

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 757 907**

51 Int. Cl.:

G10L 19/005 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **22.01.2014** E 17208127 (5)

97 Fecha y número de publicación de la concesión europea: **21.08.2019** EP 3333848

54 Título: **Encubrimiento de pérdida de trama de audio**

30 Prioridad:

05.02.2013 US 201361760814 P

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

30.04.2020

73 Titular/es:

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

72 Inventor/es:

BRUHN, STEFAN

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 757 907 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

Encubrimiento de pérdida de trama de audio

5 **Campo técnico**

La invención se refiere generalmente a un método para encubrir una trama de audio perdida de una señal de audio recibida. La invención también se refiere a un decodificador configurado para encubrir una trama de audio perdida de una señal de audio codificada recibida. La invención, adicionalmente, se refiere a un receptor que comprende un decodificador, y a un programa de ordenador y a un producto de programa de ordenador.

Antecedentes

Un sistema de comunicación de audio convencional transmite señales de audio y habla en tramas, que significa que el lado de envío primero dispone la señal de sonido en segmentos cortos, es decir, tramas de señal de sonido, de, por ejemplo, 20-40 ms, que después se codifican y se transmiten como una unidad lógica en, por ejemplo, un paquete de transmisión. Un decodificador en el lado de recepción decodifica cada una de esas unidades y reconstruye las correspondientes tramas de señal de audio, que a su vez son finalmente emitidas como una secuencia continua de muestras de señal de audio reconstruida.

Antes de la codificación, una conversión de analógico a digital (A/D) puede convertir la señal analógica de habla o audio procedente de un micrófono en una secuencia de muestras de señal digital de audio. A la inversa, en el extremo de recepción, un paso de conversión D/A final típicamente convierte la secuencia de muestras de señal digital de audio reconstruida en una señal analógica, continua en el tiempo, para reproducir en un altavoz.

Sin embargo, un sistema de transmisión convencional para señales de audio y habla puede sufrir errores de transmisión, que puede llevar a una situación en la que una o varias de las tramas transmitidas no están disponibles en el lado receptor para la reconstrucción. En ese caso, el decodificador tiene que generar una señal sustituta para cada trama no disponible. Esto puede realizarse por una llamada unidad de encubrimiento de pérdida de trama de audio en el decodificador en el lado receptor. El objetivo del encubrimiento de pérdida de trama es hacer que la pérdida de trama sea lo más inaudible posible, y, por consiguiente, mitigar el impacto de la pérdida de trama en la calidad de la señal reconstruida.

Los métodos de encubrimiento de pérdida de trama convencionales pueden depender de la estructura o la arquitectura del códec, por ejemplo, repitiendo parámetros de códec recibidos anteriormente. Dichas técnicas de repetición de parámetros son claramente dependientes en los parámetros específicos del códec usado, y pueden no ser fácilmente aplicables a otros códecs con una estructura diferente. Los métodos de encubrimiento de pérdida de tramas actuales pueden, por ejemplo, congelar y extrapolar parámetros de una trama recibida anteriormente con el fin de generar una trama de sustitución para la trama perdida. Los códecs lineares predictivos estandarizados AMR y AMR-WB son códecs de habla paramétrica que congelan los parámetros recibidos anteriores o usan alguna extrapolación de los mismos para la decodificación. En esencia, el principio es tener un modelo dado para codificar /decodificar y aplicar el mismo modelo con parámetros congelados o extrapolados.

Muchos códecs de audio aplican una técnica de dominio frecuencia de codificación, que implica aplicar un modelo de codificación en un parámetro espectral después de una transformada de dominio frecuencia. El decodificador reconstruye el espectro de señales desde los parámetros recibidos y transforma el espectro de vuelta a una señal de tiempo. Típicamente, la señal de tiempo se reconstruye trama por trama, y las tramas se combinan por técnicas de agregar y superponer y potenciar otros procesamientos adicionales para formar la señal final reconstruida. El encubrimiento de pérdida de trama de audio correspondiente aplica el mismo, o al menos un modelo de decodificación similar para tramas perdidas, donde los parámetros de dominio frecuencia de una trama recibida anteriormente son congelados o adecuadamente extrapolados y entonces usados en la conversión de dominio frecuencia a tiempo.

"Frame erasure concealment using sinusoidal analysis-synthesis and its application to MDCT-based codecs", por Parikh, et al., el documento ICASSP 2000, presenta un algoritmo de encubrimiento de borrado de trama basado en una síntesis de análisis sinusoidal. Cuando una trama se pierde, se realiza un análisis sinusoidal de la señal anterior decodificada y amortiguada. El análisis da un conjunto de sinusoides, que se usan para sintetizar la forma de onda correspondiente a la trama perdida.

Sin embargo, los métodos de encubrimiento de pérdida de trama de audio convencionales pueden sufrir de perjuicios de calidad, por ejemplo, desde la congelación de parámetros y la técnica de extrapolación y replicación del mismo modelo de decodificador para las tramas perdidas no siempre puede garantizar una evolución suave y fiel de la señal desde las tramas de señales decodificadas con anterioridad a la trama perdida. Esto puede llevar a discontinuidades de señal audibles con un impacto de calidad correspondiente. De este modo, el encubrimiento de pérdida de trama de audio con perjuicio de calidad reducido es deseable y necesitado.

Sumario

El objeto de realizaciones de la presente invención es para dirigir al menos algunos de los problemas mencionados anteriormente, y este objeto y otros se consiguen por el método y los arreglos según las reivindicaciones adjuntadas independientes, y por las realizaciones según las reivindicaciones dependientes.

Según un primer aspecto, un método de encubrimiento de pérdida de trama se divulga según la reivindicación 1.

Según un segundo aspecto, un aparato se configura para implementar un método de encubrimiento de pérdida de trama tal como se describe en la reivindicación 5.

El aparato puede estar comprendido en un decodificador de audio.

El decodificador puede estar implementado en un dispositivo, tal como, por ejemplo, un teléfono móvil.

Según un tercer aspecto, las realizaciones proporcionan un programa de ordenador que se define para encubrir una trama de audio perdida, donde el programa de ordenador comprende instrucciones que cuando se ejecutan por un procesador hace que el procesador encubra una trama de audio perdida, de acuerdo con el primer aspecto.

Según un cuarto aspecto, las realizaciones proporcionan un producto de programa de ordenador que comprende un medio de ordenador legible que almacena un programa de ordenador según el tercer aspecto mencionado anteriormente.

Las ventajas de las realizaciones descritas aquí son proporcionar un método de encubrimiento de pérdida de trama que permite mitigar el impacto audible de pérdida de trama en la transmisión de señales de audio, por ejemplo, de habla codificada. Una ventaja general es proporcionar una evolución suave y fiel de la señal reconstruida para una trama perdida, donde el impacto audible de las pérdidas de trama es muy reducido en comparación a técnicas convencionales.

Características y ventajas adicionales de las enseñanzas en las realizaciones de la presente aplicación quedarán claras después de leer la siguiente descripción y los dibujos que acompañan.

Breve descripción de los dibujos

Las realizaciones serán descritas más detalladamente y en referencia a los dibujos que se acompañan, en los que:

la figura 1 ilustra una función típica de ventana;

la figura 2 ilustra una función específica de ventana;

la figura 3 exhibe un ejemplo de un espectro de magnitud de una función de ventana;

la figura 4 ilustra un espectro de línea de una señal ejemplar sinusoidal con la frecuencia f_k ;

la figura 5 muestra un espectro de una señal sinusoidal de ventana con la frecuencia f_k ;

la figura 6 ilustra unas barras correspondientes a la magnitud de los puntos de cuadrícula de un DFT, basado en una trama de análisis;

la figura 7 ilustra una parábola que encaja a través de los puntos de cuadrícula DFT;

la figura 8 es un diagrama de flujo de un método según realizaciones;

las figuras 9 y 10 ambas ilustran un decodificador según realizaciones; y

la figura 11 ilustra un programa de ordenador y un producto de programa de ordenador, según realizaciones.

Descripción detallada

En lo siguiente, las realizaciones de la invención serán descritas más detalladamente. Para el objetivo de la explicación y no limitación, se divulgan detalles específicos, tal como escenarios y técnicas particulares, con el fin de proporcionar un pleno entendimiento.

Lo que, es más, es aparente que el método ejemplar y los dispositivos descritos más abajo pueden ser implementados, al menos en parte, por el uso de un funcionamiento de equipo lógico informático en conjunto con un microprocesador programado u ordenador de objetivo general, y/o usando un circuito integrado específico aplicado

(ASIC). Además, las realizaciones también pueden, al menos en parte, ser implementadas como un producto de programa de ordenador o en un sistema que comprende un procesador de ordenador y una memoria acoplada al procesador, donde la memoria está codificada con uno o más programas que pueden realizar las funciones divulgadas aquí.

5 Un concepto de las realizaciones descritas aquí en lo sucesivo comprende un encubrimiento de una trama de audio perdida mediante:

10 - la realización de un análisis sinusoidal de al menos parte de una señal de audio previamente recibida o reconstruida, donde el análisis sinusoidal implica identificar frecuencias de componentes sinusoidales de la señal de audio;

15 - la aplicación de un modelo sinusoidal en un segmento de la señal de audio previamente recibida o reconstruida, donde dicho segmento se usa como una trama prototipo con el fin de crear una trama de sustitución para una trama perdida, y

20 - la creación de la trama de sustitución que implica la evolución temporal de las componentes sinusoidales de la trama prototipo, hasta la instancia de tiempo de la trama de audio perdida, en respuesta a las frecuencias identificadas correspondientes.

Análisis sinusoidal

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

$$s(n) = \sum_{k=1}^K a_k \cdot \cos(2\pi \frac{f_k}{f_s} \cdot n + \varphi_k) \tag{6.1}$$

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

35 Es importante encontrar tantas frecuencias exactas de las sinusoides como sea posible. Mientras una señal sinusoidal ideal tendría un espectro de línea con frecuencias de línea f_k, encontrando que sus valores reales tendrían que requerir en principio medidas de tiempo infinitas. Entonces, en la práctica es difícil encontrar estas frecuencias, desde que solo pueden estimarse basadas en un periodo corto de medición, que corresponde al segmento de señal utilizado para el análisis sinusoidal según las realizaciones descritas aquí; este segmento de señal es referido después aquí como una trama de análisis. Otra dificultad es que la señal puede en la práctica variar en el tiempo, que significa que los parámetros de la ecuación anterior varían con el tiempo. Entonces, por un lado, es deseable utilizar una trama larga de análisis haciendo la medición más precisa; por otro lado, un periodo corto de medición tendría se necesitaría con el fin de enfrentarse mejor a las variaciones de señal posibles. Una buena compensación es utilizar una longitud de trama de análisis con el fin de, por ejemplo, 20-40 ms.

45 Según una realización preferida, las frecuencias de las sinusoides f_k se identifican por un análisis de dominio frecuencia de la trama de análisis. Con este fin, la trama de análisis se transforma en un dominio frecuencia, por ejemplo, por medio de DFT (Transformada Fourier Discreta) o DCT (Transformada de Coseno Discreta), o una transformada de dominio frecuencia similar. En el caso de que se use un DFT de la trama de análisis, el espectro es dado por:

$$X(m) = DFT(w(n) \cdot x(n)) = \sum_{n=0}^{L-1} e^{-j\frac{2\pi}{L}mn} \cdot w(n) \cdot x(n) \tag{6.2}$$

50 En esta ecuación, w(n) denota la función de ventana con la que la longitud de la trama de análisis L se extrae y se pesa.

55 La figura 1 ilustra una función de ventana típica, es decir, una ventana rectangular que es igual a 1 para n ∈ [0...L-1] y de otro modo 0. Se asume que los índices de tiempo de la señal de audio recibida anteriormente están establecidos de tal manera que la trama prototipo se referencia por los índices de tiempo n=0...L-1. Otras funciones de ventana que pueden ser más adecuadas para análisis espectrales son, por ejemplo, Hamming, Hanning, Kaiser o

Blackman.

La figura 2 ilustra una función de ventana más útil, que es una combinación de la ventana Hamming y la ventana rectangular. La ventana ilustrada en la figura 2 tiene una forma de borde ascendente como la mitad izquierda de una ventana Hamming de longitud L1 y una forma de borde descendente como la mitad derecha de una ventana Hamming de longitud L1 y entre los bordes 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 . La precisión de esta aproximación es, sin embargo, limitada por el espaciado

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

de frecuencia del DFT. Con el DFT con un bloqueo de longitud L la precisión está limitada a $\frac{f_s}{2L}$. Sin embargo, este nivel de precisión puede ser muy bajo en el alcance del método según las realizaciones descritas aquí, y una precisión mejorada puede obtenerse basada en los resultados de la siguiente consideración:

El espectro de la trama de análisis de ventana es dado por la circunvolución del espectro de la función de ventana con el espectro de línea de una señal sinusoidal modelo $S(\Omega)$, después muestreada en los puntos de cuadrícula del DFT:

$$X(m) = \int_{2\pi} \delta(\Omega - m \cdot \frac{2\pi}{L}) \cdot (W(\Omega) * S(\Omega)) \cdot d\Omega \quad (6.3)$$

Utilizando la expresión del espectro de la señal modelo sinusoidal, esto puede escribirse como:

$$X(m) = \frac{1}{2} \int_{2\pi} \delta(\Omega - m \cdot \frac{2\pi}{L}) \cdot \sum_{k=1}^K a_k \cdot \left(W(\Omega + 2\pi \frac{f_k}{f_s}) \cdot e^{-j\phi_k} + W(\Omega - 2\pi \frac{f_k}{f_s}) \cdot e^{j\phi_k} \right) \cdot d\Omega \quad (6.4)$$

Entonces, el espectro de muestra viene dado por:

$$X(m) = \frac{1}{2} \sum_{k=1}^K a_k \cdot \left(W(2\pi(\frac{m}{L} + \frac{f_k}{f_s})) \cdot e^{-j\phi_k} + W(2\pi(\frac{m}{L} - \frac{f_k}{f_s})) \cdot e^{j\phi_k} \right) \quad (6.5)$$

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

Basándose en esto, los picos observados en el espectro de magnitud de la trama vástago de análisis desde una señal sinusoidal de ventana con K sinusoides, donde las verdaderas frecuencias de senoide se encuentran en las proximidades de los picos. Además, la identificación de frecuencias de componentes sinusoidales puede además implicar identificar frecuencias en las proximidades de los picos del espectro con relación a la transformada usada de dominio frecuencia.

Si se asume que m_k es un índice DFT (punto de cuadrícula) del pico $k^{\text{ésimo}}$ observado, entonces la frecuencia correspondiente es $\hat{f}_k = \frac{m_k}{L} \cdot f_s$ que puede considerarse una aproximación de la verdadera frecuencia sinusoidal f_k . Se puede asumir que la verdadera frecuencia de senoide f_k se sitúa dentro del intervalo

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

Por claridad, se aprecia que la circunvolución del espectro de la función de ventana con el espectro del espectro de línea de la señal modelo sinusoidal puede entenderse como una superposición de versiones de frecuencia cambiadas del espectro de función de ventana, en donde las frecuencias cambiadas son las frecuencias de las sinusoides. Esta superposición es entonces muestreada en los puntos de cuadrícula DFT. La circunvolución del espectro de la función de ventana con el espectro del espectro de línea de la señal modelo sinusoidal se ilustran en las figuras 3 - figura 7, de las que la figura 3 muestra un ejemplo del espectro de magnitud de la función de ventana, y la figura 4 el espectro de magnitud (espectro de línea) de una señal sinusoidal de ejemplo con una única senoide con una frecuencia f_k . La figura 5 muestra el espectro de magnitud de la señal sinusoidal de ventana que replica y superpone el espectro de ventana de frecuencias cambiadas en las frecuencias de la senoide, y las barras en la figura 6 corresponden a la magnitud de los puntos de cuadrícula del DFT de la senoide de ventana que se obtienen calculando el DFT de la trama de análisis. Aprecie que todos los espectros son periódicos con el parámetro de frecuencia normalizado Ω donde $\Omega=2\pi$ que corresponde a la frecuencia de muestreo f_s .

Basado en la discusión anterior, y basado en la ilustración en la figura 6, se puede encontrar una mejor aproximación de las frecuencias sinusoidales reales incrementando la resolución de la búsqueda, de tal manera que es mayor que la resolución de frecuencia de la transformada de dominio frecuencia usada.

5 Además, la identificación de frecuencias de componentes sinusoidales es preferiblemente realizada con una resolución más alta que la resolución de frecuencia de la transformada de dominio frecuencia usada, y la identificación puede además implicar interpolación.

10 Una manera ejemplar preferida para encontrar una mejor aproximación de las frecuencias f_k de las sinusoides es aplicar interpolación parabólica. Un enfoque es ajustar parábolas a través de puntos de cuadrícula del espectro de magnitud DFT que rodea los picos y calcular las respectivas frecuencias que pertenecen a la máxima parábola, y una decisión adecuada ejemplar para el orden de las parábolas es 2. En más detalle, puede aplicarse el siguiente procedimiento:

15 1) Identificar los picos del DFT de la trama de análisis de ventana. La búsqueda de pico entregará el número de picos K y los correspondientes índices DFT de los picos. La búsqueda de pico puede estar hecha típicamente en el espectro de magnitud DFT o en el espectro de magnitud logarítmico DFT.

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

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

25 La figura 7 ilustra el ajuste de la parábola a través de puntos de cuadrícula DFT P_1, P_2 y P_3 .

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

30 Aplicar un modelo sinusoidal

La aplicación de un modelo sinusoidal con el fin de realizar una operación de encubrimiento de pérdida de trama según las realizaciones puede describirse como sigue:

35 En el caso de que un segmento dado de la señal codificada no pueda ser reconstruida por el decodificador desde que la información codificada correspondiente no está disponible, es decir, desde que una trama se ha perdido, una parte disponible de la señal anterior a este segmento puede ser usada como una trama prototipo. Si $y(n)$ con $n=0\dots N-1$ es el segmento no disponible para el cual una trama de sustitución $z(n)$ tiene que ser generado, y $y(n)$ con $n=0$ es la señal decodificada anteriormente disponible, una trama prototipo de la señal disponible de longitud L y primer índice n_{-1} se extrae con una función de ventana $x(n)$ y se transforma en un dominio frecuencia, por ejemplo por medio de DFT:

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

45 La función de ventana puede ser una de las funciones de ventana descritas anteriormente en el análisis sinusoidal. Preferiblemente, con el fin de ahorrar complejidad numérica, la trama transformada de dominio frecuencia debería ser idéntica a la usada durante el análisis sinusoidal.

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

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

55 Esta expresión también fue utilizada en la parte de análisis y es descrita en detalle anteriormente.

Después, hay que darse cuenta de que el espectro de la función de ventana usada tiene solo una contribución significativa en un rango de frecuencia cerca de cero. Como se ilustra en la figura 3 el espectro de magnitud de la función de ventana es grande para frecuencias cerca de cero y pequeñas sin embargo (dentro del rango de

frecuencia normalizado desde $-\pi$ a π , que corresponde a la mitad de la frecuencia muestreada. Entonces, como una aproximación, se asume 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 números positivos pequeños. En particular, una aproximación del espectro de función de ventana se usa de tal forma que, para cada k , las contribuciones de los espectros de ventana cambiados en la expresión anterior estrictamente no se superponen. Entonces, en la ecuación anterior para cada índice de frecuencia siempre hay solo al máximo la contribución desde un sumando, es decir, desde un espectro de ventana cambiado. Esto significa que la expresión anterior se reduce a la siguiente expresión aproximada:

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

para m no negativos $\in M_k$ y para cada k . Aquí, M_k denota el intervalo entero

$$M_k = \left[\text{redondeo}\left(\frac{f_k}{f_s} \cdot L\right) - m_{\min,k}, \text{redondeo}\left(\frac{f_k}{f_s} \cdot L\right) + m_{\max,k} \right]$$

donde $m_{\min,k}$ y $m_{\max,k}$ llenan la restricción explicada anteriormente de tal forma que los intervalos no se superponen. Una decisión adecuada para $m_{\min,k}$ y $m_{\max,k}$ es establecerlos en un valor entero pequeño, por ejemplo $\delta = 3$. Si, sin embargo, los índices DFT en relación a dos frecuencias sinusoidales vecinas f_k y f_{k+1} son menos que 2δ , entonces δ

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

se establece en de tal forma que se asegura que los intervalos no se superponen. La función suelo(.) es el entero más cercano al argumento de función que es menor o igual a él.

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

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

Por consiguiente, el espectro DFT del modelo sinusoidal evolucionado viene dado por:

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

Aplicar otra vez la aproximación según la cual los espectros de función de ventana cambiados no se superponen da:

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

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

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

$$\theta_k = 2\pi \cdot \frac{f_k}{f_s} n_{-1}, \text{ para cada } m \in M_k.$$

Entonces, la trama de sustitución puede calcularse con la siguiente expresión:

$$Z(n) = \text{IDFT}\{Z(m)\} \text{ con } Z(m) = Y(m) \cdot e^{j\theta_k}$$

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

Una realización específica se dirige a la aleatoriedad de fase para índices DFT que no pertenecen a ningún intervalo M_k . Como se describe anteriormente, los intervalos $M_k, k=1 \dots K$ tienen que estar establecidos de tal forma que no están estrictamente sin superponerse lo que se hace usando algún parámetro δ que controla el tamaño de los

intervalos. Puede ocurrir que δ sea pequeño en relación a la distancia de frecuencia de dos sinusoides vecinas. Entonces, en ese caso ocurre que hay un espacio entre dos intervalos. Por consiguiente, para los correspondientes índices DFT m , no hay cambio de fase correspondiente a la expresión anterior $Z(m) = Y(m) \cdot e^{i\theta_k}$ se defina. Una decisión adecuada según esta realización es aleatorizar la fase para estos índices, flexibilizar $Z(m) = Y(m) \cdot e^{i2\pi \text{rand}(\cdot)}$, donde la función $\text{rand}(\cdot)$ devuelve algún número aleatorio.

Basado en lo anterior, la figura 8 es un diagrama de flujo que ilustra un método de encubrimiento de pérdida de trama de audio ejemplar según las realizaciones:

10 En el paso 81, se realiza un análisis sinusoidal de una parte de una señal de audio reconstruida o recibida anteriormente, donde el análisis sinusoidal implica identificar frecuencias de componentes sinusoidales, es decir, sinusoides, de la señal de audio. Después, en el paso 82, un modelo sinusoidal se aplica en un segmento de la señal de audio reconstruida o recibida anteriormente, donde dicho segmento se usa como una trama prototipo con el fin de crear una trama de sustitución para una trama de audio perdida, y en el paso 83 se crea la trama de sustitución para la trama de audio perdida, que implica la evolución del tiempo de componentes sinusoidales, es decir, sinusoides, de la trama prototipo, hasta la instancia de tiempo de la trama de audio perdida, en respuesta a las frecuencias identificadas correspondientes.

20 Según una realización más, se asume que la señal de audio se compone de un número limitado de componentes sinusoidales individuales, y que al análisis sinusoidal se realiza en el dominio frecuencia. Además, la identificación de las frecuencias de componentes sinusoidales puede implicar identificar frecuencias en las proximidades de los picos de un espectro en relación a la transformada de dominio frecuencia usada.

25 Según una realización ejemplar, la identificación de frecuencias de componentes sinusoidales se realiza con una resolución más alta que la resolución de la transformada de dominio frecuencia usada, y la identificación puede además implicar interpolación, por ejemplo, de tipo parabólico.

30 Según una realización ejemplar, el método comprende extraer una trama prototipo de una señal previamente recibida o reconstruida disponible usando una función de ventana, y donde la trama prototipo extraída puede ser transformada en un dominio frecuencia.

35 Una realización más implica una aproximación de un espectro de la función de ventana, de tal forma que el espectro de la trama de sustitución se compone estrictamente de porciones no superpuestas del espectro de función de ventana aproximada.

40 Según una realización ejemplar más, el método comprende componentes sinusoidales de evolución de tiempo de un espectro de frecuencia de una trama prototipo avanzando la fase de las 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 cambiar un coeficiente espectral de la trama prototipo incluida en un intervalo M_k en las proximidades de la senoide k por un cambio proporcional en la fase a la frecuencia sinusoidal f_k y a la diferencia de tiempo entre la trama de audio perdida y la trama prototipo.

45 Una realización más comprende cambiar la fase de un coeficiente espectral de la trama prototipo que no pertenece a una senoide identificada por una fase aleatoria, o cambiar 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.

50 Una realización más implica una transformada de dominio frecuencia inversa del espectro de frecuencia de la trama prototipo.

Más específicamente, el método de encubrimiento de la pérdida de la trama de audio según a una realización más puede implicar los siguientes pasos:

55 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 y_{-1} desde la señal disponible previamente sintetizada y calcular el DFT de esa trama.

60 3) Calcular el cambio de fase 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.

4) Para cada senoide k avanzar la fase de la trama prototipo DFT con θ_k selectivamente para los índices DFT relacionados con unas proximidades alrededor de la frecuencia de senoide f_k .

65 5) Calcular el DFT inverso del espectro obtenido 4).

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

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 a algún instante anterior en el tiempo.

c) La suposición de una aproximación del espectro de una función de ventana de tal forma que el espectro de la trama de sustitución puede ser construido por porciones no superpuestas de espectros de función de ventana de frecuencia cambiada, las frecuencias cambiadas siendo las frecuencias de sinusoides.

La figura 9 es un diagrama esquemático de bloque que ilustra un decodificador ejemplar 1 configurado para realizar un método de encubrimiento de pérdida de trama de audio según las realizaciones. El decodificador ilustrado comprende uno o más procesadores 11 y equipo lógico informático adecuado con almacenamiento o memoria adecuada 12. La señal de audio codificada entrante se recibe por una entrada (IN), a la que el procesador 11 y la memoria 12 están conectados. La señal de audio decodificada y reconstruida obtenida del equipo lógico informático es emitida desde la salida (OUT). Un decodificador ejemplar se configura para encubrir una trama de audio perdida de una señal de audio recibida, y comprende un procesador 11 y memoria 12, donde la memoria contiene instrucciones ejecutables por el procesador 11, y donde el decodificador 1 se configura para:

- realizar un análisis sinusoidal de una parte de una señal de audio previamente recibida o reconstruida, donde el análisis sinusoidal implica identificar frecuencias de componentes sinusoidales de la señal de audio;

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

- crear la trama de sustitución para la trama de audio perdida evolucionando en el tiempo componentes sinusoidales de la trama prototipo, hasta el instante en el tiempo de la trama de audio perdida, en respuesta a las frecuencias identificadas correspondientes.

Según una realización más del decodificador, el modelo sinusoidal aplicado asume que la señal de audio está compuesta de un número limitado de componentes sinusoidales individuales, y la identificación de frecuencias de componentes sinusoidales de la señal de audio puede además comprender una interpolación parabólica.

Según una realización más, el decodificador está configurado para extraer una trama prototipo desde una señal disponible previamente recibida o reconstruida utilizando una función de ventana, y para transformar la trama prototipo extraído en un dominio frecuencia.

Según otra realización más, el decodificador está configurado para evolucionar en el tiempo componentes sinusoidales de un espectro de frecuencia de una trama prototipo avanzando la fase de las 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 para crear la trama de sustitución realizando una transformada de frecuencia inversa del espectro de frecuencia.

Un decodificador según una realización alternativa se ilustra en la figura 10a, tal que comprende una unidad de salida configurada para recibir una señal de audio codificada. La figura ilustra el encubrimiento de pérdida de trama por una unidad de encubrimiento de pérdida de trama lógica 13, donde el decodificador 1 está configurado para implementar un encubrimiento de una trama de audio perdida según las realizaciones descritas anteriormente. La unidad de encubrimiento de pérdida de trama lógica 13 está mejor ilustrada en la figura 10b, y comprende medios adecuados para encubrir una trama de audio perdida, es decir, medios 14 para realizar un análisis sinusoidal de una parte de una señal de audio previamente recibida o reconstruida, donde el análisis sinusoidal implica identificar frecuencias de componentes sinusoidales de la señal de audio, medios 15 para aplicar un modelo sinusoidal en un segmento de la señal de audio previamente recibida o reconstruida, donde dicho segmento es utilizado como una trama prototipo con el fin de crear una trama de sustitución para una trama de audio perdida, y medios 16 para crear la trama de sustitución para la trama de audio perdida evolucionando en el tiempo componentes sinusoidales de la trama prototipo, hasta el instante en el tiempo de la trama de audio perdida, en respuesta a las frecuencias identificadas correspondientes.

Las unidades y medios incluidos en el decodificador ilustrado en las figuras pueden ser implementados al menos en parte en un equipo físico informático, y hay numerosas variantes de elementos de circuitos que pueden ser utilizados y combinados para lograr las funciones de las unidades del decodificador. Dichas variantes se abarcan en las realizaciones. Un ejemplo particular de implementación de equipo físico informático del decodificador es la implementación en un procesador de señal digital (DSP) en el equipo físico informático y un circuito tecnológico integrado, incluyendo ambos, el circuito electrónico de objetivo general y el circuito de aplicación específica.

- 5 Un programa de ordenador según las realizaciones de la presente invención comprende instrucciones que cuando se ejecutan por un procesador hace que el procesador realice un método según un método descrito en conexión con la figura 8. La figura 11 ilustra un producto de programa de ordenador 9 según las realizaciones, en la forma de una memoria no volátil, por ejemplo, un EEPROM (Memoria Solo Legible Programable Borrable Eléctricamente), una memoria flash o un disco duro. El producto de programa de ordenador comprende un medio de ordenador legible almacenando un programa de ordenador 91, que comprende módulos de programa de ordenador 91 a, b, c, d que cuando se ejecutan en un decodificador 1 hace que el procesador del decodificador realice los pasos según la figura 8.
- 10 Un decodificador según las realizaciones de esta invención puede ser utilizado, por ejemplo, en un receptor para un dispositivo móvil, por ejemplo, un teléfono móvil o un ordenador portátil, o en un receptor para un dispositivo estático, por ejemplo, un ordenador personal.
- 15 Las ventajas de las realizaciones descritas aquí son para proporcionar un método de encubrimiento de pérdida de trama que permite mitigar el impacto audible de pérdida de trama en la transmisión de señales de audio, por ejemplo de habla codificada. Una ventaja general es proporcionar una evolución suave y fiel de la señal reconstruida para una trama perdida, donde el impacto audible de pérdidas de trama es muy reducido en comparación a técnicas convencionales.
- 20 Debe entenderse que la decisión de módulos y unidades que interaccionan, así como el nombramiento de las unidades solo son con objetivo ejemplar, y pueden ser configuradas en una pluralidad de maneras alternativas con el fin de ser capaces de ejecutar las acciones del proceso divulgado. También debería observarse que las unidades o módulos descritos en esta divulgación deben ser consideradas como entidades lógicas y no con necesidades como entidades físicas separadas. Se apreciará que el alcance de la tecnología descrita aquí abarca totalmente otras realizaciones que pueden resultar obvias para aquellos expertos en la técnica, y que por consiguiente el
- 25 alcance de esta divulgación no se ha de limitar.

REIVINDICACIONES

1. Un método de encubrimiento de pérdida de trama, en el que un segmento procedente de una señal de audio previamente sintetizada se utiliza como una trama prototipo para crear una trama de sustitución para una trama de audio perdida, comprendiendo el método:
- transformar la trama prototipo en un dominio frecuencia;
 - aplicar un modelo sinusoidal a la trama prototipo para identificar la frecuencia de una componente sinusoidal de la señal de audio;
 - calcular un cambio de fase θ_k para la componente sinusoidal;
 - cambiar θ_k la fase de la componente sinusoidal;
 - crear la trama de sustitución realizando una transformada de frecuencia inversa de un espectro de frecuencia de la trama prototipo;
- caracterizado porque:
- cambiar la fase de la componente sinusoidal comprende cambiar θ_k una fase de todos los coeficientes espectrales en la trama prototipo incluida en un intervalo M_k alrededor de una sinusoide k ;
 - las fases de coeficientes espectrales que no son cambiadas de fase son aleatorizadas; y
 - un espectro de magnitud de la trama prototipo se mantiene sin cambios.
2. El método de encubrimiento de pérdida de trama según la reivindicación 1, en el que el cambio de fase θ_k depende de la frecuencia sinusoidal f_k y un cambio de tiempo entre la trama prototipo y la trama perdida.
3. El método de encubrimiento de pérdida de trama según la reivindicación 1, en el que la identificación de la frecuencia de una componente sinusoidal implica además identificar frecuencias en las proximidades de picos del espectro relacionado con una transformada del dominio frecuencia utilizado.
4. El método de encubrimiento de pérdida de trama según la reivindicación 1, en el que la identificación de la frecuencia de una componente sinusoidal se realiza con mayor resolución que la resolución de frecuencia de la transformada del dominio frecuencia utilizado.
5. Un aparato (13) para crear una trama de sustitución para una trama de audio perdida, comprendiendo el aparato:
- medios para generar una trama prototipo desde un segmento de una señal de audio previamente sintetizada;
 - medios para transformar la trama prototipo en un dominio frecuencia;
 - medios para aplicar un modelo sinusoidal a la trama prototipo para identificar la frecuencia de una componente sinusoidal de la señal de audio;
 - medios para calcular un cambio de fase θ_k para la componente sinusoidal;
 - medios para cambiar θ_k la fase de la componente sinusoidal;
 - medios para crear la trama de sustitución realizando una transformada de frecuencia inversa de un espectro de frecuencia de la trama prototipo;
- caracterizado porque:
- cambiar la fase de la componente sinusoidal comprende cambiar θ_k una fase de todos los coeficientes espectrales en la trama prototipo incluida en un intervalo M_k alrededor de una sinusoide k ,
 - las fases de los coeficientes espectrales que no son cambiadas de fase son aleatorizadas, y
 - un espectro de magnitud de la trama prototipo se mantiene sin cambios.
6. El aparato según la reivindicación 5, en el que el cambio de fase θ_k depende de la frecuencia sinusoidal f_k y un cambio de tiempo entre la trama prototipo y la trama perdida.

7. El aparato según la reivindicación 5, en el que la identificación de la frecuencia de una componente sinusoidal implica además identificar frecuencias en las proximidades de picos del espectro relacionado con una transformada del dominio frecuencia utilizado.

5 8. El aparato según la reivindicación 5, en el que la identificación de la frecuencia de una componente sinusoidal se realiza con mayor resolución que la resolución de frecuencia de la transformada del dominio frecuencia utilizado.

10 9. Un decodificador de audio (1) que comprende el aparato según una cualquiera de las reivindicaciones 5 a 8.

10 10. Un dispositivo que comprende el decodificador de audio según la reivindicación 9.

15 11. Un programa de ordenador (91) que comprende instrucciones que, cuando se ejecutan en al menos un procesador, causa que al menos un procesador lleve a cabo el método según una cualquiera de las reivindicaciones 1 a 4.

12. Un portador de datos legible informáticamente que almacena el programa de ordenador (91) según la reivindicación 11.

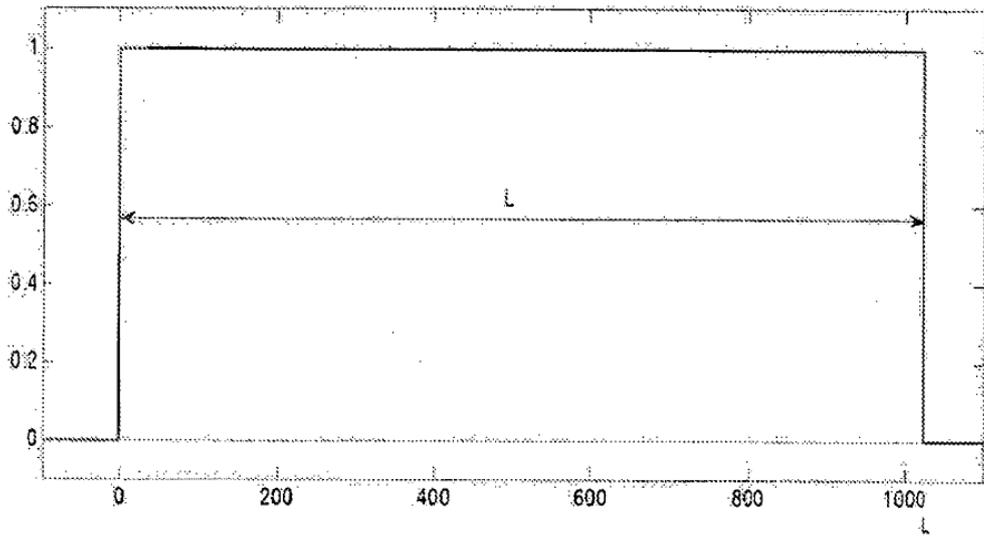


Fig. 1

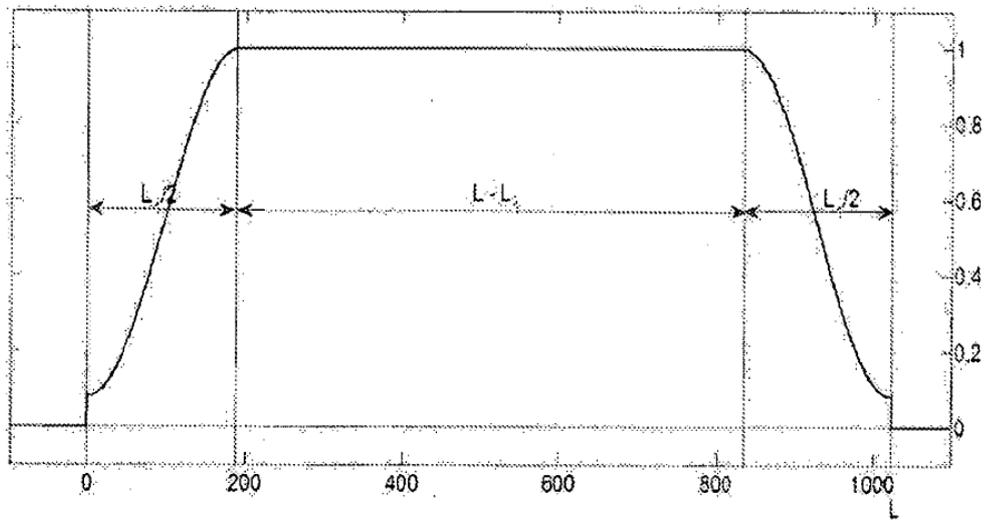


Fig. 2

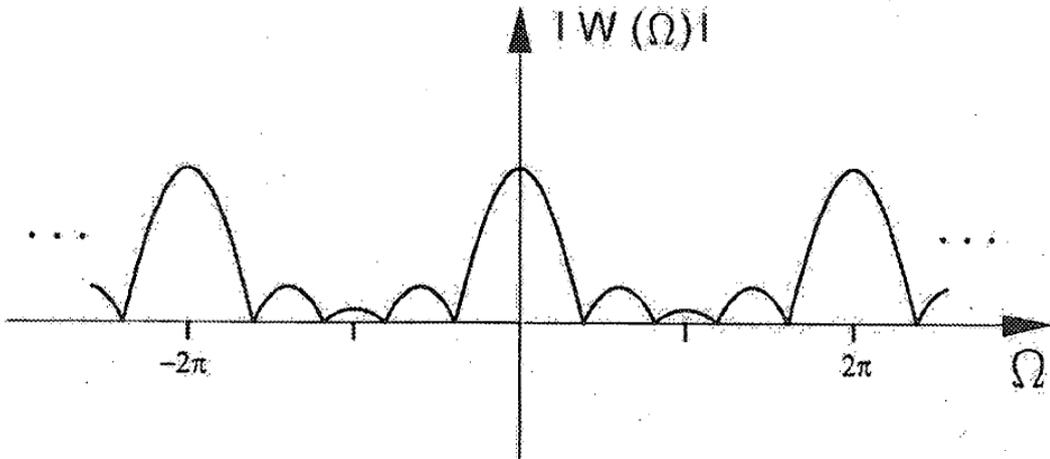


FIG. 3

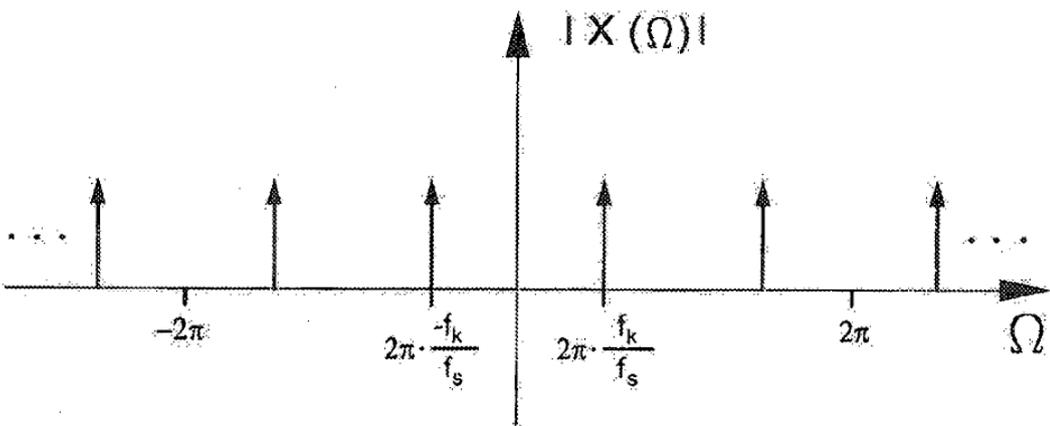
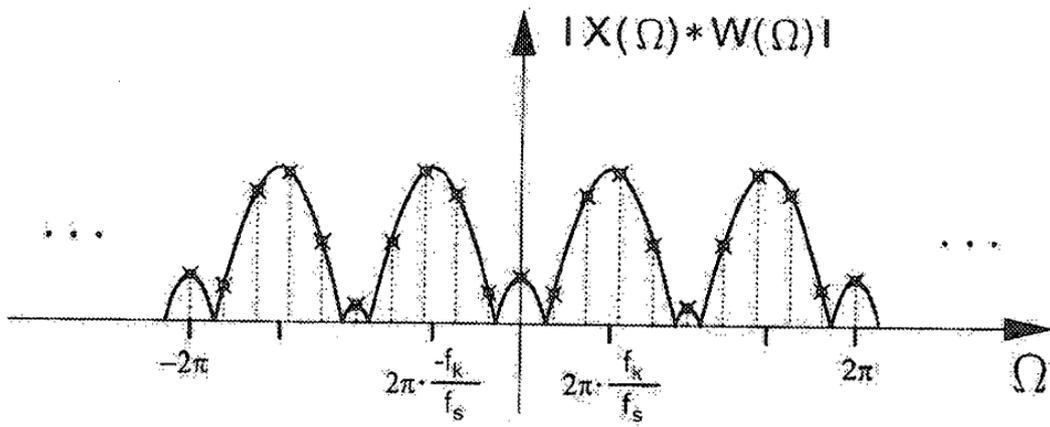
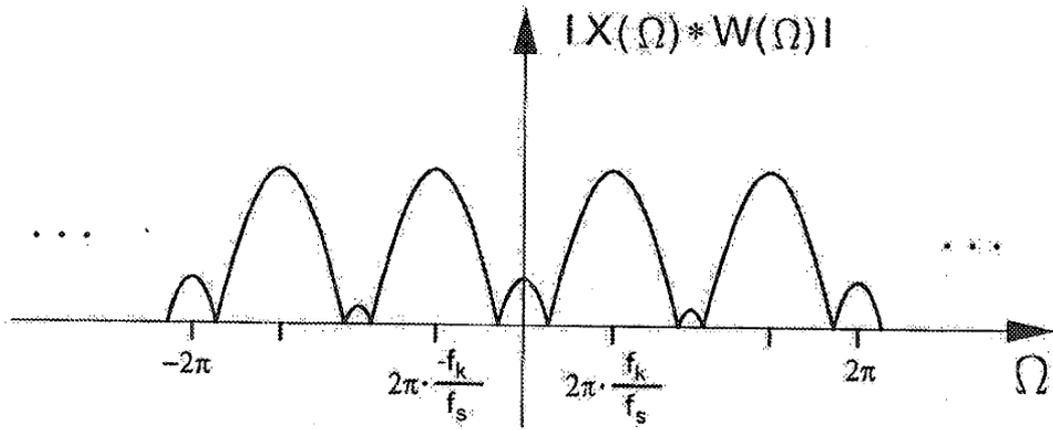


FIG. 4



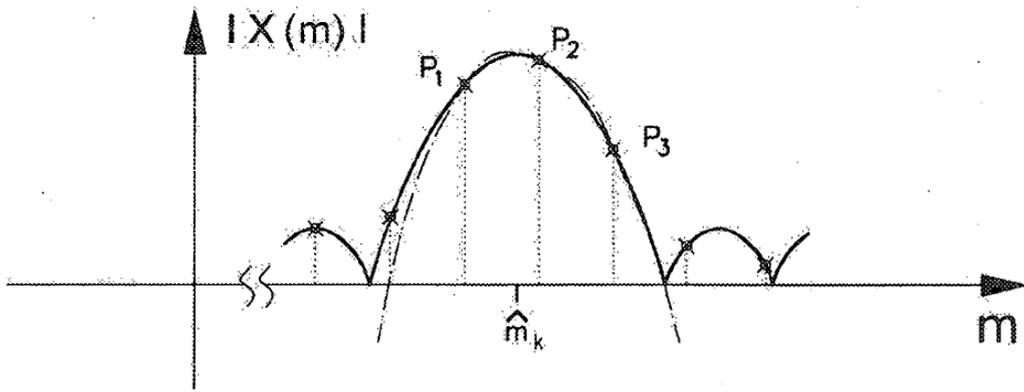


FIG. 7

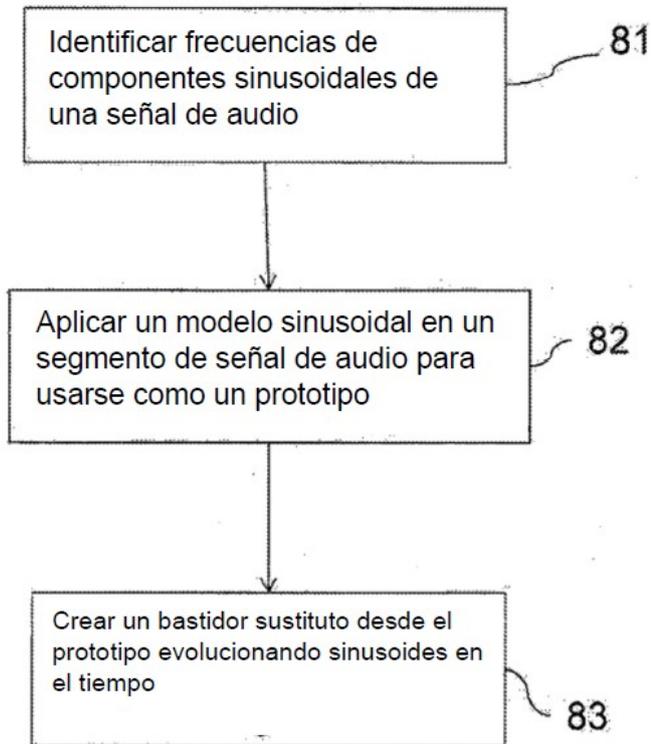


Fig. 8

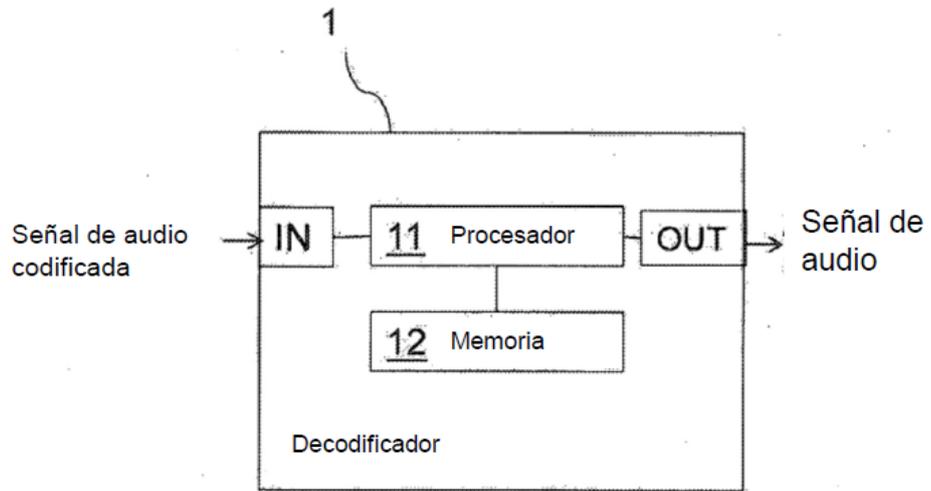


Fig. 9

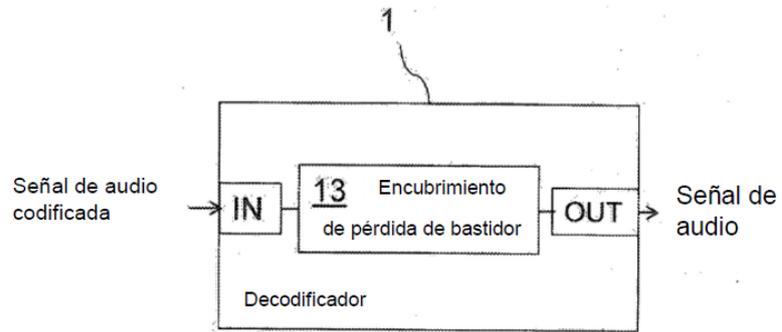


Fig. 10a

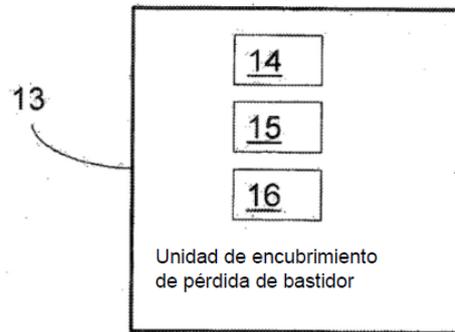


Fig. 10b

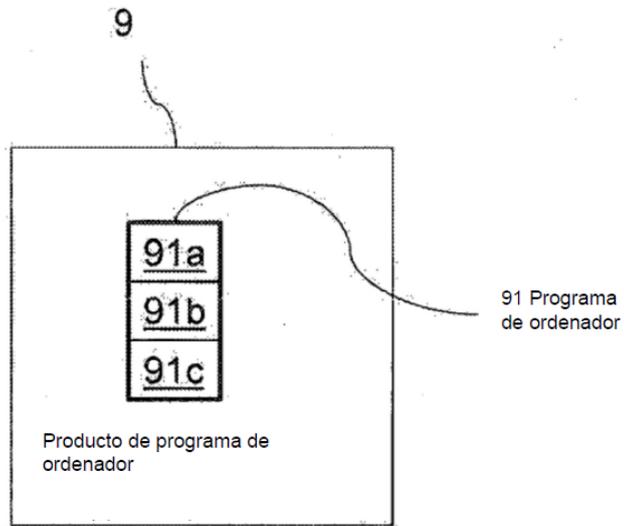


Fig. 11