



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS
ESPAÑA



⑪ Número de publicación: **2 992 880**

⑮ Int. Cl.:

G10L 19/028 (2013.01)

G10L 21/038 (2013.01)

⑫

TRADUCCIÓN DE PATENTE EUROPEA

T3

⑯ Fecha de presentación y número de la solicitud europea: **24.07.2015 E 18209545 (5)**

⑯ Fecha y número de publicación de la concesión europea: **09.10.2024 EP 3471094**

⑭ Título: **Aparato y procedimiento para generar una señal de audio mejorada mediante el uso de relleno de ruido independiente**

⑯ Prioridad:

28.07.2014 EP 14178777

⑯ Fecha de publicación y mención en BOPI de la traducción de la patente:
19.12.2024

⑯ Titular/es:

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

⑯ Inventor/es:

**DISCH, SASCHA;
GEIGER, RALF;
NIEDERMEIER, ANDREAS;
NEUSINGER, MATTHIAS;
SCHMIDT, KONSTANTIN;
WILDE, STEPHAN;
SCHUBERT, BENJAMIN y
NEUKAM, CHRISTIAN**

⑯ Agente/Representante:

PONTI & PARTNERS, S.L.P.

ES 2 992 880 T3

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

DESCRIPCIÓN

Aparato y procedimiento para generar una señal de audio mejorada mediante el uso de relleno de ruido independiente

5

Especificación

[0001] La solicitud se refiere al procesamiento de señales y, en particular, al procesamiento de señales de audio.

10

[0002] La codificación perceptual de señales de audio con fines de reducción de datos para el almacenamiento eficaz o la transmisión de estas señales es una práctica ampliamente utilizada. En particular, cuando se tienen que obtener tasas de bits mínimas, la codificación empleada lleva a una reducción de la calidad de audio que con frecuencia es causada primariamente por una limitación en el lado del codificador del ancho de banda de la señal de audio que se va a transmitir. En códecs contemporáneos, existen procedimientos bien conocidos para la restauración de las señales del lado del decodificador mediante extensión de ancho de banda de señales de audio (BWE), por ejemplo, replicación de banda espectral (SBR).

[0003] En una codificación de baja tasa de bits, se emplea con frecuencia también el denominado relleno de ruido. Prominentes regiones espectrales que se han cuantificado a cero debido a estrictas restricciones de tasas de bits se rellenan con ruido sintético en el decodificador.

[0004] Usualmente, ambas técnicas se combinan en aplicaciones de codificación de bajas tasas de bits. Además, existen soluciones integradas tales como relleno inteligente de intervalos (*Intelligent Gap Filling*, IGF) que combinan codificación de audio, relleno de ruido y relleno de intervalo espectral.

[0005] Sin embargo, todos estos procedimientos tienen en común que, en una primera etapa, la banda de base o señal de audio central se reconstruye mediante el uso de decodificación en forma de onda y relleno de ruido y, en una segunda etapa, el procesamiento de BWE o IGF se lleva a cabo mediante el uso de la señal fácilmente reconstruida. Esto lleva al hecho de que los mismos valores de ruido que fueron llenados en la banda de base por relleno de ruido durante la reconstrucción se usan para regenerar las partes que faltan en la banda alta (en BWE) o para el relleno de los intervalos espectrales restantes (en IGF). Mediante el uso de ruido altamente correlacionado para reconstruir múltiples regiones espectrales en BWE o IGF puede llevar a alteraciones perceptuales.

35 [0006] Puntos relevantes en el estado de la técnica comprenden

- SBR como un post-procesador a decodificación en forma de onda [1-3]
- AAC PNS [4]
- relleno de ruido MPEG-D USAC [5]
- 40 • G.719 y G.722.1C [6]
- MPEG-H 3D IGF [8]

[0007] Los siguientes documentos y solicitudes de patente describen procedimientos que se consideran relevantes para la solicitud:

45

- [1] M. Dietz, L. Liljeryd, K. Kjörling y O. Kunz, "Spectral Band Replication, a novel approach in audio coding", en 112th AES Convention, Múnich, Alemania, 2002.
- [2] S. Meltzer, R. Böhm y F. Henn, "SBR enhanced audio codecs for digital broadcasting such as "Digital Radio Mondiale" (DRM)", en 112th AES Convention, Múnich, Alemania, 2002.
- 50 [3] T. Ziegler, A. Ehret, P. Ekstrand y M. Lutzky, "Enhancing mp3 with SBR: Features and Capabilities of the new mp3PRO Algorithm", en 112th AES Convention, Múnich, Alemania, 2002.
- [4] J. Herre, D. Schulz, Extending the MPEG-4 AAC Codec by Perceptual Noise Substitution, Audio Engineering Society 104th Convention, Preprint 4720, Ámsterdam, Holanda, 1998.
- [5] Solicitud de patente europea EP2304720 USAC relleno de ruido
- 55 [6] ITU-T Recomendaciones G.719 y G.221C
- [7] EP 2704142
- [8] EP 13177350

[0008] Las señales de audio procesadas con estos procedimientos sufren de artefactos tales como asperezas, distorsiones de modulación y un timbre percibido como desagradable, en particular a una baja tasa de bits y, en consecuencia, bajo ancho de banda y/o la aparición de agujeros espectrales en el intervalo LF. La razón de ello es, como se explica más abajo, primariamente el hecho de que los componentes reconstruidos del espectro extendido o de intervalo relleno se basan en una o más copias directas que contienen ruido de la banda de base. Las modulaciones temporales que resultan de dicha correlación no deseada en ruido reconstruido son audibles de 65 una manera perturbadora como aspereza perceptual o distorsión objetable. Todos los procedimientos existentes

como mp3+SBR, AAC+SBR, USAC, G.719 y G.722.1C y también MPEG-H 3D IGF realizan primero una decodificación central completa que incluye el relleno de ruido antes de llenar los intervalos espectrales o la banda alta con datos espectrales copiados o especulares del núcleo.

5 [0009] El documento US 2014/0188464 A1 describe un aparato para generar una señal ampliada de ancho de banda. El aparato incluye una unidad de procesamiento anti-escasez para llevar a cabo un procesamiento anti-escasez en un espectro de baja frecuencia. Además, el aparato comprende una unidad de decodificación de extensión de alta frecuencia de dominio de frecuencia para llevar a cabo la codificación de extensión de alta frecuencia en el dominio de frecuencia en el espectro de baja frecuencia donde se lleva a cabo el procesamiento 10 anti-escasez.

[0010] El documento EP 2 704 142 A1 describe un aparato para reproducir una señal de audio basada en los primeros datos que representan una versión codificada de una primera porción de la señal de audio en una primera banda de frecuencia y segundos datos que representan información lateral en una segunda porción de la señal de 15 audio en una segunda banda de frecuencia, comprendiendo la segunda banda de frecuencia frecuencias más elevadas que la primera banda de frecuencia. El aparato comprende un primer reproductor configurado para reproducir la primera porción de la señal de audio basada en los primeros datos. Un proveedor está configurado para proporcionar una señal de parche en la segunda banda de frecuencia, donde la señal de parche está al menos parcialmente decorrelacionada con respecto a la primera porción de la señal de audio o es al menos parcialmente 20 una versión decorrelacionada de la primera porción de la señal de audio, que se ha desplazado a la segunda banda de frecuencia. Un segundo reproductor está configurado para reproducir la segunda porción de la señal de audio en la segunda banda de frecuencia basada en los segundos datos y la señal de parche. Un combinador está configurado para combinar la primera porción reproducida de la señal de audio y la señal de parche antes de que la 25 segunda porción de la señal de audio sea reproducida por el segundo reproductor o para combinar la primera porción reproducida de la señal de audio y la segunda porción reproducida de la señal de audio.

[0011] Es un objeto de la presente invención proporcionar un concepto mejorado de generación de una señal mejorada.

30 [0012] Este objeto se logra por medio de un aparato para generar una señal de audio mejorada según la reivindicación 1, un procedimiento para generar una señal de audio mejorada según la reivindicación 9, un sistema de codificación y decodificación según la reivindicación 10, un procedimiento de codificación y decodificación según la reivindicación 11 o un programa informático según la reivindicación 12.

35 [0013] La presente invención se basa en el hallazgo de que una mejora significativa de la calidad de audio de una señal mejorada generada por extensión del ancho de banda o relleno de intervalo inteligente o cualquier otra vía de generación de una señal mejorada que tiene valores espectrales para una regiónpectral de mejora no contenida en una señal de entrada se obtiene generando primeros valores de ruido para una región de relleno de ruido en una regiónpectral de la fuente de la señal de entrada y, a continuación, generando segundos valores de 40 ruido independientes para una región de ruidos en el destino o la región de destino, es decir, en la región de mejora que ahora tiene valores de ruido, es decir, los segundos valores de ruido que son independientes de los primeros valores de ruido.

45 [0014] De este modo, el problema de la técnica anterior que tiene ruido dependiente en la banda de base y la banda de mejora debido al mapeo de los valores espectrales se elimina y los problemas relacionados con artefactos tales como aspereza, distorsiones de la modulación y un timbre percibido como desagradable se eliminan en particular a bajas tasas de bits.

50 [0015] En otras palabras, el relleno de ruido de segundos valores de ruido decorrelacionados de los primeros valores de ruido, es decir, valores de ruido que son al menos parcialmente independientes de los primeros valores de ruido asegura que los artefactos no se produzcan o al menos se reduzcan con respecto a la técnica anterior. Así, el procesamiento de la técnica anterior de los valores espectrales de relleno de ruido en la banda de base por una extensión del ancho de banda directo u operación de relleno de intervalo inteligente no decorrelaciona el ruido de la banda de base, sino que solo cambia el nivel, por ejemplo. Sin embargo, introduciendo valores de ruido 55 decorrelacionados en la banda de fuente, por un lado, y en la banda de destino por el otro lado, preferentemente derivados de un proceso de ruidos separado, proporciona los mejores resultados. Sin embargo, incluso la introducción de valores de ruido no completamente decorrelacionados o no completamente independientes, pero al menos parcialmente decorrelacionados como por un valor de decorrelación de 0,5 o menos cuando el valor de decorrelación de cero indica completamente decorrelacionado, mejora el problema de correlación completa de la 60 técnica anterior.

[0016] De este modo, las realizaciones se refieren a una combinación de decodificación en forma de ondas, extensión del ancho de banda o relleno del intervalo y relleno de ruidos en un decodificador perceptual.

65 [0017] Otras ventajas son que, contrariamente a los conceptos ya existentes, se evita la aparición de

distorsiones de señales y artefactos de aspereza perceptual, que corrientemente son típicos para calcular extensiones de ancho de banda o relleno de intervalos posteriores a la decodificación en forma de ondas y relleno de ruido.

5 [0018] Esto se debe, en algunas realizaciones, a un cambio en el orden de las etapas de procesamiento mencionadas. Se prefiere llevar a cabo la extensión del ancho de banda o el relleno del intervalo directamente después de la decodificación en forma de ondas y también se prefiere computar el relleno de ruido posteriormente en la señal ya reconstruida mediante el uso de ruido no correlacionado.

10 [0019] En realizaciones adicionales, la decodificación en forma de ondas y relleno de ruido se puede llevar a cabo en un orden tradicional y también corriente abajo en el procesamiento, los valores de ruido se pueden reemplazar por ruido no correlacionado apropiadamente escalado.

15 [0020] Así, la presente invención se refiere a problemas que se producen debido a una operación de copia o una operación especular en espectros rellenos de ruido mediante el desplazamiento de la etapa de relleno de ruido a un extremo de una cadena de procesamiento y mediante el uso de ruido no correlacionado para la aplicación de parches o relleno de intervalos.

20 [0021] Posteriormente, se tratan las realizaciones preferidas de la presente invención con respecto a los dibujos adjuntos, donde:

- Fig. 1a ilustra un aparato para la codificación de una señal de audio;
- Fig. 1b ilustra un decodificador para la decodificación de una señal de audio codificada que coincide con el codificador de la Fig. 1a;
- 25 Fig. 2a ilustra una implementación preferida del decodificador;
- Fig. 2b ilustra una implementación preferida del codificador;
- Fig. 3a ilustra una representación esquemática de un espectro generado por el decodificador de dominio espectral de la Fig. 1b;
- Fig. 3b ilustra una tabla que indica la relación entre factores de escala para bandas de factores de escala y energías para las bandas de reconstrucción e información de relleno de ruido para una banda de relleno de ruido;
- 30 Fig. 4a ilustra la funcionalidad del codificador del dominio espectral para aplicar la selección de porciones espectrales en el primer y el segundo conjunto de porciones espectrales;
- Fig. 4b ilustra una implementación de la funcionalidad de la Fig. 4a;
- Fig. 5a ilustra una funcionalidad de un codificador MDCT;
- 35 Fig. 5b ilustra una funcionalidad del decodificador con una tecnología MDCT;
- Fig. 5c ilustra una implementación del regenerador de frecuencia;
- Fig. 6 ilustra un diagrama de bloques de un aparato para la generación de una señal mejorada según la presente invención;
- Fig. 7 ilustra un flujo de señales de relleno de ruido independiente dirigido por una información de selección en un decodificador según una realización de la presente invención;
- 40 Fig. 8 ilustra un flujo de señales de un relleno de ruido independiente implementado por medio de un orden intercambiado de relleno de intervalo o extensión de ancho de banda y relleno de ruido en un decodificador;
- Fig. 9 ilustra un diagrama de flujo de un procedimiento de ejemplo que no se abarca en la redacción de las reivindicaciones, pero se tiene en cuenta como útil para la comprensión de la invención;
- 45 Fig. 10 ilustra un diagrama de flujo de un procedimiento según una realización de la presente invención;
- Fig. 11 ilustra un diagrama de flujo para explicar una escalada de valores aleatorios;
- Fig. 12 ilustra un diagrama de flujo que ilustra una inclusión de la presente invención en una extensión general de ancho de banda o un procedimiento de relleno de intervalo;
- Fig. 13a ilustra un codificador con un cálculo de los parámetros de extensión de ancho de banda; y
- 50 Fig. 13b ilustra un decodificador con una extensión de ancho de banda implementado como un post-procesador más que un procedimiento integrado como en la Fig. 1a o 1b.

[0022] La Fig. 6 ilustra un aparato para generar una señal mejorada tal como una señal de audio a partir de una señal de entrada que también puede ser una señal de audio. La señal mejorada tiene valores espectrales para una región espectral de mejora, donde los valores espectrales para la región espectral de mejora no están contenidos en la señal de entrada original en una señal de entrada 600. El aparato comprende un mapeador 602 para mapear una región espectral de la fuente de la señal de entrada en una región de destino en la región espectral de mejora, donde la región espectral de la fuente comprende una región de relleno de ruido.

60 [0023] Por otra parte, el aparato comprende un rellenador de ruido 604 configurado para generar primeros valores de ruido para la región de relleno de ruido en la región espectral de la fuente de la señal de entrada y para generar segundos valores de ruido para una región de ruidos en la región de destino, donde los segundos valores de ruido, es decir, los valores de ruido en la región de destino son independientes o decorrelacionados de los primeros valores de ruido en la región de relleno de ruido.

[0024] Una realización se refiere a una situación donde el relleno del ruido se realiza realmente en la banda de base, es decir, donde los valores de ruido en la región de fuente fueron generados por el relleno del ruido. En otra alternativa, se asume que un relleno de ruido en la región de fuente no se realizó. Sin embargo, la región de fuente tiene una región de ruidos realmente rellenada con ruido como valores espectrales ilustrativamente codificados como valores espectrales por la fuente o el codificador central. El mapeo de este ruido como región de fuente a la región de mejora también genera ruido dependiente en la fuente y las regiones de destino. Con el fin de abordar este tema, el rellenador de ruido solo rellena ruido en la región de destino del mapeador, es decir, genera segundos valores de ruido para la región de ruido en la región de destino, donde los segundos valores de ruido se decorrelacionan de primeros valores de ruido en la región de fuente. Este reemplazo o relleno de ruido también puede tener lugar ya sea en un búfer mosaico de fuente o puede tener lugar en el destino propiamente dicho. La región de ruido puede ser identificada por el clasificador ya sea analizando la región de fuente o analizando la región de destino.

[0025] Para este fin, se hace referencia a la Fig. 3A. La Fig. 3A ilustra como región de relleno como una banda de factor de escala 301 en la señal de entrada y el rellenador de ruido genera los primeros valores espectrales de ruido en esta banda de relleno de ruido 301 en una operación de decodificación de la señal de entrada.

[0026] Por otra parte, esta banda de relleno de ruido 301 se mapea a una región de destino, es decir, según la técnica anterior, los valores de ruido generados se mapean a la región de destino y, en consecuencia, la región de destino tendrá un ruido dependiente o correlacionado con la región de fuente.

[0027] Según la presente invención, sin embargo, el rellenador de ruido 604 de la Fig. 6 genera segundos valores de ruido para una región de ruidos en el destino o la región de destino, donde los segundos valores de ruido se decorrelacionan o no correlacionan o independiente de los primeros valores de ruido en la banda de relleno de ruido 301 de la Fig. 3A.

[0028] En general, el relleno de ruido y el mapeador para mapear la región espectral de la fuente a una región de destino pueden estar incluidos dentro de un regenerador de alta frecuencia como se ilustra en el contexto de las Figs. 1A a 5C ilustrativamente dentro de un relleno de intervalo integrado o pueden ser implementados como un post-procesador como se ilustra en la Fig. 13b y el codificador correspondiente en la Fig. 13A.

[0029] En general, una señal de entrada se somete a una cuantificación inversa 700 o cualquier otro procesamiento de decodificador predefinido adicional 700 que significa que, en la salida del bloque 700, se obtiene la señal de entrada de la Fig. 6, de modo que la entrada en el bloque de relleno de ruido del codificador central o el bloque del rellenador de ruido 704 es la entrada 600 de la Fig. 6. El mapeador en la Fig. 6 corresponde al bloque de relleno de intervalo o extensión de ancho de banda 602 y el bloque de relleno de ruido independiente 702 también está incluido dentro del rellenador de ruido 604 de la Fig. 6. Así, los bloques 704 y 702 se incluyen ambos en el bloque de rellenador de ruido 604 de la Fig. 6 y el bloque 704 genera los así llamados primeros valores de ruido para una región de ruidos en la región de relleno de ruido y el bloque 702 genera los segundos valores de ruido para una región de ruidos en el destino o región de destino, que se deriva de la región de relleno de ruido en la banda de base por extensión de ancho de banda realizada por el mapeador o bloque de relleno de intervalo o extensión de ancho de banda 602. Por otra parte, como se trata más abajo, la operación de relleno de ruido independiente realizada por el bloque 702 está controlada por un vector de control PHI ilustrado por una línea de control 706.

45

1. Etapa: Identificación del ruido

[0030] En una primera etapa, se identifican todas las líneas espectrales que representan ruido en una trama de audio transmitida. El proceso de identificación puede ser controlado por conocimiento de posiciones de ruido transmitido ya existente usado por el relleno de ruido [4][5] o se puede identificar con un clasificador adicional. El resultado de la identificación de la línea de ruido es un vector que contienen ceros y unos, donde una posición con un uno indica una línea espectral que representa ruido.

[0031] En términos matemáticos, este procedimiento se puede describir como:

55 $\hat{X} \in \mathbb{C}^N$ es un espectro transmitido y recuantificado después del relleno de ruido [4][5] de una señal en ventana codificada por transformada de la longitud $N \in \mathbb{N}$, $m \in \mathbb{N}, 0 < m \leq N$, es la línea de detención de todo el proceso de decodificación.

[0032] El clasificador C_0 determina líneas espectrales donde se usa el relleno de ruido [4][5] en la región central:

$$C_0: \mathbb{C}^N \rightarrow \{0,1\}^m,$$

$$\varphi[i] := C_0(\tilde{X})[i] := \begin{cases} 1 & \text{en } \tilde{X}[i] \text{ también se utilizó relleno, } 0 \leq i < m \leq N, \\ 0 & \text{de ruido} \end{cases}$$

y el resultado $\varphi \in \{0,1\}^m$ es un vector de longitud m.

- 5 [0033] Un clasificador C_1 adicional puede identificar líneas adicionales en \tilde{X} que representa ruido. Este clasificador se puede describir como:

$$C_1: \mathbb{C}^N \times \{0,1\}^m \rightarrow \{0,1\}^m,$$

$$\varphi[i] := C_1(\tilde{X}, \varphi)[i] := \begin{cases} 1 & \text{si } \varphi[i] = 1 \vee \tilde{X}[i] \text{ se clasificó como ruido} \\ 0 & \text{también} \end{cases}, 0 \leq i < m \leq N.$$

- 10 [0034] Despues del proceso de identificación, el vector de indicación de ruido $\varphi \in \{0,1\}^m$ se define como:

$$\varphi[i] = \begin{cases} 1 & \text{la línea espectral } \tilde{X}[i] \text{ se identifica como una línea de ruido} \\ 0 & \text{la línea espectral } \tilde{X}[i] \text{ no se identifica como una línea de ruido} \end{cases}, 0 \leq i < m \leq N.$$

2. Etapa: ruido independiente

- 15 [0035] En la segunda etapa, se selecciona una región específica del espectro transmitido y se copia a un mosaico de fuente. Dentro de este mosaico de fuente, el ruido identificado se reemplaza por ruido aleatorio. La energía del ruido aleatorio insertado se ajusta a la misma energía del ruido original en el mosaico de fuente.

- 20 [0036] En términos matemáticos, este procedimiento se puede describir como:

n, $n < m$, es la línea de inicio es para la copia del proceso, descrito en la etapa 3. $\tilde{X}_{st} \subset \tilde{X}$ es una parte continua de un espectro transmitido \tilde{X} , que representa un mosaico de fuente de longitud $v < n$, que contiene las líneas espectrales $l_k, l_{k+1}, \dots, l_{k+v-1}$ de \tilde{X} , donde k es el índice de la primera línea espectral en el mosaico de fuente \tilde{X}_{st} , de modo que $\tilde{X}_{st}[i] = l_{k+i}, 0 \leq i < v$. Por otra parte, $\varphi' \subset \varphi$, de modo que $\varphi'[i] = \varphi[k+i], 0 \leq i < v$.

- [0037] El ruido identificado es reemplazado ahora por ruido sintético generado aleatoriamente. Con el fin de mantener la energía espectral en el mismo nivel, la energía E de ruido indicada por φ se calcula primero:

$$E := \sum_{i=0}^{v-1} \varphi'[i] |\tilde{X}_{st}[i]|^2.$$

- [0038] Si $E = 0$ omite el reemplazo del ruido independiente para el mosaico de fuente \tilde{X}_{st} , entonces reemplazar el ruido indicado por φ' :

$$\hat{X}_{st}''[i] := \begin{cases} r[i], & \varphi'[i] = 1 \\ \hat{X}_{st}'[i], & \varphi'[i] = 0 \end{cases}, \quad 0 \leq i < v,$$

donde $r[i] \in \mathbb{C}$ es un número aleatorio para todos $0 \leq i < v$.

5 [0039] A continuación, calcular la energía E' de los números aleatorios insertados:

$$E' := \sum_{i=0}^{v-1} \varphi[i] |\hat{X}_{st}'[i]|^2.$$

10 [0040] Si $E' > 0$ calcula un factor g , fijar entonces $g = 0$.

$$g := \sqrt{\frac{E}{E'}}.$$

[0041] Con g , reescalar el ruido reemplazado:

$$\hat{X}_{st}'''[i] := \begin{cases} g \hat{X}_{st}'[i], & \varphi'[i] = 1 \\ \hat{X}_{st}'[i], & \varphi'[i] = 0 \end{cases}, \quad 0 \leq i < v.$$

15 [0042] Despues del reemplazo del ruido, el mosaico de fuente \hat{X}_{st}''' contiene líneas de ruido que son independientes de las líneas de ruido en \hat{X} .

20 3. Etapa: copiado

[0043] El mosaico de fuente $\hat{X}_{st}'''[i]$ se mapea a su región de destino en \hat{X} :

$$\hat{X}[c+i] = \hat{X}_{st}'''[i], \quad 0 \leq i < v, \quad c \geq n, \quad c+i < m < N,$$

25 o, si se usa el esquema IGF [8]:

$$\hat{X}[c+i] = \begin{cases} \hat{X}_{st}'''[i], & \hat{X}[c+i] = 0 \\ \hat{X}[c+i], & \hat{X}[c+i] < 0 \end{cases}, \quad 0 \leq i < v, \quad c \geq n, \quad c+i < m < N.$$

30 [0044] La Fig. 8 ilustra una realización donde, después de cualquier post-procesamiento como la decodificación del dominio espectral ilustrado en el bloque 112 en la Fig. 1B o en la realización del post-procesador ilustrado por el bloque 1326 en la Fig. 13B, la señal de entrada se somete a un relleno de intervalo o extensión de ancho de banda primero, es decir, se somete a una operación de mapeo primero y, a continuación, se lleva a cabo un relleno de ruido independiente, es decir, dentro del espectro completo.

35 [0045] El proceso descrito en el contexto anterior de la Fig. 7 se puede llevar a cabo en una orientación in situ, de modo que el búfer intermedio \hat{X}_{st}''' no es necesario. En consecuencia, se adapta el orden de ejecución.

[0046] Ejecutar la primera etapa como se describe en el contexto de la Fig. 7, de nuevo el conjunto de líneas 40 espectrales $k, k+1, \dots, k+v-1$ de \hat{X} son la región de fuente. Realizar:

2. Etapa: copiado

[0047]

45

$$\hat{X}[c+i] = \hat{X}[k+i], \quad 0 \leq i < v, \quad c \geq n, \quad 0 < k+i < n, \quad c+i < m < N,$$

o, si se usa el esquema IGF [8]:

$$\hat{X}[c+i] = \begin{cases} \hat{X}[k+i], & \hat{X}[c+i] = 0 \\ \hat{X}[c+i], & \hat{X}[c+i] <> 0 \end{cases}$$

$$0 \leq i < v, c \geq n, \quad 0 < k+i < n, \quad c+i < m < N.$$

3. Etapa: relleno de ruido independiente

[0048] Realizar el relleno de ruido restante hasta n y calcular la energía de líneas espectrales de ruido en la región de fuente $k, k+1, \dots, k+v-1$:

$$E := \sum_{i=0}^{v-1} \varphi[k+i] |\hat{X}[k+i]|^2.$$

[0049] Realizar el relleno de ruido independiente en el relleno de intervalo o la región espectral BWE:

$$\hat{X}[c+i] := \begin{cases} r[i], & \varphi[k+i] = 1 \\ \hat{X}[c+i], & \varphi[k+i] = 0 \end{cases}, \quad 0 \leq i < v,$$

donde $r[i], 0 \leq i < v$ es de nuevo un conjunto de números aleatorios.

[0050] Calcular la energía E' de los números aleatorios insertados:

$$E' := \sum_{i=0}^{v-1} \varphi[k+i] |\hat{X}[c+i]|^2.$$

[0051] De nuevo, si $E' > 0$ calcular el factor $g := 0$:

$$g := \sqrt{\frac{E}{E'}}.$$

[0052] Con g, reescalar el ruido de reemplazo:

$$\hat{X}[c+i] := \begin{cases} g\hat{X}[c+i], & \varphi[k+i] = 1 \\ \hat{X}[c+i], & \varphi[k+i] = 0 \end{cases}, \quad 0 \leq i < v.$$

[0053] El relleno de ruido independiente de la invención también se puede usar en un ambiente de pares de canales estéreo. En consecuencia, el codificador calcula la representación de pares de canales apropiados, L/R o M/S, por banda de frecuencia y coeficientes de predicción opcionales. El decodificador aplica relleno de ruido independiente como se ha descrito con anterioridad a la representación apropiadamente seleccionada de los canales antes de la posterior computación de la conversión final de todas las bandas de frecuencia en la representación L/R.

[0054] La invención es aplicable o apropiada para todas las aplicaciones de audio donde el ancho de banda completo no está disponible o que usa relleno de intervalo para llenar agujeros espectrales. La invención se puede usar en la distribución o difusión del contenido de audio como, por ejemplo, con radio digital, transmisión por Internet y aplicaciones de comunicaciones por audio.

[0055] Posteriormente, se tratan realizaciones de la presente invención con respecto a las Figs. 9-12. En la etapa 900, las regiones de ruido se identifican en el intervalo de la fuente. Este procedimiento que se trató anteriormente con respecto a la "Identificación del ruido" se puede basar en la información lateral del relleno de ruido recibida por completo del lado del codificador o también se puede configurar para basarse alternativa o adicionalmente en el análisis de señales de la señal de entrada ya generada, pero sin valores espectrales para la región espectral de mejora, es decir, sin los valores espectrales para esta región espectral de mejora.

[0056] A continuación, en la etapa 902, el intervalo de la fuente que ya se había sometido a un relleno de ruido directo como se conoce en la técnica, es decir, un intervalo de fuente completo se copia a un búfer mosaico de fuente.

[0057] A continuación, en la etapa 904, los primeros valores de ruido, es decir, los valores de ruido directos generados dentro de la región de relleno de ruido de la señal de entrada se reemplazan en el búfer mosaico de fuente por valores aleatorios. A continuación, en la etapa 906, estos valores aleatorios se escalan en el búfer mosaico de fuente para obtener los segundos valores de ruido para la región de destino. A continuación, en la etapa 908, se realiza la operación de mapeo, es decir, su contenido del búfer mosaico de fuente disponible después de las etapas 904 y 906 se mapea en el intervalo de destino. De esta manera, por medio de la operación de reemplazo 904 y después de la operación de mapeo 908, se obtuvo la operación de relleno de ruido independiente en el intervalo 20 de la fuente y en el intervalo de destino.

[0058] La Fig. 10 ilustra una realización adicional de la presente invención. De nuevo, en la etapa 900, se identifica el ruido en el intervalo de la fuente. Sin embargo, la funcionalidad de esta etapa 900 es diferente de la funcionalidad de la etapa 900 en la Fig. 9, dado que la etapa 900 en la Fig. 9 puede operar en un espectro de señales de entrada que tiene valores de ruido ya recibidos, es decir, donde ya se realizó la operación de relleno de ruido.

[0059] Sin embargo, en la Fig. 10, cualquier operación de relleno de ruido en la señal de entrada no se realizó y la señal de entrada aún no tiene valores de ruido en la región de relleno de ruido en la entrada en la etapa 30 902. En la etapa 902, el intervalo de la fuente se mapea en el destino o el intervalo de destino, donde los valores de relleno de ruido no se incluyen en el intervalo de la fuente.

[0060] De esta manera, la identificación del ruido en el intervalo de la fuente en la etapa 900 se puede llevar a cabo, con respecto a la región de relleno de ruido, por identificación de valores espectrales cero en la señal y/o 35 mediante el uso de esta información lateral de relleno de ruido a partir de la señal de entrada, es decir, la información de relleno de ruido generada del lado del codificador. A continuación, en la etapa 904, la información de relleno de ruido y en particular la información de energía identifica que se lee la energía por introducir en la señal de entrada del lado del decodificador.

[0061] A continuación, como se ilustra en la etapa 1006, se lleva a cabo un relleno de ruido en el intervalo de la fuente y, posterior o concurrentemente, se lleva a cabo una etapa 1008, es decir, se insertan valores aleatorios en 40 posiciones en el intervalo de destino que fueron identificados por la etapa 900 en toda la banda o que fueron identificados mediante el uso de la banda de base o la información de señales de entrada junto con la información de mapeo, es decir, qué intervalo de fuente (de una pluralidad de) fuentes se mapea con qué intervalo de destino (de una pluralidad de) intervalos de destino.

[0062] Finalmente, los valores aleatorios insertados se escalan para obtener los valores de ruido independientes o no correlacionados o decorrelacionados.

[0063] Posteriormente, la Fig. 11 se trata con el fin de ilustrar mayor información acerca de la escala de los 50 valores de relleno de ruido en la región espectral de mejora, es decir, de cómo se obtienen, a partir de los valores aleatorios, los segundos valores de ruido.

[0064] En la etapa 1100, se obtiene una información de energía acerca del ruido en el intervalo de la fuente. 55 A continuación, se determina una información de energía a partir de los valores aleatorios, es decir, a partir de los valores generados por un proceso aleatorio o pseudoaleatorio, tal como se ilustra en la etapa 1102. Por otra parte, la etapa 1104 ilustra la forma de calcular el factor de escala, es decir, mediante el uso de la información de energía acerca del ruido en el intervalo de la fuente y mediante el uso de la información de energía acerca de los valores aleatorios. A continuación, en la etapa 1106, los valores aleatorios, es decir, de los que se calculó la energía en la 60 etapa 1102, se multiplican por el factor de escala generado por la etapa 1104. Así, el procedimiento ilustrado en la Fig. 11 corresponde al cálculo del factor de escala y ilustrado antes en una realización. Sin embargo, todos estos cálculos también se pueden llevar a cabo en un dominio logarítmico o en cualquier otro dominio y la etapa de multiplicación 1106 se puede reemplazar por una adición o sustracción en el intervalo logarítmico.

[0065] Otra referencia se hace a la Fig. 12 con el fin de ilustrar la incorporación de la presente invención

- dentro de un esquema de relleno de intervalo general inteligente o extensión de ancho de banda. En la etapa 1200, se recupera información envolvente espectral de la señal de entrada. La información envolvente espectral puede ser generada, por ejemplo, por un extractor de parámetros 1306 de la Fig. 13a y se puede proporcionar por un decodificador de parámetros 1324 de la Fig. 13b. A continuación, los segundos valores de ruido y los otros valores 5 en el intervalo de destino se escalan mediante el uso de esta información envolvente espectral como se ilustra en 1202. Posteriormente, cualquier otro post-procesamiento 1204 se puede llevar a cabo para obtener la señal mejorada del dominio de tiempo final que tiene un mayor ancho de banda en el caso de una extensión de ancho de banda o que tiene una cantidad reducida o ningún agujero espectral en el contexto de relleno de intervalo inteligente.
- 10 [0066] En este contexto, se destaca que, en particular para la realización de la Fig. 9, se pueden aplicar varias alternativas. Para una realización, la etapa 902 se lleva a cabo con el espectro completo de la señal de entrada o al menos con la porción del espectro de la señal de entrada que está por encima de la frecuencia de borde del relleno de ruido. Esta frecuencia asegura que, por debajo de cierta frecuencia, es decir, por debajo de esta frecuencia, no se realiza en absoluto un relleno de ruido.
- 15 [0067] A continuación, independientemente de cualquier información de intervalo de fuente/mapeo de intervalo de destino, todo el espectro de señales de entrada, es decir, el intervalo de fuente potencial completo se copia al búfer mosaico de fuente 902 y, a continuación, se procesa con la etapa 904 y 906 y la etapa 908, después se selecciona cierta región de fuente específicamente requerida de este búfer mosaico de fuente.
- 20 [0068] En otros ejemplos, sin embargo, solo los intervalos de fuente específicamente requeridos que solo pueden ser partes de la señal de entrada se copian al búfer mosaico de fuente individual o a varios búfers mosaicos de fuentes individuales en base a la información de intervalo de fuente/intervalo de destino incluida en la señal de entrada, es decir, asociada como información lateral con esta señal de entrada de audio. Dependiendo de esta 25 situación, la segunda alternativa, donde solo los intervalos de fuente específicamente requeridos son procesados por las etapas 902, 904, 906, se puede reducir la complejidad o al menos los requisitos de memoria en comparación con la situación donde siempre, independientemente de la situación de mapeo específica, se procesa todo el intervalo de fuente al menos por encima de la frecuencia de borde del relleno de ruido por medio de las etapas 902, 904, 906.
- 30 [0069] Posteriormente, se hace referencia a las Figs. 1a - 5c con el fin de ilustrar la implementación específica de la presente invención dentro de un regenerador de frecuencia 116, que se ubica delante del conversor de espectro-tiempo 118.
- 35 [0070] La Fig. 1a ilustra un aparato para la codificación de una señal de audio 99. La señal de audio 99 se introduce en un conversor de espectro-tiempo 100 para convertir una señal de audio que tiene una tasa de muestreo en una representación espectral 101 producida por el conversor de espectro-tiempo. El espectro 101 se introduce en un analizador espectral 102 para analizar la representación espectral 101. El analizador espectral 101 se configura para determinar un primer conjunto de primeras porciones espetrales 103 para codificar con una primera resolución espectral y un segundo conjunto diferente de segundas porciones espetrales 105 para codificar con una segunda 40 resolución espectral. La segunda resolución espectral es menor que la primera resolución espectral. El segundo conjunto de segundas porciones espetrales 105 se introduce en una calculadora de parámetros o codificador paramétrico 104 para calcular información envolvente espectral que tiene la segunda resolución espectral. Por otra parte, se proporciona un codificador de audio del dominio espectral 106 para generar una primera representación codificada 107 del primer conjunto de primeras porciones espetrales que tiene la primera resolución espectral. Por 45 otra parte, la calculadora de parámetros/codificador paramétrico 104 se configura para generar una segunda representación codificada 109 del segundo conjunto de segundas porciones espetrales. La primera representación codificada 107 y la segunda representación codificada 109 se introducen en un multiplexador de corriente de bits o formador de corriente de bits 108 y el bloque 108 finalmente produce la señal de audio codificada para transmisión o almacenamiento en un dispositivo de almacenamiento.
- 50 [0071] Típicamente, una primera porción espectral como 306 de la Fig. 3a se rodeará por dos segundas porciones espetrales como 307a, 307b. Este no es el caso en HE AAC, donde el intervalo de frecuencia del codificador central tiene banda limitada.
- 55 [0072] La Fig. 1b ilustra un decodificador que concuerda con el codificador de la Fig. 1a. La primera representación codificada 107 se introduce en un decodificador de audio del dominio espectral 112 para generar una primera representación decodificada de un primer conjunto de primeras porciones espetrales, donde la representación decodificada tiene una primera resolución espectral. Por otra parte, la segunda representación codificada 109 se introduce en un decodificador paramétrico 114 para generar una segunda representación 60 decodificada de un segundo conjunto de segundas porciones espetrales que tiene una segunda resolución espectral que es inferior a la primera resolución espectral.
- 65 [0073] El decodificador comprende además un regenerador de frecuencia 116 para regenerar una segunda porción espectral reconstruida que tiene la primera resolución espectral mediante el uso de una primera porción espectral. El regenerador de frecuencia 116 realiza una operación de relleno de mosaicos, es decir, usa un mosaico

o una porción del primer conjunto de primeras porciones espectrales y copia este primer conjunto de primeras porciones espectrales en el intervalo de reconstrucción o banda de reconstrucción que tiene la segunda porción espectral y típicamente realiza formación de envolvente espectral u otra operación como se indica por la segunda representación decodificada producida por el decodificador paramétrico 114, es decir, mediante el uso de la 5 información acerca del segundo conjunto de segundas porciones espectrales. El primer conjunto decodificado de primeras porciones espectrales y el segundo conjunto reconstruido de porciones espectrales como se indica en la salida del regenerador de frecuencia 116 en la línea 117 se introduce en un conversor de espectro-tiempo 118 configurado para convertir la primera representación decodificada y la segunda porción espectral reconstruida en una representación temporal 119, donde la representación temporal tiene cierta tasa de muestreo alta.

10

[0074] La Fig. 2b ilustra una implementación del codificador de la Fig. 1a. Una señal de entrada de audio 99 se introduce en un banco de filtro de análisis 220 correspondiente al conversor de espectro-tiempo 100 de la Fig. 1a. A continuación, una operación de formación de ruido temporal se lleva a cabo en el bloque TNS 222. En consecuencia, la entrada en el analizador espectral 102 de la Fig. 1a correspondiente a la máscara tonal de bloque 15 226 de la Fig. 2b pueden ser valores espectrales completos, cuando no se aplica la operación de formación de ruido temporal / formación de mosaicos temporales o pueden ser valores residuales espectrales, cuando se aplica la operación de TNS como se ilustra en la Fig. 2b, bloque 222. Para señales de dos canales o señales de múltiples canales, se puede llevar a cabo adicionalmente una codificación de canal conjunto 228, de modo que el codificador del dominio espectral 106 de la Fig. 1a pueda comprender el bloque codificador del canal conjunto 228. Por otra 20 parte, se proporciona un codificador de entropía 232 para realizar una compresión de datos sin pérdidas que también es una porción del codificador del dominio espectral 106 de la Fig. 1a.

[0075] El analizador espectral / máscara tonal 226 separa la salida de TNS bloque 222 en la banda central y los componentes tonales correspondientes al primer conjunto de primeras porciones espectrales 103 y los 25 componentes residuales correspondientes al segundo conjunto de segundas porciones espectrales 105 de la Fig. 1a. El bloque 224 indicado como codificación de extracción de parámetros IGF corresponde al codificador paramétrico 104 de la Fig. 1a y el multiplexador de la corriente de bits 230 corresponde al multiplexador de la corriente de bits 108 de la Fig. 1a.

[0076] De preferencia, el banco de filtros de análisis 222 se implementa como MDCT (banco de filtro de transformada de coseno discreta modificada) y la MDCT se usa para transformar la señal 99 en un dominio de tiempo-frecuencia con la transformada de coseno discreta modificada que actúa como la herramienta de análisis de frecuencia.

[0077] El analizador espectral 226 aplica preferentemente una máscara de tonalidad. Esta etapa de estimación de máscara de tonalidad se usa para separar los componentes tonales de los componentes de tipo ruido en la señal. Esto permite que el codificador central 228 codifique todos los componentes tonales con un módulo psicoacústico. La etapa de estimación de la máscara de tonalidad se puede implementar en numerosas vías diferentes y preferentemente se implementa de modo similar en su funcionalidad a la etapa de estimación de rastreo 40 sinusoidal usada en una modelación sinusoidal y de ruido para codificación de habla/audio [8, 9] o un codificador de audio basado en un modelo de HILN descrito en [10]. De preferencia, se usa una implementación que es fácil de implementar sin necesidad de mantener las trayectorias de nacimiento-muerte, pero también se puede usar cualquier otra tonalidad o detector de ruido.

[0078] El módulo IGF calcula la similitud que existe entre una región de fuente y una región de destino. La región de destino estará representada por el espectro de la región de fuente. La medición de la similitud entre la región de fuente y la región de destino se realiza mediante el uso de una estrategia de correlación cruzada. La 45 región de destino se divide en mosaicos de frecuencia *nTar* que no se superponen. Para cada mosaico en la

región de destino, se crean mosaicos de fuente *nSrc* a partir de una frecuencia de inicio fija. Estos mosaicos de 50 fuente se superponen con un factor de entre 0 y 1, donde 0 significa el 0 % de superposición y 1 significa el 100 % de superposición. Cada uno de estos mosaicos de fuente se correlaciona con el mosaico de destino en diferentes retardos para hallar el mosaico de fuente que concuerde mejor con el mosaico de destino. El número de mosaicos

de mejor concordancia se almacena en *tileNum[idx_tar]*, en cuyo retardo se correlaciona mejor con el blanco, se almacena en *xcorr_lag[idx_tar][idx_src]* y el signo de la correlación se almacena en 55 *xcorr_sign[idx_tar][idx_src]*. En el caso de que la correlación sea altamente negativa, el mosaico de fuente necesita ser multiplicado por -1 antes del proceso de relleno de mosaicos en el decodificador. El módulo IGF también cuida de no sobrescribir los componentes tonales en el espectro, ya que los componentes tonales se preservan mediante el uso de la máscara de tonalidad. Se usa un parámetro de tipo banda para almacenar la energía de la región de destino que nos permite reconstruir el espectro de forma precisa.

60

[0079] Este procedimiento tiene ciertas ventajas con respecto al clásico SBR [1] porque la rejilla de armónicos de una señal multitonal es preservada por el codificador central, mientras que solo los intervalos entre las sinusoides se llenan con el "ruido formado" de mejor concordancia de la región de fuente. Otra ventaja de este sistema comparado con ASR (*Accurate Spectral Replacement*, reemplazo espectral preciso) [2-4] es la ausencia de una etapa de síntesis de señales que crea las porciones importantes de la señal en el decodificador. En vez de ello, esta tarea es realizada por el codificador central, que permite la preservación de importantes componentes del espectro. Otra ventaja del sistema propuesto es la continua escalabilidad que ofrecen las características.

Justamente el uso de `tileNum[Idx_Tar]` y `xcorr_lag = 0` para cada mosaico se denomina concordancia de granularidad gruesa y se puede usar para bajas tasas de bits mientras que el uso de `xcorr_lag` variable para 10 cada mosaico nos permite hacer concordar mejor los espectros de destino y de fuente.

[0080] Además, se propone una técnica de estabilización de selección de mosaicos que elimina los artefactos de dominio de frecuencia como trinos y ruido musical.

15 **[0081]** En el caso de pares de canales estéreo, se aplica un procesamiento conjunto adicional. Esto es necesario porque para determinado intervalo de destino, la señal puede ser una fuente de sonido paneada altamente correlacionada. En el caso de que las regiones de fuente elegidas para esta región particular no se correlacionen bien, a pesar de que las energías concuerdan para las regiones de destino, la imagen espacial puede sufrir debido a las regiones de fuente no correlacionadas. El codificador analiza cada banda de energía de la región 20 de destino, típicamente realizando una correlación cruzada de los valores espectrales y si se excede cierto umbral, fija un indicador conjunto para esta banda de energía. En el decodificador, las bandas de energía del canal izquierdo y derecho se tratan de modo individual si no se fija este indicador estéreo conjunto. En el caso de fijar el indicador estéreo conjunto, tanto las energías como la conexión provisional se realizan en el dominio estéreo conjunto. La información estéreo conjunta para las regiones IGF se señaliza de modo similar a la información estéreo conjunta 25 para la codificación central, incluyendo un indicador que indica en el caso de una predicción si la dirección de la predicción es de mezcla descendente a residual, o viceversa.

[0082] Las energías se pueden calcular a partir de las energías transmitidas en el dominio L/R.

$$30 \quad midNrg[k] = leftNrg[k] + rightNrg[k];$$

$$sideNrg[k] = leftNrg[k] - rightNrg[k];$$

siendo k el índice de frecuencia en el dominio de la transformada.

35 **[0083]** Otra solución consiste en calcular y transmitir las energías directamente en el dominio estéreo conjunto para bandas donde el estéreo conjunto está activo, de modo que no se necesite una transformación adicional de energía en el lado del decodificador.

40 **[0084]** Los mosaicos de fuente siempre se crean según la matriz media/lateral:

$$sideTile[k] = 0,5 \cdot (leftTile[k] - rightTile[k])$$

$$sideTile[k] = 0,5 \cdot (leftTile[k] - rightTile[k])$$

45 **[0085]** Ajuste de energía:

$$midTile[k] = midTile[k] * midNrg[k];$$

$$sideTile[k] = sideTile[k] * sideNrg[k];$$

50 **[0086]** Estéreo conjunto -> transformación LR:
Si no se codifica ningún parámetro de predicción adicional:

$$leftTile[k] = midTile[k] + sideTile[k]$$

$$rightTile[k] = midTile[k] - sideTile[k]$$

[0087] Si se codifica un parámetro de predicción adicional y si la dirección señalada va de la parte media al lateral:

5

$$sideTile[k] = sideTile[k] - predictionCoeff \cdot midTile[k]$$

$$leftTile[k] = midTile[k] + sideTile[k]$$

$$rightTile[k] = midTile[k] - sideTile[k]$$

[0088] Si la dirección señalada va del lateral a la parte media:

$$midTile1[k] = midTile[k] - predictionCoeff \cdot sideTile[k]$$

$$leftTile[k] = midTile1[k] - sideTile[k]$$

$$rightTile[k] = midTile1[k] + sideTile[k]$$

10

[0089] Este procesamiento asegura que desde los mosaicos usados para regenerar regiones de destino altamente correlacionadas y regiones de destino paneadas, los canales izquierdo y derecho resultantes aún representan una fuente de sonido correlacionada y paneada incluso si las regiones de fuente no se correlacionan, 15 preservando la imagen estéreo para tales regiones.

[0090] En otras palabras, en la corriente de bits, los indicadores estéreos conjuntos se transmiten, lo que indica si se debe usar L/R o M/S como un ejemplo para la codificación estéreo conjunta general. En el decodificador, primero se decodifica la señal central como es indicado por los indicadores estéreo conjuntos para las bandas 20 centrales. En segundo lugar, la señal central se almacena en ambas representaciones L/R y M/S. Para el relleno de mosaico IGF, se elige la representación de mosaico de fuente para ajustar la representación de mosaico de destino como se indica por la información estéreo conjunta para las bandas de IGF.

[0091] *Temporal Noise Shaping* (TNS, configuración de ruido temporal) es una técnica estándar y parte de 25 AAC. TNS puede ser considerada como una extensión del esquema básico de un codificador perceptual, insertando una etapa de procesamiento opcional entre el banco de filtros y la etapa de cuantificación. La tarea principal del módulo de TNS es esconder el ruido de cuantificación producido en la región de enmascaramiento temporal de señales de tipo transitorio y, así, lleva a un esquema de codificación más eficiente. En primer lugar, TNS calcula un conjunto de coeficientes de predicción mediante el uso de "predicción directa" en el dominio de la transformada, por 30 ejemplo, MDCT. Estos coeficientes se usan a continuación para aplanar la envolvente temporal de la señal. Como la cuantificación afecta al espectro filtrado de TNS, también el ruido de la cuantificación es temporalmente plano. Al aplicar el filtrado de TNS inversa en el lado del decodificador, el ruido de cuantificación se conforma según la envolvente temporal del filtro de TNS y, por ello, el ruido de cuantificación se enmascara por el transitorio.

[0092] IGF se basa en una representación de MDCT. Para una codificación eficaz, se deben usar preferentemente bloques largos de aproximadamente 20 ms. Si la señal dentro de tal bloque largo contiene transitorios, se producen pre- y post-ecos audibles en las bandas espectrales de IGF debido al relleno de mosaicos.

[0093] Este efecto de pre-eco se reduce mediante el uso de TNS en el contexto de IGF. Aquí, se usa TNS 40 como una herramienta de configuración de mosaicos temporales (TTS) cuando la regeneración espectral en el decodificador se lleva a cabo en la señal residual de TNS. Los coeficientes de predicción de TTS requeridos se calculan y se aplican mediante el uso del espectro total en el lado del codificador como es usual. Las frecuencias de

inicio y detención de TNS/TTS no son afectadas por la frecuencia de inicio de IGF $f_{IGFstart}$ de la herramienta de IGF. En comparación con TNS de herencia, la frecuencia de detención de TTS se aumenta hasta la frecuencia de

detención de la herramienta de IGF, que es mayor que $f_{IGFstart}$. En el lado del decodificador, se aplican de nuevo los coeficientes de TNS/TTS en el espectro total, es decir, el espectro central más el espectro regenerado más los 5 componentes tonales del mapa de tonalidad. La aplicación de TTS es necesaria para formar la envolvente temporal del espectro regenerado para hacer concordar la envolvente de la señal original. Así, se reducen los pre-ecos mostrados. Además, también conforma el ruido de cuantificación en la señal por debajo de $f_{IGFstart}$ como es usual con TNS.

10 [0094] En decodificadores heredados, la conexión provisional espectral en una señal de audio corrompe la correlación espectral en los bordes del parche y, así, altera la envolvente temporal de la señal de audio al introducir una dispersión. Así, otro beneficio de realizar el relleno de mosaico IGF en la señal residual es que, después de la aplicación del filtro de conformación, los bordes del mosaico se correlacionan perfectamente, dando como resultado una reproducción temporal más fiel de la señal.

15 [0095] En un codificador de la invención, el espectro que se sometió a filtración TNS/TTS, procesamiento de máscara de tonalidad y estimación paramétrica IGF está desprovisto de cualquier señal por encima de la frecuencia de inicio IGF excepto por los componentes tonales. Este espectro disperso está codificado ahora por el codificador central mediante el uso de los principios de la codificación aritmética y la codificación predictiva. Estos componentes 20 codificados junto con los bits de señalización forman la corriente de bits de audio.

25 [0096] La Fig. 2a ilustra la correspondiente implementación del decodificador. La corriente de bits en la Fig. 2a correspondiente a la señal de audio codificada se introduce en el demultiplexador/decodificador que se conectaría, con respecto a la Fig. 1b, con los bloques 112 y 114. El demultiplexador de la corriente de bits separa la señal de audio de entrada en la primera representación codificada 107 de la Fig. 1b y la segunda representación codificada 109 de la Fig. 1b. La primera representación codificada que tiene el primer conjunto de primeras porciones espectrales se introduce en el bloque de decodificación del canal conjunto 204 correspondiente al decodificador de dominio espectral 112 de la Fig. 1b. La segunda representación codificada se introduce en el decodificador paramétrico 114 no ilustrado en la Fig. 2a y, a continuación, se introduce en el bloque de IGF 202 30 correspondiente al regenerador de frecuencia 116 de la Fig. 1b. El primer conjunto de primeras porciones espectrales requerido para la regeneración de frecuencia se introduce en el bloque IGF 202 a través de la línea 203. Por otra parte, después de la decodificación del canal conjunto 204, se aplica la decodificación central específica en el bloque de máscara tonal 206 de modo que la salida de la máscara tonal 206 corresponde a la salida del decodificador de dominio espectral 112. A continuación, se lleva a cabo una combinación por medio del combinador 35 208, es decir, una formación de trama, donde la salida del combinador 208 tiene ahora el espectro de intervalo completo, pero aún en el dominio filtrado TNS/TTS. A continuación, en el bloque 210, se lleva a cabo una operación de TNS/TTS inversa mediante el uso de información de filtro TNS/TTS proporcionada por medio de la línea 109, es decir, la información lateral de TTS se incluye preferentemente en la primera representación codificada generada por el codificador del dominio espectral 106 que puede ser, por ejemplo, un codificador central AAC o USAC directo o 40 45 también se puede incluir en la segunda representación codificada. En la salida del bloque 210, se proporciona un espectro completo hasta la frecuencia máxima, que es la frecuencia de intervalo completo definida por la tasa de muestreo de la señal de entrada. A continuación, se lleva a cabo una conversión de espectro/tiempo en el banco de filtros de síntesis 212 para obtener finalmente la señal de salida de audio.

45 [0097] La Fig. 3a ilustra una representación esquemática del espectro. El espectro se subdivide en bandas de factor de escala SCB donde hay siete bandas de factor de escala SCB1 a SCB7 en el ejemplo ilustrado de la Fig. 3a. Las bandas de factor de escala pueden ser bandas de factor de escala AAC que se definen en AAC estándar y tienen un mayor ancho de banda que frecuencias superiores como se ilustra en la Fig. 3a esquemáticamente. Se prefiere llevar a cabo un relleno de intervalo inteligente no desde el principio del espectro, es decir, a bajas 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995 1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055 1060 1065 1070 1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150 1155 1160 1165 1170 1175 1180 1185 1190 1195 1200 1205 1210 1215 1220 1225 1230 1235 1240 1245 1250 1255 1260 1265 1270 1275 1280 1285 1290 1295 1300 1305 1310 1315 1320 1325 1330 1335 1340 1345 1350 1355 1360 1365 1370 1375 1380 1385 1390 1395 1400 1405 1410 1415 1420 1425 1430 1435 1440 1445 1450 1455 1460 1465 1470 1475 1480 1485 1490 1495 1500 1505 1510 1515 1520 1525 1530 1535 1540 1545 1550 1555 1560 1565 1570 1575 1580 1585 1590 1595 1600 1605 1610 1615 1620 1625 1630 1635 1640 1645 1650 1655 1660 1665 1670 1675 1680 1685 1690 1695 1700 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000 2005 2010 2015 2020 2025 2030 2035 2040 2045 2050 2055 2060 2065 2070 2075 2080 2085 2090 2095 2100 2105 2110 2115 2120 2125 2130 2135 2140 2145 2150 2155 2160 2165 2170 2175 2180 2185 2190 2195 2200 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500 2505 2510 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995 3000 3005 3010 3015 3020 3025 3030 3035 3040 3045 3050 3055 3060 3065 3070 3075 3080 3085 3090 3095 3100 3105 3110 3115 3120 3125 3130 3135 3140 3145 3150 3155 3160 3165 3170 3175 3180 3185 3190 3195 3200 3205 3210 3215 3220 3225 3230 3235 3240 3245 3250 3255 3260 3265 3270 3275 3280 3285 3290 3295 3300 3305 3310 3315 3320 3325 3330 3335 3340 3345 3350 3355 3360 3365 3370 3375 3380 3385 3390 3395 3400 3405 3410 3415 3420 3425 3430 3435 3440 3445 3450 3455 3460 3465 3470 3475 3480 3485 3490 3495 3500 3505 3510 3515 3520 3525 3530 3535 3540 3545 3550 3555 3560 3565 3570 3575 3580 3585 3590 3595 3600 3605 3610 3615 3620 3625 3630 3635 3640 3645 3650 3655 3660 3665 3670 3675 3680 3685 3690 3695 3700 3705 3710 3715 3720 3725 3730 3735 3740 3745 3750 3755 3760 3765 3770 3775 3780 3785 3790 3795 3800 3805 3810 3815 3820 3825 3830 3835 3840 3845 3850 3855 3860 3865 3870 3875 3880 3885 3890 3895 3900 3905 3910 3915 3920 3925 3930 3935 3940 3945 3950 3955 3960 3965 3970 3975 3980 3985 3990 3995 4000 4005 4010 4015 4020 4025 4030 4035 4040 4045 4050 4055 4060 4065 4070 4075 4080 4085 4090 4095 4100 4105 4110 4115 4120 4125 4130 4135 4140 4145 4150 4155 4160 4165 4170 4175 4180 4185 4190 4195 4200 4205 4210 4215 4220 4225 4230 4235 4240 4245 4250 4255 4260 4265 4270 4275 4280 4285 4290 4295 4300 4305 4310 4315 4320 4325 4330 4335 4340 4345 4350 4355 4360 4365 4370 4375 4380 4385 4390 4395 4400 4405 4410 4415 4420 4425 4430 4435 4440 4445 4450 4455 4460 4465 4470 4475 4480 4485 4490 4495 4500 4505 4510 4515 4520 4525 4530 4535 4540 4545 4550 4555 4560 4565 4570 4575 4580 4585 4590 4595 4600 4605 4610 4615 4620 4625 4630 4635 4640 4645 4650 4655 4660 4665 4670 4675 4680 4685 4690 4695 4700 4705 4710 4715 4720 4725 4730 4735 4740 4745 4750 4755 4760 4765 4770 4775 4780 4785 4790 4795 4800 4805 4810 4815 4820 4825 4830 4835 4840 4845 4850 4855 4860 4865 4870 4875 4880 4885 4890 4895 4900 4905 4910 4915 4920 4925 4930 4935 4940 4945 4950 4955 4960 4965 4970 4975 4980 4985 4990 4995 5000 5005 5010 5015 5020 5025 5030 5035 5040 5045 5050 5055 5060 5065 5070 5075 5080 5085 5090 5095 5100 5105 5110 5115 5120 5125 5130 5135 5140 5145 5150 5155 5160 5165 5170 5175 5180 5185 5190 5195 5200 5205 5210 5215 5220 5225 5230 5235 5240 5245 5250 5255 5260 5265 5270 5275 5280 5285 5290 5295 5300 5305 5310 5315 5320 5325 5330 5335 5340 5345 5350 5355 5360 5365 5370 5375 5380 5385 5390 5395 5400 5405 5410 5415 5420 5425 5430 5435 5440 5445 5450 5455 5460 5465 5470 5475 5480 5485 5490 5495 5500 5505 5510 5515 5520 5525 5530 5535 5540 5545 5550 5555 5560 5565 5570 5575 5580 5585 5590 5595 5600 5605 5610 5615 5620 5625 5630 5635 5640 5645 5650 5655 5660 5665 5670 5675 5680 5685 5690 5695 5700 5705 5710 5715 5720 5725 5730 5735 5740 5745 5750 5755 5760 5765 5770 5775 5780 5785 5790 5795 5800 5805 5810 5815 5820 5825 5830 5835 5840 5845 5850 5855 5860 5865 5870 5875 5880 5885 5890 5895 5900 5905 5910 5915 5920 5925 5930 5935 5940 5945 5950 5955 5960 5965 5970 5975 5980 5985 5990 5995 6000 6005 6010 6015 6020 6025 6030 6035 6040 6045 6050 6055 6060 6065 6070 6075 6080 6085 6090 6095 6100 6105 6110 6115 6120 6125 6130 6135 6140 6145 6150 6155 6160 6165 6170 6175 6180 6185 6190 6195 6200 6205 6210 6215 6220 6225 6230 6235 6240 6245 6250 6255 6260 6265 6270 6275 6280 6285 6290 6295 6300 6305 6310 6315 6320 6325 6330 6335 6340 6345 6350 6355 6360 6365 6370 6375 6380 6385 6390 6395 6400 6405 6410 6415 6420 6425 6430 6435 6440 6445 6450 6455 6460 6465 6470 6475 6480 6485 6490 6495 6500 6505 6510 6515 6520 6525 6530 6535 6540 6545 6550 6555 6560 6565 6570 6575 6580 6585 6590 6595 6600 6605 6610 6615 6620 6625 6630 6635 6640 6645 6650 6655 6660 6665 6670 6675 6680 6685 6690 6695 6700 6705 6710 6715 6720 6725 6730 6735 6740 6745 6750 6755 6760 6765 6770 6775 6780 6785 6790 6795 6800 6805 6810 6815 6820 6825 6830 6835 6840 6845 6850 6855 6860 6865 6870 6875 6880 6885 6890 6895 6900 6905 6910 6915 6920 6925 6930 6935 6940 6945 6950 6955 6960 6965 6970 6975 6980 6985 6990 6995 7000 7005 7010 7015 7020 7025 7030 7035 7040 7045 7050 7055 7060 7065 7070 7075 7080 7085 7090 7095 7100 7105 7110 7115 7120 7125 7130 7135 7140 7145 7150 7155 7160 7165 7170 7175 7180 7185 7190 7195 7200 7205 7210 7215 7220 7225 7230 7235 7240 7245 7250 7255 7260 7265 7270 7275 7280 7285 7290 7295 7300 7305 7310 7315 7320 7325 7330 7335 7340 7345 7350 7355 7360 7365 7370 7375 7380 7385 7390 7395 7400 7405 7410 7415 7420 7425 7430 7435 7440 7445 7450 7455 7460 7465 7470 7475 7480 7485 7490 7495 7500 7505 7510 7515 7520 7525 7530 7535 7540 7545 7550 7555 7560 7565 7570 7575 7580 7585 7590 7595 7600 7605 7610 7615 7620 7625 7630 7635 7640 7645 7650 7655 7660 7665 7670 7675 7680 7685 7690 7695 7700 7705 7710 7715 7720 7725 7730 7735 7740 7745 7750 7755 7760 7765 7770 7775 7780 7785 7790 7795 7800 7805 7810 7815 7820 7825 7830 7835 7840 7845 7850 7855 7860 7865 7870 7875 7880 7885 7890 7895 7900 7905 7910 7915 7920 7925 7930 7935 7940 7945 7950 7955 7960 7965 7970 7975 7980 7985 7990 7995 8000 8005 8010 8015 8020 8025 8030 8035 8040 8045 8050 8055 8060 8065 8070 8075 8080 8085 8090 8095 8100 8105 8110 8115 8120 8125 8130 8135 8140 8145 8150 8155 8160 8165 8170 8175 8180 8185 8190 8195 8200 8205 8210 8215 8220 8225 8230 8235 8240 8245 8250 8255 8260 8265 8270 8275 8280 8285 8290 8295 8300 8305 8310 8315 8320 8325 8330 8335 8340 8345 8350 8355 8360 8365 8370 8375 8380 8385 8390 8395 8400 8405 8410 8415 8420 8425 8430 8435 8440 8445 8450 8455 8460 8465 8470 8475 8480 8485 8490 8495 8500 8505 8510 8515 8520 8525 8530 8535 8540 8545 8550 8555 8560 8565 8570 8575 8580 8585 8590 8595 8600 8605 8610 8615 8620 8625 8630 8635 8640 8645 8650 8655 8660 8665 8670 8675 8680 8685 8690 8695 8700 8705 8710 8715 8720 8725 8730 8735 8740 8745 8750 8755 8760 8765 8770 8775 8780 8785 8790 8795 8800 8805 8810 8815 8820 8825 8830 8835 8840 8845 8850 8855 8860 8865 8870 8875 8880 8885 8890 8895 8900 8905 8910 8915 8920 8925 8930 8935 8940 8945 8950 8955 8960 8965 8970 8975 8980 8985 8990 8995 9000 9005 9010 9015 9020 9025 9030 9035 9040 9045 9050 9055 9060 9065 9070 9075 9080 9085 9090 9095 9100 9105 9110 9115 9120 9125 9130 9135 9140 9145 9150 9155 9160 9165 9170 9175 9180 9185 9190 9195 9200 9205 9210 9215 9220 9225 9230 9235 9240 9245 9250 9255 9260 9265 9270 9275 9280 9285 9290 9295 9300 9305 9310 9315 9320 9325 9330 9335 9340 9345 9350 9355 9360 9365 9370 9375 9380 9385 9390 9395 9400 9405 9410 9415 9420 9425 9430 9435 9440 9445 9450 9455 9460 9465 9470 9475 9480 9485 9490 9495 9500 9505 9510 9515 9520 9525 9530 9535 9540 9545 9550 9555 9560 9565 9570 9575 9580 9585 9590 9595 9600 9605 9610 9615 9620 9625 9630 9635 9640 9645 9650 9655 9660 9665 9670 9675 9680 9685 9690 9695 9700 9705 9710 9715 9720 9725 9730 9735 9740 9745 9750 9755 9760 9765 9770 9775 9780 9785 9790 9795 9800 9805 9810 9815 9820 9825 9830 98

[0098] De preferencia, la elevada resolución se define por una codificación de tipo lineal de líneas espectrales tales como líneas MDCT, mientras que la segunda resolución o baja resolución es definida, por ejemplo, calculando solo un único valor espectral por banda de factor de escala, donde una banda de factor de escala cubre varias líneas de frecuencia. De esta manera, la segunda baja resolución es, con respecto a su resolución espectral, mucho menor que la primera o alta resolución definida por la codificación de tipo lineal típicamente aplicada por el codificador central tal como un codificador central AAC o USAC.

[0099] Con respecto al cálculo del factor de escala o la energía, la situación se ilustra en la Fig. 3b. Debido al hecho de que el codificador es un codificador central y debido al hecho de que puede haber, aunque no necesariamente, componentes del primer conjunto de porciones espectrales en cada banda, el codificador central calcula un factor de escala para cada banda no solo en el intervalo central por debajo de la frecuencia inicial de IGF

309, sino también por encima de la frecuencia inicial de IGF hasta la frecuencia máxima $f_{IGFstop}$ que es menor o igual a la mitad de la frecuencia de muestreo, es decir, $fs/2$. De esta manera, las porciones tonales codificadas 302, 304, 305, 306, 307 de la Fig. 3a y, en esta realización junto con los factores de escala SCB1 a SCB7 corresponden a los datos espectrales de alta resolución. Los datos espectrales de baja resolución se calculan a partir de la frecuencia inicial de IGF y corresponden a la información de los valores de energía E_1, E_2, E_3, E_4 , que se transmiten junto con los factores de escala SF4 a SF7.

[0100] En particular, cuando el codificador central está bajo una condición de baja tasa de bits, se puede aplicar además una operación adicional de relleno de ruido en la banda central, es decir, inferior en frecuencia que la frecuencia inicial de IGF, es decir, en bandas de factor de escala SCB1 a SCB3. En el relleno de ruido, existen varias líneas espectrales adyacentes que fueron cuantificadas a cero. Del lado del decodificador, estos valores espectrales cuantificados a cero se resintetizan y los valores espectrales resintetizados se ajustan en su magnitud mediante el uso de una energía de relleno de ruido como NF2 ilustrado en 308 en la Fig. 3b. La energía de relleno de ruido, que se puede dar en términos absolutos o en términos relativos particularmente con respecto al factor de escala como en USAC corresponde a la energía del conjunto de valores espectrales cuantificados a cero. Estas líneas espectrales de relleno de ruido también se pueden considerar que son un tercer conjunto de terceras porciones espectrales que son regeneradas por síntesis directa de relleno de ruido sin ninguna operación de IGF que se basa en la regeneración de frecuencia mediante el uso de mosaicos de frecuencia de otras frecuencias para reconstruir mosaicos mediante el uso de valores espectrales de un intervalo de fuente y la información de energía E_1, E_2, E_3, E_4 .

[0101] De preferencia, las bandas para las que se calcula la información de energía coinciden con las bandas de factor de escala. En otras realizaciones, se aplica un valor de información de energía agrupado de modo que, por ejemplo, para bandas de factores de escala 4 y 5, solo se transmite un único valor de información de energía, pero incluso en esta realización, los bordes de las bandas de reconstrucción agrupadas coinciden con los bordes de las bandas del factor de escala. Si se aplican diferentes separaciones de bandas, entonces se pueden aplicar determinados recálculos o cálculos de sincronización y esto puede tener sentido dependiendo de cierta implementación.

[0102] De preferencia, el codificador del dominio espectral 106 de la Fig. 1a es un codificador manejado psicoacústicamente ilustrado en la Fig. 4a. Típicamente, como se ilustra, por ejemplo, en MPEG2/4 AAC estándar o MPEG1/2, capa 3 estándar, la señal de audio por codificar después de ser transformada en el intervalo espectral (401 en la Fig. 4a) se dirige a un calculador de factores de escala 400. El calculador de factores de escala es controlado por un modelo psicoacústico que adicionalmente recibe la señal de audio que se va a cuantificar o que recibe como en MPEG1/2 capa 3 o MPEG AAC estándar, una representación espectral compleja de la señal de audio. El modelo psicoacústico calcula, para cada banda de factor de escala, un factor de escala que representa el umbral psicoacústico. Adicionalmente, los factores de escala se ajustan entonces, por cooperación de los bucles de iteración internos y externos bien conocidos o por cualquier otro procedimiento de codificación apropiado de modo que se satisfacen determinadas condiciones de tasas de bits. A continuación, los valores espectrales que se van a cuantificar, por un lado, y los factores de escala calculados, por otro lado, se introducen en un procesador cuantificador 404. En la operación directa del codificador de audio, los valores espectrales por cuantificar son ponderados por los factores de escala y los valores espectrales ponderados se introducen a continuación en un cuantificador fijo que tiene típicamente una funcionalidad de compresión a intervalos de amplitud superiores. A continuación, en la salida del procesador cuantificador, existen índices de cuantificación que se dirigen a un codificador de entropía que tiene típicamente una codificación específica y muy eficaz para un conjunto de índices de cuantificación cero para valores de frecuencia adyacentes o, como se denomina también en la técnica, una "ejecución" de valores cero.

[0103] En el codificador de audio de la Fig. 1a, sin embargo, el procesador cuantificador recibe típicamente información acerca de las segundas porciones espectrales del analizador espectral. De esta manera, el procesador cuantificador 404 asegura que, en la salida del procesador cuantificador 404, las segundas porciones espectrales como se identifican por el analizador espectral 102 son cero o tienen una representación reconocida por un codificador o un decodificador como una representación cero que puede codificarse de forma muy eficaz, de modo

que existen "ejecuciones" de valores cero en el espectro.

[0104] La Fig. 4b ilustra una implementación del procesador cuantificador. Los valores espetrales de MDCT pueden introducirse en un conjunto de bloque cero 410. A continuación, las segundas porciones espetrales ya se fijan en cero antes de realizar una ponderación por los factores de escala en bloque 412. En una implementación adicional, no se proporciona el bloque 410, pero la fijación a cooperación cero se lleva a cabo en bloque 418 después del bloque de ponderación 412. En otra implementación más, la fijación a operación cero también se puede llevar a cabo en una fijación a bloque cero 422 después de una cuantificación en el bloque de cuantificador 420. En esta implementación, los bloques 410 y 418 no estarían presentes. En general, se proporciona al menos uno de los bloques 410, 418, 422 dependiendo de la implementación específica.

[0105] A continuación, en la salida del bloque 422, se obtiene un espectro cuantificado correspondiente al que se ilustra en la Fig. 3a. Este espectro cuantificado se introduce a continuación en un codificador de entropía como 232 en la Fig. 2b que puede ser un codificador de Huffman o un codificador aritmético como se define, por ejemplo, en el estándar USAC.

[0106] La fijación a bloques cero 410, 418, 422, que se proporcionan alternativamente entre sí o en paralelo, es controlada por el analizador espectral 424. El analizador espectral comprende preferentemente cualquier implementación de un detector de tonalidad bien conocido o comprende cualquier tipo diferente de detectores operativos para separar un espectro en componentes que se van a codificar con alta resolución y componentes que se van a codificar con una baja resolución. Otros de tales algoritmos implementados en el analizador espectral pueden ser un detector de actividad de voz, un detector de ruidos, un detector del habla o cualquier detector que decide según la información espectral o los metadatos asociados acerca de los requisitos de resolución para diferentes porciones espetrales.

[0107] La Fig. 5a ilustra una implementación preferida del conversor de espetro-tiempo 100 de la Fig. 1a como se implementa, por ejemplo, en AAC o USAC. El conversor de espetro-tiempo 100 comprende un formador de ventanas 502 controlado por un detector de transitorios 504. Cuando el detector de transitorios 504 detecta un transitorio, entonces se señaliza un cambio de ventanas largas a ventanas cortas al formador de ventanas. El formador de ventanas 502 calcula a continuación, para superponer los bloques, tramas de ventana, donde cada trama de ventana tiene típicamente dos valores N tales como los valores 2048. A continuación, se lleva a cabo una transformación dentro de un transformador de bloques 506 y este transformador de bloques proporciona típicamente de modo adicional una pérdida, de modo que una pérdida/transformación combinada se lleva a cabo para obtener una trama espectral con valores N tales como los valores espetrales de MDCT. De esta manera, para una operación de ventana larga, la trama en la entrada del bloque 506 comprende dos valores N tales como los valores 2048 y una trama espectral, entonces tiene valores 1024. A continuación, sin embargo, se lleva a cabo una conmutación a bloques cortos cuando se llevan a cabo ocho bloques cortos, donde cada bloque corto tiene 1/8 valores de dominio temporal de ventana en comparación con una ventana larga y cada bloque espectral tiene 1/8 valores espetrales en comparación con un bloque largo. De esta manera, cuando esta pérdida se combina con una operación de superposición al 50 % del formador de ventanas, el espetro es una versión críticamente muestreada de la señal de audio del dominio de tiempo 99.

[0108] Posteriormente, se hace referencia a la Fig. 5b que ilustra una implementación específica del regenerador de frecuencia 116 y el conversor de espetro-tiempo 118 de la Fig. 1b o de la operación combinada de bloques 208, 212 de la Fig. 2a. En la Fig. 5b, se considera una banda de reconstrucción específica como una banda del factor de escala 6 de la Fig. 3a. La primera porción espectral en esta banda de reconstrucción, es decir, la primera porción espectral 306 de la Fig. 3a se introduce en el bloque de formador/ajustador de tramas 510. Por otra parte, una segunda porción espectral reconstruida para la banda del factor de escala 6 también se introduce en el formador/ajustador de la trama 510. Por otra parte, también se introduce información de energía como E3 de la Fig. 3b para una banda de factor de escala 6 en el bloque 510. La segunda porción espectral reconstruida en la banda de reconstrucción ya se generó por relleno de mosaico de frecuencia mediante el uso de un intervalo de fuente y la banda de reconstrucción se corresponde entonces con el intervalo de destino. Ahora, un ajuste de energía de la trama se lleva a cabo para obtener después finalmente la trama reconstruida completa que tiene los valores N, por ejemplo, como se obtienen en la salida del combinador 208 de la Fig. 2a. A continuación, en el bloque 512, se lleva a cabo una transformación/interpolación de bloques inversa para obtener 248 valores de dominio de tiempo para, por ejemplo, los 124 valores espetrales en la entrada del bloque 512. A continuación, se lleva a cabo una operación de ventana de síntesis en el bloque 514 que es controlada de nuevo por una indicación de ventana larga/ventana corta transmitida como información lateral en la señal de audio codificada. A continuación, en el bloque 516, se lleva a cabo una operación de superposición/adición con una trama de tiempo anterior. De preferencia, MDCT aplica un a un 50 % de superposición, de modo que, para cada nueva trama de tiempo de valores 2N, finalmente se producen valores de dominio de tiempo N. Se prefiere en gran medida una superposición del 50 % debido al hecho de que proporciona un muestreo crítico y un cruce continuo de una trama a la trama siguiente debido a la operación de superposición/adición en el bloque 516.

[0109] Como se ilustra en 301 en la Fig. 3a, se puede aplicar adicionalmente una operación de relleno de

ruido no solo por debajo de la frecuencia inicial de IGF, sino también por encima de la frecuencia inicial de IGF como para la banda de reconstrucción contemplada que coincide con la banda de factor de escala 6 de la Fig. 3a. A continuación, también se pueden introducir valores espectrales de relleno de ruido en el formador/ajustador de trama 510 y el ajuste de los valores espectrales de relleno de ruido también se puede aplicar dentro de este bloque o los 5 valores espectrales de relleno de ruido se pueden ajustar ya mediante el uso de la energía de relleno de ruido antes de ser introducidos en el formador/ajustador de trama 510.

[0110] De preferencia, se puede aplicar una operación de IGF, es decir, una operación de relleno de mosaico de frecuencia mediante el uso de valores espectrales de otras porciones en el espectro completo. De esta manera, 10 una operación de relleno de mosaico espectral no solo se puede aplicar en la banda alta por encima de una frecuencia inicial de IGF, sino que también se puede aplicar en la banda baja. Por otra parte, el relleno de ruido sin relleno de mosaico de frecuencia también se puede aplicar no solo por debajo de la frecuencia inicial de IGF sino también por encima de la frecuencia inicial de IGF. Sin embargo, se encontró que la alta calidad y la codificación de audio muy eficaz se pueden obtener cuando la operación de relleno de ruido está limitada al intervalo de frecuencias 15 por debajo de frecuencia inicial de IGF y cuando la operación de relleno de mosaico de frecuencia está restringida al intervalo de frecuencias por encima de la frecuencia inicial de IGF como se ilustra en la Fig. 3a.

[0111] De preferencia, los mosaicos de destino (TT) (que tienen frecuencias mayores que la frecuencia inicial de IGF) están unidos a bordes de banda de factores de escala del codificador de tasa completa. Los mosaicos de 20 fuente (ST), de los que se extrae la información, es decir, para frecuencias inferiores a la frecuencia inicial de IGF no se unen por bordes de banda de factores de escala. El tamaño de ST debería corresponder con el tamaño de los TT asociados. Esto se ilustra mediante el uso del siguiente ejemplo. TT[0] tiene una longitud de 10 MDCT Bins. Esto corresponde exactamente a la longitud de dos SCB posteriores (como 4 + 6). A continuación, todos los ST posibles 25 que se van a correlacionar con TT[0], tienen también una longitud de 10 bins. Un segundo mosaico de destino TT[1] que es adyacente a TT[0] tiene una longitud de 15 bins I (SCB que tiene una longitud de 7 + 8). A continuación, el ST para ello tiene una longitud de 15 bins más que 10 bins para TT[0].

[0112] Si surgiera el caso de que no se pudiera hallar un TT para un ST con la longitud del mosaico de destino (cuando, por ejemplo, la longitud de TT es mayor que el intervalo de fuente disponible), entonces no se 30 calcula una correlación y el intervalo de la fuente se copia una cantidad de veces en este TT (la copia se realiza una tras otra, de modo que una línea de frecuencia para la frecuencia más baja de la segunda copia sigue -en frecuencia- a la línea de frecuencia para la máxima frecuencia de la primera copia), hasta llenar por completo el mosaico de destino TT.

[0113] Posteriormente, se hace referencia a la Fig. 5c que ilustra otra realización preferida del regenerador de 35 frecuencia 116 de la Fig. 1b o el bloque de IGF 202 de la Fig. 2a. El bloque 522 es un generador de mosaicos de frecuencia que recibe no solo una ID de banda de destino, sino que recibe adicionalmente una ID de banda de fuente. Como ejemplo, se determinó en el lado del codificador que la banda del factor de escala 3 de la Fig. 3a es muy apropiado para reconstruir la banda del factor de escala 7. De esta manera, la ID de la banda de fuente sería 2 40 y la ID de la banda de destino sería 7. En base a esta información, el generador de mosaicos de frecuencia 522 aplica una copia o una operación de relleno de mosaicos de armónicos o cualquier otra operación de relleno de mosaicos para generar la segunda porción en bruto de componentes espectrales 523. La segunda porción en bruto de componentes espectrales tiene una resolución de frecuencia idéntica a la resolución de frecuencia incluida en el primer conjunto de primeras porciones espectrales.

[0114] A continuación, la primera porción espectral de la banda de reconstrucción tal como 307 de la Fig. 3a se introduce en un formador de tramas 524 y la segunda porción en bruto 523 también se introduce en el formador de tramas 524. A continuación, la trama reconstruida se ajusta por medio del ajustador 526 mediante el uso de un factor de aumento para la banda de reconstrucción calculada por el calculador de factor de aumento 528. De modo 50 importante, sin embargo, la primera porción espectral en la trama no se ve influida por el ajustador 526, sino que solo la segunda porción en bruto para la trama de reconstrucción se ve influida por el ajustador 526. Para este fin, el calculador del factor de aumento 528 analiza la banda de fuente o la segunda porción en bruto 523 y además analiza la primera porción espectral en la banda de reconstrucción para hallar finalmente el factor de aumento 527 correcto, de modo que la energía de la trama ajustada producida por el ajustador 526 tenga la energía E4 cuando se 55 contempla una banda de factor de escala 7.

[0115] En este contexto, es muy importante evaluar la alta precisión de reconstrucción de frecuencia de la presente invención en comparación con HE-AAC. Esto se explica con respecto a la banda del factor de escala 7 en la Fig. 3a. Se asume que un codificador de la técnica anterior tal como se ilustra en la Fig. 13a detectaría la porción 60 espectral 307 que se va a codificar con una alta resolución como una "falta de armónicos". A continuación, la energía de este componente espectral será transmitida junto con una información envolvente espectral para la banda de reconstrucción como banda del factor de escala 7 al decodificador. A continuación, el decodificador recrearía el armónico que falta. Sin embargo, el valor espectral, donde el armónico que falta 307 se reconstruiría por el decodificador de la técnica anterior de la Fig. 13b estaría en la mitad de la banda 7 a una frecuencia indicada por la 65 frecuencia de reconstrucción 390. De esta manera, la presente invención evita un error de frecuencia 391 que sería

introducido por el decodificador de la técnica anterior de la Fig. 13d.

[0116] En una implementación, el analizador espectral se implementa también para calcular similitudes entre primeras porciones espectrales y segundas porciones espectrales y para determinar, en base a similitudes calculadas, para una segunda porción espectral en un intervalo de reconstrucción una primera porción espectral que concuerde con la segunda porción espectral tanto como sea posible. A continuación, en esta implementación de intervalo de fuente/intervalo de destino variable, el codificador paramétrico introducirá adicionalmente en la segunda representación codificada una información concordante que indique para cada intervalo de destino un intervalo de fuente concordante. Del lado del decodificador, esta información se usaría entonces por un generador de mosaico de frecuencia 522 de la Fig. 5c que ilustra una generación de una segunda porción en bruto 523 en base a una ID de banda de fuente y una ID de banda de destino. Por otra parte, como se ilustra en la Fig. 3a, el analizador espectral está configurado para analizar la representación espectral hasta una frecuencia de análisis máxima que es solo una pequeña cantidad por debajo de la mitad de la frecuencia de muestreo y preferentemente al menos un cuarto de la frecuencia de muestreo o típicamente superior.

15

[0117] Como se ilustra, el codificador opera sin muestrear hacia abajo y el decodificador opera sin muestrear hacia arriba. En otras palabras, el codificador de audio del dominio espectral se configura para generar una representación espectral que tiene una frecuencia Nyquist definida por la tasa de muestreo de la señal de audio de entrada original.

20

[0118] Por otra parte, como se ilustra en la Fig. 3a, el analizador espectral está configurado para analizar la representación espectral que comienza con una frecuencia inicial de relleno de intervalo y que termina con una frecuencia máxima representada por una frecuencia máxima incluida en la representación espectral, donde una porción espectral que se extiende desde una frecuencia mínima hasta la frecuencia inicial de relleno de intervalo pertenece al primer conjunto de porciones espectrales y donde otra porción espectral como 304, 305, 306, 307 que tiene valores de frecuencia por encima de la frecuencia del relleno de intervalo está incluida adicionalmente en el primer conjunto de primeras porciones espectrales.

[0119] Como se destaca, el decodificador de audio del dominio espectral 112 está configurado de modo que una frecuencia máxima representada por un valor espectral en la primera representación decodificada sea igual a una frecuencia máxima incluida en la representación temporal que tiene la tasa de muestreo donde el valor espectral para la frecuencia máxima en el primer conjunto de primeras porciones espectrales es cero o diferente de cero. De cualquier modo, para esta frecuencia máxima en el primer conjunto de componentes espectrales, existe un factor de escala para la banda del factor de escala, que se genera y transmite independientemente de si todos los valores espectrales en esta banda del factor de escala se fijan en cero o no, tal como se trata en el contexto de las Figs. 3a y 3b.

[0120] Las realizaciones del sistema de la invención mejoran las estrategias del estado de la técnica y proporcionan de ese modo una alta eficacia de compresión, ninguna o solo una pequeña molestia perceptual y un ancho de banda de audio completo incluso para bajas tasas de bits.

[0121] El sistema general consiste en

- codificación central de banda completa
- 45 • relleno de intervalo inteligente (relleno de mosaico o relleno de ruido)
- partes tonales dispersas en el núcleo seleccionadas por máscara tonal
- codificación de pares estéreo conjuntos para banda completa, incluyendo relleno de mosaicos
- TNS en mosaicos
- blanqueo espectral en el intervalo de IGF

50

[0122] Una primera etapa hacia un sistema más eficaz consiste en eliminar la necesidad de transformar datos espirales en un segundo dominio de transformada de un codificador a uno central. Como la mayoría de codecs de audio, como, por ejemplo, AAC, que usa la MDCT como transformada básica, también es de utilidad llevar a cabo el BWE en el dominio de MDCT. Un segundo requisito para el sistema de BWE sería la necesidad de preservar la rejilla tonal preservando incluso los componentes tonales HF y la calidad del audio codificado es así superior a los sistemas existentes. Para cuidar de ambos requisitos anteriormente mencionados, se ha propuesto un sistema llamado Intelligent Gap Filling (IGF). La Fig. 2b muestra el diagrama de bloques del sistema propuesto en el lado del codificador y la Fig. 2a muestra el sistema del lado del decodificador.

[0123] Posteriormente, se describe una trama de post-procesamiento con respecto a la Fig. 13a y la Fig. 13b con el fin de ilustrar que la presente invención también se puede implementar en el reconstructor de alta frecuencia 1330 en esta realización de post-procesamiento.

[0124] La Fig. 13a ilustra un diagrama esquemático de un codificador de audio para una tecnología de extensión de ancho de banda, por ejemplo, como se usa en la codificación de audio avanzada de alta eficacia (*High*

Efficiency Advanced Audio Coding, HE-AAC). Una señal de audio en la línea 1300 se introduce en un sistema de filtro que consta de un paso bajo 1302 y un paso alto 1304. La salida de la señal por el filtro de paso alto 1304 se introduce en un extractor/codificador paramétrico 1306. El extractor/codificador paramétrico 1306 está configurado para calcular y codificar parámetros tales como un parámetro de envolvente espectral, un parámetro de adición de ruido, un parámetro de armónicos que faltan o un parámetro de filtro inverso, por ejemplo. Estos parámetros extraídos se introducen en un multiplexador de corriente de bits 1308. La señal de salida de paso bajo se introduce en un procesador que comprende típicamente la funcionalidad de un muestreador descendente 1310 y un codificador central 1312. El paso bajo 1302 restringe el ancho de banda que se va a codificar a un ancho de banda significativamente menor al que ocurre en la señal de audio de entrada original en la línea 1300. Esto proporciona un aumento de codificación significativo debido al hecho de que todas las funcionalidades que se producen en el codificador central solo tienen que operar en una señal con un ancho de banda reducido. Cuando, por ejemplo, el ancho de banda de la señal de audio en la línea 1300 es de 20 kHz y cuando el filtro de paso bajo 1302 tiene ilustrativamente un ancho de banda de 4 kHz, con el fin de satisfacer el teorema de muestreo, teóricamente es suficiente con que la señal posterior al muestreador descendente tenga una frecuencia de muestreo de 8 kHz, que es una reducción sustancial de la tasa de muestreo requerida para la señal de audio 1300 que tiene que ser de al menos 40 kHz.

[0125] La Fig. 13b ilustra un diagrama esquemático de un decodificador de la extensión de ancho de banda correspondiente. El decodificador comprende un multiplexador de corrientes de bits 1320. El demultiplexador de corrientes de bits 1320 extrae una señal de entrada para un decodificador central 1322 y una señal de entrada para un decodificador paramétrico 1324. Una señal de salida del decodificador central tiene, en el ejemplo anterior, una tasa de muestreo de 8 kHz y, por ello, un ancho de banda de 4 kHz, mientras que, para una reconstrucción completa del ancho de banda, la señal de salida de un reconstructor de alta frecuencia 1330 debe estar a 20 kHz que requiere una tasa de muestreo de al menos 40 kHz. A fin de hacer esto posible, se requiere un decodificador procesador que tiene la funcionalidad de un muestreador ascendente 1325 y un banco de filtros 1326. El reconstructor de alta frecuencia 1330 recibe entonces la señal de frecuencia baja analizada producida por el banco de filtros 1326 y reconstruye el intervalo de frecuencias definido por el filtro de paso alto 1304 de la Fig. 13a mediante el uso de la representación paramétrica de la banda de alta frecuencia. El reconstructor de alta frecuencia 1330 tiene varias funcionalidades tales como la regeneración del intervalo de frecuencia superior mediante el uso del intervalo de la fuente en el intervalo de frecuencia inferior, un ajuste de envolvente espectral, una funcionalidad de adición de ruido y una funcionalidad para introducir los armónicos que faltan en el intervalo de frecuencia superior y, si se aplica y se calcula en el codificador de la Fig. 13a, una operación de filtrado inversa para contar con el hecho de que el intervalo de mayor frecuencia no es típicamente tan tonal como el intervalo de menor frecuencia. En HE-AAC, los armónicos faltantes se resintetizan en el lado del decodificador y se colocan exactamente en la mitad de una banda de reconstrucción. Así, todas las líneas de armónicos que faltan que fueron determinados en cierta banda de reconstrucción no se colocan en los valores de frecuencia donde se ubicaron en la señal original. En vez de ello, esos armónicos que faltan se colocan en frecuencias en el centro de la banda determinada. De esta manera, cuando una línea de armónicos que faltan en la señal original se colocó muy cerca de un borde de la banda de reconstrucción en la señal original, el error en frecuencia introducido al colocar esta línea de armónicos que faltan en la señal reconstruida en el centro de la banda está cerca del 50 % de la banda de reconstrucción individual, para la que se generaron y transmitieron parámetros.

[0126] Por otra parte, incluso cuando los codificadores centrales de audio típicos operen en el dominio espectral, el decodificador central genera sin embargo una señal de dominio de tiempo que es convertida de nuevo en un dominio espectral por la funcionalidad del banco de filtros 1326. Esto introduce retrasos de procesamiento adicionales, puede introducir artefactos debido al procesamiento en tandem principalmente de la transformación del dominio espectral al dominio de frecuencia y de nuevo la transformación típicamente en un dominio de frecuencia diferente y, de hecho, esto también requiere una cantidad sustancial de complejidad de computación y, así, energía eléctrica, que es específicamente un problema cuando la tecnología de extensión de ancho de banda se aplica en dispositivos móviles tales como teléfonos móviles, tabletas u ordenadores portátiles, etc.

[0127] Aunque algunos aspectos se han descrito en el contexto de un aparato para codificar o decodificar, es evidente que estos aspectos también representan una descripción del correspondiente procedimiento, donde un bloque o dispositivo corresponden a una etapa de procedimiento o a un aspecto de una etapa de procedimiento. De manera análoga, los aspectos descritos en el contexto de una etapa de procedimiento representan también una descripción del correspondiente bloque o elemento o rasgo de un aparato correspondiente. Algunas de las etapas de procedimiento, o todas ellas, pueden ejecutarse mediante un aparato de hardware (o mediante el uso de) un aparato de hardware tal como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, algunas o más de las etapas de procedimiento más importantes se pueden ejecutar mediante tal aparato.

[0128] Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención se pueden implementar en hardware o en software. La implementación se puede llevar a cabo mediante el uso de un medio de almacenamiento no transitorio tal como un medio de almacenamiento digital, por ejemplo, un disquete, una unidad de disco duro (HDD), un DVD, un Blu-Ray, un CD, una memoria ROM, una PROM y una EEPROM, una EEPROM o

una memoria FLASH, que tengan señales de control electrónicamente legibles almacenadas en ellos, que cooperan (o son capaces de cooperar) con un sistema de ordenador programable de tal manera que el procedimiento respectivo sea llevado a cabo. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

5 [0129] Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema de ordenador programable, de tal manera que se lleve a cabo uno de los procedimientos descritos en esta invención.

10 [0130] En términos generales, las realizaciones de la presente invención se pueden implementar como un producto de programa informático con un código de programa, siendo el código de programa operativo para llevar a cabo uno de los procedimientos cuando el producto de programa informático es ejecutado en un ordenador. El código de programa se puede almacenar, por ejemplo, en un soporte legible por máquina.

15 [0131] Otras realizaciones comprenden el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por máquina.

20 [0132] En otras palabras, una realización del procedimiento inventivo es, por lo tanto, un programa informático que tiene un código de programa para llevar a cabo uno de los procedimientos descritos en esta invención, cuando el programa informático se ejecuta en un ordenador.

25 [0133] Una realización adicional del procedimiento inventivo es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en él, el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención. El soporte de datos, el medio de almacenamiento digital o el medio grabado son típicamente tangibles y/o no transitorios.

30 [0134] Una realización adicional del procedimiento inventivo es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención. El flujo de datos o la secuencia de señales, puede estar configurada, por ejemplo, para ser transferida por medio de una conexión de comunicación de datos, por ejemplo, a través de Internet.

35 [0135] Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador o un dispositivo lógico programable, configurados para, o adaptados para, llevar a cabo uno de los procedimientos descritos en esta invención.

35 [0136] Una realización adicional comprende un ordenador que tiene instalado en él el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención.

40 [0137] Una realización adicional según la invención comprende un aparato o un sistema configurados para transferir (por ejemplo, electrónicamente u ópticamente) un programa informático para llevar a cabo uno de los procedimientos descritos en esta invención, a un receptor. El receptor puede ser, por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o sistema puede comprender, por ejemplo, un servidor de archivos para transferir el programa informático al receptor.

45 [0138] En algunas realizaciones, puede utilizarse un dispositivo lógico programable (por ejemplo, una disposición de puerta programable en campo) para llevar a cabo algunas, o la totalidad de, las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, una disposición de puerta programable en campo puede cooperar con un microprocesador con el fin de llevar a cabo uno de los procedimientos descritos en esta invención. En términos generales, los procedimientos se llevan a cabo preferiblemente mediante cualquier aparato de hardware.

50 [0139] Las realizaciones descritas más arriba son meramente ilustrativas de los principios de la presente invención. Se da por entendido que modificaciones y variaciones en las disposiciones y los detalles descritos en esta invención serán evidentes para otras personas expertas en la técnica. Es el propósito, por lo tanto, estar limitada solamente por los alcances de las reivindicaciones de patente inminentes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones de esta invención.

REIVINDICACIONES

1. Aparato para generar una señal de audio mejorada a partir de una señal de audio de entrada (600), donde la señal de audio mejorada tiene valores espectrales para una región espectral de mejora, no estando los 5 valores espectrales para las regiones espectrales de mejora contenidos en la señal de entrada (600), comprendiendo el aparato:

un mapeador (602) para mapear (1002) una región espectral de la fuente de la señal de entrada (600) a una región de destino en la región espectral de mejora, donde los primeros valores de ruido en una región de relleno de ruido 10 (302) de la región espectral de fuente de la señal de audio de entrada (600) no se incluyen en la región de relleno de ruido (302) de la región espectral de fuente, donde, para la región de destino, existe una identificación de región de fuente, y

donde el mapeador (602) está configurado para seleccionar la región espectral de fuente mediante el uso de la identificación de la región de fuente y para mapear la región espectral de fuente seleccionada a la región de destino;

15 y
un rellenero de ruido (604) configurado para generar segundos valores de ruido para una región de ruido en la región de destino en la región espectral de mejora, donde los segundos valores de ruido se decorrelacionan desde los primeros valores de ruido en la región de relleno de ruido (302) de la región espectral de fuente de la señal de audio de entrada (600),

20 donde la señal de audio de entrada (600) es una señal codificada que comprende parámetros de relleno de ruido para la región espectral de fuente de la señal de audio de entrada (600), y
donde el rellenero de ruido (604) está configurado para

25 la identificación (900) de ruido en la región espectral de fuente de la señal de audio de entrada (600) mediante el uso de los parámetros de relleno de ruido para la región espectral de fuente de la señal de audio de entrada (600) o mediante la identificación de valores espectrales de cero en la región espectral de fuente,

30 la generación (1006) de los primeros valores de ruido en la región de relleno de ruido (302) en la región espectral de fuente de la señal de audio de entrada (600) mediante el uso de una información de energía de los parámetros de relleno de ruido leídos (1004) a partir de la señal codificada que comprende los parámetros de relleno de ruido para la región espectral de fuente de la señal de audio de entrada (600), identificando la información de energía una energía de los primeros valores de ruido en la región de relleno de ruido (302), posteriormente o simultáneamente a la etapa de generación (1006) de los primeros valores de ruido, inserción (1008) de valores aleatorios en posiciones en la región de destino identificados mediante el uso de los 35 parámetros de relleno de ruido junto con la identificación de región de fuente, y
el escalado (1010) de los valores aleatorios insertados en las posiciones en la región de destino para obtener los segundos valores de ruido.

2. Aparato según la reivindicación 1, que comprende, además:

40 un ajustador de envolvente para ajustar (1202) los segundos valores de ruido en la región espectral de mejora mediante el uso de información de envolvente espectral incluida en la señal de entrada (600) como información lateral.

3. Aparato según una de las reivindicaciones anteriores,
45

donde el rellenero de ruido (604) está configurado para identificar las posiciones de ruido mediante el uso de un vector de identificación (706) que tiene entradas para las posiciones espectrales en la región espectral de fuente únicamente, o que tiene entradas para las posiciones espectrales en la región espectral de fuente y en la región de destino

50 4. Aparato según la reivindicación 3, donde el rellenero de ruido (604) está configurado para
calcular (1100) la información de energía sobre los primeros valores de ruido indicados por el vector de identificación (706),
55 calcular (1102) una información de energía sobre los valores aleatorios insertados en las posiciones de la región de destino,
calcular (1104) un factor de ganancia para escalar los valores aleatorios insertados en las posiciones de la región de destino, y
aplicar (1106) el factor de ganancia a los valores aleatorios insertados en las posiciones de la región de destino.

60 5. Aparato según una de las reivindicaciones anteriores,
donde el mapeador (602) está configurado para llevar a cabo una operación de relleno de intervalo para generar la 65 región de destino, comprendiendo el aparato:

un decodificador de audio del dominio espectral (112) para generar una primera representación decodificada de un primer conjunto de primeras porciones espectrales, teniendo la representación decodificada una primera resolución espectral;

- 5 un decodificador paramétrico (114) para generar una segunda representación decodificada de un segundo conjunto de segundas porciones espectrales que tiene una segunda resolución espectral que es inferior a la primera resolución espectral;
- 10 un regenerador de frecuencia (116) para regenerar una segunda porción espectral reconstruida que tiene la primera resolución espectral mediante el uso de una primera porción espectral e información envolvente espectral para la segunda porción espectral; y
- 15 10 un conversor de espectro-tiempo (118) para convertir la primera representación decodificada en la segunda porción espectral reconstruida en una representación temporal,

donde el mapeador (602) y el rellenador de ruido (604) están al menos parcialmente incluidos en el regenerador de frecuencia (116).

- 15 6. Aparato según la reivindicación 5,

donde el decodificador de audio del dominio espectral (112) está configurado para emitir una secuencia de tramas decodificadas de valores espirales, siendo una trama decodificada la primera representación decodificada, donde la trama comprende valores espirales para el primer conjunto de porciones espirales e indicaciones cero para el segundo conjunto de segundas porciones espirales,

20 donde el aparato para decodificar comprende además un combinador (208) para combinar los valores espirales generados por el regenerador de frecuencia (116) para el segundo conjunto de segundas porciones espirales y valores espirales del primer conjunto de primeras porciones espirales en una banda de reconstrucción para obtener una trama espectral reconstruida que comprende valores espirales para el primer conjunto de las primeras porciones espirales y el segundo conjunto de las segundas porciones espirales; y

25 donde el conversor de espectro-tiempo (118) está configurado para convertir la trama espectral reconstruida en la representación temporal.

- 30 7. Aparato según una de las reivindicaciones anteriores, donde el rellenador de ruido (604) está configurado para ser controlado por un vector de control (PHI), donde el vector de control (PHI) se determina para tener un "1" valor para una posición espectral donde el relleno de ruido no se llevó a cabo, o donde el vector de control (PHI) se determina para que tenga un "1" valor para una posición espectral donde una línea espectral se identifica como una línea de ruido, y valor cero para una posición espectral donde una línea espectral no se identifica como una línea de ruido.

- 35 8. Aparato según la reivindicación 7, donde el rellenador de ruido (604) está configurado para:

40 calcular (1100) la información de energía sobre los primeros valores de ruido para la región de relleno de ruido (302),

45 calcular (1102) una información de energía sobre los valores aleatorios insertados en las posiciones de la región de destino,

50 calcular (1104) un factor de escala mediante el uso de la información de energía sobre los primeros valores de ruido y la información de energía sobre los valores aleatorios insertados, y

55 multiplicar (1106) los valores aleatorios insertados en las posiciones de la región de destino por el factor de escala en el escalado (1010) de los valores aleatorios insertados.

9. Procedimiento de generación de una señal de audio mejorada a partir de una señal de audio de entrada (600), donde la señal de audio mejorada tiene valores espirales para una región espectral de mejora, no estando los valores espirales para las regiones espirales de mejora contenidos en la señal de audio de entrada (600), que comprende:

55 el mapeo (602) una región espectral de la fuente de la señal de audio de entrada (600) a una región de destino en la región espectral de mejora, donde los primeros valores de ruido en una región de relleno de ruido (302) de la región espectral de fuente de la señal de audio de entrada (600) no se incluyen en la región de relleno de ruido (302) de la región espectral de fuente, donde, para la región de destino, existe una identificación de región de fuente, y donde el mapeo (602) comprende la selección de la región espectral de fuente mediante el uso de la identificación de la región de fuente y el mapeo de la región espectral de fuente seleccionada a la región de destino,

60 donde la señal de audio de entrada (600) es una señal codificada que comprende parámetros de relleno de ruido para la región espectral de fuente de la señal de audio de entrada (600);

65 la identificación (900) de ruido en la región espectral de fuente de la señal de audio de entrada (600) mediante el uso de los parámetros de relleno de ruido para la región espectral de fuente de la señal de audio de entrada (600) o mediante la identificación de valores espirales cero en la región espectral de fuente,

la generación (1006) de los primeros valores de ruido en la región de relleno de ruido (302) en la región espectral

de fuente de la señal de audio de entrada (600) mediante el uso de una información de energía de los parámetros de relleno de ruido leídos (1004) a partir de la señal codificada que comprende los parámetros de relleno de ruido para la región espectral de fuente de la señal de audio de entrada (600), identificando la información de energía una energía de los primeros valores de ruido en la región de relleno de ruido (302),

5 posteriormente o actualmente a la etapa de generación (1006) de los primeros valores de ruido, insertar (1008) valores aleatorios en posiciones en la región de destino identificada mediante el uso de los parámetros de relleno de ruido junto con la identificación de la región fuente, y el escalado (1010) de los valores aleatorios insertados en las posiciones en la región de destino para obtener los 10 segundos valores de ruido, donde los segundos valores de ruido se decorrelacionan a partir de los primeros valores de ruido en la región de relleno de ruido (302) de la región espectral de fuente.

10. Sistema para el procesamiento de una señal de audio, que comprende:

un codificador para generar una señal codificada a partir de la señal de audio; y

15 un aparato para generar una señal de audio mejorada a partir de una señal de audio de entrada (600) según cualquiera de las reivindicaciones 1 a 8, donde la señal codificada es sometida a un procesamiento (700) con el fin de generar la señal de audio de entrada (600) en el aparato para generar la señal de audio mejorada.

11. Procedimiento para el procesamiento de una señal de audio, que comprende:

20 generar una señal codificada a partir de la señal de audio; y

un procedimiento para generar una señal de audio mejorada a partir de una señal de audio de entrada (600) según la reivindicación 9,

25 donde la señal codificada está sometida a un procesamiento predefinido (700) para generar la señal de audio de entrada (600) en el procedimiento para generar la señal de audio mejorada.

12. Programa informático para llevar a cabo, cuando se ejecuta en un ordenador, el procedimiento según la reivindicación 9 o la reivindicación 11.

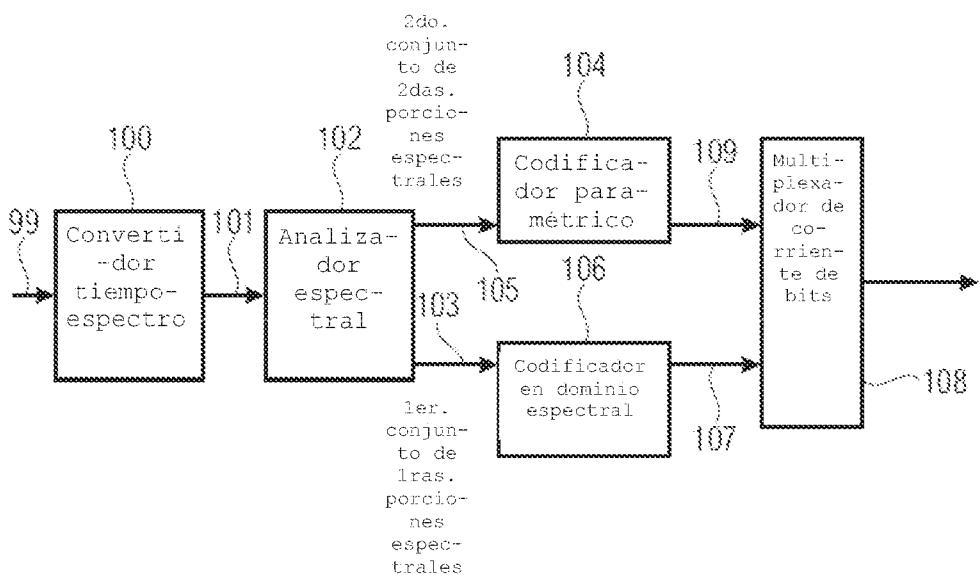


FIG 1A

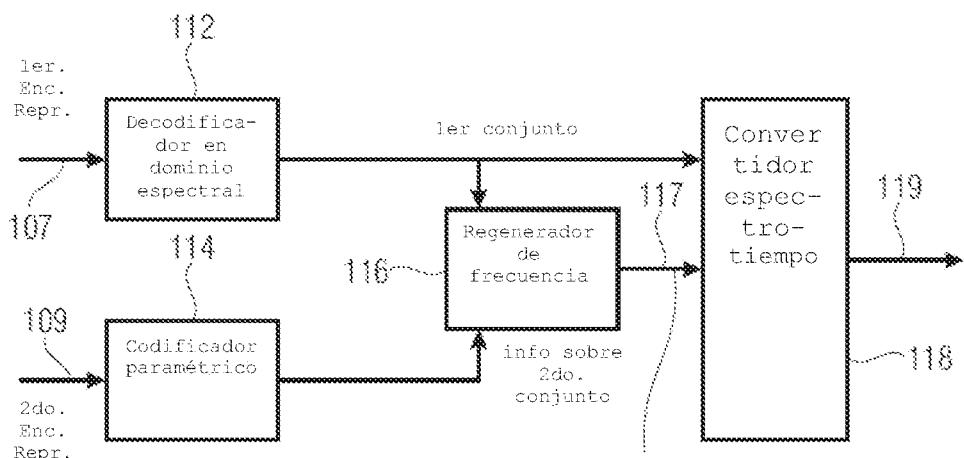


FIG 1B

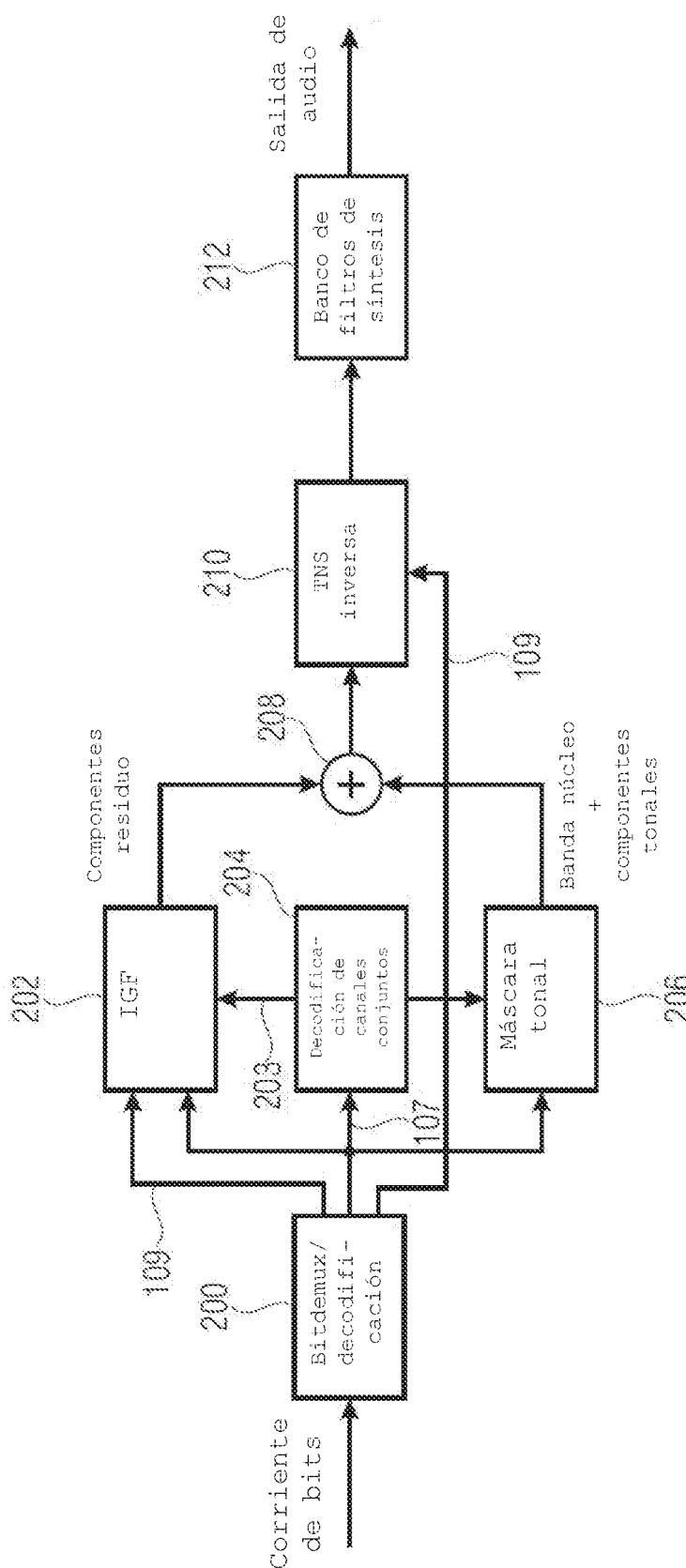


FIG 2A

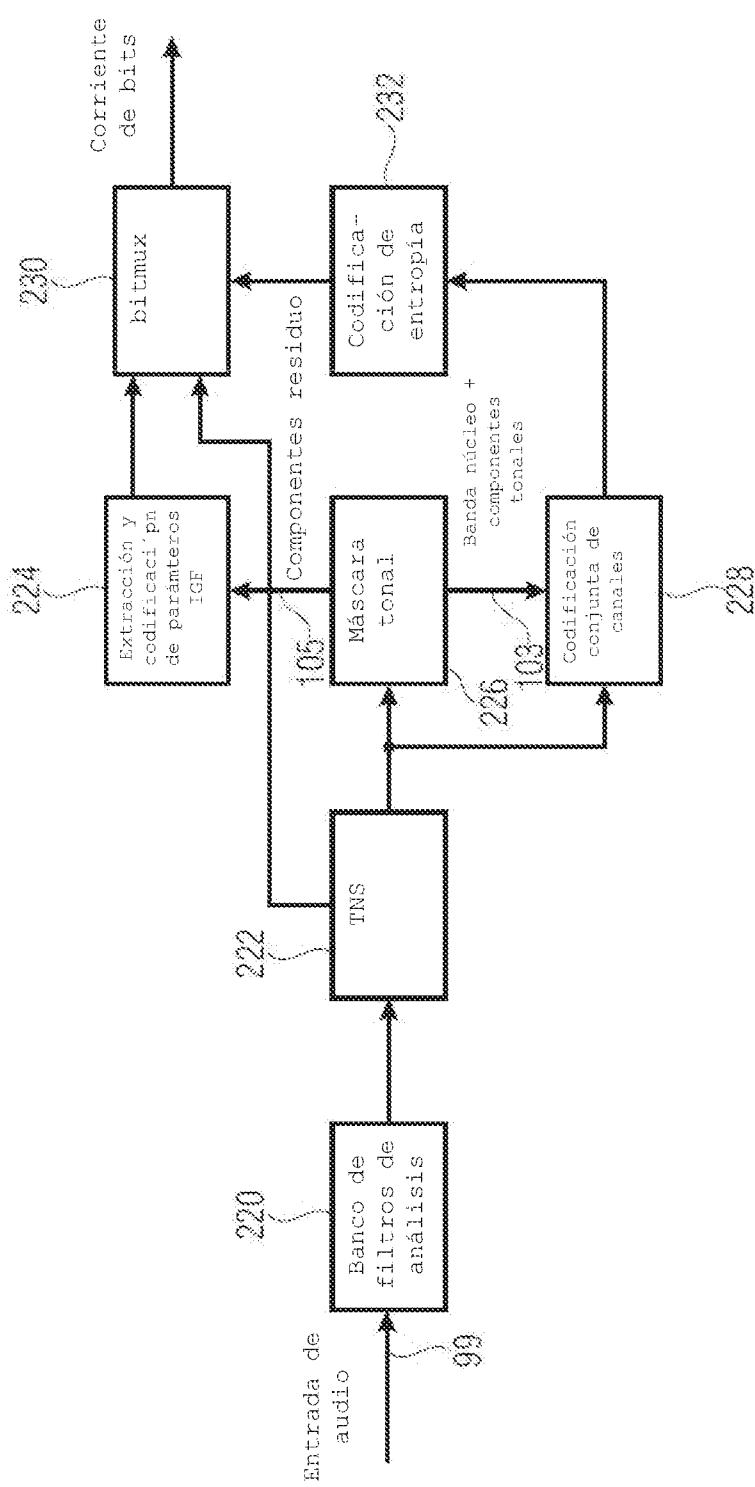


FIG 2B

ES 2 992 880 T3

- 1ra. resolución (elevada resolución) para envolvente del 1er conjunto (codificación de a líneas)
- 2da. resolución (baja resolución) para envolvente del 2do conjunto (factor de escala por SCB)

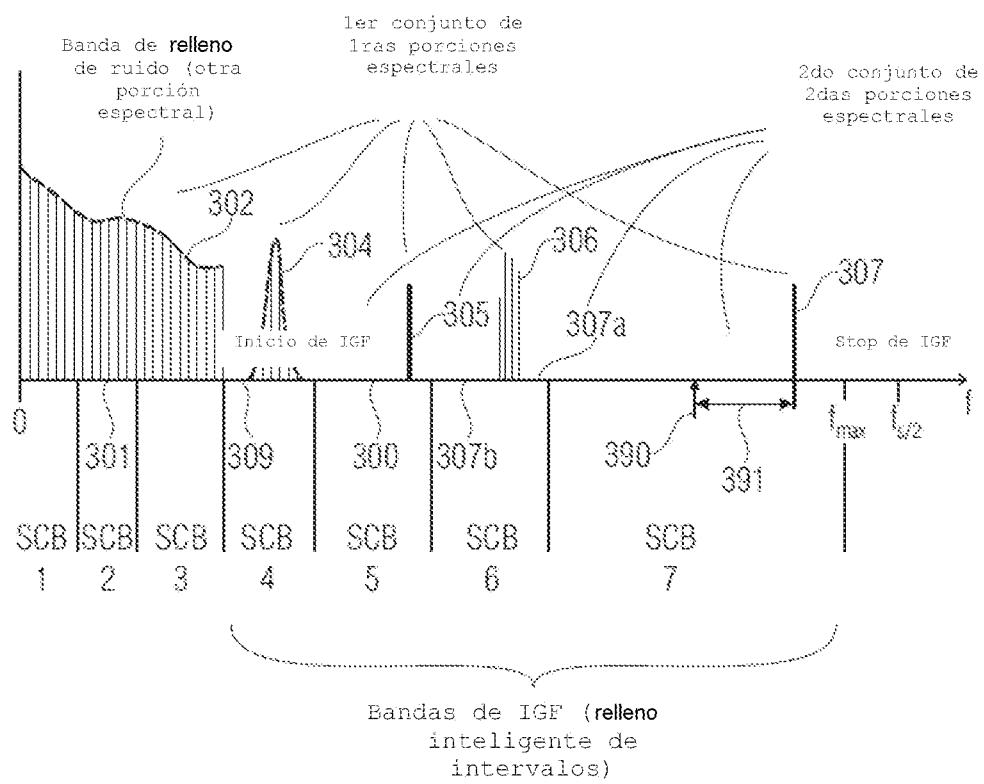


FIG 3A

SCB1	SCB2	SCB3	SCB4	SCB5	SCB6	SCB7
SF1	SF2	SF3	SF4	SF5	SF6	SF7
			E_1	E_2	E_3	E_4
	NF_2					

308 310 312

FIG 3B

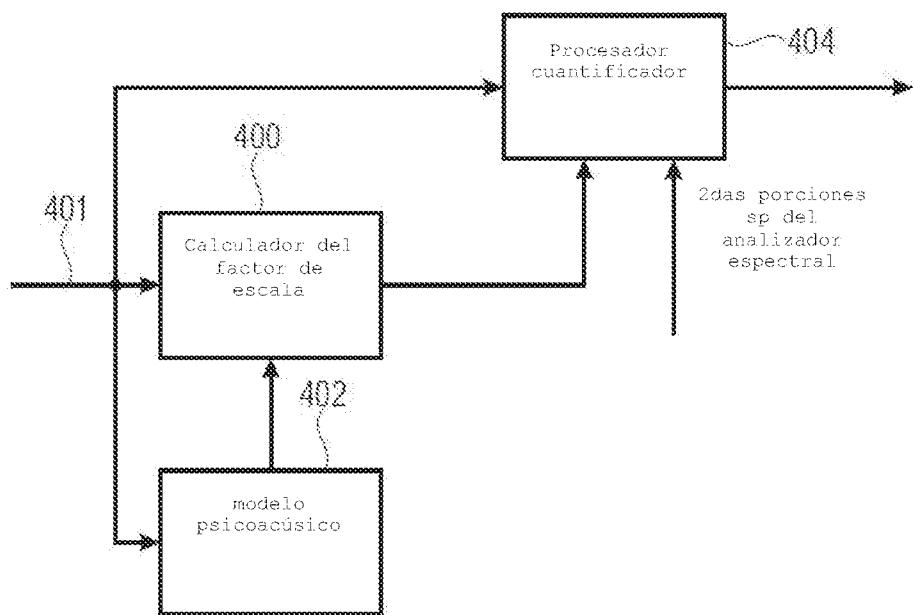


FIG 4A

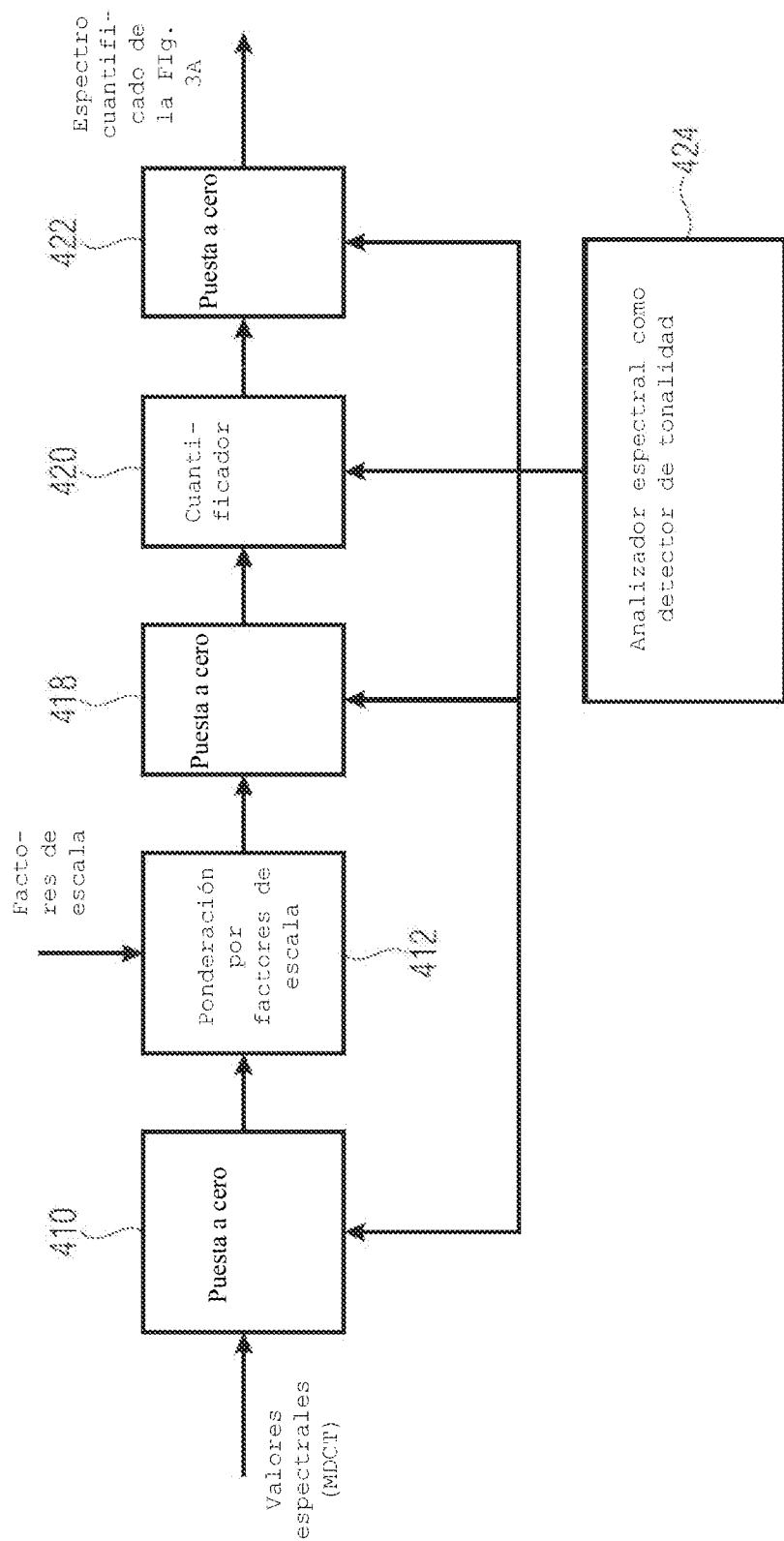


FIG. 4B

PROCESADOR CUANTIFICADOR

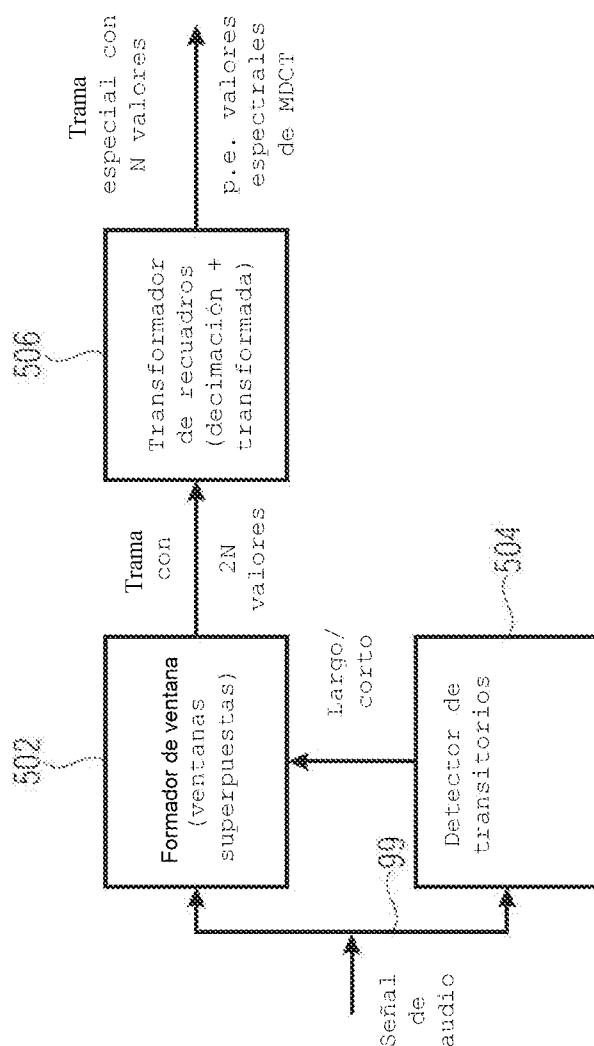


FIG 5A

(otras porciones espectrales)

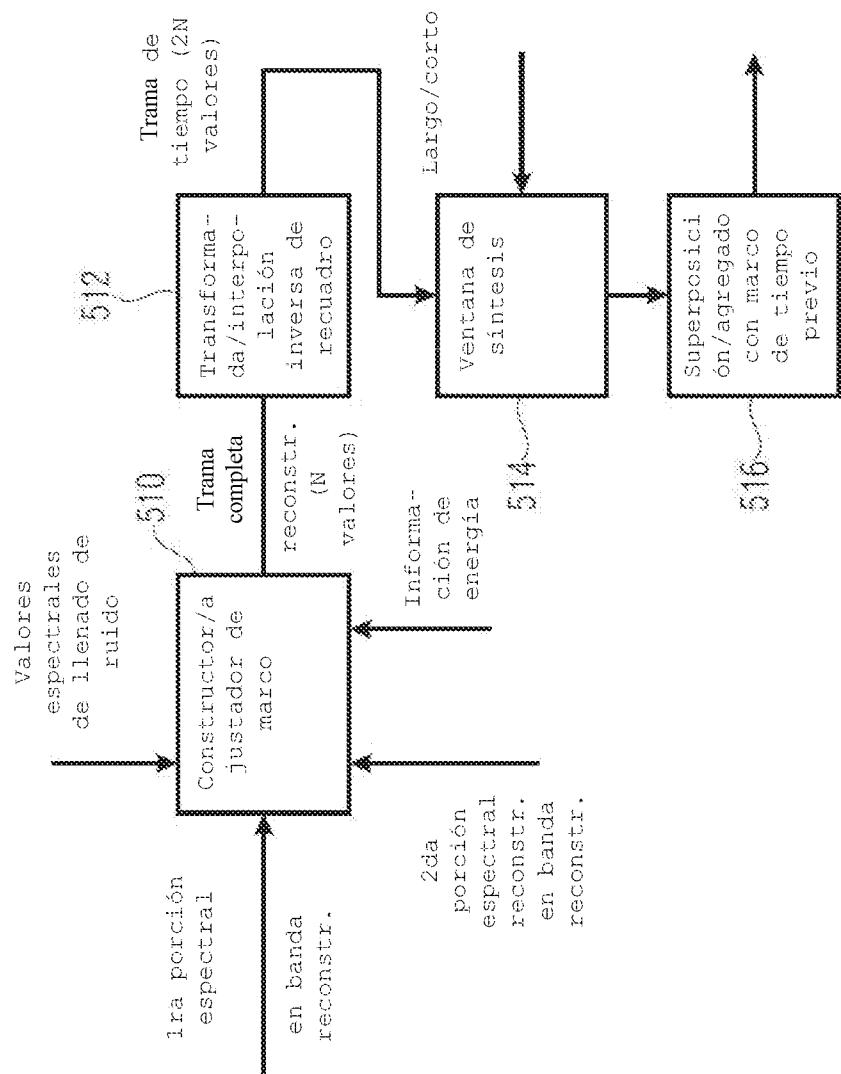
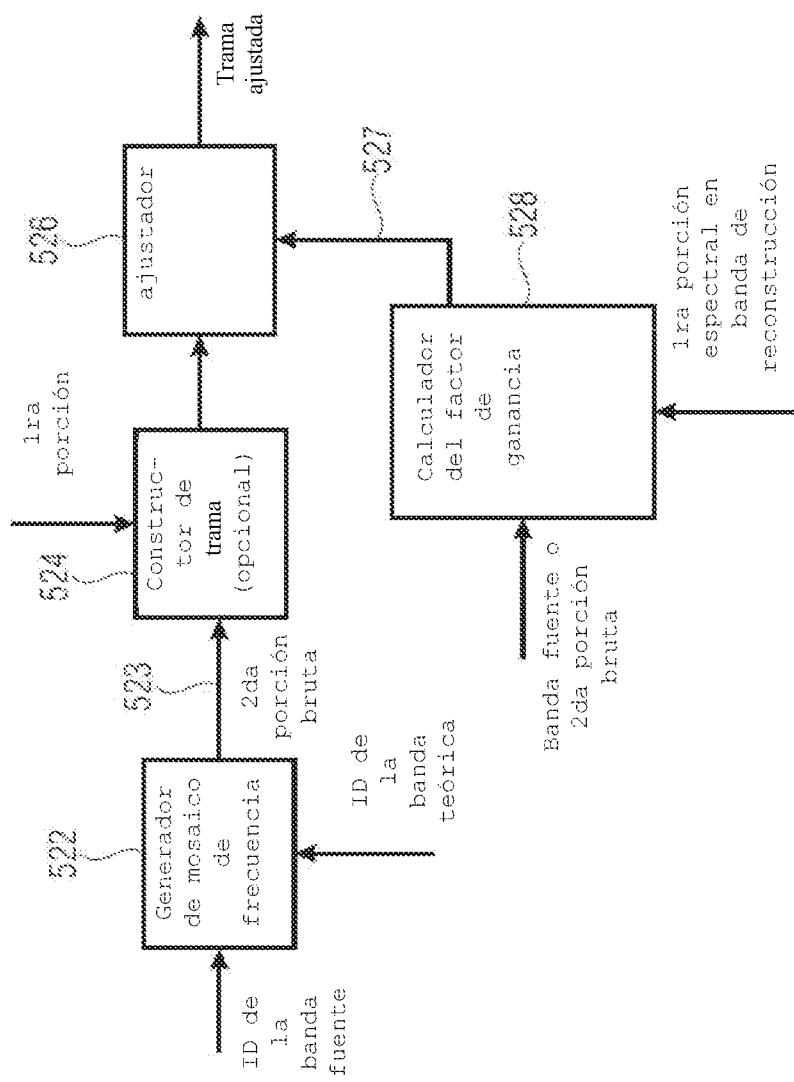


FIG 5B



५७

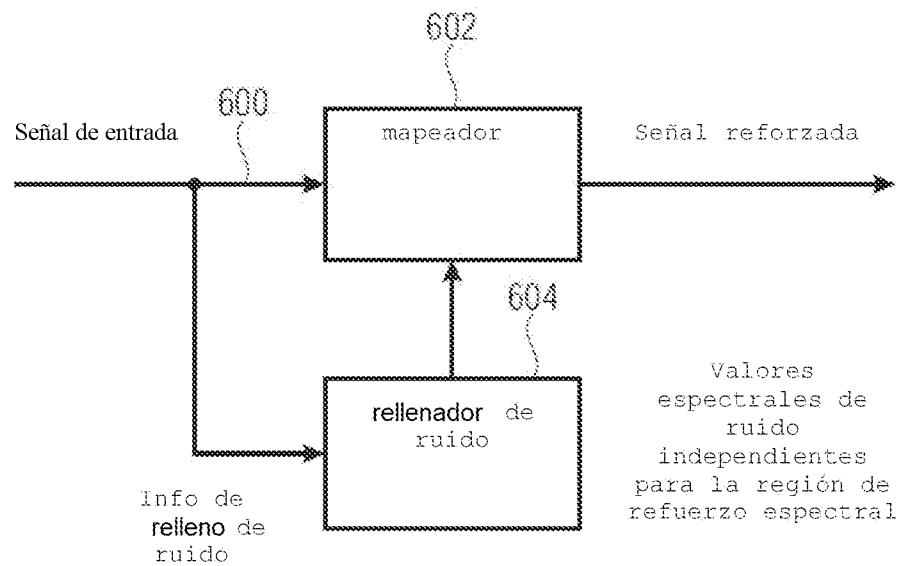
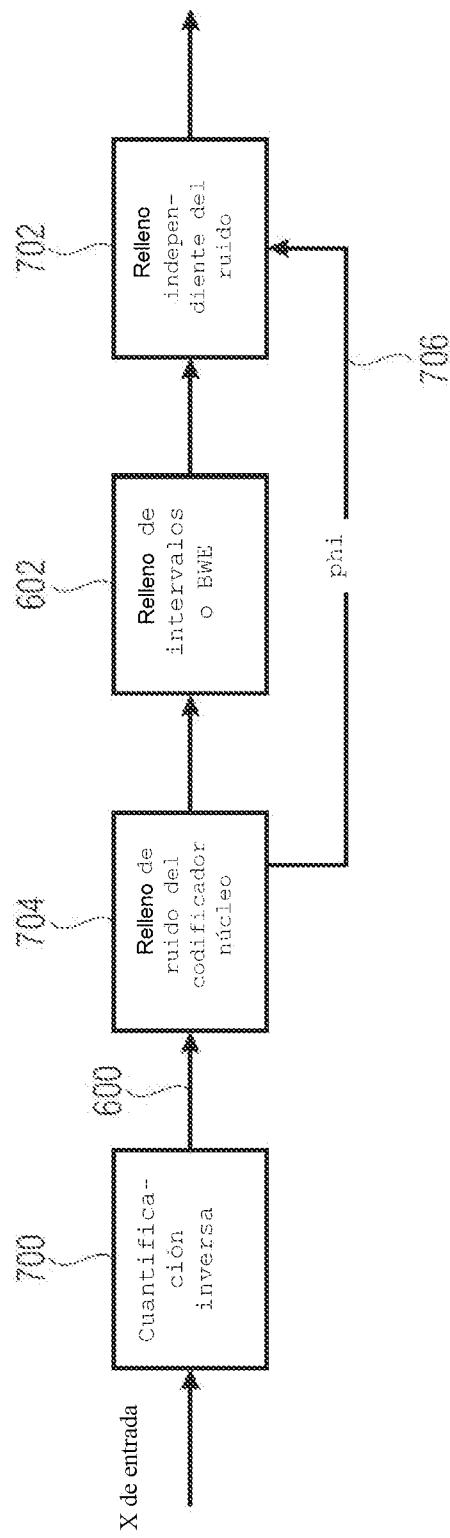
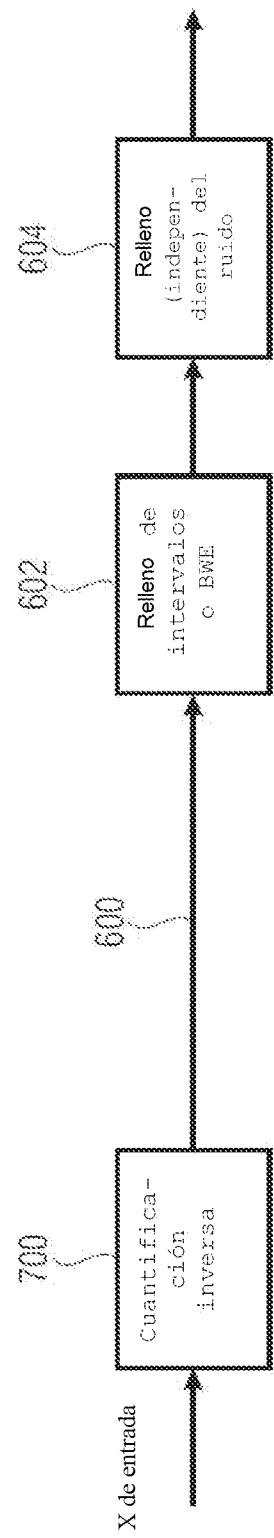


FIG 6



11



三

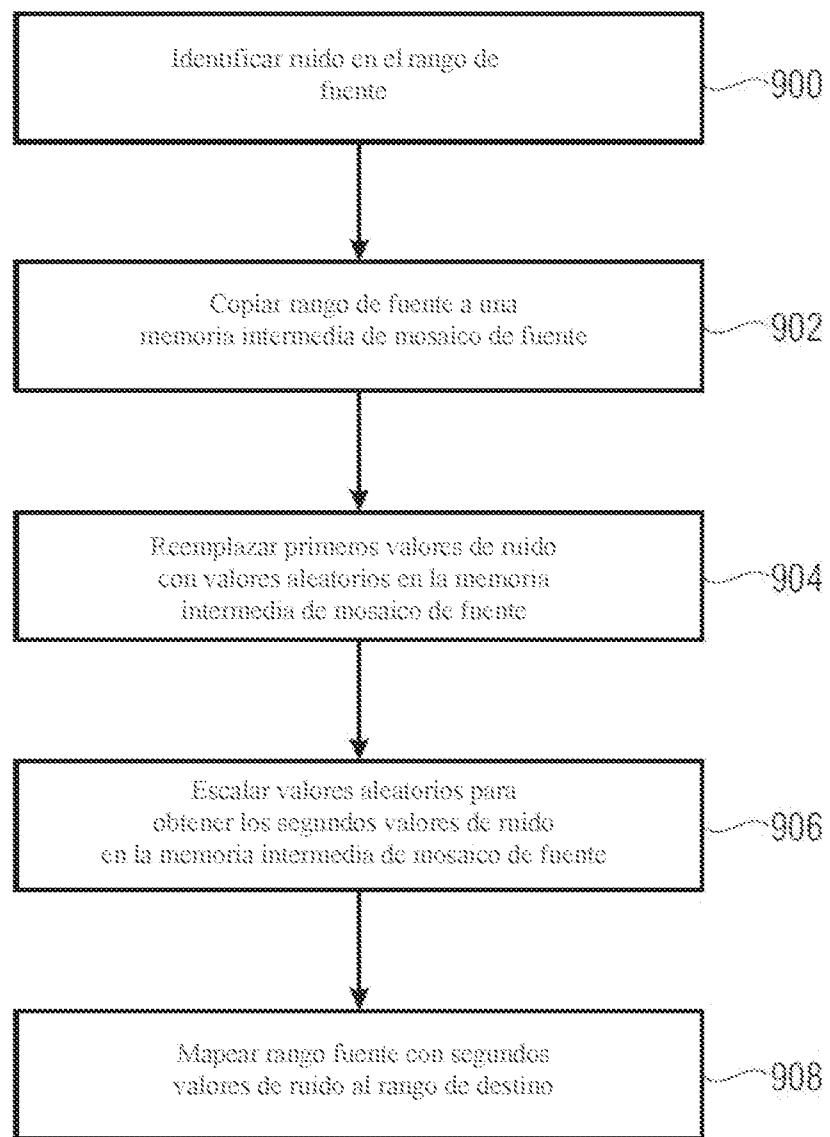


FIG 9

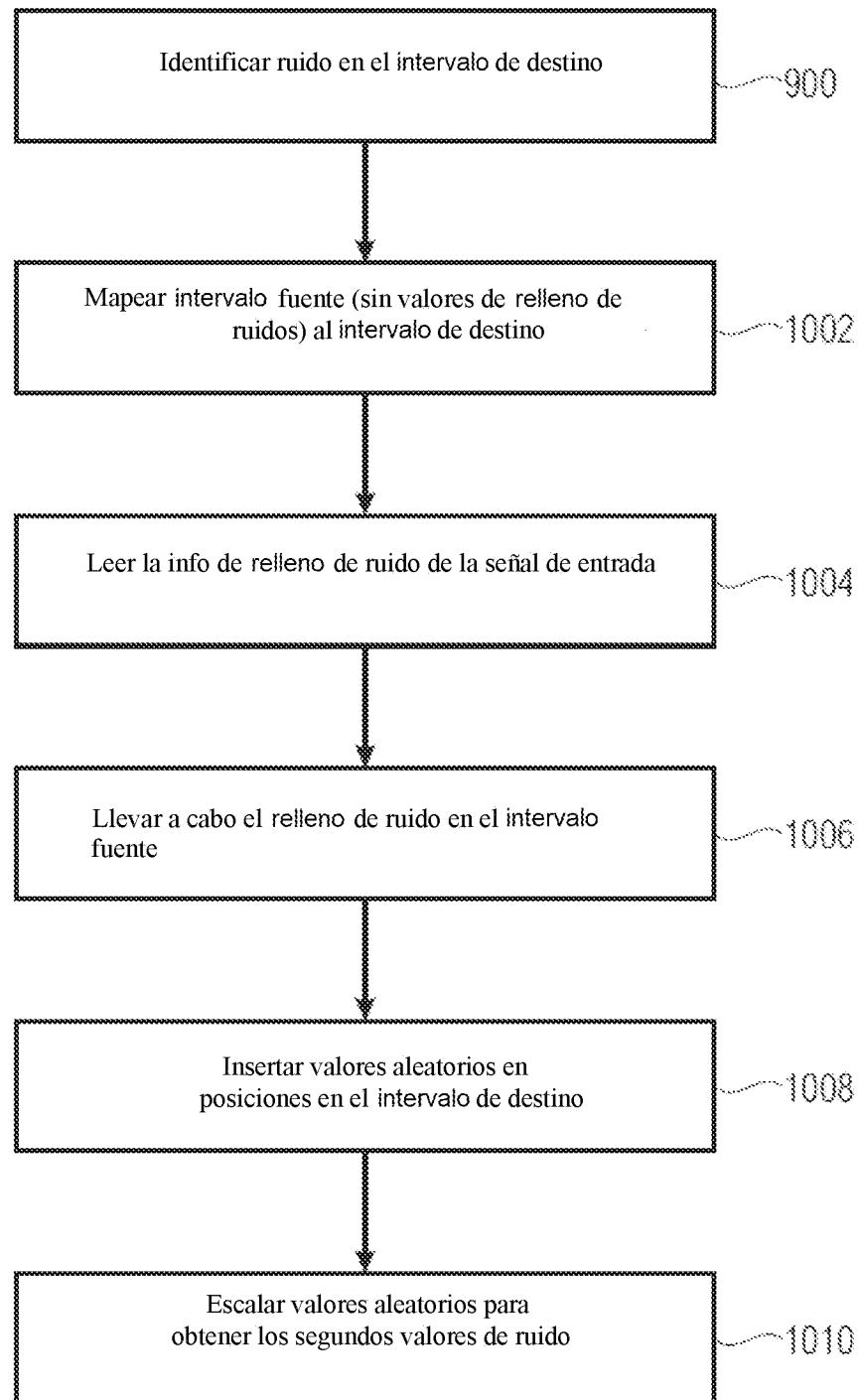


FIG 10

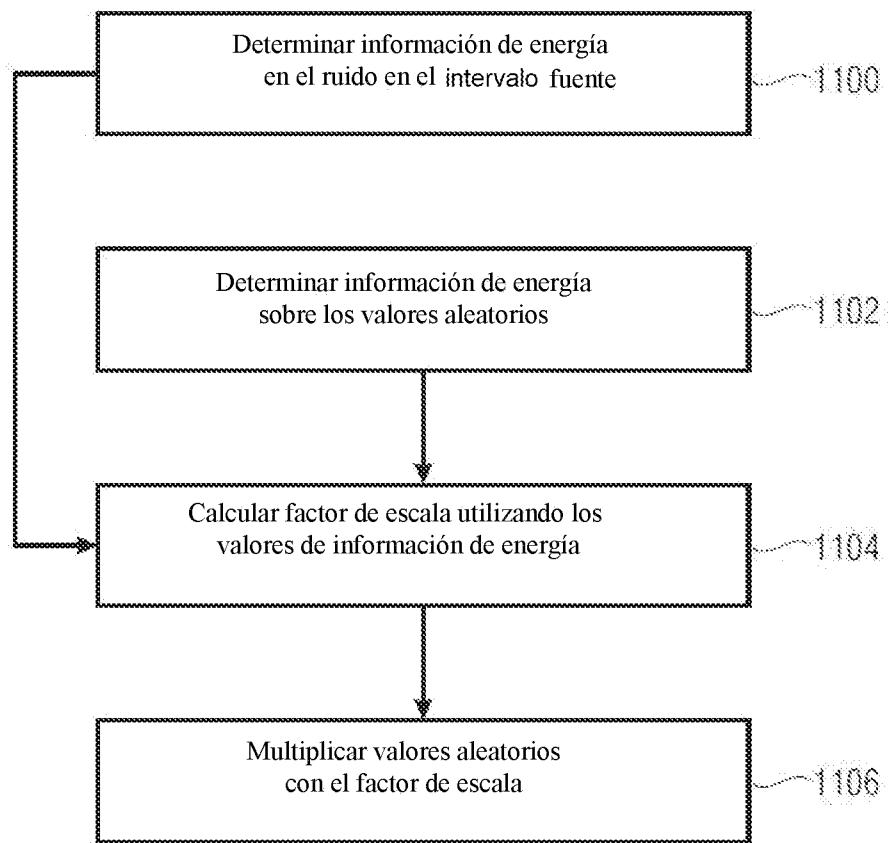


FIG 11

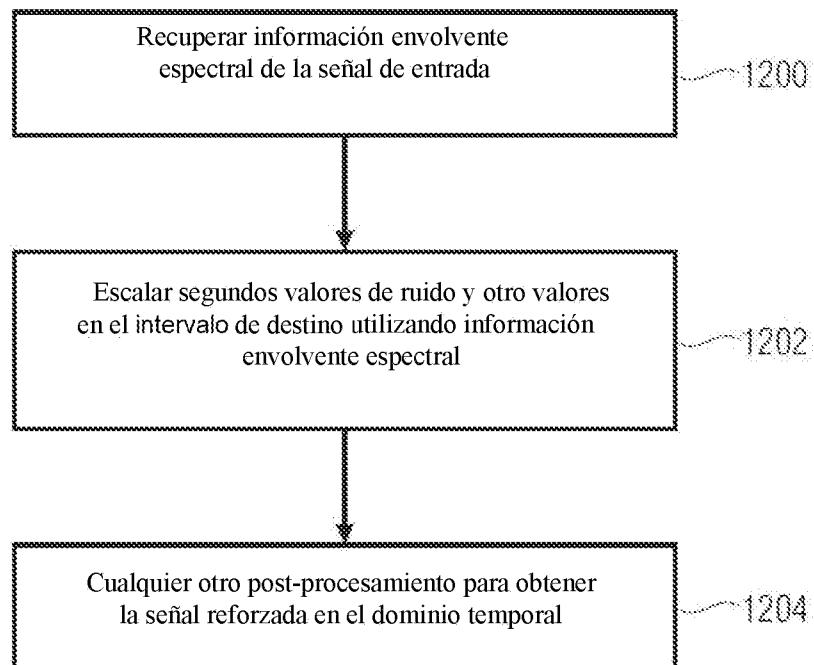


FIG 12

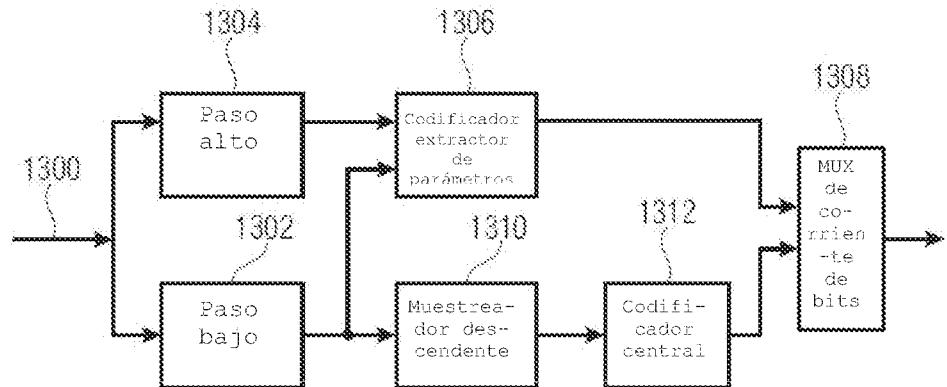


FIG 13A

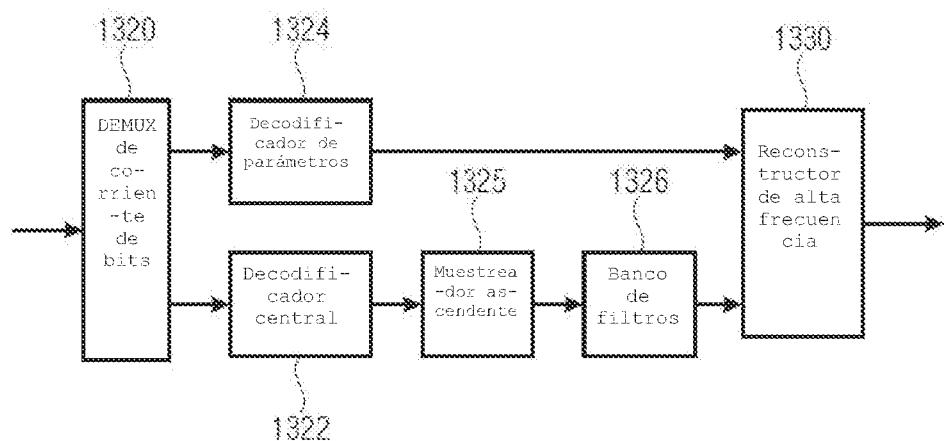


FIG 13B