

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 603 827**

51 Int. Cl.:

G10L 19/005 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.01.2014 PCT/SE2014/050068**

87 Fecha y número de publicación internacional: **14.08.2014 WO14123471**

96 Fecha de presentación y número de la solicitud europea: **22.01.2014 E 14704935 (7)**

97 Fecha y número de publicación de la concesión europea: **31.08.2016 EP 2954518**

54 Título: **Método y aparato para controlar la ocultación de pérdida de trama de audio**

30 Prioridad:

05.02.2013 US 201361761051 P
05.02.2013 US 201361760822 P
05.02.2013 US 201361760814 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
01.03.2017

73 Titular/es:

TELEFONAKTIEBOLAGET L M ERICSSON
(PUBL) (100.0%)
164 83 Stockholm, SE

72 Inventor/es:

BRUHN, STEFAN y
SVEDBERG, JONAS

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 603 827 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

Método y aparato para controlar la ocultación de pérdida de trama de audio

Campo de la técnica

5 La aplicación se refiere a los métodos y aparatos para controlar un método de ocultación para una trama de audio perdida de una señal de audio recibida.

Antecedentes

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

15 Sin embargo, dicho sistema de transmisión para voz y señales de audio puede sufrir errores de transmisión, que podrían llevar a una situación en la que una o varias de las tramas transmitidas no estén disponibles en el receptor para la reconstrucción. En ese caso, el decodificador tiene que generar una señal de sustitución para cada una de las tramas eliminadas, esto es, no disponibles. Esto se hace en la denominada unidad de pérdida de trama o de ocultación de error del decodificador de señal del extremo receptor. El propósito de la ocultación de la 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 tanto como sea posible.

20 Los métodos convencionales de ocultación de pérdida de trama pueden depender de la estructura o arquitectura del códec, por ejemplo aplicando una forma de repetición de los parámetros del códec previamente recibidos. Tales técnicas de repetición de parámetros son claramente dependientes de los parámetros específicos del códec usado y por tanto no son fácilmente aplicables a otros códec con una estructura diferente. Los métodos de ocultación de pérdida de trama pueden por ejemplo aplicar el concepto de congelación y extrapolación de parámetros de una trama recibida previamente para generar una trama de sustitución para la trama perdida.

25 Estos métodos de ocultación de trama del estado de la técnica de incorporan algunos esquemas de manejo de ráfagas de pérdida. En general, después de un número de pérdidas de trama en una fila la señal sintetizada se atenúa hasta que es completamente silenciada después de largas ráfagas de errores. Además, los parámetros de codificación que son esencialmente repetidos y extrapolados se modifican de tal manera que la atenuación se termina y los picos espectrales se aplanan.

30 Las técnicas de ocultación de pérdida de trama del estado de la técnica actual normalmente aplican el concepto de congelación y extrapolación de parámetros de una trama previamente recibida a fin de generar una trama de sustitución para la trama perdida. Muchos códec paramétricos de voz tales como los códec predictivos lineales como el AMR o AMR-WB normalmente congelan los parámetros previamente recibidos o usan alguna extrapolación de los mismos y usan el decodificador con ellos. En esencia, el principio es tener un modelo dado para codificar/decodificar y aplicar el mismo modelo con parámetros congelados o extrapolados. Las técnicas de ocultación de pérdida de trama del AMR y AMR-WB pueden ser consideradas como representativas. Se especifican en detalle en las correspondientes especificaciones de los estándares o normas.

35 Muchos códec fuera de la categoría de los códec de audio aplican para codificar técnicas del dominio de la frecuencia. Esto significa que después de alguna transformada en el dominio de la frecuencia se aplica un modelo de codificación sobre los parámetros espectrales. El decodificador reconstruye el espectro de la señal a partir de los parámetros recibidos y finalmente transforma el espectro de vuelta a una señal de tiempo. Normalmente, la señal de tiempo se reconstruye trama a trama, Tales tramas se combinan por técnicas de superposición-suma a la señal reconstruida final. Incluso en dicho caso de códec de audio, la ocultación de error del estado de la técnica normalmente aplica el mismo modelo o al menos un modelo de decodificación similar para las tramas perdidas. Los parámetros del dominio de la frecuencia procedentes de una trama previamente recibida se congelan o se extrapolan adecuadamente y luego se usan en la conversión del dominio de la frecuencia al del tiempo. Ejemplos para tales técnicas se proporcionan con los códec de audio 3GPP según los estándares 3GPP.

40 El documento US 2004/122680 de la técnica anterior describe un sistema para la ocultación de errores de trama que enseña ajustar la magnitud de la trama de sustitución dependiendo del número de tramas consecutivas perdidas.

Compendio

Las actuales soluciones del estado de la técnica de para la ocultación de pérdida de trama normalmente sufren de deficiencia de calidad. El principal problema es que la técnica de de congelación y extrapolación de parámetros y la aplicación de nuevo del mismo modelo de decodificador incluso para tramas perdidas no siempre garantiza una evolución de señal uniforme y fiel desde las tramas de señal previamente decodificadas a la trama perdida. Esto lleva normalmente a discontinuidades de señal audibles con el correspondiente impacto en la calidad.

5

Se describen nuevos esquemas para la ocultación de pérdida de trama para sistemas de transmisión de voz y de audio. Los nuevos esquemas mejoran la calidad en caso de pérdida de trama sobre la calidad alcanzable con técnicas de ocultación de trama de técnicas previas.

10 El objetivo de la presente realización es controlar un esquema de ocultación de pérdida de trama que preferiblemente sea del tipo de los nuevos métodos descritos referidos de tal modo que se alcance la mejor calidad de sonido posible de la señal reconstruida. Las realizaciones buscan optimizar esta calidad de reconstrucción tanto con respecto a las propiedades de la señal como a la distribución temporal de las pérdidas de trama. Particularmente problemáticos para la ocultación de pérdida de trama para proporcionar buena calidad son los casos cuando la señal de audio tiene propiedades fuertemente variables tales como inicios o cese de energías o si es espectralmente muy fluctuante. En ese caso los métodos de ocultación descritos pueden repetir el inicio, cese o la fluctuación espectral que lleven a grandes desviaciones respecto a la señal original y a la correspondiente calidad del servicio.

15

Otro caso problemático es si ráfagas de pérdidas de tramas ocurren en una fila. Conceptualmente, el esquema para la ocultación de la pérdida de trama según los métodos descritos puede hacer frente a tales casos, aunque como resultado puedan ocurrir molestos defectos tonales. Es otro objetivo de la presente realización mitigar tales defectos en el mayor grado posible.

20

Según un primer aspecto, se ha descrito un método para ocultar una trama de audio perdida según la reivindicación 1.

25 Según un segundo aspecto, se configura un aparato para implementar una ocultación de un nombre de trama perdida de audio, como se ha descrito en la reivindicación 12.

Según un tercer aspecto, se ha definido un programa de ordenador para ocultar una trama de audio perdida, y el programa de ordenador incluye instrucciones que al ser ejecutadas por un procesador provocan que el procesador oculte una trama de audio perdida, en acuerdo con el primer aspecto descrito más arriba.

30 Según un cuarto aspecto, un producto de programa de ordenador incluye un medio legible por ordenador que almacena un programa de ordenador según el cuarto aspecto descrito anteriormente. Una ventaja con una realización se dirige al control de los métodos de ocultación de pérdida de tramas de adaptación que permiten mitigar el impacto audible de la pérdida de trama en la transmisión de voz y señales de audio codificadas incluso más allá de la calidad alcanzada con sólo los métodos de ocultación descritos. El beneficio general de las realizaciones es proporcionar una evolución uniforme y fiel de la señal reconstruida incluso por tramas perdidas. El impacto audible de las pérdidas de trama se reduce notablemente en comparación a la utilización de técnicas del estado de la técnica.

35

Breve descripción de los dibujos

40 Para una comprensión más completa de las realizaciones ejemplares de la presente invención, se hace ahora referencia a la siguiente descripción tomada en conexión con los dibujos adjuntos en los cuales:

La Figura 1 muestra una función de ventana rectangular

La Figura 2 muestra una combinación de la ventana de Hamming con la ventana rectangular.

La Figura 3 muestra un ejemplo de un espectro de magnitud de una función de ventana.

La Figura 4 muestra un espectro de una señal sinusoidal con ventana con la frecuencia f_k .

45 La Figura 5 ilustra una línea de espectro de una señal sinusoidal ejemplar con la frecuencia f_k .

La Figura 6 ilustra barras correspondientes a la magnitud de los puntos de la cuadrícula de una DFT, basada en una trama de análisis.

La Figura 7 ilustra un ajuste de una parábola a través de los puntos de la cuadrícula de DFT P1, P2 y P3

La Figura 8 ilustra un ajuste de un lóbulo principal de un espectro de ventana.

La Figura 9 ilustra un ajuste de la función de aproximación P de lóbulo principal a través de los puntos de la cuadrícula de DFT P1 y P2.

La Figura 10 es un diagrama de flujo que ilustra un método ejemplar según las realizaciones de la invención para controlar un método de ocultación para una trama de audio perdida de una señal de audio recibida.

5 La Figura 11 es un diagrama de flujo que ilustra otro método ejemplar según las realizaciones de la invención para controlar un método de ocultación para una trama de audio perdida de una señal de audio recibida.

La Figura 12 ilustra otra realización ejemplar de la invención.

La Figura 13 muestra un ejemplo de un aparato según una realización de la invención.

La Figura 14 muestra otro ejemplo de un aparato según una realización de la invención.

10 La Figura 15 muestra otro ejemplo de un aparato según una realización de la invención.

Descripción detallada

El nuevo esquema de control para las nuevas técnicas de ocultación de trama perdida descritas implican los siguientes pasos como se muestra en la Figura 10. Se debería notar que el método se puede implementar en un controlador en un decodificador.

15 1. Detectar condiciones en las propiedades de la señal de audio previamente recibida y reconstruida o en las propiedades estadísticas de las tramas perdidas observadas para las que la sustitución de una trama perdida según los métodos descritos proporciona una calidad relativamente reducida, 101.

20 2. En el caso tal de que dicha condición sea detectada en el paso 1, modificar el elemento de los métodos según el cual el espectro de la trama de sustitución se calcula como $Z(m) = Y(m) \cdot e^{j\theta_k}$, ajustando selectivamente las fases o las magnitudes del espectro, 102.

Análisis sinusoidal

25 Un primer paso de la técnica de ocultación de pérdida de trama al que la nueva técnica de control puede aplicarse implica un análisis sinusoidal de una parte de la señal previamente recibida. El propósito de este análisis sinusoidal es encontrar las frecuencias de las sinusoides principales de esa señal, y la subyacente asunción es que la señal se compone de un número limitado de sinusoides individuales, es decir, esta es una señal multi-sinusoidal del siguiente tipo:

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

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

35 Es de principal importancia encontrar frecuencias de las sinusoides tan exactas como sea posible. Mientras una señal sinusoidal ideal tendría un espectro de línea con frecuencias de línea f_k , encontrar sus valores reales en principio requeriría un tiempo de medición infinito. Por lo tanto, hay una dificultad práctica para encontrar estas frecuencias ya que pueden sólo ser estimadas en base a un corto periodo de medición, que corresponde al segmento de señal usado para el análisis de señal sinusoidal descrito aquí; a este segmento de señal se referirá de aquí en adelante como una trama de análisis. Otra dificultad es que la señal puede en la práctica ser variable en el tiempo, lo que significa que los parámetros de la ecuación anterior varían con el tiempo. Por lo tanto, por un lado es deseable usar una trama de análisis larga, que haga la medición más precisa, por el otro lado sería necesario un periodo de medición corto a fin de hacer frente mejor a posibles variaciones de señal. Una buena solución intermedia es usar una longitud de trama de análisis del orden de por ejemplo 20-40 ms.

40 Una posibilidad preferida para identificar las frecuencias de las sinusoides f_k , es hacer un análisis en el dominio de la frecuencia de la trama de análisis. Para tal fin la trama de análisis se transforma al dominio de la frecuencia, por ejemplo mediante DFT o DCT o transformadas similares en el dominio de la frecuencia. En caso de que se use una DFT de la trama de análisis, el espectro viene dado por:

45

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

En esta ecuación $w(n)$ denota la función ventana con la cual la trama de análisis de longitud L se extrae y se pondera. Funciones de ventana típicas son por ejemplo ventanas rectangulares que son iguales a 1 para $n [0 \dots L-1]$ y de lo contrario 0 como se muestra en la Figura 1. Se supone aquí que los índices de tiempo de la señal de audio

previamente recibida están configurados de tal modo que la trama de análisis se referencia por los índices de tiempo $n=0\dots L-1$. Otras funciones de ventana que pueden ser más ajustables para el análisis espectral son, por ejemplo, ventana de Hamming, ventana de Hanning, ventana de Kaiser o ventana de Blackman, Una función de ventana que se ha encontrado particularmente útil es una combinación de la ventana de Hamming con la ventana rectangular.

5 Esta ventana tiene una forma de borde ascendente como la mitad izquierda de una ventana de Hamming de longitud $L/2$ y un borde descendente como la mitad derecha de una ventana de Hamming de longitud $L/2$ y entre los bordes ascendente y descendente la ventana es igual a 1 para la longitud de $L/2$, como se muestra en la Figura 2.

Los picos del espectro de magnitud de la trama de análisis con ventana $|X(m)|$ constituyen una aproximación de las frecuencias sinusoidales requeridas f_k . La precisión de esta aproximación está sin embargo limitada por la separación de frecuencias de la DFT. Con la DFT con longitud de bloque L la precisión se limita a:

$$\frac{f_s}{2L}$$

Los experimentos muestran que este nivel de precisión puede ser demasiado bajo en el alcance de los métodos descritos aquí. Se puede obtener una precisión mejorada basada en los resultados de la siguiente consideración:

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

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

Mediante el uso de la expresión del espectro del modelo de señal sinusoidal, esto se puede escribir 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 viene 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

$$m=0\dots L-1.$$

Basado en esta consideración se asume que los picos observados en el espectro de magnitud de la trama de análisis provienen de una señal sinusoidal con ventana con K sinusoides donde las frecuencias de sinusoides auténticas se encuentran en las inmediaciones de los picos.

Supongamos que m_k es el índice de la DFT (puntos de cuadrícula) del pico $k^{\text{ésimo}}$ observado, entonces la frecuencia correspondiente es

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

la cual puede ser considerada una aproximación de la frecuencia sinusoidal auténtica f_k . La frecuencia sinusoidal verdadera f_k se puede suponer situada en el 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 hace notar que la convolución del espectro de la función ventana con el espectro del espectro de línea del modelo sinusoidal de señal se puede entender como una superposición de versiones desplazadas en frecuencia del espectro de la función ventana, por lo que las frecuencias de desplazamiento son las frecuencias de las sinusoides. Esta superposición se muestra luego en los puntos de la cuadrícula de la DFT, Estos pasos se ilustran por las siguientes figuras. La Figura 3 muestra un ejemplo del espectro de magnitud de una función ventana. La Figura 4 muestra el espectro de magnitud (espectro de línea) de un ejemplo de señal sinusoidal con una sinusoides de frecuencia individual. La Figura 5 muestra el espectro de magnitud de la señal sinusoidal con ventana que replica y superpone los espectros de ventana desplazada en frecuencia a las frecuencias de la sinusoides. Las

barras en la Figura 6 corresponden a la magnitud de los puntos de la cuadrícula de la DFT de la senoide con ventana que se obtienen calculando la DFT de la trama de análisis. Debería observarse que todos los espectros son periódicos con el parámetro de frecuencia normalizada Ω donde $\Omega = 2\pi$ que corresponde a la frecuencia de muestreo f_s .

- 5 La descripción anterior y la ilustración de la figura 6 sugieren que una mejor aproximación de las frecuencias sinusoidales auténticas sólo se puede encontrar aumentando la resolución de la búsqueda sobre la resolución de frecuencia de la transformada en el dominio de la frecuencia usada.

Una forma preferida de encontrar mejores aproximaciones de las frecuencias f_k de las senoideas es aplicar interpolación parabólica. Tal enfoque es ajustar parábolas a través de la matriz de puntos de cuadrícula del espectro de magnitud de DFT que rodean los picos y calcular las respectivas frecuencias pertenecientes a los máximos de parábola. Una elección apropiada para el orden de la parábola es 2. Se puede aplicar en detalle el siguiente procedimiento:

- 15 1. Identificar los picos de la DFT de la trama de análisis con ventana . La búsqueda de picos dará el número de picos K y los correspondientes índices DFT de los picos. La búsqueda de picos se puede hacer normalmente en el espectro de magnitud de la DFT o en el espectro de magnitud de la DFT logarítmico.

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

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

20 El ajuste de parábola se muestra en la Figura 7.

3. Para cada una de las K parábolas calcular el índice m_k de frecuencia interpolada que corresponde al valor de q para el cual la parábola tiene su máximo. Usar $f_k = m_k \cdot f_s/L$ como aproximación para la frecuencia sinusoidal f_k .

25 La aproximación descrita proporciona buenos resultados pero puede tener algunas limitaciones ya que las parábolas no se aproximan a la forma del lóbulo principal del espectro de magnitud $|W(\Omega)|$ de la función ventana. Un esquema alternativo que hace esto es una estimación de frecuencia mejorada que usa una aproximación de lóbulo principal, descrita a continuación. La principal idea de esta alternativa es ajustar una función $P(q)$, que se aproxima al lóbulo

principal de $|W(\frac{2\pi}{L} \cdot q)|$, mediante la matriz de puntos de cuadrícula del espectro de magnitud de DFT que rodea los picos y calcular las respectivas frecuencias pertenecientes a la función máxima. La función $P(q)$ podría ser idéntica

30 al espectro de magnitud desplazado en frecuencia $|W(\frac{2\pi}{L} \cdot (q - \hat{q}))|$ de la función ventana. Por simplicidad numérica aunque debería ser más bien como ejemplo un polinomio que permita un cálculo sencillo de la función máxima. Se puede aplicar el siguiente procedimiento detallado:

1. Identificar los picos de la DFT de la trama de análisis con ventana . La búsqueda de picos dará el número de picos K y los correspondientes índices DFT de los picos. La búsqueda de picos se puede hacer normalmente en el espectro de magnitud de la DFT o en el espectro de magnitud de la DFT logarítmica.

35 2. Derivar la función $P(q)$ que aproxima el espectro de magnitud $|W(\frac{2\pi}{L} \cdot q)|$ de la función ventana o del espectro de magnitud logarítmico $\log |W(\frac{2\pi}{L} \cdot q)|$ para un intervalo dado (q_1, q_2) . La elección de la función de aproximación que aproxima el lóbulo principal del espectro de la ventana se muestra en la Figura 8.

40 3. Para cada pico k (con $k=1\dots K$) con el correspondiente índice m_k de DFT ajustar la función desplazada en frecuencia $P(q - \hat{q}_k)$ mediante los dos puntos de cuadrícula de la DFT que rodean el esperado pico real del espectro continuo de la señal sinusoidal con ventana . Por lo tanto, si $|X(m_k - 1)|$ es mayor que $|X(m_k + 1)|$ ajustar $P(q - \hat{q}_k)$ mediante los puntos.

45 $\{P_1; P_2\} = \{(m_k-1, \log(|X(m_k-1)|)); (m_k, \log(|X(m_k)|))\}$ y en otro caso a través de los puntos $\{P_1; P_2\} = \{(m_k, \log(|X(m_k)|)); (m_k+1, \log(|X(m_k+1)|))\}$. $P(q)$ se puede elegir por simplicidad para que sea un polinomio de orden 2 o 4. Esto hace de la aproximación del paso 2 un simple cálculo de regresión lineal y el cálculo de \hat{q}_k simple. El intervalo (q_1, q_2) se puede escoger para ser fijo e idéntico para todos los picos, por ejemplo $(q_1, q_2) = (-1, 1)$, o adaptativo.

En la aproximación adaptativa el intervalo se puede escoger tal que la función $P(q-\hat{q}_k)$ ajuste el lóbulo principal del espectro de la función ventana en el rango de los puntos de cuadrícula de DFT pertinentes $\{P_1; P_2\}$. El proceso de ajuste se visualiza en la Figura 9.

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 con ventana se espera que tenga su pico calcular $f_k = \hat{q}_k \cdot f_s/L$ como aproximación de la frecuencia de senoide f_k .

Hay muchos casos donde la señal transmitida es armónica lo cual significa que la señal consiste en curvas sinusoidales cuyas frecuencias son múltiplos enteros de alguna frecuencia fundamental f_0 . Este es el caso cuando una señal es muy periódica como por ejemplo la voz o los tonos sostenidos de algún instrumento musical . Esto significa que las frecuencias del modelo sinusoidal de las realizaciones no son independientes sino que más bien tienen una relación armónica y provienen de la misma frecuencia fundamental. Teniendo esta propiedad armónica en cuenta se puede mejorar por consiguiente el análisis de las frecuencias componentes sinusoidales substancialmente.

Una posibilidad de mejora se resume como sigue:

1. Comprobar si la señal es armónica. Esto puede por ejemplo hacerse evaluando la periodicidad de la señal antes de la pérdida de trama. Un método sencillo es realizar un análisis de autocorrelación de la señal. Se puede usar como un indicador el máximo de dicha función de auto correlación para algún retardo de tiempo $\tau > 0$. Si el valor de este máximo sobrepasa un umbral dado, la señal puede ser considerada armónica.

El correspondiente retardo de tiempo τ entonces corresponde al periodo de la señal que se relaciona con la

frecuencia fundamental mediante
$$f_0 = \frac{f_s}{\tau} .$$

Muchos métodos de codificación de voz predictivos lineales aplican la predicción de tono de los denominados bucles abiertos o cerrados o codificación CELP que usa libros de códigos adaptativos. La ganancia de tono y los parámetros asociados de retardo de tono derivados por dichos métodos de codificación son también indicadores útiles de si la señal es armónica y, respectivamente, para el retardo de tiempo.

Un método adicional para obtener f_0 se describe a continuación.

2. Para cada índice armónico j en el rango entero $1 \dots J_{\max}$ comprobar si hay un pico en el espectro de magnitud de la DFT (logarítmica) de la trama de análisis en las proximidades de la frecuencia armónica $f_j = j \cdot f_0$. La proximidad de f_j se puede definir como el rango delta alrededor de f_j donde delta corresponde a la resolución de frecuencia de la

DFT $\frac{f_s}{L}$, esto es, 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 el caso de que tal pico con la correspondiente frecuencia sinusoidal estimada f_k esté presente, sustituir f_k por $f_k = j \cdot f_0$.

Para el procedimiento de dos pasos dado más arriba existe también la posibilidad de comprobar si la señal es armónica y la derivación de la frecuencia fundamental implícita y posiblemente de forma iterativa sin usar necesariamente indicadores de algún método separado . Un ejemplo para dicha técnica se da de la manera siguiente:

Para cada $f_{0,p}$ fuera de un conjunto de valores candidatos $\{f_{0,1} \dots f_{0,p}\}$ aplicar el paso 2 del procedimiento, aunque sin sustituir f_k pero contando cuantos picos de la DFT están presentes en la proximidad alrededor de las frecuencias armónicas, esto es, los múltiplos enteros de $f_{0,p}$. Identificar la frecuencia fundamental $f_{0,pmax}$ para la cual se obtiene el mayor número de picos en o alrededor de las frecuencias armónicas. Si este mayor número de picos excede de un umbral dado, entonces se asume que la señal es armónica. En ese caso se puede asumir que $f_{0,pmax}$ sea la frecuencia fundamental con la que se ejecuta después el paso 2 que lleva a frecuencias sinusoidales mejoradas f_k . Una alternativa más preferible es sin embargo optimizar en primer lugar la frecuencia fundamental f_0 basada en las frecuencias de pico f_k que se haya encontrado que coinciden con frecuencias armónicas. Asumir un conjunto de M armónicos, esto es múltiplos enteros $\{n_1 \dots n_M\}$ de alguna frecuencia fundamental que se haya encontrado que coincide con algún conjunto de M picos espectrales en las frecuencias $f_{k(m)}$, $m = 1 \dots M$, a continuación la frecuencia fundamental subyacente (optimizada) $f_{0,opt}$ se puede calcular para minimizar el error entre las frecuencias de

armónicos y las frecuencias de picos espectrales. Si el error a minimizar es el error cuadrático medio

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

entonces la frecuencia fundamental óptima se calcula como

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

5 El conjunto inicial de valores candidatos $\{f_{0,1} \dots f_{0,P}\}$ se puede obtener a partir de las frecuencias de los picos de la DFT o de las frecuencias sinusoidales estimadas f_k .

Una posibilidad adicional para mejorar la precisión de las frecuencias sinusoidales estimadas f_k es considerar su evolución temporal. Para tal fin, las estimaciones de las frecuencias sinusoidales de múltiples tramas de análisis se pueden combinar por ejemplo por medio de promedio o predicción. Antes de promediar o predecir un pico se puede aplicar seguimiento que conecte los picos espectrales estimados a las mismas sinusoides subyacentes respectivas.

10 Aplicación del modelo sinusoidal

La aplicación de un modelo sinusoidal a fin de realizar una operación de ocultación de pérdida de trama aquí descrita se puede describir como sigue:

15 Se asume que un segmento dado de la señal codificada no puede ser reconstruido por el decodificador ya que la correspondiente información codificada no está disponible. Además se asume que una parte de la señal previa a este segmento está disponible. Supongamos que $y(n)$ con $n=0 \dots N-1$ sea el segmento no disponible para el que una trama de sustitución $z(n)$ tiene que ser generada y que $y(n)$ con $n < 0$ sea la señal decodificada previamente disponible. Entonces, en un primer paso una trama prototipo de la señal disponible de longitud L e índice de inicio n_{-1} se extrae con una función ventana $w(n)$ y se transforma al dominio de frecuencia, por ejemplo mediante la DFT

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

20 La función ventana puede ser una de las funciones ventana descritas más arriba en el análisis sinusoidal. Preferiblemente, a fin de reducir la complejidad numérica, la trama transformada al dominio de la frecuencia debería ser idéntica a la usada durante el análisis sinusoidal.

En un siguiente paso se aplica la asunción de modelo sinusoidal. Según eso la DFT de la trama prototipo se puede escribir como sigue:

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

25 El siguiente paso es darse cuenta que el espectro de la función de ventana usada tiene sólo una contribución significativa en un rango de frecuencia cercano a cero. Como se muestra en la Figura 3 el espectro de magnitud de la función ventana es grande para frecuencias cercanas a cero y pequeño en caso contrario (en el rango de frecuencia normalizado de $-\pi$ a π , correspondiente a la mitad de la frecuencia de muestreo). Por lo tanto, se asume como aproximación que el espectro de ventana $W(m)$ es distinto de cero sólo para un intervalo $M = [-m_{min}, m_{max}]$, siendo m_{min} y m_{max} números positivos pequeños. En particular, una aproximación del espectro de la función ventana se usa de tal modo que para cada k las contribuciones de los espectros de ventana desplazados en la expresión anterior son estrictamente no superpuestas. Por lo tanto en la ecuación anterior para cada índice de frecuencia hay siempre sólo como máximo la contribución de un sumando, esto es, de un espectro de ventana desplazado. Esto
35 significa que la expresión anterior se reduce a la siguiente expresión aproximada:

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

Aquí, M_k denota el intervalo entero $M_k = [\text{redondear}(\frac{f_k}{f_s} \cdot L) - m_{min,k}, \text{redondear}(\frac{f_k}{f_s} \cdot L) + m_{max,k}]$ donde $m_{min,k}$ y $m_{max,k}$ cumplen la restricción anteriormente explicada de tal modo que los intervalos no se solapen. Una elección adecuada para $m_{min,k}$ y $m_{max,k}$ es ajustarlos a un valor entero pequeño δ , por ejemplo $\delta=3$. Si en cambio los

índices DFT relacionados a dos frecuencias sinusoidales adyacentes f_k y f_{k+1} son menores que 2δ , entonces δ se fija

a
$$\text{suelo} \left(\frac{\left(\text{redondear} \left(\frac{f_{k+1}}{f_s} \right) - \text{redondear} \left(\frac{f_k}{f_s} \right) \right)}{2} \right)$$
 de tal modo que se asegure que los intervalos no se están solapando. La función suelo (\cdot) es el entero más cercano al argumento de la función que es más pequeño o igual a éste.

5 El siguiente paso según la realización es aplicar el modelo sinusoidal según la expresión anterior y evolucionar sus K sinusoides en el tiempo. La asunción de que el índice de tiempo del segmento eliminado comparado con el índice de tiempo de la trama prototipo difiere en n_{-1} muestras significa que las fases de las sinusoides se adelantan en

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

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

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

10 Aplicando de nuevo la aproximación según la cual los espectros de ventana desplazados no se solapan, se obtiene

$$\hat{Y}_\delta(m) = \frac{a_k}{2} \cdot W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j(\varphi_k + \theta_k)} \text{ para } m \in M_k \text{ no negativo y para cada } k$$

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

15 en $\theta_k = 2\pi \cdot \frac{f_k}{f_s} n_{-1}$, para cada $m \in M_k$. Por tanto, los coeficientes del espectro de frecuencia de la trama prototipo en la proximidad de cada senoide están desplazados proporcionalmente a la frecuencia sinusoidal f_k y a la diferencia de tiempo entre la trama de audio perdida y la trama prototipo n_{-1} .

Por lo tanto, según la realización la trama de sustitución se puede calcular por la siguiente expresión:

$$z(n) = IDTF\{Z(m)\} \text{ con } Z(m) = Y(m) \cdot e^{j\theta_k} \text{ para } m \in M_k \text{ no negativas y para cada } k$$

20 Una realización específica se ocupa de hacer aleatoria la fase para índices DFT no pertenecientes a ningún intervalo M_k . Como se describió anteriormente, los intervalos M_k $k=1\dots K$ tienen que ser ajustados de tal modo que estén estrictamente no solapados lo que se hace usando algunos parámetros δ que controlan el tamaño de los intervalos. Puede ocurrir que δ sea pequeño en relación a la distancia de frecuencia de dos sinusoides vecinas. Por lo tanto, en ese caso ocurre que hay un espacio entre dos intervalos. Por consiguiente, para los correspondientes índices m de DFT no se define desplazamiento de fase según la anterior expresión $Z(m) = Y(m) \cdot e^{j\theta_k}$. Una elección adecuada según esta realización es aleatorizar la fase para estos índices, haciendo $Z(m) = Y(m) \cdot e^{j2\pi \text{aleat}(\cdot)}$, donde la función aleat(\cdot) devuelve algún número aleatorio.

30 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, esto es 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 usar pequeños intervalos lleva a una mejor calidad. Este descubrimiento lleva a una mejora adicional según la cual el tamaño del intervalo se adapta según las propiedades de la señal. Una realización es usar un detector de tonalidad o periodicidad. Si este detector identifica la señal como tonal, el parámetro δ que controla el tamaño de intervalo se ajusta a un valor relativamente grande. En caso contrario, el parámetro δ se ajusta a un valor relativamente menor.

Basado en lo anterior, los métodos de ocultación de pérdida de trama implican los siguientes pasos:

1. Analizar un segmento de la señal disponible, previamente sintetizada para obtener las frecuencias sinusoidales f_k constituyentes de un modelo sinusoidal, utilizando opcionalmente una estimación de frecuencia mejorada.
2. Extraer una trama prototipo y_{-1} de la señal disponible previamente sintetizada y calcular la DFT de esa trama.
- 30 3. Calcular el desfase θ_k para cada senoide k en respuesta a la frecuencia sinusoidal f_k y el avance de tiempo n_{-1} entre la trama prototipo y la trama de sustitución. Opcionalmente en este paso el tamaño del intervalo M puede haber sido adaptado en respuesta a la tonalidad de la señal de audio.

4. Para cada senoide k avanzar la fase de la DFT de la trama prototipo en θ_k selectivamente para los índices DFT relacionados a un entorno alrededor de la frecuencia de senoide f_k .

5. Calcular la DFT inversa del espectro obtenido en el paso 4.

Análisis y detección de la propiedad de pérdida de trama y señal

5 Los métodos descritos anteriormente están basados en la asunción de que las propiedades de la señal de audio no cambian significativamente durante el corto periodo de tiempo entre la trama de señal previamente recibida y reconstruida y una trama perdida. En ese caso es una muy buena elección conservar el espectro de magnitud de la trama previamente reconstruida y hacer evolucionar las fases de los componentes principales sinusoidales detectados en la señal previamente reconstruida. Hay sin embargo casos donde esta asunción es errónea como por ejemplo los transitorios con cambios de energía repentinos o cambios espectrales repentinos.

10 Una primera realización de un detector de transitorios según la invención puede por consiguiente estar basada en variaciones de energía dentro de la señal previamente reconstruida. Este método, mostrado en la Figura 11, calcula la energía en la parte izquierda y parte derecha de alguna trama 113 de análisis. La trama de análisis puede ser idéntica a la trama usada para el análisis sinusoidal descrito anteriormente. Una parte (ya sea izquierda o derecha) de la trama de análisis puede ser respectivamente la primera mitad o la última de la trama de análisis o por ejemplo el primer o respectivamente el último cuarto de la trama de análisis, 110. El respectivo cálculo de energía se hace sumando los cuadrados de las muestras en esas tramas parciales:

$$E_{izda} = \sum_{n=0}^{N_{parc}-1} y^2(n - n_{izda}), \text{ y } E_{dcha} = \sum_{n=0}^{N_{parc}-1} y^2(n - n_{dcha})$$

20 Aquí $Y(n)$ denota la trama de análisis, n_{izda} y n_{dcha} denotan los respectivos índices de comienzo de las tramas parciales que son ambas de tamaño N_{parc} .

Ahora las energías de las tramas parciales izquierda y derecha se usan para la detección de una discontinuidad de señal. Esto se hace calculando la relación

$$R_{i/d} = \frac{E_{izda}}{E_{dcha}}$$

25 Una discontinuidad con un cese repentino de energía (cese) se puede detectar si la tasa $R_{i/d}$ excede algún umbral (por ejemplo 10), 115. De manera similar una discontinuidad con un incremento repentino de energía (inicio) se puede detectar si la tasa $R_{i/d}$ está por debajo de algún otro umbral (por ejemplo 0,1), 117.

30 En el contexto de los métodos de ocultación anteriormente descritos se ha encontrado que la relación de energía anteriormente definida puede en muchos casos ser un indicador demasiado insensible. En particular en señales reales y especialmente en música hay casos donde de repente surge un tono a alguna frecuencia mientras algún otro tono a alguna otra frecuencia se detiene de repente. Analizando una trama de señal con la relación de energía anteriormente definida llevaría en cualquier caso a una detección errónea resultado de al menos uno de los tonos ya que el indicador es insensible a diferentes frecuencias.

35 Una solución a este problema se describe en la siguiente realización. La detección de transitorios se hace ahora en el plano de tiempo de la frecuencia. La trama de análisis se divide de nuevo en una trama parcial izquierda y derecha, 110. Aunque ahora, estas dos tramas parciales son (después de una adecuada aplicación de ventana por ejemplo con una ventana de Hamming, 111) transformadas al dominio de la frecuencia, por ejemplo mediante una DFT 112 de N_{part} puntos.

$$Y_{izda}(m) = DFT\{y(n - n_{izda})\}_{N_{parc}}$$

y

$$40 \quad Y_{dcha}(m) = DFT\{y(n - n_{dcha})\}_{N_{parc}}, \text{ con } m = 0 \dots N_{parc} - 1.$$

Ahora la detección de transitorios se puede hacer selectivamente en frecuencia para cada contenedor DFT con índice m . Usando las potencias de los espectros de magnitud trama parcial izquierda y derecha, para cada índice m de la DFT se puede calcular una relación de energía respectiva 113 como

$$R_{i/r}(m) = \frac{|Y_{izda}(m)|^2}{|Y_{dcha}(m)|^2}$$

Los experimentos muestran que la detección de transitorios selectivos en frecuencia con resolución de contenedores DFT es relativamente imprecisa debido a fluctuaciones estadísticas (errores de estimación). Se ha encontrado que la calidad de la operación se mejora bastante cuando se hace la detección de transitorios selectivos en frecuencia en forma de bandas de frecuencia. Supongamos que $l_k = [m_{k-1} + 1, \dots, m_k]$ especifica el intervalo k^{th} , $k = 1 \dots K$, que abarca los contenedores DFT desde $m_{k-1} + 1$ hasta m_k , entonces estos intervalos definen K bandas de frecuencia. La detección de transitorios selectiva en grupos de frecuencia puede ahora estar basada en la relación a modo de banda entre las respectivas energías de banda de las tramas parciales izquierda y derecha.

$$R_{\text{left,band}}(k) = \frac{\sum_{m \in l_k} |V_{\text{izda}}(m)|^2}{\sum_{m \in l_k} |V_{\text{dcha}}(m)|^2}$$

Es preciso señalar que el intervalo $l_k = [m_{k-1} + 1, \dots, m_k]$ corresponde a la banda de frecuencia

$$B_k = \left[\frac{m_{k-1} + 1}{N_{\text{parc}}} \cdot f_s, \dots, \frac{m_k}{N_{\text{parc}}} \cdot f_s \right],$$

donde f_s denota la frecuencia de muestreo de audio.

El límite de banda de frecuencia inferior más bajo m_0 puede ser fijado a 0 pero puede también ser fijado a un índice DFT correspondiente a una frecuencia mayor a fin de mitigar los errores de estimación que aumentan con

frecuencias menores. El límite de banda de frecuencia superior más alto se puede fijar a $\frac{N_{\text{parc}}}{2}$ pero se elige preferiblemente para corresponder a alguna frecuencia más baja en la que un transitorio aún tenga un efecto audible importante.

Una elección adecuada para estos tamaños o anchuras de bandas de frecuencia es hacerlos de igual tamaño con por ejemplo una anchura de varios 100 Hz. Otra opción preferible es hacer que las anchuras de las bandas de frecuencia sigan el tamaño de las bandas auditivas humanas críticas, esto es, relacionarlas a la resolución de frecuencia del sistema auditivo. Esto significa hacer aproximadamente las anchuras de banda iguales para frecuencias de hasta 1 kHz y aumentarlas exponencialmente por encima de 1 kHz. El aumento exponencial significa por ejemplo doblar el ancho de banda de frecuencia cuando se incrementa el índice de banda k .

Como se describió en la primera realización del detector de transitorios que se basó en una relación de energía de dos tramas parciales, cualquiera de las relaciones relacionadas a energías de banda o energías de contenedores DFT de dos tramas parciales son comparadas con ciertos umbrales. Se han usado un umbral superior respectivo 115 para la detección del cese (selectivo en frecuencia) y un umbral inferior respectivo 117 para la detección del inicio (selectivo en frecuencia).

Un indicador dependiente adicional de la señal de audio que es adecuado para una adaptación del método de ocultación de pérdida de trama puede estar basado en los parámetros de códec transmitidos al decodificador. Por ejemplo, el códec puede ser un códec multi-modo como el ITU-T G.718. Tal códec puede usar modos de códec concretos para diferentes tipos de señal y un cambio del modo de códec en una trama poco tiempo antes de que la pérdida de trama pueda ser considerada como un indicador de un transitorio.

Otro indicador útil para la adaptación de la ocultación de la pérdida de trama es un parámetro de códec relacionado con la propiedad de sonorización y la señal transmitida. La sonorización se relaciona con voz muy periódica que se genera por una excitación periódica glotal del tracto vocal humano.

Un indicador adicional es si el contenido de la señal se estima que es música o voz. Tal indicador se puede obtener a partir de un clasificador de señal que puede ser normalmente parte del códec. En caso de que el códec realice tal clasificación y tome una decisión de clasificación correspondiente disponible como un parámetro de codificación para el decodificador, este parámetro es preferiblemente usado como un indicador de contenido de señal para ser usado para adaptar el método de ocultación de pérdida de trama.

Otro indicador que es preferiblemente usado para la adaptación del método de ocultación de pérdida de trama es la explosividad de las pérdidas de trama. La explosividad de las pérdidas de trama significa que ocurran varias pérdidas de trama en una fila, haciendo difícil para el método de ocultación de pérdida de trama usar partes de señal recientemente decodificadas válidas para su operación. Un indicador del estado de la técnica es el número $n_{\text{ráfaga}}$ de pérdidas de trama observadas en una fila. Este contador se incrementa una vez con cada pérdida de trama y se reinicia a cero una vez se recibe una trama válida. Este indicador se usa también en el contexto de las presentes realizaciones ejemplares de la invención.

Adaptación del método de ocultación de pérdida de trama

En el caso de que los pasos llevados a cabo 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.

Mientras el cálculo original del espectro de la trama de sustitución se hace según la expresión $Z(m) = Y(m) \cdot e^{j\theta_k}$ ahora se introduce 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 se modifica la fase con un componente de fase adicional $\vartheta(m)$. Esto lleva al siguiente cálculo modificado de la trama de sustitución:

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

Cabe señalar que los métodos de ocultación de pérdida de trama originales (no adaptados) se usan si $\alpha(m) = 1$, $\beta(m) = 1$, and $\vartheta(m) = 0$. Estos respectivos valores son por lo tanto los de por defecto.

El objetivo general al introducir adaptaciones de magnitud es evitar defectos audibles del método de ocultación de pérdida de trama. Tales defectos pueden ser sonidos musicales o tonales o extraños sonidos que se generan de las repeticiones de sonidos transitorios. Tales defectos llevarían a su vez a degradaciones de calidad, cuya prevención es el objetivo de las adaptaciones descritas. Una forma adecuada para tales adaptaciones es modificar el espectro de magnitud de la trama de sustitución a un grado adecuado.

La Figura 12 muestra una modificación de la realización del método de ocultación. La adaptación de magnitud, 123, se hace preferiblemente si el contador $n_{ráfaga}$ de pérdida de ráfaga excede algún umbral $thr_{ráfaga}$, 121, por ejemplo $thr_{ráfaga} = 3$. En este caso un valor menor que 1 se usa para el factor de atenuación, por ejemplo $\alpha(m) = 0.1$.

No obstante se ha encontrado que es beneficioso realizar la atenuación con un grado que aumente gradualmente. Una realización preferida que cumple esto es definir un parámetro logarítmico que especifique un aumento logarítmico en la atenuación por trama, *att_per_frame* ("at_por_trama"). Entonces, en caso de que el contador de ráfaga exceda el umbral el factor de atenuación que aumenta gradualmente se calcula como:

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

Aquí la constante c es una mera constante de escalado que permite especificar el parámetro *att_per_frame* por ejemplo en decibelios (dB).

Una adaptación adicional preferida se hace en respuesta al indicador si se estima que la señal es música o voz. Para contenido de música en comparación con contenido de voz es preferible aumentar el umbral $thr_{ráfaga}$ y disminuir la atenuación por trama. Esto es equivalente a realizar la adaptación del método de ocultación de pérdida de trama con un menor grado. El trasfondo de este tipo de adaptación es que la música es generalmente menos sensible a mayores ráfagas de pérdidas que la voz. Por lo tanto, el original, esto es el método de ocultación de pérdida de trama no modificado es preferible aún para este caso, al menos para un mayor número de pérdidas de trama en una fila.

Una adaptación adicional del método de ocultación en lo referente al factor de atenuación de magnitud se hace preferiblemente en caso de que se haya detectado un transitorio basado en que el indicador $R_{i/d, banda}(k)$ o de manera alternativa $R_{i/d}(m)$ o $R_{i/d}$ han superado un umbral, 122. En ese caso una acción de adaptación adecuada, 125, es modificar el segundo factor de atenuación de magnitud $\beta(m)$ de tal modo que la atenuación total sea controlada por el producto de los dos factores $\alpha(m)$.

$\beta(m)$ se fija en respuesta a un transitorio indicado. En caso de que se detecte un cese el factor $\beta(m)$ se elige preferiblemente para reflejar la disminución de energía del cese. Una elección adecuada es fijar $\beta(m)$ al cambio de ganancia detectado:

$$\beta(m) = \sqrt{R_{i/d, banda}(k)}, \text{ para } m \in I_k, k = 1 \dots K.$$

En caso de que se detecte un inicio se ha encontrado ventajoso en su lugar limitar el incremento de energía en la trama de sustitución. En ese caso el factor se puede fijar a algún valor fijo de por ejemplo 1, que significa que no hay atenuación pero tampoco ninguna amplificación.

Cabe señalar de lo anterior que el factor de atenuación de magnitud se aplica preferiblemente de manera selectiva en frecuencia, esto es con factores calculados individualmente para cada banda de frecuencia. En caso de que el enfoque de banda no se use, los factores de atenuación de magnitud correspondientes se pueden obtener aún de manera analógica. $\beta(m)$ puede entonces ser fijado individualmente para cada contenedor DFT en caso de que se use detección de transitorio selectiva en frecuencia sobre el nivel de contenedor DFT. O, en caso de que no se use indicación de transitorio selectiva en frecuencia en absoluto $\beta(m)$ puede ser globalmente idéntico para todo m .

Una adaptación preferida adicional del factor de atenuación de magnitud se hace en combinación con una modificación de la fase mediante el componente 127 de fase adicional $\vartheta(m)$. En el caso de que se use tal modificación de fase para una m dada, el factor de atenuación $\beta(m)$ se reduce incluso más. Preferiblemente, incluso el grado de modificación de fase se tiene en cuenta. Si la modificación de fase es sólo moderada, $\beta(m)$ sólo se

escala ligeramente hacia abajo, mientras que si la modificación de fase es fuerte, se escala hacia abajo en un grado mayor.

5 El objetivo general al introducir la adaptación de fase es evitar tonalidades demasiado fuertes o periodicidad de señal en las tramas de sustitución generadas, lo que a su vez llevaría a degradaciones de calidad. Una forma adecuada para tales adaptaciones es aleatorizar o hacer oscilar la fase a un grado adecuado

Tal oscilación de fase se logra si el componente de fase adicional $\vartheta(m)$ se fija a un valor aleatorio escalado con algún factor de control; $\vartheta(m) = a(m) \cdot \text{aleat}(\cdot)$.

El valor aleatorio obtenido por la función $\text{aleat}(\cdot)$ es por ejemplo generado por algún generador de números pseudoaleatorios. Se asume aquí que esto proporciona un número aleatorio en el intervalo $[0, 2\pi]$.

10 El factor de escalado $a(m)$ en la ecuación anterior controla el grado por el que la fase original θ_k es hecha oscilar. Las siguientes realizaciones abordan la adaptación de fase mediante el control de este factor de escalado. El control del factor de escalado se hace de forma análoga con el control de los factores de modificación de magnitud descritos anteriormente.

15 Según una primera realización el factor de escalado $a(m)$ se adapta en respuesta al contador de pérdida de ráfaga. Si el contador de pérdida de ráfaga $n_{\text{ráfaga}}$ excede algún umbral $\text{thr}_{\text{ráfaga}}$, por ejemplo $\text{thr}_{\text{ráfaga}} = 3$, se usa un valor mayor que 0, por ejemplo $a(m) = 0,2$.

20 No obstante se ha encontrado que es beneficioso realizar la oscilación con un aumento gradual de grado. Una realización preferida que logra esto es definir un parámetro que especifique un aumento en la oscilación por trama, *dith_increase_per_frame* ("incr_oscilac_por_trama"). Entonces en caso de que el contador de ráfaga exceda el umbral el factor de control de oscilación que aumenta gradualmente se calcula por:

$$a(m) = \text{incr_oscilac_por_trama} \left(n_{\text{ráfaga}} - \text{thr}_{\text{ráfaga}} \right).$$

Cabe destacar en la fórmula anterior que $a(m)$ tiene que limitarse a un valor máximo de 1 para el que se alcanza la oscilación completa de fase.

25 Cabe destacar que el valor de umbral de pérdida de ráfaga $\text{thr}_{\text{ráfaga}}$ usado para iniciar la oscilación de la fase puede ser el mismo umbral que el usado para la atenuación de magnitud. Sin embargo, se puede obtener mejor calidad fijando estos umbrales a valores individualmente óptimos, lo que generalmente significa que estos umbrales pueden ser diferentes.

30 Una adaptación preferida adicional se hace en respuesta al indicador si la señal es estimada como música o voz. Para contenido de música en comparación con contenido de voz es preferible aumentar el umbral $\text{thr}_{\text{ráfaga}}$ que significa que la oscilación de fase para música en comparación con la voz se hace sólo en caso de más tramas perdidas en una fila. Esto es equivalente a realizar la adaptación del método de ocultación de pérdida de trama para música con un grado inferior. El trasfondo de este tipo de adaptación es que la música es generalmente menos sensible a ráfagas de pérdidas más largas que la voz. Por lo tanto, el método original, esto es, el método de ocultación de pérdida de trama no modificado es aún preferible para este caso, al menos para un mayor número de
35 pérdidas de trama en una fila.

Una realización preferida adicional es adaptar la oscilación de fase en respuesta a un transitorio detectado. En ese caso un grado más fuerte de oscilación de fase se puede usar para los m contenedores DFT para los que se indica un transitorio o bien para ese contenedor, los contenedores DFT de la correspondiente banda de frecuencia o bien de la trama entera

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

45 En caso de que los métodos que usan una estimación de frecuencia mejorada como se ha descrito anteriormente no se den cuenta otra posibilidad de adaptación para el método de ocultación de pérdida de trama que optimiza la calidad para señales de voz sonorizadas es cambiar a algún otro método de ocultación de pérdida de trama que específicamente esté diseñado y optimizado para voz en lugar de para señales de audio generales que contengan música y voz. En ese caso, el indicador de que la señal comprende una señal de voz sonorizada se usa para seleccionar otro esquema de ocultación de pérdida de trama optimizado para voz en lugar de los esquemas descritos anteriormente.

50 Las realizaciones se aplican a un controlador de un decodificador, como se muestra en la Figura 13. La Figura 13 es un diagrama de bloques esquemático de un decodificador según las realizaciones. El decodificador 130 comprende una unidad de entrada 132 configurada para recibir una señal de audio codificada. La figura muestra la ocultación de pérdida de trama por una unidad 134 de ocultación de pérdida de trama lógica, lo que indica que el decodificador se configura para implementar una ocultación de una trama de audio perdida, según las realizaciones descritas anteriormente. Además el decodificador comprende un controlador 136 para implementar las realizaciones descritas

anteriormente. El controlador 136 se configura para detectar las condiciones en las propiedades de la señal de audio previamente recibida y reconstruida o en las propiedades estadísticas de las pérdidas observadas de trama para las que la sustitución de una trama perdida según los métodos descritos proporciona una calidad relativamente reducida. En caso de que se detecte tal condición, el controlador 136 se configura para modificar el elemento de los métodos de ocultación según el cual el espectro de la trama de sustitución se calcula por $Z(m) = Y(m) \cdot e^{j\theta_k}$ ajustando selectivamente las fases o las magnitudes de espectro. La detección se puede realizar por una unidad detectora 146 y la modificación se puede realizar por una unidad modificadora 148 como se muestra en la Figura 14.

El decodificador con sus unidades incluidas se podría implementar en hardware. Hay numerosas variantes de elementos de circuito que se pueden usar y combinar para alcanzar las funciones de las unidades del decodificador. Tales variantes son abarcadas por las realizaciones. Ejemplos particulares de implementación en hardware del decodificador es la implementación en hardware y tecnología de circuito integrado del procesador digital de señal (DSP), que incluye tanto circuitos electrónicos de propósito general como circuitos de aplicación específica.

El decodificador 150 aquí descrito podría ser implementado alternativamente por ejemplo como se muestra en la Figura 15, esto es por uno o más de un procesador 154 y software adecuado 155 con almacenamiento o memoria adecuado 156 por tanto, a fin de reconstruir la señal digital, lo que incluye realizar ocultación de pérdida de trama según las realizaciones aquí descritas, como se muestra en la Figura 13. La señal de audio codificado entrante se recibe por una entrada (ENTRADA) 152, a la que el procesador 154 y la memoria 156 están conectados. La señal de audio decodificada y reconstruida obtenida del software se saca por la salida (SALIDA) 158.

La tecnología descrita anteriormente se puede usar por ejemplo en un receptor, el cual se puede usar en un dispositivo móvil (por ejemplo teléfono móvil, portátil) o en un dispositivo fijo tal como un ordenador personal.

Debe entenderse que la elección de unidades o módulos que interactúan, así como los nombres de las unidades son sólo con propósito ejemplar, y puede ser configurado en una pluralidad de formas alternativas a fin de poder ejecutar las acciones de proceso expuestas.

Debería también destacarse que las unidades o módulos descritos en esta descripción han de considerarse como entidades lógicas y no necesariamente como entidades físicas separadas. Se apreciará que el alcance de la tecnología aquí descrita abarca completamente otras realizaciones que puedan resultar obvias para aquellos expertos en la técnica, y que el alcance de esta descripción por consiguiente no se ha de limitar.

La referencia a un elemento en singular no está dirigido a significar "uno y sólo uno" a menos que explícitamente así sea declarado, sino más bien "uno o más". Todas las equivalencias estructurales y funcionales a los elementos de las realizaciones anteriormente descritas que son conocidas por los expertos corrientes en la técnica se incorporan expresamente aquí por referencia y están destinados a ser abarcados por la presente. Por otra parte no es necesario que un dispositivo o método abarque todos y cada uno de los problemas tratados de resolver por la tecnología aquí descrita, para que puedan ser abarcados por la presente.

En la descripción anterior, con propósito de explicación y no de limitación se han descrito detalles específicos tales como arquitecturas particulares, interfaces, técnicas, etc., a fin de proporcionar una comprensión completa de la tecnología descrita. No obstante, será evidente para los expertos en la técnica que la tecnología descrita se puede poner en práctica en otras realizaciones y o combinaciones de realizaciones que partan de estos detalles específicos. Es decir, los expertos en la técnica serán capaces de diseñar varias disposiciones que, aunque no explícitamente descritas o mostradas aquí, encarnan los principios de la tecnología descrita. En algunas ocasiones, descripciones detalladas de dispositivos, circuitos, y métodos bien conocidos se omiten a fin de no oscurecer la descripción de la tecnología descrita con detalles innecesarios. Todas las afirmaciones que aquí recitan principios, aspectos, y realizaciones de la tecnología descrita, así como sus ejemplos específicos, se destinan a abarcar tanto equivalentes estructurales de los mismos. Además, se pretende que tales equivalentes incluyan tanto equivalentes actualmente conocidos así como equivalentes desarrollados en el futuro, por ejemplo, cualesquiera elementos desarrollados que realicen la misma función, independientemente de la estructura.

Así, por ejemplo, se apreciará por los expertos en la técnica que las figuras de este documento pueden representar vistas conceptuales de circuitos ilustrativos u otras unidades funcionales que realizan los principios de la tecnología, y o varios procesos que se pueden representar sustancialmente en un medio legible por ordenador y ser ejecutados por un ordenador o procesador, incluso aunque tal ordenador o procesador pueda no haberse mostrado explícitamente en las figuras.

Las funciones de los diversos elementos, que incluyen bloques funcionales se pueden proporcionar mediante el uso de hardware tal como hardware de circuito y/o hardware capaz de ejecutar software en forma de instrucciones codificadas almacenadas en un medio legible por ordenador. Así, tales funciones y bloques funcionales mostrados se han de entender como que son o bien implementados por hardware y/o implementados por ordenador, y por tanto implementados por máquina.

Las realizaciones descritas anteriormente han de ser entendidas como unos pocos ejemplos ilustrativos de la presente invención. Se entenderá por aquellos expertos en la técnica que se pueden hacer varias modificaciones, combinaciones y cambios a las realizaciones sin salir del alcance de la presente invención. En particular, soluciones

parciales diferentes en las diferentes realizaciones se pueden combinar en otras configuraciones, donde técnicamente sean posibles.

REIVINDICACIONES

1. Un método para controlar un método de ocultación para una trama de audio perdida de una señal de audio recibida, comprendiendo el método :
- 5 - detectar (101, 122) en una propiedad de una señal de audio previamente recibida y reconstruida una condición transitoria que podría llevar a una calidad de reconstrucción subóptima, cuando un método de ocultación original se usa para crear una trama de sustitución, y
- modificar (102, 125) el método de ocultación original mediante ajuste selectivo de una magnitud del espectro de un espectro de una trama de sustitución, cuando se detecta la condición transitoria .
- 10 - además detectar (101, 121) en una propiedad estadística de pérdidas de trama observadas una segunda condición que podría llevar a una calidad de reconstrucción subóptima, cuando el método de ocultación original se usa para crear la trama de sustitución, y
- además modificar (102, 123, 127) el método de ocultación original mediante el ajuste selectivo de la magnitud del espectro del espectro de la trama de sustitución, cuando se detecta la segunda condición.
2. El método según la reivindicación 1, en donde el método de ocultación original comprende:
- 15 - extraer un segmento de una señal de audio previamente recibida o reconstruida,
- en donde dicho segmento se usa como una trama prototipo.
- aplicar un modelo sinusoidal a la trama prototipo para obtener frecuencias sinusoidales del modelo sinusoidal; y
- hacer evolucionar en el tiempo las sinusoides obtenidas para crear la trama de sustitución.
- 20 3. El método según a la reivindicación 2, en donde la evolución en el tiempo comprende avanzar la fase de los coeficientes espectrales relacionados a las sinusoides obtenidas (k) en θ_k y en donde el cálculo del espectro de la trama de sustitución se realiza según la expresión $Z(m) = Y(m) \cdot e^{j\theta_k}$, en la que Y(m) es una representación en el dominio de la frecuencia de la trama prototipo.
4. El método según cualquiera de las reivindicaciones 1 a 3, en donde la condición transitoria comprende un cese detectado.
- 25 5. El método según cualquiera de las reivindicaciones 1 a 4, en donde la detección de transitorio se realiza selectivamente en frecuencia en forma de banda de frecuencia.
6. El método según la reivindicación 4 o 5, en donde el ajuste selectivo de la magnitud del espectro de la trama de sustitución se realiza selectivamente en banda de frecuencia en respuesta a un transitorio detectado en la banda de frecuencia.
- 30 7. El método según cualquiera de las reivindicaciones precedentes, en donde la segunda condición es una ocurrencia de varias pérdidas de trama consecutivas.
8. El método según a la reivindicación 7, en donde la magnitud del espectro se ajusta en respuesta a varias pérdidas de trama consecutivas detectadas mediante la realización de una atenuación con un grado gradualmente creciente .
- 35 9. El método según cualquiera de las reivindicaciones precedentes, en donde el método de ocultación original se modifica además mediante el ajuste selectivo de una fase del espectro de la trama de sustitución, cuando se detecta la segunda condición.
10. El método según la reivindicación 9, en donde ajustar la fase del espectro de la trama de sustitución comprende aleatorizar o hacer oscilar el espectro de fase.
- 40 11. El método según la reivindicación 10, en donde el espectro de fase se ajusta mediante la realización de la oscilación con un grado gradualmente creciente.
12. Un aparato que comprende:
- un procesador (154), y
- una memoria (156) que almacena instrucciones (155) que, cuando se ejecutan por el procesador, llevan al aparato a:
- 45 - detectar en una propiedad de una señal de audio previamente recibida y reconstruida una condición transitoria que podría llevar a una calidad de reconstrucción subóptima cuando se usa un método de ocultación original para crear una trama de sustitución.

- modificar el método de ocultación original, cuando se detecta la condición transitoria , mediante el ajuste selectivo de una magnitud del espectro de un espectro de una trama de sustitución;
 - además detectar en una propiedad estadística de pérdidas de trama observadas una segunda condición que podría llevar a una calidad de reconstrucción subóptima cuando el método de ocultación original se usa para crear la trama de sustitución; y
- 5
- además modificar el método de ocultación original, cuando se detecta la segunda condición, mediante el ajuste selectivo de la magnitud del espectro del espectro de la trama de sustitución.
13. El aparato según la reivindicación 12, en donde cuando se crea la trama de sustitución utilizando el método de ocultación original el aparato es llevado a:
- 10
- extraer un segmento de una señal de audio previamente recibida o reconstruida.
en donde dicho segmento se usa como una trama prototipo.
 - aplicar un modelo sinusoidal a la trama prototipo para obtener frecuencias sinusoidales del modelo sinusoidal; y
 - hacer evolucionar en el tiempo sinusoides obtenidas para crear la trama de sustitución.
14. El aparato según la reivindicación 13, en donde la evolución en el tiempo se realiza mediante el avance de la fase de los coeficientes espectrales relacionados con las sinusoides obtenidas (k) por θ_k y en donde el cálculo del espectro de la trama de sustitución se realiza según la expresión $Z(m) = Y(m) \cdot e^{j\theta_k}$ en donde Y(m) es una representación en el dominio de frecuencia de la trama prototipo.
- 15
15. El aparato según la reivindicación 12-14 que comprende además un detector de transitorio, en donde el detector de transitorio se configura para realizar una detección de transitorio selectivamente en frecuencia en forma de bandas de frecuencia.
- 20
16. El aparato según la reivindicación 15, en donde el ajuste selectivo de la magnitud del espectro de la trama de sustitución se realiza selectivamente en banda de frecuencia en respuesta a un transitorio detectado en la banda de frecuencia.
17. El aparato según cualquiera de las reivindicaciones 12-16, en donde la segunda condición es una ocurrencia de varias pérdidas de trama consecutivas.
- 25
18. El aparato según la reivindicación 17, en donde se ajusta una magnitud del espectro en respuesta a varias pérdidas de trama consecutivas detectadas realizando una atenuación con grado gradualmente creciente.
19. El aparato conforme a cualquiera de las reivindicaciones 12-18, en donde el aparato se configura para además modificar el método de ocultación original, cuando la segunda condición se detecta, mediante el ajuste selectivo de una fase del espectro de la trama de sustitución.
- 30
20. El aparato según la reivindicación 12, en donde el aparato es un decodificador en un dispositivo móvil.
21. Un programa de ordenador (155) que comprende unidades de código legibles por un ordenador que cuando se ejecutan en un aparato llevan al aparato a:
- detectar (101) en una propiedad de una señal de audio previamente recibida y reconstruida una condición transitoria que podría llevar a una calidad de reconstrucción subóptima cuando un método de ocultación original se usa para crear una trama de sustitución; y
 - modificar (102) el método de ocultación original, cuando la condición transitoria se detecta, mediante el ajuste selectivo de una magnitud del espectro de un espectro de sustitución de trama.
- 35
- detectar además en una propiedad estadística de las pérdidas de trama observadas una segunda condición que podría llevar a una calidad de reconstrucción subóptima cuando se usa el método de ocultación original para crear la trama de sustitución; y
 - modificar además el método de ocultación original, cuando se detecta la segunda condición, mediante el ajuste selectivo de la magnitud del espectro del espectro de sustitución de trama.
- 40
22. Un programa de ordenador (156), que comprende un medio legible por un ordenador y un programa de ordenador (155) según la reivindicación 22 almacenado en el medio legible por el ordenador.
- 45

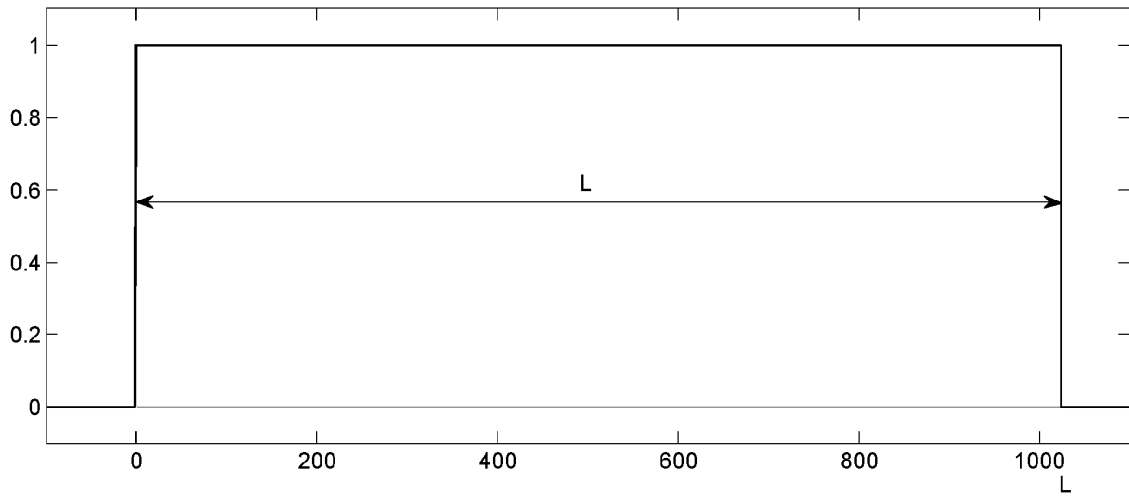


FIG. 1

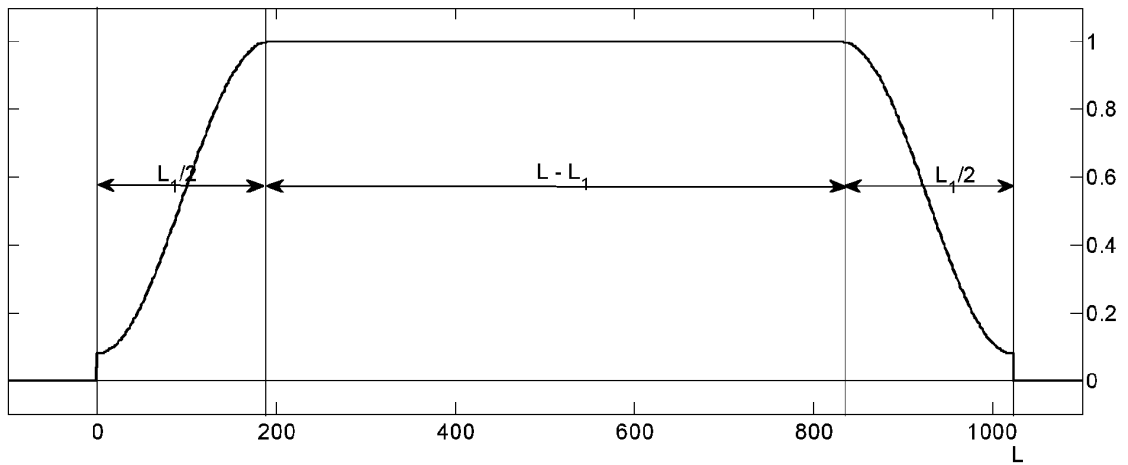


FIG. 2

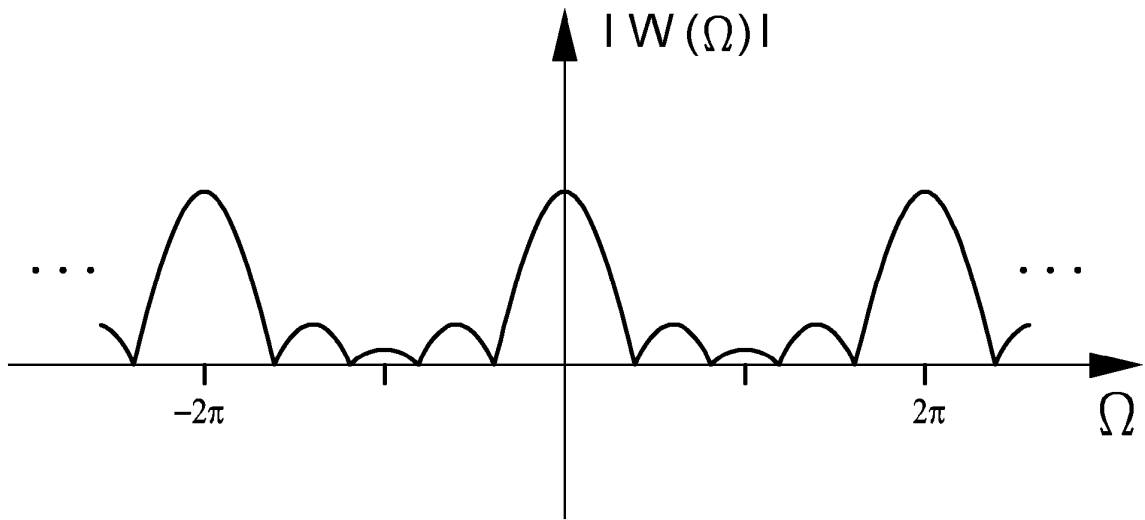


FIG. 3

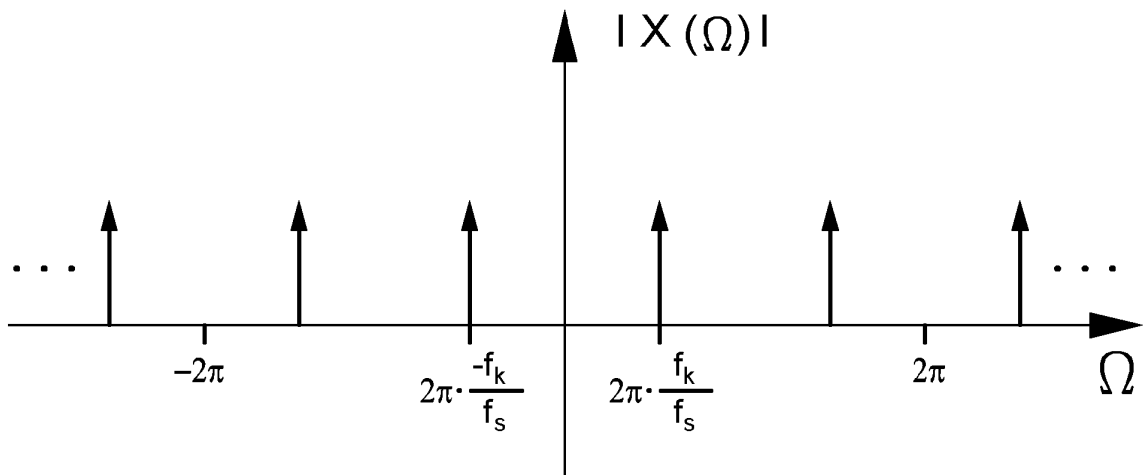


FIG. 4

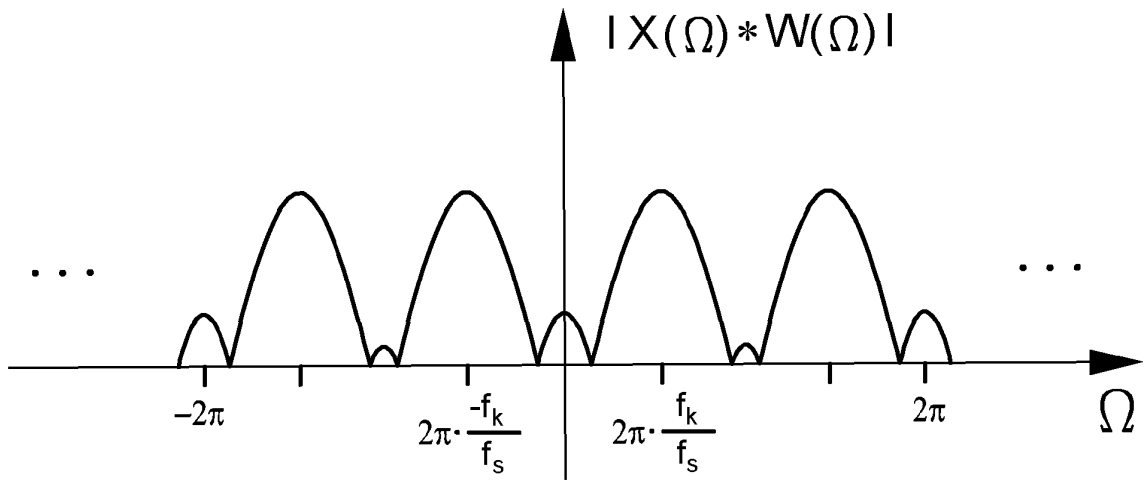


FIG. 5

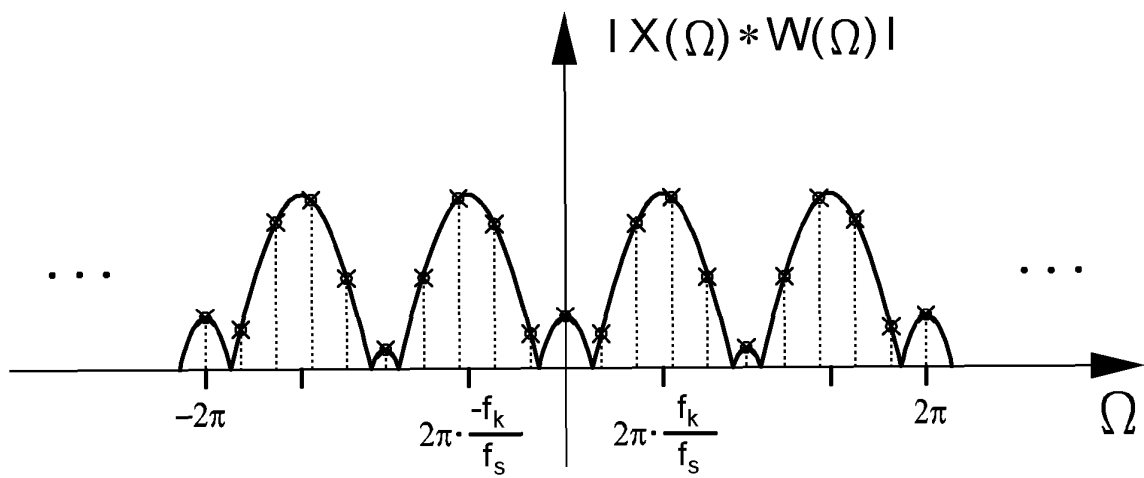
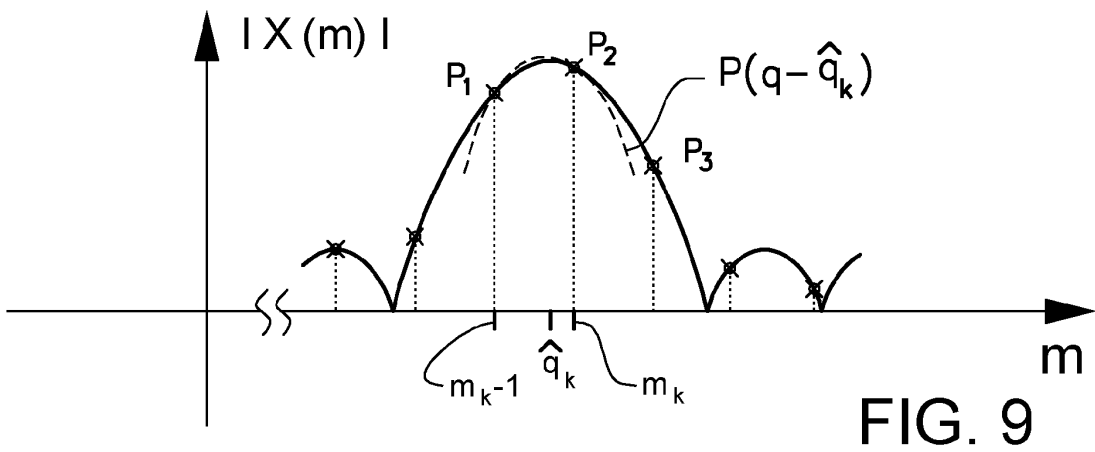
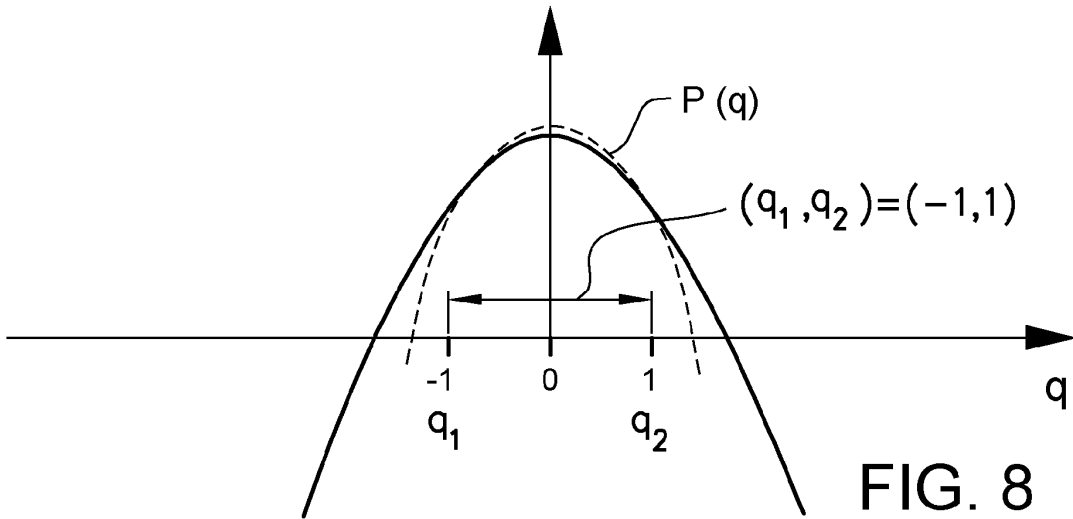
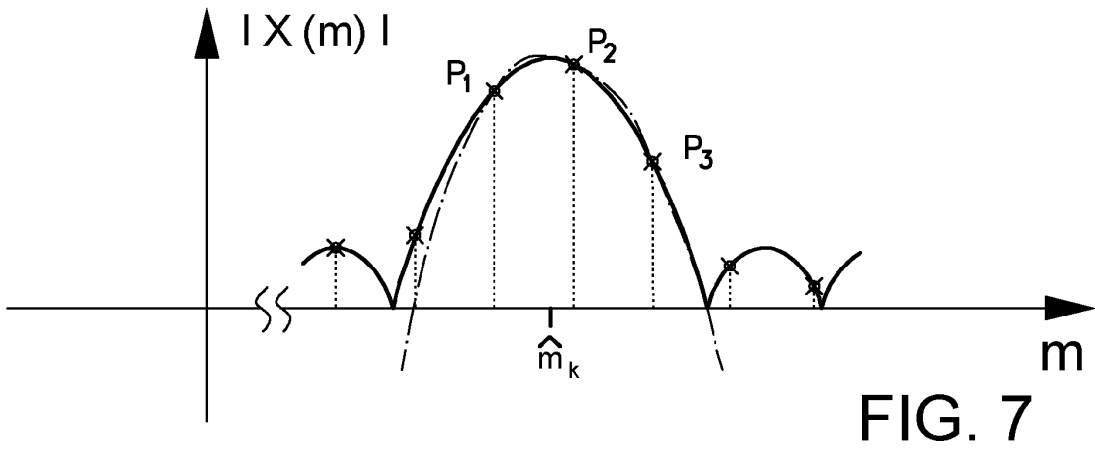


FIG. 6



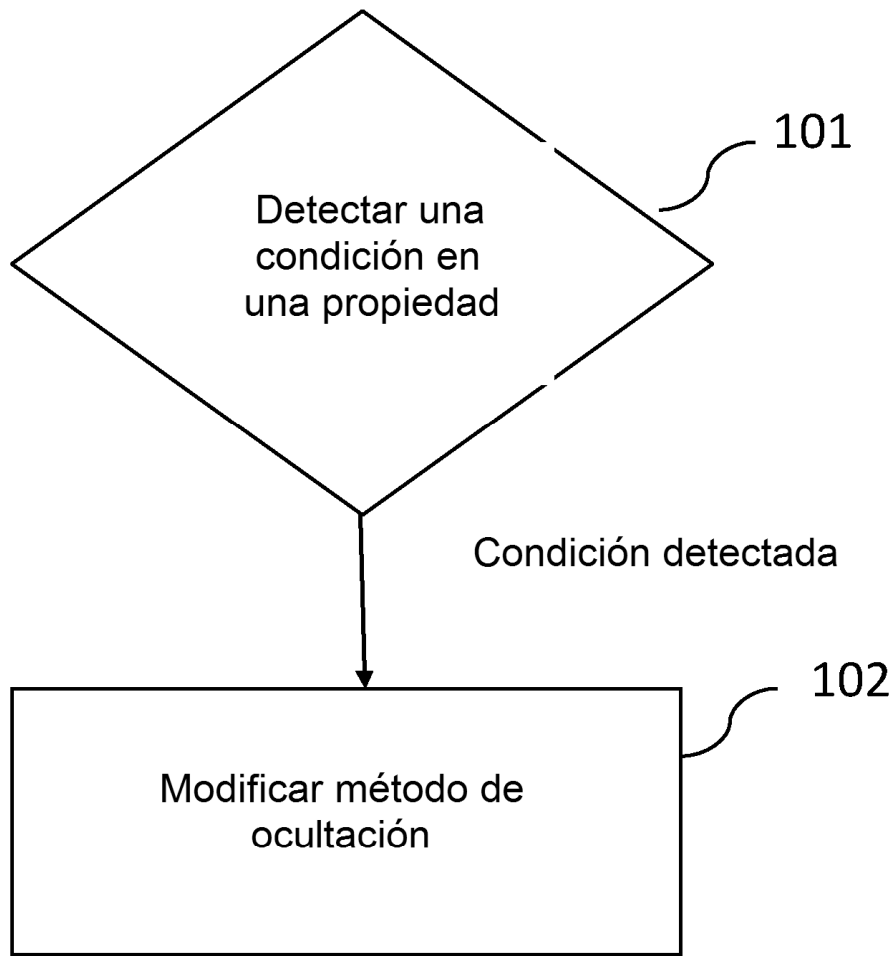


FIG. 10

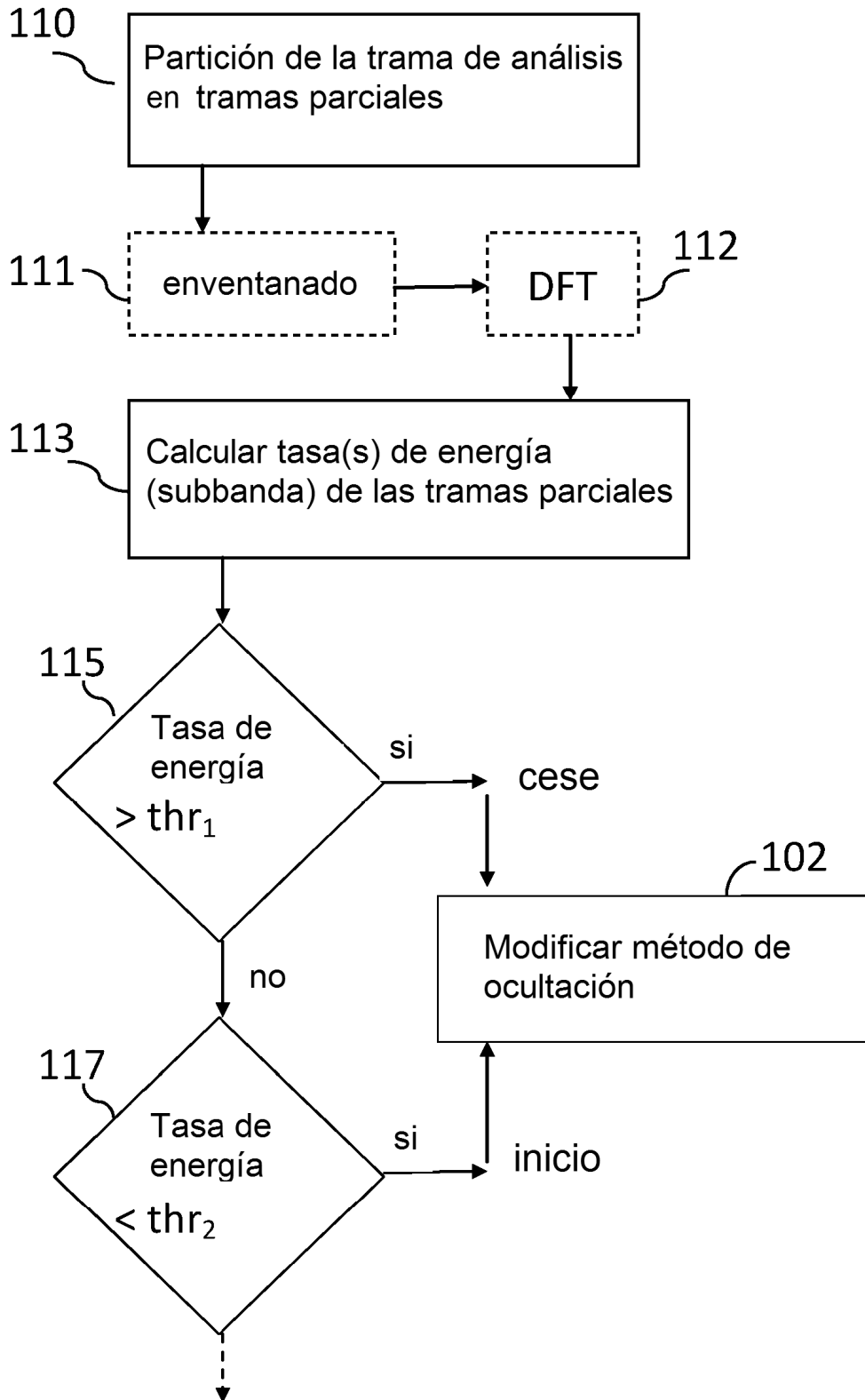


FIG. 11

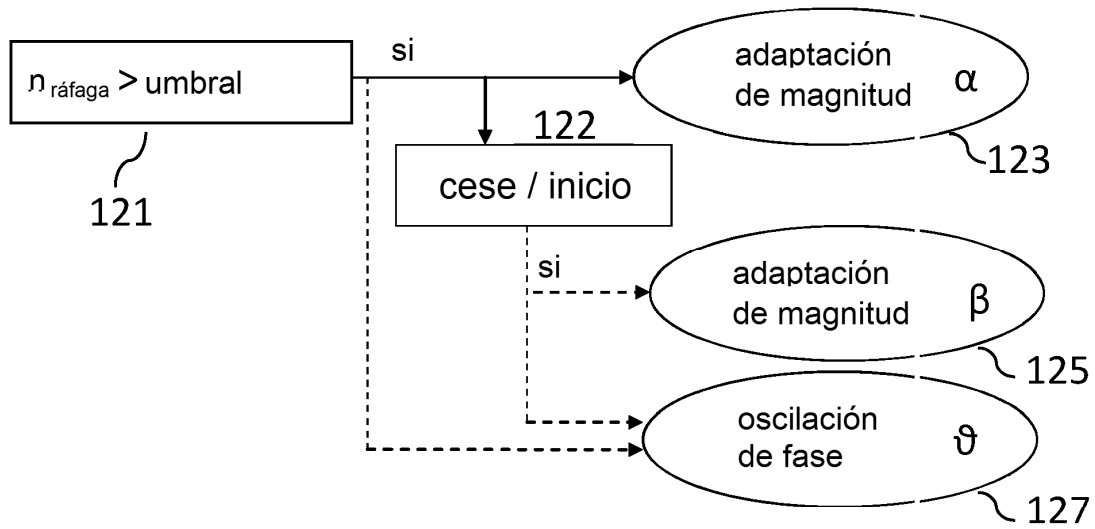


FIG. 12

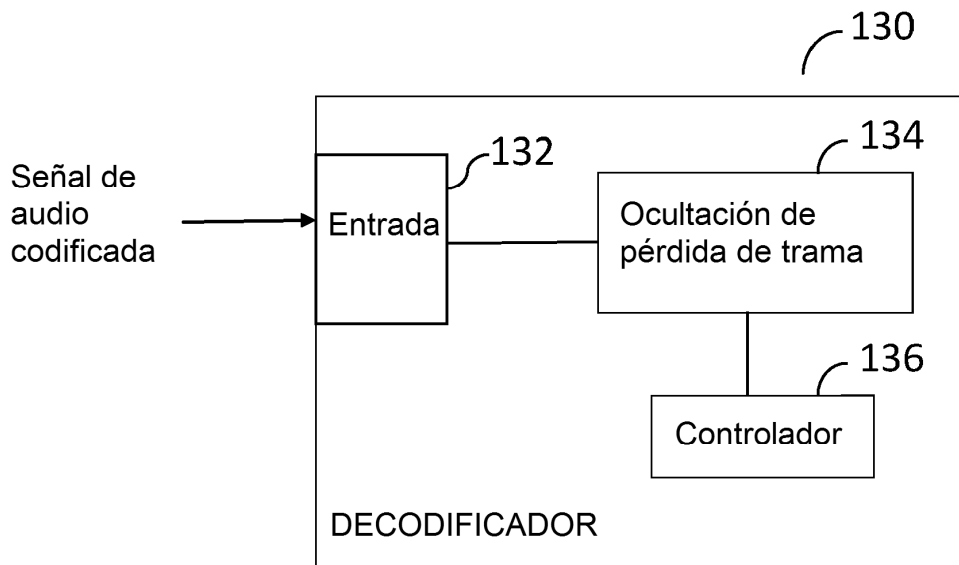


FIG. 13

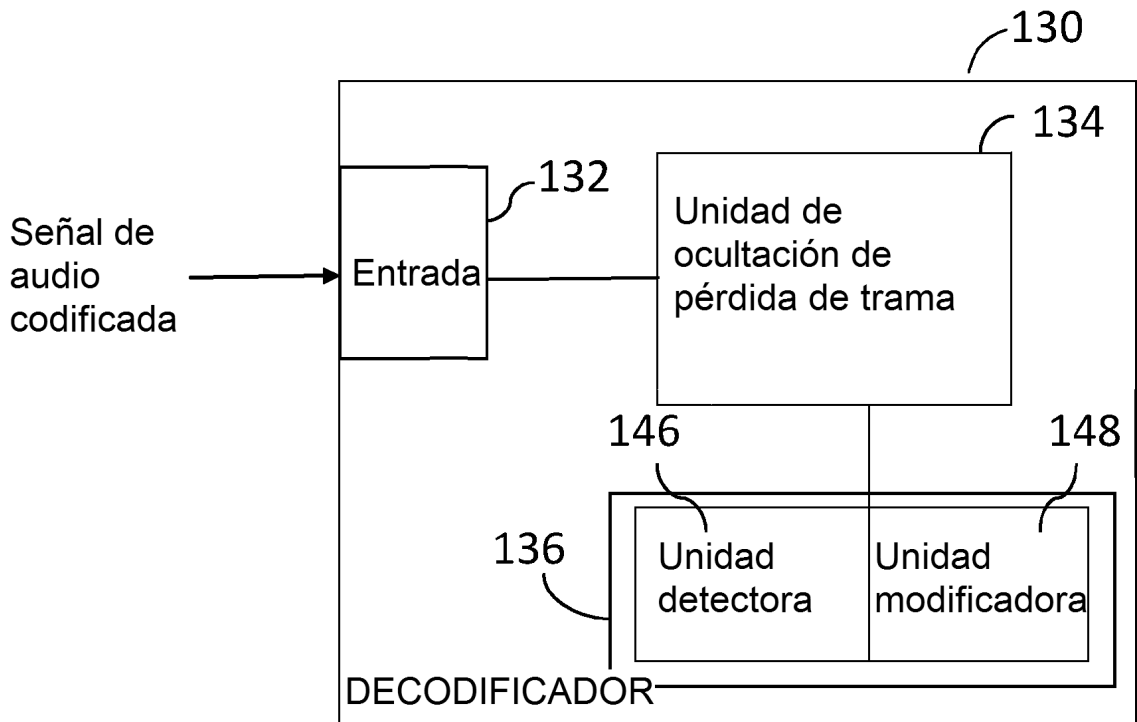


FIG. 14

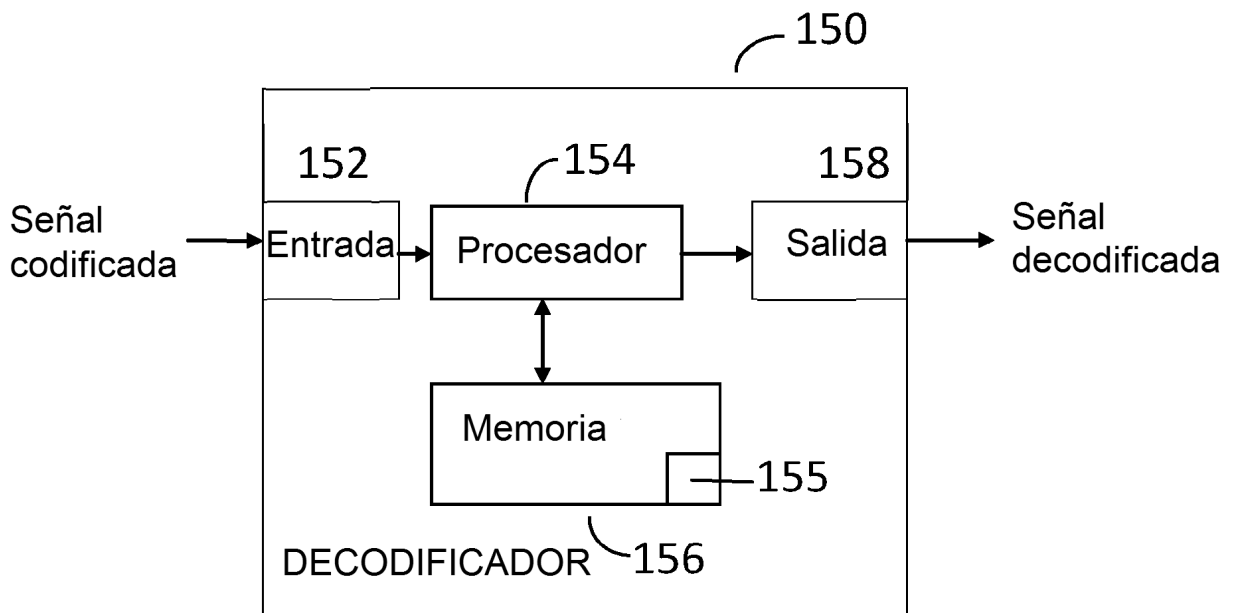


FIG. 15