

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 374 640**

51 Int. Cl.:
G10L 19/00 (2006.01)
G10L 19/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **09776859 .2**
96 Fecha de presentación: **26.06.2009**
97 Número de publicación de la solicitud: **2304720**
97 Fecha de publicación de la solicitud: **06.04.2011**

54 Título: **RELLENADOR DE RUIDO, CALCULADOR DE PARÁMETRO DE RELLENO DE RUIDO, MÉTODO PARA PROPORCIONAR UN PARÁMETRO DE RELLENO DE RUIDO, MÉTODO PARA PROPORCIONAR UNA REPRESENTACIÓN ESPECTRAL RELLENADA CON RUIDO DE UNA SEÑAL DE AUDIO, PROGRAMA INFORMÁTICO CORRESPONDIENTE Y SEÑAL DE AUDIO CODIFICADA.**

30 Prioridad:
11.07.2008 US 79872
08.10.2008 US 103820

45 Fecha de publicación de la mención BOPI:
20.02.2012

45 Fecha de la publicación del folleto de la patente:
20.02.2012

73 Titular/es:
**Fraunhofer-Gesellschaft zur Förderung der
angewandten Forschung e.V.
Hansastraße 27c
80686 München, DE**

72 Inventor/es:
**RETTLEBACH, Nikolaus;
GRILL, Bernhard;
FUCHS, Guillaume;
GEYERSBERGER, Stefan;
MULTRUS, Markus;
POPP, Harald;
HERRE, Jürgen;
WABNIK, Stefan;
SCHULLER, Gerald y
HIRSCHFELD, Jens**

74 Agente: **Arizti Acha, Monica**

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).

ES 2 374 640 T3

DESCRIPCIÓN

Rellenador de ruido, calculador de parámetro de relleno de ruido, método para proporcionar un parámetro de relleno de ruido, método para proporcionar una representación espectral rellenada con ruido de una señal de audio, programa informático correspondiente y señal de audio codificada.

5 Antecedentes de la invención

Realizaciones según la invención se refieren a un relleno de ruido para proporcionar una representación espectral rellenada con ruido de una señal de audio sobre la base de una representación espectral de entrada de la señal de audio, a un calculador de parámetro de relleno de ruido de una señal de audio, a una representación de señal de audio codificada que representa una señal de audio, a un método para proporcionar una representación espectral rellenada con ruido de una señal de audio, a un método para proporcionar un parámetro de relleno de ruido sobre la base de una representación espectral cuantificada de una señal de audio, y a programas informáticos para implementar dichos métodos.

En lo que sigue se describirán algunos escenarios en los cuales se pueden aplicar ventajosamente realizaciones según la invención. Muchos codificadores de señal de audio en el dominio de la frecuencia se basan en la idea de que algunas regiones de frecuencia o regiones espectrales (por ejemplo, líneas de frecuencia o líneas espectrales proporcionadas por una conversión de dominio del tiempo al dominio de la frecuencia), son más importantes que otras regiones espectrales. Por consiguiente, las regiones espectrales de alta importancia psicoacústica normalmente se codifican con mayor precisión que las regiones espectrales de inferior relevancia psicoacústica. La relevancia psicoacústica de las diferentes regiones espectrales puede calcularse, por ejemplo, usando un modelo psicoacústico que tiene en cuenta el enmascaramiento de regiones espectrales más débiles por picos espectrales fuertes adyacentes.

Si hay un deseo de reducir la tasa de transmisión de bits de una señal de audio codificada hacia un nivel bajo, algunas regiones espectrales se cuantifican con una precisión muy baja (por ejemplo, precisión de sólo un bit, o precisión de dos bits). Por consiguiente, muchas de las regiones espectrales cuantificadas con baja precisión se cuantifican a cero. Así, a bajas tasas de transmisión de bits los codificadores de audio basados en transformación son propensos a diferentes artefactos y especialmente a artefactos que se originan a partir de líneas de frecuencia cuantificadas a cero. En realidad, la cuantificación aproximada de valores espectrales en codificación de audio de baja tasa de transmisión de bits podría conducir a un espectro muy esparcido después de la cuantificación inversa, ya que muchas líneas espectrales podrían haberse cuantificado a cero. Estos huecos de frecuencia en la señal reconstruida producen indeseables artefactos de sonido. Puede hacer que el sonido reproducido sea demasiado agudo o inestable (*birdies*) cuando los huecos de frecuencia en el espectro se mueven de trama a trama.

El relleno de ruido es un medio para enmascarar estos artefactos rellenando, en el lado del decodificador, las bandas o los coeficientes cuantificados en cero, con ruido aleatorio. La energía del ruido insertado es un parámetro calculado y transmitido por el codificador.

Se conocen diferentes conceptos de relleno de ruido. Por ejemplo, el denominado AMR-WR+ combina relleno de ruido y una transformada discreta de Fourier (DFT), tal como se describe por ejemplo en la referencia [1]. Además, la norma internacional ITU-T G.729.1 define un concepto que combina relleno de ruido y transformada de coseno discreta modificada (MDCT). En la referencia [2] se describen detalles.

El documento WO 02/091363 A da a conocer un códec de audio en el dominio de transformación, en el que bandas de frecuencia menos relevantes desde el punto de vista perceptivo seleccionadas de una señal de audio original se sustituyen durante la codificación por un parámetro de relleno de ruido. Tras la decodificación, dicho parámetro de relleno de ruido se usa para rellenar dichas bandas menos relevantes desde el punto de vista perceptivo con ruido aleatorio, cuya amplitud se ajusta según la energía original (raíz cuadrática media) de dichas bandas. Por tanto se evitan artefactos de conmutación de ancho de banda audibles debido a la no reconstrucción de bandas cuantificadas a cero durante la codificación debido a grandes restricciones de presupuesto.

45 Aspectos adicionales relativos al relleno de ruido se describen en la solicitud de patente internacional PCT/IB2002/001388 de Koninklijke Philips Electronics N.V. (véase la referencia [3]).

De todos modos, los conceptos de relleno de ruido convencionales dan como resultado distorsiones audibles.

En vista de esta discusión, hay un deseo de crear un concepto de relleno de ruido que proporcione una mejor impresión auditiva.

50 Sumario de la invención

Una realización según la invención crea un relleno de ruido para proporcionar una representación espectral rellenada con ruido de una señal de audio sobre la base de una representación espectral de entrada de la señal de audio. El relleno de ruido comprende un identificador de región espectral configurado para identificar regiones espectrales (por ejemplo, líneas espectrales, o intervalos espectrales) de la representación espectral de entrada cuantificadas a cero y separadas de regiones espectrales distintas de cero (por ejemplo, líneas espectrales o intervalos espectrales) de la

representación espectral de entrada por al menos una región espectral intermedia, para obtener regiones espectrales identificadas. El rellenador de ruido también comprende un insertador de ruido configurado para introducir selectivamente ruido dentro de las regiones espectrales identificadas (por ejemplo, líneas espectrales o intervalos espectrales) para obtener la representación espectral rellena con ruido de la señal de audio.

5 Esta realización de la presente invención se basa en el hallazgo de que las componentes tonales de la representación espectral de una señal de audio normalmente se ven degradadas, en términos de impresión auditiva, si se aplica un relleno de ruido en la proximidad inmediata de tales componentes tonales. Por consiguiente, se ha hallado que se puede obtener una mejor impresión auditiva de una señal de audio rellena con ruido si el relleno de ruido sólo se aplica a regiones espectrales que están separadas de tales regiones tonales espectrales distintas de cero. Por consiguiente, las
10 componentes tonales del espectro de señal de audio (que no están cuantificadas a cero en la representación espectral cuantificada introducida al rellenador de ruido) permanecen audibles (es decir, no quedan afectadas por ruido adyacente muy cercano), mientras que todavía se evita eficazmente la presencia de grandes huecos espectrales.

En una realización preferida, el identificador de región espectral está configurado para identificar, como regiones espectrales identificadas, líneas espectrales de la representación espectral de entrada, que están cuantificadas a cero y que comprenden al menos un primer número predeterminado de líneas espectrales vecinas de frecuencia inferior cuantificadas a cero y al menos un segundo número predeterminado de líneas espectrales vecinas de frecuencia superior cuantificadas a cero, como regiones espectrales identificadas, en el que el primer número predeterminado es mayor o igual a uno y en el que el segundo número predeterminado es mayor o igual a uno. En esta realización, el insertador de ruido está configurado para introducir selectivamente ruido dentro de las líneas espectrales identificadas mientras que deja sin afectar por el relleno con ruido líneas espectrales cuantificadas a un valor distinto de cero y líneas espectrales cuantificadas a cero, pero que no tienen el primer número predeterminado de líneas espectrales vecinas de frecuencia inferior cuantificadas a cero, o el segundo número predeterminado de líneas espectrales vecinas de frecuencia superior cuantificadas a cero. Así, el relleno de ruido es selectivo por el hecho de que se introduce ruido sólo en líneas espectrales que están cuantificadas a cero y que están separadas de líneas cuantificadas a un valor distinto de cero, tanto en una dirección espectral ascendente como en una dirección espectral descendente, por ejemplo por el primer número predeterminado de líneas espectrales vecinas de frecuencia inferior cuantificadas a cero y por el segundo número predeterminado de líneas espectrales vecinas de frecuencia superior cuantificadas a cero.

En una realización preferida, el primer número predeterminado es igual al segundo número predeterminado, de manera que una separación mínima en la dirección de frecuencia ascendente desde líneas cuantificadas a un valor distinto de cero es igual a una separación mínima en la dirección de frecuencia descendente desde líneas cuantificadas a un valor distinto de cero.

En una realización preferida, el rellenador de ruido está configurado para introducir ruido sólo en regiones espectrales en una parte superior de la representación espectral de la señal de audio, mientras que deja una parte inferior de la representación espectral de la señal de audio sin afectar por el relleno de ruido. Un concepto de este tipo es útil ya que usualmente las frecuencias más altas son menos importantes desde el punto de vista perceptivo que las bajas frecuencias. Los valores cuantificados a cero también aparecen mayormente en la segunda mitad de los espectros (es decir, para altas frecuencias). También agregar ruido en las altas frecuencias es menos propenso a obtener un restitución de sonido final ruidosa.

En una realización preferida, el identificador de región espectral está configurado para sumar valores de intensidad de cuantificador (por ejemplo, valores de energía o valores de amplitud) de regiones espectrales en la proximidad espectral a ambos lados predeterminada de una región espectral dada (es decir, la proximidad espectral que se extiende tanto hacia las frecuencias inferiores como hacia las superiores), para obtener un valor suma, y para evaluar el valor suma para decidir si la región espectral dada es una región espectral identificada o no. Se ha hallado que un valor suma de energías de un espectro cuantificado sobre la proximidad espectral a ambos lados de una región espectral dada es una magnitud significativa para decidir si se debe aplicar relleno de ruido a la región espectral dada.

En otra realización preferida, el identificador de región espectral está configurado para explorar un rango de regiones espectrales de la representación espectral de entrada para detectar secuencias contiguas de regiones espectrales cuantificadas a cero, y para reconocer una o más regiones espectrales centrales (es decir, regiones espectrales no límite) de tales secuencias contiguas detectadas como regiones espectrales identificadas.

50 Se ha hallado que una detección de una cierta longitud de recorrido ("*run-length*") de regiones espectrales cuantificadas a cero, es una tarea que puede implementarse con complejidad computacional particularmente baja. Para identificar tal secuencia contigua de regiones espectrales, es posible decidir si todas las regiones espectrales dentro de esta secuencia de regiones espectrales están cuantificadas a cero, lo cual puede realizarse usando un algoritmo o circuito relativamente simple. Si se halla que tal secuencia contigua de regiones espectrales está cuantificada a cero, una o más regiones espectrales interiores de la secuencia (que están separadas suficientemente de regiones espectrales fuera de la presente secuencia de regiones espectrales) son tratadas como regiones espectrales identificadas. Así, explorando a través de un rango de regiones espectrales (por ejemplo, seleccionando a continuación diferentes secuencias desplazadas de regiones espectrales), se puede hacer un análisis eficaz de la representación espectral, para identificar regiones espectrales cuantificadas a cero y separadas de regiones espectrales cuantificadas a un valor distinto de cero por una distancia mínima predeterminada.

Otra realización según la invención crea un calculador de parámetro de relleno de ruido para proporcionar un parámetro de relleno de ruido sobre la base de una representación espectral cuantificada de una señal de audio. El calculador de parámetro de relleno de ruido comprende un identificador de región espectral configurado para identificar regiones espectrales de la representación espectral cuantificada cuantificadas a cero y separadas de regiones espectrales distintas de cero de la representación espectral cuantificada por al menos una región espectral intermedia, para obtener regiones espectrales identificadas. El calculador de parámetro de relleno de ruido también comprende un calculador de valor de ruido configurado para considerar selectivamente errores de cuantificación de las regiones espectrales identificadas para un cálculo del parámetro de relleno de ruido. El calculador de parámetro de relleno de ruido se basa en la idea clave de que es deseable restringir un relleno de ruido en el lado del decodificador a regiones espectrales que están separadas de regiones espectrales tonales (cuantificadas a un valor distinto de cero), y que consecuentemente el parámetro de ruido debe calcularse en el lado del codificador, tomando en consideración este concepto. Por consiguiente, se obtiene un parámetro de relleno de ruido que es particularmente bien adecuado para el concepto de decodificador descrito anteriormente. También se ha hallado que regiones espectrales, que están cuantificadas a cero, pero que están muy cerca de regiones espectrales cuantificadas a un valor distinto de cero, con frecuencia no reflejan un contenido de audio verdaderamente de tipo ruido, sino que más bien están fuertemente correlacionadas con una región espectral tonal adyacente (cuantificada a un valor distinto de cero). Por consiguiente, se ha hallado que generalmente no es deseable considerar el error de cuantificación de regiones espectrales, que están cerca de regiones espectrales cuantificadas a un valor distinto de cero para un cálculo de un parámetro de relleno de ruido, porque esto normalmente daría como resultado una fuerte sobreestimación de ruido, dando como resultado de ese modo una representación espectral reconstruida demasiado ruidosa.

Así, el concepto de cálculo de parámetro de relleno de ruido descrito en el presente documento puede usarse en combinación con el concepto de relleno de ruido descrito anteriormente e incluso en combinación con conceptos de relleno de ruido convencionales.

En realizaciones preferidas, el concepto para la identificación de regiones espectrales, que se ha comentado con respecto al relleno de ruido, también puede aplicarse en combinación con el calculador de parámetro de relleno de ruido.

En una realización preferida adicional, el calculador de valor de ruido está configurado para considerar una energía real de error de cuantificación de las regiones espectrales identificadas para el cálculo del parámetro de relleno de ruido. Se ha hallado que la consideración de un error de cuantificación real (en lugar de un error de cuantificación estimado o un error de cuantificación promedio) normalmente conlleva mejores resultados, porque el error de cuantificación real normalmente se desvía del error de cuantificación estadísticamente esperado.

En una realización preferida adicional, el calculador de valor de ruido está configurado para enfatizar una energía de error de cuantificación no tonal distribuida sobre una pluralidad de regiones espectrales identificadas en relación a una energía de error de cuantificación tonal concentrada en una sola región espectral. Este concepto se basa en el hallazgo de que un ruido de banda ancha no tonal, cuya energía promedio se sitúa por debajo de un umbral de cuantificación y que por lo tanto está cuantificado a cero, es desde el punto de vista perceptivo mucho más relevante para el relleno de ruido que una única componente de audio tonal, cuya intensidad se sitúa por debajo del umbral de cuantificación, incluso si el ruido de banda ancha no tonal cuantificado a cero y la componente tonal cuantificada a cero estuviesen ambos cuantificados a cero. La razón es que el relleno de ruido generando un ruido aleatorio en el decodificador puede modelar ruido de banda ancha no tonal ausente en la representación espectral cuantificada pero no componentes tonales ausentes. Así, un énfasis de las componentes de ruido no tonales cuantificadas a cero sobre las componentes tonales cuantificadas a cero, lleva consigo una reconstrucción de sonido más realista. Esto también se debe al hecho de que una impresión auditiva humana se degrada mucho más por la presencia de un hueco espectral (por ejemplo, en forma de ausencia de un ruido de banda ancha cuantificado a cero) que por la ausencia de un pequeño pico espectral cuantificado a cero. Una componente tonal puede concentrarse en una sola línea espectral, o puede esparcirse sobre varias líneas espectrales contiguas (por ejemplo, $i-1$, i , $i+1$). Una región espectral puede comprender, por ejemplo, una o más líneas espectrales.

En una realización preferida, el calculador de valor de ruido está configurado para calcular una suma de energías de error de cuantificación logaritmizadas de las regiones espectrales identificadas para obtener el parámetro de relleno de ruido. Calculando la suma de energías de error de cuantificación logaritmizadas de las regiones espectrales identificadas, el énfasis relativo descrito anteriormente de regiones espectrales no tonales cuantificadas a cero sobre regiones tonales cuantificadas a cero, puede obtenerse de una manera eficaz.

Otra realización según la invención crea una representación de señal de audio codificada, para representar una señal de audio. La representación de señal de audio codificada comprende una representación en el dominio espectral cuantificada codificada de la señal de audio y un parámetro de relleno de ruido codificado. El parámetro de relleno de ruido representa un error de cuantificación de las regiones espectrales de la representación en el dominio espectral cuantificadas a cero y separadas de regiones espectrales de la representación en el dominio espectral cuantificadas a un valor distinto de cero por al menos un número predeterminado de regiones espectrales intermedias. La representación de señal de audio codificada descrita anteriormente puede usarse por el relleno de ruido comentado anteriormente y puede obtenerse usando el calculador de parámetro de relleno de ruido comentado anteriormente. La representación de señal de audio codificada permite una reconstrucción de la señal de audio con calidad de audio

particularmente buena porque el parámetro de relleno de ruido refleja selectivamente el error de cuantificación de la representación en el dominio espectral cuantificada para aquellas regiones espectrales en las que está presente una información de ruido significativa y que deben considerarse selectivamente para un relleno de ruido en el lado del decodificador.

5 Otra realización según la invención crea un método para proporcionar una representación rellena con ruido de una señal de audio.

Otra realización adicional según la invención crea un método para proporcionar un parámetro de relleno de ruido sobre la base de una representación espectral cuantificada de una señal de audio.

10 Otra realización adicional según la invención crea un programa informático para implementar los métodos mencionados anteriormente.

Breve descripción de los dibujos

A continuación se describirán realizaciones según la invención, tomando como referencia las figuras adjuntas, en las que:

15 la figura 1 muestra un diagrama de bloques esquemático de un relleno de ruido, según una realización de la invención;

la figura 2 muestra un diagrama de bloques esquemático de un decodificador de señal de audio que comprende el relleno de ruido según la presente invención;

la figura 3 muestra un pseudocódigo de programa para implementar la funcionalidad del relleno de ruido de la figura 1;

20 la figura 4 muestra una representación gráfica de una identificación de regiones espectrales, que puede realizarse en el relleno de ruido según el figura 1;

la figura 5 muestra un diagrama de bloques esquemático de un calculador de parámetro de relleno de ruido según una realización de la invención;

25 la figura 6 muestra un pseudocódigo de programa para implementar la funcionalidad del calculador de parámetro de relleno de ruido según la figura 5;

la figura 7 muestra un diagrama de flujo de un método para proporcionar una representación espectral rellena con ruido de una señal de audio sobre la base de una representación espectral de entrada de la señal de audio;

la figura 8 muestra un diagrama de flujo de un método para proporcionar un parámetro de relleno de ruido sobre la base de una representación espectral cuantificada de una señal de audio; y

30 la figura 9 muestra una representación gráfica de una representación de señal de audio, según una realización de la invención.

Rellenador de ruido según las figuras 1–4

35 La figura 1 muestra un diagrama de bloques esquemático de un relleno 100 de ruido, según una realización de la invención. El relleno 100 de ruido está configurado para recibir una representación 110 espectral de entrada de una señal de audio, por ejemplo en forma de coeficientes espectrales decodificados (que pueden, por ejemplo, estar cuantificados o inversamente cuantificados). El relleno 100 de ruido también está configurado para proporcionar una representación 112 espectral rellena con ruido de la señal de audio sobre la base de la representación 110 espectral de entrada.

40 El relleno 100 de ruido comprende un identificador 120 de región espectral, que está configurado para identificar regiones espectrales de la representación 110 espectral de entrada separadas de regiones espectrales distintas de cero de la representación 110 espectral de entrada por al menos una región espectral intermedia, para obtener una información 122 que indica las regiones espectrales identificadas. El relleno 100 de ruido también comprende un insertador 130 de ruido, que está configurado para introducir selectivamente ruido dentro de las regiones espectrales identificadas (descritas mediante la información 122), para obtener la representación 112 espectral rellena con ruido de la señal de audio.

45 Con relación a la funcionalidad del relleno 100 de ruido, en general se puede decir que el relleno 100 de ruido rellena selectivamente regiones espectrales (por ejemplo, líneas espectrales o intervalos espectrales) de la representación 110 espectral de entrada con ruido, por ejemplo reemplazando valores espectrales de líneas espectrales cuantificadas a cero por valores espectrales de reemplazo que describen un ruido. De esta manera se pueden rellenar huecos espectrales o vacíos espectrales dentro de la representación 110 espectral de entrada, los cuales pueden aparecer, por ejemplo, a partir de una cuantificación aproximada de la representación 110 espectral de entrada. Sin

embargo, el rellenedor 100 de ruido no introduce ruido dentro de todas las líneas espectrales cuantificadas a cero (es decir, líneas espectrales, cuyos valores espectrales están cuantificados a cero). En cambio, el rellenedor 100 de ruido sólo introduce ruido en aquellas líneas espectrales cuantificadas a cero que comprenden una suficiente distancia desde cualquier línea espectral cuantificada a un valor distinto de cero. De esta manera, el relleno de ruido no rellena completamente huecos espectrales o vacíos espectrales, sino que mantiene una distancia espectral de al menos una región espectral (o de al menos cualquier otro número predeterminado de regiones espectrales) entre aquellas líneas espectrales en las que se introduce un ruido y líneas espectrales cuantificadas a un valor distinto de cero. Así, se mantiene una distancia espectral entre ruido de relleno, introducido en la representación espectral, y líneas espectrales cuantificadas a un valor distinto de cero, de tal manera que las líneas espectrales psicoacústicamente relevantes (que no están cuantificadas a cero en la representación espectral de entrada de la señal de audio) pueden distinguirse claramente (debido a la distancia espectral del número predeterminado de una o más regiones espectrales) del ruido de relleno introducido dentro del espectro por el rellenedor de ruido. Por consiguiente, se puede percibir claramente el contenido de audio psicoacústicamente más relevante (representado por valores de línea espectral distintos de cero en la representación 110 espectral de entrada), mientras se evitan grandes huecos espectrales. Esto se debe al hecho de que el relleno de ruido se omite selectivamente en la proximidad de líneas espectrales de la representación espectral de entrada cuantificadas a un valor distinto de cero, mientras que se ejecuta relleno de ruido en las regiones centrales de huecos espectrales o vacíos espectrales.

En lo que sigue se describirá un entorno de aplicación para el rellenedor 100 de ruido haciendo referencia a la figura 2. La figura 2 muestra un diagrama de bloques esquemático de un decodificador 200 de señal de audio, según una realización de la invención. El decodificador 200 de señal de audio comprende, como componente clave, el rellenedor 100 de ruido. El decodificador 200 de señal de audio también comprende un decodificador 210 de coeficiente espectral, que está configurado para recibir una representación 212 de señal de audio codificada y para proporcionar una representación 214 decodificada, opcionalmente inversamente cuantificada, de coeficientes espectrales de la señal de audio codificada. El decodificador 210 de coeficiente espectral puede comprender, por ejemplo, un decodificador de entropía (por ejemplo, decodificador aritmético o decodificador de longitud de recorrido) y, opcionalmente, un cuantificador inverso para derivar la representación 214 decodificada de los coeficientes espectrales (por ejemplo, en forma de coeficientes inversamente cuantificados) a partir de la representación 212 de señal de audio codificada. El rellenedor 100 de ruido está configurado para recibir la representación 214 decodificada de coeficientes espectrales (que está opcionalmente inversamente cuantificada) como la representación 110 espectral de entrada de la señal de audio.

El decodificador 200 de señal de audio también comprende un extractor 220 de factor de ruido, que está configurado para extraer una información 222 de factor de ruido de la representación 212 de señal de audio codificada y para proporcionar la información 222 de factor de ruido al rellenedor 100 de ruido. El decodificador 200 de señal de audio también comprende un reconformador 230 de espectro, que está configurado para recibir una representación 232 de espectro reconstruida desde el rellenedor 100 de ruido. La representación 232 de espectro reconstruida puede ser, por ejemplo, igual a la representación 112 espectral rellena con ruido proporcionada por el rellenedor de ruido. El reconformador 230 de espectro, que puede ser considerado opcional, está configurado para proporcionar una información 234 de espectro sobre la base de la representación 232 de espectro reconstruida. El decodificador 200 de señal de audio además comprende un conversor 240 de dominio espectral a dominio del tiempo, que recibe la información 234 de espectro proporcionada por el reconformador 230 de espectro o, en ausencia del reconformador 230 de espectro, la representación 232 de espectro reconstruida, y para proporcionar basándose en ello, una representación 242 de señal de audio en el dominio del tiempo. El conversor 240 de dominio espectral a dominio del tiempo puede estar configurado, por ejemplo, para realizar una transformada de coseno discreta modificada inversa (IMDCT).

En una realización preferida, el relleno de ruido en el lado del decodificador comprende las siguientes etapas (o sigue las siguientes etapas):

1. decodificar el suelo de ruido;
2. decodificar los valores cuantificados de las líneas de frecuencia;
3. detectar las regiones espectrales en la parte seleccionada de los espectros donde una longitud de recorrido de ceros es mayor que un tamaño mínimo de longitud de recorrido; y
4. aplicar un signo generado aleatoriamente al suelo de ruido decodificado para cada una de las líneas dentro de las regiones seleccionadas.

El suelo de ruido se decodifica como sigue:

$$nf_decoded = 0,0625*(8-index).$$

Las regiones espectrales detectadas se seleccionan, por ejemplo, de la misma manera en que se hace en lado del codificador (que se describirá más adelante).

Un ruido gaussiano sin memoria en el dominio de MDCT se genera por un espectro con la misma amplitud para todas las líneas pero con signos aleatorios. Así, para cada una de las líneas dentro de las regiones seleccionadas, el

decodificador genera un signo aleatorio (-1 ó +1) y lo aplica al suelo de ruido decodificado. Sin embargo, se pueden aplicar también otros métodos para proporcionar una contribución de ruido.

En lo que sigue se describirán algunos detalles tomando como referencia las figuras 1, 2, 3 y 4, en las que la figura 3 muestra un pseudocódigo de programa de un algoritmo para relleno de ruido en el lado del decodificador, que puede realizarse mediante el relleno 100 de ruido, y en las que la figura 4 muestra una representación gráfica del relleno con ruido.

Para empezar, la decodificación del suelo de ruido puede realizarse mediante el extractor 220 de factor de ruido, que recibe, por ejemplo, un índice de factor de ruido (al que también se designa de manera abreviada "index") y proporciona basándose en ello, el valor 222 de factor de ruido decodificado (también designado con "nf_decoded"). El índice de factor de ruido puede codificarse, por ejemplo, usando tres o cuatro bits, y puede ser, por ejemplo, un valor entero en el rango entre 0 y 7, o un valor entero en un rango entre 0 y 15.

Los valores cuantificados de las líneas de frecuencia (también designados como "líneas espectrales" o "intervalos espectrales") pueden proporcionarse por el decodificador 210 de coeficiente espectral. Por consiguiente, se obtienen valores de línea espectral (a los que también se designa como "coeficientes espectrales) cuantificados (u opcionalmente, inversamente cuantificados), que se designan como "quantized (x(i))". Aquí, i designa un índice de frecuencia de los valores de línea espectral.

A continuación se detectan regiones espectrales mediante el relleno 100 de ruido en una parte seleccionada de los espectros (por ejemplo, en una parte superior del espectro partiendo de un índice de frecuencia i de línea espectral predeterminada) donde una longitud de recorrido de ceros (es decir, de valores de línea espectral cuantificados, cuantificados a cero) es mayor que un tamaño de longitud de recorrido mínimo. La detección de tales regiones espectrales se realiza por una primera parte 310 del algoritmo 300 de la figura 3. Como se puede ver a partir de la primera parte 310 del algoritmo 300, un conjunto R de regiones detectadas se inicializa para ser un conjunto vacío en el inicio del algoritmo ($R = \{\}$).

En el ejemplo del algoritmo de la figura 3, se fija una longitud de recorrido mínima en un valor fijo de 8, pero naturalmente se puede elegir cualquier otro valor.

A continuación, se determina para una pluralidad de líneas espectrales bajo consideración (designadas por la variable continua "line index") si cada una de esas líneas espectrales bajo consideración comprende un entorno a ambos lados de líneas espectrales cuantificadas a cero (y si la línea espectral bajo consideración está en sí misma cuantificada a cero). Por ejemplo, todas las líneas en la segunda mitad de los espectros pueden considerarse sucesivamente como líneas bajo consideración, designándose una línea que está actualmente bajo consideración por "line index". Para una línea bajo consideración designada por "line index", se calcula una suma de coeficientes espectrales cuantificados "quantized(x(i))" en un entorno que oscila desde un índice de frecuencia de línea espectral de "line index - (MinimalRunLength)/2" hasta un índice de frecuencia de línea espectral de "line index + MinimalRunLength)/2". Si se halla que la suma de los valores de línea espectral en dicho entorno de la línea espectral actualmente bajo consideración (que tiene el índice de frecuencia de línea espectral "line index") es cero, entonces, la línea espectral actualmente bajo consideración (o más precisamente, el índice de frecuencia de línea espectral "line index" de la misma) se agrega al conjunto R de regiones detectadas (o líneas espectrales detectadas). Por consiguiente, si el índice de frecuencia de línea espectral de una línea espectral se agrega al conjunto R, esto significa que las líneas espectrales que tienen índices de línea entre "line index - MinimalRunLength)/2" y "line index + MinimalRunLength)/2", comprenden todas valores de línea espectral cuantificados a cero.

Por consiguiente, en la primera parte 310 del pseudocódigo 310 de programa, se obtiene un conjunto R de índices de frecuencia de línea espectral "line index", que enumera aquellas (y sólo aquellas) líneas espectrales de la parte espectral bajo consideración que están "suficientemente" separadas (es decir, por al menos MinimalRunLength/2 líneas) de cualquier línea espectral cuantificada a un valor distinto de cero.

La detección de tal región está ilustrada en la figura 4, que muestra una representación gráfica 400 de un espectro. Una abscisa 410 describe una frecuencia de líneas espectrales en términos de un índice de frecuencia de línea espectral "line index". Una ordenada 412 describe una intensidad (por ejemplo, amplitud o energía) de las líneas espectrales. Como se puede ver, la parte del espectro ilustrada en la representación gráfica 400 comprende cuatro líneas 420a, 420b, 420c y 420d espectrales, cuantificadas a un valor distinto de cero. Además, entre las líneas 420c y 420d espectrales, hay 11 líneas 422a-422k espectrales cuantificadas a cero. Además, se supone que una línea espectral sólo se considera que está suficientemente separada de una línea espectral cuantificada a un valor distinto de cero si hay al menos cuatro líneas espectrales cuantificadas a cero entre la línea espectral actualmente bajo consideración y cualquier otra línea espectral cuantificada a un valor distinto de cero (y naturalmente, si la línea espectral actualmente bajo consideración está en sí misma cuantificada a cero). Sin embargo, cuando se considera la línea 422a espectral, se hallará que la línea 422a espectral es inmediatamente adyacente a la línea 422c espectral, que no está cuantificada a cero, de manera que el índice de frecuencia de línea espectral de la línea 422a espectral no formará parte del conjunto R calculado según la primera parte 310 del algoritmo 300. De manera similar, se hallará que las líneas 422b, 422c y 422d espectrales no están separadas suficientemente de cualquier línea espectral cuantificada a un valor distinto de cero, de manera que los índices de frecuencia de línea espectral de las líneas 422b a 422d espectrales tampoco

5 formarán parte del conjunto R. En cambio, se reconocerá que la línea 422e espectral está separada suficientemente de cualquier línea espectral cuantificada a un valor distinto de cero, porque la línea 422e espectral es una línea de centro (o, más generalmente, una línea central), de una secuencia de 9 líneas espectrales contiguas todas cuantificadas a cero. Por consiguiente, un índice de frecuencia de línea espectral de la línea 422e espectral formará parte del conjunto R calculado en la primera parte 310 del algoritmo 300. Lo mismo es válido para las líneas 422f y 422g espectrales, de manera que los índices de frecuencia de línea espectral de las líneas 422f y 422g espectrales formarán parte del conjunto R determinado en la primera parte 310 del algoritmo 300, ya que las líneas 422f y 422g espectrales están separadas suficientemente de cualquier línea 420a, 420b y 420c espectral de frecuencia inferior, cuantificadas a un valor distinto de cero y de cualquier línea espectral de frecuencia superior cuantificada a un valor distinto de cero. Por otro lado, las líneas 422h, 422i, 422j y 422k espectrales no formarán parte del conjunto R, porque dichas líneas espectrales están ubicadas demasiado cerca, en términos de frecuencia, al lado de la línea 420d espectral cuantificada a un valor distinto de cero.

15 Por consiguiente, el conjunto R no comprenderá índices de frecuencia de línea espectral de las líneas espectrales 420a, 420b, 420c, 420d, porque dichas líneas espectrales están cuantificadas a un valor distinto de cero. Además, los índices de frecuencia de línea espectral de las líneas 422a, 422b, 422c, 422d, 422h, 422i, 422j y 422k espectrales no formarán parte del conjunto R porque dichas líneas espectrales están ubicadas demasiado cerca al lado de las líneas 420a, 420b, 420c y 420d espectrales. En cambio, los índices de frecuencia de línea espectral de las líneas 422e, 422f, 422g espectrales estarán incluidos en el conjunto R, porque dichas líneas espectrales están en sí mismas cuantificadas a cero y están separadas suficientemente de cualquier línea espectral adyacente distinta de cero.

20 El algoritmo 300 también comprende una segunda parte 320 de decodificación del suelo de ruido, en el que un índice de valor de ruido ("index" en la parte 320 de código de programa) se convierte en un valor de cifra de ruido decodificado ("nf_decoded" en el programa 300).

25 El código 300 de programa también comprende una tercera parte 330 de relleno de las líneas espectrales identificadas, es decir líneas espectrales cuyos índices de frecuencia i de línea espectral están en el conjunto R, con ruido. Con este propósito, los valores espectrales de las líneas espectrales identificadas (designados por ejemplo, con $x(i)$, donde la variable continua i adopta de manera subsiguiente todos los índices de frecuencia de línea espectral incluidos en el conjunto R) se fijan a valores de relleno de ruido. Los valores de relleno de ruido se obtienen, por ejemplo, multiplicando el valor de relleno de ruido decodificado ("nf_decoded") por un número aleatorio o un número pseudoaleatorio (designado con "random(-1,+1)"), donde el número aleatorio o pseudoaleatorio puede adoptar, por ejemplo, aleatoriamente o seudoaleatoriamente los números -1 y $+1$. Sin embargo, naturalmente es posible una provisión diferente de un ruido aleatorio o pseudoaleatorio.

30 El relleno de ruido también se ilustra en la figura 4. Como se puede ver en la figura 4, los valores espectrales cero de las líneas 422e, 422f y 422g espectrales se reemplazan por valores de relleno de ruido (ilustrados por líneas discontinuas en la figura 4).

35 Calculador de parámetro de relleno de ruido según las figuras 5 y 6

La figura 5 muestra un diagrama de bloques esquemático de un calculador 500 de parámetro de relleno de ruido. El calculador de parámetro de relleno de ruido está configurado para obtener una representación 510 espectral cuantificada de una señal de audio y para proporcionar, basándose en ello, un parámetro 512 de relleno de ruido. El calculador 500 de parámetro de relleno de ruido comprende un identificador 520 de región espectral, que está configurado para recibir la representación 510 espectral cuantificada de la señal de audio y para identificar regiones espectrales (por ejemplo, líneas espectrales) de la representación 510 espectral cuantificada separadas de regiones espectrales distintas de cero de la representación 510 espectral cuantificada por al menos una región espectral intermedia (por ejemplo línea espectral), para obtener una información 522 que describe regiones espectrales identificadas (por ejemplo líneas espectrales identificadas). El calculador 500 de parámetro de relleno de ruido comprende además un calculador 530 de valor de ruido configurado para recibir una información 532 de error de cuantificación y para proporcionar el parámetro 512 de relleno de ruido. Con este propósito, el calculador de valor de ruido está configurado para considerar selectivamente errores de cuantificación de las regiones espectrales identificadas, descritas mediante la información 522, para un cálculo del parámetro 512 de relleno de ruido.

50 La información 532 de error de cuantificación puede por ejemplo ser idéntica a una información de energía (o información de intensidad) que describe energías (o intensidades) de aquellas líneas espectrales que están cuantificadas a cero en la representación 510 espectral cuantificada.

55 El calculador 500 de parámetro de relleno de ruido puede comprender opcionalmente un cuantificador 540, que está configurado para recibir una representación 542 espectral no cuantificada de una señal de audio y para proporcionar la representación 510 espectral cuantificada de la señal de audio. El cuantificador 540 puede tener una resolución de cuantificación ajustable, que puede ser, por ejemplo, individualmente ajustable por línea espectral, o por banda espectral (por ejemplo, dependiendo de una relevancia psicoacústica de las líneas espectrales o bandas espectrales, obtenida usando un modelo psicoacústico). La funcionalidad del cuantificador de resolución variable puede ser igual a la funcionalidad descrita en las normas internacionales ISO/IEC 13818-7 e ISO/IEC 14496-3. En particular, el cuantificador 540 puede ajustarse de manera que haya vacíos espectrales o huecos espectrales en la representación 510 espectral

cuantificada de la señal de audio, por ejemplo, regiones contiguas de líneas espectrales adyacentes cuantificadas a cero.

Asimismo, la representación 542 espectral no cuantificada puede servir como la información 532 de error de cuantificación, o la información 532 de error de cuantificación puede derivarse a partir de la representación 542 espectral no cuantificada.

En lo que sigue se describirá en detalle la funcionalidad del cálculo del parámetro de relleno de ruido, que puede realizarse mediante el calculador 500 de parámetro de relleno de ruido. En el cálculo del parámetro de relleno de ruido en el lado del codificador, el relleno de ruido preferiblemente se aplica en el dominio de cuantificación. De esta manera, el ruido introducido se conforma más tarde mediante el filtro inverso relevante psicoacústico. La energía del ruido introducido por el decodificador se calcula y se codifica en el lado del codificador siguiendo las siguientes etapas:

1. obtener los valores cuantificados de las líneas de frecuencia;
2. seleccionar sólo una parte de los espectros;
3. detectar las regiones espectrales en la parte seleccionada de los espectros donde una longitud de recorrido de ceros es mayor que un tamaño de longitud de recorrido mínimo;
4. calcular la media geométrica del error de cuantificación sobre las regiones previamente detectadas; y
5. cuantificar uniformemente la media geométrica con 3 bits.

Con relación a la primera etapa, los valores cuantificados de las líneas de frecuencia pueden obtenerse usando el cuantificador 540. Los valores cuantificados de las líneas de frecuencia por lo tanto están representados por la representación 510 espectral cuantificada.

Con respecto a la segunda etapa, que puede considerarse opcional, ha de indicarse que el cálculo del relleno de ruido preferiblemente se realiza sobre la base de una parte de alta frecuencia de los espectros. En una realización preferida, la energía del ruido (llamada suelo de ruido) se calcula sólo en la segunda mitad de los espectros, es decir, para las altas frecuencias (pero no para las frecuencias inferiores). En realidad, usualmente las altas frecuencias (parte superior del espectro) son menos importantes desde el punto de vista perceptivo que las bajas frecuencias, y los valores cuantificados a cero aparecen mayormente en la segunda mitad de los espectros. Asimismo, agregar el ruido en las altas frecuencias es menos propenso a obtener una restitución de sonido final ruidosa.

Con respecto a la tercera etapa, restringiendo el relleno de ruido en las regiones espectrales donde ocurre una longitud de recorrido de valores cuantificados a cero, se evita que el relleno de ruido afecte demasiado a los valores distintos de cero. De esta manera, el relleno de ruido no se aplica en la proximidad de los valores distintos de cero, y la tonalidad original de estas líneas entonces se preserva mejor. El tamaño de longitud de recorrido mínimo se fija en 8 en una realización preferida. Esto significa que las 8 líneas que rodean a un valor distinto de cero no se ven afectadas por el relleno de ruido (y consecuentemente no se consideran para el cálculo de un valor de ruido).

Con respecto a la cuarta etapa, el error de cuantificación en el dominio cuantificado está en el rango $[-0,5; 0,5]$, y se supone que está uniformemente distribuido. La energía de errores de cuantificación de las regiones detectadas se promedia en el dominio logarítmico (es decir, media geométrica). El suelo de ruido, nf , se calcula entonces como sigue:

$$nf = \text{power}(10, \text{sum}(\log_{10}(E(x(i))))/(2*n))$$

En lo anterior, $\text{sum}()$ es la suma de las energías logarítmicas, $\log_{10}(E())$, de las líneas individuales $x(i)$ dentro de las regiones detectadas, y n el número de líneas dentro de estas regiones. El suelo de ruido, nf , está entre 0 y 0,5. Tal cálculo permite tener en cuenta la planeidad espectral original de los valores cero, y luego obtener información acerca de sus características de tonalidad/ruido.

Si los valores cero son muy tonales, el suelo de ruido (calculado en el aparato 500) tenderá a cero, y se agregará un suelo de ruido bajo al decodificador (por ejemplo, al decodificador 100, 200 descrito arriba). Si los valores cero son realmente ruidosos, el suelo de ruido será alto, y el relleno de ruido puede verse como una codificación altamente paramétrica de las líneas espectrales cero, tal como PNS (sustitución de ruido perceptivo) (véase también referencia [4]).

Con respecto a la quinta etapa, el índice de cuantificación ("index") del suelo de ruido, se calcula entonces como sigue:

$$\text{index} = \max(0, \min(7, \text{int}(8 - 16 * nf)))$$

El índice se transmite, por ejemplo, en 3 bits.

En lo que sigue se describirá el algoritmo para calcular el parámetro de relleno de ruido tomando como referencia la figura 6, que muestra un pseudocódigo 600 de programa de un algoritmo tal para obtener el parámetro de relleno de ruido, según una realización de la invención. El algoritmo 600 comprende una primera parte 610 de detección de

5 regiones que deben ser consideradas para calcular el parámetro de relleno de ruido. Las regiones identificadas (por ejemplo, líneas espectrales) se describen por el conjunto R, que por ejemplo puede comprender índices de frecuencia de línea espectral ("line index") de líneas espectrales identificadas. Pueden identificarse líneas espectrales que en sí mismas están cuantificadas a cero y que además están separadas suficientemente de cualquier otra línea espectral cuantificada a un valor distinto de cero.

10 La primera parte 610 del programa 600 puede ser idéntica a la primera parte 310 del programa 300. Por consiguiente, la representación espectral cuantificada ("quantized $x(i)$ ") usada en el algoritmo 600 puede ser, por ejemplo, idéntica a la representación espectral cuantificada ("quantized $x(i)$ ") usada en el algoritmo 300 en el lado del decodificador. En otras palabras, la representación espectral cuantificada usada en el lado del codificador puede transmitirse, en una forma codificada, al decodificador en un sistema de transmisión que comprende un codificador y un decodificador.

15 El algoritmo 600 comprende una segunda parte 620 de cálculo del suelo de ruido. En el cálculo del suelo de ruido, sólo se consideran aquellas regiones espectrales (o líneas espectrales) descritas por el conjunto R calculado en la primera parte 610 del algoritmo 600. Como puede verse, el valor de relleno de ruido n_f se inicializa primero a cero. El número de líneas espectrales (n) consideradas se inicializa primero a cero. A continuación, las energías de todas las líneas espectrales, cuyos índices de línea están incluidos en el conjunto R, se suman, habiéndose logaritmizado las energías de las líneas espectrales antes de la suma. Por ejemplo, se puede sumar un logaritmo de base 10 (\log_{10}) de las energías ($E(x(i))$) de las líneas espectrales. Ha de indicarse aquí que la energía real de las líneas espectrales antes de la cuantificación (designada como "E o energy $x(i)$ ") se suma en forma logaritmizada. También se cuenta el número de líneas espectrales consideradas. Así, después de la ejecución de la segunda parte 620 del algoritmo 600, la variable n_f indica una suma logarítmica de energías de las líneas espectrales identificadas antes de la cuantificación, y la variable n describe el número de líneas espectrales identificadas.

20

El algoritmo 600 también comprende una tercera parte 630 de cuantificación del valor n_f , es decir, la suma logarítmica de las líneas espectrales identificadas. Se puede usar una ecuación de correlación tal como se describió anteriormente o tal como se muestra en la figura 6.

25 Método según la figura 7

La figura 7 muestra un diagrama de flujo de un método para proporcionar una representación espectral rellena con ruido de una señal de audio sobre la base de una representación espectral de entrada de la señal de audio. El método 700 de la figura 7 comprende una etapa 710 de identificar regiones espectrales de una representación espectral de entrada de una señal de audio separadas de las regiones espectrales de valores distintos de cero de la representación espectral de entrada por al menos una región espectral intermedia, para obtener regiones espectrales identificadas. El método 700 también comprende una etapa 720 de introducir selectivamente ruido dentro de las regiones identificadas, para obtener una representación espectral rellena con ruido de la señal de audio.

30

El método 700 puede complementarse por cualquiera de las características y funcionalidades descritas en el presente documento con referencia al relleno de ruido de la invención.

35 Método según la figura 8

La figura 8 muestra un diagrama de flujo de un método para proporcionar un parámetro de relleno de ruido sobre la base de una representación espectral cuantificada de una señal de audio. El método 800 comprende una etapa 810 de identificar regiones espectrales de la representación espectral cuantificada de una señal de audio separadas de las regiones espectrales de valores distintos de cero de la representación espectral cuantificada por al menos una región espectral intermedia, para obtener regiones espectrales identificadas. El método 800 también comprende una etapa 820 de considerar selectivamente los errores de cuantificación de las regiones identificadas para el cálculo del parámetro de relleno.

40

El método 800 puede complementarse por cualquiera de las características y funcionalidades descritas en el presente documento con respecto al calculador de parámetro de relleno de ruido.

45 Representación de señal de audio según la figura 9

La figura 9 muestra una representación gráfica de una representación de señal de audio según una realización de la invención. La representación 900 de señal de audio puede formar la base, por ejemplo, para la representación 110 espectral de entrada. La representación 900 de señal de audio también puede hacerse cargo de la funcionalidad de la representación 212 de señal de audio codificada. La representación 900 de señal de audio puede obtenerse usando el calculador 500 de parámetro de relleno de ruido, pudiendo comprender la representación 900 de señal de audio, por ejemplo, la representación 510 espectral cuantificada de la señal de audio y el parámetro 512 de relleno de ruido, por ejemplo, ambos en forma codificada.

50

En otras palabras, la representación 900 de señal de audio codificada puede representar una señal de audio. La representación 900 de señal de audio codificada comprende una representación en el dominio espectral cuantificada codificada de la señal de audio y también un parámetro de relleno de ruido codificado. El parámetro de relleno de ruido representa un error de cuantificación de regiones espectrales de la representación en el dominio espectral cuantificadas

55

a cero y separadas de regiones espectrales de la representación en el dominio espectral cuantificadas a valores distinto de ceros, por al menos una región espectral intermedia.

Naturalmente, la representación 900 de señal de audio puede complementarse por cualquier información descrita anteriormente.

5 Alternativas de implementación

Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, con señales de control legibles electrónicamente almacenadas en los mismos, que actúan conjuntamente (o pueden actuar conjuntamente) con un sistema informático programable de tal manera que se realice el respectivo método.

Algunas realizaciones según la invención comprenden un soporte de datos con señales de control legibles electrónicamente, que pueden actuar conjuntamente con un sistema informático programable, de tal manera que se realice uno de los métodos descritos en el presente documento.

Generalmente, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, siendo el código de programa operativo para realizar uno de los métodos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede estar almacenado, por ejemplo, en un soporte legible por máquina.

Otras realizaciones comprenden el programa informático para realizar uno de los métodos descritos en el presente documento, almacenado en un soporte legible por máquina.

En otras palabras, una realización del método de la invención es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los métodos descritos en el presente documento, cuando el programa informático se ejecuta en un ordenador.

Una realización adicional de los métodos de la invención es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital) que comprende el programa informático para realizar uno de los métodos descritos en el presente documento.

Una realización adicional del método de la invención es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para realizar uno de los métodos descritos en el presente documento. El flujo de datos o la secuencia de señales pueden configurarse, por ejemplo, para transferirse a través de una conexión de comunicación de datos, por ejemplo, a través de Internet.

Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado para o adaptado para realizar uno de los métodos descritos en el presente documento.

Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para realizar uno de los métodos descritos en el presente documento.

En algunas realizaciones se puede usar un dispositivo lógico programable (por ejemplo una disposición de puertas programables en campo) para realizar algunas o todas las funcionalidades de los métodos descritos en el presente documento. En algunas realizaciones, una disposición de puertas programables en campo puede actuar conjuntamente con un microprocesador para realizar uno de los métodos descritos en el presente documento.

Conclusión

Para resumir lo anterior, la presente invención mejora la herramienta de codificación de audio "relleno de ruido" considerando la señal de audio y las características de señal decodificada tanto cuando se calculan parámetros de relleno de ruido en el lado del codificador, como cuando se aplica el ruido en el lado del decodificador. En una realización de la invención, se estima la tonalidad/ruido de las líneas espectrales cuantificadas a cero y se usa para la estimación del suelo de ruido. Este suelo de ruido se transmite entonces al decodificador que aplica el relleno de ruido a los valores cuantificados a cero que aparecen en regiones específicas de los espectros. Estas regiones se seleccionan basándose en las características de los espectros decodificados.

Con relación al contexto de la invención, puede observarse que la invención se ha aplicado a una codificación basada en transformación que usa una cuantificación escalar sobre MDCT. Los coeficientes de MDCT están previamente normalizados mediante una curva calculada basándose en indicadores perceptivos. La curva se deduce a partir de una fase de análisis de LPC (codificación de predicción lineal) previa ponderando los coeficientes de LPC, tal como se hace en el modo TCX de AMR-WB+ (véase la referencia [1]). A partir de los coeficientes ponderados, se diseña un filtro de ponderación perceptivo y se aplica antes de la MDCT. El filtro inverso también se aplica en el lado del decodificador después de la MDCT inversa. Este filtro de ponderación perceptiva inversa conforma los ruidos de cuantificación de una manera que minimiza o enmascara el ruido percibido.

5 En las realizaciones según la invención se superan las desventajas de la técnica anterior. El relleno de ruido se aplica convencionalmente de una manera sistemática sobre los valores cuantificados a cero considerando sólo un umbral basado en la envolvente espectral, un umbral de enmascaramiento, o un umbral de energía. La técnica anterior no considera ni las características de la señal de entrada ni las características de la señal decodificada. Así, un aparato convencional puede introducir artefactos adicionales indeseables, especialmente artefactos de ruido, y cancela las ventajas de una herramienta de este tipo.

En cambio, las realizaciones de la invención permiten un relleno de ruido mejorado con artefactos reducidos, tal como se comentó anteriormente.

Referencias:

- 10 [1] "Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec", 3GPP TS 26.290 V6.3.0, 2005-06, Technical Specification
- [2] Ragot *et al.*, "ITU-T G.729.1: AN 8-32 Kbit/S Scalable Coder Interoperable with G.729 for Wideband Telephony and Voice Over IP", Vol. 4, ICASSP 07, 15-20 abril de 2007
- 15 [3] "AUDIO CODING", solicitud internacional n.º: PCT/IB2002/001388, Solicitante: KONINKLIJKE PHILIPS ELECTRONICS N.V. [NL/NL]; Groenewoudseweg 1 NL-5621 BA Eindhoven (NL). Inventores: TAORI, Rakesh; Prof Holstlaan 6 NL-5656 AA Eindhoven (NL) y VAN DE PAR, Steven, L., J., D., E.; Prof. Holstlaan 6 NL-5656 AA Eindhoven (NL).
- [4] Generic Coding of Moving Pictures and Associated Audio: Advanced Audio Coding. International Standard 13818-7, ISO/IEC JTC1/SC29/WG11 Moving Pictures Expert Group, 1997.

REIVINDICACIONES

1. Rellenador (100) de ruido para proporcionar una representación (112) espectral rellenada con ruido de una señal de audio sobre la base de una representación (110) espectral de entrada de la señal de audio, comprendiendo el relleno de ruido:
 - 5 un identificador (120) de región espectral configurado para identificar regiones (422e, 422f, 422g) espectrales de la representación (110) espectral de entrada cuantificadas a cero y separadas de regiones (420a, 420b, 420c, 420d) espectrales distintas de cero de la representación (110) espectral de entrada por al menos una región (422a, 422b, 422c, 422d, 422h, 422i, 422j, 422k) espectral intermedia, para obtener regiones (422e, 422f, 422g) espectrales identificadas; y
 - 10 un insertador (130) de ruido configurado para introducir selectivamente ruido dentro de las regiones (422e, 422f, 422g) espectrales identificadas para obtener la representación (112) espectral rellenada con ruido de la señal de audio.
2. Rellenador (100) de ruido según la reivindicación 1, en el que el identificador (120) de región espectral está configurado para identificar, como regiones espectrales identificadas, líneas (422e, 422f, 422g) espectrales de la representación (110) espectral de entrada, que están cuantificadas a cero y que comprenden al menos un primer número predeterminado de líneas (422a, 422b, 422c, 422d; 422b, 422c, 422d, 422e, 422c, 422d, 422e, 422f)
 - 15 espectrales vecinas de frecuencia inferior cuantificadas a cero y al menos un segundo número predeterminado (4) de líneas (422f, 422g, 422h, 422i; 422g, 422h, 422i, 422j; 422h, 422i, 422j, 422k) espectrales vecinas de frecuencia superior cuantificadas a cero, como regiones espectrales identificadas;
 - en el que el primer número predeterminado es mayor o igual a 1, y en el que el segundo número predeterminado es mayor o igual a 1; y
 - 20 en el que el insertador (130) de ruido está configurado para introducir selectivamente ruido dentro de las líneas (422e, 422f, 422g) espectrales identificadas mientras que deja sin afectar por el relleno con ruido las líneas (420a, 420b, 420c, 420d) espectrales cuantificadas a un valor distinto de cero y las líneas (422a, 422b, 422c, 422d, 422h, 422i, 422j, 422k) espectrales cuantificadas a cero, pero que no tienen el primer número predeterminado de líneas espectrales vecinas de frecuencia inferior cuantificadas a cero, o el segundo número predeterminado de líneas espectrales vecinas de frecuencia superior cuantificadas a cero.
3. Rellenador (100) de ruido según la reivindicación 2, en el que el primer número predeterminado es igual al segundo número predeterminado.
4. Rellenador (100) de ruido según una de las reivindicaciones 1 a 3, en el que el relleno de ruido está configurado para introducir ruido sólo dentro de regiones espectrales en una parte superior de la representación (110)
 - 30 espectral de entrada de la señal de audio mientras que deja una parte inferior de la representación (110) espectral de entrada de la señal de audio sin afectar por el relleno con ruido.
5. Rellenador (100) de ruido según una de las reivindicaciones 1 a 4, en el que el identificador (120) de región espectral está configurado para sumar valores de intensidad cuantificados (quantized (x(i))) de regiones espectrales en la proximidad espectral a ambos lados predeterminada de una región espectral (i) dada, para obtener un valor suma (E),
 - 35 y para evaluar el valor suma (E) para decidir si la región espectral (i) dada es una región espectral identificada o no.
6. Rellenador (100) de ruido según una de las reivindicaciones 1 a 5, en el que el identificador (120) de región espectral está configurado para explorar un rango de regiones espectrales de la representación (110) espectral de entrada para detectar secuencias (422a a 422i; 422b a 422j; 422c a 422k) contiguas de regiones espectrales cuantificadas a cero, y para reconocer una o más regiones (422e, 422f, 422g) espectrales centrales de las secuencias
 - 40 contiguas detectadas como regiones espectrales identificadas.
7. Calculador (500) de parámetro de relleno de ruido para proporcionar un parámetro (512) de relleno de ruido sobre la base de una representación (510) espectral cuantificada de una señal de audio, comprendiendo el calculador de parámetro de relleno de ruido:
 - 45 un identificador (520) de región espectral configurado para identificar regiones (422e, 422f, 422g) espectrales de la representación (510) espectral cuantificada cuantificadas a cero y separadas de las regiones (420a, 420b, 420c, 420d) espectrales distintas de cero de la representación (510) espectral cuantificada por al menos una región (422a, 422b, 422c, 422d, 422h, 422i, 422j, 422k) espectral intermedia, para obtener regiones (422e, 422f, 422g) espectrales identificadas; y
 - 50 un calculador (530) de valor de ruido configurado para considerar selectivamente errores de cuantificación ((energy (x(i))) de las regiones espectrales identificadas (i) para un cálculo del parámetro (512, nf) de relleno de ruido.
8. Calculador (500) de parámetro de relleno de ruido según la reivindicación 7,
 - en el que el identificador (520) de región espectral está configurado para identificar, como regiones espectrales identificadas, líneas (422e, 422f, 422g) espectrales de la representación (510) espectral de entrada, que están

- cuantificadas a cero y que comprenden al menos un primer número predeterminado de líneas (422a, 422b, 422c, 422d; 422b, 422c, 422d, 422e, 422c, 422d, 422e, 422f) espectrales vecinas de frecuencia inferior cuantificadas a cero y al menos un segundo número predeterminado de líneas (422f, 422g, 422h, 422i; 422g, 422h, 422i, 422j; 422h, 422i, 422j, 422k) espectrales vecinas de frecuencia superior cuantificadas a cero, como regiones espectrales identificadas;
- 5 en el que el primer número predeterminado es mayor o igual a 1, y en el que el segundo número predeterminado es mayor o igual a 1; y
- en el que el calculador (520) de valor de ruido está configurado para considerar selectivamente errores de cuantificación de las regiones espectrales identificadas (i) para un cálculo del parámetro de relleno de ruido mientras que deja las líneas (420a, 420b, 420c, 420d) espectrales cuantificadas a un valor distinto de cero y las líneas (422a, 422b, 422c, 10 422d, 422h, 422i, 422j, 422k) espectrales cuantificadas a cero, pero que no tienen el primer número predeterminado (4) de líneas espectrales vecinas de frecuencia inferior cuantificadas a cero, o el segundo número predeterminado de líneas espectrales vecinas de frecuencia superior cuantificadas a cero, fuera de consideración para el cálculo del parámetro de relleno de ruido.
9. Calculador (500) de parámetro de relleno de ruido según una de las reivindicaciones 7 a 8, en el que el 15 calculador (530) de valor de ruido está configurado para considerar energías real (energy(x(i))) de los errores de cuantificación de las regiones espectrales identificadas (i) para el cálculo del parámetro de relleno de ruido (512, nf, nf_index).
10. Calculador (500) de parámetro de relleno de ruido según una de las reivindicaciones 7 a 9, en el que el 20 calculador (530) de valor de ruido está configurado para enfatizar una energía de error de cuantificación no tonal (energy(x(i))) distribuida sobre una pluralidad de regiones espectrales identificadas en relación a una energía de error de cuantificación tonal concentrada en una sola región espectral o en una pluralidad de líneas espectrales contiguas.
11. Calculador (500) de parámetro de relleno de ruido según una de las reivindicaciones 7 a 10, en el que el 25 calculador (530) de valor de ruido está configurado para calcular una suma de energías de error de cuantificación logaritmizadas (log10(energy(x(i)))) de las regiones espectrales identificadas (i), para obtener el parámetro de relleno de ruido (512, nf, nf_index).
12. Representación (900) de señal de audio codificada que representa una señal de audio, comprendiendo la 30 representación de señal de audio codificada:
una representación en el dominio espectral cuantificada codificada de la señal de audio; y
un parámetro de relleno de ruido codificado;
- 30 en la que el parámetro de relleno de ruido representa un error de cuantificación de regiones espectrales de la representación en el dominio espectral cuantificada a cero y separada de regiones espectrales de la representación en el dominio espectral cuantificada a un valor distinto de cero, por al menos una región espectral intermedia.
13. Método (700) para proporcionar una representación espectral rellenada con ruido de una señal de audio sobre 35 la base de una representación espectral de entrada de la señal de audio, comprendiendo el método:
identificar (710) regiones espectrales de la representación espectral de entrada cuantificadas a cero y separadas de regiones espectrales distintas de cero de la representación espectral de entrada por al menos una región espectral intermedia, para obtener regiones espectrales identificadas; e
introducir selectivamente (720) ruido dentro de las regiones espectrales identificadas para obtener la representación espectral rellenada con ruido de la señal de audio.
- 40 14. Método (800) para proporcionar un parámetro de relleno de ruido sobre la base de una representación espectral cuantificada de una señal de audio, comprendiendo el método:
identificar (810) regiones espectrales de la representación espectral cuantificada cuantificadas a cero y separadas de regiones espectrales distintas de cero de la representación espectral cuantificada por al menos una región espectral intermedia para obtener regiones espectrales identificadas; y
- 45 considerar selectivamente (820) errores de cuantificación de las regiones espectrales identificadas para un cálculo del parámetro de relleno de ruido.
15. Programa informático para realizar el método según la reivindicación 13 ó 14, cuando el programa informático se ejecuta en un ordenador.

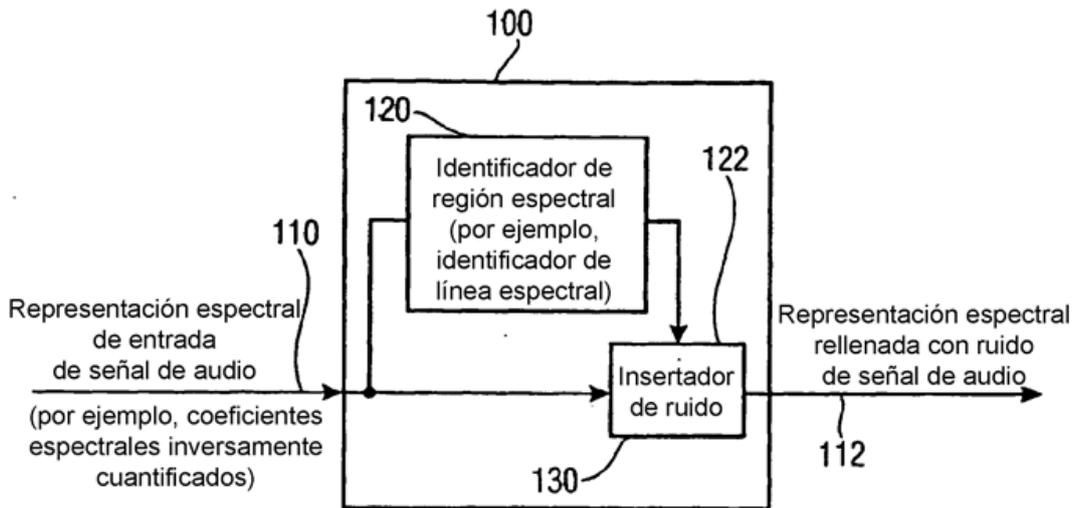
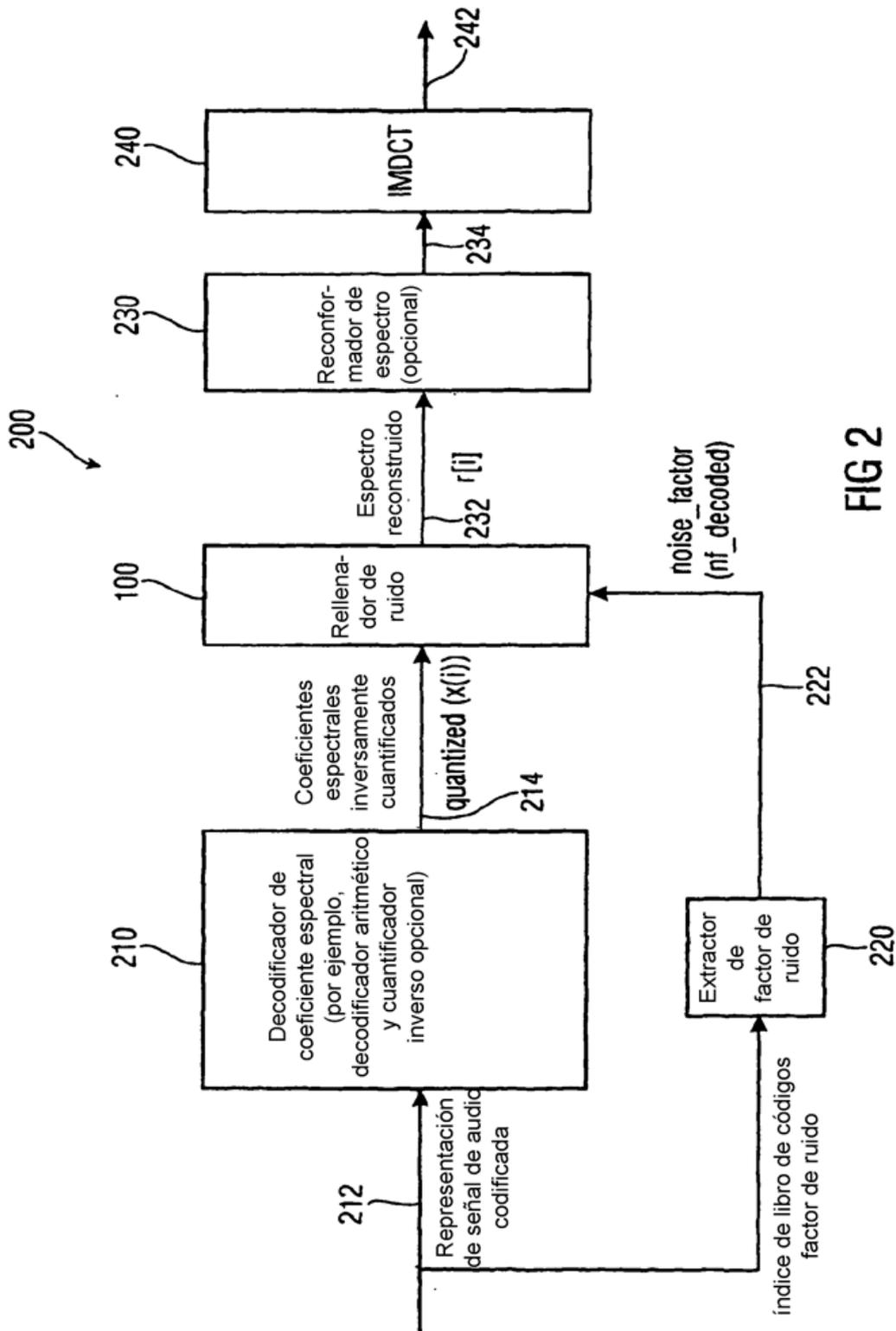


FIG 1



300



```

% Pseudocódigo relleno de ruido lado del decodificador
% Detectar regiones
R = {};
MinimalRunLength = 8;
Para (todas las líneas en la segunda mitad de los espectros)
    E=0

    Para (i de line index - (MinimalRunLength)/2 a line
index + (MinimalRunLength)/2)
        E=E+quantized(x(i));
    Fin para
    Si E==0
        R = {R, line index};
    Fin si
Fin para

310 {
% Decodificar suelo de ruido
nf_decoded = 0.0625*(8-index);

330 {
% Rellenar con ruido
Para (todas las i en R)
    x(i) = random (-1,+1)*nf_decoded;
Fin para

% Fin pseudocódigo relleno de ruido lado del docodificador

```

FIG 3

500

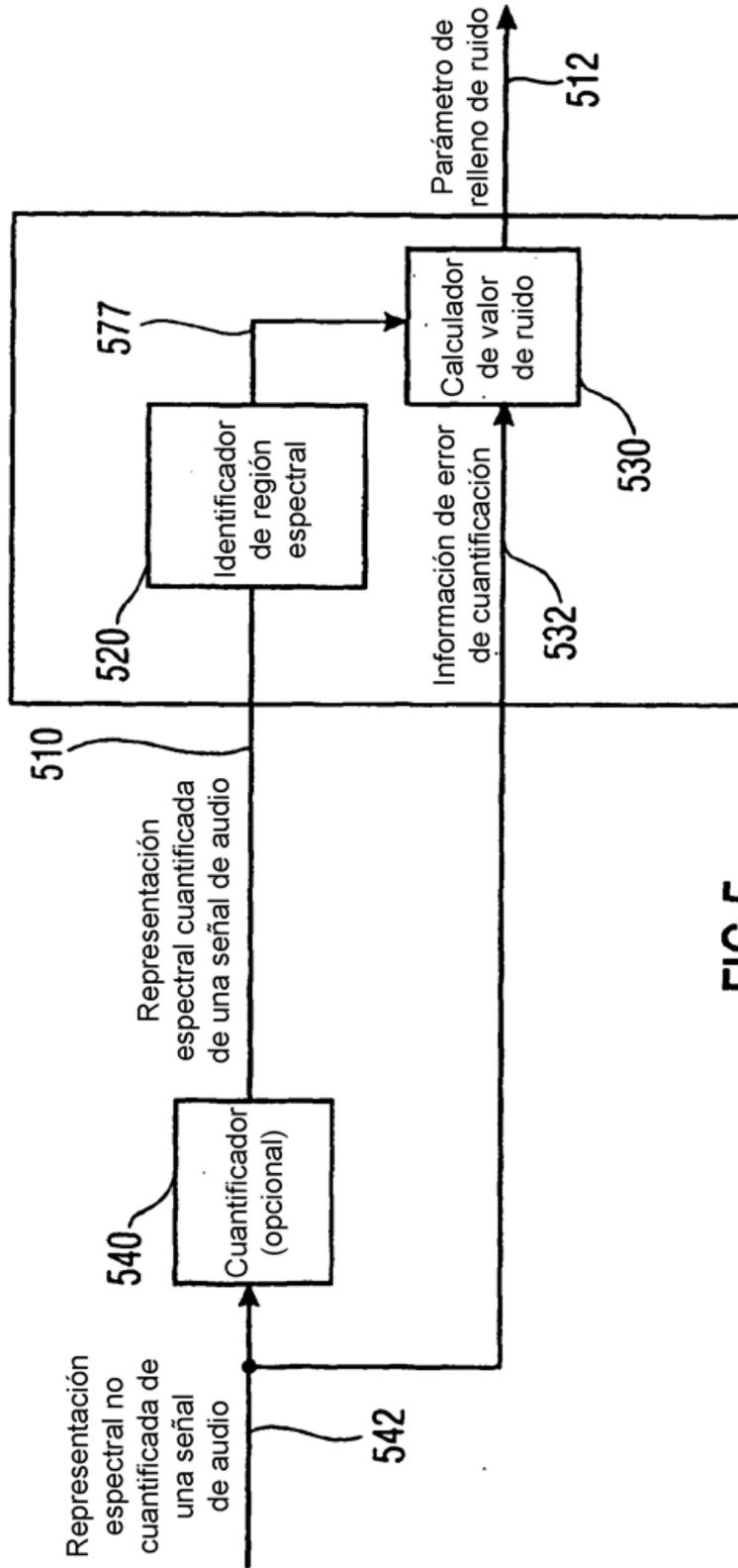


FIG 5

600
↙

```

% Pseudocódigo relleno de ruido lado del codificador

% Detectar regiones
R = {};
MinimalRunLength = 8;
Para (todas las líneas en la segunda mitad de los espectros)
  E=0
  Para(i de line index - (MinimalRunLength)/2 a line
index + (MinimalRunLength)/2)
    E=E+quantized(x(i));
  Fin para
  Si E==0
    R = {R, line index};
  Fin si
Fin para

%Calcular suelo de ruido
nf = 0;
n=0;
Para (todas las i en R)
  nf = E+log10(energy(x(i)));
  n = n+1;
Fin para

% Cuantif. nf
index=max(0, min (7, int (8-16*power (10, (nf/ (2*n))))));

% Fin del pseudocódigo relleno de ruido lado del codificador

```

610 {

620 {

630 {

FIG 6

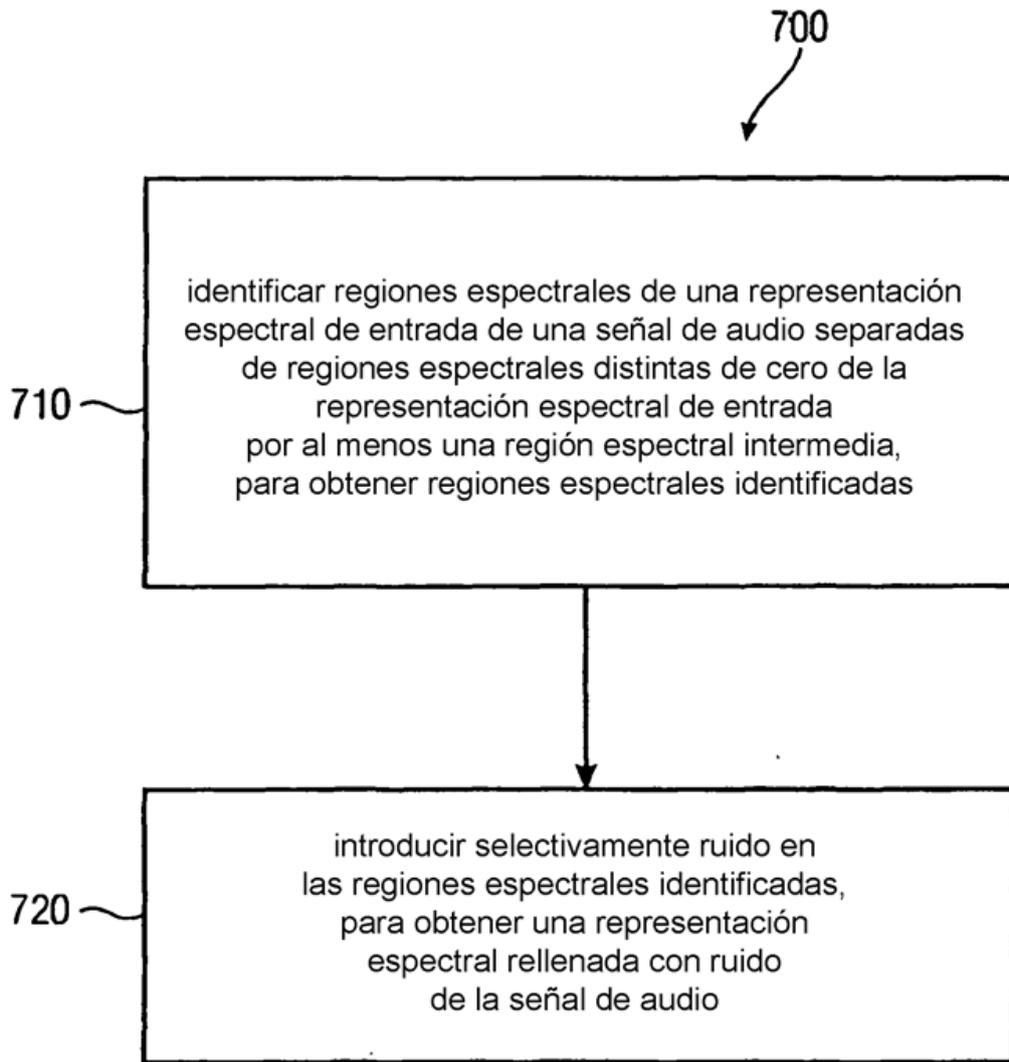


FIG 7

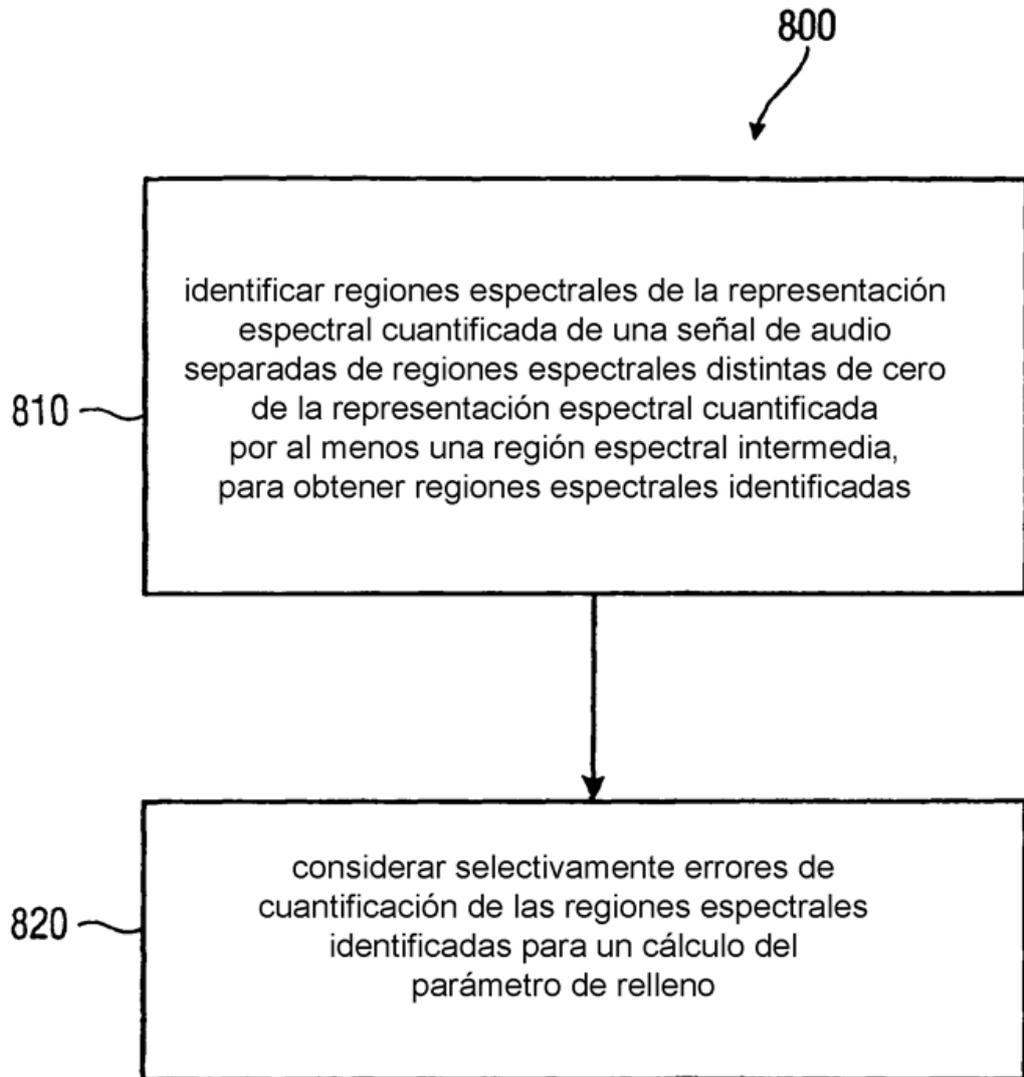


FIG 8

900
↙

Representación de señal de audio:

{

parámetro de relleno de ruido;

codificar representación en el dominio espectral

}

FIG 9