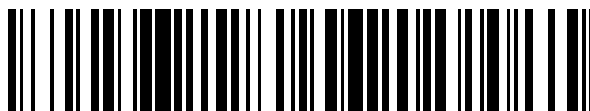


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 774 492**

51 Int. Cl.:

**G10L 19/005** (2013.01)

G10L 25/90 (2013.01)

G10L 19/08 (2013.01)

G10L 19/02 (2013.01)

G10L 19/12 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **27.10.2014** **E 17201222 (1)**

97 Fecha y número de publicación de la concesión europea: **04.12.2019** **EP 3336841**

54 Título: **Decodificador de audio y método para proporcionar una información de audio decodificada usando un ocultamiento de error que modifica una señal de excitación de dominio de tiempo**

30 Prioridad:

**31.10.2013 EP 13191133**

**28.07.2014 EP 14178825**

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

**21.07.2020**

73 Titular/es:

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

72 Inventor/es:

**LECOMTE, JÉRÉMIE**

74 Agente/Representante:

**ARIZTI ACHA, Monica**

ES 2 774 492 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

Decodificador de audio y método para proporcionar una información de audio decodificada usando un ocultamiento de error que modifica una señal de excitación de dominio de tiempo

5

**Campo técnico**

Las realizaciones de acuerdo con la invención crean decodificadores de audio para proporcionar una información de audio decodificada basándose en una información de audio codificada.

10

Algunas realizaciones según la invención crean métodos para proporcionar una información de audio decodificada sobre la base de una información de audio codificada.

15

Algunas realizaciones según la invención crean programas de ordenador para la realización de uno de dichos métodos.

Algunas realizaciones según la invención se refieren a un ocultamiento de dominio de tiempo para un códec de dominio de transformada.

20

**Antecedentes de la invención**

En los últimos años, ha habido una creciente demanda de transmisión y almacenamiento digital de contenidos de audio. Sin embargo, los contenidos de audio con frecuencia son transmitidos sobre canales no confiables, lo que trae el riesgo de que las unidades de datos (por ejemplo, paquetes) que comprenden una o más tramas de audio (por ejemplo, en forma de una representación codificada, como por ejemplo una representación de dominio de tiempo codificada o una representación de dominio de frecuencia codificada) se pierdan. En algunas situaciones, será posible requerir una repetición (reenvío) de las tramas de audio perdidas (o de unidades de datos, como paquetes, que comprenden una o más tramas de audio perdidas). Sin embargo, esto típicamente producirá un retardo sustancial, y, por lo tanto, requerirá un extenso almacenamiento temporario (buffering) de tramas de audio. En otros casos, es casi imposible requerir una repetición de tramas de audio perdidas.

25

30

A fin de obtener una buena, o al menos aceptable, calidad de audio dado el caso que las tramas de audio se pierdan sin proporcionar el extenso almacenamiento temporario (lo que consumiría una gran cantidad de memoria, y lo que, además, degradaría sustancialmente las capacidades en tiempo real de la codificación de audio), es deseable contar con conceptos para manejar una pérdida de una o más tramas de audio. En particular, es deseable contar con conceptos que produzcan una buena calidad de audio, o por lo menos, una calidad de audio aceptable, incluso, en el caso de que las tramas de audio se pierdan.

35

En el pasado, se han desarrollado algunos conceptos de ocultamiento de error, que pueden emplearse en diferentes conceptos de codificación de audio.

40

En lo que sigue, se describirá un concepto de codificación de audio convencional.

En el estándar 3gpp TS26.290, se explica una decodificación de excitación codificada transformada (decodificación TCX [excitación codificada transformada, conforme a sus siglas en inglés]) con ocultamiento de error. En lo que sigue, se proporcionarán algunas explicaciones, que se basan en la sección de "Síntesis de señal y decodificación de modo TCX" en la referencia [1].

45

Un decodificador TCX de acuerdo con la Norma Internacional 3gpp TS 26.290 se muestra en las Figs. 7 y 8, en el que las Figs. 7 y 8 muestran diagramas de bloques del decodificador TCX. Sin embargo, la figura 7 muestra aquellos bloques funcionales que son pertinentes para la decodificación TCX en una operación normal, o en un caso de una pérdida de paquetes parcial. En contraste, la Fig. 8 muestra el procesamiento pertinente de la decodificación TCX en el caso de ocultación de borrado de paquete TCX-256.

50

En otras palabras, las Figs. 7 y 8 muestran un diagrama de bloques del decodificador TCX que incluye los siguientes casos:

55

Caso 1 (figura 8): Ocultación de borrado de paquetes en TCX-256 cuando la longitud de trama TCX es de 256 muestras y el paquete relacionado está perdido, es decir,  $BFI\_TCX = (1)$ ; y

60

Caso 2 (Fig.7): decodificación TCX normal, posiblemente, con pérdidas de paquetes parciales.

A continuación, se proporcionarán algunas explicaciones en relación con las Figs. 7 y 8.

Tal como se menciona, la figura 7 muestra un diagrama de bloques de un decodificador TCX que realiza una decodificación TCX en operación normal, o, en el caso de pérdida de paquete parcial. El decodificador TCX 700 según la figura 7 recibe parámetros 710 específicos de TCX y proporciona, sobre su base, la información 712, 714 de audio decodificada.

5 El decodificador 700 de audio comprende un demultiplexor "DEMUX TCX 720", que está configurado para recibir los parámetros 710 específicos de TCX y la información "BFI\_TCX". El demultiplexor 720 separa los parámetros 710 específicos de TCX, y proporciona una información 722 de excitación codificada, una información 724 de relleno de ruido codificada, y una información 726 de ganancia global codificada. El decodificador 700 de audio comprende un  
10 decodificador 730 de excitación, que está configurado para recibir la información 722 de excitación codificada, la información 724 de relleno de ruido codificada y la información 726 de ganancia global codificada, al igual que cierta información adicional (por ejemplo, una bandera de tasa de bits "tasa\_bits\_bandera", una información "BFI\_TCX" y una información de longitud de trama TCX. El decodificador 730 de excitación proporciona, sobre su base, una señal 728 de excitación de dominio de tiempo (también designada con "x"). El decodificador 730 de excitación comprende  
15 un procesador 732 de información de excitación, que demultiplexa la información 722 de excitación codificada y decodifica los parámetros de cuantificación algebraica. El procesador 732 de información de excitación proporciona una señal 734 de excitación intermedia, que, típicamente, se encuentra en una representación de dominio de frecuencia, y que se designa con Y. El codificador 730 de excitación además comprende un inyector 736 de ruido, que está configurado para inyectar ruido en subbandas no cuantificadas, a fin de derivar una señal 738 de excitación  
20 rellena con ruido de la señal 734 de excitación intermedia. La señal 738 de excitación rellena con ruido típicamente se encuentra en el dominio de frecuencia, y se designa con Z. El inyector 736 de ruido recibe una información 742 de intensidad de ruido desde un decodificador 740 de nivel de relleno de ruido. El decodificador de excitación además comprende un desénfasis 744 de baja frecuencia adaptativo, que está configurado para realizar una operación de desénfasis de baja frecuencia sobre la base de la señal 738 de excitación rellena con ruido, de manera de obtener una señal 746 de excitación procesada, que se encuentra aún en el dominio de frecuencia, y que se designa con X'. El decodificador 730 de excitación además comprende un transformador 748 de dominio de  
25 frecuencia a dominio de tiempo, que está configurado para recibir la señal 746 de excitación procesada y para proporcionar, sobre su base, una señal 750 de excitación de dominio de tiempo, que se asocia con una cierta porción de tiempo representada por un conjunto de parámetros de excitación de dominio de frecuencia (por ejemplo, de la señal 746 de excitación procesada). El decodificador 730 de excitación además comprende un escalador 752, que está configurado para escalar la señal 750 de excitación de dominio de tiempo a fin de obtener una señal 754 de excitación de dominio de tiempo escalada. El escalador 752 recibe una información 756 de ganancia global desde un decodificador 758 de ganancia global, donde, en respuesta, el decodificador 758 de ganancia global recibe la información 726 de ganancia global codificada. El decodificador 730 de excitación además comprende una síntesis  
30 760 de superposición y adición, que recibe las señales 754 de excitación de dominio de tiempo escaladas asociadas con una pluralidad de porciones de tiempo. La síntesis 760 de superposición y adición realiza una operación de superposición y adición (que puede incluir una operación de ventaneado) sobre la base de las señales 754 de excitación de dominio de tiempo escaladas, de modo de obtener una señal 728 de excitación de dominio de tiempo temporalmente combinada durante un período más largo en el tiempo (más largo que los períodos en el tiempo para los cuales se proporcionan las señales 750, 754 de excitación de dominio de tiempo individuales).

El decodificador 700 de audio además comprende una síntesis 770 de codificación predictiva lineal (LPC, conforme a sus siglas en inglés), que recibe la señal 728 de excitación de dominio de tiempo provista por la síntesis 760 de superposición y adición y uno o más coeficientes de codificación predictiva lineal (LPC) que definen una función de  
45 filtro de síntesis 772 de codificación predictiva lineal (LPC). La síntesis de codificación predictiva lineal (LPC) 770, por ejemplo, puede comprender un primer 774 filtro, que, por ejemplo, puede filtrar por síntesis la señal 728 de excitación de dominio de tiempo, a fin de obtener la señal de audio decodificada 712. Opcionalmente, la síntesis 770 de codificación predictiva lineal (LPC) puede comprender además un segundo 772 filtro de síntesis que está configurado para filtrar por síntesis la señal de salida del primer 774 filtro usando otra función de filtro de síntesis, de modo de obtener la señal 714 de audio decodificada.

A continuación, se describirá la decodificación TCX en el caso de una ocultación de borrado de paquetes TCX-256. La figura 8 muestra un diagrama de bloques del decodificador TCX, en este caso.

55 El ocultamiento 800 de borrado de paquetes recibe una información 810 de altura, que se designa además con "altura\_tcX", y que se obtiene a partir de una trama TCX decodificada previa. Por ejemplo, la información 810 de altura puede obtenerse usando un estimador 747 de altura dominante desde la señal 746 de excitación procesada en el decodificador 730 de excitación (durante la decodificación "normal"). Aún más, el ocultamiento 800 de borrado de paquetes recibe parámetros 812 de codificación predictiva lineal (LPC), que pueden representar una función de  
60 filtro de síntesis de codificación predictiva lineal (LPC). Los parámetros 812 de codificación predictiva lineal (LPC), por ejemplo, pueden ser idénticos a los parámetros 772 de codificación predictiva lineal (LPC). En consecuencia, el ocultamiento 800 de borrado de paquetes puede ser configurado para proporcionar, sobre la base de la información 810 de altura y los parámetros 812 de codificación predictiva lineal (LPC), una señal de ocultamiento 814 de error, que puede considerarse una información de audio de ocultamiento de error. El ocultamiento 800 de borrado de

paquetes comprende un búfer 820 de excitación, que, por ejemplo, puede almacenar temporalmente una excitación previa. El búfer 820 de excitación, por ejemplo, puede hacer uso del libro de códigos adaptativo ACELP [predicción lineal excitada por libro de códigos adaptativo, conforme a sus siglas en inglés], y puede proporcionar una señal 822 de excitación. El ocultamiento 800 de borrado de paquetes puede comprender adicionalmente un primer 824 filtro, una función de filtro que puede definirse como se muestra en la figura 8. Por lo tanto, el primer 824 filtro puede filtrar la señal 822 de excitación sobre la base de los parámetros 812 de codificación predictiva lineal (LPC), de modo de obtener una versión 826 filtrada de la señal 822 de excitación. El ocultamiento de borrado de paquetes además comprende un limitador 828 de amplitud, que puede limitar una amplitud de la señal 826 de excitación filtrada sobre la base de información objetivo o información de nivel  $rms_{wsyn}$ . Aún más, el ocultamiento 800 de borrado de paquetes puede comprender un segundo 832 filtro, que puede estar configurado para recibir la señal 830 de excitación filtrada de limitada amplitud desde el limitador 822 de amplitud y para proporcionar, sobre su base, la señal de ocultamiento 814 de error. Una función de filtro del segundo 832 filtro, por ejemplo, puede definirse como se muestra en la figura 8.

15 En lo que sigue, se describirán algunos detalles respecto de la decodificación y el ocultamiento de error.

En el Caso 1 (ocultación de borrado de paquetes en TCX-256), no hay información disponible para la decodificación de la trama TCX de 256 muestras. La síntesis de TCX se halla mediante el procesamiento de la excitación pasada retardada por  $T$ , donde  $T=tono\_tcx$  es una demora de tono estimada en la trama TCX previamente decodificada, por un filtro no lineal aproximadamente equivalente a  $1/A(z)$ . Se usa un filtro no lineal en lugar de  $1/A(z)$  para evitar clics en la síntesis. Este filtro se descompone en 3 pasos.

**Paso 1:** filtración mediante:

$$\frac{\hat{A}(z/\gamma)}{\hat{A}(z)} \cdot \frac{1}{1-\alpha z^{-1}}$$

para cartografiar la excitación retardada por  $T$  en el dominio objetivo de TCX;

**Step 2:** applying a limiter (the magnitude is limited to  $\pm rms_{wsyn}$ )

**Paso 3:** la filtración mediante:

$$\frac{1-\alpha z^{-1}}{\hat{A}(z/\gamma)}$$

para hallar la síntesis. Obsérvese que la memoria intermedia **OVLP\_TCX** se establece en cero, en este caso.

### Decodificación de los parámetros VQ algebraicos.

En el Caso 2, la decodificación de TCX involucra la decodificación de los parámetros VQ algebraicos que describen cada bloque cuantificado  $B'_k$  del espectro escalado  $X'$ , donde  $X'$  es como se describe en la Etapa 2 de la Sección 5.3.5.7 de 3gpp TS 26.290. Recuérdese que  $X'$  tiene dimensión  $N$ , donde  $N = 288, 576$  y  $1152$  para TCX-256, 512 y 1024, respectivamente, y que cada bloque  $B'_k$  tiene dimensión 8. El número  $K$  de bloques  $B'_k$  es, por lo tanto, 36, 72 y 144 para TCX-256, 512 y 1024, respectivamente. Los parámetros VQ algebraicos para cada bloque  $B'_k$  se describen el Paso 5 de la Sección 5.3.5.7. Para cada bloque  $B'_k$ , tres grupos de índices binarios son enviados por el codificador:

a) el índice de libro de códigos  $n_k$ , transmitido en código unario como se describe en el Etapa 5 de la Sección 5.3.5.7;

b) la serie  $l_k$  de un punto de rejilla seleccionado  $c$  en un denominado *libro de códigos base*, que indica la permutación que debe aplicarse a un *líder* específico (véase la Etapa 5 de la Sección 5.3.5.7) para obtener un punto de rejilla  $c$ ;

c) y, si el bloque cuantificado  $B'_k$  (un punto de rejilla) no se presentó en el libro de códigos base, los 8 índices del vector de índice de extensión de Voronoi  $k$  calculado en el subpaso V1 del Paso 5 en la Sección; a partir de índices de extensión de Voronoi, puede computarse un vector de extensión  $z$  como en la referencia [1] de 3gpp TS 26.290. El número de bits en cada componente del vector de índice  $k$  es provisto por el orden de extensión  $r$ , que puede obtenerse a partir del valor de código unario de índice  $n_k$ . El factor de escala  $M$  de la extensión de Voronoi es provisto por  $M = 2^r$ .

A continuación, a partir del factor de escala  $M$ , el Vector de extensión de Voronoi  $z$  (un punto de rejilla en  $RE_\beta$ ) y el punto de rejilla  $c$  en el libro de códigos base (también, un punto de rejilla en  $RE_\beta$ ), cada bloque escalado cuantificado  $B'_k$  puede computarse como:

$$\hat{B}'_k = M \mathbf{c} + \mathbf{z}$$

5 Cuando no hay extensión de Voronoi (es decir,  $n_k < 5$ ,  $M=1$  y  $\mathbf{z}=0$ ), el libro de códigos base es o bien el libro de códigos  $Q_0$ ,  $Q_2$ ,  $Q_3$  o  $Q_4$  de la referencia [1] de 3gpp TS 26.290. Entonces no se requieren bits para transmitir el vector  $\mathbf{k}$ . De lo contrario, cuando se usa la extensión de Voronoi debido a que  $B'_k$  es suficientemente grande, entonces solo  $Q_3$  o  $Q_4$  de la referencia [1] se usan como un libro de códigos base. La selección de  $Q_3$  o  $Q_4$  está implícita en el valor de índice de libro de códigos  $n_k$ , como se describe en el Etapa 5 de la Sección 5.3.5.7.

10 **Estimación del valor de altura dominante.**

La estimación del tono dominante se realiza de modo tal que la siguiente trama a decodificarse puede extrapolarse apropiadamente si corresponde a TCX-256, y si el paquete relacionado está perdido. Esta estimación se sustenta en la asunción de que el pico de máxima magnitud en el espectro del objetivo TCX corresponde a la altura dominante.

15 La búsqueda de la máxima  $M$  es restringida a una frecuencia inferior a  $F_s/64$  kHz

$$M = \max_{i=1..N/32} (X'_{2i})^2 + (X'_{2i+1})^2$$

20 y el índice mínimo  $1 \leq i_{\max} \leq N/32$  de modo tal que  $(X'_{2i})^2 + (X'_{2i+1})^2 = M$  también es hallado. A continuación, la altura dominante se estima en el número de muestras como  $T_{\text{est}} = N / i_{\max}$  (este valor puede no ser entero). Recuérdese que el tono dominante se calcula para la ocultación de borrado de paquetes en TCX-256. A fin de evitar problemas de almacenamiento temporario (el almacenamiento en memoria intermedia de excitación se limita a 256 muestras), si  $T_{\text{est}} > 256$  muestras, *altura\_tcx* se establece en 256; de lo contrario, si  $T_{\text{est}} \leq 256$ , se evita período de altura múltiple en 256 muestras estableciendo *altura\_tcx* en

25

$$\text{pitch\_tcx} = \max \{ \lfloor n T_{\text{est}} \rfloor \mid n \text{ integer} > 0 \text{ and } n T_{\text{est}} \leq 256 \}$$

donde  $\lfloor \cdot \rfloor$  denota el redondeo al entero más cercano hacia  $-\infty$ .

30 En lo que sigue, se describirán brevemente algunos conceptos convencionales adicionales.

En ISO\_IEC\_DIS\_23003-3 (referencia [3]), se explica una decodificación TCX que emplea MDCT en el contexto del Códec Unificado de Voz y Audio.

35 En el estado de la técnica de AAC [codificación avanzada de audio, conforme a sus siglas en inglés] (conferir, por ejemplo, referencia [4]), solo se describe un modo de interpolación. Según la referencia [4], el decodificador núcleo AAC incluye una función de ocultamiento que incrementa el retardo del decodificador por una trama.

40 En la Patente Europea EP 1207519 B1 (referencia [5]), se describe la provisión de un decodificador de voz y método de compensación de error capaz de lograr el mejoramiento adicional para la voz decodificada en una trama en la cual se detectar un error. Según la patente, un parámetro de codificación de voz incluye información de modo que expresa rasgos de cada segmento corto (trama) de voz. El codificador de voz calcula de manera adaptativa los parámetros de demora y los parámetros de ganancia utilizados para la decodificación de voz según la información de modo. Aún más, el decodificador de voz controla de manera adaptativa la relación de ganancia de excitación adaptativa y la ganancia de excitación fijada según la información de modo. Además, el concepto según la patente comprende el control adaptativo de los parámetros de ganancia de excitación adaptativa y los parámetros de ganancia de excitación fijada utilizados para la decodificación de voz según valores de parámetros de ganancia decodificada en una unidad de decodificación normal en la cual no se detecta error, inmediatamente después de una unidad de decodificación cuyos datos codificados son detectados con un error.

50

Se conoce, según la patente US8.255.207B2, un método de ocultamiento de error usando una repetición de un ciclo de altura que comprende muestras de excitación.

55 En vista de la técnica anterior, existe la necesidad de hallar un mejoramiento adicional del ocultamiento de error, que proporcione una mejor impresión auditiva.

**3. Sumario de la invención.**

60 La invención proporciona un decodificador de audio según la reivindicación 1, un método según la reivindicación 2 y un programa de ordenador para llevar a cabo el método según la reivindicación 3.

Todos los casos siguientes de la palabra "realizaciones" o de la palabra "aspecto" si se refieren a las implementaciones que no comprenden todas las características de las reivindicaciones independientes, o equivalentes de las mismas, deben considerarse como ejemplos adicionales.

5 Una realización según la invención crea un decodificador de audio para proporcionar una información de audio decodificada sobre la base de una información de audio codificada. El decodificador de audio comprende un ocultamiento de error configurado para proporcionar una información de audio de ocultamiento de error para el ocultamiento de una pérdida de una trama de audio (o más de una pérdida de trama) tras una trama de audio codificada en una representación de dominio de frecuencia, usando una señal de excitación de dominio de tiempo.

10 Esta realización según la invención se sustenta en el hallazgo de que puede obtenerse un ocultamiento de error mejorado al proporcionar la información de audio de ocultamiento de error sobre la base de una señal de excitación de dominio de tiempo, incluso, si la trama de audio que precede una trama de audio perdida está codificada en una representación de dominio de frecuencia. En otras palabras, se ha reconocido que una calidad de un ocultamiento de error típicamente es mejor si el ocultamiento de error se realiza sobre la base de una señal de excitación de dominio de tiempo, cuando se compara con un ocultamiento de error realizado en un dominio de frecuencia, de modo tal que vale la pena la conmutación a un ocultamiento de error de dominio de tiempo, usando una señal de excitación de dominio de tiempo, aún si el contenido de audio que precede la trama de audio perdida está codificado en el dominio de frecuencia (es decir, en una representación de dominio de frecuencia). Esto es válido, por ejemplo, para una señal monofónica y, en su mayoría, para voz.

Por consiguiente, la presente invención permite obtener un buen ocultamiento de error, aún si la trama de audio que precede la trama de audio perdida está codificada en el dominio de frecuencia (es decir, en una representación de dominio de frecuencia).

25 En una realización preferida, la representación de dominio de frecuencia comprende una representación codificada de una pluralidad de valores espectrales y una representación codificada de una pluralidad de factores de escala para la escala de los valores espectrales, o el decodificador de audio está configurado para derivar una pluralidad de factores de escala para la escala de los valores espectrales a partir de una representación codificada de parámetros de codificación predictiva lineal (LPC). Esto podría realizarse usando FDNS (Forma de Ruido de Dominio de frecuencia). Sin embargo, se ha hallado que es conveniente derivar la señal de excitación de dominio de tiempo (que puede servir como una excitación para una síntesis de codificación predictiva lineal (LPC)) aún si la trama de audio que precede la trama de audio perdida es originalmente codificada en la representación de dominio de frecuencia que comprende información sustancialmente diferente (a decir, una representación codificada de una pluralidad de valores espectrales en una representación codificada de una pluralidad de factores de escala para la escala de los valores espectrales). Por ejemplo, en el caso de TCX, no enviamos factores de escala (desde un codificador hacia un decodificador), sino la codificación predictiva lineal (LPC), y luego, en el decodificador, transformamos la codificación predictiva lineal (LPC) en una representación de factor de escala para los bins de la Transformada de coseno discreta modificada (MDCT). En otras palabras, en el caso de TCX, enviamos el coeficiente de codificación predictiva lineal (LPC), y luego, en el decodificador, transformamos dichos coeficientes de codificación predictiva lineal (LPC) en una representación de factor de escala para TCX en USAC o en AMR-WB+ donde no hay ningún factor de escala.

45 En una realización preferida, el decodificador de audio comprende un núcleo de decodificador de dominio de frecuencia configurado para la aplicación de una escala sobre la base de factores de escala, a una pluralidad de valores espectrales derivados de la representación de dominio de frecuencia. En este caso, el ocultamiento de error está configurado para proporcionar la información de audio de ocultamiento de error para el ocultamiento de una pérdida de una trama de audio tras una trama de audio codificada en la representación de dominio de frecuencia que comprende una pluralidad de factores de escala codificados usando una señal de excitación de dominio de tiempo derivada de la representación de dominio de frecuencia. Esta realización según la invención se sustenta en el hallazgo de que la derivación de la señal de excitación de dominio de tiempo a partir de la arriba mencionada representación de dominio de frecuencia típicamente proporciona un mejor resultado de ocultamiento de error en comparación con un ocultamiento de error realizado directamente en el dominio de frecuencia. Por ejemplo, la señal de excitación es creada sobre la base de la síntesis de la trama previa; entonces, no importa realmente si la trama previa es una trama de dominio de frecuencia (MDCT (Transformada de coseno discreta modificada), FFT (Transformada de Fourier Rápida, conforme a sus siglas en inglés...)) o una trama de dominio de tiempo. Sin embargo, pueden observarse ventajas particulares si la trama previa era un dominio de frecuencia. Aún más, debe observarse que se logran resultados particularmente buenos, por ejemplo, para señal monofónica como voz. A modo de otro ejemplo, los factores de escala podrían ser transmitidos como coeficientes de codificación predictiva lineal (LPC), por ejemplo, usando una representación polinomial que es luego convertida en factores de escala del lado del decodificador.

En una realización preferida, el decodificador de audio comprende un núcleo de decodificador de dominio de frecuencia configurado para derivar una representación de señal de audio de dominio de tiempo a partir de la

representación de dominio de frecuencia sin el uso de una señal de excitación de dominio de tiempo como una cantidad intermedia para la trama de audio codificada en la representación de dominio de frecuencia. En otras palabras, se ha hallado que el uso de una señal de excitación de dominio de tiempo para un ocultamiento de error es conveniente incluso si la trama de audio que precede la trama de audio perdida está codificada en un modo de frecuencia "real" que no usa ninguna señal de excitación de dominio de tiempo como una cantidad intermedia (y que, en consecuencia, no se sustenta en una síntesis de codificación predictiva lineal (LPC)).

En una realización preferida, el ocultamiento de error está configurado para obtener la señal de excitación de dominio de tiempo sobre la base de la trama de audio codificada en la representación de dominio de frecuencia que precede una trama de audio perdida. En este caso, el ocultamiento de error está configurado para proporcionar la información de audio de ocultamiento de error para el ocultamiento de la trama de audio perdida usando dicha señal de excitación de dominio de tiempo. En otras palabras, se ha reconocido que la señal de excitación de dominio de tiempo, que se usa para el ocultamiento de error, debería derivar de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida, ya que esta señal de excitación de dominio de tiempo derivada de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida proporciona una buena representación de un contenido de audio de la trama de audio que precede la trama de audio perdida, de modo tal que el ocultamiento de error pueda ser realizado con moderado esfuerzo y buena exactitud.

En una realización preferida, el ocultamiento de error está configurado para realizar un análisis de codificación predictiva lineal (LPC) sobre la base de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida, a fin de obtener un conjunto de parámetros de codificación de predicción lineal y la señal de excitación de dominio de tiempo que representa un contenido de audio de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida. Se ha hallado que es conveniente el esfuerzo de realizar un análisis de codificación predictiva lineal (LPC), para derivar los parámetros de codificación de predicción lineal y la señal de excitación de dominio de tiempo, Aún si la trama de audio que precede la trama de audio perdida está codificada en una representación de dominio de frecuencia (que no contiene ningún parámetro de codificación de predicción lineal y ninguna representación de una señal de excitación de dominio de tiempo), debido a que puede obtenerse una información de audio de ocultamiento de error de buena calidad para muchas señales de audio de entrada sobre la base de dicha señal de excitación de dominio de tiempo. Como alternativa, la ocultación de error puede configurarse para realizar un análisis de LPC basándose en la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida, para obtener la señal de excitación de dominio de tiempo que representa un contenido de audio de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida. Además, alternativamente, el decodificador de audio puede estar configurado para obtener un conjunto de parámetros de codificación de predicción lineal usando una estimación de parámetros de codificación de predicción lineal, o el decodificador de audio puede estar configurado para obtener un conjunto de parámetros de codificación de predicción lineal sobre la base de un conjunto de factores de escala usando una transformada. En otras palabras, los parámetros de codificación predictiva lineal (LPC) pueden obtenerse usando la estimación de parámetros de codificación predictiva lineal (LPC). Esto podría realizarse o bien mediante el ventaneado/autocorr/levinsondurbin sobre la base de la trama de audio codificada en la representación de dominio de frecuencia o mediante la transformación a partir del factor de escala previo directamente a la representación de codificación predictiva lineal (LPC).

En una realización preferida, el ocultamiento de error está configurado para obtener una información de altura (o retardo) que describe una altura de la trama de audio codificada en el dominio de frecuencia que precede la trama de audio perdida, y para proporcionar la información de audio de ocultamiento de error según la información de altura. Al considerar la información de altura, puede lograrse que la información de audio de ocultamiento de error (que típicamente es una señal de audio de ocultamiento de error que cubre la duración temporal de por lo menos una trama de audio perdida) se adapte bien al contenido de audio real.

En una realización preferida, el ocultamiento de error está configurado para obtener la información de altura sobre la base de la señal de excitación de dominio de tiempo derivada de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida. Se ha hallado que una derivación de la información de altura a partir de la señal de excitación de dominio de tiempo acarrea una alta exactitud. Aún más, se ha hallado que es conveniente si la información de altura se adapta bien a la señal de excitación de dominio de tiempo, ya que la información de altura se usa para una modificación de la señal de excitación de dominio de tiempo. Al derivar la información de altura de la señal de excitación de dominio de tiempo, puede lograrse dicha relación cercana.

En una realización preferida, el ocultamiento de error está configurado para evaluar una correlación cruzada de la señal de excitación de dominio de tiempo, a fin de determinar una información de altura aproximada. Aún más, el ocultamiento de error puede estar configurado para refinar la información de altura aproximada usando una búsqueda de circuito cerrado alrededor de una altura determinada por la información de altura aproximada. En consecuencia, puede lograrse una información de altura altamente exacta con moderado esfuerzo computacional.

En una realización preferida, el ocultamiento de error del decodificador de audio puede estar configurado para obtener una información de altura sobre la base de una información lateral de la información de audio codificada.

- 5 En una realización preferida, el ocultamiento de error puede estar configurado para obtener una información de altura sobre la base de una información de altura disponible para una trama de audio decodificada previamente.

En una realización preferida, el ocultamiento de error está configurado para obtener una información de altura sobre la base de una búsqueda de altura realizada sobre una señal de dominio de tiempo o sobre una señal residual.

- 10 En otras palabras, la altura puede ser transmitida como información lateral o podría provenir también de la trama previa si hay LTP, por ejemplo. La información de altura además podría ser transmitida en la corriente de bits si está disponible en el codificador. Podemos hacer opcionalmente la búsqueda de altura en la señal de dominio de tiempo directamente o en la residual, lo que habitualmente proporciona mejores resultados sobre la residual (señal de excitación de dominio de tiempo).

En una realización preferida, el ocultamiento de error está configurado para copiar un ciclo de altura de la señal de excitación de dominio de tiempo derivada de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida una vez o múltiples veces, a fin de obtener una señal de excitación para una síntesis de la señal de audio de ocultamiento de error. Mediante el copiado de la señal de excitación de dominio de tiempo una vez o múltiples veces, puede lograrse que el componente determinista (es decir, sustancialmente periódico) de la información de ocultamiento de error audio se obtenga con buena exactitud, y sea una buena continuación del componente determinista (por ejemplo, sustancialmente periódico) del contenido de audio de la trama de audio que precede la trama de audio perdida.

- 25 En una realización preferida, el ocultamiento de error está configurado para filtrar en paso bajo el ciclo de altura de la señal de excitación de dominio de tiempo derivada de la representación de dominio de frecuencia de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida usando un filtro dependiente de la tasa de muestreo, cuyo ancho de banda depende de una tasa de muestreo de la trama de audio codificada en una representación de dominio de frecuencia. En consecuencia, la señal de excitación de dominio de tiempo puede ser adaptada para un ancho de banda de audio disponible, lo que produce una buena impresión auditiva de la información de audio de ocultamiento de error. Por ejemplo, se prefiere el paso bajo solo sobre la primera trama perdida, y preferentemente, además, el paso bajo solo si la señal no es 100 % estable. Sin embargo, debe observarse que la filtración de paso bajo es opcional, y puede realizarse solo sobre el primer ciclo de altura. Por ejemplo, el filtro puede depender de la tasa de muestreo, de modo tal que la frecuencia de corte sea independiente del ancho de banda.

En una realización preferida, el ocultamiento de error está configurado para predecir una altura en un extremo de una trama perdida a fin de adaptar la señal de excitación de dominio de tiempo, o una o más de sus copias, a la altura predicha. En consecuencia, los cambios de altura esperados durante la trama de audio perdida pueden ser considerados. En consecuencia, se evitan fallas en una transición entre la información de audio de ocultamiento de error y una información de audio de una trama apropiadamente decodificada tras una o más tramas de audio perdidas (o por lo menos se reducen, ya que es solo una trama predicha, no la real). Por ejemplo, la adaptación va desde la última buena altura hasta aquella predicha. Esto se realiza por medio de la resincronización de pulsos [7].

- 45 En una realización preferida, el ocultamiento de error está configurado para combinar una señal de excitación de dominio de tiempo extrapolada y una señal de ruido, a fin de obtener una señal de entrada para una síntesis de codificación predictiva lineal (LPC). En este caso, el ocultamiento de error está configurado para realizar la síntesis de codificación predictiva lineal (LPC), donde la síntesis de codificación predictiva lineal (LPC) está configurada para filtrar la señal de entrada de la síntesis de codificación predictiva lineal (LPC) según parámetros de codificación de predicción lineal, a fin de obtener la información de audio de ocultamiento de error. En consecuencia, tanto un componente determinista (por ejemplo, aproximadamente periódico) del contenido de audio como un componente de tipo ruido del contenido de audio pueden ser considerados. Por lo tanto, se logra que la información de audio de ocultamiento de error comprenda una impresión auditiva "natural".

- 55 En una realización preferida, el ocultamiento de error está configurado para computar una ganancia de la señal de excitación de dominio de tiempo extrapolada, que se usa para obtener la señal de entrada para la síntesis de codificación predictiva lineal (LPC), usando una correlación en el dominio de tiempo que se realiza sobre la base de una representación de dominio de tiempo de la trama de audio codificada en el dominio de frecuencia que precede la trama de audio perdida, donde se establece un retardo de correlación dependiente de una información de altura obtenida sobre la base de la señal de excitación de dominio de tiempo. En otras palabras, una intensidad de un componente periódico se determina dentro de la trama de audio que precede la trama de audio perdida, y esta intensidad determinada del componente periódico se usa para obtener la información de audio de ocultamiento de error. Sin embargo, se ha hallado que la arriba mencionada computación de la intensidad del componente periódico



proporciona resultados particularmente buenos, ya que se considera la señal de audio de dominio de tiempo real de la trama de audio que precede la trama de audio perdida. Alternativamente, puede obtenerse una correlación en el dominio de excitación o directamente en el dominio de tiempo a fin de obtener la información de altura. Sin embargo, hay también diferentes posibilidades, según la realización utilizada. En una realización, la información de altura  
5 podría ser solo la altura obtenida desde la ltp de última trama, o la altura que es transmitida como información lateral o aquella calculada.

En una realización preferida, el ocultamiento de error está configurado para el filtro de paso alto de la señal de ruido que está combinada con la señal de excitación de dominio de tiempo extrapolada. Se ha hallado que la filtración de  
10 paso alto de la señal de ruido (que típicamente es ingresada en la síntesis de codificación predictiva lineal (LPC)) logra una impresión auditiva natural. Por ejemplo, la característica de paso alto puede cambiar con la cantidad de trama perdida, tras una cierta cantidad de pérdida de trama ya no puede haber paso alto. La característica de paso alto además puede depender de la tasa de muestreo con la que se ejecuta el decodificador. Por ejemplo, el paso  
15 alto depende de la tasa de muestreo, y la característica de filtro puede cambiar en función del tiempo (sobre pérdida de trama consecutiva). La característica de paso alto además puede cambiar opcionalmente sobre pérdida de trama consecutiva, de modo tal que tras una cierta cantidad de pérdida de trama, ya no hay filtración, para solo obtener el ruido de forma de banda completa de modo de obtener un buen ruido de confort cercano al ruido de fondo.

En una realización preferida, el ocultamiento de error está configurado para cambiar selectivamente la forma  
20 espectral de la señal (562) de ruido usando el filtro de preénfasis donde la señal de ruido se combina con la señal de excitación de dominio de tiempo extrapolada si la trama de audio codificada en una representación de dominio de frecuencia que precede la trama de audio perdida es una trama de audio con voz o comprende un inicio. Se ha hallado que la impresión auditiva de la información de audio de ocultamiento de error puede ser mejorada por dicho concepto. Por ejemplo, en algún caso, es mejor disminuir las ganancias y la forma, y en algún lugar, es mejor  
25 incrementarlas.

En una realización preferida, el ocultamiento de error está configurado para computar una ganancia de la señal de ruido según una correlación en el dominio de tiempo, que se realiza sobre la base de una representación de dominio  
30 de tiempo de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida. Se ha hallado que dicha determinación de la ganancia de la señal de ruido proporciona resultados particularmente exactos, ya que puede considerarse la señal de audio de dominio de tiempo real asociada con la trama de audio que precede la trama de audio perdida. Usando este concepto, es posible poder obtener una energía de la trama ocultada cercana a la energía de la trama buena previa. Por ejemplo, la ganancia para la señal de ruido puede ser generada midiendo la energía del resultado: excitación de señal de entrada – excitación sobre la base de  
35 altura generada.

En una realización preferida, el ocultamiento de error está configurado para modificar una señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida, a fin de obtener la información de audio de ocultamiento de error. Se ha hallado que la modificación de la señal de  
40 excitación de dominio de tiempo permite la adaptación de la señal de excitación de dominio de tiempo a una evolución temporal deseada. Por ejemplo, la modificación de la señal de excitación de dominio de tiempo permite el “desvanecimiento saliente” del componente determinista (por ejemplo, sustancialmente periódico) del contenido de audio en la información de audio de ocultamiento de error. Aún más, la modificación de la señal de excitación de dominio de tiempo además permite adaptar la señal de excitación de dominio de tiempo a una variación de altura (estimada o esperada). Esto permite el ajuste de las características de la información de audio de ocultamiento de  
45 error en función del tiempo.

En una realización preferida, el ocultamiento de error está configurado para usar una o más copias modificadas de la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una  
50 trama de audio perdida, a fin de obtener la información de ocultamiento de error. Las copias modificadas de la señal de excitación de dominio de tiempo pueden obtenerse con un esfuerzo moderado, y la modificación puede ser realizada usando un simple algoritmo. En consecuencia, las características deseadas de la información de audio de ocultamiento de error pueden lograrse con esfuerzo moderado.

En una realización preferida, el ocultamiento de error está configurado para modificar la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, a fin de reducir un componente periódico de la información de audio de ocultamiento de error en función del tiempo. En consecuencia, puede considerarse que la correlación entre el contenido de audio de la trama de audio que precede la trama de audio perdida y el contenido de audio de una o más tramas de audio  
60 perdidas disminuye en función del tiempo. Además, puede evitarse causar una impresión auditiva no natural por una larga preservación de un componente periódico de la información de audio de ocultamiento de error.

En una realización preferida, el ocultamiento de error está configurado para escalar la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden la trama de audio perdida, o

una o más de sus copias, de manera de modificar la señal de excitación de dominio de tiempo. Se ha hallado que la operación de escala puede realizarse con poco esfuerzo, donde la señal de excitación de dominio de tiempo escalada típicamente proporciona una buena información de audio de ocultamiento de error.

- 5 En una realización preferida, el ocultamiento de error está configurado para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias. Por consiguiente, puede lograrse un desvanecimiento saliente del componente periódico dentro de la información de audio de ocultamiento de error.
- 10 En una realización preferida, el ocultamiento de error está configurado para ajustar una velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, según uno o más parámetros de una o más tramas de audio que preceden la trama de audio perdida, y/o según una cantidad de tramas de audio perdidas consecutivas. Por consiguiente, es posible ajustar la velocidad a la cual el
- 15 componente determinista (por ejemplo, por lo menos aproximadamente periódico) es desvanecido en forma saliente en la información de audio de ocultamiento de error. La velocidad del desvanecimiento saliente puede ser adaptada a características específicas del contenido de audio, que típicamente pueden observarse a partir de uno o más parámetros de una o más tramas de audio que preceden la trama de audio perdida. Alternativamente, o además, la cantidad de tramas de audio perdidas consecutivas puede ser considerada cuando se determina la velocidad
- 20 utilizada para el desvanecimiento saliente del componente determinista (por ejemplo, por lo menos aproximadamente periódico) de la información de audio de ocultamiento de error, que ayuda a adaptar el ocultamiento de error a la situación específica. Por ejemplo, la ganancia de la parte tonal y la ganancia de la parte ruidosa pueden desvanecerse en forma saliente de manera separada. La ganancia para la parte tonal puede converger a cero tras una cierta cantidad de pérdida de trama, mientras que la ganancia de ruido puede converger a
- 25 la ganancia determinada para alcanzar un cierto ruido de confort.

En una realización preferida, el ocultamiento de error está configurado para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, según

30 una longitud de un período de altura de la señal de excitación de dominio de tiempo, de modo tal que una entrada de señal de excitación de dominio de tiempo en una síntesis de codificación predictiva lineal (LPC) es desvanecida en forma saliente más rápidamente para señales que tienen una longitud más corta del período de altura en comparación con señales que tienen una longitud más larga del período de altura. Por consiguiente, puede evitarse que las señales que tienen una longitud más corta del período de altura sean repetidas con demasiada frecuencia

35 con alta intensidad, ya que esto típicamente resultará en una impresión auditiva no natural. En consecuencia, una calidad general de la información de audio de ocultamiento de error puede ser mejorada.

En una realización preferida, el ocultamiento de error está configurado para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, según

40 un resultado de un análisis de altura o una predicción de altura, de modo tal que un componente determinista de la entrada de señal de excitación de dominio de tiempo en una síntesis de codificación predictiva lineal (LPC) se desvanece en forma saliente más rápidamente para señales que tienen un mayor cambio de altura por unidad de tiempo en comparación con señales que tienen un cambio de altura menor por unidad de tiempo, y/o de modo tal

45 que un componente determinista de la entrada de señal de excitación de dominio de tiempo en una síntesis de codificación predictiva lineal (LPC) se desvanece en forma saliente más rápidamente para señales para las cuales una predicción de altura falla en comparación con señales para las cuales la predicción de altura tiene éxito. Por consiguiente, el desvanecimiento saliente puede hacerse más rápidamente para señales en las cuales hay una gran incertidumbre de la altura en comparación con señales para las cuales hay una menor incertidumbre de la altura. Sin embargo, mediante el desvanecimiento saliente de un componente determinista más rápidamente para señales que comprenden una incertidumbre comparativamente grande de la altura, pueden evitarse fallas audibles, o al menos pueden reducirse sustancialmente.

En una realización preferida, el ocultamiento de error está configurado para la escala de tiempo de la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, según una predicción de una altura para el tiempo de una o más tramas de audio perdidas. Por consiguiente, la señal de excitación de dominio de tiempo puede ser adaptada a una altura variable, de modo tal que la información de audio de ocultamiento de error comprende una impresión auditiva más

55 natural.

60 En una realización preferida, el ocultamiento de error está configurado para proporcionar la información de audio de ocultamiento de error para un tiempo que es mayor que una duración temporal de una o más tramas de audio perdidas. Por consiguiente, es posible realizar una operación de superposición y adición sobre la base de la información de audio de ocultamiento de error, que ayuda a reducir las fallas de bloqueo.

5 En una realización preferida, el ocultamiento de error está configurado para realizar una superposición y adición de la información de audio de ocultamiento de error y de una representación de dominio de tiempo de una o más tramas de audio apropiadamente recibidas tras una o más tramas de audio perdidas. En consecuencia, es posible evitar (o al menos reducir) las fallas de bloqueo.

10 En una realización preferida, el ocultamiento de error está configurado para derivar la información de audio de ocultamiento de error sobre la base de por lo menos tres ventanas o tramas parcialmente superpuestas que preceden una trama de audio perdida o una ventana perdida. Por consiguiente, la información de audio de ocultamiento de error puede ser obtenida con buena exactitud, aún para modos de codificación en los cuales más de dos tramas (o ventanas) son superpuestas (donde dicha superposición puede ayudar a reducir una demora).

15 Otra realización según la invención crea un método para proporcionar una información de audio decodificada sobre la base de una información de audio codificada. El método comprende la provisión de una información de audio de ocultamiento de error para el ocultamiento de una pérdida de una trama de audio tras una trama de audio codificada en una representación de dominio de frecuencia usando una señal de excitación de dominio de tiempo. Este método se sustenta en las mismas consideraciones que el arriba mencionado decodificador de audio.

20 Aún otra realización según la invención crea un programa de ordenador para la realización de dicho método cuando el programa de ordenador se ejecuta en un ordenador.

25 Otra realización según la invención crea un decodificador de audio para proporcionar una información de audio decodificada sobre la base de una información de audio codificada. El decodificador de audio comprende un ocultamiento de error configurado para proporcionar una información de audio de ocultamiento de error para el ocultamiento de una pérdida de una trama de audio. El ocultamiento de error está configurado para modificar una señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida, a fin de obtener la información de audio de ocultamiento de error.

30 Esta realización según la invención se sustenta en la idea de que un ocultamiento de error con una buena calidad de audio puede obtenerse sobre la base de una señal de excitación de dominio de tiempo, donde una modificación de la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida permite una adaptación de la información de audio de ocultamiento de error a cambios esperados (o predichos) del contenido de audio durante la trama perdida. Por consiguiente, pueden evitarse las fallas, y en particular, una impresión auditiva no natural, que serían causadas por un uso no cambiado de la señal de excitación de dominio de tiempo. En consecuencia, se logra una provisión mejorada de una información de audio de ocultamiento de error, de modo tal que las tramas de audio perdidas pueden ser ocultadas con resultados mejorados.

40 En una realización preferida, el ocultamiento de error está configurado para usar una o más copias modificadas de la señal de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, a fin de obtener la información de ocultamiento de error. Al usar una o más copias modificadas de la señal de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, puede lograrse una buena calidad de la información de audio de ocultamiento de error, con poco esfuerzo computacional.

45 En una realización preferida, el ocultamiento de error está configurado para modificar la señal de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, a fin de reducir un componente periódico de la información de audio de ocultamiento de error en función del tiempo. Al reducir el componente periódico de la información de audio de ocultamiento de error en función del tiempo, puede evitarse una conservación artificialmente larga de un sonido determinista (por ejemplo, aproximadamente periódico), lo que ayuda a hacer natural el sonido de la información de audio de ocultamiento de error.

55 En una realización preferida, el ocultamiento de error está configurado para escalar la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden la trama de audio perdida, o una o más de sus copias, de manera de modificar la señal de excitación de dominio de tiempo. La escala de la señal de excitación de dominio de tiempo constituye una manera particularmente eficiente para variar la información de audio de ocultamiento de error en función del tiempo.

60 En una realización preferida, el ocultamiento de error está configurado para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias. Se ha hallado que la reducción gradual de la ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, permite obtener una señal de

excitación de dominio de tiempo para la provisión de la información de audio de ocultamiento de error, de modo tal que los componentes deterministas (por ejemplo, componentes al menos aproximadamente periódicos) son desvanecidos en forma saliente. Por ejemplo, puede no haber solo una ganancia. Por ejemplo, se podría tener una ganancia para la parte tonal (también referida como la parte aproximadamente periódica), y una ganancia para la parte de ruido. Ambas excitaciones (o componentes de excitación) pueden ser atenuadas separadamente con diferente factor de velocidad, y luego, las dos excitaciones resultantes (o componentes de excitación) pueden combinarse antes de alimentarse a la codificación predictiva lineal (LPC) para la síntesis. En el caso de no tener ninguna estimación de ruido de fondo, los factores de desvanecimiento saliente para el ruido y para la parte tonal pueden ser similares, y entonces, se podría tener solo una aplicación de desvanecimiento saliente sobre los resultados de las dos excitaciones, multiplicada con su propia ganancia y combinados entre sí.

Por lo tanto, puede evitarse que la información de audio de ocultamiento de error comprenda un componente de audio determinista (por ejemplo, por lo menos aproximadamente periódico) temporalmente extendido, que típicamente proporcionaría una impresión auditiva no natural.

En una realización preferida, el ocultamiento de error está configurado para ajustar una velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, según uno o más parámetros de una o más tramas de audio que preceden la trama de audio perdida, y/o según una cantidad de tramas de audio perdidas consecutivas. Por lo tanto, la velocidad del desvanecimiento saliente del componente determinista (por ejemplo, por lo menos aproximadamente periódico) en la información de audio de ocultamiento de error puede adaptarse a la situación específica, con moderado esfuerzo computacional. Debido a que la señal de excitación de dominio de tiempo utilizada para la provisión de la información de audio de ocultamiento de error es típicamente una versión escalada (escalada usando la ganancia mencionada con anterioridad) de la señal de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden la trama de audio perdida, una variación de dicha ganancia (utilizada para derivar la señal de excitación de dominio de tiempo para la provisión de la información de audio de ocultamiento de error) constituye un simple, aunque eficaz método para adaptar la información de audio de ocultamiento de error a las necesidades específicas. Sin embargo, la velocidad del desvanecimiento saliente además es controlable con muy poco esfuerzo.

En una realización preferida, el ocultamiento de error está configurado para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, según una longitud de un período de altura de la señal de excitación de dominio de tiempo, de modo tal que una entrada de señal de excitación de dominio de tiempo en una síntesis de codificación predictiva lineal (LPC) es desvanecida en forma saliente más rápidamente para señales que tienen una longitud más corta del período de altura en comparación con señales que tienen una longitud más larga del período de altura. Por consiguiente, el desvanecimiento en forma saliente se realiza en forma más veloz para señales que tienen una longitud más corta del período de altura, lo que evita que un período de altura sea copiado demasiadas veces (lo que habitualmente lograría una impresión auditiva no natural).

En una realización preferida, el ocultamiento de error está configurado para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, según un resultado de un análisis de altura o una predicción de altura, de modo tal que un componente determinista de una entrada de señal de excitación de dominio de tiempo en una síntesis de codificación predictiva lineal (LPC) es desvanecida en forma saliente más rápidamente para señales que tienen un mayor cambio de altura por unidad de tiempo, en comparación con señales que tienen un menor cambio de altura por unidad de tiempo, y/o de modo tal que un componente determinista de una entrada de señal de excitación de dominio de tiempo en una síntesis de codificación predictiva lineal (LPC) es desvanecida en forma saliente más rápidamente para señales para las cuales falla una predicción de altura en comparación con señales para las cuales la predicción de altura tiene éxito. Por consiguiente, un componente determinista (por ejemplo, por lo menos aproximadamente periódico) es desvanecido en forma saliente más rápidamente para señales para las cuales hay una mayor incertidumbre de la altura (donde un mayor cambio de altura por unidad de tiempo, o incluso, un fallo de la predicción de altura, indica una incertidumbre comparativamente grande de la altura). En consecuencia, pueden evitarse las fallas, que surgirían de la provisión de una información de audio de ocultamiento de error altamente determinista en una situación en la cual la altura real es incierta.

En una realización preferida, el ocultamiento de error está configurado para la escala de tiempo de la señal de excitación de dominio de tiempo obtenida para (o sobre la base de) una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, según una predicción de una altura para el tiempo de las una o más tramas de audio perdidas. Por consiguiente, la señal de excitación de dominio de tiempo, que se usa para la provisión de la información de audio de ocultamiento de error, es modificada (en comparación con la señal de excitación de dominio de tiempo obtenida para (o sobre la base de) una o más tramas de audio que preceden una

trama de audio perdida, de modo tal que la altura de la señal de excitación de dominio de tiempo sigue los requerimientos de un período de tiempo de la trama de audio perdida. En consecuencia, puede mejorarse la impresión auditiva, que puede lograrse por la información de audio de ocultamiento de error.

5 En una realización preferida, el ocultamiento de error está configurado para obtener una señal de excitación de dominio de tiempo, que ha sido utilizada para la decodificación de una o más tramas de audio que preceden la trama de audio perdida, y para la modificación de dicha señal de excitación de dominio de tiempo, que ha sido utilizada para la decodificación de una o más tramas de audio que preceden la trama de audio perdida, a fin de obtener una señal de excitación de dominio de tiempo modificada. En este caso, el ocultamiento de dominio de tiempo está  
10 configurado para proporcionar la información de audio de ocultamiento de error sobre la base de la señal de audio de dominio de tiempo modificada. Por consiguiente, es posible reutilizar una señal de excitación de dominio de tiempo, que ya ha sido utilizada para decodificar una o más tramas de audio que preceden la trama de audio perdida. En consecuencia, puede mantenerse un esfuerzo computacional muy pequeño, si la señal de excitación de dominio de tiempo ya ha sido adquirida para la decodificación de una o más tramas de audio que preceden la trama de audio perdida.  
15

En una realización preferida, el ocultamiento de error está configurado para obtener una información de altura, que ha sido utilizada para la decodificación de una o más tramas de audio que preceden la trama de audio perdida. En este caso, el ocultamiento de error además está configurado para proporcionar la información de audio de  
20 ocultamiento de error según dicha información de altura. Por consiguiente, la información de altura previamente utilizada puede ser reusada, lo que evita un esfuerzo computacional para un nuevo cómputo de la información de altura. Por lo tanto, el ocultamiento de error es en particular computacionalmente eficiente. Por ejemplo, en el caso de ACELP, tenemos 4 demoras de altura y ganancias por trama. Podemos usar las últimas dos tramas para poder predecir la altura al final de la trama que tenemos que ocultar.  
25

A continuación, comparamos con el códec de dominio de frecuencia previamente descrito donde solo se derivan una o dos alturas por trama (podemos tener más de dos, si bien esto agregaría mucha complejidad para una ganancia no muy grande en calidad). En el caso de un códec de conmutación que es, por ejemplo, ACELP – FD – pérdida, entonces, tenemos una precisión de altura mucho mejor, ya que la altura se transmite en la corriente de bits y se  
30 basa en la señal de entrada original (no en la decodificada, como se hace en el decodificador). En el caso de alta tasa de bits, por ejemplo, podemos además enviar una información de demora de altura y ganancia, o información LTP, por trama codificada de dominio de frecuencia.

En una realización preferida, el ocultamiento de error del decodificador de audio puede estar configurado para  
35 obtener una información de altura sobre la base de una información lateral de la información de audio codificada.

En una realización preferida, el ocultamiento de error puede estar configurado para obtener una información de altura sobre la base de una información de altura disponible para una trama de audio decodificada previamente.

40 En una realización preferida, el ocultamiento de error está configurado para obtener una información de altura sobre la base de una búsqueda de altura realizada sobre una señal de dominio de tiempo o sobre una señal residual.

En otras palabras, la altura puede ser transmitida como información lateral o podría provenir también de la trama previa si hay LTP, por ejemplo. La información de altura además podría ser transmitida en la corriente de bits si está  
45 disponible en el codificador. Podemos hacer opcionalmente la búsqueda de altura en la señal de dominio de tiempo directamente o en la residual, lo que habitualmente proporciona mejores resultados sobre la residual (señal de excitación de dominio de tiempo).

En una realización preferida, el ocultamiento de error está configurado para obtener un conjunto de coeficientes de predicción lineal, que se han utilizado para decodificar una o más tramas de audio que preceden la trama de audio  
50 perdida. En este caso, el ocultamiento de error está configurado para proporcionar la información de audio de ocultamiento de error según dicho conjunto de coeficientes de predicción lineal. En consecuencia, la eficiencia del ocultamiento de error se incrementa por medio de la reutilización de información previamente generada (o previamente decodificada), por ejemplo, el conjunto previamente utilizado de coeficientes de predicción lineal. En  
55 consecuencia, se evita la alta complejidad computacional innecesaria.

En una realización preferida, el ocultamiento de error está configurado para extrapolar un nuevo conjunto de coeficientes de predicción lineal sobre la base del conjunto de coeficientes de predicción lineal, que se han utilizado para decodificar una o más tramas de audio que preceden la trama de audio perdida. En este caso, el ocultamiento  
60 de error está configurado para usar el nuevo conjunto de coeficientes de predicción lineal a fin de proporcionar la información de ocultamiento de error. Al derivar el nuevo conjunto de coeficientes de predicción lineal, utilizados para proporcionar la información de audio de ocultamiento de error, a partir de un conjunto de coeficientes de predicción lineal previamente utilizados usando una extrapolación, puede evitarse un recálculo completo de los coeficientes de predicción lineal, lo que ayuda a mantener el esfuerzo computacional razonablemente bajo. Aún

más, al realizar una extrapolación sobre la base del conjunto previamente utilizado de coeficientes de predicción lineal, puede garantizarse que el nuevo conjunto de coeficientes de predicción lineal es al menos similar al conjunto previamente utilizado de coeficientes de predicción lineal, lo que ayuda a evitar discontinuidades cuando se proporciona la información de ocultamiento de error. Por ejemplo, tras una cierta cantidad de pérdida de trama, tendemos a estimar la forma de la codificación predictiva lineal (LPC) de ruido de fondo. La velocidad de esta convergencia, por ejemplo, puede depender de la característica de señal.

En una realización preferida, el ocultamiento de error está configurado para obtener una información acerca de una intensidad de un componente de señal determinista en una o más tramas de audio que preceden una trama de audio perdida. En este caso, el ocultamiento de error está configurado para comparar la información acerca de una intensidad de un componente de señal determinista en una o más tramas de audio que preceden una trama de audio perdida con un valor de umbral, a fin de decidir si ingresar un componente determinista de una señal de excitación de dominio de tiempo en una síntesis de codificación predictiva lineal (LPC) (síntesis sobre la base del coeficiente de predicción lineal), o si ingresar solo un componente de ruido de una señal de excitación de dominio de tiempo en la síntesis de codificación predictiva lineal (LPC). Por consiguiente, es posible omitir la provisión de un componente determinista (por ejemplo, por lo menos aproximadamente periódico) de la información de audio de ocultamiento de error en el caso que solo haya una pequeña contribución de señal determinista dentro de una o más tramas que preceden la trama de audio perdida. Se ha hallado que esto ayuda a obtener una buena impresión auditiva.

En una realización preferida, el ocultamiento de error está configurado para obtener una información de altura que describe una altura de la trama de audio que precede la trama de audio perdida, y proporcionar la información de audio de ocultamiento de error según la información de altura. Por consiguiente, es posible adaptar la altura de la información de ocultamiento de error a la altura de la trama de audio que precede la trama de audio perdida. Por consiguiente, se evitan las discontinuidades, y puede lograrse una impresión auditiva natural.

En una realización preferida, el ocultamiento de error está configurado para obtener la información de altura sobre la base de la señal de excitación de dominio de tiempo asociada con la trama de audio que precede la trama de audio perdida. Se ha hallado que la información de altura obtenida sobre la base de la señal de excitación de dominio de tiempo es particularmente confiable, y además, se adapta muy bien al procesamiento de la señal de excitación de dominio de tiempo.

En una realización preferida, el ocultamiento de error está configurado para evaluar una correlación cruzada de la señal de excitación de dominio de tiempo (o, alternativamente, de una señal de audio de dominio de tiempo), a fin de determinar una información de altura aproximada, y refinar la información de altura aproximada usando una búsqueda de circuito cerrado alrededor de una altura determinada (o descrita) por la información de altura aproximada. Se ha hallado que este concepto permite obtener una muy precisa información de altura con moderado esfuerzo computacional. En otras palabras, en algunos códecs, hacemos la búsqueda de altura directamente sobre la señal de dominio de tiempo, mientras que, en algunos otros, hacemos la búsqueda de altura sobre la señal de excitación de dominio de tiempo.

En una realización preferida, el ocultamiento de error está configurado para obtener la información de altura para la provisión de la información de audio de ocultamiento de error sobre la base de una información de altura previamente computada, que se usó para la decodificación de una o más tramas de audio que preceden la trama de audio perdida, y sobre la base de una evaluación de una correlación cruzada de la señal de excitación de dominio de tiempo, que se modifica a fin de obtener una señal de excitación de dominio de tiempo modificada para la provisión de la información de audio de ocultamiento de error. Se ha hallado que la consideración tanto de la información de altura previamente computada como de la información de altura obtenida sobre la base de la señal de excitación de dominio de tiempo (usando una correlación cruzada) mejora la confiabilidad de la información de altura, y en consecuencia, ayuda a evitar fallas y/o discontinuidades.

En una realización preferida, el ocultamiento de error está configurado para seleccionar un pico de la correlación cruzada, de una pluralidad de picos de la correlación cruzada, como un pico que representa una altura según la información de altura previamente computada, de modo tal que se selecciona un pico que representa una altura que está más cercana a la altura representada por la información de altura previamente computada. Por consiguiente, pueden superarse posibles ambigüedades de la correlación cruzada, que, por ejemplo, pueden producir múltiples picos. La información de altura previamente computada es así utilizada para seleccionar el pico "apropiado" de la correlación cruzada, lo que ayuda a incrementar de manera sustancial la confiabilidad. Por otra parte, la señal de excitación de dominio de tiempo real se considera principalmente para la determinación de altura, lo que proporciona una buena exactitud (que es sustancialmente mejor que una exactitud que se obtiene sobre la base de solo la información de altura previamente computada).

En una realización preferida, el ocultamiento de error del decodificador de audio puede estar configurado para obtener una información de altura sobre la base de una información lateral de la información de audio codificada.

En una realización preferida, el ocultamiento de error puede estar configurado para obtener una información de altura sobre la base de una información de altura disponible para una trama de audio decodificada previamente.

- 5 En una realización preferida, el ocultamiento de error está configurado para obtener una información de altura sobre la base de una búsqueda de altura realizada sobre una señal de dominio de tiempo o sobre una señal residual.

10 En otras palabras, la altura puede ser transmitida como información lateral o podría provenir también de la trama previa si hay LTP, por ejemplo. La información de altura además podría ser transmitida en la corriente de bits si está disponible en el codificador. Podemos hacer opcionalmente la búsqueda de altura en la señal de dominio de tiempo directamente o en la residual, lo que habitualmente proporciona mejores resultados sobre la residual (señal de excitación de dominio de tiempo).

15 En una realización preferida, el ocultamiento de error está configurado para copiar un ciclo de altura de la señal de excitación de dominio de tiempo asociada con la trama de audio que precede la trama de audio perdida una vez o múltiples veces, a fin de obtener una señal de excitación (o por lo menos uno de sus componentes deterministas) para una síntesis de la información de audio de ocultamiento de error. Al copiar el ciclo de altura de la señal de excitación de dominio de tiempo asociada con la trama de audio que precede la trama de audio perdida una vez o múltiples veces, y al modificar dichas una o más copias usando un algoritmo de modificación comparativamente simple, la señal de excitación (o por lo menos sus componentes deterministas) para la síntesis de la información de audio de ocultamiento de error puede ser obtenida con poco esfuerzo computacional. Sin embargo, la reutilización de la señal de excitación de dominio de tiempo asociada con la trama de audio que precede la trama de audio perdida (al copiar dicha señal de excitación de dominio de tiempo) evita discontinuidades audibles.

25 En una realización preferida, el ocultamiento de error está configurado para el filtro paso bajo del ciclo de altura de la señal de excitación de dominio de tiempo asociada con la trama de audio que precede la trama de audio perdida usando un filtro dependiente de la tasa de muestreo, cuyo ancho de banda depende de una tasa de muestreo de la trama de audio codificada en una representación de dominio de frecuencia. Por consiguiente, la señal de excitación de dominio de tiempo se adapta a un ancho de banda de señal del decodificador de audio, que da como resultado una buena reproducción del contenido de audio. Para detalles y mejoramientos opcionales, se hace referencia, por ejemplo, a las explicaciones anteriores.

35 Por ejemplo, se prefiere el paso bajo de solo la primera trama perdida, y preferentemente, además, hacemos el paso bajo solo si la señal no es sin voz. Sin embargo, debe observarse que la filtración de paso bajo es opcional. Adicionalmente, el filtro puede ser dependiente de la tasa de muestreo, de modo tal que la frecuencia de corte es independiente del ancho de banda.

40 En una realización preferida, el ocultamiento de error está configurado para predecir una altura en un extremo de una trama perdida. En este caso, el ocultamiento de error está configurado para adaptar la señal de excitación de dominio de tiempo, o una o más de sus copias, a la altura predicha. Al modificar la señal de excitación de dominio de tiempo, de modo tal que la señal de excitación de dominio de tiempo que se usa realmente para la provisión de la información de audio de ocultamiento de error es modificada con respecto a la señal de excitación de dominio de tiempo asociada con una trama de audio que precede la trama de audio perdida, pueden considerarse cambios de altura esperados (o predichos) durante la trama de audio perdida, de modo tal que la información de audio de ocultamiento de error se adapta bien a la evolución real (o por lo menos a la evolución esperada o predicha) del contenido de audio. Por ejemplo, la adaptación va desde la última buena altura hasta aquella predicha. Esto se realiza por medio de la resincronización de pulsos [7].

50 En una realización preferida, el ocultamiento de error está configurado para combinar una señal de excitación de dominio de tiempo extrapolada y una señal de ruido, a fin de obtener una señal de entrada para una síntesis de codificación predictiva lineal (LPC). En este caso, el ocultamiento de error está configurado para realizar la síntesis de codificación predictiva lineal (LPC), donde la síntesis de codificación predictiva lineal (LPC) está configurada para filtrar la señal de entrada de la síntesis de codificación predictiva lineal (LPC) según parámetros de codificación de predicción lineal, a fin de obtener la información de audio de ocultamiento de error. Al combinar la señal de excitación de dominio de tiempo extrapolada (que es típicamente una versión modificada de la señal de excitación de dominio de tiempo derivada para una o más tramas de audio que preceden la trama de audio perdida) y una señal de ruido, pueden considerarse tanto componentes deterministas (por ejemplo, aproximadamente periódicos) como componentes de ruido del contenido de audio, en el ocultamiento de error. Por lo tanto, puede lograrse que la información de audio de ocultamiento de error proporcione una impresión auditiva que es similar a la impresión auditiva provista por las tramas que preceden la trama perdida.

Además, al combinar una señal de excitación de dominio de tiempo y una señal de ruido, a fin de obtener la señal de entrada para la síntesis de codificación predictiva lineal (LPC) (que puede considerarse una señal de excitación de dominio de tiempo combinada), es posible variar un porcentaje del componente determinista de la señal de audio de

5 entrada para la síntesis de codificación predictiva lineal (LPC), mientras que se mantiene una energía (de la señal de entrada de la síntesis de codificación predictiva lineal (LPC), o incluso, de la señal de salida de la síntesis de codificación predictiva lineal (LPC)). En consecuencia, es posible variar las características de la información de audio de ocultamiento de error (por ejemplo, las características de tonalidad), sin cambiar sustancialmente una energía o volumen de la señal de audio de ocultamiento de error, de modo tal que es posible modificar la señal de excitación de dominio de tiempo sin causar distorsiones audibles inaceptables.

10 Una realización según la invención crea un método para proporcionar una información de audio decodificada sobre la base de una información de audio codificada. El método comprende la provisión de una información de audio de ocultamiento de error para el ocultamiento de una pérdida de una trama de audio. La provisión de la información de audio de ocultamiento de error comprende la modificación de una señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida, a fin de obtener la información de audio de ocultamiento de error.

15 Este método se sustenta en las mismas consideraciones que el decodificador de audio descrito anteriormente.

Una realización adicional según la invención crea un programa de ordenador para la realización de dicho método, cuando el programa de ordenador se ejecuta en un ordenador.

## 20 **Breve descripción de las figuras**

Las realizaciones de la presente invención se describirán en lo que sigue con referencia a las figuras adjuntas, en las cuales:

25 La figura 1 muestra un diagrama de bloques esquemático de un decodificador de audio, según una realización de la invención;

la figura 2 muestra un diagrama de bloques esquemático de un decodificador de audio, según otra realización de la presente invención;

30 la figura 3 muestra un diagrama de bloques esquemático de un decodificador de audio, según otra realización de la presente invención;

35 la figura 4 muestra un diagrama de bloques esquemático de un decodificador de audio, según otra realización de la presente invención;

la figura 5 muestra un diagrama de bloques esquemático de un ocultamiento de dominio de tiempo para un codificador de transformada;

40 la figura 6 muestra un diagrama de bloques esquemático de un ocultamiento de dominio de tiempo para un códec de conmutación;

la figura 7 muestra un diagrama de bloques de un decodificador TCX que realiza una decodificación TCX en operación normal o en el caso de pérdida de paquete parcial;

45 Fig. 8 muestra un diagrama de bloques esquemático de un decodificador TCX que realiza una decodificación TCX en el caso de ocultación de borrado de paquetes TCX-256;

50 la figura 9 muestra un diagrama de flujo de un método para proporcionar una información de audio decodificada sobre la base de una información de audio codificada, según una realización de la presente invención; y

la figura 10 muestra un diagrama de flujo de un método para proporcionar una información de audio decodificada sobre la base de una información de audio codificada, según otra realización de la presente invención;

55 la figura 11 muestra un diagrama de bloques esquemático de un decodificador de audio, según otra realización de la presente invención.

## **Descripción detallada de las realizaciones.**

### 60 **1. Decodificador de audio según la figura 1**

La figura 1 muestra un diagrama de bloques esquemático de un decodificador 100 de audio, según una realización de la presente invención. El decodificador 100 de audio recibe una información 110 de audio codificada, que, por ejemplo, puede comprender una trama de audio codificada en una representación de dominio de frecuencia. La



información de audio codificada, por ejemplo, puede ser recibida por medio de un canal no confiable, de modo tal que se produce una pérdida de trama de vez en cuando. El decodificador 100 de audio además proporciona, sobre la base de la información 110 de audio codificada, la información 112 de audio decodificada.

5 El decodificador 100 de audio puede comprender una decodificación/procesamiento 120, que proporciona la información de audio decodificada sobre la base de la información de audio codificada en ausencia de una pérdida de trama.

10 El decodificador 100 de audio adicionalmente comprende un ocultamiento 130 de error, que proporciona una información de audio de ocultamiento de error. El ocultamiento 130 de error está configurado para proporcionar la información 132 de audio de ocultamiento de error para el ocultamiento de una pérdida de una trama de audio tras una trama de audio codificada en la representación de dominio de frecuencia, usando una señal de excitación de dominio de tiempo.

15 En otras palabras, la decodificación/procesamiento 120 puede proporcionar una información 122 de audio decodificada para tramas de audio que son codificadas en la forma de una representación de dominio de frecuencia, es decir, en la forma de una representación codificada, cuyos valores codificados describen intensidades en diferentes bins de frecuencia. En otras palabras, la decodificación/procesamiento 120, por ejemplo, puede comprender un decodificador de audio de dominio de frecuencia, que deriva un conjunto de valores espectrales de la información 110 de audio codificada y realiza una transformada de dominio de frecuencia a dominio de tiempo, para de ese modo derivar una representación de dominio de tiempo que constituye la información 122 de audio decodificada, o que forma la base para la provisión de la información 122 de audio decodificada en el caso de que haya procesamiento posterior adicional.

25 Sin embargo, el ocultamiento 130 de error no realiza el ocultamiento de error en el dominio de frecuencia, sino que, en cambio, usa una señal de excitación de dominio de tiempo, que, por ejemplo, puede servir para excitar un filtro de síntesis, por ejemplo, un filtro de síntesis de codificación predictiva lineal (LPC), que proporciona una representación de dominio de tiempo de una señal de audio (por ejemplo, la información de audio de ocultamiento de error) sobre la base de la señal de excitación de dominio de tiempo, y además, sobre la base de coeficientes de filtro de codificación predictiva lineal (LPC) (coeficientes de filtro de codificación de predicción lineal).

30 Por consiguiente, el ocultamiento 130 de error proporciona la información 132 de audio de ocultamiento de error, que, por ejemplo, puede ser una señal de audio de dominio de tiempo, para tramas de audio perdidas, donde la señal de excitación de dominio de tiempo utilizada por el ocultamiento 130 de error puede sustentarse en una o más tramas de audio previas apropiadamente recibidas (que preceden la trama de audio perdida), que son codificadas en la forma de una representación de dominio de frecuencia, o puede derivar de ellas. Como conclusión, el decodificador 100 de audio puede realizar un ocultamiento de error (es decir, proporcionar una información 132 de audio de ocultamiento de error), que reduce una degradación de una calidad de audio debido a la pérdida de una trama de audio sobre la base de una información de audio codificada, donde por lo menos algunas tramas de audio son codificadas en una representación de dominio de frecuencia. Se ha hallado que la realización del ocultamiento de error usando una señal de excitación de dominio de tiempo, aun si una trama tras una trama de audio codificada en la representación de dominio de frecuencia apropiadamente recibida está perdida, acarrea una mejorada calidad de audio en comparación con un ocultamiento de error que se realiza en el dominio de frecuencia (por ejemplo, usando una representación de dominio de frecuencia de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida). Esto se debe a que puede lograrse una transición llana entre la información de audio decodificada asociada con la trama de audio que precede la trama de audio perdida apropiadamente recibida, y la información de audio de ocultamiento de error asociada con la trama de audio perdida, usando una señal de excitación de dominio de tiempo, ya que la síntesis de señal, que se realiza habitualmente sobre la base de la señal de excitación de dominio de tiempo, ayuda a evitar discontinuidades. Por lo tanto, puede lograrse una buena (o por lo menos aceptable) impresión auditiva, usando el decodificador 100 de audio, aun si se pierde una trama de audio que sigue a una trama de audio codificada en la representación de dominio de frecuencia apropiadamente recibida. Por ejemplo, el enfoque de dominio de tiempo produce un mejoramiento sobre la señal monofónica, como voz, ya que está más cerca de lo que se hace en el caso del ocultamiento de códec de voz. El uso de la codificación predictiva lineal (LPC) ayuda a evitar las discontinuidades, y proporciona una mejor forma de las tramas.

Aún más, debe observarse que el decodificador 100 de audio puede ser suplementado por cualquiera de los rasgos y las funcionalidades que se describen en lo que sigue, ya sea individualmente, ya sea tomados en combinación.

## 60 **2. Decodificador de audio según la figura 2**

La figura 2 muestra un diagrama de bloques esquemático de un decodificador 200 de audio según una realización de la presente invención. El decodificador 200 de audio está configurado para recibir una información 210 de audio codificada y para proporcionar, sobre su base, una información 220 de audio decodificada. La información 210 de

audio codificada, por ejemplo, puede adoptar la forma de una secuencia de tramas de audio codificadas en una representación de dominio de tiempo, codificadas en una representación de dominio de frecuencia, o codificadas tanto en una representación de dominio de tiempo como en una representación de dominio de frecuencia. En otras palabras, todas las tramas de la información 210 de audio codificada pueden estar codificadas en una representación de dominio de frecuencia, o todas las tramas de la información 210 de audio codificada pueden estar codificadas en una representación de dominio de tiempo (por ejemplo, en la forma de una señal de excitación de dominio de tiempo codificada y parámetros de síntesis de señal codificados, por ejemplo, parámetros de codificación predictiva lineal (LPC)). Alternativamente, algunas tramas de la información de audio codificada pueden estar codificadas en una representación de dominio de frecuencia, y algunas otras tramas de la información de audio codificada pueden estar codificadas en una representación de dominio de tiempo, por ejemplo, si el decodificador 200 de audio es un decodificador de audio de conmutación que puede conmutar entre diferentes modos de decodificación. La información 220 de audio decodificada, por ejemplo, puede ser una representación de dominio de tiempo de uno o más canales de audio.

15 El decodificador 200 de audio puede comprender habitualmente una decodificación/procesamiento 220, que, por ejemplo, puede proporcionar una información 232 de audio decodificada para tramas de audio que son apropiadamente recibidas. En otras palabras, la decodificación/procesamiento 230 puede realizar una decodificación de dominio de frecuencia (por ejemplo, una decodificación de tipo AAC [codificación avanzada de audio], o similar) sobre la base de una o más tramas de audio codificadas, codificadas en una representación de dominio de frecuencia. Alternativamente, o además, la decodificación/procesamiento 230 puede estar configurado para realizar una la decodificación en el dominio de tiempo (o la decodificación en el dominio de predicción lineal) sobre la base de una o más tramas de audio codificadas, codificadas en una representación de dominio de tiempo (o, en otras palabras, en una representación de dominio de predicción lineal), por ejemplo, una decodificación de predicción lineal excitada por TCX (TCX = excitación codificada transformada) o una decodificación de ACELP (decodificación de predicción lineal excitada por libro de códigos adaptativo). Opcionalmente, la decodificación/procesamiento 230 puede estar configurada para conmutar entre diferentes modos de decodificación.

El decodificador 200 de audio adicionalmente comprende un ocultamiento 240 de error, que está configurado para proporcionar una información 242 de audio de ocultamiento de error para una o más tramas de audio perdidas. El ocultamiento 240 de error está configurado para proporcionar la información 242 de audio de ocultamiento de error para el ocultamiento de una pérdida de una trama de audio (o aún, una pérdida de múltiples tramas de audio). El ocultamiento 240 de error está configurado para modificar una señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida, a fin de obtener la información 242 de audio de ocultamiento de error. En otras palabras, el ocultamiento 240 de error puede obtener (o derivar) una señal de excitación de dominio de tiempo para (o sobre la base de) una o más tramas de audio codificadas que preceden una trama de audio perdida, y puede modificar dicha señal de excitación de dominio de tiempo, que es obtenida para (o sobre la base de) una o más tramas de audio apropiadamente recibidas que preceden una trama de audio perdida, de modo de obtener (por medio de la modificación) una señal de excitación de dominio de tiempo que se usa para proporcionar la información 242 de audio de ocultamiento de error. En otras palabras, la señal de excitación de dominio de tiempo modificada puede usarse como una entrada (o como un componente de una entrada) para una síntesis (por ejemplo, la síntesis de codificación predictiva lineal (LPC)) de la información de audio de ocultamiento de error asociada con la trama de audio perdida (o incluso, con múltiples tramas de audio perdidas). Al proporcionar la información 242 de audio de ocultamiento de error sobre la base de la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio apropiadamente recibidas que preceden la trama de audio perdida, pueden evitarse discontinuidades audibles. Por otra parte, al modificar la señal de excitación de dominio de tiempo derivada para (o a partir de) una o más tramas de audio que preceden la trama de audio perdida, y al proporcionar la información de audio de ocultamiento de error sobre la base de la señal de excitación de dominio de tiempo modificada, es posible considerar la variación de las características del contenido de audio (por ejemplo, un cambio de altura), y además es posible evitar una impresión auditiva no natural (por ejemplo, mediante el “desvanecimiento saliente” de un componente de señal determinista (por ejemplo, por lo menos aproximadamente periódico)). Por lo tanto, puede lograrse que la información 242 de audio de ocultamiento de error comprenda cierta similitud con la información 232 de audio decodificada obtenida sobre la base de tramas de audio apropiadamente decodificadas que preceden la trama de audio perdida, y puede lograrse aunque la información 242 de audio de ocultamiento de error comprenda un contenido de audio algo diferente cuando se compara con la información 232 de audio decodificada asociada con la trama de audio que precede la trama de audio perdida mediante cierta modificación de la señal de excitación de dominio de tiempo. La modificación de la señal de excitación de dominio de tiempo usada para la provisión de la información de audio de ocultamiento de error (asociada con la trama de audio perdida), por ejemplo, puede comprender una escala de amplitud o una escala de tiempo. Sin embargo, son posibles otros tipos de modificaciones (o incluso una combinación de una escala de amplitud y una escala de tiempo), donde, preferentemente, debe permanecer un cierto grado de relación entre la señal de excitación de dominio de tiempo obtenida (como una información de entrada) por el ocultamiento de error y la señal de excitación de dominio de tiempo modificada.

Como conclusión, el decodificador 200 de audio permite la provisión de la información 242 de audio de ocultamiento

de error, de modo tal que la información de audio de ocultamiento de error proporciona una buena impresión auditiva, aún en el caso de que una o más tramas de audio se pierdan. El ocultamiento de error se realiza sobre la base de una señal de excitación de dominio de tiempo, donde se considera una variación de las características de señal del contenido de audio durante la trama de audio perdida mediante la modificación de la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida.

Aún más, debe observarse que el decodificador 200 de audio puede ser suplementado por cualquiera de los rasgos y las funcionalidades que se describen en esta solicitud, ya sea individualmente, ya sea en combinación.

### 10 **3. Decodificador de audio de acuerdo con la Fig. 3.**

La figura 3 muestra un diagrama de bloques esquemático de un decodificador 300 de audio, según otra realización de la presente invención.

15 El decodificador 300 de audio está configurado para recibir una información 310 de audio codificada y para proporcionar, sobre su base, una información 312 de audio decodificada. El decodificador 300 de audio comprende un analizador 320 de corriente de bits, que puede además ser designado como un “desformateador de corriente de bits” o “analizador sintáctico de corriente de bits”. El analizador 320 de corriente de bits recibe la información 310 de audio codificada y proporciona, sobre su base, una representación 322 de dominio de frecuencia y posiblemente, información 324 de control adicional. La representación 322 de dominio de frecuencia, por ejemplo, puede comprender valores 326 espectrales codificados, factores 328 de escala codificados y, opcionalmente, una información 330 lateral adicional que, por ejemplo, puede controlar etapas de procesamiento específicas, por ejemplo, un relleno de ruido, un procesamiento intermedio o un procesamiento posterior. El decodificador 300 de audio además comprende una decodificación 340 de valor espectral que está configurada para recibir los valores 25 326 espectrales codificados, y para proporcionar, sobre su base, un conjunto 342 de valores espectrales decodificados. El decodificador 300 de audio además puede comprender una decodificación 350 de factor de escala, que puede estar configurada para recibir los factores 328 de escala codificados y para proporcionar, sobre su base, un conjunto 352 de factores de escala decodificados.

30 Alternativamente a la decodificación de factor de escala, puede usarse una conversión 354 de codificación predictiva lineal (LPC) a factor de escala, por ejemplo, en el caso de que la información de audio codificada comprenda una información de codificación predictiva lineal (LPC) codificada, en lugar de una información de factor de escala. Sin embargo, en algunos modos de codificación (por ejemplo, en el modo de codificación TCX del decodificador USAC de audio o en el decodificador de audio EVS), puede utilizarse un conjunto de coeficientes de codificación predictiva lineal (LPC) para derivar un conjunto de factores de escala del lado del decodificador de audio. Esta funcionalidad puede lograrse por medio de la conversión 354 de codificación predictiva lineal (LPC) a factor de escala.

El decodificador 300 de audio puede además comprender un escalador 360, que puede estar configurado para aplicar el conjunto 352 de factores escalados al conjunto 342 de valores espectrales, de modo de obtener un conjunto de valores 362 espectrales decodificados escalados. Por ejemplo, una primera banda de frecuencia que comprende múltiples valores 342 espectrales decodificados puede escalarse usando un primer factor de escala, y una segunda banda de frecuencia que comprende múltiples valores 342 espectrales decodificados pueden escalarse usando un segundo factor de escala. Por consiguiente, se obtiene el conjunto de valores 362 espectrales decodificados escalados. El decodificador 300 de audio puede comprender adicionalmente un procesamiento 366 45 opcional, que puede aplicar cierto procesamiento a los valores 362 espectrales decodificados escalados. Por ejemplo, el procesamiento 366 opcional puede comprender un relleno de ruido o alguna otra operación.

El decodificador 300 de audio además comprende una transformada 370 de dominio de frecuencia a dominio de tiempo, que está configurado para recibir los valores 362 espectrales decodificados escalados, o una versión procesada 368 de estos, y para proporcionar una representación 372 de dominio de tiempo asociada con un conjunto de valores 362 espectrales decodificados escalados. Por ejemplo, la transformada 370 de dominio de frecuencia a dominio de tiempo puede proporcionar una representación 372 de dominio de tiempo, que está asociada con una trama o subtrama del contenido de audio. Por ejemplo, la transformada de dominio de frecuencia a dominio de tiempo puede recibir un conjunto de coeficientes de Transformada de coseno discreta modificada (MDCT) (que pueden considerarse valores espectrales decodificados escalados) y proporcionar, sobre su base, un bloque de muestras de dominio de tiempo, que pueden formar la representación 372 de dominio de tiempo.

El decodificador 300 de audio puede comprender opcionalmente un posprocesamiento 376, que puede recibir la representación 372 de dominio de tiempo y modificar en cierta forma la representación 372 de dominio de tiempo, de modo de obtener una versión 378 posprocesada de la representación 372 de dominio de tiempo.

El decodificador 300 de audio además comprende un ocultamiento 380 de error que, por ejemplo, puede recibir la representación 372 de dominio de tiempo de la transformada 370 de dominio de frecuencia a dominio de tiempo y que, por ejemplo, puede proporcionar una información 382 de audio de ocultamiento de error para una o más tramas

de audio perdidas. En otras palabras, si una trama de audio se pierde, de modo tal que, por ejemplo, no hay disponibles valores 326 espectrales codificados para dicha trama de audio (o subtrama de audio), el ocultamiento 380 de error puede proporcionar la información de audio de ocultamiento de error sobre la base de la representación 372 de dominio de tiempo asociada con una o más tramas de audio que preceden la trama de audio perdida. La información de audio de ocultamiento de error puede ser típicamente una representación de dominio de tiempo de un contenido de audio.

Debe observarse que el ocultamiento 380 de error, por ejemplo, puede realizar la funcionalidad del ocultamiento 130 de error que se describe con anterioridad. Además, el ocultamiento 380 de error, por ejemplo, puede comprender la funcionalidad del ocultamiento 500 de error descrita con referencia a la figura 5. Sin embargo, en términos generales, el ocultamiento 380 de error puede comprender cualquiera de los rasgos y de las funcionalidades que se describen con respecto al ocultamiento de error en el presente documento.

Con respecto al ocultamiento de error, debe observarse que el ocultamiento de error no sucede al mismo tiempo que la decodificación de trama. Por ejemplo, si la trama  $n$  es buena, entonces, hacemos una decodificación normal, y al final, guardamos alguna variable que ayudará si tenemos que ocultar la siguiente trama, entonces, si  $n+1$  se pierde, llamamos a la función de ocultamiento proporcionando la variable que proviene de la trama buena previa. Además, actualizaremos algunas variables a fin de ayudar para la siguiente pérdida de trama o con la recuperación para la siguiente trama buena.

El decodificador 300 de audio además comprende una combinación 390 de señales, que está configurada para recibir la representación 372 de dominio de tiempo (o la representación de dominio de tiempo posprocesada 378 en el caso de que haya un posprocesamiento 376). Aún más, la combinación 390 de señales puede recibir la información 382 de audio de ocultamiento de error, que habitualmente es además una representación de dominio de tiempo de una señal de audio de ocultamiento de error provista para una trama de audio perdida. La combinación 390 de señales, por ejemplo, puede combinar representaciones de dominio de tiempo asociadas con subsiguientes tramas de audio. En el caso de que haya subsiguientes tramas de audio apropiadamente decodificadas, la combinación 390 de señales puede combinar (por ejemplo, superposición y adición) representaciones de dominio de tiempo asociadas con las subsiguientes tramas de audio apropiadamente decodificadas. Sin embargo, si una trama de audio se pierde, la combinación 390 de señales puede combinar (por ejemplo, superposición y adición) la representación de dominio de tiempo asociada con la trama de audio apropiadamente decodificada que precede la trama de audio perdida, y la información de audio de ocultamiento de error asociada con la trama de audio perdida, de modo de obtener una llana transición entre la trama de audio apropiadamente recibida y la trama de audio perdida. De manera similar, la combinación 390 de señales puede estar configurada para combinar (por ejemplo, superposición y adición) la información de audio de ocultamiento de error asociada con la trama de audio perdida y la representación de dominio de tiempo asociada con otra trama de audio apropiadamente decodificada tras la trama de audio perdida (u otra información de audio de ocultamiento de error asociada con otra trama de audio perdida, en el caso de que múltiples tramas de audio consecutivas estén perdidas).

Por consiguiente, la combinación 390 de señales puede proporcionar una información 312 de audio decodificada, de modo tal de proporcionar la representación 372 de dominio de tiempo, o una versión 378 posprocesada de esta, para tramas de audio apropiadamente decodificadas, y de modo tal que la información 382 de audio de ocultamiento de error sea provista para tramas de audio perdidas, donde habitualmente se realiza una operación de superposición y adición entre la información de audio (sin consideración de si se proporciona por una transformada 370 de dominio de frecuencia a dominio de tiempo o por el ocultamiento 380 de error) de subsiguientes tramas de audio. Debido a que algunos códecs tienen cierto aliasing sobre la parte de superposición y adición que debe ser cancelada, opcionalmente, podemos crear cierto aliasing artificial sobre la mitad de la trama que hemos creado para realizar la adición de superposición.

Debe observarse que la funcionalidad del decodificador 300 de audio es similar a la funcionalidad del decodificador 100 de audio según la figura 1, donde los detalles adicionales se muestran en la figura 3. Aún más, debe observarse que el decodificador 300 de audio según la figura 3 puede ser suplementado por cualquiera de los rasgos y de las funcionalidades que se describen en el presente documento. En particular, el ocultamiento 380 de error puede ser suplementado por cualquiera de los rasgos y de las funcionalidades que se describen en el presente documento con respecto al ocultamiento de error.

#### **4. Decodificador 400 de audio según la figura 4.**

La figura 4 muestra un decodificador 400 de audio según otra realización de la presente invención. El decodificador 400 de audio está configurado para recibir una información de audio codificada y para proporcionar, sobre su base, una información 412 de audio decodificada. El decodificador 400 de audio, por ejemplo, puede estar configurado para recibir una información 410 de audio codificada, donde diferentes tramas de audio son codificadas usando diferentes modos de codificación. Por ejemplo, el decodificador 400 de audio puede ser considerado un decodificador de audio de múltiples modos o un decodificador de audio "de conmutación". Por ejemplo, algunas de

las tramas de audio pueden ser codificadas usando una representación de dominio de frecuencia, donde la información de audio codificada comprende una representación codificada de valores espectrales (por ejemplo, valores FFT (Transformada de Fourier Rápida) o valores MDCT (Transformada de coseno discreta modificada)) y factores de escala que representan una escala de diferentes bandas de frecuencia. Aún más, la información 410 de audio codificada puede además comprender una “representación de dominio de tiempo” de tramas de audio, o una “representación de dominio de predicción lineal” de múltiples tramas de audio. La “representación de dominio de codificación de predicción lineal” (también designada brevemente como “representación de codificación predictiva lineal (LPC)”), por ejemplo, puede comprender una representación codificada de una señal de excitación, y una representación codificada de parámetros de codificación predictiva lineal (LPC) (parámetros de codificación de predicción lineal), donde los parámetros de codificación de predicción lineal describen, por ejemplo, un filtro de síntesis de codificación de predicción lineal, que se utiliza para reconstruir una señal de audio sobre la base de la señal de excitación de dominio de tiempo.

En lo que sigue, se describirán algunos detalles del decodificador 400 de audio.

El decodificador 400 de audio comprende un analizador 420 de corriente de bits que, por ejemplo, puede analizar la información 410 de audio codificada y extraer, de la información 410 de audio codificada, una representación 422 de dominio de frecuencia, que comprende, por ejemplo, valores espectrales codificados, factores de escala codificados y, opcionalmente, una información lateral adicional. El analizador 420 de corriente de bits además puede estar configurado para extraer una representación 424 de dominio de codificación de predicción lineal, que, por ejemplo, puede comprender una excitación 426 codificada y coeficientes 428 de predicción lineal codificada (que también pueden considerarse parámetros de predicción lineal codificada). Aún más, el analizador de corriente de bits puede opcionalmente extraer información lateral, que puede ser usada para el control de etapas de procesamiento adicionales, a partir de la información de audio codificada.

El decodificador 400 de audio comprende una vía 430 de codificación de dominio de frecuencia, que, por ejemplo, puede ser sustancialmente idéntica a la vía de codificación del decodificador 300 de audio según la figura 3. En otras palabras, la vía 430 de codificación de dominio de frecuencia puede comprender una decodificación 340 de valor espectral, una decodificación 350 de factor de escala, un escalador 360, un procesamiento 366 opcional, una transformada 370 de dominio de frecuencia a dominio de tiempo, un posprocesamiento opcional 376 y un ocultamiento 380 de error, como se describe con anterioridad con referencia a la figura 3.

El decodificador 400 de audio puede además comprender una vía 440 de decodificación de dominio de predicción lineal (que puede considerarse además una vía de decodificación de dominio de tiempo, ya que la síntesis de codificación predictiva lineal (LPC) se realiza en el dominio de tiempo). La vía de decodificación de dominio de predicción lineal comprende una decodificación 450 de excitación, que recibe la excitación 426 codificada provista por el analizador 420 de corriente de bits y proporciona, sobre su base, una excitación 452 decodificada (que puede adoptar la forma de una señal de excitación de dominio de tiempo decodificada). Por ejemplo, la decodificación 450 de excitación puede recibir una información de excitación codificada transformada codificada, y puede proporcionar, sobre su base, una señal de excitación de dominio de tiempo decodificada. Por lo tanto, la decodificación 450 de excitación, por ejemplo, puede realizar una funcionalidad que es efectuada por el decodificador 730 de excitación descrito con referencia a la figura 7. Sin embargo, en forma alternativa o adicional, la decodificación 450 de excitación puede recibir una excitación de predicción lineal excitada por libro de códigos adaptativo (ACELP) codificada, y puede proporcionar la señal 452 de excitación de dominio de tiempo decodificada sobre la base de dicha información de excitación ACELP codificada.

Debe observarse que hay diferentes opciones para la decodificación de excitación. Se hace referencia, por ejemplo, a los Estándares pertinentes y a las publicaciones que definen los conceptos de codificación de Predicción lineal excitada por libro de códigos (CELP), los conceptos de codificación de Predicción lineal excitada por libro de códigos adaptativo (ACELP), las modificaciones de los conceptos de codificación de Predicción lineal excitada por libro de códigos (CELP) y de los conceptos de codificación de Predicción lineal excitada por libro de códigos adaptativo (ACELP) y el concepto de codificación de excitación codificada transformada (TCX).

La vía 440 de decodificación de dominio de predicción lineal opcionalmente comprende un procesamiento 454 en el cual una señal 456 de excitación de dominio de tiempo procesada es derivada de la señal 452 de excitación de dominio de tiempo.

La vía 440 de decodificación de dominio de predicción lineal además comprende una decodificación de coeficiente 460 de predicción lineal, que está configurada para recibir coeficientes de predicción lineal codificados y para proporcionar, sobre su base, coeficientes 462 de predicción lineal decodificados. La decodificación de coeficiente 460 de predicción lineal puede usar diferentes representaciones de un coeficiente de predicción lineal como información 428 de entrada, y puede proporcionar diferentes representaciones de los coeficientes de predicción lineal decodificados como la información 462 de salida. Para detalles, se hace referencia a diferentes documentos de Estándares en los cuales se describe una codificación y/o decodificación de coeficientes de predicción lineal.

La vía 440 de decodificación de dominio de predicción lineal opcionalmente comprende un procesamiento 464, que puede procesar los coeficientes de predicción lineal decodificados y proporcionar una versión 466 procesada de los mismos.

5 La vía 440 de decodificación de dominio de predicción lineal además comprende una síntesis 470 de codificación predictiva lineal (LPC), que está configurada para recibir la excitación 452 decodificada, o su versión 456 procesada, y los coeficientes 462 de predicción lineal decodificados, o su versión 466 procesada, y para proporcionar una señal 472 de audio de dominio de tiempo decodificada. Por ejemplo, la síntesis 470 de codificación predictiva lineal (LPC)  
10 puede estar configurada para aplicar una filtración, que es definida por los coeficientes 462 de predicción lineal decodificados (o su versión 466 procesada), a la señal 452 de excitación de dominio de tiempo decodificada, o su versión procesada, de modo tal que la señal 472 de audio de dominio de tiempo decodificada es obtenida por la filtración (filtración de síntesis) de la señal 452 (o 456) de excitación de dominio de tiempo. La vía de codificación 440 de dominio de predicción lineal puede comprender en forma opcional un posprocesamiento 474, que puede ser  
15 utilizado para refinar o ajustar las características de la señal 472 de audio de dominio de tiempo decodificada.

La vía 440 de decodificación de dominio de predicción lineal además comprende un ocultamiento 480 de error, que está configurado para recibir los coeficientes 462 de predicción lineal decodificados (o su versión 466 procesada) y la señal 452 de excitación de dominio de tiempo decodificada (o su versión 456 procesada). El ocultamiento 480 de error puede recibir opcionalmente información adicional, por ejemplo, una información de altura. El ocultamiento 480 de error puede, en consecuencia, proporcionar una información de audio de ocultamiento de error, que puede presentarse en la forma de una señal de audio de dominio de tiempo, en el caso de que una trama (o subtrama) de la información 410 de audio codificada se pierda. Por lo tanto, el ocultamiento 480 de error puede proporcionar la información 482 de audio de ocultamiento de error de modo tal que las características de la información 482 de audio de ocultamiento de error se adaptan sustancialmente a las características de una última trama de audio apropiadamente decodificada que precede la trama de audio perdida. Debe entenderse que el ocultamiento 480 de error puede comprender cualquiera de los rasgos y de las funcionalidades que se describen con respecto al ocultamiento 240 de error. Asimismo, debe observarse que el ocultamiento 480 de error puede además comprender cualquiera de los rasgos y de las funcionalidades que se describen con respecto al ocultamiento de dominio de tiempo de la figura 6.  
20  
25  
30

El decodificador 400 de audio además comprende un combinador de señal (o combinación 490 de señales), que está configurado para recibir la señal 372 de audio de dominio de tiempo decodificada (o su versión 378 posprocesada), la información 382 de audio de ocultamiento de error provista por el ocultamiento 380 de error, la señal 472 de audio de dominio de tiempo decodificada (o su versión 476 posprocesada) y la información 482 de audio de ocultamiento de error provista por el ocultamiento 480 de error. El combinador de señal 490 puede estar configurado para combinar dichas señales 372 (o 378), 382, 472 (o 476) y 482 para obtener de esta manera la información de audio decodificada 412. En particular, puede aplicarse una operación de superposición y adición por medio del combinador 490 de señal. Por consiguiente, el combinador 490 de señal puede proporcionar transiciones llanas entre subsiguientes tramas de audio para las cuales se proporciona la señal de audio de dominio de tiempo por medio de diferentes entidades (por ejemplo, por diferentes vías 430, 440 de codificación). Sin embargo, el combinador 490 de señal puede además proporcionar transiciones llanas si la señal de audio de dominio de tiempo es provista por la misma entidad (por ejemplo, transformada 370 de dominio de frecuencia a dominio de tiempo, o síntesis 470 de codificación predictiva lineal (LPC)) para tramas subsiguientes. Debido a que algunos códecs tienen cierto aliasing sobre la parte de superposición y adición que debe ser cancelada, opcionalmente, podemos crear cierto aliasing artificial sobre la mitad de la trama que hemos creado para realizar la adición de superposición. En otras palabras, puede usarse opcionalmente una compensación de solapamiento de dominio de tiempo artificial (TDAC).  
35  
40  
45

50 Además, el combinador 490 de señal puede proporcionar transiciones llanas hacia y desde tramas para las cuales se proporciona una información de audio de ocultamiento de error (que, habitualmente, además es una señal de audio de dominio de tiempo).

En síntesis, el decodificador 400 de audio permite la decodificación de tramas de audio que son codificadas en el dominio de frecuencia, y tramas de audio que son codificadas en el dominio de predicción lineal. En particular, es posible la conmutación entre el uso de la vía de codificación de dominio de frecuencia y el uso de la vía de codificación de dominio de predicción lineal según las características de la señal (por ejemplo, usando una información de señalización provista por un codificador de audio). Diferentes tipos de ocultamiento de error pueden usarse para la provisión de una información de audio de ocultamiento de error, en el caso de una pérdida de trama, según si una última trama de audio apropiadamente decodificada fue codificada en el dominio de frecuencia (o, equivalentemente, en una representación de dominio de frecuencia), o en el dominio de tiempo (o equivalentemente, en una representación de dominio de tiempo, o, equivalentemente, en un dominio de predicción lineal, o, equivalentemente, en una representación de dominio de predicción lineal).  
55  
60

**5. Ocultamiento de dominio de tiempo según la figura 5.**

La figura 5 muestra un diagrama de bloques esquemático de un ocultamiento de error según una realización de la presente invención. El ocultamiento de error según la figura 5 es designado en su totalidad como 500.

5 El ocultamiento 500 de error está configurado para recibir una señal 510 de audio de dominio de tiempo y para proporcionar, sobre su base, una información 512 de audio de ocultamiento de error, que, por ejemplo, puede adoptar la forma de una señal de audio de dominio de tiempo.

10 Debe observarse que el ocultamiento 500 de error puede, por ejemplo, ocupar el lugar del ocultamiento 130 de error, de modo tal que la información 512 de audio de ocultamiento de error puede corresponder a la información 132 de audio de ocultamiento de error. Aún más, debe observarse que el ocultamiento 500 de error puede ocupar el lugar del ocultamiento 380 de error, de modo tal que la señal 510 de audio de dominio de tiempo puede corresponder a la señal 372 de audio de dominio de tiempo (o a la señal 378 de audio de dominio de tiempo), y de modo tal que la información 512 de audio de ocultamiento de error puede corresponder a la información 382 de audio de ocultamiento de error.

15 El ocultamiento 500 de error comprende un preénfasis 520, que puede considerarse opcional. El preénfasis recibe la señal de audio de dominio de tiempo y proporciona, sobre su base, una señal 522 de audio de dominio de tiempo preenfatisada.

20 El ocultamiento 500 de error además comprende un análisis 530 de codificación predictiva lineal (LPC), que está configurada para recibir la señal 510 de audio de dominio de tiempo, o su versión 522 preenfatisada, y a fin de obtener una información de codificación 532 predictiva lineal (LPC), que puede comprender un conjunto de parámetros de codificación 532 predictiva lineal (LPC). Por ejemplo, la información de codificación predictiva lineal (LPC) puede comprender un conjunto de coeficientes de filtro de codificación predictiva lineal (LPC) (o una representación de estos) y una señal de excitación de dominio de tiempo (que se adapta para una excitación de un filtro de síntesis de codificación predictiva lineal (LPC) configurado según los coeficientes de filtro de codificación predictiva lineal (LPC), a fin de reconstruir, al menos en forma aproximada, la señal de entrada del análisis de codificación predictiva lineal (LPC).

25 El ocultamiento 500 de error además comprende una búsqueda 540 de altura, que está configurada para obtener una información 542 de altura, por ejemplo, sobre la base de una trama de audio previamente decodificada.

30 El ocultamiento 500 de error además comprende una extrapolación 550, que puede estar configurada para obtener una señal de excitación de dominio de tiempo extrapolada sobre la base del resultado del análisis de codificación predictiva lineal (LPC) (por ejemplo, sobre la base de la señal de excitación de dominio de tiempo determinada por el análisis de codificación predictiva lineal (LPC), y posiblemente, sobre la base del resultado de la búsqueda de altura.

35 El ocultamiento 500 de error además comprende una generación 560 de ruido, que proporciona una señal 562 de ruido. El ocultamiento 500 de error además comprende un combinador/desvanecedor 570, que está configurado para recibir la señal 552 de excitación de dominio de tiempo extrapolada y la señal 562 de ruido, y para proporcionar, sobre su base, una señal 572 de excitación de dominio de tiempo combinada. El combinador/desvanecedor 570 puede estar configurado para combinar la señal 552 de excitación de dominio de tiempo extrapolada y la señal 562 de ruido, donde puede realizarse un desvanecimiento, de modo tal que una contribución relativa de la señal 552 de excitación de dominio de tiempo extrapolada (que determina un componente determinista de la señal de entrada de la síntesis de codificación predictiva lineal (LPC)) disminuye en función del tiempo, mientras que una contribución relativa de la señal 562 de ruido incrementa en función del tiempo. Sin embargo, una funcionalidad diferente del combinador/desvanecedor es también posible. Además, se hace referencia a la descripción a continuación.

40 El ocultamiento 500 de error además comprende una síntesis 580 de codificación predictiva lineal (LPC), que recibe la señal 572 de excitación de dominio de tiempo combinada y que proporciona una señal 582 de audio de dominio de tiempo sobre su base. Por ejemplo, la síntesis de codificación predictiva lineal (LPC) además puede recibir coeficientes de filtro de codificación predictiva lineal (LPC) que describen un filtro de forma de codificación predictiva lineal (LPC), que se aplica a la señal 572 de excitación de dominio de tiempo combinada, a fin de derivar la señal 582 de audio de dominio de tiempo. La síntesis 580 de codificación predictiva lineal (LPC) puede, por ejemplo, usar coeficientes de codificación predictiva lineal (LPC) obtenidos sobre la base de una o más tramas de audio previamente decodificadas (por ejemplo, provistas por el análisis 530 de codificación predictiva lineal (LPC)). El ocultamiento 500 de error además comprende en desénfasis 584, que puede ser considerado opcional. El desénfasis 584 puede proporcionar una señal 586 de audio de dominio de tiempo de ocultamiento de error desenfatizada.

60 El ocultamiento 500 de error además comprende, opcionalmente, una superposición 590 y adición, que realiza una

operación de superposición y adición de las señales de audio de dominio de tiempo asociadas con subsiguientes tramas (o subtramas). Sin embargo, debe observarse que la superposición 590 y adición debe considerarse opcional, ya que el ocultamiento de error puede además usar una combinación de señales que ya es provista en el entorno del decodificador de audio. Por ejemplo, la superposición 590 y adición puede ser reemplazada por la combinación 390 de señales en el decodificador 300 de audio en algunas realizaciones.

En lo que sigue, se describirán algunos detalles adicionales respecto del ocultamiento 500 de error.

El ocultamiento 500 de error según la figura 5 cubre el contexto de un códec de dominio de transformada como AAC<sub>LC</sub> o AAC<sub>ELD</sub>. En otras palabras, el ocultamiento 500 de error se adapta bien para el uso en dicho códec de dominio de transformada (y, en particular, en dicho decodificador de audio de dominio de transformada). En el caso de un códec de transformada solamente (por ejemplo, en ausencia de una vía de decodificación de dominio de predicción lineal), se usa una señal de salida de una última trama, como un punto de inicio. Por ejemplo, una señal 372 de audio de dominio de tiempo puede usarse como un punto de inicio para el ocultamiento de error. Preferentemente, no hay disponible ninguna señal de excitación, solo una señal de dominio de tiempo de salida de (una o más) tramas previas (por ejemplo, la señal 372 de audio de dominio de tiempo).

En lo que sigue, se describirá en más detalle las subunidades y funcionalidades del ocultamiento 500 de error.

### **5.1. Análisis de codificación predictiva lineal (LPC)**

En la realización según la figura 5, todo el ocultamiento se realiza en el dominio de excitación de modo de obtener una transición más llana entre tramas consecutivas. Por lo tanto, es necesario primer hallar (o, más generalmente, obtener) un conjunto apropiado de parámetros de codificación predictiva lineal (LPC). En la realización según la figura 5, se realiza un análisis 530 de codificación predictiva lineal (LPC) sobre la pasada señal 522 de dominio de tiempo preenfatzada. Los parámetros de codificación predictiva lineal (LPC) (o coeficientes de filtro de codificación predictiva lineal (LPC)) son utilizados para realizar el análisis de codificación predictiva lineal (LPC) de la señal de síntesis pasada (por ejemplo, sobre la base de la señal 510 de audio de dominio de tiempo, o sobre la base de la señal 522 de audio de dominio de tiempo preenfatzada) a fin de obtener una señal de excitación (por ejemplo, una señal de excitación de dominio de tiempo).

### **5.2. Búsqueda de altura**

Existen diferentes enfoques para obtener la altura que va a usarse para la construcción de la nueva señal (por ejemplo, la información de audio de ocultamiento de error).

En el contexto del códec usando un filtro de LTP (filtro de predicción de largo plazo, conforme a sus siglas en inglés), como un filtro de predicción de largo plazo de la codificación avanzada de audio [AAC-LTP], si la última trama fue codificación de audio avanzada (AAC) con predicción de largo plazo (LTP), usamos esta última demora de altura de predicción de largo plazo (LTP) recibida y la correspondiente ganancia para la generación de la parte armónica. En este caso, la ganancia es utilizada para decidir si construir la parte armónica en la señal o no. Por ejemplo, si la ganancia de predicción de largo plazo (LTP) es mayor que 0,6 (o cualquier otro valor predeterminado), entonces, la información de predicción de largo plazo (LTP) es utilizada para construir la parte armónica.

Si no hay ninguna información de altura disponible de la trama previa, entonces, hay, por ejemplo, dos soluciones, que se describirán en lo que sigue.

Por ejemplo, es posible realizar una búsqueda de altura en el codificador y transmitir en la corriente de bits la demora de altura y la ganancia. Esto es similar a la predicción de largo plazo (LTP), si bien no hay ninguna aplicación de filtración (además, ninguna filtración de predicción de largo plazo (LTP) en el canal limpio).

Alternativamente, es posible realizar una búsqueda de altura en el decodificador. La búsqueda de altura de Banda ancha de múltiples velocidades adaptativa (AMR-WB) en el caso de la excitación codificada transformada (TCX) se realiza en el dominio de la transformada de Fourier Rápida (FFT). En el retardo extra bajo (ELD, conforme a sus siglas en inglés), por ejemplo, si se usó el dominio de la Transformada de coseno discreta modificada (MDCT), entonces, se perderán las fases. Por lo tanto, la búsqueda de altura se realiza preferentemente directamente en el dominio de excitación. Esto proporciona mejores resultados que la realización de la búsqueda de altura en el dominio de síntesis. La búsqueda de altura en el dominio de excitación se realiza primero con un circuito abierto por medio de una correlación cruzada normalizada. A continuación, en forma opcional, refinamos la búsqueda de altura realizando una búsqueda de circuito cerrado alrededor de la altura de circuito abierto, con una cierta delta. Debido a las limitaciones del ventaneado del retardo extra bajo (ELD), podría hallarse una altura errónea, y en consecuencia, además, verificamos que la altura hallada sea correcta, o de lo contrario, la descartamos.

Como conclusión, la altura de la última trama de audio apropiadamente decodificada que precede la trama de audio



perdida puede considerarse cuando se proporciona la información de audio de ocultamiento de error. En algunos casos, hay una información de altura disponible de la decodificación de la trama previa (es decir, la última trama que precede la trama de audio perdida). En este caso, esta altura puede ser reutilizada (posiblemente, con cierta extrapolación y una consideración de un cambio de altura en función del tiempo). Además, podemos reutilizar

5 opcionalmente la altura de más de una trama del pasado, de manera de tratar de extrapolar la altura que necesitamos al final de nuestra trama ocultada.

Además, si hay una información (por ejemplo, designada como ganancia de predicción de largo plazo) disponible, que describe una intensidad (o intensidad relativa) de un componente de señal determinista (por ejemplo, por lo

10 menos aproximadamente periódico), este valor puede ser utilizado para decidir si un componente determinista (o armónico) debe ser incluido en la información de audio de ocultamiento de error. En otras palabras, al comparar dicho valor (por ejemplo, ganancia de LTP) con un valor de umbral predeterminado, puede decidirse si una señal de excitación de dominio de tiempo derivada de una trama de audio previamente decodificada debe considerarse para la provisión de la información de audio de ocultamiento de error o no.

Si no hay información de altura disponible de la trama previa (o, más precisamente, de la decodificación de la trama previa), hay diferentes opciones. La información de altura podría transmitirse desde un codificador de audio hacia un decodificador de audio, lo que simplificaría el decodificador de audio si bien crearía una sobrecarga de tasa de bits. Alternativamente, la información de altura puede ser determinada en el decodificador de audio, por ejemplo, en el

15 dominio de excitación, es decir, sobre la base de una señal de excitación de dominio de tiempo. Por ejemplo, la señal de excitación de dominio de tiempo derivada de una trama de audio previa apropiadamente decodificada puede evaluarse a fin de identificar la información de altura por ser utilizada para la provisión de la información de audio de ocultamiento de error.

### 25 **5.3. Extrapolación de la excitación o creación de la parte armónica**

La excitación (por ejemplo, la señal de excitación de dominio de tiempo) obtenida de la trama previa (o bien solo computada para la trama perdida o ya guardada en la trama perdida previa para múltiple pérdida de trama) es utilizada para la construcción de la parte armónica (además designada como componente determinista o

30 componente aproximadamente periódico) en la excitación (por ejemplo, en la señal de entrada de la síntesis de codificación predictiva lineal (LPC)) mediante el copiado del último ciclo de altura tantas veces como sean necesarias para obtener una trama y media. A fin de ahorrar complejidad, podemos además crear una trama y media solo para la primera trama de pérdida y luego, cambiar el procesamiento para subsiguiente pérdida de trama a la mitad de la trama, y crear solo una trama para cada una. A continuación, tenemos siempre acceso a la mitad de

35 una trama de superposición.

En el caso de la primera trama perdida tras una trama buena (es decir, una trama apropiadamente decodificada), el primer ciclo de altura (por ejemplo, de la señal de excitación de dominio de tiempo obtenida sobre la base de la última trama de audio apropiadamente decodificada que precede la trama de audio perdida) es el filtro paso bajo con

40 un filtro dependiente de la tasa de muestreo (ya que el retardo extra bajo (ELD) cubre una combinación de tasa de muestreo realmente amplia – que va de núcleo AAC–ELD a AAC–ELD con SBR o AAC–ELD tasa dual SBR).

La altura en una señal de voz es casi siempre cambiante. Por lo tanto, el ocultamiento presentado anteriormente tiende a crear algunos problemas (o al menos distorsiones) en la recuperación, ya que la altura al final de la señal

45 ocultada (es decir, al final de la información de audio de ocultamiento de error) con frecuencia no coincide con la altura de la primera trama buena. Por lo tanto, opcionalmente, en algunas realizaciones, se trata de predecir la altura al final de la trama ocultada a fin de coincidir con la altura al comienzo de la trama de recuperación. Por ejemplo, se predice la altura al final de una trama perdida (que se considera una trama ocultada), donde el objetivo de la predicción es establecer la altura al final de la trama perdida (trama ocultada) a fin de aproximarse a la altura al

50 comienzo de la primera trama apropiadamente decodificada tras una o más tramas perdidas (cuya primera trama apropiadamente decodificada además se denomina “trama de recuperación”). Esto podría realizarse durante la pérdida de trama o durante la primera trama buena (es decir, durante la primera trama apropiadamente recibida). A fin de obtener incluso mejores resultados, es posible reutilizar opcionalmente algunas herramientas convencionales y adaptarlas, como la resincronización de predicción de altura y pulsos. Para detalles, se hace referencia, por

55 ejemplo, a las referencias [6] y [7].

Si se usa una predicción de largo plazo (LTP) en un códec de dominio de frecuencia, es posible usar la demora como la información de inicio acerca de la altura. Sin embargo, en algunas realizaciones, se desea además contar con una mejor granularidad a fin de poder rastrear mejor el contorno de altura. Por lo tanto, se prefiere realizar una

60 búsqueda de altura al comienzo y al final de la última buena (apropiadamente decodificada) trama. A fin de adaptar la señal a la altura en movimiento, es deseable la utilización de una resincronización de pulsos, que se presenta en el estado de la técnica.

#### **5.4. Ganancia de altura**

En algunas realizaciones, se prefiere la aplicación de una ganancia sobre la excitación previamente obtenida a fin de lograr el nivel deseado. La “ganancia de la altura”(por ejemplo, la ganancia del componente determinista de la señal de excitación de dominio de tiempo, es decir, la ganancia aplicada a una señal de excitación de dominio de tiempo derivada de una trama de audio previamente decodificada, a fin de obtener la señal de entrada de la síntesis de codificación predictiva lineal (LPC)), puede, por ejemplo, ser obtenida mediante la realización de una correlación normalizada en el dominio de tiempo al final de la última buena (por ejemplo, apropiadamente decodificada) trama. La longitud de la correlación puede ser equivalente a la longitud de dos subtramas, o puede ser cambiada adaptativamente. El retardo es equivalente a la demora de altura que se usa para la creación de la parte armónica. Podemos además opcionalmente realizar el cálculo de ganancia solo sobre la primera trama perdida y luego solo aplicar un desvanecimiento saliente (reducida ganancia) para la siguiente pérdida de trama consecutiva.

La “ganancia de altura” determinará la cantidad de tonalidad (o la cantidad de componentes de señal deterministas, por lo menos aproximadamente periódicos) que se crearán. Sin embargo, es deseable agregar cierto ruido con forma para no tener solo un tono artificial. Si obtenemos muy baja ganancia de la altura, entonces construimos una señal que consiste solo en un ruido con forma.

Como conclusión, en algunos casos, la señal de excitación de dominio de tiempo obtenida, por ejemplo, sobre la base de una trama de audio previamente decodificada, es escalada según la ganancia (por ejemplo, a fin de obtener la señal de entrada para el análisis de codificación predictiva lineal (LPC)). Por consiguiente, debido a que la señal de excitación de dominio de tiempo determina un componente de señal determinista (por lo menos aproximadamente periódico), la ganancia puede determinar una intensidad relativa de dichos componentes de señal deterministas (por lo menos aproximadamente periódicos) en la información de audio de ocultamiento de error. Además, la información de audio de ocultamiento de error puede sustentarse en un ruido, que es además formado por la síntesis de codificación predictiva lineal (LPC), de modo tal que una energía total de la información de audio de ocultamiento de error es adaptada, por lo menos hasta cierto grado, a una trama de audio apropiadamente decodificada que precede la trama de audio perdida e, idealmente, además a una trama de audio apropiadamente decodificada tras las una o más tramas de audio perdidas.

#### **5.5. Creación de la parte de ruido**

Una “innovación” es creada por un generador de ruido aleatorio. En forma opcional, este ruido es adicionalmente filtrado paso alto y opcionalmente preenfático para tramas de voz y de inicio. Como para el paso bajo de la parte armónica, este filtro (por ejemplo, el filtro paso alto) es dependiente de la tasa de muestreo. Este ruido (que es provisto, por ejemplo, por una generación 560 de ruido) será formado por la codificación predictiva lineal (LPC) (por ejemplo, por la síntesis 580 de codificación predictiva lineal (LPC) para llegar lo más cerca posible del ruido de fondo. La característica de paso alto es además opcionalmente cambiada sobre pérdida de trama consecutiva, de modo tal que sobre una cierta cantidad de una pérdida de trama, ya no hay más filtración, para solo obtener el ruido con forma de banda completa a fin de lograr un ruido de confort cercano al ruido de fondo.

Una ganancia de innovación (que, por ejemplo, puede determinar una ganancia del ruido 562 en la combinación/desvanecimiento saliente 570, es decir, una ganancia usando la señal 562 de ruido que se incluye en la señal 572 de entrada de la síntesis de codificación predictiva lineal (LPC)) es, por ejemplo, calculada mediante la eliminación de la contribución previamente computada de la altura (si existe) (por ejemplo, una versión escalada, escalada usando la “ganancia de altura”, de la señal de excitación de dominio de tiempo obtenida sobre la base de la última trama de audio apropiadamente decodificada que precede la trama de audio perdida) y la realización de una correlación al final de la última buena trama. En cuanto a la ganancia de altura, esta podría realizarse opcionalmente solo sobre la primera trama perdida, y luego, el desvanecimiento saliente, si bien, en este caso, el desvanecimiento en forma saliente podría ir o bien a 0, que resulta en un silenciamiento completado, o a un nivel de ruido estimado presente en el fondo. La longitud de la correlación es, por ejemplo, equivalente a la longitud de dos subtramas, y el retardo es equivalente a la demora de altura utilizada para la creación de la parte armónica.

Opcionalmente, esta ganancia además se multiplica por (1-“ganancia de altura”) a fin de aplicar tanta ganancia sobre el ruido de modo de alcanzar la pérdida de ganancia si la ganancia de altura no es uno. Opcionalmente, esta ganancia además es multiplicada por un factor de ruido. Este factor de ruido proviene, por ejemplo, de la trama válida previa (por ejemplo, de la última trama de audio apropiadamente decodificada que precede la trama de audio perdida).

#### **5.6. Desvanecimiento saliente**

El desvanecimiento saliente se usa en su mayoría para múltiple pérdida de trama. Sin embargo, el desvanecimiento saliente puede además ser usado en el caso de que solo se pierda una sola trama de audio.

En el caso de una múltiple pérdida de trama, los parámetros de codificación predictiva lineal (LPC) no son

recalculados. O bien, se mantiene el último computado, o se realiza el ocultamiento de codificación predictiva lineal (LPC) mediante la convergencia a una forma de fondo. En este caso, la periodicidad de la señal se converge a cero. Por ejemplo, la señal 502 de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida aún utiliza una ganancia que es gradualmente reducida en función del tiempo, mientras que la señal 562 de ruido se mantiene constante o escalada con una ganancia que es gradualmente creciente en función del tiempo, de modo tal que el peso relativo de la señal 552 de excitación de dominio de tiempo se reduce en función del tiempo en comparación con el peso relativo de la señal 562 de ruido. En consecuencia, la señal 572 de entrada de la síntesis 580 de codificación predictiva lineal (LPC) se torna cada vez más “de tipo ruido”. Por lo tanto, la “periodicidad” (o, más precisamente, el componente determinista, o componente por lo menos aproximadamente periódico de la señal 582 de salida de la síntesis 580 de codificación predictiva lineal (LPC)) se reduce en función del tiempo.

La velocidad de la convergencia según la cual la periodicidad 572 de la señal, y/o la periodicidad de la señal 582, converge a 0, depende de los parámetros de la última trama correctamente recibida (o apropiadamente decodificada) y/o de la cantidad de tramas borradas consecutivas, y es controlada por un factor de atenuación,  $\alpha$ . El factor,  $\alpha$ , es adicionalmente dependiente de la estabilidad del filtro LP. Opcionalmente, es posible alterar el factor  $\alpha$  en la relación con la longitud de altura. Si la altura (por ejemplo, una longitud periódica asociada con la altura) es realmente larga, entonces mantenemos  $\alpha$  “normal”, pero si la altura es realmente corta, habitualmente es necesario copiar una cantidad de veces la misma parte de la excitación pasada. Esto rápidamente sonará demasiado artificial, y por lo tanto, se prefiere el desvanecimiento saliente más veloz de esta señal.

Además, opcionalmente, si está disponible, podemos considerar la salida de la predicción de altura. Si se predice una altura, esto significa que la altura ya estaba cambiando en la trama previa, y entonces, cuantas más tramas perdemos, más lejos estamos de la verdad. Por lo tanto, se prefiere acelerar algo el desvanecimiento en forma saliente de la parte tonal, en este caso.

Si la predicción de altura fallara debido a que la altura cambia demasiado, esto significa que o bien los valores de altura no son realmente confiables, o que la señal es realmente impredecible. Por lo tanto, nuevamente, se prefiere realizar el desvanecimiento saliente en forma más veloz (por ejemplo, el desvanecimiento saliente de la señal 552 de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio apropiadamente decodificadas que preceden una o más tramas de audio perdidas).

### **5.7. Síntesis de codificación predictiva lineal (LPC)**

A fin de regresar al dominio de tiempo, se prefiere realizar una síntesis 580 de codificación predictiva lineal (LPC) sobre la suma de las dos excitaciones (parte tonal y parte ruidosa), seguida de un desénfasis. En otras palabras, se prefiere realizar la síntesis 580 de codificación predictiva lineal (LPC) sobre la base de una combinación pesada de una señal 552 de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio apropiadamente decodificadas que preceden la trama de audio perdida (parte tonal) y la señal 562 de ruido (parte ruidosa). Como se menciona con anterioridad, la señal 552 de excitación de dominio de tiempo puede ser modificada en comparación con la señal 532 de excitación de dominio de tiempo obtenida por el análisis 530 de codificación predictiva lineal (LPC) (además de los coeficientes de codificación predictiva lineal (LPC) que describen una característica del filtro de la síntesis de codificación predictiva lineal (LPC) utilizado para la síntesis 580 de codificación predictiva lineal (LPC)). Por ejemplo, la señal 552 de excitación de dominio de tiempo puede ser una copia escalada en tiempo de la señal 532 de excitación de dominio de tiempo obtenida por el análisis 530 de codificación predictiva lineal (LPC), donde la escala de tiempo puede ser utilizada para adaptar la altura de la señal 552 de excitación de dominio de tiempo a una altura deseada.

### **5.8. Superposición y adición**

En el caso de un códec de transformada solamente, a fin de obtener la mejor superposición y adición, creamos una señal artificial para media trama más que la trama ocultada, y podemos crear aliasing artificial sobre esta. Sin embargo, pueden aplicarse diferentes conceptos de superposición y adición.

En el contexto de la codificación avanzada de audio (AAC) o la excitación codificada transformada (TCX) regular, se aplica una superposición y adición entre la media trama extra que proviene del ocultamiento y la primera parte de la primera buena trama (podría ser la mitad o menos, para ventanas de retardo menores como AAC-LD).

En el caso especial del retardo extra bajo (ELD) para la primera trama perdida, se prefiere ejecutar el análisis tres veces a fin de obtener la apropiada contribución de las tres ventanas, y luego, para la primera trama de ocultamiento, y todas las siguientes, el análisis es ejecuta una vez más. A continuación, se realiza una síntesis de retardo extra bajo (ELD), para volver al dominio de tiempo con toda la memoria apropiada para la siguiente trama en el dominio de Transformada de coseno discreta modificada (MDCT).

Como conclusión, la señal 572 de entrada de la síntesis 580 de codificación predictiva lineal (LPC) (y/o la señal 552 de excitación de dominio de tiempo) puede proporcionarse para una duración temporal que es mayor que una duración de una trama de audio perdida. Por consiguiente, la señal 582 de salida de la síntesis 580 de codificación predictiva lineal (LPC) puede además ser provista para un período de tiempo que es mayor que una trama de audio perdida. Por consiguiente, puede realizarse una superposición y adición entre la información de audio de ocultamiento de error (que es consecuentemente obtenida para un período de tiempo más largo que una extensión temporal de la trama de audio perdida) y una información de audio decodificada provista para una trama de audio apropiadamente decodificada tras una o más tramas de audio perdidas.

En síntesis, el ocultamiento 500 de error se adapta bien al caso en el cual las tramas de audio son codificadas en el dominio de frecuencia. Aun cuando las tramas de audio son codificadas en el dominio de frecuencia, la provisión de la información de audio de ocultamiento de error se realiza sobre la base de una señal de excitación de dominio de tiempo. Se aplican diferentes modificaciones a la señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio apropiadamente decodificadas que preceden una trama de audio perdida. Por ejemplo, la señal de excitación de dominio de tiempo provista por el análisis 530 de codificación predictiva lineal (LPC) se adapta a cambios de altura, por ejemplo, usando una escala de tiempo. Aún más, la señal de excitación de dominio de tiempo provista por el análisis 530 de codificación predictiva lineal (LPC) es además modificada por una escala (aplicación de una ganancia), donde un desvanecimiento saliente del componente determinista (o tonal, o por lo menos aproximadamente periódico) puede ser realizado por el escalador/desvanecedor 570, de modo tal que la señal 572 de entrada de la síntesis 580 de codificación predictiva lineal (LPC) comprende tanto un componente que deriva de la señal de excitación de dominio de tiempo obtenida por el análisis de codificación predictiva lineal (LPC) como un componente de ruido que se basa en la señal 562 de ruido. El componente determinista de la señal 572 de entrada de la síntesis 580 de codificación predictiva lineal (LPC) sin embargo, es habitualmente modificado (por ejemplo, escala de tiempo y/o escala de amplitud) con respecto a la señal de excitación de dominio de tiempo provista por el análisis 530 de codificación predictiva lineal (LPC).

En consecuencia, la señal de excitación de dominio de tiempo puede ser adaptada a las necesidades, y se evita una impresión auditiva no natural.

### 30 **6. Ocultamiento de dominio de tiempo según la figura 6**

La figura 6 muestra un diagrama de bloques esquemático de un ocultamiento de dominio de tiempo que puede usarse para un códec de conmutación. Por ejemplo, el ocultamiento 600 de dominio de tiempo según la figura 6 puede, por ejemplo, tomar el lugar del ocultamiento 240 de error, o el lugar del ocultamiento 480 de error.

Aún más, debe observarse que la realización según la figura 6 cubre el contexto (que puede usarse dentro del contexto) de un códec de conmutación usando dominios de tiempo y frecuencia combinados, tales como USAC [Codificación unificada de voz y audio] (MPEG-D/MPEG-H) o EVS (3GPP). En otras palabras, el ocultamiento 600 de dominio de tiempo puede usarse en decodificadores de audio en los cuales hay una conmutación entre una decodificación de dominio de frecuencia y una decodificación de tiempo (o, equivalentemente, una decodificación sobre la base de coeficientes de predicción lineal).

Sin embargo, debe observarse que el ocultamiento 600 de error según la figura 6 además puede usarse en decodificadores de audio que meramente realizan una decodificación en el dominio de tiempo (o equivalentemente, en el dominio de coeficiente de predicción lineal).

En el caso de un códec conmutado (e incluso, en el caso de un códec que meramente realiza la decodificación en el dominio de coeficiente de predicción lineal), habitualmente ya tenemos la señal de excitación (por ejemplo, la señal de excitación de dominio de tiempo) que proviene de una trama previa (por ejemplo, una trama de audio apropiadamente decodificada que precede una trama de audio perdida). De lo contrario (por ejemplo, si la señal de excitación de dominio de tiempo no está disponible), es posible actuar según lo explicado en la realización según la figura 5, es decir, realizar un análisis de codificación predictiva lineal (LPC). Si la trama previa era de tipo ACELP, además ya tenemos la información de tono de las subtramas en la última trama. Si la última trama fue TCX (Excitación codificada transformada) con LTP (predicción de largo plazo), además tenemos la información de demora que proviene de la predicción de largo plazo. Y si la última trama estuvo en el dominio de frecuencia sin predicción de largo plazo (LTP), entonces la búsqueda de altura se realiza preferentemente en forma directa en el dominio de excitación (por ejemplo, sobre la base de una señal de excitación de dominio de tiempo provista por un análisis de codificación predictiva lineal (LPC)).

Si el decodificador ya utiliza algunos parámetros de codificación predictiva lineal (LPC) en el dominio de tiempo, los reutilizamos y extrapolamos un nuevo conjunto de parámetros de codificación predictiva lineal (LPC). La extrapolación de los parámetros de codificación predictiva lineal (LPC) se sustenta en la pasada codificación predictiva lineal (LPC), por ejemplo, la media de las últimas tres tramas y (opcionalmente), la forma de la codificación predictiva lineal (LPC) derivada durante la estimación de ruido DTX si DTX (transmisión discontinua) existe en el

códec.

Todo el ocultamiento se realiza en el dominio de excitación a fin de obtener una transición más llana entre tramas consecutivas.

5 En lo que sigue, se describirá en más detalle el ocultamiento 600 de error según la figura 6.

El ocultamiento 600 de error recibe una excitación 610 pasada y una información 640 de altura pasada. Aún más, el ocultamiento 600 de error proporciona una información 612 de audio de ocultamiento de error.

10 Debe observarse que la excitación 610 pasada recibida por el ocultamiento 600 de error puede, por ejemplo, corresponder a la salida 532 del análisis 530 de codificación predictiva lineal (LPC). Aún más, la pasada 640 información de altura puede, por ejemplo, corresponder a la información 542 de salida de la búsqueda 540 de altura.

15 El ocultamiento 600 de error adicionalmente comprende una extrapolación 650, que puede corresponder a la extrapolación 550, de modo tal que se hace referencia a la discusión anterior.

Además, el ocultamiento de error comprende un generador 660 de ruido, que puede corresponder al generador 560 de ruido, de modo tal que se hace referencia a la discusión anterior.

20 La extrapolación 650 proporciona una señal 652 de excitación de dominio de tiempo extrapolada, que puede corresponder a la señal 552 de excitación de dominio de tiempo extrapolada. El generador 660 de ruido proporciona una señal 662 de ruido, que corresponde a la señal 562 de ruido.

25 El ocultamiento 600 de error además comprende un combinador/desvanecedor 670, que recibe la señal 652 de excitación de dominio de tiempo extrapolada y la señal 662 de ruido y proporciona, sobre su base, una señal 672 de entrada para una síntesis 680 de codificación predictiva lineal (LPC), donde la síntesis 680 de codificación predictiva lineal (LPC) puede corresponder a la síntesis 580 de codificación predictiva lineal (LPC), de modo tal que se aplican, asimismo, las explicaciones anteriores. La síntesis 680 de codificación predictiva lineal (LPC) proporciona una señal 682 de audio de dominio de tiempo, que puede corresponder a la señal 582 de audio de dominio de tiempo. El ocultamiento de error además comprende (opcionalmente) un desénfasis 684, que puede corresponder al desénfasis 584 y que proporciona una señal 686 de audio de dominio de tiempo de ocultamiento de error desenfatisada. El ocultamiento 600 de error opcionalmente comprende una superposición 690 y adición, que puede corresponder a la superposición 590 y adición. Sin embargo, se aplican asimismo las explicaciones con respecto a la superposición 590 y adición, a la superposición 690 y adición. En otras palabras, la superposición 690 y adición puede además ser reemplazada por la superposición y adición general del decodificador de audio, de modo tal que la señal 682 de salida de la síntesis de codificación predictiva lineal (LPC) o la señal 686 de salida del desénfasis pueden considerarse la información de audio de ocultamiento de error.

40 Como conclusión, el ocultamiento 600 de error difiere sustancialmente del ocultamiento 500 de error, en términos de que el ocultamiento 600 de error obtiene directamente la información de excitación 610 pasada y la información 640 de altura pasada, de una o más tramas de audio previamente decodificadas, sin la necesidad de realizar un análisis de codificación predictiva lineal (LPC) y/o un análisis de altura. Sin embargo, debe observarse que el ocultamiento 600 de error, opcionalmente, puede comprender un análisis de codificación predictiva lineal (LPC) y/o un análisis de altura (búsqueda de altura).

En lo que sigue, se describirán en más detalle algunos rasgos del ocultamiento 600 de error. Sin embargo, debe observarse que los detalles específicos deben considerarse ejemplares, en lugar de rasgos esenciales.

### 50 **6.1. Altura pasada de búsqueda de altura**

Hay diferentes enfoques para obtener la altura por ser usada en la construcción de la nueva señal.

55 En el contexto del códec que utiliza el filtro de LTP, como AAC-LTP, si la última trama (que precede la trama perdida) fue AAC con LTP, tenemos la información de tono que proviene de la última demora de tono de LTP y la ganancia correspondiente. En este caso, usamos la ganancia a fin de decidir si queremos construir la parte armónica en la señal o no. Por ejemplo, si la ganancia de predicción de largo plazo (LTP) es superior a 0,6, entonces, usamos la información de predicción de largo plazo (LTP) para construir la parte armónica.

60 Si no tenemos ninguna información de altura disponible de la trama previa, entonces hay, por ejemplo, dos soluciones adicionales.

Una solución es realizar una búsqueda de altura en el codificador y transmitir en la corriente de bits la demora de altura y la ganancia. Esto es similar a la predicción de largo plazo (LTP), si bien no aplicamos ninguna filtración

(tampoco, ninguna filtración de predicción de largo plazo en el canal limpio).

Otra solución es realizar una búsqueda de altura en el decodificador. La búsqueda de altura de Banda ancha de múltiples velocidades adaptativa (AMR-WB) en el caso de la excitación codificada transformada (TCX) se realiza en el dominio de la transformada de Fourier Rápida (FFT). En la excitación codificada transformada (TCX), por ejemplo, usamos el dominio de la transformada de coseno discreta modificada (MDCT), entonces, perdemos las fases. Por lo tanto, la búsqueda de altura se realiza directamente en el dominio de excitación (por ejemplo, sobre la base de la señal de excitación de dominio de tiempo utilizada como la entrada de la síntesis de codificación predictiva lineal (LPC), o utilizada para derivar la entrada para la síntesis de codificación predictiva lineal (LPC)), en una realización preferida. Esto habitualmente proporciona mejores resultados que la realización de la búsqueda de altura en el dominio de síntesis (por ejemplo, sobre la base de una señal de audio de dominio de tiempo completamente decodificada).

La búsqueda de altura en el dominio de excitación (por ejemplo, sobre la base de la señal de excitación de dominio de tiempo) se realiza primero con un circuito abierto por medio de una correlación cruzada normalizada. A continuación, opcionalmente, la búsqueda de altura puede ser refinada realizando una búsqueda de circuito cerrado alrededor de la altura de circuito abierto con una cierta delta.

En implementaciones preferidas, no consideramos simplemente un valor máximo de la correlación. Si tenemos una información de altura de una trama previa no propensa al error, entonces seleccionamos la altura que corresponde a aquella de los cinco valores más altos en el dominio de correlación cruzada normalizada, aunque el más cercano a la altura de la trama previa. Entonces, se verifica además que la máxima hallada no sea una máxima errónea debido a la limitación de ventana.

Como conclusión, hay diferentes conceptos para determinar la altura, donde es computacionalmente eficiente considerar una altura pasada (es decir, altura asociada con una trama de audio previamente decodificada). Alternativamente, la información de altura puede ser transmitida desde un codificador de audio hacia un decodificador de audio. Como otra alternativa, una búsqueda de altura puede realizarse del lado del decodificador de audio, donde la determinación de altura se realiza, preferentemente, sobre la base de la señal de excitación de dominio de tiempo (es decir, en el dominio de excitación). Una búsqueda de altura de dos etapas que comprende una búsqueda de circuito abierto y una búsqueda de circuito cerrado puede realizarse a fin de obtener una particularmente confiable y precisa información de altura. Alternativamente, o además, una información de altura de una trama de audio previamente decodificada puede usarse a fin de garantizar que la búsqueda de altura proporciona un resultado confiable.

## **6.2. Extrapolación de la excitación o creación de la parte armónica**

La excitación (por ejemplo, en la forma de una señal de excitación de dominio de tiempo) obtenida de la trama previa (o bien solo computada para la trama perdida o ya guardada en la trama perdida previa para múltiple pérdida de trama) es utilizada para construir la parte armónica en la excitación (por ejemplo, la señal 662 de excitación de dominio de tiempo extrapolada) mediante el copiado del último ciclo de altura (por ejemplo, una porción 610 de la señal de excitación de dominio de tiempo, cuya duración temporal es igual a una duración de período de la altura) tantas veces como sean necesarias para obtener, por ejemplo, una y media de la trama (perdida).

A fin de obtener aún mejores resultados, es opcionalmente posible reutilizar algunas herramientas conocidas del estado de la técnica y adaptarlas. Para detalles, se hace referencia, por ejemplo, a las referencias [6] y [7].

Se ha hallado que la altura en una señal de voz es casi siempre cambiante. Se ha hallado que, por lo tanto, el ocultamiento presentado anteriormente tiende a crear algunos problemas en la recuperación, ya que la altura al final de la señal ocultada a menudo no coincide con la altura de la primera trama buena. Por lo tanto, opcionalmente, se trata de predecir la altura al final de la trama ocultada, a fin de coincidir con la altura al comienzo de la trama de recuperación. Esta funcionalidad será realizada, por ejemplo, por la extrapolación 650.

Si se usa la predicción de largo plazo (LTP) en la Excitación codificada transformada (TCX), puede usarse la demora como la información inicial acerca de la altura. Sin embargo, es deseable contar con una mejor granularidad para poder rastrear mejor el contorno de altura. Por lo tanto, se realiza una búsqueda de altura opcionalmente al comienzo y al final de la última buena trama. A fin de adaptar la señal a la altura en movimiento, puede usarse una resincronización de pulsos, que se presenta en el estado de la técnica.

Como conclusión, la extrapolación (por ejemplo, de la señal de excitación de dominio de tiempo asociada con, u obtenida sobre la base de, una última trama de audio apropiadamente decodificada que precede la trama perdida) puede comprender un copiado de una porción de tiempo de dicha señal de excitación de dominio de tiempo asociada con una trama de audio previa, donde la porción de tiempo copiada puede ser modificada según un cómputo, o una estimación, de un cambio de altura (esperado) durante la trama de audio perdida. Pueden obtenerse

diferentes conceptos para la determinación del cambio de altura.

### **6.3. Ganancia de altura**

- 5 En la realización según la figura 6, se aplica una ganancia sobre la previamente obtenida excitación a fin de alcanzar un nivel deseado. La ganancia de la altura es obtenida, por ejemplo, realizando una correlación normalizada en el dominio de tiempo al final de la última buena trama. Por ejemplo, la longitud de la correlación puede ser equivalente a la longitud de dos subtramas, y el retardo puede ser equivalente a la demora de altura utilizada para la creación de la parte armónica (por ejemplo, para el copiado de la señal de excitación de dominio de tiempo). Se ha hallado que
- 10 al hacer el cálculo de ganancia en el dominio de tiempo, se proporciona una ganancia mucho más confiable, que realizándolo en el dominio de excitación. La codificación predictiva lineal (LPC) cambia en cada trama, y entonces, la aplicación de una ganancia, calculada sobre la trama previa, sobre una señal de excitación que será procesada por otro conjunto de codificación predictiva lineal (LPC), no proporcionará la energía esperada en el dominio de tiempo.
- 15 La ganancia de la altura determina la cantidad de tonalidad que se creará, si bien además se agregará cierto ruido con forma para no tener solo un tono artificial. Si se obtiene una muy baja ganancia de altura, entonces, puede construirse una señal que consiste solo en un ruido con forma.

20 Como conclusión, una ganancia que se aplica para escalar la señal de excitación de dominio de tiempo obtenida sobre la base de la trama previa (o una señal de excitación de dominio de tiempo que es obtenida para una trama previamente decodificada, o que se asocia con la trama previamente decodificada) se ajusta de modo de determinar un valor de un componente tonal (o determinista, o por lo menos aproximadamente periódico) dentro de la señal de entrada de la síntesis 680 de codificación predictiva lineal (LPC), y, en consecuencia, dentro de la información de audio de ocultamiento de error. Dicha ganancia puede ser determinada sobre la base de una correlación, que se

25 aplica a la señal de audio de dominio de tiempo obtenida por una decodificación de la trama previamente decodificada (donde dicha señal de audio de dominio de tiempo puede ser obtenida usando una síntesis de codificación predictiva lineal (LPC) que se realiza en el curso de la decodificación).

### **6.4. Creación de la parte de ruido**

- 30 Se crea una innovación por medio de un generador 660 de ruido aleatorio. Este ruido es adicionalmente filtrado de paso alto y opcionalmente preenfatisado para tramas de voz y de inicio. La filtración de paso alto y el preénfasis, que pueden realizarse selectivamente para tramas de voz y de inicio, no se muestran explícitamente en la figura 6, si bien pueden realizarse, por ejemplo, dentro del generador 660 de ruido o dentro del combinador/desvanecedor 670.
- 35 El ruido será formado (por ejemplo, tras la combinación con la señal 652 de excitación de dominio de tiempo obtenida por la extrapolación 650) por la codificación predictiva lineal (LPC) de modo de obtener lo más cercano posible al ruido de fondo.
- 40 Por ejemplo, la ganancia de innovación puede ser calculada mediante la eliminación de la contribución previamente computada de la altura (si existe) y realizando una correlación al final de la última buena trama. La longitud de la correlación puede ser equivalente a la longitud de dos subtramas, y el retardo puede ser equivalente a la demora de altura utilizada para la creación de la parte armónica.
- 45 Opcionalmente, esta ganancia puede además ser multiplicada por  $(1 - \text{ganancia de altura})$  a fin de aplicar tanta ganancia sobre el ruido para alcanzar la pérdida de energía si la ganancia de la altura no es uno. Opcionalmente, esta ganancia además es multiplicada por un factor de ruido. Este factor de ruido puede provenir de una trama válida previa.
- 50 Como conclusión, un componente de ruido de la información de audio de ocultamiento de error se obtiene mediante la formación de ruido provisto por el generador 660 de ruido usando la síntesis 680 de codificación predictiva lineal (LPC) (y, posiblemente, el desénfasis 684). Además, puede aplicarse una filtración de paso alto adicional y/o un preénfasis. La ganancia de la contribución de ruido a la señal 672 de entrada de la síntesis 680 de codificación predictiva lineal (LPC) (además designada "ganancia de innovación") puede computarse sobre la base de la última
- 55 trama de audio apropiadamente decodificada que precede la trama de audio perdida, donde un componente determinista (o por lo menos aproximadamente periódico) puede ser eliminado de la trama de audio que precede la trama de audio perdida, y donde puede realizarse entonces una correlación a fin de determinar la intensidad (o ganancia) del componente de ruido dentro de la señal de dominio de tiempo decodificada de la trama de audio que precede la trama de audio perdida.
- 60 Opcionalmente, pueden aplicarse ciertas modificaciones adicionales a la ganancia del componente de ruido.

**6.5. Desvanecimiento saliente**

El desvanecimiento en forma saliente se usa en su mayoría para múltiples pérdidas de tramas. Sin embargo, el desvanecimiento en forma saliente puede usarse además en el caso de que solo se pierda una única trama de audio.

- 5 En el caso de múltiple pérdida de trama, los parámetros de codificación predictiva lineal (LPC) no son recalculados. O bien se mantiene el último computado, o se realiza un ocultamiento de codificación predictiva lineal (LPC) como se explica anteriormente.

10 Una periodicidad de la señal se converge a cero. La velocidad de la convergencia depende de los parámetros de la última trama correctamente recibida (o correctamente decodificada) y la cantidad de tramas borradas consecutivas (o perdidas), y se controla por un factor de atenuación,  $\alpha$ . El factor,  $\alpha$ , además depende de la estabilidad del filtro de LP. Opcionalmente, el factor  $\alpha$  puede ser alterado en relación con la longitud de altura. Por ejemplo, si la altura es realmente larga, entonces  $\alpha$  puede mantenerse normal, pero si la altura es realmente corta, puede ser conveniente (o necesario) copiar una cantidad de veces la misma parte de excitación pasada. Debido a que se ha hallado que  
15 esto rápidamente sonará demasiado artificial, la señal, por lo tanto se desvanece en forma saliente más velozmente.

Además opcionalmente, es posible considerar la salida de predicción de altura. Si se predice una altura, esto significa que la altura ya estaba cambiando en la trama previa, y entonces, cuantas más tramas se pierdan, más lejos estamos de la verdad. Por lo tanto, es deseable acelerar algo el desvanecimiento en forma saliente de la parte  
20 tonal, en este caso.

Si la predicción de altura fallara debido a que la altura cambia demasiado, esto significa que o bien los valores de altura no son realmente confiables, o que la señal es realmente impredecible. Por lo tanto, nuevamente, deberíamos realizar el desvanecimiento saliente en forma más veloz.  
25

Como conclusión, la contribución de la señal 652 de excitación de dominio de tiempo extrapolada a la señal 672 de entrada de la síntesis 680 de codificación predictiva lineal (LPC) habitualmente se reduce en función del tiempo. Esto puede lograrse, por ejemplo, reduciendo un valor de ganancia, que se aplica a la señal 652 de excitación de dominio de tiempo extrapolada, en función del tiempo. La velocidad utilizada para reducir gradualmente la ganancia aplicada para escalar la señal 552 de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida (o una o más de sus copias) se ajusta según uno o más parámetros de una o más tramas de audio (y/o según una cantidad de tramas de audio perdidas consecutivas). En particular, la longitud de altura y/o el índice al cual cambia la altura en función del tiempo, y/o la cuestión de si una predicción de altura falla o tiene éxito, pueden utilizarse para ajustar dicha velocidad.  
30  
35

**6.6. Síntesis de codificación predictiva lineal (LPC)**

A fin de retornar al dominio de tiempo, se realiza una síntesis 680 de codificación predictiva lineal (LPC) sobre la suma en general (o generalmente, la combinación pesada) de las dos excitaciones (parte 652 tonal y parte 662 ruidosa), seguida del desénfasis 684.  
40

En otras palabras, el resultado de la combinación pesada (desvanecimiento) de la señal 652 de excitación de dominio de tiempo extrapolada y la señal 662 de ruido forma una señal de excitación de dominio de tiempo combinada, que es ingresada en la síntesis 680 de codificación predictiva lineal (LPC), que, por ejemplo, puede realizar una filtración de síntesis sobre la base de dicha señal de excitación de dominio de tiempo combinada 672 según coeficientes de codificación predictiva lineal (LPC) que describen el filtro de síntesis.  
45

**6.7. Superposición y adición**

50 Debido a que no se conoce, durante el ocultamiento, el modo de la siguiente trama que llega (por ejemplo, predicción lineal excitada por libro de códigos adaptativo (ACELP), Excitación codificada transformada (TCX) o dominio de frecuencia (FD)), se prefiere preparar diferentes superposiciones por adelantado. A fin de lograr la mejor superposición y adición si la siguiente trama se encuentra en un dominio de transformada (TCX o FD), una señal artificial (por ejemplo, una información de audio de ocultamiento de error) puede, por ejemplo, crearse para la mitad de una trama más que la trama ocultada (perdida). Aún más, puede crearse aliasing artificial sobre ella (donde el aliasing artificial puede, por ejemplo, adaptarse a la superposición y adición de transformada de coseno discreta modificada inversa (MDCT)).  
55

A fin de obtener una buena superposición y adición sin discontinuidad con la futura trama en el dominio de tiempo (ACELP [Predicción lineal excitada por libro de códigos adaptativo]), hacemos como anteriormente, pero sin aliasing, de modo de poder aplicar largas ventanas de superposición y adición, o si queremos usar una ventana cuadrada, se computa la respuesta de entrada cero (ZIR) al final del almacenamiento en memoria intermedia de síntesis.  
60

Como conclusión, en un decodificador de audio de conmutación (que, por ejemplo, puede conmutar entre una



5 decodificación de predicción lineal excitada por libro de códigos adaptativo (ACELP), una decodificación de  
 10 excitación codificada transformada (TCX) y una decodificación de dominio de frecuencia (decodificación de FD)),  
 puede realizarse una superposición y adición entre la información de audio de ocultamiento de error que se  
 proporciona principalmente para una trama de audio perdida, pero además, para una cierta porción de tiempo tras la  
 trama de audio perdida, y la información de audio decodificada provista para la primera trama de audio  
 apropiadamente decodificada tras una secuencia de una o más tramas de audio perdidas. A fin de obtener una  
 apropiada superposición y adición, incluso, para modos de decodificación que acarrear un aliasing de dominio de  
 tiempo en una transición entre subsiguientes tramas de audio, puede proporcionarse una información de cancelación  
 de aliasing (por ejemplo, designada aliasing artificial). Por consiguiente, una superposición y adición entre la  
 información de audio de ocultamiento de error y la información de audio de dominio de tiempo obtenida sobre la  
 base de la primera trama de audio apropiadamente decodificada tras una trama de audio perdida, logra una  
 cancelación del aliasing.

15 Si la primera trama de audio apropiadamente decodificada tras la secuencia de una o más tramas de audio perdidas  
 es codificada en el modo de predicción lineal excitada por libro de códigos adaptativo (ACELP), puede computarse  
 una información de superposición específica, que puede sustentarse en una respuesta de entrada cero (ZIR,  
 conforme a sus siglas en inglés) de un filtro de codificación predictiva lineal (LPC).

20 Como conclusión, el ocultamiento 600 de error se adapta bien para el uso en un códec de audio de conmutación. Sin  
 embargo, el ocultamiento 600 de error puede además ser usado en un códec de audio que meramente decodifica un  
 contenido de audio codificado en un modo de excitación codificada transformada (TCX) o en un modo de predicción  
 lineal excitada por libro de códigos adaptativo (ACELP).

### 25 **6.8. Conclusión**

Debe observarse que se logra un ocultamiento de error particularmente bueno mediante el concepto arriba  
 mencionado, para la extrapolación de una señal de excitación de dominio de tiempo, la combinación del resultado de  
 la extrapolación con una señal de ruido usando un desvanecimiento (por ejemplo, un desvanecimiento cruzado), y  
 para la realización de una síntesis de codificación predictiva lineal (LPC) sobre la base de un resultado del  
 desvanecimiento cruzado.

### 30 **7. Decodificador de audio según la figura 11**

35 La figura 11 muestra un diagrama de bloques esquemático de un decodificador 1100 de audio, según una  
 realización de la presente invención.

Debe observarse que el decodificador 1100 de audio puede ser parte de un decodificador de audio de conmutación.  
 Por ejemplo, el decodificador 1100 de audio puede reemplazar la vía 440 de decodificación de dominio de predicción  
 lineal en el decodificador 400 de audio.

40 El decodificador 1100 de audio está configurado para recibir una información 1110 de audio codificada y para  
 proporcionar, sobre su base, una información 1112 de audio decodificada. La información 1110 de audio codificada  
 puede, por ejemplo, corresponder a la información 410 de audio codificada, y la información 1112 de audio  
 decodificada puede, por ejemplo, corresponder a la información 412 de audio decodificada.

45 El decodificador 1100 de audio comprende un analizador 1120 de corriente de bits, que está configurado para  
 extraer una representación 1122 codificada de un conjunto de coeficientes espectrales y una representación 1124  
 codificada de coeficientes de codificación de predicción lineal de la información 1110 de audio codificada. Sin  
 embargo, el analizador 1120 de corriente de bits puede opcionalmente extraer información adicional de la  
 información 1110 de audio codificada.

50 El decodificador 1100 de audio además comprende una decodificación 1130 de valor espectral, que está  
 configurada para proporcionar un conjunto 1132 de valores espectrales decodificados sobre la base de los  
 coeficientes 1122 espectrales codificados. Puede usarse cualquier concepto de decodificación conocido para la  
 decodificación de coeficientes espectrales.

55 El decodificador 1100 de audio además comprende un coeficiente 1140 de codificación de predicción lineal para la  
 conversión de factor de escala, que está configurado para proporcionar un conjunto 1142 de factores de escala  
 sobre la base de la representación 1124 codificada de coeficientes de codificación de predicción lineal. Por ejemplo,  
 60 el coeficiente 1142 de codificación de predicción lineal para la conversión de factor de escala puede realizar una  
 funcionalidad que se describe en el estándar USAC [Codificación unificada de voz y audio]. Por ejemplo, la  
 representación 1124 codificada de los coeficientes de codificación de predicción lineal puede comprender una  
 representación polinomial, que es decodificada y convertida en un conjunto de factores de escala por el coeficiente  
 1142 de codificación de predicción lineal para la conversión de factor de escala.

El decodificador 1100 de audio además comprende un escalador 1150, que está configurado para aplicar los factores 1142 de escala a los valores 1132 espectrales decodificados, de modo de obtener valores 1152 espectrales decodificados escalados. Aún más, el decodificador 1100 de audio comprende, opcionalmente, un procesamiento 1160, que, por ejemplo, puede corresponder al procesamiento 366 que se describe con anterioridad, donde los valores 1162 espectrales decodificados escalados procesados son obtenidos por el procesamiento 1160 opcional. El decodificador 1100 de audio además comprende una transformada 1170 de dominio de frecuencia a dominio de tiempo, que está configurada para recibir los valores 1152 espectrales decodificados escalados (que pueden corresponder a los valores 362 espectrales decodificados escalados), o los valores 1162 espectrales decodificados escalados procesados (que pueden corresponder a los valores 368 espectrales decodificados escalados procesados) y proporcionar, sobre su base, una representación 1172 de dominio de tiempo, que puede corresponder a la representación 372 de dominio de tiempo que se describe con anterioridad. El decodificador 1100 de audio además comprende un primer 1174 posprocesamiento opcional, y un segundo 1178 posprocesamiento opcional, que, por ejemplo, pueden corresponder, al menos en parte, al posprocesamiento 376 opcional mencionado con anterioridad. Por consiguiente, el decodificador 1110 de audio obtiene (opcionalmente) una versión 1179 posprocesada de la representación 1172 de audio de dominio de tiempo.

El decodificador 1100 de audio además comprende un bloque 1180 de ocultamiento de error que está configurado para recibir la representación 1172 de audio de dominio de tiempo, o una versión posprocesada de esta, y los coeficientes de codificación de predicción lineal (o bien en forma codificada, o en una forma decodificada) y proporciona, sobre su base, una información 1182 de audio de ocultamiento de error.

El bloque 1180 de ocultamiento de error está configurado para proporcionar la información 1182 de audio de ocultamiento de error para el ocultamiento de una pérdida de una trama de audio tras una trama de audio codificada en una representación de dominio de frecuencia usando una señal de excitación de dominio de tiempo, y por lo tanto, es similar al ocultamiento 380 de error y al ocultamiento 480 de error, y además, al ocultamiento 500 de error y al ocultamiento 600 de error.

Sin embargo, el bloque 1180 de ocultamiento de error comprende un análisis 1184 de codificación predictiva lineal (LPC), que es sustancialmente idéntico al análisis 530 de codificación predictiva lineal (LPC). Sin embargo, el análisis 1184 de codificación predictiva lineal (LPC) puede, opcionalmente, usar los coeficientes 1124 de codificación predictiva lineal (LPC) para facilitar el análisis (en comparación con el análisis 530 de codificación predictiva lineal (LPC)). El análisis 1134 de codificación predictiva lineal (LPC) proporciona una señal de excitación de dominio de tiempo 1186, que es sustancialmente idéntica a la señal 532 de excitación de dominio de tiempo (y además, a la señal 610 de excitación de dominio de tiempo). Aún más, el bloque 1180 de ocultamiento de error comprende un ocultamiento 1188 de error, que, por ejemplo, puede realizar la funcionalidad de los bloques 540, 550, 560, 570, 580, 584 del ocultamiento 500 de error, o que, por ejemplo, puede realizar la funcionalidad de los bloques 640, 650, 660, 670, 680, 684 del ocultamiento 600 de error. Sin embargo, el bloque 1180 de ocultamiento de error difiere levemente del ocultamiento 500 de error, y además, del ocultamiento 600 de error. Por ejemplo, el bloque 1180 de ocultamiento de error (que comprende el análisis 1184 de codificación predictiva lineal (LPC)) difiere del ocultamiento 500 de error en términos de que los coeficientes de codificación predictiva lineal (LPC) (utilizados para la síntesis 580 de codificación predictiva lineal (LPC)) no son determinados por el análisis 530 de codificación predictiva lineal (LPC), si bien son (opcionalmente) recibidos desde la corriente de bits. Asimismo, el bloque 1188 de ocultamiento de error, que comprende el análisis 1184 de codificación predictiva lineal (LPC), difiere del ocultamiento 600 de error en términos de que la "excitación 610 pasada" es obtenida por el análisis 1184 de codificación predictiva lineal (LPC), en lugar de estar disponible directamente.

El decodificador 1100 de audio además comprende una combinación 1190 de señales, que está configurada para recibir la representación 1172 de audio de dominio de tiempo, o una versión posprocesada de la misma, y además, la información 1182 de audio de ocultamiento de error (naturalmente, para subsiguientes tramas de audio), y combina dichas señales, preferentemente, usando una operación de superposición y adición, de modo de obtener la información 1112 de audio decodificada.

Para más detalles, se hace referencia a las explicaciones anteriores.

## **8. Método según la figura 9**

La figura 9 muestra un diagrama de flujo de un método para proporcionar una información de audio decodificada sobre la base de una información de audio codificada. El método 900 según la figura 9 comprende la provisión de 910 una información de audio de ocultamiento de error para el ocultamiento de una pérdida de una trama de audio tras una trama de audio codificada en una representación de dominio de frecuencia usando una señal de excitación de dominio de tiempo. El método 900 según la figura 9 se sustenta en las mismas consideraciones que el decodificador de audio según la figura 1. Aún más, debe observarse que el método 900 puede ser suplementado por cualquiera de los rasgos y de las funcionalidades que se describen en esta solicitud, o bien individualmente, o bien

en combinación.

### **9. Método según la figura 10**

5 La figura 10 muestra un diagrama de flujo de un método para proporcionar una información de audio decodificada sobre la base de una información de audio codificada. El método 1000 comprende la provisión de 1010 una información de audio de ocultamiento de error para el ocultamiento de una pérdida de una trama de audio, donde una señal de excitación de dominio de tiempo obtenida para (o sobre la base de) una o más tramas de audio que preceden una trama de audio perdida es modificada a fin de obtener la información de audio de ocultamiento de error.

El método 1000 según la figura 10 se sustenta en las mismas consideraciones que el decodificador de audio mencionado anteriormente según la figura 2.

15 Aún más, debe observarse que el método según la figura 10 puede ser suplementado por cualquiera de los rasgos y de las funcionalidades que se describen en esta solicitud, o bien individualmente, o en combinación.

### **10. Observaciones adicionales**

20 En las realizaciones descritas anteriormente, pueden manipularse múltiples pérdidas de trama de diferentes maneras. Por ejemplo, si se pierden dos o más tramas, la parte periódica de la señal de excitación de dominio de tiempo para la segunda trama perdida puede derivar de (o ser igual a) una copia de la parte tonal de la señal de excitación de dominio de tiempo asociada con la primera trama perdida. Alternativamente, la señal de excitación de dominio de tiempo para la segunda trama perdida puede sustentarse en un análisis de codificación predictiva lineal (LPC) de la señal de síntesis de la trama previa perdida. Por ejemplo, en un códec, la codificación predictiva lineal (LPC) puede ser cambiante en cada trama perdida; entonces, tiene sentido la nueva realización del análisis para cada trama perdida.

### **11. Alternativas de implementación**

30 Si bien se han descrito algunos aspectos en el contexto de un aparato, es claro que estos aspectos además representan una descripción del correspondiente método, donde un bloque o dispositivo corresponde a una etapa de método o a un rasgo de una etapa de método. En forma análoga, los aspectos descritos en el contexto de una etapa de método además representan una descripción de un correspondiente bloque o ítem o rasgo de un correspondiente aparato. Algunas o la totalidad de las etapas de método pueden ser ejecutadas por (o usando) un aparato de soporte físico, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, una o más de las etapas de método más importantes pueden ser ejecutadas por dicho aparato.

40 Según ciertos requerimientos de implementación, las realizaciones de la invención pueden ser implementadas en soporte físico o soporte lógico. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo, un disquete, un DVD (disco versátil digital), un Blu-Ray, un CD (disco compacto, conforme a sus siglas en inglés), una ROM (memoria de solo lectura, conforme a sus siglas en inglés), una PROM (memoria de solo lectura programable, conforme a sus siglas en inglés), una EPROM (memoria de solo lectura programable de borrado, conforme a sus siglas en inglés), una EEPROM (memoria de solo lectura programable de borrado electrónico, conforme a sus siglas en inglés) o una memoria FLASH, que tiene señales de control electrónicamente legibles allí almacenadas, que cooperan (o son capaces de cooperar) con un sistema de ordenador programable de manera tal de llevar a cabo el respectivo método. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

50 Algunas realizaciones según la invención comprenden un portador de datos que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema de ordenador programable, de manera de llevar a cabo uno de los métodos que se describen en el presente documento.

55 En general, las realizaciones de la presente invención pueden ser implementadas como un producto de programa de ordenador con un código de programa, donde el código de programa es operativo para llevar a cabo uno de los métodos cuando se ejecuta el producto de programa de ordenador en un ordenador. El código de programa puede ser almacenado, por ejemplo, en un portador legible en una máquina.

60 Otras realizaciones comprenden el programa de ordenador para llevar a cabo uno de los métodos que se describen en el presente documento, almacenado en un portador legible en una máquina.

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

Una realización adicional del método de la invención es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, allí grabado, el programa de ordenador para llevar a cabo uno de los métodos que se describen en el presente documento. El portador de datos, el medio de almacenamiento digital o el medio grabado son típicamente tangibles y/o no transitorios.

Una realización adicional del método de la invención es, por lo tanto, una corriente de datos o una secuencia de señales que representan el programa de ordenador para llevar a cabo uno de los métodos que se describen en el presente documento. La corriente de datos o la secuencia de señales, por ejemplo, puede estar configurada para ser transferida por medio de una conexión de comunicación de datos, por ejemplo, por medio de Internet.

Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los métodos que se describen en el presente documento.

Una realización adicional comprende un ordenador que tiene allí instalado el programa de ordenador para llevar a cabo uno de los métodos que se describen en el presente documento.

Una realización adicional según la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa de ordenador para la realización de uno de los métodos descritos en esta solicitud, a un receptor. El receptor puede ser, por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o el sistema puede comprender, por ejemplo, un servidor de archivos para transferir el programa de ordenador al receptor.

En algunas realizaciones, puede usarse un dispositivo lógico programable (por ejemplo, una matriz de puertas programables de campo) para realizar algunas o la totalidad de las funcionalidades de los métodos que se describen en el presente documento. En algunas realizaciones, una matriz de puertas programables de campo puede cooperar con un microprocesador a fin de llevar a cabo uno de los métodos que se describen en el presente documento. En general, los métodos son llevados a cabo, preferentemente, por cualquier aparato de soporte físico.

El aparato descrito en el presente documento puede ser implementado usando un aparato de soporte físico, o usando un ordenador, o empleando una combinación de un aparato de soporte físico y un ordenador.

Los métodos que se describen en esta solicitud pueden ser realizados usando un aparato de soporte físico, o usando un ordenador, o empleando una combinación de un aparato de soporte físico y un ordenador.

Las realizaciones descritas con anterioridad son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles que se describen en el presente documento serán evidentes para los expertos en la técnica. Por lo tanto, se tiene la intención de limitación solo por el alcance de las reivindicaciones de patente inminentes, y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones del presente documento.

## **12. Conclusiones**

Como conclusión, si bien cierto ocultamiento para códecs de dominio de transformada ha sido descrito en el campo, las realizaciones según la invención superan los códecs (o decodificadores) convencionales. Las realizaciones según la invención usan un cambio de dominio para el ocultamiento (dominio de frecuencia a dominio de tiempo o excitación). Por consiguiente, las realizaciones según la invención crean un ocultamiento de voz de alta calidad para decodificadores de dominio de transformada.

El modo de codificación de transformada es similar a aquel en USAC (conferir, por ejemplo, referencia [3]). Utiliza la Transformada de coseno discreta modificada (MDCT) como una transformada, y la formación de ruido espectral se logra mediante la aplicación de la envoltura espectral de codificación predictiva lineal (LPC) pesada en el dominio de frecuencia (además conocida como FDNS, "formación de ruido de dominio de frecuencia"). En otras palabras, las realizaciones según la invención pueden usarse en un decodificador de audio, que utiliza los conceptos de decodificación que se describen en el estándar USAC. Sin embargo, el concepto de ocultamiento de error divulgado en esta solicitud puede además ser usado en un decodificador de audio que es de tipo "AAC" (Codificación avanzada de audio), o en cualquier códec (o decodificador) de la familia AAC.

El concepto según la presente invención se aplica a un códec conmutado tal como USAC, al igual que a un códec de dominio de frecuencia puro. En ambos casos, el ocultamiento se realiza en el dominio de tiempo o en el dominio de excitación.

En lo que sigue, se describirán algunas ventajas y algunos rasgos del ocultamiento de dominio de tiempo (o del

ocultamiento de dominio de excitación).

La ocultación de TCX convencional, como se describe, por ejemplo, con referencia a las Figs. 7 y 8, se denomina además sustitución de ruido, no es bien adecuada para las señales de tipo voz, o incluso, para señales tonales. Las realizaciones según la invención crean un nuevo ocultamiento para un códec de dominio de transformada que se aplica en el dominio de tiempo (o en el dominio de excitación de un decodificador de codificación de predicción lineal). Es similar a un ocultamiento de tipo ACELP (predicción lineal excitada por libro de códigos adaptativo), y aumenta la calidad del ocultamiento. Se ha hallado que la información de altura es conveniente (o incluso requerida, en algunos casos) para un ocultamiento de tipo ACELP. Por lo tanto, las realizaciones según la presente invención están configuradas para hallar valores de altura confiables para la trama previa codificada en el dominio de frecuencia.

Diferentes partes y detalles se han explicado anteriormente, por ejemplo, basándose en las realizaciones de acuerdo con las Figs. 5 y 6.

Como conclusión, las realizaciones según la invención crean un ocultamiento de error que supera las soluciones convencionales.

El siguiente conjunto de aspectos reproduce el contenido originalmente presentado en la solicitud anterior. No define el alcance de protección que se define únicamente por las reivindicaciones adjuntas.

Según un primer aspecto, un decodificador 200; 400 de audio para proporcionar una información 220; 412 de audio decodificada sobre la base de una información 210;410 de audio codificada puede comprender: un ocultamiento 240; 480; 600 de error configurado para proporcionar una información 242;482;612 de audio de ocultamiento de error para el ocultamiento de una pérdida de una trama de audio, en la que el ocultamiento de error está configurado para modificar una señal 452,456;610 de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, para obtener la información de audio de ocultación de error. Según un segundo aspecto cuando se refiere de nuevo al primer aspecto, en el decodificador de audio, la ocultación de error puede estar configurada para modificar una señal 452, 456; 610 de excitación de dominio de tiempo derivada de una o más tramas de audio codificadas en una representación de dominio de frecuencia que precede una trama de audio perdida, para obtener la información de audio de ocultamiento de error.

Según un tercer aspecto cuando se refiere de nuevo al primer aspecto o al segundo aspecto, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para el uso de una o más copias modificadas de la señal 452,456;610 de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, para obtener la información 242;482;612 de ocultamiento de error.

Según un cuarto aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al tercero, en el decodificador 200;400 de audio, el ocultamiento 240;482;612 de error puede estar configurado para modificar la señal 452,456;610 de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, o una o más copias de las mismas, para reducir de ese modo un componente periódico de la información 242; 482; 612 de audio de ocultamiento de error a lo largo del tiempo.

Según un quinto aspecto cuando se refiere de nuevo a uno cualquiera del aspecto primero al cuarto, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para escalar la señal 452,456;610 de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden la trama de audio perdida, o una o más copias de las mismas, para modificar de ese modo la señal de excitación de dominio de tiempo.

Según un sexto aspecto cuando se refiere de nuevo al aspecto cuarto o quinto, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para reducir gradualmente una ganancia aplicada para escalar la señal 452,456;610 de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, o la una o más copias de la misma.

Según un séptimo aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos cuarto al sexto, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para ajustar una velocidad usada para reducir gradualmente una ganancia aplicada para escalar la señal 452,456;610 de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, o la una o más copias de las mismas, de acuerdo con uno o más parámetros de una o más tramas de audio que preceden la trama de audio perdida, y/o de acuerdo con un número de tramas de audio perdidas consecutivas.

Según un octavo aspecto cuando se refiere de nuevo al aspecto sexto o séptimo, en el decodificador 200; 400 de audio, el ocultamiento 240;480;600 de error puede estar configurado para ajustar la velocidad usada para reducir gradualmente una ganancia aplicada para escalar la señal 452,456;610 de excitación de dominio de tiempo obtenida

- 5 para una o más tramas de audio que preceden una trama de audio perdida, o la una o más copias de las mismas, de acuerdo con una longitud de un periodo de altura de la señal de excitación de dominio de tiempo, de modo que un componente determinista de una entrada de señal 672 de excitación de dominio de tiempo en una síntesis 680 de LPC es desvanecida en forma saliente más velozmente para señales que tienen una longitud más corta del periodo de altura en comparación con señales que tienen una longitud más larga del periodo de altura.
- 10 Según un noveno aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos sexto al octavo, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal 452,456;610 de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, o la una o más copias de las mismas, de acuerdo con un resultado de un análisis de altura o una predicción de altura, de modo que un componente determinista de una entrada de señal 572 de excitación de dominio de tiempo en una síntesis 580 de LPC puede estar desvanecida en forma saliente más velozmente para señales que tienen un cambio de altura mayor por unidad de tiempo, en comparación con señales que tienen un menor cambio de altura por
- 15 unidad de tiempo; y/o de modo tal que un componente determinista de una entrada de señal 572 de excitación de dominio de tiempo en una síntesis 580 de LPC puede estar desvanecida en forma saliente más velozmente para señales para las cuales falla una predicción de altura en comparación con señales para las cuales la predicción de altura tiene éxito.
- 20 Según un décimo aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al noveno, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para escalar en el tiempo la señal 452,456;610 de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida, o la una o más copias de las mismas, de acuerdo con una predicción de una altura durante un tiempo de la una o más tramas de audio perdidas.
- 25 Según un undécimo aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al décimo, en el decodificador 200;400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para obtener una señal 452,456;610 de excitación de dominio de tiempo, que se ha utilizado para decodificar una o más tramas de audio que preceden la trama de audio perdida, y para modificar dicha señal de excitación de dominio de tiempo, que se ha utilizado para decodificar una o más tramas de audio que preceden la trama de audio perdida, para obtener una señal 652, 672 de excitación de dominio de tiempo modificada y el ocultamiento de error puede estar configurado para proporcionar la información 242; 482; 612 de audio de ocultamiento de error basándose en la señal de excitación 652, 672 de dominio de tiempo modificada.
- 30 Según un duodécimo aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al undécimo, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para obtener una información de altura, que se ha utilizado para decodificar una o más tramas de audio que preceden la trama de audio perdida, y el ocultamiento de error puede estar configurado para proporcionar la información 242;482;612 de audio de ocultamiento de error de acuerdo con dicha información de altura.
- 35 Según un decimotercer aspecto cuando se refiere de nuevo al duodécimo aspecto, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para obtener la información de altura sobre la base de la señal de excitación de dominio de tiempo derivada de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida.
- 40 Según un decimocuarto aspecto cuando se refiere de nuevo al decimotercer aspecto, en el decodificador 200; 400 de audio, el ocultamiento de error puede estar configurado para evaluar una correlación cruzada de la señal de excitación de dominio de tiempo, para determinar una información de altura aproximada, y el ocultamiento de error puede estar configurado para refinar la información de altura aproximada usando una búsqueda de circuito cerrado alrededor de una altura determinada por la información de altura aproximada.
- 45 Según un decimoquinto aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al decimocuarto, en el decodificador de audio el ocultamiento de error puede estar configurado para obtener una información de altura sobre la base de una información lateral de la información de audio codificada.
- 50 Según un decimosexto aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al decimoquinto, en el decodificador de audio el ocultamiento de error puede estar configurado para obtener una información de altura sobre la base de una información de altura disponible para una trama de audio previamente decodificada.
- 55 Según un decimoséptimo aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al decimosexto, en el decodificador de audio el ocultamiento de error puede estar configurado para obtener una información de altura sobre la base de una búsqueda de altura realizada sobre una señal de dominio de tiempo o sobre una señal residual.
- 60

Según un decimoctavo aspecto cuando se refiere de nuevo a uno cualquiera del primer al decimoséptimo aspecto, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para obtener un conjunto de coeficientes 462, 466 de predicción lineal, que se han utilizado para decodificar una o más tramas de audio que preceden la trama de audio perdida, y el ocultamiento de error puede estar configurado para proporcionar la información 242;482;612 de audio de ocultamiento de error de acuerdo con dicho conjunto de coeficientes de predicción lineal.

Según un decimonoveno aspecto cuando se refiere de nuevo al decimoctavo aspecto, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para extrapolar un nuevo conjunto de coeficientes de predicción lineal basándose en el conjunto de coeficientes 462, 466 de predicción lineal, que se han utilizado para decodificar una o más tramas de audio que preceden la trama de audio perdida, y el ocultamiento de error puede estar configurado para usar el nuevo conjunto de coeficientes de predicción lineal para proporcionar la información 242; 482; 612 de audio de ocultamiento de error.

Según un vigésimo aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al decimonoveno, en el decodificador 200; 400 de audio, el ocultamiento 240;480;600 de error puede estar configurado para obtener una información acerca de una intensidad de un componente de señal determinista en una o más tramas de audio que preceden una trama de audio perdida, y el ocultamiento de error puede estar configurado para comparar la información acerca de una intensidad de un componente de señal determinista en una o más tramas de audio que preceden una trama de audio perdida, con un valor de umbral, para decidir si introducir una señal 652 de excitación de dominio de tiempo determinista con la adición de una señal 662 de excitación de dominio de tiempo de tipo ruido en una síntesis 680 de LPC, o si introducir solo una señal 662 de excitación de dominio de tiempo de ruido en la síntesis de LPC.

Según un vigésimo primer aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al vigésimo, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para obtener una información de altura que describe una altura de la trama de audio que precede la trama de audio perdida, y para proporcionar la información 242;482;612 de audio de ocultamiento de error de acuerdo con la información de altura.

Según un vigésimo segundo aspecto cuando se refiere de nuevo al vigésimo primer aspecto, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para obtener la información de altura sobre la base de la señal de excitación 452,456; 610 de dominio de tiempo asociada con la trama de audio que precede la trama de audio perdida.

Según un vigésimo tercer aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al vigésimo segundo, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para evaluar una correlación cruzada de la señal de excitación de dominio de tiempo o la señal 452,456;610 de audio de dominio de tiempo, para determinar una información de altura aproximada, y el ocultamiento de error puede estar configurado para refinar la información de altura aproximada usando una búsqueda de circuito cerrado alrededor de una altura determinada por la información de altura aproximada.

Según un vigésimo cuarto aspecto cuando se refiere de nuevo a los aspectos vigésimo segundo o vigésimo tercero, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para obtener la información de tono para proporcionar la información 242;482;612 de audio de ocultación de error basándose en una información de tono previamente calculada, que se usó para una decodificación de una o más tramas de audio que preceden la trama de audio perdida, y basándose en una evaluación de una correlación cruzada de la señal 252, 256; 610 de excitación de dominio de tiempo, que se modifica para obtener una señal 652, 672 de excitación de dominio de tiempo modificada para proporcionar la información 242; 482; 612 de audio de ocultamiento de error.

Según un vigésimo quinto aspecto cuando se refiere de nuevo al aspecto vigésimo cuarto, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para seleccionar un pico de la correlación cruzada, de una pluralidad de picos de la correlación cruzada, representando un pico un tono dependiendo de la información de tono previamente calculada, de modo tal que se elige un pico que representa un tono que está más cerca del tono representado por la información de tono previamente calculada.

Según un vigésimo sexto aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al vigésimo quinto, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para copiar un ciclo de altura de la señal 452,456;610 de excitación de dominio de tiempo asociada con la trama de audio que precede la trama de audio perdida una vez o múltiples veces, con el fin de obtener una señal de excitación 672 para una síntesis 680 de la información 242; 482; 612 de audio de ocultamiento de error.

Según un vigésimo séptimo aspecto cuando se refiere de nuevo al aspecto vigésimo sexto, en el decodificador 200;

400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para el filtro paso bajo del ciclo de altura de la señal 452,456;610 de excitación de dominio de tiempo asociada con la trama de audio que precede la trama de audio perdida usando un filtro dependiente de la tasa de muestreo, cuyo ancho de banda es dependiente de una tasa de muestreo de la trama de audio codificada en una representación de dominio de frecuencia.

5 Según un vigésimo octavo aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al vigésimo séptimo, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para predecir una altura al final de una trama perdida, y el ocultamiento de error puede estar configurado para adaptar la señal de excitación de dominio de tiempo, o una o más copias de la misma, a la altura predicha.

10 Según un vigésimo noveno aspecto cuando se refiere de nuevo a uno cualquiera de los aspectos primero al vigésimo octavo, en el decodificador 200; 400 de audio, el ocultamiento 240; 480; 600 de error puede estar configurado para combinar una señal 652 de excitación de dominio de tiempo extrapolada y una señal 662 de ruido, con el fin de obtener una señal 672 de entrada para una síntesis 680 de LPC, y el ocultamiento de error puede estar configurado para realizar la síntesis de LPC, donde la síntesis de LPC puede estar configurada para filtrar la señal de entrada de la síntesis de LPC de acuerdo con los parámetros 462,466 de codificación de predicción lineal, con el fin de obtener la información de audio de ocultamiento de error.

20 Según un trigésimo aspecto, un método 1000 para proporcionar una información de audio decodificada sobre la base de una información de audio codificada puede comprender: la provisión 1010 de una información de audio de ocultamiento de error para el ocultamiento de una pérdida de una trama de audio, donde una señal de excitación de dominio de tiempo obtenida sobre la base de una o más tramas de audio que preceden una trama de audio perdida puede ser modificada a fin de obtener la información de audio de ocultamiento de error.

25 Un trigésimo primer aspecto puede proporcionar un programa de ordenador para realizar el método según un trigésimo aspecto cuando el programa de ordenador se ejecuta en un ordenador.

#### Bibliografía

30 [1] 3GPP, "Audio codec processing functions; Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec; Transcoding functions," 2009, 3GPP TS 26.290.

[2] "MDCT-BASED CODER FOR HIGHLY ADAPTIVE SPEECH AND AUDIO CODING"; Guillaume Fuchs & al.; EUSIPCO 2009.

35 [3] ISO\_IEC\_DIS\_23003-3\_(E); Information technology - MPEG audio technologies - Part 3: Unified speech and audio coding.

40 [4] 3GPP, "General Audio Codec audio processing functions; Enhanced aacPlus general audio codec; Additional decoder tools," 2009, 3GPP TS 26.402.

[5] "Audio decoder and coding error compensating method", 2000, EP 1207519 B1

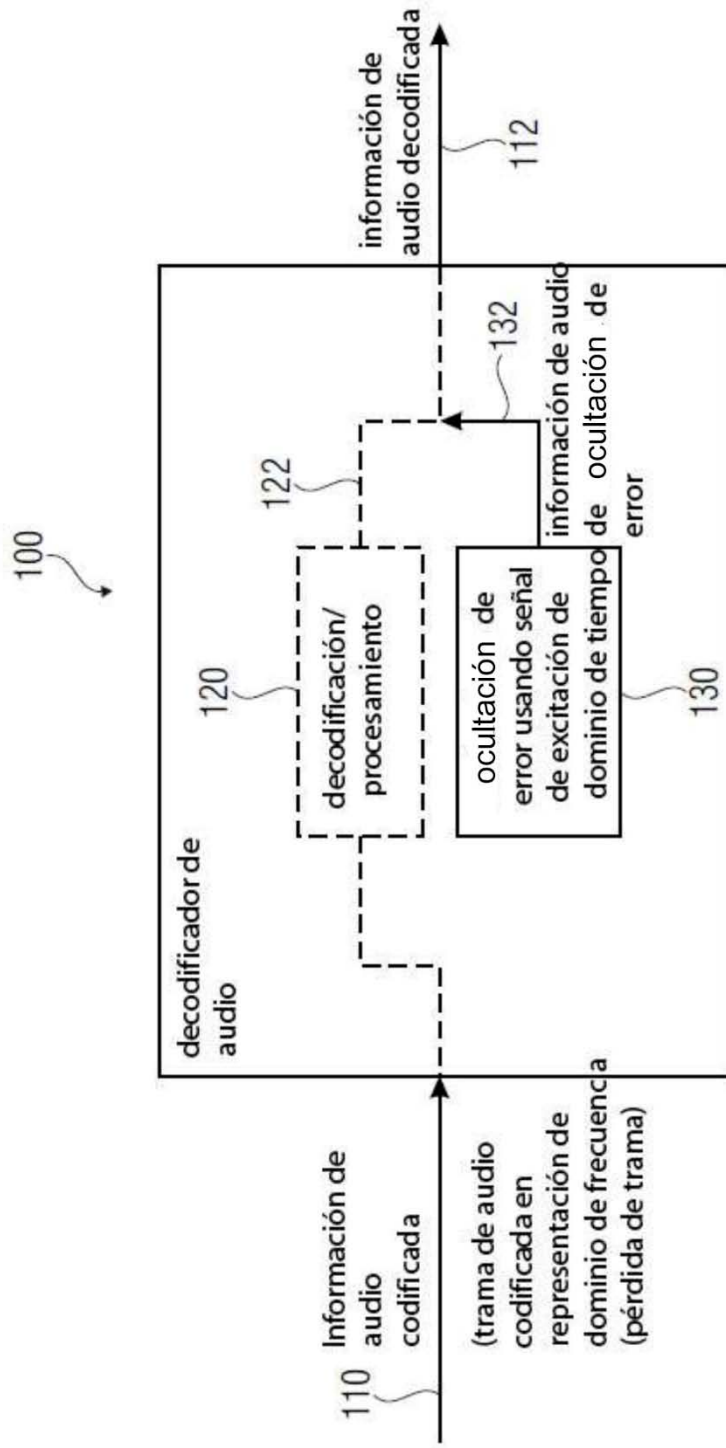
45 [6] "Apparatus and method for improved concealment of the adaptive codebook in ACELP-like concealment employing improved pitch lag estimation", 2014, PCT/EP2014/062589

[7] "Apparatus and method for improved concealment of the adaptive codebook in ACELP-like concealment employing improved pulse resynchronization", 2014, PCT/EP2014/062578



**REIVINDICACIONES**

1. Un decodificador de audio (200; 400) para proporcionar una información de audio decodificada (220; 412) basándose en una información de audio codificada (210; 410), comprendiendo el decodificador de audio:
- 5 una ocultación de error (240; 480; 600) configurada para proporcionar una información de audio de ocultación de error (242; 482; 612) para la ocultación de una pérdida de una trama de audio,
- 10 en el que la ocultación de error está configurada para modificar una señal de excitación de dominio de tiempo (452, 456; 610) obtenida para una o más tramas de audio que preceden una trama de audio perdida, para obtener la información de audio de ocultación de error;
- 15 el decodificador de audio está caracterizado porque el ocultamiento (240;480;600) de error está configurado para obtener una información acerca de una intensidad de un componente de señal determinista en una o más tramas de audio que preceden una trama de audio perdida, y
- 20 en el que el ocultamiento de error está configurado para comparar la información acerca de una intensidad de un componente de señal determinista en una o más tramas de audio que preceden una trama de audio perdida con un valor de umbral, para decidir si introducir una señal (652) de excitación de dominio de tiempo determinista con la adición de una señal (662) de excitación de dominio de tiempo de tipo ruido en una síntesis (680) de LPC, o si introducir solo una señal (662) de excitación de dominio de tiempo de ruido en la síntesis de LPC.
2. Un método (1000) para proporcionar una información de audio decodificada basándose en una información de audio codificada, comprendiendo el método:
- 25 proporcionar (1010) una información de audio de ocultación de error para la ocultación de una pérdida de una trama de audio;
- 30 en el que una señal de excitación de dominio de tiempo obtenida basándose en una o más tramas de audio que preceden una trama de audio perdida se modifica para obtener la información de audio de ocultación de error;
- 35 el método está caracterizado porque comprende obtener una información acerca de una intensidad de un componente de señal determinista en una o más tramas de audio que preceden una trama de audio perdida, y
- 40 en el que el método comprende comparar la información acerca de una intensidad de un componente de señal determinista en una o más tramas de audio que preceden una trama de audio perdida con un valor de umbral, para decidir si introducir una señal (652) de excitación de dominio de tiempo determinista con la adición de una señal (662) de excitación de dominio de tiempo de tipo ruido en una síntesis (680) de LPC, o si introducir solo una señal (662) de excitación de dominio de tiempo de ruido en la síntesis de LPC.
3. Programa de ordenador para la realización del método según la reivindicación 2, cuando el programa de ordenador se ejecuta en un ordenador.
- 45



**FIG. 1**

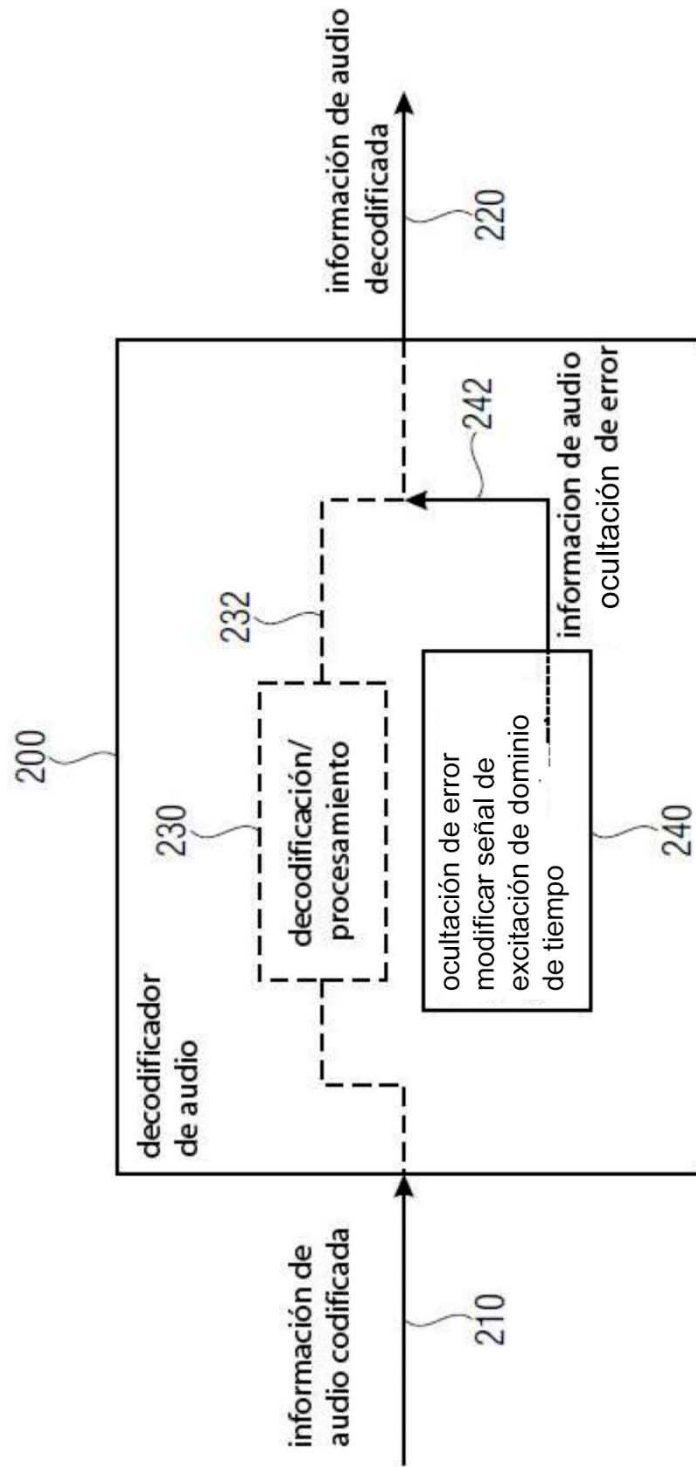


FIG. 2

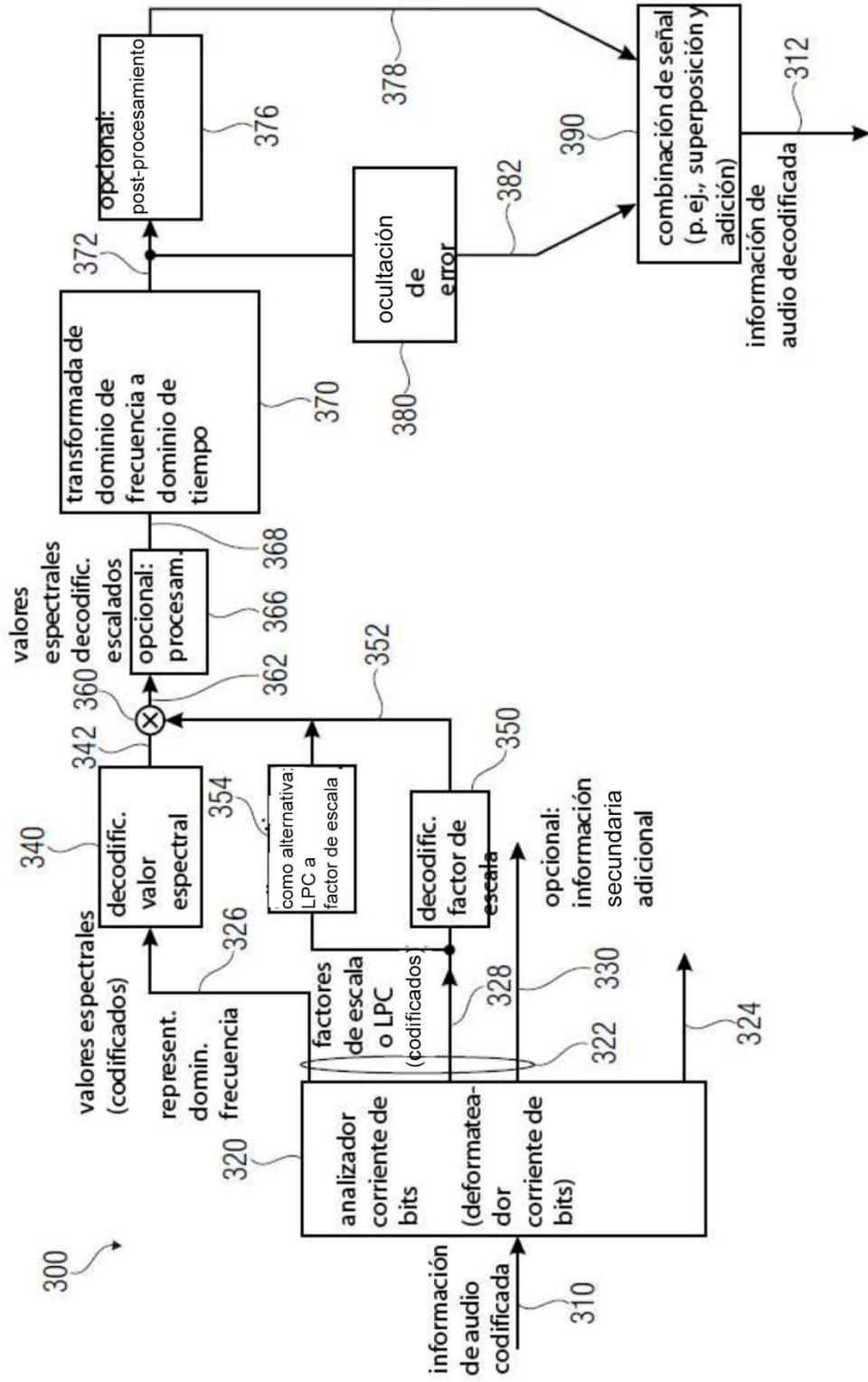


FIG. 3

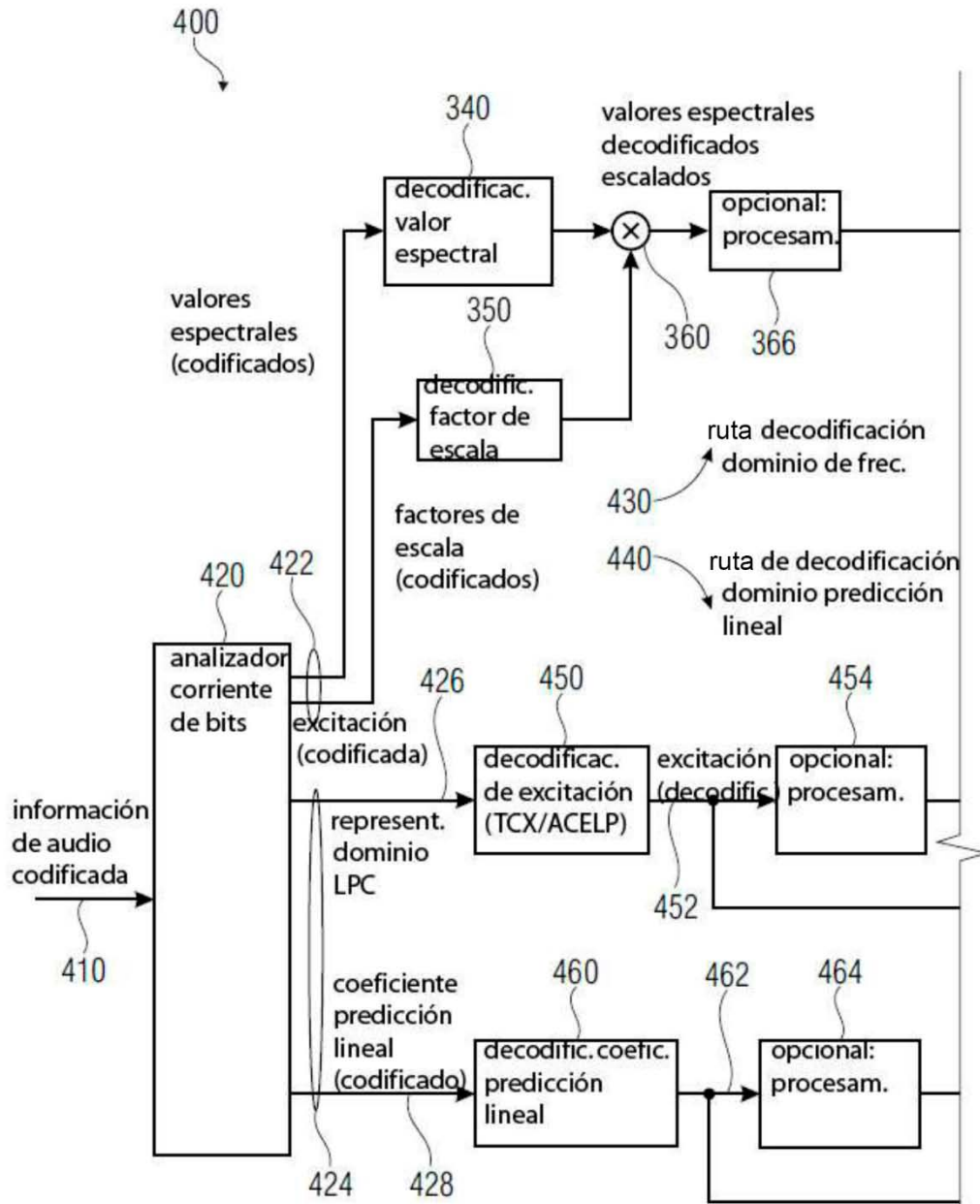


FIG 4	
FIG 4A	FIG 4B

FIG. 4A

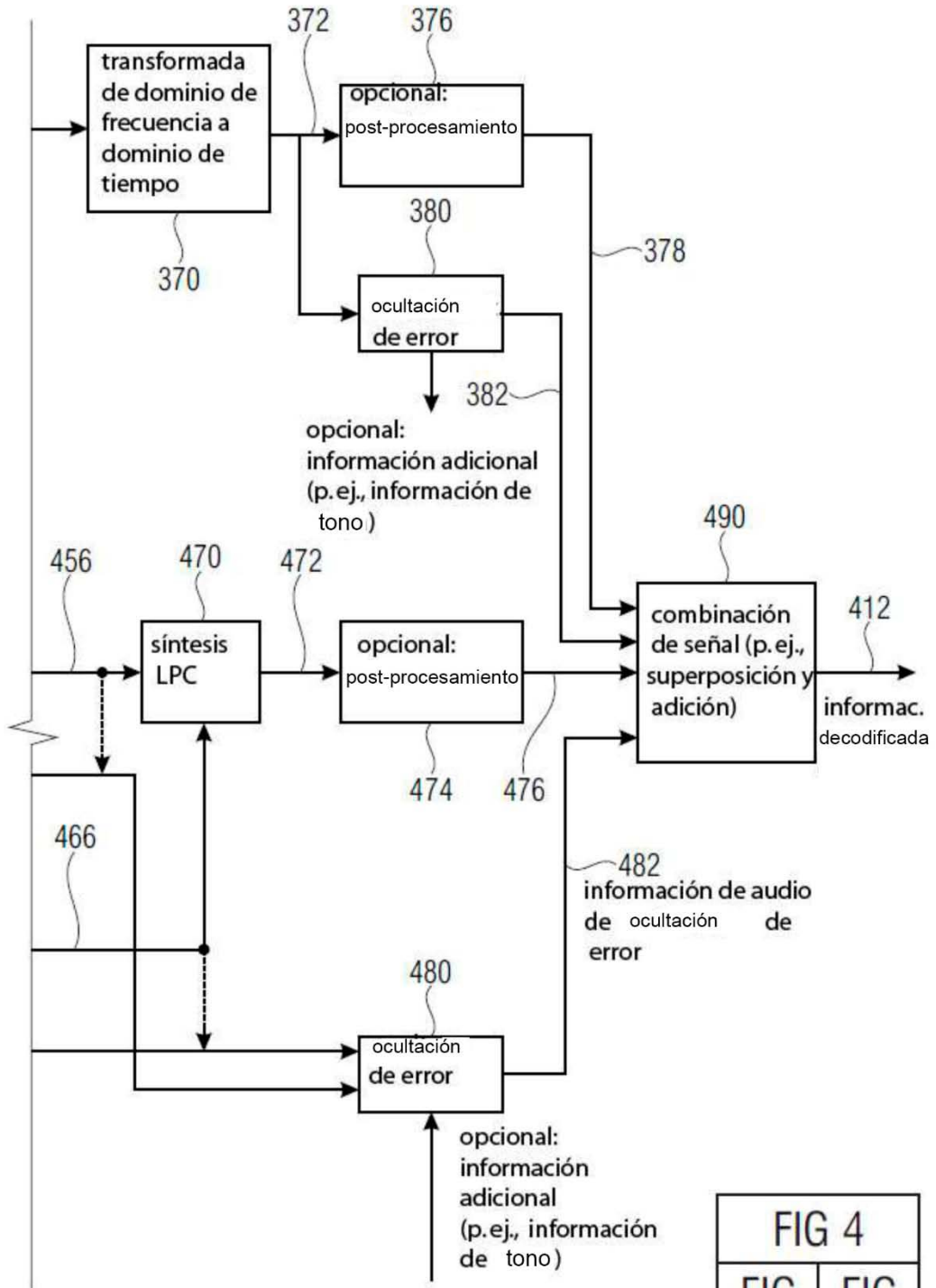
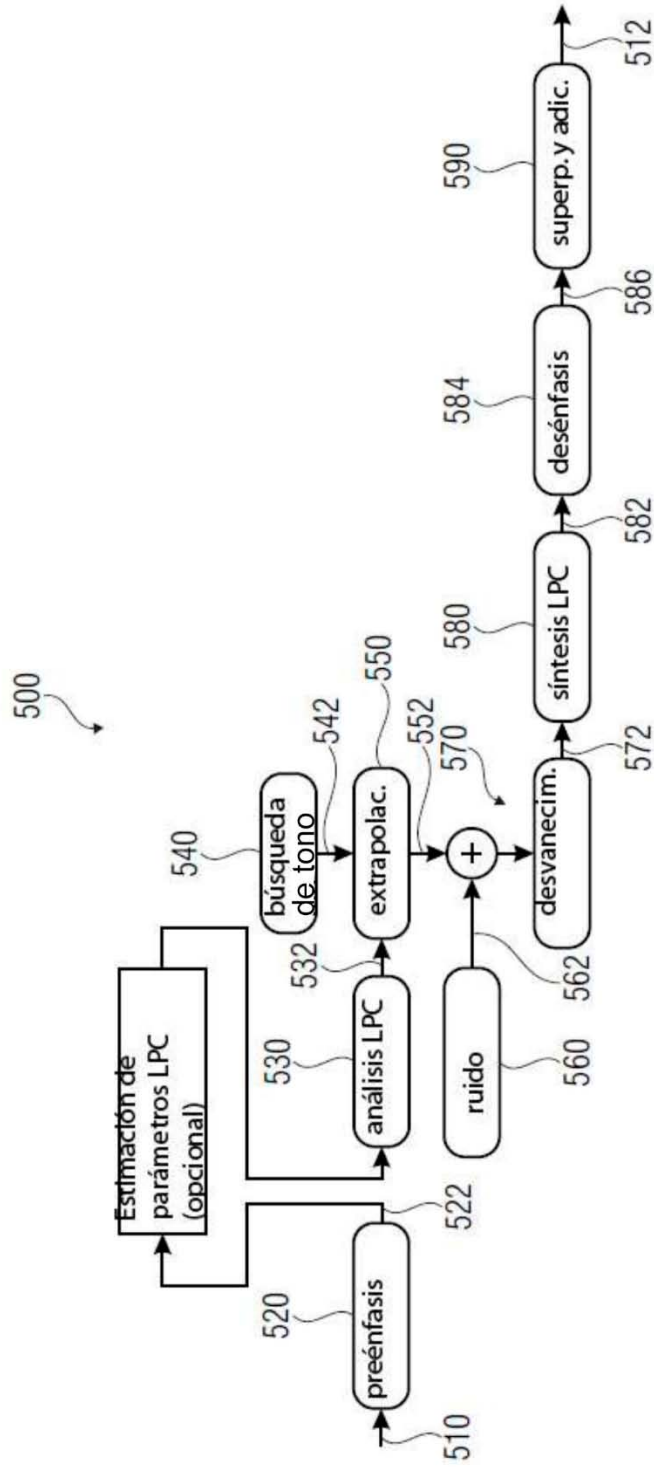


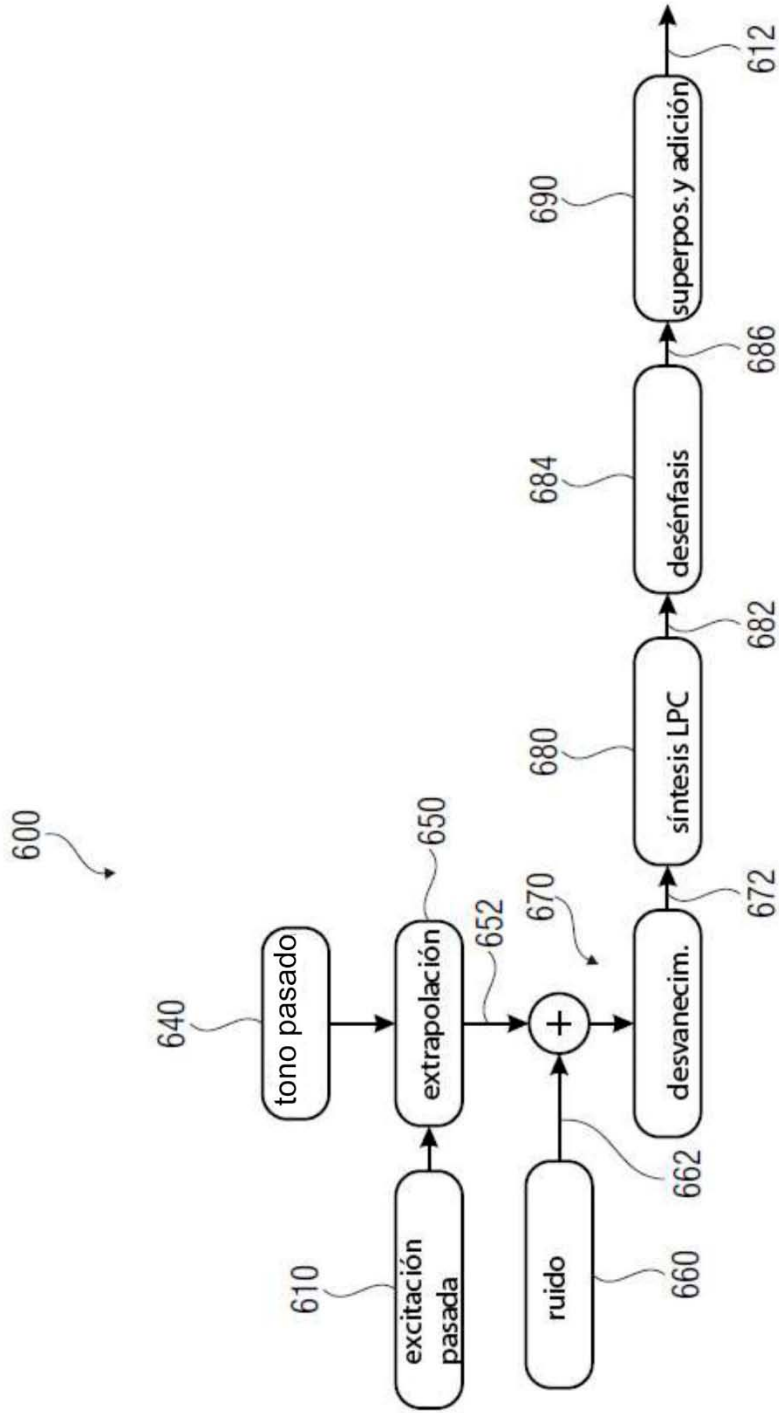
FIG. 4B

FIG 4	
FIG 4A	FIG 4B



Vista general de ocultación de dominio de tiempo para decodific. de transformada

**FIG. 5**



Vista general de ocultación de dominio de tiempo para códec de conmutación

**FIG. 6**



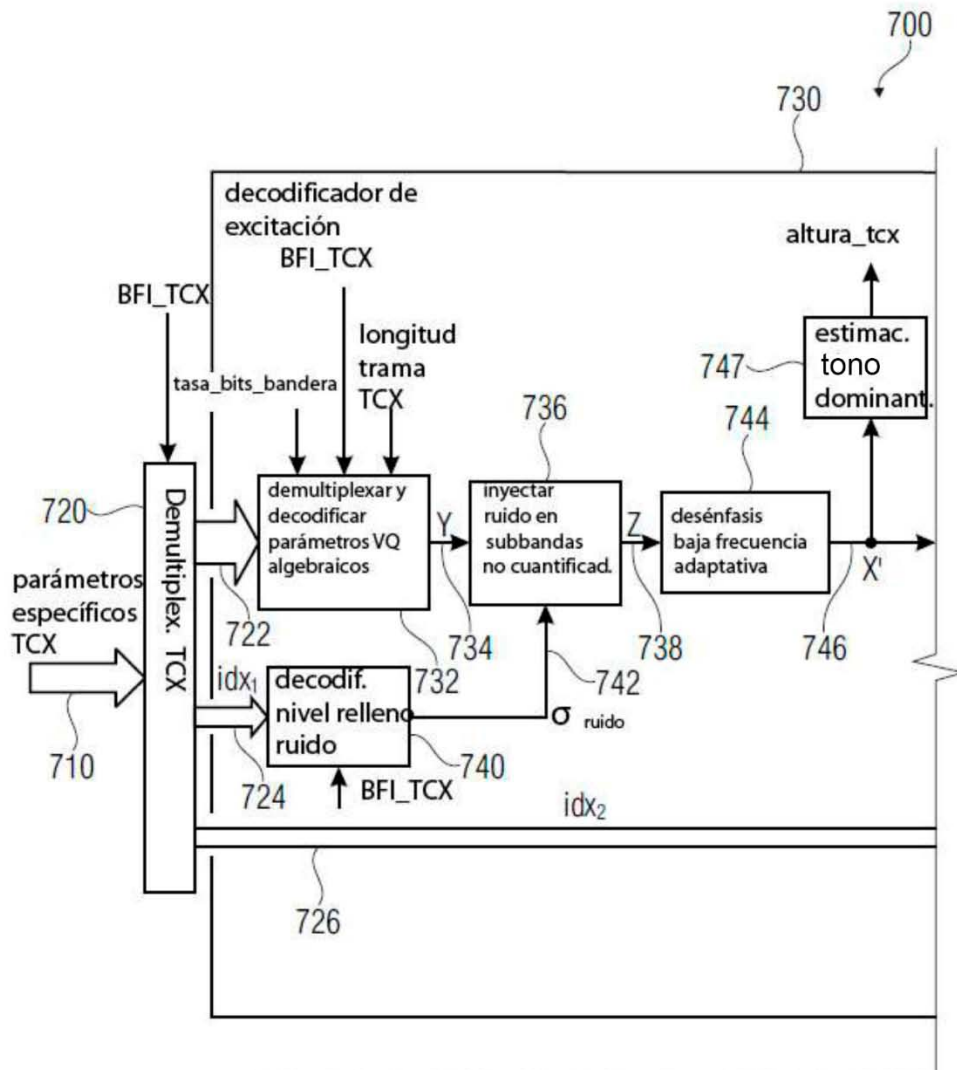


Diagrama de bloques del decodificador TCX

FIG 7	
FIG 7A	FIG 7B

**FIG. 7A**

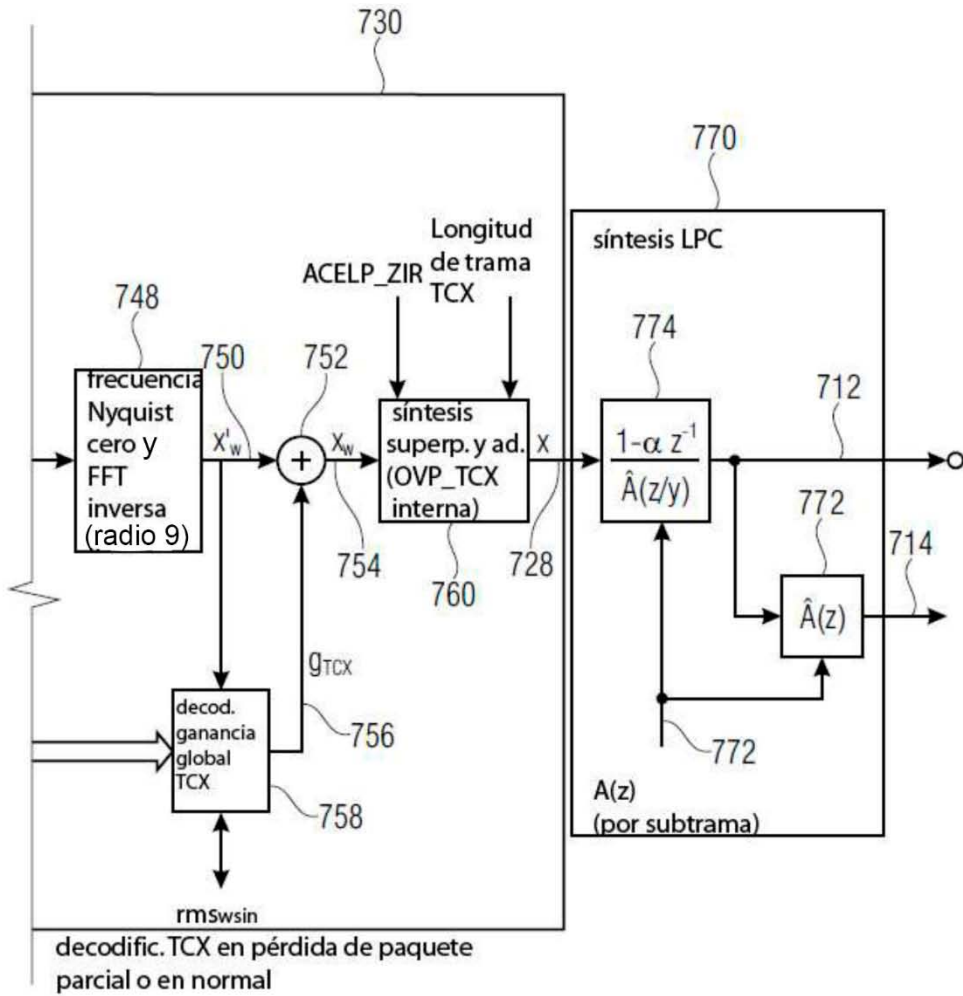
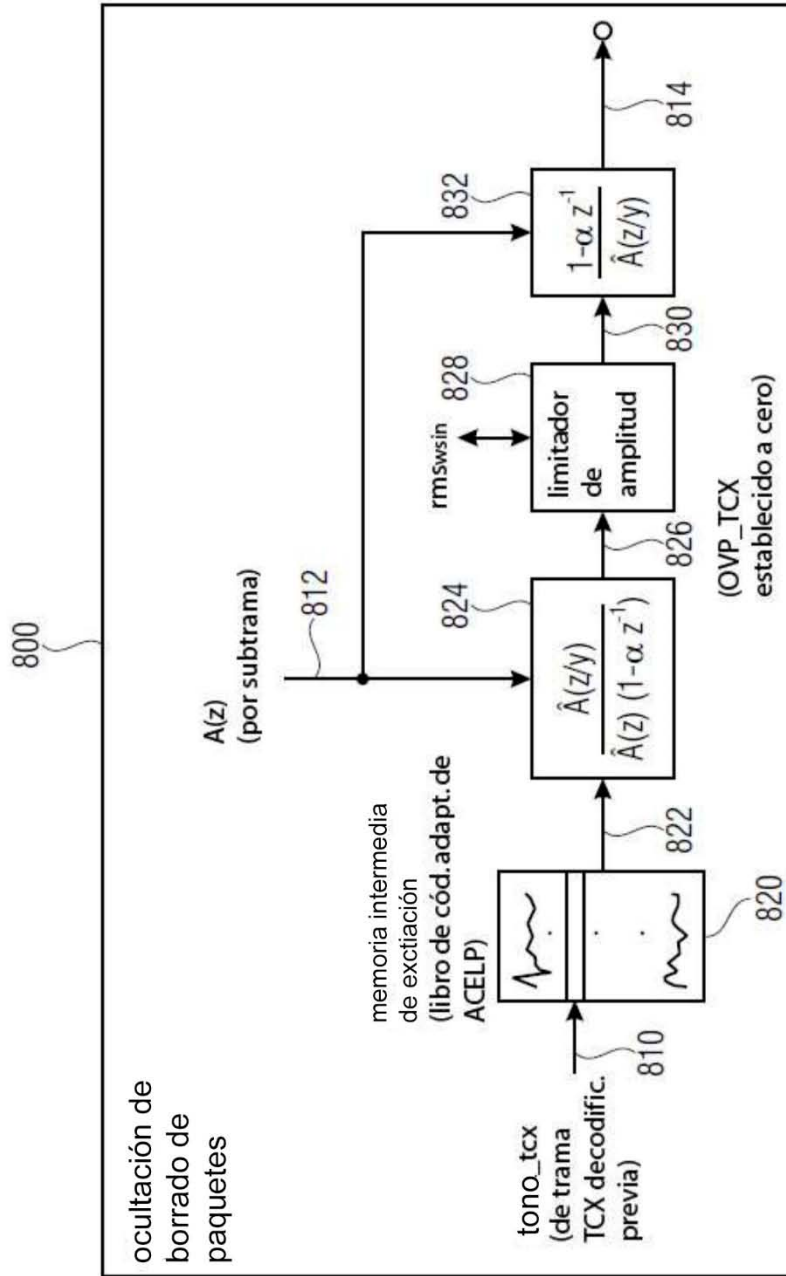


Diagrama de bloques del decodificador TCX

**FIG. 7B**

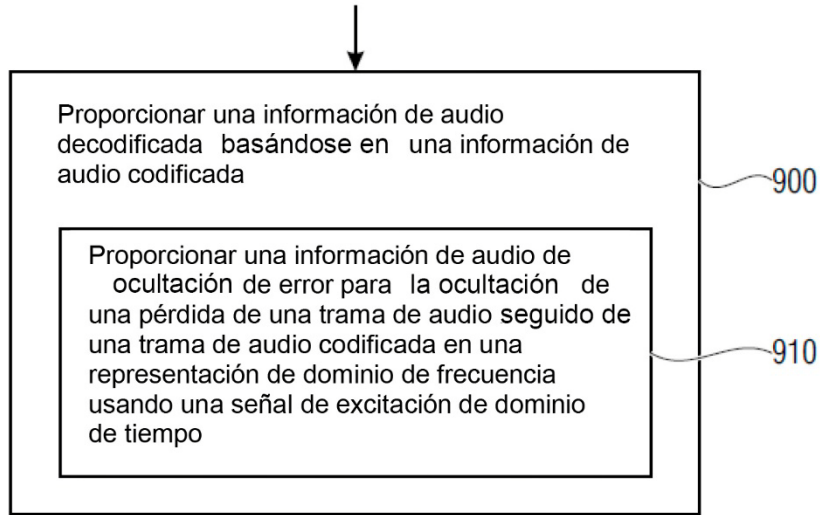
FIG 7	
FIG 7A	FIG 7B



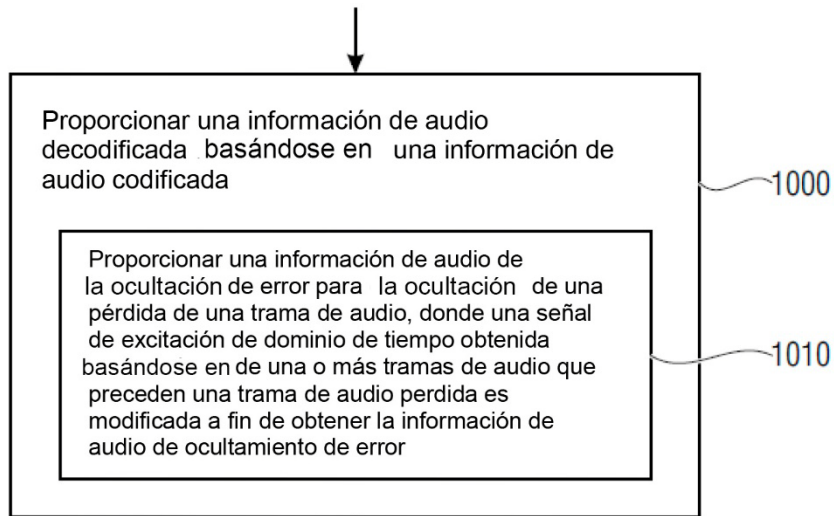
Decodificación TCX en el caso de ocultamiento de borrado de paquetes TCX-256

Diagrama de bloques del decodificador TCX

FIG. 8



**FIG. 9**



**FIG. 10**

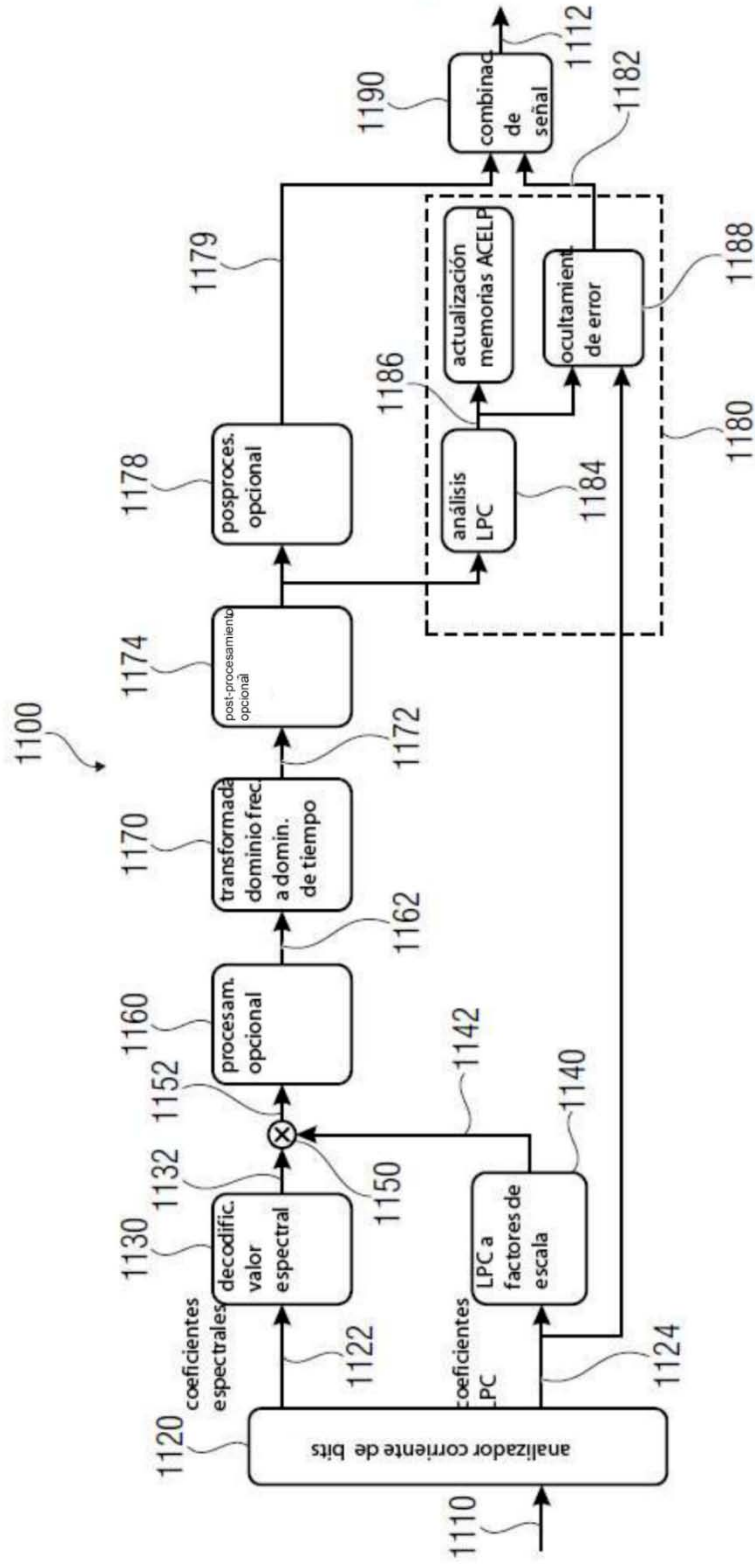


FIG. 11