

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 017 157**

51 Int. Cl.:

G10L 19/005 (2013.01)

G10L 19/022 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.05.2020** **PCT/EP2020/064394**

87 Fecha y número de publicación internacional: **17.12.2020** **WO20249380**

96 Fecha de presentación y número de la solicitud europea: **25.05.2020** **E 20728023 (1)**

97 Fecha y número de publicación de la concesión europea: **12.03.2025** **EP 3984026**

54 Título: **Ocultación de errores de subtrama de audio invertida en el tiempo**

30 Prioridad:

13.06.2019 US 201962860922 P

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

12.05.2025

73 Titular/es:

**TELEFONAKTIEBOLAGET LM ERICSSON
(PUBL) (100.00%)
164 83 Stockholm, SE**

72 Inventor/es:

**NORVELL, ERIK y
MORADI ASHOUR, CHAMRAN**

74 Agente/Representante:

ELZABURU, S.L.P

ES 3 017 157 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

Ocultación de errores de subtrama de audio invertida en el tiempo

Campo técnico

5 La presente descripción se refiere en general a comunicaciones, y más particularmente a métodos y aparatos para controlar una ocultación de pérdida de paquetes para codificación y decodificación de audio mono, estéreo o multicanal.

Antecedentes

10 Los servicios de telecomunicaciones modernos generalmente proporcionan conexiones fiables entre los usuarios finales. Sin embargo, estos servicios aún necesitan manejar condiciones de canal variables donde ocasionalmente se pueden perder paquetes de datos debido, por ejemplo, a congestión de la red o cobertura móvil deficiente. Para superar el problema de los errores de transmisión y paquetes perdidos, los servicios de telecomunicaciones pueden hacer uso de técnicas de Ocultación de Pérdida de Paquetes (PLC). En el caso de que los paquetes de datos se pierdan debido a una mala conexión, congestión de la red, etc., la información faltante de los paquetes perdidos en el lado del receptor puede sustituirse en el decodificador por una señal sintética. Las técnicas PLC a menudo pueden estar estrechamente vinculadas al decodificador, donde los estados internos pueden usarse para producir una continuación o extrapolación de señal para cubrir la pérdida de paquetes. Para un códec multimodo que tiene varios modos de funcionamiento para diferentes tipos de señales, a menudo hay varias tecnologías PLC para manejar la ocultación. Hay muchos términos diferentes utilizados para las técnicas de ocultación de pérdida de paquetes, incluyendo Ocultación de Errores de Trama (FEC), Ocultación de Pérdida de Trama (FLC) y Unidad de Ocultación de Errores (ECU).

25 Para los modos de codificación de voz basados en predicción lineal (LP), el PLC puede basarse en el ajuste de las posiciones del pulso glótico usando información estimada de tono de fin de trama y la replicación del ciclo de tono de la trama anterior [1]. La ganancia del predictor de largo plazo (LTP) converge a cero con la velocidad dependiendo del número de tramas perdidas consecutivas y la estabilidad de la última trama buena, es decir, libre de errores [2]. Los modos de codificación basados en el dominio de la frecuencia (FD) están diseñados para manejar señales generales o complejas como la música. Se pueden usar diferentes técnicas dependiendo de las características de la última trama recibida. Dicho análisis puede incluir el número de componentes tonales detectados y la periodicidad de la señal. Si la pérdida de trama ocurre durante una señal altamente periódica como el habla activa o música instrumental simple, un PLC en el dominio del tiempo, similar a PLC basado en LP, puede ser adecuado. En este caso, PLC FD puede imitar un decodificador LP estimando los parámetros LP y una señal de excitación basada en la última trama recibida [2]. En caso de que la trama perdida se produzca durante una señal no periódica o similar al ruido, la última trama recibida puede repetirse en el dominio espectral, donde los coeficientes se multiplican por una señal de signo aleatorio para reducir el sonido metálico de una señal repetida. Para una señal tonal estacionaria, se ha descubierto que resulta ventajoso usar un enfoque basado en la predicción y extrapolación de los componentes tonales detectados. Se pueden encontrar más detalles sobre las técnicas mencionadas anteriormente en [1][2][3].

40 Un método genérico de ocultación de errores que opera en el dominio de la frecuencia es la ECU de Fase (Unidad de Ocultación de Errores) [4]. La ECU de Fase es una herramienta independiente que opera en un búfer de la señal del dominio del tiempo previamente decodificada y reconstruida. El marco de trabajo de la ECU de Fase se basa en el paradigma de análisis y síntesis sinusoidal. En este método, los componentes sinusoidales de la última trama buena pueden extraerse y desplazarse en fase. Cuando se pierde una trama, las frecuencias sinusoidales se obtienen en el dominio DFT (transformada discreta de Fourier) de la síntesis decodificada anterior. Primero, se identifican los contenedores de frecuencias correspondientes al encontrar los picos del plano del espectro de magnitud. Luego, se estiman las frecuencias fraccionarias de los picos usando contenedores de frecuencia pico. Los contenedores de frecuencia correspondientes a los picos junto con los vecinos se desplazan en fase usando frecuencias fraccionarias. Para el resto de la trama, se conserva la magnitud de la síntesis anterior mientras se aleatoriza la fase. El error de ráfaga también se maneja de manera que la señal estimada se silencia suavemente al convergerla a cero. Se pueden encontrar más detalles sobre la ECU de Fase en [4].

50 El concepto de la ECU de Fase se puede usar en decodificadores que funcionan en el dominio de la frecuencia. Este concepto incluye sistemas de codificación y decodificación que realizan la decodificación en el dominio de la frecuencia, como se ilustra en la Figura 1, pero también decodificadores que realizan la decodificación en el dominio del tiempo con procesamiento adicional en el dominio de la frecuencia, como se ilustra en la Figura 2. En la Figura 1, las (sub)tramas de señal de audio de entrada en el dominio del tiempo se dividen en ventanas 100 y se transforman al dominio de la frecuencia mediante la DFT 101. Un codificador 102 realiza la codificación en el dominio de la frecuencia y proporciona parámetros codificados para la transmisión 103. Un decodificador 104 decodifica las tramas recibidas o aplica el PLC 109 en caso de pérdida de una trama. En la construcción de la trama de ocultación, el PLC puede usar una memoria 108 de tramas decodificados previamente. La trama decodificada u oculta se transforma al dominio del tiempo mediante la DFT 110 inversa, y luego la señal de audio de salida se reconstruye mediante la operación de superposición-suma 111. La Figura 2 ilustra un par de codificador y decodificador donde el decodificador aplica una transformación DFT para facilitar el procesamiento en el dominio de la frecuencia. La señal en el dominio

del tiempo recibida y decodificada primero se divide en 105 en ventanas de (sub)tramas y luego se transforma al dominio de la frecuencia mediante la DFT 106 para el procesamiento en el dominio de la frecuencia en 107 que puede realizarse antes o después del PLC 109 (en caso de pérdida de trama).

- 5 Dado que ya se produce un espectro en el dominio de la frecuencia para cada trama, la materia prima para la ECU de fase se puede obtener fácilmente simplemente almacenando el último espectro decodificado en la memoria. Sin embargo, si los espectros decodificados corresponden a tramas de la señal en el dominio del tiempo con diferentes funciones de ventana (véase la Figura 1), la eficiencia del algoritmo puede verse reducida. Esto puede suceder cuando el decodificador divide las tramas de síntesis en subtramas más cortas, por ejemplo, para manejar sonidos transitorios que requieren una resolución temporal más alta. Para lograr buenos resultados, la ECU debe producir la forma de
- 10 ventana deseada para cada trama, o puede haber artefactos de transición en cada límite de trama. Una solución es almacenar el espectro de cada trama correspondiente a una ventana determinada y aplicar la ECU sobre ellos individualmente. Otra solución podría ser almacenar un solo espectro para la ECU y corregir la ventana en el dominio del tiempo. Esto se puede implementar aplicando una ventana inversa y luego volviendo a aplicar una ventana con la forma deseada. Estas soluciones tienen algunos inconvenientes que se analizan a continuación.
- 15 Una desventaja de aplicar la ECU en el dominio de la frecuencia en subtramas individuales es que puede haber diferencias entre las subtramas que se replicarán para cada subtrama durante la trama perdida. En el caso de pérdidas de tramas consecutivas, esto puede generar un artefacto repetitivo, ya que cada subtrama puede tener una firma espectral ligeramente diferente. Otro problema es que aumenta el requisito de memoria, ya que es necesario almacenar un espectro de cada subtrama.
- 20 La solución de reorganización de ventanas, en la que se invierte y se vuelve a aplicar la ventana, supera el problema de las diferentes firmas espectrales, ya que la ECU puede basarse en una única subtrama. Sin embargo, aplicar la ventana invertida y aplicar una nueva ventana implica una división y una multiplicación para cada muestra, donde la división es una operación computacionalmente compleja y costosa. Esta solución se podría mejorar almacenando una ventana de reorganización precalculada en la memoria, pero esto aumentaría la memoria de tabla requerida. En caso
- 25 de que la ECU se aplique en una subparte del espectro, puede requerir además que se reorganice todo el espectro, ya que todo el espectro debe tener la misma forma de ventana.

Compendio

Según un primer aspecto de la invención, se proporciona un método como se define en la reivindicación 1.

- 30 Una posible ventaja es que se genera una ECU de múltiples subtramas a partir de un espectro de subtrama única aplicando una síntesis invertida en el tiempo. Esta generación puede ser adecuada para casos en los que las ventanas de subtrama son versiones invertidas en el tiempo unas de otras. La generación de todas las tramas de la ECU a partir de una única trama decodificada almacenada garantiza que las subtramas tengan una firma espectral similar, manteniendo al mismo tiempo el consumo de memoria y la complejidad computacional al mínimo.

- 35 Según un segundo aspecto de la invención, se proporciona un dispositivo decodificador como se define en la reivindicación 10.

Se proporcionan realizaciones preferidas adicionales en las reivindicaciones dependientes.

Breve descripción de los dibujos

Los dibujos adjuntos, que se incluyen para proporcionar una mejor comprensión de la descripción y se incorporan y constituyen una parte de esta solicitud, ilustran ciertas realizaciones no limitativas. En los dibujos:

- 40 La Figura 1 es un diagrama de bloques que ilustra un par de codificador y decodificador donde la codificación se realiza en el dominio DFT;
- la Figura 2 es un diagrama de bloques que ilustra un par de codificador y decodificador donde el decodificador aplica una transformación DFT para facilitar el procesamiento en el dominio de la frecuencia;
- 45 la Figura 3 es una ilustración de dos ventanas de subtrama de un decodificador, donde la ventana aplicada en la segunda subtrama es una versión invertida en el tiempo o reflejada de la ventana aplicada en la primera subtrama;
- la Figura 4 es un diagrama de bloques que ilustra un sistema codificador y decodificador que incluye un método PLC que realiza una estimación de fase y aplica síntesis ECU invertida en el tiempo usando una calculadora de fase invertida en tiempo según algunas realizaciones;
- 50 la Figura 5 es un diagrama de flujo que ilustra las operaciones de un dispositivo decodificador que realiza la síntesis de ECU invertida en tiempo según algunas realizaciones;
- la Figura 6 es una ilustración de una ventana invertida en tiempo en una senoide según algunas realizaciones;

la Figura 7 es una ilustración de cómo una ventana invertida en tiempo afecta a los coeficientes DFT en el plano complejo según algunas realizaciones;

la Figura 8 es una ilustración de  frente a la frecuencia f según algunas realizaciones;

la Figura 9 es un diagrama de bloques que ilustra un dispositivo decodificador según algunas realizaciones;

5 la Figura 10 es un diagrama de flujo que ilustra las operaciones de un dispositivo decodificador según algunas realizaciones;

la Figura 11 es un diagrama de flujo que ilustra las operaciones de un dispositivo decodificador según algunas realizaciones;

Descripción detallada

10 Los aspectos de la presente descripción se describirán ahora con más detalle a continuación con referencia a los dibujos adjuntos, en los que se muestran ejemplos de realizaciones. Sin embargo, las realizaciones pueden realizarse en muchas formas diferentes y no deben interpretarse como limitadas a las realizaciones expuestas en este documento. Más bien, estas realizaciones se proporcionan de modo que esta descripción sea minuciosa y completa, quedando definido el alcance de la invención por las reivindicaciones adjuntas.

15 También debe tenerse en cuenta que estas realizaciones no son mutuamente excluyentes. Se puede suponer tácitamente que los componentes de una realización están presentes o se usan en otra realización.

La siguiente descripción presenta diversas realizaciones de la materia objeto de la invención. Estas realizaciones se presentan como ejemplos didácticos y no deben interpretarse como limitantes del alcance de la materia objeto de la invención. Por ejemplo, ciertos detalles de las realizaciones descritas pueden modificarse, omitirse o ampliarse sin apartarse del alcance de la invención, tal como se define en las reivindicaciones.

20 La Figura 9 es un diagrama de bloques que ilustra elementos de un dispositivo 900 decodificador, que puede ser parte de un terminal móvil, un terminal de comunicación móvil, un dispositivo de comunicación inalámbrica, un terminal inalámbrico, un terminal de comunicación inalámbrica, un equipo de usuario, UE, un nodo/terminal/dispositivo de equipo de usuario, etc., configurado para proporcionar comunicación inalámbrica según realizaciones. Como se muestra, el decodificador 900 puede incluir un circuito 906 de interfaz de red (también denominado interfaz de red) configurado para proporcionar comunicaciones con otros dispositivos/entidades/funciones/etc. El decodificador 900 también puede incluir un circuito 902 procesador (también denominado procesador) acoplado operativamente al circuito 906 de interfaz de red, y un circuito 904 de memoria (también denominado memoria) acoplado operativamente al circuito procesador. El circuito 904 de memoria puede incluir un código de programa legible por ordenador que, cuando es ejecutado por el circuito 902 procesador, hace que el circuito procesador realice operaciones según realizaciones descritas en el presente documento.

Según otras realizaciones, el circuito 902 procesador puede definirse para incluir memoria, de modo que no se requiera un circuito de memoria independiente. Como se analiza en el presente documento, las operaciones del decodificador 900 pueden ser realizadas por el procesador 902 y/o la interfaz 906 de red. Por ejemplo, el procesador 902 puede controlar la interfaz 906 de red para transmitir comunicaciones a reproductores de audio multicanal y/o para recibir comunicaciones a través de la interfaz 906 de red desde uno o más nodos/entidades/servidores de red, como nodos codificadores, servidores de depósito, etc. Además, los módulos pueden almacenarse en la memoria 904, y estos módulos pueden proporcionar instrucciones de modo que cuando el procesador 902 ejecuta las instrucciones de un módulo, el procesador 902 realiza las operaciones respectivas.

40 En la descripción que sigue, se usará la notación de subtrama para describir las realizaciones. Aquí, una subtrama denota una parte de una trama más grande donde la trama más grande está compuesta por un conjunto de subtramas. Las realizaciones descritas también pueden usarse con notación de trama. En otras palabras, las subtramas pueden formar grupos de tramas que tienen la misma forma de ventana que la descrita en este documento y las subtramas no necesitan ser parte de una trama más grande.

45 Considérese un decodificador de un par de codificador y decodificador donde el método de decodificación genera espectros de frecuencia sobre una base de subtrama. Las subtramas consecutivas pueden tener la propiedad de que la forma de ventana aplicada es una versión reflejada o invertida en el tiempo una de la otra, como se ilustra en la Figura 3, donde la subtrama 2 es una versión reflejada o invertida en el tiempo de la subtrama 1. El decodificador

obtiene los espectros de las subtramas reconstruidas $\hat{X}_1(m, k)$, $\hat{X}_2(m, k)$ para cada trama m. En una realización, los espectros de subtrama pueden obtenerse a partir de una síntesis de dominio de tiempo reconstruida $\hat{x}(m, n)$, donde n es un índice de muestra. Los recuadros discontinuos en la Figura 2 indican que el procesamiento de dominio de frecuencia puede realizarse antes o después de los módulos de memoria y PLC. Los espectros pueden obtenerse multiplicando $\hat{x}(m, n)$ con las funciones de ventana de subtrama $w_1(n)$ y $w_2(n)$ y aplicando la transformada DFT de acuerdo con:

$$\hat{X}_1(m, k) = \sum_{n=0}^{N-1} \hat{x}(m, n) w_1(n) e^{-\frac{j2\pi kn}{N}}$$

$$\hat{X}_2(m, k) = \sum_{n=0}^{N-1} \hat{x}(m, n + N_{\text{paso12}}) w_2(n) e^{-\frac{j2\pi kn}{N}}$$

donde N denota la longitud de la ventana de subtrama y N_{paso12} es la distancia en muestras entre el punto de inicio de la primera y la segunda subtrama. Las funciones de ventana de subtrama $w_1(n)$ y $w_2(n)$ son versiones reflejadas o invertidas en el tiempo una de la otra. Aquí, los espectros de subtrama se obtienen a partir de una síntesis de dominio de tiempo del decodificador, similar al sistema descrito en la Figura 2. Debe notarse que las realizaciones son igualmente aplicables para un sistema donde el decodificador reconstruye los espectros de subtrama directamente, como se describe en la Figura 1. Para cada trama de audio correctamente recibida y decodificada m, el espectro correspondiente a la segunda subtrama $\hat{X}_2(m, k)$ se almacena en la memoria.

$$\hat{X}_{\text{mem}}(k) := \hat{X}_2(m, k)$$

Para las tramas recibidas correctamente, el dispositivo 900 decodificador puede proceder a realizar los pasos de procesamiento del dominio de frecuencia, realizando la transformación DFT inversa y reconstruyendo el audio de salida usando una estrategia de superposición-adición. Las tramas faltantes o dañadas pueden ser identificadas por la capa de transporte que maneja la conexión y se señalan al decodificador como una "trama defectuosa" a través de un Indicador de Trama Defectuosa (BFI), que puede tener la forma de una bandera. Cuando el dispositivo decodificador 900 detecta una trama defectuosa a través de un indicador de trama defectuosa (BFI), se activa el algoritmo PLC. PLC sigue el principio de la ECU de fase [4]. El espectro almacenado $\hat{X}_{\text{mem}}(k)$ se introduce en un algoritmo de detector de picos que detecta picos en una escala de frecuencia fraccionaria. Un conjunto de picos

$$P = \{f_i\}, i = 1, 2, \dots, N_{\text{picos}}$$

Se pueden detectar las que están representadas por su frecuencia f_i fraccionaria estimada y donde N_{picos} es el número de picos detectados. De manera similar al paradigma de codificación sinusoidal, los picos del espectro se modelan con sinusoides con una cierta amplitud, frecuencia y fase. La frecuencia fraccionaria se puede expresar como un número fraccionario de contenedores DFT, de modo que, por ejemplo, la frecuencia de Nyquist se encuentra en $f = N/2 + 1$. Cada pico se puede asociar con un número de contenedores de frecuencia que representan el pico. Estos se encuentran redondeando la frecuencia fraccionaria al entero más cercano e incluyendo los contenedores vecinos, por ejemplo, los N_{ceranos} picos en cada lado:

$$\begin{aligned} \{k_i\} &= [f_i] \\ \{G_i\} &= \{k_i - N_{\text{ceranos}}, \dots, k_i, \dots, k_i + N_{\text{ceranos}}\} \end{aligned}$$

donde $[\cdot]$ representa la operación de redondeo y G_i es el grupo de contenedores que representan el pico en la frecuencia f_i . El número N_{ceranos} es una constante de ajuste que se puede determinar al diseñar el sistema. Un N_{ceranos} proporciona una mayor precisión en cada representación de pico, pero también introduce una distancia mayor entre picos que se pueden modelar. Un valor adecuado para N_{ceranos} puede ser 1 ó 2. Los picos del espectro de ocultación $\hat{X}_{\text{ECU}}(m, k)$ se pueden formar usando estos grupos de contenedores, donde se ha aplicado un ajuste de fase a cada grupo. El ajuste de fase tiene en cuenta el cambio de fase en la senoide subyacente, suponiendo que la frecuencia permanece igual entre la última trama correctamente recibida y decodificada y la trama de ocultación. El ajuste de fase se basa en la frecuencia fraccionaria y el número de muestras entre la trama de análisis de la trama anterior y donde comenzaría la trama actual. Como se ilustra en la Figura 3, este número de muestras es N_{paso21} entre el inicio de la segunda subtrama de la última trama recibida y el inicio de la primera subtrama de la primera trama de la ECU, y $N_{\text{completas}}$ entre la primera subtrama de la última trama recibida y la primera subtrama de la primera trama ECU. Nótese que $N_{\text{completas}}$ también da la distancia entre la segunda subtrama de la última trama recibida y la segunda subtrama de la primera trama ECU.

La Figura 4 ilustra un sistema codificador y decodificador donde un bloque 109 PLC realiza una estimación de fase utilizando un estimador 112 de fase y aplica síntesis ECU invertida en el tiempo usando un calculador 113 de fase invertida en tiempo de acuerdo con realizaciones descritas a continuación.

La Figura 5 es un diagrama de flujo que ilustra los pasos de la síntesis ECU en tiempo inverso que se describe a continuación. Para la ocultación de la primera subtrama, la síntesis ECU se realiza en tiempo inverso para obtener la forma de ventana deseada. El ajuste de fase, o corrección de fase o progresión de fase (estos términos se usan indistintamente a lo largo de la descripción), para la primera subtrama para el pico i se puede escribir como

$$\Delta\phi_i = -2\phi_i - 2\pi f_i (N + N_{paso21} + (N_{perdidas} - 1)N_{completas})/N,$$

donde $N_{perdidas}$ denota el número de tramas perdidas consecutivas y ϕ_i denota la fase de la sinusoide en la frecuencia f_i . El término $(N_{perdidas} - 1)N_{completas}$ maneja la progresión de fase para errores de ráfaga, donde el paso se incrementa con la longitud de trama de la trama $N_{completas}$ completa. Para la primera trama perdida, $N_{perdidas} = 1$. Para frecuencias que están centradas en los contenedores de frecuencia del espectro, $\hat{X}_{mem}(k)$ la fase ϕ_i está disponible fácilmente con solo extraer el ángulo:

$$\phi_i = \angle \hat{X}_{mem}(k_i)$$

donde $k_i = [f_i]$.

En general, la frecuencia f_i es un número fraccionario y la fase debe estimarse en la operación 501. Un método de estimación es usar la interpolación lineal del espectro de fase.

$$\phi_i = (f_i - [f_i]) (\angle \hat{X}_{mem}([f_i]) - \angle \hat{X}_{mem}([f_i])) + \angle \hat{X}_{mem}([f_i])$$

donde $\lfloor \cdot \rfloor$ y $\lceil \cdot \rceil$ representan los operadores para redondear hacia abajo y hacia arriba respectivamente. Sin embargo, se encontró que este método de estimación era inestable. Este método de estimación requiere además dos extracciones de fase, lo que requiere la función arctan computacionalmente compleja en caso de que el espectro se represente con números complejos en la forma estándar $a + bi$. Otra estimación de fase que se encontró confiable con una complejidad computacional relativamente baja es

$$\phi_i = \angle \hat{X}_{mem}(k_i) - f_{frac}(\phi_c + \pi)$$

$$f_{frac} = f_i - k_i$$

donde f_{frac} es el error de redondeo y ϕ_c es una constante de ajuste que depende de la forma de la ventana que se aplica. Para la forma de la ventana de esta realización, se encontró que un valor adecuado era $\phi_c = 0,33$. Para otra forma de ventana, se encontró que era $\phi_c = 0,48$. En general, se espera que se pueda encontrar un valor adecuado en el rango $[0, 1, 0, 7]$.

En la operación 502 se deriva un ajuste $\Delta\Phi$ de fase invertida en el tiempo como se explicó anteriormente.

Los picos del espectro de ocultación se pueden formar aplicando el ajuste de fase al espectro almacenado en la operación 503.

$$\hat{X}_{ECU}(m, k) = (\hat{X}_{mem}(k) e^{j\Delta\phi_i})^*, k \in G_i$$

El asterisco * denota el conjugado complejo, que da una inversión temporal de la señal en la operación 504. Esto da como resultado una inversión temporal de la primera subtrama de la ECU. Debe tenerse en cuenta que también puede ser posible realizar la inversión en el dominio temporal después de la DFT inversa. Sin embargo, si $\hat{X}_{ECU}(m, k)$ solo representa una parte del espectro completo, esto requiere que el espectro restante se trate previamente, por ejemplo, mediante una inversión temporal antes del análisis de la DFT.

Los contenedores restantes de $\hat{X}_{ECU}(m, k)$, que no están ocupados por los contenedores G_i de pico, pueden denominarse espectro de ruido o componente de ruido del espectro. Pueden completarse usando los coeficientes del espectro almacenado con una fase aleatoria aplicada:

$$\hat{X}_{ECU}(m, k) = \hat{X}_{mem}(k) e^{j\phi_{aleat}}, \quad k \notin G_i$$

Donde Φ_{aleat} denota un valor de fase aleatorio. Los contenedores restantes también pueden llenarse con coeficientes espectrales que conservan una propiedad deseada de la señal, por ejemplo, la correlación con un segundo canal en un sistema decodificador multicanal. En la operación 505, el espectro $\hat{X}_{ECU}(m, k)$ de pico, donde $k \in G_i$, se combina con el espectro $\hat{X}_{ECU}(m, k)$ de ruido donde $k \notin G_i$ para formar un espectro combinado.

- 5 En las realizaciones en las que el ruido se genera en el dominio del tiempo y se transforma en ventanas, se debe realizar una inversión temporal del ruido para que coincida con la ventana de los componentes del pico y la combinación con el espectro del pico antes de aplicar la inversión temporal descrita anteriormente.

Para la generación de la segunda subtrama, que se sintetiza en tiempo normal (no invertido), se puede usar el ajuste de fase regular.

$$10 \quad \Delta\phi_i = 2\pi f_i N_{completas} N_{perdidas} / N$$

La síntesis de la ECU para la segunda subtrama se puede formar de manera similar al primer subtrama, pero omitiendo el conjugado complejo en los coeficientes de pico.

$$\hat{X}_{ECU}(m, k) = \hat{X}_{mem}(k) e^{j\Delta\phi_i}, \quad k \in G_i$$

$$\hat{X}_{ECU}(m, k) = \hat{X}_{mem}(k) e^{j\Phi_{aleat}}, \quad k \notin G_i$$

- 15 Una vez que se genera el espectro de ocultación combinado en la operación 505, el espectro de ocultación combinado se puede alimentar a los siguientes pasos de procesamiento en la operación 506, incluyendo la DFT inversa y una operación de superposición-adición que da como resultado una señal de audio de salida.

La señal de audio de salida puede transmitirse a uno o más altavoces, como altavoces, para su reproducción. Los altavoces pueden ser parte del dispositivo de decodificación, un dispositivo independiente o parte de otro dispositivo.

Derivación de la fórmula de corrección de fase para la síntesis ECU invertida en el tiempo

- 20 Suponga que la fase inicial del componente senoide es 00 y que la frecuencia de la senoide es f. La fase ϕ_1 deseada de la senoide después de avanzar N_{paso} muestras es entonces

$$\phi_1 = \phi_0 + 2\pi f N_{paso} / N$$

- 25 Para una continuación invertida en el tiempo de la senoide, la fase debe reflejarse en el eje real aplicando el conjugado complejo o simplemente tomando la fase negativa $-\phi_1$. Dado que este ángulo de fase ahora representa el punto final del marco de síntesis de la ECU, la fase debe retroceder a lo largo de la trama de análisis para llegar a la fase ϕ_2 de inicio deseada.

$$\phi_2 = -\phi_1 - 2\pi f(N - 1)/N$$

Para obtener una corrección $\Delta\phi$ de fase, es necesario restar la fase inicial, es decir,

$$30 \quad \phi_0 + \Delta\phi = \phi_2 \Rightarrow \Delta\phi = \phi_2 - \phi_0$$

Sustituyendo ϕ_2 se obtiene

$$\Delta\phi = -2\phi_0 - 2\pi f(N_{paso} + N - 1)/N$$

- 35 Para agregar progresión para pérdidas de tramas consecutivas (pérdida de ráfaga), se puede agregar un factor correspondiente a la cantidad de muestras entre los puntos de inicio de los tramas completas, $N_{desplazamiento} = (N_{perdidas} - 1)N_{completas}$. Esto proporciona la corrección de fase final.

$$\Delta\phi = -2\phi_0 - 2\pi f(N + N_{paso} - 1 + (N_{perdidas} - 1)N_{completas})/N,$$

La inversión temporal deseada se puede lograr en el dominio DFT usando un conjugado complejo junto con un desplazamiento circular de una muestra. Este desplazamiento circular se puede implementar con una corrección de fase de $2\pi k/N$ que se puede incluir en la corrección de fase final.

$$\Delta\phi = -2\phi_0 - 2\pi f(N + N_{\text{paso}} - 1 + (N_{\text{perdidas}} - 1)N_{\text{completas}})/N + 2\pi k/N$$

- 5 Para los coeficientes que representan un único pico, el intervalo de frecuencia k del desplazamiento circular se puede aproximar con la frecuencia fraccionaria $k \approx f$ y la corrección de fase se puede simplificar a

$$\begin{aligned}\Delta\phi &= -2\phi_0 - 2\pi f(N + N_{\text{paso}} - 1 + (N_{\text{perdidas}} - 1)N_{\text{completas}})/N + 2\pi f/N = \\ &= -2\phi_0 - 2\pi f(N + N_{\text{paso}} + (N_{\text{perdidas}} - 1)N_{\text{completas}})/N\end{aligned}$$

Las ventanas pueden diseñarse de tal manera que $N = N_{\text{completas}}$, en cuyo caso, de acuerdo con la invención, la expresión puede simplificarse aún más a

$$\Delta\phi = -2\phi_0 - 2\pi f(N_{\text{paso}} + N_{\text{perdidas}} \cdot N)/N$$

10

Una realización alternativa de la síntesis ECU invertida en el tiempo

En otra realización, la corrección de fase se realiza en dos pasos. La fase se avanza en el primer paso, ignorando el desajuste de la ventana.

$$\begin{aligned}\Delta\phi &= 2\pi f(N_{\text{paso}} + (N_{\text{perdidas}} - 1)N_{\text{completas}}) \\ \hat{x}_{ECU,1}(m, k) &= \hat{x}_{mem}(k)e^{j\Delta\phi_i}, \quad k \in G_i\end{aligned}$$

- 15 En un segundo paso, la inversión temporal del ventanaje se puede lograr invirtiendo la fase con $-\Phi_m$, aplicando el conjugado complejo y restaurando la fase con Φ_m :

$$\hat{x}_{ECU}(m, k) = (\hat{x}_{ECU,1}(m, k)e^{-j\Delta\phi_m})^* e^{j\Delta\phi_m}, \quad k \in G_i$$

La motivación para esta operación se puede encontrar estudiando el efecto de una ventana invertida en el tiempo en una senoide como se ilustra en la Figura 6. En la Figura 6, el gráfico superior muestra la ventana aplicada en una primera dirección, y el gráfico inferior muestra la ventana aplicada en la dirección inversa. Los tres coeficientes que representan la senoide se ilustran en la Figura 7, que ilustra cómo una ventana invertida en el tiempo afecta a los coeficientes DFT en el plano complejo. Los tres coeficientes DFT que aproximan la senoide en el gráfico superior de la Figura 6 están marcados con círculos, mientras que los coeficientes correspondientes del gráfico inferior de la Figura 6 están marcados con estrellas. El diamante denota la posición de la fase original de la senoide y la línea discontinua muestra un plano de reflejo observado a través del cual se proyectan los coeficientes de la ventana invertida en el tiempo. La ventana invertida en el tiempo da un reflejo de los coeficientes en un plano de reflejo con un ángulo Φ_m .

20

25

$$\Phi_m = \phi_0 + \phi_{\text{frac}}$$

A través de la experimentación, se encontró que Φ_{frac} podría expresarse como

$$\phi_{\text{frac}} = \pi f_{\text{frac}}$$

$$f_{\text{frac}} = f_i - k_i$$

$$k_i = \{f_i\}$$

30

donde $[\cdot]$ denota la operación de redondeo. También se encontró que Φ_ε , expresado como un ángulo positivo, se puede aproximar mediante una relación lineal con f_{frac} . En la Figura 8, el ángulo Φ_ε se expresa como una función de la frecuencia f . Al estudiar la forma de diente de sierra de la Figura 8, se encontró que una buena aproximación de Φ_ε era

$$\Phi_\varepsilon = -f_{frac}\Phi_C$$

donde Φ_C es una constante. En una realización, Φ_C se puede establecer en $\Phi_C = 0,33$, lo que produce una aproximación cercana. Dado que Φ_0 no se conoce explícitamente, una aproximación alternativa de Φ_m se puede escribir como

$$\hat{\Phi}_m = \Phi_{k_i} + \Phi_\varepsilon$$

donde Φ_{k_i} es la fase del coeficiente de pico máximo encontrado en el intervalo de frecuencia redondeado k_i después del primer paso de ajuste de fase,

$$\Phi_{k_i} = \angle \bar{X}_{ECU,1}(m, k_i)$$

La operación de alinear el plano de espejo con el eje real, aplicar el conjugado complejo y volver a invertir la fase puede entenderse como ajustar la fase de la sinusoide moldeada a una posición de fase que es neutra con respecto al conjugado complejo (0 o π), con lo que solo se invierte la forma temporal de la señal. El enfoque de dos pasos es más complejo computacionalmente que la realización descrita anteriormente. Sin embargo, las observaciones también pueden conducir a una aproximación de 00. Se puede ver en la Figura 7 que 00 puede expresarse como

$$\Phi_0 = \Phi_{k_i} + \Phi_\varepsilon - \Phi_{frac} = \Phi_{k_i} - f_{frac}(\Phi_C + \pi)$$

que es la aproximación de fase usada anteriormente.

A continuación se analizarán las operaciones del dispositivo 900 decodificador (implementado usando la estructura del diagrama de bloques de la Figura 9) con referencia al diagrama de flujo de la Figura 10 según algunas realizaciones de la invención.

Por ejemplo, los módulos pueden almacenarse en la memoria 904 de la Figura 9, y estos módulos pueden proporcionar instrucciones para que cuando las instrucciones de un módulo sean ejecutadas por el circuito 902 de procesamiento del dispositivo decodificador respectivo, el circuito 902 de procesamiento realice las operaciones respectivas del diagrama de flujo.

En la operación 1000, el circuito 902 de procesamiento genera espectros de frecuencia sobre la base de subtramas, donde las subtramas consecutivas de la señal de audio tienen una propiedad de que una forma de ventana aplicada de la primera subtrama de las subtramas consecutivas es una versión reflejada o una versión invertida en el tiempo de una segunda subtrama de las subtramas consecutivas. Por ejemplo, generar los espectros de frecuencia para cada subtrama de las dos primeras subtramas consecutivas comprende determinar:

$$\hat{X}_1(m, k) = \sum_{n=0}^{N-1} \hat{x}(m, n) w_1(n) e^{-\frac{j2\pi kn}{N}}$$

$$\hat{X}_2(m, k) = \sum_{n=0}^{N-1} \hat{x}(m, n + N_{paso12}) w_2(n) e^{-\frac{j2\pi kn}{N}}$$

donde N denota una longitud de una ventana de subtrama, la función $w_1(n)$ de ventana de subtrama es una función de ventana de subtrama para la primera subtrama $\hat{X}_1(m, k)$ de las subtramas consecutivas y $w_2(n)$ es una función de ventana de subtrama para la segunda subtrama $\hat{X}_2(m, k)$ de las subtramas consecutivas, y N_{paso12} es un número de muestras entre una primera subtrama de las dos primeras subtramas consecutivas y la segunda subtrama de las dos primeras subtramas consecutivas.

En la operación 1002, el circuito de procesamiento 902 determina si se ha recibido un indicador de trama defectuosa (BFI). El indicador de trama defectuosa proporciona una indicación de que se ha perdido o dañado una trama de audio.

En la operación 1004, el circuito 902 de procesamiento almacena, para cada trama de audio decodificado correctamente, el espectro correspondiente a la segunda subtrama en la memoria. Por ejemplo, para una trama decodificada correctamente m , el espectro correspondiente a la segunda subtrama $\hat{X}_2(m, k)$ se almacena en la memoria como $\hat{X}_{mem}(k) \equiv \hat{X}_2(m, k)$. Para tramas recibidas correctamente, el dispositivo 900 decodificador puede proceder a realizar los pasos de procesamiento en el dominio de la frecuencia, realizar la transformación DFT inversa y reconstruir el audio de salida usando una estrategia de superposición-adición como se describió anteriormente y se ilustra en la Figura 4. Obsérvese que el principio de superposición-adición es el mismo tanto para las subtramas como para las tramas. La creación de una trama requiere aplicar la superposición-adición en las subtramas, mientras que la trama de salida final es el resultado de una operación de superposición-adición entre tramas.

Cuando el circuito 902 de procesamiento detecta una trama defectuosa a través de un indicador de trama defectuosa (BFI) en la operación 1002, se realizan las operaciones 1006 a 1030 PLC.

En la operación 1006, el circuito 902 de procesamiento obtiene el espectro de señal correspondiente a la segunda subtrama de unas dos primeras subtramas consecutivas decodificadas y procesadas correctamente con anterioridad. Por ejemplo, el circuito 902 de procesamiento puede obtener el espectro de señal de la memoria 904 del dispositivo de decodificación.

En la operación 1008, el circuito 902 de procesamiento detecta picos del espectro de señal de una trama de audio recibida previamente de la señal de audio en una escala de frecuencia fraccionaria, la trama de audio recibida previamente se recibió antes de recibir el indicador de trama defectuosa.

En la operación 1010, el circuito 902 de procesamiento determina si la trama de ocultación es para la primera subtrama de dos subtramas consecutivas.

Si la trama de ocultación corresponde a la primera subtrama, en la operación 1012, el circuito 902 de procesamiento calcula la fase de cada uno de los picos. En una realización, se calcula una estimación de fase para los picos de los picos corregidos de fase invertida en el tiempo de acuerdo con:

$$\phi_i = \angle \hat{X}_{mem}(k_i) + f_{frac}(\phi_c - \pi)$$

$$f_{frac} = f_i - k_i$$

donde ϕ_i es una fase estimada en la frecuencia f_i , $\angle \hat{X}_{mem}(k_i)$ es un ángulo del espectro $\hat{X}_{mem}$, en un intervalo de frecuencia k_i , f_{frac} es un error de redondeo, ϕ_c es una constante de ajuste y k_i es $f[i]$. La constante ϕ_c de ajuste puede ser un valor en un rango entre 0,1 y 0,7.

En la operación 1014, el circuito 902 de procesamiento deriva una corrección de fase invertida en el tiempo para aplicar a los picos del espectro de señal con base en la fase estimada.

En la operación 1016, el circuito 902 de procesamiento aplica la corrección de fase invertida en el tiempo a los picos del espectro de señal para formar picos corregidos de fase invertida en el tiempo.

En la operación 1018, el circuito 902 de procesamiento aplica una inversión temporal a la subtrama de audio de ocultación. En una realización, la inversión temporal se puede aplicar aplicando un conjugado complejo a la subtrama de audio de ocultación.

En la operación 1020, el circuito 902 de procesamiento combina los picos corregidos de fase invertida en el tiempo con un espectro de ruido del espectro de señal para formar un espectro combinado de la subtrama de audio de ocultación.

Volviendo a la Figura 11, en una realización, 1016 y 1018 pueden ser realizados por el circuito 902 de procesamiento asociando cada pico con un número de contenedores de frecuencia pico en la operación 1100. El circuito 902 de procesamiento que asocia puede aplicar la corrección de fase invertida en el tiempo aplicando la corrección de fase invertida en el tiempo a cada uno del número de contenedores de frecuencia en la operación 1102. En la operación 1104, los contenedores restantes se llenan usando coeficientes del espectro de señal con una fase aleatoria aplicada.

Volviendo a la Figura 10, en la operación 1022, el circuito 902 de procesamiento genera una subtrama de audio de ocultación sintetizada basada en el espectro combinado

Si la trama de ocultación no es para la primera subtrama según lo determinado en la operación 1010, el circuito 902 de procesamiento deriva en la operación 1024 una corrección de fase no invertida en el tiempo para aplicar a los picos del espectro de señal para una segunda subtrama de ocultación de las al menos dos subtramas de ocultación consecutivas.

En la operación 1026, el circuito 902 de procesamiento aplica la corrección de fase invertida no temporal a los picos del espectro de señal para la segunda subtrama para formar picos corregidos de fase no invertida en el tiempo.

En la operación 1028, el circuito 902 de procesamiento combina los picos no corregidos en fase invertida con un espectro de ruido del espectro de señal para formar un espectro combinado para la segunda subtrama de ocultación.

- 5 En la operación 1030, el circuito 902 de procesamiento genera una segunda subtrama de audio de ocultación sintetizada con base en el espectro combinado.

Volviendo a la Figura 11, en una realización, 1026 y 1028 pueden ser realizados por el circuito 902 de procesamiento asociando cada pico con un número de contenedores de frecuencia pico en la operación 1100. El circuito 902 de procesamiento que asocia puede aplicar la corrección de fase no invertida en el tiempo aplicando la corrección de fase no invertida en el tiempo a cada uno del número de contenedores de frecuencia en la operación 1102. En la operación 1104, los contenedores restantes se llenan usando coeficientes del espectro de señal con una fase aleatoria aplicada.

10 Varias operaciones del diagrama de flujo de la Figura 10 pueden ser opcionales con respecto a algunas realizaciones de dispositivos decodificadores y métodos relacionados. Con respecto a los métodos de la realización de ejemplo 1 (que se detalla a continuación), por ejemplo, las operaciones de los bloques 1004 y 1022-1030 de la Figura 10 pueden ser opcionales. Con respecto a los métodos de la realización de ejemplo 19 (que se detalla a continuación), por ejemplo, las operaciones de los bloques 1010 y 1022-1030 de la Figura 10 pueden ser opcionales.

15 A continuación se proporcionan explicaciones de varias abreviaturas/acróminos usados en la presente descripción.

Abreviatura Explicación

DFT Transformada de Fourier Discreta

20 IDFT Transformada de Fourier Discreta Inversa

LP Predicción Lineal

PLC Ocultación de Pérdida de Paquetes

ECU Unidad de Ocultación de Errores

FEC Corrección/Ocultación de Errores de Trama

- 25 Las referencias se identifican a continuación.

[1] T. Vaillancourt, M. Jelinek, R. Salami y R. Lefebvre, "Efficient Frame Erasure Concealment in Predictive Speech Codecs using Glottal Pulse Resynchronisation", Conferencia internacional IEEE de 2007 sobre acústica, voz y procesamiento de señales - ICASSP '07, Honolulu, EH, 2007, págs. IV-1113-IV-1116.

30 [2] J. Lecomte et al., "Packet-Loss concealment technology advances in EVS", Conferencia internacional IEEE de 2015 sobre acústica, habla y procesamiento de señales (ICASSP), Brisbane, QLD, 2015, págs. 5708-5712.

[3] TS 3GPP 26.447, Códec para servicios de voz mejorados (EVS); ocultación de errores de paquetes perdidos (versión 12)

35 [4] S. Bruhn, E. Norvell, J. Svedberg y S. Sverrisson, "A novel sinusoidal approach to audio signal frame loss concealment and its application in the new evs codec standard", Conferencia internacional IEEE 2015 sobre acústica, habla y procesamiento de señales (ICASSP), Brisbane, QLD, 2015, págs. 5142-5146.

40 En general, todos los términos usados en el presente documento deben interpretarse según su significado ordinario en el campo técnico pertinente, a menos que se indique claramente un significado diferente y/o se desprenda del contexto en el que se usan. Todas las referencias a un/uno/el elemento, aparato, componente, medio, paso, etc. deben interpretarse abiertamente como una referencia a al menos una instancia del elemento, aparato, componente, medio, paso, etc., a menos que se indique explícitamente lo contrario. Los pasos de cualquier método descrito en el presente documento no tienen que realizarse en el orden exacto descrito, a menos que un paso se describa explícitamente como siguiente o anterior a otro paso y/o cuando esté implícito que un paso debe seguir o preceder a otro paso. Cualquier característica de cualquiera de las realizaciones descritas en el presente documento puede aplicarse a cualquier otra realización, siempre que sea apropiado. Asimismo, cualquier ventaja de cualquiera de las realizaciones puede aplicarse a cualquier otra realización, y viceversa. Otros objetivos, características y ventajas de las realizaciones adjuntas serán evidentes a partir de la siguiente descripción.

45 En la descripción anterior de diversas realizaciones, se debe entender que la terminología usada en este documento tiene el propósito de describir realizaciones particulares únicamente y no pretende ser limitante. A menos que se defina de otra manera, todos los términos (incluidos los términos técnicos y científicos) usados en este documento tienen el mismo significado que entiende comúnmente una persona con conocimientos ordinarios en la técnica a la que

pertenece la presente descripción. Se entenderá además que los términos, como los definidos en diccionarios de uso común, deben interpretarse como que tienen un significado que es coherente con su significado en el contexto de esta especificación y la técnica relevante y no se interpretarán en un sentido idealizado o excesivamente formal a menos que se definan expresamente de esa manera en este documento.

5 Cuando se hace referencia a un elemento como "conectado", "acoplado", "sensible" o variantes de los mismos a otro elemento, puede estar conectado, acoplado o ser sensible directamente al otro elemento o pueden estar presentes elementos intermedios. Por el contrario, cuando se hace referencia a un elemento como "conectado directamente", "acoplado directamente", "sensible directamente" o variantes de los mismos a otro elemento, no hay elementos intermedios presentes. Los números iguales se refieren a elementos iguales en todo momento. Además, "acoplado",
10 "conectado", "sensible" o variantes de los mismos tal como se usan en el presente documento pueden incluir acoplado de forma inalámbrica, conectado o sensible. Tal como se usan en el presente documento, las formas singulares "un", "una" y "el" tienen la intención de incluir también las formas plurales, a menos que el contexto indique claramente lo contrario. Es posible que no se describan en detalle funciones o construcciones bien conocidas por razones de brevedad y/o claridad. El término "y/o" incluye todas y cada una de las combinaciones de uno o más de los elementos
15 enumerados asociados.

Se entenderá que, aunque los términos primero, segundo, tercero, etc. pueden usarse en el presente documento para describir diversos elementos/operaciones, estos elementos/operaciones no deben limitarse a estos términos. Estos términos solo se usan para distinguir un elemento/operación de otro elemento/operación. Por lo tanto, un primer elemento/operación en algunas realizaciones podría denominarse un segundo elemento/operación en otras
20 realizaciones sin apartarse de las enseñanzas de la presente descripción. Los mismos números de referencia o los mismos designadores de referencia denotan los mismos elementos o elementos similares en toda la especificación.

Tal como se usan en el presente documento, los términos "comprender", "que comprende", "comprende", "incluir", "que incluye", "incluye", "tienen", "tiene", "que tiene", o variantes de los mismos son abiertos e incluyen una o más características, números enteros, elementos, pasos, componentes o funciones establecidos, pero no excluyen la
25 presencia o adición de una o más características, números enteros, elementos, pasos, componentes, funciones o grupos de los mismos. Además, tal como se usa en el presente documento, la abreviatura común "eg", que deriva de la frase latina "exempli gratia", puede utilizarse para introducir o especificar un ejemplo general o ejemplos de un elemento mencionado anteriormente, y no pretende limitar dicho elemento. La abreviatura común "ie", que deriva de la frase latina "id est", puede usarse para especificar un elemento particular a partir de una recitación más general.

30 En el presente documento se describen ejemplos de realización con referencia a diagramas de bloques y/o ilustraciones de diagramas de flujo de métodos, aparatos (sistemas y/o dispositivos) y/o productos de programas informáticos implementados por ordenador. Se entiende que un bloque de los diagramas de bloques y/o ilustraciones de diagramas de flujo, y combinaciones de bloques en los diagramas de bloques y/o ilustraciones de diagramas de flujo, se pueden implementar mediante instrucciones de programa informático que se llevan a cabo mediante uno o
35 más circuitos informáticos. Estas instrucciones de programa informático se pueden proporcionar a un circuito procesador de un circuito informático de propósito general, un circuito informático de propósito especial y/u otro circuito de procesamiento de datos programable para producir una máquina, de modo que las instrucciones, que se ejecutan a través del procesador del ordenador y/u otro aparato de procesamiento de datos programable, transformen y controlen transistores, valores almacenados en ubicaciones de memoria y otros componentes de hardware dentro de
40 dicho circuito para implementar las funciones/actos especificados en los diagramas de bloques y/o el bloque o bloques del diagrama de flujo, y de ese modo crear medios (funcionalidad) y/o estructura para implementar las funciones/actos especificados en los diagramas de bloques y/o el bloque o bloques del diagrama de flujo.

Estas instrucciones de programas informáticos también pueden almacenarse en un medio tangible legible por ordenador que pueda indicar a un ordenador u otro aparato de procesamiento de datos programable que funcione de
45 una manera particular, de modo que las instrucciones almacenadas en el medio legible por ordenador produzcan un artículo de fabricación que incluya instrucciones que implementen las funciones/actos especificados en los diagramas de bloques y/o en el bloque o bloques del diagrama de flujo. En consecuencia, las realizaciones de la presente descripción pueden incorporarse en hardware y/o en software (incluido firmware, software residente, microcódigo, etc.) que se ejecuta en un procesador, como un procesador de señales digitales, que pueden denominarse colectivamente
50 "circuito", "módulo" o variantes de los mismos.

También debe tenerse en cuenta que en algunas implementaciones alternativas, las funciones/actos indicados en los bloques pueden ocurrir fuera del orden indicado en los diagramas de flujo. Por ejemplo, dos bloques mostrados en sucesión pueden de hecho ejecutarse sustancialmente de manera concurrente o los bloques pueden a veces ejecutarse en el orden inverso, dependiendo de la funcionalidad/actos involucrados. Además, la funcionalidad de un
55 bloque dado de los diagramas de flujo y/o diagramas de bloques puede estar separada en múltiples bloques y/o la funcionalidad de dos o más bloques de los diagramas de flujo y/o diagramas de bloques puede estar al menos parcialmente integrada. Finalmente, otros bloques pueden agregarse/insertarse entre los bloques que se ilustran, y/o bloques/operaciones pueden omitirse sin alejarse del alcance de las realizaciones. Además, aunque algunos de los diagramas incluyen flechas en las rutas de comunicación para mostrar una dirección principal de comunicación, debe entenderse que la comunicación puede ocurrir en la dirección opuesta a las flechas representadas.
60

REIVINDICACIONES

1. Un método para generar una subtrama de audio de ocultación de una señal de audio en un dispositivo de decodificación, comprendiendo el método:

5 generar (1000) espectros de frecuencia sobre una base de subtrama donde subtramas consecutivas de la señal de audio tienen una propiedad de que una forma de ventana aplicada de una primera subtrama de las subtramas consecutivas es una versión reflejada o una versión invertida en el tiempo de una ventana aplicada en una segunda subtrama de las subtramas consecutivas;

detectar (1008) picos de un espectro de señal de una señal de audio recibida previamente en una escala de frecuencia fraccionaria;

10 estimar (1012, 501) una fase de cada uno de los picos;

derivar (1014, 502) un ajuste de fase para una subtrama de audio de ocultación invertida en el tiempo basado en la fase estimada, en donde el ajuste de fase $\Delta\Phi$, para los picos de la subtrama de audio de ocultación invertida en el tiempo se calcula de acuerdo con:

$$\Delta\phi = -2\phi_0 - 2\pi f(N_{\text{paso}} + N_{\text{perdidas}} \cdot N)/N,$$

15 donde Φ_0 es la fase estimada de un pico y f es una frecuencia de un pico, N_{perdidas} denota el número de tramas perdidos consecutivos, N denota la longitud de una trama completa y N_{paso} es la distancia en muestras entre el inicio de la segunda subtrama del último trama recibido y el inicio del primer subtrama de un trama de audio de ocultación;

20 aplicar (1016, 503) el ajuste de fase $\Delta\Phi$, a los picos del espectro de señal para formar picos ajustados en fase de un espectro de ocultación; y

aplicar (1018, 504) una inversión temporal mediante la aplicación de un conjugado complejo a los picos ajustados en fase del espectro de ocultación.

2. El método de la reivindicación 1 que comprende además:

25 combinar (1020) los picos ajustados de fase con un espectro de ruido del espectro de señal para formar un espectro combinado para la subtrama de audio de ocultación; y

generar (1022) una subtrama de audio de ocultación sintetizada basado en el espectro combinado.

3. El método de la reivindicación 1 o 2, en donde una trama de audio de ocultación sintetizado comprende al menos dos subtramas de ocultación consecutivas y en donde la derivación del ajuste de fase, la aplicación del ajuste de fase, la aplicación de la inversión temporal y la combinación de los picos ajustados de fase se realizan para una primera subtrama de ocultación de las al menos dos subtramas de ocultación consecutivas, comprendiendo además el método:

30 derivar (1024) un ajuste de fase para aplicar a los picos del espectro de señal para una segunda subtrama de ocultación no invertida en el tiempo de las al menos dos subtramas de ocultación consecutivas;

aplicar (1026) el ajuste de fase a los picos del espectro de señal para la segunda subtrama no invertida en el tiempo para formar picos ajustados en fase no invertida en el tiempo;

35 combinar (1028) los picos ajustados de fase invertida no temporal con un espectro de ruido del espectro de señal para formar un espectro combinado para la segunda subtrama de ocultación; y

generar (1030) una segunda subtrama de audio de ocultación sintetizada basada en el espectro combinado.

4. El método de una cualquiera de las reivindicaciones 1 a 3 que comprende además obtener (1006) el espectro de señal de la señal de audio recibida previamente de una memoria del dispositivo de decodificación.

40 5. El método de una cualquiera de las reivindicaciones 1 a 4 que comprende además asociar (1100) cada pico de los picos detectados con un número de contenedores de frecuencia de pico que representan el pico.

6. El método de la reivindicación 5, en donde para cada contenedor de frecuencia pico del número de contenedores de frecuencia pico, se aplica (1102) al contenedor de frecuencia pico uno de los ajustes de fase para una subtrama de ocultación invertida en el tiempo y el ajuste de fase para una subtrama de ocultación no invertida en el tiempo.

45

7. El método de la reivindicación 6 comprende además :

rellenar (1104) los contenedores restantes del espectro de la señal usando coeficientes del espectro de la señal almacenada, conservando los coeficientes espectrales una propiedad deseada de la señal.

8. El método de la reivindicación 7, en donde la propiedad deseada comprende la correlación con un segundo canal en un sistema decodificador multicanal.

9. El método de una cualquiera de las reivindicaciones 1 a 8, en donde la estimación de la fase de cada uno de los picos comprende:

calcular una estimación de fase para los picos de los picos ajustados por fase de acuerdo con:

$$\phi_i = \angle \hat{X}_{mem}(k_i) - f_{frac}(\phi_c + \pi)$$

$$f_{frac} = f_i - k_i$$

10. Donde Φ_i es una fase estimada en la frecuencia f_i , $\angle \hat{X}_{mem}(k_i)$ es un ángulo del espectro $\hat{X}_{mem}$ de la señal de audio recibida previamente en una frecuencia k_i , f_{frac} es un error de redondeo, Φ_c es una constante de sintonización y k_i es $[f_i]$, donde $[\cdot]$ representa la operación de redondeo.

10. Un dispositivo (900) decodificador configurado para generar una subtrama de audio de ocultación de una señal de audio, en donde el dispositivo decodificador está adaptado para:

generar espectros de frecuencia sobre una base de subtrama donde subtramas consecutivas de la señal de audio tienen una propiedad de que una forma de ventana aplicada de una primera subtrama de las subtramas consecutivas es una versión reflejada o una versión invertida en el tiempo de una ventana aplicada en una segunda subtrama de las subtramas consecutivas;

detectar picos de un espectro de señal de una señal de audio recibida previamente en una escala de frecuencia fraccionaria;

estimar una fase de cada uno de los picos;

derivar un ajuste de fase para una subtrama de audio de ocultación invertida en el tiempo con base en la fase estimada, en donde el ajuste de fase $\Delta\Phi$, para los picos de la subtrama de audio de ocultación invertida en el tiempo se calcula de acuerdo con:

$$\Delta\Phi = -2\Phi_0 - 2\pi f(N_{paso} + N_{perdidas} \cdot N)/N,$$

donde Φ_0 es la fase estimada de un pico y f es una frecuencia de un pico, $N_{perdidas}$ denota el número de tramas perdidos consecutivos, N denota la longitud de una trama completa y N_{paso} es la distancia en muestras entre el inicio de la segunda subtrama de la última trama recibida y el inicio de la primera subtrama de una trama de audio de ocultación;

aplicar el ajuste $\Delta\Phi$ de fase a los picos del espectro de la señal para formar picos ajustados en fase de un espectro de ocultación; y

aplicar una inversión temporal aplicando un conjugado complejo a los picos ajustados de fase del espectro de ocultación.

11. El dispositivo decodificador de la reivindicación 10 está adaptado además para:

combinar los picos ajustados en el tiempo con un espectro de ruido del espectro de la señal para formar un espectro combinado para la subtrama de audio de ocultación; y

generar una subtrama de audio de ocultación sintetizada con base en el espectro combinado.

12. El dispositivo decodificador según la Reivindicación 10 u 11, en donde una trama de audio de ocultación sintetizada comprende al menos dos subtramas de ocultación consecutivas y en donde la derivación del ajuste de fase, la aplicación del ajuste de fase, la aplicación de la inversión temporal y la combinación de los picos ajustados de fase se realizan para un primer subtrama de ocultación de las al menos dos subtramas de ocultación consecutivas, estando adaptado además el dispositivo decodificador para :

derivar un ajuste de fase para aplicar a los picos del espectro de señal para una segunda subtrama de ocultación no invertida en el tiempo de las al menos dos subtramas de ocultación consecutivas;

aplicar el ajuste de fase a los picos del espectro de señal para la segunda subtrama no invertido en el tiempo para formar picos ajustados en fase no invertidos en el tiempo;

combinar los picos ajustados de fase invertida no temporal con un espectro de ruido del espectro de señal para formar un espectro combinado para la segunda subtrama de ocultación; y

5 generar una segunda subtrama de audio de ocultación sintetizado con base en el espectro combinado.

13. El dispositivo decodificador de una cualquiera de las Reivindicaciones 10 a 12 adaptado además para obtener el espectro de señal de la señal de audio recibida previamente de una memoria del dispositivo decodificador.

14. El dispositivo decodificador de una cualquiera de las Reivindicaciones 10 a 13 adaptado además para asociar cada pico de los picos detectados con un número de contenedores de frecuencia de pico que representan el pico.

10 15. El dispositivo decodificador de la Reivindicación 14 está adaptado además para aplicar uno del ajuste de fase para una subtrama de ocultación invertida en el tiempo y el ajuste de fase para una subtrama de ocultación no invertida en el tiempo a cada contenedor de frecuencia pico del número de contenedores de frecuencia pico.

16. El dispositivo decodificador de la Reivindicación 15 está adaptado además para:

15 rellenar los contenedores restantes del espectro de la señal usando coeficientes del espectro de señal almacenado, conservando los coeficientes espectrales una propiedad deseada de la señal.

17. El dispositivo decodificador según la Reivindicación 16, en donde la propiedad deseada comprende la correlación con un segundo canal en un sistema decodificador multicanal.

20 18. El dispositivo decodificador según una cualquiera de las Reivindicaciones 10 a 17 adaptado para estimar la fase de cada uno de los picos calculando una estimación de fase para los picos de los picos ajustados en fase de acuerdo con:

$$\phi_i = \angle \hat{X}_{mem}(k_i) - f_{frac}(\phi_c + \pi)$$

$$f_{frac} = f_i - k_i$$

donde ϕ_i es una fase estimada en la frecuencia f_i , $\angle \hat{X}_{mem}(k_i)$ es un ángulo del espectro $\hat{X}_{mem}$ de la señal de audio recibida previamente en un intervalo de frecuencia k_i , f_{frac} es un error de redondeo, ϕ_c es una constante de ajuste, y k_i es $[f_i]$, donde $[\cdot]$ representa la operación de redondeo.

25

Figura 1

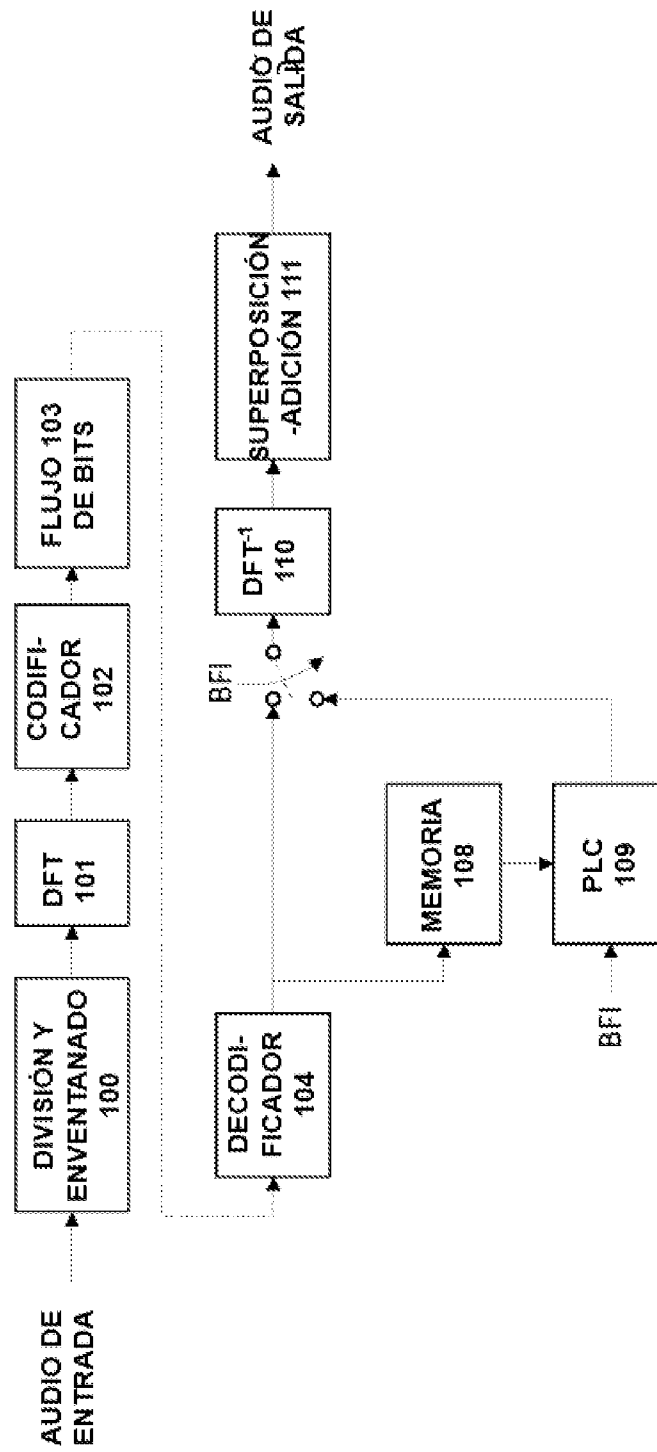


Figura 2

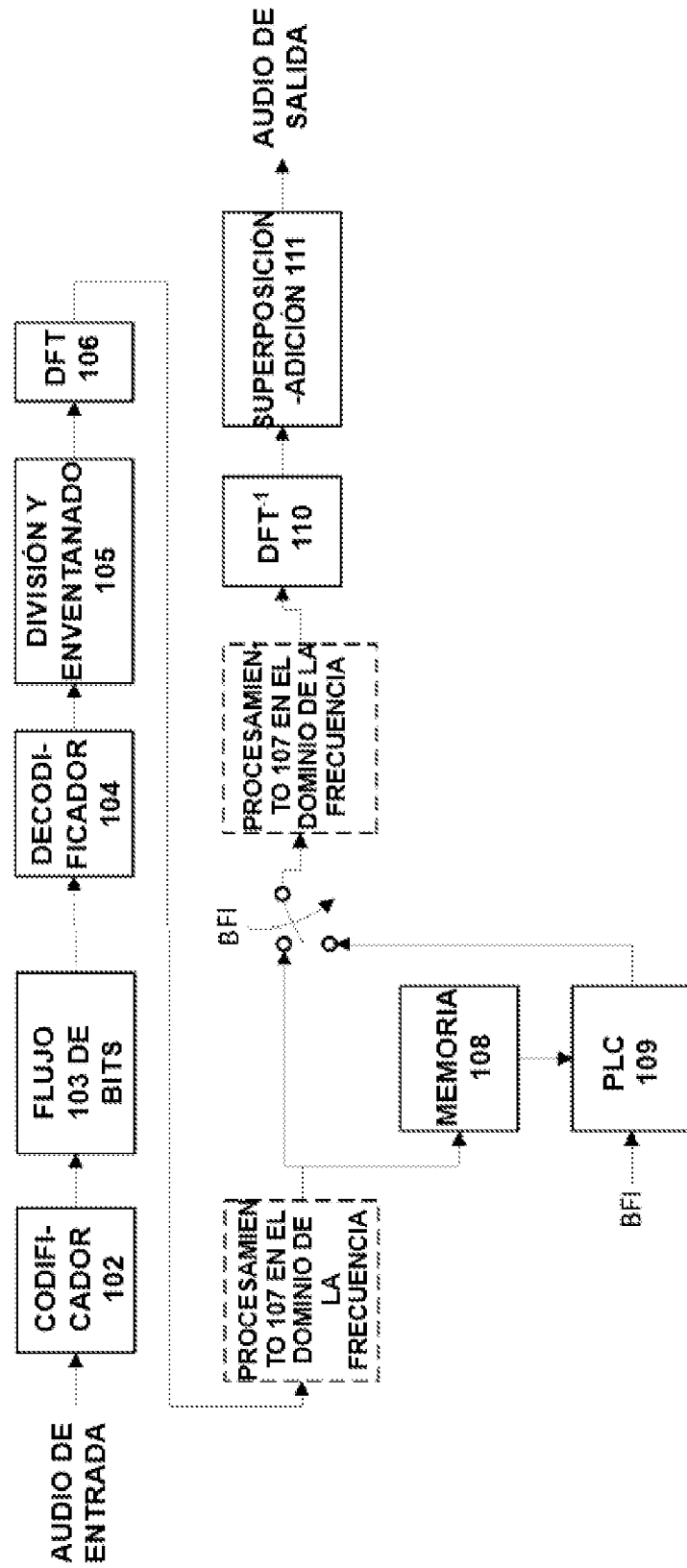


Figura 3

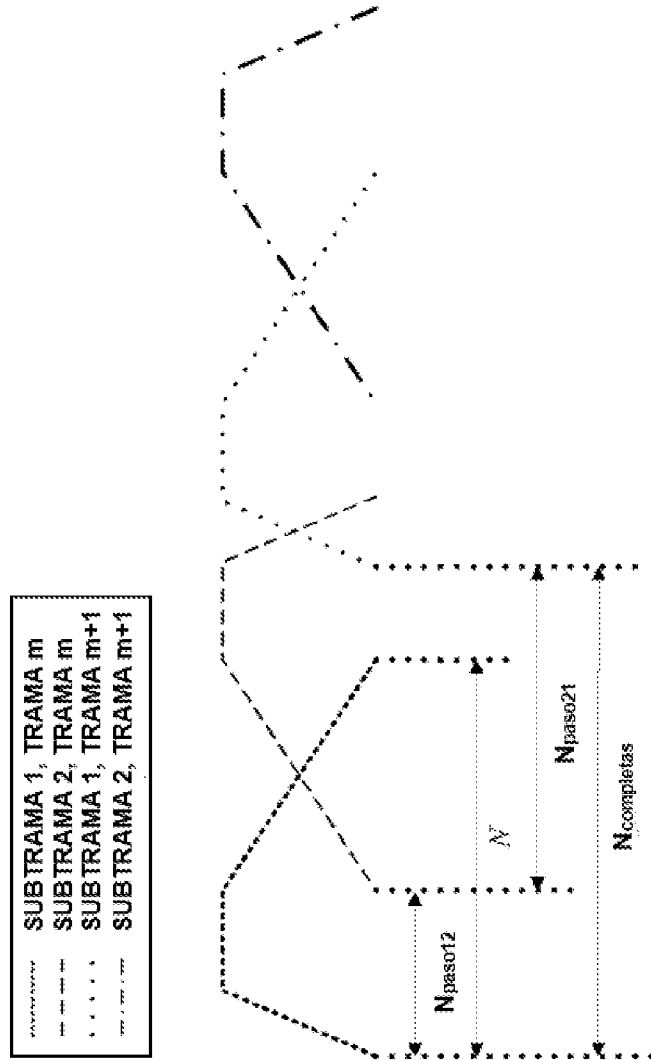


Figura 4

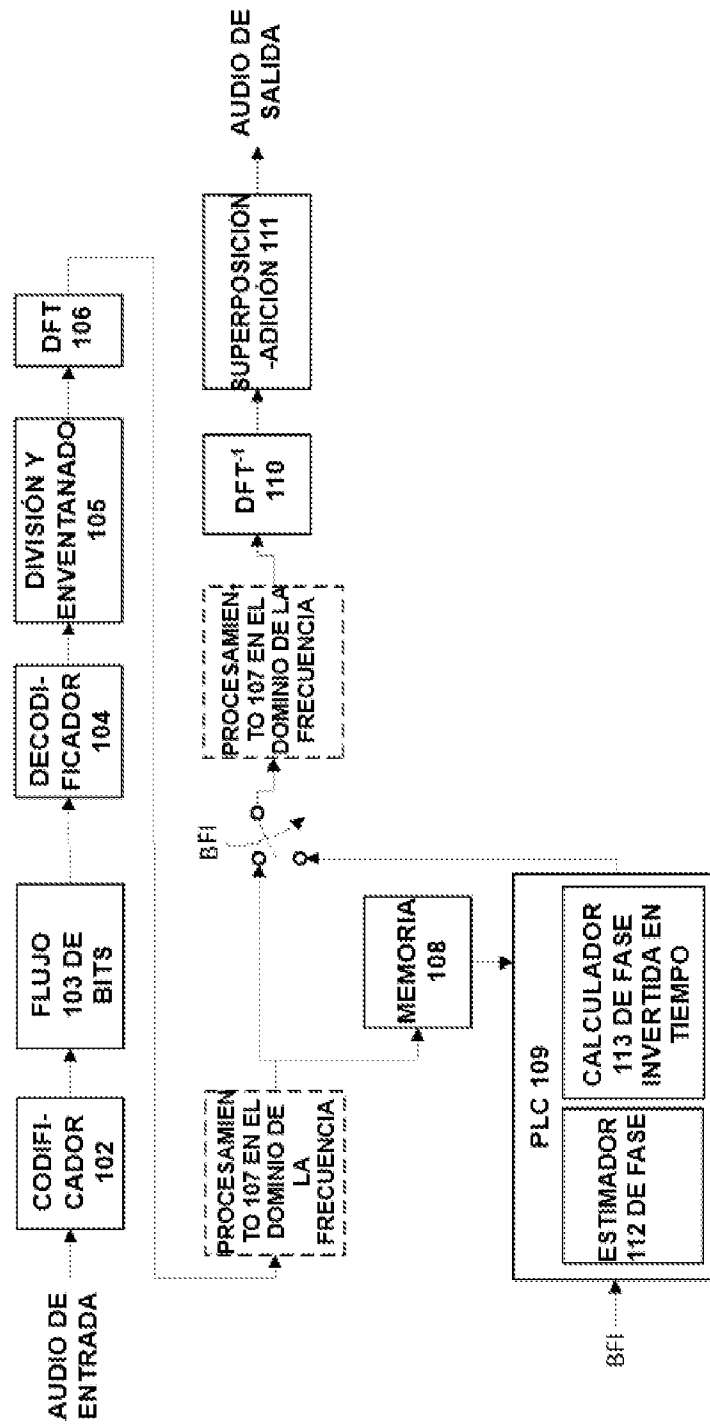


Figura 5

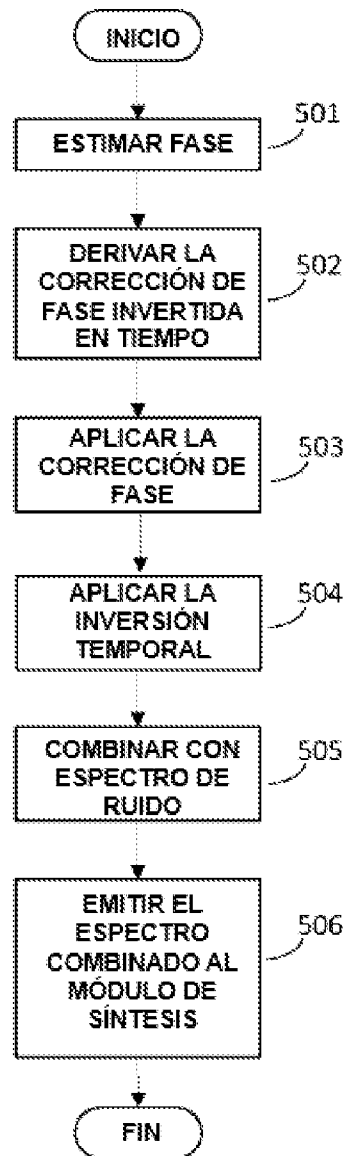


Figura 6

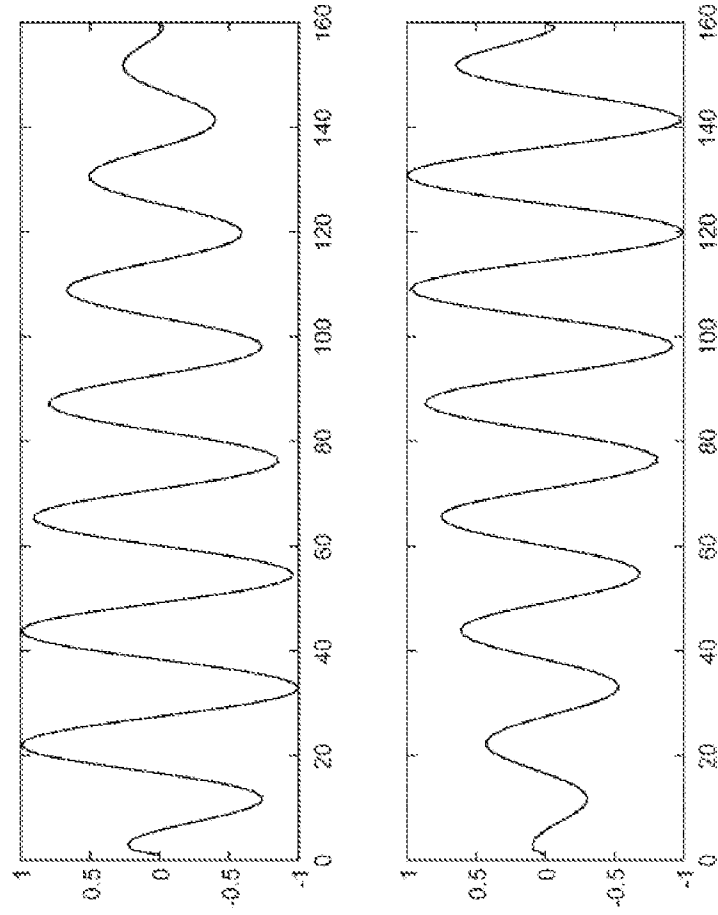


Figura 9

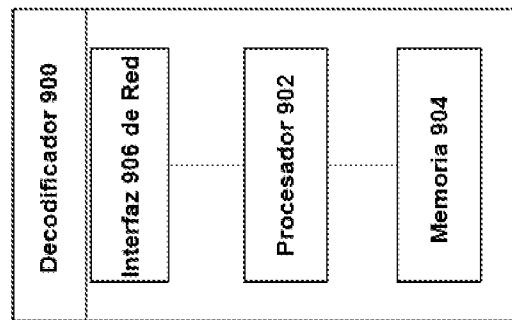
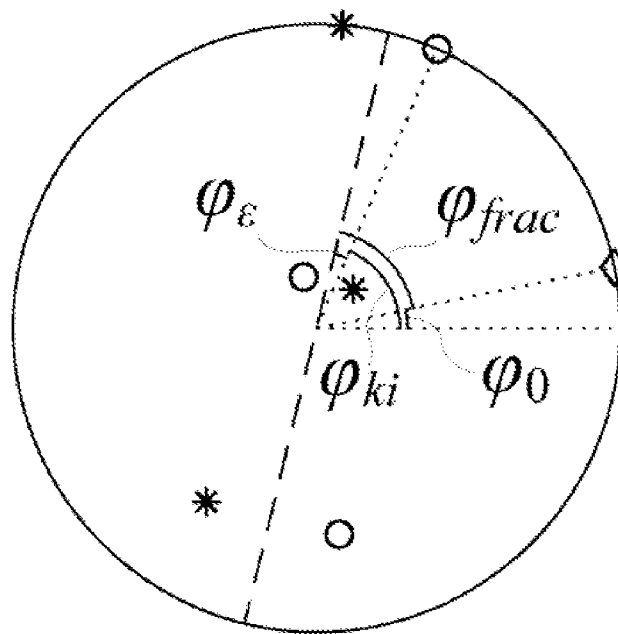


Figura 7



○	Entrada
*	Forma invertida en tiempo
— — —	Plano especular
◇	Fase

Figura 8

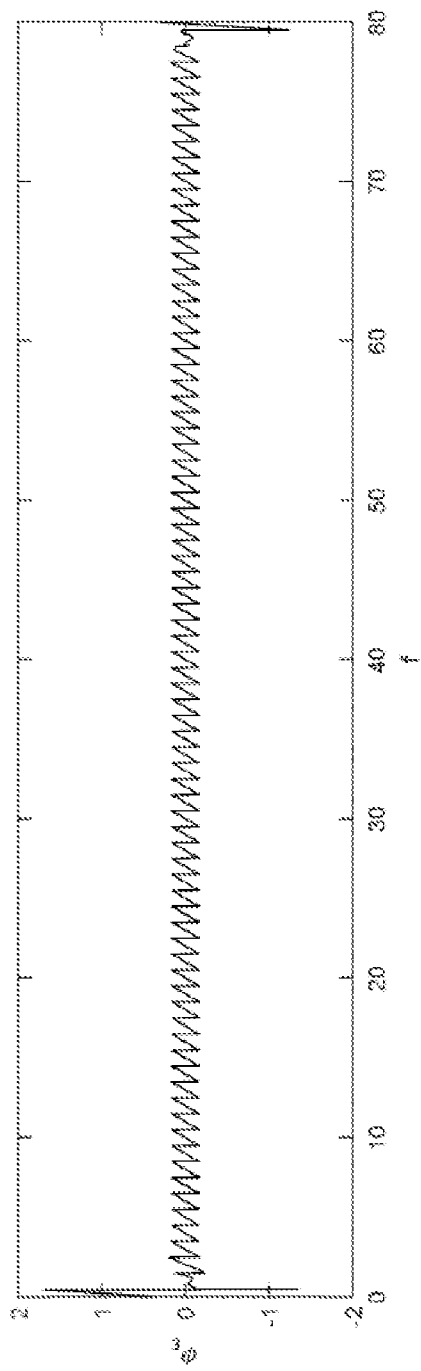


Figura 10

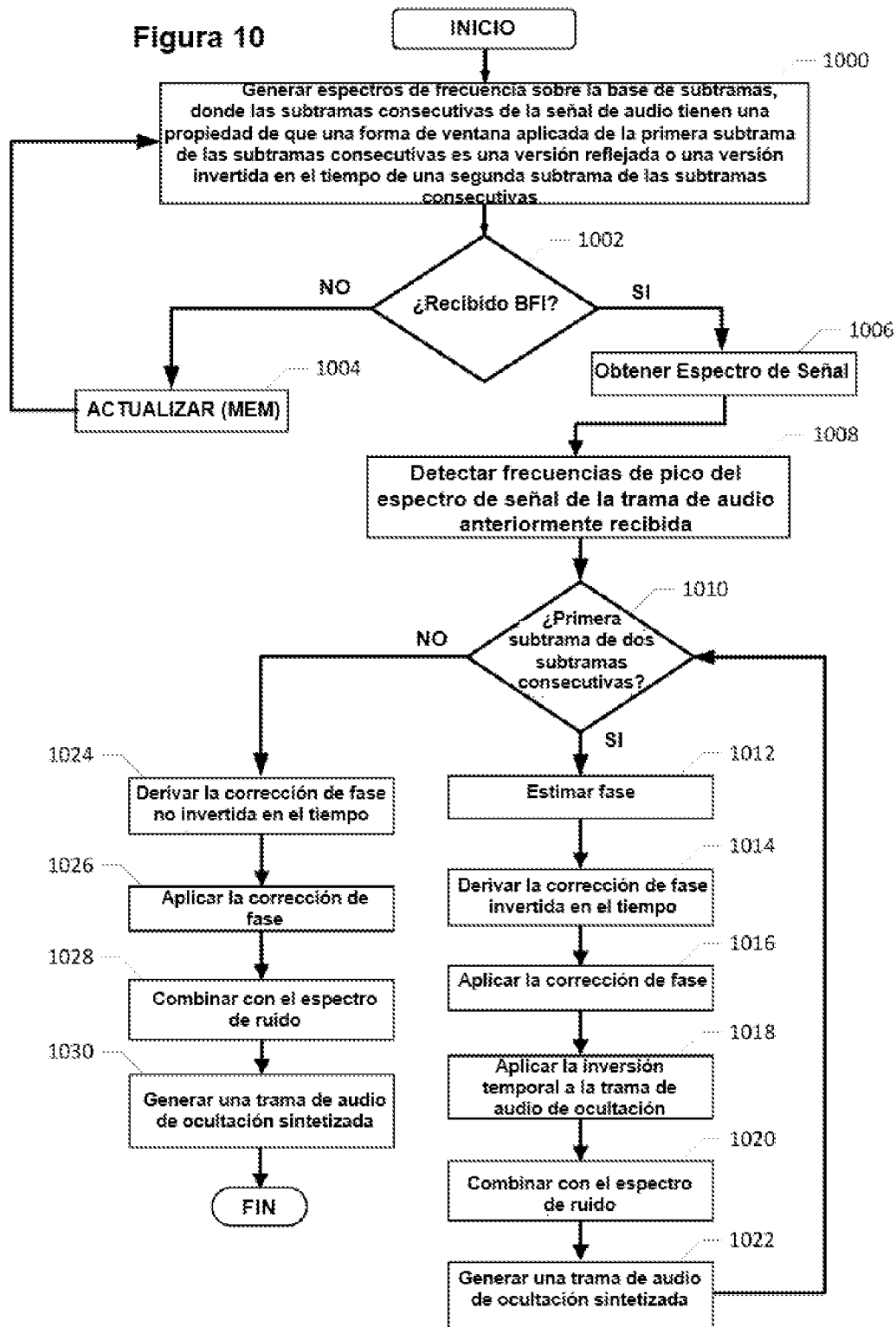


Figura 11

