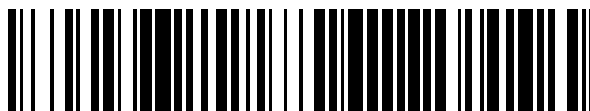


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 739 477**

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

96 Fecha de presentación y número de la solicitud europea: **27.10.2014 E 17191505 (1)**

97 Fecha y número de publicación de la concesión europea: **01.05.2019 EP 3285255**

54 Título: **Decodificador de audio y método para proporcionar una información de audio decodificada usando un ocultamiento de errores en base a una señal de excitación de dominio de tiempo**

30 Prioridad:

31.10.2013 EP 13191133

28.07.2014 EP 14178824

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

31.01.2020

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)**

**Hansastraße 27c
80686 München, DE**

72 Inventor/es:

**LECOMTE, JÉRÉMIE;
MARKOVIC, GORAN;
SCHNABEL, MICHAEL y
PIETRZYK, GRZEGORZ**

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 739 477 T3

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

DESCRIPCIÓN

Decodificador de audio y método para proporcionar una información de audio decodificada usando un ocultamiento de errores en base a una señal de excitación de dominio de tiempo

Campo técnico

Realizaciones de acuerdo con la invención crean decodificadores de audio para proporcionar una información de audio decodificada en base a una información de audio codificada.

Algunas realizaciones de acuerdo con la invención crean métodos para proporcionar una información de audio decodificada en base a una información de audio codificada.

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

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

Antecedentes de la invención

En los últimos años, ha habido una creciente demanda de transmisión y almacenamiento digital de contenidos de audio. Sin embargo, los contenidos de audio con frecuencia son transmitidos sobre canales no fiables, lo que acarrea el riesgo de que 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 una representación de dominio de tiempo codificada o una representación de dominio de frecuencia codificada) se pierdan. En algunas situaciones, sería 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 produciría un retardo sustancial, y, por lo tanto, requerirá un extenso almacenamiento temporal 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 un extenso almacenamiento temporal (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 tener conceptos para manejar una pérdida de una o más tramas de audio. En particular, es deseable tener conceptos que produzcan una buena calidad de audio, o por lo menos, una calidad de audio aceptable, incluso, en el caso de que las tramas de audio se pierdan.

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

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

En el estándar 3gpp TS26.290 se explica una decodificación de excitación codificada transformada (decodificación TCX) con ocultamiento de errores. 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].

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

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

Caso 1 (figura 8): Ocultamiento 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

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

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

Tal como se menciona, la figura 7 muestra un diagrama de bloques de un decodificador TCX que realiza una

decodificación TCX en operación normal o en el caso de pérdida de paquete parcial. El decodificador TCX 700 de acuerdo con la figura 7 recibe parámetros específicos de TCX 710 y proporciona, sobre su base, 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, así como cierta información adicional (por ejemplo, como 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, sobre su base, 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 algebraica. 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 en base a la señal de excitación rellena con ruido 738, para obtener así 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, sobre su base, 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 ventaneado) en base a las señales de excitación de dominio de tiempo escaladas 754, para obtener así una señal de excitación de dominio de tiempo temporalmente combinada 728 durante un período más largo en el tiempo (más largo que los períodos en el tiempo para los cuales se proporcionan las señales de excitación de dominio de tiempo individuales 750, 754).
- 40 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 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 así la señal de audio decodificada 714.

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

- El ocultamiento 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, el ocultamiento 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, el ocultamiento de borrado de paquetes 800 puede configurarse para proporcionar, en base a la información de tono 810 y los parámetros de LPC 812, una señal de ocultamiento de errores 814, que puede considerarse una información de audio de ocultamiento de errores. El ocultamiento de borrado de paquetes 800 comprende una memoria intermedia de excitación 820 que, por ejemplo, puede almacenar temporalmente una excitación previa. La memoria intermedia de excitación 820, por ejemplo, puede hacer uso del libro de códigos adaptativo de ACELP, y puede proporcionar una señal de excitación 822. El ocultamiento de borrado de paquetes 800 puede comprender adicionalmente un primer filtro 824, una función de filtro que puede definirse como se muestra en la figura 8. Por lo

tanto, el primer filtro 824 puede filtrar la señal de excitación 822 en base a los parámetros de LPC 812, para obtener así una versión filtrada 826 de la señal de excitación 822. El ocultamiento de borrado de paquetes además comprende un limitador de amplitud 828, que puede limitar una amplitud de la señal de excitación filtrada 826 en base a información objetivo o información de nivel rms_{wsyn} . Además, el ocultamiento de borrado de paquetes 800 puede comprender un segundo filtro 832, que puede estar configurado para recibir la señal de excitación filtrada de limitada amplitud 830 desde el limitador de amplitud 822 y para proporcionar, sobre su base, la señal de ocultamiento de errores 814. Una función de filtro del segundo filtro 832, por ejemplo, puede definirse como se muestra en la figura 8.

10 A continuación, se describirán algunos detalles respecto a la decodificación y al ocultamiento de errores.

En el Caso 1 (ocultamiento de borrado de paquetes en TCX-256), no hay información disponible para la decodificación de la trama TCX de 256 muestras. La síntesis de TCX se halla mediante el procesamiento de la excitación pasada retardada por T , donde $T=tono_tcx$ es una demora de tono estimada en la trama TCX

15 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 clics en la síntesis. Este filtro se descompone en 3 etapas.

Etapas 1: filtración mediante:

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

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

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

Etapas 3: filtración mediante:

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

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

Decodificación de los parámetros VQ algebraicos.

30 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 5.3.5.7 de 3gpp TS 26.290. Recuérdese que X' tiene dimensión N , donde $N = 288, 576$ y 1152 para TCX-256, 512 y 1024, respectivamente, y que cada bloque B'_k tiene dimensión 8. El número K de bloques B'_k es, por lo tanto, 36, 72 y 144 para TCX-256, 512 y 1024, respectivamente. Los parámetros VQ algebraicos para cada bloque B'_k se describen en la Etapa 5 de la Sección 5.3.5.7. Para cada bloque B'_k , tres grupos de índices binarios son enviados por el codificador:

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

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

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

50 A continuación, a partir del factor de escala M , el Vector de extensión de Voronoi z (un punto de rejilla en RE_8) y el punto de rejilla c en el libro de códigos base (también, un punto de rejilla en RE_8), cada bloque escalado cuantificado $\hat{B}'_k$ puede computarse como:

$$\hat{B}'_k = Mc + z$$

Cuando no hay extensión de Voronoi (es decir, $n_k < 5$, $M=1$ y $z=0$), 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

- 5 vector 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 usa 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.

Estimación del valor de tono dominante.

10 La estimación del tono dominante se realiza de modo tal que la siguiente trama por ser decodificada puede ser apropiadamente extrapolada si corresponde a TCX-256, y si el paquete relacionado está perdido. Esta estimación se basa en la asunción de que el pico de máxima magnitud en el espectro del objetivo TCX corresponde al tono dominante. La búsqueda de la máxima M es restringida a una frecuencia inferior a $F_s/64$ kHz

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

y el índice mínimo $1 \leq i_{\max} \leq N/32$ de modo tal que también se encuentra $(X'_{2i})^2 + (X'_{2i+1})^2 = M$. 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).
20 Recuérdese que el tono dominante se calcula para el ocultamiento de borrado de paquetes en TCX-256. Para evitar problemas de almacenamiento temporal (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 período de tono múltiple en 256 muestras estableciendo *tono_tcx* en

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

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

A continuación, se describirán brevemente algunos conceptos convencionales adicionales.

30 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 (conferir, 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 ocultamiento que incrementa el retardo del decodificador por una trama.

40 En la Patente Europea EP 1207519 B1 (referencia [5]), se describe proporcionar un decodificador de voz y un método de compensación de errores capaz de 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.
45 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 son detectados con un error. Se conoce de acuerdo con la solicitud de patente internacional WO 2005/078706 un método para el
50 ocultamiento de error de trama en un modo de codificación TCX usando una señal de excitación amortiguada.

En vista de la técnica anterior, existe una necesidad de una mejora adicional del ocultamiento de error que proporciona mejor impresión de audición.

55 3. Sumario de la invención

La invención proporciona decodificadores de audio de acuerdo con las reivindicaciones 1 y 2, procedimientos de acuerdo con las reivindicaciones 3 y 4, y un programa de ordenador de acuerdo con la reivindicación 5.

60 Todas las siguientes apariciones de la palabra "realización(es)", si se refiere a combinaciones de características que no comprenden todas las características definidas en las reivindicaciones independientes, se refieren a ejemplos que se presentaron originalmente pero que no representan realizaciones de la presente invención reivindicada.

Una realización de acuerdo con la invención crea un decodificador de audio para proporcionar una información de audio decodificada en base a una información de audio codificada. El decodificador de audio comprende un ocultamiento de errores configurado para proporcionar una información de audio de ocultamiento de errores para el

ocultamiento de una pérdida de una trama de audio (o más de una pérdida de trama) luego de una trama de audio codificada en una representación de dominio de frecuencia, usando una señal de excitación de dominio de tiempo. Esta realización de acuerdo con la invención se basa en el hallazgo de que puede obtenerse un ocultamiento de errores mejorado al proporcionar la información de audio de ocultamiento de errores en base a una señal de excitación de dominio de tiempo, incluso, si la trama de audio que precede una trama de audio perdida está codificada en una representación de dominio de frecuencia. En otras palabras, se ha reconocido que una calidad de un ocultamiento de errores típicamente es mejor si el ocultamiento de errores se realiza en base a una señal de excitación de dominio de tiempo, cuando se compara con un ocultamiento de errores realizado en un dominio de frecuencia, de modo tal que vale la pena la conmutación a un ocultamiento de errores 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 un buen ocultamiento de errores, 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 (Forma de Ruido de Dominio de frecuencia). Sin embargo, se ha hallado que es conveniente derivar la señal de excitación de dominio de tiempo (que puede servir como una excitación para una síntesis de LPC) incluso si la trama de audio que precede la trama de audio perdida es originalmente codificada en la representación de dominio de frecuencia que comprende información sustancialmente diferente (a decir, una representación codificada de una pluralidad de valores espectrales en una representación codificada de una pluralidad de factores de escala para la escala de los valores espectrales). Por ejemplo, en el caso de TCX, no enviamos factores de escala (desde un codificador hacia un decodificador), sino LPC, y luego, en el decodificador, transformamos la LPC en una representación de factor de escala para los bins de 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 una realización preferida, el decodificador de audio comprende un núcleo de decodificador de dominio de frecuencia configurado para la aplicación de una escala en base a factores de escala, a una pluralidad de valores espectrales derivados de la representación de dominio de frecuencia. En este caso, el ocultamiento de errores está configurado para proporcionar la información de audio de ocultamiento de errores para el ocultamiento de una pérdida de una trama de audio después 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 se basa en el hallazgo de que la derivación de la señal de excitación de dominio de tiempo a partir de la arriba mencionada representación de dominio de frecuencia típicamente proporciona un mejor resultado de ocultamiento de errores en comparación con un ocultamiento de errores realizado directamente en el dominio de frecuencia. Por ejemplo, la señal de excitación se crea en base a 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 ser transmitidos como coeficientes de LPC, por ejemplo, usando una representación polinomial que es luego convertida en factores de escala del lado del decodificador.

En una realización preferida, el decodificador de audio comprende un núcleo de decodificador de dominio de frecuencia configurado para derivar una representación de señal de audio de dominio de tiempo a partir de la representación de dominio de frecuencia sin el uso de una señal de excitación de dominio de tiempo como una cantidad intermedia para la trama de audio codificada en la representación de dominio de frecuencia. En otras palabras, se ha hallado que el uso de una señal de excitación de dominio de tiempo para un ocultamiento de errores es conveniente incluso si la trama de audio que precede la trama de audio perdida está codificada en un modo de frecuencia "real" que no usa ninguna señal de excitación de dominio de tiempo como una cantidad intermedia (y que, en consecuencia, no se basa en una síntesis de LPC).

En una realización preferida, el ocultamiento de errores está configurado para obtener la señal de excitación de

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

En una realización preferida, el ocultamiento de errores está configurado para realizar un análisis de LPC en base a 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 es conveniente 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, aun si la trama de audio que precede la trama de audio perdida está codificada en una representación de dominio de frecuencia (que no contiene ningún parámetro de codificación de predicción lineal y ninguna representación de una señal de excitación de dominio de tiempo), debido a que puede obtenerse una información de audio de ocultamiento de errores de buena calidad para muchas señales de audio de entrada en base a dicha señal de excitación de dominio de tiempo. Alternativamente, el ocultamiento de errores puede ser configurado para realizar un análisis de LPC en base a la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida, para obtener la señal de excitación de dominio de tiempo que representa un contenido de audio de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida. Además, alternativamente, el decodificador de audio puede estar configurado para obtener un conjunto de parámetros de codificación de predicción lineal usando una estimación de parámetros de codificación de predicción lineal, o el decodificador de audio puede estar configurado para obtener un conjunto de parámetros de codificación de predicción lineal en base a 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 el ventaneado/autocorr/levinson durbin en base a 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, el ocultamiento de errores está configurado 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 ocultamiento de errores de acuerdo con la información de tono. Al considerar la información de tono, puede lograrse que la información de audio de ocultamiento de errores (que típicamente es una señal de audio de ocultamiento de errores que cubre la duración temporal de por lo menos una trama de audio perdida) se adapte bien al contenido de audio real.

En una realización preferida, el ocultamiento de errores está configurado para obtener la información de tono en base a 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 conveniente 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, el ocultamiento de errores está configurado para evaluar una correlación cruzada de la señal de excitación de dominio de tiempo, para determinar una información de tono aproximada. Además, el ocultamiento de errores puede estar configurado para refinar la información de tono aproximada usando una búsqueda de circuito 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, el ocultamiento de errores del decodificador de audio puede estar configurado para obtener una información de tono en base a una información lateral de la información de audio codificada.

En una realización preferida, el ocultamiento de errores puede estar configurado para obtener una información de tono en base a una información de tono disponible para una trama de audio previamente decodificada.

En una realización preferida, el ocultamiento de errores está configurado para obtener una información de tono en base a una búsqueda de tono realizada sobre una señal de dominio de tiempo o sobre una señal residual.

En otras palabras, el tono puede ser transmitido como información lateral o podría además provenir de la trama previa si hay, por ejemplo, LTP. La información de tono además podría ser transmitida 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, el ocultamiento de errores está configurado 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 ocultamiento de errores. Mediante el copiado de la señal de excitación de dominio de tiempo una vez o múltiples veces, puede lograrse que el componente determinista (es decir, sustancialmente periódico) de la información de ocultamiento de errores 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, el ocultamiento de errores está configurado 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 ser adaptada para un ancho de banda de audio disponible, lo que produce una buena impresión auditiva de la información de audio de ocultamiento de errores. 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 el filtrado 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, el ocultamiento de errores está configurado 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 ser considerados. En consecuencia, se evitan fallos en una transición entre la información de audio de ocultamiento de errores y una información de audio de una trama apropiadamente decodificada luego de una o más tramas de audio perdidas (o por lo menos se reducen, ya que es solo una trama predicha, no la real). Por ejemplo, la adaptación va desde el último buen tono hasta el predicho. Esto se realiza por medio de la resincronización de pulsos [7].

En una realización preferida, el ocultamiento de errores está configurado 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, el ocultamiento de errores está configurado 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 ocultamiento de errores. En consecuencia, tanto un componente determinista (por ejemplo, aproximadamente periódico) del contenido de audio como un componente de tipo ruido del contenido de audio pueden ser considerados. Por lo tanto, se logra que la información de audio de ocultamiento de errores comprenda una impresión auditiva "natural".

En una realización preferida, el ocultamiento de errores está configurado para computar una ganancia de la señal de excitación de dominio de tiempo extrapolada, que se usa para obtener la señal de entrada para la síntesis de LPC, usando una correlación en el dominio de tiempo que se realiza en base a 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 obtenida en base a la señal de excitación de dominio de tiempo. En otras palabras, una intensidad de un componente periódico se determina dentro de la trama de audio que precede la trama de audio perdida, y esta intensidad determinada del componente periódico se usa para obtener la información de audio de ocultamiento de errores. Sin embargo, se ha hallado que la computación citada anteriormente de la intensidad del componente periódico proporciona resultados particularmente buenos, ya que se considera la señal de audio de dominio de tiempo real de la trama de audio que precede la trama de audio perdida. Alternativamente, puede obtenerse una correlación en el dominio de excitación o directamente en el dominio de tiempo para obtener la información de tono. Sin embargo, hay también diferentes posibilidades, de acuerdo con 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 lateral o el calculado.

En una realización preferida, el ocultamiento de errores está configurado para el filtro de paso alto de la señal de ruido que está combinada con la señal de excitación de dominio de tiempo extrapolada. Se ha hallado que la filtración de paso alto de la señal de ruido (que típicamente es ingresada 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, luego 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 luego de una cierta cantidad de pérdida de trama, ya no hay filtración, para solo obtener el ruido de forma de banda completa para obtener así un buen ruido de confort cercano al ruido de fondo.

En una realización preferida, el ocultamiento de errores está configurado para cambiar selectivamente la forma espectral de la señal de ruido (562) usando el filtro de énfasis previo en el que 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 encontrado que la impresión auditiva de la información de audio de ocultamiento de errores puede ser mejorada por dicho concepto. Por ejemplo, en algún caso, es mejor disminuir las ganancias y la forma, y en algún lugar, es mejor incrementarlas.

En una realización preferida, el ocultamiento de errores está configurado para computar una ganancia de la señal de ruido de acuerdo con una correlación en el dominio de tiempo, que se realiza en base a 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 trama de audio que precede la trama de audio perdida. Usando este concepto, es posible poder obtener una energía de la trama ocultada cercana a la energía de la trama buena previa. Por ejemplo, la ganancia para la señal de ruido puede ser generada midiendo la energía del resultado: excitación de señal de entrada-excitación en base a tono generada.

En una realización preferida, el ocultamiento de errores está configurado para modificar una señal de excitación de dominio de tiempo obtenida en base a una o más tramas de audio que preceden una trama de audio perdida, para obtener la información de audio de ocultamiento de errores. 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 "atenuación" del componente determinista (por ejemplo, sustancialmente periódico) del contenido de audio en la información de audio de ocultamiento de errores. 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 ocultamiento de errores en función del tiempo.

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

En una realización preferida, el ocultamiento de errores está configurado para modificar la señal de excitación de dominio de tiempo obtenida en base a una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, para reducir un componente periódico de la información de audio de ocultamiento de errores 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 ocultamiento de errores.

En una realización preferida, el ocultamiento de errores está configurado para escalar la señal de excitación de dominio de tiempo obtenida en base a 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 proporciona una buena información de audio de ocultamiento de errores.

En una realización preferida, el ocultamiento de errores está configurado para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida en base a una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias. Por consiguiente, puede lograrse una atenuación del componente periódico dentro de la información de audio de ocultamiento de errores.

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

acuerdo con uno o más parámetros de una o más tramas de audio que preceden la trama de audio perdida, y/o de acuerdo con una cantidad de tramas de audio perdidas consecutivas. Por consiguiente, es posible ajustar la velocidad a la cual el componente determinista (por ejemplo, por lo menos aproximadamente periódico) se atenúa en la información de audio de ocultamiento de errores. La velocidad de atenuación 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. Alternativamente, o además, la cantidad de tramas de audio perdidas consecutivas puede ser considerada cuando se determina la velocidad utilizada para la atenuación del componente determinista (por ejemplo, por lo menos aproximadamente periódico) de la información de audio de ocultamiento de errores, que ayuda a adaptar el ocultamiento de errores a la situación específica. Por ejemplo, la ganancia de la parte tonal y la ganancia de la parte ruidosa pueden atenuarse de manera separada. La ganancia para la parte tonal puede converger a cero luego 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.

En una realización preferida, el ocultamiento de errores está configurado para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida en base a una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, de acuerdo con una longitud de un período de tono de la señal de excitación de dominio de tiempo, de modo que una entrada de señal de excitación de dominio de tiempo en una síntesis de LPC se atenúa de forma más rápida para señales que tienen una longitud más corta del período de tono en comparación con señales que tienen una mayor longitud del período de tono. Por consiguiente, puede evitarse que las señales que tienen una longitud más corta del período de tono sean repetidas con demasiada frecuencia con alta intensidad, ya que esto típicamente resultará en una impresión auditiva no natural. En consecuencia, una calidad general de la información de audio de ocultamiento de errores puede ser mejorada.

En una realización preferida, el ocultamiento de errores está configurado para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida en base a una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, de acuerdo con un resultado de un análisis de tono o una predicción de tono, de modo tal que un componente determinista de la entrada de señal de excitación de dominio de tiempo en una síntesis de LPC se atenúa 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 atenúa 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, la atenuación 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 la atenuación de un componente determinista más rápidamente para señales que comprenden una incertidumbre comparativamente grande del tono, pueden evitarse fallos audibles, o al menos pueden reducirse sustancialmente.

En una realización preferida, el ocultamiento de errores está configurado para la escala de tiempo de la señal de excitación de dominio de tiempo obtenida en base a una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, de acuerdo con 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 ocultamiento de errores comprende una impresión auditiva más natural.

En una realización preferida, el ocultamiento de errores está configurado para proporcionar la información de audio de ocultamiento de errores 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 en base a la información de audio de ocultamiento de errores, que ayuda a reducir fallos de bloqueo.

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

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

Otra realización de acuerdo con la invención crea un método para proporcionar una información de audio

decodificada en base a una información de audio codificada. El método comprende proporcionar una información de audio de ocultamiento de errores para el ocultamiento de una pérdida de una trama de audio después de una trama de audio codificada en una representación de dominio de frecuencia usando una señal de excitación de dominio de tiempo. Este método se basa en las mismas consideraciones que el decodificador de audio citado anteriormente.

5 Aun otra realización de acuerdo con la invención, se crea un programa de ordenador para la realización de dicho método cuando el programa de ordenador se ejecuta en un ordenador.

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

15 Esta realización de acuerdo con la invención se basa en la idea de que un ocultamiento de errores con una buena calidad de audio puede obtenerse en base a una señal de excitación de dominio de tiempo, en el que una modificación de la señal de excitación de dominio de tiempo obtenida en base a una o más tramas de audio que preceden una trama de audio perdida permite una adaptación de la información de audio de ocultamiento de errores a cambios esperados (o predichos) del contenido de audio durante la trama perdida. Por consiguiente, pueden evitarse fallos, y en particular, una impresión auditiva no natural, que serían causados por un uso no cambiado de la señal de excitación de dominio de tiempo. En consecuencia, se logra una provisión mejorada de una información de audio de ocultamiento de errores, de modo tal que las tramas de audio perdidas pueden ocultarse con resultados mejorados.

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

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

30 En una realización preferida, el ocultamiento de errores está configurado para escalar la señal de excitación de dominio de tiempo obtenida en base a una o más tramas de audio que preceden la trama de audio perdida, o una o más de sus copias, para modificar así la señal de excitación de dominio de tiempo. El escalado de la señal de excitación de dominio de tiempo constituye una manera particularmente eficiente para variar la información de audio de ocultamiento de errores en función del tiempo.

35 En una realización preferida, el ocultamiento de errores está configurado para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias. Se ha hallado que la reducción gradual de la ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, permite obtener una señal de excitación de dominio de tiempo para la provisión de la información de audio de ocultamiento de errores, de modo tal que los componentes deterministas (por ejemplo, componentes al menos aproximadamente periódicos) se atenúan. Por ejemplo, puede no haber solo una ganancia. Por ejemplo, se podría tener una ganancia para la parte tonal (también referida como la parte aproximadamente periódica), y una ganancia para la parte de ruido. Ambas excitaciones (o componentes de excitación) pueden ser atenuadas separadamente con diferente factor de velocidad, y luego, las dos excitaciones resultantes (o componentes de excitación) pueden combinarse antes de alimentarse a la LPC para síntesis. En el caso de no tener ninguna estimación de ruido de fondo, los factores de atenuación para el ruido y para la parte tonal pueden ser similares, y entonces, se podría tener solo una aplicación de atenuación sobre los resultados de las dos excitaciones, multiplicada con su propia ganancia y combinados entre sí.

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

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

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

En una realización preferida, el ocultamiento de errores está configurado para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación de dominio de tiempo obtenida para una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, de acuerdo con un resultado de un análisis de tono o una predicción de tono, de modo tal que un componente determinista de una entrada de señal de excitación de dominio de tiempo en una síntesis de LPC se atenúa más rápidamente para señales que tienen un mayor cambio de 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 entrada de señal de excitación de dominio de tiempo en una síntesis de LPC se atenúa 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. Por consiguiente, un componente determinista (por ejemplo, por lo menos aproximadamente periódico) se atenúa más rápidamente para señales para las cuales hay una mayor incertidumbre del tono (donde un mayor cambio de tono por unidad de tiempo, o incluso, un fallo de la predicción de tono, indica una incertidumbre comparativamente grande del tono). En consecuencia, pueden evitarse los fallos, que surgirían de la provisión de una información de audio de ocultamiento de errores altamente determinista en una situación en la cual el tono real es incierto.

En una realización preferida, el ocultamiento de errores está configurado para la escala de tiempo de la señal de excitación de dominio de tiempo obtenida para (o en base a) una o más tramas de audio que preceden una trama de audio perdida, o una o más de sus copias, de acuerdo con una predicción de un tono para el tiempo de las una o más tramas de audio perdidas. Por consiguiente, la señal de excitación de dominio de tiempo, que se usa para la provisión de la información de audio de ocultamiento de errores, es modificada (en comparación con la señal de excitación de dominio de tiempo obtenida para (o en base a) una o más tramas de audio que preceden una trama de audio perdida, de modo tal que el tono de la señal de excitación de dominio de tiempo sigue los requerimientos de un período de tiempo de la trama de audio perdida. En consecuencia, puede mejorarse la impresión auditiva, que puede lograrse por la información de audio de ocultamiento de errores.

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

En una realización preferida, el ocultamiento de errores está configurado para obtener una información de tono, que

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

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

En una realización preferida, el ocultamiento de errores del decodificador de audio puede estar configurado para obtener una información de tono en base a una información lateral de la información de audio codificada.

En una realización preferida, el ocultamiento de errores puede estar configurado para obtener una información de tono en base a una información de tono disponible para una trama de audio decodificada previamente.

En una realización preferida, el ocultamiento de errores está configurado para obtener una información de tono en base a 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 lateral o podría provenir también de la trama previa si hay LTP, por ejemplo. La información de tono además podría transmitirse en la corriente de bits si está disponible en el codificador. Podemos hacer opcionalmente la búsqueda de tono en la señal de dominio de tiempo directamente o en la residual, lo que habitualmente proporciona mejores resultados sobre la residual (señal de excitación de dominio de tiempo).

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

En una realización preferida, el ocultamiento de errores está configurado para extrapolar un nuevo conjunto de coeficientes de predicción lineal en base al conjunto de coeficientes de predicción lineal, que se han utilizado para decodificar una o más tramas de audio que preceden la trama de audio perdida. En este caso, el ocultamiento de errores está configurado para usar el nuevo conjunto de coeficientes de predicción lineal para proporcionar la información de ocultamiento de errores. Al derivar el nuevo conjunto de coeficientes de predicción lineal, utilizados para proporcionar la información de audio de ocultamiento de errores, a partir de un conjunto de coeficientes de predicción lineal previamente utilizados usando una extrapolación, puede evitarse un recálculo completo de los coeficientes de predicción lineal, lo que ayuda a mantener el esfuerzo computacional razonablemente bajo. Además, al realizar una extrapolación en base al conjunto previamente utilizado de coeficientes de predicción lineal, puede garantizarse que el nuevo conjunto de coeficientes de predicción lineal es al menos similar al conjunto previamente utilizado de coeficientes de predicción lineal, lo que ayuda a evitar discontinuidades cuando se proporciona la información de ocultamiento de errores. Por ejemplo, luego de una cierta cantidad de pérdida de trama, tendemos a estimar la forma de la LPC de ruido de fondo. La velocidad de esta convergencia, por ejemplo, puede depender de la característica de señal.

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

ha hallado que esto ayuda a obtener una buena impresión auditiva.

En una realización preferida, el ocultamiento de errores está configurado para obtener una información de tono que describe un tono de la trama de audio que precede la trama de audio perdida, y proporcionar la información de audio de ocultamiento de errores de acuerdo con la información de tono. Por consiguiente, es posible adaptar el tono de la información de ocultamiento de errores al tono de la trama de audio que precede la trama de audio perdida. Por consiguiente, se evitan las discontinuidades, y puede lograrse una impresión auditiva natural.

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

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

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

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

En una realización preferida, el ocultamiento de errores del decodificador de audio puede estar configurado para obtener una información de tono en base a una información lateral de la información de audio codificada.

En una realización preferida, el ocultamiento de errores puede estar configurado para obtener una información de tono en base a una información de tono disponible para una trama de audio decodificada previamente.

En una realización preferida, el ocultamiento de errores está configurado para obtener una información de tono en base a 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 lateral, o además podría provenir de la trama previa, si hay LTP, por ejemplo. La información de tono además podría ser transmitida en la corriente de bits si está disponible en el codificador. Podemos hacer opcionalmente la búsqueda de tono sobre la señal de dominio de tiempo directamente, o sobre la residual, lo que proporciona habitualmente mejores resultados sobre la residual (señal de excitación de dominio de tiempo).

En una realización preferida, el ocultamiento de errores está configurado para copiar un ciclo de tono de la señal de excitación de dominio de tiempo asociada con la trama de audio que precede la trama de audio perdida una vez o múltiples veces, para obtener una señal de excitación (o por lo menos uno de sus componentes deterministas) para una síntesis de la información de audio de ocultamiento de errores. Al copiar el ciclo de tono de la señal de excitación de dominio de tiempo asociada con la trama de audio que precede la trama de audio perdida una vez o

múltiples veces, y al modificar dichas una o más copias usando un algoritmo de modificación comparativamente simple, la señal de excitación (o por lo menos sus componentes deterministas) para la síntesis de la información de audio de ocultamiento de errores puede ser obtenida con poco esfuerzo computacional. Sin embargo, la reutilización de la señal de excitación de dominio de tiempo asociada con la trama de audio que precede la trama de audio perdida (al copiar dicha señal de excitación de dominio de tiempo) evita discontinuidades audibles.

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

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

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

En una realización preferida, el ocultamiento de errores está configurado 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, el ocultamiento de errores está configurado para realizar la síntesis de LPC, donde 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 ocultamiento de errores. Al combinar la señal de excitación de dominio de tiempo extrapolada (que es típicamente una versión modificada de la señal de excitación de dominio de tiempo derivada para una o más tramas de audio que preceden la trama de audio perdida) y una señal de ruido, pueden considerarse tanto componentes deterministas (por ejemplo, aproximadamente periódicos) como componentes de ruido del contenido de audio, en el ocultamiento de errores. Por lo tanto, puede lograrse que la información de audio de ocultamiento de errores provea una impresión auditiva que es similar a la impresión auditiva previsto por las tramas que preceden la trama perdida.

Además, al combinar una señal de excitación de dominio de tiempo y una señal de ruido, para obtener la señal de entrada para la síntesis de LPC (que puede considerarse una señal de excitación de dominio de tiempo combinada), es posible variar un porcentaje del componente determinista de la señal de audio de entrada para la síntesis de LPC, mientras que se mantiene una energía (de la señal de entrada de la síntesis de LPC, o incluso, de la señal de salida de la síntesis de LPC). En consecuencia, es posible variar las características de la información de audio de ocultamiento de errores (por ejemplo, las características de tonalidad), sin cambiar sustancialmente una energía o volumen de la señal de audio de ocultamiento de errores, de modo tal que es posible modificar la señal de excitación de dominio de tiempo sin causar distorsiones audibles inaceptables.

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

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

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

Breve descripción de las figuras

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

- 5 La figura 1 muestra un diagrama de bloques esquemático de un decodificador de audio, de acuerdo con una realización de la invención;
- La figura 2 muestra un diagrama de bloques esquemático de un decodificador de audio, de acuerdo con otra
- 10 La figura 3 muestra un diagrama de bloques esquemático de un decodificador de audio, de acuerdo con otra realización de la presente invención;
- La figura 4 muestran un diagrama de bloques esquemático de un decodificador de audio, de acuerdo con otra
- 15 La figura 5 muestra un diagrama de bloques esquemático de un ocultamiento de dominio de tiempo para un codificador de transformada;
- 20 La figura 6 muestra un diagrama de bloques esquemático de un ocultamiento de dominio de tiempo para un códec de conmutación;
- La figura 7 muestran un diagrama de bloques de un decodificador TCX que realiza una decodificación TCX en
- 25 La figura 8 muestra un diagrama de bloques esquemático de un decodificador TCX que realiza una decodificación TCX en el caso de ocultamiento de borrado de paquetes TCX-256;
- La figura 9 muestra un diagrama de flujo de un método para proporcionar una información de audio
- 30 decodificada en base a una información de audio codificada, de acuerdo con una realización de la presente invención; y
- La figura 10 muestra un diagrama de flujo de un método para proporcionar una información de audio
- 35 decodificada en base a una información de audio codificada, de acuerdo con otra realización de la presente invención;
- La figura 11 muestra un diagrama de bloques esquemático de un decodificador de audio, de acuerdo con otra realización de la presente invención.

40 Descripción detallada de las realizaciones

1. Decodificador de audio de acuerdo con la figura 1

45 La figura 1 muestra un diagrama de bloques esquemático de un decodificador de audio 100, de acuerdo con una realización de la presente invención. 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 ser recibida por medio de un canal no fiable, de modo tal que se produce una pérdida de trama de vez en cuando. El decodificador de audio 100 además proporciona, en base a la información de audio codificada 110, la información de audio decodificada 112.

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

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

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

5 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 la provisión de la información de audio decodificada 122 en el caso de que haya procesamiento posterior adicional.

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

15 Por consiguiente, el ocultamiento de errores 130 proporciona la información de audio de ocultamiento de errores 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 el ocultamiento de errores 130 puede sustentarse en una o más tramas de audio previas apropiadamente recibidas (que preceden la trama de audio perdida), que son codificadas en la forma de una representación de dominio de frecuencia, o puede derivar de ellas. Como conclusión, el decodificador de audio 100 puede realizar un ocultamiento de errores (es decir, proporcionar una información de audio de ocultamiento de errores 132), que reduce una degradación de una calidad de audio debido a la pérdida de una trama de audio en base a una información de audio codificada, en el que por lo menos algunas tramas de audio son codificadas en una representación de dominio de frecuencia. Se ha hallado que la realización del ocultamiento de errores usando una señal de excitación de dominio de tiempo, incluso si una trama luego de una trama de audio codificada en la representación de dominio de frecuencia apropiadamente recibida está perdida, acarrea una calidad de audio mejorada en comparación con un ocultamiento de errores que se realiza en el dominio de frecuencia (por ejemplo, usando una representación de dominio de frecuencia de la trama de audio codificada en la representación de dominio de frecuencia que precede la trama de audio perdida). Esto se debe a que puede lograrse una transición llana entre la información de audio decodificada asociada con la trama de audio que precede la trama de audio perdida apropiadamente recibida, y la información de audio de ocultamiento de errores 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 en base a la señal de excitación de dominio de tiempo, ayuda a evitar discontinuidades. Por lo tanto, puede lograrse una buena (o por lo menos aceptable) impresión auditiva, usando el decodificador 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 del ocultamiento de códec de voz. El uso de la LPC ayuda a evitar las discontinuidades, y proporciona una mejor forma de las tramas.

40 Además, debe observarse que el decodificador de audio 100 puede ser suplementado por cualquiera de las características y de las funcionalidades que se describen a continuación, ya sea individualmente, ya sea tomadas en combinación.

2. Decodificador de audio de acuerdo con la figura 2

45 La figura 2 muestra un diagrama de bloques esquemático de un decodificador de audio 200 de acuerdo con una realización de la presente invención. El decodificador de audio 200 está configurado para recibir una información de audio codificada 210 y para proporcionar, sobre su base, 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 la forma de una señal de excitación de dominio de tiempo codificada y parámetros de síntesis de señal codificados, por ejemplo, parámetros de LPC). Alternativamente, algunas tramas de la información de audio codificada pueden estar codificadas en una representación de dominio de frecuencia, y algunas otras tramas de la información de audio codificada pueden estar codificadas en una representación de dominio de tiempo, por ejemplo, si el decodificador 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.

60 El decodificador de audio 200 puede comprender típicamente una decodificación/procesamiento 220, que, por ejemplo, puede proporcionar una información de audio decodificada 232 para tramas de audio que son apropiadamente recibidas. En otras palabras, la decodificación/procesamiento 230 puede realizar una decodificación de dominio de frecuencia (por ejemplo, una decodificación de tipo AAC, o similar) en base a una o más tramas de

audio codificadas, codificadas en una representación de dominio de frecuencia. Alternativamente, o además, la decodificación/procesamiento 230 puede estar configurada para realizar una decodificación en el dominio de tiempo (o decodificación en el dominio de predicción lineal) en base a 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 un ocultamiento de errores 240, que está configurado para proporcionar una información de audio de ocultamiento de errores 242 para una o más tramas de audio perdidas. El ocultamiento de errores 240 está configurado para proporcionar la información de audio de ocultamiento de errores 242 para el ocultamiento de una pérdida de una trama de audio (o incluso una pérdida de múltiples tramas de audio). El ocultamiento de errores 240 está configurado para modificar una señal de excitación de dominio de tiempo obtenida en base a una o más tramas de audio que preceden una trama de audio perdida, para obtener la información de audio de ocultamiento de errores 242. En otras palabras, el ocultamiento de errores 240 puede obtener (o derivar) una señal de excitación de dominio de tiempo para (o en base a) una o más tramas de audio codificadas que preceden una trama de audio perdida, y puede modificar dicha señal de excitación de dominio de tiempo, que es obtenida para (o en base a) una o más tramas de audio apropiadamente recibidas que preceden una trama de audio perdida, para obtener así (por medio de la modificación) una señal de excitación de dominio de tiempo que se usa para proporcionar la información de audio de ocultamiento de errores 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 ocultamiento de errores asociada con la trama de audio perdida (o incluso, con múltiples tramas de audio perdidas). Al proporcionar la información de audio de ocultamiento de errores 242 en base a la señal de excitación de dominio de tiempo obtenida en base a una o más tramas de audio apropiadamente recibidas que preceden la trama de audio perdida, pueden evitarse discontinuidades audibles. Por otra parte, al modificar la señal de excitación de dominio de tiempo derivada para (o a partir de) una o más tramas de audio que preceden la trama de audio perdida, y al proporcionar la información de audio de ocultamiento de errores en base a 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 la “atenuación” de un componente de señal determinista (por ejemplo, por lo menos aproximadamente periódico)). Por lo tanto, puede lograrse que la información de audio de ocultamiento de errores 242 comprenda cierta similitud con la información de audio decodificada 232 obtenida en base a tramas de audio apropiadamente decodificadas que preceden la trama de audio perdida, y puede lograrse aunque la información de audio de ocultamiento de errores 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 la provisión de la información de audio de ocultamiento de errores (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 el ocultamiento de errores 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 ocultamiento de errores 242, de modo tal que la información de audio de ocultamiento de errores proporciona una buena impresión auditiva, aun en el caso de que una o más tramas de audio se pierdan. El ocultamiento de errores se realiza en base a una señal de excitación de dominio de tiempo, donde se considera una variación de las características de señal del contenido de audio durante la trama de audio perdida mediante la modificación de la señal de excitación de dominio de tiempo obtenida en base a una o más tramas de audio que preceden una trama de audio perdida.

Además, debe observarse que el decodificador de audio 200 puede ser suplementado por cualquiera de las características y de las funcionalidades que se describen en esta solicitud, ya sea individualmente o en combinación.

3. Decodificador de audio de acuerdo con la figura 3

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

El decodificador de audio 300 está configurado para recibir una información de audio codificada 310 y para proporcionar, sobre su base, una información de audio decodificada 312. El decodificador de audio 300 comprende un analizador de corriente de bits 320, que puede además ser designado 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, sobre su base, 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 lateral 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, sobre su base, 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, sobre su base, un conjunto de factores de escala decodificados 352.

Alternativamente 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 USAC de audio o en el decodificador de audio EVS), puede utilizarse un conjunto de coeficientes de LPC para derivar un conjunto de factores de escala del 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 aplicar el conjunto de factores escalados 352 al conjunto de valores espectrales 342, para obtener así 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 pueden 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.

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, sobre su base, un bloque de muestras de dominio de tiempo, que pueden formar la representación de dominio de tiempo 372.

El decodificador de audio 300 puede comprender opcionalmente un post-procesamiento 376, que puede recibir la representación de dominio de tiempo 372 y modificar en cierta forma la representación de dominio de tiempo 372, para obtener así una versión post-procesada 378 de la representación de dominio de tiempo 372.

El decodificador de audio 300 además comprende un ocultamiento de errores 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 ocultamiento de errores 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), el ocultamiento de errores 380 puede proporcionar la información de audio de ocultamiento de errores en base a 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 ocultamiento de errores puede ser típicamente una representación de dominio de tiempo de un contenido de audio.

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

Con respecto al ocultamiento de errores, debe observarse que el ocultamiento de errores no sucede al mismo tiempo que la decodificación de tramas. Por ejemplo, si la trama n es buena, entonces, hacemos una decodificación normal, y al final, guardamos alguna variable que ayudará si tenemos que ocultar la siguiente trama, entonces, si $n+1$ se pierde, llamamos a la función de ocultamiento proveyendo 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.

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 ocultamiento de errores 382, que habitualmente es además una representación de dominio de tiempo de una señal de audio de ocultamiento de errores previsto para una trama de audio perdida. La combinación de señales 390, por ejemplo, puede combinar representaciones de dominio de tiempo asociadas con tramas de audio posteriores. En el caso de que haya tramas de audio posteriores apropiadamente decodificadas, la combinación de señales 390 puede combinar (por ejemplo, superposición y adición) representaciones de dominio de tiempo asociadas con las tramas de audio posteriores 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 ocultamiento de errores asociada con la trama de audio perdida, para obtener así una transición llana 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 ocultamiento de errores asociada con la trama de audio perdida y la representación de dominio de tiempo asociada con otra trama de audio apropiadamente decodificada después de la trama de audio perdida (u otra información de audio de ocultamiento de errores asociada con otra trama de audio perdida, en el caso de que múltiples tramas de audio consecutivas estén perdidas).

Por consiguiente, la combinación de señales 390 puede proporcionar una información de audio decodificada 312, de modo tal de proporcionar 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 ocultamiento de errores 382 sea previsto 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 el ocultamiento de errores 380) de tramas de audio posteriores. Debido a que algunos códecs tienen cierto solapamiento sobre la parte de superposición y adición que debe ser cancelada, 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 figura 1, en el que los detalles adicionales se muestran en la figura 3. Además, debe observarse que el decodificador de audio 300 de acuerdo con la figura 3 puede suplementarse por cualquiera de las características y de las funcionalidades que se describen en la presente solicitud. En particular, el ocultamiento de errores 380 puede suplementarse por cualquiera de las características y de las funcionalidades que se describen en esta solicitud con respecto al ocultamiento de errores.

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

La figura 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, sobre su base, 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 son codificadas 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 ser codificadas usando una representación de dominio de frecuencia, donde la información de audio codificada comprende una representación codificada de valores espectrales (por ejemplo, valores FFT 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), donde los parámetros de codificación de predicción lineal describen, por ejemplo, un filtro de síntesis de codificación de predicción lineal, que se utiliza para reconstruir una señal de audio en base a 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 lateral 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 lateral, que puede ser usada para el control de etapas de procesamiento adicionales, a partir de la información de audio codificada.

El decodificador de audio 400 comprende una vía de codificación de dominio de frecuencia 430, que, por ejemplo, puede ser sustancialmente idéntica a la vía de codificación del decodificador de audio 300 de acuerdo con la figura 3. En otras palabras, la vía 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 366 opcional, una transformada de dominio de frecuencia a dominio de tiempo 370, un post-procesamiento 376 opcional y un ocultamiento de errores 380, como se describe anteriormente con referencia a la figura 3.

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

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

La vía 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 es derivada de la señal de excitación de dominio de tiempo 452.

La vía 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, sobre su base, 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 estándares en los cuales se describe una codificación y/o decodificación de coeficientes de predicción lineal.

La vía 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 vía de decodificación de dominio de predicción lineal 440 además comprende una síntesis de LPC 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 es definida 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 es obtenida por la filtración (filtración de síntesis) de la señal de excitación de dominio de tiempo 452 (o 456). La vía de codificación de dominio de predicción lineal 440 puede comprender en forma opcional un post-procesamiento 474, que puede ser utilizado para refinar o ajustar las características de la señal de audio de dominio de tiempo decodificada 472.

La vía de decodificación de dominio de predicción lineal 440 además comprende un ocultamiento de errores 480, que está configurado 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). El ocultamiento de errores 480 puede recibir opcionalmente información adicional, por ejemplo, una información de tono. El ocultamiento de errores 480 puede, en consecuencia, proporcionar una información de audio de ocultamiento de errores, que puede presentarse en la forma de una señal de audio de dominio de tiempo, en el caso de que una trama (o subtrama) de la información de audio codificada 410 se pierda. Por lo tanto, el ocultamiento de errores 480

puede proporcionar la información de audio de ocultamiento de errores 482 de modo tal que las características de la información de audio de ocultamiento de errores 482 se adaptan sustancialmente a las características de una última trama de audio apropiadamente decodificada que precede la trama de audio perdida. Debe entenderse que el ocultamiento de errores 480 puede comprender cualquiera de las características y de las funcionalidades que se describen con respecto al ocultamiento de errores 240. Asimismo, debe observarse que el ocultamiento de errores 480 puede además comprender cualquiera de las características y de las funcionalidades que se describen con respecto al ocultamiento de dominio de tiempo de la figura 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 ocultamiento de errores 382 provista por el ocultamiento de errores 380, la señal de audio de dominio de tiempo decodificada 472 (o su versión post-procesada 476) y la información de audio de ocultamiento de errores 482 provista por el ocultamiento de errores 480. El combinador de señal 490 puede estar configurado para combinar dichas señales 372 (o 378), 382, 472 (o 476) y 482 para obtener 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 llanas entre tramas de audio posteriores para las cuales se proporciona la señal de audio de dominio de tiempo por medio de diferentes entidades (por ejemplo, por diferentes vías de codificación 430, 440). Sin embargo, el combinador de señal 490 puede además proporcionar transiciones llanas si la señal de audio de dominio de tiempo es provista 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 ser cancelada, 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 en forma opcional una compensación de solapamiento de dominio de tiempo artificial (TDAC).

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

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

5. Ocultamiento de dominio de tiempo de acuerdo con la figura 5

La figura 5 muestra un diagrama de bloques esquemático de un ocultamiento de errores de acuerdo con una realización de la presente invención. El ocultamiento de errores de acuerdo con la figura 5 se designa en su totalidad como 500.

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

Debe observarse que el ocultamiento de errores 500 puede, por ejemplo, ocupar el lugar del ocultamiento de errores 130, de modo tal que la información de audio de ocultamiento de errores 512 puede corresponder a la información de audio de ocultamiento de errores 132. Además, debe observarse que el ocultamiento de errores 500 puede ocupar el lugar del ocultamiento de errores 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 ocultamiento de errores 512 puede corresponder a la información de audio de ocultamiento de errores 382.

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

El ocultamiento de errores 500 además comprende un análisis de LPC 530, que está configurada para recibir la señal de audio de dominio de tiempo 510, o su versión preenfatisada 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 en forma aproximada, la señal de entrada del análisis de LPC).

El ocultamiento de errores 500 además comprende una búsqueda de tono 540, que está configurada para obtener una información de tono 542, por ejemplo, en base a una trama de audio previamente decodificada.

El ocultamiento de errores 500 además comprende una extrapolación 550, que puede estar configurada para obtener una señal de excitación de dominio de tiempo extrapolada en base al resultado del análisis de LPC (por ejemplo, en base a la señal de excitación de dominio de tiempo determinada por el análisis de LPC), y posiblemente, en base al resultado de la búsqueda de tono.

El ocultamiento de errores 500 además comprende una generación de ruido 560, que proporciona una señal de ruido 562. El ocultamiento de errores 500 además comprende un combinador/atenuador 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, sobre su base, una señal de excitación de dominio de tiempo combinada 572. El combinador/atenuador 570 puede estar configurado para combinar la señal de excitación de dominio de tiempo extrapolada 552 y la señal de ruido 562, en el que puede realizarse una atenuación, 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/atenuador es también posible. Además, se hace referencia a la descripción a continuación.

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

El ocultamiento de errores 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 opcional, ya que el ocultamiento de errores puede además usar una combinación de señales que ya es provista en el entorno del decodificador de audio. Por ejemplo, la superposición y adición 590 puede ser reemplazada por la combinación de señales 390 en el decodificador de audio 300 en algunas realizaciones.

A continuación, se describirán algunos detalles adicionales respecto del ocultamiento de errores 500.

El ocultamiento de errores 500 de acuerdo con la figura 5 cubre el contexto de un códec de dominio de transformada como AAC_LC o AAC_ELD. En otras palabras, el ocultamiento de errores 500 se adapta bien para el uso en dicho códec de dominio de transformada (y, en particular, en dicho decodificador de audio de dominio de transformada). En el caso de un códec de transformada solamente (por ejemplo, en ausencia de una vía de decodificación de dominio de predicción lineal), se usa una señal de salida de una última trama, como un punto de inicio. Por ejemplo, una señal de audio de dominio de tiempo 372 puede usarse como un punto de inicio para el ocultamiento de errores. 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).

A continuación, se describirá en más detalle las subunidades y funcionalidades del ocultamiento de errores 500.

5.1. Análisis de LPC

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

tiempo).

5.2. Búsqueda de tono

- 5 Existen diferentes enfoques para obtener el tono que se utilizaba para lograr la construcción de la nueva señal (por ejemplo, la información de audio de ocultamiento de errores).

10 En el contexto del códec usando un filtro de LTP (filtro de predicción de largo plazo), como un AAC-LTP, si la última trama fue 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.

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

20 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, si bien no hay ninguna aplicación de filtración (además, ninguna filtración de predicción de largo plazo (LTP) en el canal limpio).

25 Alternativamente, es posible realizar una búsqueda de tono en el decodificador. La búsqueda de tono de AMR-WB en el caso de la 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 directamente 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 circuito abierto por medio de una correlación cruzada normalizada. A continuación, opcionalmente, refinamos la búsqueda de tono realizando una búsqueda de circuito cerrado alrededor del tono de circuito abierto, con un cierto delta. Debido a las limitaciones del ventaneado de ELD, podría hallarse un tono erróneo y, en consecuencia, además, verificamos que el tono hallado sea correcto, o de lo contrario, se descarta.

30 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 ocultamiento de errores. 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 tratar de extrapolar el tono que necesitamos al final de nuestra trama oculta.

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

40 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. Alternativamente, la información de tono puede determinarse en el decodificador de audio, por ejemplo, en el dominio de excitación, es decir, en base a 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 que se utiliza para la provisión de la información de audio de ocultamiento de errores.

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

60 La excitación (por ejemplo, la señal de excitación de dominio de tiempo) obtenida de la trama previa (o bien solo computada para la trama perdida o ya guardada en la trama perdida previa para múltiple pérdida de trama) es utilizada para la construcción de la parte armónica (además designada como componente determinista o 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 la posterior 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 en base a la última trama de audio apropiadamente decodificada que precede la trama de audio perdida) es el filtro de paso bajo con un filtro dependiente de la tasa de muestreo (ya que el retardo extra bajo (ELD) cubre una combinación de tasa de muestreo realmente amplia - que va de núcleo AAC-ELD a AAC-ELD con SBR o AAC-ELD tasa dual SBR).

El tono en una señal de voz es casi siempre cambiante. Por lo tanto, el ocultamiento presentado anteriormente tiende a crear algunos problemas (o al menos distorsiones) en la recuperación, ya que el tono al final de la señal ocultada (es decir, al final de la información de audio de ocultamiento de errores) con frecuencia no coincide con el tono de la primera trama buena. Por lo tanto, opcionalmente, en algunas realizaciones, se trata de 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) para aproximarse al tono al comienzo de la primera trama apropiadamente decodificada después 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].

Si se usa una predicción de largo plazo (LTP) en un códec de dominio de frecuencia, es posible usar la demora como la información de inicio acerca 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.

5.4. Ganancia de tono

En algunas realizaciones, se prefiere la aplicación de una ganancia sobre la excitación previamente obtenida para lograr el nivel deseado. La "ganancia de 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, ser obtenida mediante la realización de una correlación normalizada en el dominio de tiempo al final de la última trama buena (por ejemplo, apropiadamente decodificada). 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 luego solo aplicar una atenuación (ganancia reducida) para la siguiente pérdida de trama consecutiva.

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

Como conclusión, en algunos casos, la señal de excitación de dominio de tiempo obtenida, por ejemplo, en base a una trama de audio previamente decodificada, se escala de acuerdo con 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 (por lo menos aproximadamente periódico), la ganancia puede determinar una intensidad relativa de dichos componentes de señal deterministas (por lo menos aproximadamente periódicos) en la información de audio de ocultamiento de errores. Además, la información de audio de ocultamiento de errores puede sustentarse en un ruido, que es además formado por la síntesis de LPC, de modo tal que una energía total de la información de audio de ocultamiento de errores es adaptada, por lo menos hasta cierto grado, a una trama de audio apropiadamente decodificada que precede la trama de audio perdida e, idealmente, además a una trama de audio apropiadamente decodificada luego de las una o más tramas de audio perdidas.

5.5. Creación de la parte de ruido

Una "innovación" se crea mediante un generador de ruido aleatorio. En forma opcional, este ruido es adicionalmente filtrado paso alto y opcionalmente preenfatisado para tramas de voz y de inicio. Como para el paso bajo de la parte armónica, este filtro (por ejemplo, el filtro de paso alto) es dependiente de la tasa de muestreo. Este ruido (que se

proporciona, por ejemplo, mediante una generación de ruido 560) estará formado por la codificación predictiva lineal (LPC) (por ejemplo, por la síntesis de LPC 580) para llegar lo más cerca posible del ruido de fondo. La característica de paso alto es además opcionalmente cambiada sobre pérdida de trama consecutiva, de modo tal que, sobre una cierta cantidad de una pérdida de trama, ya no hay más filtración, para solo obtener el ruido con forma de banda completa para lograr un ruido de confort cercano al ruido de fondo.

Una ganancia de innovación (que, por ejemplo, puede determinar una ganancia del ruido 562 en la combinación/atenuación 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), por ejemplo, se calcula mediante la eliminación de la contribución previamente computada de la tono (si existe) (por ejemplo, una versión escalada, escalada usando la "ganancia de tono", de la señal de excitación de dominio de tiempo obtenida en base a 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 solo sobre la primera trama perdida, y luego, la atenuación, si bien, en este caso, la atenuación en forma saliente podría ir o bien a 0, que resulta en un silenciamiento completado, o a un nivel de ruido estimado presente en el fondo. La longitud de la correlación es, por ejemplo, equivalente a la longitud de dos subtramas, y el retardo es equivalente a la demora de tono utilizada para la creación de la parte armónica.

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 además se multiplica 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. Atenuación

La atenuación se usa en su mayoría para múltiple pérdida de trama. Sin embargo, la atenuación puede además utilizarse en el caso de que solo se pierda una sola trama de audio.

En el caso de una múltiple pérdida de trama, los parámetros de LPC no se recalculan. O bien, se mantiene el último computado, o se realiza el ocultamiento 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 en base a 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 torna cada vez más "de tipo ruido". Por lo tanto, la "periodicidad" (o, más precisamente, el componente determinista, o componente por lo menos aproximadamente periódico de la señal de salida 582 de la síntesis de LPC 580) se reduce en función del tiempo.

La velocidad de la convergencia de acuerdo con la cual la periodicidad de la señal 572, y/o la periodicidad de la señal 582, converge a 0, depende de los parámetros de la última trama correctamente recibida (o apropiadamente decodificada) y/o de la cantidad de tramas borradas consecutivas, y es controlada por un factor de atenuación, α . El factor, α , es adicionalmente dependiente de la estabilidad del filtro de 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 largo, 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 una atenuación más veloz de esta señal.

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 la atenuación de la parte tonal, en este caso.

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 fiables, o que la señal es realmente impredecible. Por lo tanto, nuevamente, se prefiere realizar la atenuación en forma más veloz (por ejemplo, la atenuación de la señal de excitación de dominio de tiempo 552 obtenida en base a una o más tramas de audio apropiadamente decodificadas que preceden una o más tramas de audio perdidas).

5.7. Síntesis de LPC

Para regresar al dominio de tiempo, se prefiere realizar una síntesis de LPC 580 sobre a la suma de las dos excitaciones (parte tonal y parte ruidosa), seguida de un desénfasis. En otras palabras, se prefiere realizar la síntesis de LPC 580 en base a una combinación pesada de una señal de excitación de dominio de tiempo 552 obtenida en

base a una o más tramas de audio apropiadamente decodificadas que preceden la trama de audio perdida (parte tonal) y la señal de ruido 562 (parte ruidosa). Como se menciona anteriormente, la señal de excitación de dominio de tiempo 552 puede ser modificada 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.

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 la misma. Sin embargo, pueden aplicarse diferentes conceptos de superposición y adición.

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

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

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 además ser provista para un período de tiempo que es mayor que una trama de audio perdida. Por consiguiente, puede realizarse una superposición y adición entre la información de audio de ocultamiento de errores (que es consecuentemente obtenida para un período de tiempo más largo que una extensión temporal de la trama de audio perdida) y una información de audio decodificada provista para una trama de audio apropiadamente decodificada luego de una o más tramas de audio perdidas.

En resumen, el ocultamiento de errores 500 se adapta bien al caso en el cual las tramas de audio se codifican en el dominio de frecuencia. Aun cuando las tramas de audio son codificadas en el dominio de frecuencia, la provisión de la información de audio de ocultamiento de errores se realiza en base a una señal de excitación de dominio de tiempo. Se aplican diferentes modificaciones a la señal de excitación de dominio de tiempo obtenida en base a 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 también se modifica por una escala (aplicación de una ganancia), en el que una atenuación del componente determinista (o tonal, o por lo menos aproximadamente periódico) puede realizarse por el escalador/atenuador 570, de modo tal que la señal de entrada 572 de la síntesis de LPC 580 comprende tanto un componente que 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. Sin embargo, el componente determinista de la señal de entrada 572 de la síntesis de LPC 580 habitualmente se modifica (por ejemplo, escala de tiempo y/o escala de amplitud) con respecto a la señal de excitación de dominio de tiempo provista por el análisis de LPC 530.

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

6. Ocultamiento de dominio de tiempo de acuerdo con la figura 6

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

Además, debe observarse que la realización de acuerdo con la figura 6 cubre el contexto (que puede usarse dentro del contexto) de un códec de conmutación usando dominios de tiempo y frecuencia combinados, tales como USAC (MPEG-D/MPEG-H) o EVS (3GPP). En otras palabras, el ocultamiento 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 en base a coeficientes de predicción lineal).

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

- 5 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
- 10 acuerdo con la figura 5, es decir, realizar un análisis de LPC. Si la trama previa era de tipo ACELP, además ya tenemos la información de tono de las subtramas en la última trama. Si la última trama fue TCX (Excitación codificada transformada) con LTP (predicción de largo plazo), además tenemos la información de demora que proviene de la predicción de largo plazo. Y si la última trama estuvo en el dominio de frecuencia sin predicción de largo plazo (LTP), entonces la búsqueda de tono se realiza preferentemente en forma directa en el dominio de
- 15 excitación (por ejemplo, en base a una señal de excitación de dominio de tiempo provista 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 se basa 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

20 ruido DTX si DTX (transmisión discontinua) existe en el códec.

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

- 25 A continuación, se describirá en más detalle el ocultamiento de errores 600 de acuerdo con la figura 6.

El ocultamiento de errores 600 recibe una excitación pasada 610 y una información de tono pasada 640. Además, el ocultamiento de errores 600 proporciona una información de audio de ocultamiento de errores 612.

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

- 35 El ocultamiento de errores 600 adicionalmente comprende una extrapolación 650, que puede corresponder a la extrapolación 550, de modo tal que se hace referencia a la descripción anterior.

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

- 40 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.

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

- 60 Como conclusión, el ocultamiento de errores 600 difiere sustancialmente del ocultamiento de errores 500, en términos de que el ocultamiento de errores 600 obtiene directamente la información de excitación pasada 610 y la información de tono pasada 640, 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 el ocultamiento de errores 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 del ocultamiento de errores 600. Sin embargo, debe

observarse que los detalles específicos deben considerarse ejemplares, en lugar de rasgos esenciales.

6.1. Tono pasado de búsqueda de tono

- 5 Hay diferentes enfoques para obtener el tono que se utiliza en la construcción de la nueva señal.

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.

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 de largo plazo (LTP), si bien no aplicamos ninguna filtración (tampoco ninguna filtración de predicción de largo plazo en el canal limpio).

Otra solución es realizar una búsqueda de tono en el decodificador. La búsqueda de tono de AMR-WB en el caso de la TCX se realiza en el dominio de la FFT. En la TCX, por ejemplo, usamos el dominio de la MDCT, entonces, perdemos las fases. Por lo tanto, la búsqueda de tono se realiza directamente en el dominio de excitación (por ejemplo, en base a 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, en base a una señal de audio de dominio de tiempo completamente decodificada).

La búsqueda de tono en el dominio de excitación (por ejemplo, en base a la señal de excitación de dominio de tiempo) se realiza primero con un circuito abierto por medio de una correlación cruzada normalizada. A continuación, opcionalmente, la búsqueda de tono puede refinarse realizando una búsqueda de circuito cerrado alrededor del tono de circuito 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 a errores, entonces seleccionamos el tono que corresponde al 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 eficiente considerar un tono pasado (es decir, un tono asociada con una trama de audio previamente decodificada). Alternativamente, 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, en base a 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 circuito abierto y una búsqueda de circuito cerrado puede realizarse para obtener una información de tono particularmente fiable y precisa. Alternativamente 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 fiable.

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

La excitación (por ejemplo, en la forma de una señal de excitación de dominio de tiempo) obtenida de la trama previa (o bien solo computada para la trama perdida o ya guardada en la trama perdida previa para múltiple pérdida de trama) 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 período de la 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, el ocultamiento presentado anteriormente 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, mediante la extrapolación 650.

Si se usa la LTP en la 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 se presenta en el estado de la técnica.

Como conclusión, la extrapolación (por ejemplo, de la señal de excitación de dominio de tiempo asociada con, u obtenida en base a, 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 de acuerdo con un cómputo, o una estimación, de un cambio de tono (esperado) durante la trama de audio perdida. Pueden obtenerse diferentes conceptos para la determinación del cambio de tono.

6.3. Ganancia de tono

En la realización de acuerdo con la figura 6, se aplica una ganancia sobre la excitación previamente obtenida 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 trama buena. Por ejemplo, la longitud de la correlación puede ser equivalente a la longitud de dos subtramas, y el retardo puede ser equivalente a la demora de 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 fiable, 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 será procesada por otro conjunto de LPC, no proporcionará la energía esperada en el dominio de tiempo.

La ganancia del tono determina la cantidad de tonalidad que se creará, si bien además se agregará cierto ruido con forma para no tener solo un tono artificial. Si se obtiene una muy baja ganancia de tono, entonces, puede construirse una señal que consiste solo en un ruido con forma.

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

6.4. Creación de la parte de ruido

Se crea una innovación por medio de un generador de ruido aleatorio 660. Este ruido es adicionalmente filtrado de paso alto y opcionalmente preenfatisado para tramas de voz y de inicio. La filtración de paso alto y el preénfasis, que pueden realizarse selectivamente para tramas de voz y de inicio, no se muestran explícitamente en la figura 6, si bien pueden realizarse, por ejemplo, dentro del generador de ruido 660 o dentro del combinador/atenuador 670.

El ruido se conformará (por ejemplo, después de la combinación con la señal de excitación de dominio de tiempo 652 obtenida por la extrapolación 650) mediante la LPC para obtener así lo más cercano posible al ruido de fondo.

Por ejemplo, la ganancia de innovación puede calcularse mediante la eliminación de la contribución previamente computada del tono (si existe) y realizando una correlación al final de la última trama buena. 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.

Opcionalmente, esta ganancia puede además ser multiplicada 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 es multiplicada 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 ocultamiento de errores se obtiene mediante la formación de ruido proporcionado mediante el generador de ruido 660 usando la síntesis de LPC 680 (y, posiblemente, el desénfasis 684). Además, puede aplicarse una filtración de paso alto adicional y/o un preénfasis. La ganancia de la contribución de ruido a la señal de entrada 672 de la síntesis de LPC 680 (además designada "ganancia de innovación") puede computarse en base a la última trama de audio apropiadamente decodificada que precede la trama de audio perdida, en el que un componente determinista (o por lo menos aproximadamente

periódico) puede ser eliminado de la trama de audio que precede la trama de audio perdida, y 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.

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

6.5. Atenuación

- 10 La atenuación se usa en su mayoría para múltiples pérdidas de tramas. Sin embargo, la atenuación puede usarse además en el caso de que solo se pierda una única trama de audio.

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

- 15 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 mediante un factor de atenuación, α . El factor, α , además depende de la estabilidad del filtro de LP. Opcionalmente, el factor α puede alterarse en relación con la longitud de tono. Por ejemplo, si el tono es 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 atenúa más velozmente.

- 20 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 la atenuación de la parte tonal, en este caso.

- 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 fiables, o que la señal es realmente impredecible. Por lo tanto, nuevamente, deberíamos realizar la atenuación de manera más veloz.

- 30 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 señal de excitación de dominio de tiempo 552 obtenida en base a una o más tramas de audio que preceden una trama de audio perdida (o una o más de sus copias) se ajusta de acuerdo con uno o más parámetros de una o más tramas de audio (y/o de acuerdo con una cantidad de tramas de audio perdidas consecutivas). En particular, la longitud de tono y/o el índice al cual cambia el tono en función del tiempo, y/o la cuestión de si una predicción de tono falla o tiene éxito, pueden utilizarse para ajustar dicha velocidad.

6.6. Síntesis de LPC

- 45 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 ruidosa 662), seguida del desénfasis 684.

- En otras palabras, el resultado de la combinación pesada (atenuación) 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 ingresa en la síntesis de LPC 680, que, por ejemplo, puede realizar una filtración de síntesis en base a dicha señal de excitación de dominio de tiempo combinada 672 de acuerdo con coeficientes de LPC que describen el filtro de síntesis.

6.7. Superposición y adición

- 55 Debido a que no se conoce, durante el ocultamiento, 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 ocultamiento de errores) puede, por ejemplo, crearse para la mitad de una trama más que la trama ocultada (perdida). Además, puede crearse solapamiento artificial sobre la misma (en el que el solapamiento artificial puede, por ejemplo, adaptarse a la superposición y adición de MDCT).

- 60 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 largas ventanas de superposición y adición, o si queremos usar una ventana cuadrada, se computa la respuesta de entrada cero (ZIR) al final de la memoria intermedia de síntesis.

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 ocultamiento de errores que se proporciona principalmente para una trama de audio perdida, pero además, para una cierta porción de tiempo luego de la trama de audio perdida, y la información de audio decodificada proporcionada para la primera trama de audio apropiadamente decodificada después de una secuencia de una o más tramas de audio perdidas. Para obtener una apropiada superposición y adición, incluso, para modos de decodificación que acarreen un solapamiento de dominio de tiempo en una transición entre tramas de audio posteriores, puede proporcionarse una información de cancelación de solapamiento (por ejemplo, designado solapamiento artificial). Por consiguiente, una superposición y adición entre la información de audio de ocultamiento de errores y la información de audio de dominio de tiempo obtenida en base a la primera trama de audio apropiadamente decodificada luego de una trama de audio perdida, logra una cancelación del solapamiento.

Si la primera trama de audio apropiadamente decodificada luego de la secuencia de una o más tramas de audio perdidas es codificada en el modo de ACELP, puede computarse una información de superposición específica, que puede sustentarse en una respuesta de entrada cero (ZIR) de un filtro de LPC.

Como conclusión, el ocultamiento de errores 600 se adapta bien para el uso en un códec de audio de conmutación. Sin embargo, el ocultamiento de errores 600 puede además ser usado 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

Debe observarse que se logra un ocultamiento de errores particularmente bueno mediante el concepto arriba mencionado, para la extrapolación de una señal de excitación de dominio de tiempo, la combinación del resultado de la extrapolación con una señal de ruido usando una atenuación (por ejemplo, una atenuación cruzada), y para la realización de una síntesis de LPC en base a un resultado de la atenuación cruzada.

7. Decodificador de audio de acuerdo con la figura 11

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

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 vía de decodificación de dominio de predicción lineal 440 en el decodificador de audio 400.

El decodificador de audio 1100 está configurado para recibir una información de audio codificada 1110 y para proporcionar, sobre su base, 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.

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.

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 en base a los coeficientes espectrales codificados 1122. Puede usarse cualquier concepto de decodificación conocido para la decodificación de coeficientes espectrales.

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 en base a 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 el estándar USAC. Por ejemplo, la representación codificada 1124 de los coeficientes de codificación de predicción lineal puede comprender una representación polinomial, que es decodificada y convertida en un conjunto de factores de escala por el coeficiente de codificación de predicción lineal para la conversión de factor de escala 1142.

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 así 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 describe anteriormente, en el que los valores espectrales decodificados escalados procesados 1162 son obtenidos por el procesamiento 1160 opcional. 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, sobre su base, una representación de dominio de tiempo 1172, que puede corresponder a la representación de dominio de tiempo 372 que se describe anteriormente. El decodificador de audio 1100 además comprende un primer post-procesamiento 1174 opcional, y un segundo post-procesamiento 1178 opcional, que, por ejemplo, pueden corresponder, al menos en parte, al post-procesamiento 376 opcional mencionado anteriormente. 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 además comprende un bloque de ocultamiento de errores 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, sobre su base, una información de audio de ocultamiento de errores 1182.

El bloque de ocultamiento de errores 1180 está configurado para proporcionar la información de audio de ocultamiento de errores 1182 para el ocultamiento de una pérdida de una trama de audio luego 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 al ocultamiento de errores 380 y al ocultamiento de errores 480, y además, al ocultamiento de errores 500 y al ocultamiento de errores 600.

Sin embargo, el bloque de ocultamiento de errores 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, además, a la señal de excitación de dominio de tiempo 610). Además, el bloque de ocultamiento de errores 1180 comprende un ocultamiento de errores 1188, que, por ejemplo, puede realizar la funcionalidad de los bloques 540, 550, 560, 570, 580, 584 del ocultamiento de errores 500, o que, por ejemplo, puede realizar la funcionalidad de los bloques 640, 650, 660, 670, 680, 684 del ocultamiento de errores 600. Sin embargo, el bloque de ocultamiento de errores 1180 difiere levemente del ocultamiento de errores 500 y, además, del ocultamiento de errores 600. Por ejemplo, el bloque de ocultamiento de errores 1180 (que comprende el análisis de LPC 1184) difiere del ocultamiento de errores 500 en términos de que los coeficientes de LPC (utilizados para la síntesis de LPC 580) no son determinados por el análisis de LPC 530, si bien son (opcionalmente) recibidos desde la corriente de bits. Asimismo, el bloque de ocultamiento de errores 1188, que comprende el análisis de LPC 1184, difiere del ocultamiento de errores 600 en términos de que la "excitación pasada" 610 es obtenida 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, además, la información de audio de ocultamiento de errores 1182 (naturalmente, para tramas de audio posteriores), y combina dichas señales, preferentemente, usando una operación de superposición y adición, para obtener así la información de audio decodificada 1112.

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

8. Método de acuerdo con la figura 9

La figura 9 muestra un diagrama de flujo de un método para proporcionar una información de audio decodificada en base a una información de audio codificada. El método 900 de acuerdo con la figura 9 comprende la provisión de 910 una información de audio de ocultamiento de errores para el ocultamiento de una pérdida de una trama de audio luego 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 figura 9 se basa en las mismas consideraciones que el decodificador de audio de acuerdo con la figura 1. Además, debe observarse que el método 900 puede suplementarse mediante cualquiera de las características y de las funcionalidades que se describen en esta solicitud, o bien individualmente, o en combinación.

9. Método de acuerdo con la figura 10

La figura 10 muestra un diagrama de flujo de un método para proporcionar una información de audio decodificada en base a una información de audio codificada. El método 1000 comprende proporcionar 1010 una información de

audio de ocultamiento de errores para el ocultamiento de una pérdida de una trama de audio, en el que una señal de excitación de dominio de tiempo obtenida para (o en base a) una o más tramas de audio que preceden una trama de audio perdida es modificada para obtener la información de audio de ocultamiento de errores.

- 5 El método 1000 de acuerdo con la figura 10 se basa en las mismas consideraciones que el decodificador de audio arriba mencionado de acuerdo con la figura 2.

Además, debe observarse que el método de acuerdo con la figura 10 puede suplementarse mediante cualquiera de las características y de las funcionalidades que se describen en esta solicitud, o bien individualmente, o en combinación.

10. Observaciones adicionales

15 En las realizaciones descritas anteriormente, pueden manipularse múltiples pérdidas de trama de diferentes maneras. Por ejemplo, si se pierden dos o más tramas, la parte periódica de la señal de excitación de dominio de tiempo para la segunda trama perdida puede derivar de (o ser igual a) una copia de la parte tonal de la señal de excitación de dominio de tiempo asociada con la primera trama perdida. Alternativamente, la señal de excitación de dominio de tiempo para la segunda trama perdida puede sustentarse en un análisis de 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, 20 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 claro que estos aspectos además representan una descripción del correspondiente método, donde un bloque o dispositivo corresponde a una etapa de método o a un rasgo de una etapa de método. En forma análoga, los aspectos descritos en el contexto de una etapa de método además representan una descripción de un correspondiente bloque o ítem o rasgo de un correspondiente aparato. Algunas o la totalidad de las etapas de método pueden ser ejecutadas por (o usando) un aparato de soporte físico, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, una o más de las etapas de método más importantes pueden ser ejecutadas mediante dicho aparato.

De acuerdo con ciertos requerimientos de implementación, las realizaciones de la invención pueden implementarse en soporte físico o soporte lógico. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo, un disquete, un DVD, un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control electrónicamente legibles allí almacenadas, que cooperan (o son capaces de cooperar) con un sistema de ordenador programable de manera tal de realizar 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 portador de datos que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema de ordenador programable, de manera de llevar a cabo uno de los métodos que se describen en la presente solicitud.

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

Otras realizaciones comprenden el programa de ordenador para realizar uno de los métodos que se describen en la presente solicitud, almacenado en un portador legible en una máquina.

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

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

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

Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado o adaptado para realizar uno de los métodos que se describen en la presente solicitud.

5 Una realización adicional comprende un ordenador que tiene allí instalado el programa de ordenador para realizar uno de los métodos que se describen en la presente solicitud.

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

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

20 El aparato descrito en la presente solicitud puede implementarse usando un aparato de soporte físico, o usando un ordenador, o empleando una combinación de un aparato de soporte físico y un ordenador.

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

25 Las realizaciones descritas anteriormente 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 la presente solicitud 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 de la presente solicitud.

30 12. Conclusiones

35 Como conclusión, aunque cierto ocultamiento para códecs de dominio de transformada ha sido descrito 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 el ocultamiento (dominio de frecuencia a dominio de tiempo o excitación). Por consiguiente, las realizaciones de acuerdo con la invención crean un ocultamiento de voz de alta calidad para decodificadores de dominio de transformada.

40 El modo de codificación de transformada es similar a aquel en USAC (conferir, por ejemplo, referencia [3]). Utiliza la Transformada de coseno discreta modificada (MDCT) como una transformada, y la formación de ruido espectral se logra mediante la aplicación de la envoltura espectral de LPC pesada en el dominio de frecuencia (además conocida como FDNS, "formación 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 el estándar USAC. Sin embargo, el concepto de ocultamiento de errores divulgado en esta solicitud puede 45 además se usado en un decodificador de audio que es de tipo "AAC", o en cualquier códec (o decodificador) de la familia AAC.

50 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, el ocultamiento se realiza en el dominio de tiempo o en el dominio de excitación.

A continuación, se describirán algunas ventajas y algunas características del ocultamiento de dominio de tiempo (o del ocultamiento de dominio de excitación).

55 El ocultamiento de TCX convencional, como se describe, por ejemplo, con referencia a las figuras 7 y 8, además denominado sustitución de ruido, no es adecuado para las señales de tipo voz, o incluso, para señales tonales. Las realizaciones de acuerdo con la invención crean un nuevo ocultamiento para un códec de dominio de transformada que se aplica en el dominio de tiempo (o en el dominio de excitación de un decodificador de codificación de predicción lineal). Es similar a un ocultamiento de tipo ACELP, y aumenta la calidad del ocultamiento. Se ha hallado 60 que la información de tono es conveniente (o incluso requerida, en algunos casos) para un ocultamiento de tipo ACELP. Por lo tanto, las realizaciones de acuerdo con la presente invención están configuradas para hallar valores de tono fiables para la trama previa codificada en el dominio de frecuencia.

Diferentes partes y detalles se han explicado anteriormente, por ejemplo, en base a las realizaciones de acuerdo con

las figuras 5 y 6.

Para concluir, las realizaciones de acuerdo con la invención crean una ocultación de error que supera a las soluciones convencionales.

De acuerdo con un primer aspecto, un decodificador de audio 100; 300 para proporcionar una información de audio decodificada 112, 312 basándose en una información de audio codificada 110; 310 puede comprender: una ocultación de error 130; 380; 500 configurado para proporcionar una información de audio de ocultación de error 132; 382; 512 para ocultar una pérdida de una trama de audio después de una trama de audio codificada en una representación de ámbito de frecuencia 322 utilizando una señal de excitación en el ámbito del tiempo 532.

De acuerdo con un segundo aspecto haciendo referencia de nuevo al primer aspecto, el decodificador de audio 100; 300 puede comprender un núcleo de decodificador de ámbito de frecuencia 120; 340, 350, 360, 366, 370 configurado para aplicar una escala basada en factor de escala 360 a una pluralidad de valores espectrales 342 derivados de la representación de ámbito de frecuencia 322, y la ocultación de error 130; 380; 500 puede configurarse para proporcionar la información de audio de ocultación de error 132; 382; 512 para ocultar una pérdida de una trama de audio después de una trama de audio codificada en una representación de ámbito de frecuencia 322 que comprende una pluralidad de factores de escala codificados 328 que usan una señal de excitación en el ámbito del tiempo 532 derivada de la representación de ámbito de frecuencia.

De acuerdo con un tercer aspecto haciendo referencia de nuevo al primer aspecto o segundo aspecto, en el decodificador de audio 100; 300 la representación en el ámbito de la frecuencia puede comprender una representación 326 codificada de una pluralidad de valores espectrales y una representación 328 codificada de una pluralidad de factores de escala para escalar los valores espectrales, o el decodificador de audio puede configurarse para derivar una pluralidad de factores de escala para escalar los valores espectrales de una representación codificada de parámetros LPC.

De acuerdo con un cuarto aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a tercero, el decodificador de audio 100; 300 puede comprender un núcleo de decodificador de ámbito de frecuencia 120; 340, 350, 350, 366, 370 configurado para derivar una representación de señal de audio en el ámbito del tiempo 122; 372 de la representación en el ámbito de la frecuencia 322 sin usar una señal de excitación en el ámbito del tiempo como una cantidad intermedia para la trama de audio codificada en la representación en el ámbito de la frecuencia.

De acuerdo con un quinto aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a cuarto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para obtener la señal de excitación en el ámbito del tiempo 532 basándose en la trama de audio codificada en la representación en el ámbito de la frecuencia 322 que precede a una trama de audio perdida, y la ocultación de error puede configurarse para proporcionar la información de audio de ocultación de error 122; 382; 512 para ocultar la trama de audio perdida usando dicha señal de excitación en el ámbito del tiempo.

De acuerdo con un sexto aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a quinto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para realizar un análisis LPC 530 basándose en la trama de audio codificada en la representación en el ámbito de la frecuencia 322 que precede a la trama de audio perdida, para obtener un conjunto de parámetros de codificación de predicción lineal y la señal de excitación en el ámbito del tiempo 532 que representa un contenido de audio de la trama de audio codificado en la representación en el ámbito de la frecuencia que precede a la trama de audio perdida; o la ocultación de error 130; 380; 500 puede configurarse para realizar un análisis LPC 530 basándose en la trama de audio codificada en la representación en el ámbito de la frecuencia 322 que precede a la trama de audio perdida, para obtener una señal de excitación en el ámbito del tiempo 532 que representa un contenido de audio de la trama de audio codificada en la representación en el ámbito de la frecuencia que precede a la trama de audio perdida; o el decodificador de audio puede configurarse para obtener un conjunto de parámetros de codificación de predicción lineal usando una estimación de parámetro de codificación de predicción lineal; o el decodificador de audio puede configurarse para obtener un conjunto de parámetros de codificación de predicción lineal basándose en un conjunto de factores de escala usando una transformación.

De acuerdo con un séptimo aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a sexto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para obtener una información de tono 542 que describe un tono de la trama de audio codificada en la representación en el ámbito de la frecuencia que precede a la trama de audio perdida, y proporcionar la información de audio de ocultación de error 122; 382; 512 dependiendo de la información de tono.

De acuerdo con un octavo aspecto haciendo referencia de nuevo al séptimo aspecto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para obtener la información de tono 542

basándose en la señal de excitación en el ámbito del tiempo 532 derivada de la trama de audio codificada en la representación en el ámbito de la frecuencia 322 que precede a la trama de audio perdida.

- 5 De acuerdo con un noveno aspecto haciendo referencia de nuevo al octavo aspecto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 se puede configurar para evaluar una correlación cruzada de la señal de excitación en el ámbito del tiempo 532 o la señal en el ámbito del tiempo 522, para determinar una información de tono basto, y la ocultación de error puede configurarse para perfeccionar la información de tono grueso utilizando una búsqueda de bucle cerrado alrededor de un tono determinado por la información de tono grueso.
- 10 De acuerdo con un décimo aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a sexto, en el decodificador de audio, la ocultación de error puede configurarse para obtener una información de tono basándose en una información lateral de la información de audio codificada.
- 15 De acuerdo con un undécimo aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a sexto, en el decodificador de audio de acuerdo con la ocultación de error, se puede configurar para obtener una información de tono basándose en la información de tono disponible para una trama de audio decodificada previamente.
- 20 De acuerdo con un duodécimo aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a sexto, en el decodificador de audio, la ocultación de error puede configurarse para obtener una información de tono basándose en una búsqueda de tono realizada en una señal de ámbito de tiempo o en una señal residual.
- 25 De acuerdo con un decimotercer aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a duodécimo, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para copiar un ciclo de tono de la señal de excitación en el ámbito del tiempo 532 derivada de la trama de audio codificada en la representación en el ámbito de la frecuencia 322 que precede a la trama de audio perdida una o varias veces, con el fin de obtener una señal de excitación 572 para una síntesis 580 de la información de audio de ocultación de error 132; 382; 512.
- 30 De acuerdo con un decimocuarto aspecto haciendo referencia de nuevo al decimotercer aspecto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para filtrar paso bajo el ciclo de tono de la señal de excitación en el ámbito del tiempo 532 derivada de la representación en el ámbito del tiempo de la trama de audio codificada en la representación en el ámbito de la frecuencia 322 que precede a la trama de audio perdida usando un filtro dependiente de la tasa de muestreo, cuyo ancho de banda depende de la tasa de muestreo de la
- 35 trama de audio codificada en una representación en el ámbito de la frecuencia.
- 40 De acuerdo con un decimoquinto aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a decimocuarto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 se puede configurar para predecir un tono al final de una trama perdida, y la ocultación de error puede configurarse para adaptar la señal de excitación en el ámbito del tiempo 532, o una o varias copias de las mismas, al tono previsto, con el fin de obtener una señal de entrada 572 para una síntesis LPC 580.
- 45 De acuerdo con un decimosexto aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a decimoquinto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 se puede configurar para combinar una señal de excitación en el ámbito del tiempo extrapolada 552 y una señal de ruido 562, con el fin de obtener una señal de entrada 572 para una síntesis LPC 580, y la ocultación de error puede configurarse para realizar la síntesis LPC, en donde la síntesis de LPC se puede configurar para filtrar la señal de entrada 572 de la síntesis de LPC en función de los parámetros de codificación de predicción lineal, con el fin de obtener la información de audio de ocultación de error 132; 382; 512.
- 50 De acuerdo con un decimoséptimo aspecto haciendo referencia de nuevo al decimosexto aspecto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 se puede configurar para calcular una ganancia de la señal de excitación en el ámbito del tiempo extrapolado 552, que se usa para obtener la señal de entrada 572 para la síntesis LPC 580, usando una correlación en el ámbito del tiempo que se realiza basándose en una
- 55 representación en el ámbito del tiempo 122; 372; 378; 510 de la trama de audio codificada en la representación en el ámbito de la frecuencia 322 que precede a la trama de audio perdida, en donde un desfase de correlación se establece dependiendo de la información de tono obtenida basándose en la señal de excitación en el ámbito del tiempo 532, o usando una correlación en el ámbito de la excitación.
- 60 De acuerdo con un decimooctavo aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos decimosexto o decimoséptimo, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para filtrar a paso alto la señal de ruido 562 que se combina con la señal de excitación en el ámbito del tiempo extrapolada 552.

De acuerdo con un decimonoveno aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos decimotercero a decimoquinto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para cambiar la forma espectral de la señal de ruido 562 usando el filtro de énfasis previo en donde la señal de ruido se combina con la señal de excitación en el ámbito del tiempo extrapolada 552 si la trama de audio codificada en una representación en el ámbito de la frecuencia 322 precede a la trama de audio perdida es una trama de audio con voz o comprende un inicio.

De acuerdo con un vigésimo aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a decimonoveno, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para calcular una ganancia de la señal de ruido 562 en función de una correlación en el ámbito del tiempo que se realiza basándose en una representación en el ámbito del tiempo 122; 372; 378; 510 de la trama de audio codificado en la representación en el ámbito de la frecuencia 322 que precede a la trama de audio perdida.

De acuerdo con un vigésimo primer aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a vigésimo, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para modificar una señal de excitación en el ámbito del tiempo 532 obtenida basándose en una o varias tramas de audio que preceden a una trama de audio perdida, con el fin de obtener la información de audio de ocultación de error 132; 382; 512.

De acuerdo con un aspecto vigésimo segundo haciendo referencia de nuevo al aspecto vigésimo primero, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para usar una o varias copias modificadas de la señal de excitación en el ámbito del tiempo 532 obtenida basándose en una o varias tramas de audio que preceden a una trama de audio perdida, con el fin de obtener la información de ocultación de error 132; 382; 512.

De acuerdo con un aspecto vigésimo tercero haciendo referencia de nuevo a uno cualquiera de los aspectos vigésimo primero o vigésimo segundo, en el decodificador de audio 100; 300, la ocultación de error 132; 380; 500 puede configurarse para modificar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o varias tramas de audio que preceden a una trama de audio perdida, o una o varias copias de las mismas, para reducir así un componente periódico de la información de audio de ocultación de error 132; 382; 512 a lo largo del tiempo.

De acuerdo con un vigésimo cuarto aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos vigésimo primero a vigésimo tercero, en el decodificador de audio 100; 300, la ocultación de error 132; 380; 500 puede configurarse para escalar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o varias tramas de audio que preceden a la trama de audio perdida, o una o varias copias de las mismas, para modificar así la señal de excitación en el ámbito del tiempo.

De acuerdo con un vigésimo quinto aspecto haciendo referencia de nuevo al vigésimo tercer o vigésimo cuarto aspecto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para reducir gradualmente la ganancia aplicada para escalar la señal de excitación en el ámbito del tiempo 532 obtenida basándose en una o más tramas de audio que preceden a una trama de audio perdida, o la una o varias copias de las mismas.

De acuerdo con un trigésimoquinto aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos vigésimo tercero a vigésimo quinto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para ajustar una velocidad usada para reducir gradualmente la ganancia aplicada para escalar la señal de excitación en el ámbito del tiempo 532 obtenida basándose en una o más tramas de audio que preceden a una trama de audio perdida, o la una o varias copias de las mismas, en dependencia de uno o varios parámetros de una o varias tramas de audio que preceden a la trama de audio perdida, y/o en dependencia de un número de tramas de audio perdidas consecutivas.

De acuerdo con un vigésimo séptimo aspecto haciendo referencia de nuevo al vigésimo quinto o vigésimo sexto aspecto, en el decodificador de audio 100; 300, la ocultación de error puede configurarse para ajustar la velocidad usada reducir gradualmente la ganancia aplicada para escalar la señal de excitación en el ámbito del tiempo 532 obtenida basándose en una o más tramas de audio que preceden a una trama de audio perdida, o la una o varias copias de las mismas, en dependencia de la duración de un período de tono de la señal de excitación en el ámbito del tiempo 532, de tal manera que una entrada de señal de excitación en el ámbito del tiempo en una síntesis LPC se desvanece más rápido para señales que tienen una duración más corta del período de tono cuando se comparan con señales que tienen una duración más grande del período de tono.

De acuerdo con un vigésimo octavo aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos vigésimo quinto a vigésimo séptimo, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para ajustar la velocidad usada reducir gradualmente la ganancia aplicada para escalar la señal

de excitación en el ámbito del tiempo 532 obtenida basándose en una o más tramas de audio que preceden a una trama de audio perdida, o la una o varias copias de las mismas, en dependencia de un resultado de un análisis de tono 540 o una predicción de tono, de tal manera que una componente determinista de una entrada de señal de excitación en el ámbito del tiempo 572 en una síntesis LPC 580 puede desvanecerse más rápido para señales que tienen un cambio de tono mayor por unidad de tiempo en comparación con señales que tienen un cambio de tono más pequeño por unidad de tiempo, de tal manera que un componente determinista de una entrada de señal de excitación en el ámbito del tiempo 572 en una síntesis LPC 580 puede desconocerse más rápido para señales para las que falla una predicción de tono cuando se compara con señales para las cuales la predicción de tono tiene éxito.

De acuerdo con un vigésimo noveno aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos vigésimo primero a vigésimo octavo, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para escalar en tiempo la señal de excitación en el ámbito del tiempo 532 obtenida en base a uno o varias tramas de audio que preceden a una trama de audio perdida, o la una o varias copias de las mismas, en dependencia de una predicción 540 de un tono para el tiempo de una o varias tramas de audio perdidas.

De acuerdo con un trigésimo aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a vigésimo noveno, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para proporcionar la información de audio de ocultación de error 132; 382; 512 durante un tiempo que sea más largo que la duración temporal de una o varias tramas de audio perdidas.

De acuerdo con un trigésimo primer aspecto haciendo referencia de nuevo al trigésimo aspecto, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para realizar una superposición y adición 390; 590 de la información de audio de ocultación de error 132; 382; 512 y una representación en el ámbito del tiempo 122; 372, 378; 512 de una o varias tramas de audio recibidas correctamente después de una o varias tramas de audio perdidas.

De acuerdo con un trigésimosegundo aspecto haciendo referencia de nuevo a uno cualquiera de los aspectos primero a trigésimo primero, en el decodificador de audio 100; 300, la ocultación de error 130; 380; 500 puede configurarse para derivar la información de audio de ocultación de error 132; 382; 512 basándose en al menos tres tramas o ventanas parcialmente superpuestas que preceden a una trama de audio perdida o una ventana perdida.

De acuerdo con un trigésimo tercer aspecto, un método 900 para proporcionar una información de audio decodificada basándose en una información de audio codificada puede comprender: proporcionar 910 una información de audio de ocultación de error para ocultar una pérdida de una trama de audio después de una trama de audio codificada en una representación de ámbito de frecuencia utilizando una señal de excitación en el ámbito del tiempo.

Un trigésimo cuarto aspecto puede proporcionar un programa informático para realizar el método de acuerdo con un trigésimo tercer aspecto cuando el programa informático se ejecuta en un ordenador.
Bibliografía:

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

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

[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] "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

[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 (100; 300) para proporcionar una información de audio decodificada (112; 312) basándose en una información de audio codificada (110; 310), comprendiendo el decodificador de audio:

un medio de ocultación de error (130; 380; 500) configurado para proporcionar una información de audio de ocultación de error (132; 382; 512) para ocultar una pérdida de una trama de audio después de una trama de audio codificada en una representación de ámbito de frecuencia (322) utilizando una señal de excitación en el ámbito del tiempo (532);

en donde el medio de ocultación de error (130; 380; 500) está configurado para modificar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o varias tramas de audio que preceden a una trama de audio perdida, con el fin de obtener la información de audio de ocultación de error (132; 382; 512);

en donde el medio de ocultación de error (132; 380; 500) está configurado para modificar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o varias tramas de audio que preceden a una trama de audio perdida, o una o varias copias de las mismas, para reducir así un componente periódico de la información de audio de ocultación de error (132; 382; 512) a lo largo del tiempo;

en donde el medio de ocultación de error (130; 380; 500) está configurado para reducir gradualmente la ganancia aplicada para escalar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o más tramas de audio que preceden a una trama de audio perdida, o la una o varias copias de las mismas; el decodificador de audio está caracterizado porque el medio de ocultación de error está configurado para ajustar la velocidad utilizada para gradualmente reducir gradualmente la ganancia aplicada para escalar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o más tramas de audio que preceden a una trama de audio perdida, o la una o varias copias de las mismas, en dependencia de la duración de un período de tono de la señal de excitación en el ámbito del tiempo (532), de tal manera que una entrada de señal de excitación en el ámbito del tiempo en una síntesis LPC se desvanece más rápido para señales que tienen una duración más corta del período de tono cuando se comparan con señales que tienen una duración más grande del período de tono.

2. Un decodificador de audio (100; 300) para proporcionar una información de audio decodificada (112; 312) basándose en una información de audio codificada (110; 310), comprendiendo el decodificador de audio:

un medio de ocultación de error (130; 380; 500) configurado para proporcionar una información de audio de ocultación de error (132; 382; 512) para ocultar una pérdida de una trama de audio después de una trama de audio codificada en una representación de ámbito de frecuencia (322) utilizando una señal de excitación en el ámbito del tiempo (532);

en donde el medio de ocultación de error (130; 380; 500) está configurado para modificar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o varias tramas de audio que preceden a una trama de audio perdida, con el fin de obtener la información de audio de ocultación de error (132; 382; 512);

en donde el medio de ocultación de error (132; 380; 500) está configurado para modificar la señal de excitación en el ámbito del tiempo (532) obtenida en base a una o varias tramas de audio que preceden a una trama de audio perdida, o una o varias copias de las mismas, para reducir así un componente periódico de la información de audio de ocultación de error (132; 382; 512) a lo largo del tiempo, o

en donde el medio de ocultación de error (132; 380; 500) está configurado para escalar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o varias tramas de audio que preceden a la trama de audio perdida, o una o varias copias de las mismas, para modificar así la señal de excitación en el ámbito del tiempo;

el decodificador de audio está caracterizado porque el medio de ocultación de error (130; 380; 500) está configurado para ajustar la velocidad utilizada para reducir gradualmente una ganancia aplicada para escalar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o más tramas de audio que preceden a una trama de audio perdida, o la una o varias copias de las mismas, en dependencia de un resultado de un análisis de tono (540) o una predicción de tono,

de tal manera que una componente determinista de una entrada de señal de excitación en el ámbito del tiempo (572) en una síntesis LPC (580) se desvanece más rápido para señales que tienen un cambio de tono mayor por unidad de tiempo en comparación con señales que tienen un cambio de tono más pequeño por unidad de tiempo, y/o

de tal manera que un componente determinista de una entrada de señal de excitación en el ámbito del tiempo (572) en una síntesis LPC (580) se desvanece más rápido para señales para las que falla una predicción de tono cuando se compara con señales para las cuales la predicción de tono tiene éxito.

3. Un método (900) para proporcionar una información de audio decodificada basándose en una información de audio codificada, comprendiendo el método:

proporcionar (910) una información de audio de ocultación de error para ocultar una pérdida de una trama de audio después de una trama de audio codificada en una representación en el ámbito de la frecuencia utilizando

una señal de excitación en el ámbito del tiempo en donde la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o varias tramas de audio que preceden a una trama de audio perdida se modifica, con el fin de obtener la información de audio de ocultación de error (132; 382; 512);
 en donde la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o más tramas de audio que preceden a una trama de audio perdida, o una o varias copias de las mismas, se modifica para reducir así un componente periódico de la información de audio de ocultación de error (132; 382; 512) a lo largo del tiempo;
 en donde gradualmente la ganancia aplicada para escalar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o más tramas de audio que preceden a una trama de audio perdida, o la una o varias copias de las mismas, se reduce gradualmente;
 el método está caracterizado porque la velocidad utilizada para reducir gradualmente la ganancia aplicada para escalar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o más tramas de audio que preceden a una trama de audio perdida, o la una o varias copias de las mismas, se ajusta en dependencia de la duración de un período de tono de la señal de excitación en el ámbito del tiempo (532), de tal manera que una entrada de señal de excitación en el ámbito del tiempo en una síntesis LPC se desvanece más rápido para señales que tienen una duración más corta del período de tono cuando se comparan con señales que tienen una duración más grande del período de tono.

4. Un método (900) para proporcionar una información de audio decodificada basándose en una información de audio codificada, comprendiendo el método:

proporcionar (910) una información de audio de ocultación de error para ocultar una pérdida de una trama de audio después de una trama de audio codificada en una representación de ámbito de frecuencia utilizando una señal de excitación en el ámbito del tiempo;
 en donde el método comprende modificar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o varias tramas de audio que preceden a una trama de audio perdida, con el fin de obtener la información de audio de ocultación de error (132; 382; 512),
 en donde la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o más tramas de audio que preceden a una trama de audio perdida, o una o varias copias de las mismas, se modifica para reducir así un componente periódico de la información de audio de ocultación de error (132; 382; 512) a lo largo del tiempo, o
 en donde la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o más tramas de audio que preceden a la trama de audio perdida, o una o varias copias de las mismas, se escala para modificar así la señal de excitación en el ámbito del tiempo; el método está caracterizado porque la velocidad utilizada para reducir gradualmente la ganancia aplicada para escalar la señal de excitación en el ámbito del tiempo (532) obtenida basándose en una o más tramas de audio que preceden a una trama de audio perdida, o la una o varias copias de las mismas, se ajusta en dependencia de un resultado de un análisis de tono (540) o una predicción de tono, de tal manera que una componente determinista de una entrada de señal de excitación en el ámbito del tiempo (572) en una síntesis LPC (580) se desvanece más rápido para señales que tienen un cambio de tono mayor por unidad de tiempo en comparación con señales que tienen un cambio de tono más pequeño por unidad de tiempo, y/o de tal manera que un componente determinista de una entrada de señal de excitación en el ámbito del tiempo (572) en una síntesis LPC (580) se desvanece más rápido para señales para las que falla una predicción de tono cuando se compara con señales para las cuales la predicción de tono tiene éxito

5. Un programa informático para realizar el método de acuerdo con una de las reivindicaciones 3 a 4 cuando el programa informático se ejecuta en un ordenador.

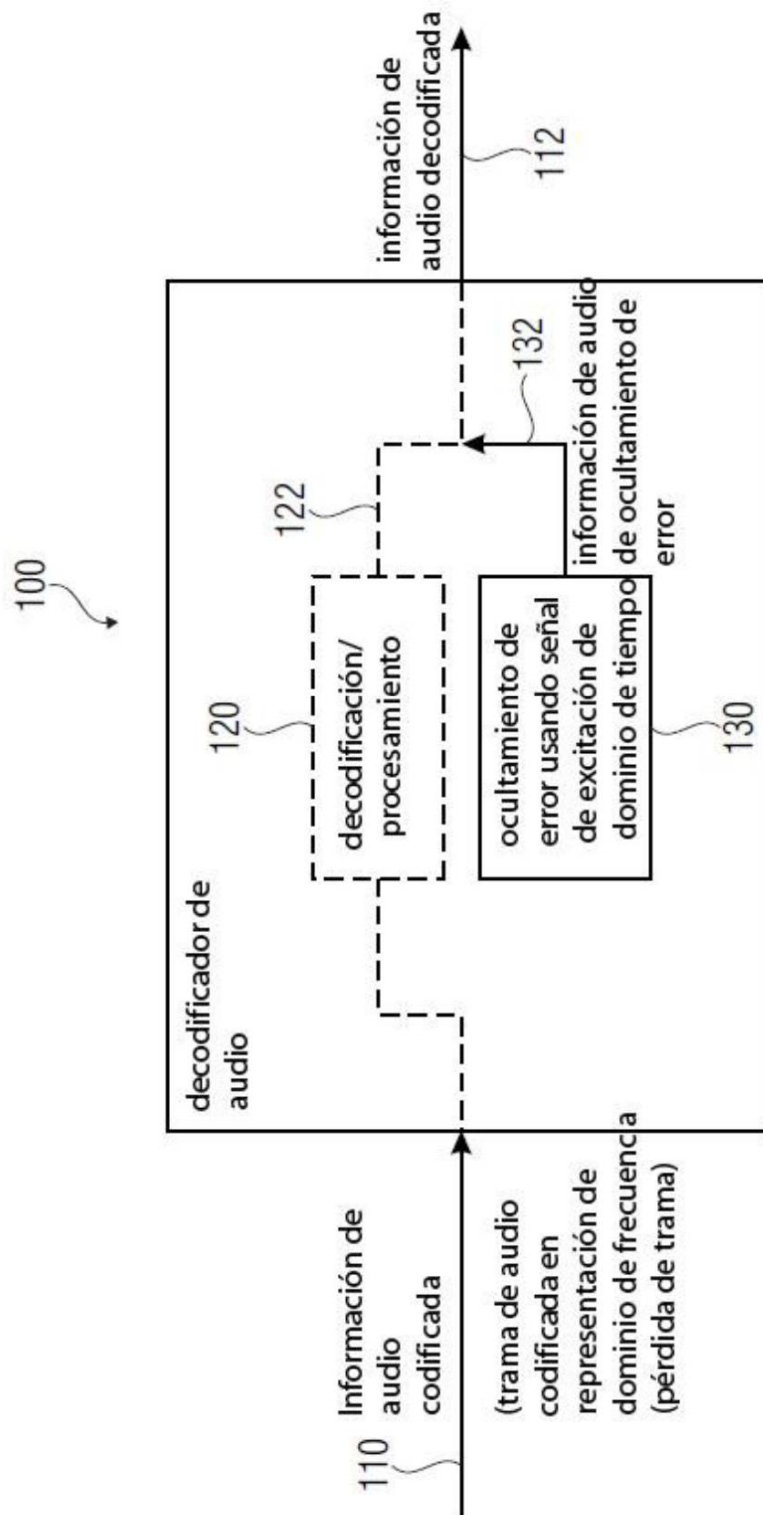


FIG. 1

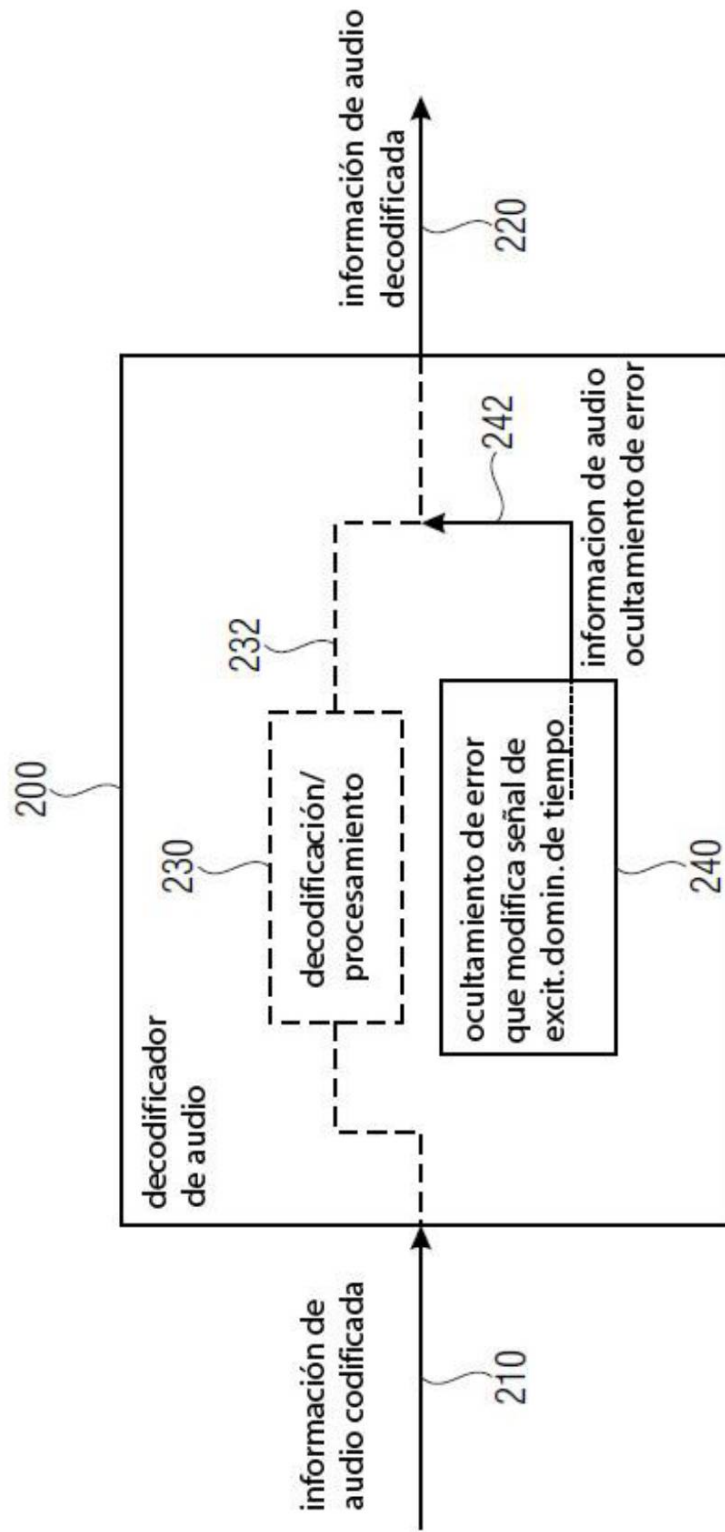


FIG. 2

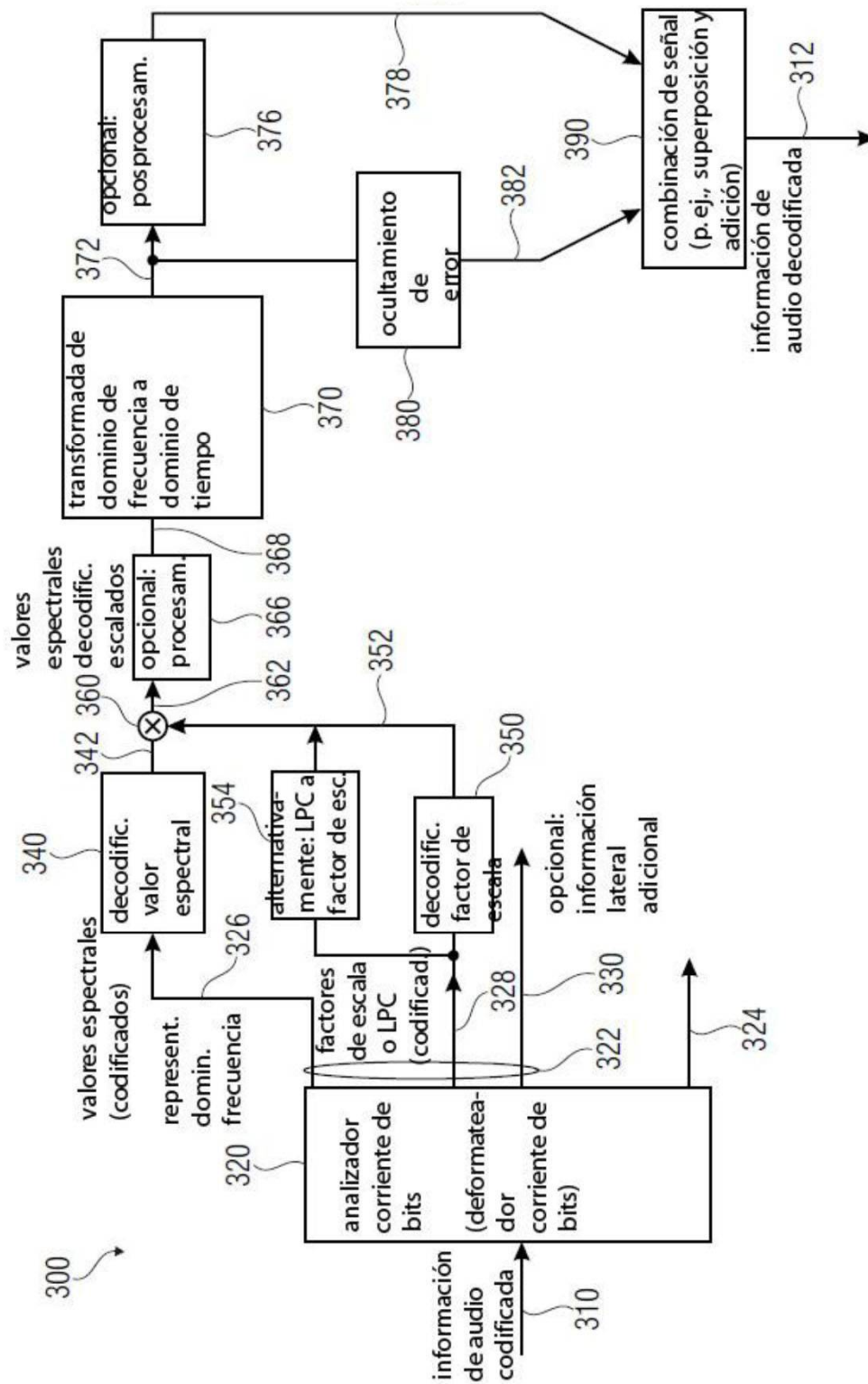
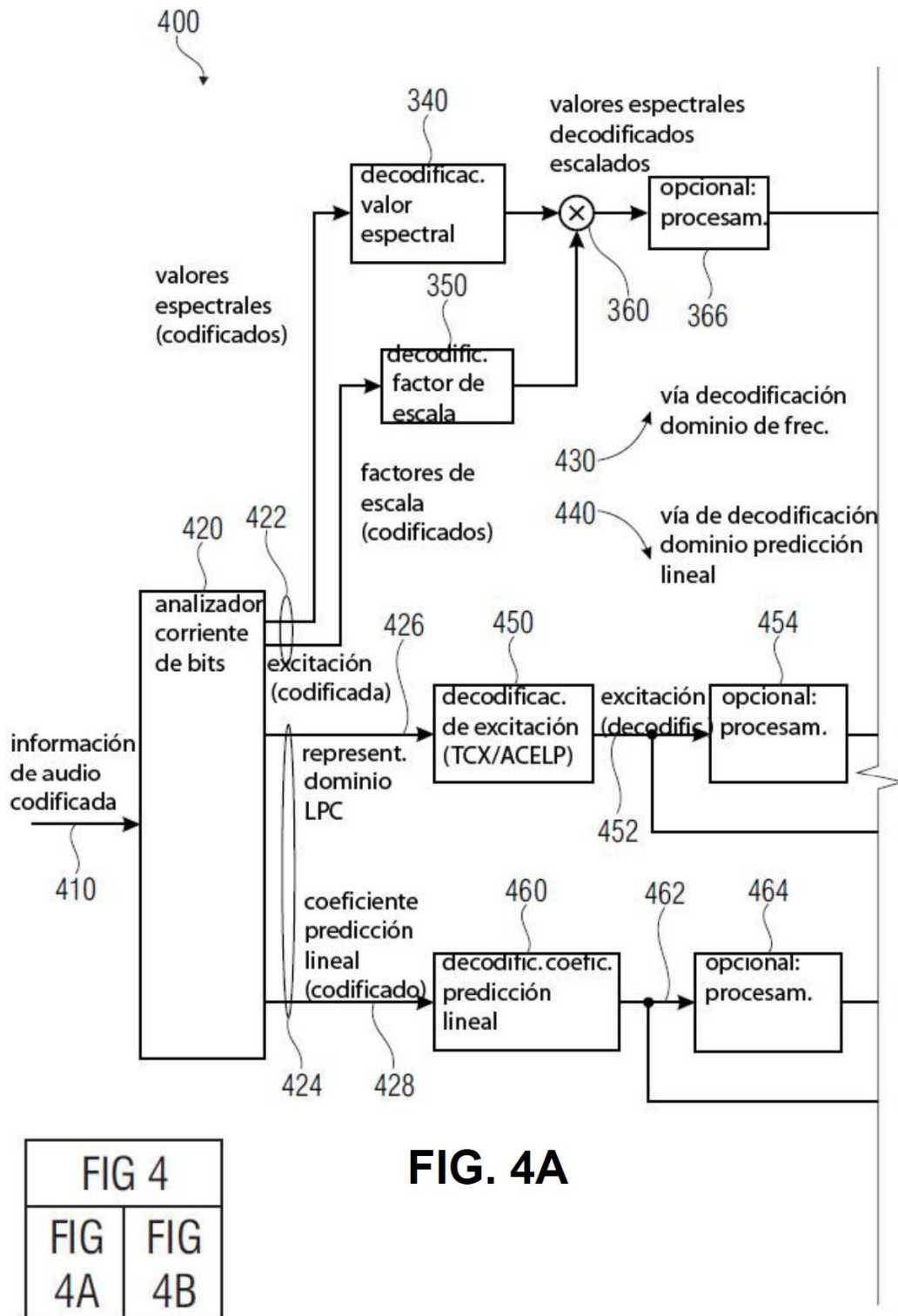


FIG. 3



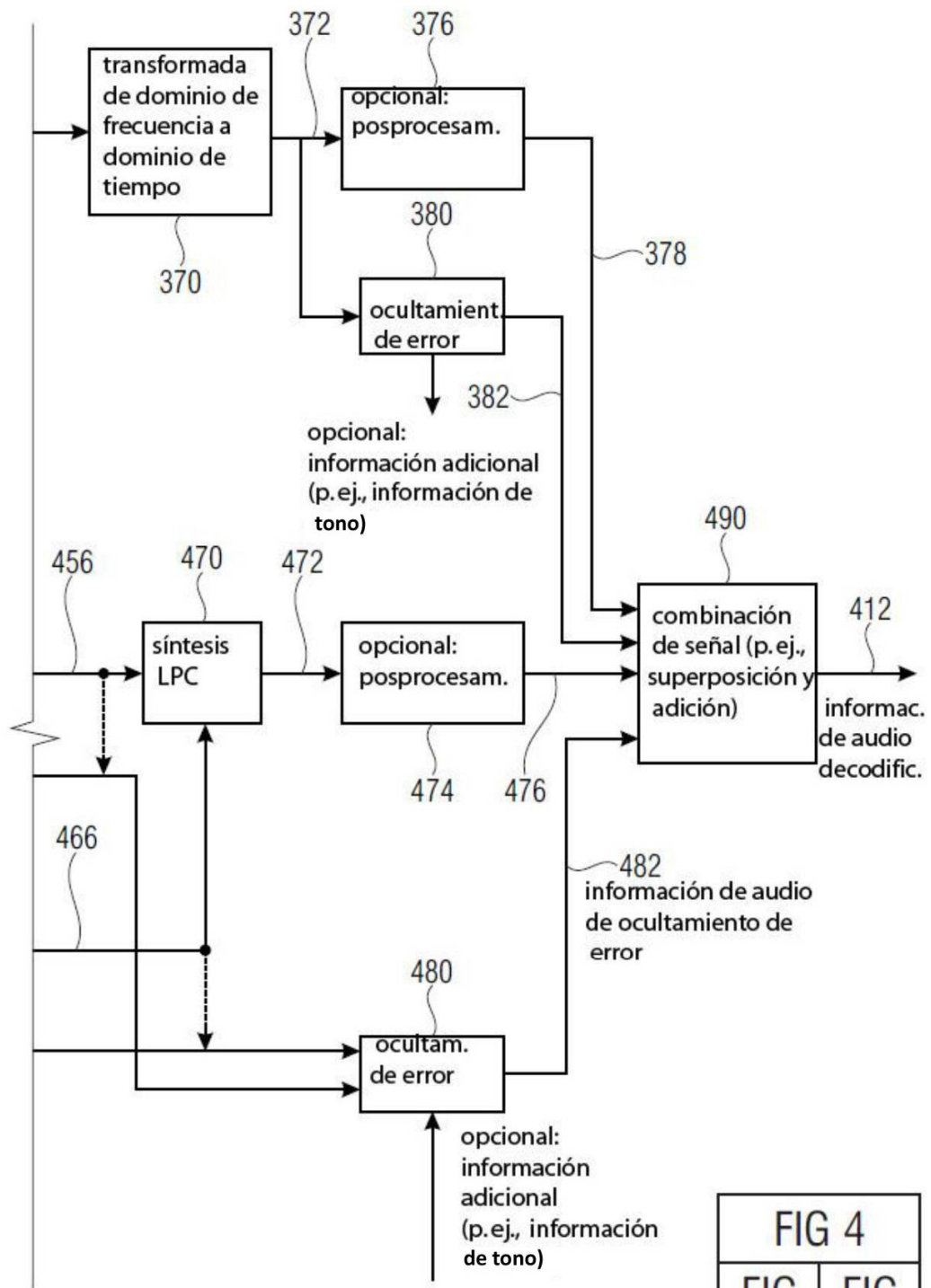
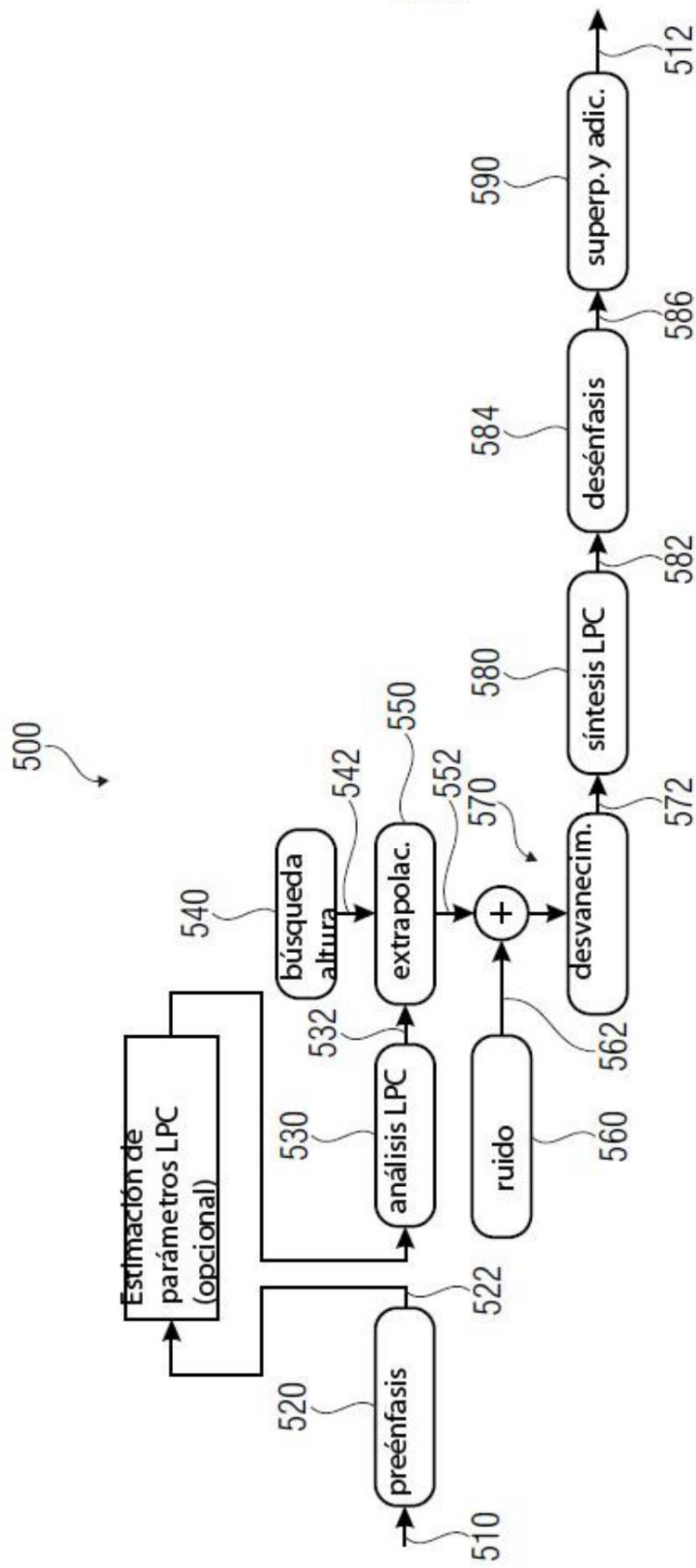


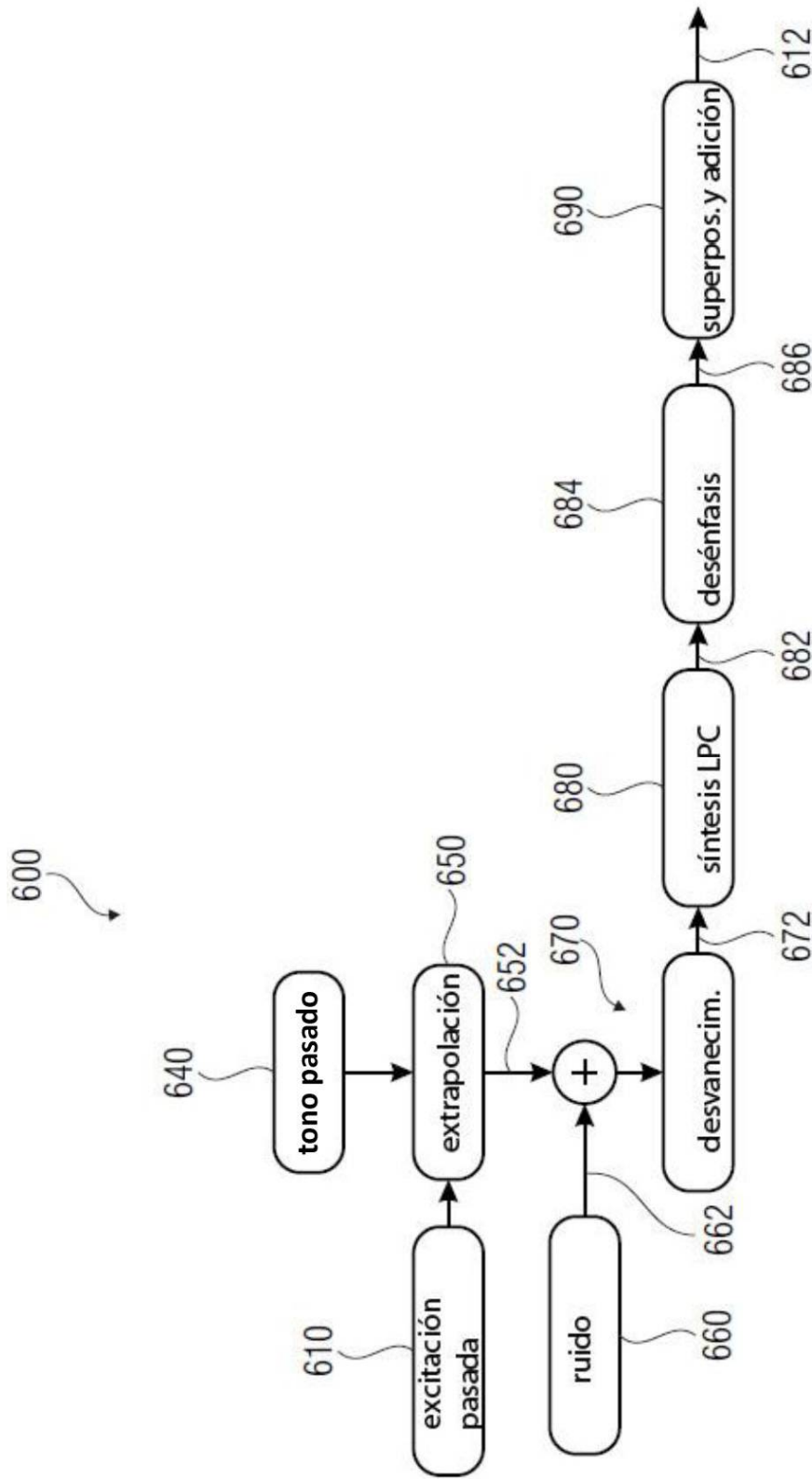
FIG. 4B

FIG 4	
FIG 4A	FIG 4B



Reseña de ocultamiento de dominio de tiempo para decodific. de transformada

FIG. 5



Reseña de ocultamiento dominio de tiempo para códec de conmutación

FIG. 6

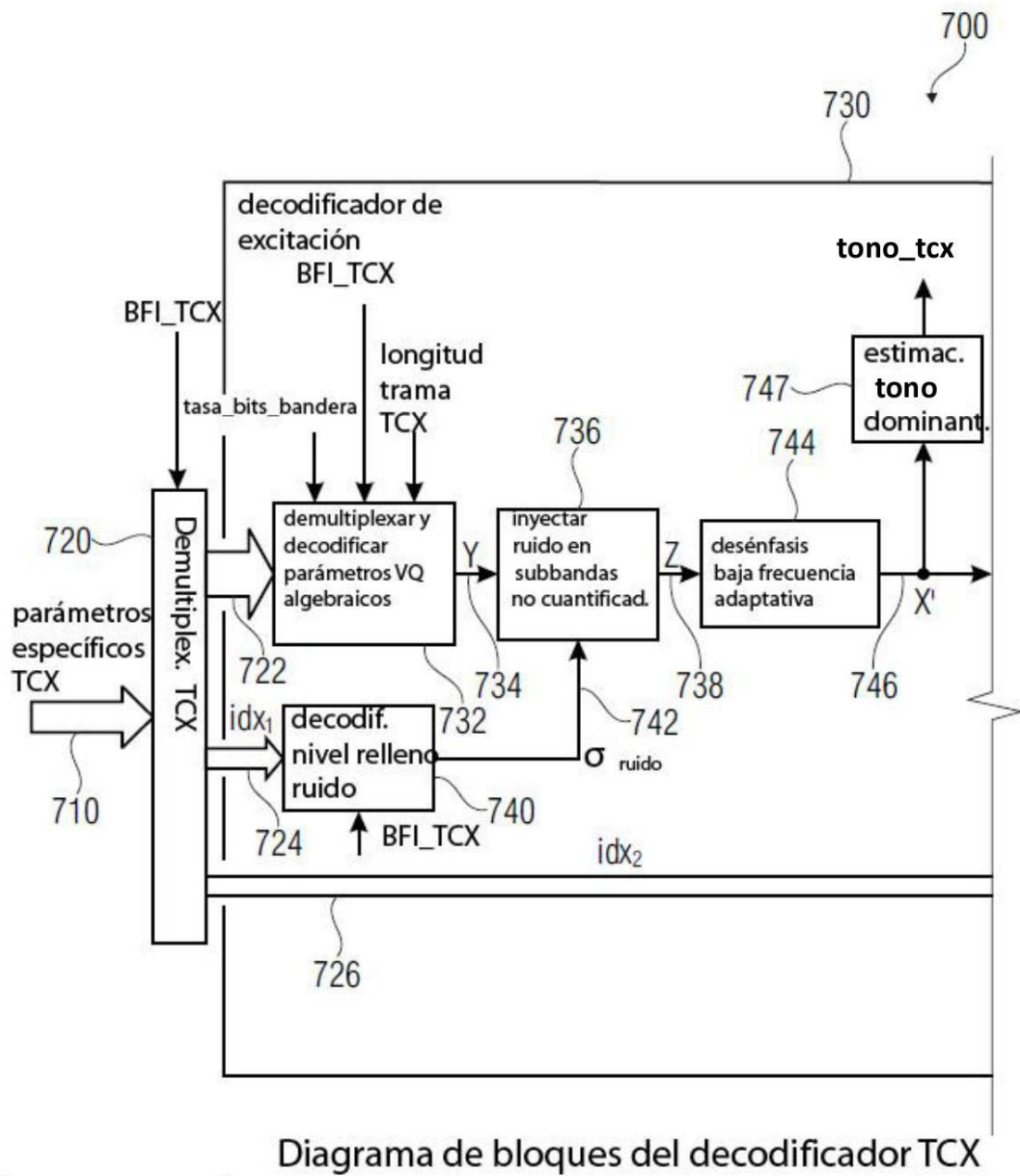


Diagrama de bloques del decodificador TCX

FIG 7	
FIG 7A	FIG 7B

FIG. 7A

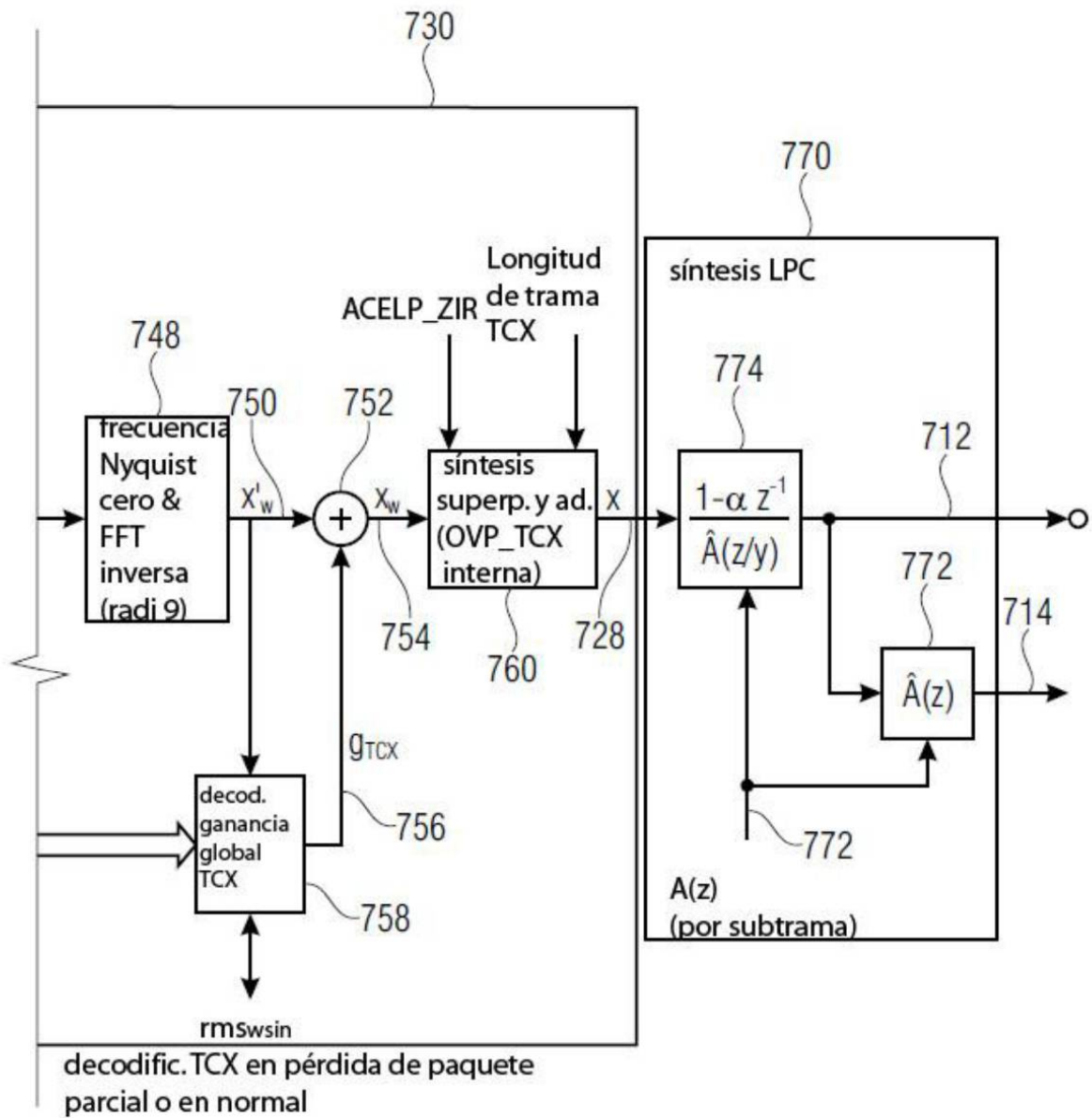
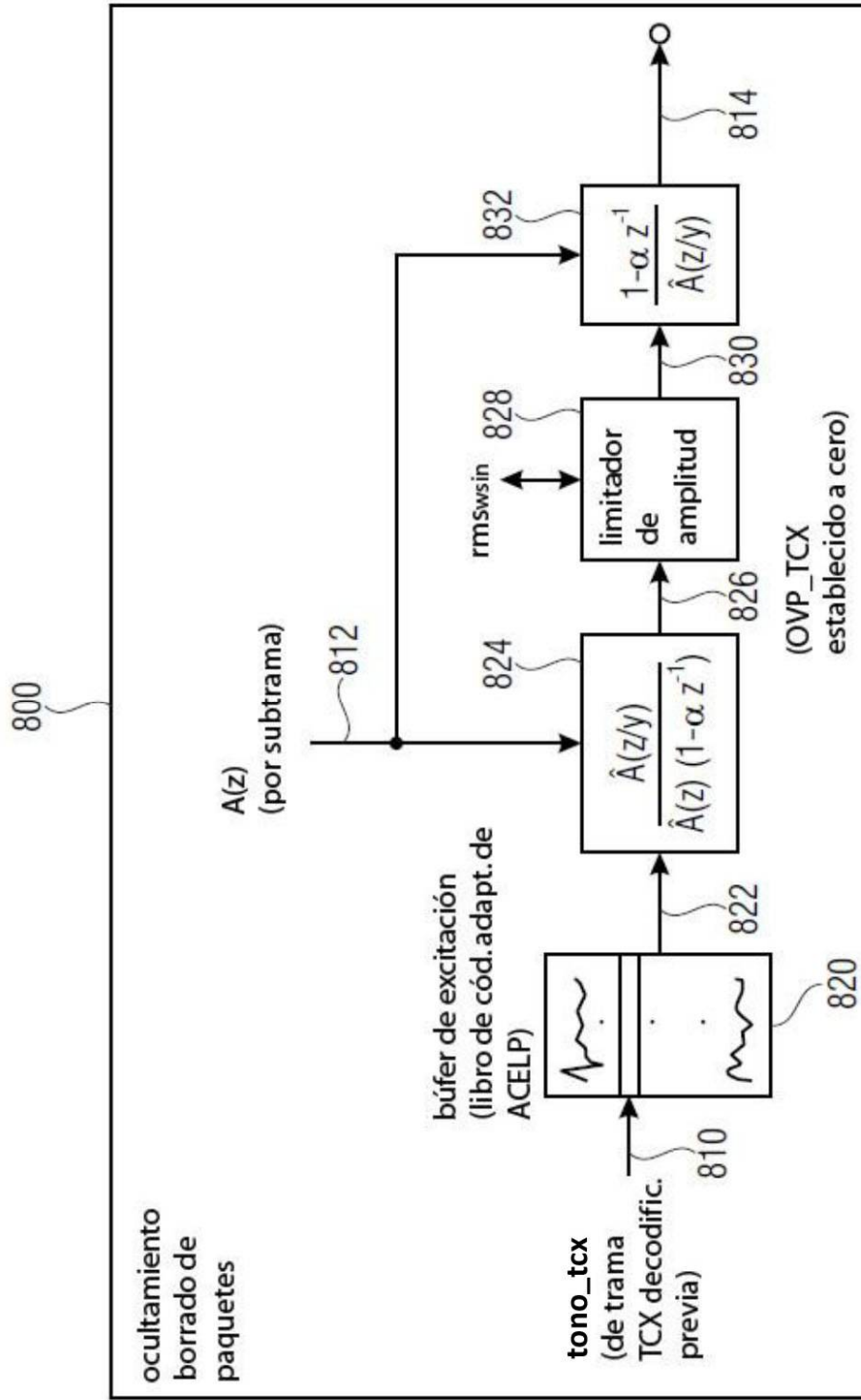


Diagrama de bloques del decodificador TCX

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

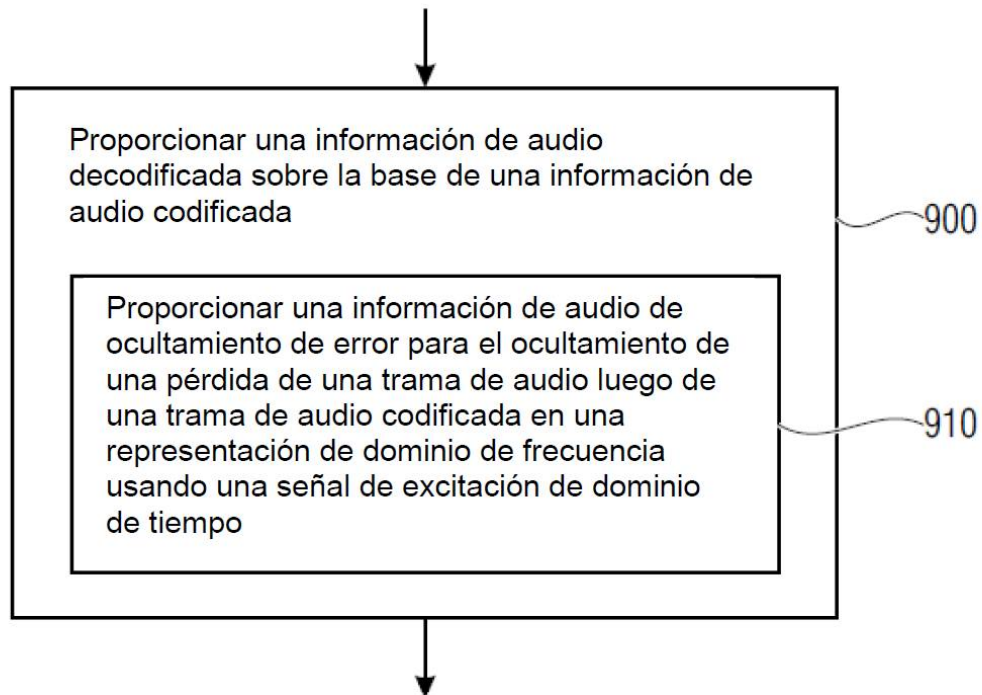


FIG. 9

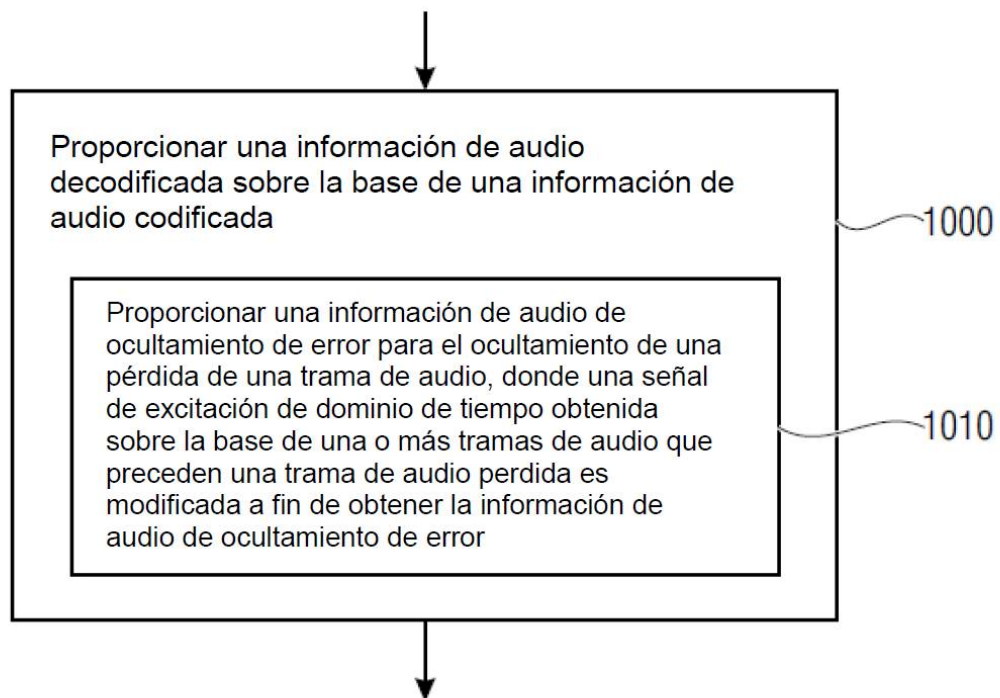


FIG. 10

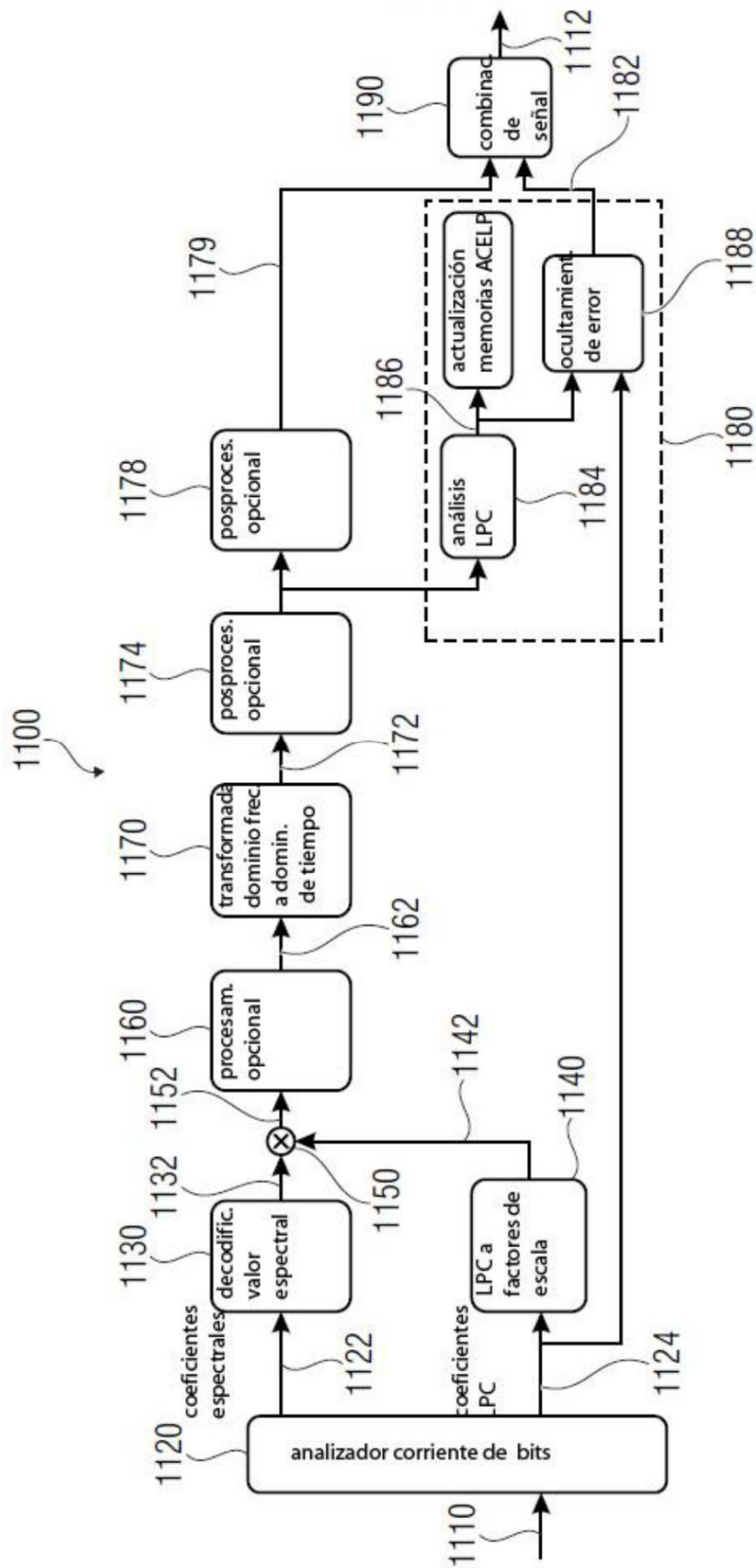


FIG. 11