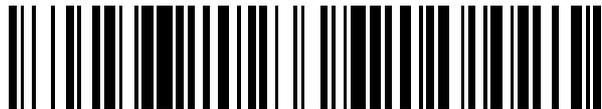


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 661 732**

51 Int. Cl.:

G10L 19/005 (2013.01)

G10L 25/90 (2013.01)

G10L 19/08 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.10.2014 PCT/EP2014/073036**

87 Fecha y número de publicación internacional: **07.05.2015 WO15063045**

96 Fecha de presentación y número de la solicitud europea: **27.10.2014 E 14789568 (4)**

97 Fecha y número de publicación de la concesión europea: **20.12.2017 EP 3063759**

54 Título: **Decodificador de audio y método para proporcionar una información de audio decodificada usando una ocultación 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:
03.04.2018

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 2 661 732 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).

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

DESCRIPCIÓN

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

15 Algunas realizaciones de acuerdo con la invención crean programas informáticos para la realización de uno de dichos métodos.

Algunas realizaciones de acuerdo con la invención se refieren a una ocultación de dominio de tiempo para un códec [codificador-decodificador] 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 se transmiten sobre canales no confiables, lo que acarrea el riesgo de que se pierdan las unidades de datos (por ejemplo, paquetes) que comprenden una o más tramas de audio (por ejemplo, en la forma de una representación codificada, como, por ejemplo, una representación de dominio de frecuencia codificada o una representación de dominio de tiempo codificada). 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 intermedio de tramas de audio. En otros casos, es casi imposible requerir una repetición de tramas de audio perdidas.

Para obtener una buena, o al menos aceptable, calidad de audio dado el caso que las tramas de audio se pierdan sin proporcionar extenso almacenamiento intermedio (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 al menos, una calidad de audio aceptable, incluso, en el caso de que las tramas de audio se pierdan.

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

A continuación, se describirá un concepto de codificación de audio convencional.

45 En la norma 3gpp TS26.290, se explica una decodificación por excitación codificada de transformada (decodificación TCX) con ocultación de error. A continuación, se proporcionarán algunas explicaciones, que se basan en la sección de "Síntesis de señal y decodificación de modo TCX" en la referencia [1].

50 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 Fig. 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.

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

Caso 1 (Fig. 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 Fig. 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 de acuerdo con la Fig. 7 recibe parámetros específicos de TCX 710 y proporciona, basándose en el mismo, la información de audio decodificada 712, 714.

5 El decodificador de audio 700 comprende un demultiplexor "DEMUX TCX 720", que está configurado para recibir los parámetros específicos de TCX 710 y la información "BFI_TCX". El demultiplexor 720 separa los parámetros específicos de TCX 710, y proporciona una información de excitación codificada 722, una información de relleno de ruido codificada 724, y una información de ganancia global codificada 726. El decodificador de audio 700 comprende un decodificador de excitación 730, que está configurado para recibir la información de excitación codificada 722, la información de relleno de ruido codificada 724 y la información de ganancia global codificada 726, 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 de excitación 730 proporciona, basándose en el mismo, una señal de excitación de dominio de tiempo 728 (también designada con "x"). El decodificador de excitación 730 comprende un procesador de información de excitación 732, que demultiplexa la información de excitación codificada 722 y decodifica los parámetros de cuantificación de vector algebraico. El procesador de información de excitación 732 proporciona una señal de excitación intermedia 734, que, típicamente, se encuentra en una representación de dominio de frecuencia, y que se designa con Y. El codificador de excitación 730 además comprende un inyector de ruido 736, que está configurado para inyectar ruido en subbandas no cuantificadas, para derivar una señal de excitación rellena con ruido 738 de la señal de excitación intermedia 734. La señal de excitación rellena con ruido 738 típicamente se encuentra en el dominio de frecuencia, y se designa con Z. El inyector de ruido 736 recibe una información de intensidad de ruido 742 desde un decodificador de nivel de relleno de ruido 740. El decodificador de excitación además comprende un desénfasis de baja frecuencia adaptativo 744, que está configurado para realizar una operación de desénfasis de baja frecuencia basándose en la señal de excitación rellena con ruido 738, para obtener de esta manera una señal de excitación procesada 746, que se encuentra aún en el dominio de frecuencia, y que se designa con X'. El decodificador de excitación 730 además comprende un transformador de dominio de frecuencia a dominio de tiempo 748, que está configurado para recibir la señal de excitación procesada 746 y para proporcionar, basándose en la misma, una señal de excitación de dominio de tiempo 750, 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 de excitación procesada 746). El decodificador de excitación 730 además comprende un escalador 752, que está configurado para escalar la señal de excitación de dominio de tiempo 750 para obtener una señal de excitación de dominio de tiempo escalada 754. El escalador 752 recibe una información de ganancia global 756 desde un decodificador de ganancia global 758, en el que, en respuesta, el decodificador de ganancia global 758 recibe la información de ganancia global codificada 726. El decodificador de excitación 730 además comprende una síntesis de superposición y adición 760, que recibe las señales de excitación de dominio de tiempo escaladas 754 asociadas con una pluralidad de porciones de tiempo. La síntesis de superposición y adición 760 realiza una operación de superposición y adición (que puede incluir una operación de generación de ventanas) basándose en las señales de excitación de dominio de tiempo escaladas 754, para obtener una señal de excitación de dominio de tiempo temporalmente combinada 728 durante un periodo más largo en el tiempo (más largo que los periodos en el tiempo para los cuales se proporcionan las señales de excitación de dominio de tiempo individuales 750, 754).

El decodificador de audio 700 además comprende una síntesis de LPC 770, que recibe la señal de excitación de dominio de tiempo 728 proporcionada por la síntesis de superposición y adición 760 y uno o más coeficientes de LPC que definen una función de filtro de síntesis de LPC 772. La síntesis de LPC 770, por ejemplo, puede comprender un primer filtro 774, que, por ejemplo, puede filtrar por síntesis la señal de excitación de dominio de tiempo 728, para obtener de esta manera la señal de audio decodificada 712. Opcionalmente, la síntesis de LPC 770 puede comprender además un segundo filtro de síntesis 772 que está configurado para filtrar por síntesis la señal de salida del primer filtro 774 usando otra función de filtro de síntesis, para obtener de esta manera la señal de audio decodificada 714.

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

La ocultación de borrado de paquetes 800 recibe una información de tono 810, que se designa además con "tono_tcx", y que se obtiene a partir de una trama TCX decodificada previa. Por ejemplo, la información de tono 810 puede obtenerse usando un estimador de tono dominante 747 desde la señal de excitación procesada 746 en el decodificador de excitación 730 (durante la decodificación "normal"). Además, la ocultación de borrado de paquetes 800 recibe parámetros de LPC 812, que pueden representar una función de filtro de síntesis de LPC. Los parámetros de LPC 812, por ejemplo, pueden ser idénticos a los parámetros de LPC 772. En consecuencia, la ocultación de borrado de paquetes 800 puede configurarse para proporcionar, basándose en la información de tono 810 y los parámetros de LPC 812, una señal de ocultación de error 814, que puede considerarse una información de audio de ocultación de error. La ocultación de borrado de paquetes 800 comprende una memoria intermedia de excitación 820, que, por ejemplo, puede almacenar en memoria intermedia una excitación previa. La memoria intermedia de excitación 820, por ejemplo, puede hacer uso del libro de códigos adaptativo ACELP, y puede

proporcionar una señal de excitación 822. La ocultación de borrado de paquetes 800 puede comprender adicionalmente un primer filtro 824, una función de filtro que puede definirse como se muestra en la Fig. 8. Por lo tanto, el primer filtro 824 puede filtrar la señal de excitación 822 basándose en los parámetros de LPC 812, para obtener una versión filtrada 826 de la señal de excitación 822. La ocultación de borrado de paquetes además
 5 comprende un limitador de amplitud 828, que puede limitar una amplitud de la señal de excitación filtrada 826 basándose en información objetivo o información de nivel rms_{wsyn} . Además, la ocultación de borrado de paquetes 800 puede comprender un segundo filtro 832, que puede estar configurado para recibir la señal de excitación filtrada limitada de amplitud 830 desde el limitador de amplitud 822 y para proporcionar, basándose en la misma, la señal de
 10 ocultación de error 814. Una función de filtro del segundo filtro 832, por ejemplo, puede definirse como se muestra en la Fig. 8.

A continuación, se describirán algunos detalles con respecto a la decodificación y a la ocultación de error.

En el Caso 1 (ocultación de borrado de paquetes en TCX-256), no hay información disponible para la decodificación
 15 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 $1/\hat{A}(z)$. Se usa un filtro no lineal en lugar de $1/\hat{A}(z)$ para evitar chasquidos en la síntesis. Este filtro se descompone en 3 etapas.

20 **Etapas 1:** filtración mediante:

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

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

25 **Etapas 2:** la aplicación de un limitador (la magnitud se limita a $\pm rms_{wsyn}$)

Etapas 3: la filtración mediante:

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

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

30 **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 \hat{B}'_k del espectro escalado X' , donde X' es como se describe en la Etapa 2 de la Sección
 35 5.3.5.7 de 3gpp TS 26.290. Recuérdesse 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 Etapa 5 de la Sección 5.3.5.7. Para cada bloque B'_k , se envían tres grupos de índices binarios por el
 40 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;

45 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 \hat{B}'_k (un punto de rejilla) no se estaba en el libro de códigos base, los 8 índices del
 50 vector de índice de extensión de Voronoi k calculado en la subetapa V1 de la Etapa 5 en la Sección; a partir de índices de extensión de Voronoi, puede calcularse 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 se proporciona 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 se proporciona por $M = 2^r$.

55 A continuación, a partir del factor de escala M , el Vector de extensión de Voronoi z (un punto de rejilla en RE_β) y el punto de rejilla c en el libro de códigos base (también, un punto de rejilla en RE_β), cada bloque escalado cuantificado \hat{B}'_k puede calcularse 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 \hat{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 tono 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 está basada en la asunción de que el pico de máxima magnitud en el espectro del objetivo TCX corresponde al tono dominante.
15 La búsqueda de la máxima M se restringe 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 se halla. A continuación, el tono 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. Para evitar problemas de almacenamiento intermedio (la memoria intermedia de excitación se limita a 256 muestras), si $T_{\text{est}} > 256$ muestras, *tono_tcx* se establece en 256; de lo contrario, si $T_{\text{est}} \leq 256$, se evita periodo de tono múltiple en 256 muestras estableciendo *tono_tcx* en

$$25 \quad \text{tono_tcx} = \max \{ \lfloor n T_{\text{est}} \rfloor \mid n \text{ entero} > 0 \text{ y } n T_{\text{est}} \leq 256 \}$$

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

30 A continuación, 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 (consúltese, por ejemplo, referencia [4]), solo se describe un modo de interpolación. De acuerdo con la referencia [4], el decodificador núcleo AAC incluye una función de ocultación 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 que pueden lograr el mejoramiento adicional para la voz decodificada en una trama en la cual se detecta un error. De acuerdo con 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 de acuerdo con la información de modo. Además, el decodificador de voz controla de manera adaptativa la relación de ganancia de excitación adaptativa y la ganancia de excitación fijada de acuerdo con la información de modo. Además, el concepto de acuerdo con 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 de acuerdo con 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 se detectan con un error.

También es conocido de acuerdo con la solicitud de patente internacional WO2005/078706 una técnica que usa una excitación de dominio de tiempo previamente almacenada obtenida a partir de una representación de dominio de frecuencia de TCX para ocultar una trama posteriormente perdida. En vista de la técnica anterior, existe una
55 necesidad de una mejora adicional de la ocultación de error, que proporciona una mejor impresión auditiva.

3. Sumario de la invención.

Una realización de acuerdo con la invención define un decodificador de audio de acuerdo con la reivindicación 1,
60 para proporcionar información de audio decodificada basándose en una información de audio codificada.

Esta realización de acuerdo con la invención está basada en el hallazgo de que puede obtenerse una ocultación de error mejorada al proporcionar la información de audio de ocultación de error basándose en 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 una ocultación de error típicamente es mejor si la ocultación de error se realiza basándose en una señal de excitación de dominio de tiempo, cuando se compara con una ocultación de error realizada en un dominio de frecuencia, de modo tal que vale la pena la conmutación a una ocultación de error de dominio de tiempo, usando una señal de excitación de dominio de tiempo, incluso 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 una buena ocultación de error, incluso 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).

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 LPC. Esto podría realizarse usando FDNS (Conformación 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 LPC) incluso si la trama de audio que precede la trama de audio perdida está 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 LPC, y luego, en el decodificador, transformamos la LPC en una representación de factor de escala para los segmentos de la MDCT. En otras palabras, en el caso de TCX, enviamos el coeficiente de LPC, y luego, en el decodificador, transformamos dichos coeficientes de LPC en una representación de factor de escala para TCX en USAC o en AMR-WB+ donde no hay ningún factor de escala en absoluto.

En una realización preferida, el decodificador de audio comprende un núcleo de decodificador de dominio de frecuencia configurado para la aplicación de una escala basándose en factores de escala, a una pluralidad de valores espectrales derivados de la representación de dominio de frecuencia. En este caso, la ocultación de error está configurada para proporcionar la información de audio de ocultación de error para la ocultación de una pérdida de una trama de audio seguido de 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 de acuerdo con la invención está basada en el hallazgo de que la derivación de la señal de excitación de dominio de tiempo a partir de la anteriormente mencionada representación de dominio de frecuencia típicamente proporciona un mejor resultado de ocultación de error en comparación con una ocultación de error realizada directamente en el dominio de frecuencia. Por ejemplo, la señal de excitación se crea basándose en la síntesis de la trama previa; entonces, no importa realmente si la trama previa es una trama de dominio de frecuencia (MDCT, FFT...) o una trama de dominio de tiempo. Sin embargo, pueden observarse ventajas particulares si la trama previa era un dominio de frecuencia. Además, 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 transmitirse como coeficientes de LPC, por ejemplo, usando una representación polinomial que a continuación se convierte en factores de escala en el lado del decodificador.

En un ejemplo, 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 una ocultación de error es ventajoso incluso si la trama de audio que precede a 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 está basada en una síntesis de LPC).

En una realización preferida, la ocultación de error está configurada para obtener la señal de excitación de dominio de tiempo basándose en la trama de audio codificada en la representación de dominio de frecuencia que precede una trama de audio perdida. En este caso, la ocultación de error está configurada para proporcionar la información de audio de ocultación de error para la ocultación 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 la ocultación de error, debería derivarse 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 la ocultación de error pueda realizarse con moderado esfuerzo y buena exactitud.

En un ejemplo, la ocultación de error está configurada 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 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 vale la pena el esfuerzo de realizar un análisis de LPC, para derivar los parámetros de codificación de predicción lineal y la señal de excitación de dominio de tiempo, incluso 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 ocultación de error de buena calidad para muchas señales de audio de entrada basándose en 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, como alternativa, 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 basándose en un conjunto de factores de escala usando una transformada. En otras palabras, los parámetros de LPC pueden obtenerse usando la estimación de parámetros de LPC. Esto podría realizarse o bien mediante la generación de ventanas/autocorr/levinson durbin basándose en 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 LPC.

En una realización preferida, la ocultación de error está configurada para obtener una información de tono (o retardo) que describe un tono 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 ocultación de error de acuerdo con la información de tono. Al considerar la información de tono, puede lograrse que la información de audio de ocultación de error (que típicamente es una señal de audio de ocultación de error que cubre la duración temporal de al menos una trama de audio perdida) se adapte bien al contenido de audio real.

En una realización preferida, la ocultación de error está configurada para obtener la información de tono basándose en 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 tono a partir de la señal de excitación de dominio de tiempo acarrea una alta exactitud. Además, se ha hallado que es ventajoso si la información de tono se adapta bien a la señal de excitación de dominio de tiempo, ya que la información de tono se usa para una modificación de la señal de excitación de dominio de tiempo. Al derivar la información de tono de la señal de excitación de dominio de tiempo, puede lograrse dicha relación cercana.

En una realización preferida, la ocultación de error está configurada para evaluar una correlación cruzada de la señal de excitación de dominio de tiempo, para determinar una información de tono aproximada. Además, la ocultación de error puede estar configurada para perfeccionar la información de tono aproximada usando una búsqueda de bucle cerrado alrededor de un tono determinado por la información de tono aproximada. En consecuencia, puede lograrse una información de tono altamente exacta con moderado esfuerzo computacional.

En una realización preferida, la ocultación de error del decodificador de audio puede estar configurada para obtener una información de tono basándose en una información secundaria de la información de audio codificada.

En una realización preferida, la ocultación de error puede estar configurada para obtener una información de tono basándose en una información de tono disponible para una trama de audio previamente decodificada.

En una realización preferida, la ocultación de error está configurada para obtener una información de tono basándose en una búsqueda de tono realizada sobre una señal de dominio de tiempo o sobre una señal residual.

En otras palabras, el tono puede transmitirse como información secundaria o podría además provenir de la trama previa si hay, por ejemplo, LTP. La información de tono además podría transmitirse en la corriente de bits si está disponible en el codificador. Se podría opcionalmente hacer la búsqueda de tono sobre la señal de dominio de tiempo directamente, o sobre la residual, que habitualmente proporciona mejores resultados sobre la residual (señal de excitación de dominio de tiempo).

En una realización preferida, la ocultación de error está configurada para copiar un ciclo de tono 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, para obtener una señal de excitación

para una síntesis de la señal de audio de ocultación 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 ocultación 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.

En una realización preferida, la ocultación de error está configurada para filtrar en paso bajo el ciclo de tono 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 adaptarse para un ancho de banda de audio disponible, lo que produce una buena impresión auditiva de la información de audio de ocultación 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 tono. 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, la ocultación de error está configurada para predecir un tono en un extremo de una trama perdida para adaptar la señal de excitación de dominio de tiempo, o una o más de sus copias, al tono predicho. En consecuencia, los cambios de tono esperados durante la trama de audio perdida pueden considerarse. En consecuencia, se evitan artefactos en una transición entre la información de audio de ocultación de error y una información de audio de una trama apropiadamente decodificada seguida de una o más tramas de audio perdidas (o al menos se reducen, ya que es solo una trama predicha, no la real). Por ejemplo, la adaptación va desde el último buen tono hasta el predicho. Esto se realiza por medio de la resincronización de pulsos [7].

En una realización preferida, la ocultación de error está configurada para combinar una señal de excitación de dominio de tiempo extrapolada y una señal de ruido, para obtener una señal de entrada para una síntesis de LPC. En este caso, la ocultación de error está configurada para realizar la síntesis de LPC, en el que la síntesis de LPC está configurada para filtrar la señal de entrada de la síntesis de LPC de acuerdo con parámetros de codificación de predicción lineal, para obtener la información de audio de ocultación de error. En consecuencia, puede considerarse tanto un componente determinista (por ejemplo, aproximadamente periódico) del contenido de audio como un componente de tipo ruido del contenido de audio. Por lo tanto, se logra que la información de audio de ocultación de error comprenda una impresión auditiva "natural".

En un ejemplo, la ocultación de error está configurada para calcular 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 LPC, usando una correlación en el dominio de tiempo que se realiza basándose en una representación de dominio de tiempo de la trama de audio codificada en el dominio de frecuencia que precede la trama de audio perdida, en el que se establece un retardo de correlación dependiente de una información de tono obtenido basándose en 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 ocultación de error. Sin embargo, se ha hallado que el cálculo anteriormente mencionado 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. Como alternativa, puede obtenerse una correlación en el dominio de excitación o directamente en el dominio de tiempo para obtener la información de tono. Sin embargo, hay también diferentes posibilidades, dependiendo de la realización utilizada. En una realización, la información de tono podría ser solo el tono obtenido desde la ltp de última trama, o el tono que se transmite como información secundaria o la calculada.

En un ejemplo, la ocultación de error está configurada 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 paso alto de la señal de ruido (que típicamente se introduce en la síntesis de LPC) logra una impresión auditiva natural. Por ejemplo, la característica de paso alto puede cambiar con la cantidad de trama perdida, seguido de 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 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 seguido de una cierta cantidad de pérdida de trama, ya no hay filtración, para solo obtener el ruido de forma de banda completa para obtener un buen ruido de confort cercano al ruido de fondo.

En un ejemplo, la ocultación de error está configurada para cambiar selectivamente la forma espectral de la señal de ruido (562) 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 ocultación de error puede mejorarse por dicho concepto. Por ejemplo, en algún caso, es mejor disminuir las ganancias y la forma, y en algún lugar, es mejor incrementarlas.

5 En un ejemplo, la ocultación de error está configurada para calcular una ganancia de la señal de ruido dependiendo de una correlación en el dominio de tiempo, que se realiza basándose en una representación de dominio 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
10 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 generarse midiendo la energía del resultado: excitación de señal de entrada - excitación basándose en tono generado.

15 En una realización preferida, la ocultación de error está configurada para modificar una señal de excitación de dominio de tiempo obtenida basándose en una o más tramas de audio que preceden una trama de audio perdida, para obtener la información de audio de ocultación de error. Se ha hallado que la modificación de la señal de 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
20 "desvanecimiento de salida" del componente determinista (por ejemplo, sustancialmente periódico) del contenido de audio en la información de audio de ocultación de error. Además, 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 tono (estimada o esperada). Esto permite el ajuste de las características de la información de audio de ocultación de error en función del tiempo.

25 En una realización preferida, la ocultación de error está configurada para usar una o más copias modificadas de la señal de excitación de dominio de tiempo obtenida basándose en una o más tramas de audio que preceden una trama de audio perdida, para obtener la información de ocultación de error. Las copias modificadas de la señal de excitación de dominio de tiempo pueden obtenerse con un esfuerzo moderado, y la modificación puede realizarse
30 usando un algoritmo sencillo. En consecuencia, las características deseadas de la información de audio de ocultación de error pueden lograrse con esfuerzo moderado.

En una realización preferida, la ocultación de error está configurada para modificar la señal de excitación de dominio de tiempo obtenida basándose en una o más tramas de audio que preceden una trama de audio perdida, o una o
35 más de sus copias, para reducir de esta manera un componente periódico de la información de audio de ocultación 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 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 ocultación de error.

40 En una realización preferida, la ocultación de error está configurada para escalar la señal de excitación de dominio de tiempo obtenida basándose en una o más tramas de audio que preceden la trama de audio perdida, o una o más de sus copias, para modificar la señal de excitación de dominio de tiempo. Se ha hallado que la operación de escala puede realizarse con poco esfuerzo, en el que la señal de excitación de dominio de tiempo escalada típicamente
45 proporciona una buena información de audio de ocultación de error.

En una realización preferida, la ocultación de error está configurada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida basándose en 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
50 desvanecimiento de salida del componente periódico dentro de la información de audio de ocultación de error.

En una realización preferida, la ocultación de error está configurada para ajustar una velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida basándose en una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, dependiendo
55 de uno o más parámetros de una o más tramas de audio que preceden la trama de audio perdida, y/o dependiendo de una cantidad de tramas de audio perdidas consecutivas. Por consiguiente, es posible ajustar la velocidad a la cual el componente determinista (por ejemplo, al menos aproximadamente periódico) se desvanece de salida en la información de audio de ocultación de error. La velocidad del desvanecimiento de salida puede adaptarse 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. Como alternativa, o además, la
60 cantidad de tramas de audio perdidas consecutivas puede considerarse cuando se determina la velocidad utilizada para el desvanecimiento de salida del componente determinista (por ejemplo, al menos aproximadamente periódico) de la información de audio de ocultación de error, que ayuda a adaptar la ocultación de error a la situación específica. Por ejemplo, la ganancia de la parte tonal y la ganancia de la parte con ruido pueden desvanecerse de

salida de manera separada. La ganancia para la parte tonal puede converger a cero seguido de una cierta cantidad de pérdida de trama, mientras que la ganancia de ruido puede converger a la ganancia determinada para alcanzar un cierto ruido de confort.

5 En una realización preferida, la ocultación de error está configurada para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida basándose en una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, dependiendo de una longitud de un periodo de tono de la señal de excitación de dominio de tiempo, de modo que una señal de excitación de dominio de tiempo introducida en una síntesis de LPC se desvanece de salida de manera más rápida para señales que tienen una longitud más corta del periodo de tono en comparación con señales que tienen una mayor longitud del periodo de tono. Por consiguiente, puede evitarse que las señales que tienen una longitud más corta del periodo de tono se repitan con demasiada frecuencia con alta intensidad, ya que esto típicamente dará como resultado una impresión auditiva no natural. En consecuencia, puede mejorarse una calidad general de la información de audio de ocultación de error.

15 En una realización preferida, la ocultación de error está configurada para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida basándose en una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, dependiendo de un resultado de un análisis de tono o una predicción de tono, de modo tal que un componente determinista de la señal de excitación de dominio de tiempo introducido en una síntesis de LPC se desvanece de salida más rápidamente para señales que tienen un mayor cambio de tono por unidad de tiempo en comparación con señales que tienen un cambio de tono menor por unidad de tiempo, y/o de modo tal que un componente determinista de la entrada de señal de excitación de dominio de tiempo en una síntesis de LPC se desvanece de salida más rápidamente para señales para las cuales una predicción de tono falla en comparación con señales para las cuales la predicción de tono tiene éxito. Por consiguiente, el desvanecimiento de salida puede hacerse más rápidamente para señales en las cuales hay una gran incertidumbre del tono en comparación con señales para las cuales hay una menor incertidumbre del tono. Sin embargo, mediante el desvanecimiento de salida de un componente determinista más rápidamente para señales que comprenden una incertidumbre comparativamente grande del tono, pueden evitarse artefactos audibles, o al menos pueden reducirse sustancialmente.

30 En una realización preferida, la ocultación de error está configurada para la escala de tiempo de la señal de excitación de dominio de tiempo obtenida basándose en una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, dependiendo de una predicción de un tono para el tiempo de una o más tramas de audio perdidas. Por consiguiente, la señal de excitación de dominio de tiempo puede adaptarse a un tono variable, de modo tal que la información de audio de ocultación de error comprende una impresión auditiva más natural.

35 En un ejemplo, la ocultación de error está configurada para proporcionar la información de audio de ocultación 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 basándose en la información de audio de ocultación de error, que ayuda a reducir los artefactos de bloqueo.

40 En un ejemplo, la ocultación de error está configurada para realizar una superposición y adición de la información de audio de ocultación de error y de una representación de dominio de tiempo de una o más tramas de audio apropiadamente recibidas seguido de una o más tramas de audio perdidas. En consecuencia, es posible evitar (o al menos reducir) los artefactos de bloqueo.

45 En un ejemplo, la ocultación de error está configurada para derivar la información de audio de ocultación de error basándose en al 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 ocultación de error puede obtenerse con buena exactitud, incluso para modos de codificación en los cuales más de dos tramas (o ventanas) se superponen (donde dicha superposición puede ayudar a reducir una demora).

50 Otra realización de acuerdo con la invención define un método de acuerdo con la reivindicación 29. Otra realización más de acuerdo con la invención define un programa informático de acuerdo con la reivindicación 30, para realizar dicho método cuando el programa informático se ejecuta en un ordenador.

Breve descripción de las figuras

60 Las realizaciones de la presente invención se describirán a continuación con referencia a las figuras adjuntas, en las cuales:

Fig. 1 muestra un diagrama de bloques esquemático de un decodificador de audio, de acuerdo con un ejemplo;

- Fig. 2 muestra un diagrama de bloques esquemático de un decodificador de audio, de acuerdo con otro ejemplo;
- Fig. 3 muestra un diagrama de bloques esquemático de un decodificador de audio, de acuerdo con una realización de la presente invención;
- 5 Fig. 4 muestra un diagrama de bloques esquemático de un decodificador de audio, de acuerdo con otra realización de la presente invención;
- Fig. 5 muestra un diagrama de bloques esquemático de una ocultación de dominio de tiempo para un codificador de transformada;
- Fig. 6 muestra un diagrama de bloques esquemático de una ocultación de dominio de tiempo para un códec de conmutación;
- 10 Fig. 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;
- 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;
- 15 Fig. 9 muestra un diagrama de flujo de un método para proporcionar una información de audio decodificada basándose en una información de audio codificada, de acuerdo con un ejemplo; y
- Fig. 10 muestra un diagrama de flujo de un método para proporcionar una información de audio decodificada basándose en una información de audio codificada, de acuerdo con otro ejemplo;
- Fig. 11 muestra un diagrama de bloques esquemático de un decodificador de audio, de acuerdo con otro ejemplo.
- 20

Descripción detallada de las realizaciones

1. Decodificador de audio de acuerdo con la Fig. 1

25 La Fig. 1 muestra un diagrama de bloques esquemático de un decodificador de audio 100, de acuerdo con un ejemplo. El decodificador de audio 100 recibe una información de audio codificada 110, 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 recibirse por medio de un canal no confiable, de modo tal que se produce una

30 pérdida de trama de vez en cuando. El decodificador de audio 100 además proporciona, basándose en la información de audio codificada 110, la información de audio decodificada 112.

El decodificador de audio 100 puede comprender una decodificación/procesamiento 120, que proporciona la información de audio decodificada basándose en la información de audio codificada en ausencia de una pérdida de

35 trama.

El decodificador de audio 100 adicionalmente comprende una ocultación de error 130, que proporciona una información de audio de ocultación de error. La ocultación de error 130 está configurada para proporcionar la información de audio de ocultación de error 132 para la ocultación de una pérdida de una trama de audio seguido de

40 una trama de audio codificada en la representación de dominio de frecuencia, usando una señal de excitación de dominio de tiempo.

En otras palabras, la decodificación/procesamiento 120 puede proporcionar una información de audio decodificada 122 para tramas de audio que se codifican en forma de una representación de dominio de frecuencia, es decir, en forma de una representación codificada, cuyos valores codificados describen intensidades en diferentes segmentos de frecuencia. En otras palabras, la decodificación/procesamiento 120, por ejemplo, puede comprender un

45 decodificador de audio de dominio de frecuencia, que deriva un conjunto de valores espectrales de la información de audio codificada 110 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 de audio decodificada 122, o que forma la base para proporcionar la información de audio decodificada 122 en el caso de que haya procesamiento

50 posterior adicional.

Sin embargo, la ocultación de error 130 no realiza la ocultación 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

55 síntesis, por ejemplo, un filtro de síntesis de LPC, que proporciona una representación de dominio de tiempo de una señal de audio (por ejemplo, la información de audio de ocultación de error) basándose en la señal de excitación de dominio de tiempo, y además, basándose en coeficientes de filtro de LPC (coeficientes de filtro de codificación de predicción lineal).

60 Por consiguiente, la ocultación de error 130 proporciona la información de audio de ocultación de error 132, que, por ejemplo, puede ser una señal de audio de dominio de tiempo, para tramas de audio perdidas, en el que la señal de excitación de dominio de tiempo utilizada por la ocultación de error 130 puede basarse en una o más tramas de audio previas apropiadamente recibidas (que preceden la trama de audio perdida), que se codifican en forma de una representación de dominio de frecuencia, o puede derivar de ellas. Como conclusión, el decodificador de audio 100

puede realizar una ocultación de error (es decir, proporcionar una información de audio de ocultación de error 132), que reduce una degradación de una calidad de audio debido a la pérdida de una trama de audio basándose en una información de audio codificada, en la que al menos algunas tramas de audio se codifican en una representación de dominio de frecuencia. Se ha hallado que la realización de la ocultación de error usando una señal de excitación de dominio de tiempo, incluso si se perdiera una trama seguida de una trama de audio codificada en la representación de dominio de frecuencia apropiadamente recibida, acarrea una mejorada calidad de audio en comparación con una ocultación 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 suave 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 ocultación 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 basándose en la señal de excitación de dominio de tiempo, ayuda a evitar discontinuidades. Por lo tanto, puede lograrse una buena (o al menos aceptable) impresión auditiva, usando el decodificador de audio 100, incluso 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 de la ocultación de códec de voz. El uso de la LPC ayuda a evitar las discontinuidades, y proporciona una mejor conformación de las tramas.

Además, debe observarse que el decodificador de audio 100 puede complementarse por cualquiera de los rasgos y las funcionalidades que se describen a continuación, ya sea individualmente o tomados en combinación.

2. Decodificador de audio de acuerdo con la Fig. 2.

La Fig. 2 muestra un diagrama de bloques esquemático de un decodificador de audio 200 de acuerdo con otro ejemplo. El decodificador de audio 200 está configurado para recibir una información de audio codificada 210 y para proporcionar, basándose en la misma, una información de audio decodificada 220. La información de audio codificada 210, 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 de audio codificada 210 pueden estar codificadas en una representación de dominio de frecuencia, o todas las tramas de la información de audio codificada 210 pueden estar codificadas en una representación de dominio de tiempo (por ejemplo, en forma de una señal de excitación de dominio de tiempo codificada y parámetros de síntesis de señal codificados, por ejemplo, parámetros de LPC). Como alternativa, 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 de audio 200 es un decodificador de audio de conmutación que puede conmutar entre diferentes modos de decodificación. La información de audio decodificada 220, por ejemplo, puede ser una representación de dominio de tiempo de uno o más canales de audio.

El decodificador de audio 200 puede comprender habitualmente una decodificación/procesamiento 230, que, por ejemplo, puede proporcionar una información de audio decodificada 232 para tramas de audio que se reciben apropiadamente. En otras palabras, la decodificación/procesamiento 230 puede realizar una decodificación de dominio de frecuencia (por ejemplo, una decodificación de tipo AAC, o similar) basándose en una o más tramas de audio codificadas, codificadas en una representación de dominio de frecuencia. Como alternativa, o además, la decodificación/procesamiento 230 puede estar configurado para realizar una decodificación en el dominio de tiempo (o la decodificación en el dominio de predicción lineal) basándose en 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 de audio 200 adicionalmente comprende una ocultación de error 240, que está configurada para proporcionar una información de audio de ocultación de error 242 para una o más tramas de audio perdidas. La ocultación de error 240 está configurada para proporcionar la información de audio de ocultación de error 242 para la ocultación de una pérdida de una trama de audio (o incluso, una pérdida de múltiples tramas de audio). La ocultación de error 240 está configurada para modificar una señal de excitación de dominio de tiempo obtenida basándose en una o más tramas de audio que preceden una trama de audio perdida, para obtener la información de audio de ocultación de error 242. En otras palabras, la ocultación de error 240 puede obtener (o derivar) una señal de excitación de dominio de tiempo para (o basándose en) 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 se obtiene para (o basándose en) una o más tramas de audio apropiadamente recibidas que preceden una trama de audio perdida,

para obtener de esta manera (por medio de la modificación) una señal de excitación de dominio de tiempo que se usa para proporcionar la información de audio de ocultación de error 242. 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 LPC) de la información de audio de ocultación de error asociada con la trama de audio perdida (o incluso, con múltiples tramas de audio perdidas). Al proporcionar la información de audio de ocultación de error 242 basándose en la señal de excitación de dominio de tiempo obtenida basándose en 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 ocultación de error basándose en 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 tono), y además es posible evitar una impresión auditiva no natural (por ejemplo, mediante el “desvanecimiento de salida” de un componente de señal determinista (por ejemplo, al menos aproximadamente periódico)). Por lo tanto, puede lograrse que la información de audio de ocultación de error 242 comprenda cierta similitud con la información de audio decodificada 232 obtenida basándose en tramas de audio apropiadamente decodificadas que preceden la trama de audio perdida, y puede lograrse aún que la información de audio de ocultación de error 242 comprenda un contenido de audio algo diferente cuando se compara con la información de audio decodificada 232 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 proporcionar la información de audio de ocultación 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), en el que, preferentemente, debe permanecer un cierto grado de relación entre la señal de excitación de dominio de tiempo obtenida (como una información de entrada) por la ocultación de error y la señal de excitación de dominio de tiempo modificada.

Como conclusión, el decodificador de audio 200 permite la provisión de la información de audio de ocultación de error 242, de modo tal que la información de audio de ocultación de error proporciona una buena impresión auditiva, incluso en el caso de que se pierda una o más tramas de audio. La ocultación de error se realiza basándose en una señal de excitación de dominio de tiempo, en la que 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 basándose en una o más tramas de audio que preceden una trama de audio perdida.

Además, debe observarse que el decodificador de audio 200 puede complementarse por cualquiera de los rasgos y las funcionalidades que se describen en el presente documento, ya sea individualmente o en combinación.

3. Decodificador de audio de acuerdo con la Fig. 3.

La Fig. 3 muestra un diagrama de bloques esquemático de un decodificador de audio 300, de acuerdo con una realización de la presente invención.

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

Como alternativa a la decodificación de factor de escala, puede usarse una conversión de LPC a factor de escala 354, por ejemplo, en el caso de que la información de audio codificada comprenda una información de 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 de audio USAC o en el decodificador de audio EVS), puede utilizarse un conjunto de coeficientes de LPC para derivar un conjunto de factores de escala en el lado del decodificador de audio. Esta funcionalidad puede lograrse por medio de la conversión de LPC a factor de escala 354.

El decodificador de audio 300 puede además comprender un escalador 360, que puede estar configurado para

- 5 aplicar el conjunto de factores escalados 352 al conjunto de valores espectrales 342, para obtener de esta manera un conjunto de valores espectrales decodificados escalados 362. Por ejemplo, una primera banda de frecuencia que comprende múltiples valores espectrales decodificados 342 puede escalarse usando un primer factor de escala, y una segunda banda de frecuencia que comprende múltiples valores espectrales decodificados 342 puede escalarse usando un segundo factor de escala. Por consiguiente, se obtiene el conjunto de valores espectrales decodificados escalados 362. El decodificador de audio 300 puede comprender adicionalmente un procesamiento opcional 366, que puede aplicar cierto procesamiento a los valores espectrales decodificados escalados 362. Por ejemplo, el procesamiento opcional 366 puede comprender un relleno de ruido o alguna otra operación.
- 10 El decodificador de audio 300 además comprende una transformada de dominio de frecuencia a dominio de tiempo 370, que está configurada para recibir los valores espectrales decodificados escalados 362, o una versión procesada 368 de los mismos, y para proporcionar una representación de dominio de tiempo 372 asociada con un conjunto de valores espectrales decodificados escalados 362. Por ejemplo, la transformada de dominio de frecuencia a dominio de tiempo 370 puede proporcionar una representación de dominio de tiempo 372, 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 MDCT (que pueden considerarse valores espectrales decodificados escalados) y proporcionar, basándose en los mismos, un bloque de muestras de dominio de tiempo, que pueden formar la representación de dominio de tiempo 372.
- 15 El decodificador de audio 300 puede comprender opcionalmente un post-procesamiento 376, que recibe la representación de dominio de tiempo 372 y modifica en cierta forma la representación de dominio de tiempo 372, para obtener una versión post-procesada 378 de la representación de dominio de tiempo 372.
- 20 El decodificador de audio 300 además comprende una ocultación de error 380 que, por ejemplo, puede recibir la representación de dominio de tiempo 372 de la transformada de dominio de frecuencia a dominio de tiempo 370 y que, por ejemplo, puede proporcionar una información de audio de ocultación de error 382 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 espectrales codificados 326 para dicha trama de audio (o subtrama de audio), la ocultación de error 380 puede proporcionar la información de audio de ocultación de error basándose en la representación de dominio de tiempo 372 asociada con una o más tramas de audio que preceden la trama de audio perdida. La información de audio de ocultación de error puede ser típicamente una representación de dominio de tiempo de un contenido de audio.
- 25 Debe observarse que la ocultación de error 380, por ejemplo, puede realizar la funcionalidad de la ocultación de error 130 que se ha descrito anteriormente. Además, la ocultación de error 380, por ejemplo, puede comprender la funcionalidad de la ocultación de error 500 descrita con referencia a la Fig. 5. Sin embargo, en términos generales, la ocultación de error 380 puede comprender cualquiera de los rasgos y de las funcionalidades que se describen con respecto a la ocultación de error en el presente documento.
- 30 Con respecto a la ocultación de error, debe observarse que la ocultación 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 ocultación que proporciona la variable que proviene de la trama buena previa. Además, actualizaremos algunas variables para ayudar para la siguiente pérdida de trama o con la recuperación para la siguiente trama buena.
- 35 El decodificador de audio 300 además comprende una combinación de señales 390, que está configurada para recibir la representación de dominio de tiempo 372 (o la representación de dominio de tiempo post-procesada 378 en el caso de que haya un post-procesamiento 376). Además, la combinación de señales 390 puede recibir la información de audio de ocultación de error 382, que habitualmente es además una representación de dominio de tiempo de una señal de audio de ocultación de error prevista para una trama de audio perdida. La combinación de señales 390, 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 de señales 390 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 de señales 390 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 ocultación de error asociada con la trama de audio perdida, para obtener de esta manera una suave transición entre la trama de audio apropiadamente recibida y la trama de audio perdida. De manera similar, la combinación de señales 390 puede estar configurada para combinar (por ejemplo, superposición y adición) la información de audio de ocultación de error asociada con la trama de audio perdida y la representación de dominio de tiempo asociada con otra trama de audio apropiadamente decodificada seguido de la trama de audio perdida (u otra información de audio de ocultación de error asociada con otra trama de audio perdida, en el caso de que se pierdan múltiples tramas de audio consecutivas).
- 40
- 45
- 50
- 55
- 60

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

Debe observarse que la funcionalidad del decodificador de audio 300 es similar a la funcionalidad del decodificador de audio 100 de acuerdo con la Fig. 1, en el que se muestran detalles adicionales en la Fig. 3. Además, debe observarse que el decodificador de audio 300 de acuerdo con la Fig. 3 puede complementarse por cualquiera de los rasgos y de las funcionalidades que se describen en el presente documento. En particular, la ocultación de error 380 puede complementarse por cualquiera de los rasgos y de las funcionalidades descritos en el presente documento con respecto a la ocultación de error.

4. Decodificador de audio 400 de acuerdo con la Fig. 4.

La Fig. 4 muestra un decodificador de audio 400 de acuerdo con otra realización de la presente invención. El decodificador de audio 400 está configurado para recibir una información de audio codificada y para proporcionar, basándose en la misma, una información de audio decodificada 412. El decodificador de audio 400, por ejemplo, puede estar configurado para recibir una información de audio codificada 410, en el que diferentes tramas de audio se codifican usando diferentes modos de codificación. Por ejemplo, el decodificador de audio 400 puede considerarse un decodificador de audio de múltiples modos o un decodificador de audio "de conmutación". Por ejemplo, algunas de las tramas de audio pueden codificarse usando una representación de dominio de frecuencia, en el que la información de audio codificada comprende una representación codificada de valores espectrales (por ejemplo, valores FFT o valores MDCT) y factores de escala que representan una escala de diferentes bandas de frecuencia. Además, la información de audio codificada 410 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 LPC"), por ejemplo, puede comprender una representación codificada de una señal de excitación, y una representación codificada de parámetros de LPC (parámetros de codificación de predicción lineal), en el que 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 basándose en la señal de excitación de dominio de tiempo.

A continuación, se describirán algunos detalles del decodificador de audio 400.

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

El decodificador de audio 400 comprende una ruta de codificación de dominio de frecuencia 430, que, por ejemplo, puede ser sustancialmente idéntica a la ruta de codificación del decodificador de audio 300 de acuerdo con la Fig. 3. En otras palabras, la ruta de codificación de dominio de frecuencia 430 puede comprender una decodificación de valor espectral 340, una decodificación de factor de escala 350, un escalador 360, un procesamiento opcional 366, una transformada de dominio de frecuencia a dominio de tiempo 370, un post-procesamiento opcional 376 y una ocultación de error 380, como se ha descrito anteriormente con referencia a la Fig. 3.

El decodificador de audio 400 puede además comprender una ruta de decodificación de dominio de predicción lineal 440 (que puede considerarse además una ruta de decodificación de dominio de tiempo, ya que la síntesis de LPC se realiza en el dominio de tiempo). La ruta de decodificación de dominio de predicción lineal comprende una decodificación de excitación 450, que recibe la excitación codificada 426 proporcionada por el analizador de corriente de bits 420 y proporciona, basándose en la misma, una excitación decodificada 452 (que puede adoptar la forma de una señal de excitación de dominio de tiempo decodificada). Por ejemplo, la decodificación de excitación 450 puede recibir una información de excitación codificada transformada codificada, y puede proporcionar,

basándose en la misma, una señal de excitación de dominio de tiempo decodificada. Por lo tanto, la decodificación de excitación 450, por ejemplo, puede realizar una funcionalidad que se realiza por el decodificador de excitación 730 descrito con referencia a la Fig. 7. Sin embargo, de manera alternativa o adicional, la decodificación de excitación 450 puede recibir una excitación ACELP codificada, y puede proporcionar la señal de excitación de dominio de tiempo 452 decodificada basándose en dicha información de excitación ACELP codificada.

Debe observarse que hay diferentes opciones para la decodificación de excitación. Se hace referencia, por ejemplo, a las Normas pertinentes y a las publicaciones que definen los conceptos de codificación de CELP, los conceptos de codificación de ACELP, las modificaciones de los conceptos de codificación CELP y de los conceptos de codificación de ACELP y el concepto de codificación de TCX.

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

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

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

La ruta de decodificación de dominio de predicción lineal 440 además comprende una síntesis de LPC (síntesis de codificación de predicción lineal) 470, que está configurada para recibir la excitación decodificada 452, o su versión procesada 456, y los coeficientes de predicción lineal decodificados 462, o su versión procesada 466, y para proporcionar una señal de audio de dominio de tiempo decodificada 472. Por ejemplo, la síntesis de LPC 470 puede estar configurada para aplicar una filtración, que se define por los coeficientes de predicción lineal decodificados 462 (o su versión procesada 466), a la señal de excitación de dominio de tiempo decodificada 452, o su versión procesada, de modo tal que la señal de audio de dominio de tiempo decodificada 472 se obtiene por la filtración (filtración de síntesis) de la señal de excitación de dominio de tiempo 452 (o 456). La ruta de decodificación de dominio de predicción lineal 440 puede comprender opcionalmente un post-procesamiento 474, que puede utilizarse para perfeccionar o ajustar las características de la señal de audio de dominio de tiempo decodificada 472.

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

El decodificador de audio 400 además comprende un combinador de señal (o combinación de señales 490), que está configurado para recibir la señal de audio de dominio de tiempo decodificada 372 (o su versión post-procesada 378), la información de audio de ocultación de error 382 proporcionada por la ocultación de error 380, la señal de audio de dominio de tiempo decodificada 472 (o su versión post-procesada 476) y la información de audio de ocultación de error 482 proporcionada por la ocultación de error 480. 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 de señal 490. Por consiguiente, el combinador de señal 490 puede proporcionar transiciones suaves 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 rutas de codificación 430, 440). Sin embargo, el combinador de señal 490 puede además proporcionar transiciones suaves si la señal de audio de dominio de tiempo

se proporciona por la misma entidad (por ejemplo, transformada de dominio de frecuencia a dominio de tiempo 370, o síntesis de LPC 470) para tramas subsiguientes. Debido a que algunos códecs tienen cierto solapamiento sobre la parte de superposición y adición que debe cancelarse, opcionalmente, podemos crear cierto solapamiento 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).

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

En síntesis, el decodificador de audio 400 permite la decodificación de tramas de audio que se codifican en el dominio de frecuencia, y tramas de audio que se codifican en el dominio de predicción lineal. En particular, es posible la conmutación entre el uso de la ruta de codificación de dominio de frecuencia y el uso de la ruta de codificación de dominio de predicción lineal dependiendo de las características de la señal (por ejemplo, usando una información de señalización proporcionada por un codificador de audio). Diferentes tipos de ocultación de error pueden usarse para proporcionar una información de audio de ocultación de error, en el caso de una pérdida de trama, dependiendo de si una última trama de audio apropiadamente decodificada se codificó 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).

5. Ocultación de dominio de tiempo de acuerdo con la Fig. 5.

La Fig. 5 muestra un diagrama de bloques esquemático de una ocultación de error de acuerdo con una realización de la presente invención. La ocultación de error de acuerdo con la Fig. 5 se designa en su totalidad como 500.

La ocultación de error 500 está configurada para recibir una señal de audio de dominio de tiempo 510 y para proporcionar, basándose en la misma, una información de audio de ocultación de error 512, que, por ejemplo, puede adoptar la forma de una señal de audio de dominio de tiempo.

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

La ocultación de error 500 comprende un preénfasis 520, que puede considerarse opcional. El preénfasis recibe la señal de audio de dominio de tiempo y proporciona, basándose en la misma, una señal de audio de dominio de tiempo pre-enfatizada 522.

La ocultación de error 500 además comprende un análisis de LPC 530, que está configurado para recibir la señal de audio de dominio de tiempo 510, o su versión pre-enfatizada 522, y para obtener una información de LPC 532, que puede comprender un conjunto de parámetros de LPC 532. Por ejemplo, la información de LPC puede comprender un conjunto de coeficientes de filtro de LPC (o una representación de los mismos) y una señal de excitación de dominio de tiempo (que se adapta para una excitación de un filtro de síntesis de LPC configurado de acuerdo con los coeficientes de filtro de LPC, para reconstruir, al menos aproximadamente, la señal de entrada del análisis de LPC).

La ocultación de error 500 además comprende una búsqueda de tono 540, que está configurada para obtener una información de tono 542, por ejemplo, basándose en una trama de audio previamente decodificada.

La ocultación de error 500 además comprende una extrapolación 550, que puede estar configurada para obtener una señal de excitación de dominio de tiempo extrapolada basándose en el resultado del análisis de LPC (por ejemplo, basándose en la señal de excitación de dominio de tiempo determinada por el análisis de LPC), y posiblemente, basándose en el resultado de la búsqueda de tono.

La ocultación de error 500 además comprende una generación de ruido 560, que proporciona una señal de ruido 562. La ocultación de error 500 además comprende un combinador/desvanecedor 570, que está configurado para recibir la señal de excitación de dominio de tiempo extrapolada 552 y la señal de ruido 562, y para proporcionar, basándose en las mismas, una señal de excitación de dominio de tiempo combinada 572. El combinador/desvanecedor 570 puede estar configurado para combinar la señal de excitación de dominio de tiempo extrapolada 552 y la señal de ruido 562, donde puede realizarse un desvanecimiento, de modo tal que una contribución relativa de la señal de excitación de dominio de tiempo extrapolada 552 (que determina un componente

determinista de la señal de entrada de la síntesis de LPC) disminuye en función del tiempo, mientras que una contribución relativa de la señal de ruido 562 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.

5 La ocultación de error 500 además comprende una síntesis de LPC 580, que recibe la señal de excitación de dominio de tiempo combinada 572 y que proporciona una señal de audio de dominio de tiempo 582 basándose en la misma. Por ejemplo, la síntesis de LPC además puede recibir coeficientes de filtro de LPC que describen un filtro de conformación de LPC, que se aplica a la señal de excitación de dominio de tiempo combinada 572, para derivar la
10 señal de audio de dominio de tiempo 582. La síntesis de LPC 580 puede, por ejemplo, usar coeficientes de LPC obtenidos basándose en una o más tramas de audio previamente decodificadas (por ejemplo, proporcionadas por el análisis de LPC 530).

15 La ocultación de error 500 además comprende un des-énfasis 584, que puede considerarse opcional. El des-énfasis 584 puede proporcionar una señal de audio de dominio de tiempo de ocultación de error des-enfatizada 586.

La ocultación de error 500 además comprende, opcionalmente, una superposición y adición 590, 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 y adición 590 debe considerarse
20 opcional, ya que la ocultación de error puede además usar una combinación de señales que ya se proporciona en el entorno del decodificador de audio. Por ejemplo, la superposición y adición 590 puede remplazarse por la combinación de señales 390 en el decodificador de audio 300 en algunas realizaciones.

A continuación, se describirán algunos detalles adicionales con respecto a la ocultación de error 500.

25 La ocultación de error 500 de acuerdo con la Fig. 5 cubre el contexto de un códec de dominio de transformada como AAC_LC o AAC_ELD. En otras palabras, la ocultación de error 500 se adapta bien para el uso en un códec de dominio de transformada de este tipo (y, en particular, en un decodificador de audio de dominio de transformada de este tipo). En el caso de un códec de transformada solamente (por ejemplo, en ausencia de una ruta 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 de audio de dominio de tiempo 372 puede usarse como un punto de inicio para la
30 ocultación 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 de audio de dominio de tiempo 372).

35 A continuación, se describirá en más detalle las subunidades y funcionalidades de la ocultación de error 500.

5.1. Análisis de LPC.

En la realización de acuerdo con la Fig. 5, toda la ocultación se realiza en el dominio de excitación para obtener una transición más suave entre tramas consecutivas. Por lo tanto, es necesario en primer lugar hallar (o, más generalmente, obtener) un conjunto apropiado de parámetros de LPC. En la realización de acuerdo con la Fig. 5, se realiza un análisis de LPC 530 sobre la señal de dominio de tiempo pre-enfatizada 522 pasada. Los parámetros de LPC (o coeficientes de filtro de LPC) se utilizan para realizar el análisis de LPC de la señal de síntesis pasada (por ejemplo, basándose en la señal de audio de dominio de tiempo 510, o basándose en la señal de audio de dominio
45 de tiempo pre-enfatizada 522) para obtener una señal de excitación (por ejemplo, una señal de excitación de dominio de tiempo).

5.2. Búsqueda de tono.

50 Existen diferentes enfoques para obtener el tono a utilizarse para lograr la construcción de la nueva señal (por ejemplo, la información de audio de ocultación de error).

En el contexto del códec usando un filtro de LTP (filtro de predicción a largo plazo), como AAC-LTP, si la última trama era AAC con LTP, usamos esta última demora de tono de LTP recibida y la correspondiente ganancia para la generación de la parte armónica. En este caso, la ganancia se utiliza para decidir si construir la parte armónica en la señal o no. Por ejemplo, si la ganancia de LTP es mayor que 0,6 (o cualquier otro valor predeterminado), entonces, la información de LTP se utiliza para construir la parte armónica.

Si no hay ninguna información de tono disponible de la trama previa, entonces, hay, por ejemplo, dos soluciones, que se describirán a continuación.

60 Por ejemplo, es posible realizar una búsqueda de tono en el codificador y transmitir en la corriente de bits la demora de tono y la ganancia. Esto es similar a la LTP, pero no hay ninguna aplicación de filtración (además, ninguna filtración de LTP en el canal limpio).

Como alternativa, es posible realizar una búsqueda de tono en el decodificador. La búsqueda de tono de AMR-WB en el caso de TCX se realiza en el dominio de FFT. En el ELD, por ejemplo, si se usó el dominio de MDCT, entonces, se perderán las fases. Por lo tanto, la búsqueda de tono se realiza preferentemente de manera directa en el dominio de excitación. Esto proporciona mejores resultados que la realización de la búsqueda de tono en el dominio de síntesis. La búsqueda de tono en el dominio de excitación se realiza primero con un bucle abierto por medio de una correlación cruzada normalizada. A continuación, opcionalmente, refinamos la búsqueda de tono realizando una búsqueda de bucle cerrado alrededor del tono de bucle abierto, con un cierto delta. Debido a las limitaciones de la generación de ventanas de ELD, podría hallarse un tono erróneo, y en consecuencia, además, verificamos que el tono hallado sea correcta, o de lo contrario, lo descartamos.

Como conclusión, el tono 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 ocultación de error. En algunos casos, hay una información de tono disponible de la decodificación de la trama previa (es decir, la última trama que precede la trama de audio perdida). En este caso, este tono puede reutilizarse (posiblemente, con cierta extrapolación y una consideración de un cambio de tono en función del tiempo). Además, podemos reutilizar opcionalmente el tono de más de una trama del pasado para intentar extrapolar el tono que necesitamos al final de nuestra trama ocultada.

Además, si hay una información (por ejemplo, designada como ganancia de predicción a largo plazo) disponible, que describe una intensidad (o intensidad relativa) de un componente de señal determinista (por ejemplo, al menos aproximadamente periódico), este valor puede utilizarse para decidir si un componente determinista (o armónico) debe ser incluido en la información de audio de ocultación de error. En otras palabras, al comparar dicho valor (por ejemplo, ganancia de LTP) con un valor umbral predeterminado, puede decidirse si una señal de excitación de dominio de tiempo derivada de una trama de audio previamente decodificada debe considerarse para proporcionar la información de audio de ocultación de error o no.

Si no hay información de tono disponible de la trama previa (o, más precisamente, de la decodificación de la trama previa), hay diferentes opciones. La información de tono 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. Como alternativa, la información de tono puede determinarse en el decodificador de audio, por ejemplo, en el dominio de excitación, es decir, basándose en 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 para identificar la información de tono para utilizarse para proporcionar la información de audio de ocultación de error.

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 calculada para la trama perdida o ya guardada en la trama perdida previa para múltiple pérdida de trama) se utiliza para la construcción de la parte armónica (también designada como componente determinista o componente aproximadamente periódico) en la excitación (por ejemplo, en la señal de entrada de la síntesis de LPC) mediante el copiado del último ciclo de tono tantas veces como sean necesarias para obtener una trama y media. Para 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 una trama de superposición.

En el caso de la primera trama perdida después de una trama buena (es decir, una trama apropiadamente decodificada), el primer ciclo de tono (por ejemplo, de la señal de excitación de dominio de tiempo obtenida basándose en la última trama de audio apropiadamente decodificada que precede la trama de audio perdida) se realiza filtro paso bajo con un filtro dependiente de la tasa de muestreo (ya que el ELD cubre una combinación de tasa de muestreo realmente amplia - que va de núcleo AAC-ELD a AAC-ELD con SBR o AAC-ELD de tasa dual SBR).

El tono en una señal de voz es casi siempre cambiante. Por lo tanto, la ocultación presentada anteriormente tiende a crear algunos problemas (o al menos distorsiones) en la recuperación, ya que el tono al final de la señal ocultada (es decir, al final de la información de audio de ocultación de error) con frecuencia no coincide con el tono de la primera trama buena. Por lo tanto, opcionalmente, en algunas realizaciones, se intenta predecir el tono al final de la trama ocultada para coincidir con el tono al comienzo de la trama de recuperación. Por ejemplo, se predice el tono al final de una trama perdida (que se considera una trama ocultada), en el que el objetivo de la predicción es establecer el tono al final de la trama perdida (trama ocultada) a aproximarse a el tono al comienzo de la primera trama apropiadamente decodificada seguido de 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). Para obtener incluso

mejores resultados, es posible reutilizar opcionalmente algunas herramientas convencionales y adaptarlas, como la resincronización de predicción de tono y pulsos. Para detalles, se hace referencia, por ejemplo, a la referencia [6] y [7].

- 5 Si se usa una predicción a largo plazo (LTP) en un códec de dominio de frecuencia, es posible usar la demora como la información de inicio acerca del tono. Sin embargo, en algunas realizaciones, se desea además contar con una mejor granularidad para poder rastrear mejor el contorno de tono. Por lo tanto, se prefiere realizar una búsqueda de tono al comienzo y al final de la última buena (apropiadamente decodificada) trama. Para adaptar la señal al tono en movimiento, es deseable la utilización de una resincronización de pulsos, que se presenta en el estado de la técnica.

10

5.4. Ganancia de tono

- 15 En algunas realizaciones, se prefiere la aplicación de una ganancia sobre la excitación previamente obtenida para lograr el nivel deseado. La “ganancia del tono” (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, para obtener la señal de entrada de la síntesis de LPC), puede, por ejemplo, obtenerse 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 cambiarse adaptativamente. El retardo es equivalente a la demora de tono 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 a continuación solo aplicar un desvanecimiento de salida (ganancia reducida) para la siguiente pérdida de trama consecutiva.

20

- 25 La “ganancia de tono” determinará la cantidad de tonalidad (o la cantidad de componentes de señal deterministas, al 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 del tono, entonces construimos una señal que consiste solo en un ruido con forma.

25

- 30 Como conclusión, en algunos casos, la señal de excitación de dominio de tiempo obtenida, por ejemplo, basándose en una trama de audio previamente decodificada, se escala dependiendo de la ganancia (por ejemplo, para obtener la señal de entrada para el análisis de LPC). Por consiguiente, debido a que la señal de excitación de dominio de tiempo determina un componente de señal determinista (al menos aproximadamente periódico), la ganancia puede determinar una intensidad relativa de dichos componentes de señal deterministas (al menos aproximadamente periódicos) en la información de audio de ocultación de error. Además, la información de audio de ocultación de error puede basarse en un ruido, que también está conformado por la síntesis de LPC, de modo tal que se adapta una energía total de la información de audio de ocultación de error, al 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 seguida de las una o más tramas de audio perdidas.

35

40 5.5. Creación de la parte de ruido

- Una “innovación” se crea por un generador de ruido aleatorio. Opcionalmente, a este ruido se realiza además opcionalmente filtrado paso alto y opcionalmente se pre-enfatiza 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 se proporciona, por ejemplo, por una generación de ruido 560) se conformará por la LPC (por ejemplo, por la síntesis de LPC 580) para llegar lo más cerca posible del ruido de fondo. La característica de paso alto se cambia también opcionalmente 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 para lograr un ruido de confort cercano al ruido de fondo.

50

- Una ganancia de innovación (que, por ejemplo, puede determinar una ganancia del ruido 562 en la combinación/desvanecimiento de salida 570, es decir, una ganancia usando la señal de ruido 562 que se incluye en la señal de entrada 572 de la síntesis de LPC) se calcula, por ejemplo, mediante la eliminación de la contribución previamente calculada del tono (si existe) (por ejemplo, una versión escalada, escalada usando la “ganancia de tono”, de la señal de excitación de dominio de tiempo obtenida basándose en 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 tono, esta podría realizarse opcionalmente únicamente sobre la primera trama perdida, y a continuación, el desvanecimiento de salida, si bien, en este caso, el desvanecimiento de salida podría ir o bien a 0, que da como resultado 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 tono utilizada para la creación de la parte armónica.

60

Opcionalmente, esta ganancia además se multiplica por (1-“ganancia de tono”) para aplicar tanta ganancia sobre el ruido de modo de alcanzar la pérdida de ganancia si la ganancia de tono no es uno. Opcionalmente, esta ganancia

se multiplica también 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 de salida

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

10 En el caso de una múltiple pérdida de trama, los parámetros de LPC no se recalculan. O bien, se mantiene el último calculado, o se realiza la ocultación de 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 de excitación de dominio de tiempo 502 obtenida basándose en una o más tramas de audio que preceden una trama de audio perdida aún utiliza una ganancia que se reduce gradualmente en función del tiempo, mientras que la señal de ruido 562 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 de excitación de dominio de tiempo 552 se reduce en función del tiempo en comparación con el peso relativo de la señal de ruido 562. En consecuencia, la señal de entrada 572 de la síntesis de LPC 580 se vuelve cada vez más “de tipo ruido”. Por lo tanto, la “periodicidad” (o, más precisamente, el componente determinista, o componente al menos aproximadamente periódico de la señal de salida 582 de la síntesis de LPC 580) se reduce en función del tiempo.

20 La velocidad de la convergencia de acuerdo con la cual la periodicidad de la señal 572, y/o la periodicidad de la señal 582, convergen 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 se controla por un factor de atenuación, α . El factor, α , es adicionalmente dependiente de la estabilidad del filtro LP. Opcionalmente, es posible alterar el factor α en la relación con la longitud de tono. Si el tono (por ejemplo, una longitud periódica asociada con el tono) es realmente larga, entonces mantenemos α “normal”, pero si el tono es realmente corto, 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 de salida más rápido de esta señal.

30 Además, opcionalmente, si está disponible, podemos considerar la salida de la predicción de tono. Si se predice un tono, esto significa que el tono 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 de salida de la parte tonal, en este caso.

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

40 5.7. Síntesis de LPC

Para regresar al dominio de tiempo, se prefiere realizar una síntesis de LPC 580 sobre la suma de las dos excitaciones (parte tonal y parte con ruido), seguido de un des-énfasis. En otras palabras, se prefiere realizar la síntesis de LPC 580 basándose en una combinación ponderada de una señal de excitación de dominio de tiempo 552 obtenida basándose en una o más tramas de audio apropiadamente decodificadas que preceden la trama de audio perdida (parte tonal) y la señal de ruido 562 (parte con ruido). Como se ha mencionado anteriormente, la señal de excitación de dominio de tiempo 552 puede modificarse en comparación con la señal de excitación de dominio de tiempo 532 obtenida por el análisis de LPC 530 (además de los coeficientes de LPC que describen una característica del filtro de la síntesis de LPC utilizado para la síntesis de LPC 580). Por ejemplo, la señal de excitación de dominio de tiempo 552 puede ser una copia escalada en tiempo de la señal de excitación de dominio de tiempo 532 obtenida por el análisis de LPC 530, en el que la escala de tiempo puede utilizarse para adaptar el tono de la señal de excitación de dominio de tiempo 552 a un tono deseado.

55 5.8. Superposición y adición.

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

60 En el contexto de la AAC o TCX regular, se aplica una superposición y adición entre la media trama adicional que proviene de la ocultación 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 ELD (retardo bajo adicional) para la primera trama perdida, se prefiere ejecutar el análisis

tres veces para obtener la contribución apropiada de las tres ventanas, y a continuación, para la primera trama de ocultación, y todas las siguientes, se ejecuta el análisis una vez más. A continuación, se realiza una síntesis de ELD, para volver al dominio de tiempo con toda la memoria apropiada para la siguiente trama en el dominio de MDCT.

5 Como conclusión, la señal de entrada 572 de la síntesis de LPC 580 (y/o la señal de excitación de dominio de tiempo 552) puede proporcionarse para una duración temporal que es mayor que una duración de una trama de audio perdida. Por consiguiente, la señal de salida 582 de la síntesis de LPC 580 puede proporcionarse además para un periodo 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 ocultación de error (que se obtiene en consecuencia para un periodo de tiempo más largo que una extensión temporal de la trama de audio perdida) y una información de audio decodificada proporcionada para una trama de audio apropiadamente decodificada seguido de una o más tramas de audio perdidas.

15 En síntesis, la ocultación de error 500 se adapta bien al caso en el cual las tramas de audio se codifican en el dominio de frecuencia. Incluso aunque las tramas de audio se codifiquen en el dominio de frecuencia, la provisión de la información de audio de ocultación de error se realiza basándose en una señal de excitación de dominio de tiempo. Se aplican diferentes modificaciones a la señal de excitación de dominio de tiempo obtenida basándose en 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 proporcionada por el análisis de LPC 530 se adapta a cambios de tono, por ejemplo, usando una escala de tiempo. Además, la señal de excitación de dominio de tiempo proporcionada por el análisis de LPC 530 se modifica además por una escala (aplicación de una ganancia), en el que un desvanecimiento de salida del componente determinista (o tonal, o al menos aproximadamente periódico) puede realizarse por el escalador/desvanecedor 570, de modo tal que la señal de entrada 572 de la síntesis de LPC 580 comprende tanto un componente que se deriva de la señal de excitación de dominio de tiempo obtenida por el análisis de LPC como un componente de ruido que se basa en la señal de ruido 562. El componente determinista de la señal de entrada 572 de la síntesis de LPC 580, sin embargo, se modifica habitualmente (por ejemplo, se escala en tiempo y/o se escala en amplitud) con respecto a la señal de excitación de dominio de tiempo proporcionada por el análisis de LPC 530.

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

6. Ocultación de dominio de tiempo de acuerdo con la Fig. 6.

35 La Fig. 6 muestra un diagrama de bloques esquemático de una ocultación de dominio de tiempo que puede usarse para un códec de conmutación. Por ejemplo, la ocultación de dominio de tiempo 600 de acuerdo con la Fig. 6 puede, por ejemplo, tomar el lugar de la ocultación de error 240, o el lugar de la ocultación de error 480.

40 Además, debe observarse que la realización de acuerdo con la Fig. 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 (MPEG-D/MPEG-H) o EVS (3GPP). En otras palabras, la ocultación de dominio de tiempo 600 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 basándose en coeficientes de predicción lineal).

45 Sin embargo, debe observarse que la ocultación de error 600 de acuerdo con la Fig. 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).

50 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 de acuerdo con la Fig. 5, es decir, realizar un análisis de LPC.

60 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 a largo plazo), además tenemos la información de demora que proviene de la predicción a largo plazo. Y si la última trama estuvo en el dominio de frecuencia sin predicción a largo plazo LTP, entonces la búsqueda de tono se realiza preferentemente de manera directa en el dominio de excitación (por ejemplo, basándose en una señal de excitación de dominio de tiempo proporcionada por un análisis de LPC).

Si el decodificador ya utiliza algunos parámetros de LPC en el dominio de tiempo, los reutilizamos y extrapolamos un

nuevo conjunto de parámetros de LPC. La extrapolación de los parámetros de LPC está basada en la LPC pasada, por ejemplo, la media de las últimas tres tramas y (opcionalmente), la forma de la LPC derivada durante la estimación de ruido DTX si existe DTX (transmisión discontinua) en el códec.

- 5 Toda la ocultación se realiza en el dominio de excitación para obtener una transición más suave entre tramas consecutivas.

A continuación, se describirá en más detalle la ocultación de error 600 de acuerdo con la Fig. 6.

- 10 La ocultación de error 600 recibe una excitación pasada 610 y una información de tono pasada 640. Además, la ocultación de error 600 proporciona una información de audio de ocultación de error 612.

- 15 Debe observarse que la excitación pasada 610 recibida por la ocultación de error 600 puede, por ejemplo, corresponder a la salida 532 del análisis de LPC 530. Además, la información de tono pasada 640 puede, por ejemplo, corresponder a la información de salida 542 de la búsqueda de tono 540.

La ocultación de error 600 adicionalmente comprende una extrapolación 650, que puede corresponder a la extrapolación 550, de modo tal que se hace referencia al análisis anterior.

- 20 Además, la ocultación de error comprende un generador de ruido 660, que puede corresponder al generador de ruido 560, de modo tal que se hace referencia al análisis anterior.

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

- 30 La ocultación de error 600 además comprende un combinador/desvanecedor 670, que recibe la señal de excitación de dominio de tiempo extrapolada 652 y la señal de ruido 662 y proporciona, basándose en las mismas, una señal de entrada 672 para una síntesis de LPC 680, en el que la síntesis de LPC 680 puede corresponder a la síntesis de LPC 580, de modo tal que se aplican también las explicaciones anteriores. La síntesis de LPC 680 proporciona una señal de audio de dominio de tiempo 682, que puede corresponder a la señal de audio de dominio de tiempo 582. La ocultación de error además comprende (opcionalmente) un des-énfasis 684, que puede corresponder al des-énfasis 584 y que proporciona una señal de audio de dominio de tiempo de ocultación de error des-enfatizada 686. La ocultación de error 600 opcionalmente comprende una superposición y adición 690, que puede corresponder a la superposición y adición 590. Sin embargo, se aplican también las explicaciones anteriores con respecto a la superposición y adición 590, a la superposición y adición 690. En otras palabras, la superposición y adición 690 puede además reemplazarse por la superposición y adición general del decodificador de audio, de modo tal que la señal de salida 682 de la síntesis de LPC o la señal de salida 686 del des-énfasis pueden considerarse la información de audio de ocultación de error.

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

A continuación, se describirán en más detalle algunos rasgos de la ocultación de error 600. Sin embargo, debe observarse que los detalles específicos deben considerarse ejemplares, en lugar de rasgos esenciales.

- 50 6.1. Tono pasado de búsqueda de tono

Hay diferentes enfoques para obtener el tono a usarse 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 para decidir si queremos construir la parte armónica en la señal o no. Por ejemplo, si la ganancia de LTP es superior a 0,6, entonces, usamos la información de LTP para construir la parte armónica.

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

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

ninguna filtración de predicción a largo plazo en el canal limpio).

Otra solución es realizar una búsqueda de tono en el decodificador. La búsqueda de tono de AMR-WB en el caso de TCX se realiza en el dominio de FFT. En TCX, por ejemplo, usamos el dominio de MDCT, entonces, perdemos las fases. Por lo tanto, la búsqueda de tono se realiza directamente en el dominio de excitación (por ejemplo, basándose en la señal de excitación de dominio de tiempo utilizada como la entrada de la síntesis de LPC, o utilizada para derivar la entrada para la síntesis de LPC), en una realización preferida. Esto habitualmente proporciona mejores resultados que la realización de la búsqueda de tono en el dominio de síntesis (por ejemplo, basándose en una señal de audio de dominio de tiempo completamente decodificada).

La búsqueda de tono en el dominio de excitación (por ejemplo, basándose en la señal de excitación de dominio de tiempo) se realiza primero con un bucle abierto por medio de una correlación cruzada normalizada. A continuación, opcionalmente, la búsqueda de tono puede perfeccionarse realizando una búsqueda de bucle cerrado alrededor del tono de bucle abierto con una cierta delta.

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

Como conclusión, hay diferentes conceptos para determinar el tono, en el que es computacionalmente eficaz considerar un tono pasado (es decir, tono asociado con una trama de audio previamente decodificada). Como alternativa, la información de tono puede transmitirse desde un codificador de audio hacia un decodificador de audio. Como otra alternativa, una búsqueda de tono puede realizarse en el lado del decodificador de audio, en el que la determinación de tono se realiza, preferentemente, basándose en la señal de excitación de dominio de tiempo (es decir, en el dominio de excitación).

Una búsqueda de tono de dos etapas que comprende una búsqueda de bucle abierto y una búsqueda de bucle cerrado puede realizarse para obtener una información de tono particularmente confiable y precisa. Como alternativa, o además, una información de tono de una trama de audio previamente decodificada puede usarse para garantizar que la búsqueda de tono proporciona un resultado confiable.

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

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

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

Si se usa la LTP en TCX, puede usarse la demora como la información inicial acerca del tono. Sin embargo, es deseable contar con una mejor granularidad para poder rastrear mejor el contorno de tono. Por lo tanto, se realiza una búsqueda de tono opcionalmente al comienzo y al final de la última buena trama. Para adaptar la señal al tono en movimiento, puede usarse una resincronización de pulsos, que está presente 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 basándose en, 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, en el que la porción de tiempo copiada puede modificarse dependiendo de un cálculo, o una estimación, de un cambio de tono (esperado) durante la trama de audio perdida. Están disponibles diferentes conceptos para la determinación del cambio de tono.

6.3. Ganancia de tono.

En la realización de acuerdo con la Fig. 6, se aplica una ganancia sobre la excitación obtenida previamente para alcanzar un nivel deseado. La ganancia del tono se obtiene, 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
 5 equivalente a la longitud de dos subtramas, y el retardo puede ser equivalente a la demora de tono 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 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 LPC cambia en cada trama, y entonces, la aplicación de una ganancia, calculada sobre la trama previa, sobre una señal de excitación que se procesará por otro conjunto de
 10 LPC, no proporcionará la energía esperada en el dominio de tiempo.

La ganancia del tono determina la cantidad de tonalidad que se creará, aunque también se agregará cierto ruido conformado para no tener solo un tono artificial. Si se obtiene una muy baja ganancia de tono, entonces, puede construirse una señal que consiste solo en un ruido conformado.
 15

Como conclusión, una ganancia que se aplica para escalar la señal de excitación de dominio de tiempo obtenida basándose en la trama previa (o una señal de excitación de dominio de tiempo que se obtiene para una trama previamente decodificada, o que se asocia con la trama previamente decodificada) se ajusta para determinar de esta manera una ponderación de un componente tonal (o determinista, o al menos aproximadamente periódico) dentro de la señal de entrada de la síntesis de LPC 680, y, en consecuencia, dentro de la información de audio de ocultación de error. Dicha ganancia puede determinarse basándose en una correlación, que se aplica a la señal de audio de dominio de tiempo obtenida por una decodificación de la trama previamente decodificada (en el que dicha señal de audio de dominio de tiempo puede obtenerse usando una síntesis de LPC que se realiza en el curso de la decodificación).
 20
 25

6.4. Creación de la parte de ruido

Se crea una innovación por medio de un generador de ruido aleatorio 660. Este ruido adicionalmente se realiza filtrado de paso alto y opcionalmente pre-enfatizado 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 Fig. 6, sino que pueden realizarse, por ejemplo, dentro del generador de ruido 660 o dentro del combinador/desvanecedor 670.
 30

El ruido se conformará (por ejemplo, seguido de la combinación con la señal de excitación de dominio de tiempo 652 obtenida por la extrapolación 650) por la LPC para obtener lo más cercano posible al ruido de fondo. Por ejemplo, la ganancia de innovación puede calcularse mediante la eliminación de la contribución previamente calculada del tono (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 tono utilizada para la creación de la parte armónica.
 35
 40

Opcionalmente, esta ganancia puede además multiplicarse por (1-ganancia de tono) para aplicar tanta ganancia sobre el ruido para alcanzar la pérdida de energía si la ganancia del tono no es uno. Opcionalmente, esta ganancia además se multiplica por un factor de ruido. Este factor de ruido puede provenir de una trama válida previa.

Como conclusión, un componente de ruido de la información de audio de ocultación de error se obtiene mediante la conformación de ruido proporcionada por el generador de ruido 660 usando la síntesis de LPC 680 (y, posiblemente, el des-énfasis 684). Además, puede aplicarse una filtración de paso alto adicional y/o un preénfasis. La ganancia de la contribución de ruido a la señal de entrada 672 de la síntesis de LPC 680 (también designada "ganancia de innovación") puede calcularse basándose en la última trama de audio apropiadamente decodificada que precede la trama de audio perdida, en el que un componente determinista (o al menos aproximadamente periódico) puede eliminarse de la trama de audio que precede la trama de audio perdida, y en el que puede realizarse entonces una correlación para 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.
 45
 50

Opcionalmente, pueden aplicarse ciertas modificaciones adicionales a la ganancia del componente de ruido.
 55

6.5. Desvanecimiento de salida

El desvanecimiento de salida se usa en su mayoría para múltiples pérdidas de tramas. Sin embargo, el desvanecimiento de salida puede usarse además en el caso de que solo se pierda una única trama de audio.
 60

En el caso de múltiple pérdida de trama, los parámetros de LPC no se recalculan. O bien se mantiene el último calculado, o se realiza una ocultación de LPC como se ha explicado anteriormente.

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, α . El factor, α , además depende de la estabilidad del filtro de LP. Opcionalmente, el factor α puede modificarse en relación con la longitud de tono. Por ejemplo, si el tono es

5 realmente largo, entonces α puede mantenerse normal, pero si el tono es realmente corto, puede ser conveniente (o necesario) copiar una cantidad de veces la misma parte de excitación pasada. Debido a que se ha hallado que esto rápidamente sonará demasiado artificial, la señal, por lo tanto se desvanece de salida más rápidamente.

Además opcionalmente, es posible considerar la salida de predicción de tono. Si se predice un tono, esto significa que el tono 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 de salida de la parte tonal, en este caso.

10

Si la predicción de tono fallara debido a que el tono cambia demasiado, esto significa que o bien los valores de tono no son realmente confiables, o que la señal es realmente impredecible. Por lo tanto, nuevamente, deberíamos realizar el desvanecimiento de salida más rápidamente.

15

Como conclusión, la contribución de la señal de excitación de dominio de tiempo extrapolada 652 a la señal de entrada 672 de la síntesis de LPC 680 habitualmente se reduce en función del tiempo. Esto puede lograrse, por ejemplo, reduciendo un valor de ganancia, que se aplica a la señal de excitación de dominio de tiempo extrapolada 652, en función del tiempo. La velocidad utilizada para reducir gradualmente la ganancia aplicada para escalar la

20 señal de excitación de dominio de tiempo 552 obtenida basándose en una o más tramas de audio que preceden una trama de audio perdida (o una o más de sus copias) se ajusta dependiendo de uno o más parámetros de una o más tramas de audio (y/o dependiendo de una cantidad de tramas de audio perdidas consecutivas). En particular, la longitud de tono y/o la tasa a la que cambia el tono en función del tiempo, y/o la cuestión de si una predicción de

25 tono falla o tiene éxito, pueden utilizarse para ajustar dicha velocidad.

6.6. Síntesis de LPC

Para retornar al dominio de tiempo, se realiza una síntesis de LPC 680 sobre la suma en general (o generalmente, la combinación ponderada) de las dos excitaciones (parte tonal 652 y parte con ruido 662), seguido del des-énfasis 684.

30

En otras palabras, el resultado de la combinación ponderada (desvanecimiento) de la señal de excitación de dominio de tiempo extrapolada 652 y la señal de ruido 662 forma una señal de excitación de dominio de tiempo combinada, que se introduce en la síntesis de LPC 680, que, por ejemplo, puede realizar una filtración de síntesis basándose en dicha señal de excitación de dominio de tiempo combinada 672 dependiendo de los coeficientes de LPC que describen el filtro de síntesis.

35

6.7. Superposición y adición

Debido a que no se conoce, durante la ocultación, el modo de la siguiente trama que llega (por ejemplo, ACELP, TCX o FD), se prefiere preparar diferentes superposiciones por adelantado. Para 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 ocultación de error) puede, por ejemplo, crearse para la mitad de una trama más que la trama ocultada (perdida). Además, puede crearse solapamiento artificial sobre ella (donde el solapamiento artificial puede, por ejemplo, adaptarse a la superposición y adición de MDCT).

40

45

Para obtener una buena superposición y adición sin discontinuidad con la futura trama en el dominio de tiempo (ACELP), hacemos como anteriormente, pero sin solapamiento, para poder aplicar ventanas de superposición y adición largas, o si queremos usar una ventana cuadrada, se calcula la respuesta de entrada a cero (ZIR) al final de la memoria intermedia de síntesis.

50

Como conclusión, en un decodificador de audio de conmutación (que, por ejemplo, puede conmutar entre una decodificación de ACELP, una decodificación de TCX y una decodificación de dominio de frecuencia (decodificación de FD)), puede realizarse una superposición y adición entre la información de audio de ocultación de error que se proporciona principalmente para una trama de audio perdida, pero también para una cierta porción de tiempo seguida de la trama de audio perdida, y la información de audio decodificada proporcionada para la primera trama de audio apropiadamente decodificada seguida de una secuencia de una o más tramas de audio perdidas. Para obtener una apropiada superposición y adición, incluso, para modos de decodificación que acarrear un solapamiento de dominio de tiempo en una transición entre subsiguientes tramas de audio, puede proporcionarse una información de cancelación de solapamiento (por ejemplo, designada como solapamiento artificial). Por consiguiente, una superposición y adición entre la información de audio de ocultación de error y la información de audio de dominio de tiempo obtenida basándose en la primera trama de audio apropiadamente decodificada seguida de una trama de audio perdida, logra una cancelación del solapamiento.

55

60

Si la primera trama de audio apropiadamente decodificada seguida de la secuencia de una o más tramas de audio perdidas se codifica en el modo de ACELP, puede calcularse una información de superposición específica, que puede basarse en una respuesta de entrada a cero (ZIR) de un filtro de LPC.

- 5 Como conclusión, la ocultación de error 600 se adapta bien para el uso en un códec de audio de conmutación. Sin embargo, la ocultación de error 600 puede además usarse en un códec de audio que meramente decodifica un contenido de audio codificado en un modo de TCX o en un modo de ACELP.

6.8. Conclusión

- 10 Debe observarse que se logra una ocultación de error particularmente buena mediante el concepto anteriormente mencionado, para extrapolar una señal de excitación de dominio de tiempo, para combinar el resultado de la extrapolación con una señal de ruido usando un desvanecimiento (por ejemplo, un desvanecimiento cruzado), y para realizar una síntesis de LPC basándose en un resultado del desvanecimiento cruzado.

15 7. Decodificador de audio de acuerdo con la Fig. 11

- 20 La Fig. 11 muestra un diagrama de bloques esquemático de un decodificador de audio 1100, de acuerdo con un ejemplo. Debe observarse que el decodificador de audio 1100 puede ser parte de un decodificador de audio de conmutación. Por ejemplo, el decodificador de audio 1100 puede reemplazar la ruta de decodificación de dominio de predicción lineal 440 en el decodificador de audio 400.

- 25 El decodificador de audio 1100 está configurado para recibir una información de audio codificada 1110 y para proporcionar, basándose en la misma, una información de audio decodificada 1112. La información de audio codificada 1110 puede, por ejemplo, corresponder a la información de audio codificada 410, y la información de audio decodificada 1112 puede, por ejemplo, corresponder a la información de audio decodificada 412.

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

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

- 40 El decodificador de audio 1100 además comprende un coeficiente de codificación de predicción lineal para la conversión de factor de escala 1140, que está configurado para proporcionar un conjunto de factores de escala 1142 basándose en la representación codificada 1124 de coeficientes de codificación de predicción lineal. Por ejemplo, el coeficiente de codificación de predicción lineal para la conversión de factor de escala 1142 puede realizar una funcionalidad que se describe en la norma USAC. Por ejemplo, la representación codificada 1124 de los coeficientes de codificación de predicción lineal puede comprender una representación polinomial, que se decodifica y convierte en un conjunto de factores de escala por el coeficiente de codificación de predicción lineal para la conversión de factor de escala 1142.

- 50 El decodificador de audio 1100 además comprende un escalador 1150, que está configurado para aplicar los factores de escala 1142 a los valores espectrales decodificados 1132, para obtener de esta manera valores espectrales decodificados escalados 1152. Además, el decodificador de audio 1100 comprende, opcionalmente, un procesamiento 1160, que, por ejemplo, puede corresponder al procesamiento 366 que se ha descrito anteriormente, en el que los valores espectrales decodificados escalados procesados 1162 se obtienen por el procesamiento opcional 1160. El decodificador de audio 1100 además comprende una transformada de dominio de frecuencia a dominio de tiempo 1170, que está configurada para recibir los valores espectrales decodificados escalados 1152 (que pueden corresponder a los valores espectrales decodificados escalados 362), o los valores espectrales decodificados escalados procesados 1162 (que pueden corresponder a los valores espectrales decodificados escalados procesados 368) y proporcionar, basándose en los mismos, una representación de dominio de tiempo 1172, que puede corresponder a la representación de dominio de tiempo 372 que se ha descrito anteriormente. El decodificador de audio 1100 además comprende un primer post-procesamiento opcional 1174, y un segundo post-procesamiento opcional 1178, que, por ejemplo, pueden corresponder, al menos en parte, al post-procesamiento opcional 376 anteriormente mencionado. Por consiguiente, el decodificador de audio 1110 obtiene (opcionalmente) una versión post-procesada 1179 de la representación de audio de dominio de tiempo 1172.

El decodificador de audio 1100 también comprende un bloque de ocultación de error 1180 que está configurado para

recibir la representación de audio de dominio de tiempo 1172, o una versión post-procesada de la misma, y los coeficientes de codificación de predicción lineal (o bien en forma codificada, o en una forma decodificada) y proporciona, basándose en los mismos, una información de audio de ocultación de error 1182.

5 El bloque de ocultación de error 1180 está configurado para proporcionar la información de audio de ocultación de error 1182 para la ocultación de una pérdida de una trama de audio seguida de 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 a la ocultación de error 380 y a la ocultación de error 480, y también a la ocultación de error 500 y a la ocultación de error 600.

10 Sin embargo, el bloque de ocultación de error 1180 comprende un análisis de LPC 1184, que es sustancialmente idéntico al análisis de LPC 530. Sin embargo, el análisis de LPC 1184 puede, opcionalmente, usar los coeficientes de LPC 1124 para facilitar el análisis (en comparación con el análisis de LPC 530). El análisis de LPC 1134 proporciona una señal de excitación de dominio de tiempo 1186, que es sustancialmente idéntica a la señal de excitación de dominio de tiempo 532 (y también a la señal de excitación de dominio de tiempo 610). Además, el bloque de ocultación de error 1180 comprende una ocultación de error 1188, que, por ejemplo, puede realizar la funcionalidad de los bloques 540, 550, 560, 570, 580, 584 de la ocultación de error 500, o que, por ejemplo, puede realizar la funcionalidad de los bloques 640, 650, 660, 670, 680, 684 de la ocultación de error 600. Sin embargo, el bloque de ocultación de error 1180 difiere ligeramente de la ocultación de error 500, y también de la ocultación de error 600. Por ejemplo, el bloque de ocultación de error 1180 (que comprende el análisis de LPC 1184) difiere de la ocultación de error 500 en términos de que los coeficientes de LPC (utilizados para la síntesis de LPC 580) no se determinan por el análisis de LPC 530, sino que se reciben (opcionalmente) desde la corriente de bits. Asimismo, el bloque de ocultación de error 1188, que comprende el análisis de LPC 1184, difiere de la ocultación de error 600 en términos de que la "excitación pasada" 610 se obtiene por el análisis de LPC 1184, en lugar de estar disponible directamente.

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

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

35 8. Método de acuerdo con la Fig. 9

La Fig. 9 muestra un diagrama de flujo de un método para proporcionar una información de audio decodificada basándose en una información de audio codificada. El método 900 de acuerdo con la Fig. 9 comprende proporcionar 910 una información de audio de ocultación de error para la ocultación de una pérdida de una trama de audio seguido de una trama de audio codificada en una representación de dominio de frecuencia usando una señal de excitación de dominio de tiempo. El método 900 de acuerdo con la Fig. 9 está basado en las mismas consideraciones que el decodificador de audio de acuerdo con la Fig. 1. Además, debe observarse que el método 900 puede complementarse por cualquiera de los rasgos y de las funcionalidades que se describen en el presente documento, o bien individualmente, o en combinación.

45 9. Método de acuerdo con la Fig. 10

La Fig. 10 muestra un diagrama de flujo de un método para proporcionar una información de audio decodificada basándose en una información de audio codificada. El método 1000 comprende proporcionar 1010 una información de audio de ocultación de error para la ocultación de una pérdida de una trama de audio, en el que una señal de excitación de dominio de tiempo obtenida para (o basándose en) una o más tramas de audio que preceden una trama de audio perdida se modifican para obtener la información de audio de ocultación de error.

El método 1000 de acuerdo con la Fig. 10 está basado en las mismas consideraciones que el decodificador de audio anteriormente mencionado de acuerdo con la Fig. 2.

Además, debe observarse que el método de acuerdo con la Fig. 10 puede complementarse por cualquiera de los rasgos y de las funcionalidades que se describen en el presente documento, o bien individualmente, o en combinación.

60 10. Observaciones adicionales

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. Como alternativa, la señal de excitación de dominio de tiempo para la segunda trama perdida puede basarse en un análisis de LPC de la señal de síntesis de la trama previa perdida. Por ejemplo, en un códec, la LPC puede ser cambiante en cada trama perdida; entonces, tiene sentido la nueva realización del análisis para cada trama perdida.

11. Alternativas de implementación

Aunque se han descrito algunos aspectos en el contexto de un aparato, es evidente 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. De manera análoga, los aspectos descritos en el contexto de una etapa de método además representan una descripción de un correspondiente bloque o elemento o rasgo de un correspondiente aparato. Algunas o la totalidad de las etapas de método pueden ejecutarse por (o usando) un aparato de hardware, 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 ejecutarse por dicho aparato.

Dependiendo de ciertos requerimientos de implementación, las realizaciones de la invención pueden implementarse en hardware o software. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo, un disquete, un DVD (disco versátil digital), un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control electrónicamente legibles almacenadas en las mismas, que cooperan (o pueden cooperar) con un sistema informático programable de manera tal que se realiza el respectivo método. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

Algunas realizaciones de acuerdo con la invención comprenden un soporte de datos que tiene señales de control electrónicamente legibles, que pueden cooperar con un sistema informático programable, de manera que se realiza uno de los métodos descritos en el presente documento.

En general, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, el código de programa es operativo para realizar uno de los métodos cuando se ejecuta el producto de programa informático en un ordenador. El código de programa puede almacenarse, por ejemplo, en un portador legible por máquina.

Otras realizaciones comprenden el programa informático para realizar uno de los métodos que se describen en el presente documento, almacenado en un portador legible por máquina.

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

Una realización adicional del método de la invención es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para llevar a cabo uno de los métodos que se describen en el presente documento. El soporte 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 informático 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, pueden estar configuradas para transferirse 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 instalado en el mismo el programa informático para llevar a cabo uno de los métodos que se describen en el presente documento.

Una realización adicional de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa informático para la realización de uno de los métodos descritos en el presente documento, 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 pueden comprender, por ejemplo, un servidor de archivos para transferir el programa informático al receptor.

En algunas realizaciones, puede usarse un dispositivo lógico programable (por ejemplo, un campo de matrices de puertas programables) para realizar algunas o la totalidad de las funcionalidades de los métodos que se describen

en el presente documento. En algunas realizaciones, un campo de matrices de puertas programables puede cooperar con un microprocesador para llevar a cabo uno de los métodos que se describen en el presente documento. En general, los métodos se realizan, preferentemente, por cualquier aparato de hardware.

- 5 El aparato descrito en el presente documento puede implementarse usando un aparato de hardware, o usando un ordenador, o empleando una combinación de un aparato de hardware y un ordenador.

Los métodos que se describen en el presente documento pueden realizarse usando un aparato de hardware, o usando un ordenador, o empleando una combinación de un aparato de hardware y un ordenador.

- 10 Las realizaciones anteriormente descritas 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

- 20 Como conclusión, aunque se ha descrito cierta ocultación para códecs de dominio de transformada en el campo, las realizaciones de acuerdo con la invención superan los códecs (o decodificadores) convencionales. Las realizaciones de acuerdo con la invención usan un cambio de dominio para la ocultación (dominio de frecuencia a dominio de tiempo o excitación). Por consiguiente, las realizaciones de acuerdo con la invención crean una ocultación de voz de alta calidad para decodificadores de dominio de transformada.

- 25 El modo de codificación de transformada es similar al de USAC (consúltese, por ejemplo, referencia [3]). Utiliza la transformada de coseno discreta modificada (MDCT) como una transformada, y la conformación de ruido espectral se logra mediante la aplicación de la envoltura espectral de LPC ponderada en el dominio de frecuencia (además conocida como FDNS, "conformación de ruido de dominio de frecuencia"). En otras palabras, las realizaciones de acuerdo con la invención pueden usarse en un decodificador de audio, que utiliza los conceptos de decodificación que se describen en la norma USAC. Sin embargo, el concepto de ocultación de error divulgado en el presente documento puede además usarse en un decodificador de audio que es de tipo "AAC" o en cualquier códec (o decodificador) de la familia AAC.

- 30 El concepto de acuerdo con 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, la ocultación se realiza en el dominio de tiempo o en el dominio de excitación.

- 35 A continuación, se describirán algunas ventajas y algunos rasgos de la ocultación de dominio de tiempo (o de la ocultación de dominio de excitación).

- 40 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 de acuerdo con la invención crean una nueva ocultación 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 una ocultación de tipo ACELP, y aumenta la calidad de la ocultación. Se ha hallado que la información de tono es ventajosa (o incluso requerida, en algunos casos) para una ocultación de tipo ACELP. Por lo tanto, las realizaciones de acuerdo con la presente invención están configuradas para hallar valores de tono confiables para la trama previa codificada en el dominio de frecuencia.

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

- 50 Como conclusión, las realizaciones de acuerdo con la invención crean una ocultación de error que supera las soluciones convencionales.

- 55 Bibliografía.

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

- 60 [2] "MDCT-BASED CODER FOR HIGHLY ADAPTIVE SPEECH AND AUDIO CODING"; Guillaume Fuchs & al.; EUSIPCO 2009.

- [3] ISO_IEC_DIS_23003-3_(E); Information technology - MPEG audio technologies - Part 3: Unified speech and

audio coding.

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

5

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

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

10

[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, 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
10 obtener la información de audio de ocultación de error; en el que la ocultación de error está configurada para modificar una señal de excitación de dominio de tiempo (452, 456; 610) derivada de una o más tramas de audio codificadas en una representación de dominio de frecuencia que preceden una trama de audio perdida, para obtener la información de audio de ocultación de error.

15 2. El decodificador de audio (200; 400) de acuerdo con la reivindicación 1, en el que la ocultación de error (240; 480; 600) está configurada para usar una o más copias modificadas de la 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 ocultación de error (242; 482; 612).

20 3. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 1 a 2, en el que la ocultación de error (240; 482; 612) está configurada para modificar la 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, o una o más copias de la misma, para reducir de esta manera un componente periódico de la información de audio de ocultación de error (242; 482; 612) en función del tiempo.

25 4. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 1 a 3, en el que la ocultación de error (240; 480; 600) está configurada para escalar la señal de excitación de dominio de tiempo (452, 456; 610) obtenida para una o más tramas de audio que preceden la trama de audio perdida, o una o más copias de la misma, para modificar de esta manera la señal de excitación de dominio de tiempo.

30 5. El decodificador de audio (200; 400) de acuerdo con la reivindicación 3 o 4, en el que la ocultación de error (240; 480; 600) está configurada para reducir gradualmente una ganancia aplicada para escalar la 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, o la una o más copias de la misma.

35 6. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 3 a 5, en el que la ocultación de error (240; 480; 600) está configurada para ajustar una velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la 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, o la una o más copias de la misma, dependiendo de uno o más parámetros de una o más tramas de audio que preceden la trama de audio perdida, y/o dependiendo de una cantidad de tramas de audio perdidas consecutivas.

40 7. El decodificador de audio (200; 400) de acuerdo con la reivindicación 5 o la reivindicación 6, en el que la ocultación de error (240; 480; 600) está configurada para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la 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, o la una o más copias de la misma, dependiendo de una longitud de un periodo de tono de la señal de excitación de dominio de tiempo, de modo tal que un componente determinista de la señal de excitación de dominio de tiempo (672) introducido en una síntesis de LPC (680) se desvanece de salida más rápidamente para señales que tienen una longitud más corta del periodo de tono en comparación con señales que tienen una longitud más larga del periodo de tono.

45 8. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 5 a 7, en el que la ocultación de error (240; 480; 600) está configurada para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la 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, o la una o más copias de la misma, dependiendo de un resultado de un análisis de tono o una predicción de tono; de modo tal que un componente determinista de la señal de excitación de dominio de tiempo (572) introducido en una síntesis de LPC (580) se desvanece de salida más rápidamente para señales que tienen un mayor cambio de
50 tono por unidad de tiempo en comparación con señales que tienen un menor cambio de tono por unidad de tiempo; y/o de modo tal que un componente determinista de una señal de excitación de dominio de tiempo (572) introducido en una síntesis de LPC (580) se desvanece de salida más rápidamente para señales para las cuales falla una predicción de tono en comparación con señales para las cuales la predicción de tono tiene éxito.

9. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 1 a 8, en el que la ocultación de error (240; 480; 600) está configurada para la escala de tiempo de la señal de excitación de dominio de tiempo (452, 456; 610) obtenida basándose en una o más tramas de audio que preceden una trama de audio perdida, o la una o más copias de la misma, dependiendo de una predicción de un tono para el tiempo de las una o más tramas de audio perdidas.
10. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 1 a 9, en el que la ocultación de error (240; 480; 600) está configurada para obtener una señal de excitación de dominio de tiempo (452, 456; 610), 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 de excitación de dominio de tiempo modificada (652, 672); y en el que la ocultación de error está configurada para proporcionar la información de audio de ocultación de error (242; 482; 612) basándose en la señal de excitación de dominio de tiempo modificada (652, 672).
11. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 1 a 10, en el que la ocultación de error (240; 480; 600) está configurada para obtener una información de tono, que se ha utilizado para decodificar una o más tramas de audio que preceden la trama de audio perdida, y en el que la ocultación de error está configurada para proporcionar la información de audio de ocultación de error (242; 482; 612) de acuerdo con dicha información de tono.
12. El decodificador de audio (200; 400) de acuerdo con la reivindicación 11, en el que la ocultación de error (240; 480; 600) está configurada para obtener la información de tono basándose en 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.
13. El decodificador de audio (200; 400) de acuerdo con la reivindicación 12, en el que la ocultación de error está configurada para evaluar una correlación cruzada de la señal de excitación de dominio de tiempo, para determinar una información de tono aproximada; y en el que la ocultación de error está configurada para perfeccionar la información de tono aproximada usando una búsqueda de bucle cerrado alrededor de un tono determinado por la información de tono aproximada.
14. El decodificador de audio de acuerdo con una de las reivindicaciones 1 a 13, en el que la ocultación de error está configurada para obtener una información de tono basándose en una información secundaria de la información de audio codificada.
15. El decodificador de audio de acuerdo con una de las reivindicaciones 1 a 14, en el que la ocultación de error está configurada para obtener una información de tono basándose en una información de tono disponible para una trama de audio decodificada previamente.
16. El decodificador de audio de acuerdo con una de las reivindicaciones 1 a 15, en el que la ocultación de error está configurada para obtener una información de tono basándose en una búsqueda de tono realizada sobre una señal de dominio de tiempo o sobre una señal residual.
17. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 1 a 16, en el que la ocultación de error (240; 480; 600) está configurada para obtener un conjunto de coeficientes de predicción lineal (462, 466), que se han utilizado para decodificar una o más tramas de audio que preceden la trama de audio perdida; y en el que la ocultación de error está configurada para proporcionar la información de audio de ocultación de error (242; 482; 612) dependiendo de dicho conjunto de coeficientes de predicción lineal.
18. El decodificador de audio (200; 400) de acuerdo con la reivindicación 17, en el que la ocultación de error (240; 480; 600) está configurada para extrapolar un nuevo conjunto de coeficientes de predicción lineal basándose en el conjunto de coeficientes de predicción lineal (462, 466), que se han utilizado para decodificar una o más tramas de audio que preceden la trama de audio perdida; y en el que la ocultación de error está configurada para usar el nuevo conjunto de coeficientes de predicción lineal para proporcionar la información de audio de ocultación de error (242; 482; 612).
19. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 1 a 18, en el que la ocultación de error (240; 480; 600) está configurada 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 en el que la ocultación de error está configurada 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 umbral, para decidir si introducir una señal de excitación de dominio de tiempo determinista (652) con la

adición de una señal de excitación de dominio de tiempo de tipo ruido (662), en una síntesis de LPC (680), o si introducir solo una señal de excitación de dominio de tiempo de ruido (662) en la síntesis de LPC.

- 5 20. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 1 a 19, en el que la ocultación de error (240; 480; 600) está configurada para obtener una información de tono que describe un tono de la trama de audio que precede la trama de audio perdida, y para proporcionar la información de audio de ocultación de error (242; 482; 612) de acuerdo con la información de tono.
- 10 21. El decodificador de audio (200; 400) de acuerdo con la reivindicación 20, en el que la ocultación de error (240; 480; 600) está configurada para obtener la información de tono basándose en la señal de excitación de dominio de tiempo (452, 456; 610) asociada con la trama de audio que precede la trama de audio perdida.
- 15 22. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 1 a 21, en el que la ocultación de error (240; 480; 600) está configurada para evaluar una correlación cruzada de la señal de excitación de dominio de tiempo o de una señal de audio de dominio de tiempo (452, 456; 610), para determinar una información de tono aproximada; y en el que la ocultación de error está configurada para perfeccionar la información de tono aproximada usando una búsqueda de bucle cerrado alrededor de un tono determinado por la información de tono aproximada.
- 20 23. El decodificador de audio (200; 400) de acuerdo con la reivindicación 21 o 22, en el que la ocultación de error (240; 480; 600) está configurada para obtener la información de tono para proporcionar la información de audio de ocultación de error (242; 482; 612) 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 de excitación de dominio de tiempo (252, 256; 610), que se modifica para obtener una señal de excitación de dominio de tiempo modificada (652, 672) para proporcionar la información de audio de ocultación de error (242; 482; 612).
- 25 24. El decodificador de audio (200; 400) de acuerdo con la reivindicación 23, en el que la ocultación de error (240; 480; 600) está configurada 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.
- 30 25. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 1 a 24, en el que la ocultación de error (240; 480; 600) está configurada para copiar un ciclo de tono de la señal de excitación de dominio de tiempo (452, 456; 610) asociada con la trama de audio que precede la trama de audio perdida, una vez o múltiples veces, para obtener una señal de excitación (672) para una síntesis (680) de la información de audio de ocultación de error (242; 482; 612).
- 35 26. El decodificador de audio (200; 400) de acuerdo con la reivindicación 25, en el que la ocultación de error (240; 480; 600) está configurada para el filtro paso bajo del ciclo de tono de la señal de excitación de dominio de tiempo (452, 456; 610) 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.
- 40 27. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 1 a 26, en el que la ocultación de error (240; 480; 600) está configurada para predecir un tono al final de una trama perdida; y en el que la ocultación de error está configurada para adaptar la señal de excitación de dominio de tiempo, o una o más copias de la misma, al tono predicho.
- 45 28. El decodificador de audio (200; 400) de acuerdo con una de las reivindicaciones 1 a 27, en el que la ocultación de error (240; 480; 600) está configurada para combinar una señal de excitación de dominio de tiempo extrapolada (652) y una señal de ruido (662), para obtener una señal de entrada (672) para una síntesis de LPC (680); y en el que la ocultación de error está configurada para realizar la síntesis de LPC; en el que la síntesis de LPC está configurada para filtrar la señal de entrada de la síntesis de LPC dependiendo de parámetros de codificación de predicción lineal (462, 466), para obtener la información de audio de ocultación de error.
- 50 29. Un método (1000) para proporcionar una información de audio decodificada basándose en una información de audio codificada, comprendiendo el método:
- 55 proporcionar (1010) una información de audio de ocultación de error para la ocultación de una pérdida de una trama de audio; en el que una señal de excitación de dominio de tiempo obtenida basándose en una o más tramas de audio que
- 60

- preceden una trama de audio perdida se modifica para obtener la información de audio de ocultación de error;
en el que el método comprende modificar una señal de excitación de dominio de tiempo (452, 456; 610) derivada de una o más tramas de audio codificadas en una representación de dominio de frecuencia que preceden una trama de audio perdida, para obtener la información de audio de ocultación de error;
- 5 en el que, para tramas de audio codificadas utilizando la representación de dominio de frecuencia, la información de audio codificada comprende una representación codificada de valores espectrales y factores de escala que representan una escala de diferentes bandas de frecuencia.
- 10 30. Un programa informático para realizar el método de acuerdo con la reivindicación 29 cuando el programa informático se ejecuta en un ordenador.

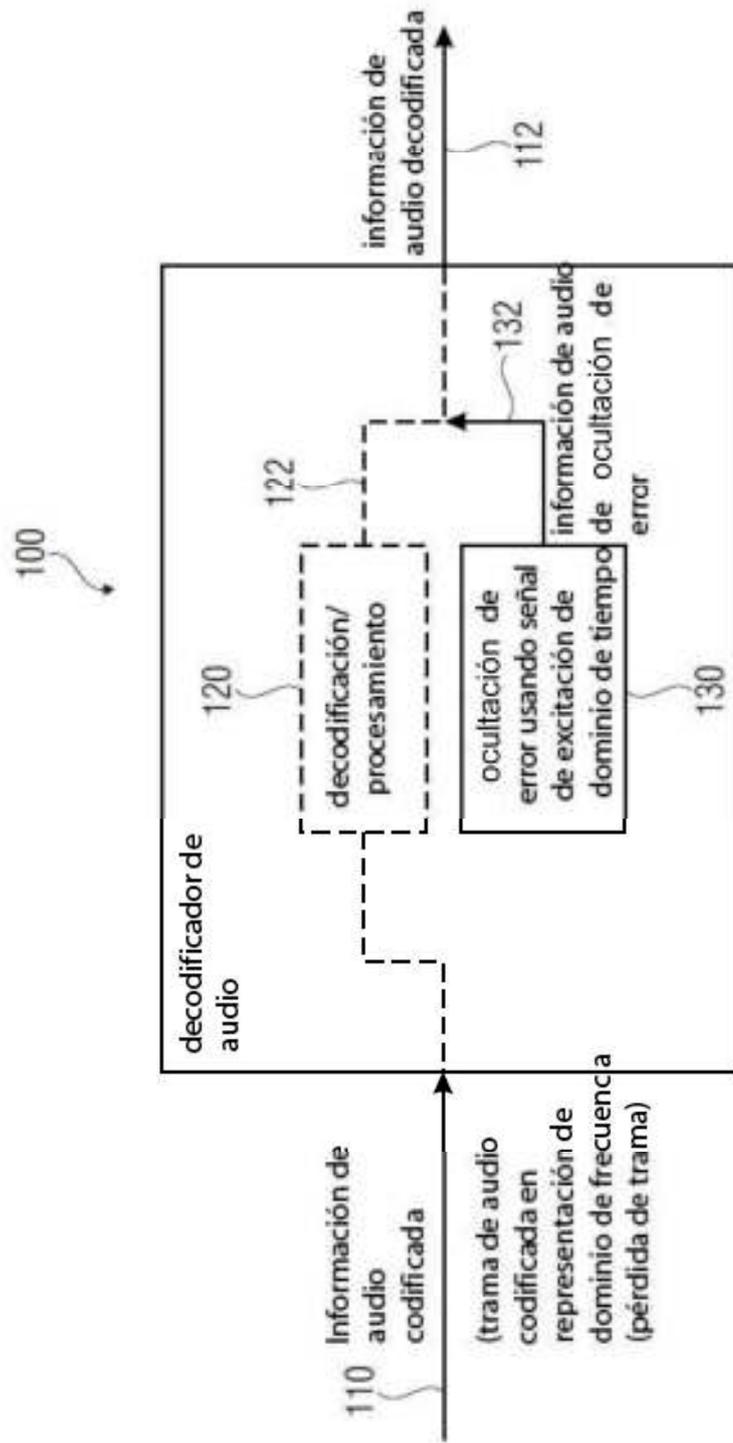


FIG. 1

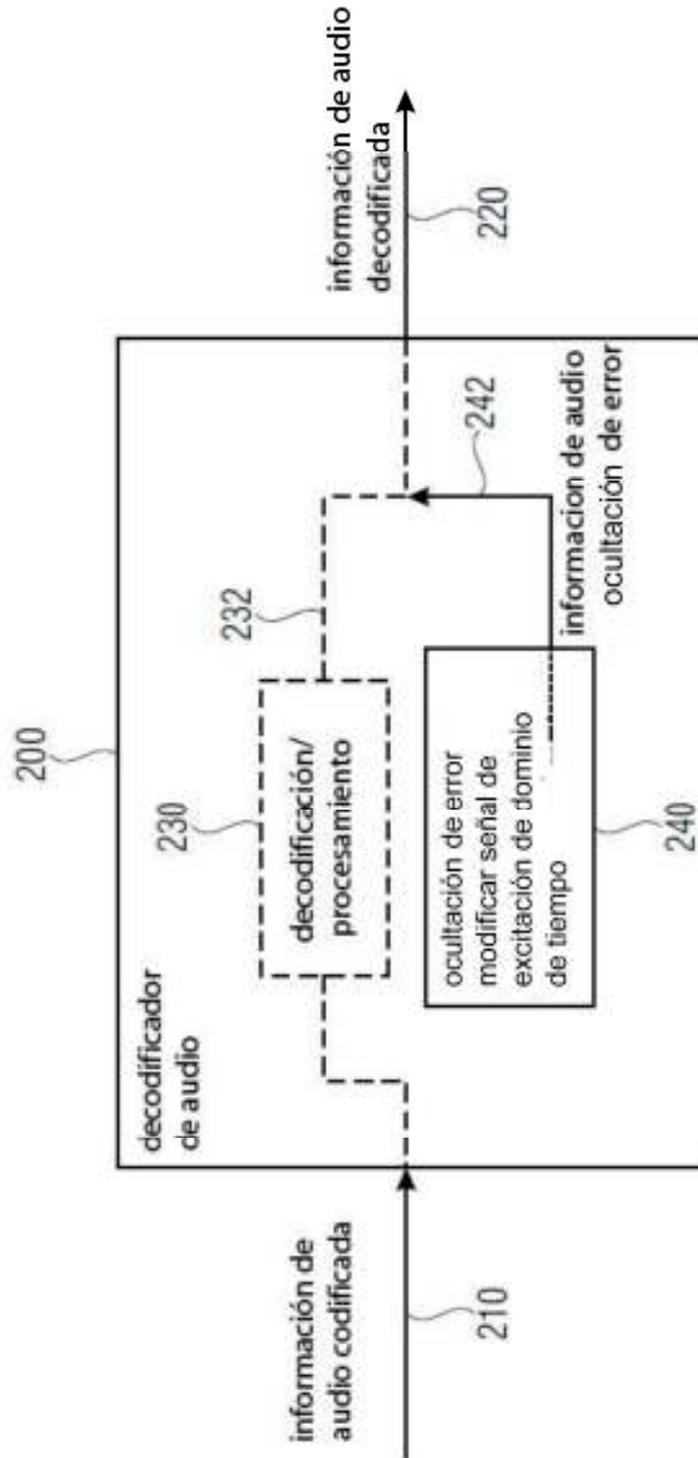


FIG. 2

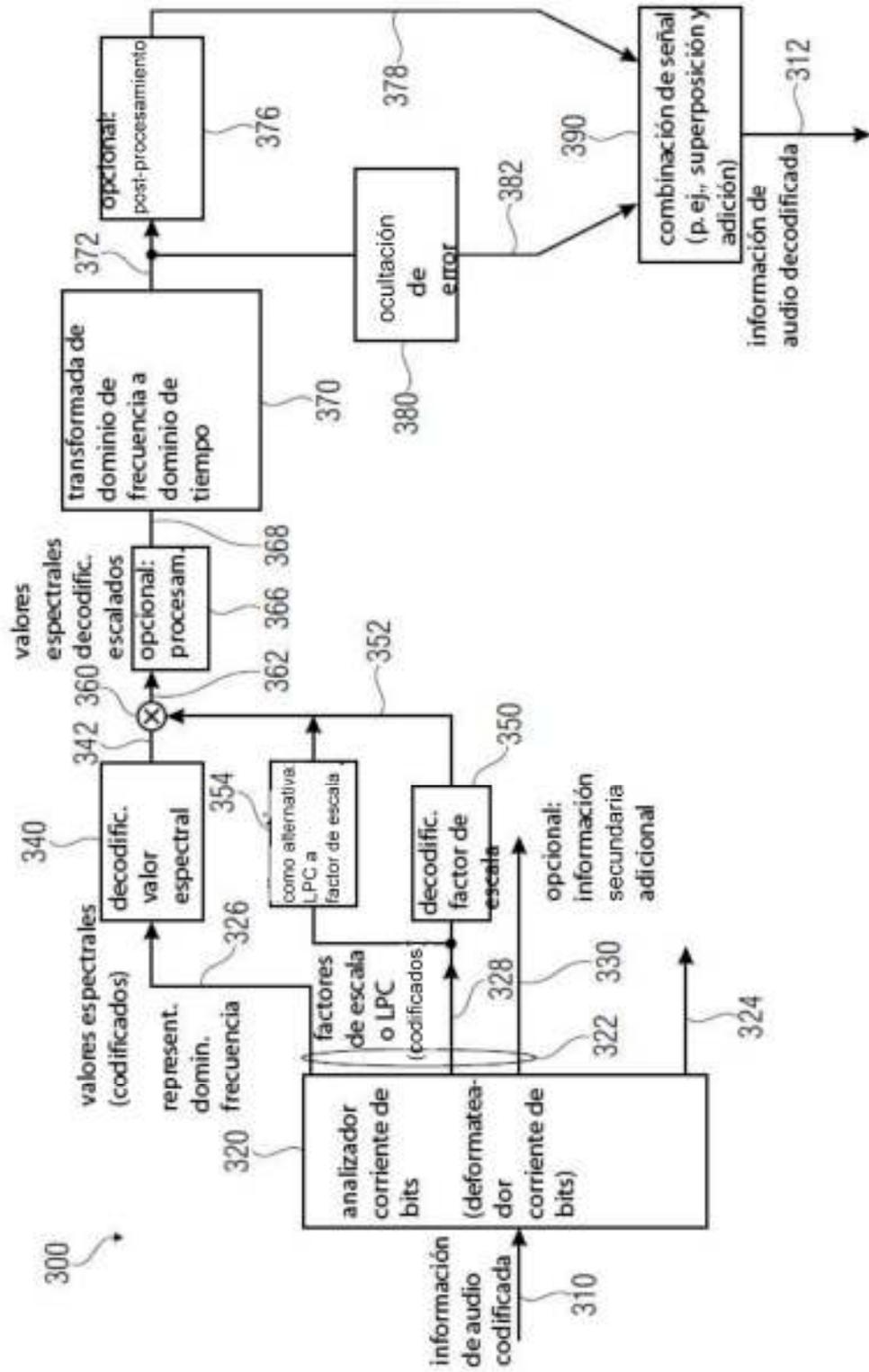


FIG. 3

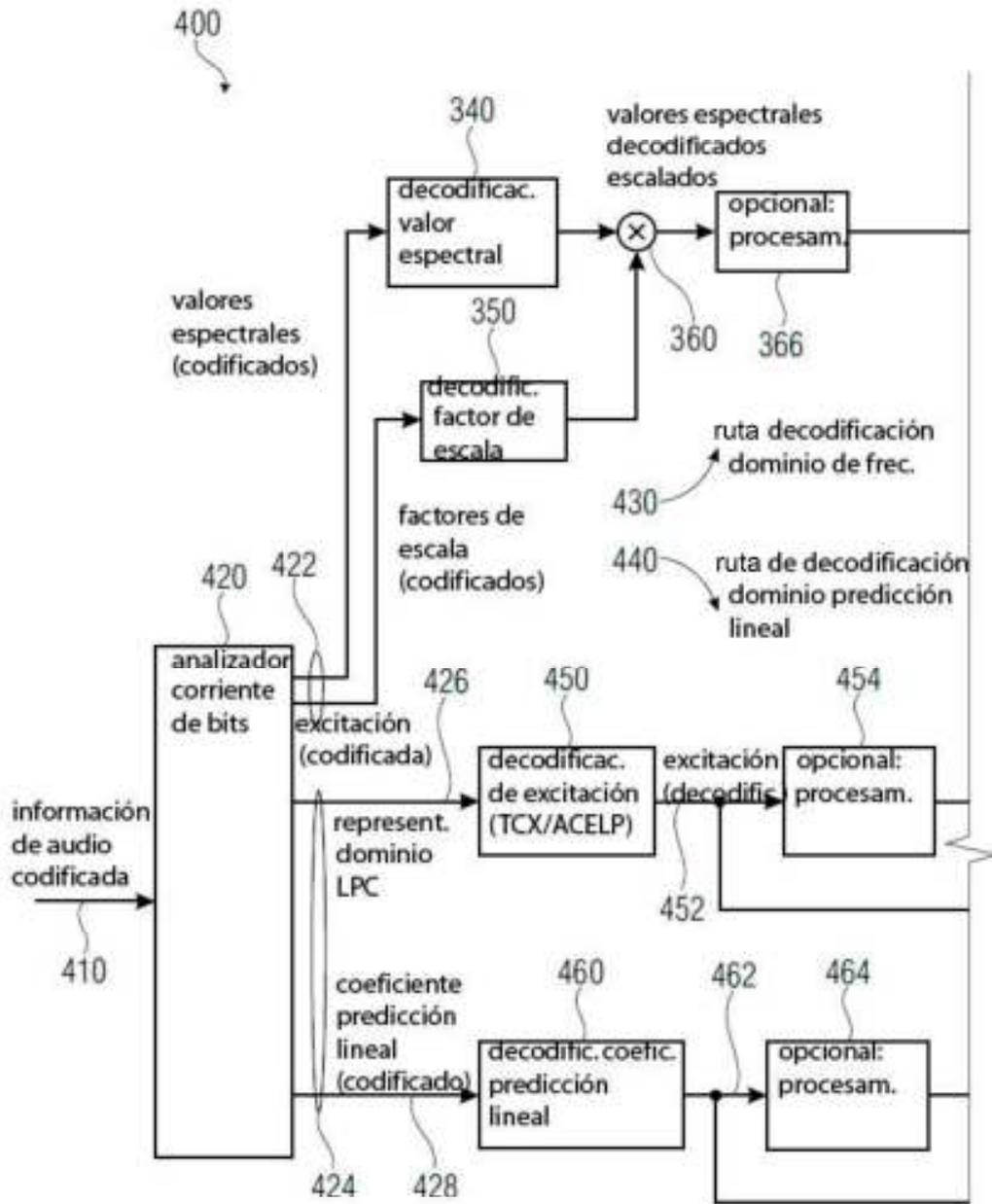


FIG 4	
FIG 4A	FIG 4B

FIG. 4A

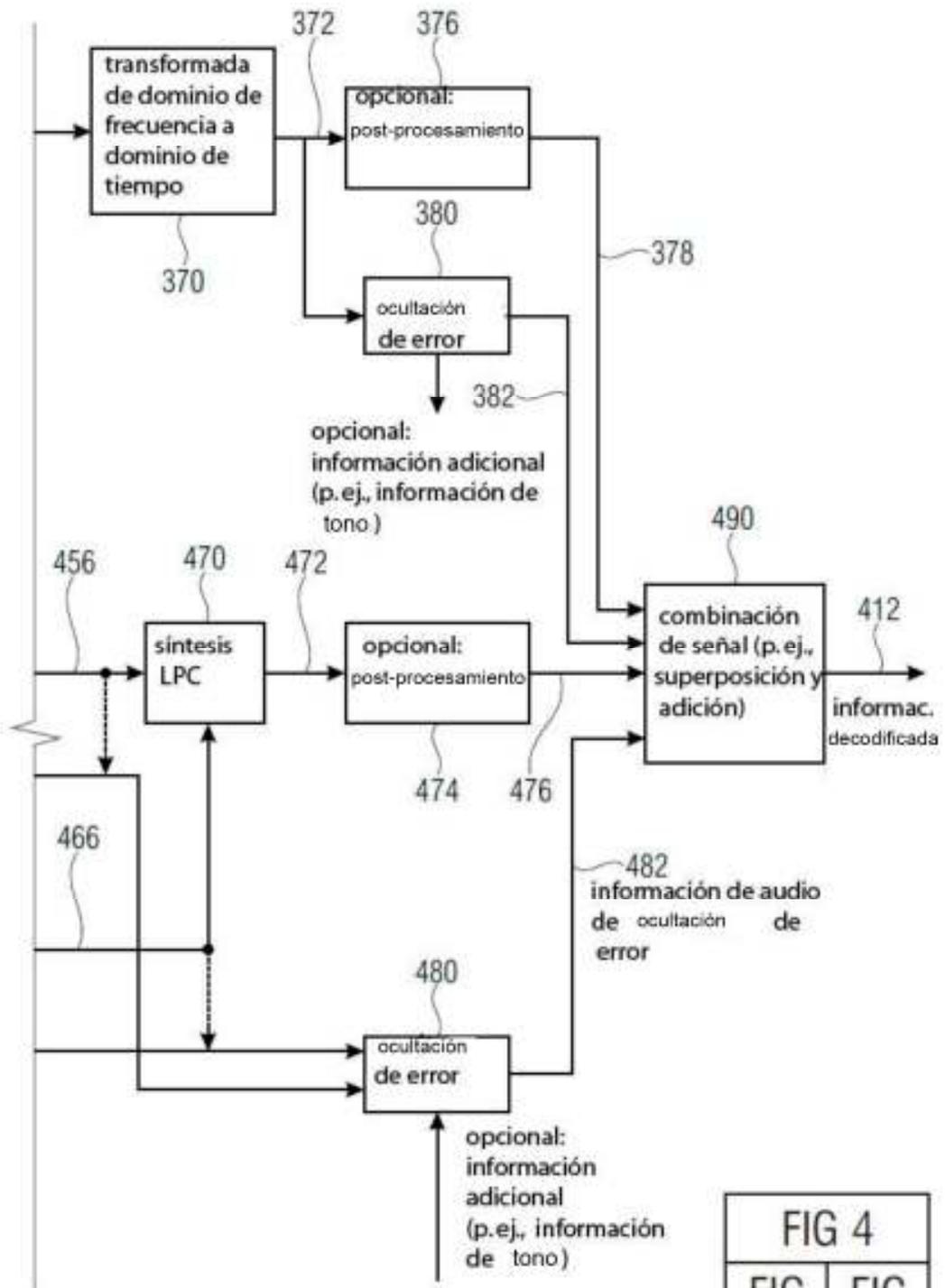
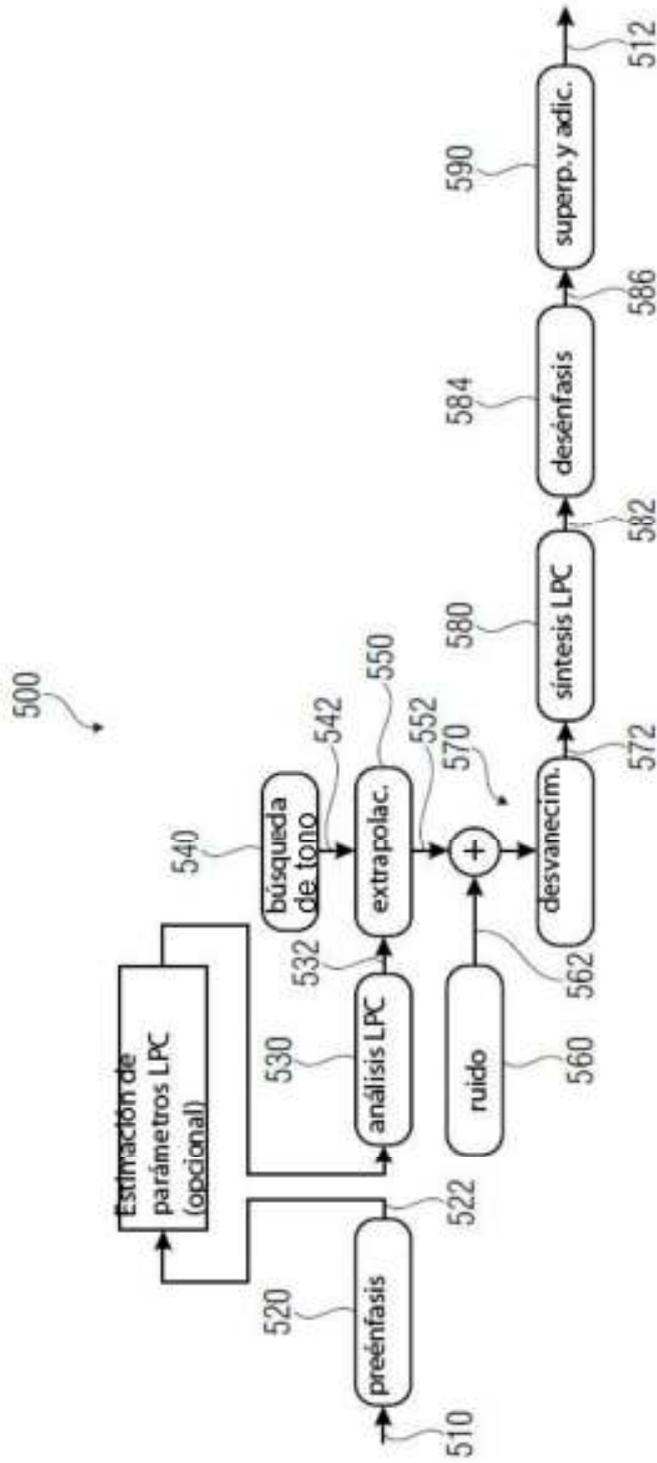


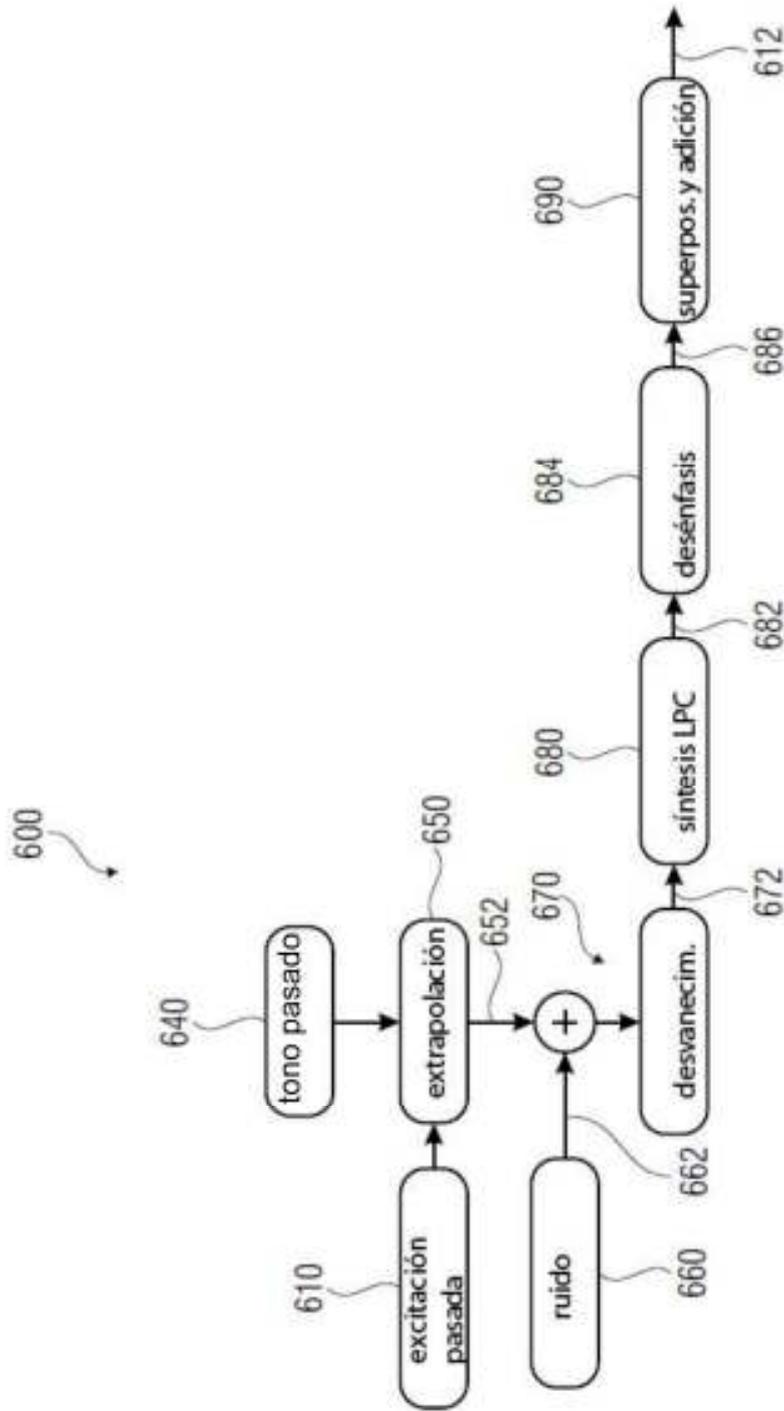
FIG. 4B

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

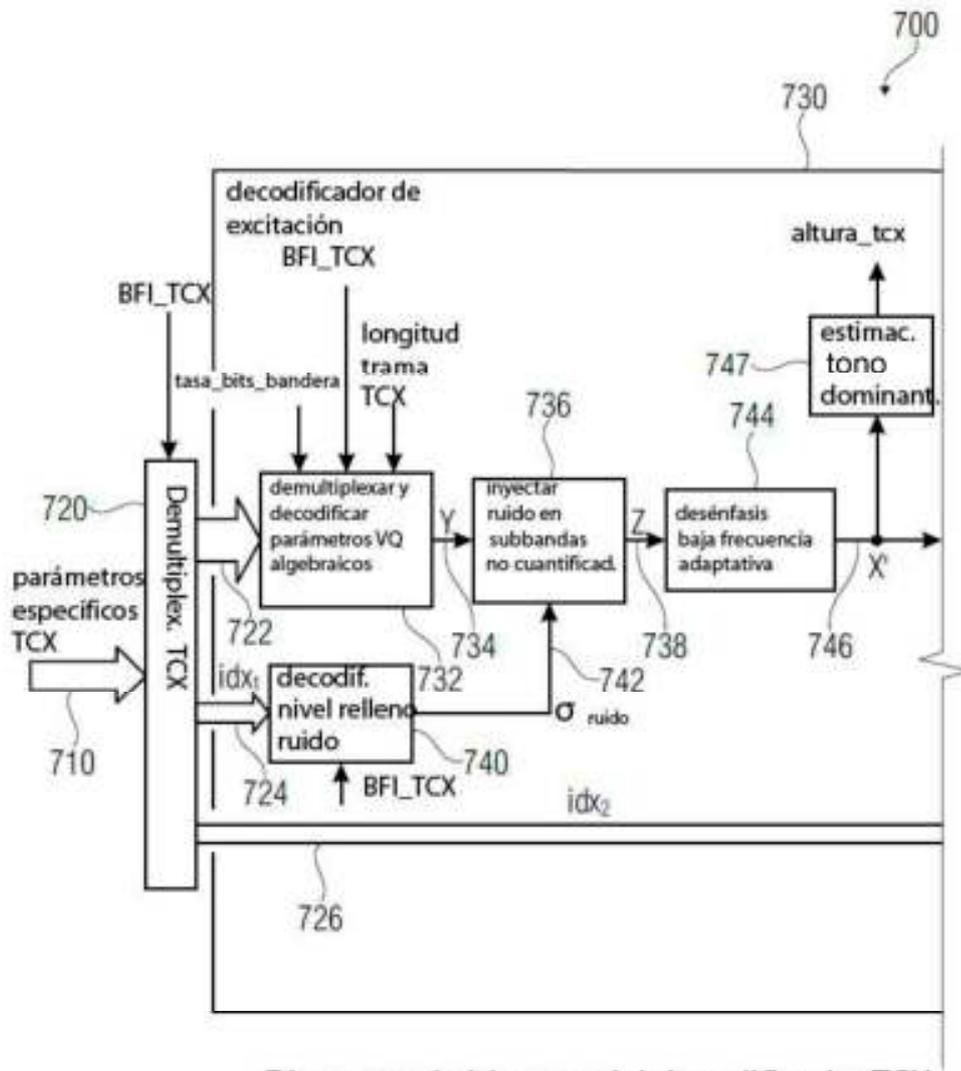


Diagrama de bloques del decodificador TCX

FIG 7	
FIG 7A	FIG 7B

FIG. 7A

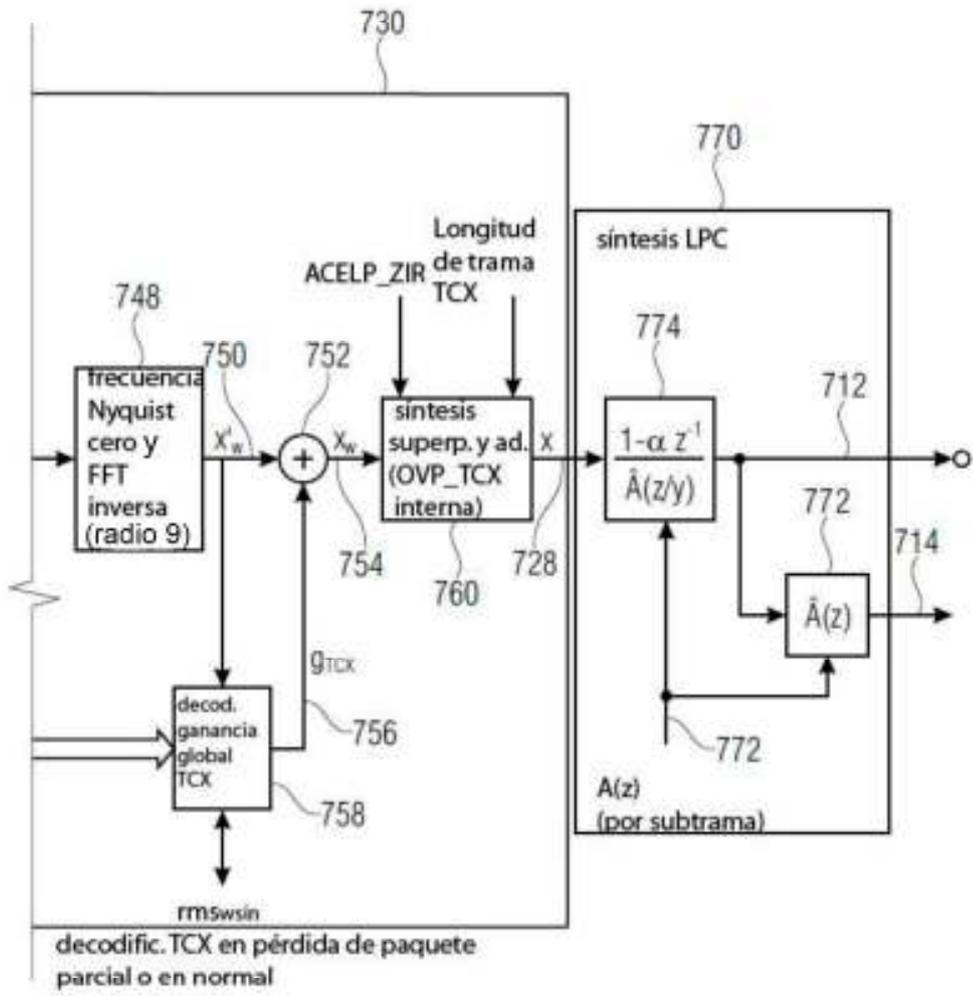
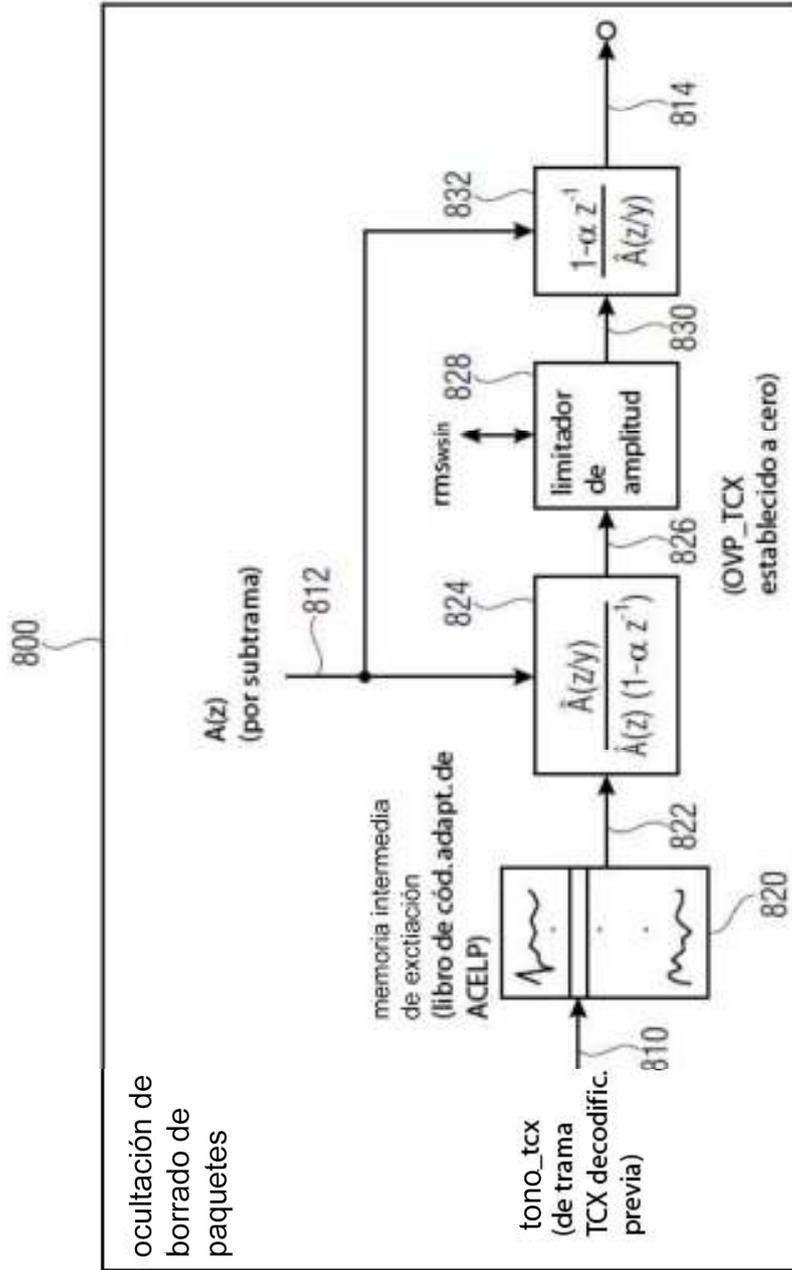


Diagrama de bloques del decodificador TCX

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

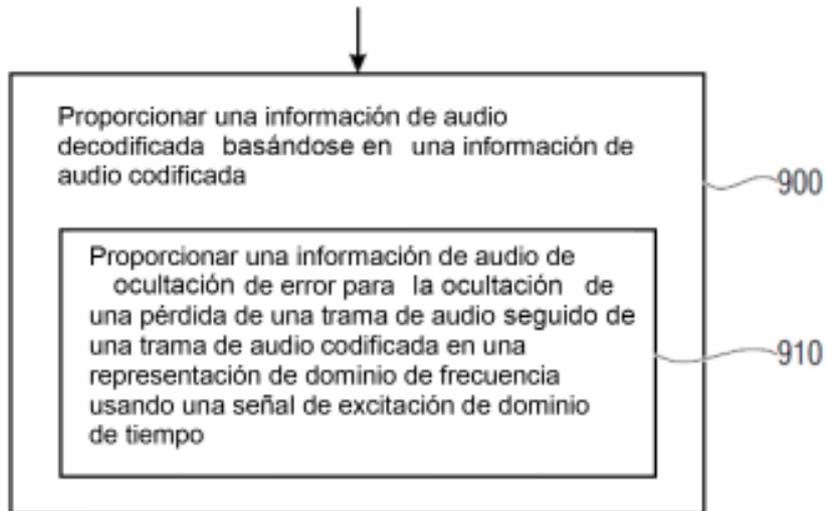


FIG. 9

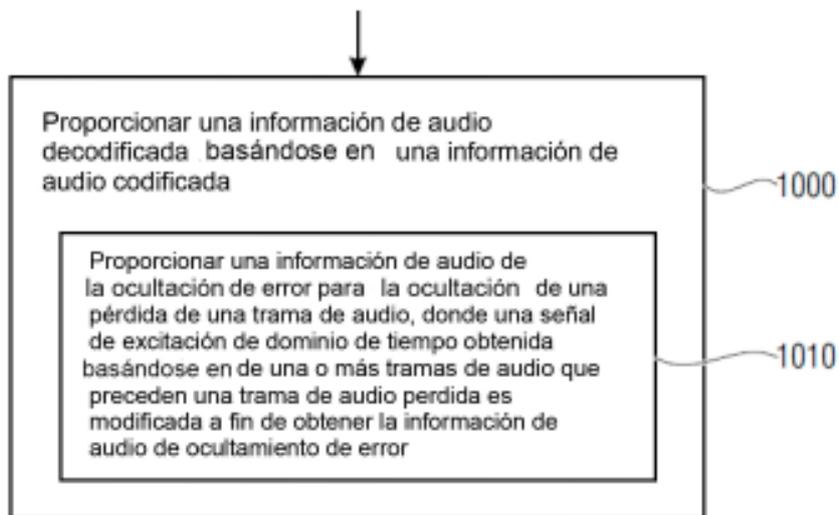


FIG. 10

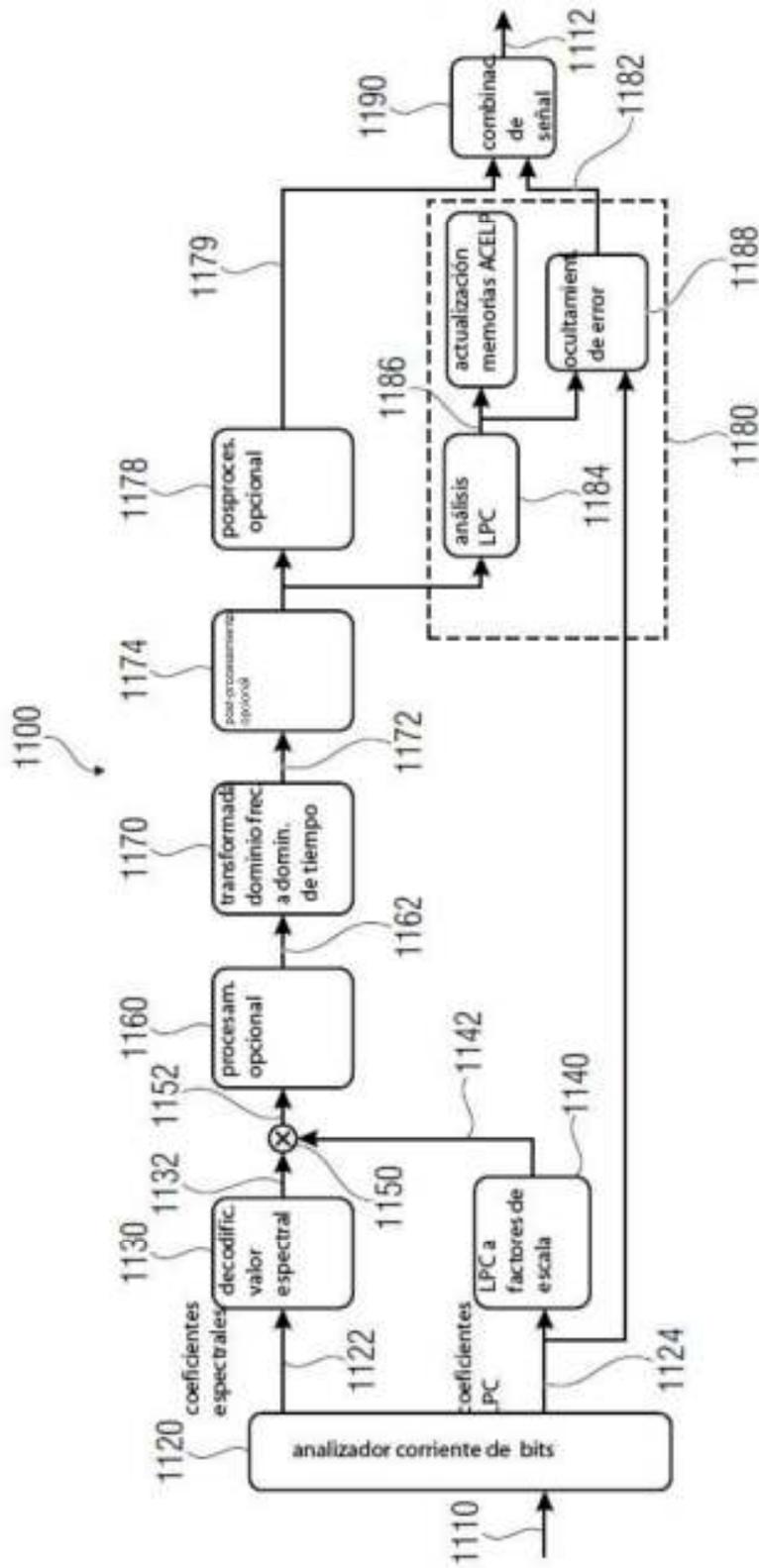


FIG. 11