

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 785 000**

51 Int. Cl.:

G10L 19/005 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **08.06.2015** **E 18167282 (5)**

97 Fecha y número de publicación de la concesión europea: **22.01.2020** **EP 3367380**

54 Título: **Gestión de errores de trama de ráfaga**

30 Prioridad:

13.06.2014 US 201462011598 P

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

02.10.2020

73 Titular/es:

**TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)
(100.0%)
164 83 Stockholm, SE**

72 Inventor/es:

BRUHN, STEFAN

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 785 000 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

Gestión de errores de trama de ráfaga

5 CAMPO TÉCNICO

Este documento se refiere a la codificación de audio y a la generación de una señal de sustitución en el receptor como un reemplazo para tramas de señal perdidas, borradas o deterioradas en caso de errores de transmisión. La técnica descrita en la presente memoria podría ser parte de un códec y/o de un decodificador, pero también podría implementarse en un módulo de mejora de la señal después de un decodificador. La técnica puede utilizarse con ventaja en un receptor.

Particularmente, las realizaciones presentadas en la presente memoria se refieren a la ocultación de pérdida de trama, y particularmente a un método y a una entidad de recepción para la ocultación de pérdida de trama.

15 ANTECEDENTES

Muchos sistemas de comunicación modernos transmiten señales de voz y audio en tramas, lo que significa que el lado emisor primero organiza la señal en segmentos o tramas cortas de, por ejemplo, 20-40 ms, que posteriormente se codifican y transmiten como una unidad lógica en, por ejemplo, un paquete de transmisión. El receptor decodifica cada una de estas unidades y reconstruye las tramas de señal correspondientes, que a su vez salen finalmente como una secuencia continua de muestras de señal reconstruidas. Antes de la codificación normalmente hay una conversión de analógico a digital (A/D), que convierte la señal de voz o audio analógica de un micrófono en una secuencia de muestras de audio. En cambio, en el extremo receptor, normalmente hay una conversión final de digital a analógico (D/A) que convierte la secuencia de muestras de señal digital reconstruidas en una señal analógica continua en el tiempo para su reproducción en altavoz.

Casi cualquier tipo de sistema de transmisión para señales de voz y audio puede sin embargo sufrir de errores de transmisión. Esto puede conducir a la situación de que una o varias de las tramas transmitidas no estén disponibles en el receptor para su reconstrucción. En ese caso, el decodificador tiene que generar una señal de sustitución para cada una de las tramas borradas, es decir, no disponibles. Esto es hecho en la llamada unidad de ocultación de error o de pérdida de trama del decodificador de señal del lado receptor. El propósito de la ocultación de pérdida de trama es hacer la pérdida de trama tan inaudible como sea posible y por lo tanto mitigar el impacto de la pérdida de trama en la calidad de la señal reconstruida todo lo posible.

Un enfoque ilustrativo para la ocultación de pérdida de trama se describe en US 2012/0010882 A1.

Un método reciente de ocultación de pérdida de trama es el llamado "et al" 'Phase ECU'. Este es un método que proporciona, particularmente, alta calidad de la señal de audio restaurada después de pérdida de paquete o trama en el caso de que la señal sea una señal de música. Hay también un método de control descrito en una aplicación previa que controla el comportamiento de un método de ocultación de pérdida de trama de tipo Phase ECU en respuesta a, por ejemplo, propiedades (estadísticas) de las pérdidas de trama.

La ruptura de las pérdidas de trama se utiliza como un indicador en el método de control en el que puede adaptarse la respuesta a un método de ocultación de pérdida de trama como Phase ECU. En términos generales, la ruptura de las pérdidas de trama significa que se producen varias pérdidas de trama seguidas, lo que dificulta que el método de ocultación de pérdida de trama utilice partes válidas de la señal decodificada recientemente para su funcionamiento. Más específicamente, un indicador típico de ruptura de pérdida de trama de vanguardia es el número n de pérdidas de trama consecutivas observadas. Este número puede mantenerse en un contador que se incrementa en uno por cada nueva pérdida de trama y que se restablece a cero tras la recepción de una trama válida.

Un método de adaptación específico de un método de ocultación de pérdida de trama como Phase ECU en respuesta a la ruptura de pérdida de trama es el ajuste selectivo de frecuencia de las fases o de las magnitudes de espectro de un espectro de la trama de sustitución $Z(m)$, siendo m un índice de frecuencia de una transformada del dominio de la frecuencia como la Transformada Discreta de Fourier (DFT). La adaptación de la magnitud se hace con un factor de atenuación $\alpha(m)$ que escala el coeficiente de la transformada de la frecuencia al índice m al incrementar el contador de ráfagas de pérdida de trama, n , hasta 0. La adaptación de la fase se realiza mediante el incremento de la aleatorización aditiva de la fase (con un incremento del componente de fase aleatoria $\vartheta(m)$) del coeficiente de la transformada de la frecuencia al índice m .

Por lo tanto, si el espectro original de la trama de sustitución del Phase ECU sigue una expresión como $Z(m) = Y(m) \cdot e^{j\theta_k}$, entonces el espectro adaptado de la trama de sustitución sigue una expresión como $Z(m) = \alpha(m) \cdot Y(m) \cdot e^{j(\theta_k + \vartheta(m))}$.

En la presente memoria la fase θ_k con $k = 1 \dots K$ es una función del índice m y de los picos espectrales K identificados por el método de Phase ECU, e $Y(m)$ es una representación del dominio de la frecuencia (espectro) de una trama de la señal de audio recibida previamente.

A pesar de las ventajas del método de adaptación descrito anteriormente del Phase ECU en condiciones de pérdida de trama de ráfaga, todavía hay defectos de calidad en el caso de una ráfaga de pérdida muy larga, por ejemplo,

5 cuando n mayor o igual a 5. En ese caso la calidad de la señal de audio reconstruida puede, por ejemplo, sufrir de artefactos tonales, a pesar de la aleatorización de fase realizada. Al mismo tiempo el incremento de la atenuación de magnitud puede reducir estos defectos audibles. Sin embargo, la atenuación de la señal puede ser percibida, para ráfagas largas de pérdida de trama, como silenciamiento o pérdidas de señal. Esto puede afectar de nuevo a la calidad general de, por ejemplo, música o el ruido ambiente de una señal de voz ya que dichas señales son sensibles a variaciones de nivel demasiado fuertes.

Por lo tanto, todavía existe la necesidad para la ocultación de pérdida de trama mejorada.

10 **Compendio**
Un objeto de las realizaciones en la presente memoria es proporcionar ocultación de pérdida de trama eficiente.

Según un primer aspecto se presenta un método para ocultación de pérdida de trama según la reivindicación 1.

15 De manera ventajosa, esto proporciona ocultación de pérdida de trama eficiente.

Según un segundo aspecto se presenta una entidad receptora para la ocultación de pérdida de trama según la reivindicación 3.

20 Debe señalarse que cualquier característica de los aspectos primero y segundo puede aplicarse a cualquier otro aspecto, donde sea apropiado. Asimismo, cualquier ventaja del primer aspecto puede igualmente aplicar al segundo aspecto, respectivamente, y viceversa. Otros objetivos, características y ventajas de las realizaciones adjuntas serán evidentes a partir de la siguiente descripción detallada, a partir de las reivindicaciones dependientes adjuntas así como de los dibujos.

25 En general, todos los términos utilizados en las reivindicaciones deben interpretarse según su significado ordinario en el campo técnico, a menos que se defina explícitamente lo contrario en la presente memoria. Todas las referencias a "un/una/el elemento, aparato, componente, medios, paso, etc." deben interpretarse abiertamente como referentes a al menos una instancia del elemento, aparato, componente, medio, paso, etc., a menos que se indique explícitamente lo contrario. Los pasos de cualquier método descrito en la presente memoria no tienen que realizarse en el orden exacto descrito, a menos que se indique explícitamente.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

35 El concepto inventivo se describe ahora, a modo de ejemplo, con referencia a los dibujos adjuntos, en los que:

La Figura 1 es un diagrama esquemático que ilustra un sistema de comunicaciones según las realizaciones;
La Figura 2 es un diagrama esquemático que muestra las unidades funcionales de una entidad receptora según una realización;

40 La Figura 3 ilustra esquemáticamente la inserción de la trama de sustitución según una realización;
La Figura 4 es un diagrama esquemático que muestra las unidades funcionales de una entidad receptora según una realización;
Las Figuras 5, 6, y 7 son diagramas de flujo de métodos según las realizaciones;

La Figura 8 es un diagrama esquemático que muestra las unidades funcionales de una entidad receptora según una realización;

45 La Figura 9 es un diagrama esquemático que muestra los módulos funcionales de una entidad receptora según una realización; y

La Figura 10 muestra un ejemplo de un producto de programa informático que comprende un medio legible por ordenador según una realización.

50 DESCRIPCIÓN DETALLADA

El concepto inventivo se describirá ahora más completamente a continuación con referencia a los dibujos adjuntos, en los que se muestran ciertas realizaciones del concepto inventivo. Este concepto inventivo puede, sin embargo, ser encarnado de muchas formas diferentes y no debe interpretarse como limitado a las realizaciones establecidas en la presente memoria; más bien, estas realizaciones son proporcionadas a modo de ejemplo para que esta descripción sea exhaustiva y completa, y transmitirá completamente el alcance del concepto inventivo a los expertos en la técnica. Los números iguales se refieren a elementos similares en toda la descripción. Cualquier paso o característica ilustrado por líneas discontinuas debe considerarse como opcional.

60 Como se señaló anteriormente, las realizaciones presentadas en la presente memoria se refieren a la ocultación de pérdida de trama, y particularmente a un método, una entidad receptora, un programa informático, y un producto de programa informático para la ocultación de pérdida de trama.

65 La Figura 1 ilustra esquemáticamente un sistema de comunicación 100 en el que una entidad transmisora (TX) 101 se comunica con una entidad receptora (RX) 103 sobre un canal 102. Se supone que el canal 102 provoca que las tramas, o paquetes, transmitidos por la entidad TX 101 a la entidad RX 103 se pierdan. La entidad receptora se supone que puede funcionar para decodificar audio, como voz o música, y para comunicarse con otros nodos o entidades, por

ejemplo, en el sistema de comunicación 100. La entidad receptora puede ser un códec, un decodificador, un dispositivo inalámbrico y/o un dispositivo estacionario; de hecho podría ser cualquier tipo de unidad en la que es deseable gestionar errores de trama de ráfaga para las señales de audio. Podría ser, por ejemplo, un teléfono inteligente, una tablet, un ordenador o cualquier otro dispositivo capaz de una comunicación cableada y/o inalámbrica y de decodificación de audio. La entidad receptora puede ser llamada, por ejemplo, nodo receptor o disposición receptora.

La Figura 2 ilustra esquemáticamente módulos funcionales de una entidad RX 200 conocida, configurada para gestionar pérdidas de trama. Un flujo de bits entrante es decodificado por el decodificador 201 para formar una señal reconstruida y si una pérdida de trama no es detectada se proporciona esta señal reconstruida como salida de la entidad RX 200. La señal reconstruida generada por el decodificador 201 también se alimenta a un buffer 202 para su almacenamiento temporal. El análisis sinusoidal de la señal de reconstrucción en el buffer es realizado por un analizador sinusoidal 203, y la evolución de fase de la señal de reconstrucción en el buffer es realizada por una unidad de evolución de fase 204 después de lo cual la señal resultante se alimenta a un sintetizador sinusoidal 205 para la generación de una señal de reconstrucción sustituta que es salida de la entidad RX 200 en caso de pérdida de trama. Se proporcionarán más detalles de las operaciones de la entidad RX 200 a continuación.

La Figura 3 en (a), (b), (c), y (d) ilustra esquemáticamente cuatro fases de un proceso de creación e inserción de una trama de sustitución en caso de pérdida de trama. La Figura 3(a) ilustra esquemáticamente partes de una señal recibida previamente 301. Se ilustra esquemáticamente una ventana en 303. La ventana se utiliza para extraer una trama, una trama 304 denominada prototipo, de la señal recibida previamente 301; la parte media de la señal recibida previamente 301 no es visible ya que es idéntica a la trama prototipo 304 donde la ventana 303 es igual a 1. La Figura 3(b) ilustra esquemáticamente el espectro de magnitud, en términos de la transformada discreta de Fourier (DFT), de la trama prototipo en la Figura 3(a), donde se identifican dos picos de frecuencia f_k y f_{k+1} . La Figura 3(c) ilustra esquemáticamente el espectro de frecuencia de la trama de sustitución generada, donde se evolucionan adecuadamente las fases alrededor de los picos y se retiene el espectro de magnitud de la trama prototipo. La Figura 3(d) ilustra esquemáticamente la trama de sustitución generada 305 que ha sido insertada.

En vista de los mecanismos descritos anteriormente para la ocultación de pérdida de trama, se ha encontrado que los artefactos tonales son causados por una periodicidad demasiado marcada y picos espectrales demasiado bruscos del espectro de la trama de sustitución, a pesar de la aleatorización.

También es notable que los mecanismos descritos junto con un método de adaptación de un método de ocultación de pérdida de trama de tipo Phase ECU también son típicos para otros métodos de ocultación de trama que generan una señal de sustitución para las tramas pérdidas bien en el dominio de la frecuencia o en el del tiempo. Por lo tanto, puede ser deseable proporcionar mecanismos genéricos para la ocultación de pérdida de trama en caso de ráfagas largas de tramas pérdidas o corruptas.

Aparte de proporcionar una eficiente ocultación de pérdida de trama, también puede ser deseable encontrar mecanismos que puedan ser implementados con mínima complejidad computacional así como con mínimos requisitos de almacenamiento.

Al menos algunas de las realizaciones descritas en la presente memoria se basan en superponer gradualmente una señal de sustitución de un método primario de ocultación de pérdida de trama con una señal de ruido, donde la frecuencia característica de la señal de ruido es una representación espectral de baja resolución de la trama de una señal correctamente recibida previamente (una "buena trama").

Se hace referencia ahora al diagrama de flujo de la Figura 6 que describe un método para la ocultación de pérdida de trama según lo realizado por una entidad receptora según una realización.

La entidad receptora se configura para, en un paso S208, añadir, en colaboración con la construcción de un espectro de la trama de sustitución para una trama pérdida, un componente de ruido a la trama de sustitución. El componente de ruido tiene una frecuencia característica correspondiente a una representación espectral de baja resolución de una señal en una trama recibida previamente.

A este respecto, si la adición en el paso S208 se realiza en el dominio de la frecuencia, el componente de ruido puede considerarse como agregado a un espectro de una trama de sustitución ya generada, y por tanto, la trama de sustitución a la que se ha añadido el componente de ruido puede considerarse como una trama de sustitución secundaria, o adicional. Así, la trama de sustitución secundaria se compone de una trama de sustitución primaria y de un componente de ruido. Estos componentes están a su vez compuestos de componentes de frecuencia.

Según una realización, el paso S208 de añadir el componente de ruido a la trama de sustitución implica confirmar que una longitud de error de ráfaga n excede un primer umbral, T1. Un ejemplo del primer umbral es establecer $T1 \geq 2$.

Se hace referencia ahora al diagrama de flujo de la Figura 7 que describe métodos para la ocultación de pérdida de trama según lo realizado por una entidad receptora según realizaciones adicionales.

5 Según una primera realización preferida, la señal de sustitución para una trama pérdida es generada por un método primario de ocultación de pérdida de trama, superpuesto con una señal de ruido. Con el aumento del número de pérdidas de trama seguidas, la señal de sustitución de la ocultación primaria de pérdida de trama se atenúa gradualmente, preferiblemente según el comportamiento de silencio del método primario de ocultación de pérdida de trama en caso de pérdida de trama de ráfaga. Al mismo tiempo, la pérdida de energía de trama debido al comportamiento de silencio del método primario de ocultación de pérdida de trama se compensa mediante la adición de una señal de ruido con similares características espectrales como una trama de una señal recibida previamente, por ejemplo, la última trama recibida correctamente.

10 Por lo tanto, el componente de ruido y el espectro de la trama de sustitución pueden escalarse con factores de escala que son dependientes del número de tramas pérdidas consecutivamente de modo que el componente de ruido se superpone gradualmente en el espectro de la trama de sustitución con una magnitud creciente como una función del número de tramas pérdidas consecutivamente.

15 Como se describirá más adelante, el espectro de la trama de sustitución puede atenuarse gradualmente por un factor de atenuación $\alpha(m)$.

20 El espectro de la trama de sustitución y el componente de ruido pueden superponerse en el dominio de la frecuencia. Alternativamente, la representación espectral de baja resolución se basa en un conjunto de parámetros de codificación predictiva lineal (LPC) y el componente de ruido puede así superponerse en el dominio del tiempo. Para obtener más información sobre cómo aplicar parámetros de PLC, véase más adelante.

25 Más específicamente, el método primario de ocultación de pérdida de trama puede ser un método de tipo Phase ECU con una característica de adaptación en respuesta a la pérdida de ráfaga como se describió anteriormente. Es decir, el componente de la trama de sustitución puede derivarse por un método primario de ocultación de pérdida de trama, como Phase ECU.

30 En ese caso la señal generada por el método primario de ocultación de pérdida de trama es de tipo $Z(m) = \alpha(m) \cdot Y(m) \cdot e^{j(\theta_k + \vartheta(m))}$, donde $\alpha(m)$ y $\vartheta(m)$ son términos de atenuación de magnitud y de aleatorización de fase. Es decir, el espectro de la trama de sustitución puede tener una fase y la fase puede superponerse con un valor de fase aleatorio $\vartheta(m)$.

35 Y, como se describió anteriormente, la fase θ_k con $k = 1 \dots K$ es una función del índice m y de los picos espectrales K identificados por el método de Phase ECU, e $Y(m)$ es una representación del dominio de la frecuencia (espectro) de una trama de la señal de audio recibida previamente.

40 Como se sugiere en la presente memoria, este espectro puede luego modificado aún más por un componente de ruido aditivo $\beta(m) \cdot e^{j\eta(m)}$, produciendo un componente combinado $\beta(m) \cdot \bar{Y}(m) \cdot e^{j\eta(m)}$, donde $Y(m)$ es una representación del espectro de magnitud de una "buena trama" recibida previamente, es decir, una trama de una, al menos relativamente, señal recibida correctamente. De este modo, el componente de ruido puede proporcionarse con un valor de fase aleatorio $\eta(m)$.

De esta manera el coeficiente espectral para el índice de espectro m sigue una expresión:

45
$$Z(m) = \alpha(m) \cdot Y(m) \cdot e^{j(\theta_k + \vartheta(m))} + \beta(m) \cdot \bar{Y}(m) \cdot e^{j\eta(m)}.$$

50 Aquí $\beta(m)$ es un factor de escala de la magnitud y $\eta(m)$ es una fase aleatoria. Por lo tanto, el componente de ruido aditivo consiste en coeficientes espectrales de fase aleatoria escalados del espectro de magnitud $Y(m)$. Según la invención, $\beta(m)$ puede elegirse de modo que compense la pérdida de energía al aplicar el factor de atenuación $\alpha(m)$ al coeficiente espectral $Y(m)$ del espectro de la trama de sustitución de la ocultación primaria de pérdida de trama. Por lo tanto, la entidad receptora puede configurarse para, en un paso opcional S204, determinar un factor de escala de magnitud $\beta(m)$ para el componente de ruido de modo que $\beta(m)$ compense la pérdida de energía resultante de aplicar el factor de atenuación $\alpha(m)$ al espectro de la trama de sustitución.

55 Bajo el supuesto de que los términos de fase aleatoria de-correlacionan los dos términos aditivos $\alpha(m) \cdot Y(m) \cdot e^{j(\theta_k + \vartheta(m))}$ y $\beta(m) \cdot \bar{Y}(m) \cdot e^{j\eta(m)}$ de la ecuación anterior, $\beta(m)$ puede, por ejemplo, determinarse como

60
$$\beta(m) = \sqrt{1 - \alpha^2(m)}.$$

65 Para evitar el problema descrito anteriormente con artefactos tonales que surgen de picos espectrales demasiado bruscos, mientras que todavía mantiene la frecuencia característica general de la señal antes de la pérdida de trama de ráfaga, la representación del espectro de magnitud $Y(m)$ es una representación de baja resolución. Se ha encontrado que una representación de baja resolución muy adecuada del espectro de magnitud se obtiene promediando del grupo de frecuencias sabio el espectro de magnitud $|Y(m)|$ de una trama de la señal recibida

5 previamente, por ejemplo, una trama recibida correctamente, una "buena" trama. La entidad receptora puede configurarse para, en un paso opcional S202a, obtener la representación de baja resolución del espectro de magnitud promediando del grupo de frecuencias sabio el espectro de magnitud de la señal en la trama recibida previamente. La representación espectral de baja resolución puede basarse en un espectro de magnitud de la señal en la trama recibida previamente.

10 Dejar $I_k = [m_{k-1} + 1, \dots, m_k]$ especificar el intervalo k^{th} , $k = 1 \dots K$, que cubre los contenedores de la DFT desde $m_{k-1} + 1$ hasta m_k , entonces estos intervalos definen K bandas de frecuencia. El promedio del grupo de frecuencias sabio para la banda k puede entonces ser hecho promediando los cuadrados de las magnitudes de los coeficientes espectrales en esa banda y calculando la raíz cuadrada de la misma:

$$15 \quad \bar{Y}_k = \sqrt{\frac{1}{|I_k|} \sum_{m \in I_k} |Y(m)|^2}$$

20 Aquí $|I_k|$ denota el tamaño del grupo de frecuencia k , es decir, el número de contenedores de frecuencia incluidos. Debe señalarse que el intervalo $I_k = [m_{k-1} + 1, \dots, m_k]$ corresponde a la banda de frecuencia

$B_k = [\frac{m_{k-1} + 1}{N} \cdot f_s, \dots, \frac{m_k}{N} \cdot f_s]$, en donde f_s denota la frecuencia de muestreo de audio y N la longitud del bloque de la transformada del dominio de la frecuencia utilizada.

25 Una opción adecuada ilustrativa para los tamaños de la banda de frecuencia o anchos es bien hacerlos del mismo tamaño con, por ejemplo, un ancho de varios 100 Hz. Otra manera ilustrativa es hacer que los anchos de la banda de frecuencia sigan el tamaño de las bandas críticas auditivas humanas, es decir, relacionarlos con la resolución de frecuencia del sistema auditivo humano. Es decir, los anchos de grupo utilizados durante el promedio del grupo de frecuencias sabio pueden seguir las bandas críticas auditivas humanas. Esto significa hacer los anchos de banda de frecuencia aproximadamente iguales para frecuencias hasta 1 kHz e incrementarlos exponencialmente por encima de 1 kHz. Incremento exponencial significa, por ejemplo, doblar el ancho de banda de frecuencia al incrementar el índice k de la banda.

35 Una realización específica ilustrativa adicional del cálculo de los coeficientes del espectro de magnitud de baja resolución \bar{Y}_k es basarlo en una multitud n de transformadas del dominio de la frecuencia de baja resolución de la señal recibida previamente. La entidad receptora puede así configurarse para, en un paso opcional S202b, obtener la representación de baja resolución de dicho espectro de magnitud promediando del grupo de frecuencias sabio una multitud n de transformadas del dominio de la frecuencia de baja resolución de la señal en la trama recibida previamente. Una opción adecuada ilustrativa de n es $n=2$.

40 Según esta realización, en primer lugar, se calculan los espectros de magnitud al cuadrado de una parte izquierda (subtrama) y de una parte derecha (subtrama) de una trama de la señal recibida previamente, por ejemplo, de la buena trama recibida más recientemente. Una trama aquí podría ser del tamaño de los segmentos de audio o tramas utilizadas en la transmisión, o una trama podría ser de algún otro tamaño, por ejemplo, un tamaño construido y utilizado por un ECU de fase, que puede construir tramas propias con diferente longitud a partir de la señal reconstruida. La longitud del bloque N_{part} de estas transformadas de baja resolución puede ser una fracción (por ejemplo, 1/4) del tamaño de trama original del método primario de ocultación de pérdida de trama. Luego, en segundo lugar, los coeficientes del espectro de magnitud de baja resolución del grupo de frecuencias sabio se calculan promediando del grupo de frecuencias sabio las magnitudes espectrales al cuadrado de las subtramas izquierda y derecha, y finalmente calculando la raíz cuadrada de las mismas:

$$50 \quad \bar{Y}_k = \sqrt{\frac{1}{2 \cdot |I_k|} \left(\sum_{m \in I_k} |Y_{left}(m)|^2 + \sum_{m \in I_k} |Y_{right}(m)|^2 \right)}$$

55 Los coeficientes del espectro de magnitud de baja resolución \bar{Y}_k son entonces obtenidos a partir de los K representantes del grupo de frecuencias:

$$60 \quad \bar{Y}(m) = \bar{Y}_k \text{ for } m \in I_k, k = 1 \dots K.$$

Existen varias ventajas con este enfoque de calcular el coeficiente del espectro de magnitud de baja resolución \bar{Y}_k ; la utilización de dos transformadas cortas del dominio de la frecuencia es preferible en términos de complejidad computacional sobre una sola transformada del dominio de la frecuencia con una mayor longitud de bloque. Además,

el promedio estabiliza la estimación del espectro, es decir, reduce las fluctuaciones estadísticas que podrían impactar la calidad alcanzable. Una ventaja específica al aplicar esta realización junto con el controlador de Phase ECU mencionado previamente, es que puede confiar en los análisis espectrales relacionados con la detección de una condición transitoria en la trama de una señal recibida previamente, la "buena trama". Esto reduce la sobrecarga computacional asociada con la invención aún más.

El objetivo de proporcionar un mecanismo con mínimos requisitos de almacenamiento también se consigue, ya que esta realización permite representar el espectro de baja resolución con solo K valores, donde K puede prácticamente ser tan bajo como, por ejemplo, 7 u 8.

Se ha encontrado además que la calidad de la señal de audio reconstruida en caso de largas ráfagas de pérdida puede mejorarse aún más si la superposición del grupo de frecuencias sabio con una señal de ruido impone un cierto grado de característica de paso bajo. Por lo tanto, puede imponerse una característica de paso bajo en la representación espectral de baja resolución.

Dicha característica evita efectivamente el desagradable ruido de alta frecuencia en la señal de sustitución. Más específicamente, esto se consigue introduciendo una atenuación adicional a través de un factor $\lambda(m)$ de la señal de ruido para frecuencias mayores. Comparado con el cálculo descrito anteriormente del factor de escala de ruido $\beta(m)$, este factor se calcula ahora según

$$\beta(m) = \lambda(m) \cdot \sqrt{1 - \alpha^2(m)} .$$

En la presente memoria el factor $\lambda(m)$ podría ser 1 para m pequeño y ser menor que 1 para m grande. Es decir, $\beta(m)$ puede determinarse como $\beta(m) = \lambda(m) \cdot \sqrt{1 - \alpha^2(m)}$, en donde $\lambda(m)$ es un factor de atenuación dependiente de la frecuencia. Por ejemplo, $\lambda(m)$ puede ser igual a 1 para m por debajo de un umbral y $\lambda(m)$ puede ser menor que 1 para m por encima de este umbral.

Debe señalarse que, preferiblemente, los factores de escala $\lambda(m)$ y $\beta(m)$, son constantes del grupo de frecuencias sabio. Esto ayuda a reducir los requisitos de complejidad y de almacenamiento. En ese caso, también el factor λ se aplica al grupo de frecuencias sabiamente según la siguiente expresión:

$$\beta_k = \lambda_k \cdot \sqrt{1 - \alpha_k^2} .$$

Se ha encontrado beneficioso establecer λ_k de modo que es 0,1 para bandas de frecuencia por encima de 8.000Hz y 0,5 para una banda de frecuencia de 4.000Hz-8.000Hz. Para bandas de frecuencia más bajas λ_k es igual a 1. También son posibles otros valores.

Se ha encontrado beneficioso además, a pesar de las ventajas de calidad del método propuesto con superposición de la señal de sustitución de un método primario de ocultación de pérdida de trama con una señal de ruido, aplicar una característica de silenciamiento para ráfagas de pérdida de trama extremadamente largas de por ejemplo, $n > 10$ (correspondiente a 200 ms o más). Por lo tanto, la entidad receptora puede configurarse para, en un paso opcional S206, aplicar un factor de atenuación de largo plazo γ a $\beta(m)$ cuando la longitud de error de ráfaga n excede un segundo umbral T2, al menos tan largo como el primer umbral T1. Según un ejemplo, $T2 \geq 10$.

En más detalle, en caso de que una síntesis sostenida de la señal de ruido pueda ser molesta para un oyente. Para solventar este problema la señal de ruido aditiva puede así ser atenuada a partir de ráfagas de pérdida de mayor tamaño que, p.ej., $n=10$. Específicamente, un factor de atenuación adicional de largo plazo γ (por ejemplo, $\gamma = 0,5$) y un umbral $thresh$ se introducen con lo que se atenúa la señal de ruido si la longitud de la ráfaga de pérdida n excede $thresh$. Esto conduce a la siguiente modificación del factor de escala de ruido:

$$\beta_\gamma(m) = \gamma^{\max(0, n - thresh)} \cdot \beta(m)$$

La característica que se consigue con esa modificación es que la señal de ruido se atenúa con $\gamma^{n - thresh}$ si n excede el umbral. Como un ejemplo, si $n = 20$ (400 ms) y $\gamma = 0,5$ y $T2 = thresh = 10$, entonces la señal de ruido se reduce a aproximadamente 1/1000.

Debe señalarse de nuevo que, la operación también puede hacerse del grupo de frecuencias sabio, como en la realización anterior.

Para resumir, según la invención, $Z(m)$ representa el espectro de una trama de sustitución y este espectro es generado por el uso de un método primario de ocultamiento de pérdida de trama, como el Phase ECU, basado en el espectro $Y(m)$ de una trama prototipo, es decir, una trama de la señal recibida previamente.

5 Para ráfagas de pérdida largas, el ECU de fase original con el controlador descrito esencialmente atenúa este espectro y aleatoriza las fases. Para n muy largo esto significa que la señal generada se silencia completamente.

10 Como se describe en la presente memoria esta atenuación es compensada añadiendo una cantidad adecuada de ruido de forma espectral. Por tanto, el nivel de la señal se mantiene esencialmente estable, incluso para $n > 5$. Para ráfagas de pérdida extremadamente largas, por ejemplo, $n > 10$, una realización implica atenuar/silenciar incluso este ruido aditivo.

15 Según una realización adicional el espectro de la señal de ruido aditivo de baja resolución $\bar{Y}(m)$ puede representarse por un conjunto de parámetros de LPC, y por tanto el espectro en este caso corresponde al espectro de un filtro de síntesis de LPC con estos parámetros de LPC como coeficientes. Dicha realización puede ser preferida si el método de PLC primario no es de tipo Phase ECU y además, por ejemplo, un método que funciona en el dominio del tiempo. En ese caso una señal de tiempo correspondiente al espectro de la señal de ruido aditivo de baja resolución $\bar{Y}(m)$ podría también, preferiblemente, ser generada en el dominio del tiempo, filtrando ruido blanco a través del filtro de síntesis con dichos coeficientes de LPC.

20 La adición del componente de ruido a la trama de sustitución como en el paso S208 puede, por ejemplo, realizarse bien en el dominio de la frecuencia o en el dominio del tiempo o en más dominios de señal equivalentes. Por ejemplo, existen dominios de señal como el filtro de espejo en cuadratura (QMF) o el dominio del filtro de sub-banda en los que los métodos primarios de ocultamiento de pérdida de trama pueden operar. En tales casos, puede preferirse generar una señal de ruido aditivo correspondiente al espectro de la señal de ruido de baja resolución $\bar{Y}(m)$ descrito en estos dominios de señal correspondientes. Aparte de las diferencias del dominio de la señal en el que se añade la señal de ruido, las anteriores realizaciones siguen siendo aplicables.

30 Se hace referencia ahora al diagrama de flujo de la Figura 5 que describe un método para la ocultación de pérdida de trama según lo realizado por una entidad receptora según una realización particular.

35 En una acción S101 se determina un componente de ruido, donde la frecuencia característica del componente de ruido es una representación espectral de baja resolución de una trama de una señal recibida previamente. El componente de ruido puede, por ejemplo, estar compuesto y denotado como $\beta(m) \cdot \bar{Y}(m) \cdot e^{j\eta(m)}$, donde $\beta(m)$ puede ser un factor de escala de magnitud y $\eta(m)$ puede ser una fase aleatoria, e $\bar{Y}(m)$ puede ser una representación del espectro magnitud de una "buena trama" recibida previamente.

40 Según la invención, en la acción S103, se determina si un número, n , de tramas pérdidas o erróneas excede un umbral. El umbral podría ser, por ejemplo, 8, 9, 10 u 11 tramas. Cuando n es menor que el umbral, el componente de ruido se añade a un espectro de la trama de sustitución Z en una acción S104. El espectro de la trama de sustitución Z puede ser derivado por un método primario de ocultación de pérdida de trama, como por ejemplo, Phase ECU. Cuando el número de tramas pérdidas n excede el umbral, se aplica un factor de atenuación y al componente de ruido.

45 El factor de atenuación puede ser constante dentro de ciertos rangos de frecuencia. Al aplicar el factor de atenuación γ , el componente de ruido se añade a un espectro de la trama de sustitución Z en la acción S104.

50 Las realizaciones descritas en la presente memoria también se refieren a una entidad receptora, o nodo receptor, que será descrito a continuación con referencia a las Figuras 4, 8 y 9. La entidad receptora será descrita brevemente para evitar una repetición innecesaria.

Una entidad receptora puede configurar para realizar una o más de las realizaciones descritas en la presente memoria.

55 La Figura 4 describe esquemáticamente módulos funcionales de una entidad receptora 400 según una realización. La entidad receptora 400 comprende un detector de pérdida de trama 401 configurado para detectar una pérdida de trama en una señal recibida a lo largo de la ruta de la señal 410. El detector de pérdida de trama comunica un generador de representación de baja resolución 402 y un generador de trama de sustitución 403. El generador de representación de baja resolución 402 se configura para generar una representación espectral de baja resolución de una señal en una trama recibida previamente. El generador de trama de sustitución 403 se configura para generar una trama de sustitución según los mecanismos conocidos, como Phase ECU. Los bloques funcionales 404 y 405 representan el escalado de las señales generadas por el generador de representación de baja resolución 402 y el generador de trama de sustitución 403, respectivamente, con los factores de escala descritos anteriormente β , γ , y α . Los bloques funcionales 406 y 407 representan la superposición de las señales así escaladas con los valores de fase descritos anteriormente η y ϑ . El bloque funcional 408 representa un sumatorio para añadir el componente de ruido así generador a la trama de sustitución. El bloque funcional 409 representa un interruptor como controlado por el detector de pérdida de trama 401 para reemplazar una trama perdida con una trama de sustitución generada. Como se señaló

anteriormente, existen muchos dominios en los que las operaciones, como la adición en el paso S208, pueden ser realizadas. Por lo tanto, cualquiera de los bloques funcionales descritos anteriormente puede configurarse para realizar las operaciones en cualquiera de estos dominios.

5 Más abajo, una entidad receptora ilustrativa 800, adaptada para permitir la realización de un método descrito anteriormente para la gestión de errores de trama de ráfaga será descrita con referencia a la Figura 8.

10 La parte de la entidad receptora que está mayormente relacionada con la solución sugerida en la presente memoria se ilustra como una disposición 800 rodeada por una línea discontinua. La disposición y posiblemente otras partes de la entidad receptora son adaptadas para permitir la realización de uno o más de los procedimientos descritos anteriormente e ilustrados, por ejemplo, en las Figuras 5, 6, y 7. La entidad receptora 800 se ilustra para comunicarse con otras entidades a través de una unidad de comunicación 802, que puede considerarse que comprende medios convencionales para una comunicación inalámbrica y/o cableada de acuerdo con un estándar o protocolo de comunicación dentro del cual es operable la entidad receptora. La disposición y/o la entidad receptora puede comprender además otras unidades funcionales 807, para proporcionar, por ejemplo, las funciones regulares de la entidad receptora, como por ejemplo, el procesamiento de la señal en colaboración con la decodificación de audio, como voz o música.

15 La parte de la disposición de la entidad receptora puede implementarse y/o describirse como sigue:

20 La disposición comprende medios de procesamiento 803, como un procesador, y una memoria 804 para almacenar instrucciones. La memoria comprende instrucciones en la forma de un programa informático 805, que cuando se ejecuta por los medios de procesamiento hace que la entidad receptora o disposición realice los métodos según los descrito en la presente memoria.

25 Una realización alternativa de la entidad receptora 800 se muestra en la Figura 9. La Figura 9 ilustra una entidad receptora 900, operable para decodificar una señal de audio.

30 Una disposición 901 puede implementarse y/o describirse esquemáticamente como sigue. La disposición 901 puede comprender una unidad de determinación 903, configurada para determinar un componente de ruido con una frecuencia característica de una representación espectral de baja resolución de una trama de una señal recibida previamente y para determinar un factor de escala de magnitud. La disposición puede comprender además una unidad de adición 904, configurada para añadir el componente de ruido a un espectro de la trama de sustitución. La disposición puede comprender además una unidad de obtención 910, configurada para obtener la representación de baja resolución del espectro de magnitud de la señal en la trama recibida previamente. La disposición puede comprender además una unidad de aplicación 911, configurada para aplicar un factor de atenuación de largo plazo. La entidad de recepción puede comprender además unidades 907 configuradas para, por ejemplo, determinar un factor de escala $\beta(m)$ para el componente de ruido. La unidad de recepción 900 comprende además una unidad de comunicación 902 que tiene un transmisor (Tx) 908 y un receptor (Rx) 909 con una funcionalidad como la de la unidad de comunicación 802. La entidad de recepción 900 comprende además una memoria 906 con una funcionalidad como la de la memoria 804.

35 Las unidades o módulos en las disposiciones descritas anteriormente podrían ser implementadas, por ejemplo, por uno o más de: un procesador o un microprocesador y software adecuado y memoria para almacenar el mismo, un Dispositivo de Lógica Programable (PLD) u otro(s) componente(s) electrónicos o circuitos de procesamiento configurados para realizar las acciones descritas anteriormente, e ilustradas, por ejemplo, en

40 La Figura 8. Es decir, las unidades o módulos en las disposiciones descritas anteriormente podrían ser implementadas por una combinación de circuitos analógicos y digitales, y/o uno o más procesadores configurados con software y/o firmware, por ejemplo, almacenado en una memoria. Uno o más de estos procesadores, así como el otro hardware digital, pueden incluirse en un solo circuito integrado de aplicación específica (ASIC), o varios procesadores y varios hardware digitales pueden distribuirse entre varios componentes separados, ya individualmente empaquetados o ensamblados en un sistema en chip (SoC).

45 La Figura 10 muestra un ejemplo de un producto de programa informático 1000 que comprende un medio legible por ordenador 1001. En este medio legible por ordenador 1001, un programa informático 1002 puede almacenarse, que el programa informático 1002 puede producir un circuito de procesamiento 803 y las entidades y dispositivos acoplados operativamente a ello, como la unidad de comunicaciones 802 y el medio de almacenamiento 804, para ejecutar los métodos según las realizaciones descritas en la presente memoria. El programa informático 1002 y/o el producto de programa informático 1001 puede así proporcionar un medio para la realización de cualquier paso como se describe en la presente memoria.

50 En el ejemplo de la Figura 10, el producto de programa informático 1001 se ilustra como un disco óptico, como un CD (disco compacto) o un DVD (disco digital versátil) o un disco Blu-Ray. El producto de programa informático 1001 también podría ser incluido como una memoria, como una memoria de acceso aleatorio (RAM), una memoria de solo lectura (ROM), una memoria de solo lectura programable borrable (EPROM), o una memoria de solo lectura

5 programable borrable eléctricamente (EEPROM) y más particularmente como un medio de almacenamiento no volátil de un dispositivo en una memoria externa como una memoria USB (Bus Serie Universal) o una memoria Flash, como una memoria Flash compacta. Así, mientras que el programa informático 1002 se muestra aquí esquemáticamente como una pista en el disco óptico representado, el programa informático 1002 puede almacenarse de cualquier manera que sea adecuada para el producto de programa informático 1001.

Algunas definiciones de posibles características y realizaciones se describen a continuación, en parte refiriéndose al diagrama de flujo de la Figura 5.

10 Un método realizado por una entidad receptora para mejorar la ocultación de pérdida de trama o la gestión de errores de trama de ráfaga, el método que comprende: en colaboración con la construcción de un espectro de la trama de sustitución Z que añade (acción 104) un componente de ruido al espectro de la trama de sustitución Z, donde la frecuencia característica del componente de ruido es una representación espectral de baja resolución de una trama de una señal recibida previamente.

15 En una posible realización, la representación espectral de baja resolución se basa en un espectro de magnitud de una trama de una señal recibida previamente. Una representación de baja resolución de un espectro de magnitud puede obtenerse, por ejemplo, promediando del grupo de frecuencias sabio del espectro de magnitud de una trama de una señal recibida previamente. De manera alternativa, una representación de baja resolución de un espectro de magnitud puede basarse en una multitud n de transformadas del dominio de la frecuencia de baja resolución de la señal recibida previamente. En una posible realización, la representación espectral de baja resolución se basa en un conjunto de parámetros de codificación lineal predictiva (LPC).

20 En una posible realización donde el espectro de la trama de sustitución Z es atenuado gradualmente por un factor de atenuación $\alpha(m)$, el método comprende la determinación de un factor de escala de magnitud $\beta(m)$ para el componente de ruido, de manera que $\beta(m)$ compense la pérdida de energía resultante de la aplicación del factor de atención $\alpha(m)$. $\beta(m)$ puede, por ejemplo, determinarse como

30
$$\beta(m) = \sqrt{1 - \alpha^2(m)}.$$

35 En una posible realización, $\beta(m)$ es derivada como $\beta(m) = \lambda(m) \cdot \sqrt{1 - \alpha^2(m)}$, en donde el factor $\lambda(m)$ es un factor de atenuación para ciertas frecuencias de la señal de ruido, por ejemplo, frecuencias más altas. $\lambda(m)$ puede ser igual a 1 para m pequeño y ser menor que 1 para m grande.

En una posible realización, los factores de escala $\alpha(m)$ y $\beta(m)$ son constantes del grupo de frecuencias sabio.

40 En una posible realización el método comprende la aplicación (acción 103) de un factor de atenuación, γ , cuando una longitud de error de ráfaga excede un umbral.

El espectro de la trama de sustitución Z puede ser derivado por un método primario de ocultación de pérdida de trama, como Phase ECU.

45 Las diferentes realizaciones pueden combinarse de cualquier manera adecuada.

50 Más abajo, será proporcionada información en realizaciones ilustrativas del método de ocultación de pérdida de trama Phase ECU, aunque el término "Phase ECU" no será mencionado explícitamente. Phase ECU ha sido mencionado en la presente memoria, por ejemplo, en términos del método primario de ocultación de pérdida de trama, para la derivación de Z antes de añadir el componente de ruido.

Un concepto de las realizaciones descritas más adelante comprende una ocultación de una trama de audio perdida:

55 - realizando un análisis sinusoidal de al menos parte de una señal de audio previamente recibida o reconstruida, en donde el análisis sinusoidal implica la identificación de frecuencias de los componentes sinusoidales de la señal de audio;

60 - aplicando un modelo sinusoidal en un segmento de la señal de audio previamente recibida o reconstruida, en donde dicho segmento se utiliza como una trama prototipo para crear una trama de sustitución para una trama perdida, y

- creando una trama de sustitución que implica la evolución en el tiempo de los componentes sinusoidales de la trama prototipo, hasta la instancia de tiempo de la trama de audio perdida, en respuesta a las correspondientes frecuencias identificadas.

Análisis sinusoidal

La ocultación de pérdida de trama según las realizaciones implica un análisis sinusoidal de una parte de la señal de audio previamente recibida o reconstruida. El propósito de este análisis sinusoidal es encontrar las frecuencias de los componentes sinusoidales principales, es decir, las sinusoides, de esa señal. Por la presente, la suposición subyacente es que la señal de audio fue generada por un modelo sinusoidal y que está compuesta de un número limitado de sinusoides individuales, es decir, que es una señal de seno múltiple del siguiente tipo:

$$s(n) = \sum_{k=1}^K a_k \cdot \cos(2\pi \frac{f_k}{f_s} \cdot n + \varphi_k)$$

En esta ecuación K es el número de sinusoides en que se supone consiste la señal. Para cada una de las sinusoides con índice k=1 ... K, a_k es la amplitud, f_k es la frecuencia, y φ_k es la fase. La frecuencia de muestreo se denomina f_s y el índice de tiempo de las muestras de la señal discreta en el tiempo s(n) n.

Puede ser beneficioso, o incluso importante, encontrar las frecuencias más exactas posibles de las sinusoides. Mientras que una señal sinusoidal ideal tendría un espectro lineal con frecuencias lineales f_k, encontrar sus verdaderos valores requeriría, en principio, un tiempo de medida infinito. Por tanto, en la práctica es difícil encontrar estas frecuencias, ya que solo pueden ser estimadas en base a un corto período de medida, que corresponde al segmento de señal utilizado para el análisis sinusoidal según las realizaciones descritas en la presente memoria; este segmento de análisis es referido en adelante como una trama de análisis. Otra dificultad es que la señal puede, en la práctica, ser variante en el tiempo, lo que significa que los parámetros de la anterior educación varían en el tiempo. Por tanto, por un lado es deseable utilizar una trama de análisis larga haciendo las medidas más precisas, por otro lado un corto período de medida sería necesario para hacer frente mejor a las posibles variaciones de la señal. Una buena compensación es utilizar una longitud de trama de análisis en el orden de, por ejemplo, 20-40 ms.

Según una realización preferida, las frecuencias de las sinusoides f_k son identificadas por un análisis del dominio de la frecuencia de la trama de análisis. Para este fin, la trama de análisis se transforma en el dominio de la frecuencia, por ejemplo, mediante la DFT (Transformada Discreta de Fourier) o la DCT (Transformada Discreta de Coseno), o una transformada del dominio de la frecuencia similar. En caso de que se utilice una DFT de la trama de análisis, el espectro X(m) en el índice de frecuencia discreta m es dado por:

$$X(m) = DFT(w(n) \cdot x(n)) = \sum_{n=0}^{L-1} e^{-j\frac{2\pi}{L}mn} \cdot w(n) \cdot x(n)$$

En esta ecuación, w(n) denota la función de ventana con la que se extrae y se pesa la trama de análisis de longitud L; j es la unidad imaginaria y e es la función exponencial.

Una función de ventana típica es una ventana rectangular que es igual a 1 para n ∈ [0...L-1] y en caso contrario 0. Se supone que los índices de tiempo de la señal de audio recibida previamente son establecidos de modo que se referencia la trama prototipo por los índices de tiempo n=0...L-1. Otras funciones de ventana que pueden ser más adecuadas para el análisis espectral son, por ejemplo, Hamming, Hanning, Kaiser or Blackman.

Otra función de ventana es una combinación de la ventana de Hamming y de la ventana rectangular. Dicha ventana puede tener una forma de flanco ascendente como la mitad izquierda de una ventana de Hamming de longitud L1 y una forma de flanco descendente como la mitad derecha de una ventana de Hamming de longitud L1 y entre los flancos ascendente y descendente la ventana es igual a 1 para la longitud de L-L1.

Los picos del espectro de magnitud de la trama de análisis de ventana |X(m)| constituyen una aproximación de las frecuencias sinusoidales requeridas f_k. Sin embargo, la precisión de esta aproximación está limitada por el espaciado

de frecuencia de la DFT. Con la DFT con una longitud de bloque L la precisión está limitada a $\frac{f_s}{2L}$.

Sin embargo, este nivel de precisión puede ser demasiado bajo en el alcance del método según las realizaciones descritas en la presente memoria, y puede obtenerse una precisión mejorada en base a los resultados de la siguiente consideración:

El espectro de la trama de análisis de ventana es dado por la convolución del espectro de la función de ventana con el espectro lineal de una señal de modelo sinusoidal $S(\Omega)$, muestreado posteriormente en los puntos de cuadrícula de la DFT:

$$X(m) = \int_{2\pi} \delta(\Omega - m \cdot \frac{2\pi}{L}) \cdot (W(\Omega) * S(\Omega)) \cdot d\Omega$$

En esta ecuación, δ representa la función delta de Dirac y el símbolo * denota la operación de convolución. Utilizando la expresión del espectro de la señal de modelo sinusoidal, esto puede escribirse como

$$X(m) = \frac{1}{2} \int_{2\pi} \delta(\Omega - m \cdot \frac{2\pi}{L}) \cdot \sum_{k=1}^K a_k \cdot \left((W(\Omega + 2\pi \frac{f_k}{f_s}) \cdot e^{-j\varphi_k} + W(\Omega - 2\pi \frac{f_k}{f_s}) \cdot e^{j\varphi_k}) \right) \cdot d\Omega$$

Por lo tanto, el espectro muestreado es dado por

$$X(m) = \frac{1}{2} \sum_{k=1}^K a_k \cdot \left((W(2\pi(\frac{m}{L} + \frac{f_k}{f_s})) \cdot e^{-j\varphi_k} + W(2\pi(\frac{m}{L} - \frac{f_k}{f_s})) \cdot e^{j\varphi_k}) \right)$$

con $m=0 \dots L-1$. En base a esto, los picos observados en el espectro de magnitud de la trama de análisis se derivan de una señal sinusoidal de ventana con K sinusoides, donde las verdaderas frecuencias de sinusoides se encuentran en las proximidades de los picos. Así, la identificación de las frecuencias de los componentes sinusoides puede implicar además la identificación de las frecuencias en las proximidades de los picos del espectro relacionado con la transformada del dominio de la frecuencia utilizada.

Si se supone que m_k es un índice de la DFT (punto de cuadrícula) del pico k^{th} observado, entonces la frecuencia

$$\hat{f}_k = \frac{m_k}{L} \cdot f_s$$

correspondiente es que puede considerarse una aproximación de la verdadera frecuencia sinusoidal f_k . Se puede suponer que la verdadera frecuencia sinusoidal f_k se encuentra dentro del intervalo:

$$\left[\left(m_k - \frac{1}{2} \right) \cdot \frac{f_s}{L}, \left(m_k + \frac{1}{2} \right) \cdot \frac{f_s}{L} \right]$$

Para mayor claridad se observa que la convolución del espectro de la función de ventana con el espectro del espectro lineal de la señal de modelo sinusoidal puede entenderse como una superposición de versiones desplazadas en frecuencia del espectro de la función de ventana, de modo que las frecuencias de desplazamiento son las frecuencias de las sinusoides. Esta superposición es luego muestreada en los puntos de cuadrícula de la DFT.

En base al debate anterior, puede encontrarse una mejor aproximación de las verdaderas frecuencias sinusoidales incrementando la resolución de la búsqueda, de modo que sea mayor que la resolución de frecuencia de la transformada del dominio de la frecuencia utilizada.

Así, la identificación de frecuencias de los componentes sinusoidales se realiza, preferiblemente, con mayor resolución que la resolución de frecuencia de la transformada del dominio de la frecuencia utilizada, y la identificación puede implicar además interpolación.

Un modo preferido ilustrativo para encontrar una mejor aproximación de las frecuencias f_k de las sinusoides es aplicar interpolación parabólica. Un enfoque es ajustar las parábolas a través de los puntos de cuadrícula del espectro de magnitud de la DFT que rodea los picos y calcular las frecuencias respectivas que pertenecen a la parábola máxima, y una opción adecuada ilustrativa para el orden de las parábolas es 2. Con más detalle, se puede aplicar el siguiente procedimiento:

1) Identificar los picos de la DFT de la trama de análisis de ventana. La búsqueda de picos entregará el número de picos K y los correspondientes índices de la DFT de los picos. La búsqueda de picos puede normalmente hacerse en el espectro de magnitud de la DFT o en el espectro de magnitud logarítmica de la DFT.

2) Para cada pico k (con $k = 1 \dots K$) con el correspondiente índice de la DFT m_k , ajustar una parábola a través de los tres puntos $\{P_1; P_2; P_3\} = \{(m_{k-1}, \log(|X(m_{k-1})|)); (m_k, \log(|X(m_k)|)); (m_{k+1}, \log(|X(m_{k+1})|))\}$, donde \log denota al operador logaritmo. Esto resulta en los coeficientes de parábola $b_k(0), b_k(1), b_k(2)$ de la parábola definida por

$$p_k(q) = \sum_{i=0}^2 b_k(i) \cdot q^i$$

3) Para cada una de las K parábolas, calcular el índice de la frecuencia interpolada \hat{m}_k correspondiente al valor de q para el que la parábola tiene su máximo, en donde se utiliza $\hat{f}_k = \hat{m}_k \cdot f_s/L$ como una aproximación para la frecuencia de sinusoides f_k .

Aplicar un modelo Sinusoidal

La aplicación de un modelo sinusoidal para realizar una operación de ocultación de pérdida de trama según las realizaciones puede describirse como sigue:

En caso de que un segmento dado de la señal codificada no pueda reconstruirse por el decodificador ya que la información de codificación correspondiente no está disponible, es decir, ya que se ha perdido una trama, una parte disponible de la señal anterior a este segmento puede utilizarse como trama prototipo. Si $y(n)$ con $n=0 \dots N-1$ es el segmento no disponible para el que tiene que generarse una trama de sustitución $z(n)$, e $y(n)$ con $n < 0$ es la señal previamente codificada disponible, una trama prototipo de la señal disponible de longitud L e índice de comienzo n_{-1} se extrae con una función de ventana $w(n)$ y se transforma en el dominio de la frecuencia, por ejemplo, mediante la DFT:

$$Y_{-1}(m) = \sum_{n=0}^{L-1} y(n - n_{-1}) \cdot w(n) \cdot e^{-j \frac{2\pi}{L} nm}$$

La función de ventana puede ser una de las funciones de ventana descritas anteriormente en el análisis sinusoidal. Preferiblemente, para ahorrar en complejidad numérica, la trama transformada al dominio de la frecuencia debe ser idéntica a la utilizada durante el análisis sinusoidal.

En un siguiente paso se aplica la suposición del modelo sinusoidal. Según la suposición del modelo sinusoidal, la DFT de la trama prototipo puede escribirse como sigue:

$$Y_{-1}(m) = \frac{1}{2} \sum_{k=1}^K a_k \cdot \left(W\left(2\pi\left(\frac{m}{L} + \frac{f_k}{f_s}\right)\right) \cdot e^{-j\varphi_k} + W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j\varphi_k} \right)$$

Esta expresión también fue utilizada en la parte de análisis y se describió con detalle anteriormente.

A continuación, se consigue que el espectro de la función de ventana utilizada tenga solo una contribución significativa en un rango de frecuencia cercano a cero. El espectro de magnitud de la función de ventana es grande para frecuencias cercanas a cero y pequeño en caso contrario (dentro del rango de frecuencias normalizado de $-\pi$ a π , correspondiente a la mitad de la frecuencia de muestreo). Por lo tanto, se supone como una aproximación que el espectro de ventana $W(m)$ no es cero para un intervalo $M = [-m_{\min}, m_{\max}]$, con m_{\min} y m_{\max} siendo pequeños números positivos. En particular, se utiliza una aproximación del espectro de la función de ventana de manera que para cada k las contribuciones del espectro de la ventana desplazada en la expresión anterior no están estrictamente superpuestas. Por lo tanto en la ecuación anterior, para cada índice de frecuencia siempre hay solo, como máximo, la contribución de un sumando, es decir, de un espectro de ventana desplazada. Esto significa que la expresión anterior

$$\tilde{Y}_{-1}(m) = \frac{a_k}{2} \cdot W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j\varphi_k}$$

se reduce a la siguiente expresión aproximada:

para $m \in M_k$ no negativo y para cada k .

5 En la presente memoria, M_k denota el intervalo entero: $M_k = [\text{round}(\frac{f_k}{f_s} \cdot L) - m_{\min,k},$
 $\text{round}(\frac{f_k}{f_s} \cdot L) + m_{\max,k}]$,

redondeo en donde $m_{\min,k}$ y $m_{\max,k}$ cumplen la restricción explicada anteriormente de manera que los intervalos no se superponen. Una opción adecuada para $m_{\min,k}$ y $m_{\max,k}$ es establecerlos a un pequeño valor entero δ , por ejemplo, $\delta = 3$. En cambio, si los índices de la DFT relacionados con las dos frecuencias sinusoidales vecinas f_k y f_{k+1} son menores que 2δ , entonces δ se establece a

10 $\text{floor}(\frac{\text{round}(\frac{f_{k+1}}{f_s} \cdot L) - \text{round}(\frac{f_k}{f_s} \cdot L)}{2})$ de manera que se garantice que los intervalos no se superpongan. La función suelo (\cdot) es el entero más cercano al argumento de la función que es más pequeño o igual a él.

15 El siguiente paso según las realizaciones es aplicar el modelo sinusoidal según la expresión anterior y evolucionar sus K sinusoides en el tiempo. La suposición de que los índices de tiempo del segmento borrado comparado con los índices de tiempo de la trama prototipo defieren por $n-1$ muestras significa que las fases de las sinusoides avanzan por

20
$$\theta_k = 2\pi \cdot \frac{f_k}{f_s} n_{-1} .$$

Por lo tanto, el espectro de la DFT del modelo sinusoidal evolucionado es dado por:

$$Y_0(m) = \frac{1}{2} \sum_{k=1}^K a_k \cdot \left(W\left(2\pi\left(\frac{m}{L} + \frac{f_k}{f_s}\right)\right) \cdot e^{-j(\varphi_k + \theta_k)} + W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j(\varphi_k + \theta_k)} \right) .$$

25 Aplicando de nuevo la aproximación, según la cual el espectro de la función de ventana desplazada no se superpone, da:

$$\bar{Y}_0(m) = \frac{a_k}{2} \cdot W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j(\varphi_k + \theta_k)}$$

30 para $m \in M_k$ no negativo y para cada k .

Comparando la DFT de la trama prototipo $Y_{-1}(m)$ con la DFT del modelo sinusoidal evolucionado $Y_0(m)$ utilizando la aproximación, se encuentra que el espectro de magnitud permanece sin cambios mientras que la fase es desplazada

$$\theta_k = 2\pi \cdot \frac{f_k}{f_s} n_{-1}$$

por , para cada $m \in M_k$.

35 Por lo tanto, la trama de sustitución puede ser calculada por la siguiente expresión: $z(n) = IDFT\{Z(m)\}$ con $Z(m) = Y(m) \cdot e^{j\theta_k}$ para $m \in M_k$ no negativo y para cada k .

40 Una realización específica aborda la aleatorización de fase para los índices de la DFT que no pertenecen a ningún intervalo M_k . Como se describió anteriormente, los intervalos $M_k, k=1 \dots K$ tienen que ser establecidos de manera que no estén estrictamente superpuestos lo que es hecho utilizando algún parámetro δ que controla el tamaño de los intervalos. Puede ocurrir que δ sea pequeño en relación con la distancia de frecuencia de dos sinusoides vecinas. Por lo tanto, en el caso de que ocurra eso hay un hueco entre dos intervalos. En consecuencia, para los índices de la DFT m correspondientes no se define desplazamiento de fase según la expresión anterior $Z(m) = Y(m) \cdot e^{j\theta_k}$. Una opción adecuada según esta realización es aleatorizar la fase para estos índices, produciendo $Z(m) = Y(m) \cdot e^{j2\pi \text{rand}(\cdot)}$, donde la función aleatoria(\cdot), devuelve un número aleatorio.

45 En un paso, se realiza un análisis sinusoidal de una parte de una señal de audio previamente recibida o reconstruida, en donde el análisis sinusoidal implica la identificación de los componentes sinusoidales, es decir, sinusoides, de la señal de audio. A continuación, en un paso, se aplica un modelo sinusoidal en un segmento de la señal de audio

5 previamente recibida o reconstruida, en donde dicho segmento se utiliza como una trama prototipo para crear una trama de sustitución para una trama de audio perdida, y en un paso se crea la trama de sustitución para la trama de audio perdida, que implica la evolución en el tiempo de los componentes sinusoidales, es decir, sinusoides, de la trama prototipo, hasta la instancia de tiempo de la trama de audio perdida, en respuesta a las correspondientes frecuencias identificadas.

10 Según otra realización, se supone que la señal de audio está compuesta de un número limitado de componentes sinusoidales individuales, y que el análisis sinusoidal se realiza en el dominio de la frecuencia. Además, la identificación de las frecuencias de los componentes sinusoidales puede implicar la identificación de frecuencias en las proximidades de los picos de un espectro relacionado con la transformada del dominio de la frecuencia utilizada.

15 Según una realización ilustrativa, la identificación de frecuencias de los componentes sinusoidales se realiza con mayor resolución que la resolución de la transformada del dominio de la frecuencia utilizada, y la identificación puede implicar además interpolación, por ejemplo, de tipo parabólico.

20 Según una realización ilustrativa, el método comprende la extracción de una trama prototipo a partir de una señal disponible previamente recibida o reconstruida utilizando una función de ventana, y en donde la trama prototipo extraída se puede transformar a un dominio de la frecuencia.

25 Una realización adicional implica una aproximación de un espectro de la función de ventana, de manera que el espectro de la trama de sustitución se compone de partes estrictamente no superpuestas del espectro aproximado de la función de ventana.

30 Según otra realización ilustrativa, el método comprende la evolución en el tiempo de componentes sinusoidales de un espectro de frecuencia de una trama prototipo avanzando la fase de los componentes sinusoidales, en respuesta a la frecuencia de cada componente sinusoidal y en respuesta a la diferencia de tiempo entre la trama de audio perdida y la trama prototipo, y el cambio de un coeficiente espectral de la trama prototipo incluido en un intervalo M_k en las proximidades de una senoide k por un desplazamiento de fase proporcional a la frecuencia sinusoidal f_k y a la diferencia de tiempo entre la trama de audio perdida y la trama prototipo.

35 Una realización adicional comprende el cambio de la fase de un coeficiente espectral de la trama prototipo que no pertenece a una senoide identificada por una fase aleatoria, o el cambio de la fase de un coeficiente espectral de la trama prototipo no incluido en ninguno de los intervalos relacionados con las proximidades de la senoide identificada por un valor aleatorio.

40 Una realización adicional implica una transformada inversa del dominio de la frecuencia del espectro de frecuencia de la trama prototipo.

Más específicamente, el método de ocultación de pérdida de trama de audio según una realización adicional puede implicar los siguientes pasos:

- 45 1) Analizar un segmento de la señal disponible, previamente sintetizada, para obtener las frecuencias sinusoidales constituyentes f_k de un modelo sinusoidal.
- 2) Extraer una trama prototipo $\gamma-1$ a partir de la señal disponible previamente sintetizada y calcular la DFT de esa trama.
- 3) Calcular el desplazamiento de fase θ_k para cada senoide k en respuesta a la frecuencia sinusoidal f_k y al avance de tiempo $n-1$ entre la trama prototipo y la trama de sustitución.
- 4) Para cada senoide k avanzar de manera selectiva la fase de la DFT de la trama prototipo con θ_k para los índices de la DFT relacionados con una proximidad alrededor de la frecuencia de senoide f_k .
- 5) Calcular la DFT inversa del espectro obtenido en 4).

55 Las realizaciones descritas anteriormente pueden además ser explicadas por las siguientes suposiciones:

- 60 a) La suposición de que la señal puede ser representada por un número limitado de sinusoides.
- b) La suposición de que la trama de sustitución está suficientemente bien representada por estas sinusoides evolucionadas en el tiempo, en comparación con algún instante de tiempo anterior.
- 65 c) La suposición de una aproximación del espectro de una función de ventana de manera que el espectro de la trama de sustitución puede construirse mediante partes no superpuestas del espectro de la función de ventana desplazada en frecuencia, siendo las frecuencias desplazadas las frecuencias de senoide.

A continuación, se presentará información sobre una elaboración más detallada del Phase ECU:

Un concepto de las realizaciones descritas más adelante comprende la ocultación de una trama de audio perdida:

- 5 - realizando un análisis sinusoidal de al menos parte de una señal de audio previamente recibida o reconstruida, en donde el análisis sinusoidal implica la identificación de frecuencias de los componentes sinusoidales de la señal de audio;
- 10 - aplicando un modelo sinusoidal sobre un segmento de la señal de audio previamente recibida o reconstruida, en donde dicho segmento se utiliza como una trama prototipo para crear una trama de sustitución para una trama perdida;
- 15 - creando la trama de sustitución para la trama de audio perdida, que implica una evolución en el tiempo de los componentes sinusoidales de la trama prototipo, hasta la instancia de tiempo de la trama de audio perdida, en base a las correspondientes frecuencias identificadas;
- 20 - realizando al menos uno de una estimación de frecuencia mejorada en la identificación de frecuencias, y una adaptación de la creación de la trama de sustitución en respuesta a la tonalidad de la señal de audio, en donde la estimación de frecuencia mejorada comprende al menos uno de una aproximación de lóbulo principal, una mejora armónica, y una mejora entre tramas.

Las realizaciones descritas aquí comprenden una estimación de frecuencia mejorada. Esto puede implementarse, por ejemplo, utilizando una aproximación de lóbulo principal, una mejora armónica, o una mejora entre tramas, y esas tres realizaciones alternativas se describen a continuación:

25 Aproximación de lóbulo principal

Surge una limitación con la interpolación parabólica descrita anteriormente de que las parábolas utilizadas no se aproximan a la forma del lóbulo principal del espectro de magnitud $|W(\Omega)|$ de la función de ventana. Como solución,

30 esta realización se ajusta a un función $P(q)$, que aproxima el lóbulo principal de $|W(\frac{2\pi}{L} \cdot q)|$ a través de los puntos de cuadrícula del espectro de magnitud de la DFT que rodean los picos y calcula las frecuencias respectivas pertenecientes a la función máxima. La función $P(q)$ podría ser idéntica al espectro de magnitud desplazado en

$$|W(\frac{2\pi}{L} \cdot (q - \hat{q}))|$$

35 frecuencia de la función de ventana. Sin embargo, por simplicidad numérica debería ser, por ejemplo, un polinomio que permita un cálculo directo de la función máxima. Se aplica el siguiente procedimiento detallado:

40 1. Identificar los picos de la DFT de la trama de análisis de ventana. La búsqueda de picos entregará el número de picos K y los índices correspondientes de la DFT de los picos. La búsqueda de picos puede normalmente hacerse en el espectro de magnitud de la DFT o en el espectro de magnitud logarítmica de la DFT.

2. Derivar la función $P(q)$ que aproxima el espectro de magnitud $|W(\frac{2\pi}{L} \cdot q)|$ de la función de

$$\log |W(\frac{2\pi}{L} \cdot q)|$$

ventana o del espectro de magnitud logarítmica para un intervalo dado (q_1, q_2) .

45 3. Para cada pico k (con $k=1 \dots K$) con el índice m_k de la DFT correspondiente ajustado a la función desplazada en frecuencia $P(q-\hat{q}_k)$ a través de dos puntos de cuadrícula de la DFT que rodean el esperado pico verdadero del espectro continuo de la señal sinusoidal de ventana. Por lo tanto, para el caso de funcionamiento con el espectro de magnitud logarítmica, si $|X(m_{k-1})|$ es mayor que $|X(m_{k+1})|$ se ajusta $P(q-\hat{q}_k)$ a través de los puntos $\{P1; P2\} = \{(m_{k-1}, \log(|X(m_{k-1})|)); (m_k, \log(|X(m_k)|))\}$ y de otro modo a través de los puntos $\{P1; P2\} = \{(m_k, \log(|X(m_k)|)); (m_{k+1}, \log(|X(m_{k+1})|))\}$. Para el ejemplo alternativo de funcionamiento con un

50 espectro de magnitud lineal en vez de uno logarítmico, si $|X(m_{k-1})|$ es mayor que $|X(m_{k+1})|$ se ajusta $\hat{P}(q-q_k)$ a través de los puntos

{P1; P2} = {(m_{k-1}, |X(m_{k-1})|); (m_k, |X(m_k)|)} y de otro modo a través de los puntos {P1; P2} = {(m_k, |X(m_k)|); (m_{k+1}, |X(m_{k+1})|)}. Por simplicidad, P(q) puede seleccionarse que sea un polinomio bien de orden 2 o 4. Esto hace que la aproximación en el paso 2 sea un simple cálculo de regresión lineal y el cálculo de \hat{q}_k directo. El intervalo (q₁, q₂) puede elegirse que sea fijo e idéntico para todos los picos, por ejemplo, (q₁, q₂) = (-1, 1), o adaptativo. En el enfoque adaptativo el intervalo puede elegirse de manera que la función P(q- \hat{q}_k) se ajusta al lóbulo principal del espectro de función de ventana en el rango de los puntos de cuadrícula de la DFT relevantes {P1; P2}.

4. Para cada uno de los K parámetros de desplazamiento de frecuencia \hat{q}_k para los que el espectro continuo de la señal sinusoidal de ventana se espera que tenga su pico calculado $\hat{f}_k = \hat{q}_k \cdot f_s/L$ como una aproximación para la frecuencia de senoide f_k .

Mejora armónica de la estimación de frecuencia

La señal transmitida puede ser armónica, lo que significa que la señal consiste en ondas sinusoidales cuyas frecuencias son múltiples enteros de alguna frecuencia fundamental f_0 . Este es el caso cuando la señal es muy periódica como, por ejemplo, para la voz sonora o para los tonos sostenidos de algún instrumento musical. Esto significa que las frecuencias del modelo sinusoidal de las realizaciones no son independientes sino que tienen una relación armónica y derivan de la misma frecuencia fundamental. Teniendo en cuenta esta propiedad armónica puede, consecuentemente, mejorar sustancialmente el análisis de las frecuencias del componente sinusoidal, y esta realización implica el siguiente procedimiento:

1. Comprobar si la señal es armónica. Esto puede, por ejemplo, ser hecho evaluando la periodicidad de la señal antes de la pérdida de trama. Un método directo es realizar un análisis de auto-correlación de la señal. El máximo de dicha función de auto-correlación para algún desfase temporal $\tau > 0$ puede utilizarse como un indicador. Si el valor de este máximo excede un umbral dado, la señal puede considerarse armónica. El desfase temporal correspondiente τ corresponde entonces al período de la señal que está relacionado con la frecuencia

$$f_0 = \frac{f_s}{\tau} .$$

fundamental a través de

Muchos métodos de codificación predictiva lineal del habla aplican la llamada predicción de tono de bucle abierto o cerrado o codificación CELP (predicción lineal excitada por código) utilizando libros de código adaptativos. La ganancia de tono y los parámetros de retraso de tono asociados derivados por dichos métodos de codificación son también indicadores útiles de si la señal es armónica y, respectivamente, para el desfase temporal.

Se describe a continuación un método adicional:

2. Para cada índice armónico j dentro del rango entero 1...J_{max} comprobar si existe un pico en el espectro de magnitud (logarítmica) de la DFT de la trama de análisis dentro de la proximidad de la frecuencia armónica $f_j = j \cdot f_0$. La proximidad de f_j puede definirse como el rango delta alrededor de f_k donde delta corresponde a la

resolución de frecuencia de la DFT $\frac{f_s}{L}$, es decir, el intervalo

$$\left[j \cdot f_0 - \frac{f_s}{2 \cdot L}, j \cdot f_0 + \frac{f_s}{2 \cdot L} \right] .$$

En caso de que dicho pico esté presente con la frecuencia sinusoidal estimada correspondiente \hat{f}_k , reemplazar \hat{f}_k por

$$\hat{f}_k = j \cdot f_0 .$$

Para el procedimiento dado anteriormente también existe la posibilidad de hacer la comprobación de si la señal es armónica y la derivación de la frecuencia fundamental implícitamente y posiblemente de una manera iterativa sin necesariamente utilizar indicadores de algún método separado. Se da un ejemplo para dicha técnica como sigue:

Para cada $f_{0,p}$ salida de un conjunto de valores candidatos {f_{0,1} ... f_{0,P}} aplicar el procedimiento 2 descrito anteriormente, aunque sin reemplazar \hat{f}_k pero contando cuantos picos de la DFT están presentes dentro de la proximidad alrededor de las frecuencias armónicas, es decir, los múltiplos enteros de $f_{0,p}$. Identificar la frecuencia fundamental $f_{0,pmax}$ para la que se obtiene el mayor número de picos en o alrededor de las frecuencias armónicas. Si este mayor número de picos excede un umbral dado, entonces se supone que la señal es armónica. En ese caso se puede suponer que $f_{0,pmax}$ es la frecuencia fundamental con la que luego se ejecuta el procedimiento 2 que conduce a frecuencias sinusoidales mejoradas \hat{f}_k . Sin embargo, una alternativa más preferible es primero optimizar la estimación de frecuencia

5 fundamental f_0 basándose en los picos de frecuencia \hat{f}_k que se han encontrado que coinciden con las frecuencias armónicas. Suponer un conjunto de M armónicos, es decir, múltiplos enteros $\{n_1 \dots n_M\}$ de alguna frecuencia fundamental que se ha encontrado que coincide con algún conjunto de M picos espectrales en las frecuencias $\hat{f}_{k(m)}$, $m = 1 \dots M$, entonces puede calcularse la estimación de frecuencia fundamental (optimizada) subyacente $f_{0,opt}$ para minimizar el error entre las frecuencias armónicas y las frecuencias de pico espectrales. Si el error a minimizar es el

$$E_2 = \sum_{m=1}^M (n_m \cdot f_0 - \hat{f}_{k(m)})^2,$$

error cuadrático medio
óptima se calcula como

entonces la estimación de frecuencia fundamental

$$f_{0,opt} = \frac{\sum_{m=1}^M n_m \cdot \hat{f}_{k(m)}}{\sum_{m=1}^M n_m^2}.$$

10 El conjunto inicial de valores candidatos $\{f_{0,1} \dots f_{0,p}\}$ puede obtenerse a partir de las frecuencias de los picos de la DFT o de las frecuencias sinusoidales estimadas \hat{f}_k .

Mejora entre tramas de la estimación de frecuencia

15 Según esta realización, se mejora la precisión de las frecuencias sinusoidales estimadas \hat{f}_k considerando su evolución temporal. Así, las estimaciones de las frecuencias sinusoidales de un múltiplo de las tramas de análisis se combinan, por ejemplo, mediante promediado o predicción. Antes de promediar o predecir, se aplica un seguimiento de picos que conecta los picos espectrales estimados con las mismas sinusoides subyacentes respectivas.

Aplicar un modelo Sinusoidal

20 La aplicación de un modelo sinusoidal para realizar una operación de ocultación de pérdida de trama según las realizaciones puede describirse como sigue:

25 En caso de que un segmento dado de la señal codificada no pueda reconstruirse por el decodificador ya que la información codificada correspondiente no está disponible, es decir, ya que se ha perdido una trama, una parte disponible de la señal anterior a este segmento puede utilizarse como trama prototipo. Si $y(n)$ con $n=0 \dots N-1$ es el segmento no disponible para el que tiene que generarse una trama de sustitución $z(n)$, e $y(n)$ con $n < 0$ es la señal previamente codificada disponible, una trama prototipo de la señal disponible de longitud L e índice de comienzo n_{-1} se extrae con una función de ventana $w(n)$ y se transforma en el dominio de la frecuencia, por ejemplo, mediante la DFT:

$$Y_{-1}(m) = \sum_{n=0}^{L-1} y(n - n_{-1}) \cdot w(n) \cdot e^{-j\frac{2\pi}{L}nm}$$

40 La función de ventana puede ser una de las funciones de ventana descritas anteriormente en el análisis sinusoidal. Preferiblemente, para ahorrar en complejidad numérica, la trama transformada al dominio de la frecuencia debe ser idéntica a la utilizada durante el análisis sinusoidal, lo que significa que la trama de análisis y la trama prototipo serán idénticas, e igualmente sus respectivas transformadas del dominio de la frecuencia.

45 En un siguiente paso se aplica la suposición del modelo sinusoidal. Según la suposición del modelo sinusoidal, la DFT de la trama prototipo puede escribirse como sigue:

$$Y_{-1}(m) = \frac{1}{2} \sum_{k=1}^K a_k \cdot \left(W\left(2\pi\left(\frac{m}{L} + \frac{f_k}{f_s}\right)\right) \cdot e^{-j\varphi_k} + W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j\varphi_k} \right) .$$

Esta expresión también fue utilizada en la parte de análisis y se describió con detalle anteriormente.

5 A continuación, se consigue que el espectro de la función de ventana utilizada tenga solo una contribución significativa en un rango de frecuencia cercano a cero. Como se señaló anteriormente, el espectro de magnitud de la función de ventana es grande para frecuencias cercanas a cero y pequeño en caso contrario (dentro del rango de frecuencias normalizado de $-\pi$ a π , correspondiente a la mitad de la frecuencia de muestreo). Por lo tanto, se supone como una aproximación que el espectro de ventana $W(m)$ no es cero solo para un intervalo $M = [-m_{min}, m_{max}]$, con m_{min} y m_{max} siendo pequeños números positivos. En particular, se utiliza una aproximación del espectro de la función de ventana de manera que para cada k las contribuciones del espectro de la ventana desplazada en la expresión anterior no están estrictamente superpuestas. Por lo tanto en la ecuación anterior, para cada índice de frecuencia siempre hay solo, como máximo, la contribución de un sumando, es decir, de un espectro de ventana desplazada. Esto significa que la expresión anterior se reduce a la siguiente expresión aproximada:

$$15 \quad \tilde{Y}_{-1}(m) = \frac{a_k}{2} \cdot W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j\varphi_k} \quad \text{para } m \in M_k \text{ no negativo y para cada } k.$$

$$M_k = \left[\text{round}\left(\frac{f_k}{f_s} \cdot L\right) - m_{min,k} , \right.$$

20 En la presente memoria, M_k denota el intervalo entero:

redondeo $\left. \left(\frac{f_k}{f_s} \cdot L \right) + m_{max,k} \right]$, donde $m_{min,k}$ y $m_{max,k}$ cumplen la restricción explicada anteriormente de manera que los intervalos no se superponen. Una opción adecuada para $m_{min,k}$ y $m_{max,k}$ es establecerlos a un pequeño valor entero δ , por ejemplo, $\delta = 3$. En cambio, si los índices de la DFT relacionados con las dos frecuencias sinusoidales vecinas f_k y f_{k+1} son menores que 2δ , entonces δ se establece a

$$25 \quad \left(\frac{\text{round}\left(\frac{f_{k+1}}{f_s} \cdot L\right) - \text{round}\left(\frac{f_k}{f_s} \cdot L\right)}{2} \right)$$

suelo de manera que se garantice que los intervalos no se superpongan. La función suelo (\cdot) es el entero más cercano al argumento de la función que es más pequeño o igual a él.

30 El siguiente paso según las realizaciones es aplicar el modelo sinusoidal según la expresión anterior y evolucionar sus K sinusoides en el tiempo. La suposición de que los índices de tiempo del segmento borrado comparado con los índices de tiempo de la trama prototipo defieren por n_{-1} muestras significa que las fases de las sinusoides avanzan por

$$35 \quad \theta_k = 2\pi \cdot \frac{f_k}{f_s} n_{-1} .$$

Por lo tanto, el espectro de la DFT del modelo sinusoidal evolucionado es dado por:

$$Y_0(m) = \frac{1}{2} \sum_{k=1}^K a_k \cdot \left(W\left(2\pi\left(\frac{m}{L} + \frac{f_k}{f_s}\right)\right) \cdot e^{-j(\varphi_k + \theta_k)} + W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j(\varphi_k + \theta_k)} \right) .$$

Aplicando de nuevo la aproximación, según la cual el espectro de la función de ventana desplazada no se superpone,

$$\tilde{Y}_0(m) = \frac{a_k}{2} \cdot W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j(\varphi_k + \theta_k)}$$

40 da: para $m \in M_k$ no negativo y para cada k . Comparando la DFT de la trama prototipo $Y_{-1}(m)$ con la DFT del modelo sinusoidal evolucionado $Y_0(m)$ utilizando la

aproximación, se encuentra que el espectro de magnitud permanece sin cambios mientras que la fase es desplazada

por $\theta_k = 2\pi \cdot \frac{f_k}{f_s} n_{-1}$, para cada $m \in M_k$. Por lo tanto, la trama de sustitución puede ser calculada por la siguiente expresión:

5 $z(n) = \text{IDFT}\{Z(m)\}$ con $Z(m) = Y(m) \cdot e^{j\theta_k}$ para $m \in M_k$ no negativo y para cada k , donde IDFT denota la inversa de la DFT.

10 Una realización específica aborda la aleatorización de fase para los índices de la DFT que no pertenecen a ningún intervalo M_k . Como se describió anteriormente, los intervalos M_k , $k=1 \dots K$ tienen que ser establecidos de manera que no estén estrictamente superpuestos lo que es hecho utilizando algún parámetro δ , que controla el tamaño de los intervalos. Puede ocurrir que δ sea pequeño en relación con la distancia de frecuencia de dos sinusoides vecinas. Por lo tanto, en el caso de que ocurra eso hay un hueco entre dos intervalos. En consecuencia, para los índices de la DFT m correspondientes no se define desplazamiento de fase según la expresión anterior $Z(m) = Y(m) \cdot e^{j\theta_k}$. Una opción adecuada según esta realización es aleatorizar la fase para estos índices, produciendo $Z(m) = Y(m) \cdot e^{j2\pi \text{rand}(\cdot)}$, donde la función aleatoria(\cdot), devuelve un número aleatorio.

15 A continuación, se describen realizaciones que adaptan el tamaño de los intervalos M_k en respuesta a la tonalidad de la señal.

20 Una realización de esta invención comprende la adaptación del tamaño de los intervalos M_k en respuesta a la tonalidad de la señal. Esta adaptación puede combinarse con la estimación de frecuencia mejorada descrita anteriormente, que utiliza, por ejemplo, una aproximación de lóbulo principal, una mejora armónica, o una mejora entre tramas. Sin embargo, una adaptación del tamaño de los intervalos M_k en respuesta a la tonalidad de la señal puede realizarse, de manera alternativa, sin ninguna estimación de frecuencia mejorada anterior.

25 Se ha encontrado beneficioso para la calidad de las señales reconstruidas optimizar el tamaño de los intervalos M_k . En particular, los intervalos deberían ser mayores si la señal es muy tonal, es decir, cuando tiene picos espectrales claros y distintos. Este es el caso, por ejemplo, cuando la señal es armónica con una clara periodicidad. En otros casos donde la señal tiene una estructura espectral menos pronunciada con máximos espectrales más amplios, se ha encontrado que el uso de pequeños intervalos conduce a una mejor calidad. Este hallazgo conduce a una mejora adicional según la cual el tamaño del intervalo se adapta según las propiedades de la señal. Una realización es utilizar una tonalidad o un detector de periodicidad. Si este detector identifica la señal como tonal, el parámetro δ que controla el tamaño del intervalo se establece a un valor relativamente grande. De lo contrario, el parámetro δ se establece a valores relativamente más pequeños.

30 Se realiza un análisis sinusoidal de una parte de una señal de audio previamente recibida o reconstruida, en donde el análisis sinusoidal implica, en un paso, la identificación de frecuencias de los componentes sinusoidales, es decir, sinusoides, de la señal de audio. En un paso, se aplica un modelo sinusoidal sobre un segmento de la señal de audio previamente recibida o reconstruida, en donde dicho segmento se utiliza como una trama prototipo para crear una trama de sustitución para una trama de audio perdida, y en un paso se crea la trama de sustitución para la trama de audio perdida, que implica la evolución en el tiempo de los componentes sinusoidales, es decir, sinusoides, de la trama prototipo, hasta la instancia de tiempo de la trama de audio perdida, en respuesta a las correspondientes frecuencias identificadas. Sin embargo, el paso de la identificación de frecuencias de los componentes sinusoidales y/o el paso de la creación de la trama de sustitución puede comprender además la realización de al menos uno de una estimación de frecuencia mejorada en la identificación de frecuencias, y una adaptación de la creación de la trama de sustitución en respuesta a la tonalidad de la señal de audio. La estimación de frecuencia mejorada comprende al menos uno de una aproximación de lóbulo principal, una mejora armónica, y una mejora entre tramas.

35 Según una realización adicional, se supone que la señal de audio está compuesta de un número limitado de componentes sinusoidales individuales.

40 Según una realización ilustrativa, el método comprende la extracción de una trama prototipo a partir de una señal disponible previamente recibida o reconstruida utilizando una función de ventana, y en donde la trama prototipo extraída puede transformarse en una representación del dominio de la frecuencia.

45 Según una primera realización alternativa, la estimación de frecuencia mejorada comprende la aproximación de la forma de un lóbulo principal de un espectro de magnitud relacionado con una función de ventana, y puede comprender además identificar uno o más picos espectrales, k , y los correspondientes índices de la transformada discreta del dominio de la frecuencia m_k asociados con una trama de análisis; derivar una función $P(q)$ que aproxima el espectro de magnitud relacionado con la función de ventana, y para cada pico, k , con un índice de la transformada discreta del dominio de la frecuencia m_k correspondiente, ajustar una función de frecuencia desplazada $P(q - q_k)$ a través de dos puntos de cuadrícula de la transformada discreta del dominio de la frecuencia que rodean un pico real esperado de un espectro continuo de una señal del modelo sinusoidal asumido asociada con la trama de análisis.

- Según una segunda realización alternativa, la estimación de frecuencia mejorada es una mejora armónica, que comprende determinar si la señal de audio es armónica, y derivar una frecuencia fundamental, si la señal es armónica. La determinación puede comprender al menos uno de la realización de un análisis de auto-correlación de la señal de audio y la utilización de un resultado de una predicción de tono de bucle cerrado, por ejemplo, la ganancia de tono. El
- 5 paso de la derivación puede comprender la utilización de un resultado adicional de una predicción de tono de bucle cerrado, por ejemplo, el desfase de tono. Según además esta segunda realización alternativa, el paso de la derivación puede comprender la comprobación de, para un índice armónico j , si existe un pico en un espectro de magnitud dentro de la proximidad de una frecuencia armónica asociada con dicho índice armónico y una frecuencia fundamental, estando asociado el espectro de magnitud con el paso de la identificación.
- 10 Según una tercera realización alternativa, la estimación de frecuencia mejorada es una mejora entre tramas, que comprende la combinación de frecuencias identificadas de dos o más tramas de la señal de audio. La combinación puede comprender un promedio y/o una predicción, y puede aplicarse un seguimiento de picos antes del promedio y/o predicción.
- 15 Según una realización, la adaptación en respuesta a la tonalidad de la señal de audio implica la adaptación de un tamaño de un intervalo M_k localizado en la proximidad de una componente sinusoidal k , dependiendo de la tonalidad de la señal de audio. Además, la adaptación del tamaño de un intervalo puede comprender el incremento del tamaño del intervalo para una señal de audio que tiene, comparativamente, más picos espectrales distintos, y la reducción del tamaño del intervalo para una señal de audio que tiene, comparativamente, picos espectrales más anchos.
- 20 El método según las realizaciones puede comprender la evolución en el tiempo de los componentes sinusoidales de un espectro de frecuencia de una trama prototipo avanzando la fase de una componente sinusoidal, en respuesta a la frecuencia de esta componente sinusoidal y en respuesta a la diferencia de tiempo entre la trama de audio perdida y la trama prototipo. Puede comprender además el cambio de un coeficiente espectral de la trama prototipo incluido en el intervalo M_k localizado en la proximidad de una senoide k por un desplazamiento de fase proporcional a la frecuencia sinusoidal f_k y a la diferencia de tiempo entre la trama de audio perdida y la trama prototipo.
- 25 Las realizaciones también pueden comprender una transformada inversa del dominio de la frecuencia del espectro de frecuencia de la trama prototipo, después de los cambios de los coeficientes espectrales descritos anteriormente.
- 30 Más específicamente, el método de ocultación de pérdida de trama de audio según una realización adicional puede implicar los siguientes pasos:
- 35 1) Analizar un segmento de la señal disponible, previamente sintetizada, para obtener las frecuencias sinusoidales constituyentes f_k de un modelo sinusoidal.
- 40 2) Extraer una trama prototipo y_{-1} a partir de la señal previamente sintetizada disponible y calcular la DFT de esa trama.
- 45 3) Calcular el desplazamiento de fase θ_k para cada senoide k en respuesta a la frecuencia sinusoidal f_k y al avance de tiempo n_{-1} entre la trama prototipo y la trama de sustitución, en donde el tamaño del intervalo M_k puede haber sido adaptado en respuesta a la tonalidad de la señal de audio.
- 50 4) Para cada senoide k avanzar la fase de la DFT de la trama prototipo con θ_k de manera selectiva para los índices de la DFT relacionados con una proximidad alrededor de la frecuencia de senoide f_k .
- 55 5) Calcular la DFT inversa del espectro obtenido en el paso 4).
- 60 Las realizaciones descritas anteriormente pueden además ser explicadas por las siguientes suposiciones:
- d) La suposición de que la señal puede ser representada por un número limitado de sinusoides.
- e) La suposición de que la trama de sustitución está suficientemente bien representada por estas sinusoides evolucionadas en el tiempo, en comparación con algún instante de tiempo anterior.
- f) La suposición de una aproximación del espectro de una función de ventana de manera que el espectro de la trama de sustitución puede construirse mediante partes no superpuestas del espectro de la función de ventana desplazada en frecuencia, siendo las frecuencias desplazadas las frecuencias de senoide.
- Lo siguiente está relacionado con un método de control para Phase ECU, que fue mencionado previamente.
- Adaptación del método de ocultación de pérdida de trama
- En caso de que los pasos realizados anteriormente indiquen una condición que sugiera una adaptación de la operación de ocultación de pérdida de trama, se modifica el cálculo del espectro de la trama de sustitución:
- 65

Mientras que el cálculo original del espectro de la trama de sustitución se hace según la expresión $Z(m) = Y(m) \cdot e^{j\theta k}$, se introduce ahora una adaptación que modifica tanto la magnitud como la fase. La magnitud se modifica mediante el escalado con dos factores $\alpha(m)$ y $\beta(m)$ y la fase se modifica con un componente de fase aditivo $\vartheta(m)$. Esto conduce al siguiente cálculo modificado de la trama de sustitución:

5

$$Z(m) = \alpha(m) \cdot \beta(m) \cdot Y(m) \cdot e^{j(\theta_k + \vartheta(m))}.$$

10

Debe señalarse que los métodos originales (no adaptados) de ocultación de pérdida de trama se utilizan si $\alpha(m) = 1$, $\beta(m) = 1$ y $\vartheta(m) = 0$. Por lo tanto, estos respectivos valores son los predeterminados.

15

El objetivo general con la introducción de las adaptaciones de magnitud es evitar los artefactos audibles del método de ocultación de pérdida de trama. Dichos artefactos pueden ser sonidos musicales o tonales o sonidos extraños que surgen a partir de repeticiones de sonidos transitorios. Dichos artefactos, a su vez, conducirían a degradaciones de la calidad, cuya prevención es el objetivo de las adaptaciones descritas. Una forma adecuada para dichas adaptaciones es modificar el espectro de magnitud de la trama de sustitución a un grado adecuado.

20

Se describirá ahora una realización de una modificación del método de ocultación. La adaptación de magnitud se hace, preferiblemente, si el contador de pérdidas de ráfaga n_{burst} excede algún umbral thr_{burst} , por ejemplo, $thr_{burst} = 3$. En ese caso se utiliza un valor menor que 1 para el factor de atención, por ejemplo, $\alpha(m) = 0,1$.

25

Sin embargo, se ha encontrado que es beneficioso realizar la atenuación con un grado creciente gradualmente. Una realización preferida que logra esto es definir un parámetro logarítmico que especifica un incremento logarítmico en la atenuación por trama, att_por_trama . Entonces, en caso de que el contador de ráfagas exceda el umbral, el factor de atenuación creciente gradualmente es calculado por

$$\alpha(m) = 10^c \cdot att_per_frame \cdot (n_{burst} - thr_{burst}).$$

30

Aquí la constante c es una mera constante de escala que permite especificar el parámetro att_por_trama , por ejemplo, en decibelios (dB).

35

Se hace una adaptación preferida adicional en respuesta al indicador si se estima que la señal es música o habla. Para el contenido musical en comparación con el habla, es preferible aumentar el umbral thr_{burst} y disminuir la atenuación por trama. Esto es equivalente a realizar la adaptación del método de ocultación de pérdida de trama con un grado inferior. Los antecedentes de este tipo de adaptación es que la música es, en general, menos sensible a ráfagas de pérdida más largas que el habla. Por lo tanto, el original, es decir, el método de ocultación de pérdida de trama no modificado todavía es preferible para este caso, al menos para un número mayor de pérdidas de trama seguidas.

40

Una adaptación adicional del método de ocultación respecto al factor de atenuación de magnitud se hace, preferiblemente, en caso de que se haya detectado un transitorio basándose en que el indicador $R_{l/r,band}(k)$ o, de manera alternativa, $R_{l/r}(m)$ o $R_{l/r}$ ha pasado un umbral. En ese caso una acción de adaptación adecuada es modificar el segundo factor de atenuación de magnitud $\beta(m)$ de manera que la atenuación total está controlada por el producto de los dos factores $\alpha(m) \cdot \beta(m)$.

45

$\beta(m)$ se establece en respuesta a un transitorio indicado. En caso de que se detecte una compensación el factor $\beta(m)$ se elige, preferiblemente, para reflejar la disminución de energía de la compensación. Una opción adecuada es establecer $\beta(m)$ al cambio de ganancia detectado:

50

$$\beta(m) = \sqrt{R_{l/r,band}(k)}, \text{ for } m \in I_k, k = 1 \dots K.$$

55

En caso de que se detecte un inicio, se encuentra ventajoso además, limitar el aumento de energía en la trama de sustitución. En ese caso el factor puede establecerse a algún valor fijo de, por ejemplo, 1, lo que significa que no hay atenuación, pero tampoco ninguna amplificación.

60

En lo anterior se debe señalar que el factor de atenuación de magnitud se aplica, preferiblemente, a la frecuencia de manera selectiva, es decir, con factores calculados individualmente para cada banda de frecuencia. En caso de que no se utilice el enfoque de banda, los correspondientes factores de atenuación de magnitud todavía pueden ser obtenidos de una manera análoga. $\beta(m)$ puede entonces establecerse individualmente para cada contenedor de la DFT en caso de que se utilice la detección transitoria selectiva de frecuencia en el nivel del contenedor de la DFT. O, en caso de que no se utilice ninguna indicación transitoria selectiva de frecuencia, $\beta(m)$ puede ser, a escala global, idéntico para todo m .

65

Una adaptación preferida adicional del factor de atenuación de magnitud se hace junto con una modificación de la fase mediante el componente de fase adicional $\vartheta(m)$. En caso de que se utilice, para un m dado, dicha modificación de fase, el factor de atenuación $\beta(m)$ se reduce aún más. Preferiblemente, incluso el grado de la modificación de fase es tenido en cuenta. Si la modificación de fase es solo moderada, $\beta(m)$ solo se reduce ligeramente, mientras que si la modificación de fase es fuerte, $\beta(m)$ se reduce en mayor medida.

El objetivo general con la introducción de adaptaciones de fase es evitar tonalidades demasiado fuertes o periodicidad de señal en las tramas de sustitución generadas, lo que a su vez conduciría a degradaciones de la calidad. Una forma adecuada para dichas adaptaciones es aleatorizar u oscilar la fase en un grado adecuado.

Dicha oscilación de fase se logra si el componente de fase adicional $\vartheta(m)$ se establece a un valor aleatorio escalado con algún factor de control: $\vartheta(m) = \alpha(m) \cdot \text{aleatoria}(\cdot)$.

El valor aleatorio obtenido por la función aleatoria(\cdot) es, por ejemplo, generado por algún generador de números pseudo aleatorios. Aquí se supone que se proporciona un número aleatorio dentro del intervalo $[0, 2\pi]$.

El factor de escala $\alpha(m)$ en la ecuación anterior controla el grado por el que se oscila la fase original θ_k . Las siguientes realizaciones abordan la adaptación de fase mediante el control de este factor de escala. El control del factor de escala se hace de una forma análoga al control de los factores de modificación de magnitud descrito anteriormente.

Según una primera realización el factor de escala $\alpha(m)$ se adapta en respuesta al contador de pérdidas de ráfaga. Si el contador de pérdidas de ráfaga n_{burst} excede algún umbral thr_{burst} , por ejemplo, $thr_{burst} = 3$, se utiliza un valor mayor que 0, por ejemplo, $\alpha(m) = 0,2$.

Sin embargo, se ha encontrado beneficioso realizar la oscilación con un grado creciente gradualmente. Una realización preferida que logra esto es definir un parámetro que especifica un incremento en la oscilación por trama, *incremento_oscilación_por_trama*. Entonces, en caso de que el contador de ráfagas exceda el umbral el factor de control de oscilación creciente gradualmente es calculado por

$$\alpha(m) = \text{dith_increase_per_frame} \cdot (n_{burst} - thr_{burst}).$$

Debe señalarse en la anterior fórmula que $\alpha(m)$ tiene que limitarse a un valor máximo de 1 para que se alcance la oscilación de fase completa.

Debe señalarse que el valor del umbral de pérdida de ráfaga thr_{burst} utilizado para la oscilación de fase inicial puede ser el mismo umbral que el utilizado para la atenuación de magnitud. Sin embargo, puede obtenerse mejor calidad estableciendo estos umbrales a valores óptimos de manera individual, lo que, en general, significa que estos umbrales pueden ser diferentes.

Se hace una adaptación preferida adicional en respuesta al indicador si se estima que la señal es música o habla. Para el contenido musical en comparación con el habla, es preferible aumentar el umbral thr_{burst} lo que significa que la oscilación de fase para la música, en comparación con el habla, se hace solo en caso de más tramas perdidas seguidas. Esto es equivalente a realizar la adaptación del método de ocultación de pérdida de trama para la música con un grado inferior. Los antecedentes de este tipo de adaptación es que la música es, en general, menos sensible a ráfagas de pérdida más largas que el habla. Por lo tanto, el original, es decir, el método de ocultación de pérdida de trama no modificado todavía es preferible para este caso, al menos para un número mayor de pérdidas de trama seguidas.

Una realización preferida adicional es adaptar la oscilación de fase en respuesta a un transitorio detectado. En ese caso puede utilizarse un mayor grado de oscilación de fase para los contenedores de la DFT m para los que es indicado un transitorio bien para ese contenedor, para los contenedores de la DFT de la banda de frecuencia correspondiente o de la trama completa.

Parte de los esquemas descritos abordan la optimización del método de ocultación de pérdida de trama para señales armónicas y, particularmente, para la voz sonora.

En caso de que los métodos que utilizan una estimación de frecuencia mejorada como se describió anteriormente no se realicen, otra posibilidad de adaptación para el método de ocultación de pérdida de trama que optimiza la calidad para las señales de voz sonora, es cambiar a algún otro método de ocultación de pérdida de trama que esté diseñado y optimizado, específicamente, para el habla más que para las señales de audio generales que contienen música y habla. En ese caso, se utiliza el indicador de que la señal comprende una señal de voz sonora para seleccionar otro esquema de ocultación de pérdida de trama optimizado para el habla antes que los esquemas descritos anteriormente.

En resumen, debe entenderse que la elección de unidades o módulos de interacción, así como la nomenclatura de las unidades son solo para fines ilustrativos, y pueden configurarse en una pluralidad de modos alternativos para poder ejecutar las acciones del proceso descrito.

5 Cabe señalar que las unidades o módulos descritos en esta descripción deben considerarse como entidades lógicas y no necesariamente como entidades físicas separadas. Se apreciará que el alcance de la tecnología descrita en la presente memoria abarca completamente otras realizaciones que resultan obvias para los expertos en la técnica, y que el alcance de esta descripción no se limita en consecuencia.

10 La referencia a un elemento en singular no pretende significar “uno y solo uno”, a menos que se indique explícitamente, sino más bien “uno o varios”. Todos los equivalentes estructurales y funcionales a los elementos de las realizaciones descritas anteriormente, que son conocidos por los expertos en la técnica, se incorporan expresamente en la presente memoria como referencia, y se pretende que se incorporen por la presente. Además, no es necesario para un dispositivo o método abarcar todos y cada uno de los problemas que se intentan resolver mediante la tecnología descrita en la presente memoria, para que se abarquen por la presente.

15 En la descripción anterior, para fines de explicación y no de limitación, se establecen detalles específicos como arquitecturas, interfaces, técnicas, etc. particulares para proporcionar una comprensión profunda de la tecnología descrita. Sin embargo, será evidente para los expertos en la técnica que la tecnología descrita puede practicarse en otras realizaciones y/o combinaciones de realizaciones que se apartan de estos detalles específicos. Es decir, los expertos en la técnica podrán idear varias disposiciones que, aunque no se describen o muestran explícitamente en la presente memoria, encarnan los principios de la tecnología descrita. En ciertas ocasiones, se omiten descripciones detalladas de dispositivos bien conocidos, circuitos, y métodos para no oscurecer la descripción de la tecnología descrita con detalles innecesarios. Todas las declaraciones en la presente memoria que recitan principios, aspectos, y realizaciones de la tecnología descrita, así como ejemplos específicos de los mismos, se pretende que abarquen tanto los equivalentes estructurales como funcionales de los mismos. Además, se pretende que dichos equivalentes incluyan tanto los equivalentes conocidos actualmente como los equivalentes desarrollados en el futuro, por ejemplo, cualesquiera elementos desarrollados que realicen la misma función, independientemente de la estructura.

20 Así, por ejemplo, será apreciado por los expertos en la técnica que las figuras en la presente memoria pueden representar vistas conceptuales de circuitos ilustrativos u otras unidades funcionales que encarnan los principios de la tecnología, y/o varios procesos que pueden, sustancialmente, representarse en un medio legible por ordenador y ser ejecutados por un ordenador o procesador, aunque dicho ordenador o procesador no se muestre explícitamente en las figuras.

25 Las funciones de los diversos elementos que incluyen bloques funcionales pueden proporcionarse a través del uso de hardware, como hardware de circuito y/o hardware capaz de ejecutar software en forma de instrucciones codificadas almacenadas en un medio legible por ordenador. Así, dichas funciones y bloques funcionales ilustrados deben entenderse como implementadas bien en hardware y/o implementadas en ordenador, y por tanto, implementadas en máquina.

30 Las realizaciones descritas anteriormente deben entenderse como algunos ejemplos ilustrativos de la presente invención. Los expertos en la materia entenderán que se pueden realizar diversas modificaciones, combinaciones y cambios a las realizaciones sin salirse del alcance de la presente invención. En particular, pueden combinarse diferentes soluciones parciales en las diferentes realizaciones en otras configuraciones, donde sea técnicamente posible.

35 El concepto inventivo ha sido descrito anteriormente, principalmente, con referencia a algunas realizaciones. Sin embargo, como aprecia fácilmente una persona experta en la técnica, son igualmente posibles otras realizaciones que las descritas anteriormente dentro del alcance del concepto inventivo, según lo definido por las reivindicaciones de patente adjuntas.

REIVINDICACIONES

- 5 1. Un método de ocultación de pérdida de trama para el gestión de error de ráfaga, siendo realizado el método por una entidad receptora, comprendiendo el método:
- 10 generar un espectro de la trama de sustitución mediante el uso de un método primario de ocultación de pérdida de trama, en donde el espectro de la trama de sustitución se basa en un espectro de una trama de una señal de audio recibida previamente;
- determinar (S101) un componente de ruido, en donde una frecuencia característica del componente de ruido es una representación espectral de baja resolución de la trama de la señal de audio recibida previamente;
- 15 determinar (S102) si un número n de tramas pérdidas o erróneas excede un umbral;
- si el número n de tramas pérdidas o erróneas es menor que el umbral, añadir (S104, S208) el componente de ruido al espectro de la trama de sustitución;
- si el número n de tramas pérdidas o erróneas excede el umbral, aplicar (S103, S206) un factor de atenuación y al componente de ruido y después de la aplicación del factor de atenuación, añadir (S104, S208) el componente de ruido al espectro de la trama de sustitución.
2. El método según la reivindicación 1, en donde el umbral es mayor que o igual a 10.
- 20 3. Una entidad receptora (103, 200, 400, 800, 900) para la ocultación de pérdida de trama, comprendiendo la entidad receptora un circuito de procesamiento (803), siendo configurado el circuito de procesamiento para provocar que la entidad receptora:
- 25 genere un espectro de la trama de sustitución mediante el uso de un método primario de ocultación de pérdida de trama, en donde el espectro de la trama de sustitución se basa en un espectro de una trama de una señal de audio recibida previamente;
- determine un componente de ruido, en donde una frecuencia característica del componente de ruido es una representación espectral de baja resolución de la trama de la señal de audio recibida previamente;
- 30 determine si un número n de tramas pérdidas o erróneas excede un umbral;
- añada el componen de ruido al espectro de la trama de sustitución, si el número n de tramas pérdidas o erróneas es menor que el umbral;
- aplique un factor de atenuación y al componente de ruido si el número n de tramas pérdidas o erróneas excede el umbral y después de aplicar el factor de atenuación, añadir el componente de ruido al espectro de la trama de sustitución.
- 35 4. La entidad receptora según la reivindicación 3, en donde el umbral es mayor que o igual a 10.

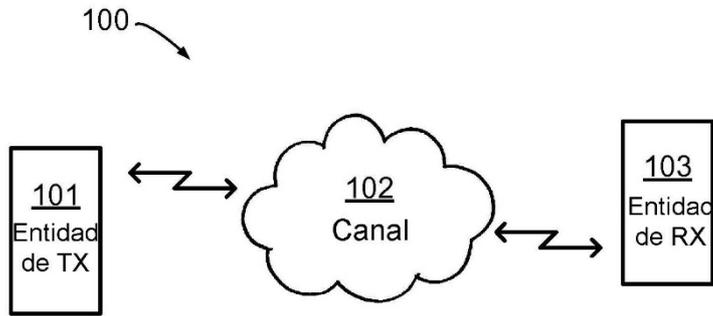


Fig. 1

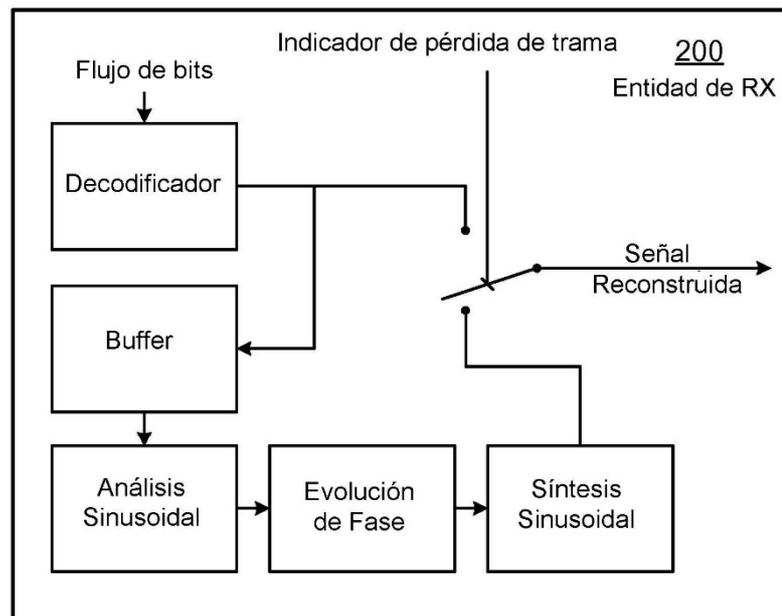


Fig. 2

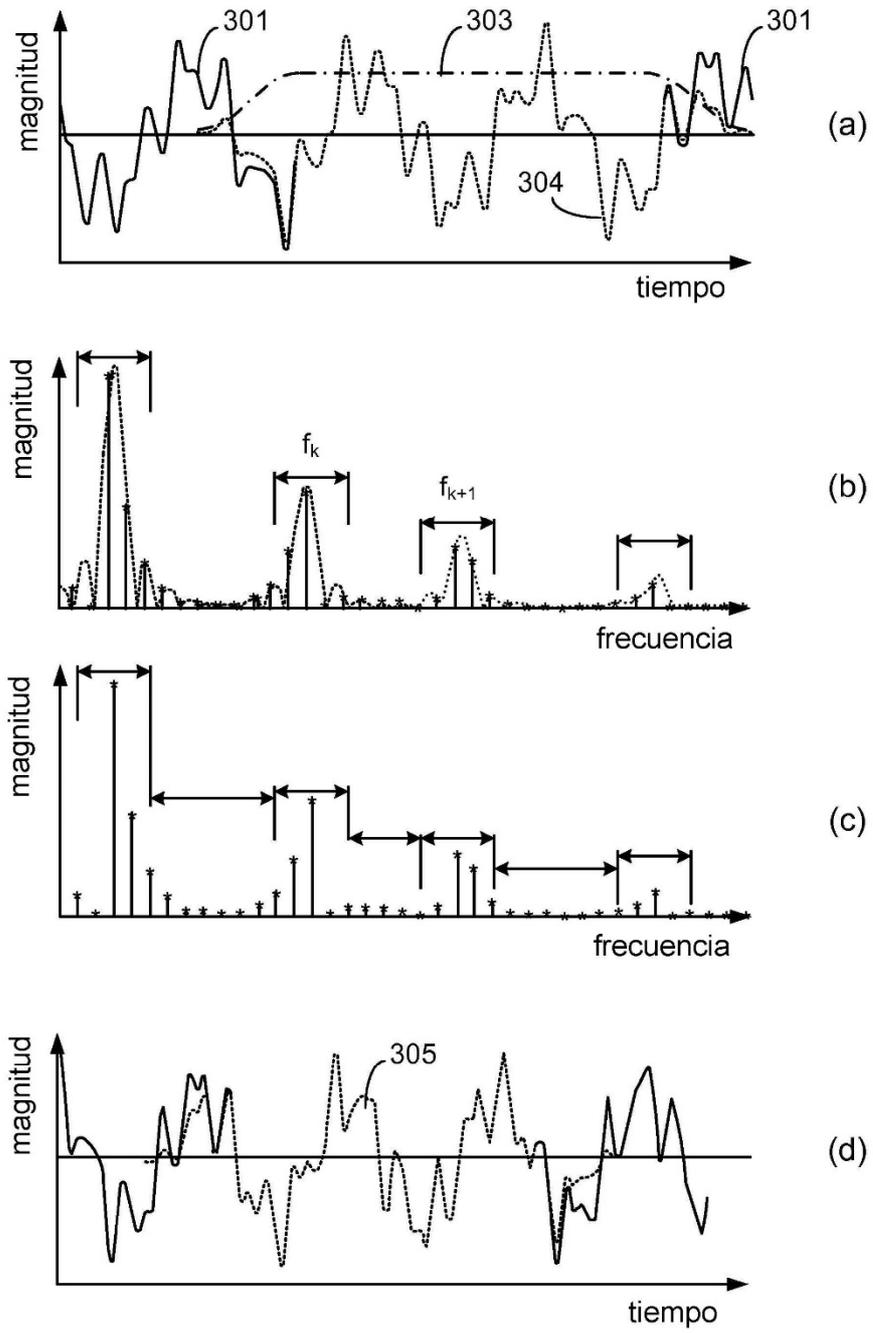


Fig. 3

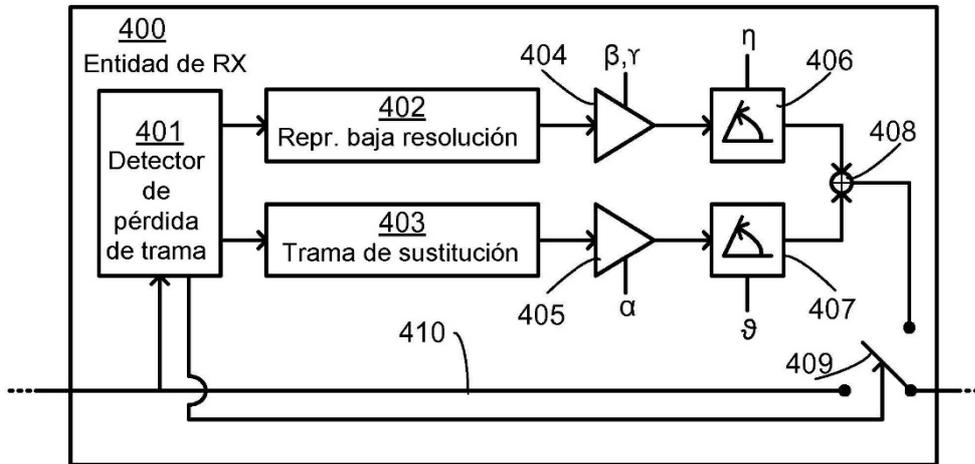


Fig. 4

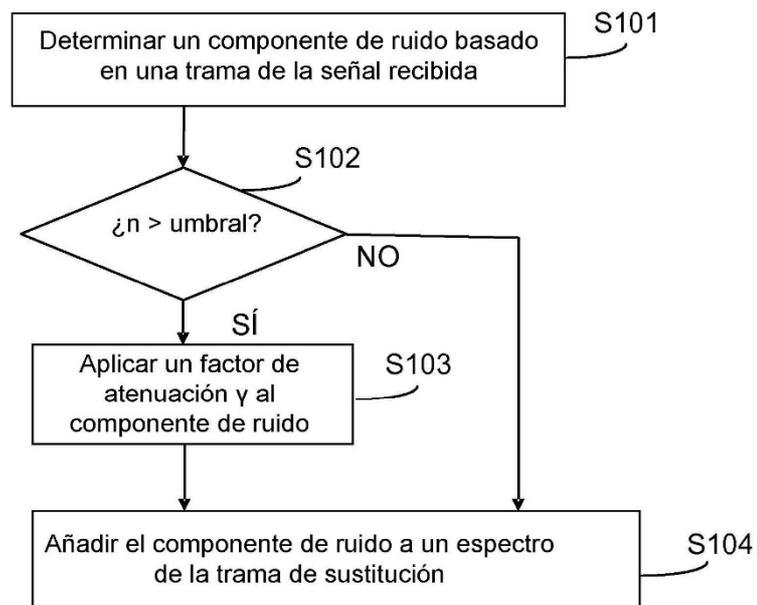


Fig. 5

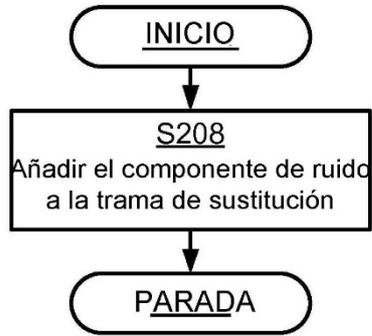


Fig. 6

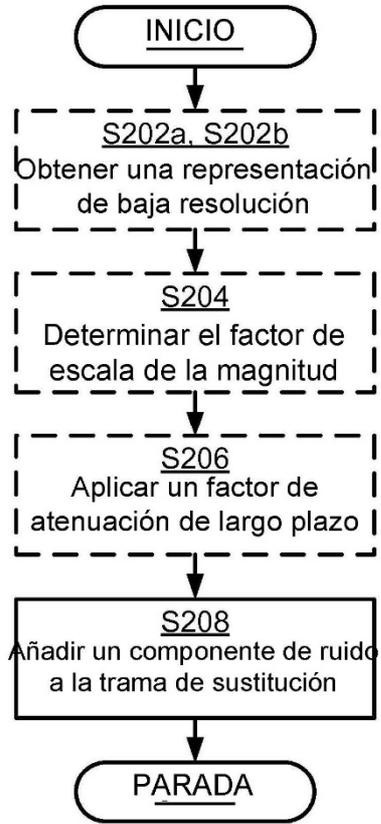


Fig. 7

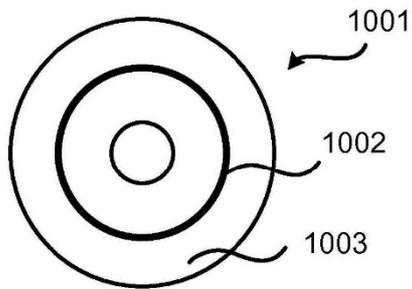


Fig. 10

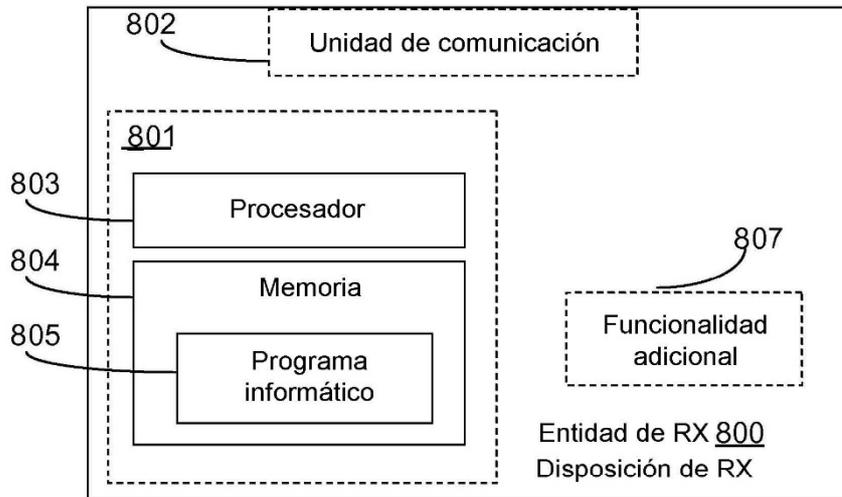


Fig. 8

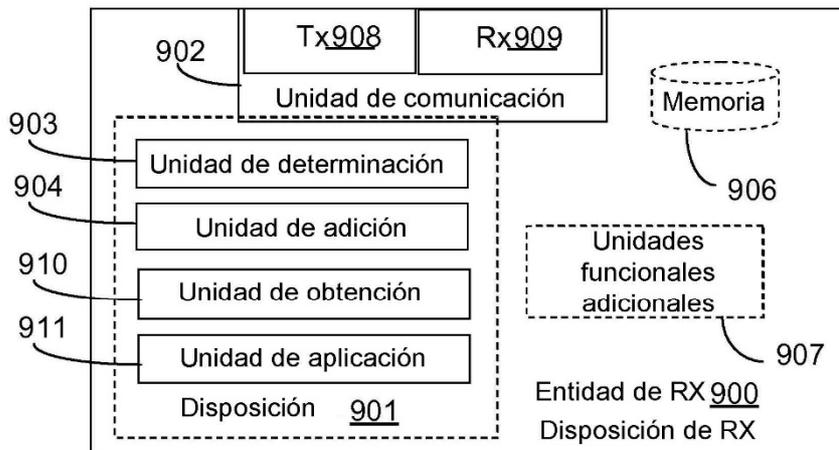


Fig. 9