

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 994 065**

51 Int. Cl.:

G10L 19/005 (2013.01)

G10L 25/90 (2013.01)

G10L 19/107 (2013.01)

G10L 19/08 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.06.2014 E 19172360 (0)**

97 Fecha y número de publicación de la concesión europea: **03.07.2024 EP 3540731**

54 Título: **Estimación del retardo del tono**

30 Prioridad:

21.06.2013 EP 13173157

05.05.2014 EP 14166990

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
16.01.2025

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.00%)**

**Hansastr. 27c
80686 München, DE**

72 Inventor/es:

**LECOMTE, JÉRÉMIE;
SCHNABEL, MICHAEL;
MARKOVIC, GORAN;
DIETZ, MARTIN y
NEUGEBAUER, BERNHARD**

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 994 065 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

Estimación del retardo del tono

5 La presente invención se refiere al procesamiento de señales de audio, en particular, al procesamiento del habla, y más particularmente, a un aparato y a un método para mejorar la ocultación del libro de códigos adaptativo en la ocultación similar a ACELP (ACELP = Predicción Lineal Excitada por Código Algebraico).

10 El procesamiento de señales de audio se hace cada vez más importante. En el campo del procesamiento de las señales de audio, las técnicas de ocultación juegan un papel importante. Cuando una trama se pierde o se daña, la información perdida de la trama perdida o dañada tiene que ser reemplazada. En el procesamiento de las señales del habla, en particular, cuando se consideran códecs de voz de ACELP o similares a ACELP; la información de tono es muy importante. Se necesitan técnicas para la predicción de tono y para la resincronización del pulso.

15 En cuanto a la reconstrucción del tono, existen diferentes técnicas de extrapolación de tono en la técnica anterior.

Una de estas técnicas es una técnica basada en la repetición. La mayor parte de los códecs del estado de la técnica aplica un enfoque simple de ocultación basada en repetición, lo que significa que el último periodo de tono recibido correctamente antes que se repita la pérdida de paquetes, hasta que llegue una trama buena y sea posible decodificar nueva información de tono a partir de la corriente de bits. O bien, se aplica una lógica de estabilidad del tono según la cual se elige un valor de tono que se ha recibido algún tiempo más antes de la pérdida de paquetes. Los códecs que siguen el enfoque basado en la repetición son, por ejemplo, el G.719 (ver [ITU08b, 8.6]), G.729 (ver [ITU12, 4.4]), AMR (ver [3GP12a, 6.2.3.1], [ITU03]), AMR-WB (ver [3GP12b, 6.2.3.4.2]) y AMR-WB + (ACELP y TCX20 (ocultación similar a ACELP) (ver [3GP09])); (AMR = Coeficiente Múltiples Adaptativos, AMR-WB = Ancho de Banda de Coeficientes Múltiples Adaptativo).

Otra técnica para la reconstrucción del tono de la técnica anterior es la derivación de tono a partir del dominio de tiempo. Para algunos códecs, el tono es necesario para la ocultación, pero no está integrado en el flujo de bits. Por ello, el tono se calcula en base a la señal del dominio de tiempo de la trama anterior con el fin de calcular el periodo de tono, que luego se mantiene constante durante la ocultación. Un códec que sigue este enfoque es, por ejemplo, el G.722, véase, en particular G.722 Apéndice 3 (ver [ITU06a, III.6.6 y III.6.7]) y G.722 Apéndice 4 (ver [ITU07, IV. 6.1.2.5]).

Otra técnica para la reconstrucción del tono de la técnica anterior se basa en la extrapolación. Algunos códecs del estado de la técnica aplican enfoques de extrapolación y ejecutan algoritmos específicos para cambiar el tono según las estimaciones de tono extrapoladas durante la pérdida de paquetes. Estos enfoques se describirán con mayor detenimiento a continuación con referencia a G.718 y G.729.1.

Inicialmente se considera G.718 (véase [ITU08a]). Una estimación del tono futuro se lleva a cabo mediante extrapolación para apoyar el módulo de resincronización de pulsos glotales. Esta información acerca del posible valor futuro del tono se utiliza para sincronizar los pulsos glotales de la excitación oculta.

La extrapolación del tono se lleva a cabo solamente si la última trama buena no era SORDA. La extrapolación del tono de G.718 se basa en la suposición de que el codificador tiene un contorno de tono suave. Dicha extrapolación se realiza en base a los retardos de tono $d_{fr}^{[i]}$ de las últimas siete subtramas antes del borrado.

En G.718, se lleva a cabo una actualización de la historia de los valores de tono flotantes después de cada trama recibida correctamente. Para este propósito, los valores de tono se actualizan solamente si el modo de núcleo es distinto de SORDO. En el caso de una trama perdida, la diferencia $\Delta_{dfr}^{[i]}$ entre los retardos de tono flotantes se calcula según la siguiente fórmula:

$$\Delta_{dfr}^{[i]} = d_{fr}^{[i]} - d_{fr}^{[i-1]} \quad \text{para } i = -1, \dots, -6 \quad (1)$$

En la fórmula (1), $d_{fr}^{[-1]}$ designa el retardo de tono de la última (es decir, la cuarta) subtrama de la trama anterior; $d_{fr}^{[-2]}$ designa el retardo de tono de la tercera subtrama de la trama anterior; etc.

Según G.718, la suma de las diferencias $\Delta_{dfr}^{[i]}$ se calcula como:

$$s_{\Delta} = \sum_{i=-6}^{-1} \Delta_{dfr}^{[i]} \quad (2)$$

5 Dado que los valores $\Delta_{dfr}^{[i]}$ pueden ser positivos o negativos, se suma la cantidad de inversiones de signo de $\Delta_{dfr}^{[i]}$ y se indica la posición de la primera inversión mediante un parámetro que se mantiene en memoria.

El parámetro f_{corr} se encuentra mediante la siguiente expresión:

$$f_{corr} = 1 - \frac{\sqrt{\sum_{i=-6}^{-1} (\Delta_{dfr}^{[-i]} - s_{\Delta})^2}}{6 \cdot d_{max}} \quad (3)$$

10 donde $d_{max} = 231$ es el retardo de tono máximo considerado.

En G.718, se encuentra una posición i_{max} , que indica la máxima diferencia absoluta según la siguiente definición:

$$15 \quad i_{max} = \left\{ \max_{i=-6}^{-1} \left(\text{abs} \left(\Delta_{dfr}^{[i]} \right) \right) \right\}$$

y una relación para esta diferencia máxima se calcula como sigue:

$$20 \quad r_{max} = \left| \frac{5 \cdot \Delta_{dfr}^{[i_{max}]}}{(s_{\Delta} - \Delta_{dfr}^{[i_{max}]})} \right| \quad (4)$$

25 Si esta relación es igual o superior a 5, entonces el tono de la cuarta subtrama de la última trama correctamente recibida se utiliza para todas las subtramas a ser ocultas. Si esta relación es superior a 5, esto significa que el algoritmo no es lo suficientemente seguro para extrapolar el tono, y la resincronización de pulso glotal no se efectuará.

30 Si r_{max} es inferior a 5, entonces se lleva a cabo un procesamiento adicional para lograr la mejor extrapolación posible. Se utilizan tres métodos diferentes para extrapolar el tono futuro. Para elegir entre los posibles algoritmos de extrapolación de tono, se calcula un parámetro de desviación f_{corr2} , que depende del factor f_{corr} y de la posición de la máxima variación de tono i_{max} . Sin embargo, inicialmente se modifica la diferencia de tono flotante media a efectos de eliminar diferencias de tono demasiado grandes de la media:

Si $f_{corr} < 0,98$ y si $i_{max} = 3$, entonces la diferencia de tono fraccional media $\bar{\Delta}_{dfr}$ se determina mediante la siguiente fórmula:

$$35 \quad \bar{\Delta}_{dfr} = \left(\frac{s_{\Delta} - \Delta_{dfr}^{[-4]} - \Delta_{dfr}^{[-5]}}{3} \right) \quad (5)$$

para eliminar las diferencias de tono relacionadas con la transición entre dos tramas.

Si $f_{corr} \geq 0,98$ o si $i_{max} \neq 3$, la diferencia de tono fraccional media $\bar{\Delta}_{dfr}$ se calcula como sigue:

$$40 \quad \bar{\Delta}_{dfr} = \frac{s_{\Delta} - \Delta_{dfr}^{[i_{max}]}}{6} \quad (6)$$

y la máxima diferencia de tono flotante se reemplaza con este nuevo valor medio:

$$\Delta_{dfr}^{[i_{max}]} = \overline{\Delta_{dfr}} \quad (7)$$

Con esta nueva media de las diferencias de tono flotantes, la desviación normalizada f_{corr2} se calcula como sigue:

$$f_{corr2} = 1 - \frac{\sqrt{\sum_{i=-1}^{I_{sf}} \left(\Delta_{dfr}^{[i]} - \overline{\Delta_{dfr}} \right)^2}}{I_{sf} \cdot d_{max}} \quad (8)$$

en el que I_{sf} es igual a 4 en el primer caso y es igual a 6 en el segundo caso.

En función de este nuevo parámetro, se efectúa una elección entre los tres métodos para extrapolar el tono futuro:

- Si $\Delta_{dfr}^{[i]}$ cambia de signo más de dos veces (esto indica una elevada variación de tono), la primera inversión de signo se halla en la última trama buena (para $i < 3$), y $f_{corr2} > 0,945$, el tono extrapolado, d_{ext} , (el tono extrapolado también lleva la designación T_{ext}) se calcula como sigue:

$$s_y = \sum_{i=-1}^{-4} \Delta_{dfr}^{[i]}$$

$$s_{xy} = \Delta_{dfr}^{[-2]} + 2 \cdot \Delta_{dfr}^{[-3]} + 3 \cdot \Delta_{dfr}^{[-4]}$$

$$d_{est} = \text{redondeado} \left[\Delta_{dfr}^{[-1]} + \left(\frac{7 \cdot s_y - 3 \cdot s_{xy}}{10} \right) \right]$$

- Si $0,945 < f_{corr2} < 0,99$ y Δ_{dfr}^i cambia de signo por lo menos una vez, la media ponderada de las diferencias fraccionales de tono se emplea para extrapolar el tono. La ponderación, f_w , de la diferencia media está relacionada con la desviación normalizada, f_{corr2} , y la posición de la primera inversión de signo se define como sigue:

$$f_w = f_{corr2} \cdot \left(\frac{i_{mem}}{7} \right)$$

El parámetro i_{mem} de la fórmula depende de la posición de la primera inversión del signo de Δ_{dfr}^i , de manera tal que $i_{mem} = 0$ si la primera inversión del signo tuvo lugar entre las dos últimas subtramas de la trama perdida, de manera tal que $i_{mem} = 1$ si la primera inversión de signo tuvo lugar entre las subtramas segunda y tercera de la trama perdida. Si la primera inversión de signo está cercana al extremo de la última trama buena, esto significa que la variación de tono fue menos estable justo antes de la trama pérdida. Por lo tanto, el factor de ponderación aplicado a la media estará cercano a 0 y el tono extrapolado d_{ext} estará cercano al tono de la cuarta subtrama de la última trama buena:

$$d_{ext} = \text{redondeado} \left[\Delta_{dfr}^{[-1]} + 4 \cdot \overline{\Delta_{dfr}} \cdot f_w \right]$$

- En caso contrario, se considera que la evolución del tono es estable y se determina el tono extrapolado d_{ext} como sigue:

$$d_{ext} = \text{redondeado} \left[d_{dfr}^{[-1]} + 4 \cdot \overline{\Delta_{dfr}} \right]$$

Después de este procesamiento, el retardo de tono está limitado entre 34 y 231 (los valores designan los retardos de tono permitidos máximo y mínimo).

Seguidamente se considera otro ejemplo de técnicas de reconstrucción de tono basadas en extrapolación, G.729.1 (véase [ITU06b]).

- 5 El G.729.1 incluye un enfoque de extrapolación de tono (Véase [Gao]), en el caso en que no pueda decodificarse información directa de ocultamiento de error (por ejemplo, información de fase). Esto sucede si se pierden dos tramas consecutivas (una supertrama consiste en cuatro tramas que pueden ser sea ACELP o TCX20). También hay tramas TCX40 o TCX80 posibles y también todas sus combinaciones.
- 10 Cuando se pierden una o más tramas en una región sonora, la información de tono anterior siempre se utiliza para reconstruir la trama perdida actual. La precisión del tono estimado actual puede influir directamente en la alineación de fase con respecto a la señal original, y es fundamental para la calidad de la reconstrucción de la trama actual perdida y la trama recibida después de la trama perdida. Usando varios retardos de tono en lugar de simplemente copiar el retraso de tono anterior se traduciría en una estimación de tono estadísticamente mejor. En el codificador
- 15 G.729.1, la extrapolación de tono para FEC (FEC = corrección directa de errores) consiste en una extrapolación lineal basado en los últimos cinco valores de tono. Los últimos cinco valores de tono son: $P(i)$, para $i = 0, 1, 2, 3, 4$, en el que $P(4)$ es el valor de tono más reciente. El modelo de extrapolación se define según:

$$P'(i) = a + i \cdot b \quad (9)$$

20

El valor de tono extrapolado para la primera subtrama en una trama perdida se define seguidamente como:

$$P'(5) = a + 5 \cdot b \quad (10)$$

25

Para determinar los coeficientes a y b , se minimiza E , en el que el error E se define según:

$$\begin{aligned} E &= \sum_{i=0}^4 [P'(i) - P(i)]^2 \\ &= \sum_{i=0}^4 [(a + b \cdot i) - P(i)]^2 \end{aligned} \quad (11)$$

Ajustando

30

$$\frac{\delta E}{\delta a} = 0 \quad \text{y} \quad \frac{\delta E}{\delta b} = 0 \quad (12)$$

se obtienen los valores de a y b :

$$a = \frac{3 \sum_{i=0}^4 P(i) - \sum_{i=0}^4 i \cdot P(i)}{5} \quad \text{y} \quad b = \frac{\sum_{i=0}^4 i \cdot P(i) - 2 \sum_{i=0}^4 P(i)}{10} \quad (13)$$

35

A continuación, se describe un concepto de ocultamiento por borrado de trama para el códec AMR-WB como se presenta en [MCZ11]. Este concepto de ocultamiento por borrado de trama se basa en la predicción lineal de tono y ganancia. En dicho documento se propone un enfoque de inter/extrapolación de tono lineal en el caso de una pérdida de trama, basada en un Criterio de Mínimos Cuadrados Medios.

40

Según este concepto de ocultación por borrado de trama, en el decodificador, cuando el tipo de la última trama válida antes de la trama borrada (la trama pasada) es el mismo que para la temprana después de la trama borrada (la trama futura), se define el tono $P(i)$, en el que: $i = -N, -N + 1, \dots, 0, 1, \dots, N + 4, N + 5$, y donde N es el número de subtramas pasadas y futuras de la trama borrada. $P(1), P(2), P(3), P(4)$ son los tonos de cuatro subtramas en la trama borrada, $P(0), P(-1), \dots, P(-N)$ son los tonos de las últimas subtramas, y $P(5), P(6), \dots, P(N + 5)$ son los tonos de las futuras subtramas. Se emplea un modelo de predicción lineal $P'(i) = a + b \cdot i$. Para $i = 1, 2, 3, 4$; $P'(1), P'(2), P'(3), P'(4)$ son los tonos predichos para la trama borrada. El Criterio MMS (MMS = Mínimo Cuadrado Medio) es tomado en cuenta para obtener los valores de los dos coeficientes predichos a y b según un enfoque de

50 interpolación. Según este enfoque, el error E se define como:

$$\begin{aligned}
E &= \sum_{-N}^0 [P'(i) - P(i)]^2 + \sum_5^{N+5} [P'(i) - P(i)]^2 \\
&= \sum_{-N}^0 [a + b \cdot i - P(i)]^2 + \sum_5^{N+5} [a + b \cdot i - P(i)]^2
\end{aligned} \tag{14a}$$

Entonces, los coeficientes a y b pueden obtenerse calculando:

5

$$\frac{\delta E}{\delta a} = 0 \quad \text{y} \quad \frac{\delta E}{\delta b} = 0 \tag{14b}$$

$$a = \frac{2 \left[\sum_{i=-N}^0 P(i) + \sum_{i=5}^{N+5} P(i) \right] \cdot (N^3 + 9N^2 + 38N + 1)}{(N+1) \cdot (4N^3 + 36N^2 + 107N - 1)} \tag{14c}$$

10

$$b = \frac{9 \left[\sum_{i=-N}^0 P(i) + \sum_{i=5}^{N+5} P(i) \right]}{1 - 107N - 36N^2 - 4N^3} \tag{14d}$$

Los retardos de tono para las cuatro últimas subtramas de la trama borrada pueden calcularse según:

$$\begin{aligned}
P'(1) &= a + b \cdot 1; & P'(2) &= a + b \cdot 2 \\
P'(3) &= a + b \cdot 3; & P'(4) &= a + b \cdot 4
\end{aligned} \tag{14e}$$

15

Se ha descubierto que $N = 4$ provee el mejor resultado. $N = 4$ significa que en la interpolación se utilizan cinco subtramas pasadas y cinco subtramas futuras.

20

Sin embargo, cuando el tipo de las tramas pasadas es diferente del tipo de las tramas futuras, por ejemplo, cuando la trama pasada tiene voz, pero la trama futura no tiene voz, se utilizan solamente los tonos con voz de las tramas pasadas o futuras para predecir los tonos de la trama borrada, empleándose el enfoque de extrapolación anteriormente descrito.

25

Seguidamente se considera la resincronización de pulsos en la técnica anterior, en particular con referencia a G.718 y G.729.1. En [VJGS12] se describe un enfoque para la resincronización de pulsos

Se empieza por describir la construcción de la parte periódica de la excitación.

30

Para una ocultación de tramas borradas que siguen a una trama recibida correctamente que no sea SORDA, la parte periódica de la excitación se construye mediante la repetición del último periodo de tono filtrado de bajo pase de la trama anterior.

35

La construcción de la parte periódica se realiza utilizando una copia simple de un segmento filtrado de bajo pase de la señal de excitación tomada en el extremo de la trama anterior.

La longitud del periodo de tono se redondea al número entero más cercano:

$$T_c = \text{redondeo}(\text{último_tono}) \tag{15a}$$

40

Considerando que la longitud del último periodo de tono es T_p , entonces la longitud del segmento que se copia, T_r , puede definirse, por ejemplo, según:

$$T_r = [T_p + 0,5] \tag{15b}$$

45

La parte periódica se construye para una trama y para una trama adicional.

Por ejemplo, con M subtramas en una trama, la longitud de subtrama es: $L_{subtr} = \frac{L}{M}$.

en la que L es la longitud de la trama, también designada como L_{trama} : $L = L_{trama}$.

La figura 3 ilustra una parte periódica construida de una señal de voz.

T [0] es la ubicación del primer pulso máximo en la en la parte periódica construida de la excitación. Las posiciones de los otros pulsos están dadas por:

$$T [i] = T [0] + i T_c \quad (16a)$$

correspondiente a:

$$T [i] = T [0] + i T_r \quad (16b)$$

Después de la construcción de la parte periódica de la excitación, se realiza la resincronización de pulso global para corregir la diferencia entre la posición estimada de destino del último pulso en la trama perdida (P), y su posición real en la parte periódica construida de la excitación (T[k]).

La evolución de retardo de tono se extrapola sobre la base de los retardos de tono de las últimas siete subtramas antes de la trama perdida. Los retardos de tono en cada subtrama son:

$$p [i] = \text{redondeo} (T_c + (i + 1) \delta), 0 \leq i < M \quad (17a)$$

donde

$$\delta = \frac{T_{ext} - T_c}{M} \quad (17b)$$

y T_{ext} (también designado como d_{ext}) es el tono extrapolado como se describió anteriormente para d_{ext} .

La diferencia, designada como d, entre la suma del número total de muestras dentro de ciclos de tono con el tono constante (T_c) y la suma del número total de muestras dentro de ciclos de tono con el tono en evolución, $p[i]$, se encuentra dentro de una longitud de trama. No existe descripción en la documentación acerca de cómo encontrar d.

En el código de fuente de G.718 (Véase [ITU08a]), d se halla utilizando el siguiente algoritmo (en el que M es el número de subtramas en una trama):

```
ftmp = p[0];
i = 1;
while (ftmp < L_frame - pit_min) {
    sect = (short)(ftmp*M/L_frame);
    ftmp += p[sect];
    i++;
}
d = (short)(i*Tc - ftmp);
```

La cantidad de pulsos en la parte periódica construida dentro de una longitud de trama más el primer pulso en la trama futura es N. No hay ninguna descripción en la documentación acerca de cómo N.

En el código de fuente de G.718 (Véase [ITU08a]), se puede encontrar N según lo siguiente:

$$N = 1 + \left\lceil \frac{L_{trama}}{T_c} \right\rceil \quad (18a)$$

La posición del último pulso T[n] en la parte periódica construida de la excitación que forma parte de la trama perdida

se determina mediante lo siguiente:

$$n = \begin{cases} N - 1, & T[N - 1] < L_{trama} \\ N - 2, & T[N - 1] \geq L_{trama} \end{cases} \quad (18b)$$

5 La posición estimada del último pulso P es:

$$P = T[n] + d \quad (19a)$$

10 La posición actual de la posición del último pulso T [k] es la posición del pulso en la parte periódica construida de la excitación (inclusive la búsqueda del primer pulso después de la trama actual) más cercana a la posición estimada teórica P:

$$\forall i \quad |T[k] - P| \leq |T[i] - P|, \quad 0 \leq i < N \quad (19b)$$

15 La resincronización de pulsos glotales se lleva a cabo mediante la adición o retirada de muestras en las regiones de mínima energía de los ciclos de tono completos. La cantidad de muestras a ser añadida o eliminada se determina mediante la siguiente diferencia:

$$diff = P - T[k] \quad (19c)$$

20 Las regiones de energía mínima se determinan mediante una ventana deslizante de cinco muestras. La posición de energía mínima se ajusta en el centro de la ventana en la cual la energía está en un mínimo. La búsqueda se lleva a cabo entre dos pulsos de tono desde $T[i] + T_c / 8$ a $T[i + 1] - T_c / 4$. Hay $N_{min} = n - 1$ regiones de energía mínima.

25 Si $N_{min} = 1$, entonces hay solamente una región de energía mínima y en dicha posición se insertan o suprimen $diff$ muestras.

30 Para $N_{min} > 1$, se añade o retira una cantidad menor de muestras al inicio y más hacia el extremo de la trama. El número de muestras a ser eliminadas o añadidas entre los pulsos T [i] y T [i+1] se halla mediante la siguiente relación recursiva:

$$R[i] = \text{redondeado} \left(\frac{(i+1)^2}{2} f - \sum_{k=0}^{i-1} R[k] \right) \quad \text{con} \quad f = \frac{2|diff|}{N_{min}^2} \quad (19d)$$

35 Si $R[i] < R[i - 1]$, entonces los valores de R [i] y R [i - 1] están intercambiados.

40 El documento US 6 035 271 A da a conocer un método y un aparato para extraer información de valor de tono del habla. El método selecciona al menos tres picos más altos de una función de autocorrelación normalizada y produce una pluralidad de candidatos de frecuencia para la determinación del valor de tono. La pluralidad de candidatos de frecuencia se usa para identificar puntos estratégicos en valores de tono, y se usa además para realizar búsqueda tanto hacia adelante como hacia atrás cuando no se puede identificar fácilmente un punto estratégico.

45 El documento US 2012/072209 A1 da a conocer un dispositivo electrónico para estimar un retardo de tono. El dispositivo electrónico incluye un procesador e instrucciones ejecutables almacenadas en la memoria que es una comunicación electrónica con el procesador. El dispositivo electrónico obtiene una trama actual. El dispositivo electrónico también obtiene una señal residual basada en la trama actual.

50 El documento US 2013/041657 A1 da a conocer que un sistema y método pueden configurarse para analizar una información de audio derivada de una señal de audio. El sistema y método pueden rastrear el tono del sonido a través de la señal de audio. El rastreo de tono a través de la señal de audio puede tener en cuenta el cambio en tono mediante la determinación en tiempo individual ventanas de muestra en la duración de la señal un tono estimado y una representación de envoltura armónica en el tono estimado.

55 El documento US 6 507 814 B1 da a conocer un códec de habla multi-tasa que soporta una pluralidad de modos de tasa de bits de codificación mediante la selección de manera adaptiva de modos de tasa de bits de codificación para coincidir con las restricciones de canal de comunicación. En los modos decodificación de tasa de bits más alta, una

representación de habla precisa a través de CELP (predicción lineal excitada por código) y otros parámetros de modelado asociados se generan para una decodificación y reproducción de mayor calidad.

5 XINWEN MU ET AL, "A frame erasure concealment method based on tone and gain linear prediction for AMRWB
code", CONSUMER ELECTRONICS (ICCE), 2011 IEEE INTERNATIONAL CONFERENCE ON, IEEE, (20110109),
doi:10.1109/ICCE.2011.5722880, ISBN 978-1-4244-8711-0, páginas 815 - 816, da a conocer un método de
ocultación de error para el códec AMR-WB, usando la restricción de búsqueda de excitación en el codificador y una
predicción lineal del tono y ganancias del libro de códigos adaptivos y el libro de códigos innovador en el
10 decodificador para mejorar la calidad del habla tras un borrado de la trama. Los resultados experimentales
demuestran que el método propuesto logra una mejora del rendimiento con respecto a los métodos existentes.

"G.729 based Embedded Variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable con
G.729; G.729.1 (05/06)", ITU-T STANDARD, INTERNATIONAL TELECOMMUNICATION UNION, GENEVA; CH,
15 (20060529), n.º G.729.1 (05/06), páginas 1 - 100, da a conocer un codificador de velocidad de bits, variable, en
particular, 8-32 kbit/s escalable.

El documento US 8 255 207 B2 da a conocer un método y dispositivo para ocultar borrados de trama provocados
por tramas de una señal de sonido codificada borrada durante la transmisión desde un codificador hasta un
20 decodificador y para la recuperación del decodificador tras los borrados de trama que comprenden, en el codificador,
la determinación de los parámetros de ocultación/recuperación incluyendo al menos información de fase relacionada
con las tramas de la señal de sonido codificada. Los parámetros de ocultación/recuperación determinados en el
codificador se transmiten al decodificador y, en el decodificador, la ocultación del borrado de trama se realiza en
respuesta a los parámetros de ocultación/recuperación recibidos.

25 "ITU-T G.718 - Frame error robust narrow-band and wideband embedded variable bit-rate coding de speech and
audio from 8-32 kbit/s", (20080630), URL: <http://www.itu.int/rec/T-REC-G.718-200806-I>, (20131112), da a conocer
codificación de tasa de bits variable empotrada en banda ancha y banda estrecha robusta de error de trama.

30 MOHAMED CHIBANI ET AL, "Fast Recovery for a CELP-Like Habla Codec After a Frame Erasure", IEEE
TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING, IEEE SERVICE CENTER, NUEVA
YORK, NY, USA, (20071101), vol. 15, n.º 8, doi:10.1109/TASL.2007.907332, ISSN 1558-7916, páginas 2485 - 2495,
da a conocer una técnica que se propone para mejorar la recuperación tras un borrado de trama. La técnica consiste
en una búsqueda de excitación restringida en el codificador y un procedimiento de resincronización en el
35 decodificador. La restricción apunta a la reducción de la contribución del libro de códigos adaptivo al hacer que el
libro de códigos de innovación modele parcialmente la excitación del tono. Además, para las tramas altamente
sonoras, la información relacionada con el tono contenida en la excitación de innovación se explota en el
decodificador para acelerar la resincronización del libro de códigos adaptivo después de un borrado de trama.
Cuando se aplica al códec de banda ancha multi-tasas adaptivo (AMR-WB), el método aporta una mejora
40 significativa en el caso de borrado de tramas, a costa de una pérdida de calidad menor en comparación con el códec
estándar a la misma tasa de bits. El método no necesita retardo adicional y tiene la ventaja de mantener una
interoperabilidad completa entre el códec estándar y su versión modificada.

El objeto de la presente invención es proporcionar conceptos mejorados para el procesamiento de las señales de
45 audio, en particular, para proveer conceptos mejorados para el procesamiento del habla, y más particularmente para
proveer conceptos mejorados de ocultación.

El objeto de la presente invención se logra mediante un aparato según la reivindicación 1, mediante un método
según la reivindicación 6 y mediante un programa de ordenador según la reivindicación 7.

50 En un ejemplo, se describe un aparato para la determinación de un retardo de tono estimado. El aparato comprende
una interfaz de entrada para recibir una pluralidad de valores de retardo de tono originales, y un estimador de
retardo de tono para la estimación del retardo de tono estimado. El estimador de retardo de tono se configura para
estimar el retardo de tono estimado en función de una pluralidad de valores de retardo de tono originales y en
función de una pluralidad de valores de información, en el que para cada valor de retardo de tono original de la
55 pluralidad de valores de retardo de tono originales, un valor de información de la pluralidad de valores de
información se asigna a dicho valor de retardo de tono original.

Según un ejemplo, el estimador de retardo de tono puede estar configurado por ejemplo para estimar el retardo de
60 tono estimado en función de la pluralidad de valores de retardo de tono originales y en función de la pluralidad de
valores de ganancia de tono como la pluralidad de valores de información, en el que para cada valor de retardo de
tono original de la pluralidad de valores de retardo de tono originales, un valor de ganancia de tono de la pluralidad
de valores de ganancia de tono se asigna a dicho valor de retardo de tono original.

En un ejemplo particular, cada una de la pluralidad de valores de ganancia de tono puede, por ejemplo, ser una ganancia de libro de códigos adaptivo.

5 En un ejemplo, el estimador de retardo de tono puede estar configurado por ejemplo para estimar el retardo de tono estimado mediante la minimización de una función de error.

Según una realización de la presente invención, el estimador de retardo de tono puede, por ejemplo, configurarse para estimar el retardo de tono estimado mediante la determinación de dos parámetros a , b , mediante la minimización de la función de error

10

$$err = \sum_{i=0}^k g_p(i) \cdot ((a + b \cdot i) - P(i))^2$$

en la que a es un número real, en la que b es un número real, en la que k es un número entero con $k \geq 2$, y en la que $P(i)$ es el valor de retardo de tono original i -ésimo, en la que $g_p(i)$ es el valor de ganancia de tono i -ésimo que se está asignando al valor de retardo de tono i -ésimo $P(i)$, en donde i indica un índice de subtrama.

15

En un ejemplo, el estimador de retardo de tono puede estar configurado por ejemplo para estimar el retardo de tono estimado mediante la determinación de dos parámetros a , b , mediante la minimización de la función de error

20

$$err = \sum_{i=0}^4 g_p(i) \cdot ((a + b \cdot i) - P(i))^2$$

en la que a es un número real, en la que b es un número real, en la que $P(i)$ es el valor de retardo de tono original i -ésimo, en la que $g_p(i)$ es el valor de ganancia de tono i -ésimo que se está asignando al valor de retardo de tono i -ésimo $P(i)$, en donde i indica un índice de subtrama.

25

Según un ejemplo, el estimador de retardo de tono puede estar configurado por ejemplo para determinar el retardo de tono estimado p según $p = a \cdot i + b$.

En un ejemplo, el estimador de retardo de tono puede estar configurado por ejemplo para estimar el retardo de tono estimado en función de la pluralidad de valores de retardo de tono originales y en función de una pluralidad de valores de tiempo como la pluralidad de valores de información, en el que para cada valor de retardo de tono original de la pluralidad de valores de retardo de tono originales, un valor de tiempo de la pluralidad de valores de tiempo se asigna a dicho valor de retardo de tono original.

30

Según un ejemplo, el estimador de retardo de tono puede estar configurado por ejemplo para estimar el retardo de tono estimado mediante la minimización de una función de error.

35

En un ejemplo, el estimador de retardo de tono puede estar configurado por ejemplo para estimar el retardo de tono estimado mediante la determinación de dos parámetros a , b , mediante la minimización de la función de error

40

$$err = \sum_{i=0}^k tiempo_{pasado}(i) \cdot ((a + b \cdot i) - P(i))^2$$

en la que a es un número real, en la que b es un número real, en la que k es un número entero con $k \geq 2$, y en la que $P(i)$ es el valor de retardo de tono original i -ésimo, en la que el $tiempo_{pasado}(i)$ es el valor de tiempo i -ésimo que se está asignando al valor de retardo de tono i -ésimo $P(i)$, en donde i indica un índice de subtrama.

45

Según un ejemplo, el estimador de retardo de tono puede estar configurado por ejemplo para estimar el retardo de tono estimado mediante la determinación de dos parámetros a , b , mediante la minimización de la función de error

50

$$err = \sum_{i=0}^4 tiempo_{pasado}(i) \cdot ((a + b \cdot i) - P(i))^2$$

en la que a es un número real, en la que b es un número real, en la que $P(i)$ es el valor de retardo de tono original i -ésimo, en la que el $tiempo_{pasado}(i)$ es el valor de tiempo i -ésimo que se está asignando al valor de retardo de tono i -ésimo $P(i)$.

En un ejemplo, el estimador de retardo de tono se configura para determinar el retardo de tono estimado p según $p = a \cdot i + b$.

5 Además, se describe un método para la determinación de un retardo de tono estimado según un ejemplo. El método comprende:

- Recibir una pluralidad de valores de retardo de tono originales. Y:

10 - Estimar el retardo de tono estimado.

La estimación del retardo de tono estimado se realiza en función de una pluralidad de valores de retardo de tono originales y en función de una pluralidad de valores de información, en el que para cada valor de retardo de tono original de la pluralidad de valores de retardo de tono originales, un valor de información de la pluralidad de valores de información se asigna a dicho valor de retardo de tono original.

Además, un programa de ordenador para la implementación del método descrito anteriormente cuando se ejecuta en un ordenador o se proporciona un procesador de señales.

20 Además, se proporciona un aparato para la reconstrucción de una trama que comprende una señal de habla como una trama reconstruida, estando dicha trama reconstruida asociada con una o más tramas disponibles, siendo dicha una o más tramas disponibles al menos una de una o más tramas anteriores de la trama reconstruida y una o más tramas posteriores de la trama reconstruida, en el que la una o más tramas disponibles comprenden uno o más ciclos de tono como uno o más ciclos de tono disponibles. El aparato comprende una unidad de determinación para la determinación de una diferencia de número de muestra que indica una diferencia entre un número de muestras de uno de los uno o más ciclos de tono disponibles y un número de muestras de un primer ciclo de tono a ser reconstruido. Además, el aparato comprende un reconstructor de trama para la reconstrucción de la trama reconstruida mediante la reconstrucción, en función de la diferencia de número de muestra y en función de las muestras de dicho uno de los uno o más ciclos de tono disponibles, el primer ciclo de tono a ser reconstruido como un primer ciclo de tono reconstruido. El reconstructor de trama se configura para reconstruir la trama reconstruida, de manera tal que la trama reconstruida comprende completa o parcialmente el primer ciclo de tono reconstruido, de manera tal que la trama reconstruida comprende completa o parcialmente un segundo ciclo de tono reconstruido, y de manera tal que el número de muestras del primer ciclo de tono reconstruido difiere de un número de muestras del segundo ciclo de tono reconstruido.

35 Según un ejemplo, la unidad de determinación puede estar configurado por ejemplo para determinar una diferencia de número de muestra para cada una de una pluralidad de ciclos de tono a ser reconstruido, de manera tal que la diferencia de número de muestra de cada uno de los ciclos de tono indica una diferencia entre el número de muestras de dicho uno de los uno o más ciclos de tono disponibles y un número de muestras de dicho ciclo de tono a ser reconstruido. El reconstructor de trama puede estar configurado por ejemplo para reconstruir cada ciclo de tono de la pluralidad de ciclos de tono a ser reconstruido en función de la diferencia de número de muestra de dicho ciclo de tono a ser reconstruido y en función de las muestras de dicho uno de los uno o más ciclos de tono disponibles, para reconstruir la trama reconstruida.

45 En un ejemplo, el reconstructor de trama puede estar configurado por ejemplo para generar una trama intermedia en función de dicho uno de los del uno o más ciclos de tono disponibles. El reconstructor de trama puede estar configurado por ejemplo para modificar la trama intermedia para obtener la trama reconstruida.

50 Según un ejemplo, la unidad de determinación puede estar configurado por ejemplo para determinar un valor de diferencia de trama (d ; s) indicando cómo deben retirarse muchas muestras de la trama intermedia o cómo deben añadirse muchas muestras a la trama intermedia. Además, el reconstructor de trama puede estar configurado por ejemplo para retirar primeras muestras de la trama intermedia para obtener la trama reconstruida, cuando el valor de diferencia de trama indica que las primeras muestras deben retirarse de la trama. Además, el reconstructor de trama puede estar configurado por ejemplo para añadir segundas muestras a la trama intermedia para obtener la trama reconstruida, cuando el valor de diferencia de trama (d ; s) indica que las segundas muestras deben añadirse a la trama.

60 En un ejemplo, el reconstructor de trama puede estar configurado por ejemplo para retirar las primeras muestras de la trama intermedia cuando el valor de diferencia de trama indica que las primeras muestras deben retirarse de la trama, de manera tal que el número de primeras muestras que se retiran de la trama intermedia se indica por el valor de diferencia de trama. Además, el reconstructor de trama puede estar configurado por ejemplo para añadir las segundas muestras a la trama intermedia cuando el valor de diferencia de trama indica que las segundas muestras deben añadirse a la trama, de manera tal que el número de segundas muestras que se añaden a la trama intermedia se indica por el valor de diferencia de trama.

Según un ejemplo, la unidad de determinación puede estar configurada por ejemplo para determinar el número de diferencia de trama s de manera tal que la fórmula:

$$s = \sum_{i=0}^{M-1} (p[i] - T_r) \frac{L}{MT_r}$$

5

Se mantiene cierto, en el que L indica un número de muestras de la trama reconstruida, en la que M indica un número de subtramas de la trama reconstruida, en la que T_r indica una longitud de periodo de tono redondeado de dicho uno de los uno o más ciclos de tono disponibles, y en la que $p[i]$ indica una longitud de periodo de tono de un ciclo de tono reconstruido de la subtrama i -ésima de la trama reconstruida.

10

En un ejemplo, el reconstructor de trama puede, por ejemplo, adaptarse para generar una trama intermedia en función de dicho uno de los uno o más ciclos de tono disponibles. Además, el reconstructor de trama puede, por ejemplo, adaptarse para generar la trama intermedia de manera tal que la trama intermedia comprende un primer ciclo de tono intermedio parcial, uno o más ciclos de tono intermedios adicionales, y un segundo ciclo de tono intermedio parcial. Además, el primer ciclo de tono intermedio parcial puede, por ejemplo, depender de una o más de las muestras de dicho uno de los uno o más ciclos de tono disponibles, en el que cada uno de los uno o más ciclos de tono intermedios adicionales depende de todas de las muestras de dicho uno de los uno o más ciclos de tono disponibles, y en el que el segundo ciclo de tono intermedio parcial depende de una o más de las muestras de dicho uno de los uno o más ciclos de tono disponibles. Además, la unidad de determinación puede estar configurado por ejemplo para determinar un número de diferencia de porción de inicio que indica cómo muchas muestras deben retirarse o añadirse del primer ciclo de tono intermedio parcial, y en el que el reconstructor de trama se configura para retirar una o más primeras muestras del primer ciclo de tono intermedio parcial, o se configura para añadir una o más primeras muestras al primer ciclo de tono intermedio parcial en función del número de diferencia de porción de inicio. Además, la unidad de determinación puede estar configurado por ejemplo para determinar para cada uno de los ciclos de tono intermedios adicionales un número de diferencia de ciclo de tono que indica cómo muchas muestras deben retirarse o añadirse de dicho uno de los ciclos de tono intermedios adicionales. Además, el reconstructor de trama puede estar configurado por ejemplo para retirar una o más segundas muestras de dicho uno de los ciclos de tono intermedios adicionales, o se configura para añadir una o más segundas muestras a dicho uno de los ciclos de tono intermedios adicionales en función de dicho número de diferencia de ciclo de tono. Además, la unidad de determinación puede estar configurado por ejemplo para determinar un número de diferencia de porción final que indica cómo muchas muestras deben retirarse o añadirse del segundo ciclo de tono intermedio parcial, y en el que el reconstructor de trama se configura para retirar una o más terceras muestras del segundo ciclo de tono intermedio parcial, o se configura para añadir una o más terceras muestras al segundo ciclo de tono intermedio parcial en función del número de diferencia de porción final.

15

20

25

30

35

Según un ejemplo, el reconstructor de trama puede estar configurado por ejemplo para generar una trama intermedia en función de dicho uno de los del uno o más ciclos de tono disponibles. Además, la unidad de determinación puede, por ejemplo, adaptarse para determinar una o más porciones de señal de baja energía de la señal de habla comprendida por la trama intermedia, en el que cada una de la una o más porciones de señal de baja energía es una primera porción de señal de la señal de habla dentro de la trama intermedia, cuando la energía de la señal de habla es menor que en una segunda porción de señal de la señal de habla comprendida por la trama intermedia. Además, el reconstructor de trama puede estar configurado por ejemplo para retirar una o más muestras de al menos una de la una o más porciones de señal de baja energía de la señal de habla, o para añadir una o más muestras a al menos una de la una o más porciones de señal de baja energía de la señal de habla, para obtener la trama reconstruida.

40

45

En un ejemplo particular, el reconstructor de trama puede estar configurado por ejemplo para generar la trama intermedia, de manera tal que la trama intermedia comprende uno o más ciclos de tono reconstruidos, de manera tal que cada uno de los uno o más ciclos de tono reconstruidos depende de dicho uno de los del uno o más ciclos de tono disponibles. Además, la unidad de determinación puede estar configurado por ejemplo para determinar un número de muestras que debe retirarse de cada uno de los uno o más ciclos de tono reconstruidos. Además, la unidad de determinación puede estar configurado por ejemplo para determinar cada uno de la una o más porciones de señal de baja energía de manera tal que para cada una de la una o más porciones de señal de baja energía un número de muestras de dicha porción de señal de baja energía depende del número de muestras que debe retirarse del uno de los uno o más ciclos de tono reconstruidos, en el que dicha porción de señal de baja energía se ubica dentro de dicho uno de los uno o más ciclos de tono reconstruidos.

50

55

En un ejemplo, la unidad de determinación puede estar configurada por ejemplo para determinar una posición de una o más pulsos de la señal de habla de la trama a ser reconstruida como trama reconstruida. Además, el

60

reconstructor de trama puede estar configurado por ejemplo para reconstruir la trama reconstruida en función de la posición del uno o más pulsos de la señal de habla.

- 5 Según un ejemplo, la unidad de determinación puede estar configurado por ejemplo para determinar una posición de dos o más pulsos de la señal de habla de la trama a ser reconstruida como trama reconstruida, en el que $T[0]$ es la posición de uno de los dos o más pulsos de la señal de habla de la trama a ser reconstruida como trama reconstruida, y en el que la unidad de determinación se configura para determinar la posición ($T[i]$) de pulsos adicionales de los dos o más pulsos de la señal de habla según la fórmula:

10
$$T[i] = T[0] + i T_r$$

en la que T_r indica una longitud redondeada de dicho uno de los uno o más ciclos de tono disponibles, y en la que i es un número entero.

- 15 Según un ejemplo, la unidad de determinación puede estar configurado por ejemplo para determinar un índice k del último pulso de la señal de habla de la trama a ser reconstruida como la trama reconstruida de manera tal que

$$k = \left\lfloor \frac{L - s - T[0]}{T_r} - 1 \right\rfloor$$

- 20 en la que L indica un número de muestras de la trama reconstruida, en la que s indica el valor de diferencia de trama, en la que $T[0]$ indica una posición de un pulso de la señal de habla de la trama a ser reconstruida como la trama reconstruida, que es diferente del último pulso de la señal de habla, y en la que T_r indica una longitud redondeada de dicho uno de los uno o más ciclos de tono disponibles.

- 25 En un ejemplo, la unidad de determinación puede estar configurado por ejemplo para reconstruir la trama a ser reconstruida como la trama reconstruida mediante la determinación de un parámetro δ , en el que δ se define según la fórmula:

$$\delta = \frac{T_{ext} - T_p}{M}$$

- 30 en la que la trama a ser reconstruida como la trama reconstruida comprende M subtramas, en la que T_p indica la longitud de dicho uno de los uno o más ciclos de tono disponibles, y en la que T_{ext} indica una longitud de uno de los ciclos de tono a ser reconstruido de la trama a ser reconstruida como la trama reconstruida.

- 35 Según un ejemplo, la unidad de determinación puede estar configurado por ejemplo para reconstruir la trama reconstruida mediante la determinación de una longitud redondeada T_r de dicho uno de los uno o más ciclos de tono disponibles en base a la fórmula:

$$T_r = \lfloor T_p + 0.5 \rfloor$$

- 40 en la que T_p indica la longitud de dicho uno de los uno o más ciclos de tono disponibles.

En un ejemplo, la unidad de determinación puede estar configurado por ejemplo para reconstruir la trama reconstruida mediante la aplicación de la fórmula:

45
$$s = \delta \frac{L M + 1}{T_r} - L \left(1 - \frac{T_p}{T_r} \right)$$

- 50 en la que T_p indica la longitud de dicho uno de los uno o más ciclos de tono disponibles, en la que T_r indica una longitud redondeada de dicho uno de los uno o más ciclos de tono disponibles, en la que la trama a ser reconstruida como la trama reconstruida comprende M subtramas, en la que la trama a ser reconstruida como la trama reconstruida comprende L muestras, y en la que δ es un número real que indica una diferencia entre un número de muestras de dicho uno de los uno o más ciclos de tono disponibles y un número de muestras de uno de uno o más ciclos de tono a ser reconstruido.

Además, se proporciona un método para la reconstrucción de una trama que comprende una señal de habla como una trama reconstruida, estando dicha trama reconstruida asociada con una o más tramas disponibles, siendo dicha una o más tramas disponibles al menos una de una o más tramas anteriores de la trama reconstruida y una o más tramas posteriores de la trama reconstruida, en el que la una o más tramas disponibles comprenden uno o más ciclos de tono como uno o más ciclos de tono disponibles. El método comprende:

5 - Determinar una diferencia de número de muestra $(\Delta_0^p; \Delta_i; \Delta_{k+1}^p)$ que indica una diferencia entre un número de muestras de uno de los uno o más ciclos de tono disponibles y un número de muestras de un primer ciclo de tono a ser reconstruido. Y:

10 - Reconstruir la trama reconstruida mediante la reconstrucción, en función de la diferencia de número de muestra $(\Delta_0^p; \Delta_i; \Delta_{k+1}^p)$ y en función de las muestras de dicho uno de los uno o más ciclos de tono disponibles, el primer ciclo de tono a ser reconstruido como un primer ciclo de tono reconstruido.

15 La reconstrucción de la trama reconstruida se realiza, de manera tal que la trama reconstruida comprende completa o parcialmente el primer ciclo de tono reconstruido, de manera tal que la trama reconstruida comprende completa o parcialmente un segundo ciclo de tono reconstruido, y de manera tal que el número de muestras del primer ciclo de tono reconstruido difiere de un número de muestras del segundo ciclo de tono reconstruido.

20 Además, un programa de ordenador para implementar el método descrito anteriormente cuando se ejecuta en un ordenador o se proporciona un procesador de señales.

Además, se proporciona un sistema para la reconstrucción de una trama que comprende una señal de habla. El sistema comprende un aparato para la determinación de un retardo de tono estimado según uno de los ejemplos descritos anteriormente o descritos a continuación, y un aparato para la reconstrucción de la trama, en el que el aparato para la reconstrucción de la trama se configura para reconstruir la trama en función del retardo de tono estimado. El retardo de tono estimado es un retardo de tono de la señal de habla.

25 En un ejemplo, la trama reconstruida puede, por ejemplo, estar asociada con una o más tramas disponibles, siendo dicha una o más tramas disponibles al menos una de una o más tramas anteriores de la trama reconstruida y una o más tramas posteriores de la trama reconstruida, en el que la una o más tramas disponibles comprenden uno o más ciclos de tono como uno o más ciclos de tono disponibles. El aparato para la reconstrucción de la trama puede, por ejemplo, ser un aparato para la reconstrucción de una trama según uno de los ejemplos descritos anteriormente o descritos a continuación.

30 La presente invención se basa en el hallazgo de que la técnica anterior tiene inconvenientes significativos. Tanto G.718 (véase [ITU08a]) como G.729.1 (véase [ITU06b]) usan extrapolación de tono en el caso de una pérdida de trama. Esto es necesario, debido a que en el caso de una pérdida de trama, también se pierden los retardos de tono. Según G.718 y G.729.1, el tono se extrapola teniendo en cuenta la evolución del tono durante las últimas dos tramas. Sin embargo, el retardo de tono que se reconstruye por G.718 y G.729.1 no es muy preciso y, por ejemplo, a menudo da como resultado un retardo de tono reconstruido que difiere de manera significativa del retardo de tono real.

35 Las realizaciones de la presente invención proporcionan una reconstrucción de retardo de tono más precisa. Para este fin, en contraste a G.718 y G.729.1, algunos ejemplos tienen en cuenta la información en la confiabilidad de la información del tono.

40 Según la técnica anterior, la información del tono en el que se basa la extrapolación comprende los últimos ocho retardos de tono recibidos correctamente, para el cual el modo de codificación fue diferente de SORDA. Sin embargo, en la técnica anterior, la característica de voz podría ser bastante débil, indicada por una baja ganancia de tono (que corresponde a una baja ganancia de predicción). En la técnica anterior, en el caso de la extrapolación se basa en los retardos de tono que tienen diferentes ganancias de tono, la extrapolación no podrá generar resultados razonables o incluso fallar en absoluto y recurrirá a un simple enfoque de repetición de retardo de tono.

45 Ejemplos se basan en el hallazgo de que la razón para estas limitaciones de la técnica anterior es que en el lado del codificador, el retardo de tono se elige con respecto a maximizar la ganancia de tono con el fin de maximizar la ganancia de codificación del libro de códigos adaptivo, pero que, en el caso de que la característica del habla es débil, el retardo de tono no podría indicar la frecuencia fundamental de manera precisa, puesto que el ruido en la señal de habla provoca la estimación del retardo de tono para volverse imprecisa.

50 Por tanto, durante la ocultación, según los ejemplos, la aplicación de la extrapolación de retardo de tono se pondera en función de la confiabilidad de los retardos recibidos anteriormente usados para esta extrapolación.

Según algunos ejemplos, puede emplearse la ganancia pasada del libro de códigos adaptivos (ganancia de tono) como una medida de confiabilidad.

5 Según algunos ejemplos adicionales ponderación según qué tan lejos en el pasado, se recibieron los retardos de tono, se usa como una medida de confiabilidad. Por ejemplo, se colocan pesos altos para retardos más recientes y se colocan menos pesos para retardos que se recibieron hace más tiempo.

10 Según ejemplos, se proporcionan conceptos de predicción de tono ponderado. A diferencia de la técnica anterior, la predicción de tono proporcionado de la presente invención usa una medida de para cada uno de los retardos de tono se basa en, hacer el resultado de predicción mucho más válido y estable. Particularmente, según la invención, la ganancia de tono se usa como un indicador para la confiabilidad. Alternativa o adicionalmente, según algunos ejemplos útiles para entender la invención, el tiempo que se ha pasado tras la recepción correcta del retardo de tono puede, por ejemplo, usarse como un indicador.

15 En cuanto a la resincronización de pulsos, una de las limitaciones de la técnica anterior en cuanto a la resincronización de pulsos glotales es, que la extrapolación del tono no se tiene en cuenta, cuántos pulsos (ciclos de tono) se deben construir en la trama ocultada.

20 Según la técnica anterior, se realiza la extrapolación del tono de manera tal que solo se esperan los cambios en el tono en los bordes de las subtramas.

Según los ejemplos, cuando se realiza la resincronización de pulsos glotales, se pueden tener en cuenta los cambios de tono que son diferentes de los cambios de tono continuos.

25 G.718 y G.729.1 tienen los siguientes inconvenientes:

30 En primer lugar, en la técnica anterior, cuando se calcula d , se supone que hay un número entero de ciclos de tono dentro de la trama. Dado que d define la ubicación del último pulso en la trama oculta, la posición de la última trama no será correcta, cuando exista un número no entero de ciclos de tono dentro de la trama. Esto se describe en las figuras 6 y 7. La figura 6 ilustra una señal de voz antes de una retirada de las muestras. La figura 7 ilustra la señal de voz después de la retirada de las muestras. Por otra parte, el algoritmo utilizado por la técnica anterior para el cálculo de d es ineficiente.

35 Por otra parte, el cálculo de la técnica anterior requiere la cantidad de pulsos N en la parte periódica construida de la excitación. Esto agrega una complejidad de cálculo, no necesaria.

Además, en la técnica anterior, el cálculo de la cantidad de pulsos N en la parte periódica construida de la excitación no tiene en cuenta la ubicación del primer pulso.

40 Las señales presentadas en las figuras 4 y 5 tienen el mismo período de tono de longitud T_c .

La figura 4 ilustra una señal de voz que tiene tres pulsos dentro de una trama.

45 En cambio, la figura 5 ilustra una señal de voz que tiene solamente dos pulsos dentro de una trama.

Estos ejemplos ilustrados por las figuras 4 y 5 muestran que la cantidad de pulsos es función de la posición del primer pulso.

50 Por otra parte, según la técnica anterior, se verifica, si $T[N-1]$, que en la ubicación del pulso $N^{\text{ésimo}}$ en la parte periódica construida de la excitación se halla dentro de la longitud de la trama, aún si por definición N incluye el primer pulso en la trama siguiente.

55 Por otra parte, según la técnica anterior, no se añaden ni se retiran muestras antes del primer pulso y después del último pulso. Esto conduce al inconveniente de que podría haber un súbito cambio en la longitud del primer ciclo de tono completo, y además, esto conduce además al inconveniente de que la longitud del ciclo de tono después de último pulso podría ser mayor que la longitud del último ciclo de tono completo antes del último pulso, aun cuando el retardo de tono está disminuyendo (véanse las figuras 6 y 7).

60 Los ejemplos están basados en el hallazgo de que los pulsos $T[k] = P - diff$ y $T[n] = P - d$ no son iguales cuando:

- $d > \left\lceil \frac{T_c}{2} \right\rceil$. En este caso $diff = T_c - d$ y la cantidad de muestras eliminadas será $diff$ en lugar de d .

- $T[k]$ se halla en la trama futura y se lo lleva a la trama actual solamente después de eliminar d muestras.

- $T[n]$ es movido a la trama futura después de añadir $-d$ muestras ($d < 0$).

Esto conducirá a una posición errónea de los pulsos en la trama oculta.

5

Además, los ejemplos están basados en el hallazgo de que en la técnica anterior el valor máximo de d se limita al mínimo valor permitido para el retardo de tono codificado. Esto es una restricción que limita la presentación de otros problemas, pero también limita los cambios posibles en el tono y por lo tanto limita la resincronización de pulsos.

10

Por otra parte, los ejemplos están basados en el hallazgo de que, en la técnica anterior, la parte periódica se construye utilizando un retardo de tono número entero, y esto crea un corrimiento de frecuencia de las armónicas y una significativa degradación en la ocultación de las señales tonales con un tono constante. Esta degradación puede observarse en la figura 8, en el que la figura 8 ilustra una representación en tiempo y frecuencia de una señal de voz que se está resincronizando cuando se utiliza un retardo de tono redondeado.

15

Por otra parte, los ejemplos están basados en el hallazgo de que la mayoría de los problemas de la técnica anterior se presenta en situaciones ilustradas en los ejemplos descritos en las figuras 6 y 7, en el que se retiran d muestras.

Aquí se tiene en cuenta que no hay restricción acerca del valor máximo de d , para que el problema sea más fácilmente visible. El problema también se presenta cuando hay un límite para d , pero no es tan obviamente visible.

20

En lugar de seguir incrementando el tono, se obtendrá un incremento súbito seguido por una súbita disminución del tono. Los ejemplos están basados en el hallazgo de que esto sucede por cuanto no se retiran muestras antes y después del último puso, indirectamente también causado por el hecho de no tomarse en cuenta que el pulso se mueve dentro de la trama después de la retirada de d muestras. En este ejemplo también hay un cálculo erróneo de N .

25

Según los ejemplos, se proporcionan conceptos mejorados para la resincronización de pulsos. Los ejemplos proporcionan la ocultación de señales monofónicas, que incluyen el habla, y que se comparan de manera ventajosa con las técnicas existentes descritas en las G.718 (Véase [ITU08a]) y G.729.1 (Véase [ITU06b]). Los ejemplos proporcionados son adecuados para señales con un tono constante, como también para señales con un tono cambiante.

30

Entre otras, según los ejemplos, se proporcionan tres técnicas:

35

Según una primera técnica proporcionada por un ejemplo, se proporciona un concepto de búsqueda de pulsos, por el hecho de que a diferencia con G.718 y G.729.1, se toma en cuenta la ubicación del primer pulso en el cálculo del número de pulsos en la parte periódica construida, designado como N .

40

Según una segunda técnica proporcionada por otro ejemplo, se proporciona un algoritmo para buscar pulsos por el hecho de que, a diferencia con G.718 y G.729.1, no hay necesidad de tener en cuenta el número de pulsos en la parte periódica construida, designado como N , que tenga en cuenta la ubicación del primer pulso, y que de manera indirecta calcula el último índice de pulso en la trama oculta, designado como k .

45

Según una tercera técnica proporcionada por otro ejemplo, no es necesaria una búsqueda de pulsos. Según esta tercera técnica, se combina una construcción de la parte periódica con la retirada o adición de las muestras, con lo cual se logra una menor complejidad que las técnicas anteriores.

50

Adicional o alternativamente, algunos ejemplos proporcionan los siguientes cambios a las técnicas anteriormente descritas, así como para las técnicas de G.718 y G.729.1:

- La parte fraccional del retardo de tono puede utilizarse por ejemplo para construir la parte periódica para señales con un tono constante.

55

- El corrimiento de la ubicación prevista del último pulso en la trama oculta puede calcularse por ejemplo para un número no número entero de ciclos de trama dentro de la trama.

- Las muestras pueden ser añadidas o retiradas también antes del primer pulso y después del último pulso.

- Las muestras pueden ser también añadidas o retiradas si hay solamente un solo pulso.

60

- El número de muestras a ser eliminadas o añadidas puede variar por ejemplo linealmente, siguiendo el cambio lineal predicho en el tono.

A continuación, se describen realizaciones de la presente invención y ejemplos adicionales útiles para entender la invención con mayor detalle con referencia a las figuras, en las que:

- La figura 1 ilustra un aparato para determinar un retardo de tono estimado según un ejemplo;
- 5 la figura 2a ilustra un aparato para reconstruir una trama que comprende una señal de voz en forma de una trama reconstruida según un ejemplo;
- la figura 2b ilustra una señal de voz que comprende una pluralidad de pulso;
- la figura 2c ilustra un sistema para reconstruir una trama que comprende una señal de voz según un ejemplo;
- 10 la figura 3 ilustra una parte periódica construida de una señal de voz;
- la figura 4 ilustra una señal de voz que tiene tres pulsos dentro de una trama;
- 15 la figura 5 ilustra una señal de voz que tiene dos pulsos dentro de una trama;
- la figura 6 ilustra una señal de voz antes de una retirada de muestras;
- la figura 7 ilustra la señal de voz de la figura 6 después de la retirada de muestras;
- 20 la figura 8 ilustra una representación en tiempo y frecuencia de una señal de voz que se está resincronizando mediante la utilización de un retardo de tono redondeado;
- la figura 9 ilustra una representación en tiempo y frecuencia de una señal de voz que se está resincronizando utilizando un retardo de tono no redondeado con la parte fraccional;
- 25 la figura 10 ilustra un diagrama de retardo de tono, en el que el retardo de tono se reconstruye empleando conceptos del estado de la técnica ;
- 30 la figura 11 ilustra un diagrama de retardo de todo, en el que el retardo de tono se reconstruye según un ejemplo;
- la figura 12 ilustra una señal de voz antes de la retirada de las muestras ; y
- la figura 13 ilustra la señal de voz de la figura 12; que ilustra adicionalmente Δ_0 a Δ_3 .
- 35

La figura 1 ilustra un aparato para determinar un retardo de tono estimado según un ejemplo. El aparato comprende una interfaz 110 de entrada para recibir una pluralidad de valores de retardo de tono originales, y un estimador 120 del retardo de tono para estimar el retardo de tono estimado. El estimador 120 del retardo de tono está configurado para estimar el retardo de tono estimado en función de una pluralidad de valores de retardo de tono originales y en función de una pluralidad de valores de ganancia de información, en el que, para cada valor de retardo de tono original de la pluralidad de valores de retardo de tono originales, se asigna un valor de información de la pluralidad de valores de información a dicho un valor de retardo de tono original.

40

Según un ejemplo, el estimador 120 de retardo de tono puede estar configurado por ejemplo para estimar el retardo de tono estimado en función de la pluralidad de valores de retardo de tiempo originales y en función de una pluralidad de valores de ganancia de tono en calidad de pluralidad de valores de información, en el que para cada valor de tono de retardo de tono original de la pluralidad de valores de retardo de tono originales, se asigna un valor de ganancia de la pluralidad de valores de ganancia de tono a dicho un valor de retardo de tono original.

45

En un ejemplo particular, cada uno de entre la pluralidad de valores de ganancia de tono puede ser por ejemplo una ganancia del libro de código adaptativo.

50

En un ejemplo, el estimador 120 del retardo de tono puede estar configurado para por ejemplo estimar el retardo de tono estimado mediante la minimización de una función de error.

55

Según una realización de la invención, el estimador 120 de retardo de tono se configura para estimar el retardo de tono estimado mediante la determinación de dos parámetros a , b , mediante la minimización de la función de error

60

$$err = \sum_{i=0}^k g_p(i) \cdot ((a + b \cdot i) - P(i))^2$$

en la que a es un número real, en la que b es un número real, en la que k es un número entero con $k \geq 2$, y en la que $P(i)$ es el valor de retardo de tono original i -ésimo, en la que $g_p(i)$ es el valor de ganancia de tono i -ésimo que se está asignando al valor de retardo de tono i -ésimo $P(i)$.

60

En un ejemplo según la invención, el estimador 120 del retardo de tono puede estar configurado por ejemplo para estimar el retardo de tono estimado mediante la determinación de dos parámetros, a, b, mediante la minimización de la siguiente función de error:

$$err = \sum_{i=0}^4 g_p(i) \cdot ((a + b \cdot i) - P(i))^2$$

5

en la que a es un número real, en la que b es un número real, en la que P(i) es el valor de retardo de tono original i-ésimo, en la que $g_p(i)$ es el valor de ganancia de tono i-ésimo que se está asignando al valor de retardo de tono i-ésimo P(i).

10

Según un ejemplo, el estimador 120 del retardo de tono puede estar configurado, por ejemplo, para determinar el retardo de tono estimado p según $p = a \cdot i + b$.

En un ejemplo, el estimador 120 del retardo de tono puede estar configurado por ejemplo para estimar el retardo de tono estimado en función de la pluralidad de valores de retardo de tono originales y en función de una pluralidad de valores de tiempo como la pluralidad de valores de información, en el que para cada valor de retardo de tono original de la pluralidad de valores de tono originales, se asigna un valor de tiempo de la pluralidad del valor del tiempo a dicho un valor de retardo de tono original.

20 Según un ejemplo, el estimador 120 del retardo de tono puede estar configurado para estimar el retardo de tono estimado mediante la minimización de una función de error.

En un ejemplo, el estimador 120 del retardo de tono puede estar configurado por ejemplo para estimar el retardo de tono estimado mediante la determinación de dos parámetros a, b, en base a la minimización de la siguiente función de error:

25

$$err = \sum_{i=0}^k \text{tiempo pasado}(i) \cdot ((a + b \cdot i) - P(i))^2$$

30 en la que a es un número real, en la que b es un número real, en la que k es un número entero con un valor $k \geq 2$, y en la que P(i) es el valor de retardo de tono original i-ésimo, en la que $\text{tiempo pasado}(i)$ es el valor de tiempo i-ésimo que se asigna al valor de retardo de tono i-ésimo P(i).

Según un ejemplo, el estimador 120 del retardo de tono puede estar configurado por ejemplo para estimar el retardo de tono estimado mediante la determinación de dos parámetros a, b, por medio de la minimización de la siguiente función de error:

35

$$err = \sum_{i=0}^4 \text{tiempo pasado}(i) \cdot ((a + b \cdot i) - P(i))^2$$

40 en la que a es un número real, en la que b es un número real, en la que P(i) es el valor de retardo de tono original i-ésimo, en la que el $\text{tiempo pasado}(i)$ es el valor de tiempo i-ésimo que se asigna al valor de retardo de tono i-ésimo P(i).

En un ejemplo, el estimador 120 del retardo de tono está configurado para determinar el retardo de tono estimado p según $p = a \cdot i + b$.

45 A continuación se describen ejemplos que proporcionan una predicción ponderada del tono, haciéndose referencia a las fórmulas (20) a (24b).

En primer lugar, se describen ejemplos de predicción de tono ponderados en las que se emplea la ponderación según la ganancia de tono, haciéndose referencia a las fórmulas (20) a (22c). Según algunos de estos ejemplos, para superar el inconveniente de la técnica interior, los retardos de tono se ponderan con la ganancia de tono a efectos de llevar a cabo la predicción de tono.

50

En algunos ejemplos, la ganancia de tono puede ser la ganancia del libro de código adaptativo g_p como se define en la norma G.729 (Véase [ITU12], en particular en el capítulo 3.7.3, más particularmente en la fórmula (43)). En G.729,

la ganancia del libro de código adaptativo se determina según:

$$g_p = \frac{\sum_{n=0}^{39} x(n) y(n)}{\sum_{n=0}^{39} y(n) y(n)}$$

con la condición de que $0 \leq g_p \leq 1,2$

5 En este caso, $x(n)$ es la señal objetivo y se obtiene $y(n)$ mediante la convolución de $v(n)$ con $h(n)$ según:

$$x(n) = \sum_{i=0}^n v(i)h(n-i) \quad n = 0; \dots, 39$$

10 en el que $v(n)$ es el vector del libro del código adaptativo, y en el que $y(n)$ es el vector del libro de código adaptativo filtrado, y en el que $h(n-i)$ es una respuesta de impulso de un filtro de síntesis ponderado, como se define en G.729 (Véase [ITU12]).

De manera similar, en algunos ejemplos, la ganancia de tono puede ser la ganancia del libro del código adaptativo g_p como se define en la norma G.718 (Véase [ITU08a], en particular en el capítulo 6.8.4.1.4.1, más particularmente en la fórmula (170)). En G.718, la ganancia del libro del código adaptativo se determina según:

$$C_{CL} = \frac{\sum_{n=0}^{63} x(n) y_k(n)}{\sum_{n=0}^{63} y_k(n) y_k(n)}$$

20 en el que $x(n)$ es la señal objetivo y $y_k(n)$ es la excitación filtrada anterior en el retardo k .

Por ejemplo, véase [ITU08a], capítulo 6.8.4.1.4.1, fórmula (171), para una definición de cómo podría definirse $y_k(n)$.

25 Similarmente, en algunos ejemplos, la ganancia de tono puede ser la ganancia del libro de código adaptativo g_p definida en la norma AMR (Véase [3GP12b]), en el que la ganancia del libro de código adaptativo g_p como tono de ganancia se define según:

$$g_p = \frac{\sum_{n=0}^{63} x(n) y(n)}{\sum_{n=0}^{63} y(n) y(n)}$$

delimitado por $0 \leq g_p \leq 1,2$

30 en el que $y(n)$ es un vector del libro del código adaptativo filtrado.

En algunos ejemplos particulares, los retardos de tono pueden ser por ejemplo ponderados con la ganancia de tono, por ejemplo, antes de llevarse a cabo la predicción del tono.

35 Para esta finalidad, según un ejemplo es posible introducir una memoria intermedia de longitud 8 que contenga las ganancias de tono, que son tomadas en las mismas subtramas que los retardos de tono. En un ejemplo, la memoria intermedia puede por ejemplo utilizado ser actualizada utilizando las mismas reglas exactas que la actualización de los retardos de tono. Una posible realización consiste en actualizar ambas memorias intermedias (que contienen retardos de tono y ganancias de tono de las últimas ocho subtramas) al final de cada trama, independientemente de
40 si esta trama estaba exenta de errores o era propensa a errores.

Hay dos estrategias de predicción diferentes conocidas de la técnica anterior, que pueden ser reforzadas para utilizar la predicción ponderada de tono:

Algunos ejemplos proporcionan mejoras inventivas significativas de la estrategia de predicción del patrón de G.718. En el G.718, en el caso de una pérdida de paquete, las memorias intermedias pueden ser multiplicadas entre sí o de a elementos, a efectos de ponderar el retardo de tono con un elevado factor si la ganancia de tono asociada es elevada, y para ponderarlo con un factor bajo si la ganancia de tono asociada es baja. Después de esto, según G.718, se lleva a cabo la predicción de tono de la manera usual (Véase [ITU08a, capítulo 7.11.1.3] para mayores detalles para G.718).

Algunos ejemplos proporcionan mejoras inventivas significativas de la estrategia de predicción del patrón de G.729.1. El algoritmo utilizado en G.729.1 para predecir el tono (Véase [ITU06b] por detalles acerca de G.729.1) se modifica según los efectos a efectos de utilizar la predicción ponderada.

Según algunos ejemplos según la invención, el objetivo consiste es minimizar la siguiente función de error:

$$err = \sum_{i=0}^4 g_p(i) \cdot ((a + b \cdot i) - P(i))^2 \quad (20)$$

en el que $g_p(i)$ contiene las ganancias de tono de subtramas pasadas y $P(i)$ contiene los correspondientes retardos de tono.

En la fórmula de la invención (20), $g_p(i)$ representa el factor de ponderación. En el ejemplo anterior, cada $g_p(i)$ representa una ganancia de tono de una de las subtramas anteriores.

Seguidamente, se proporcionan ecuaciones según los ejemplos, que describen cómo derivar los factores a y b, que podrían utilizarse para predecir el retardo de tono según: $a + i \cdot b$, en el que i es el número de subtramas de la subtrama a ser predicha.

Por ejemplo, para obtener la primera subtrama predicha sobre la base de la predicción de las últimas cinco subtramas $P(0)$, ..., $P(4)$, el valor de tono predicho $P(5)$ sería:

$$P(5) = a + 5 \cdot b .$$

A efectos de derivar los coeficientes a y b, la función de error puede derivarse por ejemplo y puede ajustarse en cero:

$$\frac{\delta err}{\delta a} = 0 \quad y \quad \frac{\delta err}{\delta b} = 0 \quad (21a)$$

La técnica anterior que no da a conocer emplear la ponderación inventiva proporcionada. En particular, la técnica anterior no emplea el factor de ponderación $g_p(i)$.

Por tanto, en la técnica anterior que no emplea un factor de ponderación $g_p(i)$, la derivación de la función de error y el ajuste de la derivada de la función de error en 0 daría el siguiente resultado:

$$a = \frac{3 \sum_{i=0}^4 P(i) - \sum_{i=0}^4 i \cdot P(i)}{5} \quad y \quad b = \frac{\sum_{i=0}^4 i \cdot P(i) - 2 \sum_{i=0}^4 P(i)}{10} \quad (21b)$$

(Véase [ITU06b, 7.6.5]).

En cambio, si se utiliza el enfoque de predicción ponderada de los ejemplos previstos, por ejemplo, el enfoque de predicción ponderado de la fórmula (20) con el factor de ponderación $g_p(i)$, los valores de a y b serían:

$$a = - \frac{A + B + C + D + E}{K} \quad (22a)$$

$$b = + \frac{F + G + H + I + J}{K} \quad (22b)$$

Según un ejemplo particular, A, B, C, D; E, F, G, H, I, J y K puede tener por ejemplo los siguientes valores:

$$\begin{aligned}
A &= (3g_{p3} + 4g_{p2} + 3g_{p1})g_{p4} \cdot P(4) \\
B &= ((2g_{p2} + 2g_{p1})g_{p3} - 4g_{p3}g_{p4}) \cdot P(3) \\
C &= (-8g_{p3}g_{p4} - 3g_{p2}g_{p3} + g_{p1}g_{p2}) \cdot P(2) \\
D &= (-12g_{p1}g_{p4} - 6g_{p1}g_{p3} - 2g_{p1}g_{p2}) \cdot P(1) \\
E &= (-16g_{p0}g_{p4} - 9g_{p0}g_{p3} - 4g_{p0}g_{p2} - g_{p0}g_{p1}) \cdot P(0) \\
F &= (g_{p3} + 2g_{p2} + 3g_{p1} + 4g_{p0})g_{p4} \cdot P(4) \\
G &= ((g_{p2} + 2g_{p1} + 3g_{p0})g_{p3} - g_{p2}g_{p4}) \cdot P(3) \\
H &= (-2g_{p2}g_{p4} - g_{p2}g_{p3} + (g_{p1} + 2g_{p0})g_{p2}) \cdot P(2) \\
I &= (-3g_{p1}g_{p4} - 2g_{p1}g_{p3} - g_{p1}g_{p2} + g_{p0}g_{p1}) \cdot P(1) \\
J &= (-4g_{p0}g_{p4} - 3g_{p0}g_{p3} - 2g_{p0}g_{p2} - g_{p0}g_{p1}) \cdot P(0) \\
K &= (g_{p3} + 4g_{p2} + 9g_{p1} + 16g_{p0})g_{p4} + (g_{p2} + 4g_{p1} + 9g_{p0})g_{p3} + (g_{p1} + 4g_{p0})g_{p2} + g_{p0}g_{p1} \quad (22c)
\end{aligned}$$

Las figuras 10 y 11 muestran una superior performance de la extrapolación de tono propuesta.

5

En este caso, la figura 10 ilustra un diagrama de retardo de tono, en el que el retardo de tono se reconstruye utilizando conceptos del estado de la técnica. En cambio, en la figura 11, se ilustra un diagrama de retardo de tono en el que el retardo de tono se reconstruye según los ejemplos.

10 En particular, la figura 10 ilustra la performance de las normas de la técnica anterior G.718 y G.729.1, mientras que la figura 11 ilustra la performance del concepto proporcionado por un ejemplo.

El eje de las abscisas designa el número de subtrama. La línea 1010 continua muestra el retardo de tono del codificador que está empotrado en una corriente de bits, y que se pierde en el área 1030 del segmento gris. El eje de ordenadas izquierdo representa un eje de retardos de tono. El eje de ordenadas a la derecha representa un eje de ganancia de tono. La línea 1010 continua ilustra el retardo de tono, mientras que las líneas 1021, 1022, 1023 punteadas ilustran la ganancia de tono. El rectángulo 1030 representa la pérdida de trama. Debido a la pérdida de trama que tuvo lugar en el área 1030 del segmento gris, la información acerca del retardo de tono y de la ganancia de tono en esta área no está disponible en el lado del codificador y ha de ser reconstruida.

15

En la figura 10, el retardo de tono que se oculta mediante la norma G.718 se ilustra mediante la porción de línea de puntos y trazos 1011. El retardo de tono que se está ocultando mediante la norma G.729.1 se ilustra mediante la porción 1012 de línea continua. Puede verse claramente que mediante la utilización de la predicción de tono provista (figura 11, porción 1013 de línea continua) corresponde esencialmente al retardo de tono del codificador perdido y es por lo tanto ventajoso con respecto a las técnicas de G.718 y G.729.1.

20

A continuación se describen ejemplos en los que se utiliza la ponderación en función del tiempo transcurrido, con referencia a las fórmulas (23a) a (24b).

25

Para superar los inconvenientes de la técnica anterior, algunos ejemplos aplican una ponderación de tiempo en los retardos de tono, antes de realizar la predicción de tono. La aplicación de una ponderación de tiempo puede efectuarse mediante la minimización de esta función de error:

$$err = \sum_{i=0}^4 \text{tiempo pasado}(i) \cdot ((a + b \cdot i) - P(i))^2 \quad (23a)$$

30

en el que $\text{tiempo_pasado}(i)$ representa la inversa de la cantidad de tiempo que ha transcurrido después de haberse recibido correctamente el retardo de tono y $P(i)$ contiene los correspondientes retardos de tono.

35

Algunos ejemplos pueden, por ejemplo, aplicar ponderaciones elevadas a retardos más recientes y ponderaciones menores a retardos que se han recuperado tiempo atrás.

40

Según algunos ejemplos, en tal caso puede emplearse la fórmula (21a) para derivar a y b .

45

Para obtener la primera subtrama predicha, algunos ejemplos pueden por ejemplo llevar a cabo la predicción sobre la base de las cinco últimas subtramas, $P(0) \dots P(4)$. Por ejemplo, el valor de tono predicho $P(5)$ puede seguidamente obtenerse según:

$$P(5) = a + 5 \cdot b \quad (23b)$$

Por ejemplo, si

5

$$\text{tiempo}_{\text{pasado}} = [1/5 \ 1/4 \ 1/3 \ 1/2 \ 1]$$

(ponderación de tiempo según retardo de subtrama), esto resultaría en lo siguiente:

$$a = \frac{-3,5833 \cdot P(4) + 1,4167 \cdot P(3) + 3,0833 \cdot P(2) + 3,9167 \cdot P(1) + 4,4167 \cdot P(0)}{9,2500} \quad (24a)$$

$$b = \frac{+2,7167 \cdot P(4) + 0,2167 \cdot P(3) - 0,6167 \cdot P(2) - 1,0333 \cdot P(1) - 1,2833 \cdot P(0)}{9,2500} \quad (24b)$$

10

A continuación, se describen ejemplos que proporcionan la resincronización de pulsos.

En la figura 2a se ilustra un aparato para reconstruir una trama, que comprende una señal de voz como una trama reconstruida según un ejemplo. Dicha trama reconstruida está asociada con una o más tramas disponibles, siendo dichas una o más tramas disponibles por lo menos una de las una o más tramas anteriores de la trama reconstruida y una o más tramas subsiguientes de la trama reconstruida, en el que la una o más tramas disponibles comprenden uno o más ciclos de tono en forma de uno o más ciclo de tono disponibles.

15

El aparato comprende una unidad 210 de determinación para determinar una diferencia de número de muestras $(\Delta_0^p; \Delta_i; \Delta_{k+1}^p)$ que indica una diferencia entre un número de muestras de uno de los uno o más ciclos de tono disponibles y un número de muestras de un primer ciclo de tonos a ser reconstruido.

20

Por otra parte, el aparato comprende un reconstructor de tramas para reconstruir la trama reconstruida mediante la reconstrucción, en función de la diferencia del número de muestras $(\Delta_0^p; \Delta_i; \Delta_{k+1}^p)$ y en función de las muestras de dicho uno de los uno o más ciclos de trama disponibles, del primer ciclo de tono a ser reconstruido como primer ciclo de tono reconstruido.

25

El reconstructor 220 de tramas está configurado para reconstruir la trama reconstruida, de manera tal que la trama reconstruida comprende completa o parcialmente el primer ciclo de tono reconstruido, de manera tal que la trama reconstruida comprende completamente o parcialmente un segundo ciclo de tono reconstruido, y de manera tal que el número de muestras del primer ciclo de tono reconstruido difiere de un número de muestras del segundo ciclo de tono reconstruido.

30

La reconstrucción de un ciclo de tono se lleva a cabo reconstruyendo alguna o la totalidad de las muestras del ciclo de tono que ha de ser reconstruido. Si el ciclo de tono a ser reconstruido está completamente comprendido por una trama que se ha perdido, entonces puede ser necesario reconstruir la totalidad de la muestra del ciclo de tono. Si el ciclo de tono a ser reconstruido está solo parcialmente comprendido por la trama que se ha perdido, y si algunas de las muestras del ciclo de tono están disponibles, por ejemplo, ya que están comprendidas por otra trama entonces puede ser por ejemplo suficiente reconstruir solamente las muestras del ciclo de tono que están comprendidos por la trama que se ha perdido para reconstruir el ciclo de tono.

35

40

La figura 2b ilustra la funcionalidad del aparato de la figura 2a. En particular, la figura 2b ilustra una señal 222 de voz que comprende los pulsos 211, 212, 213, 214, 215, 216, 217.

45

Una primera porción de la señal 222 de voz está comprendida por una trama n-1. Una segunda porción de la señal 222 de voz está comprendida por una trama n. Una tercera porción de la señal 222 de voz está comprendida por la trama n+1.

50

En la figura 2b, la trama m-1 precede a la trama n y la trama n+1 sucede a la trama n. Esto significa que la trama n-1 comprende una porción de la señal de voz que tuvo lugar anteriormente en el tiempo en comparación con la porción de la señal de voz de la trama n; y la trama n+1 comprende una porción de la señal de voz que tuvo lugar más tarde en el tiempo en comparación con la porción de la señal de voz de la trama n.

55

En el ejemplo de la figura 2b se supone que la trama n se perdió o está dañada y por lo tanto, solo las tramas anteriores a la trama n ("tramas anteriores") y las tramas que suceden a la trama n ("tramas sucesivas") están

disponibles ("tramas disponibles").

5 Un ciclo de tono puede, por ejemplo, definirse como sigue: un ciclo de tono se inicia con uno de los pulsos de 211, 212, 213, etc. y termina con el pulso inmediatamente posterior en la señal de voz. Por ejemplo, los pulsos 211 y 212 definen el ciclo 201 de tono. Los pulsos 212 y 213 definen el ciclo 202 de tono. Los pulsos 213 y 214 definen el ciclo 203 de tono, etc.

10 Como alternativa pueden considerarse otras definiciones de ciclo de tono, bien conocidas por una persona experta en la materia, que emplean, por ejemplo, otros puntos de inicio y de fin del ciclo de tono.

15 En el ejemplo de la figura 2b, la trama n no se encuentra disponible en un receptor o está dañada. Por lo tanto, el receptor está consciente de los pulsos 211 y 212 y del ciclo 201 de tono de la trama n-1. Por otra parte, el receptor está consciente de los pulsos 216 y 217 y del ciclo 206 de tono de la trama n+1. Sin embargo, la trama n que comprende los pulsos 213, 214 y 215, que comprende completamente los ciclos de tono 203 y 204 y que comprende parcialmente los ciclos de tono 202 y 205, tiene que ser reconstruido.

20 Según algunos ejemplos, la trama n se puede reconstruir en función de por lo menos un ciclo de tono ("ciclos de tono disponibles") de las tramas disponibles (por ejemplo, la trama anterior n-1 o la siguiente trama n+1). Por ejemplo, las muestras del ciclo 201 de tono de la trama n-1 pueden, por ejemplo, ser cíclicamente copiadas varias veces para reconstruir las muestras de la trama perdida o dañada. Mediante el copiado repetido cíclico de las muestras del ciclo de tono, el ciclo de tono en sí mismo se copia, por ejemplo, si el ciclo de tono es c, entonces

$$\text{muestra}(x + i \cdot c) = \text{muestra}(x);$$

25 siendo i un número entero.

30 En los ejemplos, se copian las muestras desde el final de la trama n-1. La longitud de la porción de la trama n-ésima que se copia es igual a la longitud del ciclo 201 de tono (o casi igual). Pero las muestras de tanto 201 como 202 se utilizan para copiar. Esto puede ser considerado de manera especialmente cuidadosa cuando hay un solo pulso en la trama n-1-ésima. En algunos ejemplos, las muestras copiadas son modificadas.

35 Mediante el copiado cíclicamente de manera repetitiva de las muestras de un ciclo de tono, los pulsos 213, 214, 215 de la trama perdida n se mueven a posiciones equivocadas, cuando el tamaño de los ciclos de tono que están compuestos (completamente o parcialmente) por la trama perdida (n) (ciclos 202, 203, 204 y 205 de tono) se diferencia de la magnitud del ciclo de tono copiado disponibles (aquí: el ciclo 201 de tono).

40 Por ejemplo, en la figura 2b, la diferencia entre el ciclo 201 de tono y el ciclo 202 de tono está indicado por Δ_1 , la diferencia entre el ciclo 201 de tono y el ciclo 203 de tono está indicada por Δ_2 , la diferencia entre el ciclo 201 de tono y el ciclo 204 de tono está indicada por Δ_3 , y la diferencia entre el ciclo 201 de tono y el ciclo 205 de tono está indicada por Δ_4 .

45 En la figura 2b, puede verse que el ciclo 201 de tono de la trama n-1 es significativamente más grande que el ciclo 206 de tono. Por otra parte, los ciclos 202, 203, 204 y 205 de tono, que están comprendidos (parcial o totalmente) por la trama n y, cada uno de ellos es más pequeño que el ciclo 201 de tono y más grande que el ciclo 206 de tono. Por otra parte, los ciclos de tono más cercanos al ciclo de tono grande 201 (por ejemplo, el ciclo 202 de tono) son más grandes que los ciclos de tono (por ejemplo, el ciclo 205 de tono) que están más cercanos al ciclo 206 de tono pequeño.

50 En base a estos hallazgos el reconstructor 220 de tramas está configurado para reconstruir la trama reconstruida de manera tal que el número de muestras del primer ciclo de tono reconstruido decir de un número de muestras de un segundo ciclo de tono reconstruido que está parcial o totalmente comprendido por la trama reconstruida.

55 Por ejemplo, según algunos ejemplos, la reconstrucción de la trama depende de una diferencia de número de muestras que indica una diferencia entre el número de muestras de uno de los uno más en ciclos de tono disponibles (por ejemplo, el ciclo 201 de tono) y un número de muestras de un primer ciclo de tono (por ejemplo, los ciclos 202, 203, 204, 205 de tono) que han de ser reconstruidos.

60 Por ejemplo, según un ejemplo, las muestras del ciclo 201 de tono pueden ser por ejemplo copiadas de manera cíclicamente repetida.

Seguidamente, la diferencia en el número de muestras indica cuántas muestras serán suprimidas de la copia cíclicamente repetida correspondiente al primer ciclo de tono a ser reconstruido, o cuántas muestras han de ser añadidas a la copia cíclicamente repetida correspondiente al primer ciclo de tono a ser reconstruido

En la figura 2b, cada número de muestras indica cuántas muestras serán suprimidas de la copia cíclicamente repetida. Sin embargo, en otros ejemplos, el número de muestras puede indicar cuántas muestras serán añadidas a la copia cíclicamente repetida. Por ejemplo, en algunos ejemplos, es posible añadir muestras mediante la adición de muestras de amplitud cero al correspondiente ciclo de tono. En otros ejemplos, es posible añadir muestras al ciclo de tono mediante el copiado de otras muestras del ciclo de tono, por ejemplo, mediante el copiado de muestras que son vecinas a las posiciones de las muestras a ser añadidas.

Aunque anteriormente se han descrito ejemplos en los que las muestras de un ciclo de trama de una trama que precede a la trama perdida o dañada han sido copiadas de manera cíclicamente repetitiva, en otros ejemplos, las muestras de un ciclo de tono de una trama que sigue a la trama perdida o dañada son copiadas de manera cíclicamente repetida a efectos de reconstruir la trama perdida. Rigen de manera análoga los mismos principios descritos anteriormente y a continuación.

Es posible determinar una diferencia de número de muestras de este tipo para cada ciclo de tono a ser reconstruido. Seguidamente, la diferencia del número de muestras de cada ciclo de tono indica cuántas muestras serán suprimidas de la copia cíclicamente repetida correspondiente al de tono correspondiente a ser reconstruido, o cuántas muestras serán añadidas a la copia cíclicamente repetida correspondiente al ciclo de tono correspondiente a ser reconstruido.

Según un ejemplo, la unidad 210 de determinación de puede estar configurada por ejemplo para determinar una diferencia en el número de muestras para cada uno de entre una pluralidad de ciclos de tono a ser reconstruidos, de manera tal que la diferencia en el número de muestras de cada uno de los ciclos de tono indica una diferencia entre la cantidad de muestras de dicho uno de los uno o más ciclos de tono disponibles y una cantidad de muestras de dicho ciclo de tono a ser reconstruido. El reconstructor 220 de trama, puede estar configurado por ejemplo para reconstruir cada ciclo de tono de entre la pluralidad de ciclos de tono a ser reconstruidos en función de la diferencia en el número de muestras de dicho ciclo de tono a ser reconstruido y en función de la muestra de dicho uno de los uno o más ciclos de tono disponibles, a efectos de reconstruir la trama reconstruida.

En un ejemplo, el reconstructor 220 de trama, puede estar configurado por ejemplo para generar una trama intermedia en función de dicho uno de los uno o más ciclo de tono disponibles. El reconstructor 220 de tramas puede estar configurado por ejemplo para modificar la trama intermedia de manera de obtener la trama reconstruida.

Según un ejemplo, la unidad 210 de determinación puede estar configurada por ejemplo para determinar un valor de la diferencia de las tramas (d; s) que indica cuántas muestras han de ser eliminadas de la trama intermedia o cuántas muestras han de ser añadidas a la trama intermedia. Por otra parte, el reconstructor 220 de tramas puede estar configurado, por ejemplo, para eliminar primeras muestras de la trama intermedia de maneras de obtener la trama reconstruida, cuando el valor de la diferencia de tramas indica que las primeras muestras han de ser eliminadas de la trama. Por otra parte, el reconstructor 220 de tramas puede estar configurado por ejemplo para añadir segundas muestras a la trama intermedia de manera de obtener la trama reconstruida, cuando el valor de la diferencia de tramas (d; s) indica que las segundas muestras serán añadidas a la trama.

En un ejemplo, el reconstructor 220 de tramas puede estar configurado por ejemplo para eliminar las primeras muestras de la trama intermedia cuando el valor de la diferencia de las tramas indica que las primeras muestras serán eliminadas de la trama, de manera tal que el número de primeras muestras que son eliminadas de la trama intermedia está indicado por el valor de la diferencia de tramas. Por otra parte, el reconstructor 220 de trama puede estar configurado por ejemplo para añadir las segundas muestras a la trama intermedia cuando el valor de la diferencia de tramas indica que las segundas muestras serán añadidas a la trama, de manera tal que el número de segundas muestras que son añadidas a la trama intermedia está indicado por el valor de la diferencia de tramas.

Según un ejemplo, la unidad 210 de determinación puede estar configurada, por ejemplo, para determinar el número de la diferencia de tramas s de manera tal que la fórmula:

$$s = \sum_{i=0}^{M-1} (p[i] - T_r) \frac{L}{MT_r}$$

mantenga su validez en el que L indica un número de muestras de la trama reconstruida, en el que M indica un número de subtramas de la trama reconstruida, en el que T_r indica una longitud de período de tono de dicho uno de los uno o más ciclos de tono disponibles, y en el que $p[i]$ indica una longitud de período de tono de un ciclo de tono reconstruido de la subtrama i-ésima de la trama reconstruida.

En un ejemplo, el reconstructor 220 de trama puede estar adaptado por ejemplo para generar una trama intermedia en función de dicho uno de los uno o más ciclos de tono disponibles. Por otra parte, el reconstructor 220 de trama puede estar adaptado por ejemplo para generar la trama intermedia de manera tal que la trama intermedia

comprenda un primer ciclo de tono intermedio parcial, uno o más ciclos de tono intermedio adicionales, y un segundo ciclo de tono intermedio parcial. Por otra parte, el primer ciclo de tono intermedio parcial puede depender por ejemplo de una o más de las muestras de dicho uno o más ciclos de tono disponibles, en el que cada uno de los uno o más ciclos de tonos intermedios adicionales depende de la totalidad de las muestras de dicho uno de los uno o más ciclos de tono disponibles, y en el que el segundo ciclo de tono intermedio parcial depende de una o más de las muestras de dicho uno de los uno o más ciclos de tono disponibles adicionales. Por otra parte, la unidad 210 de determinación puede estar configurada por ejemplo para determinar un número de diferencia de porción inicial que indique cuántas muestras han de ser eliminadas de o añadidas en el primer ciclo de tono intermedio parcial, y en el que el reconstructor 220 de tramas está configurado para eliminar una o más primeras muestras del primer ciclo de tono intermedio parcial, o está configurado para añadir una o más primeras muestras al primer ciclo de tono intermedio parcial en función del número de diferencia de la porción inicial. Además, la unidad 210 de determinación puede estar configurada por ejemplo para determinar para cada uno de los ciclos de tonos intermedios adicionales un número de diferencia de ciclo de tono que indique cuántas muestras han de ser eliminadas o añadidas en dicho uno de los ciclos de tono intermedios adicionales. Además, el reconstructor 220 de tramas puede estar configurado por ejemplo para eliminar una o más segundas muestras de dicho uno de los ciclos de tonos intermedios adicionales, o está configurado para añadir una o más segundas muestras a dicho uno de los ciclos de tono intermedios adicionales en función de dicho número de diferencia de ciclo de tono. Por otra parte, la unidad 210 de determinación puede estar configurada por ejemplo para determinar un número de diferencia de porción extrema que indique cuántas muestras han de ser eliminadas o añadidas en el ciclo de tono intermedio parcial, y en el que el reconstructor 220 de tramas puede estar configurado por ejemplo para eliminar una o más terceras muestras del segundo ciclo de tono intermedio parcial, o está configurado para añadir una o más terceras muestras al segundo ciclo de tono intermedio parcial en función del número de diferencia de porciones extremas.

Según un ejemplo, el reconstructor 220 de tramas puede estar configurado por ejemplo para generar una trama intermedia en función de dicho uno de los uno o más el ciclo de tono disponibles. Por otra parte, la unidad 210 de determinación puede estar adaptada por ejemplo para determinar una o más porciones de señales de baja energía de la señal de voz comprendida por la trama intermedia, en el que cada una de las porciones de señal de baja energía es una primera porción de voz dentro de la trama intermedia, en el que la energía de la señal de voz es inferior a la de una segunda porción de señal de la señal de voz comprendida por la trama intermedia. Por otra parte, el reconstructor 220 de tramas puede estar configurado por ejemplo para eliminar una o más muestras desde por lo menos una no de las una o más porciones de baja energía entre porciones de señales de baja energía de la señal de voz, o para añadir una o más muestras a por lo menos una de las una o más porciones de señales de baja energía de la señal de voz, de manera obtener la trama reconstruida.

En un ejemplo particular, el reconstructor 220 de tramas puede estar configurado por ejemplo para generar la trama intermedia, de manera tal que la trama intermedia comprenda uno o más ciclos de tono reconstruidos, de manera tal que cada uno de los uno o más ciclos de tono reconstruidos dependa de dicho uno de los uno o más ciclos de tono disponibles. Por otra parte, la unidad de determinación puede por ejemplo estar configurada para determinar un número de muestras que han de ser eliminadas de cada uno de los uno o más ciclos de todo reconstruidos. Por otra parte, la unidad 210 de determinación puede estar configurada por ejemplo para determinar cada una de las una o más porciones de señal de baja energía de manera tal que para cada una de las porciones de señales de baja energía un número de muestras de dicha porción de señal de baja energía depende del número de muestras que han de ser eliminadas de uno de los uno o más ciclos de tono reconstruidos, en el que dicha porción de señal de baja energía está situada dentro de dicho uno de los uno o más ciclos de tono reconstruidos.

En un ejemplo, la unidad 210 de determinación puede estar configurada por ejemplo para determinar una posición de uno o más pulsos de la señal de voz de la trama a ser reconstruida como trama reconstruida. Además, el reconstructor 120 de tramas puede por ejemplo está configurado para reconstruir la trama reconstruida en función de la posición de los uno o más pulsos de la señal de voz.

Según un ejemplo, la unidad 210 de determinación puede estar configurada por ejemplo para determinar una posición de dos o más pulsos de la señal de voz de la trama a ser reconstruida como trama reconstruida, en el que $T[0]$ es la posición de uno de los dos o más pulsos de la señal de voz de la trama ser reconstruida como trama reconstruida, y en el que la unidad 210 de determinación estar configurada para determinar la posición ($T[i]$) de otros pulsos de los dos o más pulsos de la señal de voz según la fórmula:

$$T[i] = T[0] + i T_r$$

en el que T_r indica una longitud redondeada de dicho uno de los uno o más ciclos de tono disponibles, y en el que i es un número entero.

Según un ejemplo, la unidad 210 de determinación puede estar configurada para determinar un índice k del último pulso de la señal de voz de la trama a ser reconstruida como la trama reconstruida de manera tal que

$$k = \left\lfloor \frac{L - s - T[0]}{T_r} - 1 \right\rfloor,$$

5 en el que L indica un número de muestras de la trama reconstruida, en el que s indica el valor de la diferencia de tramas, en el que $T[0]$ indica una posición de un pulso de la señal de voz de la trama a ser reconstruida como la trama reconstruida, que es diferente del último pulso de la señal de voz, y en el que T_r indica una longitud redondeada de dicho uno de los uno o más ciclos de tono disponibles.

10 En un ejemplo, la unidad 210 de determinación puede estar configurada por ejemplo para reconstruir la trama a ser reconstruida como la trama reconstruida mediante la determinación de un parámetro δ , en el que δ se define según la fórmula:

$$\delta = \frac{T_{ext} - T_p}{M}$$

15 en el que la trama a ser reconstruida como la trama reconstruida comprende M subtramas, en el que T_p indica la longitud de dicho uno de los uno o más ciclos de tono disponibles, y en el que T_{ext} indica una longitud de uno de los ciclos de tono a ser reconstruidos de la trama a ser reconstruida como trama reconstruida.

20 Según un ejemplo, la unidad 210 de determinación puede estar configurada por ejemplo para reconstruir la trama reconstruida mediante la determinación de una longitud redondeada T_r de dicho uno de los uno o más ciclos de tono disponibles en base a la siguiente fórmula:

$$T_r = [T_p + 0,5]$$

25 en el que T_p indica la longitud de dicho uno de los uno o más ciclos de tono disponibles.

En un ejemplo, la unidad 210 de determinación puede estar configurada por ejemplo para reconstruir la trama reconstruida mediante la aplicación de la siguiente fórmula:

$$s = \delta \frac{L M + 1}{T_r} - L \left(1 - \frac{T_p}{T_r} \right)$$

30 en la que T_p indica la longitud de dicho uno de los uno o más ciclos de tono disponibles, en el que T_r indica una longitud redondeado de dicho uno de los uno o más ciclos de tono disponibles, en el que la trama a ser reconstruida como trama reconstruida comprende M subtramas, en el que la trama ser reconstruida como la trama reconstruida comprende L muestras, y en el que δ es un número real i que indica una diferencia entre un número de muestras de dicho uno de los uno o más ciclos de tono disponibles y un número de muestras de uno de los uno o más ciclos de tono a ser reconstruidos.

Seguidamente se describen ejemplos con mayor detalle.

40 A continuación, un primer grupo de ejemplos de resincronización de pulso se describe con referencia a las fórmulas (25) - (63).

45 En tales ejemplos, si no hay un cambio de tono, se utiliza el último retardo de tono sin redondeo, preservándose la parte fraccional. La parte periódica se constrúan mediante el tono no número entero e interpolación, por ejemplo, en [MTTA90]. Esto reducirá el corrimiento o desplazamiento de frecuencia de las armónicas, en comparación con la utilización del retardo de tono redondeado, con lo cual se mejora de manera significativa la ocultación de las señales tonales o de voz con tono constante.

50 La ventaja se ilustra en las figuras 8 y 9, en el que la señal representativa de tono con pérdidas de trama se oculta mediante retardo de tono fraccional respectivamente redondeado y no redondeado. En este caso, la figura 8 ilustra una representación en tiempo y frecuencia de una señal de voz que se está resincronizando mediante un retardo de tono redondo. La figura 9 ilustra una representación en tiempo frecuencia de una señal de voz que se está resincronizando mediante un retardo de tono no redondeado con la parte fraccional.

55 Habrá una mayor complicidad de cálculo cuando se utiliza la parte funcional del tono. Esto no debería influir sobre la complejidad del caso por cuanto no hay necesidad de la resincronización de pulso glotal.

Si no hay un cambio de tono predicho entonces no hay necesidad del procesamiento que se explica a continuación.

- Si se predice un cambio de tono, los ejemplos descritos con referencia a las fórmulas (25) - (63) proporcionan conceptos para determinar d , siendo la diferencia entre la suma del número total de muestras dentro de los ciclos de tono con el tono constante (T_c) y la suma del número total de muestras dentro de los ciclos de tono con el tono en evolución $p[i]$.

A continuación se define T_c como en la fórmula (15a): $T_c = \text{redondeo}(\text{último_tono})$.

- 10 Según los ejemplos, la diferencia d puede determinarse mediante un algoritmo más rápido y más preciso (enfoque de algoritmo rápido para determinar d) como se describe a continuación.

Un algoritmo de este tipo puede basarse por ejemplo en los principios siguientes:

- 15 - En cada subtrama i : $T_c - p[i]$ muestras para cada ciclo de tono (de longitud T_c) deberían ser eliminadas (o $p[i] - T_c$ añadido si $T_c - p[i] < 0$).

$$\frac{L_subfr}{T_c}$$

- Hay $\frac{L_subfr}{T_c}$ ciclos de tono en cada subtrama.

$$\frac{L_subfr}{T_c}$$

- Por lo tanto, para cada subtrama i ($T_c - p[i]$) $\frac{L_subfr}{T_c}$ muestras deberían ser eliminadas.

- 20 Según algunos ejemplos, no se lleva a cabo ningún redondeo y se utiliza un tono fraccional. Entonces:

$$p[i] = T_c + (i + 1) \delta$$

- Por lo tanto, para cada subtrama i , $(i + 1) \delta \frac{L_subfr}{T_c}$ muestras deberían ser eliminadas si $\delta < 0$ (o añadidas si $\delta > 0$).

$$d = -\delta \frac{L_subfr}{T_c} \sum_{i=1}^M i$$

- Por lo tanto, d (en el que M es el número de subtramas en una trama).

25

Según algunos otros ejemplos, se lleva a cabo un redondeo. Para el tono número entero (M es el número de subtramas en una trama), d se define como sigue:

$$d = \text{redondeo} \left(\left(M T_c - \sum_{i=0}^{M-1} p[i] \right) \frac{L_subfr}{T_c} \right) \quad (25)$$

30

Según un ejemplo, se proporciona un algoritmo para calcular d de acuerdo con:

```
ftmp = 0;
for (i=0; i < M; i++) {
    ftmp += p[i];
}
d = (short)floor((M*T_c - ftmp)*(float)L_subfr / T_c + 0.5);
```

35

En otro ejemplo, la última línea del algoritmo se reemplaza por:

$$d = (\text{short})\text{floor} (L_frame - \text{ftmp} * (\text{float}) L_subfr / T_c + 0,5);$$

40

Según los ejemplos el último pulso $T[n]$ se encuentra según:

$$n = i \mid T[i] + iT_c < L_trama \quad \wedge \quad T[i] + (i + 1)T_c \geq L_trama \quad (26)$$

Según un ejemplo, se emplea una fórmula para calcular N . Esta fórmula se obtiene a partir de la fórmula (26) según:

$$N = 1 + \left\lceil \frac{L_trama - T[0]}{T_c} \right\rceil \quad (27)$$

y el último pulso tiene seguidamente el índice $N - 1$.

5 Según esta fórmula, es posible calcular N para los ejemplos ilustrados por las figuras 4 y 5.

A continuación se describe un concepto sin búsqueda explícita del último pulso, pero que toma en cuenta las posiciones de los pulsos. Un concepto de este tipo no necesita N , el último índice de pulso en la parte periódica construida.

10

La posición real del último pulso en la parte periódica construida de la excitación ($T[k]$) determina el número de ciclos de tono completos k , en el que se retiran (o se añaden) las muestras.

15 La figura 12 ilustra una posición del último pulso $T[2]$ antes de la retirada de d muestras. En cuanto a los ejemplos descritos con respecto a las fórmulas (25) - (63), el signo 1210 de referencia designa d .

En el ejemplo de la figura 12, el índice del último pulso k es 2 y hay dos ciclos de tono completos desde las cuales deberían eliminarse las muestras.

20 Después de eliminar d desde la señal de longitud $L_trama + d$, no hay muestras de la señal original más allá de $L_trama + d$ muestras. Por lo tanto, $T[k]$ se halla dentro de $L_trama + d$ muestras y k se determina por lo tanto mediante la siguiente expresión:

$$k = i \mid T[i] < L_trama + d \leq T[i + 1] \quad (28)$$

25

A partir de la fórmula (17) y de la formula (28), se deduce que

$$T[0] + kT_c < L_trama + d \leq T[0] + (k + 1) T_c \quad (29)$$

30 Es decir:

$$\frac{L_trama + d - T[0]}{T_c} - 1 \leq k < \frac{L_trama + d - T[0]}{T_c} \quad (30)$$

De la fórmula (30) se deduce que:

35

$$k = \left\lceil \frac{L_trama + d - T[0]}{T_c} - 1 \right\rceil \quad (31)$$

40 En un códec que utiliza por ejemplo tramas de por lo menos 20 ms y, en el que la frecuencia fundamental más baja del habla es por ejemplo de por lo menos 40 Hz, en la mayoría de los casos existe por lo menos un pulso en la trama oculta que no es SORDA.

Seguidamente se describe un caso con por lo menos dos ($k \geq 1$) con referencia las fórmulas (32) - (46).

45 Supongamos que en cada ciclo de tono i -ésimo entre pulsos, se eliminarán Δ_i muestras, en el que Δ_i se define como:

$$\Delta_i = \Delta + (i - 1) a, \quad 1 \leq i \leq k, \quad (32)$$

en la que a es una variable desconocida que debe expresarse en términos de las variables conocidas.

50 Supongamos que serán eliminadas $t \Delta_0$ muestras antes del primer pulso, en el que Δ_0 se define como:

$$\Delta_0 = (\Delta - a) \frac{T[0]}{T_c} \quad (33)$$

Supongamos que Δ_{k+1} muestras serán eliminadas después del último pulso, en el que Δ_{k+1} se define como:

$$\Delta_{k+1} = (\Delta + ka) \frac{L + d - T[k]}{T_c} \quad (34)$$

5 Las dos suposiciones últimas armonizan con la fórmula (32) que tiene en cuenta la longitud de los ciclos de tono parciales primero y segundo.

Cada uno de los valores de Δ_i es una diferencia en el número de muestras. Por otra parte, Δ_0 es una diferencia en el número de muestras. Además, Δ_{k+1} es una diferencia en el número de muestras.

10

La figura 13 ilustra la señal de voz de la figura 12, y adicionalmente ilustra Δ_0 a Δ_3 . El número de muestras a ser eliminadas en cada ciclo de tono se representa esquemáticamente en el ejemplo de la figura 13, en el que $k = 2$. En cuanto a los ejemplos descritos con referencia las fórmulas (25) - (63), el signo 1210 de referencia designa d .

15 El número total de muestras a ser eliminadas, d , se relacione seguidamente con Δ_i como sigue:

$$d = \sum_{i=0}^{k+1} \Delta_i \quad (35)$$

A partir de las fórmulas (32) - (35), es posible obtener d como sigue:

$$d = (\Delta - a) \frac{T[0]}{T_c} + (\Delta + ka) \frac{L + d - T[k]}{T_c} + \sum_{i=1}^k (\Delta + (i-1)a) \quad (36)$$

La fórmula (36) es equivalente a:

$$d = \Delta \left(\frac{T[0]}{T_c} + \frac{L + d - T[k]}{T_c} + k \right) + a \left(k \frac{L + d - T[k]}{T_c} - \frac{T[0]}{T_c} + \frac{k(k-1)}{2} \right) \quad (37)$$

25

Supongamos que el último ciclo de tono completo de una trama oculta tiene una longitud $p[M - 1]$, es decir:

$$\Delta_k = T_c - p[M - 1] \quad (38)$$

30 A partir de las fórmulas (32) y (38) se deduce que:

$$\Delta = T_c - p[M - 1] - (k - 1)a \quad (39)$$

Por otra parte, de las fórmulas (37) y (39), se deduce que:

35

$$d = (T_c - p[M - 1] + (1 - k)a) \left(\frac{T[0]}{T_c} + \frac{L + d - T[k]}{T_c} + k \right) + a \left(k \frac{L + d - T[k]}{T_c} - \frac{T[0]}{T_c} + \frac{k(k-1)}{2} \right) \quad (40)$$

La fórmula (40) es equivalente a:

40

$$d = (T_c - p[M - 1]) \left(\frac{T[0]}{T_c} + \frac{L + d - T[k]}{T_c} + k \right) + a \left((1 - k) \frac{T[0]}{T_c} + (1 - k) \frac{L + d - T[k]}{T_c} + \right)$$

$$+ (1 - k)k + k \left(\frac{L + d - T[k]}{T_c} - \frac{T[0]}{T_c} + \frac{k(k-1)}{2} \right) \quad (41)$$

A partir las fórmulas (17) y (41), se deduce que:

$$d = (T_c - p[M-1]) \frac{L + d}{T_c} + a \left(-k \frac{T[0]}{T_c} + \frac{L + d - T[k]}{T_c} - \frac{k(k-1)}{2} \right) \quad (42)$$

La fórmula (42) es equivalente a:

$$dT_c = (T_c - p[M-1])(L + d) + a \left(-kT[0] + L + d - T[k] + \frac{k(1-k)}{2}T_c \right) \quad (43)$$

Por otra parte, de la fórmula (43), se deduce que:

$$a = \frac{dT_c - (T_c - p[M-1])(L + d)}{-kT[0] + L + d - T[k] + \frac{k(1-k)}{2}T_c} \quad (44)$$

La fórmula (44) es equivalente a:

$$a = \frac{p[M-1](L + d) - T_c L}{L + d - (k+1)T[0] - kT_c + \frac{k(1-k)}{2}T_c} \quad (45)$$

Además, la fórmula, (45) es equivalente a:

$$a = \frac{p[M-1](L + d) - T_c L}{L + d - (k+1)T[0] - \frac{k(1+k)}{2}T_c} \quad (46)$$

Según los ejemplos, se calcula ahora en base a las fórmulas (32)-(34), (39) y (46), cuántas muestras han de ser eliminadas o añadidas antes del primer pulso, y/o entre y/o después del último pulso.

25

En un ejemplo, las muestras se retiran o añaden en las regiones de mínima energía.

Según los ejemplos, el número de muestras a ser eliminadas puede ser redondeado, por ejemplo, mediante:

$$\begin{aligned} \Delta'_0 &= \lfloor \Delta_0 \rfloor \\ \Delta'_i &= \lfloor \Delta_i + \Delta_{i-1} - \Delta'_{i-1} \rfloor, \quad 0 < i \leq k \\ \Delta_{k+1} &= d - \sum_{i=0}^k \Delta_i \end{aligned}$$

30

Seguidamente se describe un caso con un pulso ($k = 0$) con referencia a las fórmulas (47) a (55).

Si hay exactamente un solo pulso en la trama oculta, entonces Δ_0 muestras habrán de ser eliminadas antes del pulso:

35

$$\Delta_0 = (\Delta - a) \frac{T[0]}{T_c} \quad (47)$$

en el que Δ y a son variables desconocidas que han de expresarse en términos de las variables conocidas. Δ_1 han de ser eliminadas después del pulso, en el que:

$$\Delta_1 = \Delta \frac{L + d - T[0]}{T_c} \quad (48)$$

5

En tal caso, el número total de muestras a ser eliminadas está dado por la siguiente expresión:

$$d = \Delta_0 + \Delta_1 \quad (49)$$

10 A partir de las fórmulas (47) a (49), se deduce que:

$$d = (\Delta - a) \frac{T[0]}{T_c} + \Delta \frac{L + d - T[0]}{T_c} \quad (50)$$

La fórmula (50) es equivalente a:

15

$$dT_c = \Delta (L + d) - aT[0] \quad (51)$$

Su supone que la relación entre el ciclo de tono antes del pulso y el ciclo de tono después del pulso es la mismo que la relación entre el retardo de tono en la última subtrama y la primera subtrama en la trama previamente recibida:

20

$$\frac{\Delta}{\Delta - a} = \frac{p[-1]}{p[-4]} = r \quad (52)$$

A partir de la fórmula (52), se deduce que:

25

$$a = \Delta \left(1 - \frac{1}{r} \right) \quad (53)$$

Además, a partir de las fórmulas (51) y (53), se deduce que:

$$dT_c = \Delta (L + d) - \Delta \left(1 - \frac{1}{r} \right) T[0] \quad (54)$$

30

La fórmula (54) es equivalente a:

$$\Delta = \frac{dT_c}{L + d + \left(\frac{1}{r} - 1 \right) T[0]} \quad (55)$$

35 Hay $\lfloor \Delta - a \rfloor$ muestras a ser eliminadas o adicionadas en la región de mínima energía antes del pulso y $d - \lfloor \Delta - a \rfloor$ muestras después del pulso.

A continuación, se describe un concepto simplificado según los ejemplos, que no requiere una búsqueda de (la ubicación de) los pulsos, y haciéndose referencia a las fórmulas (56) - (63).

40

$t[i]$ designa la longitud del ciclo de tono i -ésimo. Después de eliminar d muestras de la señal, se obtienen k ciclos de tono completos y un ciclo de tono parcial (hasta completo).

Por lo tanto:

45

$$\sum_{i=0}^{k-1} t[i] < L \leq \sum_{i=0}^k t[i] \quad (56)$$

Dado que se obtienen ciclos de tono de longitud $t[i]$ a partir del ciclo de tono de longitud T_c después de eliminar algunas muestras, y dado que el número total de muestras eliminadas es d , se deduce que

$$kT_c < L + d \leq (k+1)T_c \quad (57)$$

Se deduce que:

$$\frac{L+d}{T_c} - 1 \leq k < \frac{L+d}{T_c} \quad (58)$$

Además, se deduce que

$$k = \left\lceil \frac{L+d}{T_c} \right\rceil - 1 \quad (59)$$

Según los ejemplos, puede suponerse un cambio lineal en el retardo de tono:

$$t[i] = T_c - (i+1)\Delta, \quad 0 \leq i \leq k$$

En los ejemplos, $(k+1)\Delta$ muestras son eliminadas en el ciclo de tono k -ésimo.

Según los ejemplos, en la parte del ciclo de tono k -ésimo, que permanece en la trama después de eliminar las muestras,

$$\frac{L+d-kT_c}{T_c} (k+1)\Delta$$

muestras son eliminadas.

Por lo tanto, el número total de las muestras eliminadas es:

$$d = \frac{L+d-kT_c}{T_c} (k+1)\Delta + \sum_{i=0}^{k-1} (i+1)\Delta \quad (60)$$

La fórmula (60) es equivalente a:

$$d = \frac{L+d-kT_c}{T_c} (k+1)\Delta + \frac{k(k+1)}{2}\Delta \quad (61)$$

Además, la fórmula (61) es equivalente a:

$$\frac{d}{(k+1)} = \left(\frac{L+d-kT_c}{T_c} + \frac{k}{2} \right) \Delta \quad (62)$$

Por otra parte, la formula (62) es equivalente a:

$$\Delta = \frac{2dT_c}{(k+1)(2L+2d-kT_c)} \quad (63)$$

Según los ejemplos, $(i + 1) \Delta$ muestras son eliminadas en la posición de la mínima energía. No hay necesidad de conocer la ubicación de los pulsos, ya que la búsqueda de la posición de energía mínima se efectúa en la memoria intermedia circular que contiene un ciclo de tono.

5 Si la posición de energía mínima es después del primer pulso y si no han de revolverse muestras antes del primer pulso, esto se podría presentar una situación en la cual el retardo de tono evoluciona como $(T_c + \Delta), T_c, T_c, (T_c - \Delta), (T_c - 2\Delta)$ (2 ciclos de tono en la última trama recibida y tres ciclos de tono en la trama oculta). Por lo tanto, habría una discontinuidad. Una discontinuidad similar puede surgir después del último pulso, pero no en el mismo momento cuando sucede antes del primer pulso.

10 Por otra parte, la región de energía mínima aparecería más probablemente después del primer pulso, si el pulso es más cercano al inicio de la trama oculta. Si el primer pulso el más cercano al inicio de la trama oculta es más probable que el último ciclo de tono en la última trama recibida sea más grande que T_c . A efectos de reducir la posibilidad de una discontinuidad en el cambio de tono, debería utilizarse la ponderación para dar ventaja en las regiones mínimas más cercanas al inicio o al fin del ciclo de tono.

15 Según los ejemplos, se describe una implementación de los conceptos provistos, que implementa una o más o la totalidad de las siguientes etapas de método:

20 1. Almacenar, en una memoria intermedia temporal B, T_c muestras filtradas de bajo pase desde el final de la última trama recibida, buscando en paralelo la región de energía mínima. La memoria intermedia temporal se considera como una memoria intermedia circular cuando se está buscando la región de energía mínima (esto puede significar que la región de energía mínima puede consistir en unas pocas muestras del inicio y unas pocas muestras del final del ciclo de tono). La región de energía mínima puede, por ejemplo, ser la ubicación del mínimo para la ventana deslizando de longitud $\lceil (k + 1) \Delta \rceil$ muestras. Puede utilizarse por ejemplo la ponderación que puede dar por

25 ejemplo una ventaja a las regiones mínimas más cercanas al inicio del ciclo de tono.

2. Copiar las muestras de la memoria intermedia temporal B en la trama, desplazar $\lceil \Delta \rceil$ muestras a la región de energía mínima seguido. Por lo tanto, se crea un ciclo de tono con una longitud $t[0]$. Ajustar $\delta_0 = \Delta - \lceil \Delta \rceil$.

30 3. Para el ciclo de tono i -ésimo ($0 < i < k$), copiar las muestras de los ciclos de tono $(i - 1)$ ésimo, desplazar $\lceil \Delta \rceil + \lceil \delta_{i-1} \rceil$ muestras a la región de energía mínima. Ajustar $\delta_i = \delta_{i-1} - \lceil \delta_{i-1} \rceil + \Delta - \lceil \Delta \rceil$. Repetir esta etapa $k - 1$ veces.

35 4. Para la búsqueda del ciclo de tono k ésimo para la nueva región mínima en el ciclo de tono $(k - 1)$ ésimo mediante la ponderación de dar ventaja a las regiones mínimas más cercanas al final del ciclo de tono. Seguidamente copiar las muestras del ciclo de tono $(k - 1)$ ésimo, trasladar

$$d - \left\lceil \frac{k(k+1)}{2} \Delta + \frac{k(k-1)}{2} \Delta \right\rceil = d - \lceil k^2 \Delta \rceil$$

40 muestras a la región de energía mínima.

Si deben añadirse muestras, puede utilizarse el procedimiento equivalente tomando en cuenta que $d < 0$ y $\Delta < 0$ que adicionamos en total $|d|$ muestras, es decir se añaden $\lceil (k + 1) \Delta \rceil$ en el ciclo k ésimo en la posición de la mínima energía.

45 El tono fraccional puede ser utilizado a nivel de subtrama para derivar d como se describió en lo que precede en cuanto al "enfoque del algoritmo rápido para determinar d " que como de alguna manera se utilizan las longitudes de los ciclos de tono aproximados.

50 Seguidamente se describe un segundo grupo de ejemplos de resincronización de pulsos con referencia las fórmulas (64) - (113). Estos ejemplos del primer grupo emplean la definición de la fórmula (15b),

$$T_r = [T_p + 0,5]$$

55 en el que la longitud del último período de tono es T_p , y la longitud del segmento que se copia es T_r .

Si algunos parámetros utilizados por el segundo grupo de ejemplos de resincronización de pulsos no se definen a continuación, los ejemplos de la presente invención pueden emplear las definiciones provistas por estos parámetros con respecto al primer grupo de ejemplos de la resincronización de pulsos definidos en lo que precede (Véanse las

60

fórmulas (25) - (63)).

5 Algunas de las fórmulas (64) a (113) del segundo grupo de ejemplos de resincronización pueden redefinir algunos de los parámetros utilizados con respecto al primer grupo de ejemplos de resincronización de pulsos. En este caso, las definiciones redefinidas provistas rigen para los segundos ejemplos de resincronización de pulsos.

Como se describe en lo que precede, según algunos ejemplos, la parte periódica puede por ejemplo construirse para una trama y una subtrama adicional, en el que la longitud de la trama se designa como $L = L_{trama}$.

10 Por ejemplo, con M subtramas en una trama, la longitud de la subtrama es $L_{subtr} = \frac{L}{M}$.

Como ya se describió, $T[0]$ es la ubicación del primer pulso máximo en la parte periódica construida de la excitación. Las posiciones de los otros pulsos están dadas por:

$$15 \quad T[i] = T[0] + i T_r.$$

Según los ejemplos, y en función de la construcción de la parte periódica de la excitación, después de la construcción de la parte periódica de la excitación, se lleva a cabo la resincronización de pulso glotal para corregir la diferencia entre la posición teórica estimada del último pulso en la trama perdida (P), y su posición real en la parte periódica construida de la excitación ($T[k]$).

20 La posición teórica estimada del último pulso en la trama perdida (P) puede por ejemplo determinarse de manera indirecta mediante la estimación de la evolución del retardo de tono. La evolución del retardo de tono es por ejemplo extrapolado en base a los retardos de tono de las últimas siete subtramas antes de la trama perdida. Los retardos de tono en evolución en cada subtrama son:

$$25 \quad p[i] = T_p + (i + 1)\delta, 0 \leq i < M \quad (64)$$

en el que

$$30 \quad \delta = \frac{T_{ext} - T_p}{M} \quad (65)$$

y T_{ext} es el tono extrapolado e i es el índice de subtrama. La extrapolación de tono puede efectuarse por ejemplo utilizando el ajuste lineal ponderado el método de G.718 o el método de G.729.1 o cualquier otro método para la interpolación del tono, que por ejemplo tenga en cuenta una o más tramas futuras. La extrapolación de tono también puede ser no lineal. En un ejemplo, T_{ext} puede determinarse de la misma manera cómo se define T_{ext} en lo que precede.

40 La diferencia dentro de una longitud de trama entre la suma del número total de muestras con ciclos de tono con el tono en evolución ($p[i]$) y la suma del número total de muestras con ciclos de tono con el tono constante (T_p) se designa como s .

Según los ejemplos, si $T_{ext} > T_p$ entonces s muestras deberían ser añadidas a una trama, y si $T_{ext} < T_p$ entonces $-s$ muestras deberían ser eliminadas de una trama. Después de añadir o eliminar $|s|$ muestras, el último pulso en la trama oculta estará en la posición teórica estimada (P).

45 Si $T_{ext} = T_p$, no hay necesidad de añadir ni eliminar muestras dentro de una trama.

Según algunos ejemplos, la resincronización de pulso glotal se efectúa adicionando o removiendo muestras en las regiones de energía mínima de todos los ciclos de tono.

50 A continuación, se describe el parámetro de cálculo s según los ejemplos, haciéndose referencia a las fórmulas (66) a (69).

Según algunas realizaciones, la diferencia, s , puede calcularse por ejemplo en base a los siguientes principios:

55 - En cada subtrama i , deberían añadirse $p[i] - T_r$ muestras para cada ciclo de tono (de longitud T_r) (si $p[i] - T_r > 0$); (o $T_r - p[i]$ muestras deberían eliminarse si $p[i] - T_r < 0$).

Hay $\frac{L - \text{subfr}}{T_r} = \frac{L}{MT_r}$ ciclo de tono en cada subtrama.

Por lo tanto, en la subtrama i-ésima deberían eliminarse $(p[i] - T_r) \frac{L}{MT_r}$ muestras.

5 Por lo tanto, en armonía con la fórmula (64), según un ejemplo, puede calcularse S por ejemplo según formula (66):

$$s = \sum_{i=0}^{M-1} (p[i] - T_r) \frac{L}{MT_r} = \sum_{i=0}^{M-1} (T_p + (i+1)\delta - T_r) \frac{L}{MT_r} =$$

$$= \frac{L}{MT_r} \sum_{i=0}^{M-1} ((i+1)\delta + T_p - T_r) \quad (66)$$

10 La fórmula (66) es equivalente a:

$$s = \frac{L}{MT_r} \left(M(T_p - T_r) + \delta \sum_{i=0}^{M-1} (i+1) \right) = \frac{L}{MT_r} \left(M(T_p - T_r) + \delta \frac{M(M+1)}{2} \right), \quad (67)$$

en el que la fórmula (67) es equivalente a:

$$15 \quad s = \frac{L}{T_r} \left(T_p - T_r + \delta \frac{M+1}{2} \right) = \frac{L}{T_r} \delta \frac{M+1}{2} + \frac{L}{T_r} (T_p - T_r), \quad (68)$$

y en el que la fórmula (68) es equivalente a:

$$20 \quad s = \delta \frac{L}{T_r} \frac{M+1}{2} - L \left(1 - \frac{T_p}{T_r} \right) \quad (69)$$

Obsérvese que s es positivo si $T_{ext} > T_p$ y deberían añadirse muestras, y que s es negativo si $T_{ext} < T_p$ y deberían eliminarse muestras. Por lo tanto, el número de muestras a ser eliminadas o añadidas pueden designarse como $|s|$.

25 Seguidamente se describe el cálculo del índice del último pulso según los ejemplos, haciéndose referencia a las fórmulas (70) - (73).

La posición real del último pulso en la parte periódica construida de la excitación ($T[k]$) determina el número de ciclos de tono completos k , en el que las muestras son eliminadas (o añadidas).

30 En la figura 12 se ilustra una señal de voz antes de la retirada de muestras.

El ejemplo ilustrado en la figura 12, el índice del último pulso k es 2 y hay dos ciclos de tono completos de los cuales han de eliminarse las muestras. En cuanto a los ejemplos descritos con referencia a las fórmulas (64) - (113), el signo 1210 de referencia designa $|s|$.

35 Después de eliminar $|s|$ de la señal de longitud $L - s$, en el que $L = L_{trama}$, o después de añadir $|s|$ muestras a la señal de longitud $L - s$, no hay muestras de la señal original más allá de las $L - s$ muestras. Debe ser evidente que s es positivo si se han de añadir muestras y que s es negativo si han de eliminarse muestras. Por lo tanto, $L - s < L$ si se añaden muestras y $L - s > L$ sin se retiran muestras. Por lo tanto, $T[k]$ ha de estar dentro de $L - s$ muestras y k se determina por lo tanto mediante la siguiente expresión:

$$40 \quad k = i \mid T[i] < L - s \leq T[i+1] \quad (70)$$

A partir de las fórmulas (15b) y (70), se deduce que

$$45 \quad T[0] + kT_r < L - s \leq T[0] + (k+1)T_r \quad (71)$$

Es decir

$$\frac{L - s - T[0]}{T_r} - 1 \leq k < \frac{L - s - T[0]}{T_r} \quad (72)$$

5 Según un ejemplo, es posible determinar k en base a la fórmula (72) como sigue:

$$k = \left\lceil \frac{L - s - T[0]}{T_r} - 1 \right\rceil \quad (73)$$

10 Por ejemplo, en un códec que emplea tramas de por ejemplo por lo menos 20 ms, y empleándose una frecuencia fundamental más baja del habla de por lo menos 40 Hz, en la mayoría de los casos existe un pulso en la trama oculta distinta de SORDA.

A continuación, se describe el cálculo del número de muestras a ser eliminadas en la región mínimas según los ejemplos, haciéndose referencia a las fórmulas (74) a (99).

15 Puede suponerse por ejemplo que Δ_i muestras en cada ciclo de tono i-ésimo entre pulsos serán eliminadas (o añadidas) en el que Δ_i se define como:

$$\Delta_i = \Delta + (i - 1)a, 1 \leq i \leq k \quad (74)$$

20 y en la que a es una variable desconocida que puede expresarse por ejemplo en términos de las variables conocida.

Además, puede suponerse por ejemplo que se eliminarán (o añadirán) Δ_0^p muestras antes del primer pulso, en el que Δ_0^p se define como:

$$25 \quad \Delta_0^p = \Delta_0 \frac{T[0]}{T_r} = (\Delta - a) \frac{T[0]}{T_r} \quad (75)$$

Por otra parte, puede suponerse por ejemplo que Δ_{k+1}^p muestras después del último pulso serán eliminadas (o añadidas), en el que Δ_{k+1}^p se define como:

$$30 \quad \Delta_{k+1}^p = \Delta_{k+1} \frac{L - s - T[k]}{T_r} = (\Delta + ka) \frac{L - s - T[k]}{T_r} \quad (76)$$

Las dos últimas oposiciones concuerdan con la fórmula (74) que tiene en cuenta la longitud de los ciclos de tono parciales primero y segundo.

35 El número de muestras a ser eliminadas (o añadidas) en cada ciclo de tono se ha representado esquemáticamente en la figura 13, en la que k = 2. La figura 13 ilustra una representación esquemática de muestras eliminadas en cada ciclo de tono. En cuanto a los ejemplos descritos con referencia a las fórmulas (64) a (113), el signo 1210 de referencia designa |s|.

40 El número total de muestras a ser eliminadas (o añadidas), s, está relacionado con Δ_i según:

$$|s| = \Delta_0^p + \Delta_{k+1}^p + \sum_{i=1}^k \Delta_i \quad (77)$$

45 A partir de las fórmulas (74) a (77) se deduce que:

$$|s| = (\Delta - a) \frac{T[0]}{T_r} + (\Delta + ka) \frac{L - s - T[k]}{T_r} + \sum_{i=1}^k (\Delta + (i - 1)a) \quad (78)$$

La fórmula (78) es equivalente a:

$$\{s\} = (\Delta - a) \frac{T[0]}{T_r} + (\Delta + ka) \frac{L - s - T[k]}{T_r} + k\Delta + a \sum_{i=1}^k (i-1) \quad (79)$$

Además, la fórmula (79) es equivalente a:

5

$$\{s\} = (\Delta - a) \frac{T[0]}{T_r} + (\Delta + ka) \frac{L - s - T[k]}{T_r} + k\Delta + a \frac{k(k-1)}{2} \quad (80)$$

Por otra parte, la fórmula (80) es equivalente:

$$\{s\} = \Delta \left(\frac{T[0]}{T_r} + \frac{L - s - T[k]}{T_r} + k \right) + a \left(k \frac{L - s - T[k]}{T_r} - \frac{T[0]}{T_r} + \frac{k(k-1)}{2} \right) \quad (81)$$

Además, tomar en cuenta la fórmula (16b) en la fórmula (81) es equivalente a:

15

$$\{s\} = \Delta \left(\frac{L - s}{T_r} \right) + a \left(k \frac{L - s - T[k]}{T_r} - \frac{T[0]}{T_r} + \frac{k(k-1)}{2} \right) \quad (82)$$

Según los ejemplos, puede suponerse que el número de muestras a ser eliminadas (o añadidas) en el ciclo de tono completo después del último pulso está dado por:

20

$$\Delta_{k+1} = |T_r - p[M-1]| = |T_r - T_{ext}| \quad (83)$$

A partir de la fórmula (74) y de la fórmula (83), se deduce que:

$$\Delta = |T_r - T_{ext}| - ka \quad (84)$$

25

A partir de la fórmula (82) y de la fórmula (84), se deduce que:

$$\{s\} = (|T_r - T_{ext}| - ka) \left(\frac{L - s}{T_r} \right) + a \left(k \frac{L - s - T[k]}{T_r} - \frac{T[0]}{T_r} + \frac{k(k-1)}{2} \right) \quad (85)$$

La fórmula (85) es equivalente a:

30

$$\{s\} = |T_r - T_{ext}| \left(\frac{L - s}{T_r} \right) + a \left(-k \frac{L - s}{T_r} + k \frac{L - s - T[k]}{T_r} - \frac{T[0]}{T_r} + \frac{k(k-1)}{2} \right) \quad (86)$$

Además, la fórmula (86) es equivalente a:

35

$$\{s\} = |T_r - T_{ext}| \left(\frac{L - s}{T_r} \right) + a \left(-k \frac{T[k]}{T_r} - \frac{T[0]}{T_r} + \frac{k(k-1)}{2} \right) \quad (87)$$

Por otra parte, la fórmula (87) es equivalente a:

40

$$\{s\} T_r = |T_r - T_{ext}| (L - s) + a \left(-k T[k] - T[0] + \frac{k(k-1)}{2} T_r \right) \quad (88)$$

A partir de las fórmulas (16b) y (88), se deduce que:

$$\{s\} T_r = |T_r - T_{ext}| (L - s) + a \left(-k T[0] - k^2 T_r - T[0] + \frac{k(k-1)}{2} T_r \right) \quad (89)$$

45

La fórmula (89) es equivalente a:

$$|s|T_r = |T_r - T_{ext}|(L - s) + a \left(-(k + 1)T[0] - \frac{k(k + 1)}{2}T_r \right) \quad (90)$$

Además, la fórmula (90) es equivalente a:

$$|s|T_r - |T_r - T_{ext}|(L - s) = a \left(-(k + 1)T[0] - \frac{k(k + 1)}{2}T_r \right) \quad (91)$$

Por otra parte, la fórmula (91) es equivalente a:

$$|s|T_r - |T_r - T_{ext}|(L - s) = -(k + 1)a \left(T[0] + \frac{k}{2}T_r \right) \quad (92)$$

Además, la fórmula (92) es equivalente a:

$$|T_r - T_{ext}|(L - s) - |s|T_r = (k + 1)a \left(T[0] + \frac{k}{2}T_r \right) \quad (93)$$

A partir de la fórmula (93), se deduce que:

$$a = \frac{|T_r - T_{ext}|(L - s) - |s|T_r}{(k + 1) \left(T[0] + \frac{k}{2}T_r \right)} \quad (94)$$

Por lo tanto, por ejemplo, en base a la fórmula (94), según los ejemplos:

- se calcula cuántas muestras han de ser eliminadas y/o añadidas antes del primer pulso; y/o
- se calcula cuántas muestras han de ser eliminadas y/o entre pulsos; y/o
- se calcula cuántas muestras han de ser eliminadas y/o añadidas después del último pulso.

Según algunos ejemplos, las muestras pueden ser por ejemplo, eliminadas o añadidas en las regiones de energía mínima.

A partir de las fórmulas (85) y (94) se deduce que:

$$\Delta_0^p = (\Delta - a) \frac{T[0]}{T_r} = (|T_r - T_{ext}| - ka - a) \frac{T[0]}{T_r} \quad (95)$$

La fórmula (95) es equivalente a:

$$\Delta_0^p = (|T_r - T_{ext}| - (k + 1)a) \frac{T[0]}{T_r} \quad (96)$$

Además, a partir de las fórmulas (84) y (94), se deduce que:

$$\Delta_i = \Delta + (i - 1)a = |T_r - T_{ext}| - ka + (i - 1)a, 1 \leq i \leq k \quad (97)$$

La fórmula (97) es equivalente a:

$$\Delta_i = |T_r - T_{ext}| - (k + 1 - i)a, 1 \leq i \leq k \quad (98)$$

Según un ejemplo, el número de muestras a ser eliminadas después del último pulso puede calcularse en base a la fórmula (97) según:

$$\Delta_{k+1}^p = |s| - \Delta_0^p - \sum_{i=1}^k \Delta_i \quad (99)$$

5 Cabe observar que, según los ejemplos, Δ_0^p , Δ_i y Δ_{k+1}^p son positivos y que el signo de s determina si las muestras han de ser añadidas o eliminadas.

Debido a razones de complejidad, en algunos ejemplos, se desea añadir o eliminar un número entero de muestras y por lo tanto, en tales ejemplos Δ_0^p , Δ_i y Δ_{k+1}^p pueden, por ejemplo redondearse. En otros ejemplos, pueden utilizarse otros conceptos basados en la interpolación de longitud de onda, a título alternativo o adicional, para evitar el redondeo, pero con una mayor complejidad.

A continuación, se describe un algoritmo para la resincronización de pulsos según los ejemplos, haciéndose referencia a las fórmulas (100) a (113).

15 Según los ejemplos, los parámetros de entrada de un algoritmo de este tipo pueden ser, por ejemplo:

L - longitud de la trama

M - número de subtramas

20

T_p - longitud del ciclo de tono al final de la última trama recibida

T_{ext} - longitud del ciclo de tono al final de la trama oculta

25

src_exc - señal de excitación de entrada que fue creada mediante el copiado del último ciclo de tono de bajo paso filtrado de la señal de excitación desde el extremo de la última trama recibida, como descrito en lo que precede.

dst_exc - señal de excitación de salida creada a partir de src_exc mediante la utilización del algoritmo descrito en el presente documento para la resincronización del pulso.

30

Según los ejemplos, un algoritmo de este tipo puede comprender una o más o la totalidad de las etapas siguientes:

- Calcular cambio de tono por subtrama en base a la fórmula (65):

$$35 \quad \delta = \frac{T_{ext} - T_p}{M} \quad (100)$$

- Calcular el tono de inicio redondeado en base a la fórmula (15b):

$$40 \quad T_r = [T_p + 0,5] \quad (101)$$

- Calcular el número de muestras a ser añadidas (a ser eliminadas si negativo) en base a la fórmula (69):

$$s = \delta \frac{L M + 1}{T_r} - L \left(1 - \frac{T_p}{T_r} \right) \quad (102)$$

45 - Encontrar la ubicación del primer pulso máximo $T\{0\}$ entre las primeras T_r muestras en la parte periódica construida del src_exc de excitación.

- Obtener el índice del último pulso en el dst_exc de trama resincronizada en base a la fórmula (73):

$$50 \quad k = \left\lfloor \frac{L - s - T\{0\}}{T_r} - 1 \right\rfloor \quad (103)$$

- Calcular Ω - el delta de las muestras a ser añadidas o eliminadas entre ciclos consecutivos en base a la fórmula (94):

$$a = \frac{|T_r - T_{ext}|(L - s) - |s|T_r}{(k + 1)\left(T[0] + \frac{k}{2}T_r\right)} \quad (104)$$

- Calcular el número de muestras a ser añadidas o eliminadas antes del primer pulso en base a la fórmula (96):

$$\Delta_0^p = (|T_r - T_{ext}| - (k + 1)a) \frac{T[0]}{T_r} \quad (105)$$

- Redondear hacia abajo el número de muestras a ser añadidas o eliminadas antes del primer pulso y mantener en memoria la parte fraccional:

$$\Delta'_0 = \lfloor \Delta_0^p \rfloor \quad (106)$$

$$F = \Delta_0^p - \Delta'_0 \quad (107)$$

- Para cada región entre 2 pulsos, calcular el número de muestras a ser añadidas o eliminadas en base a la fórmula (98):

$$\Delta_i = |T_r - T_{ext}| - (k + 1 - i)a, 1 \leq i \leq k \quad (108)$$

- Redondear hacia abajo del número de muestras a ser añadidas o eliminadas entre 2 pulsos, tomando en cuenta la parte fraccional remanente del redondeo anterior:

$$\Delta'_i = \lfloor \Delta_i + F \rfloor \quad (109)$$

$$F = \Delta_i - \Delta'_i \quad (110)$$

- Si debido al F añadido para algún i sucede que $\Delta'_i > \Delta'_{i-1}$, conmutar los valores para Δ'_i y Δ'_{i-1} .

- Calcular el número de muestras a ser añadidas o eliminadas después del primer pulso en base a la fórmula (99):

$$\Delta'_{k+1} = \lfloor |s + 0,5| \rfloor - \sum_{i=0}^k \Delta'_i \quad (111)$$

- Seguidamente, calcular el número máximo de muestras a ser añadidas o eliminadas entre las regiones de mínima energía:

$$\Delta'_{max} = \max_i \Delta'_i = \begin{cases} \Delta'_k, \Delta'_k \geq \Delta'_{k+1} \\ \Delta'_{k+1}, \Delta'_k < \Delta'_{k+1} \end{cases} \quad (112)$$

- Hallar la ubicación del segmento de mínima energía $P_{min}[1]$ entre los primeros dos pulsos en src_exc , que tiene una longitud Δ'_{max} . Para cada segmento de energía mínima consecutivo entre dos pulsos, la posición se calcula mediante:

$$P_{min}[i] = P_{min}[1] + (i - 1)T_r, 1 < i \leq k \quad (113)$$

- Si $P_{min}[1] > T_r$ entonces calcular la ubicación del segmento de mínima energía antes del primer pulso en src_exc para lo cual se utiliza $P_{min}[0] = P_{min}[1] - T_r$. En caso contrario hallar la ubicación del segmento de mínima energía $P_{min}[0]$ antes del primer pulso en src_exc , que tiene una longitud Δ'_0 .

- Si $P_{min}[1] + kT_r < L - s$ en tal caso calcular la ubicación del segmento de mínima energía después del último pulso en src_exc para lo cual se utiliza $P_{min}[k + 1] = P_{min}[1] + kT_r$. En caso contrario hallar la ubicación del segmento de mínima energía $P_{min}[k + 1]$ después del último pulso en src_exc, que tiene una longitud Δ_{k+1} .
- 5 - Si habrá más de un pulso en la señal de excitación oculta dst_exc, t que es si k es igual a 0, limitar la búsqueda de $P_{min}[1]$ a $L - s$. $P_{min}[1]$ seguidamente apunta a la ubicación del segmento de mínima energía después del último pulso en src_exc.
- Si $s > 0$ añadir Δ_i^s muestras en la ubicación $P_{min}[i]$ para $0 \leq i \leq k + 1$ al src_exc de la señal y almacenarla
- 10 en dst_exc, en caso contrario si $s < 0$ eliminar Δ_i^s muestras en la ubicación $P_{min}[i]$ para $0 \leq i \leq k + 1$ del src_exc de la señal y almacenar en dst_exc. Hay $k + 2$ en el que las regiones se añaden o retiran.

La figura 2c ilustra un sistema para reconstruir una trama que comprende una señal de voz según un ejemplo. El sistema comprende un aparato 10 para determinar un retardo de tono estimado según uno de los ejemplos anteriormente descritos, y un aparato 200 para reconstruir la trama, en el que el aparato para reconstruir la trama está configurado para reconstruir la trama en función del retardo de tono estimado. El retardo de tono estimado es un retardo de tono de la señal de voz.

En un ejemplo, la trama reconstruida puede estar por ejemplo asociada con una o más tramas disponibles, siendo una o más tramas disponibles por lo menos una de una o más trama que preceden a la trama reconstruida y una o más tramas que siguen a la trama reconstruida, en el que la una o más tramas disponibles comprenden uno o más ciclos de tono en forma de uno o más de ciclos de tono disponibles. El aparato 200 para reconstruir la trama puede ser por ejemplo un aparato para reconstruir una trama según uno de los ejemplos anteriormente descritos.

Aunque algunos aspectos han sido descritos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del método correspondiente, en el que un recuadro o dispositivo corresponde a una etapa de método o a un rasgo de una etapa del método. De manera análoga, los aspectos descritos en el contexto de una etapa de método también representan una descripción de un recuadro, ítem o rasgo, correspondientes de un aparato correspondiente.

La señal descompuesta puede ser almacenada en un medio de almacenamiento digital, o puede ser transmitida en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión alámbrico tal como Internet.

En función de determinados requerimientos para la implementación, las realizaciones de la invención pueden ser implementadas en hardware o en software. La implementación puede llevarse a cabo mediante un medio de almacenamiento digital, por ejemplo, un disquete, un DVD, un CD, un ROM; un PROM, un EPROM, un EEPROM, o una memoria flash, que tienen señales de control electrónicamente legibles almacenadas en ellos, que cooperan (o que son capaces de cooperar) con un sistema de ordenador programable de manera tal que el respectivo método se lleve a cabo.

Algunas realizaciones según la invención comprenden un portador de datos no transitorio que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema de ordenador programable, de manera tal que se lleve a cabo uno de los métodos de la invención descritos en el presente documento.

En términos generales, las realizaciones de la presente invención pueden implementarse como un producto programa de ordenador con un código de programa, siendo el código del programa operativo para llevar a cabo uno de los métodos cuando el producto programa de ordenador es ejecutado en un ordenador. El código de programa puede por ejemplo almacenarse en un portador legible por máquina.

Otras realizaciones comprenden el programa de ordenador para llevar a cabo uno de los métodos de la invención descritos en el presente documento, almacenado en un portador legible por máquina.

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

Por lo tanto, otra realización de los métodos inventivos consiste en un portador de datos (o en un medio de almacenamiento digital, o en un medio legible por ordenador) que comprende, grabado en él, el programa de ordenador para llevar a cabo uno de los métodos descritos en el presente documento.

Por lo tanto, otra realización del método inventivo consiste en una corriente de datos o en una secuencia de señales

que representa el programa de ordenador para llevar a cabo uno de los métodos descritos en el presente documento. La corriente de datos o la secuencia de señales pueden estar configuradas por ejemplo para ser transferida por intermedio de una conexión para la comunicación de datos, por ejemplo, por intermedio de Internet.

- 5 Otra realización comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los métodos descritos en el presente documento.

Otra realización comprende un ordenador que tiene instalado en el mismo el programa de ordenador para llevar a cabo uno de los métodos descritos en el presente documento.

10

En algunas realizaciones, puede utilizarse un dispositivo lógico programable (por ejemplo, una matriz de puerta de campo programable) para llevar a cabo algunas de las funcionalidades, o todas ellas, de los métodos descritos en el presente documento. En algunas realizaciones, una matriz de puerta de campo programable puede cooperar con un microprocesador a efectos de llevar a cabo uno de los métodos descritos en el presente documento. En términos generales es preferible que los métodos sean llevados a cabo mediante cualquier aparato de hardware.

15

Las realizaciones anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se da por entendido que a la persona con pericia en la técnica le serán evidentes modificaciones y variaciones en cuanto a las disposiciones y detalles descritos en el presente documento. Por ello la finalidad es la que la invención quede limitada solamente por los alcances de las reivindicaciones de patente pendientes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en el presente documento.

20

Referencias

- 25 [3GP09] 3GPP; Technical Specification Group Services and System Aspects, Extended adaptive multi-rate - wideband (AMR-WB+) codec, 3GPP TS 26.290, 3rd Generation Partnership Project, 2009.

[3GP12a], Adaptive multi-rate (AMR) speech codec; error concealment of lost frames (release 11), 3GPP TS 26.091, 3rd Generation Partnership Project, Sep 2012.

30

[3GP12b], Speech codec speech processing functions; adaptive multi-rate - wideband (AMRWB) speech codec; error concealment of erroneous or lost frames, 3GPP TS 26.191, 3rd Generation Partnership Project, Sep 2012.

[Gao] Yang Gao, Pitch prediction for packet loss concealment, patente europea 2 002 427 B1.

35

[ITU03] ITU-T, Wideband coding of speech at around 16 kbit/s using adaptive multi-rate wideband (amr-wb), Recommendation ITU-T G.722.2, Telecommunication Standardization Sector of ITU, Jul 2003.

- 40 [ITU06a], G.722 Appendix III: A high-complexity algorithm for packet loss concealment for G.722, ITU-T Recommendation, ITU-T, Nov 2006.

[ITU06b], G.729.1: G.729-based embedded variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with g.729, Recommendation ITU-T G.729.1, Telecommunication Standardization Sector of ITU, Mayo 2006.

45

[ITU07], G.722 Appendix IV: A low-complexity algorithm for packet loss concealment with G.722, ITU-T Recommendation, ITU-T, Ag 2007.

[ITU08a], G.718: Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s, Recommendation ITU-T G.718, Telecommunication Standardization Sector of ITU, Jun 2008.

50

[ITU08b], G.719: Low-complexity, full-band audio coding for high-quality, conversational applications, Recommendation ITU-T G.719, Telecommunication Standardization Sector of ITU, Jun 2008.

- 55 [ITU12], G.729: Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (cs-acelp), Recommendation ITU-T G.729, Telecommunication Standardization Sector of ITU, Junio 2012.

[MCZ11] Xinwen Mu, Hixin Chen, y Yan Zhao, A frame erasure concealment method based on pitch and gain linear prediction for AMR-WB codec, Consumer Electronics (ICCE), 2011 IEEE International Conference on, Enero 2011, páginas 815-816.

60

[MTTA90] J.S. Marques, I. Trancoso, J.M. Tribolet, and L.B. Almeida, Improved pitch prediction with fractional delays in celp coding, Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on, 1990, páginas 665-668 vol.2.

[VJGS12] Tommy Vaillancourt, Milan Jelinek, Philippe Gournay, and Redwan Salami, Method and device for efficient frame erasure concealment in speech codecs, documento US 8.255.207 B2, 2012.

5

REIVINDICACIONES

1. Aparato para la determinación de un retardo de tono estimado, que comprende:

5 una interfaz (110) de entrada para recibir una pluralidad de valores de retardo de tono originales, y

un estimador (120) de retardo de tono para estimar el retardo de tono estimado,

10 en el que el estimador (120) de retardo de tono se configura para estimar el retardo de tono estimado en función de una pluralidad de valores de retardo de tono originales y en función de una pluralidad de valores de información, en el que para cada valor de retardo de tono original de la pluralidad de valores de retardo de tono originales, un valor de información de la pluralidad de valores de información se asigna a dicho valor de retardo de tono original,

15 en el que el estimador (120) de retardo de tono se configura para estimar el retardo de tono estimado en función de la pluralidad de valores de retardo de tono originales y en función de una pluralidad de valores de ganancia de tono como la pluralidad de valores de información,

20 en el que para cada valor de retardo de tono original de la pluralidad de valores de retardo de tono originales, un valor de ganancia de tono de la pluralidad de valores de ganancia de tono se asigna a dicho valor de retardo de tono original, en el que el estimador de retardo de tono se configura para estimar el retardo de tono estimado mediante la minimización de una función de error, en el que el estimador de retardo de tono se configura para estimar el retardo de tono estimado mediante la determinación de dos

25 parámetros a , b , mediante la minimización de una función de error

$$err = \sum_{i=0}^k g_p(i) \cdot ((a + b \cdot i) - P(i))^2$$

30 en la que a es un número real, en la que b es un número real, en la que k es un número entero, y en la que $P(i)$ es el valor de retardo de tono original i -ésimo, en la que $g_p(i)$ es el valor de ganancia de tono i -ésimo que se está asignando al valor de retardo de tono i -ésimo $P(i)$, en donde i indica un índice de subtrama, en donde $k \geq 2$.

2. Aparato según la reivindicación 1, en el que cada una de la pluralidad de valores de ganancia de tono es una ganancia de libro de códigos adaptivo.

35 3. Aparato según la reivindicación 1, en el que $k = 4$.

4. Sistema para la reconstrucción de una trama que comprende una señal de habla, comprendiendo el sistema:

40 un aparato según la reivindicación 1 para la determinación de un retardo de tono estimado, y

un aparato para la reconstrucción de la trama, en el que el aparato para la reconstrucción de la trama se configura para reconstruir la trama en función del retardo de tono estimado,

45 en el que el retardo de tono estimado es un retardo de tono de la señal de habla.

5. Sistema para la reconstrucción de una trama según la reivindicación 4,

50 en el que la trama reconstruida se asocia con una o más tramas disponibles, dicha una o más tramas disponibles estando al menos una de una o más tramas anteriores de la trama reconstruida y una o más tramas posteriores de la trama reconstruida, en el que la una o más tramas disponibles comprenden uno o más ciclos de tono como uno o más ciclos de tono disponibles, y

55 en el que el aparato para la reconstrucción de la trama comprende

una unidad (210) de determinación para la determinación de una diferencia de número de muestra que indica una diferencia entre un número de muestras de uno de los uno o más ciclos de tono disponibles y un número de muestras de un primer ciclo de tono a ser reconstruido, y

60 un reconstructor (220) de trama para la reconstrucción de la trama reconstruida mediante la reconstrucción,

5 en función de la diferencia de número de muestra y en función de las muestras de dicho uno de los uno o más ciclos de tono disponibles, el primer ciclo de tono a ser reconstruido como un primer ciclo de tono reconstruido, en el que el reconstructor (220) de trama se configura para reconstruir la trama reconstruida, de manera tal que la trama reconstruida comprende completa o parcialmente el primer ciclo de tono reconstruido, de manera tal que la trama reconstruida comprende completa o parcialmente un segundo ciclo de tono reconstruido, y de manera tal que el número de muestras del primer ciclo de tono reconstruido difiere de un número de muestras del segundo ciclo de tono reconstruido,

10 en el que la unidad (210) de determinación se configura para determinar la diferencia de número de muestra en función del retardo de tono estimado.

6. Método para la determinación de un retardo de tono estimado, que comprende:

15 recibir una pluralidad de valores de retardo de tono originales, y
estimar el retardo de tono estimado,

20 en el que la estimación del retardo de tono estimado se realiza en función de una pluralidad de valores de retardo de tono originales y en función de una pluralidad de valores de información,

en el que para cada valor de retardo de tono original de la pluralidad de valores de retardo de tono originales, un valor de información de la pluralidad de valores de información se asigna a dicho valor de retardo de tono original,

25 en el que la estimación del retardo de tono estimado se realiza en función de la pluralidad de valores de retardo de tono originales y en función de una pluralidad de valores de ganancia de tono como la pluralidad de valores de información, en el que para cada valor de retardo de tono original de la pluralidad de valores de retardo de tono originales, un valor de ganancia de tono de la pluralidad de valores de ganancia de tono se asigna a dicho valor de retardo de tono original, en el que la estimación del retardo de tono estimado se realiza mediante la determinación de dos parámetros a , b , mediante la minimización de la función de error

$$err = \sum_{i=0}^k g_p(i) \cdot ((a + b \cdot i) - P(i))^2$$

35 en la que a es un número real, en la que b es un número real, en la que k es un número entero, y en la que $P(i)$ es el valor de retardo de tono original i -ésimo, en la que $g_p(i)$ es el valor de ganancia de tono i -ésimo que se está asignando al valor de retardo de tono i -ésimo $P(i)$ en donde i indica un índice de subtrama, en donde $k \geq 2$.

40 7. Programa de ordenador adaptado para implementar el método según la reivindicación 6 cuando se ejecuta en un ordenador o procesador de señal.

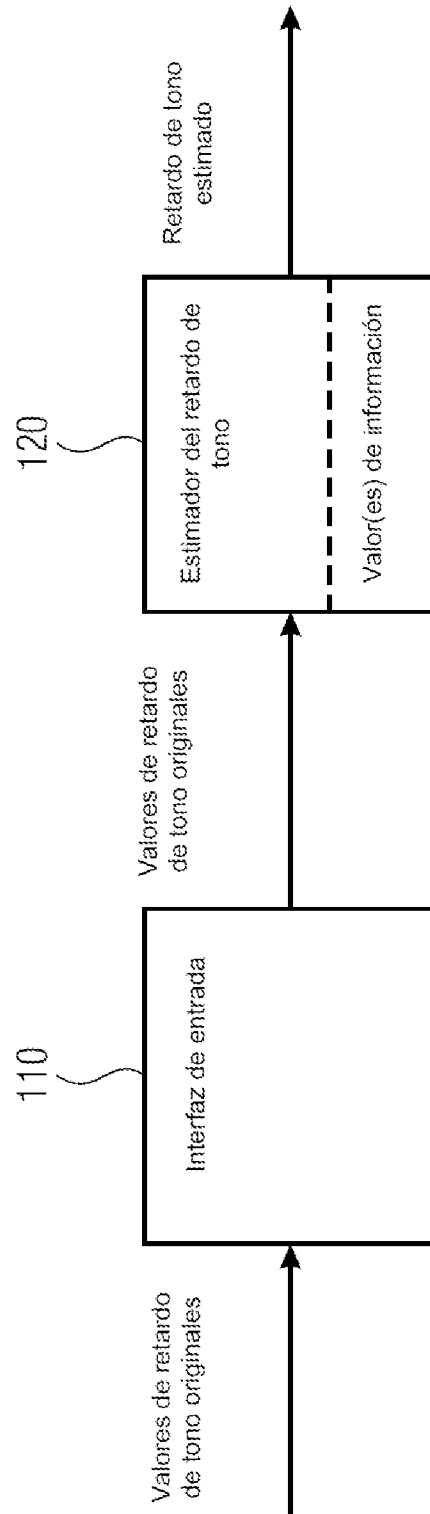


FIG 1

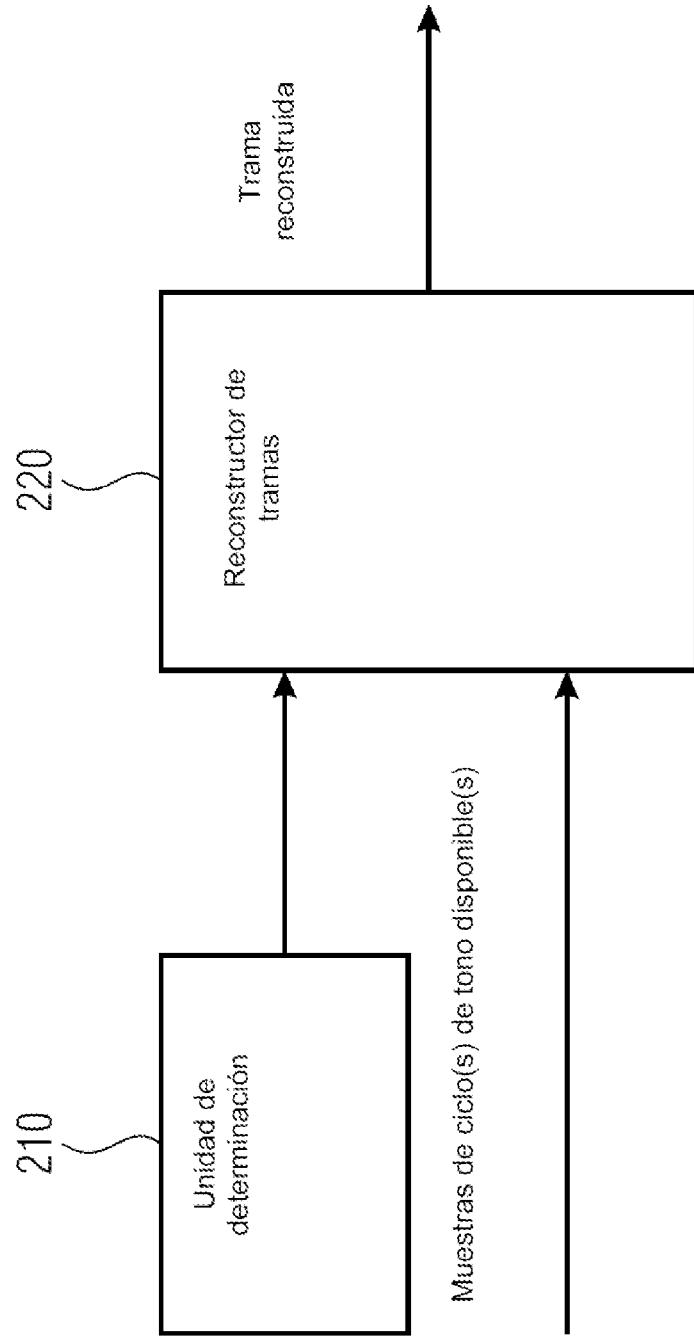


FIG 2A

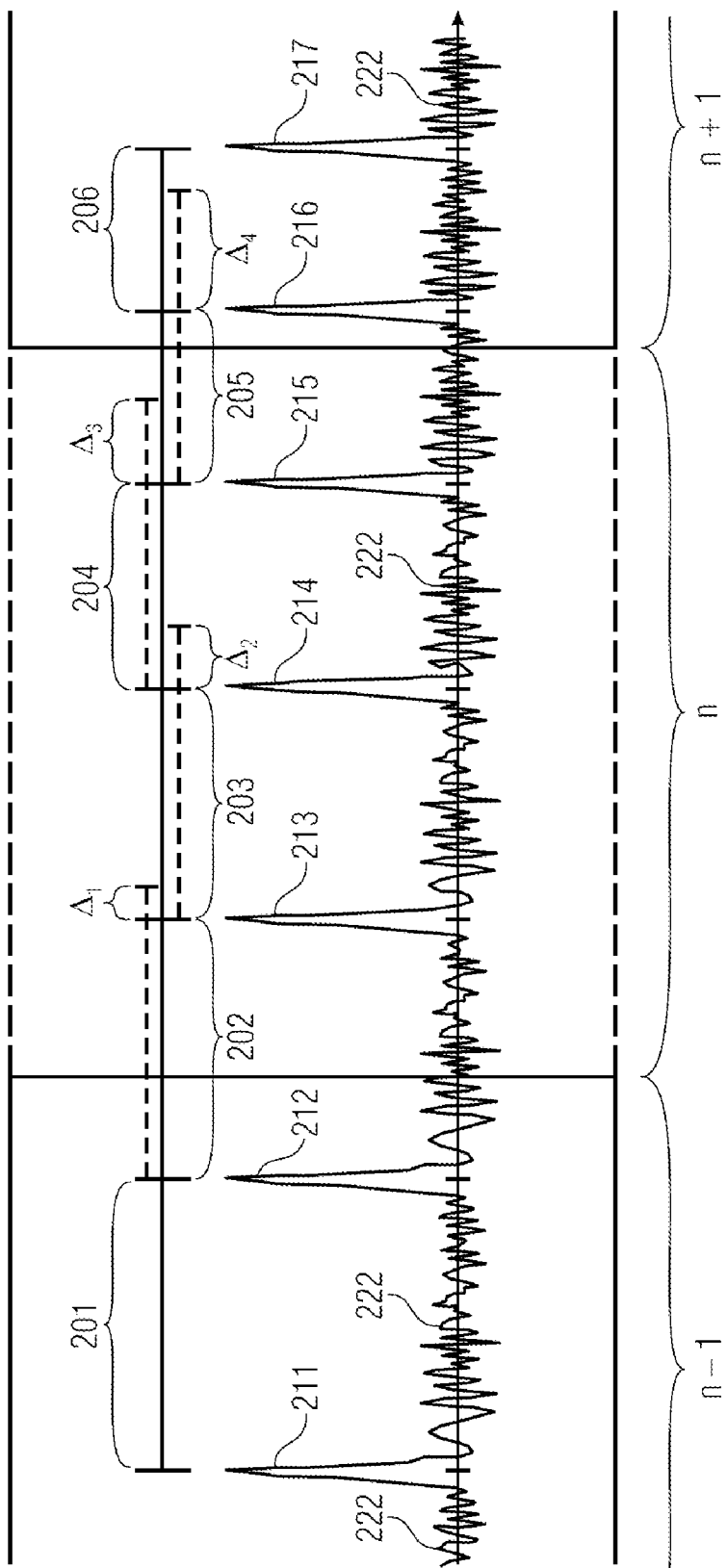


FIG 2B

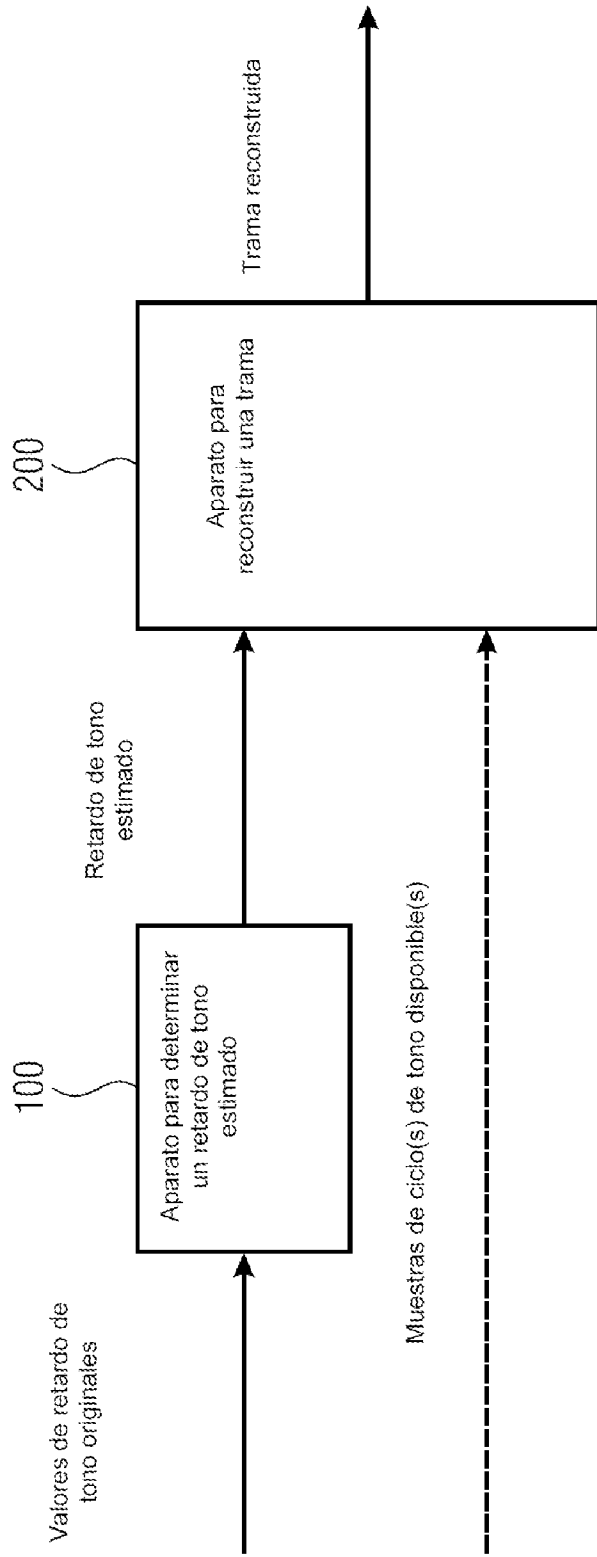


FIG 2C

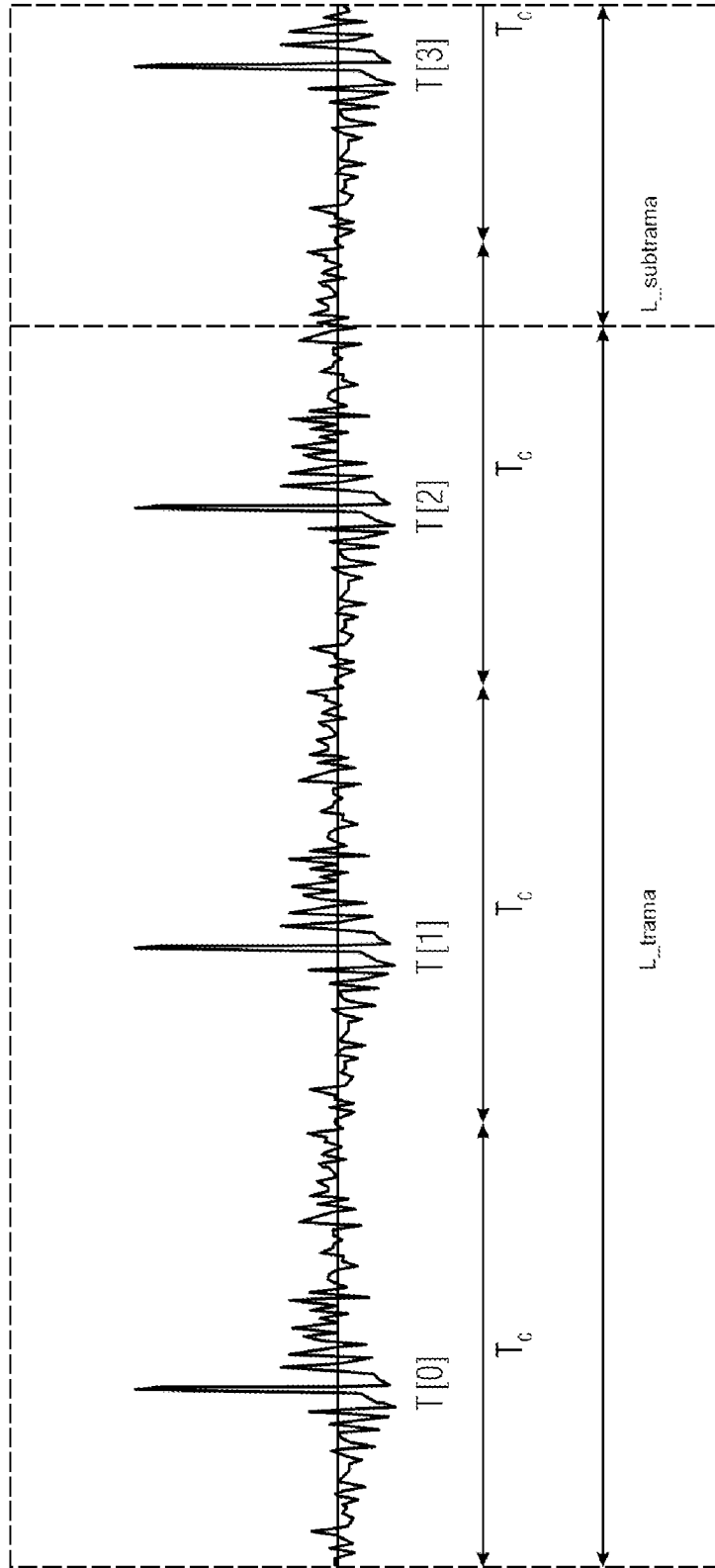


FIG 3

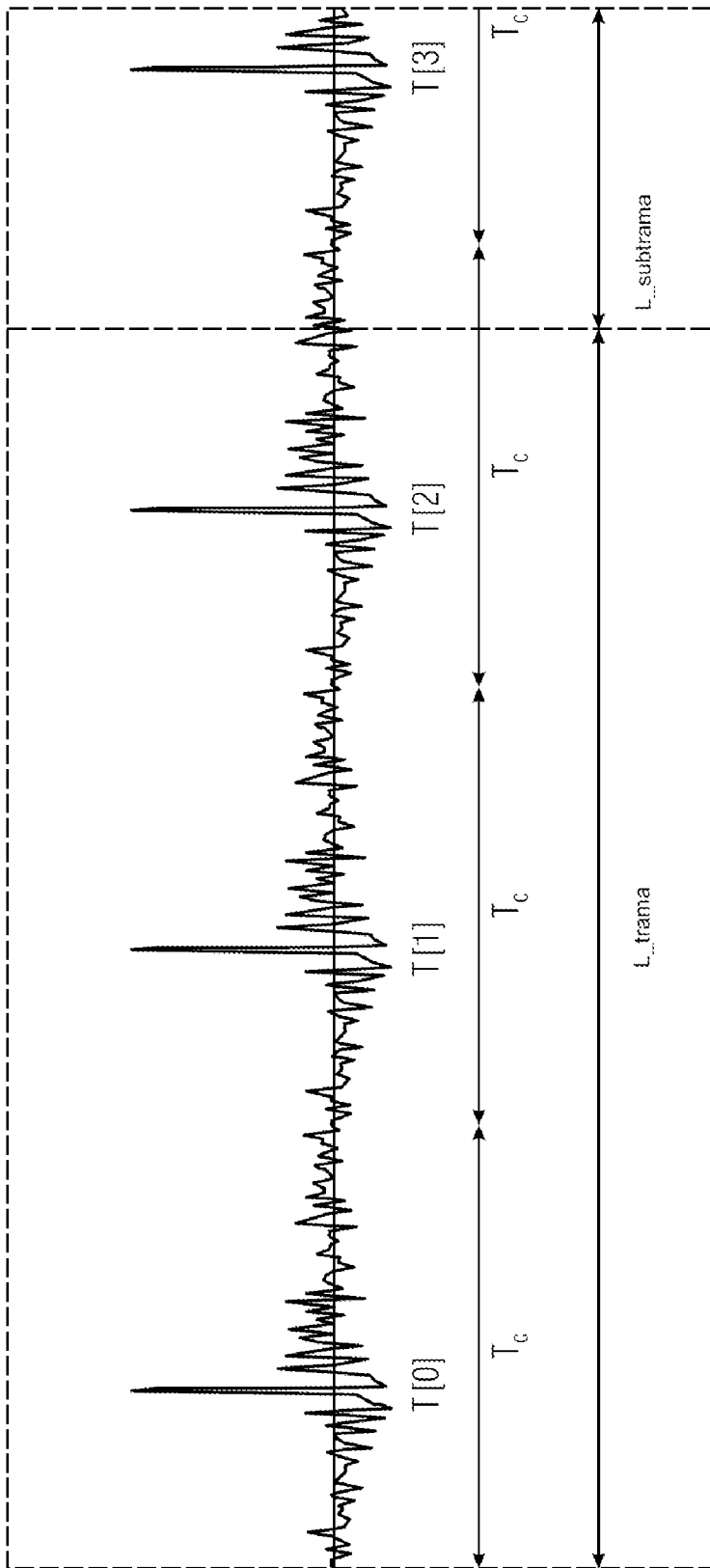


FIG 4

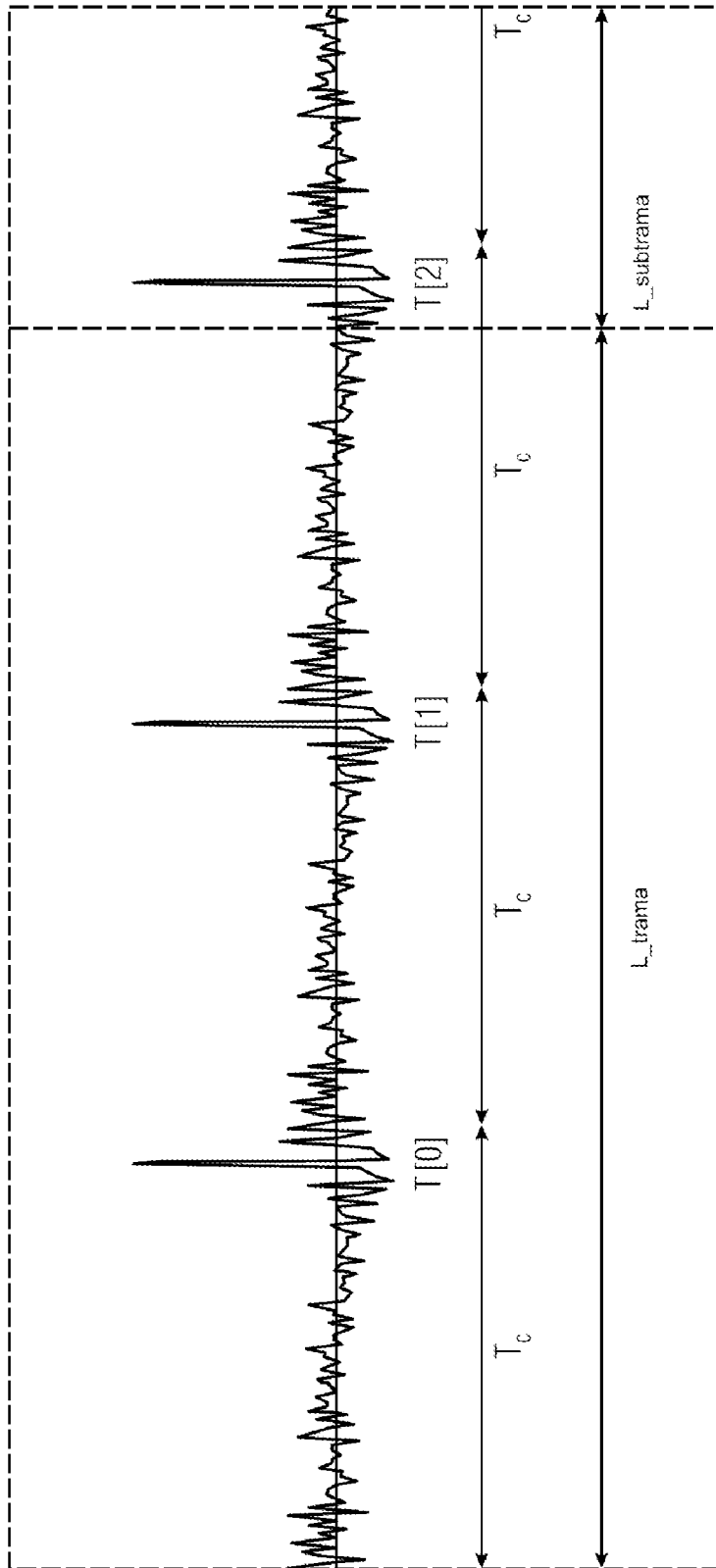


FIG 5

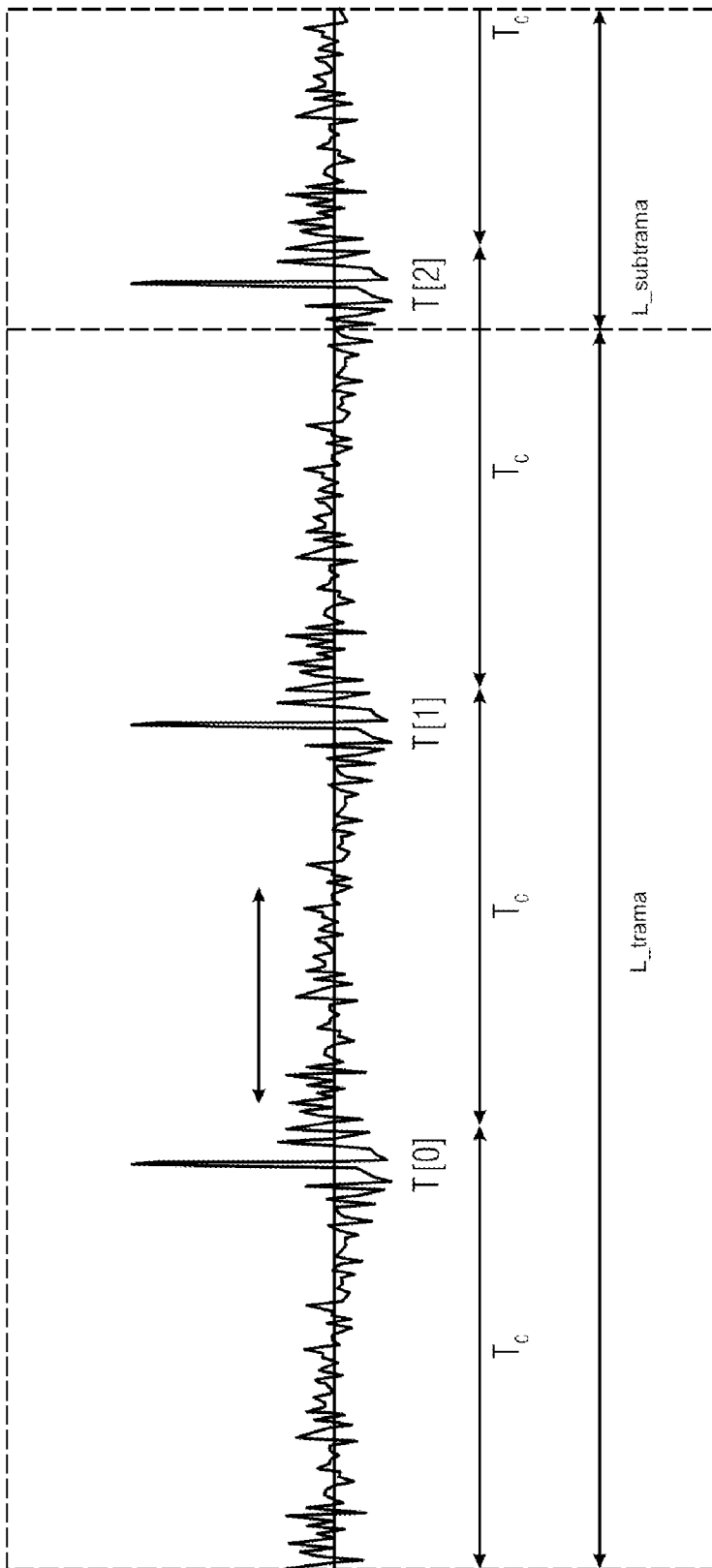


FIG 6

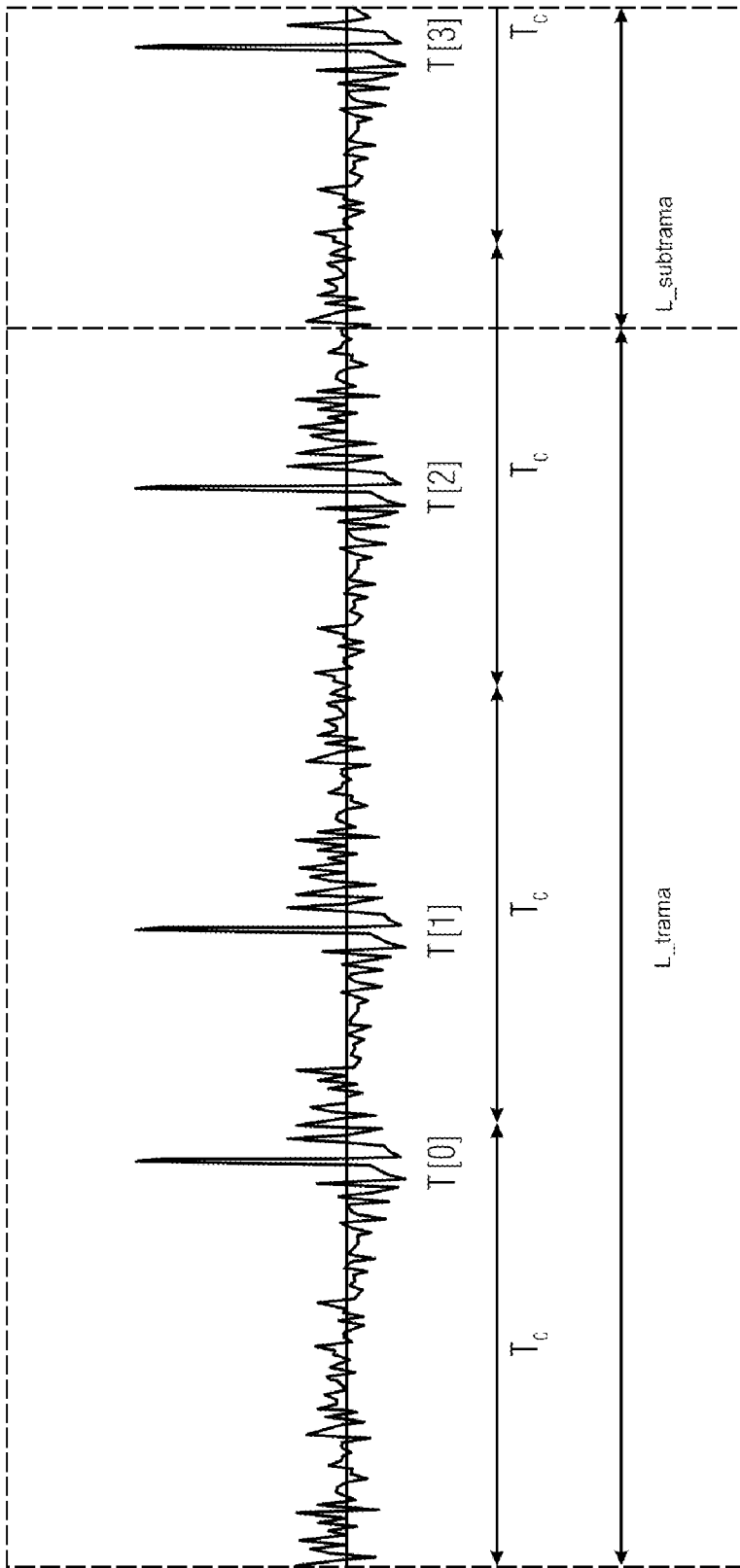


FIG 7

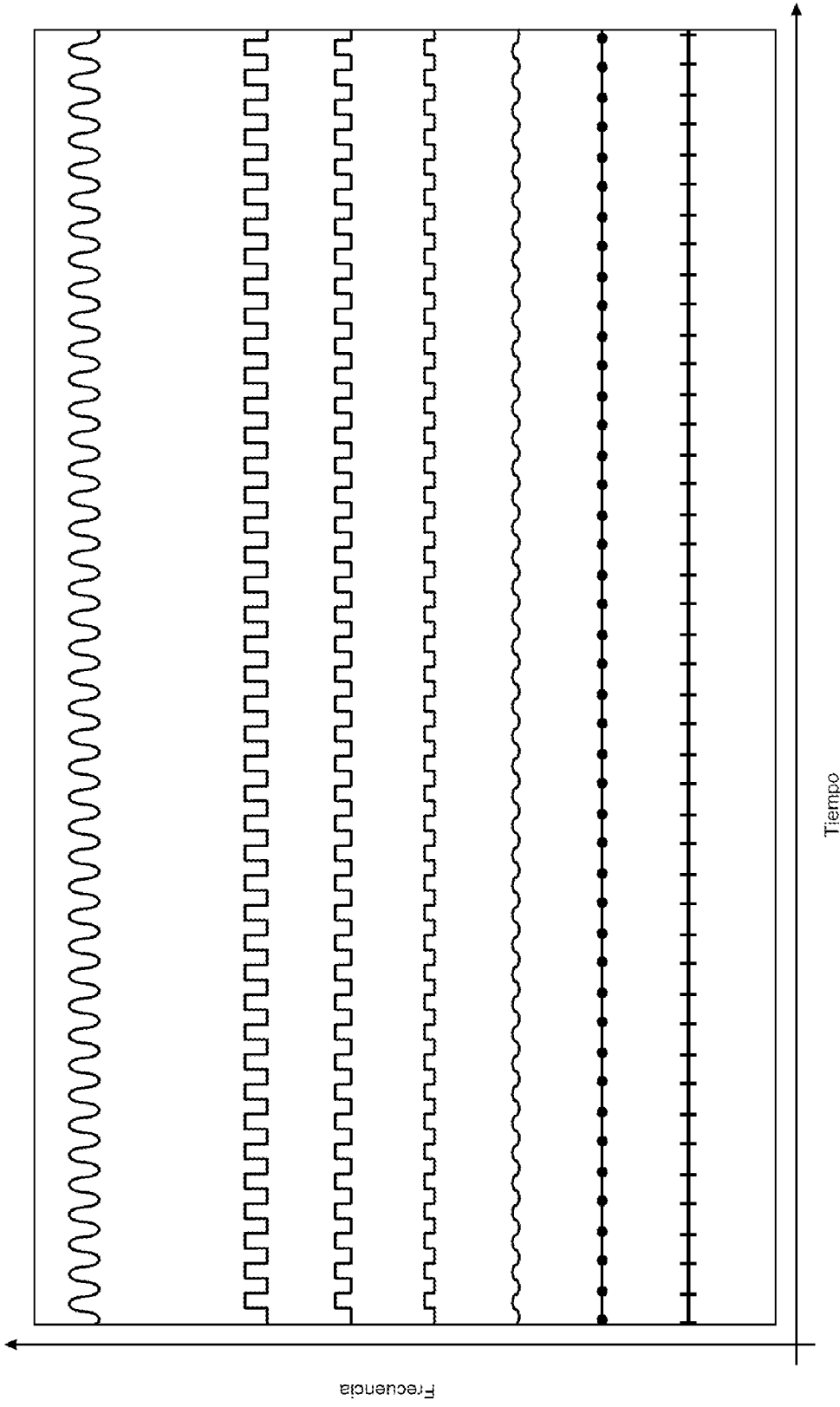


FIG 8

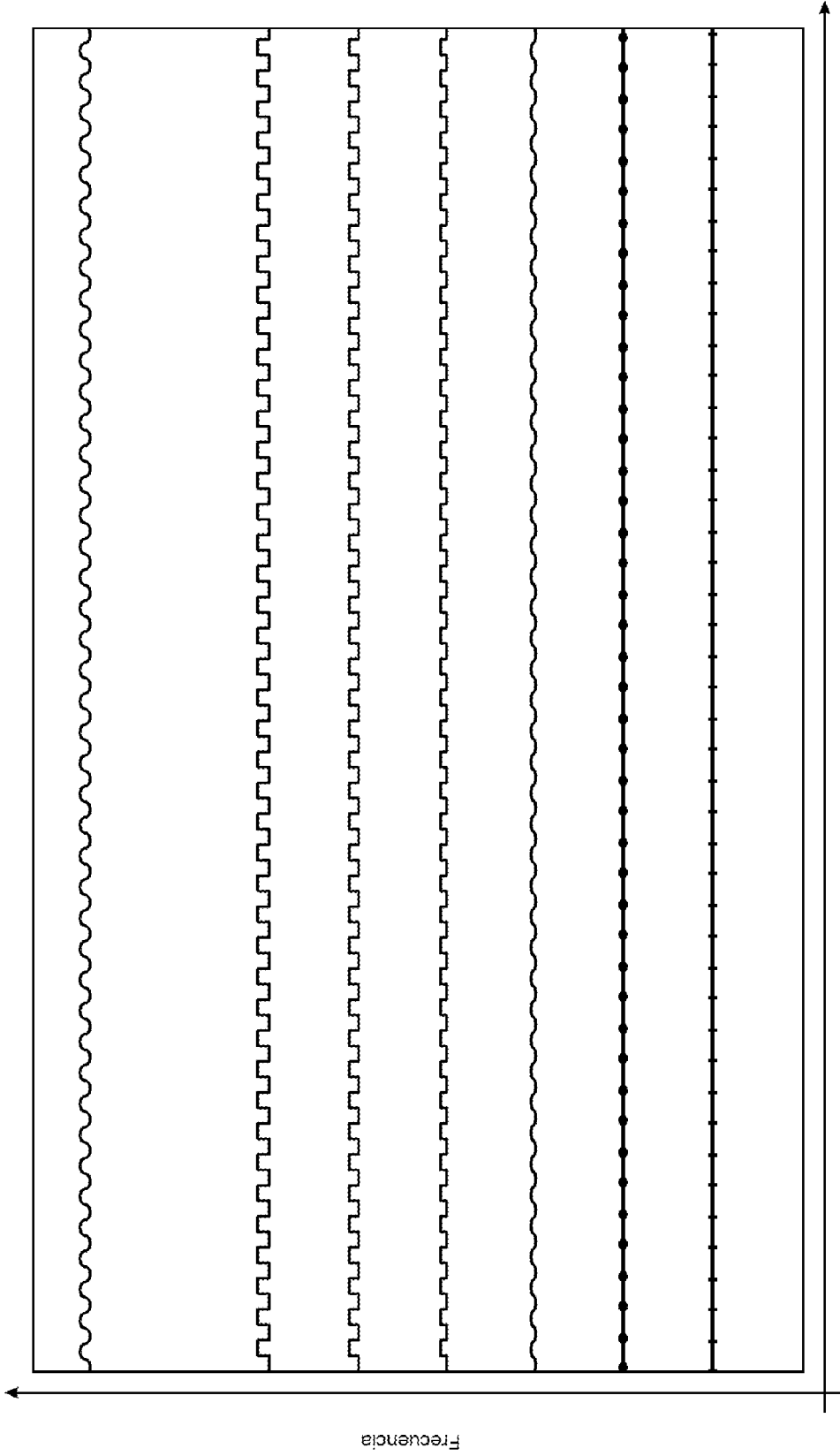


FIG 9

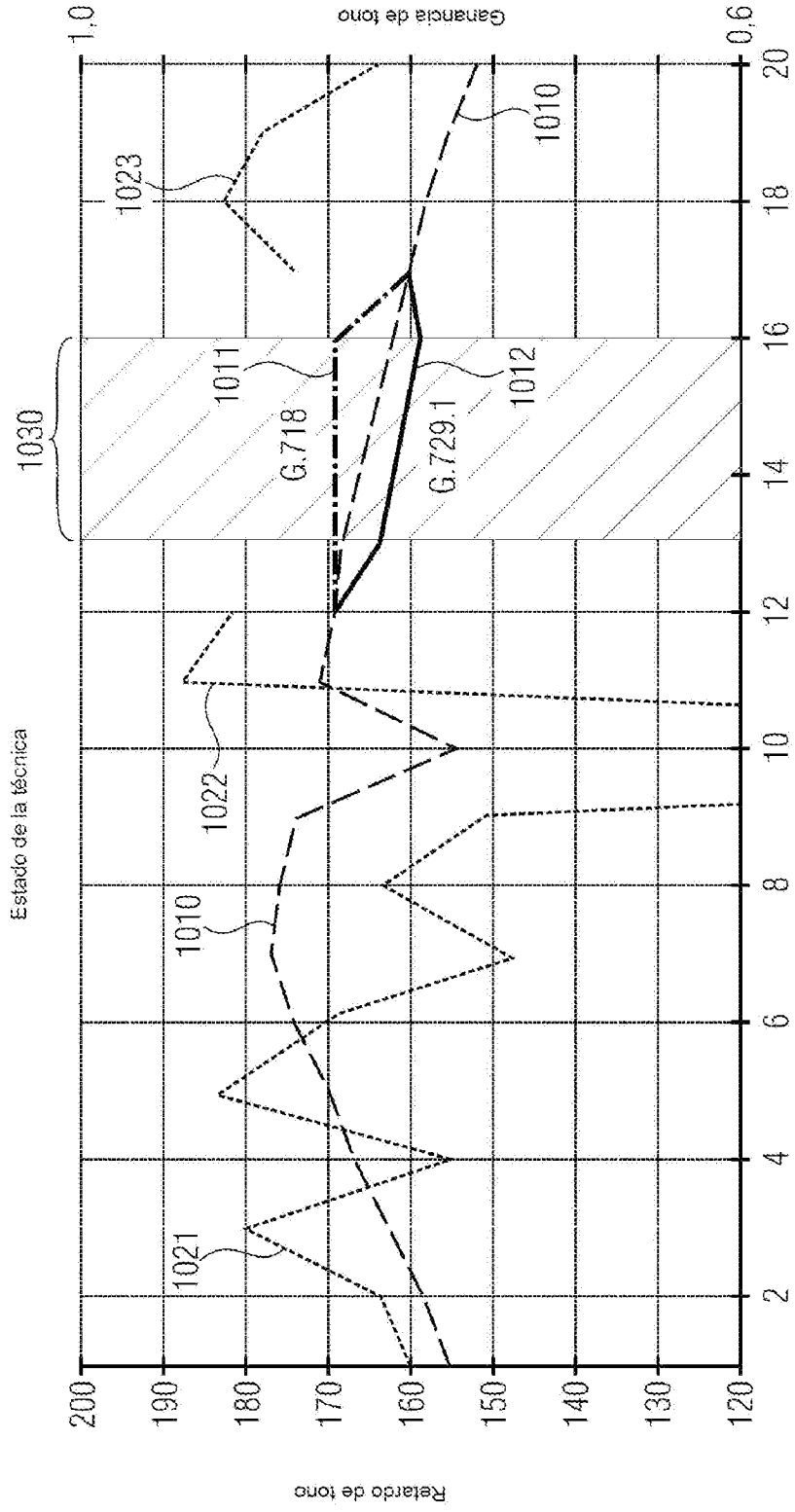


FIG 10

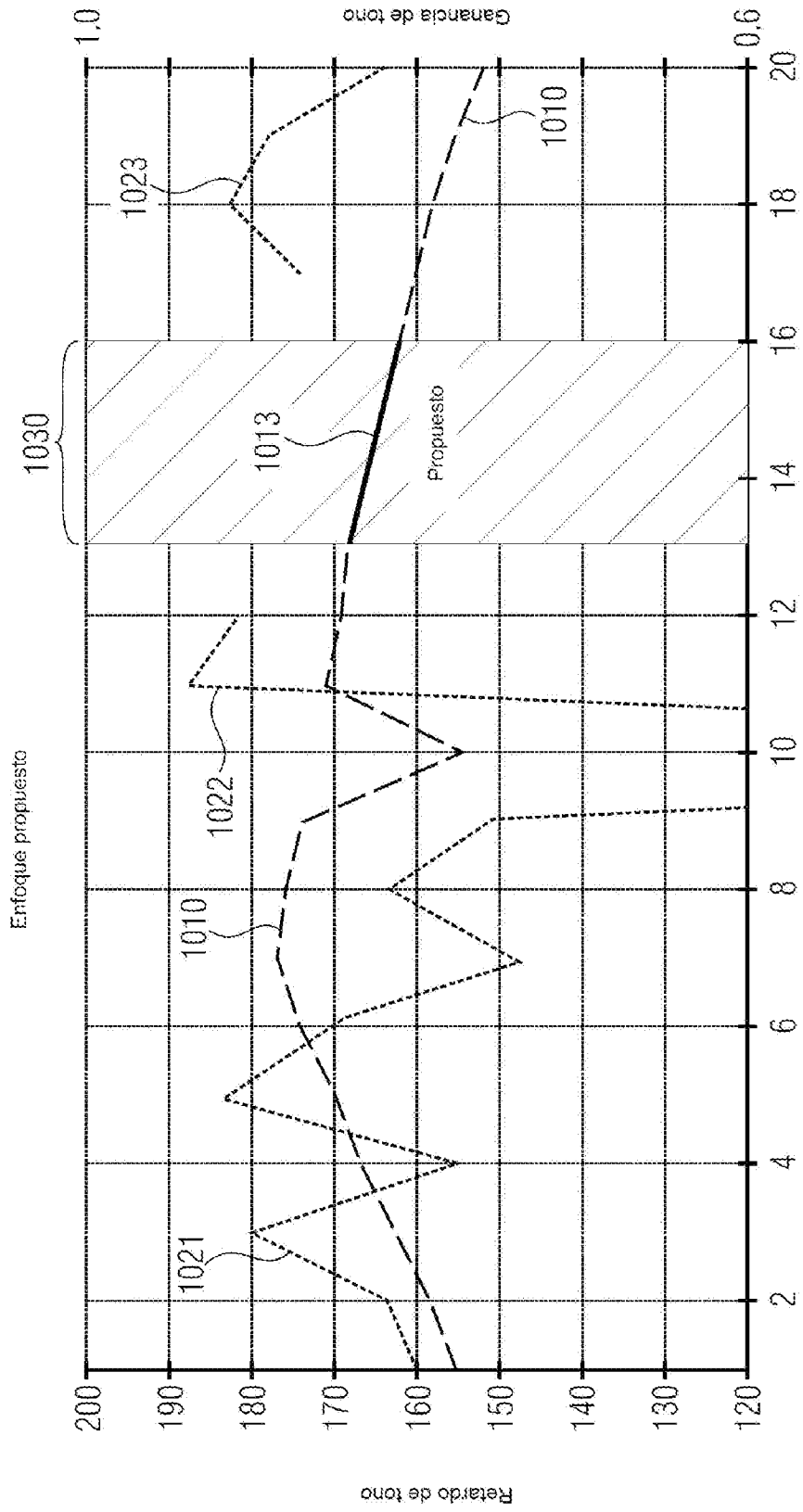


FIG 11

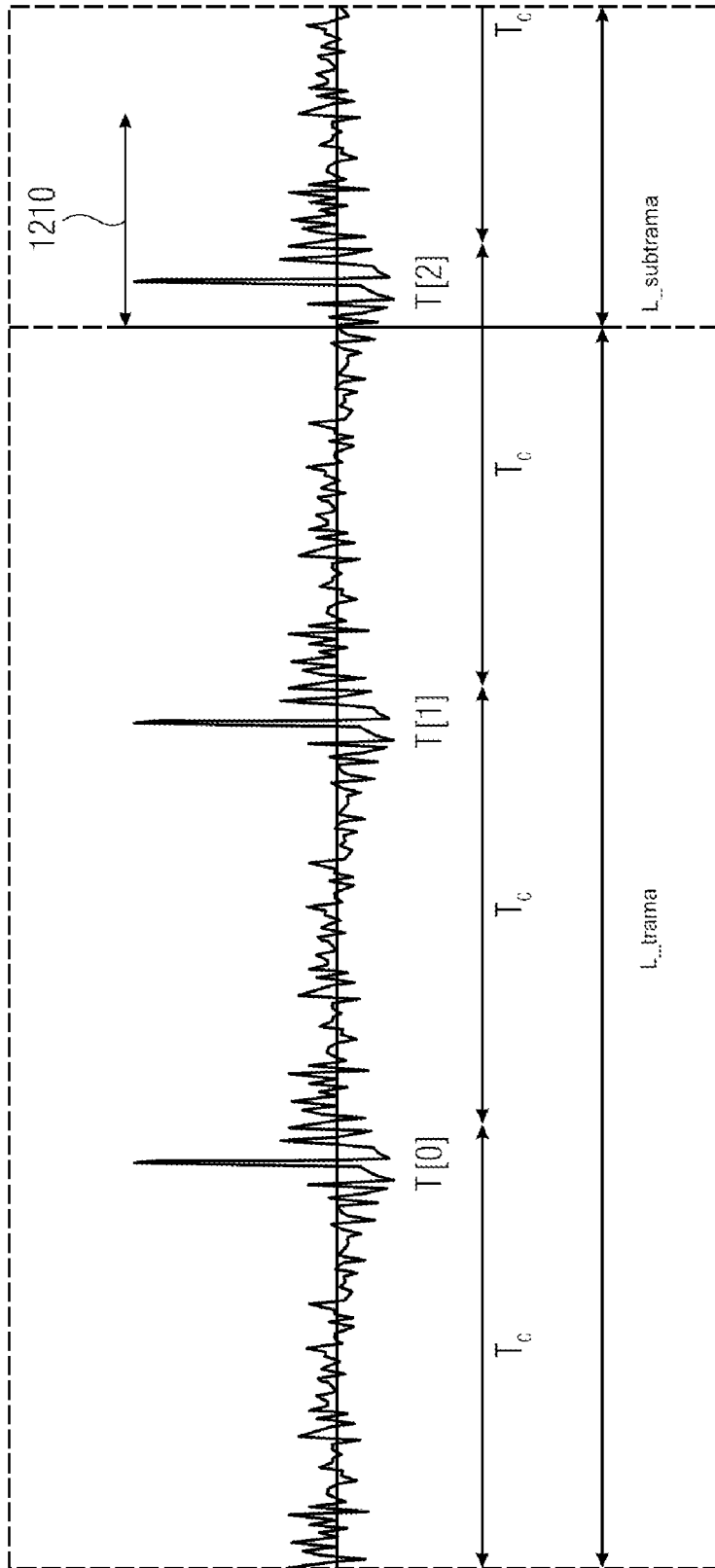


FIG 12

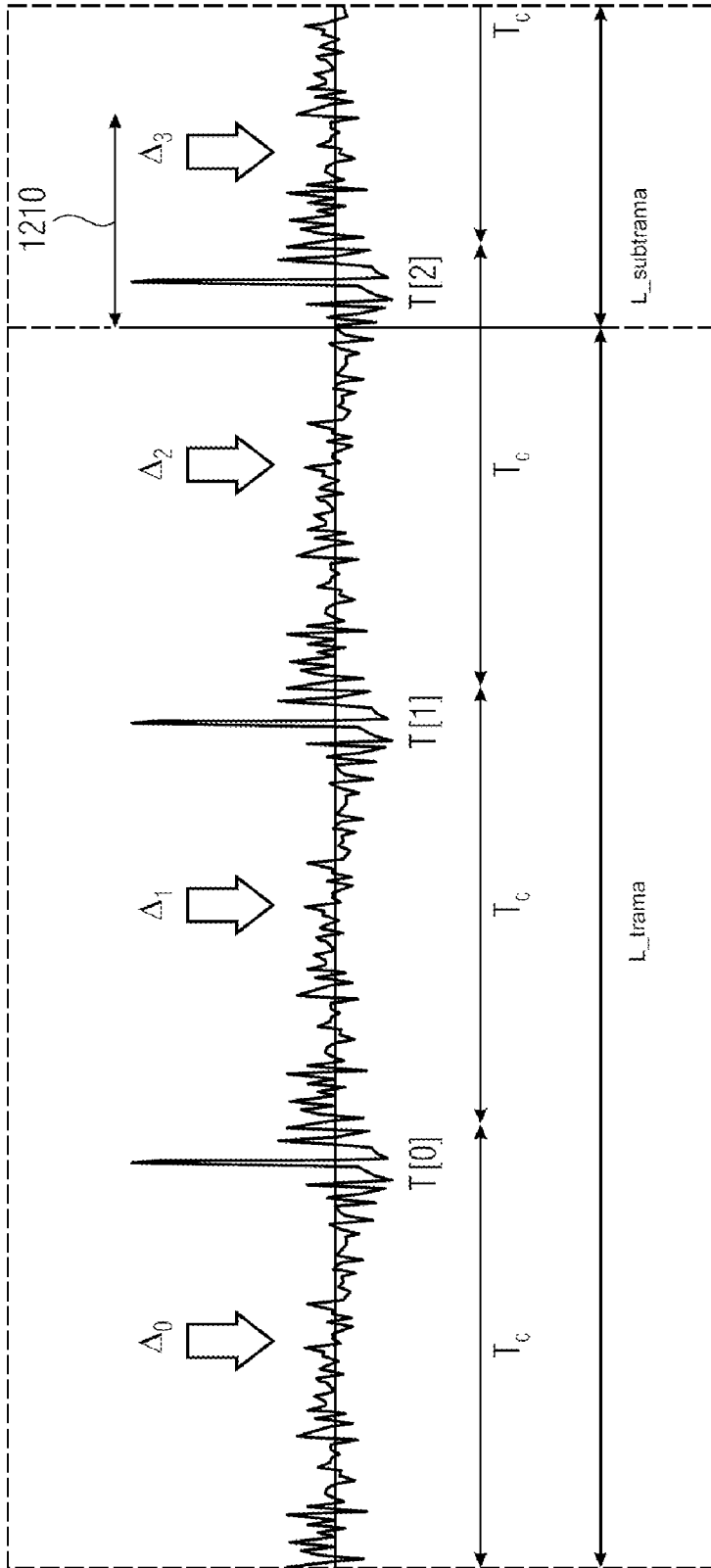


FIG 13