ruido que enmascara los componentes tonales de la manera antes descrita.

En las implementaciones específicas que se describen a continuación, la parte de la información complementaria para la ejecución del llenado con ruido dependiente de la tonalidad no agrega nada a la información complementaria existente del códec en que se utiliza el llenado con ruido. También se puede utilizar toda la información del flujo de datos que se utiliza para la reconstrucción del espectro, independientemente del llenado con ruido, para el modelado del llenado con ruido.

De acuerdo con un ejemplo de implementación, el llenado con ruido en el dispositivo de llenado con ruido 30 se realiza de la siguiente manera. Todas las líneas espectrales por encima de un índice de inicio de llenado con ruido que se cuantifican a cero son reemplazadas por un valor no cero. Esto se hace, por ejemplo, de manera aleatoria o pseudoaleatoria con una función de densidad de probabilidades espectralmente constante o utilizando parcheo de otras ubicaciones (fuentes) espectrales del espectrograma. Véase, por ejemplo, la Fig. 15. La Fig. 15 ilustra dos ejemplos correspondientes a un espectro al que se ha de someter a llenado con ruido tal como el espectro 34 o los espectros 18 del espectrograma 12 emitido por el cuantificador 108 o los espectros 164 emitidos por el cuantificador 154. El índice de partida de llenado con ruido es un índice de líneas espectrales $iFreq_0$ y $iFreq_1$ ($0 < iFreq_0 \leq iFreq_1$), donde $iFreq_0$ y $iFreq_1$ son índices de líneas espectrales predeterminados dependientes de la tasa de bits y el ancho de banda. El índice de partida de llenado con ruido es igual al índice $iStart$ ($iFreq_0 \leq iStart \leq iFreq_1$) de una línea espectral cuantificada a un valor no cero, donde todas las líneas espectrales con los índices j ($iStart < j \leq iFreq_1$) se cuantifican a cero. También se podrían transmitir diferentes valores para $iStart$, $iFreq_0$ o $iFreq_1$ en el flujo de bits para permitir la inserción de ruido de muy baja frecuencia en ciertas señales (por ej., ruido ambiental).

ES 2 796 485 T3

El ruido insertado se modela en los siguientes pasos:

- 5 1. En el dominio residual o el dominio ponderado. El modelado en el dominio residual o en el dominio ponderado ha sido extensamente descrito con respecto a las Figs. 1–14.
 - 10 2. El modelado espectral utilizando un LPC o el FDNS (modelado en el dominio de transformada utilizando la respuesta a la magnitud de LPC) ha sido descrito con respecto a las Figs. 13 y 14. El espectro también puede ser modelado empleando factores de escala (como en AAC) o utilizando cualquier otro método de modelado espectral para modelar el espectro completo, como se describe con respecto a las Figs. 9–12.
 - 15 3. El modelado opcional empleando TNS (Modelado temporal de ruido) usando un menor número de bits ha sido descrito brevemente con respecto a las Figs. 9–12
- La única información complementaria adicional necesaria para el llenado con ruido es el nivel, que se transmite utilizando 3 bits, por ejemplo.

20 Cuando se utiliza FDNS no hay necesidad de adaptarlo a un llenado con ruido específico y modela el ruido en la totalidad del espectro utilizando un menor número de bits que los factores de escala.

Se puede introducir una inclinación espectral en el ruido insertado para contrarrestar la inclinación espectral producida por la preénfasis en el modelado perceptual basado de LPC. Dado que la preénfasis representa un filtro suave de paso alto aplicado a la señal de entrada, la compensación de la inclinación puede contrarrestar esto multiplicando el equivalente de la función de transferencia de un filtro de paso bajo sutil en el espectro del ruido insertado. La inclinación espectral de esta operación de paso bajo depende del factor de preénfasis y, preferentemente, de la tasa de bits y el ancho de banda. Esto fue mencionado con referencia a la Fig. 8.

30 Por cada agujero espectral, constituido por 1 o más líneas espectrales consecutivas cuantificadas a cero, se puede modelar el ruido insertado como se ilustra en la Fig. 16. El nivel de llenado con ruido se puede encontrar en el codificador y transmitir en el flujo de bits. No hay llenado con ruido en las líneas espectrales cuantificadas no cero y aumenta en la zona de transición hasta el llenado con ruido completo. En la zona de llenado con ruido completo el nivel de llenado con ruido es igual al nivel transmitido en el flujo de bits, por ejemplo. Esto evita insertar un alto nivel de ruido en la proximidad inmediata de las líneas espectrales cuantificadas no cero que podrían potencialmente enmascarar o distorsionar los componentes tonales. Sin embargo, todas las líneas cuantificadas a cero son reemplazadas por ruido, sin dejar ningún agujero espectral.

40 El ancho de la transición depende de la tonalidad de la señal de entrada. La tonalidad se obtiene por cada trama de tiempo. En las Figs. 17a–d se ilustra a manera de ejemplo el modelado de ruido para diferentes tamaños de agujeros y anchos de transición.

La medida de tonalidad del espectro se puede basar en la información disponible en el flujo de bits:

- Ganancia de LTP
- 45 • Bandera de habilitación de reordenamiento del espectro (ver [6])
- Bandera de habilitación de TNS

50 El ancho de la transición es proporcional a la tonalidad – pequeña para las señales de tipo ruido, grande en el caso de las señales muy tonales.

En una forma de realización, el ancho de la transición es proporcional a la ganancia de LTP si la ganancia de LTP > 0. Si la ganancia de LTP es igual a 0 y el reordenamiento del espectro está habilitado, en ese caso se utiliza el ancho de la transición correspondiente a la ganancia promedio de LTP. Si el TNS está habilitado, en ese caso no hay área de transición, sino que se debe aplicar el llenado con ruido completo a todas las líneas espectrales cuantificadas a cero. Si la ganancia de LTP es igual a 0 y el TNS y el reordenamiento del espectro están deshabilitados, se utiliza un ancho de transición mínimo.

60 Si no hay información de tonalidad en el flujo de bits se puede calcular una medida de la tonalidad en la señal decodificada sin el llenado con ruido. Si no hay información de TNS, se puede calcular una medida temporal del nivel plano de la señal decodificada. No obstante, si hay información de TNS disponible, dicha medida del nivel plano se puede derivar directamente de los coeficientes de filtro de TNS, por ejemplo computando la ganancia de predicción del filtro.

ES 2 796 485 T3

En el codificador, se puede calcular el nivel de llenado con ruido preferentemente tomando en cuenta el ancho de la transición. Hay varias maneras posibles de determinar el nivel de llenado con ruido a partir del espectro cuantificado. La más sencilla es sumar la energía (cuadrado) de todas líneas del espectro normalizado de entrada en la región de llenado con ruido (es decir por encima de i_{Start}) que fuera cuantificada a cero, luego dividir esta suma por el número de esas líneas para obtener la energía promedio por línea y por último calcular un nivel de ruido cuantificado de la raíz cuadrada de la energía de línea promedio. De esta manera se deriva con eficacia el nivel de ruido del RMS de los espectros componentes cuantificados a cero. Digamos que, por ejemplo, A es la serie de índices i de líneas espectrales en que el espectro ha sido cuantificado a cero y que pertenece a cualquiera de las porciones de cero, por ej., está por encima de la frecuencia inicial y digamos que N indica el factor de escala de ruido total. Los valores del espectro que aun no han sido cuantificados se indican con y_i . Además, $restante(i)$ ha de ser una función que indica, por cada valor espectral cuantificado a cero en el índice i , el índice del valor cuantificado a cero en el extremo de baja frecuencia de la porción de cero a la cual i pertenece, y $F_i(j)$ donde $j = 0$ a $J_i - 1$ ha de indicar la función asignada, dependiendo de la tonalidad, a la porción de cero que se inicia en el índice i , donde J_i indica el ancho de esa porción de cero. Luego se puede determinar N según $N = \sqrt{\sum_i A_i^2 / cardinality(A)}$.

En la forma de realización preferida, se consideran los tamaños de los agujeros individuales así como el ancho de la transición. Para este fin, se agrupan sucesiones de líneas consecutivas cuantificadas a cero en las regiones de agujeros. A continuación se escala cada línea espectral de entrada normalizada en una región de agujero, es decir cada valor espectral de la señal original en una posición espectral dentro de cualquier porción espectral contigua de cero, mediante la función de transición, como se describiera en la sección anterior y seguidamente se calcula la suma de las energías de las líneas escaladas. Como en la forma de realización sencilla anterior, seguidamente se puede calcular el nivel de llenado con ruido del RMS de las líneas cuantificadas a cero. Aplicando la nomenclatura antes expuesta, N se puede computar según $N = \sqrt{\sum_i A(F_{restante(i)}(i - restante(i)) \cdot y_i)^2 / cardinality(A)}$.

Un problema con esta técnica es, sin embargo, que se subestima la energía espectral en las regiones de agujero pequeñas (es decir las regiones con un ancho muy inferior al doble del ancho de la transición), puesto que en el cálculo de RMS, el número de líneas espectrales de la suma por el cual se divide la suma de energía permanece inalterado. En otras palabras, cuando el espectro cuantificado exhibe principalmente regiones de agujero muy pequeñas, el nivel de llenado con ruido obtenido es más bajo cuando el espectro es ralo y solo tiene unas pocas regiones largas de agujero. Para garantizar que, en ambos casos, se encuentre un nivel similar de ruido, es ventajoso, por lo tanto, adaptar el recuento de líneas empleado en el denominador del cómputo de RMS al ancho de la transición. Es muy importante que, si el tamaño de una región de agujero es mejor que el doble del ancho de la transición, el número de líneas espectrales en esa región de agujero no se cuente como está, es decir como número entero de líneas, sino como número de línea fraccionario que es menor que el número entero de líneas. En la anterior fórmula concerniente a N , por ejemplo, se reemplazaría "cardinality(A)" por un número menor dependiendo del número de porciones "pequeñas" de cero.

Por añadidura, la compensación de la inclinación espectral en el llenado con ruido debido a la codificación perceptual basada en LPC también debe ser tomada en cuenta durante el cálculo del nivel de ruido. Más específicamente, se aplica preferentemente la inversa de la compensación de inclinación del llenado con ruido del lado del decodificador a las líneas espectrales originales no cuantificadas que fueran cuantificadas a cero antes del cómputo del nivel de ruido. En el contexto de la codificación basada en LPC que emplea preénfasis, esto implica que las líneas de más alta frecuencia se amplifican ligeramente con respecto a las líneas de frecuencia más baja antes de la estimación del nivel de ruido. Aplicando la nomenclatura antes citada, N se puede computar según $N = \sqrt{\sum_i A(F_{restante(i)}(i - restante(i)) \cdot LPF(i)^{-1} \cdot y_i)^2 / cardinality(A)}$. Como se mencionara anteriormente, dependiendo de las circunstancias, la función LPF que corresponde a la función 15 puede tener una pendiente positiva y LPF cambiada para leerse como HPF, en consecuencia. Se señala brevemente que en todas las fórmulas antes expuestas, el uso de "LPF" fijando $F_{restante}$ en una función constante para que sea todo uno, revelaría la manera en que se puede aplicar el concepto de someter al ruido que se ha de introducir en el espectro 34 a una inclinación espectral global sin el llenado de agujeros dependiente de la tonalidad.

Los cómputos posibles de N se pueden realizar en el codificador como, por ejemplo, en 108 o 154.

Por último, se encontró que cuando se cuantificaban a cero los armónicos de una señal estacionaria muy tonal, las líneas que representaban esos armónicos llevaban a un nivel de ruido relativamente alto o inestable (es decir, fluctuante en el tiempo). Esta falla se puede reducir utilizando en el cálculo del nivel de ruido la magnitud promedio de las líneas cuantificadas a cero en lugar de su RMS. Si bien esta estrategia no siempre garantiza que la energía de las líneas llenadas con ruido en el decodificador reproduzca la energía de las líneas originales en las regiones de llenado con ruido, sí garantiza que los picos espectrales en las regiones de llenado con ruido hagan solo una contribución limitada al nivel de ruido total, reduciendo así el riesgo de sobreestimación del nivel de ruido.

Por último, se señala que un codificador puede estar configurado incluso para ejecutar el llenado con ruido por completo a fin de mantenerse en línea con el decodificador, por ejemplo, para fines de análisis y síntesis.

ES 2 796 485 T3

Por consiguiente, la forma de realización precedente describe, entre otras cosas, un método adaptativo a la señal para reemplazar los ceros introducidos en el proceso de cuantificación con ruido modelado espectralmente. Se describe una extensión de llenado con ruido correspondiente a un codificador y un decodificador que satisface los requisitos antes mencionados implementando lo siguiente:

- 5 • El índice de inicio de llenado con ruido puede adaptarse al resultado de la cuantificación del espectro pero sin limitarse a cierto rango
- Se puede introducir una inclinación espectral en el ruido insertado para contrarrestar la inclinación espectral producida por el modelado perceptual de ruido
- 10 • Todas las líneas cuantificadas a cero por encima del índice de inicio de llenado con ruido son reemplazadas por ruido
- Mediante una función de transición, el ruido insertado se atenúa cerca de las líneas espectrales no cuantificadas a cero
- 15 • La función de transición depende de las características instantáneas de la señal de entrada
- La adaptación del índice de inicio de llenado con ruido, la inclinación espectral y la función de transición se pueden basar en la información disponible en el decodificador
- 20

No hay necesidad de información complementaria adicional, excepto por un nivel de llenado con ruido

Si bien se han descrito algunos aspectos en el contexto de un aparato, es obvio que estos aspectos también representan una descripción del método correspondiente, en el cual un bloque o dispositivo corresponde a un paso del método o a una característica de un paso del método. De manera análoga, los aspectos descritos en el contexto de un paso del método también representan una descripción de un bloque o ítem correspondiente o de una característica de un aparato correspondiente. Algunos o todos los pasos del método pueden ser ejecutados por medio de (o utilizando) un aparato de hardware, como por ejemplo, un microprocesador, una computadora programable o un circuito electrónico. En algunas realizaciones, uno cualquiera o más de los pasos más importantes del método pueden ser ejecutados por ese tipo de aparato.

Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden ser implementadas en hardware o en software. La implementación se puede realizar empleando un medio de almacenamiento digital, por ejemplo un disco blando, un DVD, un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene almacenadas en la misma señales de control legibles electrónicamente, que cooperan (o tienen capacidad para cooperar) con un sistema de computación programable de tal manera que se ejecute el método respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por una computadora.

Algunas realizaciones de acuerdo con la invención comprenden un transportador de datos que comprende señales de control legibles electrónicamente, con capacidad para cooperar con un sistema de computación programable de tal manera que se ejecute uno de los métodos descritos en la presente.

En general, las realizaciones de la presente invención pueden ser implementadas en forma de producto de programa de computación con un código de programa, donde el código de programa cumple la función de ejecutar uno de los métodos al ejecutarse el programa de computación en una computadora. El código de programa puede ser almacenado, por ejemplo, en un portador legible por una máquina.

Otras formas de realización comprenden el programa de computación para ejecutar uno de los métodos aquí descritos, almacenado en un portador legible por una máquina.

En otras palabras, una realización del método de la invención consiste, por lo tanto, en un programa de computación que consta de un código de programa para realizar uno de los métodos aquí descritos al ejecutarse el programa de computación en una computadora.

Otra de las realizaciones de los métodos de la invención consiste, por lo tanto, en un portador de datos (o medio de almacenamiento digital, o medio legible por computadora) que comprende, grabado en el mismo, el programa de computación para ejecutar uno de los métodos aquí descritos. El portador de datos, el medio de almacenamiento digital o el medio grabado son por lo general tangibles y no transitorios.

Otra forma de realización del método de la invención es, por lo tanto, un flujo de datos o una secuencia de señales que representa el programa de computación para ejecutar uno de los métodos aquí descritos. El flujo de datos o la secuencia de señales pueden estar configurados, por ejemplo, para ser transferida a través de una conexión de comunicación de datos, por ejemplo por la Internet.

ES 2 796 485 T3

Otra forma de realización comprende un medio de procesamiento, por ejemplo una computadora, un dispositivo lógico programable, configurado o adaptado para ejecutar uno de los métodos aquí descritos.

- 5 Otra forma de realización comprende una computadora en la que se ha instalado el programa de computación para ejecutar uno de los métodos aquí descritos.

10 Otra forma de realización de acuerdo con la invención comprende un aparato o sistema configurado para transferir (por ejemplo por vía electrónica u óptica) un programa de computación para poner en práctica uno de los métodos aquí descritos en un receptor. El receptor puede ser, por ejemplo, una computadora, un dispositivo móvil, un dispositivo de memoria y similar. El aparato o sistema puede comprender, por ejemplo, un servidor de archivos para transferir el programa de computación al receptor.

15 En algunas formas de realización, se puede utilizar un dispositivo lógico programable (por ejemplo una matriz de puertas programables en el campo) para ejecutar algunas o todas las funcionalidades de los métodos aquí descritos. En algunas formas de realización, una matriz de puertas programables en el campo puede cooperar con un microprocesador para ejecutar uno de los métodos aquí descritos. Por lo general, los métodos son ejecutados preferentemente por cualquier aparato de hardware.

20 El aparato descrito en la presente puede ser implementado empleando un aparato de hardware o empleando una computadora, o utilizando una combinación de un aparato de hardware y una computadora.

25 Los métodos aquí descritos se pueden poner en práctica empleando un aparato de hardware o empleando una computadora, o utilizando una combinación de un aparato de hardware y una computadora.

Las realizaciones precedentemente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y detalles aquí descritos han de ser evidentes para las personas con capacitación en la técnica.

30 Referencias

[1] B. G. G. F. S. G. M. M. H. P. J. H. S. W. G. S. J. H. Nikolaus Rettelbach, "Noise Filler, Noise Filling Parameter Calculator Encoded Audio Signal Representation, Methods and Computer Program". Patente US 2011/0173012 A1.

35 [2] *Extended Adaptive Multi-Rate-Wideband (AMR-WB+) codec*, 3GPP TS 26.290 V6.3.0, 2005-2006.

[3] B. G. G. F. S. G. M. M. H. P. J. H. S. W. G. S. J. H. Nikolaus Rettelbach, "Audio encoder, audio decoder, methods for encoding and decoding an audio signal, audio stream and computer program". Patente WO 2010/003556 A1.

40 [4] M. M. N. R. G. F. J. R. J. L. S. W. S. B. S. D. C. H. R. L. P. G. B. B. J. L. K. K. H. Max Neuendorf, "MPEG Unified Speech and Audio Coding – The ISO/MPEG Standard for High-Efficiency Audio Coding of all Content Types," in 132a Convención de AES, Budapest, 2012. También aparece en el Journal of the AES, vol. 61, 2013.

45 [5] M. M. M. N. a. R. G. Guillaume Fuchs, "MDCT-Based Coder for Highly Adaptive Speech and Audio Coding," in 17a Conferencia Europea sobre Procesamiento de Señales (EUSIPCO 2009), Glasgow, 2009.

[6] H. Y. K. Y. M. T. Harada Noboru, "Coding Mmethod, Decoding Method, Coding Device, Decoding Device, Program, and Recording Medium". Patente WO 2012/046685 A1.

50

REIVINDICACIONES

1. Codificador de audio que admite llenado con ruido, configurado para cuantificar y codificar un espectro de una señal de audio en un flujo de datos y
- 5 fijar y codificar en el flujo de datos un nivel de llenado con ruido espectralmente global para ejecutar llenado con ruido en el espectro de la señal de audio, de manera dependiente de una tonalidad de la señal de audio,
- 10 caracterizado porque el codificador está configurado para, al fijar y codificar el nivel de llenado con ruido espectralmente global, medir un nivel de la señal de audio dentro de porciones espectrales contiguas de cero (40) del espectro (34), espectralmente modeladas dependientes de la tonalidad de la señal de audio, utilizando
- 15 una función (48, 50) que presume un máximo en un punto interno (52) de la porción espectral contigua de cero (40) y que presenta bordes en bajada hacia el exterior (58, 60) cuya pendiente absoluta depende negativamente de la tonalidad, o
- 20 una función (48, 50) que presume un máximo en un punto interno (52) de la porción espectral contigua de cero (40), y que presenta bordes en bajada hacia el exterior (58, 60) cuyo ancho espectral (54, 56) depende positivamente de la tonalidad, o
- 25 una función constante o unimodal (48, 50) donde una integral de la cual – normalizada a una integral de 1 – sobre los cuartos externos (a, d) de la porción espectral contigua de cero (40) depende negativamente de la tonalidad o
- 30 una función establecida (80) dependiendo de un ancho de la respectiva porción espectral contigua de cero de tal manera que la función quede confinada a la respectiva porción espectral contigua de cero, y dependiendo de la tonalidad de la señal de audio de manera tal que, si la tonalidad de la señal de audio se incrementa, la función se torna más compacta en la parte interna de la respectiva porción espectral contigua de cero y se aleja de los bordes externos de la respectiva porción espectral contigua de cero.
2. Codificador de audio de acuerdo con la reivindicación 1, en el que la medida es una media cuadrática.
3. Codificador de audio de acuerdo con la reivindicación 1 o 2, en el que el codificador está configurado para cuantificar el espectro (34) utilizando un tamaño del paso de cuantificación espectralmente variante y adaptativo a la señal de acuerdo con una envolvente espectral de predicción lineal, señalar la envolvente espectral de predicción lineal a través de coeficientes de predicción lineal (162) en el flujo de datos y codificar el espectro (34) en el flujo de datos.
4. Codificador de audio de acuerdo con la reivindicación 1 o 2, en el que el codificador está configurado para cuantificar el espectro (34) utilizando un tamaño del paso de cuantificación espectralmente variante y adaptativo a la señal de acuerdo con factores de escala (112) relacionados con bandas de factores de escala (110), señalar los factores de escala en el flujo de datos y codificar el espectro (34) en el flujo de datos.
5. Codificador de audio de acuerdo con cualquiera de las reivindicaciones 1 a 4, en el que el aparato está configurado para derivar la tonalidad de un parámetro de codificación utilizado para codificar el espectro de la señal de audio.
6. Método para codificar audio que admite llenado con ruido, comprendiendo el método cuantificar y codificar un espectro de una señal de audio en un flujo de datos y fijar y codificar en el flujo de datos un nivel de llenado con ruido espectralmente global para ejecutar llenado con ruido en el espectro de la señal de audio, de manera dependiente de una tonalidad de la señal de audio,
- 55 caracterizado porque fijar y codificar el nivel de llenado con ruido espectralmente global comprende medir un nivel de la señal de audio dentro de porciones espectrales contiguas de cero (40) del espectro (34), espectralmente modeladas dependientes de la tonalidad de la señal de audio, utilizando
- 60 una función (48, 50) que presume un máximo en un punto interno (52) de la porción espectral contigua de cero (40), y que presenta bordes en bajada hacia el exterior (58, 60) cuya pendiente absoluta depende negativamente de la tonalidad, o
- una función (48, 50) que presume un máximo en un punto interno (52) de la porción espectral contigua de cero (40), y que presenta bordes en bajada hacia el exterior (58, 60) cuyo ancho espectral (54, 56) depende positivamente de la tonalidad, o

ES 2 796 485 T3

- 5 una función constante o unimodal (48, 50) donde una integral de la cual – normalizada a una integral de 1 – sobre los cuartos externos (a, d) de la porción espectral contigua de cero (40) depende negativamente de la tonalidad, o
- 10 una función establecida (80) dependiendo de un ancho de la respectiva porción espectral contigua de cero de tal manera que la función quede confinada a la respectiva porción espectral contigua de cero, y dependiendo de la tonalidad de la señal de audio de manera tal que, si la tonalidad de la señal de audio se incrementa, la función se torna más compacta en la parte interna de la respectiva porción espectral contigua de cero y se aleja de los bordes externos de la respectiva porción espectral contigua de cero.
- 15 7. Programa de computación con un código de programa para realizar, cuando se ejecuta en una computadora, un método de acuerdo con la reivindicación 6.

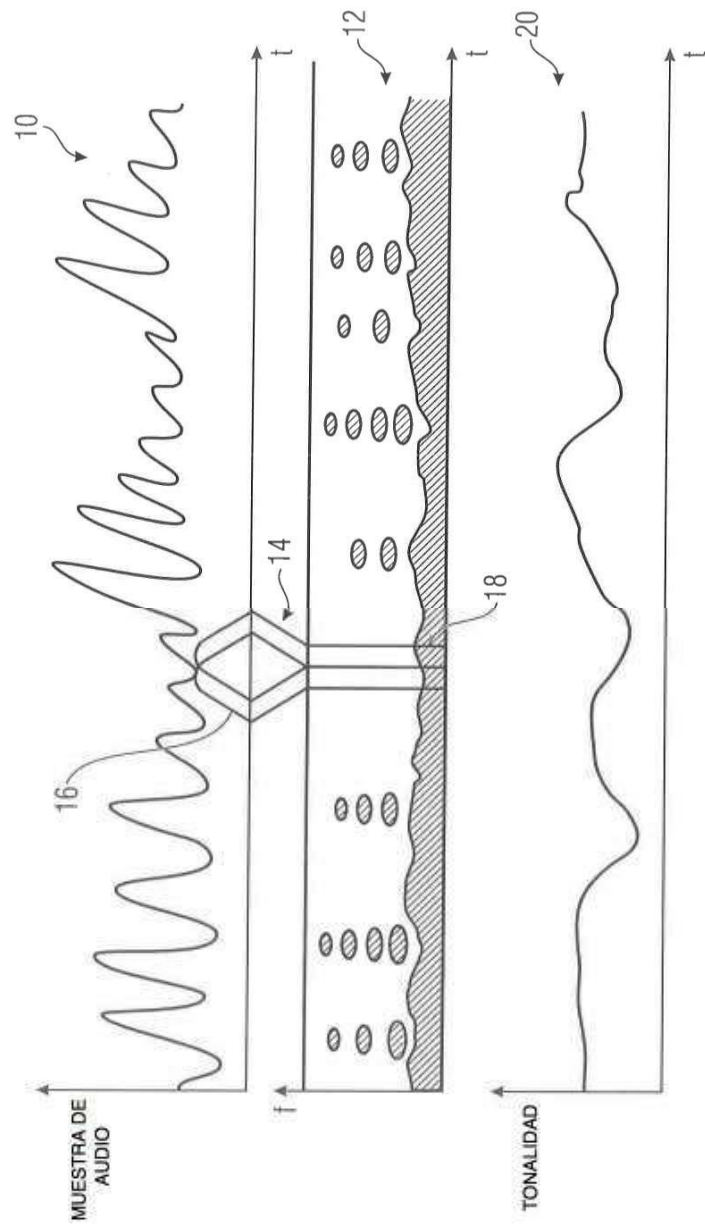


FIG 1

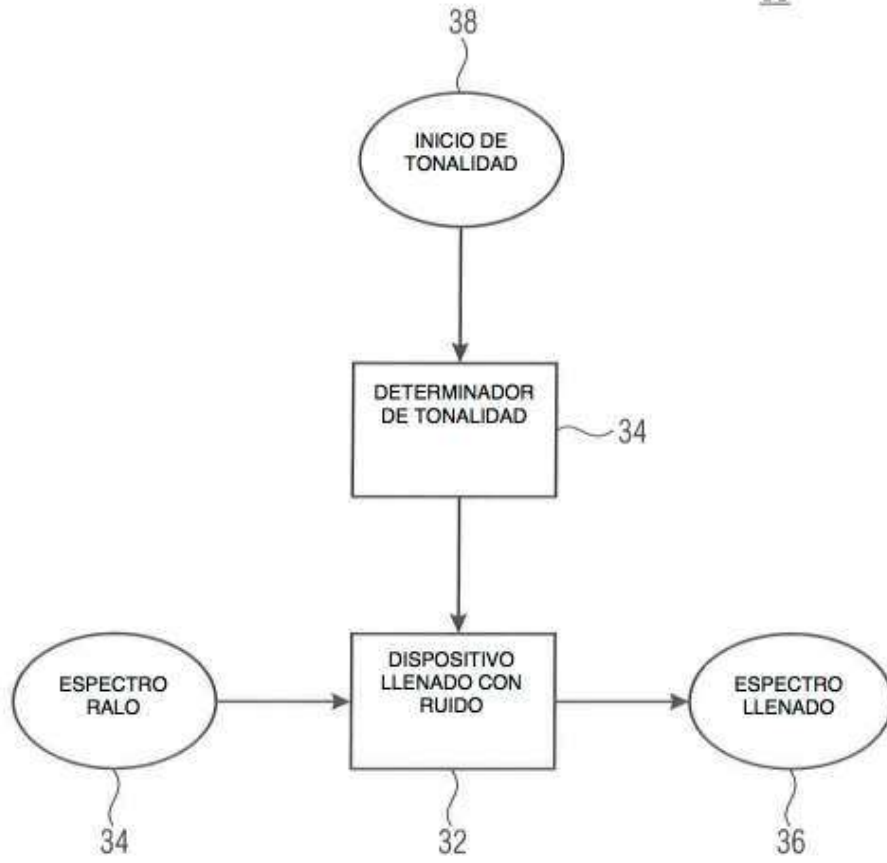


FIG 2

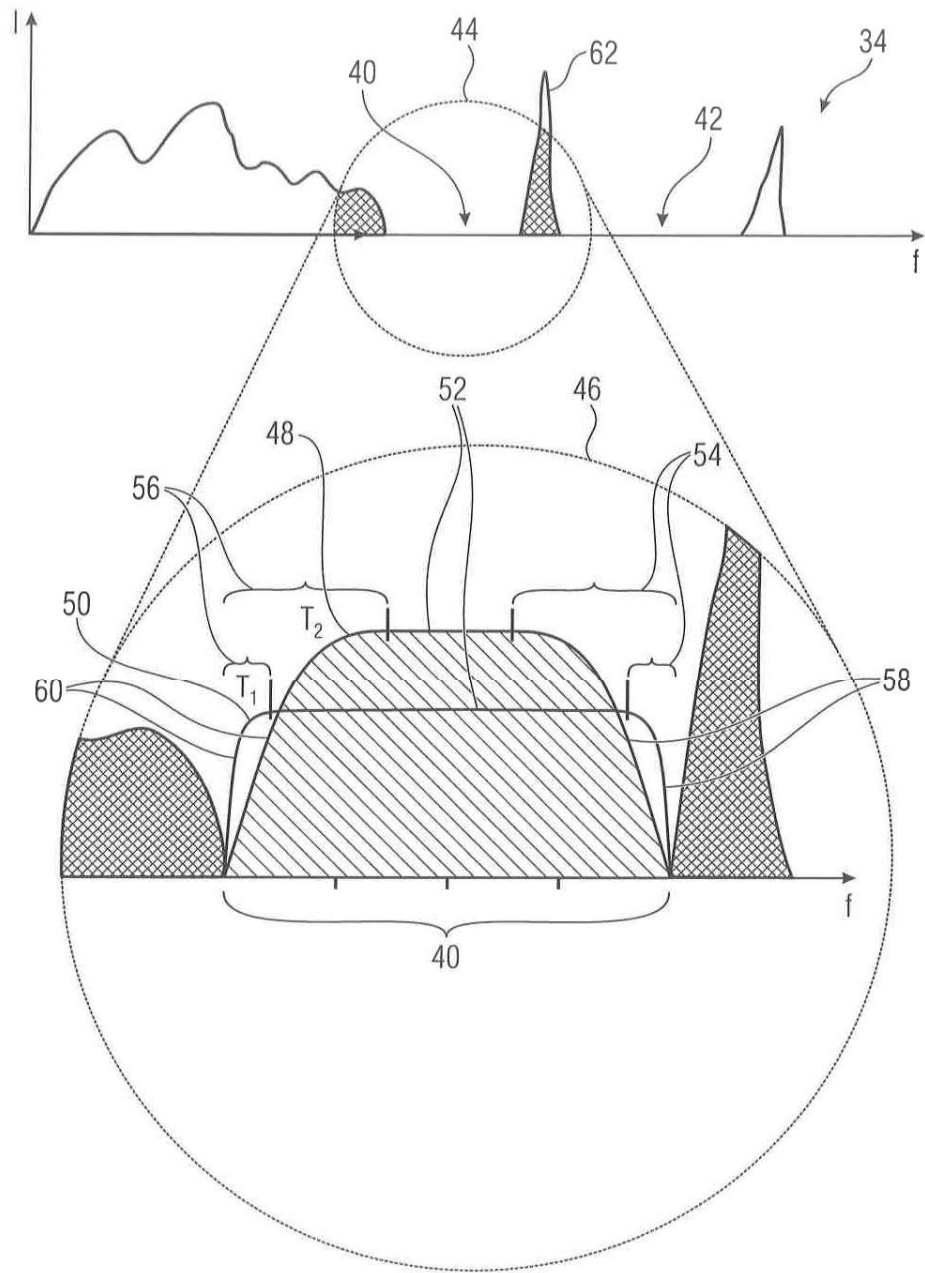


FIG 3

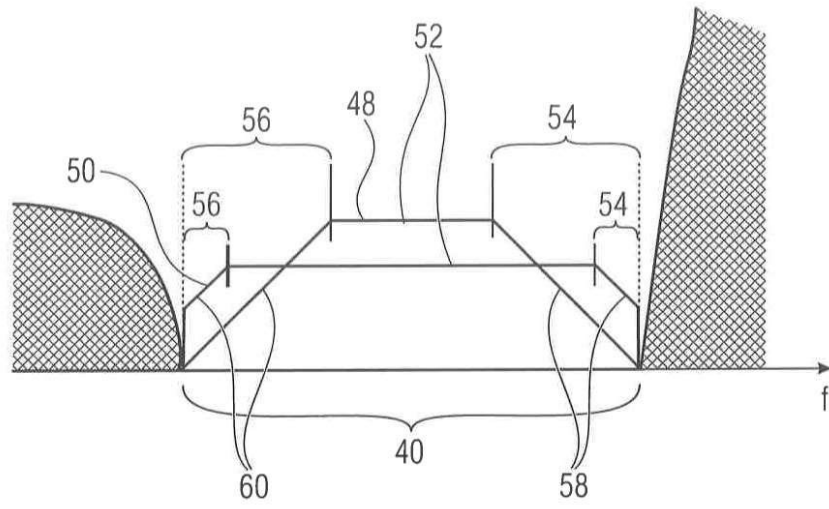


FIG 4

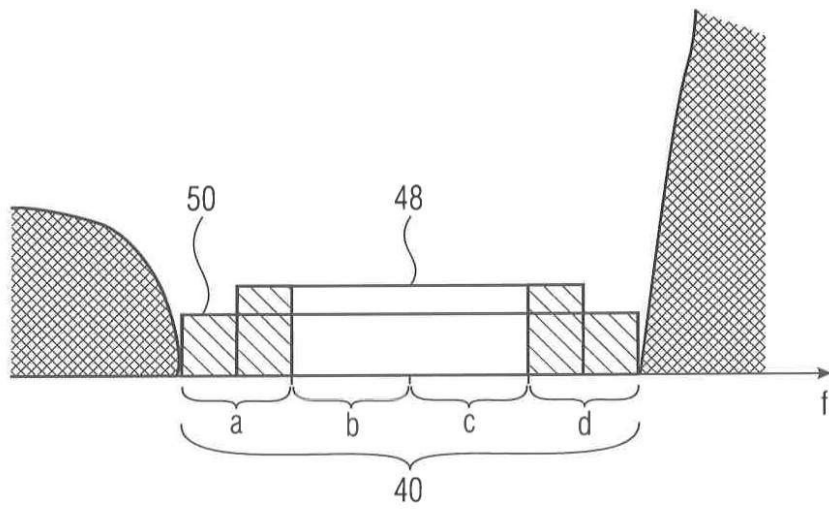


FIG 5

32



FIG 6

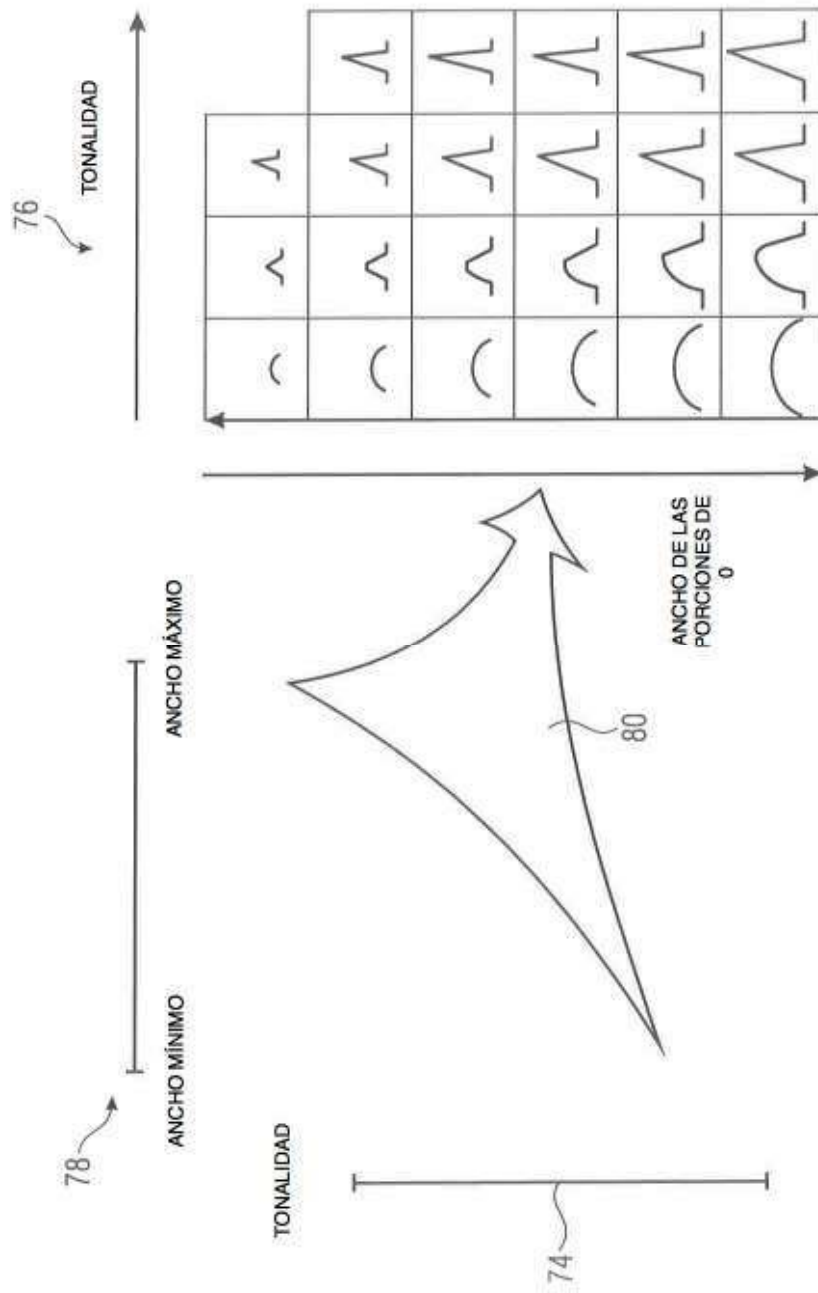


FIG 7

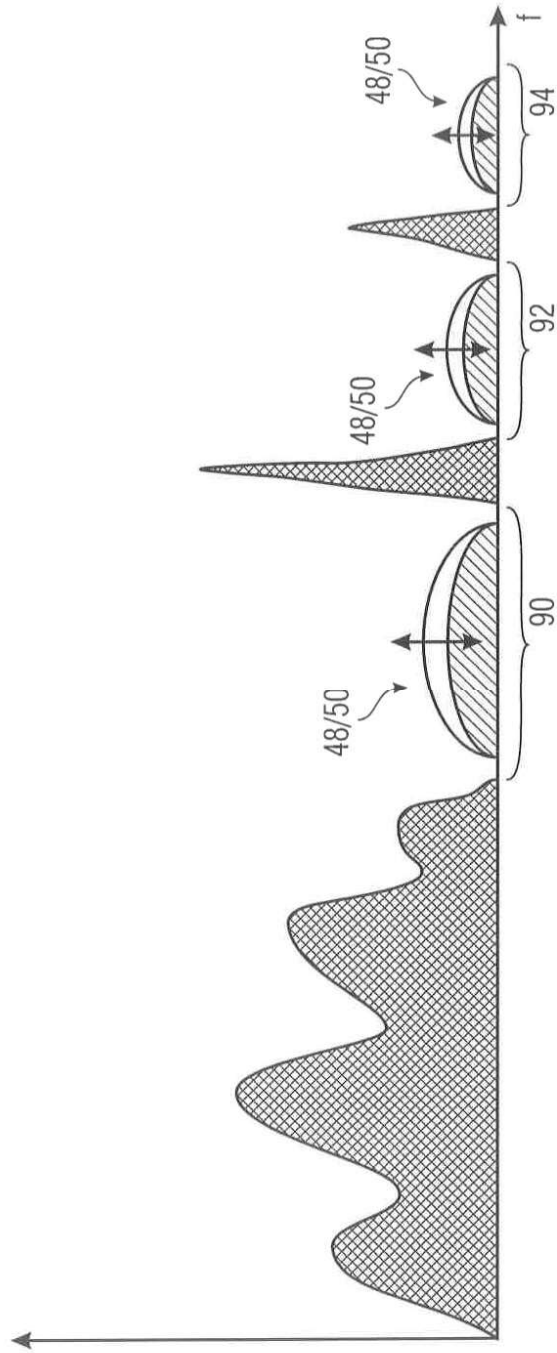


FIG 8

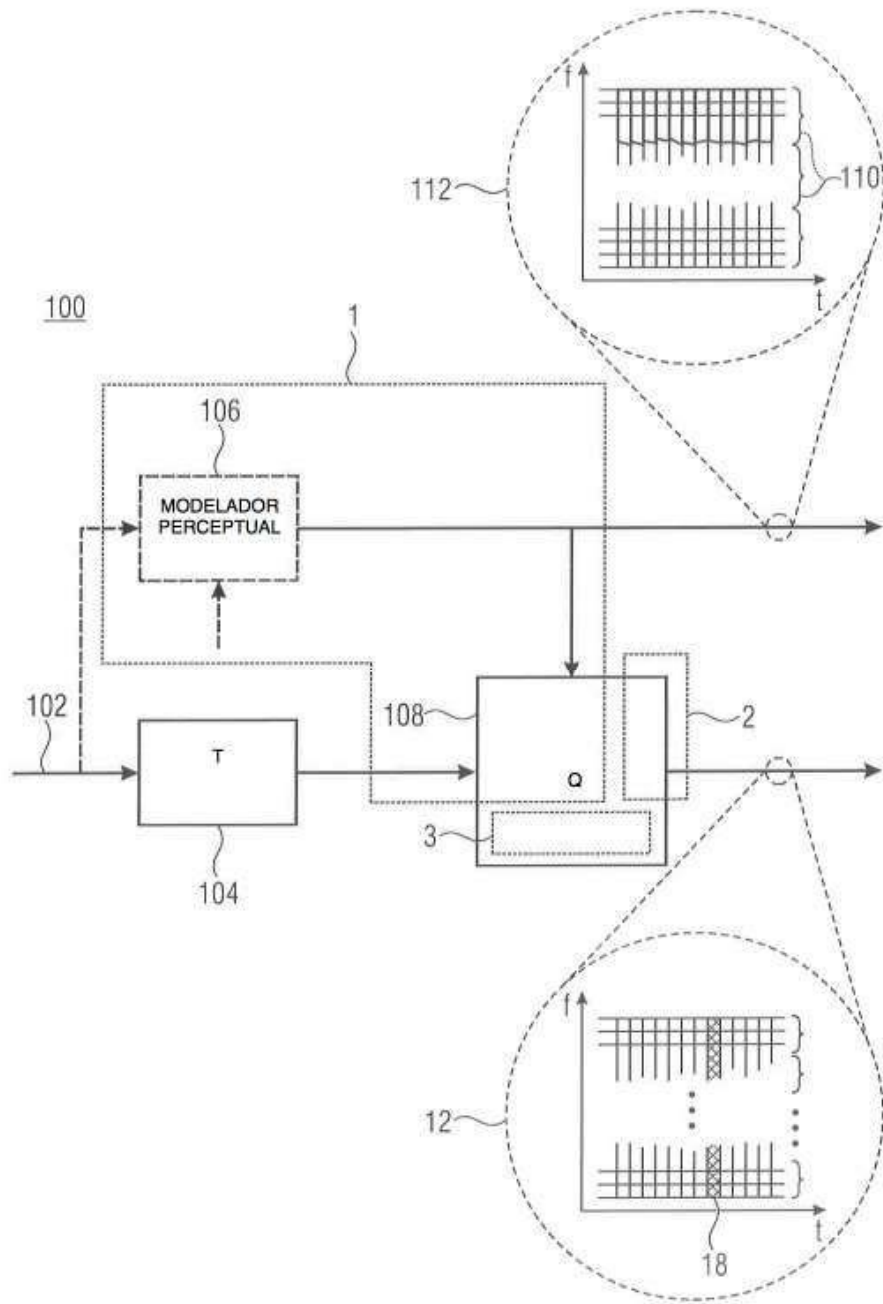


FIG 9

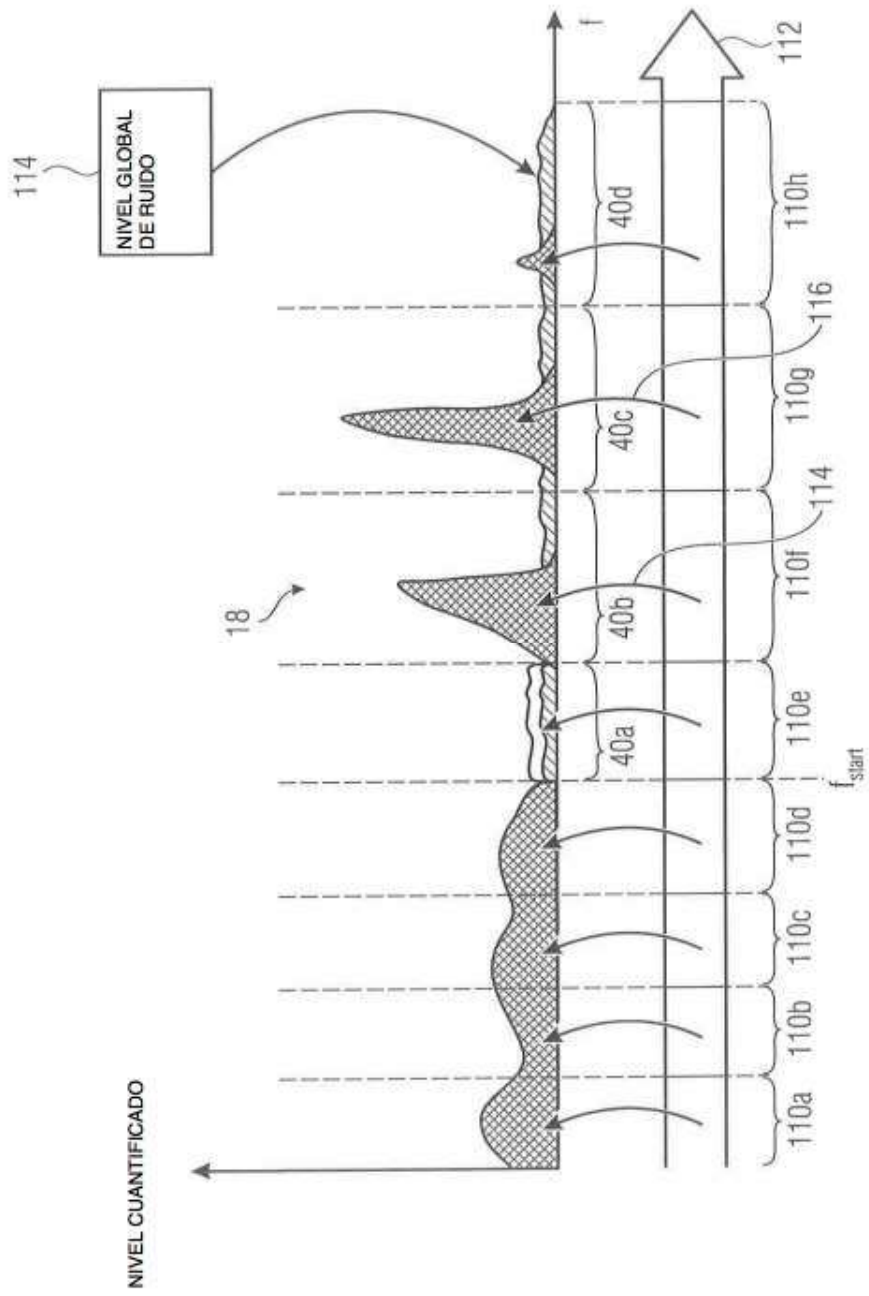


FIG 10

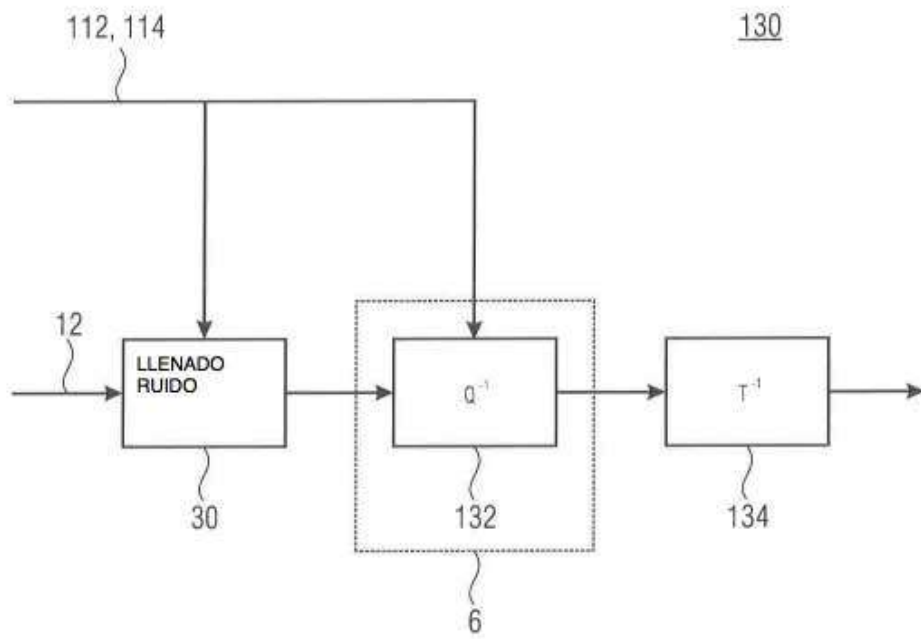


FIG 11

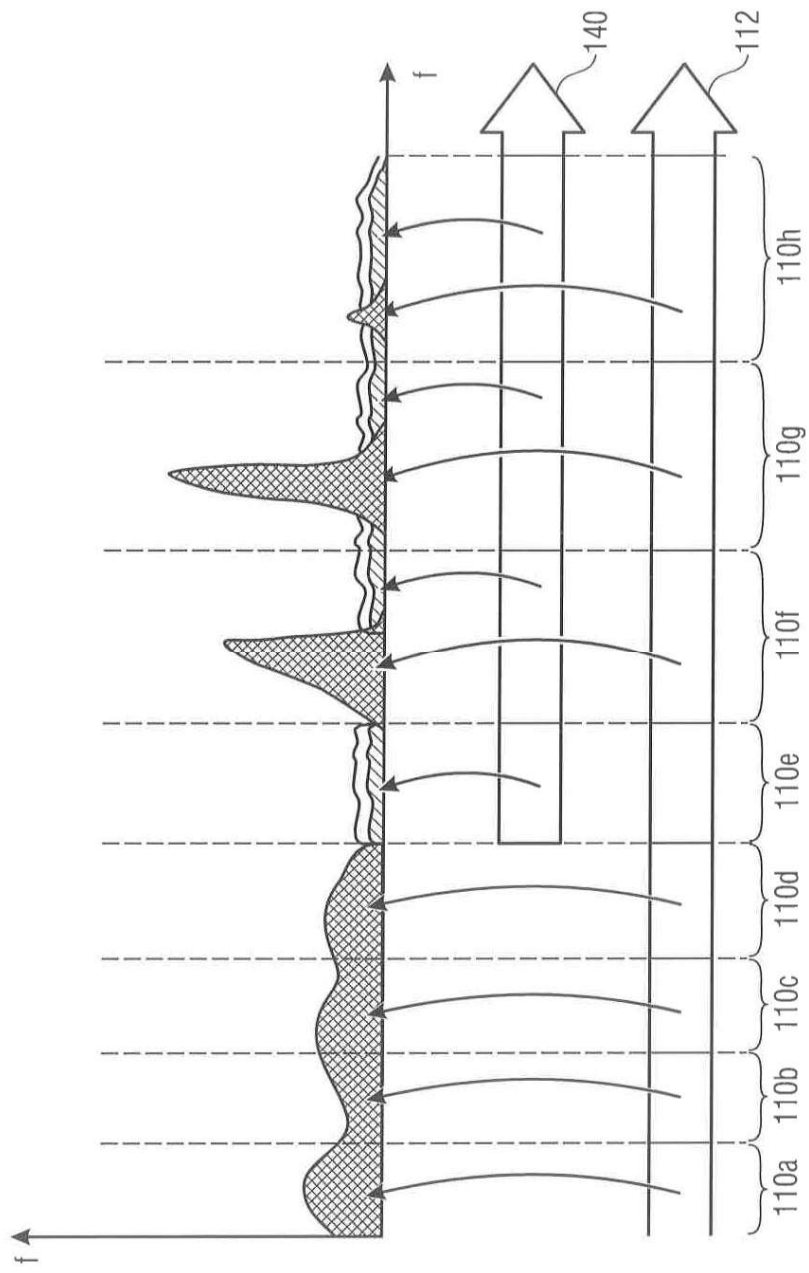


FIG 12

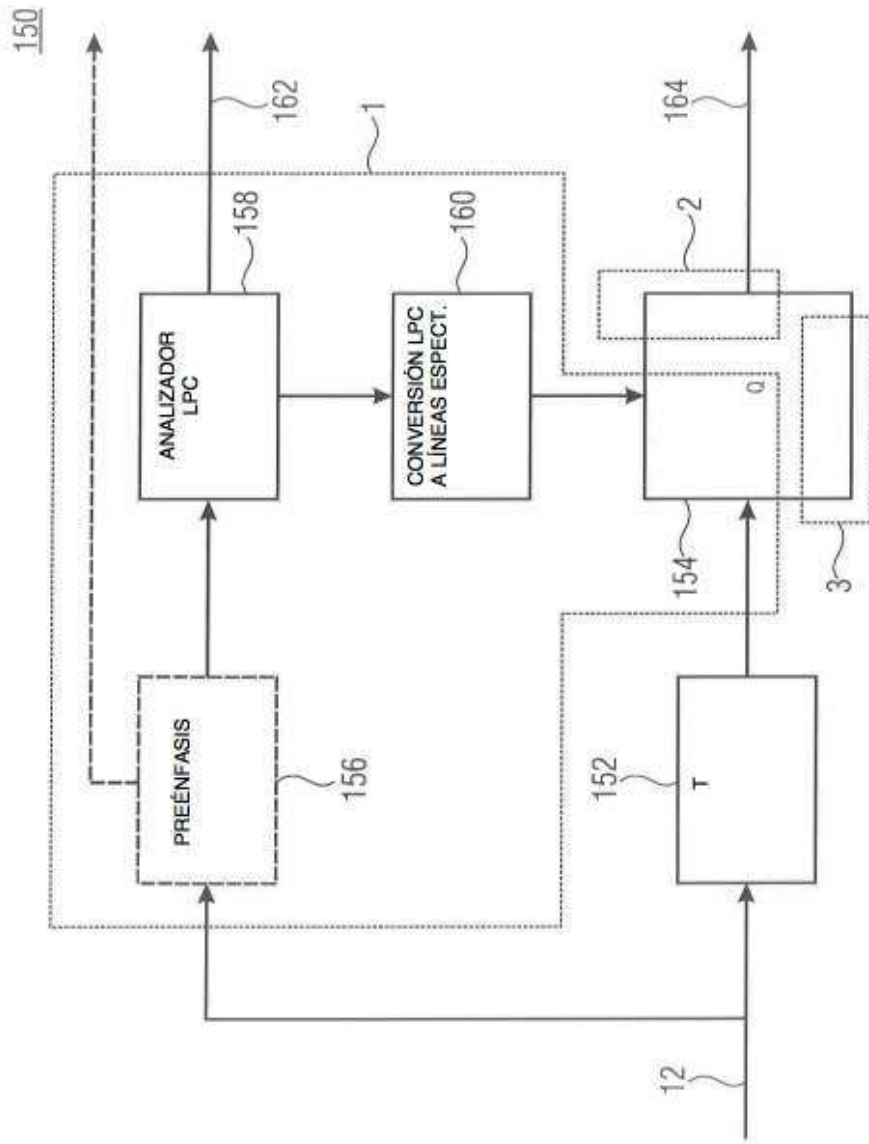


FIG 13

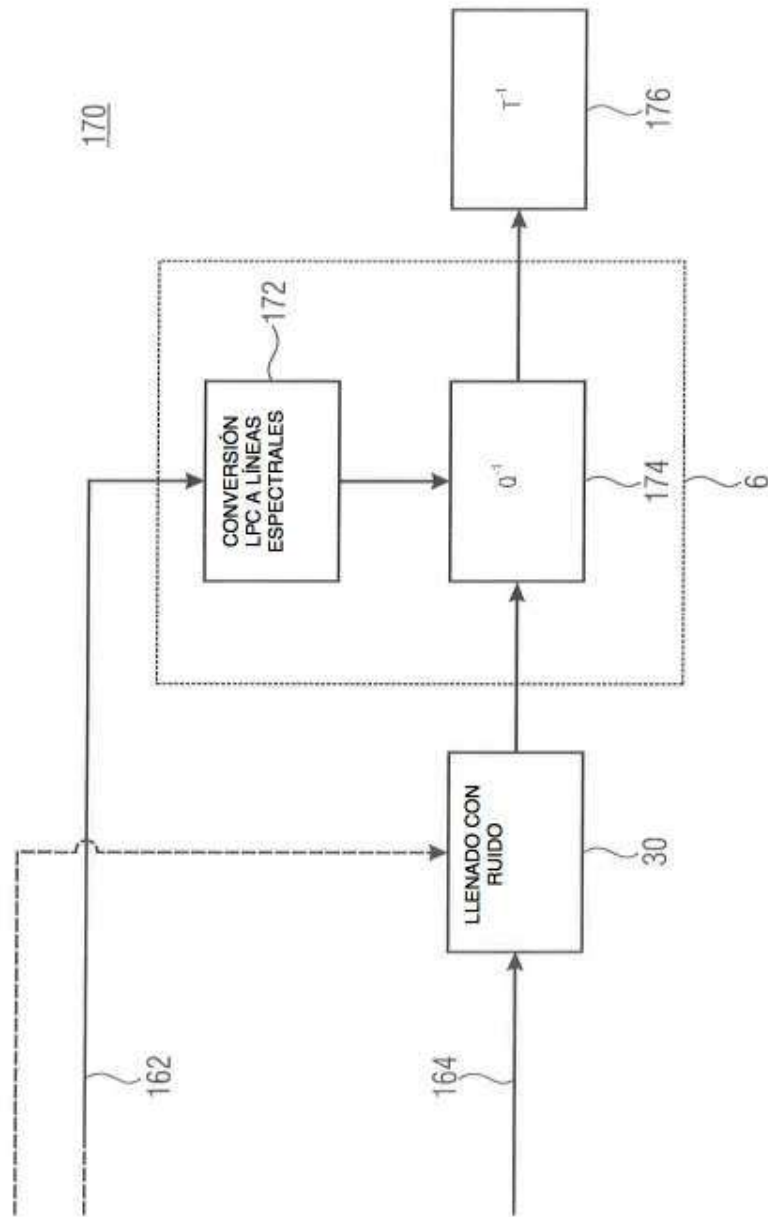


FIG 14

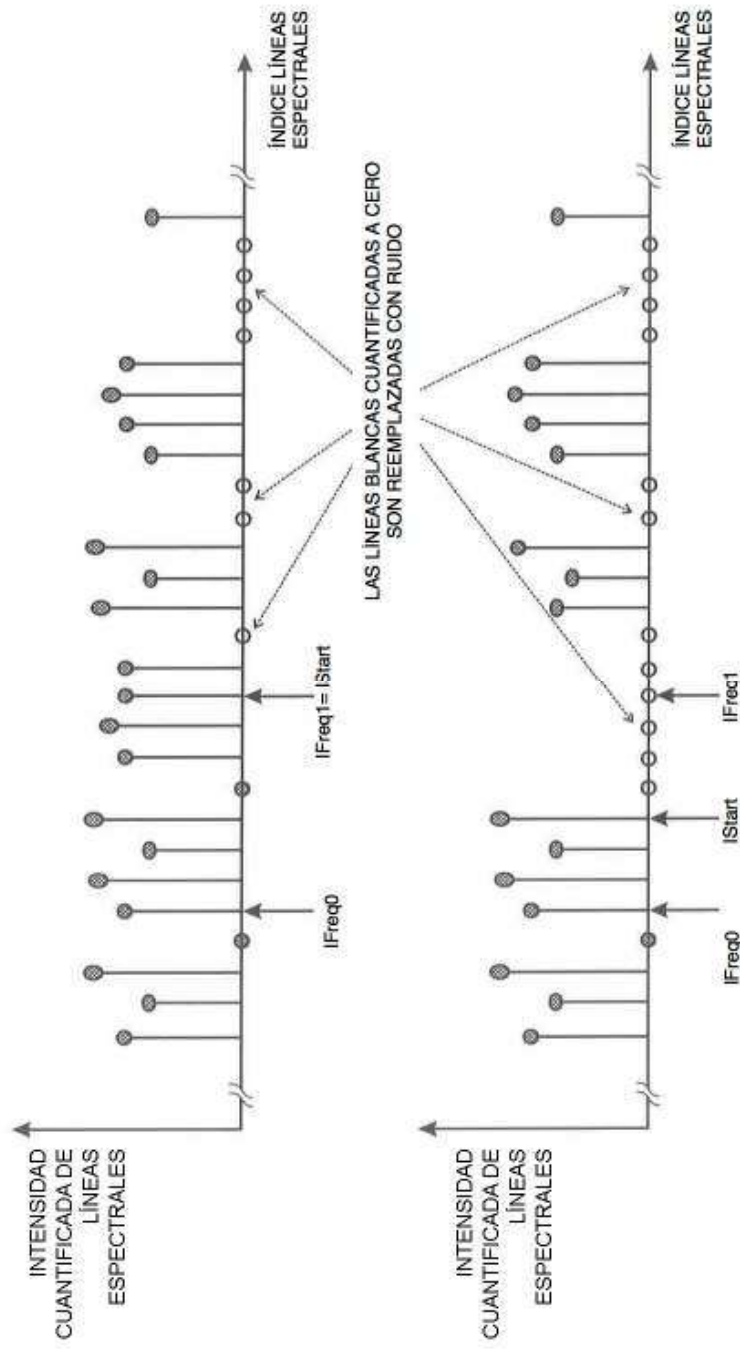


FIG 15

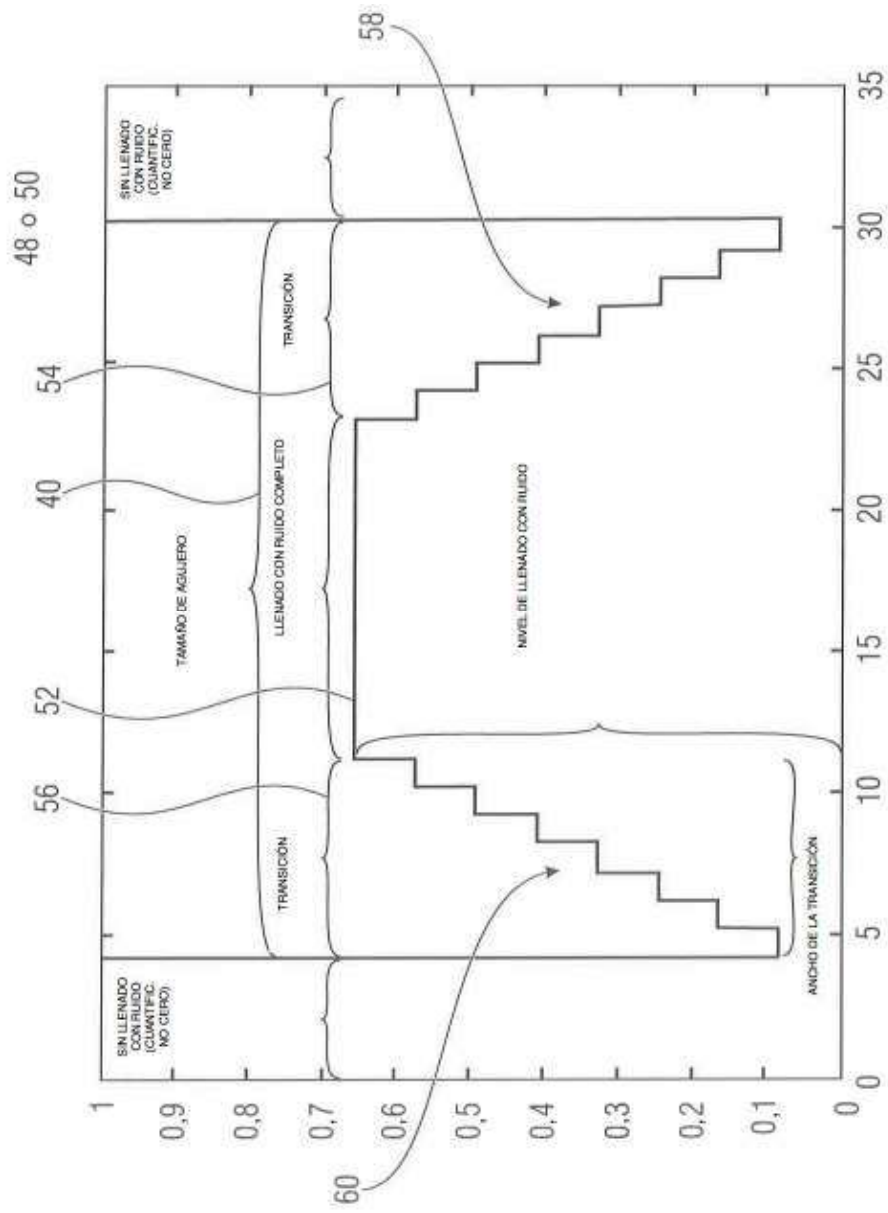
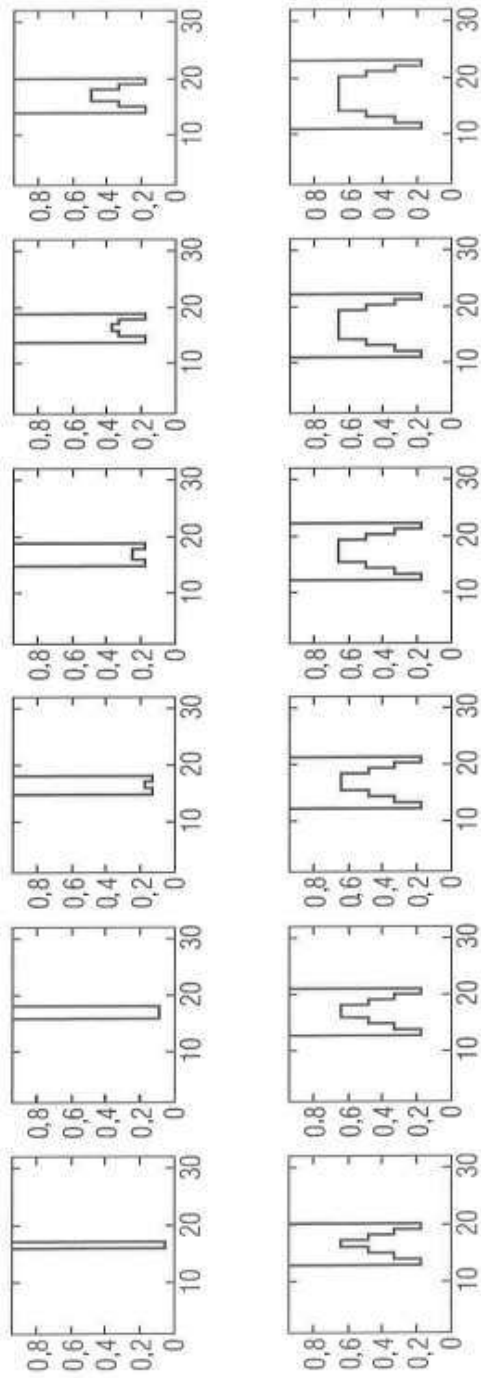
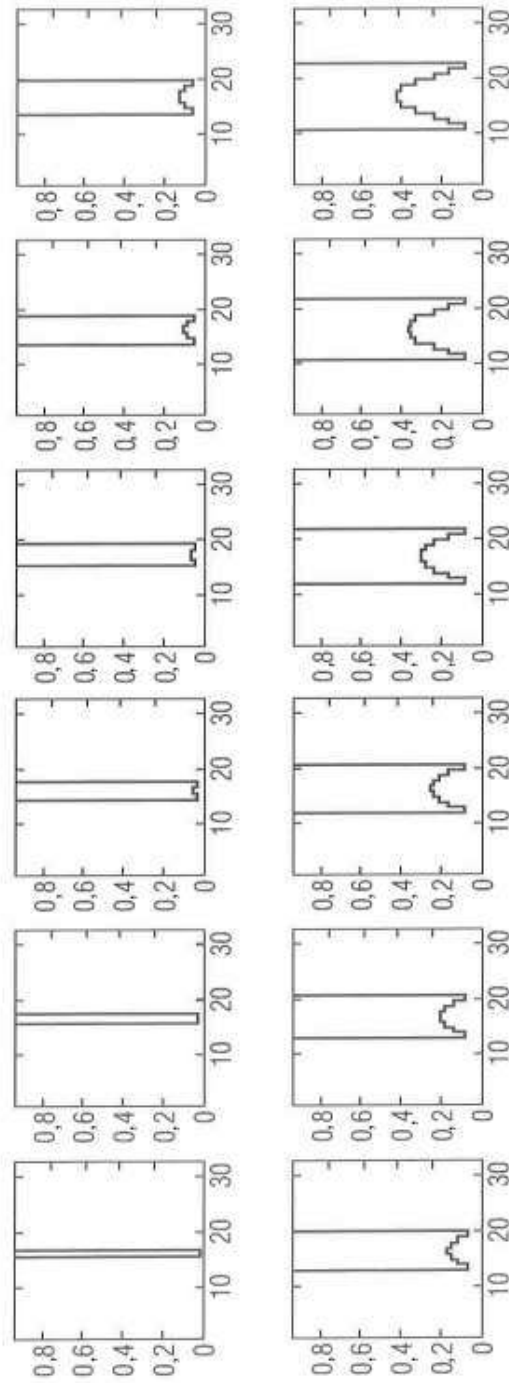


FIG 16



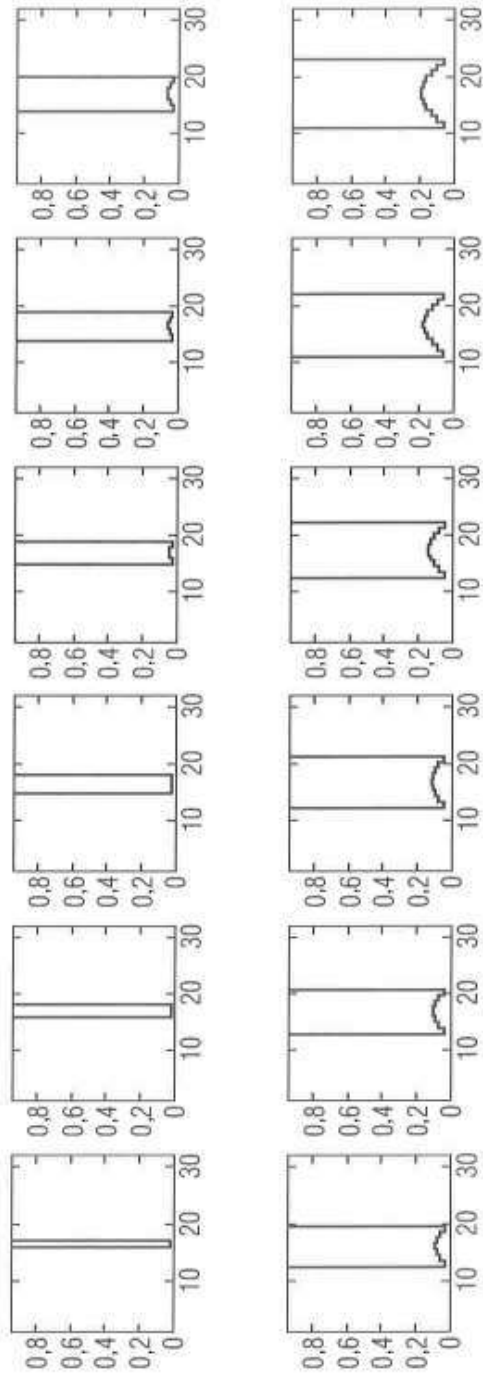
LLENADO CON RUIDO CORRESPONDIENTE A TAMAÑOS DE AGUJERO DE 1 A 12, ANCHO DE TRANSICIÓN 4

FIG 17A



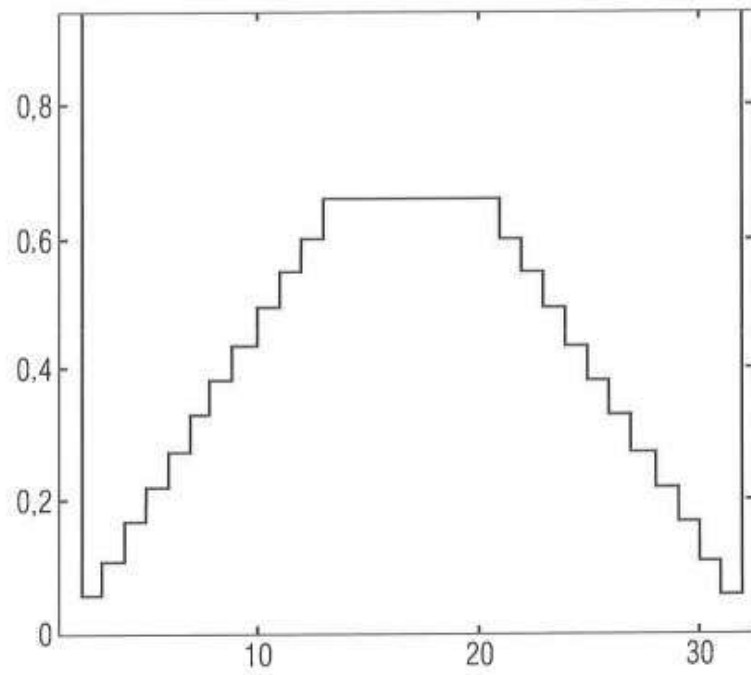
LLENADO CON RUIDO CORRESPONDIENTE A TAMAÑOS DE AGUJERO DE 1 A 12, ANCHO DE TRANSICIÓN 8

FIG 17B



LLENADO CON RUIDO CORRESPONDIENTE A TAMAÑOS DE AGUJERO DE 1 A 12, ANCHO DE TRANSICIÓN 12

FIG 17C



LLENADO CON RUIDO PARA TAMAÑOS DE AGUJERO 28, ANCHO DE TRANSICIÓN 12

FIG 17D

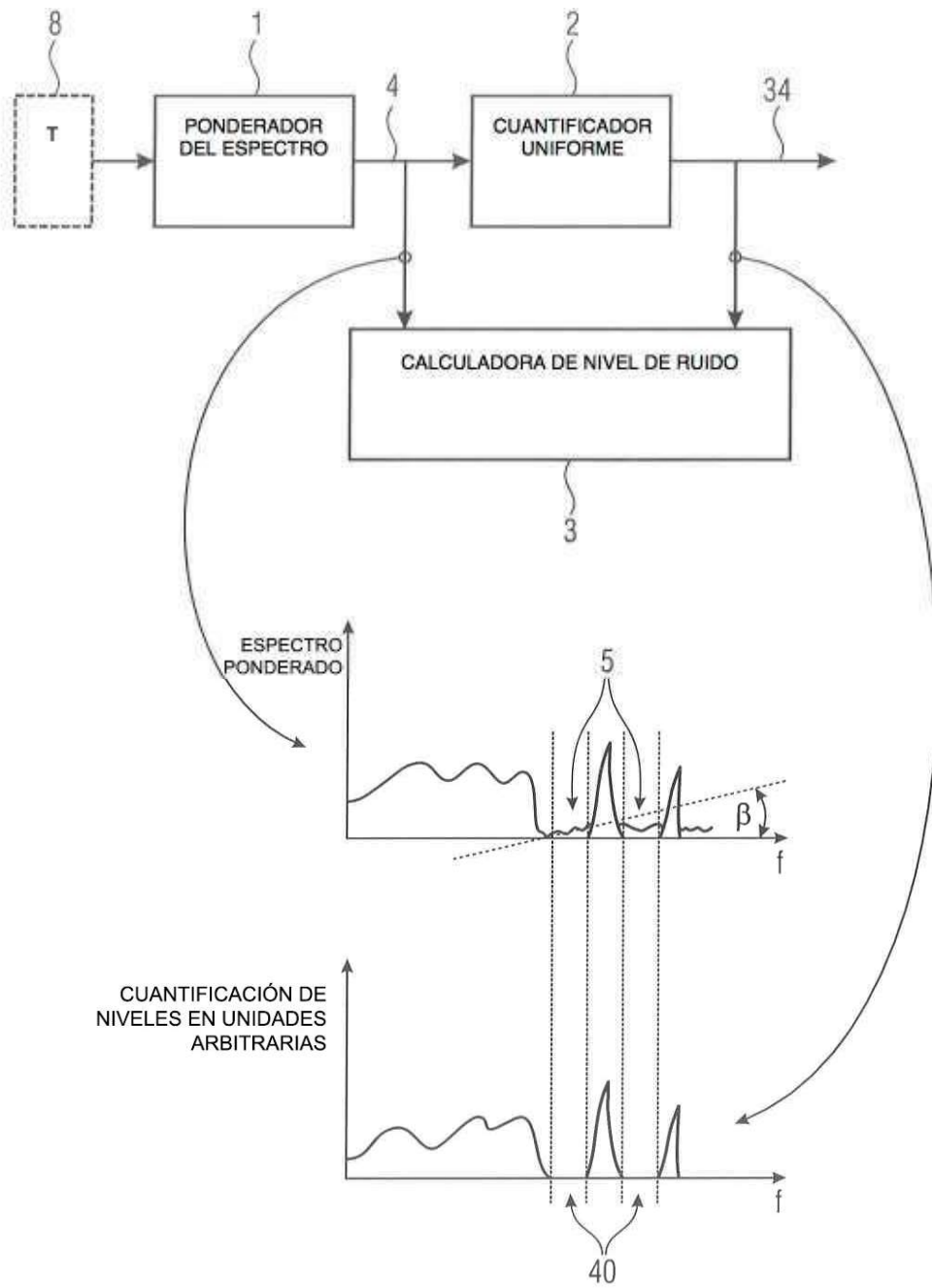


FIG 18A

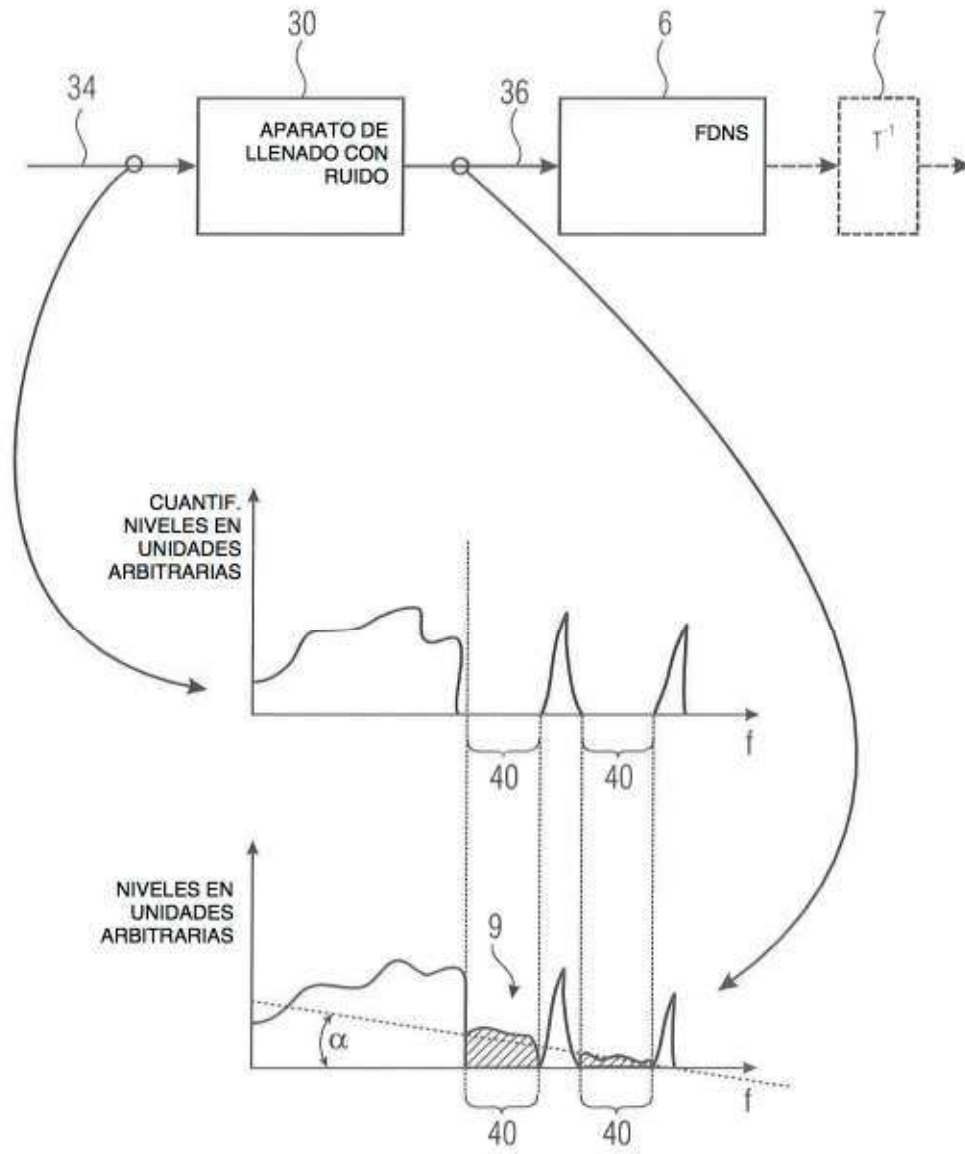


FIG 18B

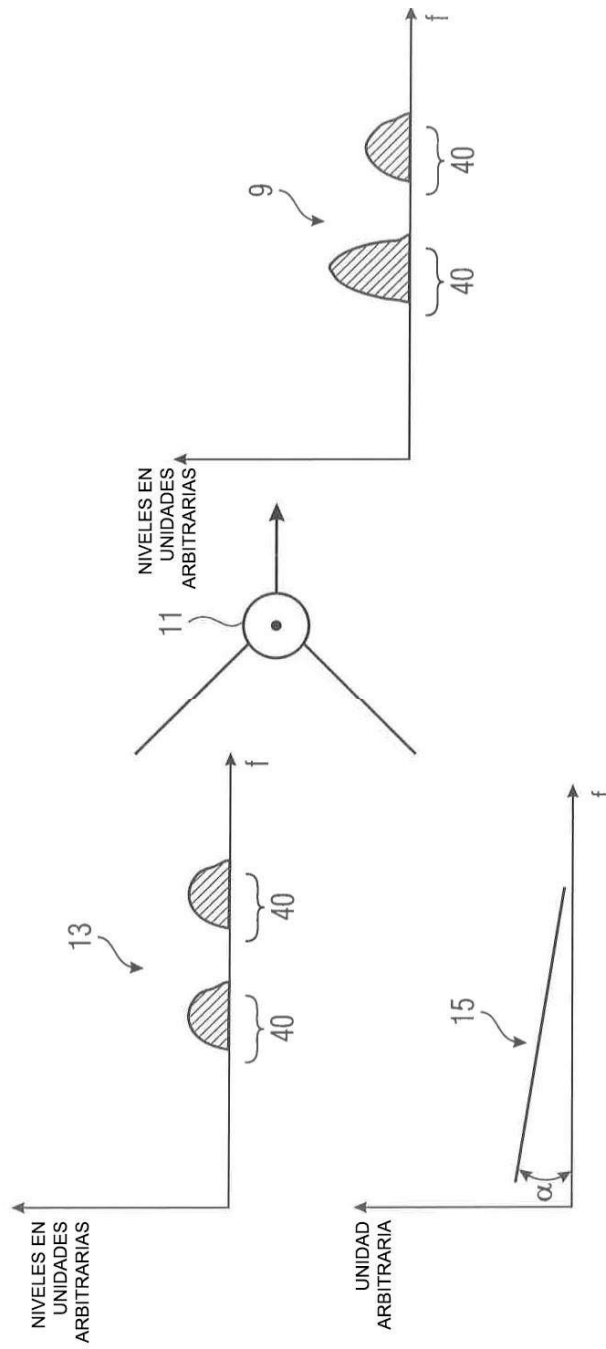


FIG 18C